

Ambienti integrati per il processing audio-video

Progetto

Walking v 1.0

**Bosi Mathieu
matr. 2480**

1. Introduzione

L'intento di questo progetto è stato quello di creare un'installazione interattiva audio-visuale che consentisse di influire sui parametri di un'animazione in tempo reale attraverso l'analisi spettrale di un segnale captato nell'ambiente riconoscendone alcune caratteristiche descrittive.

In particolare lo scopo è stato quello di riuscire a discriminare fra paesaggi sonori *lo-fi*, neutri e *hi-fi* per controllare parametricamente una scena animata evocativa rispettivamente di stati d'animo quali tristezza-pesantezza, tranquillità-routine e gioia-spensieratezza.

La scena animata è composta da elementi sempre presenti (sole, nuvole, omino, terreno, cielo) e da altri che cambiano (fiori e farfalle, casette, grattacieli e nebbia-fumo). Alcuni degli elementi sempre presenti cambiano colore (omino, cielo, terreno) o comportamento (omino).


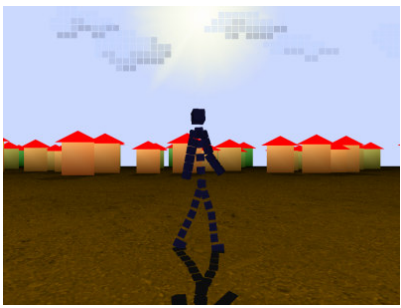

Paesaggio sonoro rilevato	Stato d'animo	Screenshot animazione	Descrizione scena
<i>lo-fi</i>	tristezza pesantezza		Omino curvo e di colore rosso scuro (sforzo-depressione) che cammina chinato, grattacieli grigi sullo sfondo, cielo e atmosfera offuscata da nebbia e fumo, terreno grigio (asfalto) colori in genere cupi e spenti.
neutro (silenzio)	tranquillità routine		Omino di colore blu scuro che cammina normalmente, casette sullo sfondo, cielo azzurro chiaro e terreno marrone (terra), colori "naturalisti".
<i>hi-fi</i>	gioia spensieratezza		Omino saltellante di colore verde chiaro e circondato da farfalle che gli volano intorno, sullo sfondo fiori rossi animati, cielo giallo caldo e terreno di colore verde (erba), colori primari e sgargianti.

Tabella 1: rapporto fra qualità del paesaggio sonoro e stati d'animo / animazioni

L'applicazione principale che si occupa della resa visiva è stata programmata nel linguaggio C++ facendo uso di una serie di librerie ausiliarie. Da questo lavoro è nato un *framework* organico che può essere riutilizzato per creare altre applicazioni audio-video interattive ad alte prestazioni.

2. Funzionamento

Il funzionamento dell'installazione è basato sul seguente processo:

- 1) Acquisizione del segnale audio ambientale
- 2) Analisi / caratterizzazione dello spettro (MAX/MSP) e trasmissione dei valori all'applicazione di *rendering* video
- 3) Traduzione delle caratteristiche estratte in un parametro "felicità" (*happiness*) compreso fra -1 e 1
- 4) Controllo a partire dalla *happiness* dei parametri dell'animazione (geometrie, curve cinematiche, colori...)

2.1 Acquisizione del segnale audio ambientale

L'acquisizione del segnale audio ambientale può avvenire utilizzando un microfono collegato ad una periferica di acquisizione audio. La soluzione ottimale sarebbe quella di disporre di una scheda audio esterna in grado di essere collegata ad un microfono direzionale che sia puntato verso l'area in cui il visitatore può interagire con l'opera.

