

UNIVERSITÁ DI PADOVA



FACOLTÁ DI INGEGNERIA

TESI DI LAUREA

Gestione di database eterogenei di HRTF per il rendering in
tempo reale di audio 3D

Management of heterogeneous HRTF databases for real time
3D audio rendering

Laureando: Boscolo Bariga Gianluca

Relatore: Ch.mo Prof. Avanzini Federico

Correlatore: Ing. Geronazzo Michele

Corso di Laurea Triennale in Ingegneria Informatica

Data Laurea: 27 novembre 2012

Anno Accademico 2012/2013

Indice

Introduzione	11
1 Suono nello spazio	13
Introduzione	13
1.1 Apparato uditivo	13
1.2 Virtualizzazione del suono	14
1.3 Elementi di Psicoacustica	16
1.3.1 Proprietà dell'azimuth	18
1.3.2 Proprietà dell'elevazione	19
1.3.3 Proprietà della distanza	20
2 HRTF e HRTF database	23
Introduzione	23
2.1 HRTF – Head Related Transfer Function	23
2.2 Database di HRTF	25
2.2.1 Il soggetto KEMAR	25
2.2.2 Sessione generica di registrazione di HRTF	26
2.3 CIPIC database	27
2.3.1 Misurazione delle HRTF CIPIC	27
2.4 Database LISTEN	28
2.4.1 Misurazione delle HRTF LISTEN	28
2.5 Database MARL	28
2.5.1 Funzioni aggiuntive di MARL	29
2.6 Compensazione delle HRTF	29
3 Pure Data	31
Introduzione	31
3.1 Introduzione a Pure Data	31
3.1.1 Utilizzo di Pure Data	32
3.2 Elementi di Programmazione in Pure Data	33
3.2.1 Oggetti	33
3.2.2 Messaggi	34
3.3 Oggetti e messaggi utilizzati	34
3.3.1 Subpatch	35
3.3.2 Blocchi send / receive	35

3.3.3	Expr if	35
3.3.4	Metro	36
3.3.5	Select	37
3.3.6	Set	37
4	Ambiente di lavoro, cw_binaural	39
	Introduzione	39
4.1	Cw_binaural	39
4.1.1	Interpolazione di HRTF	39
4.1.2	Implementazione di cw_binaural	40
4.1.3	Funzionalità di cw_binaural	41
4.1.3.1	HRTF database supportati	42
4.1.3.2	Argomenti di creazione di un oggetto cw_binaural	43
5	Creazione di una patch Pure Data per l'utilizzo di cw_binaural	45
	Introduzione	45
5.1	Premesse e scopi	45
5.1.1	Struttura del file system del progetto	47
5.1.2	Struttura del database MARL	47
5.2	Finestra principale	50
5.3	Core patch cwBinauralDemo	52
5.4	Importazione del database MARL	54
6	Considerazioni e sviluppi futuri	57
6.1	Considerazioni sull'implementazione	57
6.2	Utilizzo futuro di cw_binaural	57
	Bibliografia	59
	Ringraziamenti	61

Elenco delle figure

1.1	Apparato uditivo	14
1.2	Virtualizzazione del suono	15
1.3	Sistemi di coordinate	17
1.4	Cono di confusione	18
1.5	Variazione della risposta in base all'elevazione	19
2.1	Differenze tra HRTF di 3 soggetti per una stessa sorgente sonora	24
2.2	Il soggetto KEMAR durante una registrazione per CIPIC	25
2.3	Posizioni della sorgente in CIPIC	27
2.4	Confronto tra HRTF RAW e COMPENSATED	30
3.1	Pure Data	31
3.2	Main Window e Patch Window di Pure Data	32
3.3	Elementi di Pure Data	33
3.4	Messaggio	34
3.5	Subpatch	35
3.6	Blocchi send e receive	35
3.7	Esempio di blocco expr if	36
3.8	Esempio di utilizzo dell'oggetto Metro	36
3.9	Esempio di utilizzo dell'oggetto Select	37
3.10	Esempio di utilizzo del messaggio Set	37
4.1	Elaborazione del suono	40
4.2	Semplice utilizzo di cw_binaural con rumore bianco come sorgente sonora	41
4.3	Caricamento di HRTF LISTEN e CIPIC	42
4.4	Caricamento di un soggetto LISTEN tramite "set_hrtf_db"	43
5.1	Struttura del progetto su SVN	46
5.2	Struttura del database MARL	48
5.3	Struttura di una HRTF MARL	48
5.4	Interfaccia della demo per cw_binaural	50
5.5	Sub-patch per la selezione guidata di soggetti CI	51

5.6	Forma finale degli inlet a cw_binaural	52
5.7	Segnalazione dell'errore nel caricamento di un soggetto	53
5.8	Porzione di blocco per la selezione manuale/guidata dei soggetti LISTEN	53
5.9	File .wav di HRTF per un soggetto MARL	54
5.10	Messaggio per il caricamento di HRTF MARL	55

*“L’ingegneria non è l’arte di costruire. E’ invece l’arte di non costruire:
cioè, è l’arte di fare bene con un dollaro quello che un qualunque
pasticcione potrebbe fare con due.”*

Arthur Mellen Wellington

Introduzione

La creazione e gestione di nuovi modelli per rendere disponibile agli utenti un'esperienza di audio tridimensionale immersivo e coinvolgente è una sfida ambiziosa da affrontare ma quanto mai stimolante. Il recente aumento d'interesse verso gli scenari di realtà aumentata richiede particolare attenzione per offrire, accanto ad una visione 3D ad alta risoluzione, un audio quanto più prossimo alla vita reale.

Questo lavoro di tesi si apre con l'analisi delle caratteristiche principali dell'apparato uditivo dell'uomo, introducendo poi le principali nozioni di psicoacustica, audio virtualizzato e dei parametri necessari per costruire un modello di rendering audio che abbia le caratteristiche fondamentali. Tale sensazione d'immersione in un ambiente virtuale può essere ottenuta grazie ad una particolare tecnica denominata audio binaurale. Questa tecnica tiene conto delle varie trasformazioni che modificano l'onda sonora nel tragitto da sorgente a timpano del soggetto ascoltatore ad opera di spalle, torso e padiglione auricolare. Tutti questi fattori contribuiscono a formare la risposta impulsiva HRIR *Head Related Impulse Response* e la sua trasformata di Fourier HRTF, *Head Related Transfer Function*, esposte nel capitolo 2.

Ottenere le risposte individualizzate, fortemente dipendenti dal soggetto in ascolto, richiede apparati di acquisizione molto costosi. Nel tempo sono stati costituiti dei database di HRTF, tre di questi utilizzati in questo lavoro e descritti nelle loro caratteristiche più importanti:

- CIPIC;
- LISTEN;
- MARL.

Le operazioni di elaborazione del segnale sono state effettuate con un software specializzato in audio processing, Pure Data (capitolo 3), e con l'ausilio di un motore audio binaurale, cw-binaural (capitolo 4).

Il capitolo 5 espone come questi due strumenti abbiano portato alla creazione di una patch che, facendo uso di HRTF di tre diversi database, si propone di fungere da ausilio per esperimenti di audio spazializzato con un controllo della posizione della sorgente sonora desiderata.

CAPITOLO 1: Suono nello spazio

1. Introduzione

In questo primo capitolo vengono fornite le basi necessarie per la comprensione degli argomenti trattati in questo lavoro, a partire da una descrizione dell'apparato uditivo umano per poi descrivere la tecnica di virtualizzazione dell'audio e procedere con l'esposizione dei maggiori parametri di cui fa uso.

1.1 Apparato uditivo

Tra i cinque sensi con i quali l'uomo si rapporta con il mondo esterno, quello di interesse per lo studio e lo sviluppo di questo lavoro è l'udito. Esso permette di udire i suoni, ovvero un insieme di onde sonore liberate nell'aria da una sorgente.

L'organo fondamentale per la percezione del suono è l'orecchio: esso agisce da trasduttore nel trasformare l'energia acustica prima in energia meccanica e, successivamente, in energia elettrica inviata, tramite le terminazioni nervose, al cervello per l'elaborazione.

L'orecchio può essere suddiviso in tre parti. La pinna, ovvero il padiglione auricolare, offre una vasta superficie al fronte sonoro e permette di raccoglierne un'ampia porzione; il suono viene riflesso dal padiglione auricolare e concentrato verso il condotto uditivo; essa è fondamentale per la localizzazione del suono nello spazio, apportando importanti informazioni circa l'elevazione della sorgente sonora. Il condotto uditivo termina sull'orecchio medio: esso è composto dal timpano, che vibra in accordo con il suono che raggiunge l'orecchio. Al timpano sono collegati i tre ossicini più piccoli del corpo umano: martello, incudine e staffa. Il martello è collegato con il timpano, posto alla fine del condotto uditivo.

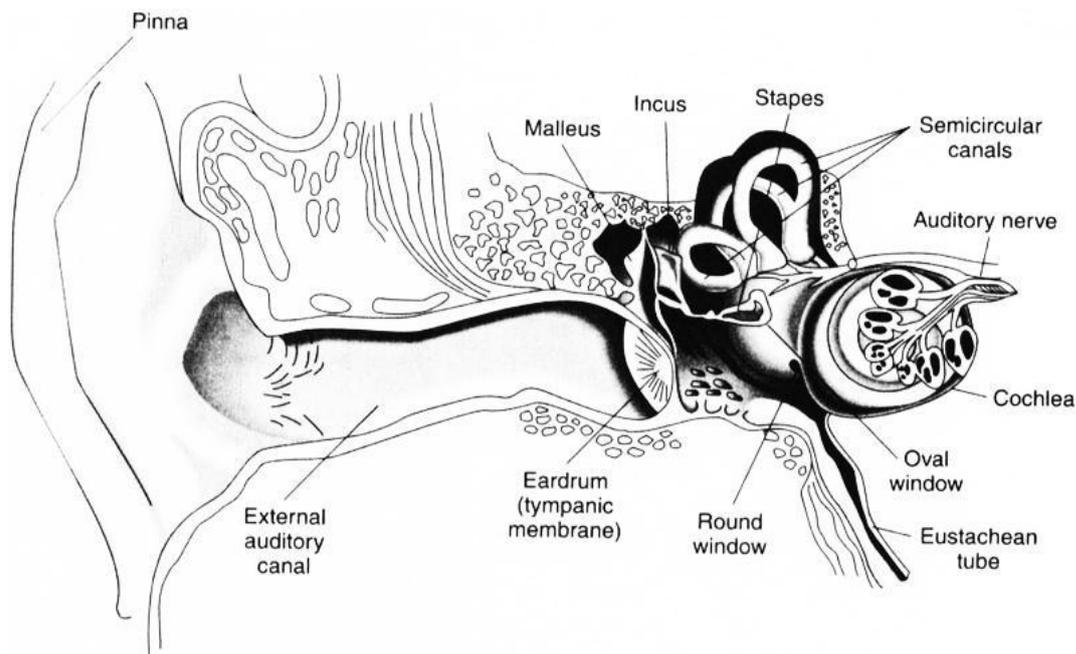


Figura 1.1: Apparato uditivo

Il movimento della membrana provoca, attraverso un complesso gioco di leve, l'amplificazione della vibrazione del timpano per la trasmissione alla coclea, ovvero al primo organo di cui è composto l'orecchio interno. Compito di questa ultima sezione è la conversione dell'energia meccanica in impulsi elettrici da inviare al cervello. La coclea è un osso a forma di chiocciola contenente del fluido che riceve la vibrazione dalla staffa e la trasporta al suo interno dove è presente il vero organo deputato alla conversione dell'energia meccanica in elettrica: l'organo del Corti. All'interno dell'organo del Corti si trova la membrana basilare che ospita una popolazione di cellule acustiche cigliate, circa 4000, che vibrano in accordo con la vibrazione del fluido. Ogni gruppo di ciglia è collegato ad una terminazione nervosa in grado di convertire la vibrazione ricevuta dal fluido in impulsi elettrici percepiti dal cervello come suoni.

Naturalmente, una singola frequenza non andrà ad eccitare una ciglia singola, ma ne ecciterà un gruppo. L'estensione delle ciglia eccitate dalla singola frequenza viene denominata banda critica ed è alla base di molti fenomeni di psicoacustica.

1.2 Virtualizzazione del suono

L'ambiente nel quale un suono viene riprodotto modifica ciò che percepiamo attraverso un fenomeno detto riverberazione. Esso è legato alla riflessione del suono da parte di ostacoli posti nell'ambiente in cui viene riprodotto e causa la propagazione di versioni ritardate o attenuate del suono stesso.

Gli elementi che modificano il suono emesso dalla sorgente da quello percepito sono, tra gli altri:

- il busto e le spalle;
- la testa del soggetto;
- la pinna dell'orecchio, ovvero la parte esterna.

Questi elementi sono, con differenze più o meno evidenti, diversi da una persona all'altra quindi un suono proveniente da una singola sorgente può essere udito in versioni diverse da soggetti distinti. Busto, testa e ambiente contribuiscono nella modifica del suono e una particolare tecnica di riproduzione audio, l'audio virtualizzato, ha lo scopo di presentare ai due timpani delle orecchie gli stessi segnali che produrrebbe un suono reale posizionato in quel punto nello spazio.

La virtualizzazione dell'audio viene ottenuta operando sul suono tramite i contributi dati da busto, testa e ambiente [1] con una tecnica presentata nel capitolo 2 e riassunta nella Figura 1.2.

