



INFORMATICA MUSICALE

UNIVERSITA' DEGLI STUDI DI CATANIA
DIPARTIMENTO DI MATEMATICA E INFORMATICA
LAUREA TRIENNALE IN INFORMATICA
A.A. 2019/20
Prof. Filippo L.M. Milotta

ID PROGETTO: 0E

TITOLO PROGETTO: Sintesi del suono granulare, per modulazione e per distorsione

AUTORE 1: Mazzari Mattia

AUTORE 2: Nasca Prospero

AUTORE 3: Pachera Giovanni

Indice

1. Obiettivi del progetto	2
1.1 Comprendere la distinzione tra tecniche di sintesi del suono lineari e non lineari	2
1.2 Analizzare per via grafica le tecniche di sintesi	2
1.3 Vantaggi e limiti della modulazione	2
2. Riferimenti Bibliografici	3
3. Argomenti Teorici Trattati	3
3.1 Sintesi Del Suono E Campi Di Applicazione	3
3.2 Tecnica Lineare: Sintesi Granulare	4
3.3 Granulazione Dei Suoni	5
3.4 Tecniche Non Lineari: Sintesi Per Modulazione	6
3.5 Tecniche Non Lineari: Sintesi Per Distorsione	8
3.6 Espansori E Compressori Di Dinamica	8

1. Obiettivi del progetto

1.1 Comprendere la distinzione tra tecniche di sintesi del suono lineari e non lineari

La sintesi del suono si può ottenere a partire da tecniche lineari e non lineari. Nel primo caso, il segnale in uscita è il risultato di una trasformazione, in termini di incremento e decremento, proporzionale all'ingresso. In manipolazioni di questo tipo, si preferisce lavorare nel dominio del tempo. Le tecniche non lineari, invece, possono modificare ampiamente la forma d'onda iniziale, in maniera indipendente dalla complessità del suono in input. In genere, queste tecniche vengono scelte per la traslazione o l'arricchimento dello spettro di un segnale.

1.2 Analizzare per via grafica le tecniche di sintesi

In base alla tecnica scelta, è possibile far variare alcuni parametri, che caratterizzano la trasformazione. Visualizzare cosa succede alla forma d'onda nel dominio del tempo, o in quello della frequenza, consente di visualizzare meglio la manipolazione a cui è soggetto il segnale in ingresso.

1.3 Vantaggi e limiti della modulazione

La modulazione è una tecnica di sintesi che permette di calcolare in modo semplice lo spettro del segnale risultante. Inoltre, con due soli oscillatori si è in grado di generare infinite componenti (ciò rappresenta un vantaggio non indifferente in termini di risorse). Di conseguenza, si hanno ampie possibilità in ambito polifonico. Di contro, è importante puntualizzare che gli spettri di suoni meno caratterizzati non sono facilmente riconoscibili (questo comporta il calcolo della FM complessa). Tra l'altro, come tutte le sintesi non lineari, non si può "indovinare" a monte come produrre un suono ben preciso: occorre conoscerlo e studiarlo.

2. Riferimenti Bibliografici

Dispense DEI-UniPD, a cura dei prof: Giovanni De Poli, Carlo Drioli e Federico Avanzini
(<http://www.dei.unipd.it/~musica/IM/signalmodels.pdf>)

Elementi di informatica musicale – Ing. Antonio Rodà
(<https://digilander.libero.it/infomus/lucidi/sintesi.pdf>)

Sintesi sonora - Tommaso Rosati
(<https://www.tommasorosati.it/wp-content/uploads/2016/03/Sintesisonora.pdf>)

Effetti con unità di ritardo: distorsione non lineare, a cura del prof. Luca A. Ludovico -
UNIMI, Facoltà di scienze matematiche, fisiche e naturali
(<https://slideplayer.it/slide/17679976/>)

Audio e multimedia – Vincenzo Lombardo, Andrea Valle

3. Argomenti Teorici Trattati

3.1 Sintesi Del Suono E Campi Di Applicazione

La sintesi del suono riguarda la generazione di suoni mediante un ben preciso procedimento di calcolo. Nel caso di segnali numerici, viene utilizzata una legge matematica per calcolare il valore dei campioni che rappresentano il suono (nota: il suono digitale è una successione di numeri, detti campioni).