2.2 Analisi e caratterizzazione dello spettro (MAX/MSP)

Attualmente l'analisi dello spettro viene eseguita da una apposita patch programmata nel linguaggio MAX/MSP. Questa patch riceve il segnale audio proveniente dalla periferica di acquisizione audio e ne estrae determinate caratteristiche. La patch fa uso di alcuni oggetti esterni per ricavare i parametri dello spettro [1]. Attualmente vengono estratte le informazioni di *brightness* dello spettro e l'intensità delle basse frequenze utilizzate relativamente per determinare il tipo di paesaggio sonoro secondo la relazione:

- *brightness* = *hi-fi*
- silenzio = neutrale
- basse frequenze = *lo-fi*

Una futura versione di *Walking* mirerà ad incorporare le funzioni di analisi dello spettro al proprio interno ed effettuerà un'analisi maggiormente organica del dato audio di ingresso. L'interfacciamento all'ingresso audio e l'analisi frequenziale tramite FFT sono attualmente già implementati nell'applicazione di *rendering* video attraverso l'uso di un'apposita libreria software [2].

2.2.1 Comunicazione fra MAX/MSP e l'applicazione di *rendering* video

La comunicazione fra la patch MAX/MSP e l'applicazione di *rendering* video avviene tramite l'utilizzo da parte MAX di un'apposita libreria di oggetti [3] progettati per la comunicazione attraverso il protocollo *Open Sound Control* [4].

Dal lato dell'applicazione di *rendering* è stata utilizzata un'ulteriore libreria software [5] che consente di ricevere i parametri inviati dalla patch.

2.3 Calcolo del parametro *happiness*

Una volta ricevuti i parametri riguardanti la *brightness* spettrale ed il contenuto in basse frequenze questi vengono normalizzati nel range [0..1] e combinati nella seguente maniera:

```

k = 0.995
felicità[n] = felicità[n-1] * k + IN_brightness * (1-k)
tristezza[n] = tristezza[n-1] * k + IN_lowfreq * (1-k)
happiness[n] = CLIP( felicità[n] - tristezza[n], -1, 1 )

```

per ogni n -esimo fotogramma dell'animazione. Le funzioni di sopra sono studiate per ottenere una certa inerzia nel cambiamento del valore finale di *happiness* e per variare fra un massimo ed un minimo relativamente lontani dal range $[-1, 1]$ in modo da ottenere una permanenza prolungata negli stati di felicità-tristezza attraverso un *clipping* fra i valori 1 e -1 della somma risultante.

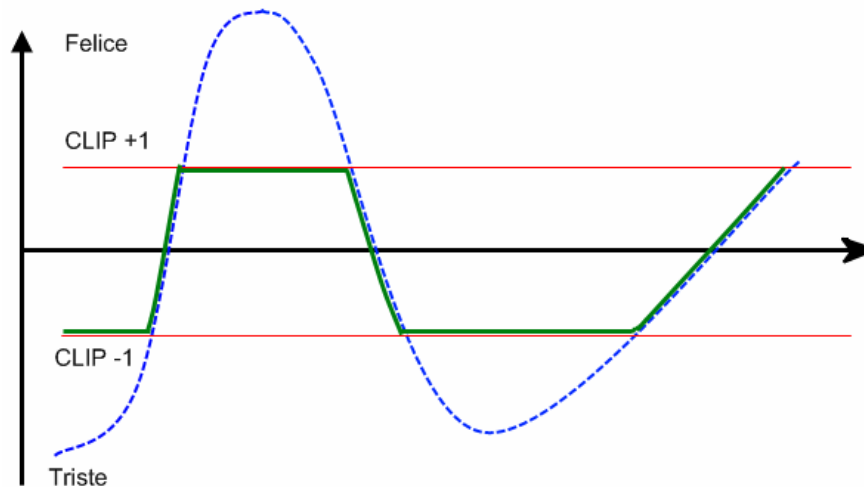


Figura 1: parametro *happiness*; la curva verde rappresenta l'effettivo valore calcolato a partire dai valori di ingresso descriventi il tipo di paesaggio sonoro

2.4 Controllo parametrico dell'animazione

A partire dal parametro *happiness* calcolato l'applicazione di *rendering* video si occupa di effettuare un *morphing* graduale fra i tre stati principali dell'animazione. Per poter ottenere questo risultato ogni elemento costituente la scena riguardo colori, geometrie e movimenti è stato parametrizzato.

Per effettuare il *morphing* si è proceduto quindi ad effettuare una interpolazione lineare fra tutti i parametri di ognuno dei tre stati fondamentali utilizzando il valore di *happiness* come parametro generale di controllo.