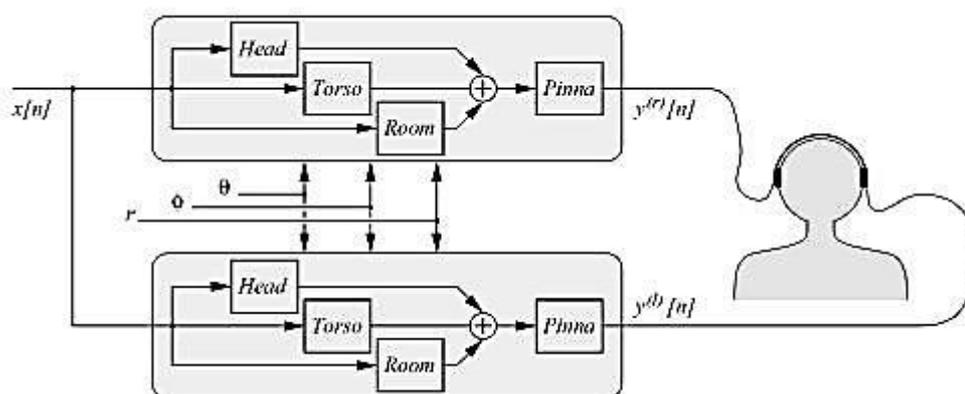


Figura 1.2: Virtualizzazione del suono

Con riferimento alla Figura 1.2:

- $x(n)$ rappresenta il suono diffuso dalla sorgente sonora;
- head, torso e room sono i contributi dati da testa, torso e stanza all'orecchio destro e sinistro;
- $y^{(r)}(n)$ e $y^{(l)}(n)$ sono il risultato dell'elaborazione che restituisce l'audio virtualizzato, riprodotto rispettivamente nell'orecchio destro e sinistro del soggetto;
- ϕ , θ ed r , ovvero azimuth, elevazione e distanza, sono parametri il cui significato sarà esposto nel paragrafo 1.3.1, 1.3.2 ed 1.3.3.

Esistono varie tecniche per ottenere l'audio virtualizzato: ad esempio la spazializzazione fisica consiste nella riproduzione del suono attraverso un sistema multi canale, con una qualche codifica per ciascun canale (quale può essere, tra le altre, la tecnica Dolby Surround)[2]. Normalmente il surround è implementato mediante audio multicanale nel quale uno o più canali audio sono destinati ad essere riprodotti da diffusori acustici posizionati, secondo prestabilite regole, alle spalle, di fronte e a lato dell'ascoltatore. Tali canali audio sono chiamati canali surround.

Questa tecnica, per rendere al meglio, deve essere utilizzata in ambienti neutri, ovvero non riverberanti e senza colorazioni, e consente una virtualizzazione del suono ottima solo nel punto predisposto per l'ascolto, producendo risultati distorti o comunque non ottimi non appena l'ascoltatore si sposta dal punto richiesto.

Un altro approccio prevede l'utilizzo di auricolari, andando a simulare tutti gli elementi che il suono naturale incontra durante il suo tragitto dalla sorgente ai nostri timpani, ovvero la forma della stanza, il corpo dell'ascoltatore e le sue parti che interagiscono con l'onda sonora (quali le pinne e il torso). Le cuffie, tuttavia, tendono a conferire alla sorgente una vicinanza eccessiva e sono afflitte da un problema detto non esternalizzazione: l'ascoltatore ha l'impressione che la fonte sonora sia posizionata all'interno della sua testa. Per risolvere questi problemi, delle basi di psicoacustica sono necessarie per comprendere la nostra abilità di localizzare sorgenti sonore e per la progettazione di sistemi di sintesi 3D efficienti.

1.3 Elementi di Psicoacustica

Con “musica” o “suono” “si fa generalmente riferimento a un complesso di processi che vanno dalla generazione di pattern sonori fino alla percezione e alla elaborazione del messaggio musicale da parte di un ascoltatore” [3]. Un suono è udito dal nostro sistema uditivo quando, come detto sopra, un'onda di pressione avente caratteristiche fisiche ben definite, va ad eccitare il timpano. Le tre sensazioni principali che accompagnano l'ascolto di un suono sono:

- altezza (pitch);
- intensità (loudness): dipende dal flusso di energia che accompagna la vibrazione ma varia anche in base alla durata del suono o alla presenza di altre sorgenti sonore;
- timbro (timbre): dipende dallo spettro di energia, la cui evoluzione temporale è un elemento fondamentale per il riconoscimento e la caratterizzazione dei suoni strumentali.

Nonostante rimangano molti aspetti sconosciuti, le caratteristiche principali del processo uditivo sono noti da molto tempo ed accurati studi psicologici hanno

stabilito quanto accuratamente possiamo dare giudizi sulla localizzazione di una sorgente sonora.

Per specificare la localizzazione di una sorgente sonora relativamente all'ascoltatore, si necessita di un sistema di coordinate. Una scelta naturale può essere quella di un sistema di riferimento centrato sulla testa del soggetto: l'asse x passante, approssimativamente, attraverso il suo orecchio destro, l'asse y diretto davanti all'ascoltatore e l'asse z che punta verso l'alto. Questi assi definiscono tre piani standard: il piano orizzontale xy , il piano frontale xz e il piano mediano yz . Chiaramente, il piano orizzontale definisce una separazione sopra - sotto, il piano frontale quella davanti - dietro mentre il piano mediano definisce una separazione destra - sinistra.

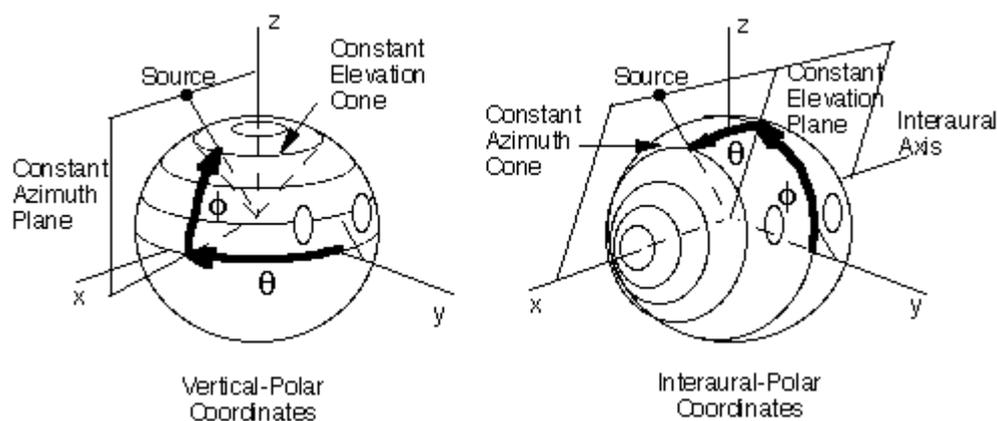


Figura 1.3: Sistemi di coordinate

Tuttavia, essendo il capo di forma sferica, solitamente si usa un sistema di riferimento sferico e qui le coordinate standard sono azimuth, elevazione e range.

Essendo possibili più modi di definire queste coordinate, esistono due sistemi di riferimento equivalenti per tali coordinate, ovvero le coordinate polari verticali e le coordinate polari interaurali.

Nel sistema a coordinate polari si misura dapprima l'angolo θ di azimuth come l'angolo tra il piano mediano e un piano verticale contenente l'asse z e, poi, si misura l'angolo ϕ di elevazione come l'angolo con il piano orizzontale. Con questa scelta, superfici con azimuth costante sono piani che intersecano l'asse z mentre superfici con costante elevazione sono coni concentrici con l'asse z .

Utilizzando invece un sistema di coordinate interaurali, si misura dapprima l'angolo di elevazione come l'angolo tra il piano orizzontale e un piano contenente la sorgente e l'asse x , che è l'asse interaurale; l'azimuth è misurato come l'angolo

con il piano mediano. Con questa scelta, le superfici a elevazione costante sono piani passanti per l'asse interaurale e le superfici con azimuth costante sono coni concentrici con l'asse interaurale.

1.3.1 Proprietà dell'azimuth

Uno dei più importanti pionieri dell'audio spazializzato fu il barone John Strutt Rayleigh. Quasi 100 anni fa sviluppò la Duplex Theory, secondo la quale l'azimuth è individuato tramite due parametri: Interaural Time Difference e Interaural Intensity Difference.

ITD (Interaural time difference) è definita come la differenza dei tempi di arrivo del fronte d'onda tra orecchio sinistro e destro, mentre IID (Interaural intensity difference) misura similmente la differenza di ampiezza tra i suoni uditi tra i due orecchi. In generale, un suono è percepito vicino all'orecchio al quale arriva per primo il fronte d'onda sonoro, ovvero ad un ampio ITD corrisponde una ampia traslazione della sorgente sonora in una posizione laterale rispetto all'ascoltatore.

E' opinione comune che il sistema uditivo utilizza ITD, IID e delle indicazioni spettrali per determinare la posizione spaziale di una sorgente in tutte le posizioni possibili, non solo in quelle appartenenti al piano orizzontale [4]. Tuttavia, mentre gli esperimenti di psicoacustica hanno dimostrato la semplice relazione lineare che intercorre tra ITD, IID e posizione laterale percepita, la relazione tra contenuto spettrale ricevuto dall'ascoltatore e posizione nello spazio (non limitata solo sul piano orizzontale) non è così semplice.

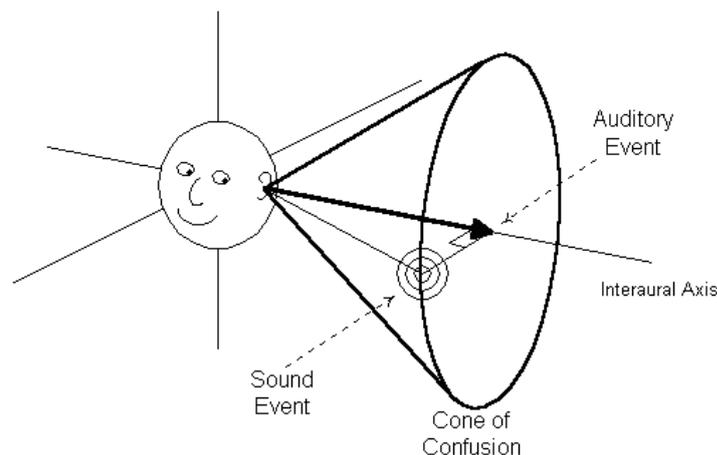


Figura 1.4: Cono di confusione

Un fenomeno interessante avviene se, utilizzando un sistema di coordinate polari interaurali e tenendo l'azimuth costante, si ottiene un valore costante di ITD. Quindi, c'è una corrispondenza tra ogni ITD e un cono di azimuth costanti, talvolta chiamato "cono di confusione". Questi coni non sono, per ragioni strutturali, individuabili con un sistema di coordinate polari verticali. In conclusione, si può notare come il solo ITD costringe la sorgente in uno dei coni di confusione e, quindi, non è sufficiente per determinare se la sorgente stessa si trovi davanti o dietro l'ascoltatore.

1.3.2 Proprietà dell'elevazione

Le nostre orecchie funzionano come delle "antenne acustiche": le sue cavità risonanti sono in grado di amplificare alcune frequenze mentre portano alcuni fenomeni di interferenza che vanno ad attenuare altre frequenze. Inoltre, la sua risposta in frequenza è direzionalmente dipendente dalla posizione della sorgente come visualizzato in Figura 1.5.

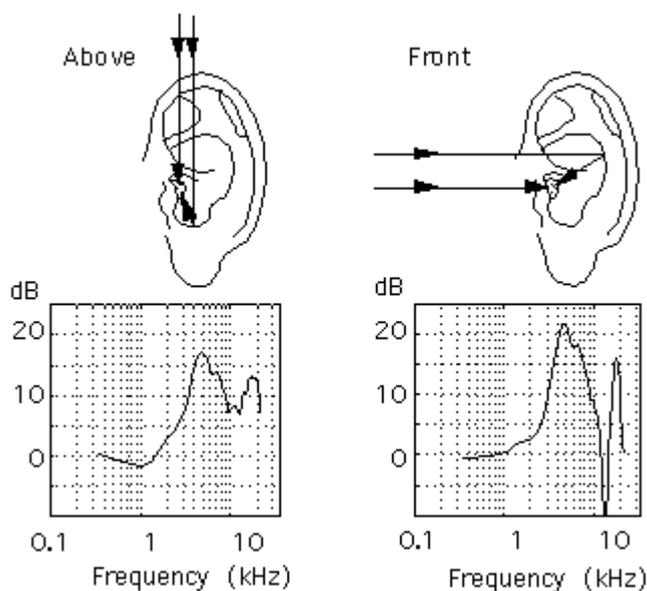


Figura 1.5: Variazione della risposta in base all'elevazione

Essendo la pinna in grado di riflettere meglio i suoni provenienti dal fronte che da sopra, l'incavo della risposta è molto più pronunciato per sorgenti poste davanti al soggetto rispetto a quelle poste superiormente. Anche se la materia è ancora oggetto di studio, si ritiene che la pinna sia la parte del sistema uditivo che fornisce più informazioni circa l'elevazione della sorgente sonora. L'importanza della pinna nell'individuazione dell'elevazione della sorgente porta, in modelli poco accurati o non specificatamente studiati sulla propria conformazione fisica, a un fenomeno detto *front-back confusion* che porta il soggetto a collocare la sorgente sonora di

fronte a sè mentre si trova dietro, e viceversa [5]. La presenza di questo fenomeno, come il cono di confusione citato in 1.3.1, è dovuto al fatto che più punti nello spazio corrispondono ad identici ITD e IID. Questa situazione si presenta tanto nelle applicazioni teoriche quanto nell'ascolto di tutti i giorni: in particolare i casi di front-back confusion sono dipendenti dal contenuto spettrale della sorgente sonora. Un espediente per individuare correttamente la posizione della sorgente anche in caso di front-back confusion è data dalle dynamic cues: infatti, dando la possibilità all'ascoltatore di muovere la testa, essi riescono a utilizzare le informazioni derivanti dalle variazioni di ITD, ILD e movimenti dei massimi e minimi presenti nel contenuto spettrale per localizzare correttamente la sorgente, prima non precisamente collocabile.[5]

1.3.3 Proprietà della distanza

In condizioni anecoiche, ovvero in un ambiente progettato per massimizzare l'abbattimento delle riflessioni del suono (anecoico deriva infatti dal greco e significa "senza eco"), la percezione della distanza di una sorgente è correlata all'intensità del segnale.

