

ABSTRACT

Il lavoro di ricerca qui presentato mira ad analizzare il contributo acustico della cuffia nel rendering di audio spazializzato. Viene descritto il setup sperimentale utilizzato in laboratorio per effettuare la registrazione delle Headphone Impulse Responses (HPIRs) di 7 diversi modelli di cuffie tramite il manichino KEMAR. L'analisi dei risultati ottenuti trova un'interpretazione grazie all'introduzione delle nozioni di base che descrivono i meccanismi messi in atto dal cervello umano per la localizzazione di una sorgente sonora nello spazio. Grazie ai risultati ottenuti dallo studio delle diverse HPIR, vengono individuate le caratteristiche che una cuffia deve soddisfare per riprodurre sul timpano la stessa pressione acustica che si misura durante l'ascolto reale a campo diffuso (FEC).

SOMMARIO

0. PRAFAZIONE

1. INTRODUZIONE

- 1.1 UN EXCURSUS STORICO
- 1.2 3D AUDIO
- 1.3 AUDIO TRANSAURALE E BINAURALE

2. CUFFIE

- 2.1 TASSONOMIA
 - 2.1.1 TRASDUTTORE
 - 2.1.2 POSIZIONAMENTO
 - 2.1.3 ISOLAMENTO ACUSTICO
- 2.2 CRITERIO FEC
 - 2.2.1 ASCOLTO FREE FIELD
 - 2.2.2 ASCOLTO TRAMITE HEADPHONE

3. MISURAZIONI

- 3.1 POSIZIONE DI ACQUISIZIONE
- 3.2 CATENA DI ACQUISIZIONE
- 3.3 CATENA DI RIPRODUZIONE
- 3.4 REGISTRAZIONE HpIR

4. RISULTATI

- 4.1 HpTF
- 4.2 HpTF STANDARD DEVIATION
- 4.3 ILD
- 4.4 PHASE
- 4.5 DIFFERENZE DI PHASE INTERAURALI

5. CONCLUSIONI

PREFAZIONE

Il lavoro qui presentato si inserisce nell'ambito di ricerca sull'audio binaurale.

L'idea alla base dell'audio binaurale è che ogni evento sonoro che percepiamo, è racchiuso a livello percettivo da due segnali di pressione sonora (uno per ogni orecchio) che tramite l'apparato uditivo vengono interpretati dal cervello. Se fosse quindi possibile registrare questi segnali con una coppia di microfoni all'interno di ogni orecchio molto prossimi al timpano, saremmo in grado di acquisire, e quindi riproporre in un secondo momento, tutte quelle informazioni spaziali che il nostro cervello utilizza per codificare la posizione delle sorgenti sonore che ci circondano.

Per tale motivo, le tecniche base di ripresa binaurale sono effettuate tramite una coppia di microfoni miniaturizzati inseriti nel canale uditivo di un ascoltatore o utilizzando un manichino (Head and Torso Simulation HATS). Tale manichino, munito di microfoni posizionati all'interno delle orecchie, riproduce le caratteristiche umane medie di busto e testa, dal punto di vista anatomico e dei coefficienti di riflessione e assorbimento della pelle. Il materiale così acquisito, dopo un'adeguata equalizzazione, viene riprodotto tramite una coppia di altoparlanti o un paio di cuffie.

Entro quali limiti il sistema di riproduzione può però essere considerato trasparente al materiale audio veicolato?

Scopo della sperimentazione condotta è studiare i contributi apportati dalla cuffie, così da individuare dei cluster di cuffie maggiormente idonee al rendering di audio binaurale e di conseguenza fornire tale materiale per un successivo lavoro di compensazione.

Nel primo capitolo di questo elaborato viene presentato l'ambito di ricerca dell'audio 3D e come esso costituisca un'importante evoluzione rispetto le classiche tecniche stereofoniche. Vengono spiegati i meccanismi con i quali il cervello interpreta gli stimoli sonori secondo differenze interaurali di tempo e spettro, così da fornire al lettore gli strumenti per una lettura critica dei risultati esposti al capitolo quattro.

Sono inoltre discusse le due tecniche di riproduzione binaurale tramite coppia di altoparlanti o di cuffie, sottolineando vantaggi e limiti di ciascuna, e per quale motivo si sia scelto di seguire la seconda.

Le principali caratteristiche di una cuffia e il suo funzionamento sono presentati nel secondo capitolo. Una classificazione secondo il tipo di trasduttore (dinamiche ed elettrostatiche), del livello di isolamento dai rumori esterni (chiusa, aperta e semi aperta) e dalla posizione del padiglione rispetto l'orecchio esterno (circaurale, supraurale, auricolari ed in-ear) estende la comprensione dei vari modelli presi in analisi. Viene quindi presentato un modello trasmissivo per descrivere la riproduzione della registrazione binaurale tramite cuffie.

All'inizio del terzo capitolo, è analizzata l'influenza della posizione dei microfoni rispetto al canale uditivo, sull'informazione spaziale acquisibile tramite le HPIR. Viene quindi descritto il setup della catena hardware utilizzata durante gli esperimenti, analizzando con attenzione le prestazioni di ogni componente.

Il dati ottenuti e i risultati della loro elaborazione sono presentati nel quarto capitolo. Particolare attenzione è posta all'analisi di quelle caratteristiche che potrebbero compromettere una corretta esternalizzazione del segnale binaurale.

Nell'ultimo capitolo sono evidenziati i risultati di particolare interesse emersi durante l'elaborazione dei dati e a fronte degli stessi sono proposti alcuni sviluppi. Viene discusso l'uso di Head Related Transfer Function (HRTF) individuali e sono presentati alcuni cenni sulla compensazione delle HPIR.

1 INTRODUZIONE

1.1 UN EXCURSUS STORICO

Il desiderio di ricreare un panorama sonoro quanto più realistico e fedele alla nostra percezione uditiva accompagna l'ingegneria del suono fin da suoi albori.

I primi esperimenti per la messa a punto di sistemi di registrazione orientati alla riproduzione mediante altoparlanti risalgono agli inizi '20 , ma già negli anni '30 i laboratori Bell Labs in USA e l'ingegnere Alan Blumlein in Europa sviluppano le prime tecniche di riproduzione stereofonica.

La tecnica stereofonica prevede che il suono sia registrato e riprodotto su due canali distinti: destra e sinistra. La scena sonora ricostruita dall'ascoltatore in fase di riproduzione è compresa tra le 2 sorgenti ed è così possibile ottenere una buona localizzazione sul piano orizzontale. Un notevole sviluppo rispetto la precedente tecnica monofonica (un solo canale in riproduzione), ma lo scenario uditivo prodotto risulta ancora piccolo e monodimensionale se confrontato con l'esperienza sensitiva che ci offre ogni giorno il nostro apparato uditivo.

Con l'obiettivo di ottenere sempre un maggior coinvolgimento ed immersione dell'ascoltatore nell'evento sonoro, numerose tecniche multicanale sono da sempre sperimentate.

Nel 1939 gli ingegneri RCA e Disney svilupparono un sistema di acquisizione a 9 tracce per la registrazione dei segnali provenienti da 33 microfoni sparsi tra gli elementi di un'orchestra sinfonica. Qualche anno più tardi durante l'esposizione Universale di Bruxelles del 1958 gli ingegneri della Philips affiancarono Edgar Varèse nella realizzazione di un padiglione munito di 350 altoparlanti per un'installazione artistica di suoni luci e immagini.

A partire dagli anni '60 - '80 la diffusione dei computer fornisce nuovi strumenti agli ingegneri dell'audio e apre nuovi orizzonti alla spazializzazione del suono. Riacquistano così interesse progetti che erano stati abbandonati perché troppo complessi da realizzare o troppo costosi.

1.2 3D AUDIO

Per spazializzazione del suono si intende il controllo della mappa sonora che l'ascoltatore ricostruisce a partire dai segnali acustici che investono l'apparato uditivo. La nostra vita di tutti i giorni comporta molte esperienze uditive di tipo spaziale : si tratta di una consapevolezza che riguarda la posizione degli oggetti, la loro dimensione e il loro movimento nello spazio che circonda l'ascoltatore.

Per modellizzare lo spazio attorno all'ascoltatore si utilizzano tre direzioni principali : fronte-dietro, sinistra-destra, sopra-sotto. Si individuano così il piano frontale, il piano mediano e il piano orizzontale, e nel loro punto di intersezione è situata la testa dell'ascoltatore. (Figura 1.1)

L'orecchio posto sullo stesso lato della sorgente è detto ipsilaterale mentre l'orecchio che si trova dal lato opposto è detto controlaterale.

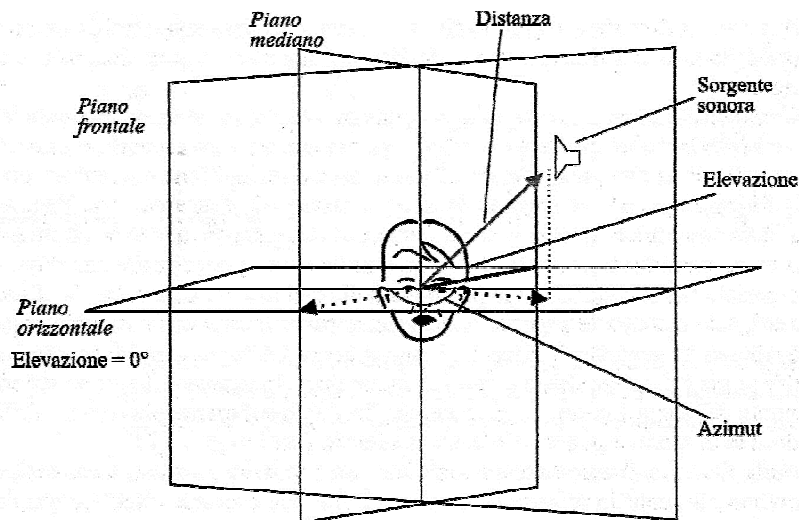


Figura 1.1 Posizionamento di un oggetto rispetto all'ascoltatore

Gli studi condotti sull'ascolto direzionale hanno individuato due meccanismi alla base della spazializzazione delle sorgenti sonore : ITD e ILD.

L' *interaural time difference* (ITD) viene rilevata quando una sorgente non si trova nel piano mediano. In questo caso l'orecchio contro laterale riceve lo stimolo sonoro in ritardo rispetto l'orecchio ipso laterale. L'uso della ITD e' molto preciso e permette di localizzare una sorgente sonora entro un solo grado, ma dal momento che la testa rappresenta un'ombra acustica per le alte e medie frequenze, questo meccanismo è efficace soltanto sulla componente del segnale a basse frequenze. (Figura 1.2)

L' *interaural level difference* (ILD) rappresenta invece una differenza spettrale tra i due segnali che raggiungono le orecchie. Questa differenza è frutto di un'azione di filtraggio ad opera dell'ombra acustica dalla testa e dall'azione del padiglione auricolare. L'orecchio esterno, infatti, a causa della creazione di risonanze e riflessioni filtra il segnale audio che raggiunge il timpano. Questo filtraggio, frutto dell'iterazione tra il suono diretto che entra nel canale uditivo, e quello che viene riflesso sulle pieghe del padiglione arricchisce il segnale sonoro di informazioni utili per la sua collocazione spaziale. (Figura 1.2) (binaural technique 9.5 figure?)

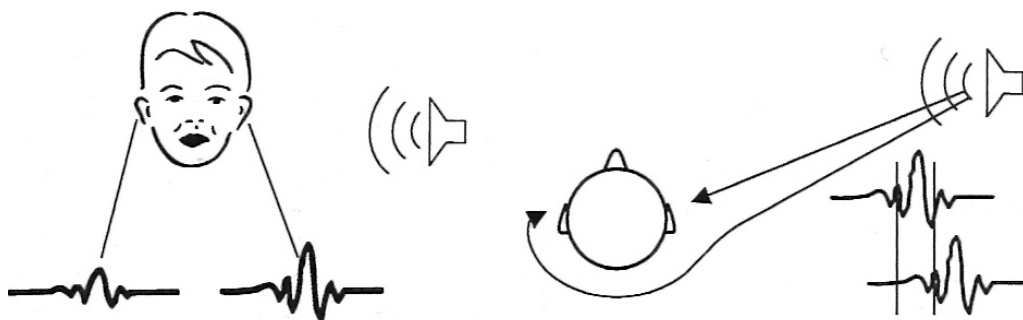


Figura 1.2 Differenza interaurale di intensità ILD (**sinistra**) e differenza interaurale di tempo ITD (**destra**)

Tutti gli indizi sul filtraggio spettrale prodotti da testa, padiglione, spalle e busto possono essere riassunti nella Head Related transfer function (HRTF).

