

UNIVERSITÀ DEGLI STUDI DI PADOVA

Dipartimento di Ingegneria dell'Informazione

Corso di Laurea Vecchio Ordinamento in  
Ingegneria Informatica

TESI DI LAUREA

**VALUTAZIONE PARAMETRICA DI UN  
MODELLO STRUTTURALE DI ORECCHIO  
ESTERNO PER IL RENDERING SPAZIALE DEL  
SUONO**

**Relatore:** Prof. Federico Avanzini

**Correlatore:** Dott. Simone Spagnol

**Laureando:** Sandro Scaiella

Data Laurea: 14 Aprile 2014

Anno Accademico  
2012-2013



*A tutta la mia famiglia e  
a una persona davvero speciale e insostituibile:  
Ciki.  
Ed infine, a Kiki, che è stata la mia seconda mamma.*



# Abstract

In questo lavoro viene trattato il tema della sintesi di suoni spazializzati. Gli effetti acustici cui è sottoposto il suono nel tragitto dalla sorgente alle orecchie dell'ascoltatore sono riassunti in una specifica funzione di trasferimento (Head Related Transfer Function, HRTF), di cui la PRTF (Pinna Related Transfer Function) rappresenta il solo contributo dato dall'orecchio esterno. A partire da un modello di PRTF sintetica preesistente, vengono apportate delle modifiche al blocco riflettente di tale modello, lasciando al contrario invariato il blocco risonante. Viene quindi progettato ed implementato un esperimento psicoacustico con l'obbiettivo di testare la bontà di tale modello per sorgenti sonore sul piano mediano e trarre utili considerazioni sui parametri testati. Si è quindi reso necessario realizzare un test audiometrico, allo stesso tempo accurato e veloce, da sottoporre ai soggetti testati per accertarsi dell'assenza di problemi uditivi tali da compromettere l'esito dell'esperimento psicoacustico. Si analizzano infine i risultati relativi ad ogni soggetto testato e si ottiene la combinazione dei parametri che meglio si adatta ad una buona percezione dell'elevazione, ottenendo così utili indicazioni per il miglioramento dei modelli strutturali sintetici utilizzati per l'ascolto binaurale in cuffia.

**Parole chiave:** audio 3D, audio binaurale, HRTF, spazializzazione.



# Indice

<b>Abstract</b>	<b>iii</b>
<b>Introduzione</b>	<b>xiii</b>
<b>1 Il suono nello spazio</b>	<b>1</b>
1.1 Apparato uditivo . . . . .	1
1.2 Il suono . . . . .	2
1.3 Sistemi di riferimento . . . . .	3
1.4 Percezione del suono nello spazio . . . . .	4
1.4.1 Audio binaurale . . . . .	4
1.4.2 La testa . . . . .	5
1.4.3 L'orecchio esterno . . . . .	6
1.4.4 Busto e spalla . . . . .	7
1.4.5 Teoria Duplex . . . . .	9
1.4.6 Percezione dell'elevazione . . . . .	9
1.4.7 Head Related Transfer Function . . . . .	10
1.5 Rendering di audio 3D . . . . .	13
1.5.1 Modello strutturale . . . . .	15
<b>2 Modellazione di PRTF sintetiche</b>	<b>17</b>
2.1 Stato dell'arte . . . . .	17
2.2 Modello strutturale di PRTF . . . . .	20
2.2.1 Filtri . . . . .	24
2.2.2 Geometria della pinna . . . . .	26
2.3 Modifiche al modello strutturale di PRTF . . . . .	30
2.3.1 Filtri . . . . .	30
2.3.1.1 Blocco risonante . . . . .	30
2.3.1.2 Blocco riflettente . . . . .	33
2.3.2 Condizioni sperimentali sui parametri dei notch . . . . .	35
2.3.2.1 Studi precedenti . . . . .	35
2.3.2.2 Scelta dei parametri . . . . .	38

<b>3</b>	<b>Design Sperimentale</b>	<b>45</b>
3.1	Apparato . . . . .	45
3.1.1	Software . . . . .	45
3.1.2	Hardware . . . . .	45
3.1.2.1	Cuffie . . . . .	45
3.1.2.2	Scheda audio esterna . . . . .	46
3.1.2.3	Fonometro . . . . .	47
3.2	Test audiometrico . . . . .	47
3.2.1	Procedura “Maximum Likelihood” . . . . .	48
3.2.2	Protocollo . . . . .	49
3.2.3	Calibrazione e rappresentazione grafica dell’hearing threshold . . . . .	53
3.2.3.1	Brevi note di audiologia . . . . .	54
3.3	Test psicoacustico . . . . .	55
3.3.1	Soggetti . . . . .	55
3.3.2	Stimoli . . . . .	56
3.3.3	Condizioni . . . . .	56
3.3.4	Acquisizione immagine della pinna e tracciamento . . . . .	57
3.3.5	Procedura . . . . .	60
3.3.5.1	GUI soggetti . . . . .	60
3.3.5.2	GUI esperimento . . . . .	62
<b>4</b>	<b>Risultati</b>	<b>65</b>
4.1	Considerazioni preliminari . . . . .	65
4.2	Soggetti . . . . .	68
4.2.1	Soggetto 20 . . . . .	68
4.2.2	Soggetto 21 . . . . .	71
4.2.3	Soggetto 22 . . . . .	74
4.2.4	Soggetto 24 . . . . .	77
4.2.5	Soggetto 26 . . . . .	80
4.2.6	Soggetto 27 . . . . .	83
4.2.7	Soggetto 28 . . . . .	86
4.2.8	Soggetto 30 . . . . .	89
4.2.9	Soggetto 32 . . . . .	92
4.2.10	Soggetto 33 . . . . .	95
4.2.11	Soggetto 35 . . . . .	98
4.2.12	Soggetto 36 . . . . .	101
4.3	Condizione globale migliore e considerazioni finali . . . . .	104
<b>5</b>	<b>Conclusioni e sviluppi futuri</b>	<b>109</b>
	<b>Appendici</b>	<b>115</b>
<b>A</b>	<b>Grafici</b>	<b>115</b>

<b>B</b>	<b>Audiogrammi</b>	<b>123</b>
<b>C</b>	<b>Codice Matlab</b>	<b>129</b>
C.1	Funzione <i>exp1_model_reflection</i> . . . . .	129
C.2	Funzione <i>exp1_model_resonance</i> . . . . .	130
C.3	Funzione <i>exp1_notching</i> . . . . .	132
C.4	Funzione <i>peaking2</i> . . . . .	132
C.5	Funzione <i>prtf</i> . . . . .	133
C.6	Funzione <i>create_subject</i> . . . . .	134
C.7	Funzione <i>contour_trace</i> . . . . .	135
C.8	Funzione <i>stimulus3</i> . . . . .	137
C.9	Funzione <i>subjects</i> . . . . .	137
C.10	Funzione <i>experiment1</i> . . . . .	153
C.11	Funzione <i>MLP</i> . . . . .	160



# Elenco delle figure

1.1	Sistema uditivo . . . . .	2
1.2	Sistemi di riferimento . . . . .	4
1.3	Interaural time difference e interaural level difference . . . . .	5
1.4	Stima dell'ITD nel caso di testa sferica e di una sorgente sonora distante . . . . .	6
1.5	I sei modi di risonanza della pinna identificati da Shaw . . . . .	7
1.6	Effetti del busto e della spalla . . . . .	8
1.7	Cono di confusione . . . . .	10
1.8	Esempi di HRTF . . . . .	11
1.9	Schema a blocchi di un sistema di rendering di audio 3D per mezzo di cuffie . . . . .	14
1.10	Modello strutturale della HRTF proposto da Brown e Duda . . . . .	15
1.11	Sistema di riproduzione di audio 3D basato su un modello strutturale di HRTF . . . . .	16
2.1	Individuazione nelle HRIR dell'interazione tra suono e ostacoli . . . . .	19
2.2	Anatomia della pinna . . . . .	19
2.3	Modello di PRTF proposto da Satarzadeh . . . . .	20
2.4	Sistema di coordinate polari interaurali e regione di validità del modello . . . . .	21
2.5	Diagramma di flusso dell'algorithmo di separazione . . . . .	22
2.6	Risonanze per differenti elevazioni . . . . .	23
2.7	Notch spettrali per differenti elevazioni . . . . .	23
2.8	Box plot e media delle profondità del primo notch tra soggetti CIPIC . . . . .	24
2.9	Modello strutturale della PRTF . . . . .	24
2.10	Confronto tra i due filtri Peak IIR del secondo ordine . . . . .	26
2.11	Sistema di coordinate polari 2D e punti di riflessione . . . . .	27
2.12	Punti di riflessione su Soggetto 134 CIPIC . . . . .	28
2.13	Tracciamento contorni della pinna . . . . .	28
2.14	Punti di riflessione in coordinate polari . . . . .	29
2.15	Modello strutturale di HRTF . . . . .	29
2.16	PRTF originali e sintetiche . . . . .	29
2.17	Prima e seconda risonanza in cascata con e senza filtri passa basso e passa alto . . . . .	31
2.18	Primo e secondo picco per le 7 elevazioni . . . . .	32
2.19	Primo, secondo e terzo notch in cascata per le elevazioni 0° . . . . .	34
2.20	PRTF sintetiche per il Soggetto 30 per l'elevazione 0° . . . . .	34
2.21	PRTF per il Soggetto 21, ad elevazione 0° e con le condizioni 1, 2 e 3 . . . . .	40
2.22	PRTF per il Soggetto 21, ad elevazione 0° e con le condizioni 4, 5 e 6 . . . . .	41

2.23	PRTF per il Soggetto 21, ad elevazione 0° e con la condizione 7, 8 e 9 . . . . .	42
2.24	PRTF per il Soggetto 21, ad elevazione 0° e con le condizioni 10, 11 e 12 . . . . .	43
3.1	Sennheiser HDA 200 . . . . .	46
3.2	Sennheiser HDA 200: risposta in frequenza . . . . .	47
3.3	Roland Edirol UA-101 . . . . .	47
3.4	MLP: interfaccia grafica . . . . .	48
3.5	Funzione logistica . . . . .	49
3.6	Setup per la calibrazione delle cuffie: Cuffia, sostegno e fonometro. . . . .	53
3.7	Audiogramma soggetto con ID 20. . . . .	55
3.8	Setup per l'acquisizione dell'immagine della pinna. . . . .	57
3.9	Immagine di pinna acquisita tramite fotocamera. . . . .	58
3.10	Tracciamento dei 3 contorni sulla pinna del soggetto. . . . .	59
3.11	Cabina silente Sound Station Pro 45 by Puma s.r.l.. . . . .	60
3.12	Interfaccia GUI contenente il database con la lista dei soggetti testati. . . . .	61
3.13	Interfaccia GUI utilizzata dal soggetto per indicare le posizioni rilevate. . . . .	61
3.14	Finestra di avviso di creazione dei filtri in corso. . . . .	63
3.15	Finestra di countdown mostrata durante la pausa tra blocchi di stimoli. . . . .	64
4.1	Scatterplot del Soggetto ID 20 . . . . .	69
4.2	Scatterplot del Soggetto ID 21 . . . . .	72
4.3	Scatterplot del Soggetto ID 22 . . . . .	75
4.4	Scatterplot del Soggetto ID 24 . . . . .	78
4.5	Scatterplot del Soggetto ID 26 . . . . .	81
4.6	Scatterplot del Soggetto ID 27 . . . . .	84
4.7	Scatterplot del Soggetto ID 28 . . . . .	87
4.8	Scatterplot del Soggetto ID 30 . . . . .	90
4.9	Scatterplot del Soggetto ID 32 . . . . .	93
4.10	Scatterplot del Soggetto ID 33 . . . . .	96
4.11	Scatterplot del Soggetto ID 35 . . . . .	99
4.12	Scatterplot del Soggetto ID 36 . . . . .	102
A.1	Cascata dei 3 notch a elevazioni -45° e -30° per il Soggetto 30 . . . . .	116
A.2	Cascata dei 3 notch a elevazioni -15° e 15° per il Soggetto 30 . . . . .	117
A.3	Cascata dei 3 notch a elevazioni 30° e 45° per il Soggetto 30 . . . . .	118
A.4	PRTF sintetiche a elevazioni -45° e -30° per il Soggetto 30 . . . . .	119
A.5	PRTF sintetiche a elevazioni -15° e 15° per il Soggetto 30 . . . . .	120
A.6	PRTF sintetiche a elevazioni 30° e 45° per il Soggetto 30 . . . . .	121
B.1	Audiogrammi dei Soggetti 20 e 21 . . . . .	124
B.2	Audiogrammi dei Soggetti 22 e 24 . . . . .	125
B.3	Audiogrammi dei Soggetti 27 e 28 . . . . .	126
B.4	Audiogrammi dei Soggetti 30 e 33 . . . . .	127

# Elenco delle tabelle

2.1	Parametri per il primo filtro peak . . . . .	31
2.2	Parametri per il secondo filtro peak . . . . .	32
2.3	Condizioni sperimentali . . . . .	39
3.1	RETSPLs per le cuffie Sennheiser HDA 200 . . . . .	54
3.2	Soggetti esperimento . . . . .	56
4.1	Classifica delle condizioni per il Soggetto 20 . . . . .	68
4.2	Dati statistici di interesse per il Soggetto 20 . . . . .	70
4.3	Classifica delle condizioni per il Soggetto 21 . . . . .	71
4.4	Dati statistici di interesse per il Soggetto 21 . . . . .	73
4.5	Classifica delle condizioni per il Soggetto 22 . . . . .	74
4.6	Dati statistici di interesse per il Soggetto 22 . . . . .	76
4.7	Classifica delle condizioni per il Soggetto 24 . . . . .	77
4.8	Dati statistici di interesse per il Soggetto 24 . . . . .	79
4.9	Classifica delle condizioni per il Soggetto 26 . . . . .	80
4.10	Dati statistici di interesse per il Soggetto 26 . . . . .	82
4.11	Classifica delle condizioni per il Soggetto 27 . . . . .	83
4.12	Dati statistici di interesse per il Soggetto 27 . . . . .	85
4.13	Classifica delle condizioni per il Soggetto 28 . . . . .	86
4.14	Dati statistici di interesse per il Soggetto 28 . . . . .	88
4.15	Classifica delle condizioni per il Soggetto 30 . . . . .	89
4.16	Dati statistici di interesse per il Soggetto 30 . . . . .	91
4.17	Classifica delle condizioni per il Soggetto 32 . . . . .	92
4.18	Dati statistici di interesse per il Soggetto 32 . . . . .	94
4.19	Classifica delle condizioni per il Soggetto 33 . . . . .	95
4.20	Dati statistici di interesse per il Soggetto 33 . . . . .	97
4.21	Classifica delle condizioni per il Soggetto 35 . . . . .	98
4.22	Dati statistici di interesse per il Soggetto 35 . . . . .	100
4.23	Classifica delle condizioni per il Soggetto 36 . . . . .	101
4.24	Dati statistici di interesse per il Soggetto 36 . . . . .	103
4.25	Classifica globale delle condizioni per valore del <i>mean_angle_error</i> . . . . .	104
4.26	Classifica globale delle condizioni per numero di <i>up/down confusion</i> . . . . .	105
4.27	Classifica globale delle condizioni in base alla formula ( <i>regression_slope</i> + <i>R2</i> ) . . . . .	106



# Introduzione

L'illusione che un determinato suono sia posizionato in una determinata locazione virtuale rappresenta l'obiettivo principale di qualsiasi sistema di rendering di audio 3D. Tale obiettivo può essere raggiunto sia per mezzo di altoparlanti sia di cuffie. L'idea di base, sottesa al citato sistema, è quella di presentare ai due timpani dell'ascoltatore gli stessi segnali che produrrebbe una sorgente reale posizionata in quel particolare punto dello spazio di cui si vuol dare la percezione.

Il presente lavoro di tesi è dedicato allo studio della percezione dell'elevazione di sorgenti sonore localizzate sul piano mediano, ovvero equidistanti dalle due orecchie dell'ascoltatore; lo stimolo sonoro è sottoposto ai soggetti mediante cuffie. L'idea di fondo è quella di studiare e modellare un sistema che permetta, una volta ricavate alcune caratteristiche antropometriche relative all'ascoltatore, di ricostruire le HRTF (*Head Related Transfer Function*) individualizzate. Esse rappresentano dei filtri che modellano gli effetti sulla propagazione del suono dovuti alla testa, all'orecchio esterno, al busto ed alla spalla. Questo lavoro di tesi rientra all'interno di tale ambito di studi, soffermandosi esclusivamente sul contributo dato dall'orecchio esterno (*Pinna Related Transfer Function*, PRTF), i cui effetti, come vedremo nel corso della trattazione, permettono di discriminare l'elevazione di una sorgente sonora.

Al fine di valutare i parametri di un modello di PRTF sintetica relazionati alle caratteristiche antropometriche dell'orecchio esterno del singolo ascoltatore, si è resa necessaria la progettazione e l'implementazione di un esperimento psicoacustico, ottenendo, per ogni soggetto, la combinazione di parametri che meglio si adatta ad una percezione spaziale realistica.

Nel Capitolo 1 sono descritte le caratteristiche del sistema uditivo umano, vengono analizzati i principali effetti sulla propagazione del suono dovuti alle diverse parti anatomiche (testa, orecchio esterno, busto e spalla), si definisce l'HRTF, viene introdotta la Teoria Duplex ed infine sono presentati i sistemi utilizzati nella resa dell'audio 3D.

Nel Capitolo 2 si descrivono gli studi alla base della modellazione del contributo dell'orecchio esterno, riportando alcuni lavori precedenti, presenti in letteratura, sulla sintesi di PRTF sintetiche. Viene quindi descritto il modello di PRTF sintetica alla base della presente tesi ed infine il modello modificato testato con l'esperimento psicoacustico.

Nel Capitolo 3 viene descritto il design dell'esperimento psicoacustico e del test audiometrico cui sono sottoposti i soggetti, presentando le diverse fasi del test di localizzazione ed il protocollo sperimentale seguito dai medesimi soggetti testati.

Nel Capitolo 4 vengono presentati, per ogni singolo soggetto, i risultati del test psicoacustico, ricavando infine la combinazione di parametri antropometrici migliore per la percezione dell'elevazione della sorgente sonora.

In Appendice, quali ulteriori elementi conoscitivi, sono riportati gli audiogrammi di alcuni dei soggetti testati, ricavati per mezzo del test audiometrico, e due serie di grafici rappresentanti la PRTF sintetica globale nonché il suo solo blocco riflettente. Infine viene riportato il codice Matlab con i principali algoritmi realizzati per l'implementazione del test psicoacustico e del test audiometrico.



# Capitolo 1

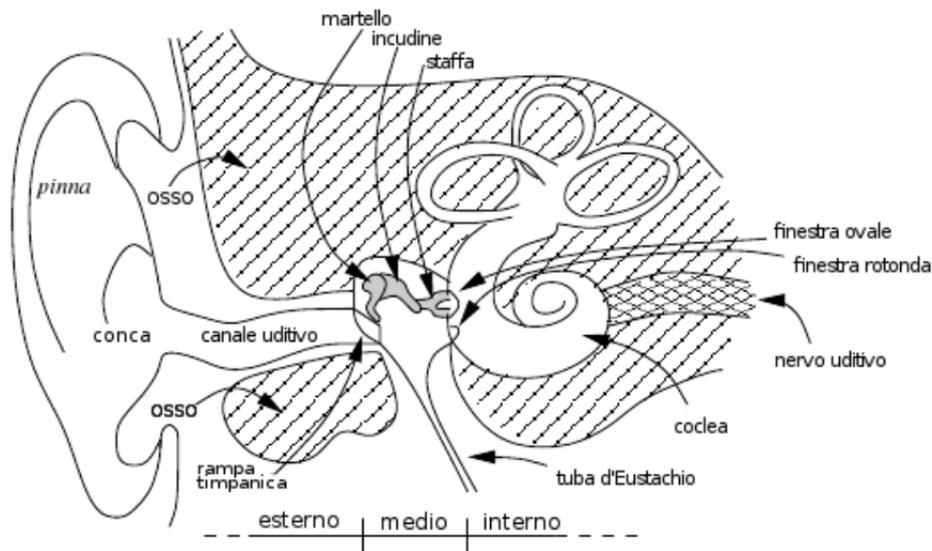
## Il suono nello spazio

In questo capitolo vengono introdotti i concetti alla base di questo lavoro di tesi. Nella Sezione 1.1 viene descritto il sistema uditivo; nella Sezione 1.2 viene data una definizione di suono; nella Sezione 1.3 vengono descritti i sistemi di riferimento maggiormente utilizzati nella letteratura; nella Sezione 1.4 vengono descritti i principali effetti delle parti anatomiche sulla propagazione del suono ed infine nella Sezione 1.5 vengono presentati i sistemi utilizzati per la resa dell'audio 3D.

### 1.1 Apparato uditivo

Come descritto dettagliatamente in [25] e [38], l'orecchio umano è diviso in tre parti: l'*orecchio esterno*, l'*orecchio medio* e l'*orecchio interno*. In Figura 1.1 è raffigurato il sistema uditivo. L'orecchio esterno è formato dal padiglione auricolare (detto anche pinna), dal canale uditivo esterno detto anche meato e dal timpano. Il padiglione auricolare consente di catturare i segnali ad alta frequenza (superiori a 3 kHz) che altrimenti non riuscirebbero ad entrare nel condotto uditivo. Esso consente di individuare la provenienza del suono per mezzo delle sue pieghe; infatti, tali pieghe causano la riflessione delle onde sonore che giungono all'imboccatura del canale uditivo tra loro sfasate a causa dei differenti percorsi seguiti. Da tali sfasamenti e dalle variazioni spettrali conseguenti, il sistema nervoso centrale ottiene indicazioni sulla direzione di provenienza del suono. Il canale uditivo collega il padiglione auricolare con il timpano. La sua frequenza di risonanza è intorno ai 2 kHz, quindi le frequenze intorno a tale valore vengono amplificate ed arrivano in modo più efficace al timpano. Quest'ultimo è una membrana sottile la cui funzione è quella di trasmettere le vibrazioni ai tre ossicini dell'orecchio medio.

L'orecchio medio consente di propagare ed amplificare il suono per mezzo della cassa timpanica che si trova oltre il timpano e dei tre ossicini, *martello*, *incudine* e *staffa*. La forza che giunge dalla vibrazione del timpano agisce sul martello e si trasmette alla base della staffa per mezzo dell'incudine. La staffa esercita una pressione meccanica sulla *finestra ovale*, la cui superficie è di molto inferiore a quella del timpano, e sui cui agisce una forza amplificata in modo da percepire anche i segnali deboli. L'orecchio medio permette di proteggere la coclea da suoni intensi: il muscolo timpanico irrigidendosi fa sì che il timpano non si deformi troppo ed allo



**Figura 1.1:** Sistema uditivo

stesso tempo il muscolo stapedio allontana la staffa dalla finestra ovale, riducendo in questo modo il trasferimento di vibrazione. Per mezzo della *tromba di Eustachio* che congiunge la cassa timpanica con la faringe, la pressione ai due lati della membrana del timpano viene equilibrata e la membrana ritorna in posizione subito dopo la vibrazione.

L'orecchio interno è costituito dalla *coclea* e dal *labirinto*, quest'ultimo regola l'equilibrio. La coclea si presenta come un tubo a forma di chiocciola suddiviso in tre condotti o rampe: *rampa vestibolare*, *rampa timpanica* e *rampa media*. Essi sono divisi da due membrane di cui la più spessa è chiamata *membrana basilare*, mentre la più sottile è chiamata *membrana di Reissner*. La rampa vestibolare e quella timpanica sono riempite con un fluido chiamato *perilinf*, quest'ultimo in contatto diretto con il liquido cefalorachidiano contenuto nella cavità del cervello. La rampa media è invece riempita di un fluido chiamato *endolinf*, in contatto con il sistema vestibolare. Le oscillazioni vengono quindi trasmesse dalla staffa alla perilinf e quindi alla membrana basilare. Su quest'ultima è posto l'*organo di Corti*, composto da oltre 20.000 *cellule ciliate*. Esso si occupa della trasduzione da segnale acustico a segnale elettrochimico, che viene inviato al cervello per mezzo delle cellule nervose che fanno parte del *nervo vestibolo cocleare*. Il segnale giunge quindi nell'area acustica della corteccia celebrale, e successivamente al lobo temporale del cervello in cui si ha la decodifica dell'impulso elettrico in percezione del suono.

## 1.2 Il suono

Per suono si intende quel fenomeno meccanico dato da una perturbazione di un mezzo di trasmissione (in genere l'aria) e che abbia caratteristiche tali da essere percepito dall'orecchio umano.

L'equazione che regola la propagazione del suono è quella di d'Alembert:

$$\nabla^2 p(\mathbf{x}, t) = \frac{1}{c^2} \frac{\partial^2 p}{\partial t^2}(\mathbf{x}, t) \quad (1.1)$$

dove  $x$  rappresenta le coordinate Euclidee nello spazio e  $p$  è la pressione acustica.  $\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}$  indica l'operatore Laplaciano 3-dimensionale e  $c$  indica la velocità di propagazione del suono<sup>1</sup>.

Nel momento in cui l'aria viene perturbata, il valore di pressione non risulta più costante, ma varia da punto a punto, aumentando dove le molecole sono compresse e diminuendo dove le molecole sono espanse [38]. Tali variazioni di pressione giungono ai nostri timpani ed attraverso l'orecchio medio e quello interno vengono convertite in segnali elettrici ed inviati al cervello sotto forma di modello neuronale. Il mezzo di propagazione interferisce con l'informazione sonora causando l'assorbimento di parte delle onde sonore, in maniera diversa a seconda delle caratteristiche di temperatura e umidità del mezzo stesso. Anche l'ambiente nel quale il suono si propaga produce versioni ritardate ed attenuate del suono stesso a causa dell'interazione dell'onda con pareti ed ostacoli, fenomeno a cui viene dato il nome di riverbero. Infine anche le diverse parti del corpo interferiscono con la propagazione del suono, causando i fenomeni di cui parleremo nella sezione 1.4.

### 1.3 Sistemi di riferimento

Introduciamo dapprima tre piani tra loro ortogonali:

- *piano mediano;*
- *piano frontale;*
- *piano orizzontale.*

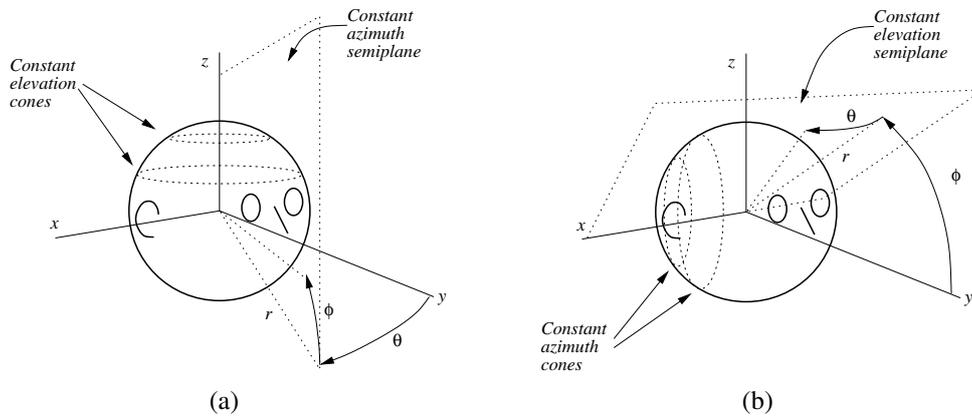
Il piano mediano è quel piano sul quale si dispongono le sorgenti sonore (assunte puntiformi) che sono equidistanti dalle due orecchie, separando in questo modo il lato sinistro dal lato destro. Il piano orizzontale è il piano su cui giacciono le orecchie, separando la parte superiore della testa da quella inferiore. Il piano frontale divide la testa dell'ascoltatore nel senso del torso e separa quindi fronte e retro [25].

Per rappresentare spazialmente una sorgente sonora posta attorno all'ascoltatore, in letteratura sono utilizzati principalmente due sistemi di riferimento polari:

- *verticale polare;*
- *interaurale polare.*

---

<sup>1</sup>In aria secca ed alla temperatura di 20°C la velocità del suono è di 343.2 m/s



**Figura 1.2:** Sistemi di riferimento: (a) verticale-polare, (b) interaurale polare

Il primo è raffigurato in Figura 1.2(a). In questo sistema l'azimut  $\theta$  corrisponde all'angolo tra il piano  $yz$ , piano mediano, e il semipiano iniziante dall'asse verticale  $z$  e contenente la sorgente sonora ( $\theta \in [-180^\circ, +180^\circ]$ ). Su questo semipiano, l'elevazione  $\phi$  corrisponde all'angolo tra la semiretta di intersezione con il piano orizzontale  $xy$  e la semiretta iniziante dall'origine degli assi e contenente la sorgente sonora ( $\phi \in [-90^\circ, 90^\circ]$ ). In questo sistema di coordinate, le superfici ad azimut costante sono dei semipiani aventi l'asse verticale  $z$  come retta delimitatrice, mentre le superfici ad elevazione costante sono coni concentrici attorno allo stesso asse.

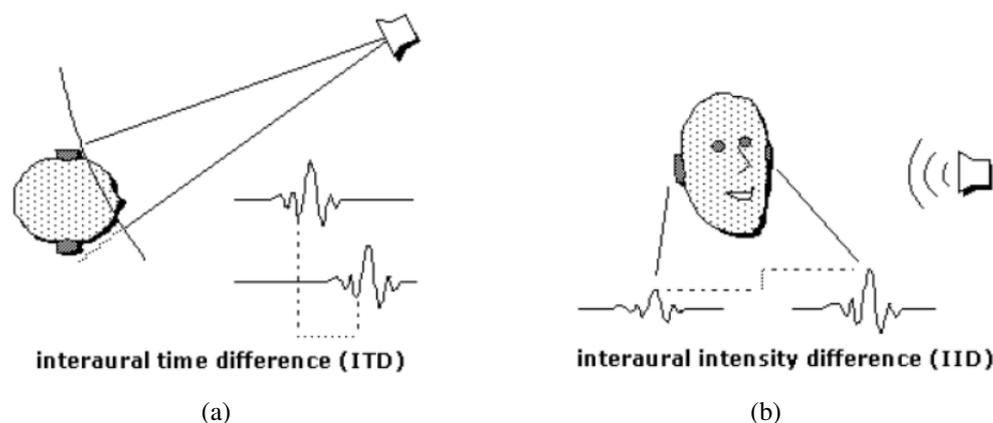
Il secondo è raffigurato in Figura 1.2(b). In questo sistema l'elevazione  $\phi$  corrisponde all'angolo tra il piano orizzontale  $xy$  e il semipiano iniziante dall'asse interaurale  $x$  e contenente la sorgente sonora ( $\phi \in [-180^\circ, 180^\circ]$ ). Su questo piano, l'azimut corrisponde all'angolo tra la semiretta di intersezione con il piano mediano  $yz$  e la semiretta iniziante dall'origine degli assi e contenente la sorgente sonora ( $\theta \in [-90^\circ, 90^\circ]$ ). In questo sistema di coordinate, le superfici ad elevazione costante sono dei semipiani aventi l'asse interaurale come retta delimitatrice, mentre le superfici ad azimut costante sono coni concentrici attorno all'asse  $x$ .

In entrambi i sistemi,  $r$  rappresenta la distanza euclidea tra l'origine degli assi e la sorgente sonora.

## 1.4 Percezione del suono nello spazio

### 1.4.1 Audio binaurale

Il suono prodotto da una sorgente è soggetto a diverse trasformazioni nel cammino verso le orecchie dell'ascoltatore, tra cui le perdite di energia isofrequenziali dovute alla distanza percorsa e quelle dipendenti dalla frequenza del suono, dovute quest'ultime agli oggetti schermanti che l'onda incontra durante la propagazione. I segnali che arrivano alle due orecchie dell'ascoltatore saranno quindi tra loro diversi anche a causa della presenza della testa. Proprio in base a tale differenza dei flussi sonori, l'ascoltatore sarà in grado di discriminare direzione e distanza



**Figura 1.3:** *Interaural time difference e interaural level difference*

della sorgenti, entro soglie di errore accettabili. Tale concetto è alla base della teoria sull'*audio binaurale*.

Vediamo in dettaglio l'effetto di filtraggio provocato dalla testa, dall'orecchio esterno, dal busto e dalla spalla, parti del corpo con cui interferisce il suono lungo il tragitto per raggiungere i timpani.

## 1.4.2 La testa

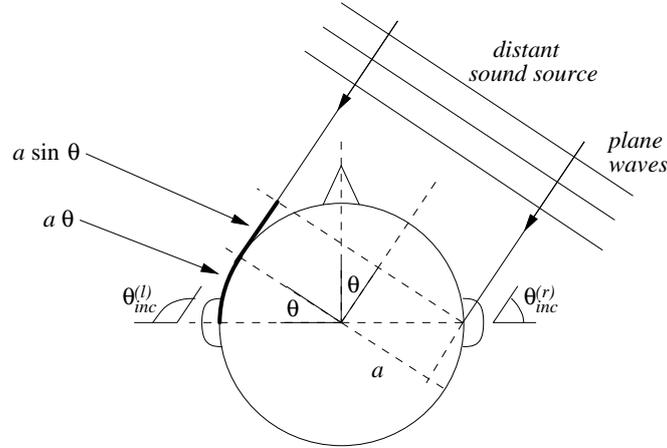
La testa in acustica è intesa come un corpo rigido sul cui asse interaurale sono ancorate, alla stessa altezza, le orecchie. Essa agisce da ostacolo alla propagazione delle onde sonore, provocando i seguenti due effetti principali:

- ritardare il tempo di arrivo dell'onda all'orecchio più distante, dal momento che l'onda deve percorrere una distanza maggiore per raggiungerlo. La differenza tra i tempi di arrivo alle due orecchie viene indicata come *interaural time difference* (ITD), mostrata in Figura 1.3(a);
- attenuare l'intensità sonora dell'onda che giunge all'orecchio più lontano, realizzando un effetto d'ombra. La differenza tra le intensità sonore alle due orecchie viene indicata come *interaural level difference* (ILD), mostrata in Figura 1.3(b).

Assumendo l'ipotesi di testa sferica e sorgente sonora sufficientemente distante in modo da considerare planari le onde che raggiungono la testa, l'ITD può essere ricavata dividendo la distanza extra che l'onda deve percorrere per raggiungere l'orecchio più distante ( $\Delta x$ ) per la velocità del suono ( $c$ ). Da semplici considerazioni geometriche, come mostrato in Figura 1.4, otteniamo:

$$ITD = \frac{a}{c}(\theta + \sin \theta) \quad (1.2)$$

dove  $a$  indica il raggio della testa,  $\theta$  indica l'angolo di azimut in radianti e  $c$  è la velocità del suono in m/s. Quando la sorgente si trova direttamente di fronte alla testa ( $\theta = 0$ ) l'ITD assume



**Figura 1.4:** Stima dell'ITD nel caso di testa sferica e di una sorgente sonora distante

il valore di zero, mentre quando la sorgente si trova sul piano frontale ( $\theta = \pi/2 \text{ rad}$ ) essa assume il suo valore massimo pari a  $a/c(\pi/2 + 1)$ .

Mentre l'ITD viene considerata in prima approssimazione indipendente dalla frequenza, l'ILD è fortemente influenzata dalla frequenza delle onde sonore: a basse frequenze alle due orecchie giunge una pressione sonora di differenza trascurabile l'una dall'altra, mentre, ad alte frequenze, la differenza diventa significativa. Considerando la funzione di trasferimento di una testa ideale sferica di raggio  $a$  sottoposta ad una sorgente sonora posta ad una distanza  $r > a$  dal centro della sfera, otteniamo:

$$H_{sphere}(\rho, \theta_{inc}, \mu) = -\frac{\rho}{\mu} e^{-i\mu\rho} \sum_{\mu=0}^{+\infty} (2m+1) P_m(\cos \theta_{inc}) \frac{h_m(\mu\rho)}{h'_m(\mu)} \quad (1.3)$$

dove  $P_m$  e  $h_m$  rappresentano rispettivamente il polinomio di Legendre di ordine m-esimo e la funzione sferica di Hankel di ordine m-esimo,  $\theta_{inc}$  indica l'angolo di incidenza,  $\mu = \omega a/c$  rappresenta la frequenza normalizzata e  $\rho = r/a$  è la distanza normalizzata.

A basse frequenze la funzione di trasferimento non è dipendente dalla direzione del suono e l'ampiezza è essenzialmente unitaria per ogni angolo di incidenza. Ad alte frequenze ( $\mu > 1$ ) è possibile invece notare una dipendenza dall'angolo di incidenza; nel caso di incidenza normale ( $\theta_{inc} = 0$ ) si osserva un guadagno di 6 dB (situazione riconducibile a quella di un'onda piana che incide perpendicolarmente con una superficie piana rigida) e all'aumentare dell'angolo di incidenza tale guadagno diminuisce. La risposta minima non si presenta esattamente nel punto opposto alla sorgente sonora ( $\theta_{inc} = \pi \text{ rad}$ ), poichè in esso si nota l'effetto *bright spot* dovuto alla somma costruttiva delle onde che viaggiano nelle diverse direzioni intorno alla testa.

### 1.4.3 L'orecchio esterno

Come visto in precedenza, prima di arrivare al timpano, il suono deve incontrare la pinna ed il canale uditivo. Quest'ultimo si comporta da risonatore ad una dimensione ed è approssimabile

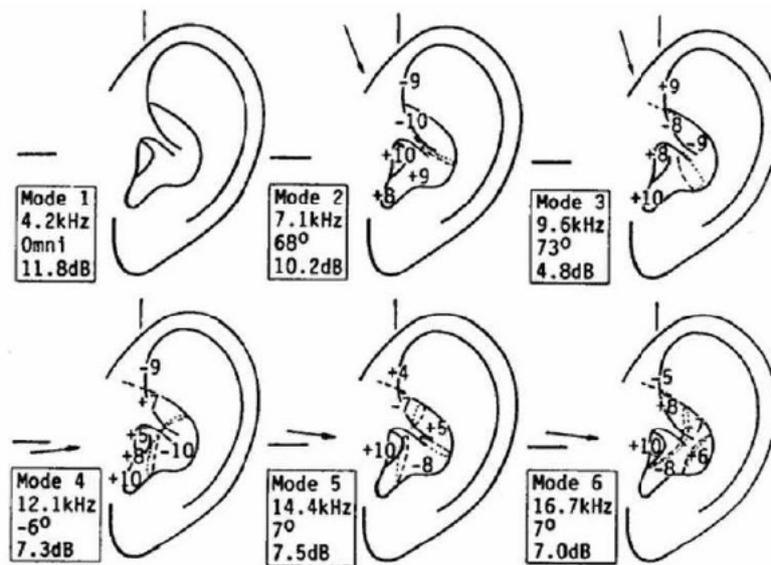


Figura 1.5: I sei modi di risonanza della pinna identificati da Shaw

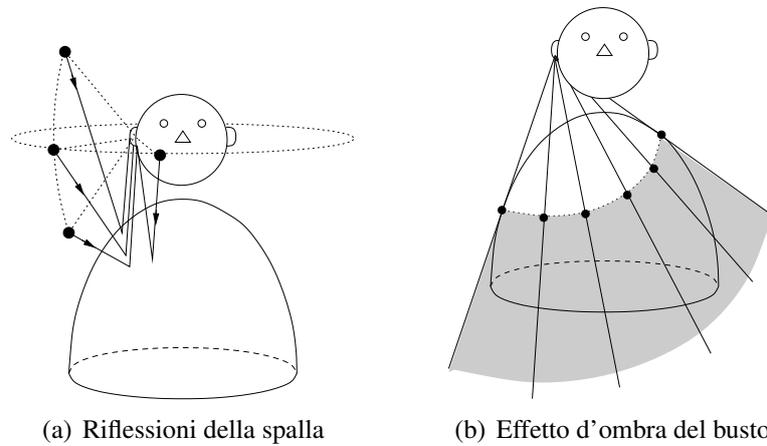
con un cilindro di larghezza costante, con pareti ad alta impedenza acustica. La pinna presenta delle sporgenze e delle cavità che tendono a sopprimere o ad accentuare alcune frequenze; in particolare essa agisce come un filtro la cui funzione di trasferimento dipende dalla direzione della sorgente sonora rispetto all'orecchio. Dalla sorgente al canale uditivo, il suono può arrivare tramite due percorsi: un percorso diretto ed uno indiretto soggetto a riflessioni della pinna. Se i segnali dai due cammini arrivano in fase il risultato è un aumento dell'energia sonora (interferenza costruttiva), mentre, se il segnale riflesso è fuori fase rispetto a quello diretto si generano delle interferenze distruttive, provocando un'attenuazione dell'energia sonora. Quando la differenza nella lunghezza dei due cammini è pari a metà lunghezza d'onda si genera la massima interferenza a cui viene dato il nome di *pinna notch*. Tali notch, come vedremo successivamente, sono indicatori essenziali per la percezione dell'elevazione.

Per quel che riguarda le risonanze, Shaw in [34] ha identificato sei modi di risonanza della pinna, eccitati con direzioni preferenziali diverse: una risonanza omnidirezionale che si verifica attorno alla frequenza di 4.2 kHz (modo 1), due risonanze verticali attorno alle frequenze di 7.1 e 9.6 kHz (modi 2 e 3) ed infine tre risonanze orizzontali attorno alle frequenze di 12.2, 14.4 e 16.7 kHz (modi 4, 5, 6). I modi di Shaw sono rappresentati in Figura 1.5.

La pinna ha un effetto trascurabile per frequenze al di sotto dei 3 kHz, dal momento che, in tale range, le lunghezze d'onda sono superiori alla dimensione della pinna e quindi non influenzate da essa.

#### 1.4.4 Busto e spalla

Il busto e la spalla influenzano le onde sonore in due aspetti principali:



**Figura 1.6:** *Effetti del busto e della spalla*

- forniscono riflessioni aggiuntive che si sommano al suono diretto;
- forniscono un effetto di shadowing per suoni provenienti dal basso.

Considerando la risposta impulsiva ad un orecchio, si osserverebbe un impulso iniziale seguito da una serie di impulsi successivi causati dalle riflessioni sulle spalle. Il ritardo tra il suono diretto e i raggi riflessi non varia molto se la posizione della sorgente sonora varia su una circonferenza nel piano orizzontale (soprattutto se il raggio è grande rispetto a quello della testa); il ritardo invece varia notevolmente se la sorgente sonora si muove verticalmente (in particolare il massimo ritardo per gli impulsi riflessi si ottiene quando la sorgente è posta esattamente sopra l'ascoltatore). Nel dominio della frequenza le riflessioni del busto agiscono come un comb filter, introducendo dei notch periodici nello spettro. Le frequenze in cui si verificano i notch sono inversamente proporzionali ai ritardi e così si produce un pattern che varia con l'elevazione della sorgente. La frequenza del notch più basso corrisponde al ritardo più lungo. I ritardi più lunghi di un sesto di un millisecondo produrranno uno o più notch al di sotto dei 3 kHz, che è approssimativamente la frequenza più bassa in cui gli effetti dell'orecchio esterno iniziano ad essere evidenti. Al calare dell'elevazione si raggiunge un punto di incidenza al di sotto del quale le riflessioni del busto scompaiono mentre emergono dei fenomeni di shadowing dovuti ad un vero e proprio cono d'ombra che avvolge l'orecchio più distante. Come discusso precedentemente per la testa, un ostacolo incide maggiormente per i suoni ad alta frequenza (approssimativamente superiori a 3 kHz nel caso di busto e spalle), mentre per quelli a bassa frequenza la diffrazione e lo scattering garantiscono comunque il raggiungimento dell'orecchio controlaterale. Sebbene gli effetti acustici di busto e spalle non siano intensi come quelli introdotti dall'orecchio esterno, sono comunque importanti perché appaiono a frequenze più basse, dove i segnali sonori tipici hanno la maggior parte della loro energia e dove la risposta della pinna è essenzialmente piatta. In termini di intervalli di frequenza gli effetti previsti dal busto sono quindi complementari a quelli forniti dall'orecchio esterno. In Figura 1.6 sono schematizzati gli effetti della spalla e del busto.

### 1.4.5 Teoria Duplex

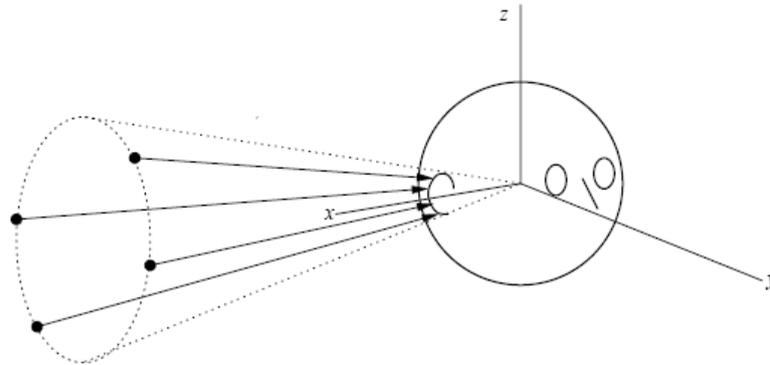
Abbiamo precedentemente introdotto i concetti di ITD e ILD, parametri chiave su cui si basa la Teoria Duplex sviluppata nel 1907 dal fisico Lord Rayleigh. Essa permette di spiegare in che modo l'uomo riesca a discriminare suoni posizionati nel piano orizzontale. Nello studiare i mezzi attraverso cui un ascoltatore riesce a percepire la direzione di un suono, Rayleigh trovò di fondamentale importanza i parametri di ITD e ILD, dovuti al ruolo attivo della testa nella differenziazione dei due flussi sonori in ingresso ai canali uditivi.

