

Relazione progetto Computer Music

Jacopo Lovatello – M. 862780 – CP. 10546576

Abstract

Il progetto realizzato per il corso di Computer Music vuole essere un applicativo in grado di sintetizzare in tempo reale texture musicali ricevendo come input parametri emotivi esprimibili anche da un'utenza priva di conoscenze tecniche e teoriche nei campi di sound design e composizione musicale.

Considerando la necessità di uno strumento di sintesi efficace affiancato a una grammatica utile per la composizione algoritmica, si è optato per il linguaggio SuperCollider per l'intera realizzazione del progetto.

La comunicazione tra l'algoritmo di composizione e i motori di sintesi è affidata al protocollo MIDI per merito della sua semplicità di utilizzo e della sua ampia diffusione, così che sia possibile riutilizzare l'algoritmo di composizione con diversi motori di sintesi e viceversa.

Con lo scopo di ottenere un timbro semplicemente assimilabile anche da un'utenza meno esperta, è stata scelta la sintesi sottrattiva, tramite modello sorgente – filtro, come tecnica di sintesi principale, la quale è stata arricchita da un processamento non lineare del segnale in grado di modificare il contenuto frequenziale del suono generato tenendo in considerazione i parametri emotivi specificati dall'utente.

Per la composizione algoritmica è stato implementato un automata stocastico utile generare in tempo reale una linea melodica non ripetitiva né completamente casuale, ma che tiene conto di regole di composizione musicale per il calcolo delle tabelle di probabilità.

Music Emotion Detection come strumento per la sintesi

Con lo scopo di realizzare un applicativo utilizzabile da utenti senza conoscenze in ambito musicale, sono stati scelti come valori di input parametri emotivi, questo perché anche un'utenza meno esperta è in grado, intuitivamente, di associare una determinata emozione a un brano musicale.

Considerando la necessità di modificare parametri emotivi in tempo reale sia per il design del suono che per l'aspetto compositivo si è optato per un approccio dimensionale basato sullo spazio di Valence e Arousal.

L'approccio dimensionale basa la sua descrizione delle diverse emozioni su un set ridotto di dimensioni che corrisponde alla rappresentazione umana interna delle emozioni.

Lo spazio utilizzato nell'applicativo utilizza due feature: Valence, che indica la positività, o negatività, dello stato emotivo provato, e Arousal che identifica il livello di attivazione dell'emozione.

Le emozioni vengono dunque rappresentate come punti nello spazio di Valence e Arousal e, per far corrispondere ciò all'immissione dei parametri da parte dell'utente, è stato utilizzato uno slider bidimensionale in grado di rappresentare lo spazio sopra illustrato.

I valori di Valence e Arousal specificati vengono successivamente trasformati in parametri e feature di basso e medio livello da fornire ai motori di sintesi e all'algoritmo di composizione che li utilizzeranno per generare la texture musicale.

Motori di sintesi

La generazione timbrica, in grado di rispecchiare i parametri emotivi indicati dall'utente, è stata affidata a 4 diversi motori di sintesi: due sorgenti sonore e due modificatori di sorgenti sonore.

Per sorgente sonora si intende uno strumento dedicato alla creazione di un segnale audio.

Nelle tecniche di sintesi classiche, derivate dalle tecnologie analogiche, le sorgenti sono realizzate tramite oscillatori: dispositivi atti alla generazione di forme d'onda periodiche a una determinata frequenza e ampiezza.

Per modificatore di sorgente sonora si intende uno strumento in grado di modificare, sotto diversi aspetti, un segnale audio ricevuto in ingresso.

I principali modificatori di sorgente sonora si suddividono in sistemi lineari e sistemi non lineari: la differenza principale tra le due tipologie sta nella capacità di aggiungere armoniche, inizialmente assenti nel segnale originale, da parte dei sistemi non lineari, mentre i primi sono soltanto in grado di alterare il contenuto frequenziale già presente nel segnale originale.

Nell'applicativo le sorgenti sonore sono state realizzate seguendo la classica tecnica della sintesi sottrattiva, in grado, attraverso la generazione di un timbro ricco processato da sistemi lineari, di ottenere suoni facilmente prevedibili nel loro contenuto spettrale e quindi adatti a una struttura melodica.

I modificatori di sorgenti sonore implementati sono invece caratterizzati da un comportamento non lineare in grado di arricchire e degradare il segnale audio tenendo in considerazione gli effetti necessari a realizzare un timbro utile a rappresentare diversi contenuti emotivi.