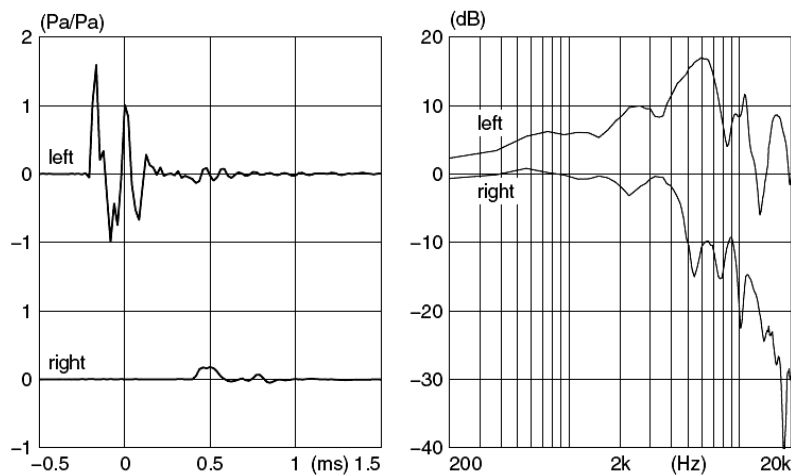


Figura 1.3 HRTF nel dominio del tempo (**sinistra**) e nel dominio della frequenza (**destra**)

La HRTF descrive quindi tutti i cambiamenti di ampiezza, fase e forma d'onda che occorrono alle nostre orecchie a causa della presenza del soggetto nello scenario acustico.

Secondo il modello introdotto nel capitolo 2.2 per descrivere la situazione di propagazione del suono a campo libero (free-field), la HRTF viene definita come il rapporto tra la pressione sonora sul timpano, e una pressione sonora utilizzata come riferimento misurata al centro della testa in assenza della stessa :

$$HRTF(\angle) = \frac{P_{ear}(\angle)}{P_{ref}}$$

Come indicato nell'espressione, tale rapporto è funzione dell'angolo di incidenza dell'onda sonora.

Gli esperimenti condotti e le conoscenze acquisite sulla collocazione delle sorgenti sonore nello spazio ad opera del cervello, ha offerto le conoscenze su cui avviare un lavoro di ricerca sulla spazializzazione artificiale del suono.

1.3 AUDIO TRANSAURALE E BINAURALE

Un'importante classificazione si basa sul metodo utilizzato per la riproduzione del materiale audio. Si distingue l'audio binaurale che fa uso di cuffie (headphones), e l'audio transaurale che impiega un sistema di altoparlanti a due canali.

Dal momento che tutte e due le tecniche sono oggetto di ricerca è evidente come non siano una migliore dell'altra ma presentino entrambe vantaggi e svantaggi che le fanno preferire per un'applicazione piuttosto che un'altra.

E' innanzitutto evidente, come con l'uso delle cuffie, l'orecchio destro sia interessato solo dal materiale audio del canale destro, e viceversa il sinistro. Questo permette un assoluto controllo da parte del progettista del sistema audio 3D sul materiale audio ascoltato evitando fenomeni di cross-talk che sono invece presenti in una riproduzione transaurale. (Figura 1.4)

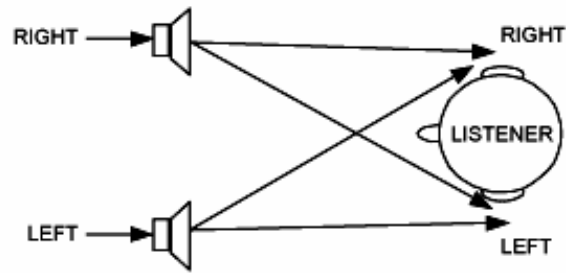


Figura 1.4 Fenomeni di *crosstalk* nell'ascolto transaurale

Il grado di comfort offerto da una cuffia (soprattutto se utilizzata per lunghi periodi) può essere molto scarso se confrontato alla sensazione di naturalezza e apertura offerta da un ascolto a campo diffuso (tramite loudspeakers).

La libertà di movimento associata all'ascolto transaurale può però compromettere la resa del sistema audio 3D: se si considera l'ITD, risulta evidente come un movimento dell'ascoltatore verso un altoparlante costituisca una distorsione non prevedibile dal progettista e che quindi falserà la corretta collocazione della sorgente. Viene così a individuarsi una regione spaziale limitata, detta *sweet spot*, all'interno della quale le informazioni spaziali riprodotte saranno correttamente interpretate dal soggetto. Al di fuori di essa, invece, la resa del sistema non è prevedibile ma in generale non porterà alla corretta spazializzazione.

E' invece ragionevole considerare che l'ITD nella riproduzione tramite cuffie sia minimo, visto che entrambi i segnali percorrono la stessa distanza (il canale uditivo).

Si noti inoltre che grazie all'uso di cuffie, è possibile isolare l'ascoltatore dalle caratteristiche acustiche dell'ambiente fisico nel quale si effettua la riproduzione. Al contrario un ascolto diffuso tramite altoparlanti ne è influenzato e acquisisce le informazioni acustiche dell'ambiente. Tale limite può risultare particolarmente pesante in quelle applicazioni di realtà aumentata, dove lo scenario virtuale si sovrappone allo quello reale.

Infine il setup di riproduzione tramite cuffie costituisce una soluzione trasportabile e di basso costo: due aspetti da tenere in considerazione nell'ottica di applicazioni consumer.

2 CUFFIE

2.1 TASSONOMIA

2.1.1 TRASDUTTORE

A seconda del principio di funzionamento del trasduttore vengono distinte cuffie dinamiche ed elettrostatiche.

In una cuffia dinamica un magnete (tipicamente Ferrite o Neodinomio) è utilizzato per generare un campo magnetico stazionario.

Il trasduttore costituito da una membrana sottile e solidale ad una bobina immersa in questo campo magnetico, viene messo in movimento dalle variazioni di campo magnetico indotto dalla corrente che attraversa la bobina. (Figura 2.1)

Il principio è del tutto analogo a quello degli speaker e dei microfoni dinamici.

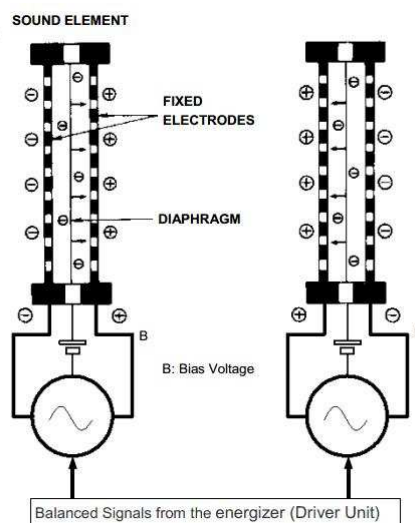


Figura 2.1 Trasduttore dinamico (**sinistra**) , modello di funzionamento di un trasduttore elettrostatico (**destra**)

In una cuffia elettrostatica il diaframma è una sottile membrana caricata elettricamente e posta tra due armature metalliche perforate. A queste due armature è applicato il segnale elettrico da riprodurre.

La membrana carica viene messa in movimento in funzione della polarità e dell' intensità del campo elettrico prodotto: l'aria viene così forzata tramite i fori delle armature a generare l'onda sonora.

L'assenza di parti in movimento ad eccezione della membrana e lo spessore di pochi μm della stessa conferiscono alle cuffie elettrostatiche un'estesa risposta in frequenza ben oltre i 20.000Hz portando così le attenuazioni da estremi di banda fuori dal range dell'udito umano.

Per pilotare tale tipologia di cuffie è necessario un amplificatore, per fornire le tensioni (0.1 - 1 kV) necessarie alle armature per generare il campo magnetico. (Figura 2.1)

2.1.2 POSIZIONAMENTO

Un altro parametro che caratterizza le cuffie è il posizionamento rispetto l'orecchio esterno . Si distinguono cuffie *extraurali*, *circumaurali*, *sovraurali*, *auricolari* e *in-ear*.

Le cuffie *extraurali* sono composte da una coppia di speaker full range posti a pochi cm davanti le orecchie. L'assenza di strutture riflettenti permette al suono di svilupparsi senza ostacoli ottenendo una propagazione dell'onda sonora fino all'orecchio esterno priva di rifrazioni e riflessioni. Essendo una tipologia di cuffia molto poco diffusa, ne è presentata un'immagine ad esempio (Figura 2.2)



Figura 2.2 Cuffia extraurale AKG K1000

Nelle cuffie *circumaurali* i cuscinetti circondano completamente l'orecchio esterno racchiudendola. A seconda della pressione che esercitano sulla testa possono sigillare l'orecchio così da isolarlo dai rumori esterni.

Nelle cuffie *sovraurali* invece, i cuscinetti poggiano sulla pinna. Sono di dimensioni inferiori rispetto le circumaurali e offrono una minore attenuazione ai rumori ambientali. (Figura 2.3)



Figura 2.3 Modello di cuffia *sovraurale* (**sinistra**) e *circumaurale* (**destra**)

Gli *auricolari*, o *earbuds*, sono cuffie di ridotte dimensioni che si collocano all'interno della pinna, al di fuori dell'ingresso del canale uditivo. Non hanno significative capacità di isolamento dai rumori esterni e spesso la fedeltà di riproduzione viene compromessa a favore della portabilità e del contenimento dei costi.

Anche gli *in-ear* sono auricolari di ridotte dimensioni che vengono però inseriti direttamente nel canale uditivo. Questo design offre un isolamento dai rumori esterni superiore agli *earbud*.

2.1.3 ISOLAMENTO ACUSTICO

A seconda dell'isolamento acustico offerto rispetto i rumori ambientali si distinguono cuffie aperte, semi aperte e chiuse.

L'isolamento delle cuffie aperte è ridotto se non nullo. Offrono un ascolto naturale e simile all'ascolto tramite loudspeaker.

Le cuffie chiuse hanno invece il padiglione chiuso e alcuni modelli offrono un'attenuazione di 30 dB.

Le cuffie semiaperte, infine, sono una via di mezzo tra aperte e chiuse.

2.2 CRITERIO FEC

Per una corretta riproduzione del materiale audio binaurale sperimentando il senso di esternalizzazione ad essa associato, è di fondamentale importanza che i segnali che raggiungono il timpano dell'ascoltatore siano quanto più simili a ciò che si ascolterebbe in una situazione free field. Essendo le registrazioni effettuate tramite microfoni posti all'entrata del canale uditivo bloccato, si rende necessaria un'analisi su come la diversa impedenza offerta dal canale aperto in fase di ascolto, influenzi la Headphone transfer function (HpTF).

Per agevolare la trattazione si fa uso del modello trasmissivo proposto in [1].

2.2.1 ASCOLTO FREE FIELD

Il modello trasmissivo che descrive l'ascolto free-field è il seguente:

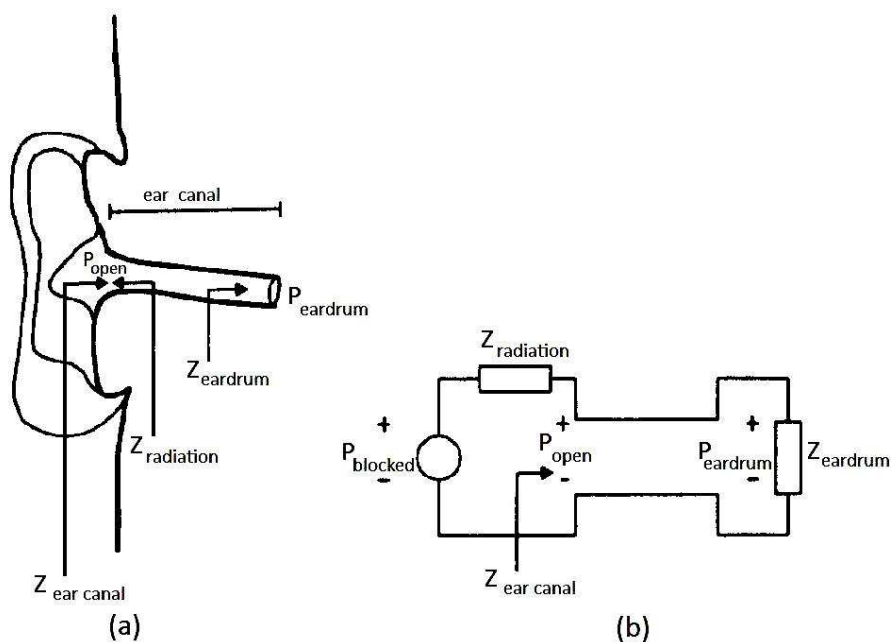


Figura 2.4 Anatomia e modello trasmissivo di ascolto in free-field

L'input del sistema viene rappresentato con un generatore equivalente di Thévenin dove $P_{blocked}$ è la pressione a circuito aperto, e $Z_{radiation}$ l'impedenza vista dal canale uditivo in situazione free-field.