L'individuazione della distanza della sorgente sonora è più complicata rispetto all'individuazione dell'azimuth e dell'elevazione. Il parametro maggiormente utilizzato per l'individuazione di questo parametro è, come detto sopra, l'intensità con i relativi incrementi e diminuzioni: grazie all'esperienza, l'individuo impara ad associare tali variazioni alla posizione fisica della sorgente. [6]

Quando la sorgente si trova lontana dal soggetto (nel *far field* ovvero approssimativamente a più di 1,5 m dal centro della testa del soggetto) si può assumere che i segnali direzionali siano indipendenti dalla distanza. Infatti, per grandi distanze, il fronte d'onda è approssimabile ad un fronte d'onda planare e quindi non è possibile ricavare i parametri utili per una corretta collocazione.

Avvicinando gradualmente la sorgente vicino al soggetto (entrando così nel *near field*, ovvero a una distanza inferiore a 1,5m) si ha che l'ITD (Interaural Time Difference) resta quasi indipendente dalla distanza mentre l'ILD aumenta fortemente nell'intero spettro e in modo molto marcato alle basse frequenze.

Per caratterizzare il contributo della distanza nel modificare il suono presentato alle pinne dell'ascoltatore, nel lavoro condotto da Spagnol et al. del Sound and Music Computing Group, si è fatto ricorso a funzioni di trasferimento sferiche (STF) in cui si assume che l'ITD non venga modificato dalla distanza mentre si interviene sull'ILD, il tutto tramite una funzione

$$H(\mu, \theta_{inc})$$

con μ parametro che tiene conto della distanza della sorgente mentre θ_{inc} è l'angolo di incidenza. Per i dettagli di tale lavoro si faccia riferimento a [6].

CAPITOLO 2: Hrtf e Hrtf database

2. Introduzione

In questo capitolo viene esposto il concetto di HRTF ossia Head Related Transfer Function, strumento indispensabile per lo svolgimento di questo lavoro e per gli studi nei campi di audio 3D e binaurale. Segue poi una trattazione sui 3 database di HRTF utilizzati con particolare riferimento alle procedure di registrazione delle risposte e all'individuazione dei punti nello spazio in cui effettuare tali misurazioni.

2.1 HRTF – Head Related Transfer Function

Mentre è ragionevole predire la relazione tra ITD, IID e posizione nel piano, è meno intuitivo come il contenuto spettrale e la posizione spaziale possano essere collegate matematicamente. Come primo passo per comprendere questo collegamento, molti ricercatori hanno effettuato la misurazione dei contributi di pinna, torso e stanza sul suono di una sorgente situata a diverse posizioni dello spazio per poi determinare la risposta impulsiva di ciascun orecchio. Queste misurazioni sono conosciute con il nome di Head Related Impulse Response (HRIR) e riassumono la dipendenza del filtraggio di un suono con ostacoli quali la testa, il torso e la pinna. In questo modo i ricercatori calcolano la risposta impulsiva di un orecchio per poi analizzare e scoprire in che modo vengono percepiti i suoni a relative posizioni.

Formalmente, una singola HRIR è definita come una risposta impulsiva di un singolo orecchio (destro o sinistro) a un suono, misurata da uno specifico punto nello spazio. La convoluzione di una sorgente arbitraria con la HRIR specifica della posizione converte il suono emesso in quello percepito dall'utente con il suo sistema uditivo:

$$Output(t) = Input(t) * HRIR(t)$$

dove con Input si intende il suono emesso mentre con output si intende il suono percepito. Tale relazione è valida se la HRIR è misurata idealmente nella posizione del timpano e se la riproduzione avviene nella stessa posizione, ovvero con i loudspeaker della cuffia posizionati idealmente nella medesima posizione di registrazione.

Tipicamente, le HRIR sono misurate su persone o manichini per entrambe le orecchie ad angoli prefissati sull'ascoltatore. Pertanto, le HRIR sono rappresentazioni della risposta impulsiva e naturalmente rappresentate nel dominio del tempo; la rappresentazione nel dominio della frequenza è ottenuta attraverso una trasformata di Fourier ed è denominata Head Related Transfer Function (HRTF). Essendo la HRTF la trasformata di una risposta impulsiva, essa è funzione di trasferimento e pertanto tra output e input sussiste la relazione:

$$HRTF(f) = Output(f) / Input(f)$$

La complessità di un sistema di spazializzazione dipende dalla lunghezza delle HRIR e ci sono stati quindi diversi tentativi di ridurne le lunghezze preservando il contenuto e le importanti informazioni spaziali.

Le HRTF e le HRIR variano significativamente da persona a persona e gravi distorsioni nella percezione dei suoni (in modo particolare tra sorgenti posizionate davanti o dietro) si possono registrare quando un soggetto ascolta audio spazializzato facendo uso di HRTF non registrate e sintetizzate sulla propria conformazione fisica.

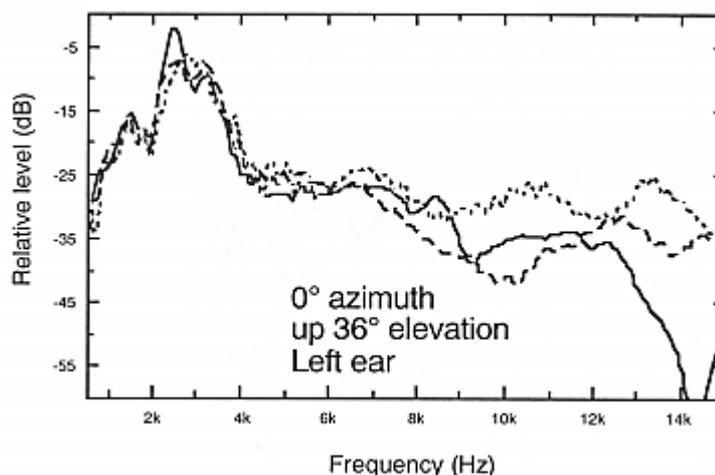


Figura 2.1: Differenze tra HRTF di tre soggetti per una stessa sorgente nella stessa posizione [7]

2.2 Database di HRTF

Storicamente, questi argomenti sono stati oggetto di studio da parte di molte università e gruppi di ricerca sparse in tutto il mondo. Un database uniforme di HRTF consente lo studio delle differenze tra diversi soggetti e la relazione tra contenuto spettrale e caratteristiche antropometriche del soggetto.

Pertanto, ogni laboratorio ha costruito il proprio database di HRTF ottenuto attraverso misurazioni, su soggetti umani o manichini, ed in generale con metodi diversi di registrazione. I database cui si fa riferimento in questo lavoro sono CIPIC, LISTEN e MARL.

2.2.1 Il soggetto KEMAR

I manichini KEMAR, utilizzati in molte occasioni durante le registrazioni di HRTF citate in questo lavoro, sono simulatori di testa e torso con antropometria basata sulle misure standard di uomini e donne. Il manichino ha lo scopo di simulare i cambiamenti che occorrono ad un'onda sonora nel momento in cui entra in contatto con il soggetto, come ad esempio la rifrazione e la diffrazione sulle pinne.

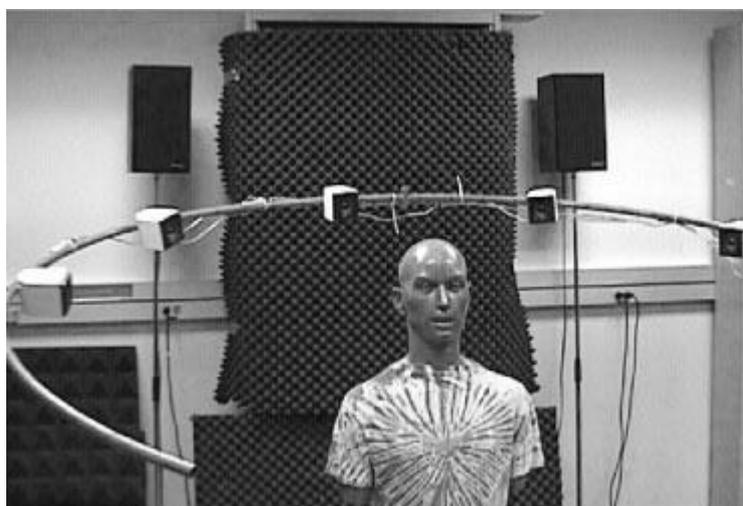


Figura 2.2: il soggetto KEMAR durante una registrazione per CIPIC

Ne esistono svariati modelli adatti allo scopo degli esperimenti: per le misurazioni CIPIC, ad esempio, sono stati utilizzati i modelli KEMAR con pinne larghe o piccole. Questi simulatori si adattano molto bene alle registrazioni di HRTF in quanto rimangono sempre immobili (cosa non perfettamente verificata per i soggetti umani che si muovono, seppur in misura molto piccola, durante le

misurazioni). Infatti, una piccola variazione nella posizione della testa del soggetto può invalidare l'intera misurazione di HRTF e, per evitare tale eventualità, durante la misurazione ci si serve di sistemi di tracciamento del soggetto e della sua posizione.

Il soggetto KEMAR si rivela anche particolarmente utile perché consente la registrazione separata delle IR di busto e testa senza orecchie (HAT, Head And Torso Simulator) o solo delle orecchie. Inoltre, la sottrazione dei vari contributi dati, ad esempio, da busto e orecchie, non sono immediate.

2.2.2 Sessione generica di registrazione di HRTF

La conoscenza delle funzioni HRTF consente di riprodurre situazioni di ascolto tridimensionali attraverso l'uso di due soli sistemi di diffusione del suono. In particolare, l'applicazione di queste funzioni fornisce risultati soddisfacenti solo se il sistema di riproduzione approssima quanto meglio la condizione di binauralità. Tale condizione viene raggiunta completamente quando la riproduzione avviene in cuffia o, in modo ancora migliore, quando il soggetto è posto in una camera anecoica, con la testa fissa in un'unica posizione e con sonde di riproduzione acustica inserite nei canali uditivi. Questa condizione ha dato luogo ad un certo sviluppo della tecnica di registrazione binaurale che consiste nella ripresa del suono attraverso l'uso di due microfoni all'interno di una testa artificiale o all'interno dei condotti uditivi del soggetto.

Tali funzioni HRTF vengono ricavate campionando lo spazio attraverso la registrazione di impulsi temporali lungo una griglia di azimuth ed elevazioni. Il suono viene ripreso dai microfoni posti nei condotti uditivi ed elaborato per fornire un numero finito di risposte impulsive. Tali risposte vengono poi utilizzate dal sistema di elaborazione (DSP - *Digital Signal Processing*) e riproduzione del suono per ricostruire un qualsiasi posizionamento attraverso la tecnica dell'interpolazione.

Indipendentemente dal fatto che le risposte vengano calcolate per un soggetto artificiale o umano, viene registrato in una camera anecoica un segnale che rappresenta, in prima approssimazione, la risposta del padiglione dell'orecchio esterno stimolato con un segnale impulsivo riprodotto da un altoparlante posto a circa un metro di distanza. La rilevazione comprende generalmente la registrazione simultanea della risposta destra e sinistra (risposta impulsiva binaurale) e viene ripetuta tante volte quante sono le posizioni relative tra sorgente e rilevatore che si desidera memorizzare.

2.3 CIPIC database

Questo database di HRTF è stato pubblicato nel 2001 dal Center for Image Processing and Integrated Computing (CIPIC) della University of California. Il CIPIC Interface Laboratory ha misurato le HRTF con una risoluzione spaziale molto ampia per più di 90 soggetti e, con la release 1.0, ha reso pubblico un sottoinsieme di 45 di queste, incluse le risposte in frequenza di due soggetti KEMAR con pinna larga o piccola [8]. Escludendo quindi i due soggetti KEMAR, i 43 soggetti umani resi pubblici (27 uomini e 16 donne) sono studenti o visitatori del CIPIC Interface Laboratory. L'ultimo aggiornamento del database risale al 2004, con aggiornamenti per rendere compatibili le funzioni con la nuova versione di Matlab.

2.3.1 Misurazione delle HRTF CIPIC

Tutte le HRTF sono state misurate sui soggetti seduti al centro di una sfera di raggio 1m e con gli assi centrati sull'asse interaurale del soggetto. La posizione della testa del soggetto non è stata forzata nella posizione corretta avendo tuttavia la possibilità di monitorare la posizione tenuta grazie ad un monitor di controllo.