Considerando un segnale sinusoidale a bassa frequenza (fino a 1.5 kHz circa) si ha che la sua lunghezza d'onda è maggiore del raggio della testa, dunque la ITD si riduce ad una differenza di fase tra due segnali sinusoidali, utile per discriminare la direzione del suono nel piano orizzontale. Al contrario la ILD risulta trascurabile dal momento che a basse frequenze la funzione di trasferimento della testa risulta piatta. Considerando invece un segnale sinusoidale ad alta frequenza (da 1.5 a 10 kHz), con lunghezza d'onda dell'ordine di centimetri, si ottiene un ITD maggiore di un periodo del segnale. Essendo l'orecchio umano sensibile soltanto alla fase e non alla differenza temporale assoluta, la ITD risulta poco utile nel suddetto campo di frequenze, a meno che non si consideri l'onset del suono. Al contrario, la ILD diventa non più trascurabile a causa dell'effetto schermante causato dalla testa che provoca un aumento di valore di tale parametro, considerando la diminuzione di energia del suono che giunge all'orecchio più distante.

Purtroppo l'informazione che si ottiene dai parametri ITD e ILD può essere ambigua, poichè sorgenti sonore posizionate in qualsiasi punto di una superficie conica che punta verso l'orecchio dell'ascoltatore producono gli stessi valori per i due parametri. In realtà tali parametri non saranno mai identici a causa delle asimmetrie facciali e alla presenza delle orecchie. Queste superfici vengono di solito indicate con il termine di *coni di confusione* (mostrate in Figura 1.7) ed estendono il concetto di *front-back confusion*. Quest'ultimo è un tipo di errore molto comune in cui una sorgente posizionata di fronte all'ascoltatore ad un certo azimut  $\theta$ , ed una seconda sorgente posizionata sul retro ad azimut pari a  $\pi - \theta$ , presentano valori molto simili di ITD e ILD e di conseguenza gli ascoltatori posizionano sul retro sorgenti poste davanti e viceversa. La Teoria Duplex quindi non spiega in alcun modo come si possa determinare la posizione del suono al di fuori del piano dell'azimut. Infatti, nel caso di una sorgente posta sul piano mediano, i valori di ITD e ILD sono prossimi allo zero e quindi l'apporto dato da tali parametri è nullo. Allo stesso tempo, l'uomo è in grado di discriminare suoni posti su tale piano e da ciò ne consegue che altre caratteristiche utili alla distinzione entrano in gioco. Da un punto di vista fisico, il suono che arriva alle due orecchie è uguale e quindi le informazioni utili per discriminare la posizione della fonte sonora vanno ricercate nello spettro del segnale. Purtroppo, la relazione tra spettro del segnale e posizione del suono nello spazio non è semplice come la relazione con ITD/ILD.

### 1.4.6 Percezione dell'elevazione

La percezione della direzione di un suono sul piano mediano risulta più difficile rispetto a quella sul piano orizzontale. Infatti, se per la percezione sul piano orizzontale il *localization blur* (la differenza minima di angolazione tra due posizioni spaziali perchè si possa avvertire una diffe-



**Figura 1.7:** *Cono di confusione*

renza di posizione) assume valori di circa  $1^\circ$ - $2^\circ$ , per il piano mediano tale soglia raggiunge valori di  $4^\circ$  o superiori, come indicato in [8]. Questo è sia dovuto al fatto che sul piano orizzontale i valori di ITD e ILD sono prossimi allo zero, perdendo quindi la loro utilità come indicatori di posizione, e sia a causa della necessità del contributo in alta frequenza (superiore a 4-5 kHz) per una corretta localizzazione verticale, come indicato in [20] e [4].

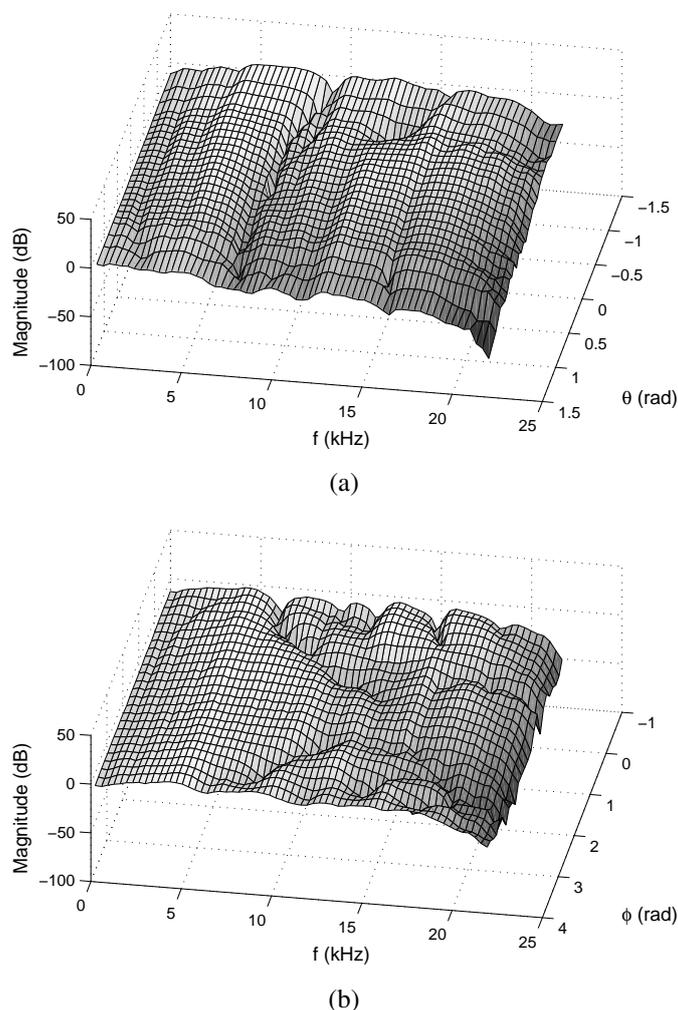
La pinna svolge un ruolo importantissimo nella localizzazione verticale, come è possibile verificare confrontando giudizi di localizzazione effettuati sia in condizioni normali che in assenza della pinna o con quest'ultima parzialmente ostruita. Un indicatore essenziale nella discriminazione dell'elevazione di un suono è il movimento di picchi e notch all'interno delle HRTF di un ascoltatore. La loro frequenza centrale, larghezza di banda e ampiezza varia in funzione della posizione della sorgente e dell'ascoltatore, come indicato in [35]; all'aumentare e al diminuire dell'elevazione, essi rispettivamente crescono o diminuiscono in frequenza, diventando meno pronunciati in profondità per suoni provenienti dall'alto, come mostrato in [20].

Tuttavia affinché si abbia un'accurata percezione dell'elevazione, è necessario che il suono contenga consistente energia nelle alte frequenze, poichè, come abbiamo precedentemente visto, la pinna ha dimensioni ridotte e lunghezze d'onda maggiori di essa non sono influenzate.

### 1.4.7 Head Related Transfer Function

Le HRTF sono dei filtri che modellano gli effetti sulla propagazione del suono dovuti alla testa, all'orecchio esterno, al busto e alla spalla. Infatti, gli effetti delle diverse parti del corpo sono lineari, e possono quindi essere descritti da funzioni di trasferimento e combinati assieme. La pressione sonora prodotta da una sorgente nel canale uditivo è determinata esclusivamente dalla risposta impulsiva nel tragitto dell'onda sonora dalla sorgente stessa al timpano, chiamata *Head Related Impulse Response* (HRIR), la cui trasformata di Laplace è proprio la HRTF. Esiste una HRTF riferita all'orecchio destro ed una all'orecchio sinistro, che, a meno di perfette simmetrie, sono tra loro diverse.

Per una sorgente acustica situata, rispetto all'ascoltatore, nel punto di coordinate  $(r, \theta, \phi)$ , avremo due HRTF, indicate con  $H^{(l)}(r, \theta, \phi, \omega)$  e  $H^{(r)}(r, \theta, \phi, \omega)$ , dove  $\theta$  è l'azimut,  $\phi$  è l'elevazione e  $\omega$  è la frequenza angolare e  $r$  è la distanza, riferite rispettivamente all'orecchio sinistro e



**Figura 1.8:** Esempi di HRTF (a) nel piano  $xy$  ( $\theta \in [-\pi/2, \pi/2]$ ,  $\phi = 0$ ) e (b) nel piano  $yz$  ( $\theta = 0$ ,  $\phi \in [-\pi/4, \pi]$ )

destro. Nel caso di sorgente in *campo lontano* ( $r > 1$  m), otteniamo le seguenti formule:

$$H^{(l)}(\theta, \phi, \omega) = \frac{\Phi^{(l)}(\theta, \phi, \omega)}{\Phi_f(\omega)}, \quad H^{(r)}(\theta, \phi, \omega) = \frac{\Phi^{(r)}(\theta, \phi, \omega)}{\Phi_f(\omega)} \quad (1.4)$$

La HRTF è data quindi dal rapporto tra il livello di pressione sonora (SPL)  $\Phi^{(l),(r)}(\theta, \phi, \omega)$  in prossimità del timpano e l'SPL in campo libero al centro della testa  $\Phi_f(\omega)$ , come se l'ascoltatore fosse assente.

In Figura 1.8 sono mostrati due esempi della risposta in ampiezza di HRTF, misurate sul piano orizzontale e sul piano mediano.

Le HRTF trovate per via sperimentale possono essere approssimate con filtri a fase minima; l'assunzione di fase minima ci consente di specificare in maniera univoca la fase della HRTF a

partire dalla sua risposta in ampiezza, ed inoltre ci permette di separare le informazioni sull'ITD dalle specifiche dei filtri che approssimano le HRTF.

Considerando l'andamento della risposta in ampiezza delle HRTF sul piano mediano al variare dell'angolo di elevazione, per tutte le curve si trova un massimo posizionato intorno ai 4 kHz, dovuto alla risonanza del condotto uditivo che non cambia con la posizione della sorgente al variare dell'elevazione, e, per frequenze superiori ai 5-6 kHz, si manifesta un avvallamento (notch) principale nella risposta in frequenza dovuto all'effetto della pinna. La frequenza centrale di tale notch varia al variare dell'elevazione, risultando di circa 6 kHz per basse elevazioni e di 10 kHz per alte elevazioni.

Considerando invece l'andamento della risposta in ampiezza delle HRTF, riferite ad esempio all'orecchio destro, sul piano orizzontale al variare dell'azimut si trova che, quando la sorgente è ad azimut di  $90^\circ$ , il livello è più elevato, mentre quando la sorgente è ad azimut di  $270^\circ$  il livello è più basso, questo perchè nel primo caso l'orecchio destro è ipsilaterale, si trova cioè nello stesso lato della sorgente, mentre nel secondo caso l'orecchio destro è controlaterale, si trova cioè nel lato opposto rispetto alla sorgente. Le stesse considerazioni valgono per HRTF riferite all'orecchio sinistro.

Tipicamente le HRTF sono misurate empiricamente attraverso l'uso di microfoni posti all'ingresso o all'uscita del canale uditivo nelle orecchie dell'ascoltatore reale o di un manichino (*dummy head*). Tali misurazioni avvengono di solito in una camera anecoica <sup>2</sup>, in cui il soggetto si colloca in un punto preciso della stanza dove udirà un segnale di stimolo prodotto da un altoparlante che viene via via posizionato a diverse angolazioni di azimut ed elevazione. Quando si eseguono delle misurazioni utilizzando un manichino KEMAR <sup>3</sup>, per registrare le risposte allo stimolo, si sfruttano i microfoni incapsulati all'imbocco degli ipotetici canali uditivi. Nel caso in cui il soggetto sia un individuo umano, questo viene attrezzato con dei piccoli microfoni sonda installati all'imbocco del canale uditivo di entrambe le orecchie, così da poter captare le risposte agli stimoli sonori. Alla registrazione segue una fase di post-processing al fine di rimuovere i contributi delle funzioni di trasferimento dell'apparecchiatura di misurazione (come il microfono e l'altoparlante). Sebbene tali misure consentano una codifica posizionale molto accurata, il costo computazionale risulta piuttosto elevato, così come il costo in termini di apparecchiature e infrastrutture.

Le HRTF possono anche essere determinate da modelli matematici, tra cui quelli più semplici sono derivati dalla soluzione di equazioni acustiche delle onde che investono una sfera rigida (la testa). Si determina la pressione presente in due punti della sfera che rappresentano le orecchie per differenti frequenze ed angoli di elevazioni calcolando così le HRTF destre e sinistre. Ad esempio in [2], gli autori propongono un modello fisico detto *physical snowman model* (pupazzo di neve) basato su una schematizzazione della testa e del torso fatta con due sfere rigide sovrapposte. Tecniche più elaborate consentono di calcolare le HRTF direttamente dalla superficie della testa senza orecchie, costruita attraverso il *Boundary Element Method* (BEM) sotto forma di una mesh di elementi discreti, come mostrato in [23].

<sup>2</sup>La camera anecoica è un ambiente di laboratorio strutturato in modo da ridurre il più possibile la riflessione dei segnali sulle pareti.

<sup>3</sup>Il KEMAR è un modello artificiale dotato di busto e testa utilizzato nella ricerca acustica da cui è possibile rimuovere con facilità la pinna.

Infine troviamo i modelli determinati a partire dalle HRTF misurate in cui vengono analizzate le caratteristiche più significative ed indipendenti dall'ascoltatore. L'analisi delle HRTF in genere viene eseguita per mezzo di analisi statistiche, tra cui una piuttosto usata è basata sull'Analisi alle Componenti Principali (PCA) [24].

Per quanto riguarda le differenze tra HRTF teoriche e misurate, si ha che mentre le HRTF teoriche sono prive di rumore, il *signal to noise ratio* (SNR) di quelle misurate sembra essere una funzione della posizione nello spazio. Lo SNR delle sorgenti ipsilaterali è generalmente più elevato rispetto alle sorgenti controlaterali, questo perchè l'orecchio controlaterale riceve meno energia rispetto a quello ipsilaterale. Inoltre le HRTF misurate risultano più complesse rispetto a quelle teoriche, presentando più picchi e avvallamenti nel modulo, dovuti all'interazione del suono con il torso e la pinna, fenomeni non considerati nel modello sferico.

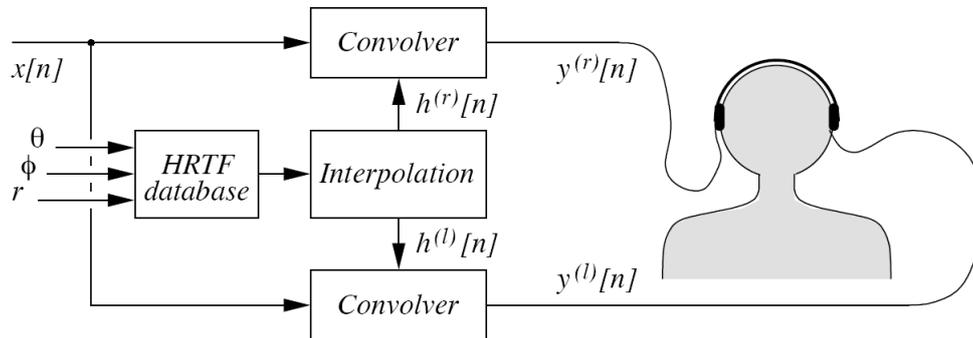
## 1.5 Rendering di audio 3D

Con suono virtuale 3D si intende la capacità di ricreare nell'ascoltatore l'illusione che un determinato suono sia posizionato in una determinata locazione virtuale, presentando ai due timpani dell'ascoltatore gli stessi segnali che produrrebbe una sorgente reale posizionata in quel particolare punto dello spazio. Come indicato in [25]:

L'obiettivo ideale per il progettista di un sistema di audio 3D è un controllo completo sullo spazio uditivo percepito da un ascoltatore: in altre parole, l'operatore di un sistema audio 3D dovrebbe essere in grado di predire, entro il livello ammesso dalla statistica e ignorando di conseguenza le specificità individuali, la mappa sonora che l'ascoltatore inferirà usando il sistema di audio 3D.

Nella resa del suono spazializzato (*auralizzazione*), bisogna considerare l'ambiente da simulare, la posizione dell'ascoltatore e quella della sorgente e calcolare le possibili propagazioni dell'onda sonora nell'ambiente, costruendo infine dei filtri digitali che rappresentano i possibili ritardi e le possibili attenuazioni che il suono subisce nel suo cammino verso l'ascoltatore lungo le direzioni di propagazione. Infatti, come visto precedentemente, il suono che arriva all'ascoltatore è dato dalla somma del suono diretto, ovvero il suono che non subisce riflessione, il suono riflesso che arriva entro poche decine di millisecondi quando le riflessioni sono ancora poche, ed il suono riverberato dovuto alle riflessioni sulle pareti dell'ambiente circostante. Infine, combinando i suoni emessi dalla sorgente con tali filtri digitali, è possibile riprodurre il suono spazializzato. I filtri digitali corrispondono alle HRIR per orecchio sinistro e destro introdotte precedentemente.

La riproduzione di suono spazializzato può avvenire sia per mezzo di altoparlanti (audio *transaurale*) che di cuffie (audio *binaurale*). Nella prima soluzione troviamo i sistemi *stereo* a due canali ed i sistemi *multicanale*. Il sistema stereo è la più semplice tecnologia a prevedere audio spazializzato. I segnali destro e sinistro vengono inviati ai corrispondenti altoparlanti, e nel caso di segnali uguali (altoparlanti "in fase") e se l'ascoltatore è equidistante dai due altoparlanti, allora quest'ultimo percepirà un "suono fantasma" localizzato in una posizione intermedia rispetto ai due speaker. In tale sistema è possibile muovere la sorgente da un altoparlante all'altro. Tuttavia



**Figura 1.9:** Schema a blocchi di un sistema di rendering di audio 3D per mezzo di cuffie

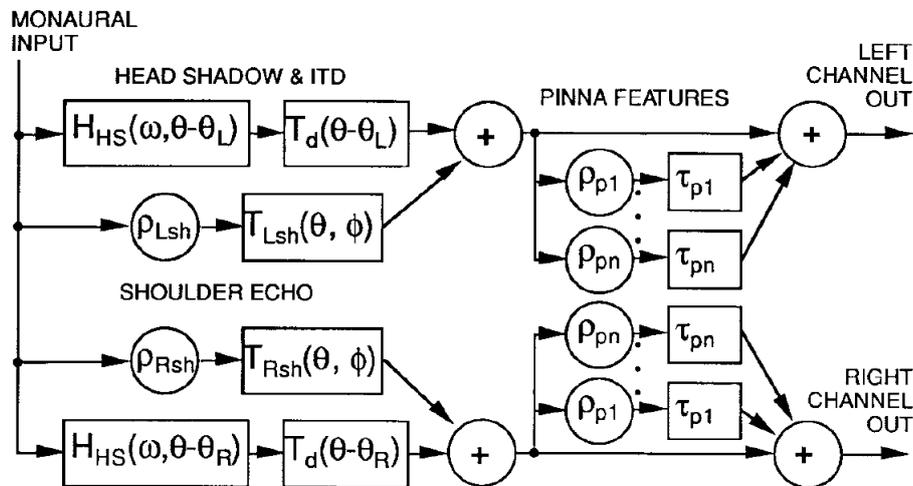
il suono dei due altoparlanti arriva ad entrambe le orecchie (*crosstalking*), con differenze di tempo e di intensità a causa del maggior cammino che l'onda deve seguire per raggiungere l'orecchio più lontano e all'effetto schermante della testa. Per questo motivo è necessario che l'ascoltatore si trovi alla distanza corretta dai due altoparlanti (*sweet spot*) altrimenti non viene percepita la sorgente “fantasma” pensata dal progettista del suono. Per quanto riguarda invece i sistemi multicanali, l'idea di base è quella di riservare un canale separato per ogni direzione. Il sistema più diffuso è il Dolby 5.1, composto da 5 canali (3 frontali e 2 surround) più un canale per i suoni a bassa frequenza da 20 a 120 Hz (denominato LFE, *low frequency effects*). Anche tale sistema ha lo svantaggio di essere influenzato dalla posizione degli altoparlanti e, come nel caso del sistema stereo, solo una piccola zona spaziale ed un'unica orientazione dell'ascoltatore permettono la creazione di un corretto segnale binaurale ai timpani. Sia il sistema stereo che quello multicanale risentono dell'effetto riverberante dell'ambiente in cui avviene la riproduzione.

Al contrario, le cuffie permettono di eliminare il contributo dell'ambiente in cui avviene la riproduzione e l'effetto di *crosstalking*, permettendo così al progettista del suono un totale controllo del campo sonoro. Per produrre effetti di audio 3D in cuffia è necessario fare uso delle funzioni di trasferimento della testa (HRTF), viste in precedenza. Dato un segnale anecoico ed una posizione virtuale ( $\theta, \phi, r$ ) che si intende ricreare, vengono sintetizzati i segnali destro e sinistro attraverso due fasi:

1. il segnale anecoico viene ritardato di una certa quantità, in modo da introdurre l'ITD desiderato;
2. il segnale anecoico viene convoluto con le risposte all'impulso riferite all'orecchio destro e a quello sinistro.

Infine i due segnali vengono inviati agli appositi canali destro e sinistro. In questo modo la sensazione percepita dall'ascoltatore sarà quella di una sorgente virtuale nella posizione spaziale indicata dalla coppia di HRTF utilizzata (*sintesi binaurale*). In Figura 1.9 è rappresentato tale schema.

Sfortunatamente le HRTF non sono universali e quindi per ogni ascoltatore si deve procedere con delle misurazioni ad hoc che richiedono una strumentazione adatta e una procedura di post-processing elaborata. Per queste ragioni si ricorre, come abbiamo precedentemente visto, a

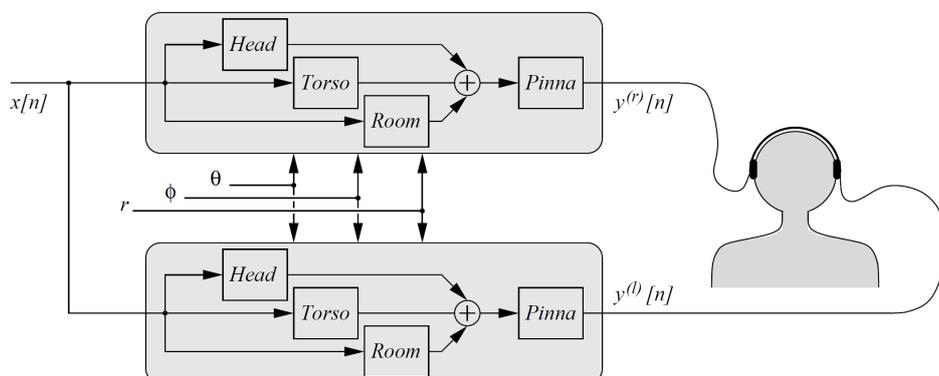


**Figura 1.10:** Modello strutturale della HRTF proposto da Brown e Duda

HRTF generiche misurate su di un manichino o su altri individui. Una serie di esperimenti furono condotti da Wenzel in [41] in modo da valutare l'efficacia di HRTF non individuali. Sebbene la percezione sul piano orizzontale fosse accurata, l'uso di HRTF non individuali produceva un alto tasso di errori di *front/back confusion* (suoni percepiti nell'emisfero sbagliato). In [7], Begault ha confrontato gli effetti di HRTF generiche ed individuali sul parlato sia in condizioni statiche che dinamiche. I risultati mostravano che la localizzazione sul piano orizzontale in condizioni statiche con HRTF generiche era meno accurata rispetto a quella con HRTF individuali e che i movimenti della testa erano cruciali per ridurre gli angoli di errore nel piano verticale ed il numero di *front/back confusion*. Quindi sebbene l'impiego di HRTF non individuali rappresenti una tecnica semplice ed economica per riprodurre audio spazializzato in cuffia, esso può comportare artefatti percettivi che degradano l'accuratezza nella localizzazione, sensazioni di estraneità e il sorgere di un particolare fenomeno, detto *internalizzazione*, per cui le sorgenti sono localizzate all'interno della testa. Particolari criticità si hanno con gli indicatori dell'elevazione che non possono essere contraddistinti da caratteristiche spettrali generalizzate, dimostrando il ruolo chiave dell'antropometria nella caratterizzazione delle HRTF individuali.

### 1.5.1 Modello strutturale

Il rendering di suoni spazializzati basato su HRTF può avvenire in diversi modi. Sono state proposte approssimazioni delle HRTF complessive basate su funzioni razionali (*modelli polo/zero*) o su espansioni in serie di componenti principali, come in [24] o di armoniche, come in [13]. Tuttavia la complessità dei coefficienti dei filtri e dell'insieme dei pesi derivanti dalle analisi le rende inadatte per un uso in real-time. In [9], gli autori Brown e Duda propongono un metodo per approssimare le HRTF mediante l'uso di filtri digitali. L'idea innovativa alla base di tale modello è la possibilità di combinare in maniera additiva i contributi dati alle HRTF dalle singole parti anatomiche, che, come visto in precedenza, sono lineari. Gli effetti del torso, della testa e della pinna sono trattati separatamente e poi combinati in modo da ottenere la HRTF comples-



**Figura 1.11:** Sistema di riproduzione di audio 3D basato su un modello strutturale di HRTF

siva. I parametri dei filtri digitali associati ad ogni blocco anatomico possono essere relazionati a misure antropometriche effettuate sull'ascoltatore. In Figura 1.10 è rappresentato il modello strutturale completo proposto da Brown e Duda, mentre in Figura 1.11 è raffigurato un generico sistema audio binaurale basato su un modello strutturale.

Si ottiene in tal modo un modello di HRTF allo stesso tempo economico ed adattabile allo specifico ascoltatore. In esso è possibile incorporare anche gli effetti legati all'ambiente di riproduzione.

In questa tesi ci siamo occupati esclusivamente del contributo dato dalla pinna, posizionando le sorgenti da testare tutte sul piano mediano, dove il contributo dato dalla testa è trascurabile, ed eliminando, tramite finestramento applicato sulle HRIR, il contributo dato dal busto e spalla. Vedremo nel capitolo successivo come ricavare da semplici misure antropometriche i valori dei parametri del filtro associato alla pinna.

# Capitolo 2

## Modellazione di PRTF sintetiche

In questo capitolo vengono descritti gli studi alla base della modellazione del contributo della nostra pinna ed i parametri sperimentali utilizzati. Nella Sezione 2.1 sono riportati alcuni lavori precedenti, presenti in letteratura, sulla sintesi di PRTF sintetiche. Nella Sezione 2.2 è descritto dettagliatamente il modello di PRTF sintetica alla base della presente tesi. Infine, nella Sezione 2.3 viene presentato il modello realizzato per gli esperimenti psicoacustici effettuati in questo lavoro di tesi.

### 2.1 Stato dell'arte

Secondo la cosiddetta *modellazione strutturale*, descritta in [9], l'HRTF può essere scomposta in contributi tra loro indipendenti:

- *Testa*;
- *Busto e Spalla*;
- *Padiglione auricolare (Pinna)*.

Possiamo quindi considerare la pinna come un oggetto indipendente dalle altre parti anatomiche e modellare il contributo separatamente. La risposta della pinna viene indicata con PRTF (Pinna Related Transfer Function).

Una volta fissata la direzione della sorgente relativa all'ascoltatore, le differenze più marcate tra le HRTF sono dovute alle caratteristiche della pinna, quest'ultima diversa da persona a persona. La pinna ha un ruolo fondamentale nel determinare il contenuto in frequenza delle HRTF grazie a due fenomeni acustici principali:

1. riflessioni lungo i bordi della pinna. Secondo Batteau [5], le onde sonore vengono tipicamente riflesse dall'orecchio esterno finché la loro lunghezza d'onda è abbastanza piccola rispetto alle dimensioni della pinna stessa, e l'interferenza causata dalle onde provenienti direttamente dalla sorgente e quelle riflesse fa sì che, nella parte delle alte frequenze, compaiano delle acute attenuazioni (notch) nello spettro del segnale ricevuto con periodicità inversamente proporzionale al tempo di ritardo di ogni riflessione;

2. risonanze nelle cavità della pinna. Come discusso da Shaw [34], poichè la conca si comporta come un risonatore, alcune bande di frequenze, sia delle onde dirette che di quelle riflesse, vengono amplificate in modo significativo in base all'elevazione della sorgente.

Di conseguenza, la parte della HRTF dovuta al contributo della pinna presenta nella sua ampiezza una sequenza di picchi e notch.

Nella letteratura troviamo molti studi sulla sintesi delle PRTF. Le osservazioni di Batteau hanno portato ad un modello di pinna [39] a doppio cammino con l'obiettivo di simulare i percorsi, diretto e riflesso, che l'onda sonora percorre. Pur fornendo risultati soddisfacenti, tale modello tende a sovrastimare il numero di notch presenti nello spettro.

Un simile approccio fu adottato da Barreto in [14] con l'introduzione di quattro percorsi paralleli che rappresentano diversi cammini che l'onda sonora segue in ragione delle riflessioni causate dalla pinna. Ognuno di questi tragitti è caratterizzato da un tempo di ritardo e da un fattore d'ampiezza che indica la perdita di energia. Inoltre è stato introdotto nel modello un blocco risonatore con lo scopo di simulare il comportamento da risonatore delle cavità della pinna. I parametri del modello sono stati associati ad otto caratteristiche antropometriche. Purtroppo questo approccio, per la sua applicazione, richiede l'uso di uno scanner 3D, disponibile solo in alcuni laboratori di ricerca.

Un altro approccio è quello proposto da Raykar [30]. Egli ha riscontrato che le riflessioni di testa, torso, pinna e ginocchia possono essere osservate sia nel tempo che in frequenza. Nel dominio del tempo, come mostrato in Figura 2.1, possiamo notare tre creste date dall'interazione del suono con gli ostacoli incontrati nel suo tragitto: la diffrazione della testa, gli effetti della pinna, la riflessione del torso e la riflessione delle ginocchia. Grazie a robuste tecniche basate sui residui delle predizioni lineari delle HRIR, gli autori sono riusciti ad estrarre le frequenze dei notch dovuti alla pinna e relazionarli con la forma stessa della pinna. Conoscendo il ritardo temporale  $t_d$ , dipendente dall'elevazione, tra l'onda diretta e l'onda riflessa, è possibile calcolare il punto di riflessione ad una distanza dal canale uditivo pari a:

$$d(\phi) = \frac{ct_d(\phi)}{2} \quad (2.1)$$

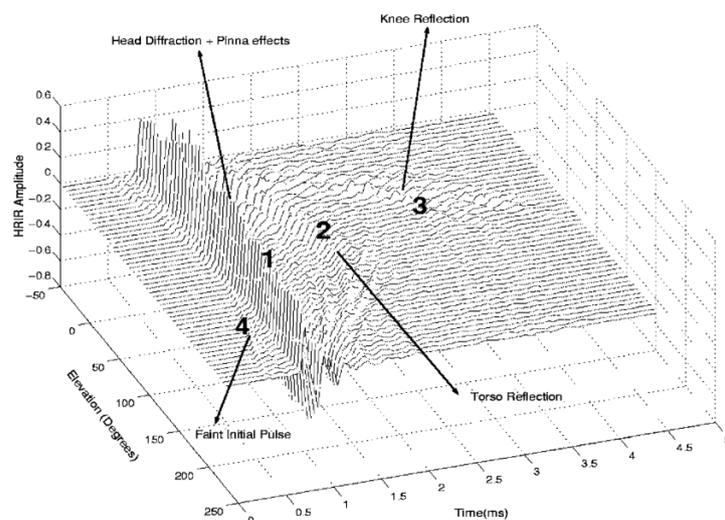
con  $c$  velocità del suono (circa 343 m/s). Considerando un coefficiente di riflessione positivo, ogni frequenza estratta  $f_0$  è la prima di una serie periodica:

$$f_n = \frac{(2n+1)}{2t_d(\phi)} = \frac{c(2n+1)}{4d(\phi)} \quad n=0,1,\dots, \quad (2.2)$$

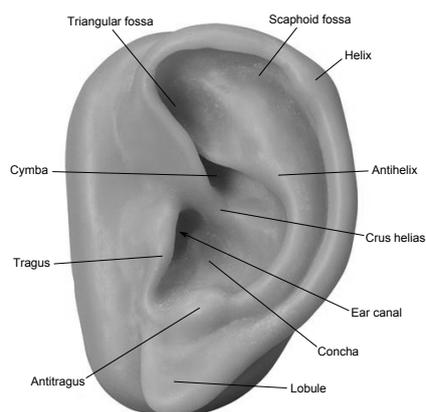
$$f_0(\phi) = \frac{c}{4d(\phi)} \quad (2.3)$$

La distanza  $d(\phi)$  è stata proiettata sull'immagine 2-D della pinna, di cui una rappresentazione è data in Figura 2.2 con indicate le principali superfici di riflessione: l'elice (*Rim* o *Helix*), l'antielice (*Antihelix*), la conca (*Concha*) e il *Crus Helias*. I punti di riflessione trovati dagli autori risultano cadere sul *crus helias* e sulla parete della conca.

Un altro contributo importante è stato introdotto da Satarzadeh [31]. In questo lavoro, le PRTF con elevazione intorno allo zero sono state sintetizzate attraverso un modello composto da



**Figura 2.1:** Individuazione nelle HRIR dell'interazione tra suono e ostacoli



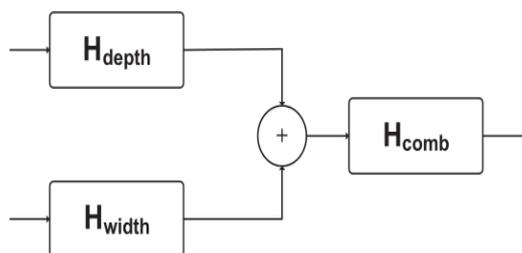
**Figura 2.2:** Anatomia della pinna

due filtri passa banda del secondo ordine e da un filtro a pettine. Questi servono a rappresentare rispettivamente le due maggiori risonanze (modi risonanti 1 e 4 di Shaw) e la riflessione principale. I due filtri passa banda del secondo ordine ed il filtro a pettine sono interconnessi come in Figura 2.3. Il filtro a pettine è della forma  $[1 + \rho \exp(-st_d)]$ , con  $\rho$  coefficiente di riflessione e  $t_d$  tempo di ritardo tra l'onda diretta e quella riflessa, calcolato a partire dalla frequenza  $f_0$  stimata dalla spaziatura di notch consecutivi nella PRTF:

$$t_d = \frac{1}{2f_0} \quad (2.4)$$

se  $\rho > 0$ , oppure:

$$t_d = \frac{1}{f_0} \quad (2.5)$$



**Figura 2.3:** Modello di PRTF proposto da Satarzadeh

se  $\rho < 0$ .

Per quel che concerne la parte risonante, la conca viene approssimata con un cilindro, dove la profondità e la larghezza del cilindro definiscono univocamente la profondità della risonanza, mentre la larghezza della risonanza è correlata al tempo di ritardo  $t_d$  a seconda che la superficie riflettente primaria sia la conca o l'elice. Quindi, secondo Satarzadeh, nel caso di conca rappresentabile con un cilindro che presenti una superficie riflettente dominante, sia essa la conca o l'elice, il suo filtro di basso ordine basato sull'antropometria offre una buona approssimazione sia per PRTF composte da molti notch sia per PRTF con pochi avvallamenti.

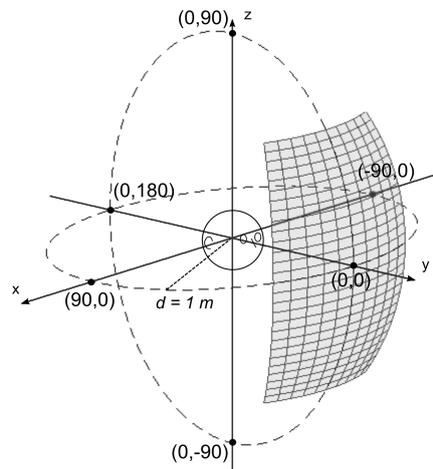
Nel Paragrafo successivo viene presentata la modellazione realizzata da Spagnol, Geronazzo e Avanzini, alla base del lavoro di questa tesi.

## 2.2 Modello strutturale di PRTF

Partendo dal lavoro di Satarzadeh, in [15] gli autori hanno realizzato un algoritmo che permette di scomporre la PRTF in due componenti: componente risonante e componente riflessiva. Sono state considerate le PRTF di 20 dei 45 soggetti presenti nel CIPIC, database di pubblico dominio contenente HRIR misurate per 1250 direzioni spaziali<sup>1</sup>. Di questi 20 soggetti è infatti disponibile l'immagine della pinna. Prendendo come sistema di riferimento il sistema di coordinate polari interaurali rappresentato in Figura 2.4, gli autori si sono focalizzati sulle HRIR sul piano mediano (angolo di azimut  $\theta = 0$ ) con elevazione da  $-45$  a  $90$  gradi a passi di  $5.625^\circ$ . La scelta di considerare il piano mediano nasce dall'osservazione che, come indicato dagli autori, fissata l'elevazione, variazioni di azimut fino ad un massimo di  $\Delta\theta = 30^\circ$  causano, nelle PRTF, variazioni spettrali non rilevanti. In questa regione le variazioni dovrebbero essere, quindi, dipendenti solo dall'elevazione. Per quanto riguarda il limite superiore di elevazione considerata ( $\phi = 45^\circ$ ), la scelta degli autori è dovuta all'alto tasso di incertezza nel giudicare elevazioni  $\phi > 45^\circ$ , come mostrato in [8] e in [28], e alla scarsa presenza, nelle PRTF, di profondi notch per tali elevazioni. La regione di validità del modello proposto dagli autori è raffigurata in Figura 2.4.

Le HRIR considerate sono state inizialmente finestrate con una finestra di Hann in modo da non considerare il segnale dopo 1 ms, eliminando così dalle HRIR gli effetti dovuti al torso, spalla e ginocchia. Per ciò che concerne la testa invece, considerando che la risposta in frequenza

<sup>1</sup><http://interface.cipic.ucdavis.edu/sound/hrtf.html>



**Figura 2.4:** Sistema di coordinate polari interaurali e regione di validità del modello

di una testa sferica senza pinne è piatta, non è stato preso nessun accorgimento. Dalle HRIR così finestrate nel dominio del tempo, si è passati al dominio della frequenza ottenendo le PRTF. L'algoritmo sviluppato, il cui diagramma di flusso è mostrato in Figura 2.5, è basato sull'idea di compensare iterativamente lo spettro del modulo della PRTF con un filtro multi-notch fino a non rilevare più notch significativi. Una volta raggiunta l'approssimazione desiderata, per esempio all'iterazione  $n$ -esima, lo spettro della componente  $H_{res}^{(n)}$  conterrà la componente risonante, mentre la combinazione dei filtri multi-notch fornirà la componente riflessiva  $H_{refl}^{(n)}$ . I dettagli dell'algoritmo sono descritti in [15].

L'algoritmo produce una decomposizione realistica, almeno nell'intervallo di frequenza 3-18 kHz. Per mezzo di un sistema di identificazione basato su un modello ARMA del sesto ordine, gli autori sono riusciti a ricavare le frequenze centrali delle risonanze. Come è possibile vedere in Figura 2.6, dove sono raffigurate le risonanze ottenute per il Soggetto 010 e il Soggetto 165 del CIPIC, vi sono due principali aree con forte guadagno. La prima, intorno ai 4 kHz, rimane costante per tutte le elevazioni e per tutti i soggetti del database. Essa contiene il modo omnidirezionale 1 di Shaw e la sua larghezza di banda apparentemente aumenta al crescere dell'elevazione. Dalla figura ricaviamo che una seconda risonanza interferisce all'interno di questo range di frequenza, in particolare il modo 2 di Shaw che appare intorno ai 7 kHz. La seconda area differisce tra soggetti sia in ampiezza che nella forma ed è più marcata a basse elevazioni tra 12 e 18 kHz, range di frequenza associato con i modi orizzontali 4,5 e 6 di Shaw. Le risonanze a più alte frequenze sono irrilevanti, poichè al limite dell'intervallo di frequenze udibili.

In Figura 2.7 è rappresentata la componente riflessa dei soggetti CIPIC 010 e 048. In particolare sono evidenziate le frequenze dei notch, al variare dell'elevazione. Rispetto alla componente risonante, notiamo che le frequenze dei notch sono dipendenti dall'elevazione e dalla forma della pinna. Per sorgenti sopra la testa dell'ascoltatore non notiamo notch nella PRTF, mentre, al diminuire dell'elevazione, vi è la comparsa di notch con profondità e larghezza di banda dipendente proprio da quest'ultima. Per il rilevamento dei notch si è usato l'algoritmo di tracciamento delle parziali McAulay-Quatieri, usato originariamente per raggruppare le parziali di sinusoidi tra più

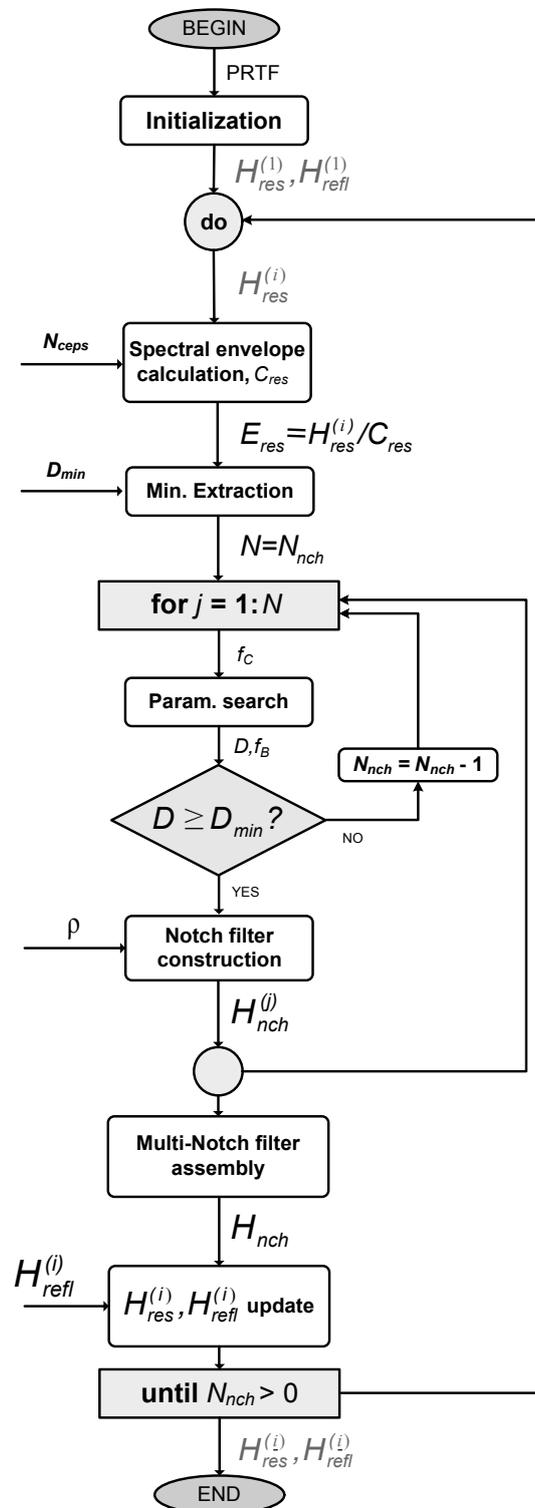
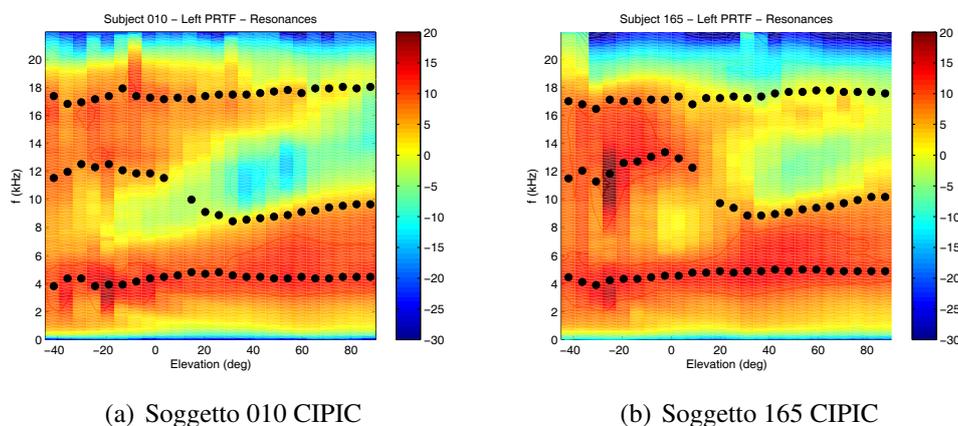
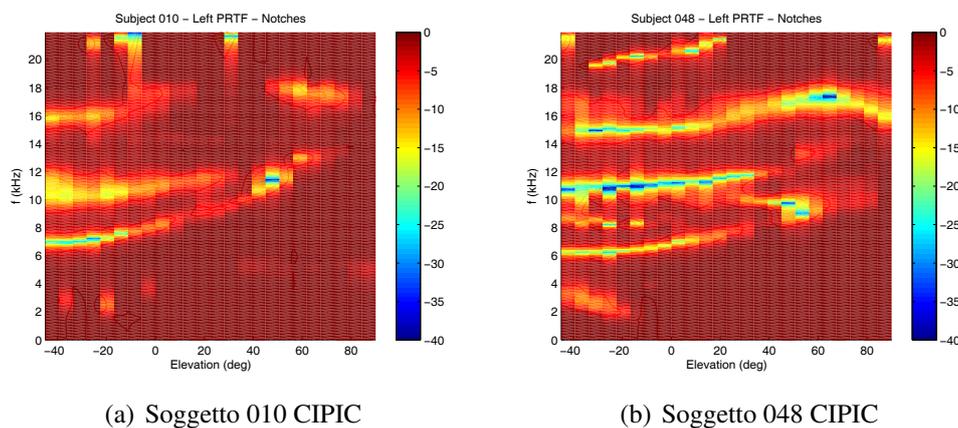


Figura 2.5: Diagramma di flusso dell' algoritmo di separazione



**Figura 2.6:** Risonanze per differenti elevazioni

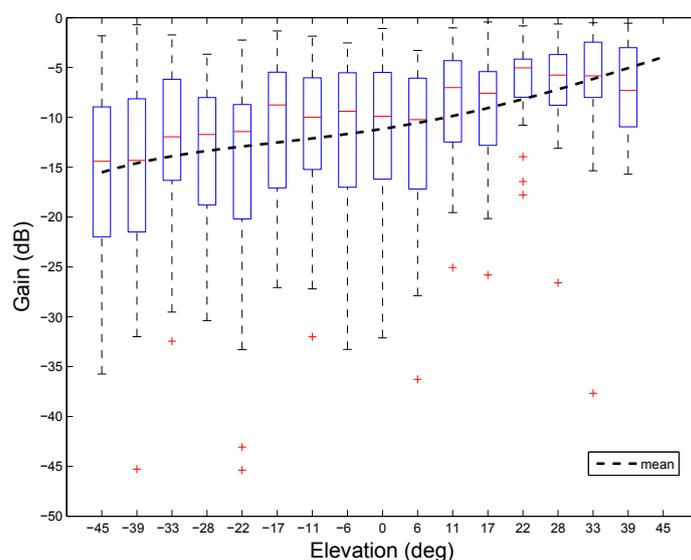


**Figura 2.7:** Notch spettrali per differenti elevazioni

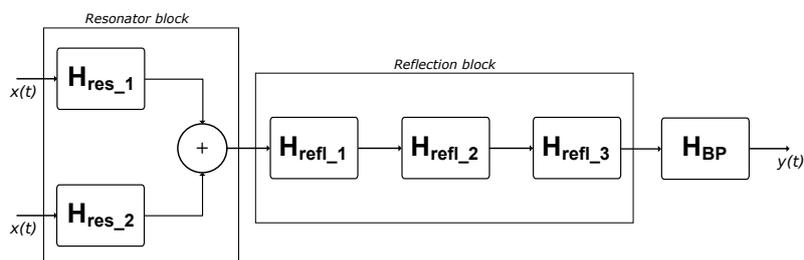
finestre temporali contigue, in funzione della loro posizione spaziale. Esso è stato modificato, in modo da sostituire il tempo con l'elevazione e le parziali con i notch. Dai risultati di tale algoritmo sono risultate, per la gran parte dei soggetti presi in considerazione, tre diverse tracce nelle PRTF, scartando dal computo le tracce al di fuori dell'intervallo  $4 \div 14$  kHz, e quelle che non presentano un avvallamento superiore a 5dB. In Figura 2.8 è rappresentato il box plot e la media delle profondità della prima traccia notch estratta, con valori mediati tra i soggetti CIPIC considerati nel modello. Vediamo che i valori di profondità variano in funzione dell'elevazione. All'aumentare dell'elevazione si notano valori di profondità minori. Essi inoltre variano da soggetto a soggetto e ad oggi non sono stati trovati legami tra la profondità dei notch e le misure antropometriche.