La funzione di trasferimento del suono dal campo diffuso al timpano è espressa dal rapporto tra la pressione misurata sul timpano $P_{eardrum}$ e la pressione sonora che si misurerebbe al centro della testa in assenza del soggetto P_{ref} :

$$\frac{P_{eardrum}(\angle)}{P_{ref}} = \frac{P_{blocked}(\angle)}{P_{ref}} \cdot \frac{P_{open}}{P_{blocked}} \cdot \frac{P_{eardrum}}{P_{open}} \quad (2.1)$$

Si introduce il rapporto misurato all'entrata del canale tra la pressione a canale aperto e canale chiuso. Seguendo la nomenclatura introdotta in Figura 2.4 tale rapporto è espresso da:

$$\frac{P_{open}}{P_{blocked}} = \frac{Z_{earcanal}}{Z_{radiation} + Z_{earcanal}} \quad (2.2)$$

2.2.2 ASCOLTO TRAMITE HEADPHONE

Il modello trasmissivo che descrive l'ascolto tramite headphone è il seguente:

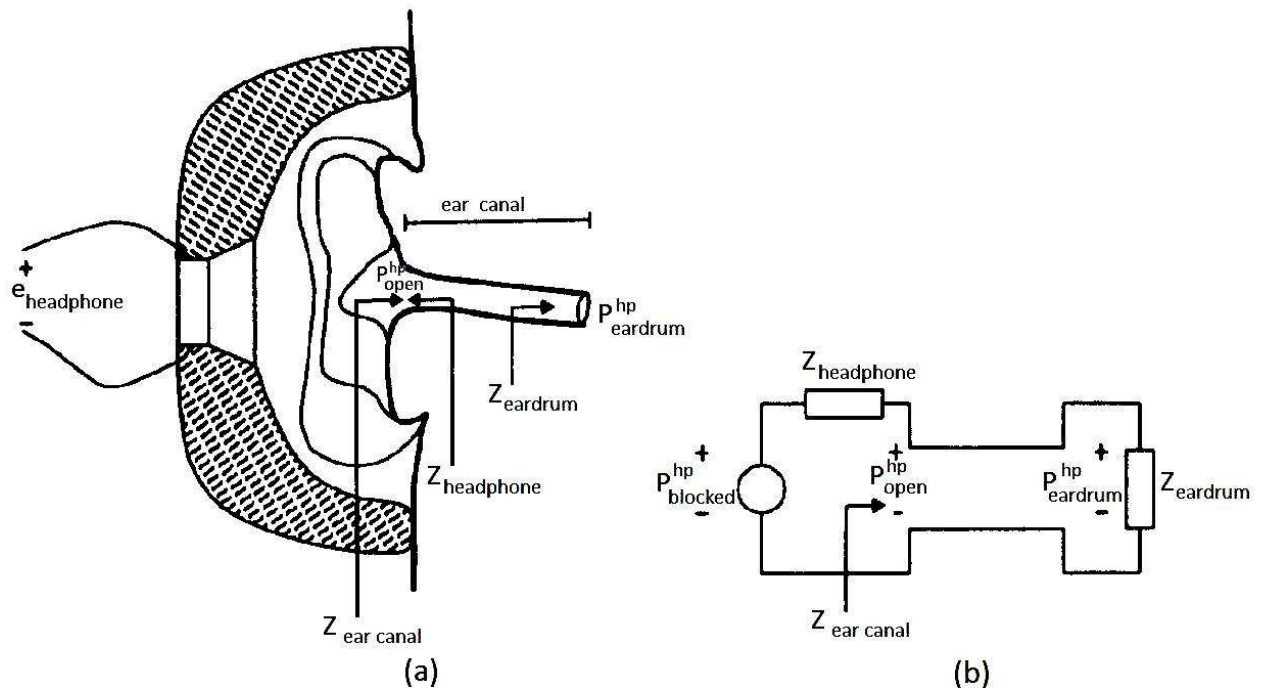


Figura 2.5 Anatomia e modello trasmissivo di ascolto tramite headphone

Anche questa situazione viene rappresentata tramite un generatore equivalente di Thévenin costituito dalla pressione all'entrata del canale uditivo bloccato $P_{blocked}^{hp}$ e l'impedenza $Z_{headphone}$. Si fa notare come le impedenze dei due modelli equivalenti di Thévenin individuati siano diverse.

La funzione di trasferimento che lega la tensione ai capi della cuffia $E_{headphone}$ e la pressione sonora sul timpano $P_{eardrum}^{hp}$ è espressa da:

$$\frac{P_{eardrum}^{hp}}{E_{headphone}} = \frac{P_{blocked}^{hp}}{E_{headphone}} \cdot \frac{P_{open}^{hp}}{P_{blocked}^{hp}} \cdot \frac{P_{eardrum}^{hp}}{P_{open}^{hp}} \quad (2.3)$$

Possiamo ora esprimere il rapporto della pressione all'entrata del canale aperto e chiuso come :

$$\frac{P_{open}^{hp}}{P_{blocked}^{hp}} = \frac{Z_{earcanal}}{Z_{headphone} + Z_{earcanal}} \quad (2.4)$$

Al fine di valutare il mismatch di impedenza all'entrata del canale, associato all'uso della cuffia, Møller [3] introduce la pressure division ratio (PDR) come il rapporto tra (2.2) e (2.4) .

$$PDR = \frac{Z_{earcanal} + Z_{headphone}}{Z_{earcanal} + Z_{radiation}} \quad (2.5)$$

Le cuffie per le quali il rapporto di pressione all'entrata del canale (2.4) è uguale a quello misurato all'entrata del canale in free-field (2.2) sono dette FEC : free air equivalent coupling.

L'uso di queste cuffie in applicazioni di audio binaurali è particolarmente conveniente dal momento che il filtro di equalizzazione utilizzato per linearizzarne la risposta in frequenza, dovrà compensare soltanto la risposta in frequenza della cuffia. Al contrario, per una cuffia che non soddisfa il requisito di FEC, il filtro di equalizzazione dovrà compensare anche il mismatch di impedenza causato dalla stessa (PDR).

Cuffie che presentano un PDR prossimo all'unità sono le cuffie extraurali e alcuni modelli sia open back che closed back come la SENNHEISER HDA200 utilizzata nella ricerca qui presentata [4].

3 MISURAZIONI

I modelli utilizzati nell’esperimento e le loro caratteristiche sono descritti nella tabella 3.1.

Come si può notare ogni tipologia di cuffia è stata rappresentata da almeno un modello, eccezione fatta per le cuffie extraurali (poco diffuse) e a trasduttore di tipo elettrostatico (molto costose vista l’attrezzatura richiesta per il funzionamento).

BRAND	MODEL	TRANCDUCER TYPE	ACOUSTIC COUPLING	EARCUP DESIGN	FREQUENCY RANGE	IMPEDANCE	SENSITIVITY
AKG	240 mk2	DINAMICO	CIRCUMAURALE	APERTA	15-25000 Hz	55 Ohm	91 dB/mW
AKG	271 mk2	DINAMICO	CIRCUMAURALE	CHIUSA	16-28000 Hz	55 Ohm	91 dB/mW
AKG	501	DINAMICO	CIRCUMAURALE	APERTA	16-30000 Hz	120 Ohm	94 dB/mW
SONY	MDR ZX100	DINAMICO	SUPRAURALE	CHIUSA	12-22000 Hz	24 Ohm	100 dB/mW
SENNHEISER	HDA200	DINAMICO	CIRCUMAURALE	CHIUSA	20-20000 Hz	40 Ohm	
BEYERDYNAMIC	DT770M	DINAMICO	CIRCUMAURALE	CHIUSA	5-30000 Hz	80 Ohm	105 dB/mW
APPLE	EARBUDS	DINAMICO	EARBUD	/	20-20000 Hz	23 Ohm	109 dB/mW

Tabella 3.1 Modelli e caratteristiche delle cuffie utilizzate

3.1 POSIZIONE DI ACQUISIZIONE

Quando un segnale audio è registrato, sia con un manichino che con una persona su cui sono posizionati dei microfoni, il panorama sonoro viene trasformato in due segnali, uno per ogni orecchio. E’ quindi un prerequisito di fondamentale importanza che i due segnali acquisiti contengano la massima informazione spaziale possibile, e ugualmente che essa non venga alterata dal processo di registrazione.

Questa problematica viene più volte affrontata nella letteratura [3] [5] [6] , e risulta per lo più condiviso che esistono diversi posti all’interno del canale uditivo, oltre ovviamente al timpano, dove l’informazione spaziale è disponibile.

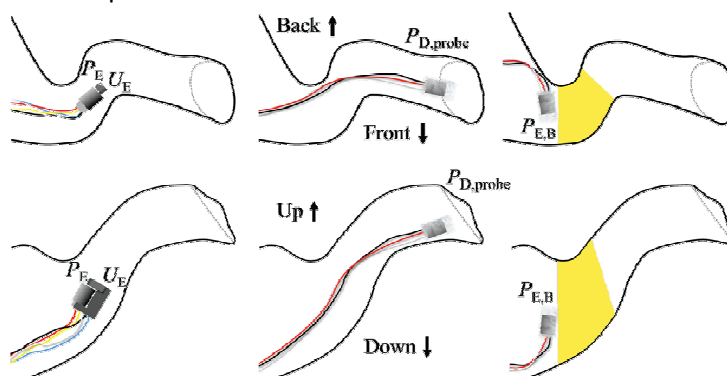


Figura 3.1 Nell’immagine sono presentati le 3 principali tecniche di misura. Sinistra: misura all’entrata del canale aperto. Centrale : misura eseguita in prossimità del timpano. Destra: misura all’entrata del canale bloccato P = pressione U = velocità. L’immagine è originariamente presentata in [2]

In [3] [5] [6] viene dimostrato come sia possibile acquisire il segnale all'entrata del canale uditivo e per di più a canale bloccato, senza che l'informazione spaziale sia eccessivamente compromessa. Si noti che la perdita dell'apporto del canale uditivo sull'esternalizzazione del suono, è compensata dal punto di vista pratico dalla possibilità di condurre le registrazioni più agevolmente con microfoni più grandi, e soprattutto dalla possibilità di utilizzare HATS senza il canale uditivo.

In [3] viene inoltre dimostrato come la registrazione all'uscita del canale aperto consenta registrazioni caratterizzate da un maggior grado di generalità. Come si può vedere in figura 3.2, risulta che la deviazione standard all'entrata del canale aperto è maggiore rispetto quella misurata al timpano, o all'entrata del canale bloccato. Tra entrata del canale bloccato e timpano, le differenze sono minori ma comunque a favore del primo.

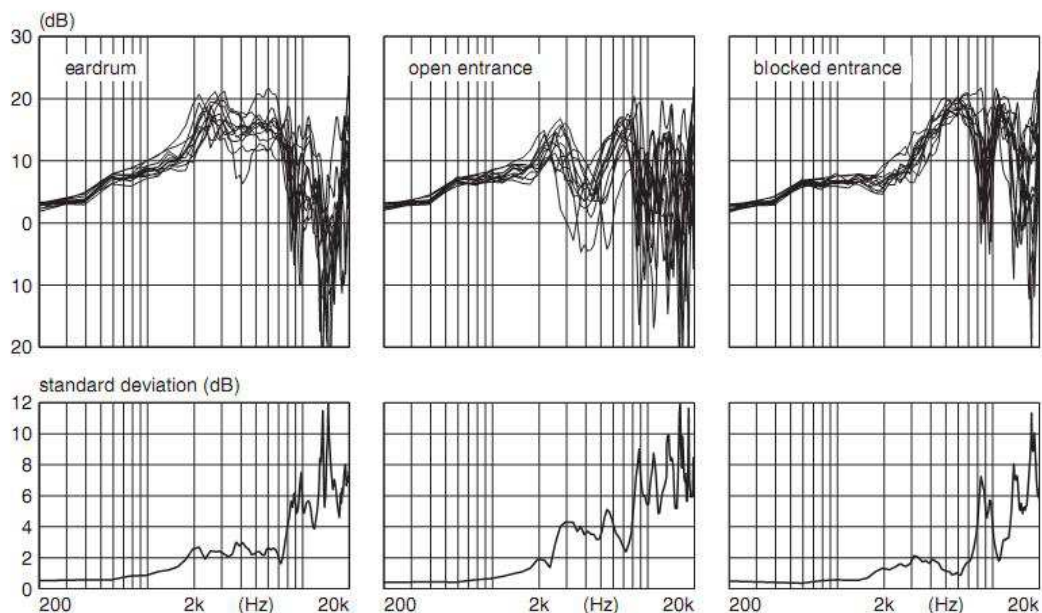


Figura 3.2 HRTFs misurata all'orecchio destro per 12 soggetti diversi. La misura è stata effettuata in prossimità del timpano (**sinistra**), all'entrata del canale uditivo aperto (**centro**) e all'entrata del canale bloccato (**destra**)

Possiamo quindi affermare che le HRTF misurate all'entrata del canale bloccato, sebbene comportino una parziale perdita di informazione riferita al contributo del canale uditivo, sono quelle che presentano una minore variabilità e quindi un maggiore range di utilizzo. Risulta ovvero più generica e meno influenzata dal soggetto (HATS o persona fisica) utilizzato per la misura.