La posizione della sorgente è specificata dall'angolo di azimuth θ e dall'angolo di elevazione ϕ nelle coordinate polari interaurali (vedi Figura 2.3). L'elevazione è stata uniformemente suddivisa in passi di $360/64 = 5.625^\circ$ da -45° a $+30.625^\circ$. Per ottenere una distribuzione uniforme dei punti sulla sfera, gli azimuth sono stati posti a -80° , -65° , -55° , da -45° a $+45^\circ$ in passi di $+5^\circ$, a $+55^\circ$, $+65^\circ$ e infine $+80^\circ$. Questa configurazione porta quindi a 1250 punti di posizionamento possibile della sorgente.

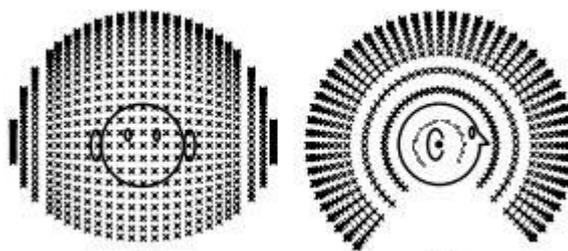


Figura 2.3: posizioni della sorgente in CIPIC

A fungere da sorgente sonora nelle varie posizioni specificate qui sopra fu un altoparlante Bose Acoustimass (con un diametro di 5-8 cm). All'entrata del canale uditivo del soggetto è stato inserito un microfono Etymotich Research ER-7C allo scopo di registrare le risposte in frequenza, digitalizzate con una frequenza di 44.1 kHz. Lo stimolo sonoro utilizzato è costituito da un segnale Golay-code. In questo modo è stata ottenuta la risposta in frequenza RAW ovvero grezza.

Tuttavia, tutte le HRTF del database CIPIC sono rese disponibili nella versione COMPENSATED, ottenuta tramite una particolare elaborazione della versione RAW, descritta nel paragrafo 2.6.

2.4 Database LISTEN

L'Institute for Research and Coordination Acoustic/Music (IRCAM), in collaborazione con AKG Acoustics, ha rilasciato nel 2003 un set di misurazioni di HRIR facenti parte del progetto di ricerca Listen. Esso è composto dalle HRIR di 51 soggetti. A differenza di CIPIC, LISTEN fa uso delle coordinate vertical - polar[9].

2.4.2 Misurazione delle HRTF LISTEN

Le misurazioni delle risposte sono state eseguite in una camera anecoica di circa 320 m³ coperta con uno strato spesso 1,1 m di lana di vetro atta ad assorbire i suoni con frequenza superiore ai 75 Hz. Il movimento della sorgente sonora fu assicurato da una struttura metallica che, ruotando e spostandosi, garantiva la totale copertura delle posizioni previste.

La posizione del soggetto era supervisionata da un controller per garantire l'integrità delle misurazioni: la conferma di corretta posizione della testa era data da un sistema di head tracking. In modo simile alle misurazioni CIPIC, un altoparlante simulava la sorgente sonora in movimento mentre il microfono, all'interno del condotto uditivo del soggetto, andava a misurare la risposta.

Per le misurazioni LISTEN si è fatto uso di un loudspeaker TANNOY system 600, giudicato un buon compromesso tra qualità e peso, e di uno stimolo costituito da una *logarithmic sine sweep*, ovvero di un segnale sinusoidale la cui frequenza aumenta di un fattore fissato per ogni unità di tempo.[10]

A differenza di CIPIC, LISTEN mette a disposizione sia le versioni RAW che COMPENSATED delle risposte, sia in formato .wav che .mat.

Le elevazioni sono state suddivise in 10 punti che partono da -45° per finire a 90°, con incrementi verticali di 15°. Le posizioni di azimuth variano da un massimo di 24 (incrementi di 15° nella posizione di 0°) ad un solo punto nella posizione di 90°. Le risposte impulsive sono dunque disponibili quindi per ciascun soggetto in 512 punti, ciascuno con il proprio valore di ITD.

2.5 Database MARL

Ogni database pubblicamente disponibile è stato misurato facendo uso di differenti standard, come ad esempio diversi incrementi di azimuth ed elevazione, differenti range angolari, sample rate, stimoli sonori utilizzati e metodi di compensazione. Per offrire una soluzione a questo problema, il Music and Audio

Research Laboratory della New York University ha presentato nel 2011 un database proveniente dall'unione dei 4 più utilizzati database di HRIR in circolazione, per un totale di 113 soggetti:

- da S001_marl-nyu a S051_mar-nyu provenienti da LISTEN;
- da S052_marl-nyu a S096_marl-nyu provenienti da CIPIC;
- da S097_marl-nyu a S111_marl-nyu provenienti da FIU;
- i soggetti S112_marl-nyu e S113_marl-nyu provenienti dal KEMAR-MIT.

Il FIU database è un set di misurazioni rilasciate dalla Florida University nel 2010 per 15 soggetti umani in 6 elevazioni e 12 punti di azimuth. Il KEMAR-MIT, invece, è costituito di due HRIR catturate nel 1995 presso il Massachusetts Institute of Technology con degli speaker posizionati ad una distanza di 1,4 m dal manichino KEMAR con 14 diverse posizioni di elevazione e azimuth varianti da 72 a 1 in base all'elevazione considerata.

2.5.1 Funzioni aggiuntive di MARL

MARL mette a disposizione degli utenti funzioni MATLAB atte a rendere il database versatile ed adattabile agli scopi tra i più vari. In questo lavoro si è utilizzata principalmente `exportAudio.m` che consente la traduzione dell'intero database dal formato `.mat` a `.wav` delle HRTF.

2.6 Compensazione delle HRTF

Le risposte RAW devono essere compensate per correggere le limitazioni dei microfoni ed altoparlanti utilizzati durante la misurazione. La correzione di base utilizzata è fatta rimuovendo il soggetto dal punto di misurazione, ponendo il microfono al centro della sfera e misurando la free field response (FFR). La risposta HRTF compensata è quindi ottenuta come rapporto tra la trasformata di Fourier della HRIR RAW e quella misurata a campo libero:

$$HRTF_{compensated}(f) = HRTF_{raw}(f) / FFR(f)$$

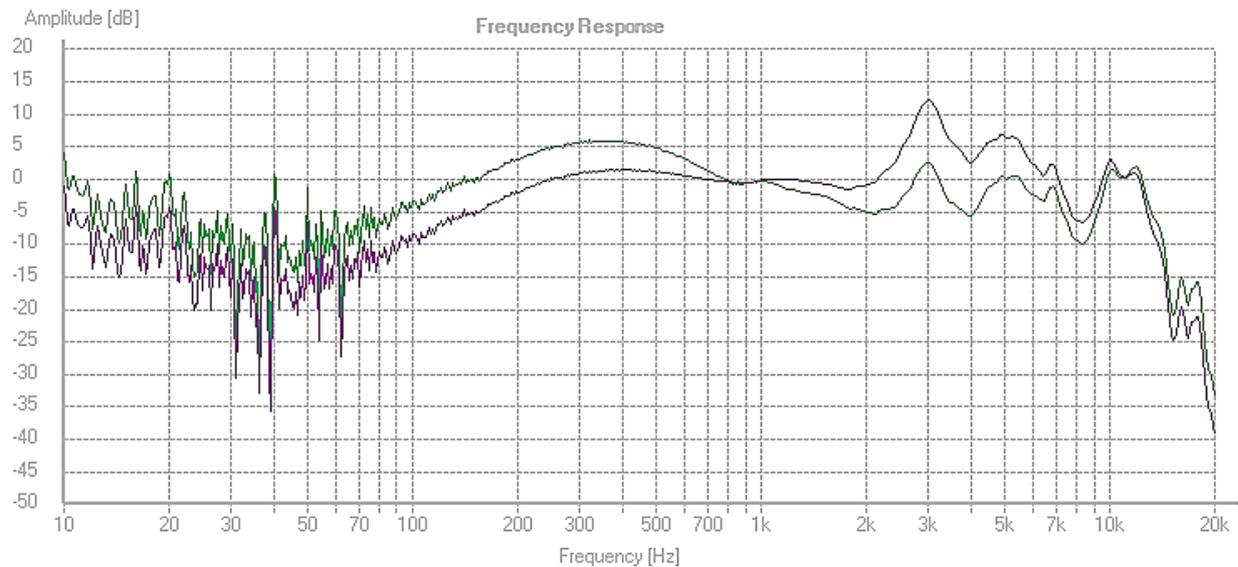


Figura 2.4: confronto tra HRTF RAW (viola) e COMPENSATED (verde) [11]

La HRIR risultante descrive quindi la pressione sonora all'entrata del condotto uditivo, comprendente tutte le trasformazioni acustiche operate dal corpo del soggetto stesso. In conclusione, queste correzioni sono indispensabili, ma tuttavia non ancora sufficienti, per riprodurre i suoni binaurali con un timbro naturale, indipendentemente dalla direzione e dalla colorazione spettrale della sorgente.

CAPITOLO 3: Pure Data

3 Introduzione

In questo capitolo viene data un'introduzione a Pure Data, il software utilizzato per il lavoro di questa tesi. Si farà particolare riferimento agli strumenti utilizzati cercando di dare una panoramica veloce ma tuttavia esauriente allo scopo di comprendere i contenuti dei capitoli successivi.

3.1 Introduzione a Pure Data

Pure Data è un ambiente di programmazione grafica in tempo reale per processare audio e video [12]. Pure Data fu sviluppato dall'ingegnere Miller Puckette negli anni '90 per la creazione di musica interattiva a computer e lavori multimediali. Mentre Puckette è l'autore principale del programma, attorno a questo software si è creata una larga base di sviluppatori che lavorano per offrire sempre nuove espansioni e possibilità. È rilasciato con una licenza simile a BSD, che ne garantisce l'assoluta libertà nella consultazione e modifica a proprio piacimento del software.



Figura 3.1: Pure Data

Pure Data è comunemente utilizzato per esecuzioni di musica dal vivo, interfacce con sensori, effetti sonori, analisi sonore approfondite, interfacciamenti con sensori, robot e siti web. Essendo tutte queste diverse tipologie di media trattate come dati digitali all'interno del programma, esistono molte opportunità incrociate di sintesi tra di essi. Il suono può, ad esempio, essere utilizzato per manipolare un video per poi essere messo a disposizione via streaming su Internet

per un altro computer che lo potrebbe utilizzare per altre operazioni di modifica o controllo.

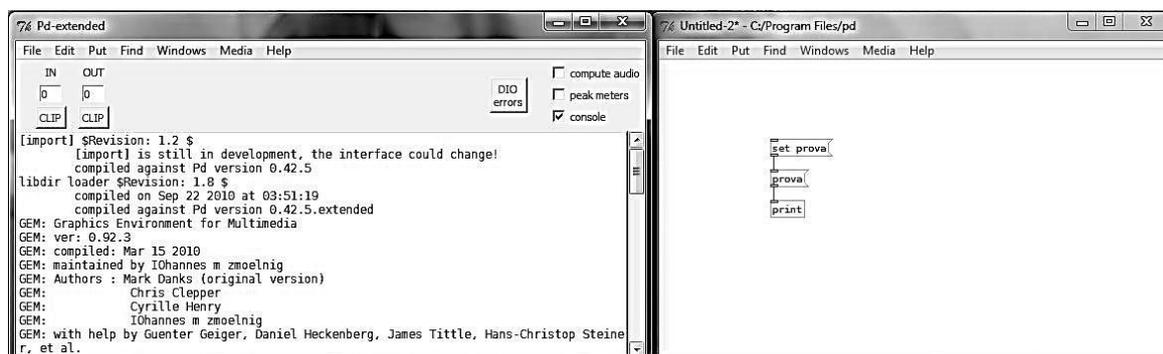


Figura 3.2: Main Window e Patch Window di Pure Data

3.1.1 Utilizzo di Pure Data

Pd è definito come un ambiente di programmazione grafico real-time per il trattamento dell'audio. Esso utilizza degli oggetti grafici che l'utente posiziona e collega sullo schermo. Non si tratta, quindi, di una programmazione basata sulla scrittura di un codice ma della creazione di un flow chart. Gli oggetti inseriti nel programma svolgono svariate funzionalità. Alcuni blocchi sono utili per la gestione dei segnali audio altri per la supervisione dei segnali di controllo. Essi, inoltre, per interagire fra di loro sono collegati attraverso delle linee. L'analogia con il mondo reale è lampante: i dispositivi fisici utilizzati prima dell'avvento del computer music, sono rappresentati dagli oggetti che si posizionano sullo schermo e i cavi sono raffigurati dai collegamenti tra l'entità virtuali. Proprio per questo Pd è definito come un linguaggio di programmazione orientato al flusso di dati (datastream-oriented programming language).

Uno degli aspetti fondamentali di Pure data risiede nell'esecuzione real-time. I cambiamenti effettuati in Pd sono istantanei; il comportamento non è uguale ai classici linguaggi di programmazione nei quali il codice deve prima essere processato (es. compilato) prima di poter essere utilizzato. Pd si comporta come uno strumento musicale classico: l'esecutore sente istantaneamente la modifica del suono, a seguito di un suo intervento. Questo rende Pd utilizzabile anche in performance live.

Pure Data è attualmente disponibile per l'installazione su Windows, Linux, Mac OS X e iOS.