Questo tipo di manipolazione del segnale acustico si serve di due tipi di modello per rappresentare il suono.

Il primo, detto di sorgente, simula con un mezzo informatico il naturale processo fisico di generazione del suono.

Il secondo, detto di segnale, modella la forma d'onda che viene percepita dall'ascoltatore, specificando il dominio in cui viene studiata. Le sintesi che riguardano questo tipo di modello sono le seguenti:

- Sintesi per generazione: il suono viene prodotto tramite un processo generativo a partire dai parametri $p_1 - p_n$ (sintesi granulare).
- Sintesi per trasformazione: i parametri vengono prima sottoposti ad un processo generativo semplice e poi elaborati da uno o più processi di trasformazione (tecniche lineari e non lineari).
- Sintesi per interazione: si ha la generazione di un suono a partire dall'interazione tra un processo generativo semplice (eccitatore) e un processo di trasformazione (risonatore).

I campi di applicazione più importanti sono la simulazione dei suoni prodotti dagli strumenti musicali tradizionali e la generazione di suoni rispetto alle esigenze del musicista durante la composizione di brani.

I parametri del suono che possono essere gestiti sono l'altezza, l'intensità, la durata, il timbro e la localizzazione spaziale. Il modo in cui vengono scelti e combinati dipenderà dalla tecnica utilizzata.

3.2 Tecnica Lineare: Sintesi Granulare

“Ogni suono, 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 tempo più o meno brevi.”

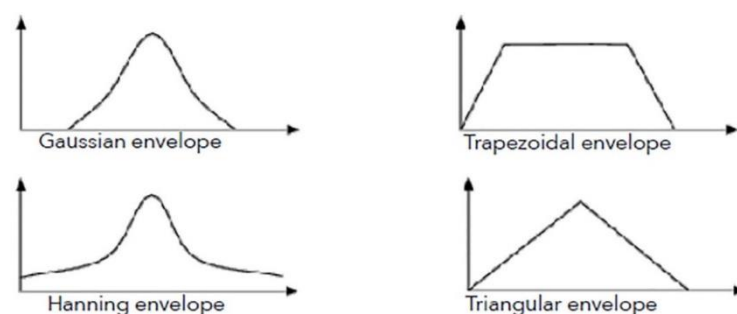
Tale affermazione appartiene a Xenakis, uno degli esponenti delle avanguardie musicali del Novecento, nonché uno dei primi ad applicare la sintesi granulare nel mondo della musica (1959).

Questa tecnica lineare di sintesi sonora si serve di più suoni elementari per costruirne uno complesso. Ciò è possibile mediante una successione di suoni di breve durata (tra 1 e 100 millisecondi), chiamati grani.

Questi ultimi possono essere combinati o riprodotti in sequenza o sovrapposti a velocità, fase, volume variabili. Il risultato che si ottiene non è un unico tono, ma un insieme di suoni.

È possibile ricavare i grani da campioni audio o da suoni di sintesi. Durante la fase di riproduzione, possono essere impostati diversi parametri, tra cui:

- L'involuppo. I tipi più comunemente scelti sono:



- La densità dei grani (numero di grani riprodotti nello stesso momento). Bassi valori di densità mettono in risalto i singoli suoni. Quelli alti, al contrario, generano delle sovrapposizioni tra i vari grani, producendo suoni continui o bande.