3.2 CATENA DI ACQUISIZIONE

Il sistema di riproduzione e di acquisizione è stato organizzato per introdurre la minore distorsione possibile, così da indagare le sole caratteristiche del sistema in studio prescindendo dalle attrezzature impiegate per lo scopo.

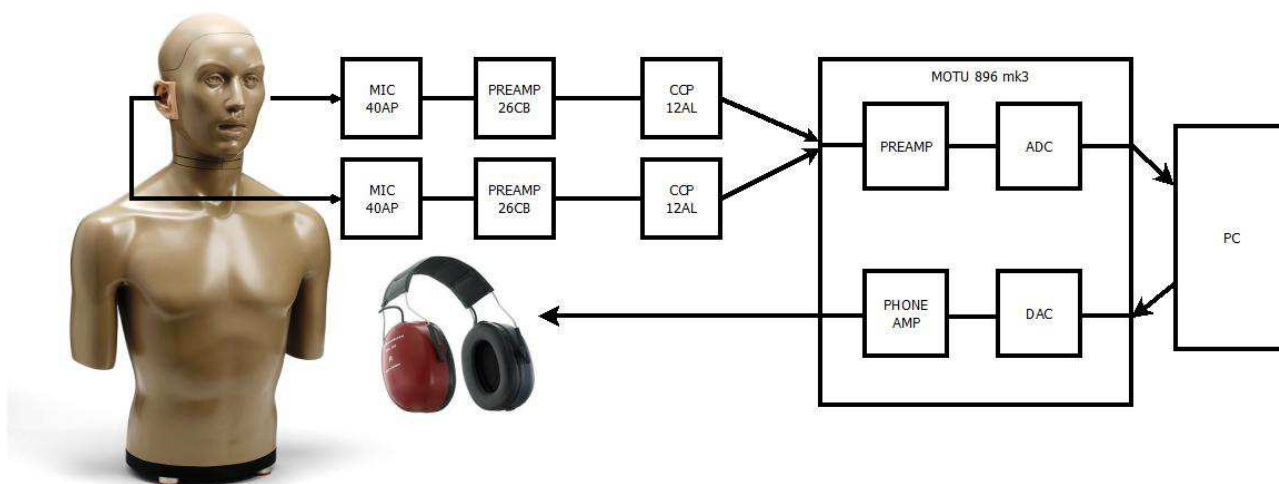


Figura 3.3 Diagramma a blocchi della catena di acquisizione

Lo stimolo sonoro (sinesweep esponenziale figura 3.4) è stato riprodotto tramite una DAW (Digital Audio Workstation) connessa ad una scheda audio Motu 896mk3¹. Al fine di introdurre il minor errore sistematico nell'acquisizione delle HpIR, abbiamo prima confrontato la risposta in frequenza del canale di acquisizione delle 2 interfacce presenti in laboratorio RME fireface800 e Motu896mk3, rilevando una minor distorsione nella seconda. (leggasi linearità in frequenza, larghezza di banda, THD)

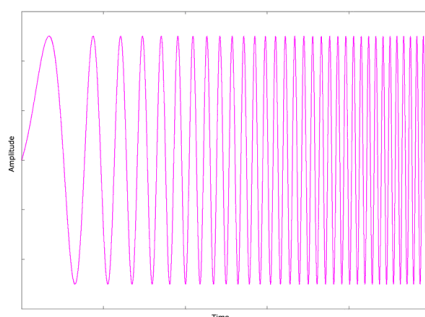


Figura 3.4 Sinesweep logaritmico

Le registrazioni sono state effettuate ad una sample rate di 48 kHz e word length di 24 bit. La prima, in accordo al teorema del campionamento di Nyquist-Shannon, ci permette di convertire correttamente frequenze appartenenti alla banda audio (20-20000 Hz) mentre la bit depth scelta ci garantisce un SNR di 146 dB.

¹ <http://www.motu.com/products/motuaudio/896mk3>

Il manichino Kemar 45BM ², cuore dell'esperimento, prodotto dalla G.R.A.S. è una simulazione di testa e torso (HATS) basato sulle caratteristiche antropomorfe medie di uomini e donne (worldwide). Dotato di 2 microfoni posti all'entrata dei canali uditivi, è utilizzato per registrare le alterazioni dei fronti d'onda a causa delle riflessioni e rifrazioni sul suo corpo antropomorfo. Con l'obiettivo di approssimare al meglio le HRTF di un soggetto umano, e quindi anche del contributo dell'orecchio esterno, il manichino è munito di 2 orecchie artificiali. E' così possibile condurre esperimenti ripetibili e di valenza generica [3].

I microfoni utilizzati G.R.A.S. 40AP ³ sono microfoni a condensatore dalle ridotte dimensioni (0.5 inch) per generiche applicazioni audio. Sono caratterizzati da un diagramma polare omnidirezionale, un'elevata sensibilità di 50 mV/Pa e un'estesa risposta in frequenza di ± 2.0 dB tra 3.15 Hz - 10 kHz e ± 1.0 dB tra 12.5 Hz - 7.5 kHz. (Figura 3.5)

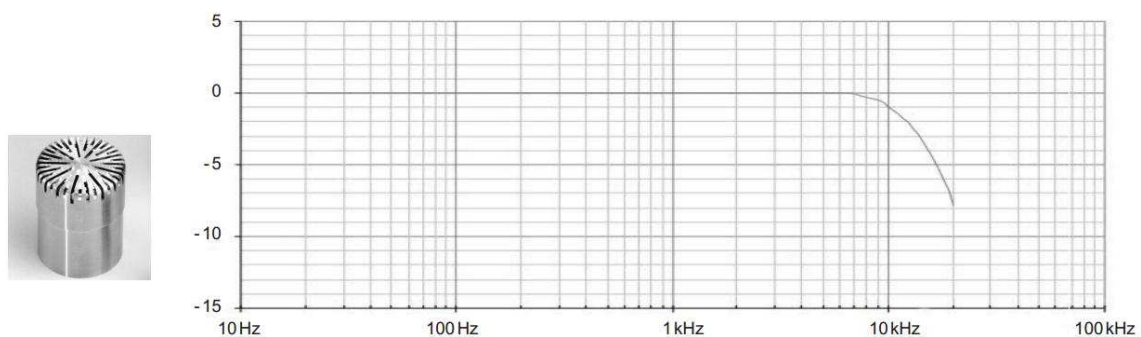


Figura 3.5 Microfono G.R.A.S 40AP e relativa risposta in frequenza

Anche il preamplificatore G.R.A.S. 26CB ⁴ utilizzato per adattare l'impedenza tra capsula microfonica e scheda audio è caratterizzato da un'ampia banda passante (2 Hz - 200 kHz a ± 0.2 dB) garantendo così una distorsione minima, nella pratica trascurabile. Il preamplificatore grazie ai valori di impedenza di ingresso $20\text{ G}\Omega$ e impedenza di uscita $<50\ \Omega$ svolge un ruolo chiave per evitare effetti di carico sul microfono che ne modificherebbero la risposta e minimizza inoltre l'influenza di capacità parassite ed interferenze. Infine si segnala che il preamplificatore introduce un'attenuazione sul segnale di -0.25 dB che sarà poi compensata tramite la sezione di preamplificazione pre ADC nella scheda Motu.

I preamplificatori utilizzati necessitano per un corretto funzionamento di una corrente costante di alimentazione compresa tra i 2mA e 20 mA (valori tipici sono di 4mA). Considerata la banda passante del modulo utilizzato a tal scopo, il G.R.A.S. 12AL ⁵, di 1Hz-200kHz a -3dB la sua distorsione in frequenza è minima e quindi trascurabile. Il modulo inoltre può essere alimentato tramite alimentatore esterno DC o batterie.

Durante gli esperimenti è stata utilizzata l'alimentazione a batterie perché al momento della messa in funzione il segnale acquisito tramite la catena risultava affetto da un disturbo costante e non correlato al segnale riprodotto dalle cuffie. Analizzando lo spettro in frequenza del rumore si è

² <http://www.gras.dk/00012/00330/>

³ <http://www.gras.dk/00012/00013/00029/00119/>

⁴ <http://www.gras.dk/00012/00045/00046/00048/>

⁵ <http://www.gras.dk/00012/00054/00089/00090/>

notato come esso presentasse importanti picchi ai multipli di 50 Hz. Individuando quindi nell'alimentatore la sorgente del disturbo abbiamo continuato le registrazioni alimentato il modulo tramite batterie alcaline.

Figura 3.6 Spettro del segnale acquisito alimentando GRAS 12AL tramite tensione di rete (**sinistra**) e tramite batterie alcaline (**destra**)

Continuando lungo la catena di acquisizione audio si sono utilizzati i preamplificatori che equipaggiano l'interfaccia audio Motu 896mk3 per aumentare l'ampiezza del segnale proveniente dai microfoni di 20 dB così da ottimizzare la SNR del segnale e ottenere una quindi una conversione analogico/digitale più accurata.

Il livello di ingresso del segnale utilizzato per la calibrazione (test tone a 500 Hz) è stato tenuto a circa -15 dBFS.

3.3 CATENA DI RIPRODUZIONE

La catena di riproduzione dello stimolo sonoro è stata organizzata nel modo seguente: Dopo esserci assicurati della bontà dell'uscita cuffie della scheda audio (grafico) abbiamo utilizzato quest'ultima per inviare lo stimolo alle diverse cuffie tramite un'apposita prolunga audio all'interno della cabina silente.

Tutte le registrazioni sono infatti state eseguite all'interno della cabina silente PUMA PRO45⁶ per garantire un isolamento dai rumori ambientali che al contrario avrebbero falsato le misurazioni.

In accordo allo standard IEC 60268-7 si è scelto di riprodurre lo sinesweep a un volume di 94 dB SPL e dal momento che la sensibilità (dB/mW) delle cuffie è diversa da modello a modello si è resa necessaria una calibrazione individuale. Si è proceduto come segue:



Figura 3.7 Sistema di calibrazione

⁶ http://www.pumasrl.it/images/cataloghi/Data_sheet_PRO45S_medical_ENG.pdf

Tramite il pc si è inviato alla cuffia il test tone a 500 Hz.

La pressione sonora in uscita dalla cuffia misurata tramite un fonometro è stata regolata tramite il guadagno del preamplificatore per le cuffie fino a leggere nel fonometro 94 dB SPL.

Per riprodurre le condizioni di utilizzo delle cuffie durante l'esperimento, si è realizzato un sostegno (Figura 3.8) per accoppiare acusticamente cuffie e fonometro evitando così che una parte della pressione prodotta andasse dispersa falsando così la misura.



Figura 3.8 Sostegno utilizzato per la calibrazione

Così facendo ad ogni cuffia si è associato il guadagno (Tabella 3.2) necessario ad ottenere in uscita sempre la stessa pressione sonora di 94dB SPL.

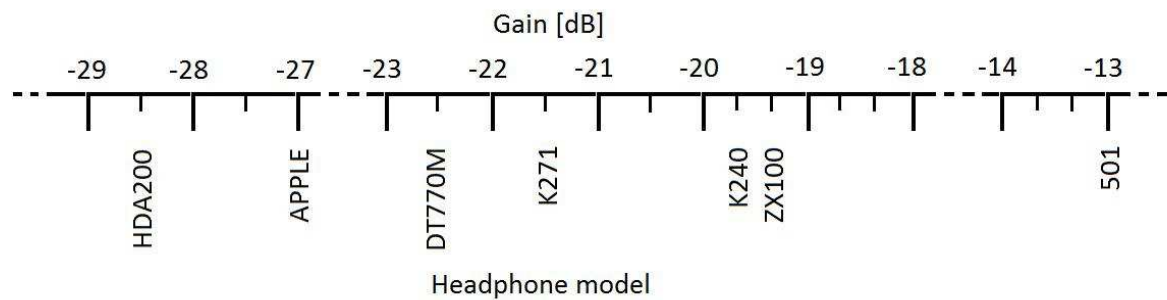


Tabella 3.2 Sono rappresentati i guadagni di ogni cuffia, per ottenere un test tone a 94 dB SPL

3.4 REGISTRAZIONE HpIR

Una volta completata la fase preliminare di set up e calibrazione della catena audio si è proceduto con la registrazione delle HpIR.