3.2 Elementi di programmazione in Pure Data

All'apertura del programma, viene aperta la finestra principale di PD che ha la forma di una console di controllo. In essa vengono visualizzati vari messaggi tra i quali, all'apertura, quelli di configurazione dell'audio e di caricamento delle librerie aggiuntive. Per poter procedere alla programmazione occorre aprire una nuova finestra, la patch window, ovvero l'ambiente di programmazione vero e proprio. Essa può trovarsi in due modalità: edit mode o exec mode. Il primo consente di inserire tutti gli elementi all'interno della finestra mentre il secondo è necessario per gestire la patch quando questa è in azione.

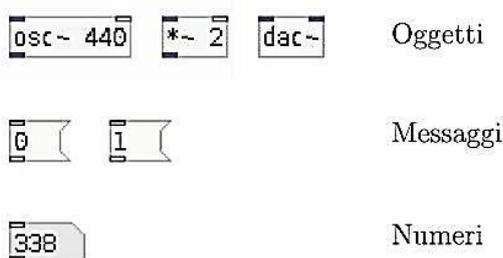


Figura 3.3: Elementi di Pure Data

La finestra di patch è il luogo che permette la realizzazione degli algoritmi di Pd. Essendo un ambiente grafico, la finestra si riempirà di entità di varia natura, dette scatole (box). Queste scatole sono di quattro tipi [13]:

- oggetti;
- messaggi;
- GUI;
- commenti.

3.2.1 Oggetti

Gli elementi fondamentali della programmazione in Pure Data sono gli oggetti, rappresentati dalle object box, caratterizzate dalla forma rettangolare e dalla presenza di entrate (inlets), nella parte superiore, e di uscite (outlets), nella parte inferiore.

Un oggetto può creare o processare dati oppure segnale audio e ricevere attraverso gli inlets messaggi, liste o uscite di altri oggetti. I collegamenti tra tutte le object box avvengono tramite cavi di connessione che si realizzano tendendo premuto il tasto sinistro del mouse a partire dall'outlet mittente e rilasciandolo non appena si è raggiunto l'inlet di arrivo.

Gli oggetti insistono su dati o su segnali audio. Graficamente, gli oggetti – audio si distinguono da quelli per i dati perchè per il primo tipo l'atomo (ovvero il nome) termina sempre con il segno tilde e gli inlet e outlet che processano segnale sono scuri. I cavi che trasportano segnale sono inoltre più spessi di quelli che trasportano dati.

3.2.2 Messaggi

Le message box hanno forma rettangolare con il lato destro rientrante. Possono contenere stringhe di caratteri, numeri, liste o variabili e vengono attivati con un click del mouse quando è in run mode oppure alla ricezione di un altro messaggio o di particolari azioni.

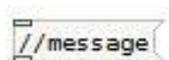


Figura 3.4: Messaggio

Un modo utile per attivare un messaggio è quello di utilizzare un bang, uno dei più importanti oggetti di Pd. La sua funzione è quella di innescare un'azione o una sequenza di azioni. Un altro oggetto molto comune nelle patch di Pd è la number box che consente di inviare messaggi numerici con il vantaggio di poter variare i numeri all'interno della box stessa con la pressione e il trascinarsi del mouse verso il basso per diminuire e verso l'alto per aumentare il numero stesso.

3.3 Oggetti e messaggi utilizzati

Nel lavoro si è fatto uso, come si avrà modo di vedere nel capitolo 6, dei seguenti oggetti e funzioni messi a disposizione da Pure Data:

- subpatch;
- blocchi send / receive;
- expr if;
- metro;
- select;
- set

illustrati brevemente qui di seguito.

3.3.1 Subpatch

Complicando ed aggiungendo funzionalità alla propria patch si rischia di compromettere la leggibilità a terzi utenti del proprio programma. Per risolvere questo problema Pure Data mette a disposizione le subpatch che consentono di dividere la programmazione in più patch windows tramite un oggetto con la sintassi



```
pd <nome_subpatch>
```

Figura 3.5: Subpatch

La subpatch viene quindi aperta con un click sull'object corrispondente.

3.3.2 Blocchi send / receive

Di default, le connessioni tra blocchi in Pure Data avviene attraverso i cavi, disegnati col mouse dall'outlet della sorgente all'inlet della destinazione. Sempre per risolvere il problema della leggibilità e, come nel mio caso, inviare dati tra pagine diverse di patch e subpatch, si può fare ricorso agli objects send e receive, attraverso la sintassi in Figura 3.6.



```
send <nome_variabile> receive <nome_variabile>
```

Figura 3.6: blocchi send e receive

Una funzionalità simile può essere ottenuta operando nella scheda opzioni dell'oggetto number o comment del quale si desidera rendere disponibile il valore come variabile locale: infatti in tale scheda è disponibile assegnare il nome col quale un altro blocco number o comment può fare riferimento ad esso (equivalente a send) oppure indicare il nome della variabile che deve essere ricevuta e visualizzata in tale blocco (equivalente a receive).

3.3.3 Expr if

expr è un versatile strumento che permette di raggruppare in un unico oggetto una serie di operazioni. Con Pd può capitare spesso di dover effettuare una sequenza di operazioni matematiche, attraverso expr si possono combinare insieme in modo da non occupare spazio e da rendere più chiara per il programmatore la sequenza delle operazioni stesse. Per poter funzionare con dei dati in ingresso expr necessita di simboli per rappresentare delle variabili, che, a

differenza delle variabili delle message box si scrivono in questa forma: \$xn dove x è il tipo della variabile e n è il numero ordinale che la rappresenta.

Le variabili possono essere di tipo intero, decimale, simbolo e rispettivamente avranno la forma \$in, \$fn e \$sn.

La forma di expr più utilizzata in questo lavoro è quella if che costituisce l'operatore condizionale in Pure Data. La sua sintassi è data da:

```
expr if($f1>10, $f1+1, $f1-1)
```

Figura 3.7: Esempio di blocco expr if

All'interno della condizione e del codice da eseguire nel caso in cui essa sia verificata o meno, si fa riferimento alle variabili disponibili all'ingresso del blocco expr tramite gli inlet, nell'ordine corretto.

3.3.4 Metro

Come accennato in precedenza, il Bang è un oggetto molto utilizzato nella programmazione in Pure Data come impulso per l'inizio di qualsiasi tipo di elaborazione. Tuttavia, in certi contesti, per rendere l'esecuzione della patch più autonoma viene messo a disposizione l'oggetto Metro, con la seguente sintassi:

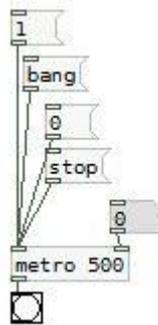


Figura 3.8: esempio di utilizzo dell'oggetto Metro

che accetta come parametro l'intervallo tra un bang e l'altro espresso in millisecondi. Nel primo inlet accetta due tipi di messaggi:

- "1" o "bang" avviano l'oggetto metro;
- "0" o "stop" fermano la generazione di bang.

Al secondo inlet di destra viene invece accettato, sotto forma di Number, il rate di generazione dei bang come parametro, sempre espresso in millisecondi.

3.3.5 Select

Nella sua forma più semplice, [select] controlla i suoi input confrontandoli con le costanti fornite dal programmatore all'interno dell'oggetto:

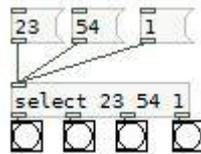


Figura 3.9: Esempio di utilizzo dell'oggetto Select

Se l'ingresso corrisponde con l'n-esimo parametro del select, l'oggetto produce un bang sull'n-esimo outlet. Nel caso non ci fosse nessun match, il select produce un bang nell'ultimo outlet.

3.3.6 Set

A differenza degli oggetti utilizzati in precedenza, questo prevede l'uso di un messaggio. Il set viene usato con la seguente sintassi:



Figura 3.10: Esempio di utilizzo del messaggio Set

il suo outlet è collegato ad un messaggio nel quale andrà a settare il valore. Tale valore può essere un valore prefissato dal programmatore oppure un valore presente nell'inlet ed indicato con la stessa sintassi delle variabili utilizzate per l'oggetto Select.

CAPITOLO 4: Ambiente di lavoro, cw_binaural

4. INTRODUZIONE

In questo capitolo viene presentato il motore binaurale utilizzato in Pure Data per gli esperimenti di audio 3D attraverso l'uso dei database di HRTF presentati nei capitoli precedenti.

4.1 cw_binaural

David Doukhan e Anne Sédès, entrambi dell'Università di Parigi, presentarono nel 2009 un nuovo external per Pure Data di nome cw_binaural. Essi presentarono “un nuovo external di sintesi binaurale per Pure Data adatto al rendering di suoni 3D tramite auricolari”[2]. Esso si interfaccia con l'utente in modo semplice e si propone come mezzo utile per la ricerca nel campo dell'audio binaurale.

4.1.1 Interpolazione di HRTF

Essendo azimuth ed elevazione dei valori continui, fornire l'intero set di HRTF per ogni singola posizione della sorgente sarebbe stato infattibile e per questo un sistema di sintesi binaurale deve essere in grado di interpolare le HRTF mancanti da quelle a sua disposizione nel database. Considerando sorgenti in movimento, una transizione brutale da una HRTF a un'altra sarebbe risultata come un'artefatto per l'ascoltatore simulando il movimento non lineare ma discreto della sorgente da un punto all'altro nello spazio: di conseguenza una interpolazione tra le risposte HRTF non fu ignorata da Doukhan e Sédès nella progettazione di cw_binaural.

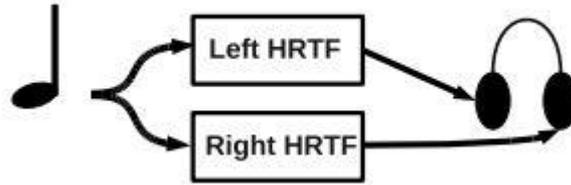


Figura 4.1: Il suono in ingresso a cw-binaural è processato con le HRTF destra e sinistra e udibile dall'utente tramite gli auricolari

Lo stesso problema è stato affrontato da diversi external che si proponevano lo stesso obiettivo di cw_binaural, adottando diverse soluzioni:

- earplug: adottò un'interpolazione lineare basata sull'uso della HRTF più vicina alla posizione richiesta. Tuttavia il risultato era di scarsa qualità, producendo il movimento a scatti di cui si accennava sopra;
- spat: usò un metodo basato sulla scomposizione della HRTF in una componente di fase minima e in una componente all-pass, procedendo poi ad un'approssimazione della componente all pass in un ritardo puro:

$$HRTF(f) = H_{min} \cdot H_{ap}$$

con H_{min} e H_{ap} componenti di fase minima e all-pass [2]. La struttura risultante consiste nell'uso di un ritardo corrispondente all'ITD (Interaural Time Delay) stimato dalla HRTF, in cascata a un filtro corrispondente alla componente di fase minima. Si procede poi alle interpolazioni separate delle due componenti. Questo è il metodo utilizzato anche in cw_binaural.

4.1.2 Implementazione di cw_binaural

Gli obiettivi perseguiti nell'implementazione di questo external furono principalmente due, ovvero creare un'interfaccia user-friendly e fornire un external il più possibile personalizzabile per la sintesi di audio 3d.

L'object cw_binaural presenta 3 inlet:

- la sorgente sonora;
- azimuth;
- elevazione.

Dispone inoltre di 2 outlet ovvero le uscite per l'altoparlante destro e sinistro delle cuffie. Propone un controllo intuitivo, ispirandosi al suo predecessore earplug, ma tuttavia cercando di coprire alcune sue grosse lacune quali:

- possibilità d'uso anche in assenza di HRTF;
- uso di più insiemi di HRTF simultaneamente;

- uso di diversi database di HRTF con numeri diversi di misurazioni;
- controllo di azimuth ed elevazione tramite un inlet numerico. I range di angoli in ingresso accettati dall'external sono:
 - $[-90^{\circ}, +90^{\circ}]$ per l'elevazione;
 - $[-180, +180[$ per l'azimuth.

Nel caso gli angoli in ingresso all'external non fossero compresi nel range indicato, viene invocato un metodo con lo scopo di normalizzare tale angolo in modo da includerlo in tale sistema di coordinate.

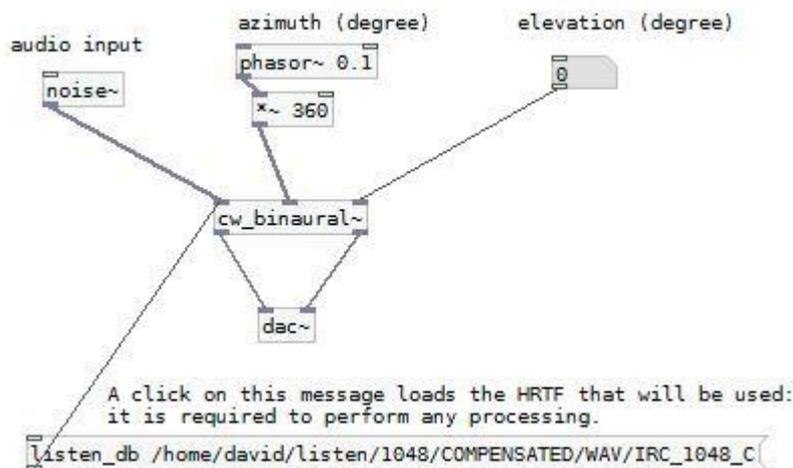


Figura 4.2: Semplice utilizzo di cw_binaural con rumore bianco come sorgente sonora

4.1.3 Funzionalità di cw_binaural

L'archivio che contiene l'external cw_binaural contiene anche diversi tutorial atti ad evidenziare a pieno le potenzialità dello stesso. Una prima applicazione è illustrata facendo riferimento alla figura 4.2: dopo aver aggiunto alla path di Pure Data la cartella in cui è contenuto il codice C++ dell'external, è possibile creare l'oggetto "cw_binaural~". Nel suo primo inlet viene posto una sorgente di rumore bianco, molto utilizzata in questo tipo di esperimenti. Il rumore bianco è un particolare tipo di rumore caratterizzato dall'assenza di periodicità nel tempo e da ampiezza costante su tutto lo spettro di frequenze. Nel secondo e terzo inlet, come già accennato sopra, stanno l'azimuth e l'elevation. Si noti in particolare come l'azimuth venga fatto variare grazie all'oggetto phasor~.