Ogni geometria della scena è calcolata in maniera completamente sintetica ed è riprodotta per mezzo di un'apposita libreria software [6] che consente l'interfacciamento con le funzionalità del sistema operativo e con quelle della scheda video utilizzando le librerie *OpenGL* [7].

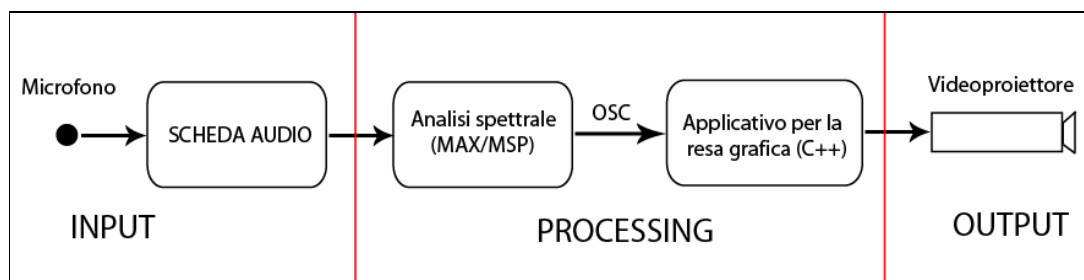
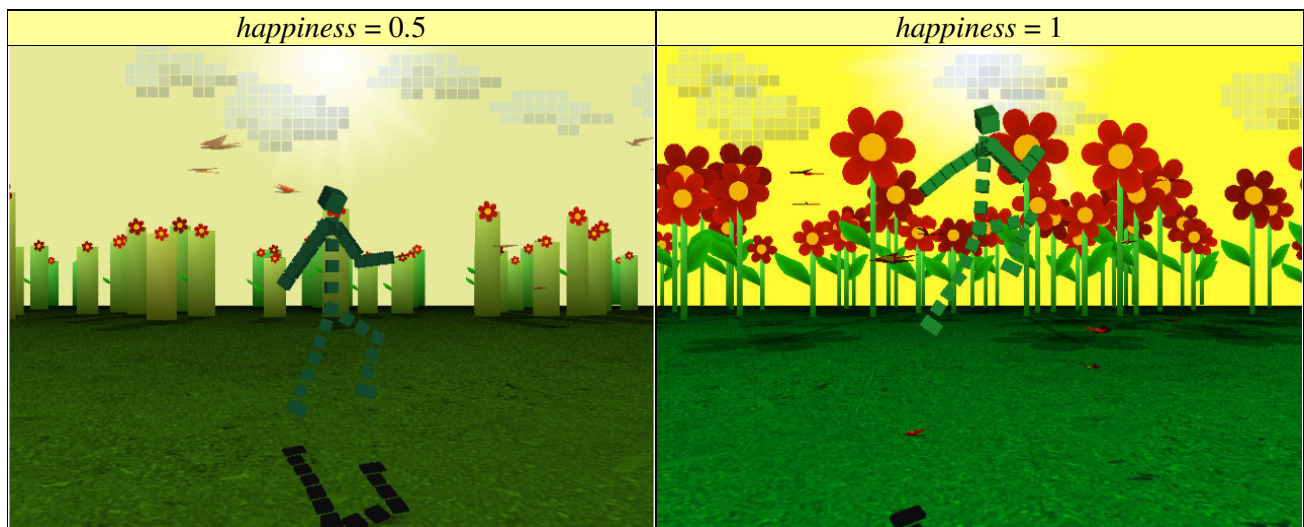
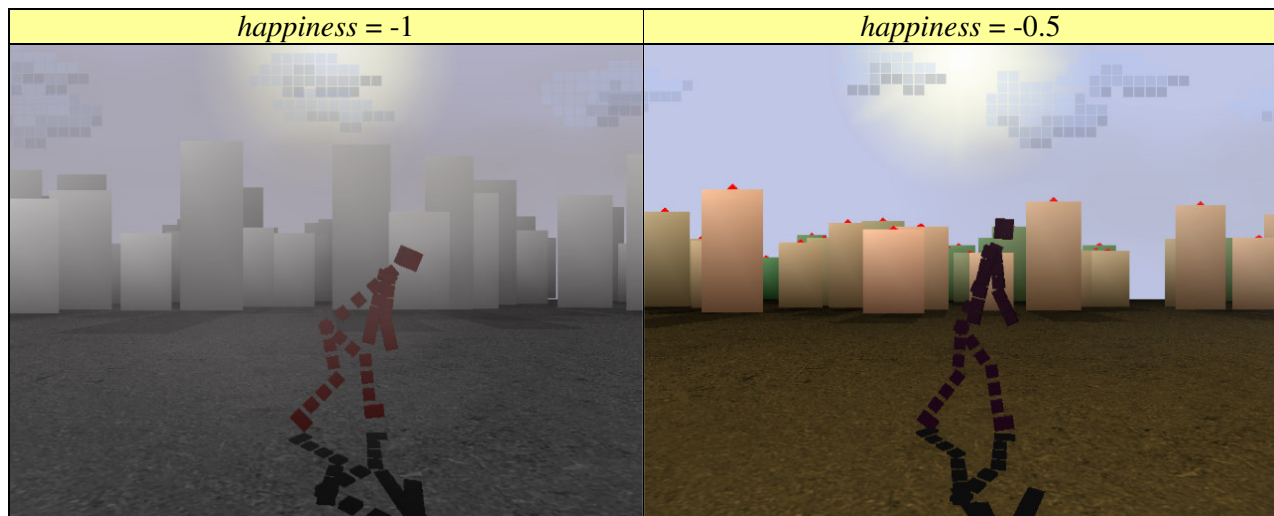