Il procedimento utilizzato è il seguente:

- Per ogni cuffia si è impostato il guadagno letto in fase di calibrazione per ottenere in uscita un livello di 94 dBSPL.
- Per eccitare il sistema, tramite DAW si è inviato alle cuffie uno sinesweep logaritmico di frequenze da 20 a 20000 Hz.
- I segnali acquisiti dai due microfoni del manichino Kemar, sono digitalizzati tramite l'interfaccia MOTU e memorizzati in una traccia audio stereo con codifica PCM.

Al fine di indagare quanto le Headphone transfer function (HpTF) siano robuste rispetto il posizionamento della cuffia sulla testa, per ciascun modello sono state effettuate 10 registrazioni intervallate da riposizionamenti. La problematica è discussa in [6]. In [7] è mostrato come un'equalizzazione per compensare la HpTF registrata per una sola misura, quindi senza riposizionamenti, conduca a risultati peggiori di una risposta non compensata.

Tutte le misurazioni sono state effettuate all'interno della cabina silente PUMA PRO45 per garantire una sufficiente attenuazione dai rumori ambientali che avrebbero altrimenti falsato i risultati ottenuti (in particolare modo per le cuffie aperte).

Dopo aver acquisito tutte le HpIR RAW, si è proceduto alla loro memorizzazione su formato standard come proposto in [8].

Si fa presente infine, che dopo aver prodotto le HpTF di tutte le cuffie si è notato come il guadagno del canale destro risultava mediamente inferiore a quello sinistro di 2dB. Al fine di individuarne la causa si è ripercorsa la catena di acquisizione dai microfoni del manichino fino al sezione di preamplificazione della scheda audio. La causa del mismatch di amplificazione è così stata individuata nel diverso guadagno dei due preamplificatori G.R.A.S. 26CB. Per una corretta lettura dei risultati si è quindi applicato al canale destro un offset di + 2 dB.

4 RISULTATI

Secondo il procedimento descritto sono state raccolte 10 risposte al logaritmico sinesweep per ognuna delle 7 cuffie.

Si è innanzitutto provveduto alla memorizzazione secondo il database proposto in [8]; successivamente per l'elaborazione e la visualizzazione dei risultati sono stati sviluppati degli script nell'ambiente Matlab®. (Script in Appendice)

4.1 HpTF

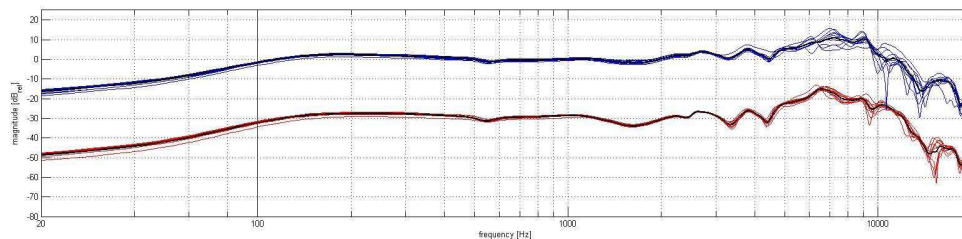
Nella prima figura sono riportate le funzioni di trasferimento di ogni cuffia.

Con una linea sottile sono visualizzate le HpTF relative ad ogni posizionamento, e un tratto più spesso ne evidenzia la media.

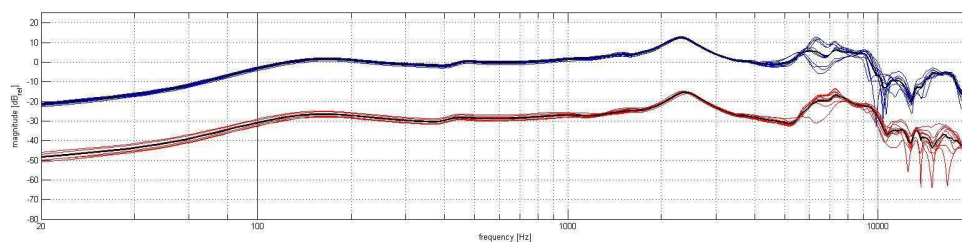
Tutte le HpTF presentano un roll off alle basse frequenze a partire dai 100-200 Hz.

Al di sopra dei 3kHz la funzione di trasferimento di tutte le cuffie, presenta dei peak e dei notch. L'entità di tali distorsioni non è uguale in tutte le cuffie, bensì alcune si avvicinano meglio di altre ad una risposta piatta.

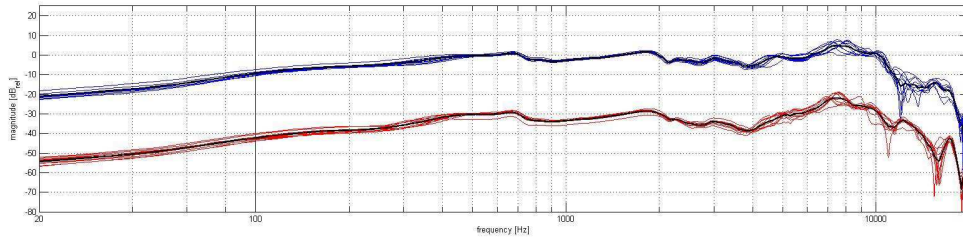
K240



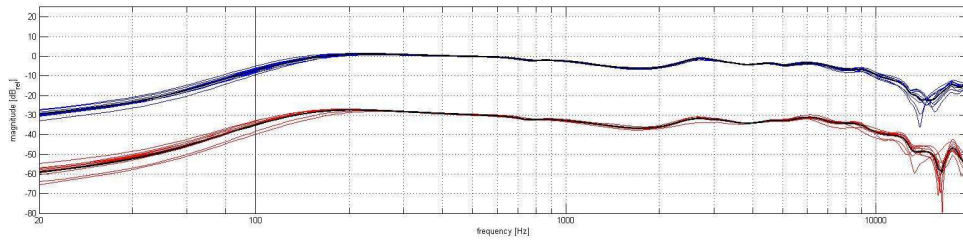
ZX100



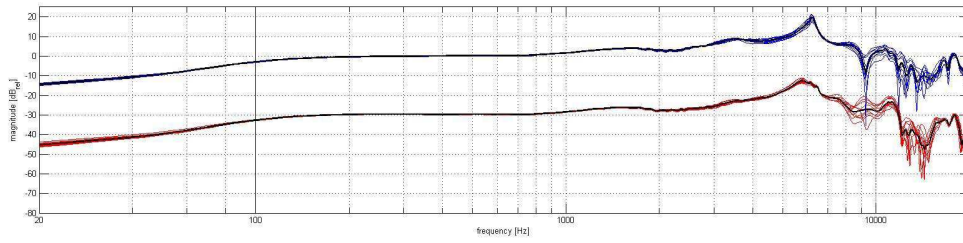
K271



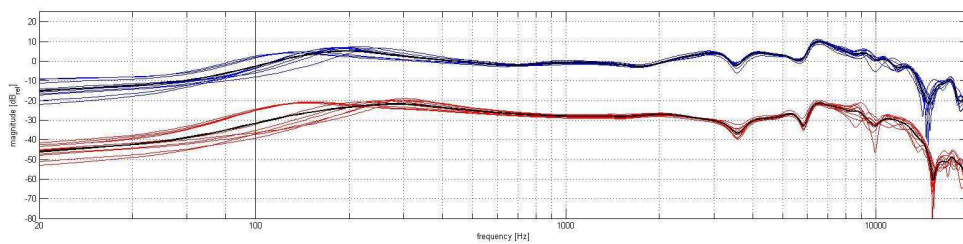
HDA 200



501



DT770



EARBUD

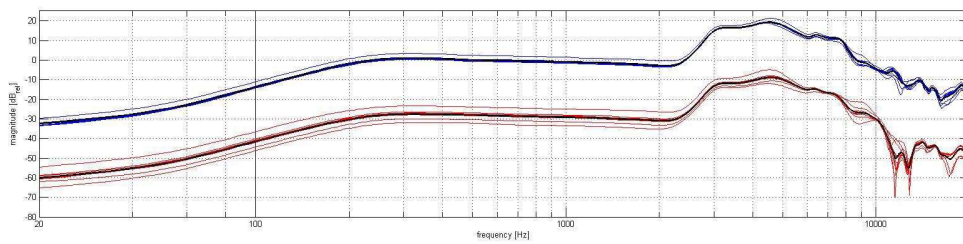


Figure 4.1 HpTFs. Per convenienza di esposizione il canale destro è traslato di -30dB.

4.2 HpTF STANDARD DEVIATION

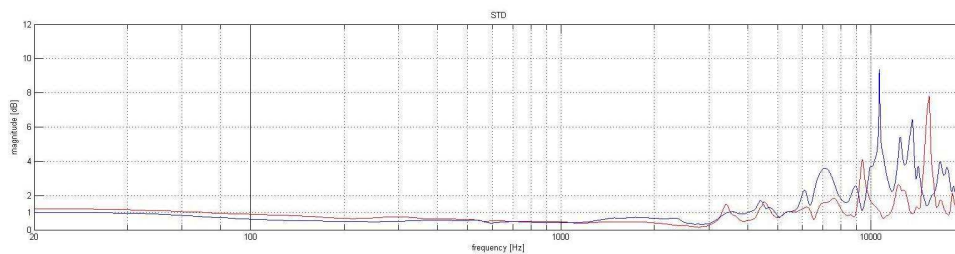
Un altro parametro analizzato, è la robustezza della risposta in frequenza della cuffia rispetto ai riposizionamenti sulla testa.

Come si è già potuto notare dai precedenti grafici, le HpTF presentano una variabilità particolarmente pronunciata in alcuni range di frequenze.

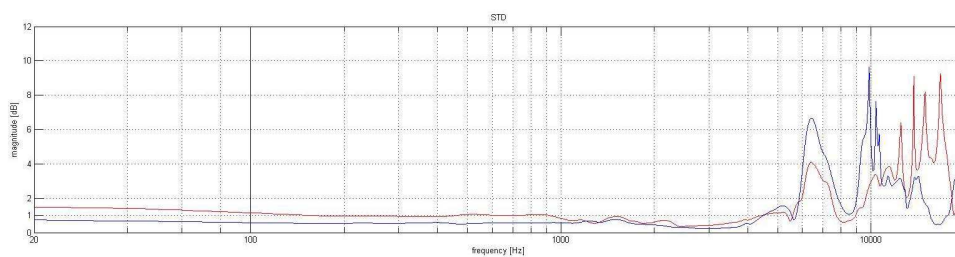
Per una lettura più agevole, i grafici che seguono rappresentano la deviazione standard sui 10 riposizionamenti per ogni cuffia.

Si nota come la variabilità delle HpRT associata ai riposizionamenti, è limitata entro alcuni dB per le frequenze inferiori a 7kHz. Per le frequenze superiori al contrario tale variabilità aumenta considerevolmente indipendentemente dalla tipologia di cuffia.