Il caricamento dell'HRTF viene fatto tramite invio di un messaggio al primo inlet. La sintassi per il caricamento è molto semplice per soggetti appartenenti ai database CIPIC e LISTEN, naturalmente supportati da cw_binaural. Infatti l'individuazione del tipo di database viene fatto, alternativamente, con le parole

chiave “cipic_db” o “listen_db”. Occorre poi individuare la cartella all’interno della quale è salvata la risposta del soggetto e inserirla come parametro all’interno dello stesso messaggio. Un click del mouse sul messaggio o un bang sul suo inlet scatena l’evento di caricamento dell’HRTF e quindi l’inizio dell’esperienza dell’ascolto di audio virtualizzato con cw_binaural.

L’external mette inoltre a disposizione una funzione Matlab denominata cipicdb2wav.m: essa ha lo scopo di convertire l’intero database CIPIC dal formato .mat con il quale è fornito dai CIPIC Laboratories in formato .wav, supportato da cw_binaural.

cipicdb2wav.m accetta come parametri d’ingresso il path dove è collocato il database CIPIC in formato .mat e il path dove si desidera salvare la versione convertita in .wav. Partendo poi dal .mat, effettua la scomposizione della risposta nelle varie posizioni previste dallo standard CIPIC (specificate in 2.3.1) creando il nuovo database con la stessa struttura di cartelle ma con le risposte memorizzate in formato .wav e quindi pronte all’utilizzo in cw_binaural.

4.1.3.1 HRTF database supportati

Cw_binaural provvede al supporto di ogni database, assumendo che ogni risposta sia salvata in forma di file .wav il cui nome contenga l’azimuth e l’elevazione espressa in gradi. In particolare e, come già detto sopra, i database LISTEN e CIPIC sono agevolati nell’uso grazie all’esistenza delle parole chiave “listen_db” e “cipic_db”. La sintassi del messaggio da inviare all’external per caricare una HRTF è, quindi:

```
listen_db /yourpath/1048/COMPENSATED/WAV/IRC_1048_C  
cipic_db /yourpath/subject_003
```

Figura 4.3: Caricamento di HRTF LISTEN e CIPIC

Il caricamento di un database generico avviene invece tramite l’invio del messaggio “set_hrtf_db” che dispone di 4 argomenti:

- il path alla directory contenente il database da utilizzare;
- una codifica su come estrarre dal nome dei file .wav i gradi di elevazione e azimuth;
- true o false a seconda che l’azimuth sia o meno fornito prima dell’elevazione nel nome del file .wav;
- true se il database utilizza coordinate polari verticali, false se invece utilizza coordinate polari interaurali.

A titolo di esempio, nella figura 4.4 è indicato il messaggio equivalente per caricare una risposta in formato LISTEN.

```
set_hrtf_db /yourpath/1048/COMPENSATED/WAV/IRC_1048_C
IRC_[0-9]+_C_R[0-9]+_T([-+]?[0-9]*//.?[0-9]+)_P([-+]?[0-9]*/
/.?[0-9]+)//.wav true true
```

Figura 4.4: caricamento di un soggetto LISTEN tramite “set_hrtf_db”

In set_hrtf_db si va a definire la composizione standard del nome dei file .wav ossia viene indicata la posizione alla quale è contenuto il nome del soggetto, l'angolo di azimuth, di elevazione e altri parametri. La composizione del messaggio sarà esplicitata completamente per il set del database MARL in 5.4.

4.1.3.2 Argomenti di creazione di un oggetto cw_binaural

Ogni istanza dell'external ha 3 argomenti di creazione con relativi valori di default.

Il primo è la lunghezza della risposta usata per il filtraggio, il cui valore di default è 128.

Il secondo argomento indica il metodo di filtraggio, il cui valore di default è 'RIFF'. Con questo metodo il filtraggio avviene nel dominio del tempo: questo comporta un costo computazionale considerevole ma latenza nulla. L'altro metodo di filtraggio supportato è 'FFT' che ha costo computazionale più basso ma latenza non più nulla.

Il terzo argomento, se settato a 'NODELAY', indica che l'interpolazione è eseguita direttamente sull'HRIR, comportando così minori costi computazionali ma minore qualità dal punto di vista percettivo: essa corrisponde alla tecnica adottata da *earplug* per la gestione del delay, ovvero di non eseguire la fase di scomposizione in filtro a fase minima e all-pass. In alternativa, si può procedere ad una scomposizione dell'HRIR in componenti di minimal phase e pure delay. L'utente può scegliere tra gli algoritmi disponibili per effettuare tale scomposizione:

- Hermite4: effettua l'interpolazione hermitiana su quattro punti;
- 6points: effettua l'interpolazione hermitiana su sei punti;
- linear: effettua l'interpolazione lineare su due punti;
- nofractional: non effettua nessuna interpolazione, si limita ad utilizzare i valori relativi alle posizioni spaziali disponibili.

Il valore di default per questo parametro è 'Hermite4'. Anche 6points è un buon metodo ma con un costo computazionale più alto. linear e nofractional invece sono sconsigliati ed utilizzabili solo per scopi didattici.

CAPITOLO 5: Creazione di una patch per l'utilizzo di cw_binaural

5 Introduzione

In questo capitolo viene esposto il risultato di questo lavoro, con l'illustrazione dei vari obiettivi raggiunti e della struttura della patch Pure Data progettata.

5.1 Premesse e scopi

Scopo di questo lavoro era arrivare alla creazione di una patch Pure Data utile dimostrativa dell'utilizzo dell'external cw_binaural. Essa si propone di essere semplice ed intuitiva, ponendo i comandi disponibili all'utente in modo ordinato e sequenziale.

I requisiti principali sono:

- Possibilità di interagire con HRTF appartenenti ai database CIPIC, LISTEN e MARL. Il caricamento di queste risposte deve avvenire in modo più o meno automatizzato.
- Possibilità di scelta della sorgente sonora con cui effettuare l'esperimento scegliendo rumore bianco o un qualsiasi file .wav presente nell'elaboratore.
- Possibilità di scegliere in modo manuale la posizione della sorgente nello spazio oppure di impostare il suo movimento o nella variabile azimuth o nella variabile elevation o in entrambe.
- In generale, arrivare al minor numero di object / messages / links in modo da rendere il più leggibile possibile la patch all'utente.

Inoltre, sempre come parte del lavoro di tesi, si è proceduto all'organizzazione dei database in forme simili per quanto riguarda la struttura delle directory: questo

vale in particolare per il database MARL, riorganizzato con una struttura delle cartelle simile a quella del database LISTEN esposta in 5.1.2.

Lo sviluppo è avvenuto tramite il sistema di controllo versione SVN. Le cartelle di cui si è fatto più utilizzo durante il lavoro sono organizzate come indicato in figura 5.1.

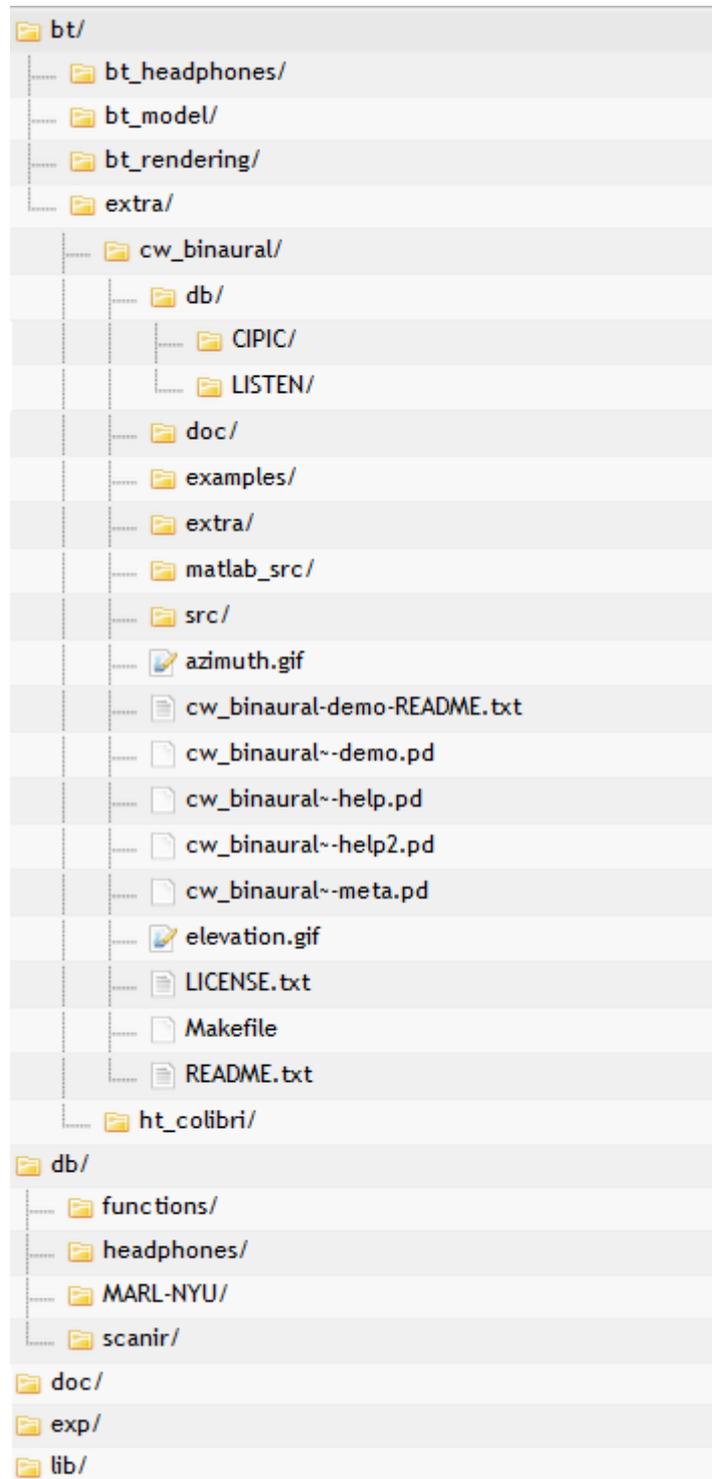


Figura 5.1: Struttura del progetto su SVN

5.1.1 Struttura del file system del progetto

Con riferimento alla figura 5.1, si procede ora ad una esposizione dei vari elementi di cui si compone il lavoro qui esposto:

- *db* contiene il database MARL nella sub-directory *MARL-NYU/* con la nuova struttura esposta in 5.1.2. Il database è posto in questa directory e non nella cartella in cui sono contenuti CIPIC e LISTEN perchè MARL costituisce un database di HRTF molto vario nella sua composizione (essendo ottenuto dall'unione di più archivi) e quindi da questa posizione è reso disponibile per l'utilizzo con altri progetti;
- *bt*, acronimo di Binaural Technologies, è la cartella che contiene al suo interno il lavoro di questa tesi e di altre attività di ricerca del gruppo SMC dell'Università di Padova;
- *extra/cw_binaural* è la cartella che contiene:
 - */db* contiene al suo interno i database CIPIC e LISTEN, entrambi in formato .wav e pronti all'utilizzo in *cw_binaural*;
 - */doc* contiene la presentazione della patch *cw_binaural* in formato .pdf a cura degli autori Doukhan e Sédès, citata in Bibliografia alla voce [2];
 - */examples* contiene patch di esempio atte alla comprensione delle funzionalità di base di *cw_binaural* come l'utilizzo di database di HRTF e gli argomenti di creazione di una istanza in Pure Data;
 - */matlab_src* contiene *cipicdb2wav.m*, funzione matlab la cui utilità è stata esposta in 4.1.3;
 - */src* contiene i sorgenti C++ della patch *cw_binaural*;
 - *cw_binaural~demo.pd* è il file Pure Data che contiene la patch sviluppata in questo lavoro ed esposta a partire da 5.2. Sempre in questa directory è disponibile un readme con indicazioni per l'utilizzo della patch;
 - *cw_binaural~help.pd* e *cw_binaural~help2.pd* sono patch di esempio fornite dai programmatori di *cw_binaural*.

5.1.2 Struttura del database MARL

Come accennato in 5.1, il database MARL è stato oggetto di una riorganizzazione della sua struttura in cartelle per essere simile a quella utilizzata dal database LISTEN.



Figura 5.2: Struttura del database MARL

All'interno di ciascun soggetto, si trovano due sottocartelle:

- *COMPENSATED/* per la risposta compensata;
- *RAW/* per la risposta non post-processata.