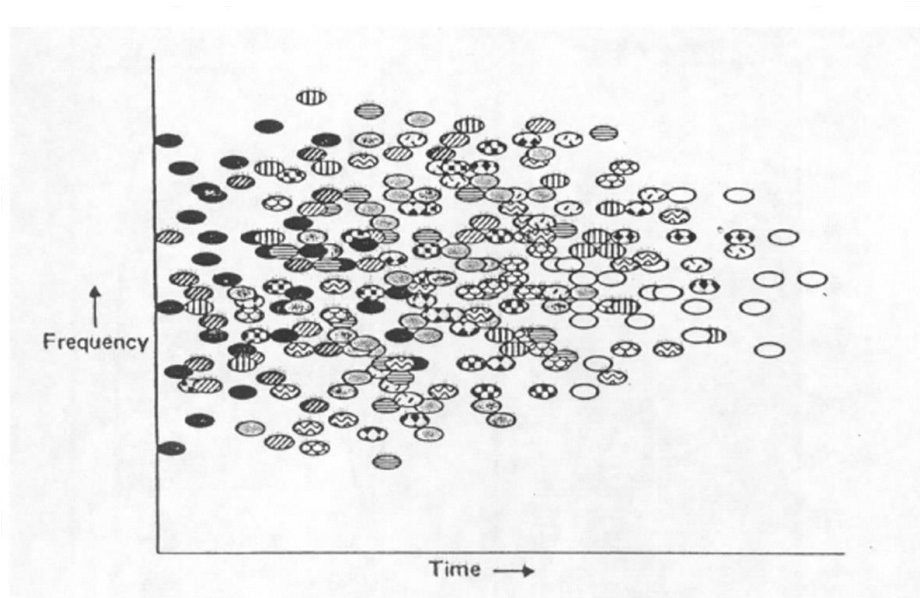
- Posizione nello spazio dei grani. Essi, infatti, possono essere distribuiti in diversi canali di riproduzione.
- Larghezza di banda della nuvola. La frequenza di ogni grano viene definita a partire dai limiti di minimo e di massimo entro i quali i grani sono dispersi. Quando essi coincidono, si ottiene una linea continua che, dal punto di vista sonoro, corrisponde ad un suono di altezza determinata.

Successivamente, i grani possono essere riprodotti in tre modi: sincrono (la distanza tra i grani è costante); quasi sincrono (la distanza tra i grani è quasi costante); asincrono (la distanza tra i grani non è costante).

3.3 Granulazione Dei Suoni

Esistono due diverse forme di sintesi granulare:

- granulazione di suoni sintetici (sintesi astratta). I grani vengono rappresentati attraverso forme d'onda arbitrarie, il cui inviluppo e la cui ampiezza sono funzioni gaussiane. Essi si distribuiscono sul piano cartesiano in forma di nuvola, in cui l'asse delle ordinate è legato alla frequenza, mentre sulle ascisse si trova il tempo.



- Granulazione di suoni reali (suoni pre-campionati sotto forma di file, oppure ottenuti in tempo reale tramite microfono). In questo caso, lo strumento utilizzato prende il nome di oscillatore, il quale viene controllato da un inviluppo.

Da un punto di vista pratico, la granulazione del suono si ottiene curando con precisione l'allineamento temporale dei grani, con l'obiettivo di evitare fenomeni di discontinuità di fase, che potrebbero produrre effetti acustici poco gradevoli.

Grazie ad una sovrapposizione parziale dei grani ("Overlap") si possono riprodurre suoni reali in modo preciso, avendo la possibilità di modificare in seguito le loro caratteristiche dinamiche.

Ogni grano viene moltiplicato per una finestra temporale, grazie alla quale si generano effetti di fade-in e fade-out.

I grani più lunghi manterranno l'identità timbrica del suono di partenza, mentre quelli più corti assumeranno un carattere impulsivo.

Questa tecnica di sintesi può essere ricondotta a scenari di vita reali. Per esempio, basti pensare all'applauso di una platea, di fronte all'esecuzione di un pianista. Dapprima, verrà udito solo qualche battito di mani, fino a raggiungere un suono decisamente più complesso: il cervello non tenderà a sommare i singoli suoni, ma percepirà un suono nuovo, come il risultato di battiti in posizioni diverse, con diversi volumi e profili d'inviluppo.