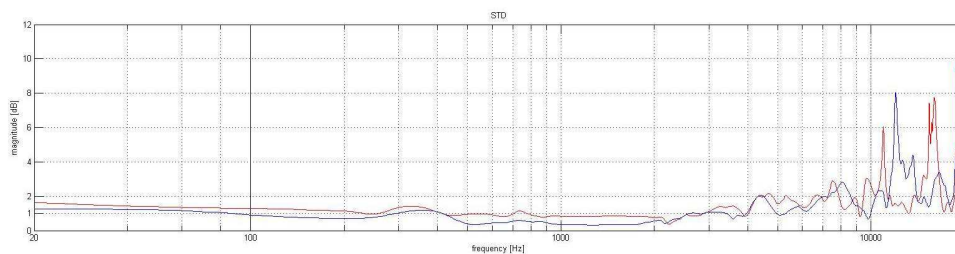
K240



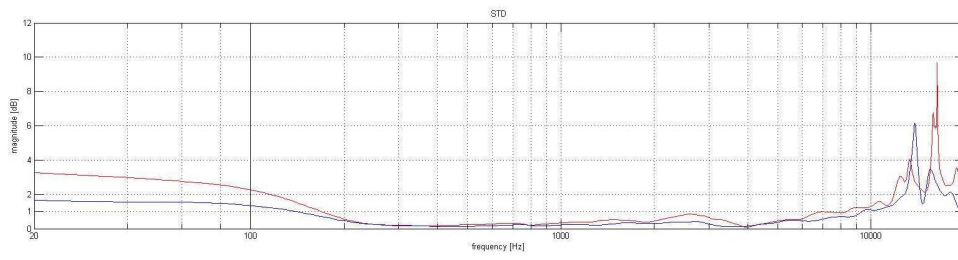
ZX100



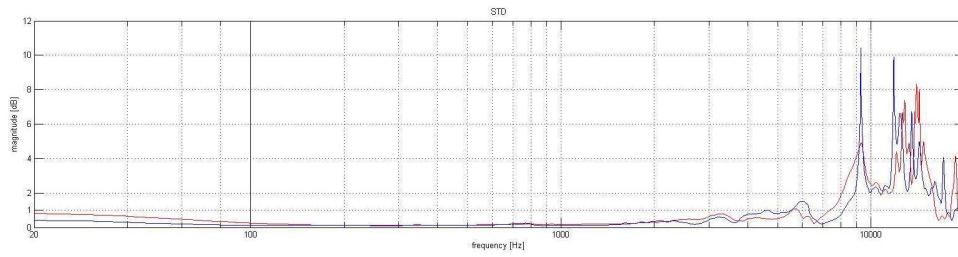
K271



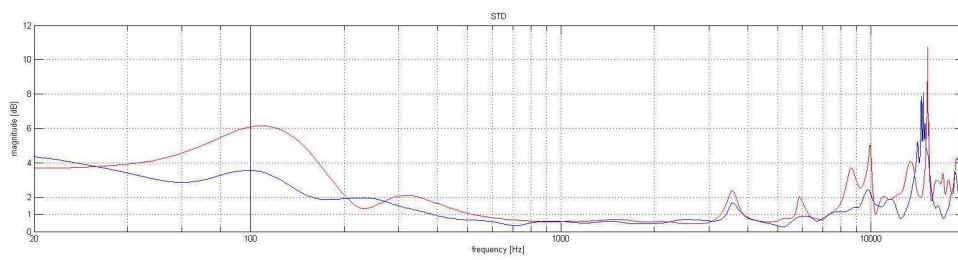
HDA 200



501



DT770



EARBUD

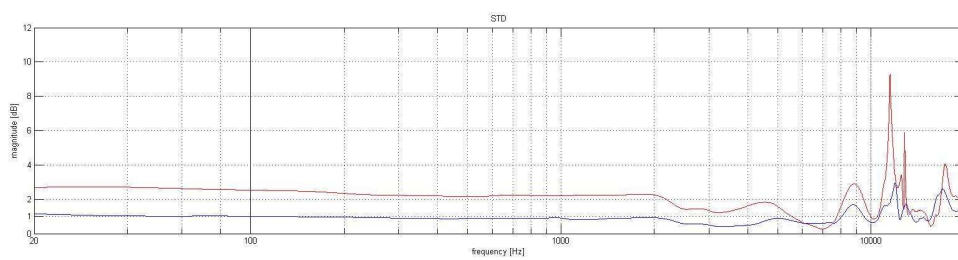
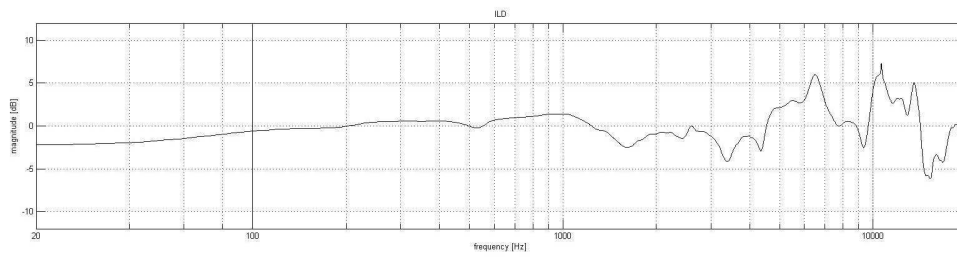


Figure 4.2 Deviazione standard delle HpTF legate a risposionamenti

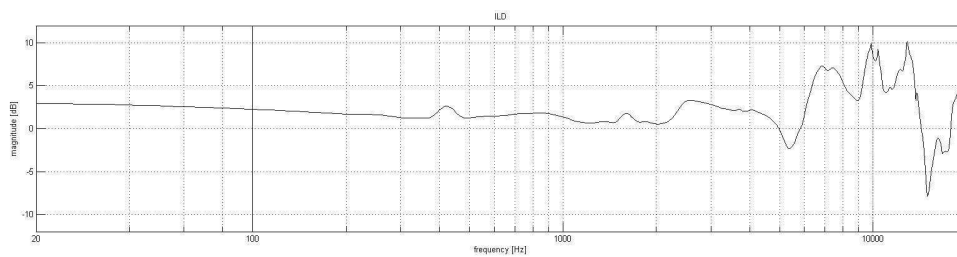
4.3 ILD

Come anticipato nel primo capitolo, si è voluto indagare la distorsione introdotta dalle cuffie sugli indicatori utilizzati dal cervello per la ricostruzione dello scenario audio 3D. I grafici qui riportati presentano la ILD tra la media delle HpTF destra e sinistra.

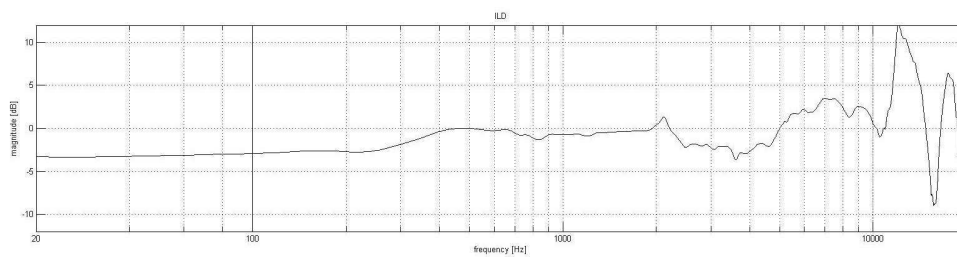
K240



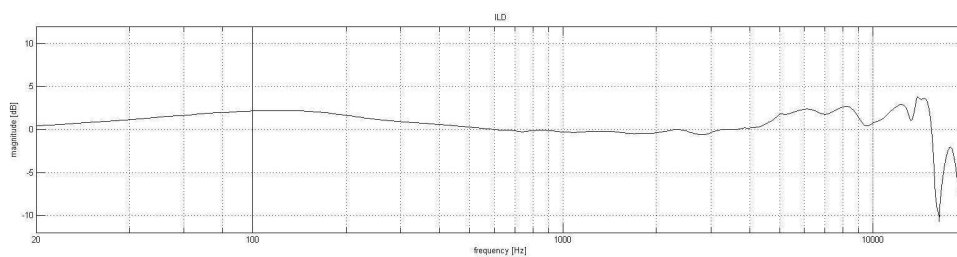
ZX100



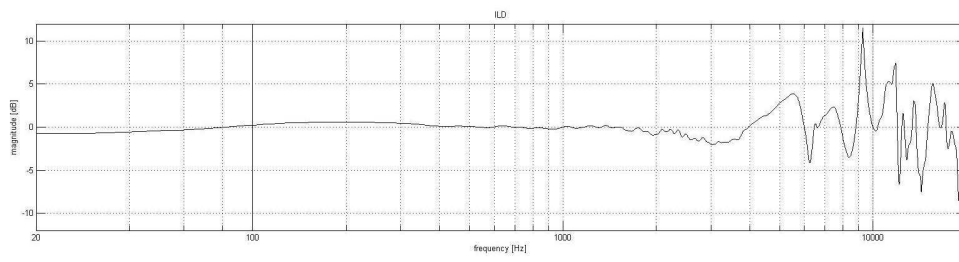
K271



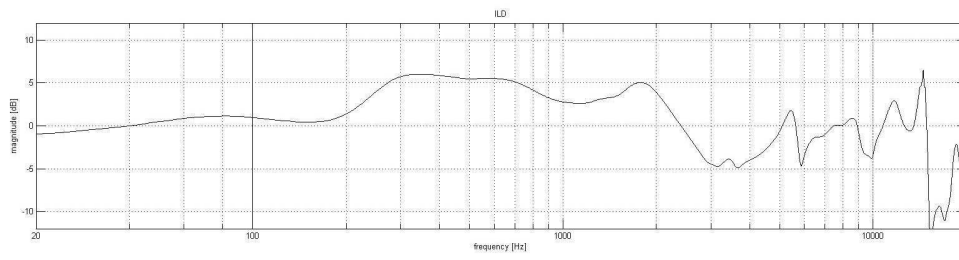
HDA 200



501



DT770



EARBUD

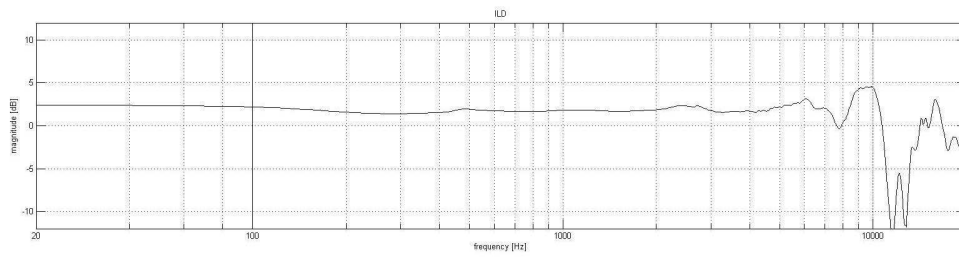
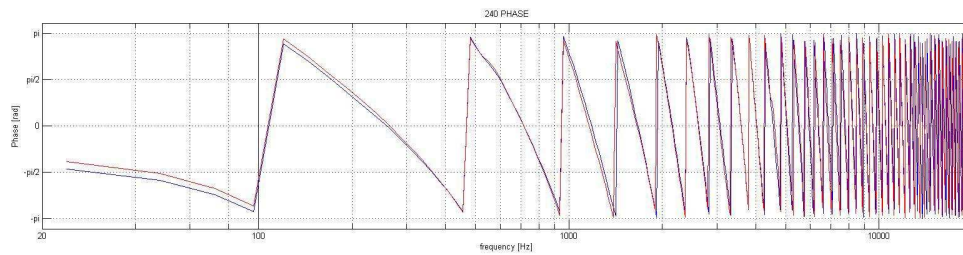


Figure 4.3 ILD tra la media delle HpTF destra e sinistra

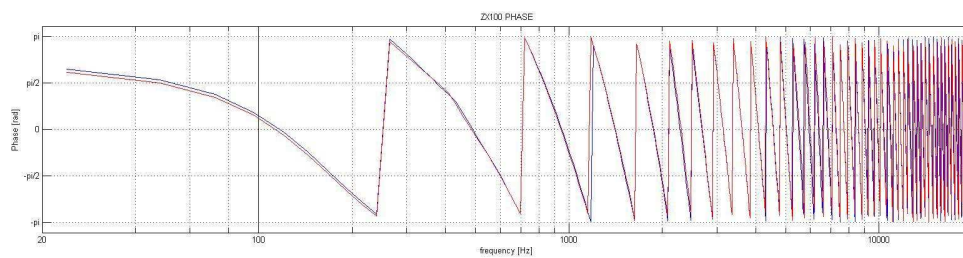
4.4 PHASE

Si è voluto indagare la distorsione di fase introdotta dalle cuffie sulla riproduzione del segnale. Ogni scostamento della fase dall'andamento lineare con la frequenza, rappresenta una distorsione a causa della quale, componenti del segnale a frequenza diversa subiscono ritardi diversi.