I motori di sintesi sono stati implementati in SuperCollider tramite SynthDef: definizione che contiene le specifiche sulle UGen (funzioni in grado di generare segnali audio o di controllo) e sulle loro interconnessioni.

Di seguito i motori di sintesi nel dettaglio.

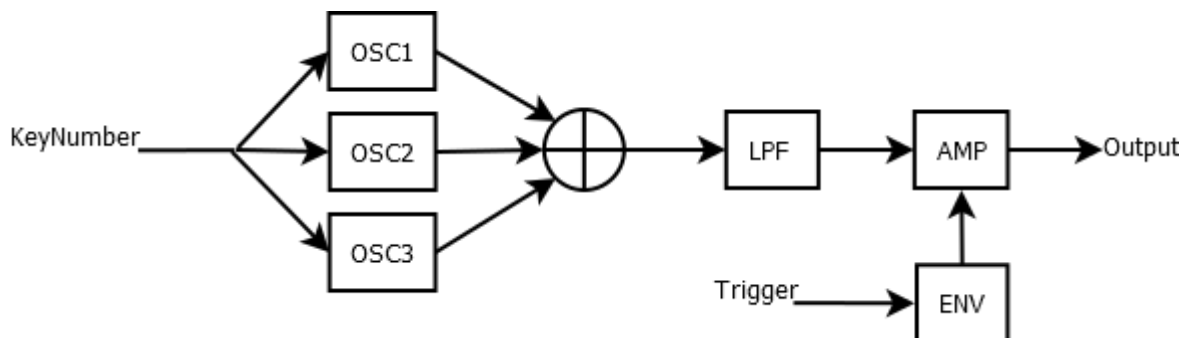
Sorgenti sonore

- **Melody:** motore di sintesi utilizzato per la creazione della linea melodica della texture musicale, in particolare è composto da 3 oscillatori a dente di sega i quali segnali sommati vengono processati in primo luogo da un filtro passa basso e in seguito da un amplificatore. I 3 oscillatori ricevono, dall'algoritmo di composizione automatica, il medesimo valore di Key Number il quale viene leggermente modificato, diversamente per ogni oscillatore, dal parametro detune con lo scopo di ottenere un suono più ricco e movimentato per merito dei battimenti, utile nel caso l'utente abbia specificato un valore di Arousal alto. La frequenza di taglio del filtro passa basso è posta 3 ottave sopra il valore di intonazione degli oscillatori con l'intento smorzare le armoniche classiche dell'onda a dente di sega ottenendo così un suono più tenue e adatto a strutture melodiche. Per fornire un'articolazione temporale all'output, il segnale uscente dal filtro viene processato da un amplificatore la cui ampiezza nel tempo è scolpita da un inviluppo attivato dall'algoritmo di composizione automatica e i cui valori di durata, attacco e decadimento sono modificati a seconda dei parametri di Valence e Arousal specificati: per alti valori di Arousal si avrà un inviluppo più percussivo.

Di questo motore di sintesi vengono richiamate due diverse istanze: la prima per lo sviluppo melodico, la seconda che agisce da bordone per dare una maggiore identità tonale al fraseggio.

Per l'implementazione in SuperCollider delle componenti sopra elencate sono state utilizzate le seguenti UGen:

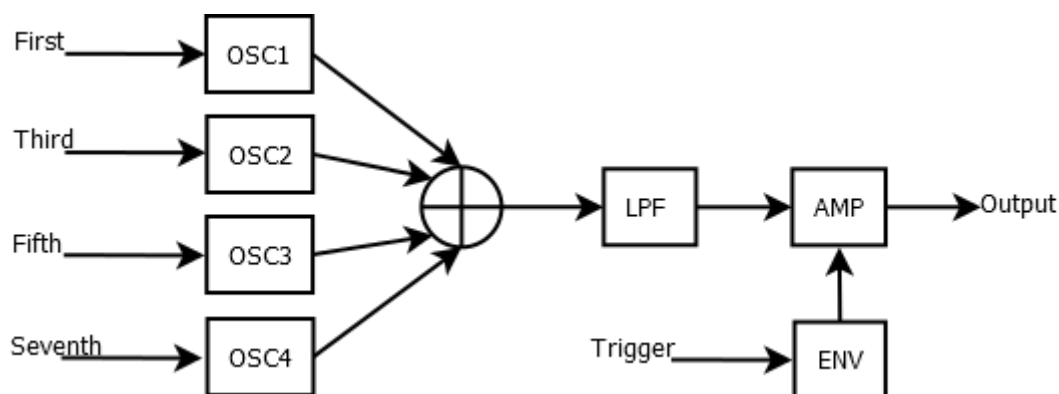
- LFSaw: oscillatore a dente di sega limitato in banda;
- LPF: filtro passa basso del secondo ordine;
- EnvGen: generatore di inviluppo utilizzato tramite specifica Env.perc per inviluppi composti dalle fasi di attacco e decadimento.