3.4 Tecniche Non Lineari: Sintesi Per Modulazione

Le trasformazioni lineari non sono in grado di modificare le frequenze delle componenti in ingresso. Per questo motivo, si ricorre a tecniche non lineari, capaci di cambiare abilmente la natura di un suono in input. Questa variazione non avviene proporzionalmente alla complessità dei suoni coinvolti, ma segue una legge del tutto diversa.

Gli effetti principali alla base della non linearità riguardano l'arricchimento dello spettro e la traslazione dello stesso. Il primo deriva da una distorsione non lineare del suono, tale da poterne modificare la brillantezza; il secondo, invece, si ottiene moltiplicando per una sinusoide (detta portante), in modo da spostare il suono intorno alla frequenza del segnale portante.

Con il passare del tempo, si è riusciti a includere questi due effetti in un'unica tecnica, detta sintesi per modulazione. Essa si verifica quando uno dei parametri di un oscillatore (fase, frequenza e ampiezza) viene modificato rispetto ad un altro segnale.

Il segnale trasformato prende il nome di portante ("carrier"), mentre quello da cui dipende la modulazione si dice modulante ("modulator").

La modulazione può avvenire in ampiezza e in frequenza: due segnali di controllo (modulanti) modificano periodicamente i segnali originali (portanti).

Tremolo e vibrato costituiscono due esempi di questo processo. Nel primo caso, la variazione viene fatta al di sotto della soglia di udibilità (sotto i 20 Hz): la modulazione viene così percepita come una trasformazione prettamente espressiva. Nel secondo, invece, se la frequenza della modulante rientra nel campo uditivo, il risultato ottenuto è un cambiamento qualitativo del timbro (a partire da una modifica dello spettro).

Se il segnale modulante controlla l'ampiezza della portante, si possono avere due tipi di modulazione in ampiezza: quella ad anello ("Ring Modulation", RM) e quella d'ampiezza ("Amplitude Modulation", AM).

La modulazione ad anello si ha se il segnale modulante è bipolare, cioè si estende tra un massimo positivo e uno negativo. Il segnale risultante avrà uno spettro complesso formato da due frequenze, pari a $C - M$ e $C + M$, dette bande laterali.

Al contrario, la modulazione d'ampiezza lavora con segnali modulanti unipolari, vale a dire compresi completamente tra valori positivi. In questo caso, lo spettro del segnale risultante conserva anche la componente frequenziale della portante, oltre alla somma e alla differenza di portante e modulante. Se la differenza è minore di zero si ha un'inversione di fase, cioè la frequenza in questione diventerà positiva.

Durante lo studio delle caratteristiche spettrali vanno presi in considerazione diversi parametri. Il primo è il rapporto tra le frequenze di modulante e portante e viene definito come C:M ratio. Esso rappresenta un indicatore dell'armonicità dello spettro: se si ha un valore intero lo spettro sarà armonico. In generale, più è semplice la frazione, più sono vicini gli armonici risultanti. Quando la C:M ratio è quasi intera, si ottiene un suono che viene percepito in maniera meno artificiale, in quanto simula le inarmonicità presenti negli strumenti acustici.

L'altro fattore fondamentale è l'ampiezza della modulante, in grado di condizionare direttamente l'energia delle bande laterali. Dal rapporto tra l'ampiezza della modulante e quello della portante, è possibile ricavare l'indice di modulazione, un altro parametro descrittivo di questo tipo di trasformazione.