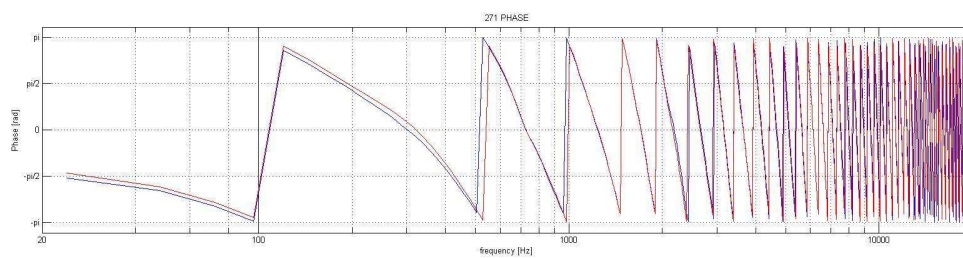
K240



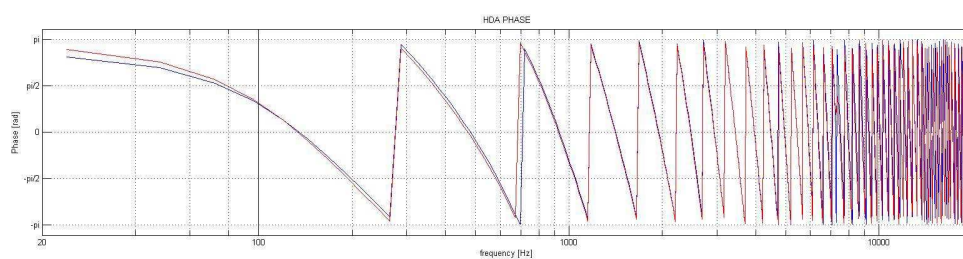
ZX100



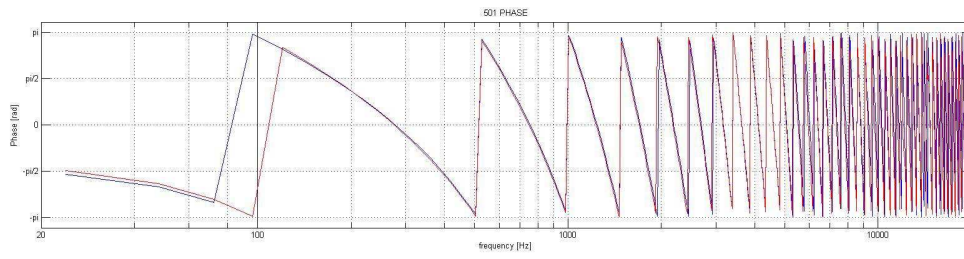
K271



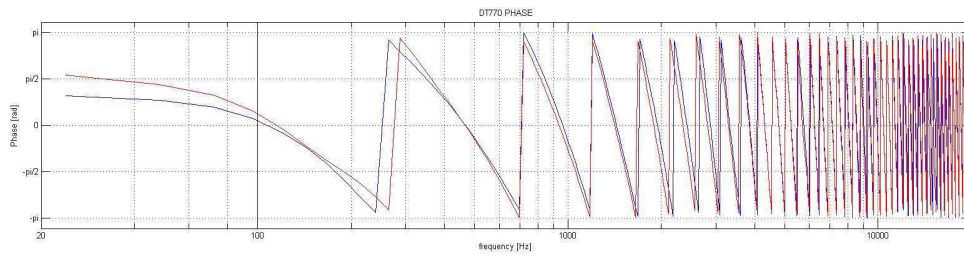
HDA 200



501



DT770



EARBUD

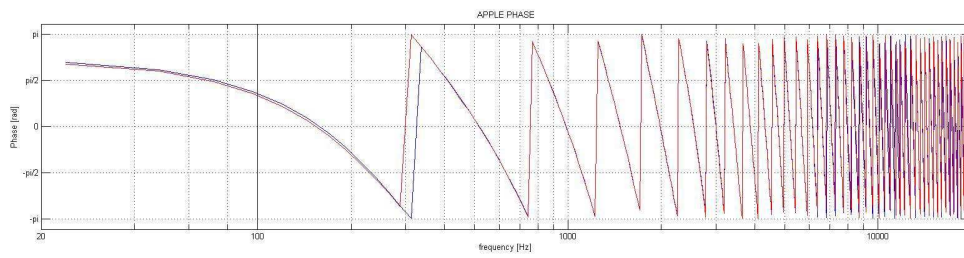


GRAFICO 4 PHASE

Figure 4.4 Plot della fase

4.5 DIFFERENZE DI PHASE INTERAURALI

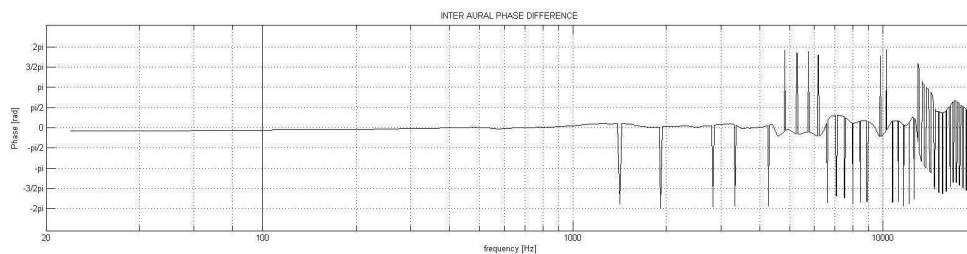
Sebbene l'orecchio non sia sensibile a distorsioni di fase assolute sul singolo segnale, se ne vuole proporre un'ulteriore elaborazione.

In un sistema la distorsione di fase è indice del ritardo tra il segnale di uscita rispetto quello in ingresso.

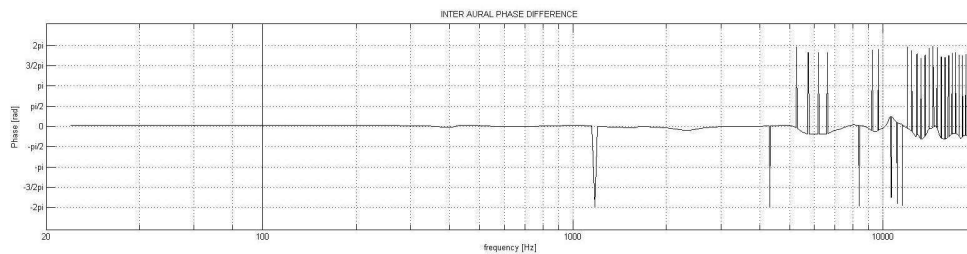
Negli esperimenti condotti l'input del sistema è rappresentato da uno sinesweep monofonico: entrambi gli altoparlanti delle cuffie sono quindi interessati dallo stesso segnale.

Per tale motivo differenze nella distorsione di fase tra il canale destro e sinistro sono indice di una diversa velocità di propagazione tra i due canali per determinate frequenze.

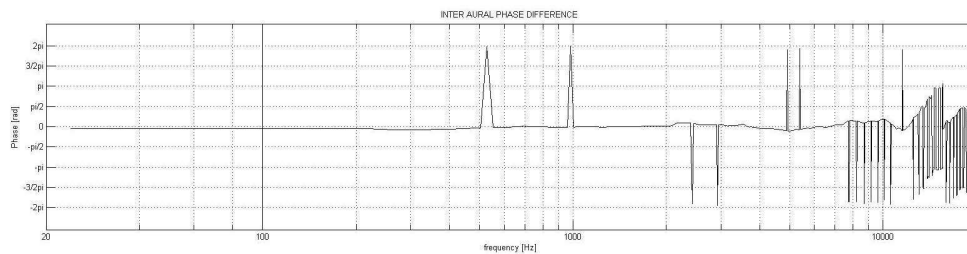
K240



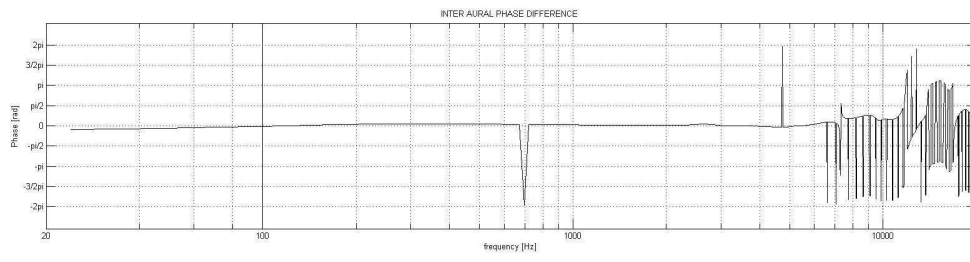
ZX100



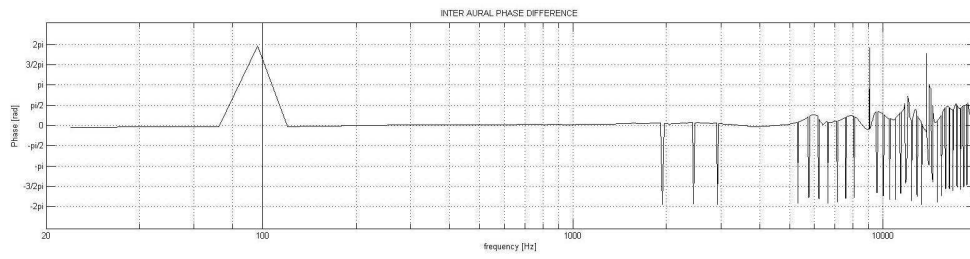
K271



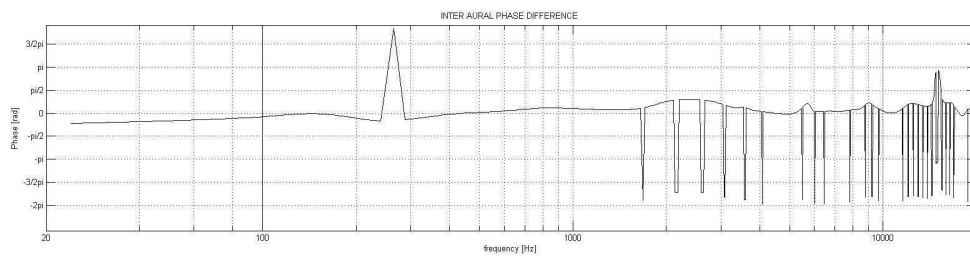
HDA200



501



DT770



EARBUD

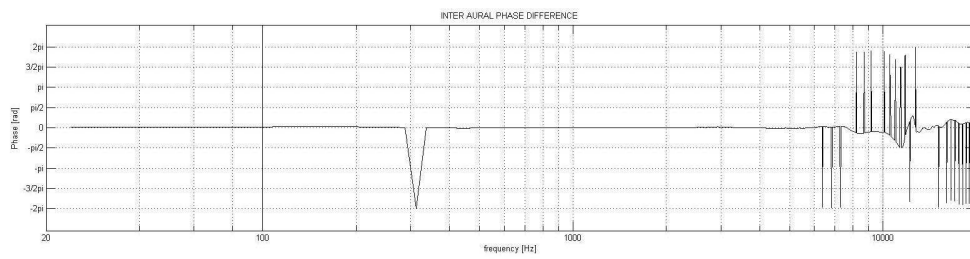


Figure 4.5 Plot dello sfasamento relativo tra segnale destro e sinistro

5 CONCLUSIONI

Le HpTFs di 7 diversi modelli di cuffie sono state acquisite tramite manichino KEMAR per indagarne il contributo nel veicolare segnali audio binaurali.

Si è evidenziato come la riproduzione tramite headphones introduca distorsioni di diversa natura:

Tutte le cuffie analizzate presentano un comportamento in frequenza approssimabile con un filtro passa banda, che si discosta dall'idealità di una risposta a guadagno costante (flat) nella banda audio. Le rispettive frequenze di taglio sono attorno ai 150Hz e 9kHz.

Associata ai riposizionamenti delle cuffie sulla testa, si rileva una variabilità della HpTF. Tale deviazione da una risposta media, è particolarmente pronunciata per le frequenze superiori ai 7 kHz ed è presente indipendentemente dalla tipologia di cuffia

L'ILD si discosta per tutte le cuffie dall'idealità di 0dB presentando un andamento per lo più costante e compreso tra ± 4 dB per le frequenze inferiori a 3kHz, mentre aumenta considerevolmente per le frequenze superiori (anche ± 12 dB).

Queste distorsioni concorrono, assieme ad altre, a un'errata riproduzione del segnale audio binaurale. Per ripristinare la corretta esternalizzazione della sorgente sonora si rende quindi necessario un intervento di compensazione della HpTF.

Sviluppi futuri includono l'uso delle HpTF registrate per testare diversi metodi di equalizzazione[9].

Essendo però l'audio binaurale un'esperienza altamente soggettiva, quanto sono accettabili le generalizzazioni introdotte dal manichino KEMAR in fase di registrazione delle HpIR?

In [10] è evidenziato come le variazioni nelle HRTFs possono essere di entità paragonabile a quelle introdotte dalle HpTF. Risulta perciò evidente che un'equalizzazione della HpTF garantirà una corretta riproduzione del materiale audio solo per il soggetto sul quale sono state acquisite le HpIR. Inoltre, come evidenziato in questo studio e in [10] e [11], le differenze tra HpTF di modelli di cuffie diverse possono essere molto significative (± 10 dB).

Si conclude perciò, che un'esatta riproduzione binaurale è possibile solo per il binomio soggetto-cuffia utilizzato l'acquisizione delle HpIR.

A causa delle difficoltà pratiche nel seguire questa strada, una direzione di ricerca che trae origine da questo studio preliminare, verte allo sviluppo di modelli basati su caratteristiche fisiche anatomiche del soggetto e della cuffia

Alcuni scenari di ricerca, prevedono l'uso di dispositivi quali head-tracker per rilevare i movimenti del soggetto e modificarne di conseguenza la mappa sonora riprodotta.