- Harmony: il Synth in analisi viene utilizzato per la creazione dell'armonia del brano, come nel caso precedente segue il paradigma classico della sintesi sottrattiva, l'unica differenza sta nel numero degli oscillatori, ora 4, ognuno utilizzato per la creazione di una voce dell'accordo.

Le sorgenti ricevono 4 diversi valori di MIDI Key Number a seconda dell'accordo da creare, il segnale da loro generato segue lo stesso processamento del synth Melody tipico della sintesi sottrattiva: prima avviene un filtraggio spettrale tramite un filtro passa basso, in seguito viene fornita un'articolazione temporale al segnale tramite un amplificatore il cui valore ampiezza nel tempo è scolpito da un inviluppo la cui forma (attacco, decadimento e durata) è modificata a seconda dei valori di Valence e Arousal indicati dall'utente.

Per l'implementazione in SuperCollider delle componenti sopracitate sono state utilizzate le medesime UGen del motore di sintesi melodico.



Modificatori di sorgenti sonore

Lo sviluppo armonico e quello melodico generati dai due motori sopra citati vengono processati, in ordine di esposizione, dai seguenti modificatori di segnale:

- **Distorsion:** sistema non lineare a due canali, pulito e distorto, in grado di arricchire il contenuto spettrale del segnale ricevuto come input tramite distorsione.
L'effetto è controllato da 3 parametri: dry, utile a indicare l'ampiezza del segnale pulito, wet, che incide sulla presenza del segnale distorto e clip che determina la quantità di distorsione applicata.
Tramite questo effetto è possibile aggiungere diverse armoniche non presenti inizialmente nel segnale originale che contribuiscono alla degradazione dello stesso, utile nel caso si vogliano rappresentare emozioni negative con valore di Valence basso.
Per l'implementazione in SuperCollider delle componenti sopra descritte è stata utilizzata la funzione `softclip` in grado di distorcere non linearmente in segnale fornito.
- **Reverb:** riverbero utilizzato per degradare ed arricchire il timbro processato dal distorsore. Il segnale in ingresso viene innanzitutto processato da un controllo di tono, utile per scurire il timbro nel caso di valori di Valence bassi, implementato tramite un filtro passa banda controllato dai parametri di cutoff (frequenza di taglio) e quality factor (la larghezza della banda passante).
Il segnale uscente dal filtro passa banda viene processato dall'effetto di riverbero vero e proprio, controllato dai parametri di roomsize, utile a rendere il suono più ricco e confuso, dampening, in grado di smorzare le alte frequenze generate dal riverbero, e mix tra segnale pulito e quello processato.
Per l'implementazione in SuperCollider delle componenti sopra citate sono state utilizzate le seguenti UGen:
 - BPF: filtro passa banda Butterworth del secondo ordine utilizzato per il controllo di tono;
 - FreeVerb: effetto di riverbero basato sul riverberatore di Schroeder.

Algoritmo di composizione automatica

L'algoritmo realizzato per la composizione automatica fa uso di automata stocastici per la selezione dell'intervallo, specificato in standard MIDI, tra l'attuale nota e la successiva dello sviluppo melodico.

Per comportamento stocastico si intende una procedura che comprende decisioni casuali in grado di generare eventi basandosi su tabelle di probabilità utili a pesare l'avvenimento di un determinato evento rispetto ad altri, ad essi si contrappongono i comportamenti deterministici, il cui comportamento è fisso e potenzialmente molto complesso.

Per automata si intende una procedura il cui output è determinato dallo stato attuale e dall'input, esso può essere deterministico o stocastico a seconda della presenza o meno di comportamenti casuali sia nell'assegnazione di un valore a un determinato stato che nel passaggio da uno stato all'altro, che rendono imprevedibile il comportamento complessivo dell'automata; contrariamente, in un automata deterministico, è sufficiente essere a conoscenza dello stato e dell'input iniziali per predire il suo intero sviluppo.