Nella modulazione in frequenza (FM) la modulante modifica la frequenza della portante. La frequenza di C subisce delle variazioni, che renderanno il suono più acuto (l'output di M è positivo, quindi C aumenta) o più negativo (l'output di M è negativo, dunque C diminuisce). Per quantificare la massima escursione in frequenza subita dalla portante (misurata in Hz), si ricorre ad un altro parametro detto frequenza di deviazione di picco. Il grande vantaggio della FM, rispetto ad AM e RM, dipende dal fatto che essa permette di creare degli spettri anche molto complessi attraverso una tecnica di grande semplicità computazionale. In linea del tutto teorica, la modulazione FM consente di generare una serie infinita di bande laterali, ma in realtà quelle timbricamente significative sono poche. Per avere un'idea di quante bande possano essere tali, si ricorre al rapporto tra la deviazione e la frequenza della modulante, detto indice di modulazione.

$$I = \frac{D}{M}$$

Assumendo D e M costanti (in quanto legati a modulante e portante), l'indice fornisce una misura del numero di bande laterali dello spettro. In questa relazione è racchiusa un'altra importante informazione: la ricchezza dello spettro (cioè il numero di bande laterali) è proporzionale all'ampiezza del segnale modulante; la sua natura (cioè la posizione di queste bande), invece, dipende dalla relazione tra portante e modulante.

3.5 Tecniche Non Lineari: Sintesi Per Distorsione

La sintesi per distorsione ("waveshaping") rappresenta un altro esempio di tecniche di sintesi non lineari. Essa permette di arricchire con diverse armoniche il segnale di partenza (in genere una sinusoide), che viene fatto passare attraverso un blocco distorcente, in grado di modificarlo e creare altre componenti spettrali.

Basti pensare, ad esempio, al clipping: esso produce in uscita una quadratura del segnale di ingresso, alterando la forma d'onda dello stesso.

Questo blocco è descritto da una funzione distorcente definita come somma di polinomi di Chebishev. Essi vengono appositamente scelti per poter essere in grado di prevedere il numero di armoniche che avrà il segnale in uscita (cosa che, in generale, non è facile da prevedere a priori). La funzione distorcente dipende solo dal valore istantaneo dell'ingresso e, nel caso della sintesi, assume la seguente forma:

$$s(t) = F[x(t)] = F[I \cos(2\pi ft)]$$

dove I rappresenta l'ampiezza della sinusoide $x(t)$, scelta come ingresso.

Essa si applica attraverso una tabella con valori già calcolati: in corrispondenza dell'ampiezza del segnale in ingresso viene riportato il valore dell'ampiezza assunta dall'uscita, ottenuta da questa trasformazione.

Supponendo che l'ampiezza I vari tra 1 e -1 nella sua forma normalizzata, la funzione distorcente contenuta in tabella andrà a mappare tutti i valori in entrata in un nuovo intervallo, sempre compreso tra -1 e 1.

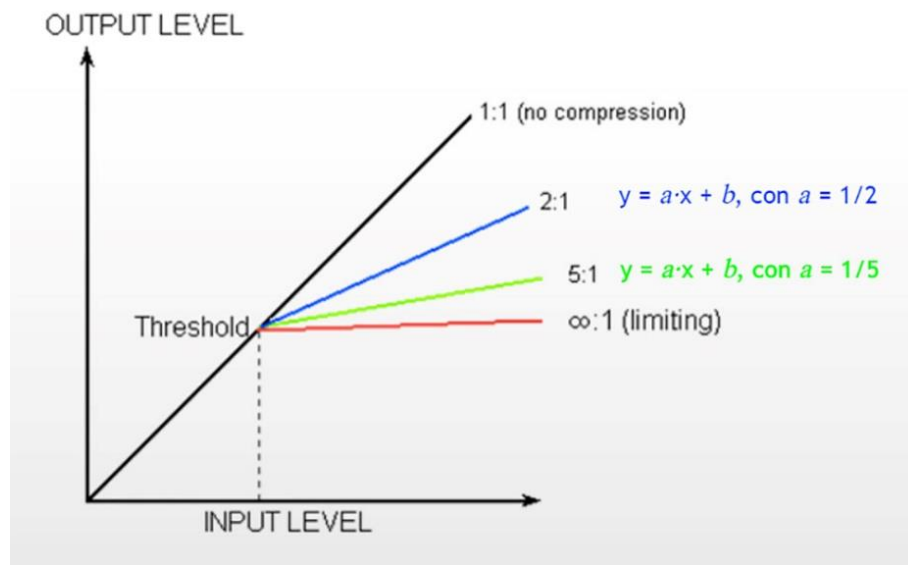
Come accennato in precedenza, una funzione distorcente può generare, in via del tutto teorica, infinite armoniche, ma se la legge è polinomiale di grado N , vengono prodotte solo le prime N armoniche. Questo consente di conoscere l'ampiezza dell'armonica k -esima. Chiaramente, se l'ampiezza del segnale di ingresso subisce una variazione, cambierà anche lo spettro prodotto per distorsione (si può pensare a questa dinamica come ad una contrazione o espansione della funzione stessa).