Proprio partendo da queste osservazioni, gli autori hanno realizzato un modello strutturale di pinna costituito da due principali blocchi:

- *Blocco risonante;*



**Figura 2.8:** Box plot e media delle profondità del primo notch tra soggetti CIPIC



**Figura 2.9:** Modello strutturale della PRTF

- *Blocco riflettente.*

Il blocco risonante è costituito da due filtri peak del secondo ordine in parallelo con all'ingresso lo stesso segnale. L'uscita dei due filtri viene sommata per poi essere inviata all'ingresso del blocco riflettente, costituito dalla cascata di 3 filtri notch del secondo ordine. Opzionalmente è possibile aggiungere alla fine dello schema un filtro passa banda  $H_{BP}$ , utile in caso di interconnessione con altri blocchi (ad es. la testa). In Figura 2.9 è mostrato lo schema del modello strutturale.

### 2.2.1 Filtri

Per ciò che riguarda i filtri notch, i tre parametri necessari sono:  $f_C$ ,  $D$ ,  $f_B$ , dove con  $f_C$  indichiamo la frequenza centrale del notch, con  $D$  la profondità e con  $f_B$  la larghezza di banda. Avendo

a disposizione questi dati è stato possibile costruire un filtro del secondo ordine del tipo:

$$H_{nch} = \frac{1 + (1+k)\frac{H_0}{2} + l(1-k)z^{-1} + (-k - (1+k)\frac{H_0}{2})z^{-2}}{1 + l(1-k)z^{-1} - kz^{-2}} \quad (2.6)$$

con

$$k = \frac{\tan\left(\pi\frac{f_B}{f_s}\right) - V_0}{\tan\left(\pi\frac{f_B}{f_s}\right) + V_0} \quad (2.7)$$

$$l = -\cos\left(2\pi\frac{f_C}{f_s}\right) \quad (2.8)$$

$$V_0 = 10^{\frac{D}{20}} \quad (2.9)$$

$$H_0 = V_0 - 1 \quad (2.10)$$

dove  $f_s$  indica la frequenza di campionamento.

Per quel che riguarda i filtri risonanti, gli autori hanno utilizzato un filtro peak IIR del secondo ordine. Con i dati di  $f_C$ , frequenza centrale,  $G$ , guadagno e  $f_B$ , larghezza di banda del picco, ottenuti dall'algoritmo di separazione di cui sopra, è possibile costruire per il primo picco un filtro risonante duale di quello utilizzato per i notch:

$$H_{res}^1(z) = \frac{1 + \frac{H_0}{2}(1+k) + d(1-k)z^{-1} + (-k - \frac{H_0}{2}(1+k))z^{-2}}{1 + d(1-k)z^{-1} - kz^{-2}} \quad (2.11)$$

$$d = -\cos\left(2\pi\frac{f_C}{f_s}\right) \quad (2.12)$$

$$V_0 = 10^{\frac{G}{20}} \quad (2.13)$$

$$H_0 = V_0 - 1 \quad (2.14)$$

$$k = \frac{\tan\left(\pi\frac{f_B}{f_s}\right) - 1}{\tan\left(\pi\frac{f_B}{f_s}\right) + 1} \quad (2.15)$$

mentre per il secondo picco un filtro del tipo :

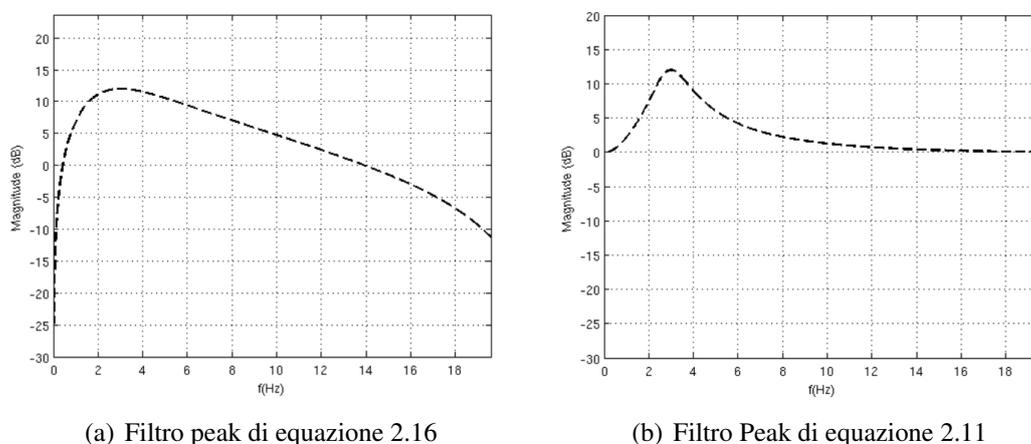
$$H_{res}^2(z) = \frac{V_0(1-h)(1-z^{-2})}{1 + 2dhz^{-1} + (2h-1)z^{-2}} \quad (2.16)$$

con

$$h = \frac{1}{1 + \tan\left(\pi\frac{f_B}{f_s}\right)} \quad (2.17)$$

$$d = -\cos\left(2\pi\frac{f_C}{f_s}\right) \quad (2.18)$$

$$V_0 = 10^{\frac{G}{20}} \quad (2.19)$$



**Figura 2.10:** Confronto tra i due filtri Peak IIR del secondo ordine

La scelta di due filtri diversi nasce dalla constatazione che un filtro del secondo tipo attenua il modulo fuori della frequenza di banda e questo potrebbe essere un problema per le basse frequenze. Per ovviare ad esso, si è scelto di utilizzare per il primo picco un filtro peak che non modifichi il contributo spettrale al di fuori della frequenza di banda.

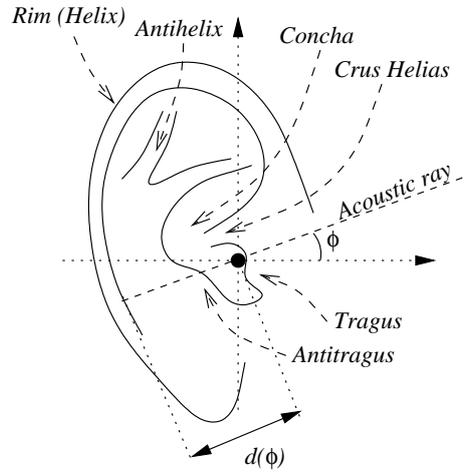
Le risposte in ampiezza di questi due filtri sono indicati in Figura 2.10. Con l'uso di tali filtri i risultati che si ottengono sono soddisfacenti, con una buona somiglianza tra PRTF reali e PRTF ricostruite. L'approssimazione è buona per frequenze fino 14 kHz, diventando scarsa per frequenze superiori, ma, dal momento che le frequenze interessanti a fini psicoacustici sono molto basse, tale discrepanza tra le PRTF non influisce. Il modello realizzato produce risultati soddisfacenti per PRTF con una o due risonanze principali e notch non troppo pronunciati.

## 2.2.2 Geometria della pinna

Gli autori in [36, 37] hanno provato a stabilire una relazione tra frequenze centrali dei notch e geometria della pinna. A questo fine hanno utilizzato una procedura di tipo *ray-tracing*, possibile dal momento che la lunghezza d'onda del suono è piccola se confrontata alla dimensione della pinna all'interno della banda di frequenze analizzata. I modelli di riflessione usualmente assumono coefficienti di riflessione positivi. Al contrario, gli autori, per la loro procedura, hanno assunto coefficienti di riflessione negativi, come ricavato dal lavoro di Satarzadeh [31]. Come punto di partenza per la misura delle distanze è stato scelto il canale uditivo. Indicando con  $t_d(\phi)$  il tempo di ritardo tra l'onda diretta e l'onda riflessa, con  $\phi$  indicante l'elevazione, il punto di riflessione è ad una distanza pari a:

$$d(\phi) = \frac{ct_d(\phi)}{2} \quad (2.20)$$

con  $c$  velocità del suono. Il fattore 2 è dovuto al fatto che il suono deve coprire 2 volte la distanza tra area di riflessione sulla pinna e il canale uditivo. Sostituendo il tempo con la frequenza e



**Figura 2.11:** Sistema di coordinate polari 2D e punti di riflessione

considerando il caso di riflessione negativa si ha:

$$f(\phi) = \frac{n+1}{t(\phi)} = \frac{c(n+1)}{2d(\phi)} \quad n = 0, 1, 2, \dots \quad (2.21)$$

Considerando di quest'ultima serie la prima frequenza  $f_0$ , otteniamo:

$$f_0(\phi) = \frac{c}{2d(\phi)} \quad (2.22)$$

da cui, con la relazione inversa, è possibile calcolare la distanza del punto di riflessione:

$$d(\phi) = \frac{c}{2f_0(\phi)} \quad (2.23)$$

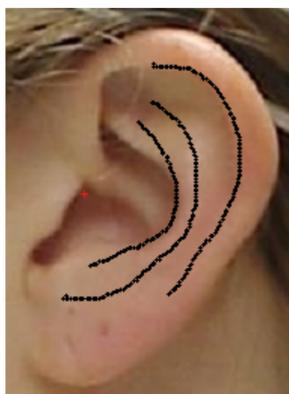
Considerando il sistema di coordinate polari 2-D rappresentato in Figura 2.11, ogni notch è mappato nel punto  $(d(\phi), \pi + \phi)$ .

Dai punti così trovati, gli autori hanno ottenuto una buona corrispondenza tra punti di riflessione e geometria della pinna, come è possibile vedere in Figura 2.12. Dall'immagine è possibile notare come la traccia più vicina al canale uditivo segua la parete della conca del soggetto per tutte le elevazioni, la traccia intermedia cada nell'area tra la conca e l'elice ed infine la traccia più esterna segua la forma dell'elice. Proprio a partire da tali risultati, gli autori in [16] hanno realizzato la procedura inversa rispetto a quanto esposto sopra: a partire dall'immagine di pinna sono stati tracciati i tre contorni più rilevanti, mostrati in Figura 2.13 dove, partendo dal contorno più esterno, abbiamo:

- $C_1/C_2$ : bordo dell'elice o parete interna dell'elice nel caso di orecchio sporgente;
- $C_4$ : antielice;
- $C_3$ : bordo esterno della conca.



**Figura 2.12:** *Punti di riflessione su Soggetto 134 CIPIC*



**Figura 2.13:** *Tracciamento contorni della pinna*

Questi sono salvati come una sequenza di pixel e, attraverso calcoli trigonometrici, trasformati in una coppia di coordinate polari  $(d, \phi)$  rispetto al canale uditivo, come è mostrato in Figura 2.14. Invertendo la formula 2.23 si ricavano le frequenze centrali dei notch ad ogni elevazione e per ognuna delle tre tracce. Per ciò che riguarda i parametri di larghezza di banda e profondità, non è ancora stata trovata una relazione con la forma della pinna. Gli autori hanno utilizzato per questi due parametri i valori medi tra tutti i soggetti considerati per ogni traccia e per ogni elevazione. Anche per le risonanze ci si è riferiti a valori medi per i parametri di larghezza di banda e altezza. Si è così realizzato il modello strutturale di HRTF mostrato in Figura 2.15. L'unico parametro indipendente usato dal blocco relativo alla pinna è l'elevazione grazie al quale calcolare le funzioni polinomiali associate alla frequenza centrale, alla larghezza di banda e all'altezza delle risonanze e alla larghezza di banda e alla profondità dei notch. Infatti solo le frequenze centrali dei notch sono legate alla forma della pinna e quindi la funzione polinomiale ad esse associata può essere calcolata offline. Come visto sopra, la parte risonante è realizzata con due filtri peak del secondo ordine in parallelo, mentre la parte riflettente è costituita dalla cascata di tre filtri notch del secondo ordine. Per i dettagli sui filtri vedere la sezione 2.2.1. In Figura 2.16 è mostrato un confronto tra PRTF originali e sintetiche per il soggetto 048 del CIPIC.

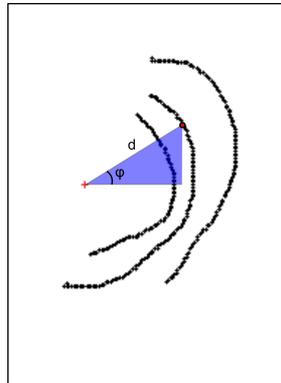


Figura 2.14: Punti di riflessione in coordinate polari

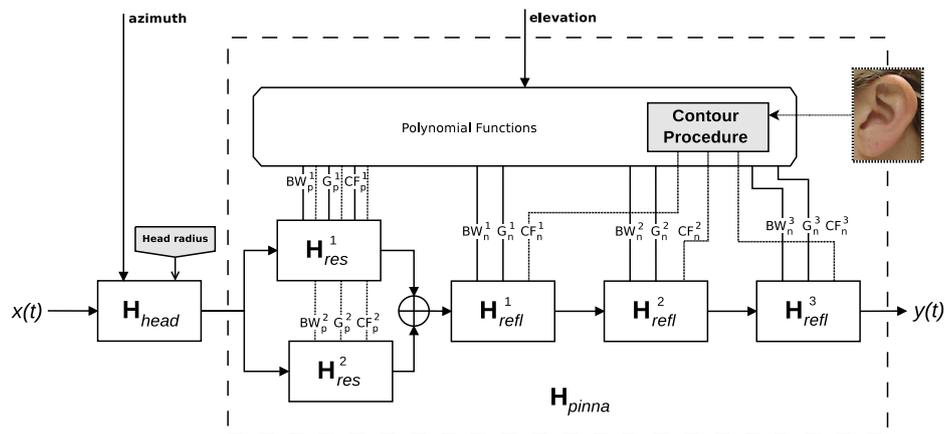
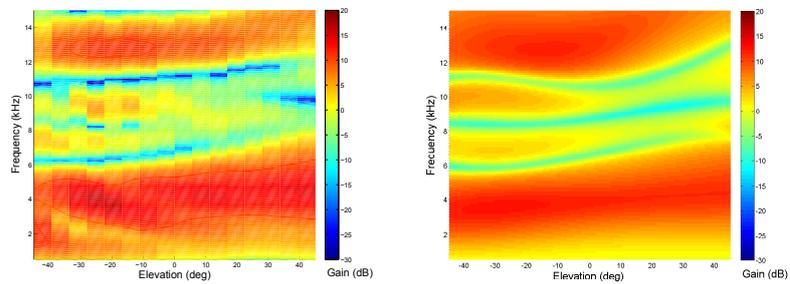


Figura 2.15: Modello strutturale di HRTF



(a) Soggetto 048 CIPIC: PRTF originali  
 (b) Soggetto 048 CIPIC: PRTF sintetiche

Figura 2.16: PRTF originali e sintetiche

## 2.3 Modifiche al modello strutturale di PRTF

Con il filtro IIR del secondo ordine, utilizzato per la realizzazione dei notch (Eq. 2.6), non tutte le combinazioni di parametri sono possibili: per un notch particolarmente stretto e profondo, il filtro produrrà una risposta in frequenza più larga e meno cava. Inoltre, se ad un notch stretto e appuntito è affiancato un altro notch, il filtro del secondo ordine potrebbe costruire una risposta in frequenza non corrispondente alla PRTF originaria nell'intervallo di frequenza tra i due notch. La larghezza di banda di un notch potrebbe, in questi casi, interferire con quella di notch adiacenti. Da qui l'idea di utilizzare, in fase di sintesi, un filter design che forzi la larghezza di banda del notch a restare uguale alle specifiche scelte in modo da migliorare il blocco riflettente sintetizzato, al più con un peggioramento nell'accuratezza della profondità dei notch.

Lo studio di questa tesi è rivolto principalmente al blocco riflettente del modello di cui sopra. Infatti il modello strutturale di pinna analizzato in questo lavoro è costituito da un blocco risonante, ripreso interamente dal modello precedente, e da un blocco riflettente realizzato non più con filtri peak IIR del secondo ordine, bensì utilizzando la funzione di filter design a disposizione tra le funzioni di Matlab. A differenza dello studio originario, le elevazioni considerate nel nostro studio variano tra  $\phi = -45^\circ$  e  $\phi = 45^\circ$  a passo di  $15^\circ$  anziché di  $5.695^\circ$ , data la difficoltà, da parte dei soggetti, di differenziare posizioni così ravvicinate. La scelta di tale regione sul piano mediano segue le medesime ragioni del modello già descritto.

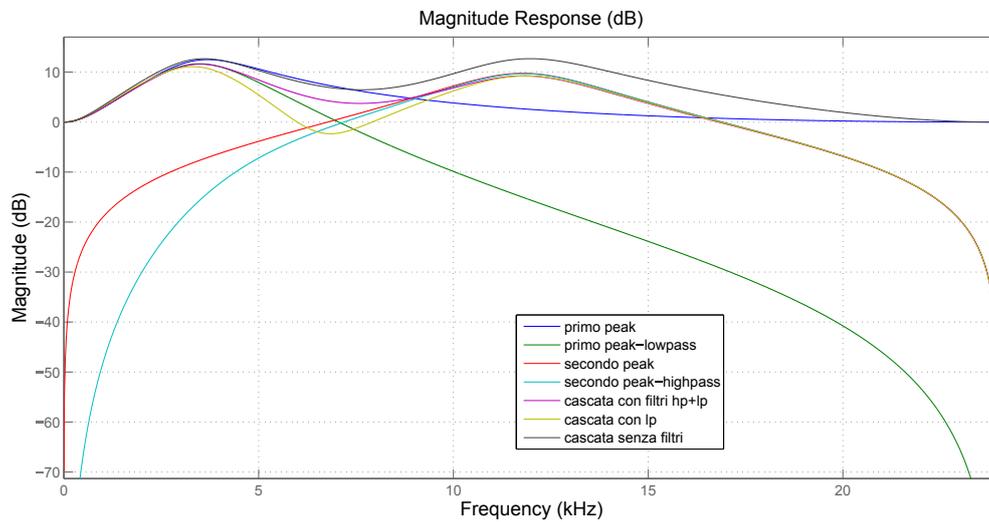
Cominciamo quindi con la descrizione dei filtri utilizzati in questa nuova versione del modello strutturale per poi approfondire la scelta dei parametri di notch testati.

### 2.3.1 Filtri

#### 2.3.1.1 Blocco risonante

Per quel che riguarda il blocco risonante abbiamo utilizzato due filtri del secondo ordine. Come in [15], per modellare la prima risonanza abbiamo utilizzato un filtro del tipo in Eq. 2.11. Come già visto, le ragioni che hanno portato alla scelta di tale filtro per la prima risonanza sono dovute al fatto che esso non modifica l'intervallo al di fuori della frequenza di banda. Tuttavia, affinché la seconda risonanza rispettasse la specifica di altezza stabilita, è stato necessario aggiungere un filtro passa basso alla prima risonanza in modo che quest'ultima non interferisse con la seconda risonanza. Per la seconda risonanza, si è utilizzato, come in [15], un filtro del secondo ordine IIR del tipo in equazione 2.16. Alla seconda risonanza abbiamo applicato un filtro passa alto in modo da non modificare la risposta in frequenza tra i due picchi. In Figura 2.17 è mostrato l'effetto dei filtri passa basso e passa alto sulla risposta dei singoli picchi e sulla relativa cascata.

Per i 3 parametri dei filtri risonanti, frequenza centrale del picco  $f_c$ , guadagno  $G$  e larghezza di banda  $f_B$ , i valori utilizzati sono valori medi sui soggetti CIPIC. Poiché i parametri disponibili non erano associati alle elevazioni prese in considerazione nel presente studio, abbiamo applicato un'interpolazione ai dati, ricavando i valori presenti in Tabella 2.1 per il primo picco e in Tabella 2.2 per il secondo picco. In Figura 2.18 è riportato il blocco riflettente costituito dai due peak in cascata, per le diverse elevazioni testate.



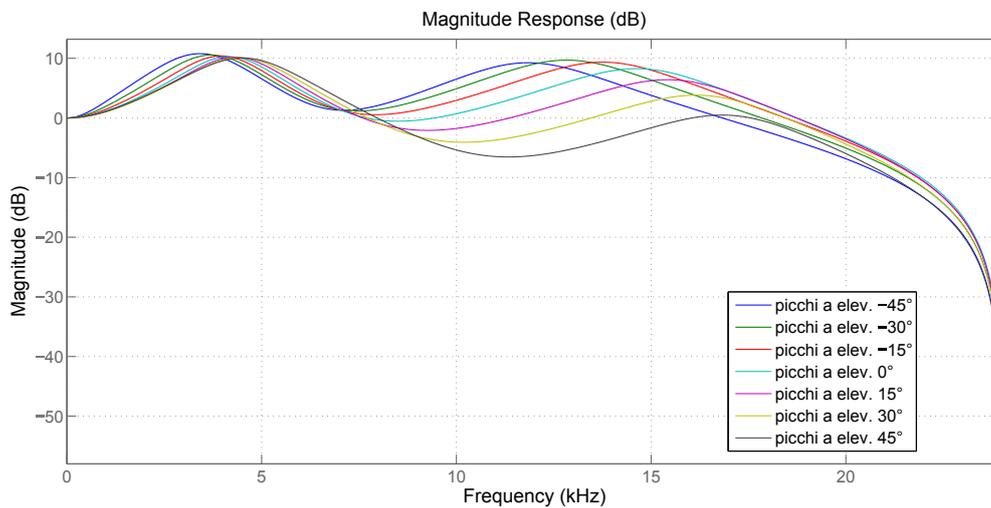
**Figura 2.17:** Prima e seconda risonanza in cascata con e senza filtri passa basso e passa alto

Elevazione (gradi)	Frequenza (Hz)	Altezza (dB SPL)	BW (kHz)
-45	3.391	12.5	2.733
-30	3.619	12.4	2.508
-15	3.822	12.2	2.383
0	3.999	12.1	2.359
15	4.148	11.8	2.440
30	4.272	11.6	2.621
45	4.371	11.3	2.902

**Tabella 2.1:** Parametri per il primo filtro peak alle diverse elevazioni:  $f_c$  frequenza centrale picco,  $G$  guadagno,  $f_B$  larghezza di banda

Elevazione (gradi)	Frequenza (kHz)	Altezza (dB SPL)	BW (kHz)
-45	10.830	9.7	3.682
-30	11.757	10	3.739
-15	12.613	9.6	3.795
0	13.410	8.5	3.849
15	14.133	6.7	3.899
30	14.797	4.1	3.948
45	15.390	0.9	3.995

**Tabella 2.2:** Parametri per il secondo filtro peak alle diverse elevazioni:  $f_c$  frequenza centrale picco,  $G$  guadagno,  $f_B$  larghezza di banda



**Figura 2.18:** Primo e secondo picco per le 7 elevazioni

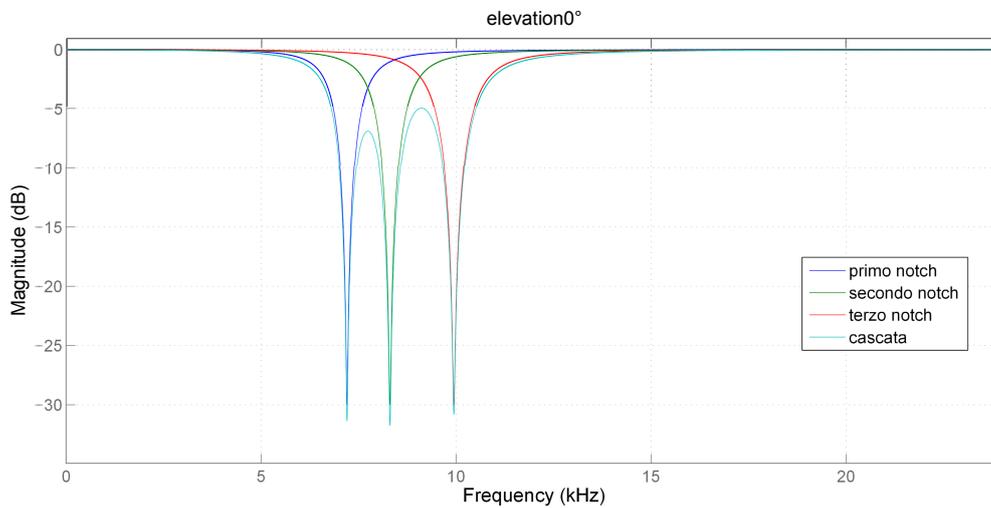
### 2.3.1.2 Blocco riflettente

Per quel che riguarda la parte riflettente, abbiamo utilizzato, per la creazione dei notch, la funzione `fdesign.parameq` disponibile in Matlab. Essa permette di creare filtri equalizzatori parametrici, quindi con la possibilità di settare frequenza centrale, profondità e larghezza di banda del filtro. Come abbiamo visto nella Sezione 2.2.1, i filtri notch utilizzati dagli autori presentano delle imperfezioni, in particolare per notch profondi e stretti posti in cascata. Con l'uso della funzione `fdesign.parameq` riusciamo ad aver maggiore controllo sulla forma di ogni filtro, in particolare aumentando l'ordine del filtro stesso. Di contro però, un aumento dell'ordine del filtro si traduce in un filtro di tipo brick wall (a muro di mattoni). Filtri di questo tipo possono operare tagli di frequenza con pendenze d'attenuazione molto ripide che, nel caso di cascate di notch, peggiorano il risultato finale. Per tale motivo si è scelto di optare per un ordine di filtro basso. La funzione `fdesign.parameq` viene da noi utilizzata con le seguenti specifiche come parametri d'ingresso:

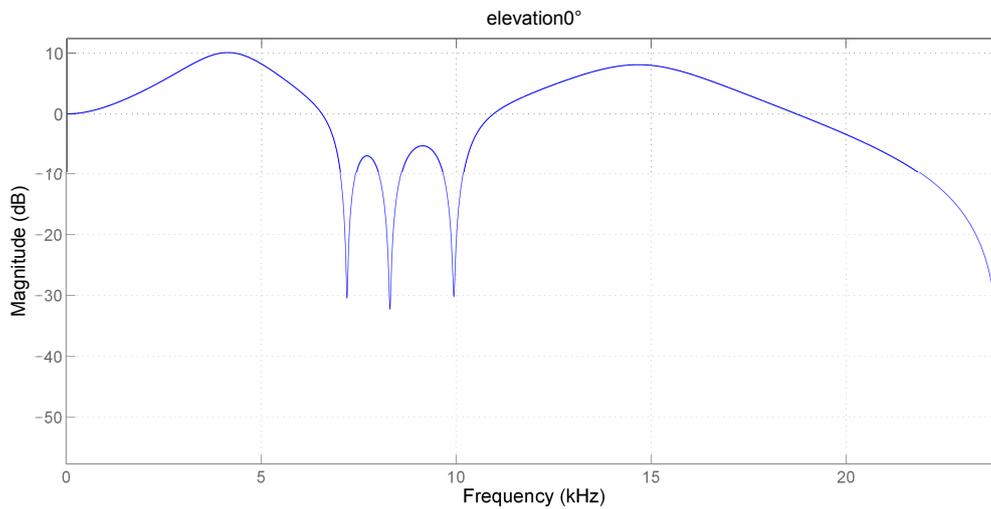
- $N$ : ordine del filtro;
- $F0$ : frequenza centrale;
- $BW$ : larghezza di banda;
- $G_{ref}$ : profondità di riferimento (dB);
- $G0$ : profondità alla frequenza centrale (dB);
- $GBW$ : profondità alla quale la larghezza di banda è misurata (dB).

Per i parametri  $N$ ,  $G_{ref}$ ,  $GBW$  i valori scelti sono, rispettivamente, 2, 0 dB, -3 dB. Gli ultimi due valori indicano l'utilizzo di larghezze di banda a -3 dB assoluti, ovvero rispetto a 0 dB. I valori di frequenza centrale dipendono dalla pinna del soggetto (frequenze centrali estratte dai 3 contorni principali), mentre profondità e larghezza di banda dalle condizioni che testeremo nell'esperimento psicoacustico e di cui parleremo nella sezione successiva. In Figura 2.19 sono mostrati i 3 notch in cascata estratti da uno dei soggetti testati nel presente studio (Soggetto 30), per l'elevazione di  $0^\circ$ , con profondità a -30 dB e larghezza di banda relativa alla frequenza centrale pari a  $0.15f_c$ . In Appendice A sono riportati i grafici con le restanti elevazioni.

Dall'unione del blocco risonante e del blocco riflettente si ottengono le PRTF sintetiche. In Figura 2.20 è riportata la PRTF sintetica per il Soggetto 30 per l'elevazione di  $0^\circ$ , fissata la profondità e la larghezza di banda del notch rispettivamente a -30 dB e  $0.15f_c$ , con  $f_c$  frequenza centrale del notch. In Appendice A sono riportati i grafici con le restanti elevazioni.



**Figura 2.19:** Primo, secondo e terzo notch in cascata per le elevazioni  $0^\circ$



**Figura 2.20:** PRTF sintetica per il Soggetto 30 con profondità e larghezza di banda del notch rispettivamente a  $-30$  dB e  $0.15f_c$  per l'elevazione di  $0^\circ$

## 2.3.2 Condizioni sperimentali sui parametri dei notch

### 2.3.2.1 Studi precedenti

In [20], Hebrank e Wright hanno mostrato che la localizzazione sul piano mediano è possibile grazie ai profondi notch spettrali generati dal filtraggio della pinna. Attraverso 3 esperimenti gli autori hanno definito il range di frequenze contenente gli indicatori necessari per la localizzazione sul piano mediano, misurando la risposta dei soggetti testati a rumori con larghezza di banda modificata da filtri passa basso e passa alto. Dai risultati ottenuti, il range di frequenze preso in considerazione è stato 3.8–16 kHz. Inoltre facendo localizzare ai soggetti rumori filtrati in modo da avere caratteristiche di tipo passa basso, passa alto, elimina banda e passa banda, si è ricavato, per ogni direzione, in base alle risposte dei soggetti, il possibile filtraggio operato dall'orecchio esterno. Quest'ultima ipotesi è stata verificata misurando la risposta di pinne artificiali a rumori posizionati sul piano mediano. I risultati finali di tale studio sono di seguito elencati:

- Un notch di 1 ottava con frequenza di cut-off più bassa tra 4 e 8 kHz è responsabile della sensazione di suono frontale;
- Un picco di 1/4 di ottava tra 7 e 9 kHz è associato alla sensazione di suono posto sopra la testa;
- Un piccolo picco tra 10 e 12 kHz, con energia decrescente sia sotto che sopra, è associato alla sensazione di suono posto dietro la testa.

Un contributo importante dato dagli autori è stato quello di ricavare la relazione tra posizione dei notch, nella risposta in frequenza dell'orecchio esterno alle diverse elevazioni, e tempo di ritardo tra onda diretta e onda riflessa del suono che giunge all'orecchio.

In [10] Buchlein, utilizzando, come stimoli estratti di musica, voce umana e rumori, ha indagato la capacità di percepire picchi e notch di guadagno e larghezza di banda variabile. Inizialmente l'autore ha testato la percezione di estratti di musica e voci filtrati con notch di circa un ottava di larghezza di banda, con profondità di 5 e 10 dB. I risultati ottenuti mostrano che lo stimolo con notch di 10 dB è percepito facilmente dalla gran parte dei soggetti a tutte le frequenze fino ad 11 kHz. Più difficile invece la percezione di stimoli con notch di 5 dB. In questo caso, solo alla frequenza di 300 Hz i soggetti hanno percepito il notch. In un secondo esperimento sono stati testati notch con profondità di 20 dB e con larghezza di banda pari a  $\Delta f/f = 0.55, 0.35, 0.2$ . I notch con larghezza di banda pari a  $\Delta f/f = 0.55$  sono percepiti da tutti i soggetti, in particolare per frequenze tra 3-5 kHz, mentre i notch con larghezza di banda pari a  $\Delta f/f = 0.35$  sono udibili solo tra 3-5 kHz. I notch con larghezza di banda  $\Delta f/f = 0.2$ , al contrario, non sono udibili in nessun range di frequenza. L'autore quindi si sofferma sulla percezione dei picco, mostrando che essi sono più facilmente udibili anche quando si presentano con forma stretta e con altezze basse. Tali risultati si replicano anche utilizzando, come stimolo di test, i rumori. Anche in questo caso i test sono stati eseguiti con picchi e notch di uguale larghezza di banda e con guadagno di 15, 10 e 5 dB. Mentre i picchi sono sempre udibili, i notch, a 15 dB di profondità sono udibili soprattutto nel range 3-5 kHz, a 10 dB di profondità invece risultano udibili solo a

frequenze inferiori a 1 kHz. Infine i notch di 5 dB, come negli esperimenti precedenti, non sono udibili.

In [3] e [11], gli autori Pinto e Lopez-Poveda mostrano come l'abilità di rilevare notch ad alta frequenza vari tra gli ascoltatori, e dipenda dalla larghezza di banda del notch così come dal livello e dalla durata dello stimolo. Lo studio parte dal presupposto che il nervo uditivo (AN) sia il solo mezzo di trasmissione dell'informazione acustica al sistema uditivo centrale. Di conseguenza, l'abilità di percepire notch in alta frequenza per la localizzazione del suono richiede che essi siano adeguatamente rappresentati nella risposta del nervo uditivo. Le fibre uditive vengono classificate in due tipi in base allo *spontaneous rate*; in particolare, abbiamo le fibre con HSR (high spontaneous rate,  $> 15$  spikes/s), che presentano una bassa soglia ( $< 10$  dB SPL), e le fibre con LSR (low spontaneous rate), con soglia più alta ( $> 15$  dB SPL). Ogni fibra del nervo uditivo è più sensibile ad una particolare frequenza (frequenza caratteristica o CF) e risponde ad un ristretto range di frequenze attorno ad essa. Ogni fibra può quindi essere vista come uno speciale filtro con una certa CF e la popolazione di fibre del nervo uditivo può essere vista come un banco di filtri con frequenze caratteristiche che spaziano all'interno del range di frequenze udibili. In figura 6 del citato articolo viene mostrata la soglia di profondità del notch in funzione del livello dello stimolo. Le larghezze di banda considerate sono: 1000, 2000, 4000 Hz. In generale, la soglia di profondità del notch incrementa al diminuire della larghezza di banda del notch. Infatti, un incremento significativo nella soglia di profondità del notch si ha per il notch con larghezza di banda pari a 1000 Hz. Dalla figura si evince che per notch con larghezza di banda 2000 o 4000 Hz la soglia resta al di sotto dei 10 dB, al variare del livello dello stimolo. In conclusione, il rilevamento di notch con larghezze di banda uguali o superiori a 2000 Hz risulta sempre possibile a tutti i livelli, mentre il rilevamento di notch con larghezza di banda di 1000 Hz risulta chiaramente più difficile, particolarmente ad alti livelli dello stimolo. La spiegazione data dagli autori è che la larghezza di banda delle fibre del nervo uditivo, con frequenza caratteristica vicina alla frequenza centrale del notch, a bassi livelli dello stimolo sia intorno ai 1000 Hz ma aumenti al crescere del livello dello stimolo. Ad alti livelli, le fibre in questione risponderanno all'energia anche in regioni di frequenza adiacenti alla banda di 1000 Hz del notch. Questo rende più difficile discriminare uno stimolo piatto da uno con un notch anche per notch piuttosto profondi. Inoltre, la discriminazione spettrale risente del numero di fibre con frequenza caratteristica all'interno della banda del notch. All'aumentare della banda del notch aumentano anche le fibre contenute in tale banda. Questo spiega il perchè, per notch con banda molto larga, la soglia sia piuttosto bassa. Un risultato interessante trovato dagli autori è anche il legame tra soglia della profondità del notch e livello dello stimolo. Si tratta di una funzione non monotonica poichè la soglia aumenta all'aumentare del livello, fino ad un massimo in prossimità del livello di 70 – 80 dB SPL, per poi diminuire per livelli superiori.

Risultati simili sono stati trovati da Moore in [27]. Nel citato studio l'autore analizza l'abilità dei soggetti di rilevare e discriminare picchi e notch con frequenze centrali di 1 e 8 kHz, utilizzando, come stimoli, rumori. Da tale studio otteniamo informazioni sulla soglia di profondità del notch in funzione della sua larghezza di banda. Le larghezze di banda prese in esame sono 0.125, 0.25 e 0.5, tutte relative alla frequenza centrale del notch. Come mostrato in Figura 2 in [27], la soglia aumenta al diminuire della larghezza di banda. Considerando la soglia per la frequenza di 8000 Hz, osserviamo che, con larghezze di banda pari a  $0.125f_c$ , essa non risulta

misurabile per tutti i soggetti testati. Al contrario, tutti i soggetti riescono a rilevare notch con larghezze di banda di  $0.25f_c$  e  $0.5f_c$ . Le soglie variano tra 10 e 20 dB per larghezze di banda di  $0.25f_c$  e diminuiscono sensibilmente per larghezze di banda di  $0.5f_c$ . Alcuni dei notch introdotti dalla pinna sono anche più ampi di quelli usati nello studio di Moore. Infatti, la larghezza di banda e la profondità dei notch nell'HRTF variano tra gli ascoltatori, come mostrato da Shaw in [33]. Ad esempio, dallo studio di Hebrank e Wright risulta che uno dei maggiori indicatori per il rilevamento di suoni nel piano mediano in direzione frontale è un notch di larghezza di banda di  $0.7f_c$ . In base ai risultati di Moore, un tale notch dovrebbe essere rilevato da tutti i soggetti. Cambiamenti nella posizione di un suono sul piano mediano si traducono in spostamenti nella frequenza centrale dei notch introdotti dalla pinna. Dagli esperimenti di Moore estrapoliamo che discriminazioni del centro frequenza di notch risentono meno della larghezza di banda. Infatti, tutti i soggetti sono in grado di rilevare cambiamenti nel centro frequenza di notch di larghezza di banda di  $0.125f_c$  così come sono in grado di rilevare cambiamenti nella profondità dei notch della medesima larghezza di banda. Inoltre dallo studio si evince che la soglia per il rilevamento di un cambio di profondità di un notch è maggiore della soglia per il rilevamento del cambio di altezza di un picco, e che la prima incrementa molto più della seconda al diminuire della larghezza di banda.

In [29], gli autori Poon e Brugge descrivono i risultati di esperimenti nei quali sono state registrate le risposte di singole fibre del nervo uditivo a rumori a banda larga, precedentemente filtrati da una funzione con forma che approssima quella dei notch presenti nelle HRTF. Gli esperimenti sono stati effettuati sui gatti. I dati sui gatti, come osservato dagli autori, sono simili a quelli ottenuti sugli umani. Infatti, confrontando le HRTF dei gatti con quelle degli umani, notiamo la presenza in entrambe di notch spettrali tra 5 – 40 kHz. Dai risultati si ricava che le fibre del nervo uditivo sono molto sensibili alla presenza dei notch con frequenza centrale prossima alla frequenza caratteristica delle fibre. Il notch sintetico utilizzato per i loro esperimenti ha profondità di 35 dB e frequenza centrale variabile tra 5 – 35 kHz. Sono state testate circa 161 fibre del nervo uditivo, con frequenza caratteristica tra 0.1 – 40 kHz. Circa il 54% di tali fibre presentavano un alto tasso di scarica ( $HSR > 20$  spikes/s), mentre il restante 46% presentavano un basso LSR. Il rapporto tra tasso di scarica e livello dello stimolo è approssimabile con una funzione monotonica. Superata la soglia della *tuning curve*, al di sotto della quale la fibra presenta *spontaneous rate*, il tasso di scarica cresce all'aumentare del livello dello stimolo. Si nota altresì un'ampia riduzione nel tasso di scarica nella zona della frequenza caratteristica della fibra quando la frequenza del notch è vicina ad essa, scendendo, in alcuni casi, al di sotto dello *spontaneous rate* che si avrebbe in assenza di stimolo acustico. Gli autori hanno stimato anche il livello dello stimolo, sopra la soglia della fibra, al quale il notch produce la maggiore riduzione del tasso di scarica, ricavandone un valore compreso tra 15 – 35 dB. Riepilogando: la massima riduzione nel tasso di scarica si ha quando la frequenza centrale del notch è vicina alla frequenza caratteristica della fibra del nervo uditivo, e quando il livello del notch è 15 – 55 dB al di sopra del soglia della *tuning curve*. Infine, nel considerare la relazione tra la posizione del notch e la *threshold tuning curve* delle fibre uditive, i due autori hanno trovato che sovrapponendo il notch alla *tuning curve* della fibra uditiva, di frequenza caratteristica prossima a quella del notch, si ha un ottimo fit utilizzando come profondità del notch il valore di 35 dB. La conclusione quindi degli autori è che l'informazione contenuta nella frequenza e nell'intensità di rilevanti notch

nelle HRTF venga trasmessa al sistema nervoso centrale attraverso il tasso di scarica delle fibre uditive.

In [22], gli autori hanno effettuato test di localizzazione nel piano mediano utilizzando un modello parametrico di HRTF. Tale modello è costituito solo da alcuni dei picchi e notch estratti dalle HRTF originali. Gli autori hanno trovato che per la  $pHRTF(all)$ , HRTF parametrica composta da tutti i picchi e notch rilevanti, l'elevazione percepita è accurata tanto quanto quella percepita utilizzando le HRTF misurate sul singolo individuo. Al contrario, si hanno scarsi risultati considerando  $pHRTF$  costituite da un solo notch o da un solo picco. Per le  $pHRTF$  costituite solo dai primi due notch rilevanti (N1 e N2), estratti dalle HRTF a partire dalla frequenza di 4 kHz, i risultati mostrano un'accuratezza nella percezione dell'elevazione paragonabile a quella che si avrebbe con le HRTF originali. Questo implica che i notch N1 e N2 assumono un ruolo importantissimo nella percezione dell'elevazione di un suono nel piano mediano. Dai risultati è evidente anche il ruolo del secondo picco per la percezione di elevazioni sopra la testa. Per quel che riguarda il primo picco (P1), pur non essendo legato alla percezione dell'elevazione, può essere visto come un importante riferimento per il sistema uditivo umano nell'analisi dei primi due notch. Inoltre anche in questo lavoro, come già verificato da altri studi, risulta che la frequenza centrale, la profondità e la forma dei notch N1 ed N2 cambiano in maniera considerevole al variare dell'elevazione del suono.

### 2.3.2.2 Scelta dei parametri

Dalle osservazioni delle analisi citate abbiamo ricavato utili indicazioni sulla scelta dei nostri parametri sperimentali. I 3 parametri associati al notch sono:

- $f_c$ : frequenza centrale del notch;
- $G$ : gain o profondità del notch;
- $BW$ : larghezza di banda del notch calcolata a -3 dB SPL.