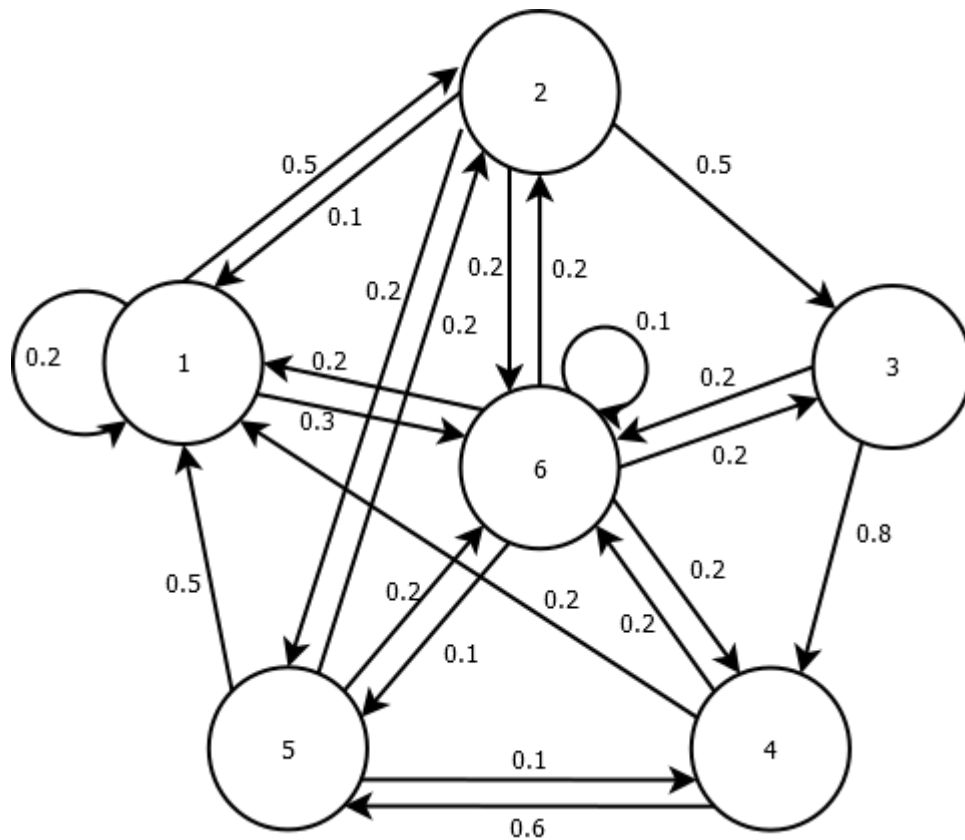
A ogni automata finito è possibile associare un diagramma di stato contenente un nodo per ogni stato e un arco per ogni coppia stato – input, nel caso di un automata stocastico come quello realizzato per l'applicativo, l'input è rappresentato da un valore casuale estratto da una tabella di probabilità.

Gli automata vengono utilizzati nella composizione algoritmica per stabilire diversi parametri musicali all'interno dei singoli stati così da ottenere una sequenza di valori utili per la generazione di una melodia. Essi possono essere inoltre utilizzati per la specifica di parametri dedicati alla generazione timbrica del suono (e.g. indici di modulazione, frequenze di taglio dei filtri).

Nell'applicativo realizzato è stato utilizzato un automata stocastico con lo scopo di rendere la texture musicale non ripetitiva, nonostante ciò si è cercato di non ottenere un comportamento completamente casuale, ma uno sviluppo melodico influenzato da regole di composizione musicale implementate tramite adeguate tabelle di probabilità.

La casualità data dal comportamento stocastico e l'implementazione in tempo reale dell'algoritmo hanno permesso di ottenere buoni risultati anche con l'utilizzo di un automata relativamente ridotto.

Di seguito lo schema dell'automata implementato.



Stato	Intervallo	Ampiezza
1	4	1
2	-1	1
3	3	1
4	1	1
5	4	1
6	0	0

Ogni stato contiene l'intervallo tra l'attuale nota e la successiva e il suo valore di ampiezza: come deducibile dalla tabella, lo stato 6, contenente valore di ampiezza pari a 0, è stato inserito con lo scopo di implementare la pausa musicale.

L'algoritmo calcola un nuovo stato, e quindi un nuovo valore di MIDI Key Number per lo sviluppo melodico, periodicamente, con un intervallo temporale specificato dal parametro tatum, attraverso la funzione Task di SuperCollider; il valore estratto ogni 4 iterazioni viene inoltre utilizzato per la costituzione dello sviluppo armonico: tramite un assegnamento stocastico viene calcolata la voce, a scelta fra prima, terza, quinta e settima, a cui assegnare il valore estratto dallo sviluppo melodico, le restanti voci dell'accordo vengono calcolate di conseguenza.

La sequenza melodica viene estratta da un array contenente 3 ottave di una scala in un determinato modo musicale scelto a seconda del valore di Valence specificato.