3.6 Espansori E Compressori Di Dinamica

Se la funzione è del tipo: $f(x) = ax$, non si verifica alcuna distorsione: l'uscita avrà lo stesso andamento dell'ingresso, a meno di un fattore moltiplicativo (maggiore o minore di 1, cioè legato a espansori o compressori di dinamica). In questo caso, viene meno il comportamento non lineare: l'uscita viene incrementata o decrementata rispetto all'ingresso in modo proporzionale.

Una concreta applicazione di questo processo può essere individuata in due tecnologie: l'espansore e il compressore.

Il primo è un dispositivo che, per certi valori di ampiezza del segnale di ingresso, fa variare il segnale in uscita ($a > 1$). Esso viene utilizzato per ampliare la gamma dinamica di un dato segnale, percepito come piatto.



Il compressore, invece, è un dispositivo che fa variare in minor misura il segnale in uscita ($a < 1$). Contrariamente al caso precedente, esso viene scelto per limitare la gamma dinamica di un segnale, in modo che i suoi picchi non generino effetti di distorsione.

Entrambi i dispositivi possono essere utilizzati per migliorare i processi di registrazione e riproduzione del suono. In genere, il suono in ingresso al registratore viene sottoposto ad una compressione in ampiezza, per evitare il clipping. Dopo aver ridotto opportunamente la gamma dinamica, si passa alla fase successiva: durante la riproduzione, interviene l'espansore, che riporta il segnale ai livelli originali.

Per lavorare con leggi di questo tipo, si definiscono diversi indici di riferimento:

- soglia di compressione: stabilisce il livello di volume a partire dal quale si vuole che il compressore intervenga (viene espressa in dB);
- rapporto di compressione ("ratio"): decide di quanto debba essere compresso il segnale in input che supera la soglia. Quando la "ratio" non è del tipo $x:1$ (come nel caso della compressione), ci si riferisce all'espansore ($1:x$).
- Tempo di attacco: espresso in millisecondi, indica quanto rapidamente debba agire il compressore, non appena il segnale supera la soglia. Un valore medio lungo permette di agire in maniera più morbida sul suono, ma potrebbe comportare un eccessivo superamento della soglia. Di contro, comprimerlo troppo rapidamente potrebbe chiudere in modo brusco l'attacco di alcuni strumenti.
- Tempo di rilascio: generalmente più lento del precedente, stabilisce dopo quanto tempo debba terminare l'azione del compressore. Un tempo di rilascio breve

permette al dispositivo di seguire le variazioni di volume del segnale in modo fedele, andando tuttavia incontro a possibili distorsioni o oscillazioni dello stesso. Un rilascio lungo, invece, risulta più naturale, ma non deve agire su altri suoni, dopo che quello soggetto a compressione è rientrato sotto soglia.

- Gain: rappresenta il guadagno da applicare al segnale compresso in uscita, scelto in modo tale da compensare eventuali perdite di volume.

In conclusione, è possibile affermare che la sintesi sonora rappresenta uno strumento molto potente per creare un suono, a partire da leggi matematiche ben precise. In base all'effetto che si vuole ottenere a partire da un certo segnale in ingresso, si opterà per tecniche di sintesi lineari o non lineari, tenendo a mente che le trasformazioni non lineari sono capaci di alterare in gran misura la forma d'onda di partenza.