Tuttavia, il database MARL viene fornito nella sola versione RAW, perciò la cartella *COMPENSATED* e le sue sub-folders sono vuote: questa scelta è stata fatta nella previsione che possano essere rese pubbliche versioni *COMPENSATED* di tale database o che vengano calcolate o ricavate dalle versioni RAW già disponibili.

Ulteriori due sottocartelle, *MAT/* e *WAV/*, contengono, rispettivamente, le risposte in formato .mat e wav.

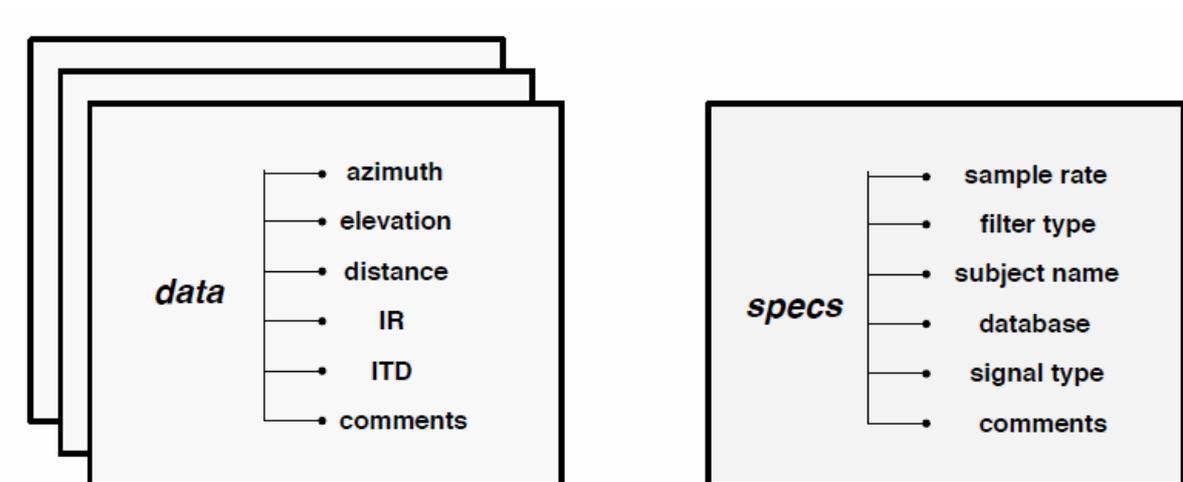


Figura 5.3: Struttura di una HRTF MARL [14]

La figura 5.3 illustra il contenuto del file .mat di una HRTF MARL. Nello specifico:

- *data*: ogni risposta impulsiva è memorizzata in questo array. Ogni elemento in questo array è una structure che contiene al suo interno le

informazioni riguardante tale risposta impulsiva, quali azimuth ed elevazione.

- specs: contiene tutte le informazioni generali riguardanti un particolare database di HRTF, quali ad esempio il sample rate, i filtri utilizzati e il nome del database.

5.2 Finestra principale

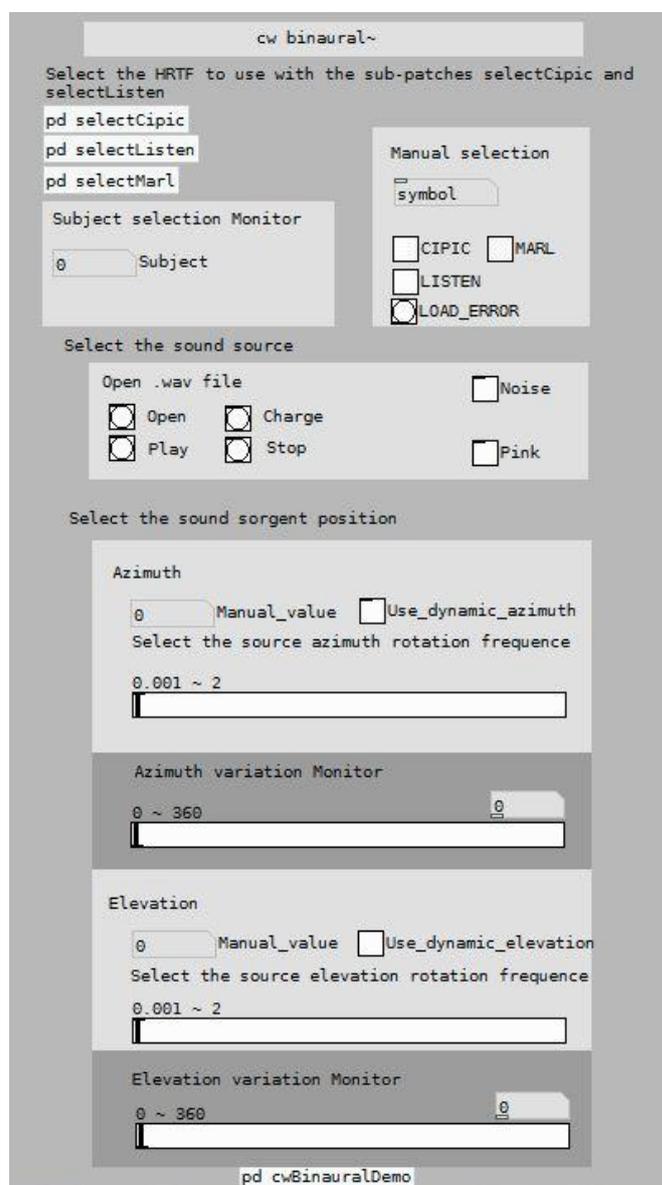


Figura 5.4: Interfaccia della demo per cw_binaural

Punto di partenza è stata la versione di cui si è parlato nel capitolo precedente e in figura 4.2; esso è stato poi esteso e, dopo varie aggiunte e modifiche, si è giunti alla finestra principale con l'aspetto di figura 5.4.

La finestra principale si compone di un menù sequenziale nel quale è possibile settare i vari parametri per l'esperimento.

La selezione del soggetto e quindi del database può essere effettuata tramite due modalità: guidata o manuale. Nella modalità guidata si ha a disposizione una sub-patch per ogni database, nel quale sono a disposizione degli oggetti bang che,

una volta premuti, consentono il caricamento della risposta desiderata. Un esempio di tale sub-patch è in figura 5.5 per il database CIPIC

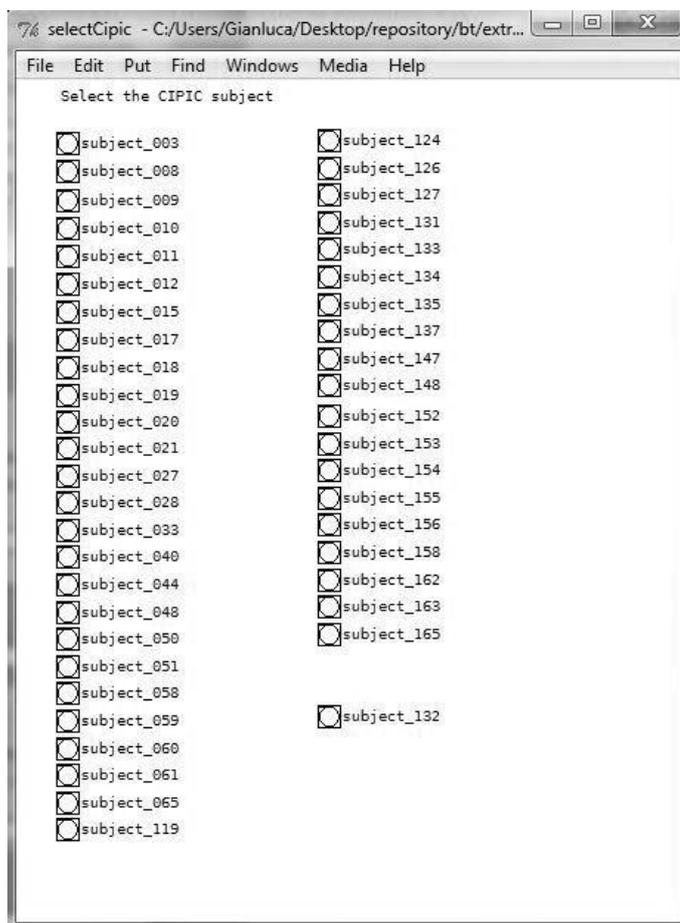


Figura 5.5: Sub-patch per la selezione guidata di soggetti CIPIC

Nel caso si volesse, invece, procedere ad una selezione manuale del soggetto, si ha a disposizione nella finestra principale un oggetto symbol all'interno del quale digitare il numero di soggetto desiderato (comprensivo di eventuali zeri tra le prime cifre a sinistra) e tre oggetti toggle per la selezione del database cui fa riferimento il numero di soggetto. A titolo di controllo, al di sotto di tali caselle, è stato aggiunto un oggetto bang (si veda, nella figura 5.4, il bang con etichetta LOAD_ERROR) che segnala un eventuale errore nella combinazione numero soggetto / database e quindi un'operazione di caricamento di risposta non andata a buon fine. Sia per il caricamento manuale che guidato, è a disposizione un monitor di controllo che visualizza il numero del soggetto correntemente caricato.

La sezione successiva propone la scelta della sorgente sonora da utilizzare; essa può essere di 3 tipi:

- Noise: corrisponde all'omonimo oggetto Pure Data e riproduce rumore bianco uniformemente distribuito;

- Pink: simile a noise, ma non uniformemente distribuito: si limita infatti alle frequenze più basse. In pratica, può essere ottenuto da Noise attraverso un filtro passa basso;
- File .wav.

La terza ed ultima sezione consente invece il controllo dell'azimuth e dell'elevation sia in modo automatico che in modo manuale. Per entrambe le scelte è a disposizione, nell'ordine, una casella number in cui andare ad inserire il valore numerico oppure, a lato, un oggetto toggle che una volta premuto avvia la rotazione automatica della sorgente nella variabile che si sta modificando. In una delle ultime modifiche alla patch si è deciso di aggiungere la facoltà di selezionare la velocità di rotazione della sorgente: questo è stato ottenuto con l'aggiunta di un hslider con la selezione della frequenza, in Hz, di rotazione. Due monitor consentono di visualizzare lo stato attuale delle due variabili in termini di gradi e sono costituiti da due oggetti hSlider.

5.3 Core Patch cwBinauralDemo

Si procede ora ad una descrizione sommaria della struttura della sub-patch cwBinauralDemo. Il nucleo originario, ovvero quello proposto nella figura 4.2, si è evoluto fino alla forma illustrata in figura 5.6.

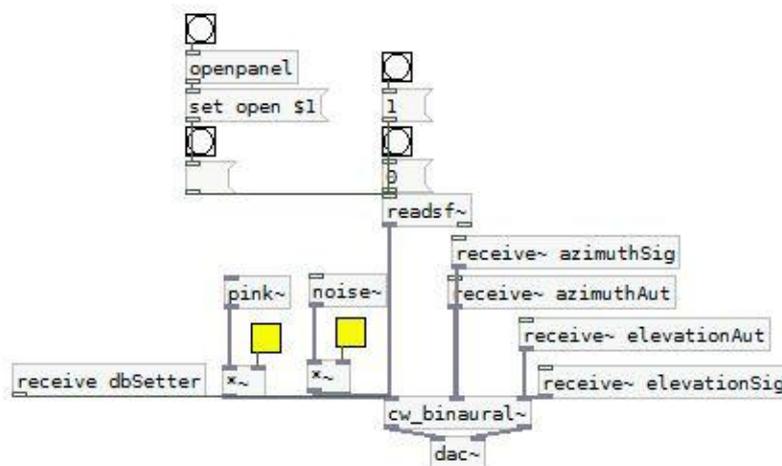


Figura 5.6: forma finale degli inlet a cw_binaural

Qui si può notare il largo utilizzo che si è fatto dei collegamenti wireless per l'invio di variabili all'interno della patch senza l'utilizzo di cavi, allo scopo di rendere il più leggibile e leggero l'intero applicativo. Infatti:

- dbSetter è la variabile che riceve il messaggio per il caricamento del soggetto dal database e il cui processo di composizione sarà chiarito più avanti in questa sezione;

- azimuthSig e azimuthAut sono le variabili numeriche corrispondenti, rispettivamente, alla composizione manuale ed automatica dell'azimuth (cosa analoga vale per elevationAut e elevationSig).

La gestione della selezione dei soggetti e del relativo monitor è effettuata tramite controlli incrociati che inviano bang alle rispettive caselle presenti nella finestra principale, facendo anche qui uso delle variabili locali e dei blocchi send / receive. Il bang di errore segnalato nel caso di una combinazione non consistente tra soggetto e database è generato tramite la composizione mostrata in Figura 5.7. Da notare come, con l'oggetto print, tale messaggio venga anche stampato a video nella main window di Pure Data.

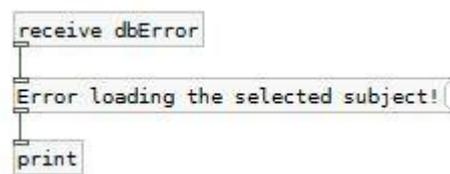


Figura 5.7: Segnalazione dell'errore nel caricamento di un soggetto

La selezione dei soggetti avviene, infine, tramite la combinazione mostrata in Figura 5.8 e mostrata qui in versione leggermente semplificata per consentirne una più veloce comprensione.