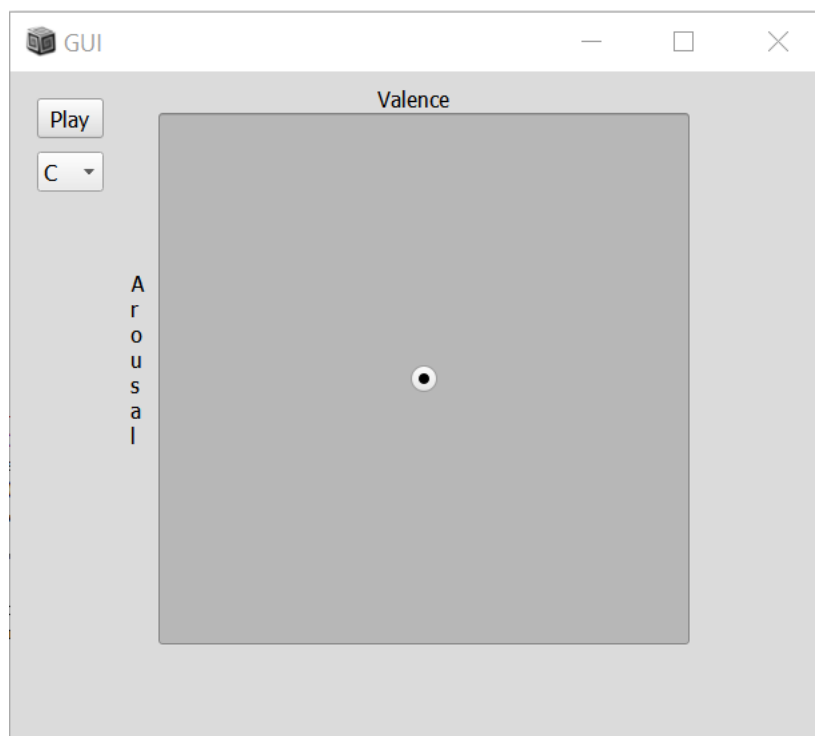
Secondo la teoria del Modal Envelope è possibile assegnare un diverso modo musicale per parametri di Valence differenti, ottenendo diversi riscontri emotivi; ciò è stato implementato modificando la scala da cui l'algoritmo estrae le note a seconda del valore di Valence attualmente specificato dall'utente, mappando linearmente i 7 diversi modi musicali (lidio, ionico, misolidio,

dorico, eolio, frigio e locrio) sulla dimensione di Valence. Al fine di rendere maggiormente presente la caratteristica tonale del fraseggio, e di conseguenza ottenere una migliore resa emotiva per i diversi modi utilizzati, è stata richiamata una seconda istanza del Synth Melody agente da bordone intonato sulla tonica.

GUI e interazione

Tramite l'interfaccia grafica l'utente può attivare e disattivare la generazione della texture, tramite bottone Play/Stop, sceglierne la tonalità, attraverso il menu a tendina dedicato, e agire sui valori di Valence e Arousal tramite slider bidimensionale.

Si è optato per un'implementazione minimalistica dell'interfaccia basata solo su SuperCollider per permettere un semplice utilizzo su diverse macchine, senza la necessità di installare molteplici software; sono stati inseriti solamente i controlli sopraelencati per permettere un semplice utilizzo anche a un'utenza meno esperta, priva di conoscenze musicali e di sound design.



I parametri di Valence e Arousal influenzano i motori di sintesi e l'algoritmo di composizione come segue:

- Valence:
 - Frequenza di taglio del filtro passa banda utilizzato per il controllo di tono del riverbero: un valore di Valence minore corrisponde a una frequenza di taglio più bassa che contribuisce all'ottenimento di un suono più cupo e scuro.
 - Roomsize e mix del riverbero: un valore di Valence minore corrisponde a un suono più riverberato e confuso.

- Wet, Dry e Mix del distorsore: un valore di Valence minore corrisponde a un suono più distorto e degradato.
 - Selezione del modo musicale (Modal Envelope) della texture: permette di mappare linearmente i diversi modi in relazione al parametro di Valence.
- Arousal:
- Tatum: parametro che influenza la velocità di esecuzione della texture, un valore di Arousal alto determinerà un'esecuzione più veloce.
 - Lunghezza e forma degli involuppi delle sorgenti sonore per renderli coerenti con il valore di tatum e per ottenere involuppi più percussivi per valori di Arousal alti.
 - Detune delle sorgenti sonore: per valori alti di Arousal vengono leggermente stonati gli oscillatori introducendo battimenti in grado di rendere il suono più ricco e movimentato.