L'unico parametro associato alle misurazioni antropometriche del soggetto è  $f_c$ , la frequenza centrale del notch. Tali valori sono ricavati per mezzo di una GUI sviluppata in Matlab con cui effettuiamo il tracciamento dei contorni sulla pinna e, da tali contorni, ricaviamo le frequenze centrali dei notch associati ai punti di riflessione, come mostrato nella Sezione 2.2.2. La descrizione di tale procedura sarà sviluppata dettagliatamente nel capitolo successivo.

La scelta di larghezze di banda di  $0.25f_c$  e fissa a 2000 Hz è dovuta alla constatazione che tali valori corrispondono alle minime larghezze di banda, relative e fisse, alle quali la maggior parte dei soggetti riesce a rilevare un notch alle frequenze di nostro interesse, come mostrato rispettivamente in [27] e in [3]. Anche se la larghezza di  $0.15f_c$  risulta di difficile percezione in base alla letteratura, si veda ad esempio Moore in [27], la scelta di includerla nei parametri testati ha prodotto, come vedremo successivamente, interessanti risultati. La decisione di testare anche tale larghezza di banda nasce dalla constatazione che a tale valore si ha meno sovrapposizione tra notch affiancati, ottenendo quindi una PRTF simile a quella originale.

Per quanto concerne la profondità, abbiamo deciso di utilizzare 2 valori:

Condizione	Guadagno (dB)	Banda	Nr.Notch
1	-10	0.15	3
2	-10	0.15	2
3	-10	0.25	3
4	-10	0.25	2
5	-10	2000	3
6	-10	2000	2
7	-30	0.15	3
8	-30	0.15	2
9	-30	0.25	3
10	-30	0.25	2
11	-30	2000	3
12	-30	2000	2

**Tabella 2.3:** Condizioni sperimentali

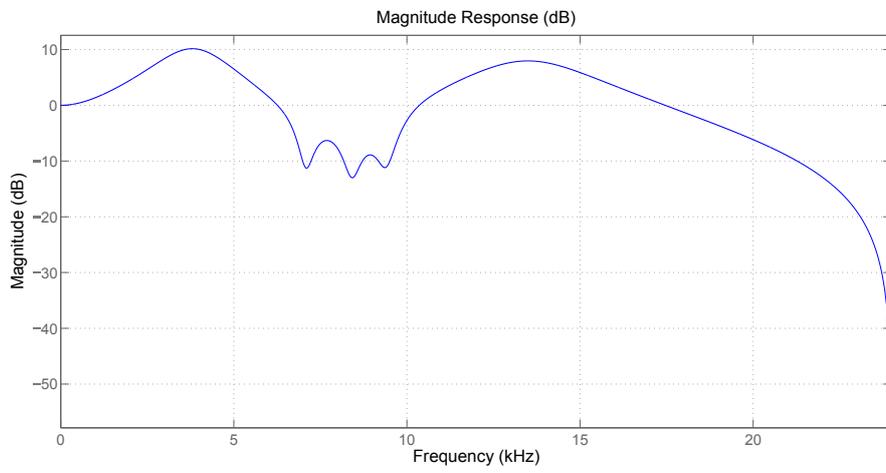
1. 10 dB: la minima soglia alla quale, con larghezza di banda di  $0.25f_c$ , gran parte dei soggetti riesce a rilevare il notch, come mostrato in [27], e, allo stesso tempo, la soglia sotto la quale, con larghezze di banda fisse a 2000 e 4000 Hz, si distribuiscono le soglie di profondità dei notch, come mostrato in [3] e [11];
2. 30 dB: la scelta di tale valore deriva dall'ottimo fit trovato tra un notch con profondità di 35 dB, di poco superiore al valore di 30 dB da noi scelto, e le tuning curve delle fibre uditive, come mostrato in [29].

Poniamo in evidenza che i valori di profondità da noi testati figurano tra quelli presenti nel box plot in Figura 2.8.

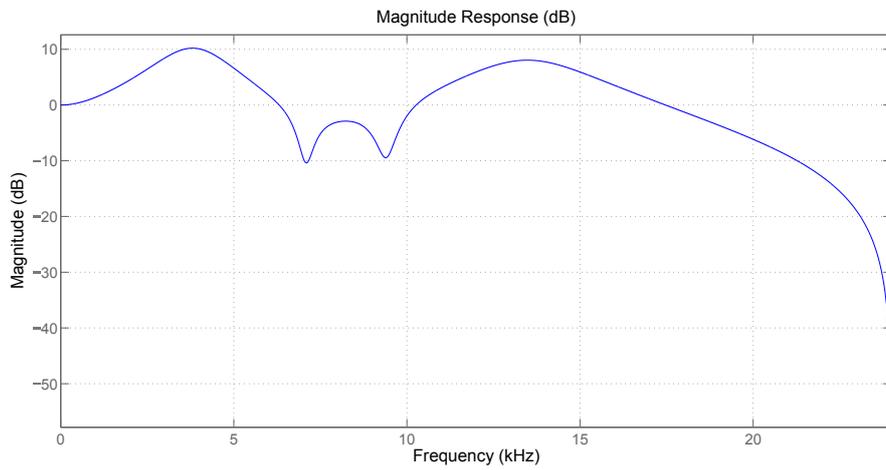
Infine, l'ultimo parametro variabile delle nostre PRTF sintetiche è il numero di notch presenti. Abbiamo testato PRTF costituite da 2 peak e 3 notch e PRTF composte da 2 peak e 2 notch, eliminando quello centrale. Tale decisione è in accordo con i risultati ottenuti in [22], dove, come abbiamo già detto, la presenza di due soli notch è risultata sufficiente per una localizzazione accurata sul piano mediano.

Nella Tabella 2.3 sono riepilogate le 12 condizioni sperimentali testate.

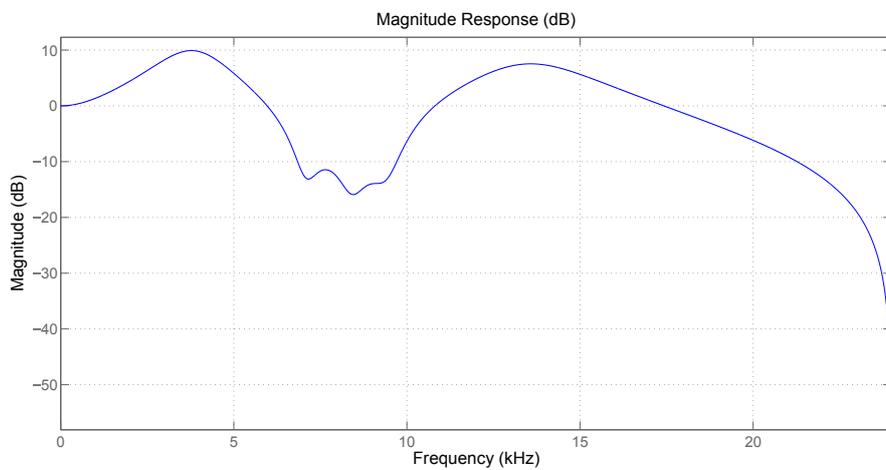
In Figura 2.21, 2.22, 2.23, 2.24, sono raffigurate, a titolo esemplificativo, le PRTF sintetiche associate al Soggetto 21, ognuna legata ad una condizione e considerate per una elevazione di  $0^\circ$ .



(a) Condizione 1, elevazione 0°

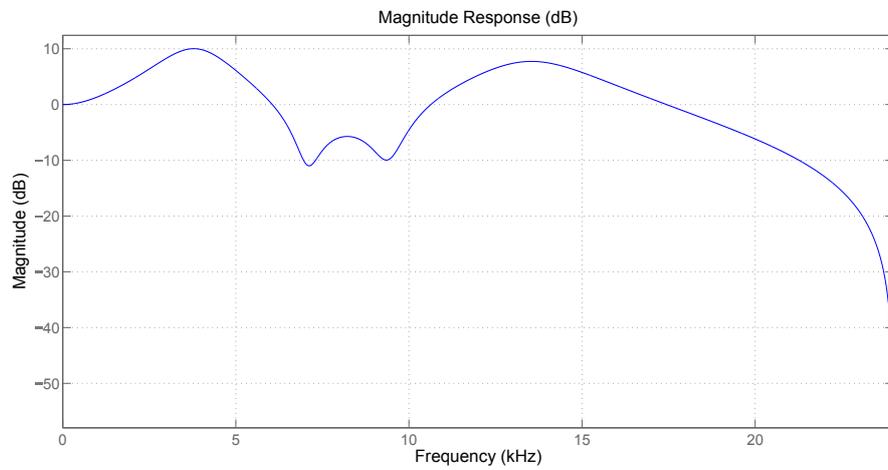


(b) Condizione 2, elevazione 0°

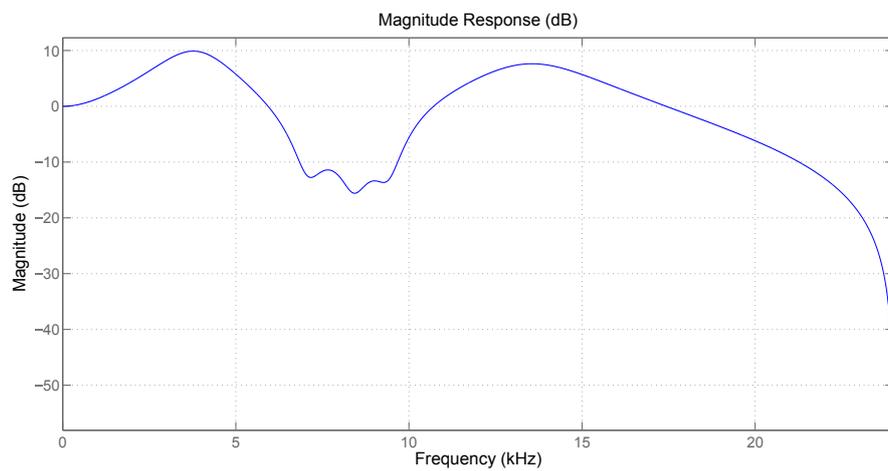


(c) Condizione 3, elevazione 0°

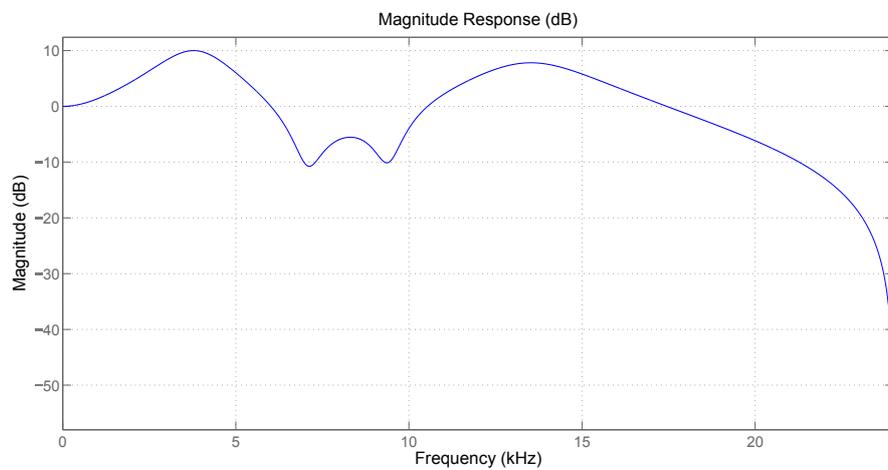
**Figura 2.21:** PRTF per il Soggetto 21, ad elevazione 0° e con le condizioni 1, 2 e 3



(a) Condizione 4, elevazione 0°

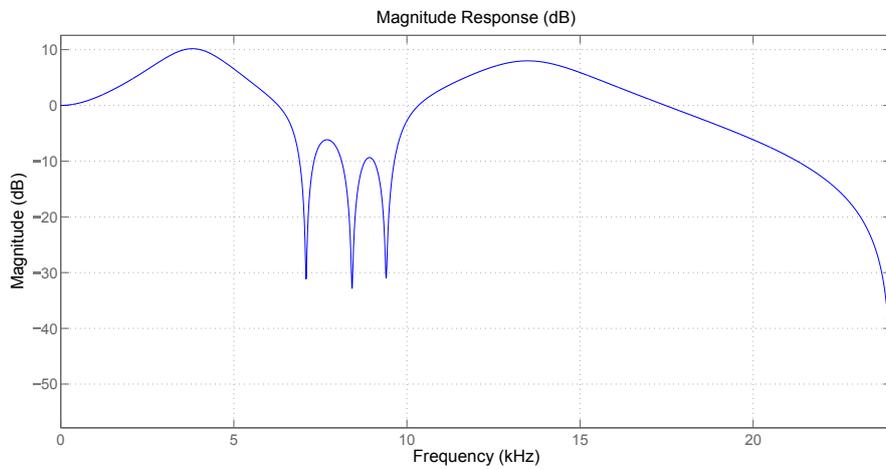


(b) Condizione 5, elevazione 0°

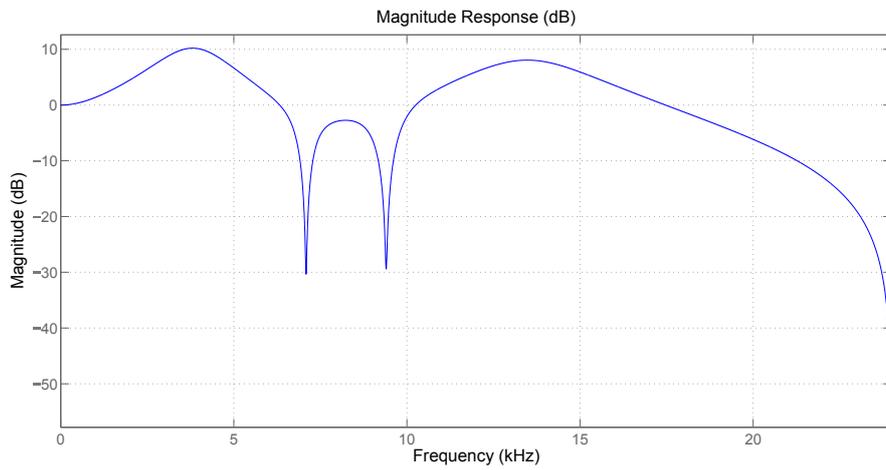


(c) Condizione 6, elevazione 0°

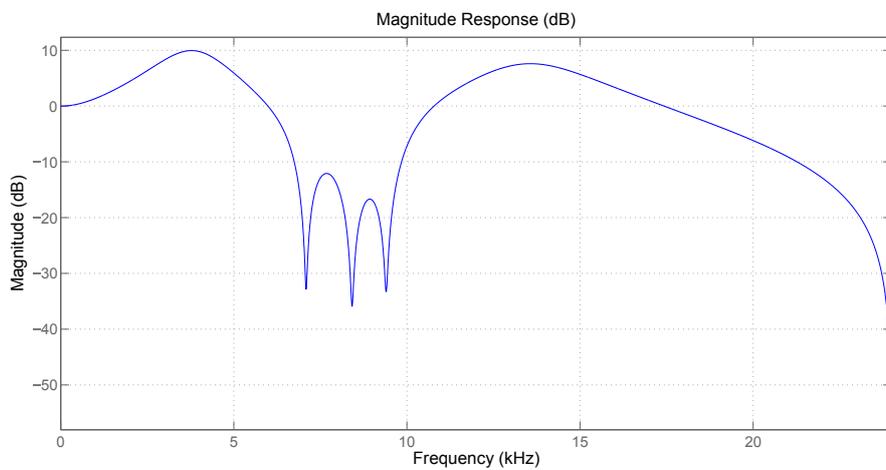
**Figura 2.22:** PRTF per il Soggetto 21, ad elevazione 0° e con le condizioni 4, 5 e 6



(a) Condizione 7, elevazione 0°

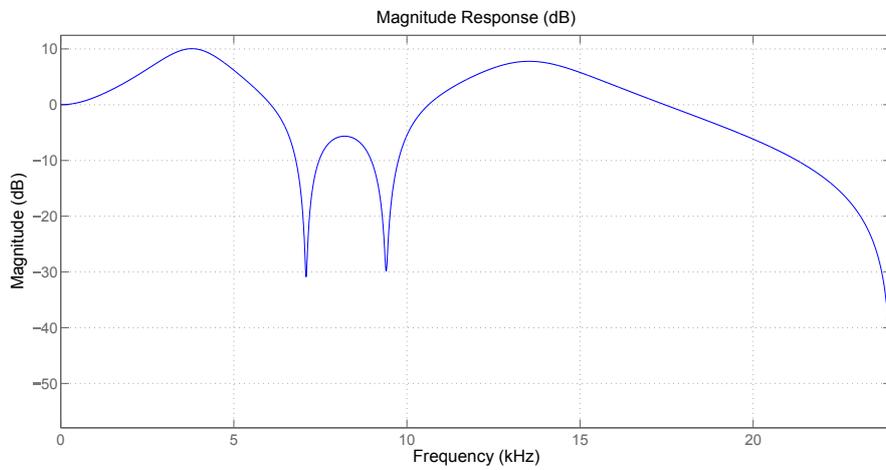


(b) Condizione 8, elevazione 0°

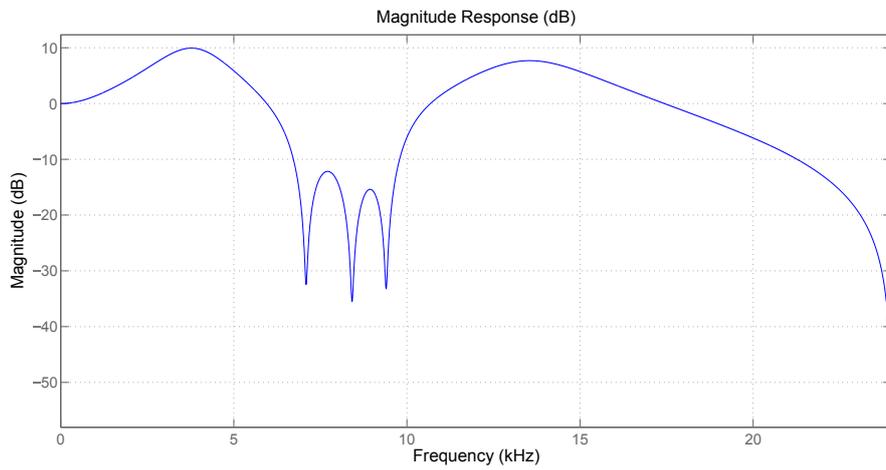


(c) Condizione 9, elevazione 0°

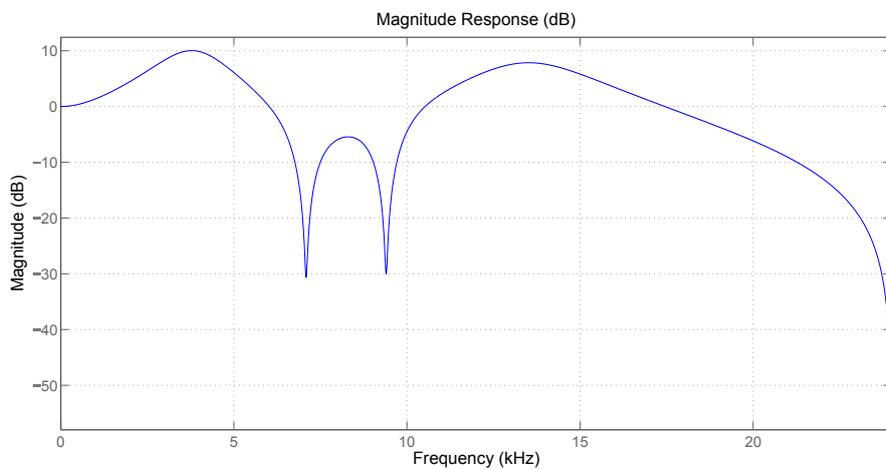
**Figura 2.23:** PRTF per il Soggetto 21, ad elevazione 0° e con la condizione 7, 8 e 9



(a) Condizione 10, elevazione  $0^\circ$



(b) Condizione 11, elevazione  $0^\circ$



(c) Condizione 12, elevazione  $0^\circ$

**Figura 2.24:** PRTF per il Soggetto 21, ad elevazione  $0^\circ$  e con le condizioni 10, 11 e 12



# Capitolo 3

## Design Sperimentale

In questo capitolo viene descritto il design dell'esperimento. La Sezione 3.1 contiene una descrizione dell'apparato, sia software che hardware, utilizzato per la realizzazione dei test. La Sezione 3.2 descrive il test audiometrico sottoposto ai soggetti per accertarsi dell'assenza di problemi uditivi tali da compromettere l'esito dell'esperimento. La Sezione 3.3 descrive le fasi del test di localizzazione e il protocollo sperimentale seguito dai soggetti testati.

### 3.1 Apparato

#### 3.1.1 Software

Matlab è il software utilizzato per la realizzazione dell'intero esperimento. Esso è un linguaggio ad alto livello utilizzato in diversi ambiti, tra cui l'elaborazione dei segnali. Matlab consente di analizzare dati, sviluppare algoritmi e creare modelli. Esso integra il calcolo, la visualizzazione e la programmazione. Inoltre, esso mette a disposizione un ambiente grafico, chiamato GUIDE (GUI Design Environment), utile per la realizzazione di semplici ed immediate interfacce grafiche. Di grande utilità risultano due toolbox integrati in Matlab: un toolbox statistico (*Statistics Toolbox*) e un toolbox relativo alla manipolazione delle immagini (*Image Processing Toolbox*). Tale linguaggio è risultato altresì adatto per l'analisi e la rappresentazione dei dati.

#### 3.1.2 Hardware

##### 3.1.2.1 Cuffie

Sia per il test audiometrico che per l'esperimento psicoacustico abbiamo fatto uso delle cuffie Sennheiser HDA 200 mostrate in Figura 3.1. Si tratta di cuffie circumaurali, di tipo dinamico e chiuse, con i cuscinetti che circondano completamente l'orecchio. Sono in grado di isolare parzialmente l'orecchio da rumori esterni. La risposta in frequenza di tali cuffie è 20-20000 Hz, come mostrato in Figura 3.2, con un'impedenza di 40 Ohm. Come evidenziato da un lavoro precedente [32], la riproduzione tramite cuffie introduce distorsioni di varia natura. In particolare



**Figura 3.1:** *Sennheiser HDA 200*

tutte le cuffie analizzate nel lavoro di cui sopra, tra cui anche le Sennheiser HDA 200, presentano un comportamento in frequenza assimilabile a un filtro passa banda, che si discosta da una risposta a guadagno costante ideale. Le frequenze di taglio sono attorno ai 150 Hz e 9 KHz. Inoltre, associata ai riposizionamenti della cuffia sulla testa, si rileva anche una variabilità della HpTF (Headphones Transfer Function), come mostrato nel grafico in basso in Figura 3.2, dove viene rappresentata la deviazione standard su 10 riposizionamenti. Tale variabilità è particolarmente pronunciata in alcuni range di frequenze. Essa è limitata entro alcuni dB per le frequenze inferiori a 7 KHz, mentre, per frequenze superiori, tale variabilità aumenta considerevolmente. Ovviamente tutto questo si traduce in un'equalizzazione non voluta del segnale binaurale. Si è reso quindi necessario un intervento di compensazione delle cuffie. Da sottolineare che la risposta in frequenza delle cuffie Sennheiser HDA 200 è stata calcolata su di un manichino KEMAR senza pinna. In realtà un'esatta riproduzione binaurale è possibile solo per il binomio soggetto-cuffia su cui sono state acquisite le HpIR. Di conseguenza la compensazione delle cuffie valutata sul KEMAR è solo una compensazione media su frequenze sotto i 10 KHz.

### 3.1.2.2 Scheda audio esterna

La scheda audio esterna utilizzata è una Roland Edirol AudioCapture UA-101 mostrata in Figura 3.3. Si tratta di un'interfaccia audio Hi-Speed con connessione USB 2.0. Essa è in grado di gestire contemporaneamente 10 ingressi e 10 uscite alla qualità di 24-bit 96kHz. Grazie alla tecnologia Hi-Speed si garantiscono livelli di latenza estremamente bassi. La scheda è in grado di lavorare a 44.1, 48, 88.2, 96, 192 kHz. Tutti i canali sono bilanciati sia in ingresso che in uscita.

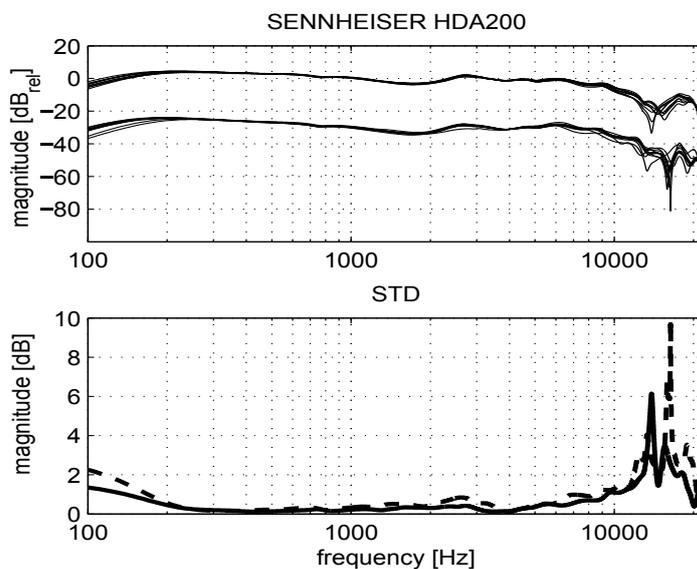


Figura 3.2: Sennheiser HDA 200: risposta in frequenza



Figura 3.3: Roland Edirol UA-101

### 3.1.2.3 Fonometro

Il fonometro utilizzato è un Larson & Davis 824, analizzatore digitale conforme alla IEC-651, IEC-804 tipo 1, IEC 61672 tipo 1, con rilievo contemporaneo di 48 parametri fonometrici e costanti di tempo Fast, Slow, Impulse, Picco, Leq parallele e con pesature A, C e Lin contemporanee, il tutto con una linearità dinamica superiore ai 105 dB.

## 3.2 Test audiometrico

Per il test audiometrico, necessario per verificare che i soggetti sottoposti all'esperimento non avessero problemi uditivi, abbiamo utilizzato, seppur con modifiche apportate dal sottoscritto, il toolbox Matlab MLP [17], sviluppato da Massimo Grassi e Alessandro Soranzo. Il toolbox MLP permette di stimare la soglia di udibilità per mezzo della procedura *Maximum Likelihood (ML)*

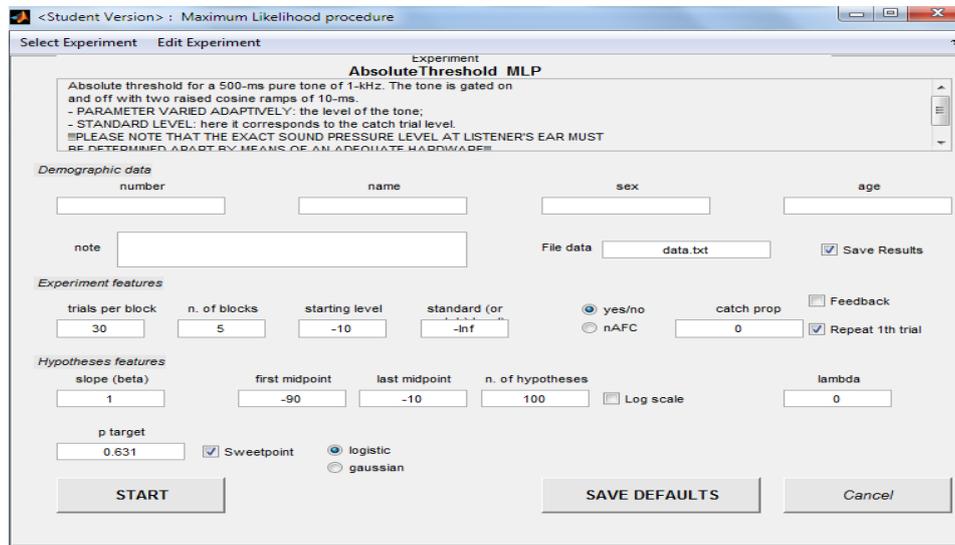


Figura 3.4: MLP: interfaccia grafica

proposta da David Green [18]. Tale procedura permette di calcolare il livello di soglia uditiva con un buon compromesso tra accuratezza e rapidità. L'MLP è il primo software ad implementare tale procedura. Esso è dotato di interfaccia grafica (vedi Figura 3.4) ed è scaricabile<sup>1</sup> gratuitamente.

### 3.2.1 Procedura “Maximum Likelihood”

Come indicato in [18], la procedura di *maximum likelihood* proposta da D. Green è basata su due concetti chiave: funzione psicometrica e *p-target*.

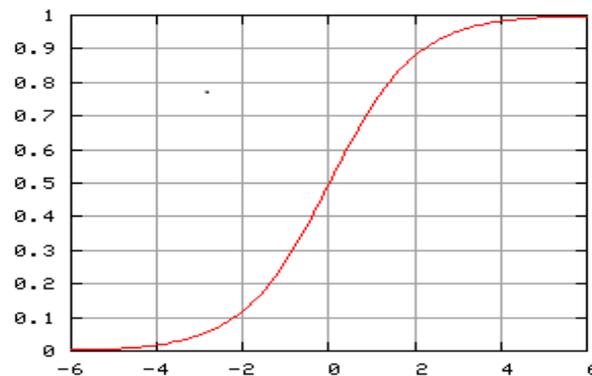
Per funzione psicometrica, nel nostro caso funzione logistica come in Figura 3.5, si intende una funzione del tipo:

$$P(yes) = \alpha + (1 - \alpha - \lambda) \frac{1}{1 + e^{-k(x-m)}} \quad (3.1)$$

dove  $\alpha$  è il false-alarm rate, la probabilità cioè di avere risposta positiva dal soggetto testato anche in assenza di stimolo,  $m$  è la media della funzione logistica (il punto dove la funzione logistica ha valore 0.5),  $\lambda$  è la probabilità di errori di attenzione, e  $k$  è una costante che indica la pendenza della funzione psicometrica. Gli errori di attenzione e i *false-alarm* possono compromettere l'accuratezza nella rilevazione della soglia uditiva poichè, come nel caso dei *false-alarm*, il soggetto dichiara di aver udito uno stimolo sebbene il software non ne abbia inviato alcuno in cuffia. Allo stesso modo, nel caso di errori di attenzione, il soggetto può dichiarare di non aver udito nulla anche al primo stimolo, quello con livello più alto ed udibile da tutti i soggetti senza problemi uditivi. In questo modo la soglia uditiva ricavata risulta seriamente compromessa.

Per *p-target* intendiamo un punto arbitrario sulla funzione psicometrica incluso tra l'estremo inferiore e superiore. Esso corrisponde al livello dello stimolo che minimizza la varianza della stima della soglia uditiva. Esso è quindi utilizzato per ricavare il livello dello stimolo da inviare

<sup>1</sup><http://www.psy.unipd.it/~grassi/psychoacoustics.zip>



**Figura 3.5:** Funzione logistica

al soggetto ad ogni trial, una volta scelta la funzione psicometrica con maggiore *likelihood* tra le ipotesi disponibili. Ovviamente l'ipotesi di funzione psicometrica scelta dipende dalle risposte date dal soggetto agli stimoli già presentati. Il valore ricavato dagli studi di D. Green [18] è 0.6 ovvero il 60% lungo la funzione psicometrica. Quest'ultimo valore ottimizza la stima della soglia uditiva, poichè la varianza associata con la stima di questo *p-target* è minore della varianza associata con altri *p-target*.

Lo sperimentatore ipotizza una serie di funzioni psicometriche da associare al soggetto. Dopo ogni risposta del soggetto allo stimolo, la procedura calcola, per ogni ipotesi, il *likelihood*. Quindi viene selezionata quella che ha maggiore *likelihood*. Il livello dello stimolo successivo da presentare al soggetto viene ricavato dal punto *p-target* sull'ipotesi selezionata. Il *likelihood* per ogni ipotesi viene calcolato dopo ogni trial. Alla fine della procedura sarà selezionata l'ipotesi con maggiore *likelihood*, ovvero quella che più si avvicina a quella personale del soggetto e da essa si ricaverà la soglia uditiva. Si veda in proposito l'articolo di Green [18], dove è presentato un esempio di funzionamento di tale procedura.

### 3.2.2 Protocollo

Il protocollo utilizzato è di tipo adattivo, ovvero i livelli dello stimolo da inviare in cuffia durante i trial dipendono dalla risposta del soggetto. Procedure di tipo adattivo riescono a massimizzare il rapporto tra il numero di stimoli presentati vicino alla soglia uditiva e il numero di stimoli lontani da tale soglia. Essa è anche parametrica, dato che all'inizio della procedura è lo sperimentatore che decide i valori di alcuni parametri. Infine la procedura è del tipo *yes/no*. Al soggetto è richiesto di rispondere *yes*, tramite inserimento da tastiera del carattere '1', se ha udito lo stimolo, e *no*, digitando da tastiera il carattere '0', in caso contrario.

Per far partire l'esperimento bisogna digitare MLP dal prompt di Matlab. Comparirà quindi l'interfaccia di Figura 3.4. L'interfaccia presenta 3 sezioni.

La prima sezione, denominata *Demographic data*, contiene i seguenti campi:

- *Number*;

- *Name*;
- *Sex*;
- *Age*;
- *File data*.

Nel campo *Number* inseriamo a piacimento un numero da associare al soggetto e nel campo *File data* indichiamo il nome del file che conterrà tutti i risultati finali del test. Nei restanti campi, inseriamo i dati del soggetto da testare.

La seconda sezione, denominata *Experiment features*, contiene i seguenti campi:

- *Trials per block*;
- *N. of blocks*;
- *Starting level*;
- *yes/no, nAFC*;
- *Catch prop.*

Il campo *Trials per block* indica quanti stimoli inviare al soggetto, per ogni frequenza, al fine di ricavare la soglia uditiva finale. Gli studi di Green [18] hanno mostrato che con un valore di 12 trial si possono avere ottimi risultati:

The mean threshold estimates appears to be sensible, and a standard deviation of 3 dB at most could be achieved if the threshold estimate were based on only 12 trials.

Abbiamo preferito, per i nostri test, utilizzare un numero di trial pari a 20 poichè, come considerato da Massimo Grassi e Alessandro Soranzo, le stime di Green appaiono piuttosto ottimistiche. Un valore pari a 20 trial è sembrato un buon compromesso tra accuratezza e rapidità. Il campo *N. of blocks* indica quanti blocchi di trial utilizzare. Per ragioni di rapidità, abbiamo preferito utilizzare un solo blocco, dato che, da precedenti prove, la stima della soglia uditiva non sembrava subire rilevanti cambiamenti. Il campo *Starting level* indica il livello dello stimolo iniziale. Abbiamo posto come livello iniziale il valore -10 dB FS, in modo che esso venga udito da tutti i soggetti. Vedremo successivamente a quale valore, in scala dB assoluta, corrisponda tale livello. In base agli studi di Green, il livello iniziale dello stimolo non influenza di tanto la stima finale della soglia uditiva. Nella seconda sezione è necessario anche scegliere il tipo di test da effettuare tra le opzioni *yes/no* e *nAFC*. Nel nostro caso scegliamo *yes/no*. Infine, l'ultimo campo da riempire è *Catch prop.* L'inserimento di tale parametro è dovuto al fatto che, come dimostrato da Green, la procedura di *maximum likelihood* sottostima il *false-alarm* rate. Per ridurre tale sottostima si introducono dei *catch trials* durante la stima della soglia uditiva. Durante i *catch trials*, il livello dello stimolo è settato a zero o al minimo livello del range scelto. Nel calcolo del *likelihood* delle varie ipotesi, utilizziamo anche le risposte date durante i *catch trials*. Ovviamente ci aspetteremmo che la risposta sia negativa, pochè il livello dello stimolo è zero, a meno che

il soggetto non stia dando una risposta condizionata errata. Lo sperimentatore ha quindi la possibilità di inserire in tale campo la proporzione di *catch trials* all'interno degli stimoli presentati. Si raccomanda di tenere il *catch trials* rate a circa il 20% del numero totale di trial.

La terza sezione, denominata *Hypotheses features*, contiene i seguenti campi:

- *Slope*;
- *First midpoint*;
- *Last midpoint*;
- *N. of hypotheses*;
- *Lambda*.

Il campo *Slope* indica la pendenza delle funzioni psicometriche logistiche. La procedura di *maximum likelihood* utilizza una serie di funzioni psicometriche (*hypotheses*), tutte con la stessa pendenza. Esse però differiscono poichè si distribuiscono lungo un certo range del livello degli stimoli. Purtroppo lo sperimentatore non può sapere con certezza lo slope della funzione psicometrica associata al soggetto. Tuttavia, gli studi di Green [18] dimostrano che:

[...]the match between the observer's psychometric function and that used in the analysis is important, but not critical to the success of the maximum-likelihood procedure.

Abbiamo quindi scelto per i nostri test  $k=0.5$  per simulare il range della funzione psicometrica come trovato in altri studi [40].

Con i successivi tre campi *First midpoint*, *Last midpoint*, *N. of hypotheses* scegliamo il range e il numero di ipotesi. Il range delle ipotesi deve coprire di molto il range dei livelli degli stimoli dove pensiamo si trovi la soglia uditiva del soggetto. Di conseguenza, il valore del campo *First midpoint* deve trovarsi di molto sotto rispetto a tale soglia, mentre il valore del campo *Last midpoint* deve trovarsi ben sopra rispetto a tale soglia. Il campo *N. of hypotheses* indica quante ipotesi di funzioni psicometriche scegliamo per il nostro esperimento. Ovviamente, maggiore è il numero di ipotesi e più vicina sarà la soglia uditiva trovata a quello reale del soggetto. Purtroppo, aumentando di tanto il numero di ipotesi, aumenterà di conseguenza anche il carico computazionale. Abbiamo quindi scelto di utilizzare un numero di ipotesi pari a 100. Infine, nel campo *Lambda* indichiamo la percentuale di errori di attenzione. Per il nostro esperimento abbiamo scelto di non considerare errori di attenzione e quindi di fissare tale valore a 0.

Gli ultimi parametri da inserire prima di far partire il test sono i seguenti:

- *P target*;
- *Logistic/Gaussian*.

Il campo *P target* si riferisce al punto sulla funzione psicometrica di cui abbiamo discusso sopra. Inseriamo quindi come valore 0.6. Infine scegliamo *Logistic* come tipo di funzione psicometrica.

A questo punto, cliccando sul pulsante *Start*, facciamo partire il test audiometrico. Per il test abbiamo utilizzato le cuffie Sennheiser HDA 200, le cui caratteristiche sono descritte nella

sezione 3.1.2.1. Con la modifica al codice da me effettuata, la procedura di *maximum likelihood* viene ripetuta per le seguenti frequenze, caratterizzanti i test audiometrici:

125, 250, 500, 1000, 2000, 3000, 4000, 6000, 8000 Hz

Per ogni frequenza testata, il primo stimolo inviato in cuffia risulta chiaramente udibile da tutti i soggetti. A fine presentazione di ogni stimolo, sul monitor compare un messaggio in cui si invita il soggetto a digitare su tastiera '1' se lo stimolo è stato udito, '0' se al contrario lo stimolo non è stato udito. Si attende quindi la risposta del soggetto. In base alla risposta del soggetto, il *maximum likelihood* viene calcolato, e il successivo stimolo viene inviato dopo circa 1 sec. Dopo 20 trial, parametro da noi inserito nell'interfaccia grafica iniziale, viene mostrata la soglia uditiva stimata per quella determinata frequenza. Il tutto viene ripetuto anche per le altre frequenze. Tale procedura viene effettuata prima sull'orecchio sinistro e dopo su quello destro. L'intero test, considerando orecchio sinistro e orecchio destro, dura circa 30 minuti. A fine procedura, MLP mostra a schermo l'audiogramma del soggetto (Figura 3.7), salva due file di dati e un file di formato .fig contenente l'audiogramma del soggetto. Il primo file di dati, il cui nome corrisponde a quello indicato dallo sperimentatore nel campo *file data*, contiene i seguenti campi:

- *Subject number*;
- *Name*;
- *Sex*;
- *Age*;
- *Note*;
- *Block number*;
- *Trial number*;
- *Level stimulus*;
- *Subject's response*;
- *Threshold (per ogni trial)*;
- $\gamma$ : false-alarm rate.

In questo file le risposte del soggetto vengono indicate con '1', per risposta affermativa, e '0' per risposta negativa. Il secondo file contiene il *threshold* del soggetto per ogni frequenza. Nella Tabella 3.2 è indicata la lista dei soggetti testati.



**Figura 3.6:** Setup per la calibrazione delle cuffie: Cuffia, sostegno e fonometro.

### 3.2.3 Calibrazione e rappresentazione grafica dell'hearing threshold

I valori di soglia uditiva ottenuti a fine procedura sono in dB FS (full scale). Si è reso necessario quindi trasformare i valori in dB assoluti. Per effettuare tale conversione abbiamo effettuato una calibrazione del toolbox per mezzo di un fonometro. Lo script per la calibrazione è denominato `calibration.m`. Esso accetta, come parametro di ingresso, la frequenza per la quale calibrare il livello dello stimolo in uscita. Tale calibrazione viene effettuata inviando allo script, come parametro d'ingresso, la frequenza di 1000 Hz. Successivamente, per mezzo della funzione `GenerateTone`, viene sintetizzato un suono con tale frequenza e attenuato al livello di -10 dB FS. Tale suono viene quindi riprodotto e per mezzo del fonometro misuriamo a quanti dB SPL corrispondono -10 dB FS, una volta fissato all'uscita del pc un volume confortevole per i soggetti che si sottopongono al test che, nel nostro caso, è pari al 25% del volume totale. Il valore rilevato dal fonometro è pari a 78.6 dB SPL. In Figura 3.6 è mostrato il setup utilizzato per la calibrazione delle cuffie.

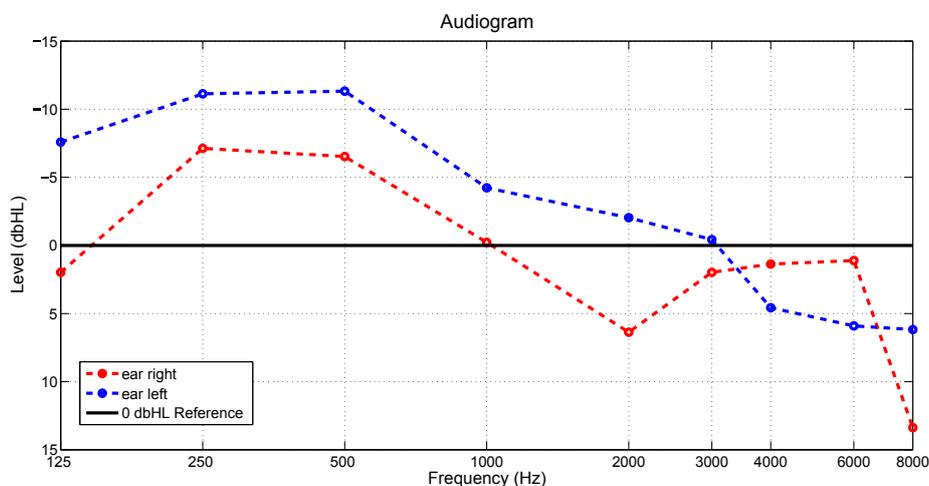
La differenza tra il dB FS utilizzato dal toolbox e il dB SPL del fonometro è quindi in questo caso 88.6 dB:  $78.6 \text{ dB SPL} - (-10 \text{ dB FS})$ . In altre parole, utilizzando come offset il valore ricavato di 88.6 dB SPL, se il *threshold* trovato dal toolbox risultasse -40 dB FS, in dB SPL quest'ultimo sarebbe pari a 48.6. Convertire tutti i valori in dB SPL è necessario per confrontare la soglia uditiva trovata con la soglia minima di udibilità. Sovrapponendo la soglia minima di udibilità con la soglia uditiva trovata ricaviamo se il soggetto in questione ha o meno carenze uditive. Affinchè il soggetto non presenti problemi uditivi, la soglia uditiva trovata non dovrebbe essere, per almeno una delle due orecchie, al di sopra della soglia minima più di 25 dB, come indicato in ANSI 3.6-1996 (Specifications for Audiometers). Infatti, essendo la localizzazione sul piano mediano un processo monoaurale, come indicato in [21], il superamento di tale soglia per una delle due orecchie non influisce sulla localizzazione delle sorgenti. I risultati confermano che tutti i soggetti testati sono idonei a partecipare al test di localizzazione.