APPENDICE

SCRIPT MATLAB

1. Function per il calcolo della HpTF dalla risposta all'impulso del sistema

```
1. function [f,frGi1,frGi2]= rep_plot(y)
2. % Evaluation of HpTF for a single
3. % measurement.
4. % input
5. %     y=sine sweep response
6. % output:
7. %     f=sample frequency
8. %     frGi1=HpTF of left channel
9. %     frGi2=HpTF of right channel
10. %
11.
12. [u, Fs_u] = wavread('sweep1.wav');
13.
14. delay = 7290;
15. length_input_signal = 47700;
16.
17. Ts=1/Fs_u; % sampling time [s]
18.
19. uf=u(1:length_input_signal);
20. yf=y;
21.
22. % Windos
23. yf1 = yf(delay:delay-1+length_input_signal,1);
24. yf2 = yf(delay:delay-1+length_input_signal,2);
25.
26. % define data for sys id
27. z1 = iddata(yf1, uf, Ts);
28. z2 = iddata(yf2, uf, Ts);
29.
30. %% Spectral analysis ETFE
31.
32. % freq resp id using ETFE
33. g1 = etfe(z1, 1000, 5000);
34. g2 = etfe(z2, 1000, 5000);
35.
36. w = g1.Frequency;
37. f = w/2/pi;
38.
39. frGi1 = squeeze(g1.ResponseData);
40. frGi2 = squeeze(g2.ResponseData);
41.
42. % Plot etfe results
43. %
44. % +13.4dB normalize HpIR to recording level @ 500Hz
45. % + 2.0dB match gap between L/R preamp output level
46. % -30.0dB shift left HpIR for reading convenience
47. %
48.
49. semilogx(f, 20*log10(abs(frGi1))+18,'b')
50. hold on
51. semilogx(f, 20*log10(abs(frGi2))+2-30+18,'r')
52. ylabel('magnitudo [dB_{rel}]');
53. xlabel('frequency [Hz]')
54. x=[20,100, 1000, 10000]; % define the x values where you want to have a tick
55. set(gca,'XTick',x); % Apply the ticks to the current axes
56. % Define the tick labels based on the user-defined format
57. set(gca,'XTickLabel', arrayfun(@(v) sprintf('%d',v), x, 'UniformOutput', false) );
58.
59. grid on;
60. title('BEYERDINAMIC DT770M')
61. xlim([20 20000])
62. ylim([-80 25])
63.
64. frGi1 = 20*log10(abs(frGi1))+18;
65. frGi2 = 20*log10(abs(frGi2))+2 +18;
```


2. Metodo main per il plot delle 10 HpTF , calcolo e plot di Standard deviation e ILD

```
1. clc;
2. close all;
3. clear all;
4. %frGil Left
5. %frGi2 Right
6. figure(1)
7. subplot(3,1,1)
8. %single reposition plots
9. [y, Fs_y] = wavread('\271\rep1.wav');
10. [f,frGil(1,:),frGi2(1,:)] = rep_plot(y);
11. [y, Fs_y] = wavread('\271\rep2.wav');
12. [f,frGil(2,:),frGi2(2,:)] = rep_plot(y);
13. [y, Fs_y] = wavread('\271\rep3.wav');
14. [f,frGil(3,:),frGi2(3,:)] = rep_plot(y);
15. [y, Fs_y] = wavread('\271\rep4.wav');
16. [f,frGil(4,:),frGi2(4,:)] = rep_plot(y);
17. [y, Fs_y] = wavread('\271\rep5.wav');
18. [f,frGil(5,:),frGi2(5,:)] = rep_plot(y);
19. [y, Fs_y] = wavread('\271\rep6.wav');
20. [f,frGil(6,:),frGi2(6,:)] = rep_plot(y);
21. [y, Fs_y] = wavread('\271\rep7.wav');
22. [f,frGil(7,:),frGi2(7,:)] = rep_plot(y);
23. [y, Fs_y] = wavread('\271\rep8.wav');
24. [f,frGil(8,:),frGi2(8,:)] = rep_plot(y);
25. [y, Fs_y] = wavread('\271\rep9.wav');
26. [f,frGil(9,:),frGi2(9,:)] = rep_plot(y);
27. [y, Fs_y] = wavread('\271\rep10.wav');
28. [f,frGil(10,:),frGi2(10,:)] = rep_plot(y);
29.
30. %mean plot
31. mean_frGil = mean(frGil);
32. semilogx(f, mean_frGil, 'k', 'LineWidth', 2)
33. hold on
34. mean_frGi2 = mean(frGi2);
35. semilogx(f, mean_frGi2-30, 'k', 'LineWidth', 2)
36.
37. %standard deviation plot
38. subplot(3,1,2)
39. std_frGi2 = std(frGi2);
40. semilogx(f, std_frGi2, 'r', 'LineWidth', 1)
41. hold on
42. std_frGil = std(frGil);
43. semilogx(f, std_frGil, 'b', 'LineWidth', 1)
44. ylabel('magnitudo [dB]');
45. xlabel('frequency [Hz]')
46. x=[20,100, 1000, 10000]; % define the x values where you want to have a tick
47. set(gca, 'XTick', x); % Apply the ticks to the current axes
48. % Define the tick labels based on the user-defined format
49. set(gca, 'XTickLabel', arrayfun(@(v) sprintf('%d', v), x, 'UniformOutput', false));
50.
51. y=[0,1,2,4,6,8,10,12];
52. set(gca, 'YTick', y);
53. grid on;
54. title('STD')
55. xlim([20 20000])
56. ylim([0 12])
57. %   ILD : interaural level difference
58. %   RIGHT-LEFT
59. %   (positive values when right signal is higher than left one)
60. subplot(3,1,3)
61. ILD = (mean_frGi2 - mean_frGil);
62. semilogx(f, ILD, 'k', 'LineWidth', 1);
63. ylabel('magnitudo [dB]');
64. xlabel('frequency [Hz]')
65. x=[20,100, 1000, 10000]; % define the x values where you want to have a tick
66. set(gca, 'XTick', x); % Apply the ticks to the current axes
67. set(gca, 'XTickLabel', arrayfun(@(v) sprintf('%d', v), x, 'UniformOutput', false));
68. grid on;
69. title('ILD');
70. xlim([20 20000]);
71. ylim([-12 +12]);
72. y=[-15 , -10 , -5, 0, 5, 10, 15 ];
73. set(gca, 'YTick', y);
```

3. Function per il calcolo della fase della risposta all'impulso del sistema

```
1. function [f, phGi1,phGi2] = phase_plot(y)
2. % Evaluation of PHASE for a single
3. % measurement.
4. % input
5. %     y=sine sweep response
6. % output:
7. %     f=sample frequency
8. %     phGi1=Phase of left channel
9. %     phGi2=Phase of right channel
10. %
11.
12. %% Prepare data for sys id
13. [u, Fs_u] = wavread('sweep1.wav');
14. delay = 7290;
15. length_input_signal = 47700;
16.
17. Ts=1/Fs_u; % sampling time [s]
18. uf=u(1:length_input_signal);
19. yf=y;
20.
21. % window
22. yf1 = yf(delay:delay-1+length_input_signal,1);
23. yf2 = yf(delay:delay-1+length_input_signal,2);
24. % define data for sys id
25. z1 = iddata(yf1, uf, Ts);
26. z2 = iddata(yf2, uf, Ts);
27.
28. %% Spectral analysis ETFE
29.
30. % freq resp id using ETFE
31. g1 = etfe(z1, 1000, 1000)
32. g2 = etfe(z2, 1000, 1000)
33.
34. w = g1.Frequency;
35. f = w/2/pi;
36.
37. phGi1 = squeeze(g1.ResponseData);
38. phGi2 = squeeze(g2.ResponseData);
```

4. Metodo main per il plot della fase e delle differenze di fase binaurali

```
1. clc;
2. close all;
3. clear all;
4. %phGi1 Left
5. %phGi2 Right
6. figure(1)
7. %single reposition plot
8. [y, Fs_y] = wavread('.\DT770\rep1.wav');
9. [f,phGi1(1,:),phGi2(1,:)] = phase_plot(y);
10. [y, Fs_y] = wavread('.\DT770\rep2.wav');
11. [f,phGi1(2,:),phGi2(2,:)] = phase_plot(y);
12. [y, Fs_y] = wavread('.\DT770\rep3.wav');
13. [f,phGi1(3,:),phGi2(3,:)] = phase_plot(y);
14. [y, Fs_y] = wavread('.\DT770\rep4.wav');
15. [f,phGi1(4,:),phGi2(4,:)] = phase_plot(y);
16. [y, Fs_y] = wavread('.\DT770\rep5.wav');
17. [f,phGi1(5,:),phGi2(5,:)] = phase_plot(y);
18. [y, Fs_y] = wavread('.\DT770\rep6.wav');
19. [f,phGi1(6,:),phGi2(6,:)] = phase_plot(y);
20. [y, Fs_y] = wavread('.\DT770\rep7.wav');
21. [f,phGi1(7,:),phGi2(7,:)] = phase_plot(y);
22. [y, Fs_y] = wavread('.\DT770\rep8.wav');
23. [f,phGi1(8,:),phGi2(8,:)] = phase_plot(y);
24. [y, Fs_y] = wavread('.\DT770\rep9.wav');
25. [f,phGi1(9,:),phGi2(9,:)] = phase_plot(y);
26. [y, Fs_y] = wavread('.\DT770\rep10.wav');
27. [f,phGi1(10,:),phGi2(10,:)] = phase_plot(y);
28.
29. %
30. % Phase plot of repositions mean
```

```

31. %
32. subplot(2,1,1)
33. mean_phGi1 = mean(phGi1);
34. semilogx(f, angle(mean_phGi1),'b')
35. hold on
36. mean_phGi2 = mean(phGi2);
37. semilogx(f, angle(mean_phGi2),'r')
38. title(' DT770 PHASE ')
39. xlabel('frequency [Hz]')
40. x=[20,100, 1000, 10000]; % define the x values where you want to have a tick
41. set(gca,'XTick',x); % Apply the ticks to the current axes
42. % Define the tick labels based on the user-defined format
43. set(gca,'XTickLabel', arrayfun(@(v) sprintf('%d',v), x, 'UniformOutput', false) );
44.
45. ylabel('Phase [rad]');
46. set(gca,'YTick',-pi:pi/2:pi)
47. set(gca,'YTickLabel',{'-pi','-pi/2','0','pi/2','pi'})
48. grid on;
49. ylim([-3.5 3.5])
50. xlim([20 20000])
51.
52. %
53. % Phase difference beetwen L and R channel
54. % L and R means are used [left-right]
55. %
56.
57. subplot(2,1,2)
58. semilogx(f,angle(mean_phGi1)-angle(mean_phGi2),'k');
59. grid on;
60. title('INTER AURAL PHASE DIFFERENCE ')
61.
62. xlim([20 20000])
63. xlabel('frequency [Hz]')
64. x=[20,100, 1000, 10000]; % define the x values where you want to have a tick
65. set(gca,'XTick',x); % Apply the ticks to the current axes
66. % Define the tick labels based on the user-defined format
67. set(gca,'XTickLabel', arrayfun(@(v) sprintf('%d',v), x, 'UniformOutput', false) );
68. ylabel('Phase [rad]');
69. set(gca,'YTick',-pi*2:pi/2:pi*2)
70. set(gca,'YTickLabel',{'-2pi','-3/2pi','-pi','-i/2','0','pi/2','pi','3/2pi','2pi'})

```

Bibliografia :

1. H. Møller : "Fundamentals of binaural technology". Appl Acoust 36:171–218 (1992)
2. M. Hiipakka, T. Kinnari, V. Pulkki : "Estimating head-related transfer functions of human subjects from pressure–velocity measurements". J. Acoust. Soc. Am. 131 (5), 2012
3. D. Hammershøi and H. Møller : "Binaural Technique – Basic Methods for Recording, Synthesis, and Reproduction" (1996)
4. R. Nishimura, P. Mokhtari, H. Takemoto, H. Kato: "An Attempt to Calibrate Headphones for Reproduction of Sound Pressure at the Eardrum". IEEE Transactions on Audio, Speech & Language Processing
5. D. Hammershøi , H. Møller "Sound transmission to and within the human ear canal. J Acoust Soc Amer 100:408–427
6. K.I. Mcanally and R.L. Martin : "Variability in the headphone-to-ear-canal transfer function". In J. Audio Eng. Soc.
7. A. Kulkarni and H. S. Colburn : "Variability in the characterization of the headphone transfer-function " In J. Audio Eng. Soc.
8. M. Geronazzo, F. Granza, S. Spagnol, and F. Avanzini : "A standardized repository of Head-Related and Headphone Impulse Response data" J. Audio Eng. Soc. (2013)
9. Z. Schärer , A. Lindau : " Evaluation of Equalization Methods for Binaural Signals". J. Audio Eng. Soc. (2009)
10. H. Møller, D. Hammershøi, CB Jensen and MF Sørensen, "Transfer characteristics of headphones measured on human ears" J. Audio Eng. Soc
11. D. Pralong and S. Carlile : "The role of individualized headphone calibration for the generation of high fidelity virtual auditory space" J.Acoust. Soc. Am.
12. Bruce Bartlett : "Tecniche stereofoniche di microfononaggio" Hoepli
13. V. Lombardo , A. Valle : " Audio e multimedia" Apogeo
14. Everest F.Alton : "Manuale di Acustica" Hoepli