Figura 2: pipeline generale del sistema

Parametro *happiness* - *morphing* fra le diverse scene:



Bosi Mathieu



6

3. Organizzazione dell'installazione

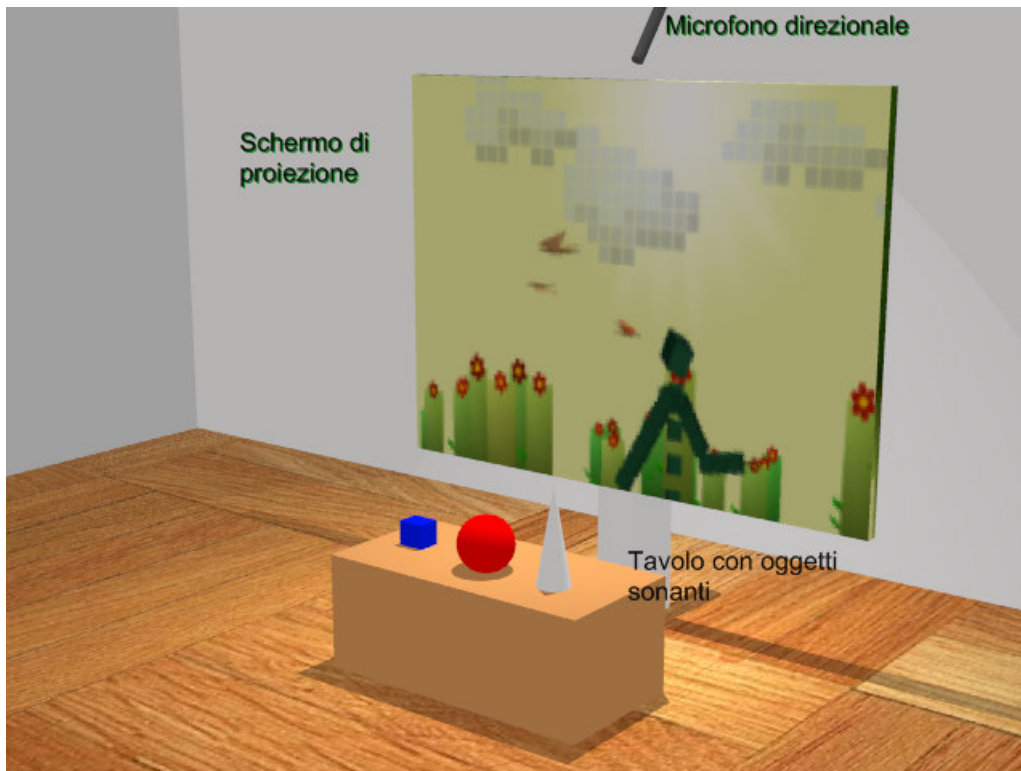


Figura 4: esempio di disposizione dell'installazione *Walking*

3.1 Setup dell'applicazione

Attualmente per funzionare *Walking* necessita di un PC dotato di sistema operativo *Microsoft Windows* e di una scheda video che supporti almeno lo standard *OpenGL 1.1*. Come già sopra accennato al fine di un corretto funzionamento dell'analisi audio è auspicabile disporre di una scheda audio e di un microfono di buona qualità. Per avviare l'applicazione si deve:

- entrare nella cartella dell'applicazione, attivare la patch MAX/MSP regolando eventualmente il guadagno di ingresso
- lanciare il file eseguibile

Per uscire dall'applicazione è sufficiente premere il tasto *ESC*.

3.2 Possibili modalità di interazione

Riguardo alla modalità di interazione è possibile per il visitatore sia produrre suoni con appositi oggetti sonanti messi a disposizione che riprodurre un tipo di paesaggio sonoro, sia consentire allo stesso di utilizzare la propria voce per imitare i diversi tipi di suono. Gli oggetti sonanti possono ad esempio essere:

Panorama sonoro *hi-fi*:

- un carillon azionabile dal visitatore (suoni nitidi armonici percussivi ritmici)
- strumenti percussivi come un sonaglio o un triangolo (suoni brillanti e leggeri)

Panorama sonoro *lo-fi*:

- un trapano (suono monotono e fastidioso)

- un piccolo motore a scoppio