Frequenza (Hz)	RETSPL (dB SPL)
1000	5.5
2000	4.5
3000	2.5
4000	9.5
6000	17
8000	17.5
500	11
250	18
125	30.5

**Tabella 3.1:** *Reference equivalent threshold sound pressure levels per cuffie Sennheiser HDA 200*

### 3.2.3.1 Brevi note di audiologia

In audiologia, per indicare la capacità uditiva di una persona, si utilizza il decibel audiometrico, dB HL (decibel Hearing Level), che esprime in dB il Livello Uditivo e indica la capacità uditiva della persona rispetto agli standard normali. Per dB HL si definisce a zero il valore di minima pressione sonora (in dB SPL) udibile a tutte le frequenze, cioè lo 0 dB HL rappresenta, per una data frequenza, la soglia uditiva di un soggetto normoudente. Le soglie uditive standard di un soggetto normoudente in dB SPL, cioè i livelli di pressione sonora, sono indicati nello standard ANSI S3.6-1996, per una gamma di frequenze -da 125 Hz a 8000 Hz- che comprende le frequenze più significative che caratterizzano i suoni ambientali quotidiani e la voce. Tali livelli standard sono in relazione a tutti i tipi di cuffie comunemente usate per gli esami audiometrici. In base a queste soglie di udibilità, in audiologia clinica si trasformano i dB SPL in dB HL sottraendo ai dB SPL dette soglie di udibilità definite per ogni frequenza. Per quanto riguarda il nostro test audiometrico, in Tabella 3.1 sono riportate le soglie di udibilità per le Sennheiser HDA 200, ricavate da ISO 389-8:2004 [1]. In questo caso, per un tono puro di frequenza 1000 Hz, il valore di 0 dB HL corrisponde al valore di 5.5 dB SPL: questo implica che per un suono a questa frequenza di intensità pari a un determinato valore dB SPL, per ottenere il corrispondente livello dB HL, occorre sottrarre il valore 5.5 al dB SPL, cioè un suono di intensità pari a 40 dB SPL corrisponde ad un livello di 34.5 dB HL. Dall'esame audiometrico si costruisce l'audiogramma, che è il grafico delle soglie di udibilità di un orecchio in funzione della frequenza del suono. Il grafico dell'audiogramma ha sull'asse delle ascisse le frequenze dei toni puri utilizzati nell'esame tonale e sull'asse delle ordinate i valori delle soglie di udibilità, in dB HL, per quelle frequenze. Le soglie di udibilità standard, in dB SPL, regolamentate dallo standard ANSI S3.6-1996, corrispondono tutte a 0 dB HL. Un valore maggiore di 0 dB HL significa che, per una data frequenza, il suono udibile a quella frequenza ha un livello sonoro maggiore rispetto al livello standard di riferimento per un normoudente, quindi indica una perdita uditiva. Ad esempio un valore di 20 dB HL indica che la persona ha una perdita uditiva di 20 dB. L'udito normale è considerato relativo a livelli compresi tra -10 dB HL e 25 dB HL. In Figura 3.7 è riportato un esempio di audiogramma ottenuto alla fine della procedura MLP. Da notare, purtroppo, un deterioramento



**Figura 3.7:** Audiogramma soggetto con ID 20.

di performance alle frequenze 3000 e 4000 Hz dovute alla non linearità della risposta delle cuffie sopra e a queste frequenze, come indicato nelle Sezione 3.1.2.1, ed una non affidabilità per le frequenze 6000 e 8000 Hz.

Per la calibrazione delle cuffie Sennheiser HDA 200 abbiamo fatto riferimento all'ISO 389-8:2004 [1], contenente i RETSPLs relativi a questa cuffia, indicati in Tabella 3.1. Per RETSPLs si intende *reference equivalent threshold sound pressure levels*. Sono le intensità fisiche necessarie per un normoudente a raggiungere la soglia uditiva. Esse variano da cuffia a cuffia. Ognuna di esse ad ogni frequenza è il suono più flebile che possa essere sentito; di conseguenza ognuno di queste soglie corrisponde a 0 dB HL. Per le Sennheiser, ad esempio, sono necessari 30.5 dB per avere 0 dB HL ad una frequenza di 125 Hz, 18 dB per 0 dB HL a 250 Hz, 11 dB per 0 dB HL a 500 Hz, etc... Quindi abbiamo che a 500 Hz dove il riferimento è 11 dB SPL per lo 0 dB HL, 60 dB HL risultano in:

$$60 + 11 = 71 \text{ dB SPL}$$

## 3.3 Test psicoacustico

### 3.3.1 Soggetti

I soggetti testati, sia di sesso maschile che femminile, sono stati scelti dopo essere stati sottoposti al test audiometrico di cui sopra. Tutti hanno riportato, almeno per una delle due orecchie, un soglia uditiva inferiore a 25 dB HL. Il test audiometrico e il test psicoacustico di localizzazione sono stati effettuati in giorni differenti al fine di non affaticare i soggetti. L'età dei soggetti variava tra i 23 e i 41 anni. Solo alcuni dei soggetti avevano precedentemente avuto esperienze con il suono spazializzato e nessuno è stato sottoposto a sessioni di training. La Tabella 3.2 contiene la lista dei soggetti che hanno partecipato all'esperimento.

ID	Sesso	Età
20	M	33
21	M	28
22	M	28
24	M	39
26	F	32
27	F	33
28	M	37
30	M	41
32	M	27
33	F	23
35	M	23
36	M	23

**Tabella 3.2:** *Soggetti esperimento*

### 3.3.2 Stimoli

Ogni stimolo è costituito da un treno di 3 impulsi di rumore gaussiano, ognuno della durata di 300 ms. Esso è presentato ad un livello di 60 dbA misurati a livello della pinna sulle cuffie. Gli impulsi sono separati da pause di 250 ms. A tal proposito, è stato mostrato che:

sounds having energy for all of their frequency bands, such as white noise, are easier to localize than sounds having a sparse spectral content (sine tone) [12].

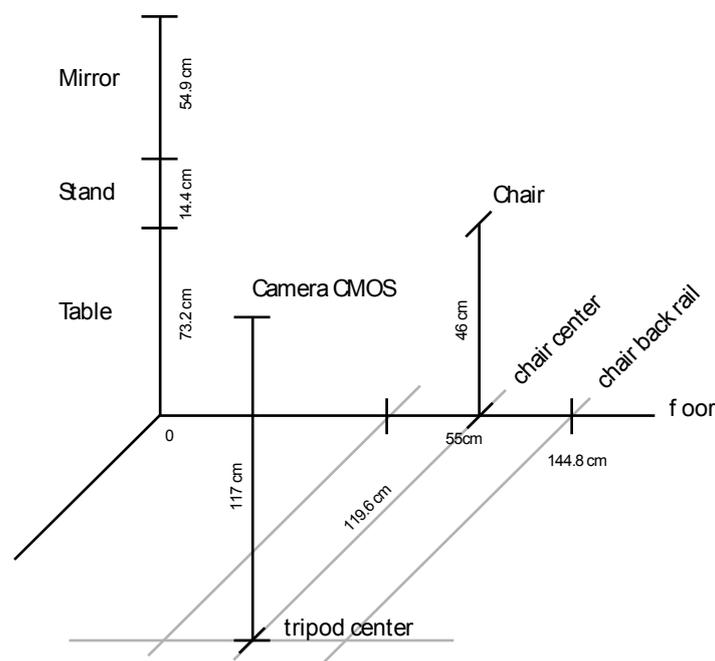
I rumori sono stati finestrati con una finestra di Hann in modo da ottenere fade-in e fade-out. I rumori di breve durata permettono di prevenire, durante l'esperimento, movimenti della testa, quindi risultano adatti a esperimenti di tipo statico.

Lo stimolo finale presentato al soggetto, come vedremo nella Sezione 3.3.5.2, è ottenuto tramite convoluzione con una delle nostre PRTF sintetiche e filtraggio per la compensazione delle cuffie utilizzate per l'esperimento, argomento trattato nella sezione 3.1.2.1.

### 3.3.3 Condizioni

Le condizioni testate sono 12, tutte riferite a parametri dei notch delle PRTF. Come riportato in Tabella 2.3, per ogni condizione abbiamo 3 campi: profondità, larghezza di banda e numero di notch. Le profondità prese in esame in dB sono: -10, -30. Per ciò che concerne la larghezza di banda, distinguiamo larghezze di banda relative alla frequenza centrale del notch e larghezza di banda fissa. Quelle relative alla frequenza centrale sono:  $0.15f_c$ ,  $0.25f_c$ , mentre la larghezza di banda fissa è stata posta a 2000 Hz. Abbiamo infine condizioni caratterizzate dalla presenza di 3 notch e condizioni che ne presentano solo 2 (togliendo quello centrale).

Le ragioni che hanno portato a tale scelta di parametri derivano da analisi di studi precedenti e sono descritte nella Sezione 2.3.2.



**Figura 3.8:** Setup per l'acquisizione dell'immagine della pinna.

Nel corso del test di localizzazione ogni stimolo viene presentato 4 diverse volte per ogni posizione e condizione. Tale scelta permette di evitare che i risultati finali siano influenzati da possibili errori di attenzione ed inoltre permette di considerare un possibile effetto *learning*, grazie al quale il soggetto testato, nel corso dell'esperimento, diviene aduso ai suoni spazializzati.

Considerando 7 elevazioni e 12 condizioni, otteniamo un totale di:

$$7 \text{elevazioni} * 12 \text{condizioni} * 4 \text{ripetizioni} = 336 \text{stimoli}$$

Le elevazioni testate in gradi sono: [-45,-30,-15,0,15,30,45], tutte sul piano frontale e medio, espresse utilizzando il sistema di coordinate verticali polari. Le ragioni che hanno portato a considerare tale regione spaziale sono state descritte nella Sezione 2.2. Gli stimoli vengono presentati in 4 blocchi da 84 stimoli ciascuno ed in ogni blocco vengono presentati tutti gli stimoli associati ad una ripetizione. L'ordine degli stimoli è una permutazione random. A fine di ogni blocco, al soggetto viene data una pausa di tre minuti, durante la quale è libero di restare nella sala silenziosa oppure di uscire.

### 3.3.4 Acquisizione immagine della pinna e tracciamento

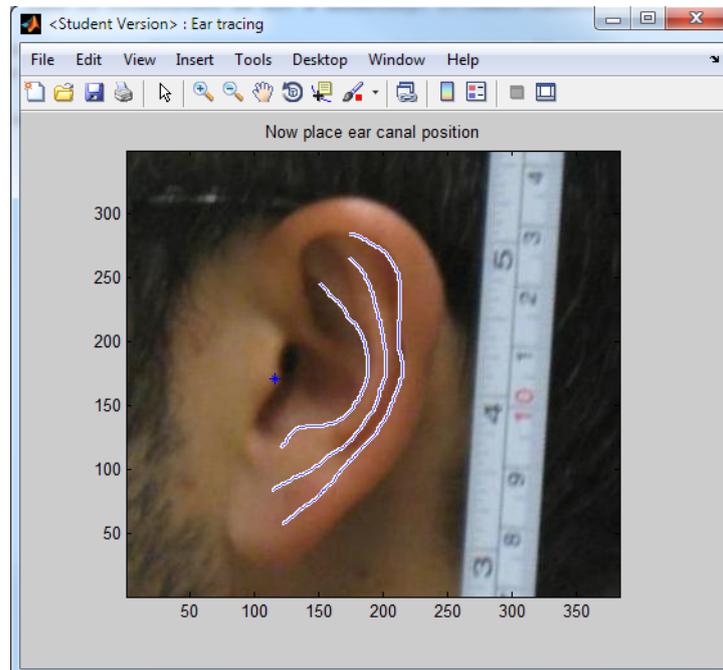
I soggetti che partecipano all'esperimento vengono fatti sedere all'interno di una stanza. All'interno di essa abbiamo creato un setup controllato al fine di acquisire un'immagine, mostrato in Figura 3.8. Al soggetto viene chiesto di centrare la propria testa rispetto ad una linea verticale sovrapposta ad uno specchio davanti ad esso e alternativamente di chiudere un occhio e muovere la testa in modo da sovrapporre l'occhio chiuso alla linea dritta. In questo modo vengono fissate,



**Figura 3.9:** Immagine di pinna acquisita tramite fotocamera.

riferite al soggetto, due delle tre dimensioni dello spazio, in particolare lo *yaw* e il *roll*. Quindi per mezzo di una fotocamera posta su di un treppiede acquisiamo l'immagine della pinna, solo dopo aver verificato, per mezzo di una livella, l'orizzontalità della fotocamera. La macchina fotografica utilizzata è una Canon Digital IXUS 85 IS con lunghezza focale 6,2-18,6 mm (equivalente a 35 – 105 mm in formato 35 mm). Il sensore CMOS della fotocamera è perpendicolare alla pinna del soggetto. Come è possibile vedere in una immagine di pinna fotografata (Figura 3.9), è stato posto un righello vicino all'orecchio del soggetto in modo da aver un fattore di conversione pixel-metri. L'immagine così acquisita viene manipolata in Matlab per mezzo di una GUI precedentemente sviluppata in [6]. Per mezzo di tale GUI è possibile scalare l'immagine, ruotarla, ed applicare il fattore di conversione pixel-metri. Grazie alla rotazione è possibile allineare orizzontalmente il tragus con la punta del naso del soggetto. Grazie a questo lavoro di post-produzione fissiamo la terza dimensione (*pitch*). A questo punto l'immagine della pinna è pronta per poterne tracciare i contorni per mezzo di una GUI realizzata in Matlab. A tal fine ci siamo serviti di un tablet *Wacom* largo abbastanza in modo da facilitare il tracciamento della pinna. Vengono tracciati i 3 contorni su cui avvengono le riflessioni dell'onda sonora come indicato in figura 3.10. Si parte con il tracciamento più esterno effettuato sull'elice, per poi continuare con il tracciamento dell'antielice e della conca. Infine indichiamo la posizione del canale uditivo (*ear canal*). Vengono quindi calcolate le distanze tra il canale uditivo e i contorni di riflessione. Tali dati vengono salvati all'interno di un file *.mat* associato all'immagine della pinna. Tale file contiene:

- il vettore  $F_0$ , di dimensioni  $3 \times 7$ , con le frequenze estratte dai 3 contorni, relativamente alle 7 elevazioni testate. Ogni riga del vettore  $F_0$  è associata ad un contorno: la prima riga rappresenta le frequenze relative al contorno più esterno, la seconda rappresenta le frequenze relative al contorno centrale, mentre l'ultima riga è associata al contorno più interno. Ogni colonna è invece associata ad una delle 7 elevazioni testate;
- i vettori  $C_x$  e  $C_y$  contenenti i punti tracciati sui 3 contorni;



**Figura 3.10:** Tracciamento dei 3 contorni sulla pinna del soggetto.

- *earcanal* contenente la posizione del canale uditivo;
- *dist*, vettore con le distanze tra il canale uditivo e i punti sui contorni;
- *pixelToMeterFactor* contenente il fattore di conversione pixel-metri.

L'immagine della pinna in formato *.jpg* ed il relativo file *.mat*, contenente i dati di cui sopra, vengono salvate all'interno della cartella *img*.



**Figura 3.11:** *Cabina silente Sound Station Pro 45 by Puma s.r.l..*

### 3.3.5 Procedura

I test, sia quello di localizzazione che quello audiometrico, sono stati effettuati all'interno di una cabina silente<sup>2</sup> del tipo rappresentato in Figura 3.11. I soggetti sono stati istruiti sulle fasi dell'esperimento ed è stata mostrata loro l'interfaccia grafica su cui indicare la posizione degli stimoli percepita. L'interfaccia grafica è stata realizzata tramite l'ambiente grafico GUIDE messo a disposizione da Matlab. La Figura 3.12 mostra la GUI di Matlab utilizzata nella gestione del database dei soggetti del test, mentre la Figura 3.13 mostra la GUI tramite cui i soggetti testati indicano la posizione della sorgente da loro rilevata. Al soggetto vengono fatte indossare le cuffie Sennheiser HDA 200 e l'esperimento ha inizio digitando *StartMain* dal prompt di Matlab. Vediamo ora in particolare le 2 GUI che compaiono nel corso dell'esperimento.

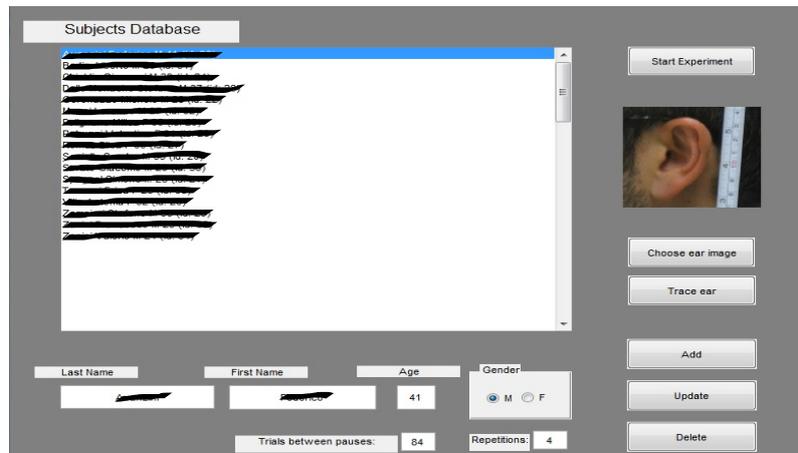
#### 3.3.5.1 GUI soggetti

La GUI in Figura 3.12 è quella che compare inizialmente. Essa viene lanciata da `startMain.m` che a sua volta chiama `subject.m`. Essa è utilizzata per la gestione del database dei soggetti ed è costituita dai seguenti *push button*:

- *Start experiment;*
- *Choose ear image;*
- *Trace ear;*
- *Add;*

---

<sup>2</sup>Sound Station Pro 45 by Puma s.r.l



**Figura 3.12:** *Interfaccia GUI contenente il database con la lista dei soggetti testati.*



**Figura 3.13:** *Interfaccia GUI utilizzata dal soggetto per indicare le posizioni rilevate.*

- *Update;*
- *Delete.*

Con il tasto *Start experiment* lo sperimentatore lancia la GUI in Figura 3.13. Tale tasto deve essere premuto solo dopo aver selezionato il soggetto da testare dalla lista dei soggetti, se già presente, e aver tracciato, cliccando sul tasto *Trace ear*, i contorni della pinna come indicato nella sezione 3.3.4. Nel caso si trattasse di un nuovo soggetto non presente nel database, sarà necessario compilare i campi:

- *Last name;*
- *First name;*
- *Age;*

- *Gender*.

relativi al nuovo soggetto e cliccare sul tasto *Add*. In questo modo il nuovo soggetto è inserito all'interno del database. Il tasto *Delete* serve ad eliminare un soggetto dal database ed il tasto *Update* è utilizzato per aggiornare i dati del soggetto. Il tasto *Choose ear image* è utilizzato per associare un soggetto all'immagine della sua pinna, precedentemente fotografata come indicato nella sezione 3.3.4. Il campo *Trials between pauses* è utilizzato per indicare ogni quanti stimoli è inserita una pausa nel corso dell'esperimento ed in particolare, per il nostro esperimento, tale valore è fissato a 84. Il campo *Repetitions* indica il numero di ripetizioni di ogni stimolo per ogni condizione considerata. Come visto in precedenza, tale valore è fissato a 4. Una volta riempiti tutti i campi è necessario cliccare sul tasto *Start experiment*. Allo script Matlab realizzato per la creazione dell'interfaccia GUI esperimento vengono passati come parametri:

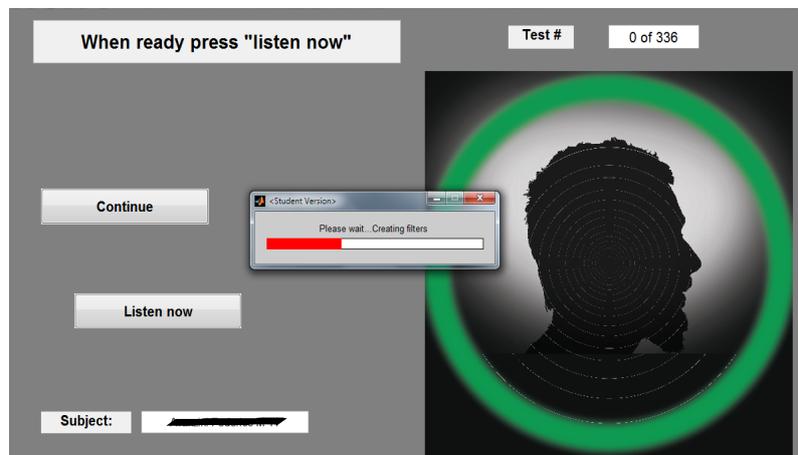
- *ID* del soggetto da testare;
- Il valore presente nel campo *Repetitions*;
- Il valore presente nel campo *Trial between pauses*.

I dati sono passati tra le GUI utilizzando le funzioni `setappdata(...)` e `getappdata(...)`.

### 3.3.5.2 GUI esperimento

La GUI esperimento (Figura 3.13) viene richiamata dalla GUI soggetti cliccando sul tasto *Start experiment* presente in quest'ultima. Servendosi dei parametri che gli sono stati passati, lo script alla base dell'interfaccia GUI esperimento carica il vettore *F0*, salvato alla fine della procedura di tracciamento dei contorni, relativo all'ID del soggetto testato. Ricordiamo che tale vettore contiene, associati ai contorni estratti sulla pinna, le frequenze dei 3 notch relativamente alle 7 elevazioni testate. Quindi lancia lo script `create_subject` passandogli, come parametri, il vettore *F0* caricato e l'ID del soggetto. Compare a questo punto la finestra *Creating filters* (Figura 3.14), avvisandoci che è in corso la preparazione dei filtri sintetici. Lo script `create_subject` è utilizzato per la creazione dei filtri con cui realizzare gli stimoli da inviare al soggetto; in particolare, lo script, per ognuna delle 7 elevazioni, svolge le seguenti funzioni:

- crea, per ognuna delle 12 elevazioni, il vettore delle larghezze di banda, il vettore delle profondità e quello delle frequenze centrali dei notch;
- lancia lo script `prtf`, tramite cui viene creata la PRTF sintetica associata alla condizione in esame;
- lancia lo script `stimulus3`, tramite cui viene creato uno stimolo di 3 impulsi di rumore gaussiano;
- effettua il filtraggio della PRTF creata con lo stimolo ottenuto al passo precedente;
- effettua il filtraggio per la compensazione delle cuffie;



**Figura 3.14:** Finestra di avviso di creazione dei filtri in corso.

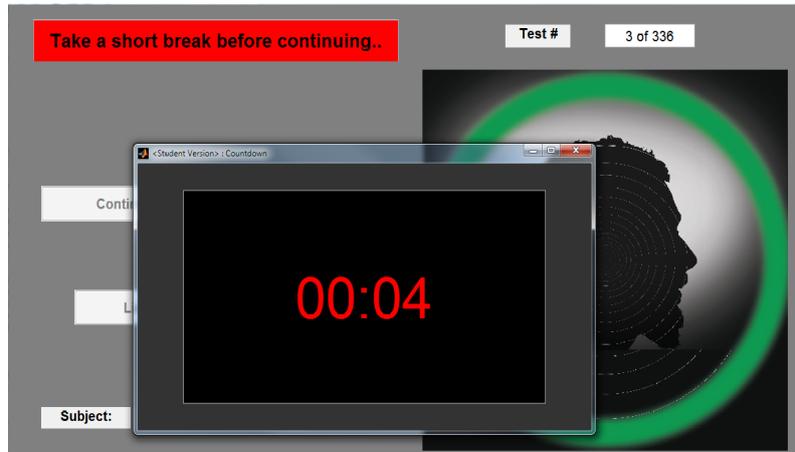
- crea, all'interno della cartella `subjects`, la cartella `subjectID` con ID relativo al soggetto testato. All'interno di essa vengono create due ulteriori sottocartelle: la cartella `filters` contenente 84 filtri in file di formato `.mat` e la cartella `stimoli.wav` contenente appunto gli stimoli in formato `.wav`. All'interno della cartella relativa al soggetto testato, viene creato anche il file `blockTest.mat` contenente l'ordine di presentazione degli stimoli.

La GUI esperimento contiene i seguenti due campi:

- *Test*;
- *Subject*.

Il campo *Test* indica quanti stimoli sono già stati presentati al soggetto. Il campo *Subject* indica il nome, cognome, sesso ed età del soggetto testato. All'interno della GUI troviamo il tasto *Listen now* utilizzato inizialmente per far partire la presentazione degli stimoli e poi disabilitato. La GUI contiene anche un'immagine di profilo di una testa circondata da una corona circolare di colore verde. Essa rappresenta approssimativamente un insieme di punti equidistanti dal centro della testa sul piano mediano. Durante l'esperimento al soggetto viene chiesto di indicare la posizione della sorgente rilevata cliccando con il mouse su un punto all'interno della corona verde, poichè gli stimoli che presentiamo sono tutti sul piano mediano, ignorando la distanza della sorgente percepita rispetto al centro della testa. Al soggetto non viene data nessuna indicazione delle elevazioni testate. Egli è libero di scegliere qualsiasi punto sulla circonferenza. Infine abbiamo un *text box* su cui compaiono le indicazioni sulle varie fasi dell'esperimento e su come il soggetto debba riportare il suo feedback all'interno dell'interfaccia. Oltre al testo, anche il colore del *text box* cambia. Le varie combinazioni di colore e testo sono le seguenti:

- Colore grigio *When ready press listen now*;
- Colore giallo *Sound playing*;



**Figura 3.15:** Finestra di countdown mostrata durante la pausa tra blocchi di stimoli.

- Colore verde *Choose elevation inside the green circle;*
- Colore rosso *Loading next sound.*

L'esperimento comincia cliccando sul tasto *Listen now*, si attende la presentazione dello stimolo durante la quale compaiono i due messaggi *Loading next sound* e *Sound playing*, ed al messaggio *Choose elevation inside the green circle* si attende il feedback del soggetto. Il soggetto, utilizzando un comune mouse, indica sull'immagine presente nell'interfaccia grafica un punto, nella direzione in cui crede sia arrivato il suono, esclusivamente sulla corona verde raffigurante il piano mediano. Subito dopo il click, un nuovo stimolo viene inviato in cuffia. Dopo la presentazione di ogni blocco da 84 stimoli, una finestra di countdown compare (Figura 3.15). Al soggetto viene concessa una pausa di tre minuti prima di proseguire con l'esperimento. Nell'arco del test abbiamo inserito 4 pause, dato comunque modificabile dall'interfaccia soggetti. Al termine del test i dati con le risposte del soggetto vengono salvate all'interno della cartella *results*, in un file di tipo *.mat* e nominato con l'ID del soggetto. Da segnalare che alcuni soggetti hanno preferito effettuare il test ad occhi aperti, mentre altri ad occhi chiusi. Non sono state date indicazioni al riguardo da parte dello sperimentatore. Quest'ultimo, durante l'intera fase dell'esperimento, ha la possibilità di monitorare il soggetto testato da fuori la cabina. L'intero test di localizzazione dura, considerando le 4 pause inserite, circa 45 minuti.

# Capitolo 4

## Risultati

### 4.1 Considerazioni preliminari

Prima di approfondire i risultati dei nostri esperimenti per ogni singolo soggetto, presentiamo alcune utili considerazioni emerse durante lo studio dei dati ed indichiamo il tipo di analisi utilizzato.

I due più comuni tipi di errori sono:

1. *front/back confusion*: l'ascoltatore indica la posizione della sorgente di un suono nell'emisfero sbagliato; in particolare, un suono proveniente da davanti viene percepito come posto dietro l'ascoltatore e viceversa;
2. *up/down confusion*: un suono posto sotto il livello degli occhi viene percepito sopra e viceversa; in particolare, si ha *up/down confusion* quando la posizione ideale del suono è al di sopra del piano orizzontale mentre la posizione percepita è nel semipiano inferiore al di sotto dell'elevazione di  $-15^\circ$  e, viceversa, quando la posizione ideale del suono è al di sotto del piano orizzontale mentre la posizione percepita è nel semipiano superiore al di sopra dell'elevazione di  $15^\circ$ .

La presenza di errori di tipo *front/back* è giustificata dal fatto che le sorgenti utilizzate per il nostro esperimento sono posizionate tutte sul piano mediano, dove i valori di ITD (*interaural time difference*) e ILD (*interaural level difference*) sono prossimi allo zero. Di conseguenza l'ITD e l'ILD non possono essere utilizzati dall'ascoltatore come indicatori della posizione del suono nello spazio. Da studi precedenti [42] risulta che la localizzazione di sorgenti virtuali produce un tasso di errori di *front/back* superiore alla localizzazione di sorgenti reali. Abbiamo quindi deciso di correggere tali tipi di errori, prima di effettuare l'analisi dei risultati. La correzione consiste nel mappare, con simmetria assiale, tutti i valori posti nell'emisfero posteriore nell'emisfero frontale. Tale procedura è una pratica ormai comune, come mostrato in [42]. In realtà per il nostro esperimento parleremo piuttosto di *front-back reversal* poichè i suoni prodotti sono tutti sul piano frontale e quindi l'unico caso di confusione possibile si ha quando il soggetto posiziona sul retro un suono frontale, non potendo mai presentarsi il caso contrario.

Per quanto riguarda gli errori di tipo *up/down*, o anche errori di quadrante, essi non sono stati corretti ma considerati nella stima della condizione migliore globale.

Oltre ai due già citati errori, per ogni risposta del soggetto abbiamo calcolato l'errore di elevazione tra posizione ideale del suono (*target*) e posizione percepita. Tale errore può considerarsi come errore di precisione.

Per ciò che concerne l'analisi dei risultati, per mezzo dello script `plot_results`, per ogni soggetto vengono calcolati, divisi per condizione e memorizzati in un file denominato `conditions_id_nr.mat`, dove `nr` indica l'ID del soggetto, i seguenti dati:

- *mean\_angle\_error*: media degli angoli di errore per tutte le 7 elevazioni;
- *angle\_of\_error*: angolo di errore medio per ognuna delle 7 elevazioni;
- *standard\_deviation*: deviazione standard dell'errore per ognuna delle 7 elevazioni;
- *front\_back\_reversal\_rate*: percentuale di errori *front/back* sul totale degli stimoli;
- *up\_down\_confusion\_rate*: percentuale di errori *up/down* sul totale degli stimoli;
- *regression\_slope*: pendenza della retta di regressione per i dati risultanti alla luce dell'applicazione della tecnica di analisi iterativa, in seguito illustrata;
- *rsquared* o *R2*: valore che indica quanto vicini sono i dati alla retta di regressione trovata alla fine della tecnica di analisi iterativa;
- *regression\_bias*: intercetta della retta di regressione;
- *polar\_error\_rate*: percentuale di *polar\_errors* (di seguito definiti).

I campi *mean\_angle\_error*, *angle\_of\_error*, *standard\_deviation*, *front\_back\_reversal\_rate* e infine *up\_down\_confusion\_rate* si riferiscono all'analisi di primo ordine. In quest'ultima, per ogni condizione, calcoliamo la probabilità di errori di *front\_back\_reversal* e di *up\_down\_confusion*, l'angolo di errore medio (*angle\_of\_error*) e la deviazione standard (*standard\_deviation*) per ognuna delle 7 elevazioni, ed infine la media degli angoli di errore (*mean\_angle\_error*).

I campi *regression\_slope*, *rsquared*, *regression\_bias*, *polar\_error\_rate* si riferiscono tutti ad una tecnica di analisi iterativa da noi effettuata, basata sulla procedura contenuta in [26]. In tale studio, nella parte relativa all'analisi dei risultati, viene applicata una procedura iterativa per mezzo della quale gli autori cercano di eliminare le risposte fuori *target* (*outliers*) e di isolare la principale concentrazione di risposte per calcolarne il miglior predittore lineare. Inizialmente, con la totalità dei dati a disposizione, viene calcolata la regressione tra risposte date dal soggetto e *target*. Successivamente, i valori che distano dalla retta di regressione più di una certa distanza  $\Delta$ , fissata dagli autori a  $40^\circ$ , vengono temporaneamente eliminati e viene ricalcolata la regressione lineare con i nuovi dati scremati. Tale procedura viene ripetuta 2-3 volte ed i dati eliminati in una prima fase vengono considerati nuovamente per il calcolo della distanza rispetto alla nuova retta di regressione, rientrando tra le risposte "buone" in caso di distanza  $< \Delta$  o nuovamente rigettate in caso contrario. Viene infine calcolata la distanza tra le risposte date dal soggetto e la

retta di regressione trovata a fine procedura con pendenza indicata nel campo *regression\_slope*, e le risposte per le quali tale distanza è  $> 45^\circ$  vengono considerate *polar errors* ed eliminate dal grafico dei risultati. Il tasso di *polar errors* eliminati è riportato nel campo *polar\_error\_rate*. Nel campo *rsquared* è riportato il coefficiente di determinazione ottenuto a fine procedura. Infine nel campo *regression\_bias* è indicata l'intercetta della retta di regressione ottenuta a fine procedura. Come vedremo successivamente nei risultati di ogni singolo soggetto, tale analisi permette di ottenere grafici in cui è possibile estrapolare dalle risposte un trend, anche per soggetti i cui risultati non sembrano seguire alcuna tendenza.

Abbiamo anche testato i risultati di valori di elevazione completamente casuali inseriti da un generatore di numeri random distribuiti uniformemente nell'intervallo  $[-90^\circ, 90^\circ]$ . In questo caso l'errore medio di elevazione converge ad un valore di  $49^\circ$  e il quinto percentile ad un valore intorno ai  $40^\circ$ . Quest'ultimo valore è stato considerato, nelle nostre analisi, come soglia dell'errore medio di elevazione, per singola condizione, oltre la quale il soggetto risulta non attendibile, almeno per quella determinata condizione.

Ad ogni soggetto è associata la classifica delle condizioni sulla base della media degli angoli di errore medi ad esse relativo. Ovviamente, minore risulta tale valore e più in alto in classifica si posiziona la condizione. Le condizioni con *mean\_angle\_error*  $> 40^\circ$  vengono riportate in rosso poichè, come visto sopra, in tali condizioni il soggetto risulta inaffidabile e di conseguenza in esse riceve una penalizzazione.

Per ogni soggetto rappresentiamo 12 grafici, uno per ogni condizione, dove in ascissa troviamo l'elevazione reale e in ordinata l'elevazione percepita e, all'interno di tale grafico, le risposte date dal soggetto vengono raffigurate con quadratini di colore rosso. La retta di colore blu indica la retta ideale (bisettrice) e la retta di colore nero indica la retta di regressione che si ottiene alla fine della tecnica di analisi iterativa.

Condizione	mean_angle_error (°)
8	16.84
10	17.03
6	17.9
11	19.44
2	20.34
7	20.52
3	22.76
12	23.64
4	24.78
9	26.58
1	26.94
5	26.99

**Tabella 4.1:** *Classifica delle condizioni per il Soggetto 20*

## 4.2 Soggetti

### 4.2.1 Soggetto 20

Il Soggetto 20, come mostrato in Figura 4.1, ottiene buoni risultati in tutte le condizioni ad eccezione delle condizioni 1 e 5 in cui le medie degli angoli di errore per tutte le 7 elevazioni risultano le peggiori con valori rispettivamente di  $26.94^\circ$  e  $26.99^\circ$ , con punte di errore medio di elevazione di  $58.22^\circ$  e  $32.32^\circ$  rispettivamente per l'elevazione di  $-45^\circ$  e  $-15^\circ$  per la condizione 1. In tale condizione ottiene anche il peggior risultato per quanto riguarda il numero di *up/down confusion* con un valore pari ad 8. Le condizioni migliori risultano in ordine 8, 10, 6, 11 con valori della media degli angoli di errore per tutte le 7 elevazioni molto vicini tra loro, rispettivamente di  $16.8^\circ$ ,  $17^\circ$ ,  $17.9^\circ$  e  $19.4^\circ$ . Anche il numero di *up/down confusion* risulta piuttosto basso con valori che oscillano tra 1 e 3. Nella condizione 8 è possibile notare che il soggetto percepisce meglio le elevazioni di  $-45^\circ$ ,  $-30^\circ$ ,  $-15^\circ$  e  $45^\circ$  con angolo di errore medio rispettivamente di  $19.64^\circ$ ,  $7.36^\circ$ ,  $8.79^\circ$  e  $6.7^\circ$  e deviazione standard pari rispettivamente a  $16.17^\circ$ ,  $8.03^\circ$ ,  $13.38^\circ$  e  $7.84^\circ$ , mentre le elevazioni di  $0^\circ$ ,  $15^\circ$  e  $30^\circ$  presentano un angolo di errore medio pari rispettivamente a  $22.1^\circ$ ,  $31.43^\circ$  e  $21.84^\circ$  e deviazione standard di  $8.98^\circ$ ,  $42.1^\circ$  e  $27.11^\circ$ . In essa la pendenza della retta di regressione è pari a 0.82. Infine, notiamo che in 10 condizioni su 12 la retta di regressione si attesta su valori superiori a 0.6 e che il *front\_back\_reversal\_rate* è pari allo 0%, dal momento che tutti i suoni sono percepiti di fronte. Dai risultati si evince che, considerando le 2 migliori condizioni, il Soggetto riesce a localizzare meglio con 2 notch anziché 3, di profondità pari a  $-30$  dB e con larghezza di banda sia di  $0.15f_c$  che  $0.25f_c$ . Il dato che meglio differenzia le condizioni migliori rispetto a tutte le altre è il numero di notch presenti nella PRTE, essendo uguale come dato tra le condizioni migliori. Il Soggetto 20 risulta tra i migliori soggetti testati. In Tabella 4.1 è riportata la classifica delle condizioni, mentre in Tabella 4.2 sono riportati i dati statistici di interesse per le 12 condizioni.

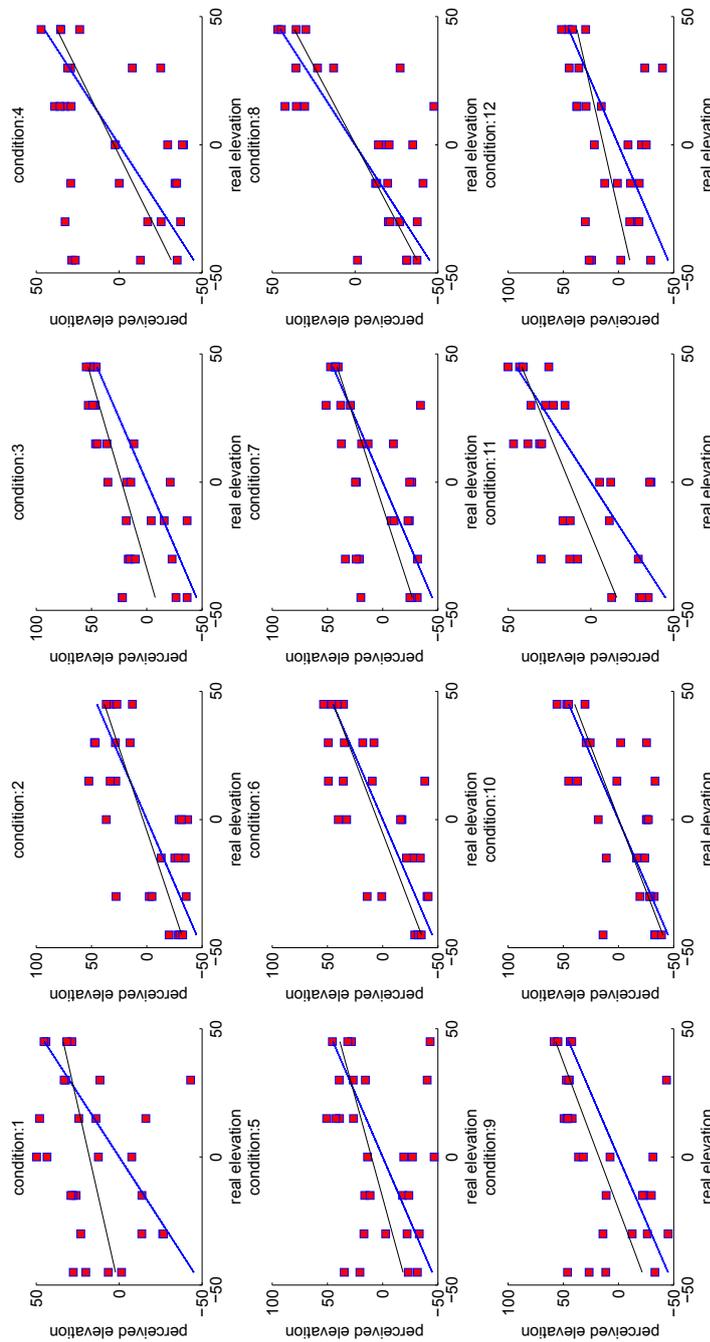


Figura 4.1: Scatterplot del Soggetto ID 20

Cond.	mean_angle_error (°)	reversal_rate (%)	up_down_rate (%)	reg._slope	R2	reg._bias	polar_error_rate (%)
1	26.9	0	28.5	0.35	0.21	17.9	3.5
2	20.3	0	3.5	0.77	0.6	3.3	3.57
3	22.7	0	17.8	0.67	0.63	22.7	3.5
4	24.7	0	17.8	0.76	0.59	3	14.2
5	26.9	0	21.4	0.63	0.47	10.1	14.2
6	17.9	0	3.5	0.88	0.66	4.6	3.5
7	20.5	0	17.8	0.75	0.59	7.2	10.1
8	16.8	0	7.1	0.82	0.7	-0.3	7.1
9	26.5	0	10.7	0.87	0.68	17.6	14.2
10	17	0	7.1	0.88	0.73	-0.3	10.7
11	19.4	0	10.7	0.63	0.61	12.9	7.1
12	23.6	0	17.8	0.53	0.39	13.6	7.1

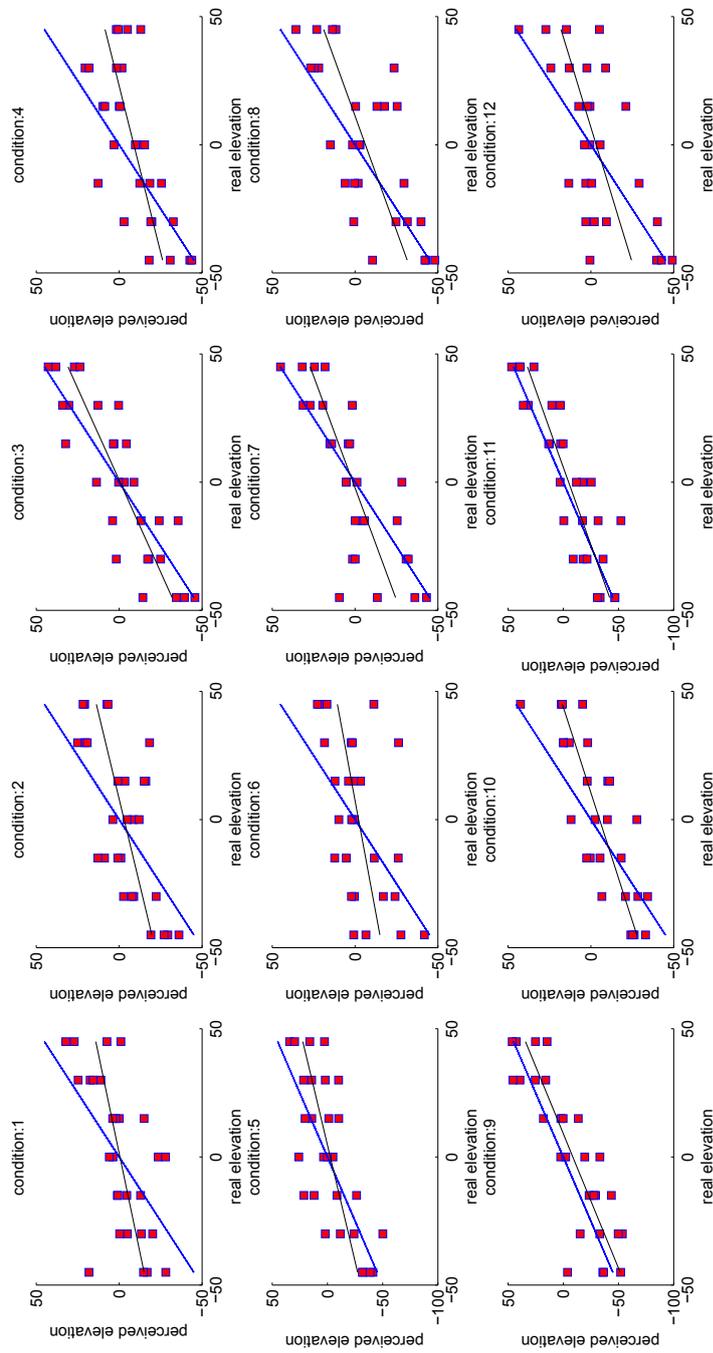
**Tabella 4.2:** *Dati statistici di interesse per il Soggetto 20*

Condizione	mean_angle_error (°)
11	11.55
3	12.32
7	13.35
9	14.45
8	16.39
5	16.63
10	16.7
12	17.76
4	17.86
2	19.58
1	19.98
6	20.07

**Tabella 4.3:** *Classifica delle condizioni per il Soggetto 21*

#### 4.2.2 Soggetto 21

Il Soggetto 21, come è mostrato in Figura 4.2, ottiene buoni risultati in tutte le condizioni; in particolare, le migliori risultano, in ordine decrescente, la 11, 3, 7 e 9 con medie di angoli di errore per tutte le 7 elevazioni al di sotto di  $15^\circ$ . Nella condizione migliore il soggetto percepisce meglio le elevazioni di  $-45^\circ$ ,  $15^\circ$  e  $45^\circ$  con media degli angoli di errore pari rispettivamente a  $7.32^\circ$ ,  $7.96^\circ$  e  $7.88^\circ$  e deviazione standard rispettivamente di  $8.52^\circ$ ,  $6.75^\circ$  e  $8.39^\circ$ , mentre le elevazioni di  $-30^\circ$ ,  $-15^\circ$ ,  $0^\circ$ ,  $30^\circ$  presentano medie pari rispettivamente a  $11.89^\circ$ ,  $17.67^\circ$ ,  $14.39^\circ$  e  $13.76^\circ$  e deviazione standard pari rispettivamente a  $11.27^\circ$ ,  $21.9^\circ$ ,  $11.87^\circ$  e  $16.29^\circ$ . Nelle condizioni migliori sono del tutto assenti gli errori di *up/down confusion*, assumendo comunque valori che oscillano tra 1 e 3 per tutte le altre condizioni. La condizione peggiore risulta la 6 con media dell'angolo di errore per tutte le 7 elevazioni pari a  $20.07^\circ$  e numero di *up/down confusion* pari a 1. In tale condizione le elevazioni di  $45^\circ$ ,  $30^\circ$  e  $-45^\circ$  presentano un alto valore di angolo di errore medio, rispettivamente di  $32.51^\circ$ ,  $30.96^\circ$  e  $26.15^\circ$ , con deviazione standard tra  $16^\circ$  e  $19^\circ$ . Da notare come il soggetto percepisca tutti i suoni come provenienti dal retro, presentando quindi un *front\_back\_reversal\_rate* pari al 100%. Le rette di regressione assumono valori che variano tra 0.28 e 0.95, quest'ultimo associato alla condizione migliore 11. Dai risultati si evince che il soggetto riesce a localizzare meglio con PRTF costituite da 3 notch anzichè 2, con profondità di -30 dB, ad eccezione della condizione 3 dove la profondità è -10, con larghezze di banda pari a 2000 Hz o anche  $0.25f_c$ . Il dato che differenzia meglio le condizioni migliori rispetto a tutte le altre è il numero di notch presenti nella PRTF, essendo uguale come dato tra le condizioni migliori. Il Soggetto 21 risulta tra i migliori soggetti testati. In Tabella 4.3 è riportata la classifica delle condizioni, mentre in Tabella 4.4 sono riportati i dati statistici di interesse per le 12 condizioni.