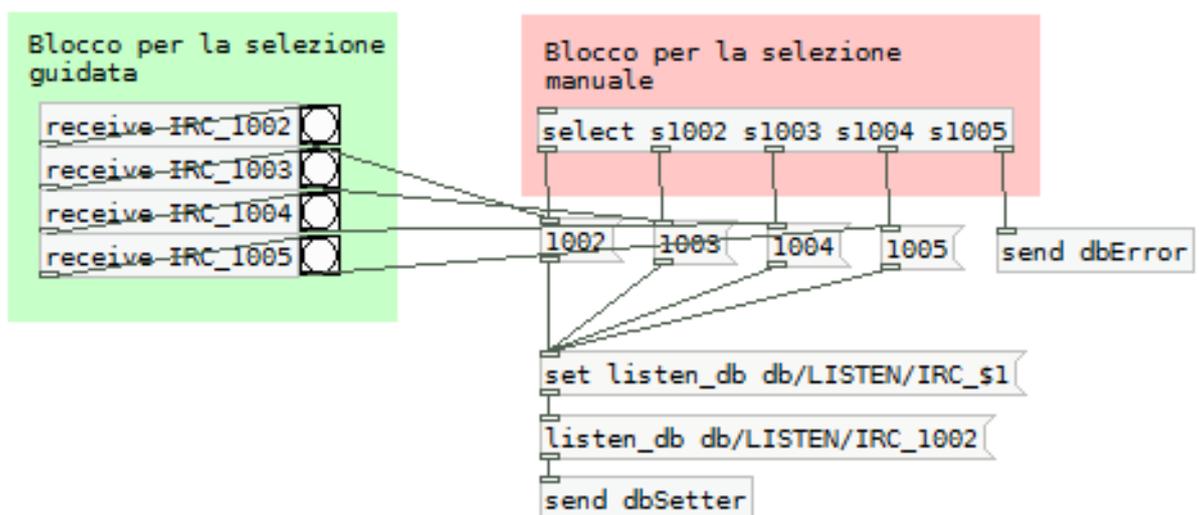


Figura 5.8: Porzione di blocco per la selezione manuale / guidata dei soggetti LISTEN

La composizione del messaggio da inviare a dbSetter avviene tramite concatenazione con un messaggio set standard. In alto a sinistra sono presenti i receive relativi ai bang presenti nelle subpatch per la selezione guidata dei soggetti dei database. Essi inviano un bang al numero del soggetto da concatenare al set. La selezione manuale, invece, avviene tramite un oggetto select, al cui ingresso (per semplicità qui tralasciato) sta il contenuto della casella di selezione manuale presente nella finestra principale. Come noto, il select produce un bang all'uscita corrispondente, componendo così, in modo simile alla selezione guidata, il messaggio da inviare tramite dbSetter. Un numero di soggetto errato viene segnalato all'ultima casella del select e inviato, tramite la variabile dbError, al bang presente nella finestra principale che segnala un errore nel caricamento di un soggetto.

5.4 Importazione del database MARL

Cw_binaural consente l'utilizzo di qualsiasi database di HRTF a patto che esso sia fornito tramite file .wav e che nel nome del file sia contenuta l'indicazione dell'angolo di azimuth e di elevazione. Il database MARL viene fornito dall'University of New York in formato .mat ma, come detto precedentemente, viene messa a disposizione la funzione exportAudio.m che consente la traduzione dell'intero db in formato .wav. Una volta effettuato questo passaggio, si ottiene l'intero database con nomi dei file nel formato illustrato in Figura 5.9.



Figura 5.9: File .wav di HRTF per un soggetto MARL

La struttura tipica di un nome di file .wav in formato MARL è data da:

S<numero_soggetto>_Az<angolo_di_azimuth>_El<angolo_di_elevazione>.wav

Da tale struttura si nota fin da subito che risponde ai requisiti per l'utilizzo in cw_binaural. Si ha, quindi, che il messaggio per il caricamento del database è del tipo illustrato in Figura 5.10.

```
set_hrtf_db ../../../../db/MARL-NYU/S001/RAW/WAV  
S+[0-9]+_Az([-+]?[0-9]+)_El([+]?[0-9]+)//.wav true true
```

Figura 5.10: Messaggio per il caricamento di HRTF MARL

il messaggio `set_hrtf_db` è così composto:

- `../../..` indicano che il database MARL si trova a 3 livelli superiori nella gerarchia di directory;
- `db/MARL-NYU/S001/RAW/WAV` accede alla cartella in cui sono contenute le HRTF RAW relative al soggetto S001;
- `S` indica la prima lettera fissa che compone tutti i file `.wav` in formato `wav`. Segue la prima parte variabile, concatenata ad `S` tramite l'operatore "+": tra parentesi quadre `[0-9]` è indicato che il soggetto è espresso da un numero.
- Segue la definizione dell'azimuth: concatenata alla parte precedente del messaggio sempre tramite "+", la parte fissa del nome del file che definisce l'azimuth è `"_Az"`. La parte variabile contenente i gradi è indicata dall'espressione

$$([-+]?[0-9]+)$$

dove `[-+]?` indica che l'angolo è indicato in quantità sia positive che negative e `[0-9]`, come in precedenza, indica che in quella posizione sarà contenuto un valore numerico;

- La definizione dell'elevazione è simile a quella dell'azimuth: varia solo la parte fissa che ora è `"_El"`;
- il primo `true` indica che l'azimuth è fornito per primo all'interno del nome del file;
- il secondo `true` indica che MARL fa uso delle coordinate polari verticali.

CAPITOLO 6: Considerazioni e sviluppi futuri

6.1 Considerazioni sull'implementazione

Alla conclusione di questo lavoro, la principale considerazione che si può fare circa un eventuale utilizzo dell'external cw-binaural consiste nella possibilità di utilizzare la combinazione in cascata di più componenti strutturali a comporre la risposta impulsiva completa. Risulterebbe così possibile la concatenazione della risposta impulsiva dell'ambiente di ascolto con la risposta del busto del soggetto e quella del suo padiglione auricolare, trattate e modellate separatamente. In questo modo diventerebbe più agevole la creazione di risposte *ad hoc* personalizzate per le caratteristiche di un certo soggetto. Anche la risposta della cuffia utilizzata durante l'ascolto elimina la colorazione indesiderata sull'audio dovuta al particolare modello di cuffie utilizzate.

6.2 Utilizzo futuro di cw-binaural

L'utilizzo futuro di cw-binaural pone diverse prospettive di estensione e svariate direzioni applicative. Nell'ordine, tale external sarà ampiamente utilizzato negli esperimenti di psicoacustica per studiare le differenze tra HRTF di diversi soggetti e quindi su come un suono venga udito da soggetti diversi. Altro campo di applicazione è certamente il rendering di audio binaurale: allegando ad ogni risposta disponibile nei database utilizzati le caratteristiche fisiche del soggetto considerato, posso trovare, per qualunque ascoltatore volesse fare esperienza di audio 3d, la risposta più adatta da utilizzare, pur non disponendo della sua HRTF individuale.

Più in generale, cw-binaural potrà essere utilizzato per esperimenti di spazializzazione del suono e le numerose applicazioni quali l'home video per la visione di film, l'ascolto di concerti e la realtà virtuale, solo per citarne alcuni.

Un aumento dell'accuratezza nella precisione non può prescindere da un aumento dei soggetti presenti nei database utilizzati e, possibilmente, l'aggiunta di

interi nuovi database pubblicati da qualche laboratorio di ricerca e facilmente utilizzabile nell'external con la funzionalità "set_hrtf db".

In conclusione, la creazione della patch di demo oggetto di questo lavoro di tesi si è posta l'obiettivo di creare un'interfaccia per esperimenti di psicoacustica e audio spazializzato versatile e facilmente maneggiabile, impiegabile per un semplice ascolto di come opera una HRIR appena registrata e per la simulazione di ambienti dinamici come una sorgente in rotazione attorno ad un soggetto nello spazio tridimensionale.

Bibliografia

- [1] M. Geronazzo, S. Spagnol, F. Avanzini. *Customized 3D sound for innovative interaction design*. In Proc. of the Computer Human Interaction Italy 2011 Conference, Alghero. Settembre 2011.
- [2] D. Doukhan, A. Sédès. *CW_binaural: A binaural synthesis external for Pure Data*. Pure Data Convention 2009 , Weimar. Luglio 2009.
- [3] C. Drioli, N. Orio. *Elementi di Acustica e Psicoacustica*. Disponibile all'indirizzo <http://www.dei.unipd.it/~musica/IM/cap2.pdf>, 17/10/2012.
- [4] C. Cheng, G. Wakefield. *Introduction to Head-Related Transfer Functions (HRTFs): Representations of HRTFs in Time, Frequency and Space*. Journal of Audio Engineering, vol. 49, no. 4, pp. 231-249. Aprile 2001
- [5] F. Avanzini, G. De Poli. *Sound in space*. Disponibile all'indirizzo http://www.dei.unipd.it/~musica/IM06/Dispense06/4_soundinspace.pdf, 17/10/2012.
- [6] M. Geronazzo, S. Spagnol, F. Avanzini. *Hearing distance: a low-cost model for near-field binaural effects*. In Proc. of the European Signal Processing Conference, Bucharest. Agosto 2012.
- [7] E. Giordani. *Elementi di spazializzazione del suono*. Disponibile all'indirizzo http://xoomer.virgilio.it/e.giordani/docs.spat_a.pdf, 17/10/2012.
- [8] V. R. Algazi, R. O. Duda, D. M. Thompson. *The CIPIC HRTF database*. In Proc of the 2001 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics. Ottobre 2001.
- [9] <http://recherche.ircam.fr/equipes/salles/listen/>, 17/10/2012.
- [10] Swen, Muller. *Transfer-Function Measurement with Sweeps*. Journal of Audio Engineering, vol. 49, issue 6, pp. 443-471. Giugno 2001.
- [11] <http://www.head-fi.org/t/111193/orthodynamic-roundup/18825>, 17/10/2012.
- [12] <http://puredata.info>, 17/10/2012.

-
- [13] F. Bianchi. *Inventare il suono con Pure Data*. Disponibile all'indirizzo http://www.webprog.it/fileguida/Inventare_il_suono.pdf, 17/10/2012.
- [14] A. Andreopoulou, A. Roginska. *Documentation for the MARL-NYU file format*. In Proc. of the 131st Audio Engineering Society Convention, New York. Ottobre 2011.

Ringraziamenti

Giunto a un traguardo così importante come la laurea, così intenso e frutto d'impegno e numerosi sacrifici, formulare dei ringraziamenti risulta difficile sia per la quantità di persone che mi hanno supportato durante questi quattro anni che per il rischio di dimenticare qualcuno tra i tanti amici, educatori ed in generale compagni di vita che hanno contribuito, ciascuno a suo modo, a formare quella persona che oggi mi trovo ad essere.

Prima di tutto, il riconoscimento più grande va ai miei genitori che costituiscono da sempre per me un modello incomparabile ed insostituibile. A loro devo anche il supporto che non mi è mai mancato in questi anni anche nei momenti più difficili quando sembrava più facile lasciare tutto e quindi grande parte di quello che rappresenta questa laurea è merito loro.

Si uniscono alla famiglia i miei fratelli Elena e Federico che non hanno mai rinunciato ad essere interessati ed attivi nel tenere alto il morale in tutte le varie occasioni di esami e prove. Viene poi la mia fidanzata Paola che con affetto mi ha sempre motivato anche quando gli obiettivi di breve periodo non riuscivano come previsto e con immancabile pazienza, specie in questo periodo di tesi e ultimi esami, ha saputo sopportare le mie lunghe assenze per studio e preparazione delle prove.

Come promesso, non posso dimenticare Nico, Ciro ed in generale tutta la famiglia Patrian che, grazie al lavoro estivo, mi hanno dato la possibilità sì di proseguire negli studi, ma anche di vivere con allegria e amicizia l'estate e concedendomi la chance di vivere l'esperienza lavorativa con una passione ed un affiatamento che difficilmente credo mi si riproporrà in futuro.

Gli amici, che per ragioni di spazio non posso nominare uno a uno, hanno dato il loro contributo, ciascuno in modo particolare, e ringrazio specialmente quella ristretta cerchia che, arrivati a mantenere questo rapporto alla mia età, mi danno la sicurezza di poter coltivare ancora l'amicizia nelle numerose prospettive che mi si porranno da ora in avanti.

Un particolare ringraziamento va anche al gruppo di Informatica Musicale ed in particolare all'ingegnere Geronazzo Michele e il professore Avanzini Federico

che con pazienza mi hanno aiutato a completare questo lavoro e mi hanno fatto scoprire un ambito di studi molto interessante e che spero potrò, in qualche modo, approfondire e ampliare nei miei progetti di studio futuri.

Ultimi, ma non per importanza, tutti i miei parenti che costituiscono la “famiglia allargata” di zii e cugini ed in modo speciale alla mia nonna Nina che, con inesauribile affetto ed attenzione ha seguito la mia carriera universitaria e sicuramente spero di rendere ulteriormente orgogliosa di me negli anni a venire.

In conclusione, un caro ricordo dei nonni Amleto e Bianca, che con mio grande dispiacere non possono gioire con me oggi di questo importante traguardo ma che hanno vissuto con me, presenti ai miei primi esami in modo fisico e, in questi ultimi due anni, con una presenza mai estinta nonostante l'assenza materiale. Oltre che ai miei genitori, dedico a loro questo lavoro di tesi, come segno di ringraziamento per tutto ciò che sono stati per me e dell'affetto che mi lega, nonostante tutto, ancora oggi alle loro persone.