Un'opzione per un'installazione di più semplice realizzazione potrebbe prevedere un sistema che proponga al visitatore una serie di bottoni i quali attivino dei suoni preregistrati appartenenti alle categorie di *soundscape* riconosciute dall'applicazione, ad esempio un cinguettio di uccellini o lo scrosciare delle onde del mare per il panorama *hi-fi* e una strada trafficata o i rumori di un cantiere per il panorama *lo-fi*. In generale è comunque importante che i suoni possano essere riconducibili ad esperienze quotidiane.

4. Considerazioni teoriche sul rapporto opera-visitatore

L'opera rappresenta un esempio di immagine *attratta* dal suono in cui la scena visuale non è costante ma bensì fissata dai suoni di ambiente. Questo che spinge il visitatore ad esplorare alcune caratteristiche del suono per scoprire come il sistema risponde agli input. Egli ricava informazioni a partire dalla rappresentazione di stati d'animo, come l'umore triste o felice convogliati dall'animazione. In questa ottica l'opera, pur rimanendo in un contesto di intrattenimento, diviene uno spunto di riflessione sull'influenza che il panorama sonoro ha sulla qualità della nostra vita quotidiana.

L'opera mette in comunicazione il carattere intimo/interiore dell'udito con il carattere esteriore della visione.

Riferimenti web alle risorse utilizzate per lo sviluppo di *Walking*

- [1] Tristan Jehan's MAX/MSP externals: <http://web.media.mit.edu/~tristan/maxmsp.html>
- [2] *FmodEx* API - Audio I/O and processing: <http://www.fmod.org/>
- [3] Emmanuel Flety and Rémy Muller's *Open Sound Control* externals for MAX/MSP:
<http://recherche.ircam.fr/equipes/temps-reel/movement/muller/static.php?page=opensoundcontrol>
- [4] *Open Sound Control* (OSC): <http://www.cnmat.berkeley.edu/OpenSoundControl/>
- [5] Ross Bencina's *OSCPack* API: <http://www.audiomulch.com/~rossb/code/oscpack/>
- [6] *Simple DirectMedia Layer* API: <http://www.libsdl.org/>
- [7] *OpenGL* API: <http://www.opengl.org/>