**Figura 4.2:** Scatterplot del Soggetto ID 21

Cond.	mean_angle_error (°)	reversal_rate (%)	up_down_rate (%)	reg._slope	R2	reg._bias	polar_error_rate (%)
1	19.98	100	7.14	0.32	0.36	-0.46	0
2	19.58	100	10.71	0.37	0.47	-2.99	0
3	12.32	100	0	0.69	0.74	-0.56	0
4	17.86	100	0	0.38	0.49	-8.86	0
5	16.64	100	3.57	0.55	0.54	-2.59	0
6	20.07	100	3.57	0.28	0.29	-2.17	0
7	13.35	100	0	0.57	0.61	1.29	0
8	16.39	100	10.71	0.56	0.53	-6.3	0
9	14.45	100	0	0.95	0.80	-8.72	3.57
10	16.7	100	0	0.5	0.65	-5.3	0
11	11.55	100	0	0.82	0.78	-4.85	0
12	17.76	100	3.57	0.47	0.44	-3.27	0

**Tabella 4.4:** Dati statistici di interesse per il Soggetto 21

Condizione	mean_angle_error (°)
7	15.93
11	16.63
10	16.71
5	17.47
12	17.86
4	18.88
8	19.29
9	19.32
1	19.86
3	20.41
2	21.16
6	23.41

**Tabella 4.5:** *Classifica delle condizioni per il Soggetto 22*

### 4.2.3 Soggetto 22

Il Soggetto 22, come mostrato in Figura 4.3, ottiene ottimi risultati in tutte le condizioni; le condizioni migliori, in ordine decrescente, sono 7, 11, 10, 5, con valore della media degli angoli di errore per tutte le 7 elevazioni compreso tra  $15.93^\circ$  e  $17.47^\circ$ . Nella condizione migliore, gli angoli di errore medio per le 7 elevazioni variano tra  $11.52^\circ$  e  $21.41^\circ$ , quest'ultimo riferito alla elevazione di  $-15^\circ$  dove la deviazione standard si attesta intorno a  $21.90^\circ$ , mentre il primo valore si riferisce alla elevazione meglio percepita che in questo caso è  $-30^\circ$  dove la deviazione standard si riduce al valore di  $4.8^\circ$ . Non notiamo invece differenze significative nella percezione delle elevazioni  $-45^\circ$ ,  $0^\circ$ ,  $15^\circ$ ,  $30^\circ$  e  $45^\circ$ . La pendenza della retta di regressione assume l'ottimo valore di 1.08. Anche la condizione 11 presenta ottimi valori con pendenza di regressione pari a 0.93, e medie degli angoli di elevazione tutte inferiori a  $20^\circ$ , ad eccezione dell'elevazione di  $15^\circ$  con media pari a  $28.26^\circ$ . La deviazione standard resta al di sotto di  $20^\circ$  per tutte le elevazioni tranne che per  $15^\circ$  dove assume il valore di  $26.97^\circ$ . Il numero di *polar errors* in tale condizione è pari a 0, laddove nella migliore condizione esso è pari a 2. La condizione peggiore risulta la 6 con media degli angoli di errore per tutte le 7 elevazioni pari a  $23.41^\circ$  e con le elevazioni del semipiano superiore che presentano errori compresi tra  $28.9^\circ$  e  $32.4^\circ$  e deviazione standard tra  $23.27^\circ$  e  $30.95^\circ$ . Tale soggetto presenta un numero di errori di *up/down confusion* compreso tra 0 e 3, con il valore 0 associato alla condizione migliore. Il numero di *front\_back* varia tra 10 e 20, con la condizione migliore che ne presenta 13. Le rette di regressione assumono in generale valori piuttosto alti con 9 condizioni su 12 che presentano valori al di sopra di 0.86. Dai risultati si evince che il soggetto percepisce meglio le elevazioni con PRTF costituite da 3 notch anziché 2, tranne nel caso della condizione 10 che ne contiene 2, con profondità di  $-30$  dB e con larghezze di banda sia di 2000 Hz che  $0.15f_c$ . Il dato che differenzia meglio le condizioni migliori rispetto a tutte le altre è la profondità dei notch, risultando uguale come dato tra le condizioni migliori. Il Soggetto 22 risulta tra i migliori soggetti testati. In Tabella 4.5 è riportata la classifica delle condizioni, mentre in Tabella 4.6 sono riportati i dati statistici di interesse per le 12 condizioni.

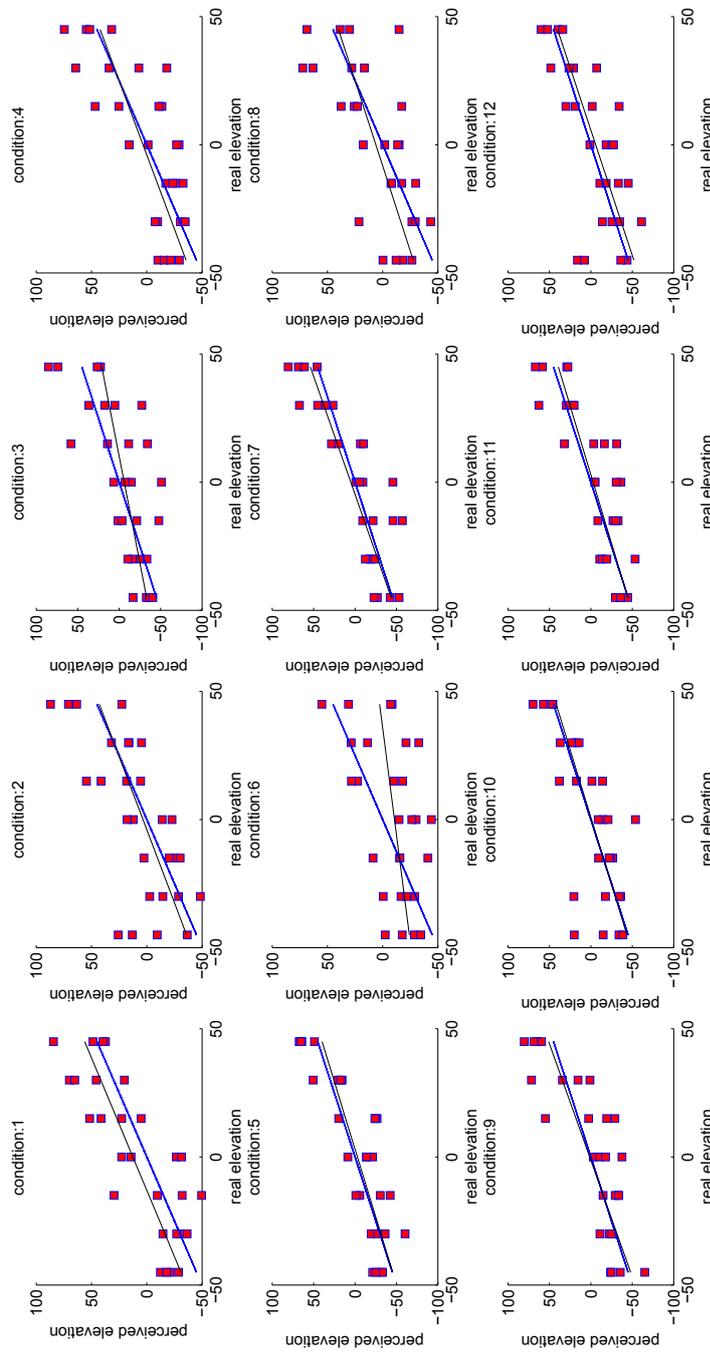


Figura 4.3: Scatterplot del Soggetto ID 22

Cond.	mean_angle_error (°)	reversal_rate (%)	up_down_rate (%)	reg._slope	R2	reg._bias	polar_error_rate (%)
1	19.96	35.71	3.57	0.96	0.70	12.59	3.57
2	21.16	35.71	3.57	0.87	0.62	3.43	7.14
3	20.41	53.57	7.14	0.6	0.58	-6.12	10.71
4	18.88	71.42	3.57	0.86	0.62	3.39	3.57
5	17.47	50	10.71	0.94	0.66	-2.87	0
6	23.41	39.28	10.71	0.29	0.17	-10.79	3.57
7	15.93	46.42	0	1.08	0.81	5.14	7.14
8	19.29	53.57	7.14	0.75	0.57	6.12	3.57
9	19.32	57.14	7.14	1.09	0.7	1.37	3.57
10	16.71	53.57	7.14	0.94	0.79	-0.84	10.71
11	16.63	64.28	7.14	0.93	0.71	-2.78	0
12	17.86	57.14	7.14	1.02	0.77	-5.98	7.14

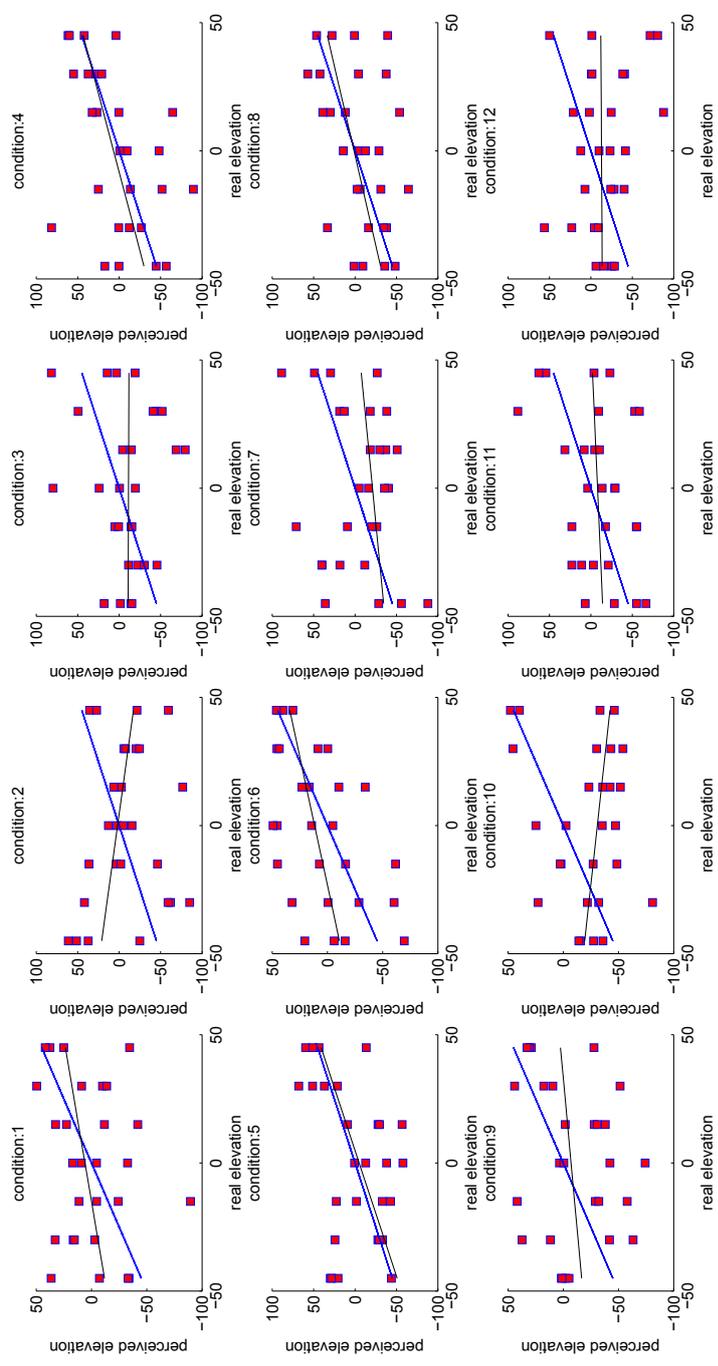
**Tabella 4.6:** *Dati statistici di interesse per il Soggetto 22*

Condizione	mean_angle_error (°)
6	25.39
8	26.08
4	26.60
5	28.63
1	30.38
11	32.42
9	35.12
10	36.24
3	36.33
7	37.77
12	39.40
2	41.01

**Tabella 4.7:** *Classifica delle condizioni per il Soggetto 24*

#### 4.2.4 Soggetto 24

Il Soggetto 24, come mostrato in Figura 4.4, ottiene buoni risultati solo in 5 condizioni; le 4 migliori, in ordine decrescente, sono la 6, 8, 4, 5, con media degli angoli di errore per tutte le 7 elevazioni compresa tra 25° e 29°. La migliore risulta la condizione 6, con le elevazioni del semipiano superiore percepite meglio di quelle del semipiano inferiore. Infatti, per le elevazioni tra 0° e 45° l'angolo di errore medio varia tra 5° e 28°, rispettivamente per l'elevazione di 0° e 45°, e con deviazione standard compresa tra 6° e 26°, rispettivamente per l'elevazione di 45° e 0°. Al contrario, alle elevazioni tra -45° e -15° è associato un angolo di errore medio che varia tra 32° e 39.5°, rispettivamente per l'elevazione di -15° e -45°, con deviazione standard che varia tra 37° e 44.7°, rispettivamente per l'elevazione di -45° e -15°. In tale condizione il numero di *front\_back\_reversal* è pari a 12, il numero di *up/down confusion* e di *polar errors* è pari a 4 e la pendenza della retta di regressione è pari a 0.49. Da notare anche il buon risultato trovato con la condizione 8, in cui la media degli angoli di errore per tutte le 7 elevazioni risulta pari a 26.08°, il numero di *front\_back\_reversal* è 11, il numero di *up/down confusion* è 4, il numero di *polar errors* è pari a 5 ed infine la pendenza della retta di regressione è pari a 0.71. Per le condizioni 5 e 8 la tecnica di analisi iterativa ha prodotto importanti miglioramenti sulla pendenza della retta di regressione, pur causando un *polar\_error\_rate* pari al 25%. La condizione peggiore risulta la 2 in cui la media degli angoli di errore per tutte le 7 elevazioni supera il valore della soglia fissato a 40° e quindi in tale condizione il soggetto risulta inattendibile. Notiamo anche la presenza di condizioni con rette di regressione negative, indicando quest'ultimo dato che il soggetto tende ad invertire, nella localizzazione, il semipiano superiore con quello inferiore e viceversa. Dai risultati si evince che il soggetto percepisce meglio con PRTF costituite da soli 2 notch, con profondità indifferentemente di -10 dB e -30 dB e con larghezza di banda sia di 2000 Hz che di  $0.15f_c$ . Quindi è il numero di notch presenti il dato più caratteristico che differenzia le condizioni migliori rispetto a tutte le altre. In Tabella 4.7 è riportata la classifica delle condizioni, mentre in Tabella 4.8 sono riportati i dati statistici di interesse per le 12 condizioni.



**Figura 4.4:** Scatterplot del Soggetto ID 24

Cond.	mean_angle_error (°)	reversal_rate (%)	up_down_rate (%)	reg._slope	R2	reg._bias	polar_error_rate (%)
1	30.38	21.42	21.42	0.39	0.23	5.91	14.28
2	41.01	25	35.71	-0.42	0.29	1.89	25
3	36.33	25	25	-0.01	0	-11.35	17.85
4	26.60	25	14.28	0.81	0.65	6.73	21.42
5	28.63	35.71	28.57	1.02	0.64	-5.01	25
6	25.39	42.85	14.28	0.49	0.43	11.57	14.28
7	37.77	35.71	42.85	0.29	0.12	-21.03	28.57
8	26.08	39.28	14.28	0.71	0.54	1.26	17.85
9	35.12	28.57	25	0.21	0.08	-7.11	21.42
10	36.24	53.57	35.71	-0.25	0.2	-30.88	21.42
11	32.42	39.28	17.85	0.13	0.03	-8.16	21.42
12	39.40	35.71	28.57	0.01	0	-12.89	17.85

**Tabella 4.8:** Dati statistici di interesse per il Soggetto 24

Condizione	mean_angle_error (°)
5	32.15
9	36.3
8	38.42
1	39.59
2	41.32
3	46.63
4	42.09
6	40.8
7	45.52
10	48.62
11	45.81
12	53.17

**Tabella 4.9:** *Classifica delle condizioni per il Soggetto 26*

#### 4.2.5 Soggetto 26

Il Soggetto 26, come mostrato in Figura 4.5, ottiene risultati scadenti su tutte le condizioni, ad eccezione della 5. In quest'ultima la media degli angoli di errore per tutte le 7 elevazioni si attesta sul valore di 32.15°, con le elevazioni del semipiano superiore percepite meglio rispetto a quelle inferiori. Infatti, per le elevazioni di 15°, 30° e 45°, l'angolo di errore medio varia tra 17° e 23.4°, il primo associato all'elevazione di 15° ed il secondo a quella di 45° e la deviazione standard varia tra 25° e 28°, associata all'elevazione di 15° e 45° rispettivamente, mentre per le elevazioni di 0°, -15°, -30°, -45°, l'angolo di errore medio varia tra 34° e 49.36°, il primo riferito all'elevazione di -30° ed il secondo a quella di -45° e la deviazione standard varia tra 25.57° e 36.84°, sempre riferite a -30° e a -45°. In tale condizione il numero di *front\_back\_reversal* è pari a 13, il numero di *up/down confusion* è pari a 5, il numero di *polar errors* è pari a 4 ed infine la pendenza della retta di regressione è pari a 0.62. In tutte le altre condizioni i risultati ottenuti non mostrano alcuna tendenza ed inoltre ad 8 condizioni su 12 viene assegnato un valore pari ad *inf* poichè in esse la media degli angoli di errore per tutte le 7 elevazioni supera la soglia di 40°, risultando quindi il soggetto, in tali condizioni, inaffidabile. Dall'unica condizione significativa ricaviamo che il soggetto percepisce meglio le elevazioni con PRTF costituite da 3 notch, con profondità pari a -10 dB e larghezza di banda pari a 2000 Hz. Il soggetto, come vedremo successivamente, non rientra tra quelli selezionati per il calcolo della condizione globale migliore. In Tabella 4.9 è riportata la classifica delle condizioni, mentre in Tabella 4.10 sono riportati i dati statistici di interesse per le 12 condizioni.

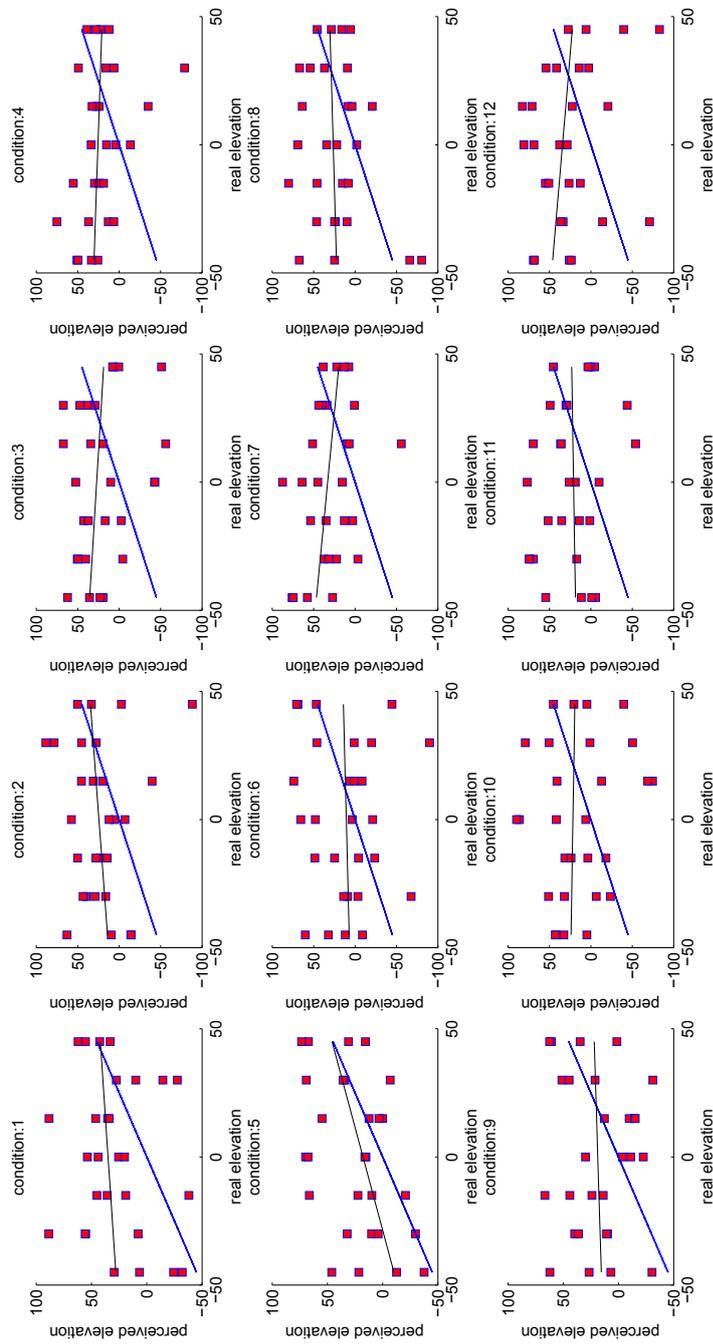


Figura 4.5: Scatterplot del Soggetto ID 26

Cond.	mean_angle_error (°)	reversal_rate (%)	up_down_rate (%)	reg._slope	R2	reg._bias	polar_error_rate (%)
1	39.59	64.28	28.57	0.15	0.07	35.01	25
2	41.32	46.42	35.71	0.22	0.1	24.35	17.85
3	46.63	57.14	42.85	-0.18	0.09	27.12	17.85
4	42.09	46.42	42.85	-0.1	0.03	25.56	10.71
5	32.15	46.42	17.85	0.62	0.44	17.75	14.28
6	40.8	39.28	25	0.07	0	10.51	28.57
7	45.52	39.28	35.71	-0.29	0.18	33.03	10.71
8	38.42	39.28	32.14	0.08	0.01	26.34	14.28
9	36.30	32.14	32.14	0.07	0	18.69	14.28
10	48.62	39.28	39.28	-0.05	0	21.56	28.57
11	45.81	67.85	32.14	0.04	0	20.8	25
12	53.17	42.85	42.85	-0.26	0.16	34.19	25

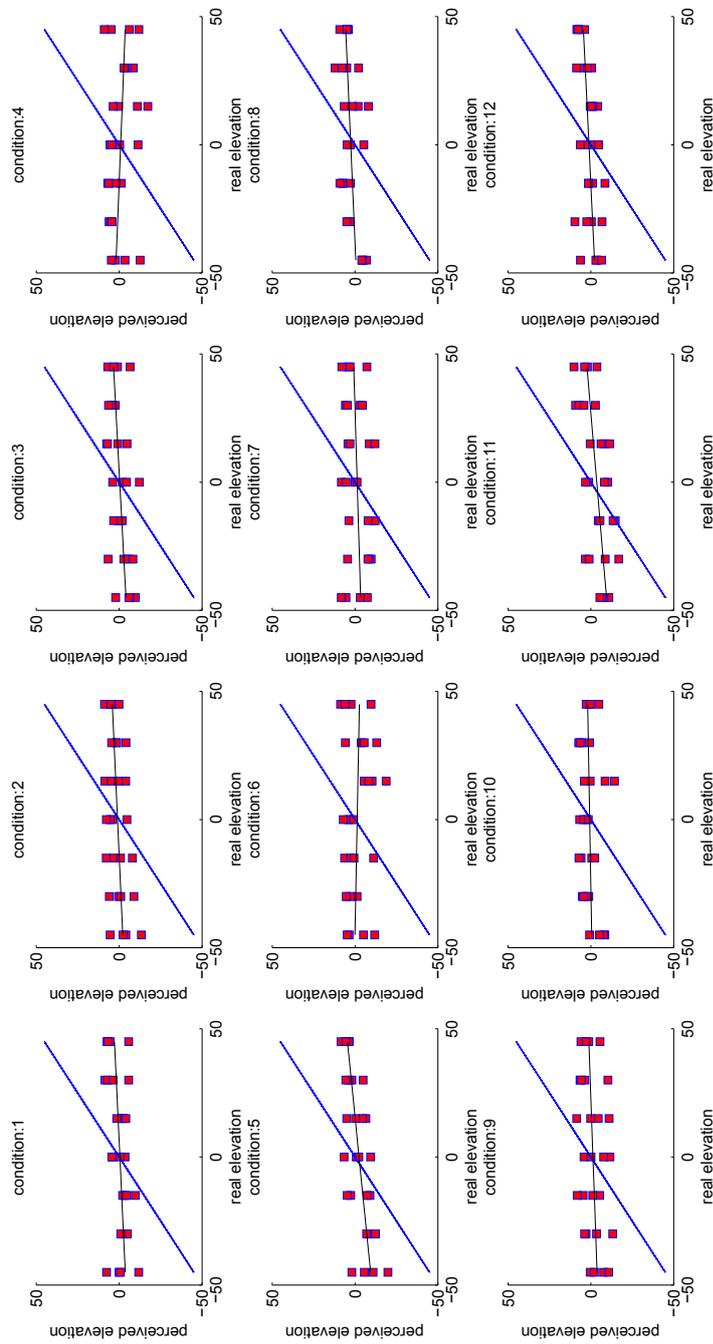
**Tabella 4.10:** Dati statistici di interesse per il Soggetto 26

Condizione	mean_angle_error (°)
5	22.69
11	23.14
1	24.09
3	24.28
12	24.69
7	24.69
2	24.82
8	25.54
9	25.57
10	26.3
6	28.13
4	29.09

**Tabella 4.11:** *Classifica delle condizioni per il Soggetto 27*

#### 4.2.6 Soggetto 27

Il Soggetto 27, come mostrato in Figura 4.6, ottiene risultati discreti in quasi tutte le condizioni. Per tale soggetto è necessario sottolineare, come da lui stesso indicato, la difficoltà riscontrata con il sistema di puntamento da noi utilizzato; in particolare, il soggetto ha avuto grosse difficoltà nel proiettare l'elevazione percepita su di un grafico basato sull'elevazione espressa in gradi. Infatti, le rette di regressione relative alle 12 condizioni sono tutte comprese tra -0.06 e 0.15 poichè le risposte date dal soggetto si posizionano quasi tutte tra -15° e 15°, pur differenziandosi all'interno di tale range in base alla percezione di elevazione del soggetto. Le condizioni migliori risultano la 5 e la 11. Nella condizione 5, la media degli angoli di errore è pari a 22.69°, con angoli di errore medio per le elevazioni di -15°, 0° e 15° tra 4.79° e 16.6°, il primo riferito all'elevazione di 0° ed il secondo a quella di 15° e deviazione standard tra 5.06° e 6.92°, rispettivamente per l'elevazione di 15° e -15° ed angoli di errore medio per le elevazioni di -45°, -30°, 30° e 45° tra 20.16° e 39.73°, il primo riferito all'elevazione di -30° ed il secondo a quella di 45° e deviazione standard tra 2.31° e 9.17°, rispettivamente per l'elevazione di 45° e -45°. Per la condizione 11, la media degli angoli di errore è pari a 23.14°, con angoli di errore medio per le elevazioni di -15°, 0° e 15° tra 5.43° e 21.41° e deviazione standard tra 4.99° e 6.7°, ed angoli di errore medio per le elevazioni di -45°, -30°, 30° e 45° tra 24.66° e 41.92° e deviazione standard tra 2.33° e 9.29°. In tutte le condizioni sono assenti errori di *front.back.reversal* e *polar errors*, mentre gli errori di *up/down confusion* sono presenti in numero pari ad 1 solo per la condizione 4 e 6. La pendenza della retta di regressione per le condizioni migliori 5 e 11 assume valori pari rispettivamente a 0.15 e 0.13. Dai risultati ottenuti si evince che il soggetto percepisce meglio l'elevazione con PRTF costituite da 3 notch anzichè 2, con profondità indifferentemente di -10 dB e -30 dB e con larghezza di banda pari a 2000 Hz. Quindi è il numero di notch presenti insieme con la larghezza di banda il dato più caratteristico che differenzia le condizioni migliori rispetto a tutte le altre. In Tabella 4.11 è riportata la classifica delle condizioni, mentre in Tabella 4.12 sono riportati i dati statistici di interesse per le 12 condizioni.



**Figura 4.6:** Scatterplot del Soggetto ID 27

Cond.	mean_angle_error (°)	reversal_rate (%)	up_down_rate (%)	reg._slope	R2	reg._bias	polar_error_rate (%)
1	24.09	0	0	0.07	0.16	-0.44	0
2	24.82	0	0	0.07	0.14	1	0
3	24.28	0	0	0.08	0.2	-0.35	0
4	29.09	0	3.57	-0.06	0.06	-0.92	0
5	22.69	0	0	0.15	0.43	-2.42	0
6	28.13	0	3.57	-0.02	0.01	-1.3	0
7	24.69	0	0	0.04	0.04	-1.3	0
8	25.54	0	0	0.06	0.15	2.56	0
9	25.57	0	0	0.05	0.06	-1.27	0
10	26.3	0	0	0.02	0.02	0.65	0
11	23.14	0	0	0.13	0.3	-3.58	0
12	24.69	0	0	0.07	0.19	1.13	0

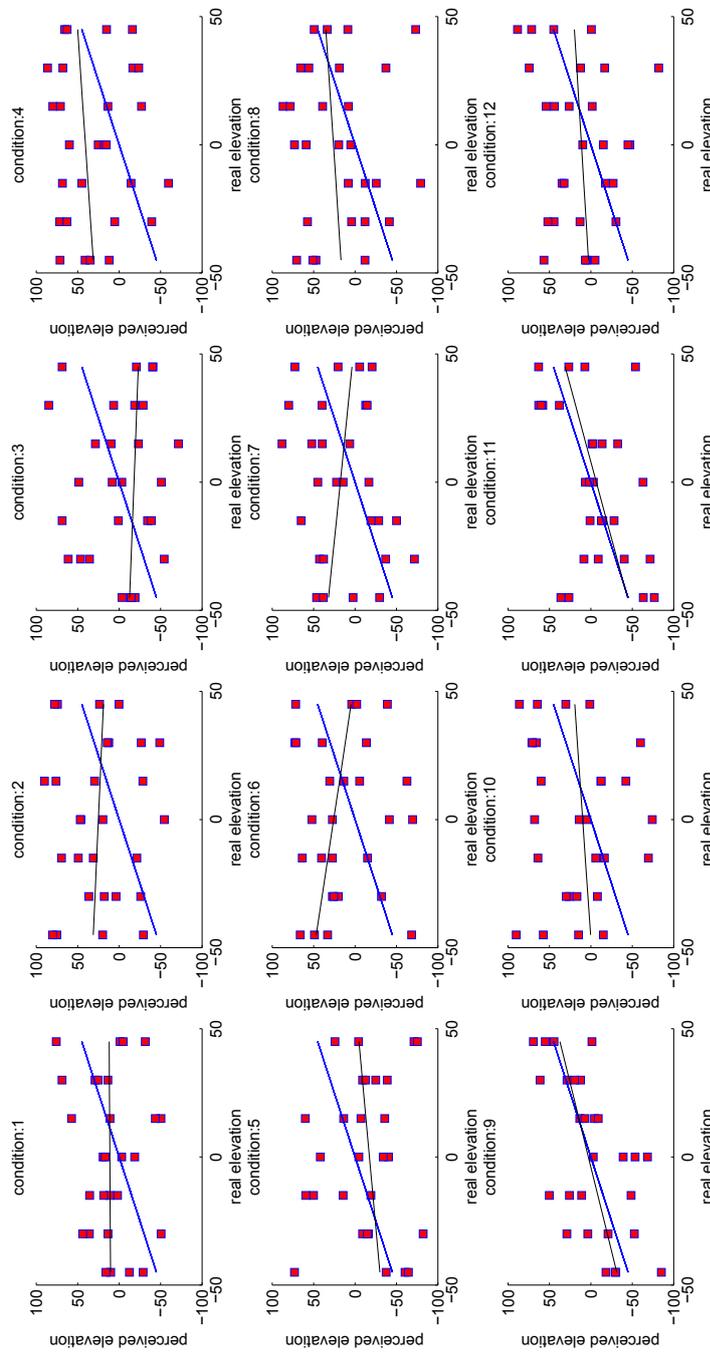
**Tabella 4.12:** Dati statistici di interesse per il Soggetto 27

Condizione	mean_angle_error (°)
9	26.62
11	28.64
1	35.31
12	39.41
7	39.89
2	47.79
3	43.96
4	49.04
5	42.44
6	46.06
8	43.61
10	46.4

**Tabella 4.13:** *Classifica delle condizioni per il Soggetto 28*

#### 4.2.7 Soggetto 28

Il Soggetto 28, come mostrato in Figura 4.7, ottiene buoni risultati in sole 2 condizioni, pur ottenendo una media degli angoli di errore per tutte le 7 elevazioni inferiore alla soglia di 40° in 5 condizioni su 12. Le condizioni migliori risultano la 9 e la 11, con media degli angoli di errore rispettivamente pari a 26.62° e 28.64°. Per ciò che riguarda la condizione 9, notiamo che il soggetto percepisce meglio le elevazioni superiori rispetto a quelle inferiori; in particolare, per le elevazioni di 15°, 30° e 45°, l'angolo di errore medio varia tra 13.26° e 20.32°, il primo riferito all'elevazione di 15° ed il secondo a quella di 45°, con deviazione standard che varia tra 10.25° e 30.6°, rispettivamente per le elevazioni di 15° e 45° mentre per le elevazioni di -45°, -30°, -15° e 0°, l'angolo di errore medio varia tra 24.21° e 41.4°, il primo riferito all'elevazione di -45° ed il secondo a quella di -15° con deviazione standard tra 28° e 42°, rispettivamente per le elevazioni di 0° e -15°. Il numero di *front\_back\_reversal* è pari a 19, quello di *up/down confusion* è pari a 3, il numero di *polar errors* è pari a 5 e la pendenza della retta di regressione è pari a 0.76. Nella condizione 11 notiamo che il soggetto percepisce meglio le elevazioni tra -30° e 30° con l'angolo di errore medio che varia tra 7.93° e 27.82° e la deviazione standard che varia tra 11.29° e 35.29°, mentre per le elevazioni di -45° e 45° l'angolo di errore medio varia tra 43.3° e 50.6° e la deviazione standard tra 48.91° e 58.89°. In tale condizione il numero di *front\_back\_reversal* è pari a 13, il numero di *up/down confusion* e di *polar errors* è pari a 4 ed infine la pendenza della retta di regressione è pari a 0.83. In tutte le altre condizioni non è riscontrabile alcuna tendenza ed inoltre in 4 condizioni su 12 la retta di regressione presenta pendenza negativa, indicando la confusione, nella localizzazione, tra semipiano superiore e semipiano inferiore. Dai dati si evince che il soggetto percepisce meglio con PRTF costituite da 3 notch, con profondità di -30 dB e larghezza di banda indifferentemente di 2000 Hz e  $0.25f_c$ . Quindi è il numero di notch presenti insieme con la profondità il dato più caratteristico che differenzia le condizioni migliori rispetto a tutte le altre. In Tabella 4.13 è riportata la classifica delle condizioni, mentre in Tabella 4.14 sono riportati i dati statistici di interesse per le 12 condizioni.



**Figura 4.7:** Scatterplot del Soggetto ID 28

Cond.	mean_angle_error (°)	reversal_rate (%)	up_down_rate (%)	reg._slope	R2	reg._bias	polar_error_rate (%)
1	35.31	42.85	28.57	0.01	0	11.19	21.42
2	47.79	42.85	39.28	-0.13	0.06	25.06	42.85
3	43.96	57.14	39.28	-0.11	0.04	-17.92	32.14
4	49.04	39.28	39.28	0.2	0.07	40.61	25
5	42.44	57.14	28.57	0.28	0.14	-17.37	28.57
6	46.06	42.85	39.28	-0.46	0.39	25.74	32.14
7	39.89	50	21.42	-0.31	0.14	17.74	28.57
8	43.61	50	21.42	0.2	0.05	26	32.14
9	26.62	67.85	10.71	0.76	0.54	2.93	17.85
10	46.4	42.85	28.57	0.21	0.09	9.85	50
11	28.64	46.42	14.28	0.83	0.56	-6.66	14.28
12	39.41	42.85	25	0.18	0.04	11.31	28.57

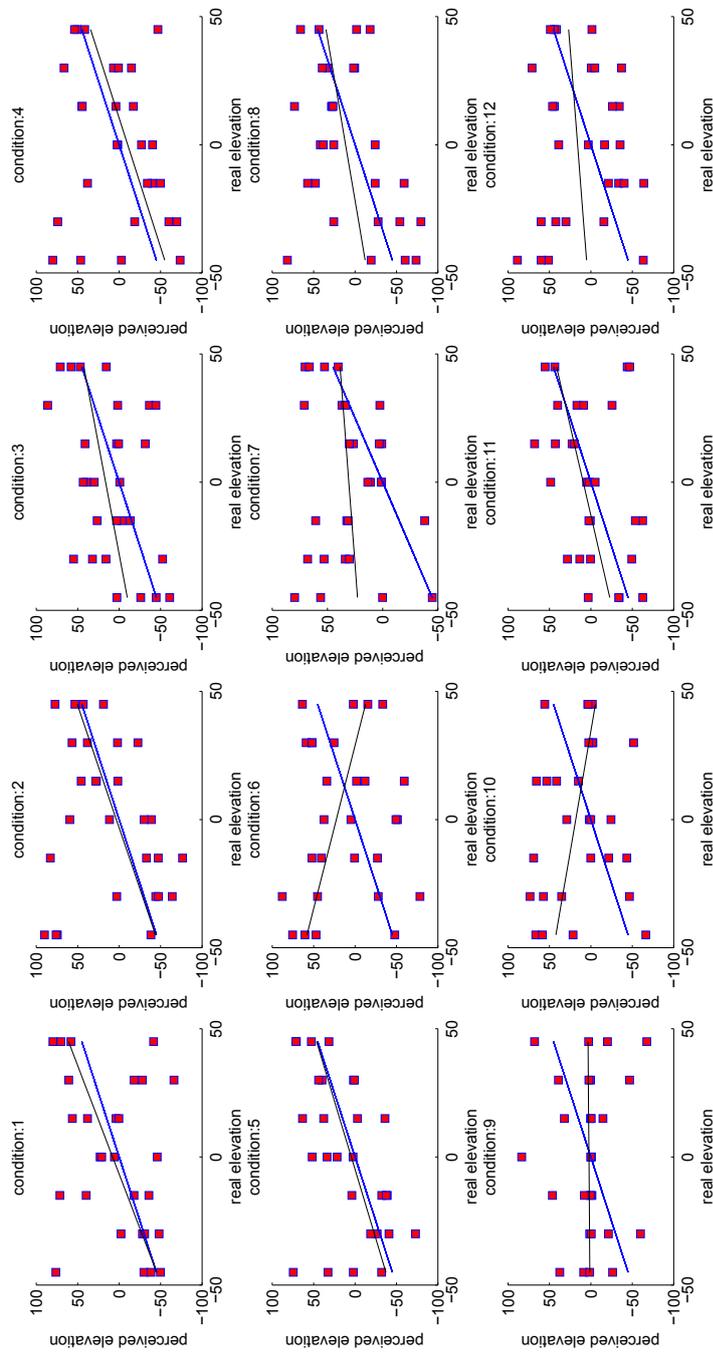
**Tabella 4.14:** *Dati statistici di interesse per il Soggetto 28*

Condizione	mean_angle_error (°)
11	28.75
5	29.01
3	31.41
1	33.58
8	33.73
9	34.33
7	35.19
4	36.58
2	38.7
6	45.45
10	42.65
12	41.46

**Tabella 4.15:** Classifica delle condizioni per il Soggetto 30

#### 4.2.8 Soggetto 30

Il Soggetto 30, come mostrato in Figura 4.8, ottiene ottimi risultati in 7 condizioni su 12, superando la soglia di  $40^\circ$  in sole 3 condizioni. Le condizioni migliori risultano, in ordine decrescente, la 11, 5, 3, con media degli angoli di errore per tutte le 7 elevazioni rispettivamente di  $28.75^\circ$ ,  $29.01^\circ$ ,  $31.41^\circ$ . Nella condizione migliore il soggetto percepisce meglio l'elevazione di  $-45^\circ$  e quelle tra  $-15^\circ$  e  $30^\circ$ , con angolo di errore medio che varia tra  $15.2^\circ$  e  $29.79^\circ$ , riferiti alla elevazione di  $0^\circ$  e  $-15^\circ$  e deviazione standard tra  $22.28^\circ$  e  $34.6^\circ$ , rispettivamente per le elevazioni di  $15^\circ$  e  $-15^\circ$ , mentre per le elevazioni di  $-30^\circ$  e  $45^\circ$  l'angolo di errore medio assume rispettivamente il valore di  $37.84^\circ$  e  $48.18^\circ$  e la deviazione standard  $33.7^\circ$  e  $54.92^\circ$ . In tale condizione il numero di *front\_back\_reversal* è pari a 2, il numero di *up/down confusion* è pari a 4, il numero di *polar errors* è pari a 6 ed infine la retta di regressione presenta un valore pari a 0.71. Nella condizione 5 il soggetto percepisce meglio le elevazioni tra  $-30^\circ$  e  $0^\circ$  e tra  $30^\circ$  e  $45^\circ$  con angolo di errore medio che varia tra  $17.06^\circ$  e  $27.33^\circ$  e deviazione standard compresa tra  $18.9^\circ$  e  $23.85^\circ$ . Invece, per le elevazioni di  $-45^\circ$  e  $15^\circ$ , l'angolo di errore medio assume rispettivamente valore pari a  $64.29^\circ$  e  $23.51^\circ$  con deviazione standard pari a  $45.36^\circ$  e  $43.98^\circ$ . In tale condizione il numero di *front\_back\_reversal* è pari a 0, quello di *up/down confusion* è pari a 3, il numero di *polar errors* è pari a 5 e la retta di regressione assume il valore di 0.93. Da notare che in 2 delle 3 condizioni in cui viene superata la soglia di  $40^\circ$ , precisamente le condizioni 6 e 10, il soggetto tende ad invertire il semipiano superiore con quello inferiore e viceversa. L'applicazione della tecnica di analisi iterativa ha prodotto importanti miglioramenti sulla pendenza della retta di regressione per le condizioni 1, 2, 4, con *polar\_error\_rate* pari rispettivamente a 25.57% e 25%. Dai risultati si evince che il soggetto percepisce meglio l'elevazione con PRTF costituite da 3 notch, con profondità sia di  $-10$  dB che  $-30$  dB e con larghezza di banda pari a 2000 Hz. Quindi è il numero di notch presenti insieme con la larghezza di banda il dato più caratteristico che differenzia le condizioni migliori rispetto a tutte le altre. In Tabella 4.15 è riportata la classifica delle condizioni, mentre in Tabella 4.16 sono riportati i dati statistici di interesse per le 12 condizioni.



**Figura 4.8:** Scatterplot del Soggetto ID 30

Cond.	mean_angle_error (°)	reversal_rate (%)	up_down_rate (%)	reg._slope	R2	reg._bias	polar_error_rate (%)
1	33.58	3.57	25	1.18	0.83	7.85	28.57
2	38.7	3.57	17.85	1.04	0.7	3.31	25
3	31.41	10.71	25	0.58	0.38	16.32	25
4	36.48	0	21.42	0.99	0.55	-10.29	25
5	29.01	0	10.71	0.93	0.63	4.52	17.85
6	45.45	10.71	35.71	-0.78	0.65	22.64	39.28
7	35.19	0	32.14	0.17	0.04	30.59	10.71
8	33.73	7.14	17.85	0.52	0.26	11.37	32.14
9	34.33	3.57	17.85	0.02	0	2.24	17.85
10	42.65	7.14	28.57	-0.52	0.43	18.25	25
11	28.75	7.14	14.28	0.71	0.63	9.2	21.42
12	41.46	0	32.14	0.23	0.06	15.93	46.42

**Tabella 4.16:** Dati statistici di interesse per il Soggetto 30

Condizione	mean_angle_error (°)
10	33.43
8	36.28
5	39.12
1	40.58
2	47.5
3	44.54
4	44.04
6	42.96
7	47.5
9	47.34
11	52.6
12	45.17

**Tabella 4.17:** *Classifica delle condizioni per il Soggetto 32*

#### 4.2.9 Soggetto 32

Il Soggetto 32, come mostrato in Figura 4.9, ottiene scarsi risultati in tutte le condizioni, ad eccezione della condizione 10, risultando inaffidabile in 9 condizioni su 12 poiché in esse la media degli angoli di errore per tutte le 7 elevazioni supera la soglia di 40°. Nella condizione 10, con media associata pari a 33.43°, il soggetto percepisce meglio le elevazioni tra -15° e 15° con angolo di errore medio che varia tra 16.41° e 27.38°, associati rispettivamente all'elevazione di -15° e 0°, con deviazione standard che varia tra 19.28° e 32.65° riferiti alle elevazioni di 0° e 15°. Le elevazioni di -45°, -30°, 30° e 45° presentano un angolo di errore medio che varia tra 36.99° e 49.13°, il primo associato all'elevazione di -30° ed il secondo all'elevazione di -45°, con deviazione standard che varia tra 28.28° e 52.07°, associati rispettivamente all'elevazione di -45° e 30°. In tale condizione il numero di *front\_back\_reversal* è pari a 5, quello di *up/down\_confusion* a 8, il numero di *polar\_errors* è pari a 5 ed infine la retta di regressione assume il valore di 0.38. In tutte le altre condizioni non è possibile estrapolare alcuna tendenza dai dati. Dai risultati si evince che il soggetto percepisce meglio le elevazioni con PRTF costituite da 2 notch, con profondità di -30 dB e larghezza di banda pari a  $0.25f_c$ . Il soggetto, come vedremo successivamente, non rientra tra quelli selezionati per il calcolo della condizione globale migliore. In Tabella 4.17 è riportata la classifica delle condizioni, mentre in Tabella 4.18 sono riportati i dati statistici di interesse per le 12 condizioni.

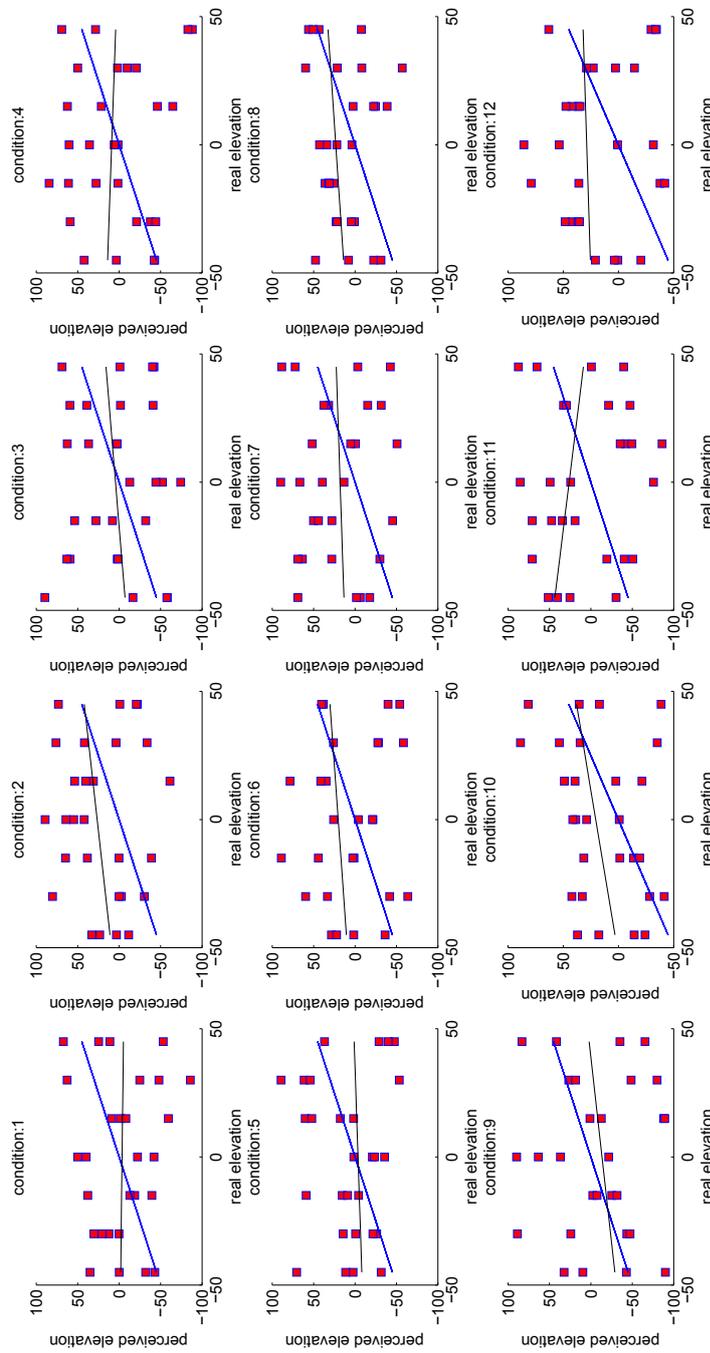


Figura 4.9: Scatterplot del Soggetto ID 32

Cond.	mean_angle_error (°)	reversal_rate (%)	up_down_rate (%)	reg._slope	R2	reg._bias	polar_error_rate (%)
1	40.58	21.42	32.14	-0.03	0	-3.52	21.42
2	47.5	21.42	32.14	0.34	0.19	26.47	28.57
3	44.54	28.57	28.57	0.25	0.11	4.37	53.57
4	44.04	21.42	35.71	-0.1	0.02	-0.1	50
5	39.12	28.57	25	0.1	0.01	-3.62	32.14
6	42.96	17.85	39.28	0.22	0.08	20.28	39.28
7	47.5	28.57	39.28	0.1	0.01	17.98	39.28
8	36.28	10.71	39.28	0.2	0.08	23.2	17.85
9	47.34	50	32.14	0.34	0.17	-13.47	46.42
10	33.43	17.85	28.57	0.38	0.18	20.53	17.85
11	52.6	28.57	53.57	-0.38	0.33	26.16	50
12	45.17	14.28	35.71	0.07	0.01	28.67	35.71

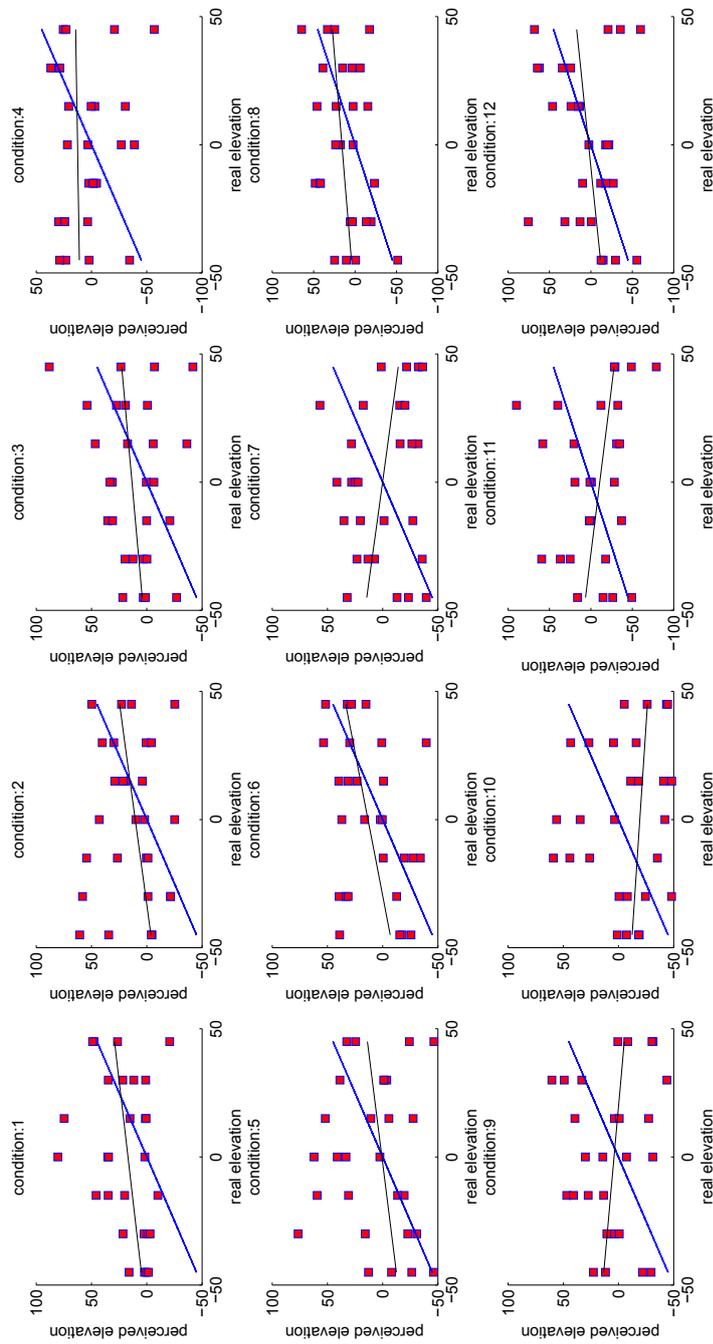
**Tabella 4.18:** Dati statistici di interesse per il Soggetto 32

Condizione	mean_angle_error (°)
6	26.09
8	28.67
12	29.49
4	30.46
2	30.67
1	31.4
3	32.14
5	33.65
7	37.16
9	37.26
10	39.56
11	40.76

**Tabella 4.19:** *Classifica delle condizioni per il Soggetto 33*

#### 4.2.10 Soggetto 33

Il Soggetto 33, come mostrato in Figura 4.10, ottiene discreti risultati in poche condizioni, risultando inaffidabile nella condizione 11 in cui la media degli angoli di errore per tutte le 7 elevazioni supera la soglia di  $40^\circ$ . Le condizioni migliori risultano, in ordine decrescente, la 6, 8 e 12 con media pari rispettivamente a  $26.09^\circ$ ,  $28.67^\circ$  e  $29.49^\circ$ . Nella condizione migliore il soggetto percepisce meglio le elevazioni tra  $-15^\circ$  e  $45^\circ$ , con angolo di errore medio che assume valori tra  $12.58^\circ$  e  $30.57^\circ$ , dove il primo è associato all'elevazione di  $-15^\circ$  e il secondo all'elevazione di  $30^\circ$ , e deviazione standard con valori tra  $14.49^\circ$  e  $40.02^\circ$ , riferiti rispettivamente alle elevazioni di  $-15^\circ$  e  $30^\circ$ . Per le elevazioni di  $-45^\circ$  e  $-30^\circ$  l'angolo di errore assume rispettivamente il valore di  $40.04^\circ$  e  $52.7^\circ$  con deviazione standard pari a  $29.58^\circ$  e  $23.86^\circ$ . In tale condizione il numero di *front\_back\_reversal* è pari a 25, il numero di *up/down confusion* è pari a 5, il numero di *polar errors* a 2 ed infine la retta di regressione assume il valore di 0.44. Da notare l'alto numero di *front\_back\_reversal* per tutte le condizioni, ad indicare che il soggetto percepisce la maggior parte dei suoni come provenienti da dietro, e rette di regressione con pendenze negative per 4 condizioni, ad indicare che in esse il soggetto inverte il semipiano superiore con quello inferiore e viceversa. Le restanti condizioni non presentano dati da cui sia possibile estrapolare una chiara tendenza. Dai risultati si evince che il soggetto percepisce meglio con PRTF costituite da 2 notch, con profondità di -10 dB e larghezza di banda pari a 2000 Hz. In Tabella 4.19 è riportata la classifica delle condizioni, mentre in Tabella 4.20 sono riportati i dati statistici di interesse per le 12 condizioni.



**Figura 4.10:** Scatterplot del Soggetto ID 33

Cond.	mean_angle_error (°)	reversal_rate (%)	up_down_rate (%)	reg._slope	R2	reg._bias	polar_error_rate (%)
1	31.4	78.57	21.42	0.26	0.21	17.01	10.71
2	30.67	85.71	21.42	0.31	0.19	10.4	14.28
3	32.14	64.28	21.42	0.2	0.08	13.4	10.71
4	30.46	85.71	28.57	0.03	0	12.64	10.71
5	33.65	82.14	25	0.29	0.13	0.69	17.85
6	26.09	89.28	17.85	0.44	0.3	13.25	7.14
7	37.16	85.71	42.85	-0.13	0.16	0.09	10.71
8	28.67	82.14	21.42	0.25	0.1	15.81	3.57
9	37.26	75	28.57	-0.2	0.07	3.96	7.14
10	39.56	85.71	35.71	-0.15	0.06	-19.39	21.42
11	40.76	75	39.28	-0.38	0.22	-10.83	21.42
12	29.49	75	17.85	0.31	0.13	2.69	21.42

**Tabella 4.20:** Dati statistici di interesse per il Soggetto 33

Condizione	mean_angle_error (°)
1	35.18
7	36.3
4	38.37
2	42.41
3	47.58
5	47.24
6	40.41
8	40.75
9	43.43
10	45.7
11	45.55
12	46.65

**Tabella 4.21:** *Classifica delle condizioni per il Soggetto 35*

#### 4.2.11 Soggetto 35

Il Soggetto 35, come mostrato in Figura 4.11, ottiene scarsi risultati in tutte le condizioni e, ad eccezione delle condizioni 1, 7, 4, risulta inaffidabile in 9 di esse in quanto la media degli angoli di errore per tutte le 7 elevazioni è superiore a  $40^\circ$ . L'unica condizione in cui è possibile estrapolare una tendenza minima nelle risposte è la condizione 7 con media degli angoli di errore per tutte le 7 elevazioni pari a  $36.6^\circ$ . In tale condizione il soggetto percepisce meglio le elevazioni di  $-15^\circ$ ,  $0^\circ$  e  $30^\circ$ , con angolo di errore medio che varia tra  $11.06^\circ$  e  $30.54^\circ$ , il primo associato all'elevazione di  $0^\circ$  ed il secondo a quella di  $30^\circ$ , con deviazione standard che varia tra  $15.81^\circ$  e  $44.83^\circ$ , riferiti rispettivamente all'elevazione di  $0^\circ$  e  $30^\circ$ . Il numero di *front\_back\_reversal* risulta pari a 9, quello di *up/down confusion* pari a 6, il numero di *polar errors* è pari a 8 ed infine la retta di regressione assume il valore di 0.25. Si evince quindi che il soggetto percepisce meglio l'elevazione con PRTF costituite da 3 notch, con profondità di -30 dB e larghezza di banda pari a  $0.15f_c$ . Il soggetto, come vedremo successivamente, non rientra tra quelli selezionati per il calcolo della condizione globale migliore. In Tabella 4.21 è riportata la classifica delle condizioni, mentre in Tabella 4.22 sono riportati i dati statistici di interesse per le 12 condizioni.

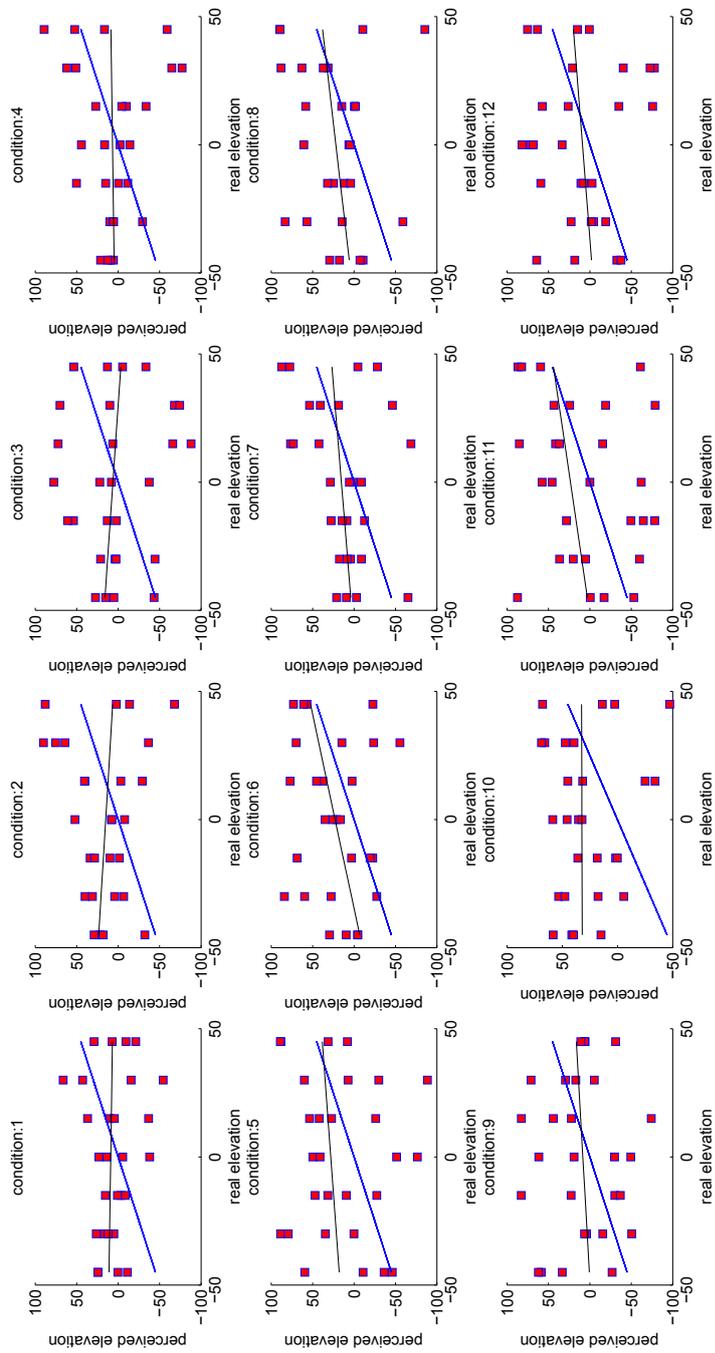


Figura 4.11: Scatterplot del Soggetto ID 35

Cond.	mean_angle_error (°)	reversal_rate (%)	up_down_rate (%)	reg._slope	R2	reg._bias	polar_error_rate (%)
1	35.18	28.57	32.14	-0.04	0	8.95	10.71
2	42.41	28.57	35.71	-0.18	0.06	15	25
3	47.58	32.14	35.71	-0.21	0.25	6.36	39.28
4	38.37	28.57	25	0.04	0	6.75	17.85
5	47.24	39.28	32.14	0.22	0.08	28.06	42.85
6	40.41	35.71	28.57	0.66	0.48	22.85	21.42
7	36.3	32.14	21.42	0.25	0.14	15.25	28.57
8	40.75	25	25	0.35	0.17	21.87	28.57
9	43.43	28.57	25	0.17	0.07	8.41	35.71
10	45.7	32.14	39.28	0	0	32.16	10.71
11	45.55	35.71	28.57	0.45	0.31	23.4	42.85
12	46.65	28.57	32.14	0.24	0.17	8.82	39.28

**Tabella 4.22:** Dati statistici di interesse per il Soggetto 35

Condizione	mean_angle_error (°)
1	32.92
12	35.19
7	38.04
4	39.76
2	41.29
3	40.9
5	46.2
6	41.16
8	45.04
9	49.98
10	42.39
11	48.28

**Tabella 4.23:** *Classifica delle condizioni per il Soggetto 36*

#### 4.2.12 Soggetto 36

Il Soggetto 36, come mostrato in Figura 4.12, ottiene buoni risultati solo in 4 condizioni, risultando inaffidabile nelle restanti 8 in cui la media degli angoli di errore per tutte le 7 elevazioni supera la soglia di  $40^\circ$ . Le condizioni migliori risultano, in ordine decrescente, la 1, 12, 7 e 4 con media degli angoli di errore per tutte le 7 elevazioni pari rispettivamente a  $32.92^\circ$ ,  $35.19^\circ$ ,  $38.04^\circ$  e  $39.76^\circ$ . Nella condizione migliore, il soggetto percepisce meglio le elevazioni di  $-45^\circ$ ,  $0^\circ$ ,  $15^\circ$ ,  $45^\circ$  con angolo di errore medio che varia tra  $20.36^\circ$  e  $28.61^\circ$ , il primo riferito all'elevazione di  $45^\circ$  ed il secondo a quella di  $0^\circ$ , con deviazione standard che varia tra  $27.3^\circ$  e  $38.3^\circ$  associata rispettivamente all'elevazione di  $45^\circ$  e  $-45^\circ$ . In tale condizione, il numero di *front\_back\_reversal* e di *polar errors* è pari a 9, quello di *up/down confusion* è pari a 6 ed infine la retta di regressione presenta pendenza con valore di 1.37. Per le condizioni 7 e 12 è possibile estrapolare una tendenza nei dati solo dopo l'applicazione della tecnica di analisi iterativa, ottenendo pendenze della retta di regressione pari rispettivamente a 0.97 e 0.77 ma, purtroppo, con un numero di *polar errors* piuttosto alto (12). I risultati di tutte le restanti condizioni non presentano dati significativi. Si evince che il soggetto percepisce meglio l'elevazione con PRTF costituite da 3 notch, con profondità di -10 dB e larghezza di banda pari a  $0.15f_c$ . Il soggetto, come vedremo successivamente, non rientra tra quelli selezionati per il calcolo della condizione globale migliore. In Tabella 4.23 è riportata la classifica delle condizioni, mentre in Tabella 4.24 sono riportati i dati statistici di interesse per le 12 condizioni.

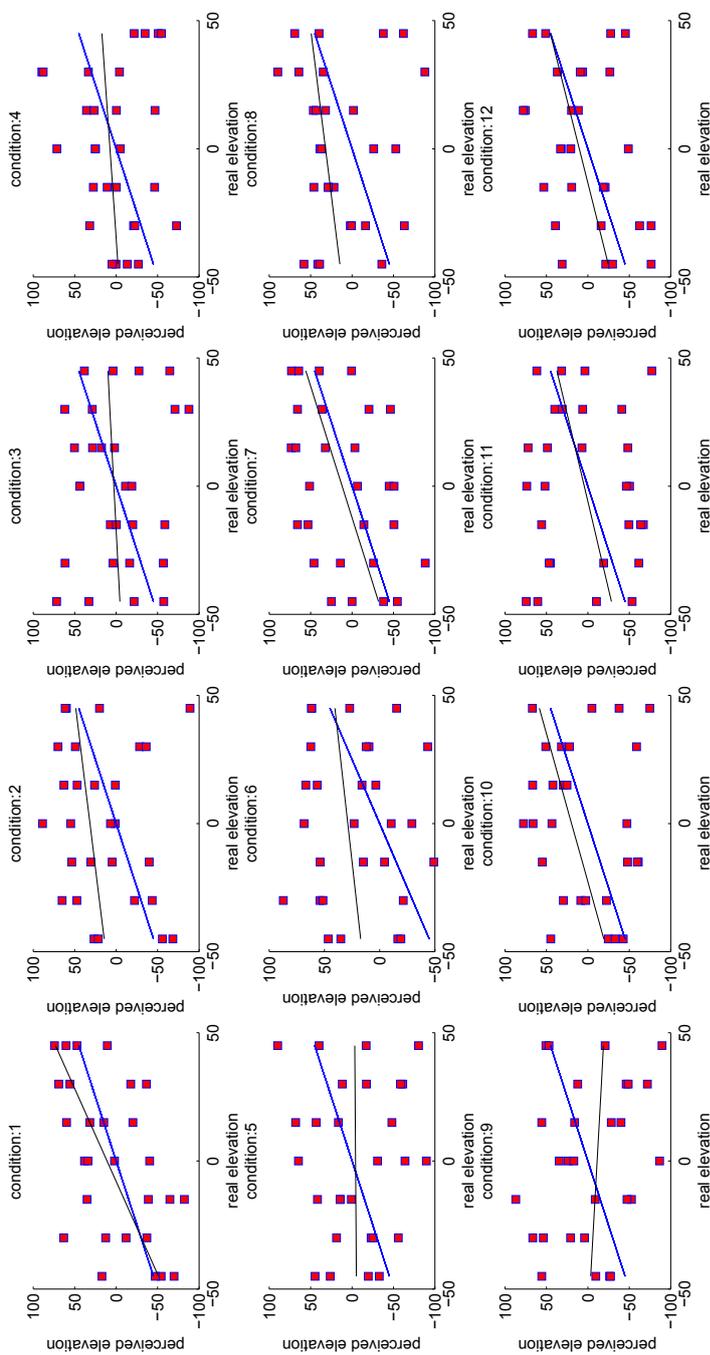


Figura 4.12: Scatterplot del Soggetto ID 36

Cond.	mean_angle_error (°)	reversal_rate (%)	up_down_rate (%)	reg._slope	R2	reg._bias	polar_error_rate (%)
1	32.92	32.14	21.42	1.37	0.86	11.01	32.14
2	41.29	46.42	32.14	0.38	0.17	31.58	28.57
3	40.9	42.85	25	0.16	0.05	2.72	35.71
4	39.76	42.85	25	0.21	0.08	7.45	32.14
5	46.2	67.85	35.71	0.02	0	-4.34	42.85
6	41.16	35.71	28.57	0.25	0.08	28.82	17.85
7	38.04	57.14	21.42	0.97	0.74	11.92	42.85
8	45.04	57.14	35.71	0.38	0.24	31.96	28.57
9	49.98	64.28	42.85	-0.17	0.04	-11.22	39.28
10	42.39	53.57	21.42	0.86	0.72	19.59	42.85
11	48.28	53.57	28.57	0.73	0.59	4.32	53.57
12	35.19	53.57	25	0.77	0.64	10.25	42.85

**Tabella 4.24:** Dati statistici di interesse per il Soggetto 36

Condizione	mean_angle_error (°)
11	25.17
8	26.27
5	27.19
9	27.41
1	27.71
3	27.95
7	28.06
6	29.06
4	29.15
12	29.21
10	30.2
2	30.51

**Tabella 4.25:** *Classifica globale delle condizioni per valore del mean\_angle\_error*

### 4.3 Condizione globale migliore e considerazioni finali

Per ottenere la condizione globale migliore tra le 12 abbiamo dapprima selezionato i soggetti che presentavano almeno 5 condizioni con media degli angoli di errore per tutte le 7 elevazioni inferiore alla soglia di 40° e successivamente, per ogni condizione, è stato calcolato:

- il valore medio degli angoli di errore medi associati ai soggetti selezionati, considerando anche, per quest'ultimo valore, i valori oltre soglia ( $mean\_angle\_error > 40^\circ$ );
- il numero di *up/down confusion*;
- il valore dato dalla seguente formula:  $(regression\_slope\_original + R2\_original)$ , con  $regression\_slope\_original$  e  $R2\_original$  riferiti ai dati originari;

I soggetti selezionati risultano: [20, 21, 22, 24, 27, 28, 30, 33]. I restanti soggetti sono stati scartati in quanto la loro percezione dell'elevazione è stata giudicata inaffidabile dal momento che, in almeno 8 condizioni, presentavano dati senza alcuna tendenza significativa; includere anche tali soggetti nel calcolo della classifica delle condizioni avrebbe sicuramente reso i risultati meno attendibili. Nelle Tabelle 4.25, 4.26, 4.27 sono mostrate rispettivamente le classifiche delle condizioni in base al valore del  $mean\_angle\_error$ , al numero di *up/down confusion* ed al valore della formula  $(regression\_slope\_original + R2\_original)$  mediato sugli 8 soggetti. Confrontando i risultati delle 3 classifiche risulta evidente che la condizione 11 è la migliore. Tale condizione, come è possibile verificare in ognuna delle classifiche legate al singolo soggetto, occupa, per 5 degli 8 soggetti selezionati, o la prima o la seconda posizione ed in essa solo il Soggetto 33 risulta inaffidabile. Inoltre si può ancora notare che nella condizione 11 i soggetti selezionati presentano un basso numero di *polar errors* che varia tra un minimo di 0 per i Soggetti 21, 22 e 27 ed un massimo di 6 per i Soggetti 24 e 30. Quest'ultimo dato indica che per tale condizione l'uso della tecnica di analisi iterativa porta solo ad un trascurabile miglioramento nei risultati,

Condizione	<i>up/down confusion</i>
8	28
9	28
11	29
4	36
5	36
6	36
2	37
12	37
1	38
3	38
10	40
7	44

**Tabella 4.26:** *Classifica globale delle condizioni per numero di up/down confusion*

a dimostrazione del fatto che in essa le risposte dei soggetti sono ottime anche senza alcuna scrematura. Per questo motivo si è scelto di utilizzare, nella terza classifica, i valori di pendenza della retta di regressione e di coefficiente di determinazione riferiti ai dati originari.

Dalla prima e terza classifica notiamo che le condizioni dispari ottengono un risultato migliore rispetto alle condizioni pari, ad eccezione della condizione 8 che occupa le prime posizioni in tutte e tre le classifiche. Inoltre le condizioni 8, 9, 11 occupano, nelle tre classifiche, le posizioni più alte. Questo indica che tendenzialmente i soggetti testati riescono a percepire meglio l'elevazione con PRTF costituite da 3 notch e con profondità di -30 dB. Considerando le classifiche associate ad ognuno degli 8 soggetti selezionati, notiamo che per le condizioni vincenti comprese tra la 7 e la 12, è il numero di notch l'elemento comune tra esse; infatti, i soggetti con condizione vincente compresa tra la 7 e la 12 sono il Soggetto 20, 21, 22, 28, 30, con condizione prima in classifica rispettivamente pari a 8, 11, 7, 9, 11. Tranne la condizione 8, tutte le altre hanno come elemento in comune il numero di notch pari a 3, mentre la larghezza di banda varia tra 2000 Hz,  $0.25f_c$  e  $0.15f_c$ . Se consideriamo, per tali soggetti, le risposte in frequenza dei filtri associate alle condizioni 7, 9, 11 (mostrate in Appendice), esse sono tra loro simili anche con larghezze di banda differenti. Segue che la larghezza di banda, per le condizioni vincenti comprese tra la 7 e la 12, tutte con profondità del notch di -30 dB, è un elemento trascurabile. Al contrario, per le condizioni vincenti comprese tra la 1 e la 6, tutte con profondità del notch di -10 dB, l'elemento comune tra esse è la larghezza di banda. Infatti, i soggetti 24, 27 e 33 presentano rispettivamente come condizione vincente la 6, 5 e 6, tutte con larghezza di banda pari a 2000 Hz e con numero di notch sia di 2 che di 3. Considerando le risposte in frequenza per le condizioni comprese tra la 1 e la 6, notiamo che, a differenza di quelle tra la 7 e la 12, anche per condizioni con uguale numero di notch, la larghezza di banda diventa un elemento non trascurabile nella differenziazione. Da tutto questo possiamo ricavare che, per condizione vincente con notch a -30 dB di profondità (condizioni comprese tra la 7 e la 12), l'elemento più influente è il numero di notch, mentre, per condizione vincente con notch a -10 dB (condizioni comprese tra la 1 e la 6), esso è

Condizione	$mean\_ (regression\_slope + R2)$
11	0.82
9	0.7
8	0.69
3	0.65
5	0.65
7	0.59
1	0.52
4	0.51
6	0.48
10	0.48
2	0.47
12	0.42

**Tabella 4.27:** *Classifica globale delle condizioni in base alla formula  $(regression\_slope + R2)$  mediata sugli 8 soggetti*

la larghezza di banda. Ovviamente tutto ciò dipende dalle diverse profondità di notch utilizzate per le condizioni che fanno sì che, per profondità di -30 dB, la larghezza di banda non influisca più di tanto, mentre, per profondità di notch pari a -10 dB, essa sia un elemento determinante. Per quanto riguarda quindi la larghezza di banda, non è possibile ricavare alcun dato comune tra le prime condizioni. La condizione migliore dei nostri test presenta una larghezza di banda fissa a 2000 kHz, un numero di notch pari a 3 e una profondità pari a -30 dB. Le condizioni peggiori risultano invece la 10 e la 2, entrambe accomunate dall'assenza del notch centrale.

Dal nostro esperimento risultano quindi essenziali per una buona percezione dell'elevazione:

- una profondità per i notch sicuramente superiore a -10 dB, risultato questo in accordo con quelli trovati da Moore in [27] dove, ricordiamo, risultava che le soglie dei notch variavano tra 10 e 20 dB per larghezze di banda di  $0.25f_c$ , ed in accordo anche con i risultati di Poon e Brugge in [29] in cui si è utilizzato un notch con profondità pari a -35 dB, di poco superiore alla profondità di -30 dB considerata nel nostro esperimento;
- la presenza di un secondo notch in una banda di frequenza tra 7000 Hz e 10000 Hz circa, risultato questo in accordo con quello di Iida in [22], in cui i primi due soli notch nella PRTF risultavano essenziali per una localizzazione paragonabile a quella con PRTF originali;
- una larghezza di banda tale da non causare la sovrapposizione dei notch tra di loro e larga abbastanza in modo che il notch sia rilevato facilmente. La larghezza di banda migliore è risultata di 2000 Hz, in accordo con i risultati di Pinto e Lopez-Poveda in [3] e [11]. Le prime posizioni occupate dalla condizione 9 indicano gli ottimi risultati trovati anche considerando una larghezza di banda pari a  $0.25f_c$ , in accordo con lo studio di Moore in [27].

Da brevi interviste sottoposte ai soggetti a fine esperimento si è riscontrata, per alcuni di loro, la difficoltà nel dover proiettare all'esterno, su di una corona circolare, come nell'interfaccia gra-

fica utilizzata dal soggetto durante l'esperimento, un suono sentito all'interno della propria testa. Questo purtroppo è dovuto alla difficoltà di esternalizzare la fonte sonora, problema che potrebbe essere attenuato mediante l'utilizzo di riverbero e compensazione individuale delle cuffie.

Infine, dai risultati appare netta la differenza tra gli 8 soggetti senza alcuna esperienza con il suono spazializzato (Soggetti 24, 26, 27, 28, 32, 33, 35, 36) e soggetti esperti (Soggetti 20, 21, 22, 30).



## Capitolo 5

### Conclusioni e sviluppi futuri

Gli obiettivi principali di questo lavoro di tesi sono riassunti nei due punti seguenti:

- verificare, attraverso un esperimento psicoacustico, l'accuratezza e la funzionalità del modello proposto in [15], alla base del modello successivamente modificato;
- analizzare, attraverso un esperimento psicoacustico, l'influenza della profondità, della larghezza di banda e del numero dei notch presenti nelle PRTF (Pinna Related Transfer Function), riguardo alla percezione dell'elevazione di un suono posizionato sul piano mediano.

Gran parte del lavoro di tesi è stato dedicato alla progettazione ed all'implementazione del test psicoacustico. A tal fine è stato dapprima fatto un lavoro di ricerca di studi realizzati in precedenza da altri autori, consultando in particolar modo le pubblicazioni presenti in due riviste di ricerca scientifica: *Journal of the Acoustic Society of America*<sup>1</sup> e *Journal of the Audio Engineering Society*<sup>2</sup>. Successivamente è stato implementato un esperimento di tipo controllato, con la verifica di 12 condizioni sperimentali, ugualmente sottoposte a tutti i soggetti testati, solo in parte scelti casualmente, dal momento che 4 dei soggetti avevano precedentemente avuto esperienze con il suono spazializzato. Nel Capitolo 3 è descritto il design di tale esperimento, mentre in Appendice sono riportati i codici scritti in ambiente Matlab con le funzioni e gli algoritmi utilizzati per l'implementazione del test. Anche la realizzazione del test audiometrico ha richiesto lo studio di lavori precedenti e la ricerca degli standard ISO per la calibrazione delle cuffie usate durante il test. Il test audiometrico all'uopo approntato per la presente situazione sperimentale, pur non ambendo ad avere le capacità diagnostiche e predittive dei più comuni test in uso da parte della medicina specialistica, risulta un valido strumento in grado di segnalare la presenza di gravi disturbi uditivi, tali da compromettere l'esito del test psicoacustico.

Il modello di PRTF proposto è parametrizzato sulle caratteristiche antropometriche della pinna dei soggetti, recuperate mediante il tracciamento dei contorni principali su di un'immagine della pinna catturata per mezzo di una fotocamera. Tale sistema permette di ricostruire le PRTF

---

<sup>1</sup><http://acousticalsociety.org/>

<sup>2</sup><http://www.aes.org/>

individualizzate in modo che, una volta scelte le coordinate spaziali della sorgente sonora, la percezione spaziale del soggetto sia realistica. Quest'approccio è sicuramente meno costoso rispetto al dover misurare le HRIR (Head Related Impulse Response) associate allo specifico soggetto per poi convolverle nel tempo con la fonte sonora, al fine di ottenere l'effetto 3D. È altresì da evidenziare che non tutti i laboratori dispongono degli strumenti di misura adatti e che la procedura stessa di misurazione risulta piuttosto lunga e non esente da difficoltà. Per la sintesi delle PRTF presentata in questa tesi si è fatto uso di un sistema economico costituito solo di un pc, del software Matlab, di un tablet e di un paio di cuffie, adatto alle caratteristiche di velocità e leggerezza richieste da un sistema *real time* con cui impiegare il nostro modello.

Dal nostro esperimento abbiamo ottenuto ottimi riscontri sulla bontà del modello, in particolare considerando i risultati dei soggetti che avevano precedentemente avuto esperienza con il suono spazializzato, dettagliatamente analizzati nella Capitolo 4, cui si rimanda per una disamina degli elementi di maggiore rilievo. Si pone in evidenza il fatto che, nell'analizzare le 2 diverse profondità del notch e le 3 diverse larghezze di banda testate, abbiamo ricavato per tali parametri i valori generici che meglio si adattano ad una buona percezione dell'elevazione. Tali parametri dei notch infatti, a differenza delle frequenze centrali, non risultano, nell'attuale stato dell'arte, associabili alle caratteristiche antropometriche della pinna e quindi al momento non individualizzabili.

Nel corso del lavoro di tesi sono emerse alcune criticità, la cui risoluzione potrebbe migliorare l'efficacia di tale modellazione alla luce di sviluppi futuri. In primo luogo, le differenze notate tra soggetti senza alcuna esperienza di suoni spazializzati e soggetti esperti rende necessaria una sessione di training preliminare al test psicoacustico. Infatti, oltre alla fase iniziale di *debriefing* in cui vengono comunicate al soggetto testato le istruzioni per svolgere correttamente l'esperimento, è necessario far ascoltare, possibilmente in una data anteriore a quella in cui viene effettuato il test, demo contenenti suoni spazializzati per le diverse elevazioni. In questo modo si introduce il soggetto al concetto di direzione del suono, minimizzando così gli svantaggi di usare soggetti nuovi a questo tipo di esperimento. In secondo luogo, la difficoltà riscontrata da un soggetto con il metodo di puntamento da noi utilizzato, consistente nel proiettare l'elevazione percepita su di un grafico basato sull'elevazione espressa in gradi, pone la necessità di migliorare tale metodo, anche alla luce degli studi presenti nella letteratura riguardanti tale aspetto.

Interessante a riguardo lo studio di Haber in [19], nell'ambito del quale l'autore confronta vari metodi di puntamento tra cui: puntare con parti del corpo (naso, petto, dita), con estensioni del corpo (bastoni, bacchette), con puntatori (*dial pointer*), con metodi verbali come il *clockface* ed infine metodi basati sul disegno della traiettoria (*drawing*). Dallo studio sopra citato è emerso che il puntamento con le tre parti del corpo non differisce molto da quello con estensioni del corpo, considerando l'errore medio. Questi 5 metodi risultano migliori di quelli basati su *dial pointer*, i quali a loro volta appaiono migliori di quelli verbali e di quelli basati su disegno. Possiamo dividere questi metodi in tre classi con un errore medio di  $7.6^\circ$  per metodi basati sul corpo e su estensioni del corpo,  $11.4^\circ$  per quelli basati su puntatori,  $14^\circ$  per il *clockface/drawing*.

Sebbene tali risultati siano stati ottenuti testando soggetti non vedenti, secondo l'autore una performance non soddisfacente con metodi verbali come *clockface* potrebbe essere predetta anche in soggetti vedenti, essendo in questo caso irrilevante l'unica differenza significativa tra le due tipologie di gruppi che è la facoltà, da parte dei vedenti, di ricalibrazione visuo motoria tra

la loro percezione dell'ubicazione spaziale dello stimolo sonoro e dove essi vedono la loro risposta, che di contro influisce sulle performance ottenute mediante l'utilizzo degli altri strumenti di rilevazione in precedenza presentati. Per tali ragioni, nell'ottica di sviluppi futuri, si potrebbe ricorrere ad un metodo di puntamento con dita o estensioni delle stesse, rilevando la posizione indicata per mezzo di PhaceSpace, sistema di *motion capture* attualmente già disponibile in laboratorio.

Ulteriori aspetti da considerare per studi futuri sono i seguenti:

- la relazione tra componenti risonanti della PRTF e la forma delle cavità della pinna. Ricordiamo infatti che nel presente studio si è utilizzato il medesimo blocco risonante per tutti i soggetti, con valori mediati sui soggetti CIPIC;
- l'utilizzo di larghezze di banda e di profondità di notch variabili per le diverse elevazioni, poichè, come indicato in [35], questi parametri si modificano in funzione della posizione della sorgente e dell'ascoltatore. Nel presente lavoro ogni singola condizione è invece stata testata per le 7 elevazioni con parametri fissi.

Per rendere la modellazione più adatta e gestibile in ambiente *real time* sarebbe possibile realizzare una procedura di estrazione automatica dei contorni della pinna, possibilmente a partire da un'immagine 3D della stessa, utile nei casi di orecchie sporgenti dove la pinna non cade su un piano parallelo al piano mediano.

Relativamente alla percezione dell'elevazione, nell'ottica di future implementazioni sperimentali, sarebbe altresì interessante effettuare un confronto tra i risultati che si avrebbero con le PRTF da noi sintetizzate, considerando, per i parametri di profondità, larghezza di banda e numero di notch, la condizione migliore ottenuta dai risultati della presente tesi, con quelli che si otterrebbero utilizzando le HRTF misurate direttamente sul soggetto oppure determinate da modelli matematici (BEM).

Da ultimo si sottolinea come le manipolazioni del suono sviluppate in questo lavoro di tesi possano trovare impiego in tutti quei settori che richiedono l'uso di suono spazializzato: videogame, cinema, musica, sistemi per videoconferenza, strumenti musicali elettronici, aviazione etc. In questi contesti infatti il suono assume un ruolo sempre più importante nel restituire la sensazione di immersione in ambiente virtuale, rendendo più suggestiva (si parla in questi casi di realtà aumentata) e realistica la fruizione dell'informazione acustica fornita all'utente. A tale scopo è possibile integrare il modello di PRTF realizzato nel presente lavoro con il modello della testa e delle altre parti del corpo, in linea con l'idea alla base del modello strutturale.



# Appendici

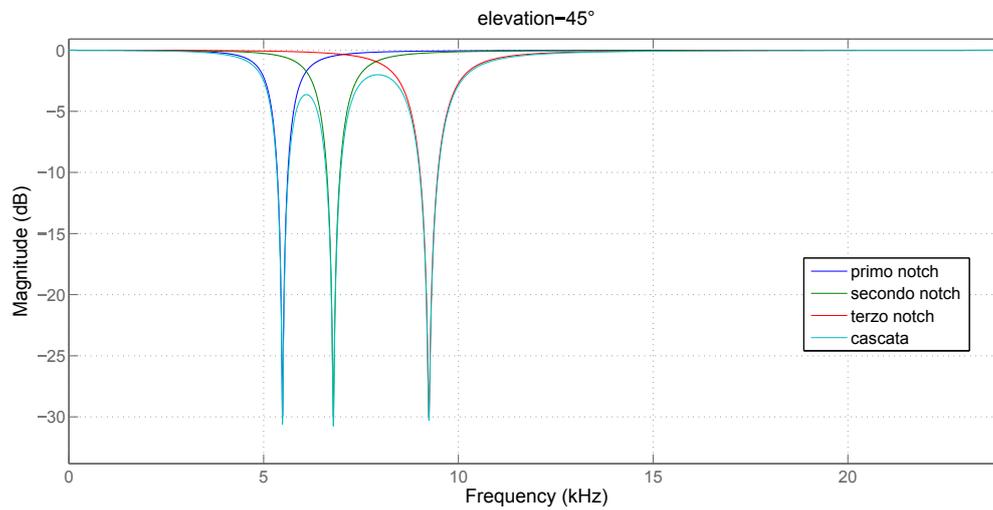
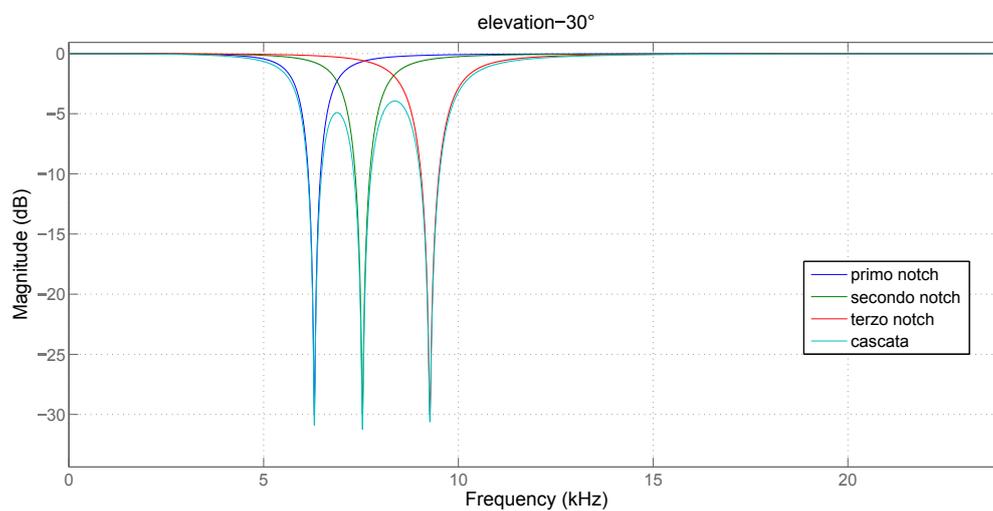


# Appendice A

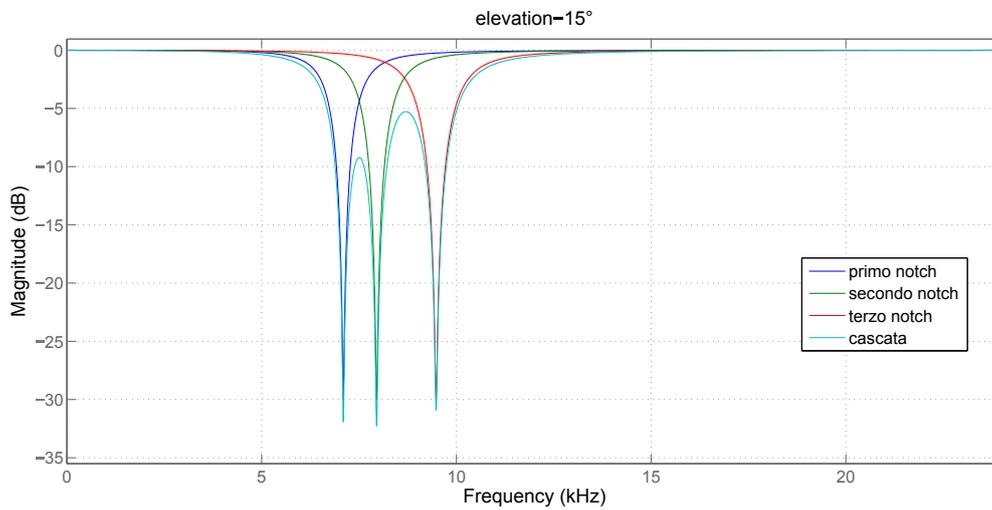
## Grafici

In quest'appendice riportiamo, a titolo esemplificativo, due serie di grafici riferiti al Soggetto 30 (soggetto scelto a caso):

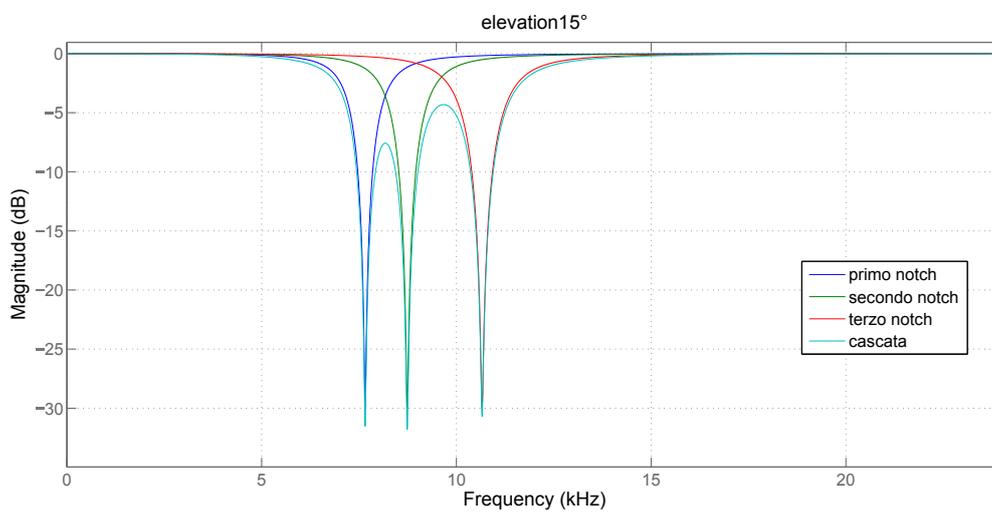
- nei primi sei grafici sono rappresentati i blocchi riflettenti delle PRTF sintetiche, costituiti da 3 notch in cascata (un notch per ogni contorno estratto dalla pinna, come descritto nel Capitolo 2), nel caso di notch con profondità -30 dB e larghezza di banda relativa alla frequenza centrale pari a  $0.15f_c$ , per le elevazioni  $-45^\circ$ ,  $-30^\circ$ ,  $-15^\circ$ ,  $15^\circ$ ,  $30^\circ$  e  $45^\circ$ . Il grafico relativo all'elevazione di  $0^\circ$  è riportato nel Capitolo 2;
- nei restanti grafici sono rappresentate le PRTF sintetiche date dall'unione del blocco riflettente, nel caso di notch con profondità -30 dB e larghezza di banda relativa alla frequenza centrale pari a  $0.15f_c$ , e di quello risonante, quest'ultimo costituito dal parallelo di due picchi. I sei grafici si riferiscono alle elevazioni  $-45^\circ$ ,  $-30^\circ$ ,  $-15^\circ$ ,  $15^\circ$ ,  $30^\circ$  e  $45^\circ$ . Il grafico relativo all'elevazione di  $0^\circ$  è riportato nel Capitolo 2.

(a) Elevazione  $-45^\circ$ (b) Elevazione  $-30^\circ$ 

**Figura A.1:** Primo, secondo e terzo notch in cascata con profondita  $-30$  dB e larghezza di banda di  $0.15f_c$  per le elevazione  $-45^\circ$  e  $-30^\circ$  per il Soggetto 30

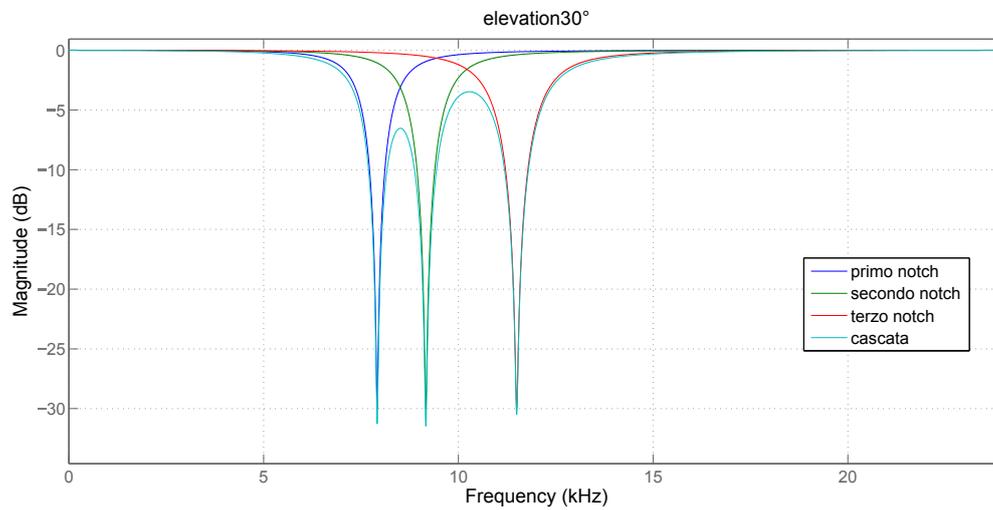


(a) Elevazione  $-15^\circ$

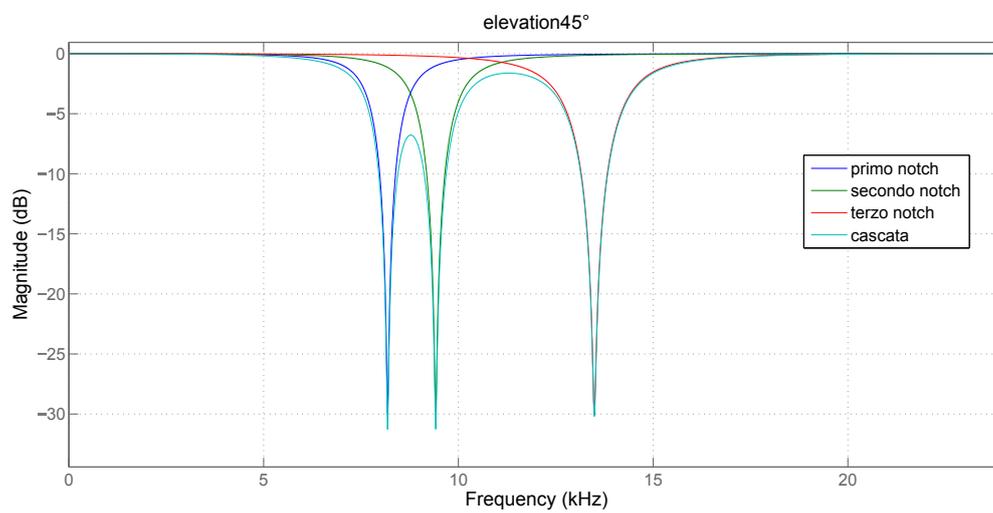


(b) Elevazione  $15^\circ$

**Figura A.2:** Primo, secondo e terzo notch in cascata con profondita  $-30$  dB e larghezza di banda di  $0.15f_c$  per le elevazione  $-15^\circ$  e  $15^\circ$  per il Soggetto 30

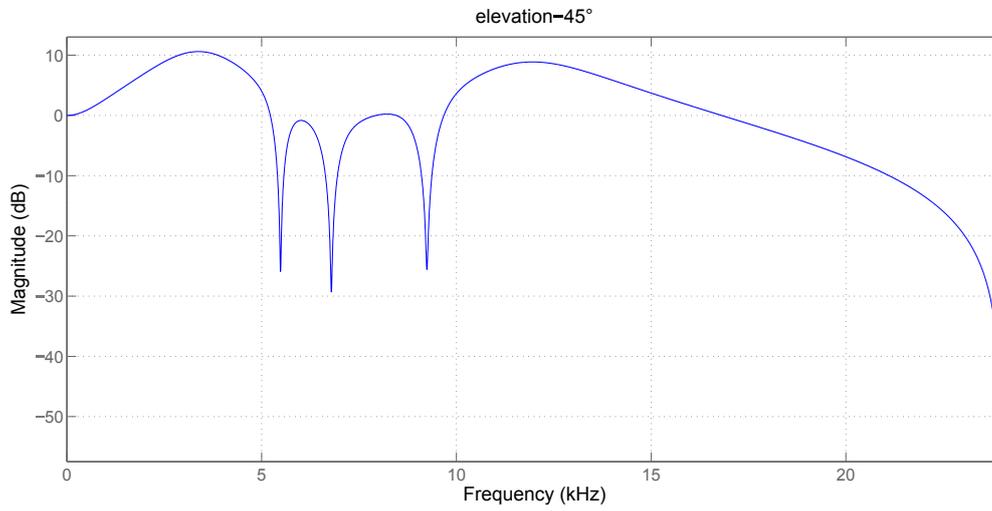


(a) Elevazione 30°

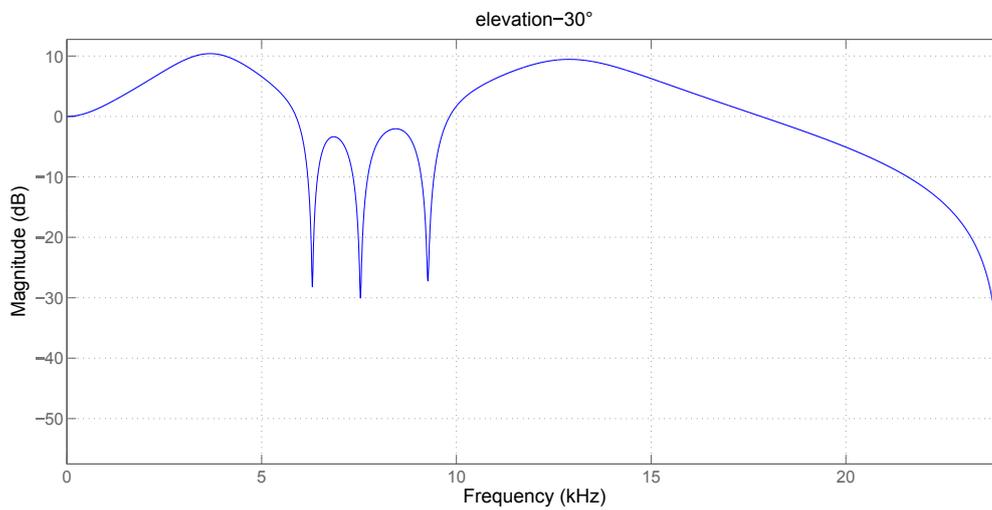


(b) Elevazione 45°

**Figura A.3:** Primo, secondo e terzo notch in cascata con profondità -30 dB e larghezza di banda di  $0.15f_c$  per le elevazione 30° e 45° per il Soggetto 30

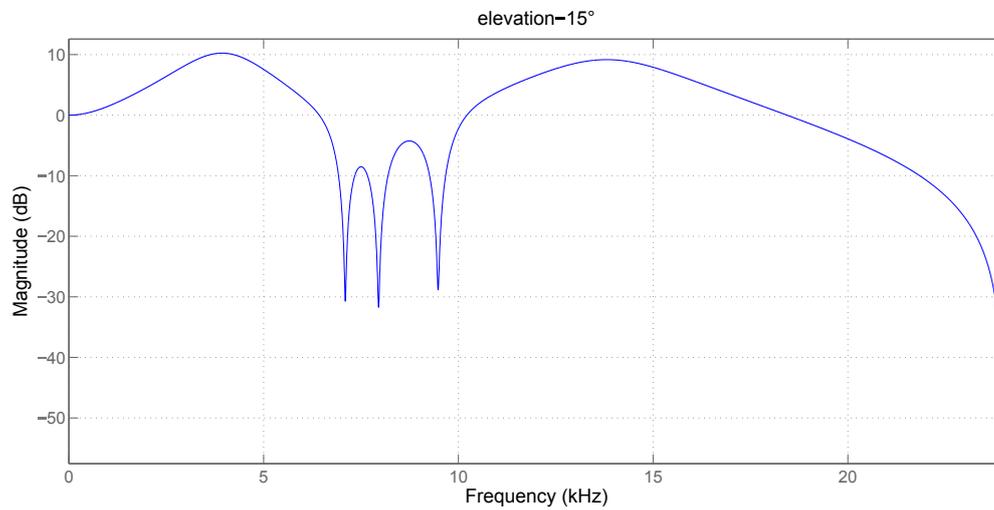
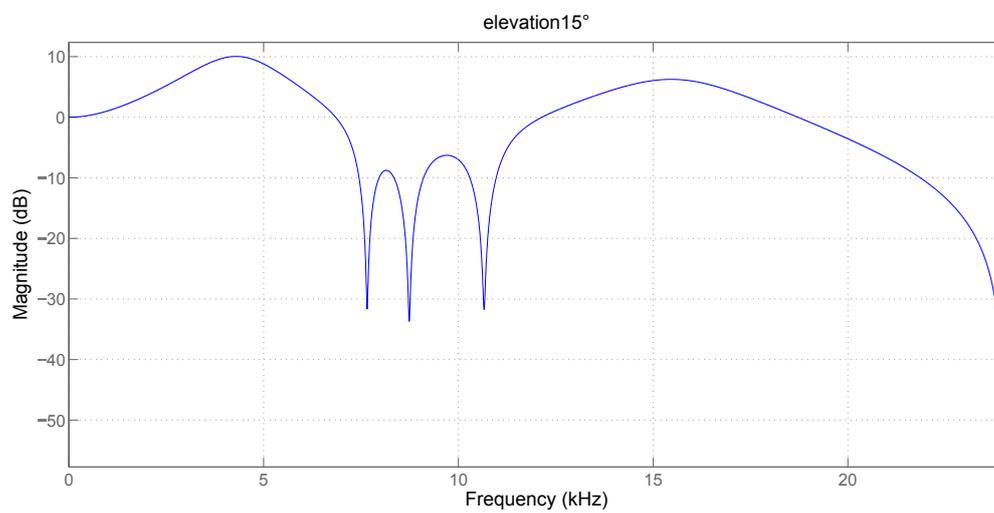


(a) Elevazione  $-45^\circ$

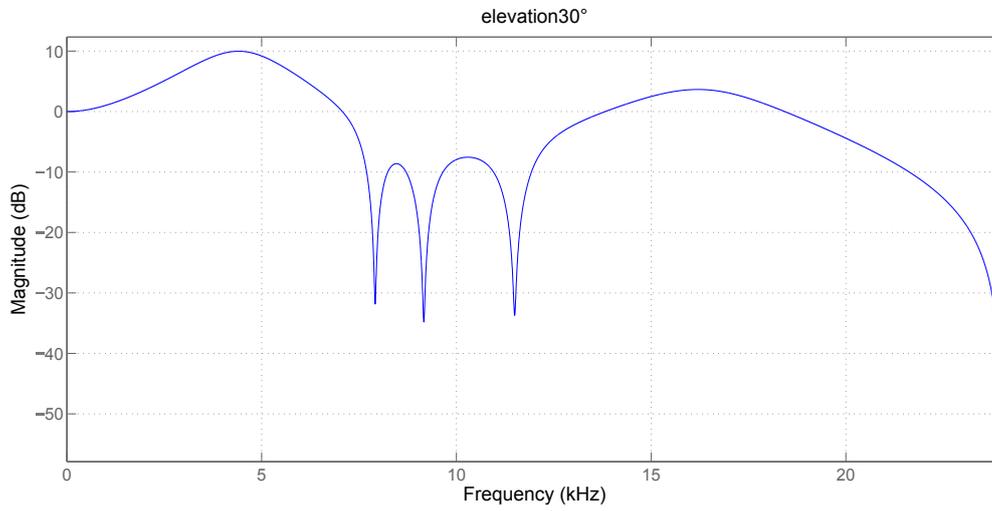


(b) Elevazione  $-30^\circ$

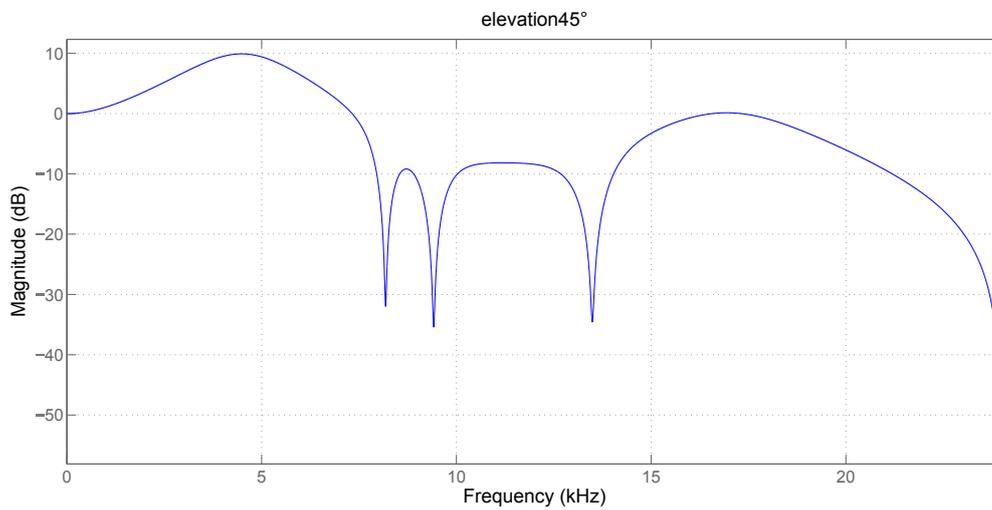
**Figura A.4:** *PRTF sintetiche per il Soggetto 30 con profondità e larghezza di banda del notch rispettivamente a  $-30$  dB e  $0.15f_c$  per elevazioni di  $-45^\circ$  e  $-30^\circ$*

(a) Elevazione  $-15^\circ$ (b) Elevazione  $15^\circ$ 

**Figura A.5:** *PRTF sintetiche per il Soggetto 30 con profondità e larghezza di banda del notch rispettivamente a -30 dB e  $0.15f_c$  per elevazioni di  $-15^\circ$  e  $15^\circ$*



(a) Elevazione 30°



(b) Elevazione 45°

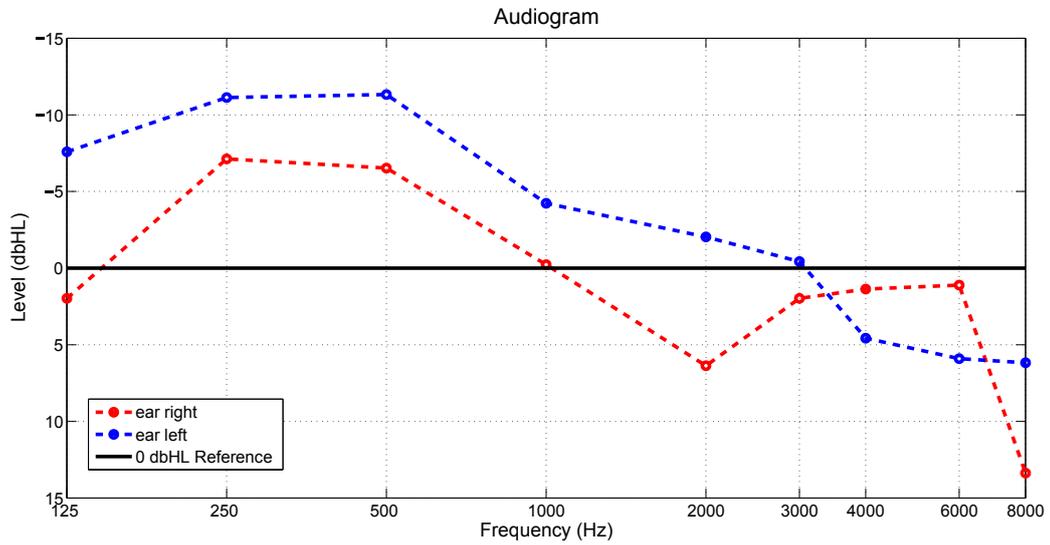
**Figura A.6:** *PRTF sintetiche per il Soggetto 30 con profondità e larghezza di banda del notch rispettivamente a -30 dB e  $0.15f_c$  per elevazioni di 30° e 45°*



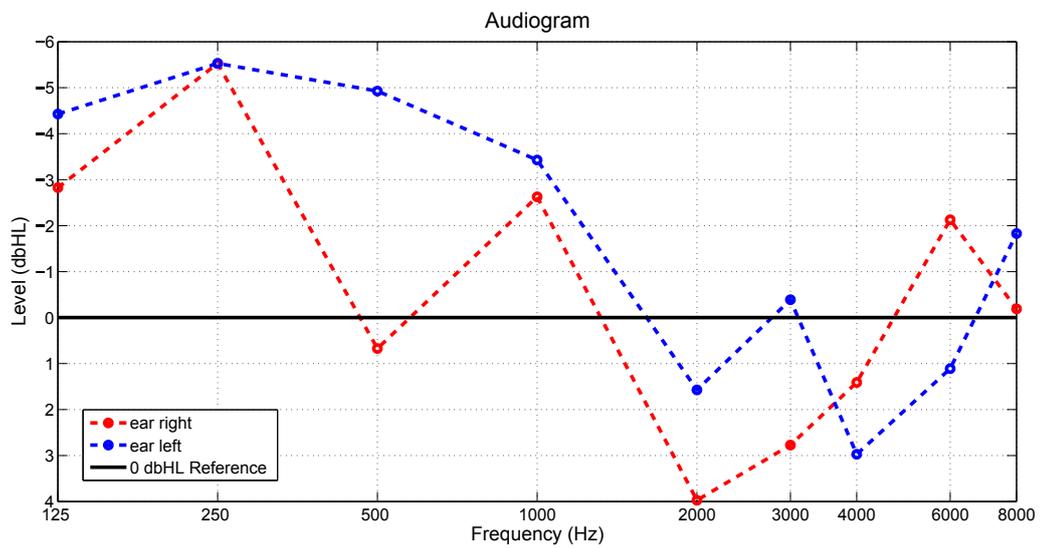
# **Appendice B**

## **Audiogrammi**

In quest'appendice sono riportati gli audiogrammi relativi agli 8 soggetti migliori tra quelli testati. Per una descrizione dettagliata di tali soggetti si rimanda al Capitolo 4. Nel Capitolo 3, oltre ad una breve introduzione sui concetti di audiologia, sono descritte le fasi seguite per la rappresentazione degli audiogrammi, informazioni utili per una adeguata comprensione dei grafici seguenti.

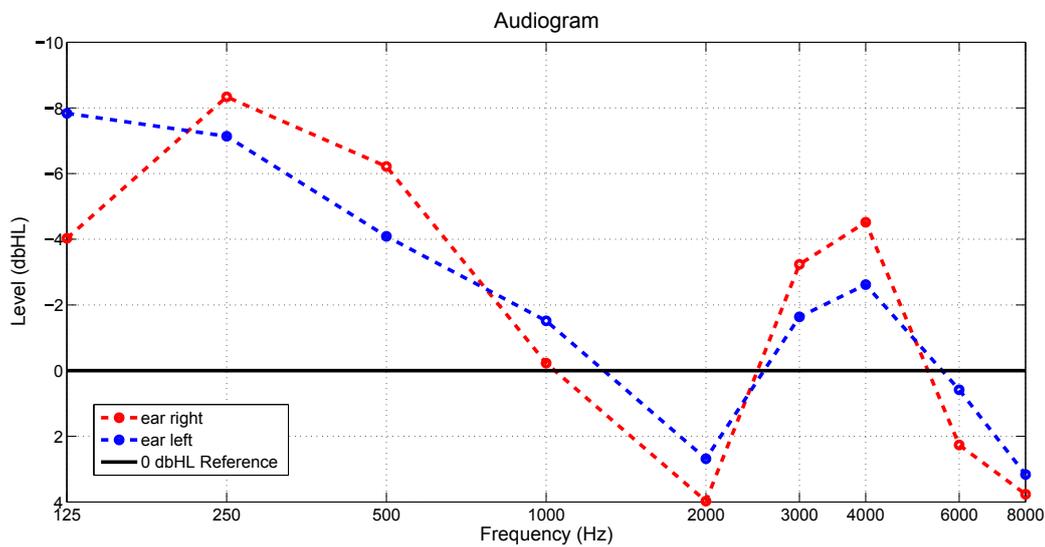


(a) Soggetto 20

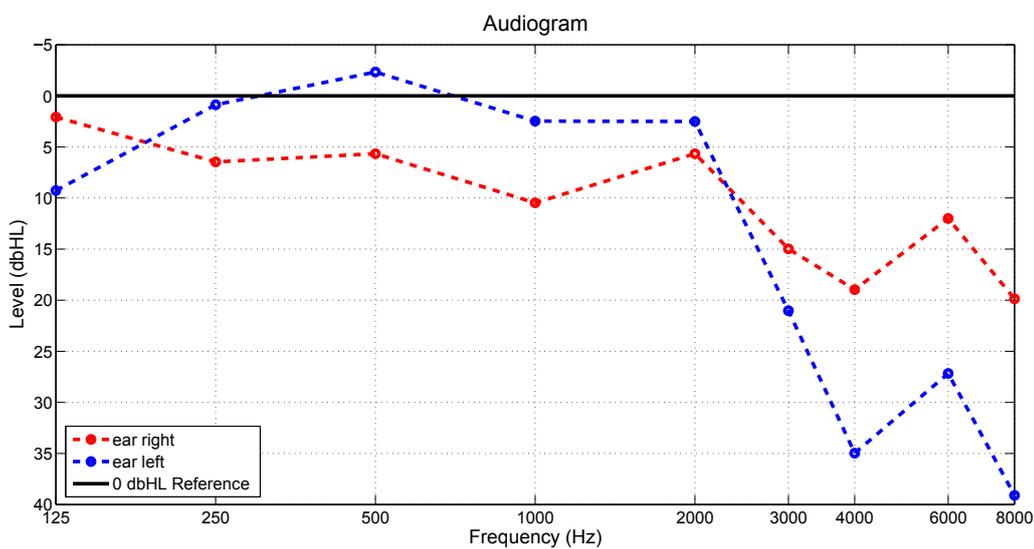


(b) Soggetto 21

**Figura B.1:** Audiogrammi dei Soggetti 20 e 21

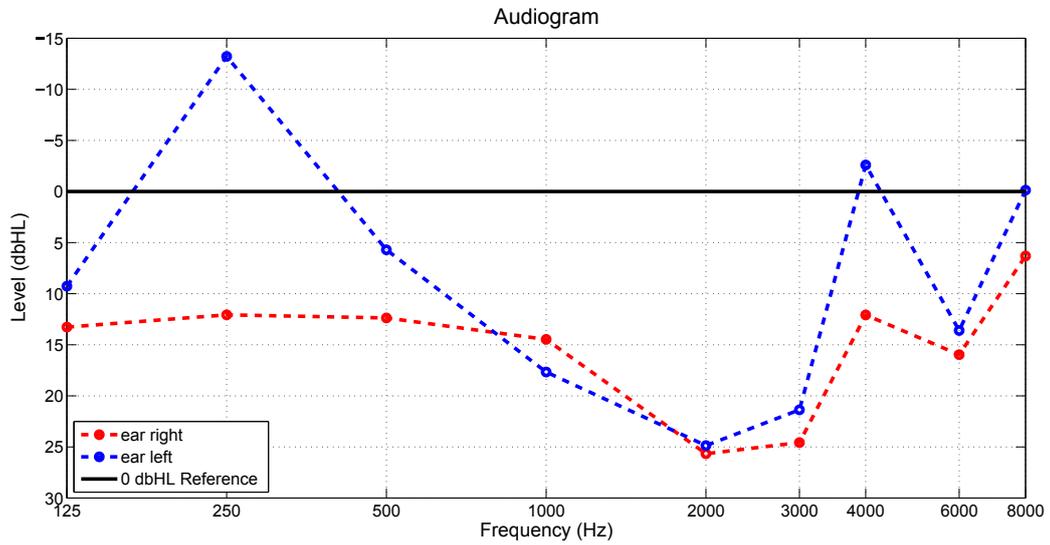


(a) Soggetto 22

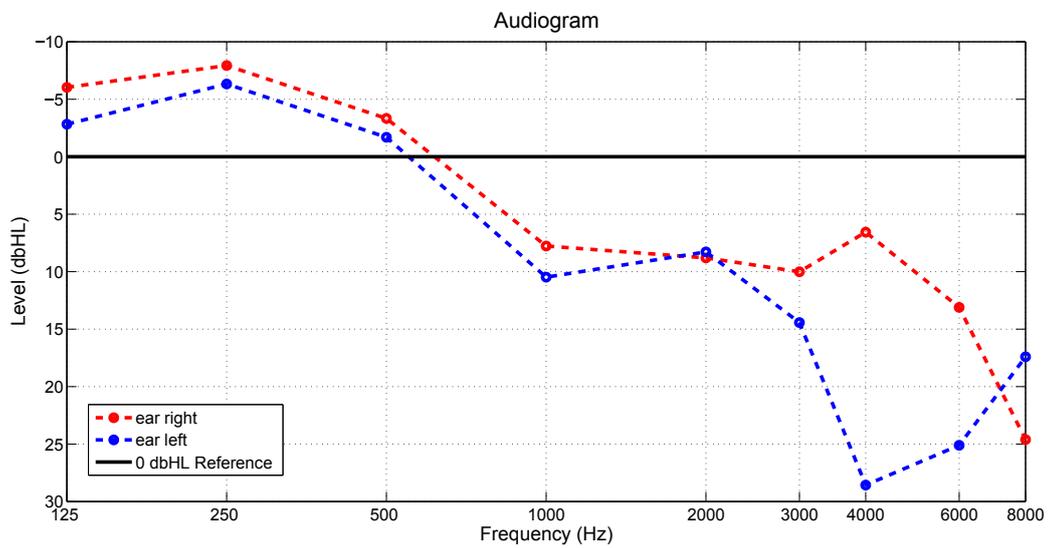


(b) Soggetto 24

**Figura B.2:** Audiogrammi dei Soggetti 22 e 24

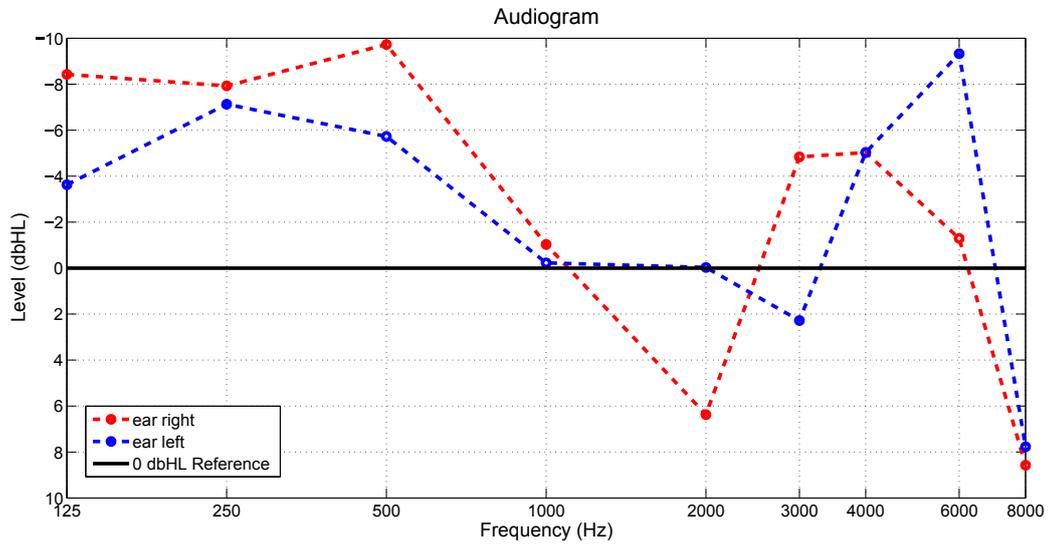


(a) Soggetto 27

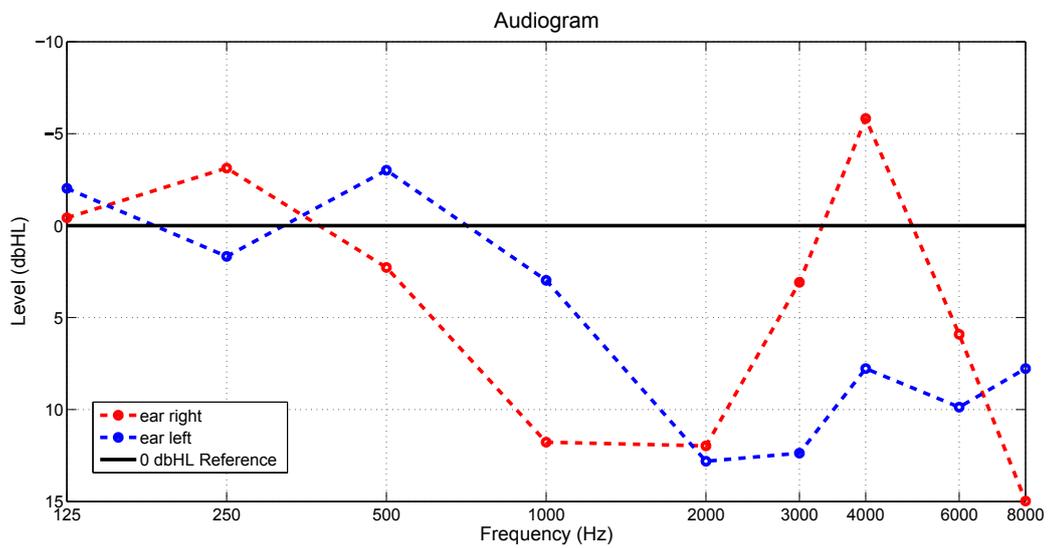


(b) Soggetto 28

**Figura B.3:** Audiogrammi dei Soggetti 27 e 28



(a) Soggetto 30



(b) Soggetto 33

**Figura B.4:** Audiogrammi dei Soggetti 30 e 33



# Appendice C

## Codice Matlab

### C.1 Funzione *expl\_model\_reflection*

```
function [Hsynth, Asynth,HR,fr] = expl_model_reflection(fc, a, fb)

% IN:
% fc: vettore di frequenze centrali -> notch
% a: vettore di ampiezze corrispondenti a vettore fc fc(i) -> a(i)
% fb: vettore di bande passanti corrispondenti a vettore fc fc(i) -> fb(i)

% OUT:
% Hsynth: risposta filtro
% Asynth: ampiezza segnale in dB
% HR: filter object

NFFT = 512;
fs = 48000;
num_track = size(fc,2);
FR = [];
AR = [];
BR = [];
HR = [];
fr=[];

for j = 1 : num_track
    if (fc(j) ~= 0)
        FR = [FR fc(j)];
        AR = [AR a(j)];
        BR = [BR fb(j)];
    end
end

%Calcolo della cascata dei notch
for k = 1 : length(FR)
```

```

    [f] = exp1_notching(AR(k), FR(k), BR(k), fs);
    fr=[fr f];
    % Notch in serie
    if k==1
        HR=[f];
    else
        HR=cascade(HR, f);
    end
end

%Calcolo della risposta in frequenza e in ampiezza del filtro
[Hsynth, f_synth] = freqz(HR, NFFT/2, fs);
Asynth = 20*log10(abs(Hsynth));

end

```

## C.2 Funzione *exp1\_model\_resonance*

```

function [Hsynth, Asynth, b_res, a_res] = exp1_model_resonance(fc, a, fb)

% IN:
% fc: vettore di frequenze centrali -> peak
% a: vettore di ampiezze corrispondenti al vettore fc fc(i) -> a(i)
% fb: vettore di bande passanti corrispondenti a vettore fc fc(i) -> fb(i)

% OUT:
% Hsynth: risposta filtro
% Asynth: ampiezza segnale in dB
% b_res: coeff. numeratore filtro
% a_res: coeff. denominatore filtro

NFFT = 512;
fs = 48000;
num_track = size(fc,2);
FR = [];
AR = [];
BR = [];

for j = 1 : num_track
    if ( fc(j) ~= 0 && a(j) ~= 0 && fb(j) ~= 0 )
        FR = [FR fc(j)];
        AR = [AR a(j)];
        BR = [BR fb(j)];
    end
end

b_res = [];
a_res = [];

```

```

%Calcolo della singola risonanza (k=1) o delle risonanze in parallelo (k=2)
for k = 1 : length(FR)
    % Una sola risonanza
    if (length(FR) == 1)
        [B,A] = peaking2(AR(k), FR(k), BR(k), fs);
    % Due risonanze
    else
        if (k == 1)
            cut_f = (FR(k)+BR(k)/2)/(fs/2);
            [blow, alow] = butter(2, cut_f, 'low'); % filtro passa basso
                ...applicato al primo peak
            [B,A] = peaking2(AR(k), FR(k), BR(k), fs); % prima risonanza
        else
            cut_f_high = (FR(k-1)-BR(k-1)/2)/(fs/2);
            [bhigh, ahigh] = butter(2, cut_f_high, 'high'); % filtro passa
                ...alto applicato al secondo peak
            Wo = FR(k)/(fs/2); BW = BR(k)/(fs/2); % seconda risonanza
            [B,A] = iirpeak(Wo,BW);
            B = B * 10^(AR(k)/20);
        end
    end
end

%- Parallelo
if (k == 1)
    b_res = conv(B,blow);
    a_res = conv(A,alow);
else
    B=conv(B,bhigh);
    A=conv(A,ahigh);
    b_res = conv(b_res,A) + conv(B,a_res);
    a_res = conv(a_res,A);
end
end

%Calcolo della risposta in frequenza e in ampiezza del filtro
[Hsynth,f_synth] = freqz(b_res,a_res,NFFT/2,fs);

if (length(FR) == 2)
    [blow, alow] = butter(2, 0.9, 'low');
    [Hlp, flp] = freqz(blow, alow, NFFT/2,fs);
    Hsynth = Hsynth .* Hlp;
else
    [blow, alow] = butter(2, 0.7, 'low');
    [Hlp, flp] = freqz(blow, alow, NFFT/2,fs);
    Hsynth = Hsynth .* Hlp;
end

Asynth = 20*log10(abs(Hsynth));

end

```

### C.3 Funzione *exp1\_notching*

```
function [f] = exp1_notching(G,fc,fb,fs)

% IN:
%      G:      profondita del notch
%      fc:     frequenza centrale del notch
%      fb:     larghezza di banda del notch
%      fs:     frequenza di campionamento

% OUT:
%      f:      filtro notch

% Parametri di fdesign.parmeq:
%
%      N:      Order
%      F0:     Center Frequency
%      BW:     Bandwidth
%      Gref:   Reference gain
%      G0:     Center frequency gain
%      GBW:    Bandwidth gain
%      fs:     Sampling Frequency

d=fdesign.parmeq('N,F0,BW,Gref,G0,GBW',2,fc,fb,0,G,-3,fs);

%return values
f=design(d);

end
```

### C.4 Funzione *peaking2*

```
function [b, a] = peaking2(G, fc, fb, fs)

% IN:
%      G:      guadagno del peak
%      fc:     frequenza centrale del peak
%      fb:     larghezza di banda del peak
%      fs:     frequenza di campionamento

% OUT:
%      b:      coeff. numeratore del filtro
%      a:      coeff. denominatore del filtro

d = -cos(2*pi*(fc/fs));
V0 = 10^(G/20);
H0 = V0 - 1;

%Calcolo coefficienti del filtro
```

```

if (G <= 0)
    a = (tan(pi*fb/fs) - V0)/(tan(pi*fb/fs) + V0);
else
    a = (tan(pi*fb/fs) - 1)/(tan(pi*fb/fs) + 1);
end

b0 = 1 + (1 + a)*(H0/2);
b1 = d*(1 - a);
b2 = -a - (1 + a)*(H0/2);
a1 = d*(1 - a);
a2 = -a;

%return values
a = [ 1, a1, a2];
b = [ b0, b1, b2];

```

## C.5 Funzione *prtf*

```

function [ Prtf ] = prtf(fc_peak,a_peak,fb_peak,fc_notch,a_notch,fb_notch)

% IN:
% fc_peak: vettore frequenze centrali -> peak
% a_peak: vettore ampiezze fc_peak(i) -> a_peak(i)
% fb_peak: vettore di bande passanti fc_peak(i) -> fb_peak(i)
% fc_notch: vettore di frequenze centrali -> notch
% a_notch: vettore di ampiezze fc_notch fc_notch(i) -> a_notch(i)
% fb_notch: vettore di bande passanti fc_notch(i) -> fb_notch(i)

% OUT:
% Prtf: cascata del modello di risonanza e di quello riflessivo

[H1,A1,b,a]=exp1_model_resonance(fc_peak,a_peak,fb_peak); % risonanza
Pr=dfilt.df2(b,a); % trasforma output modello risonanza in filter object
[H,A,Nr]=exp1_model_reflection(fc_notch,a_notch,fb_notch); %riflessione
Prtf=cascade(Pr,Nr); % modello prtf dalla cascata dei due modelli creati
end

```

## C.6 Funzione *create\_subject*

```

function []=create_subject(F0,id)

%IN:
%      F0 :    vettore frequenze centrali notch
%      id :    id soggetto

load('peak_first_interp1.mat'); % carica matrice con i dati del primo peak
load('peak_second_interp.mat'); % carica matrice con dati del secondo peak
load('equalization_filter.mat'); % carica filtro di equalizzazione headphone
load('conditions.mat'); % carica matrice delle condizioni sperimentali
cd subjects;
mkdir(['subject_' int2str(id)],'filters'); % creo directory subject_id e
                                         ...filters contenente i filtri peak+notch
mkdir(['subject_' int2str(id)],'stimoli_wav'); % creo directory
                                         ...stimoli_wav contenente stimoli gia filtrati da inviare in cuffia
cd ..;
h = waitbar(0,'Please wait...Creating filters');

for i=1:7
    c=-45+(i-1)*15;
    for j=1:12
        if conditions(j,4)==1
            BW=[conditions(j,2),conditions(j,2),conditions(j,2)]; % bandwidth
                                                    ...nel caso di banda fissa
        else
            BW=conditions(j,2)*F0(:,7-(i-1))'; % bandwidth nel caso di
                                                    ...banda relativa alla frequenza centrale
        end

        if conditions(j,3)==3
            fr_centrali=F0(:,7-(i-1))';
            depth=[conditions(j,1),conditions(j,1),conditions(j,1)]; %vettore
                ...dei guadagni nel caso della condizione con 3 notch
        else
            fr_centrali=F0([1,3],7-(i-1))';
            depth=[conditions(j,1),conditions(j,1)]; % vettore dei guadagni
                ...nel caso della condizione con 2 notch
            BW=BW([1,3]); % vettore bandwidth per la condizione con 2 notch
        end

        [f]=prtf([new_peak(i,1),new_peak_second(i,1)],[new_peak(i,2), ...
            new_peak_second(i,2)],[new_peak(i,3), ...
            new_peak_second(i,3)],fr_centrali,depth,BW); %calcolo prtf
        [l]=stimulus3(0.3);
        y=filter(f,1); % filtraggio dello stimolo con prtf
        max_val = 1.05*max(max(abs(y)));
        y = y/max_val;
        y=fftfilt(data.IR,[y';zeros(length(data.IR)-1,1)]);
        wavwrite(y,48000,32,['subjects\' ['subject_' int2str(id)]...

```

```

        '\stimoli_wav\' ['cond_' int2str(j) '_elev_' int2str(c)]]);
    save(['subjects\' ['subject_' int2str(id)] '\filters\'...
        ['cond_' int2str(j) '_elev_' int2str(c),'.mat']], 'f');
    waitbar(((i-1)*12+j)/84);
end
end
close(h);
end

```

## C.7 Funzione *contour\_trace*

```

function [Cx, Cy] = contour_trace(imageName, num_contours)

% IN:
% imageName:   nome del file contenente l'immagine della pinna
% num_contours: numero di contorni da estrarre da immagine pinna

% OUT:
% Cx: vettori colonna corrispondenti alle coordinate x di ogni contorno
% Cy: vettori colonna corrispondenti alle coordinate y di ogni contorno

f = figure;
set(f,'name','Ear tracing','numbertitle','off');
if ( num_contours > 1 )
    plural = 's';
else
    plural = '';
end
img = imread(imageName);
image(flipdim(img,1));
set(gca,'YDir','normal');
truesize(f);

% Tracciamento dei contorni
title(['Draw ' int2str(num_contours) ' line' plural ...
    ' starting from the most external one']);
Cx = zeros(1000,num_contours);
Cy = zeros(1000,num_contours);
for k = 1:num_contours
    try
        fhand = imfreehand('Closed', false);
        if ( ~isempty(fhand) )
            c = fhand.getPosition;
            [rows,cols] = size(c);
            Cx(1:rows,k) = c(:,1);
            Cy(1:rows,k) = c(:,2);
        else
            return;
        end
    end
end

```

```

    catch ME
        uiwait(msgbox('Tracing aborted.', ...
            'Abort','none'));
        return;
    end
end

% Scelta punto corrispondente all'entrata del canale uditivo
title(['Now place ear canal position']);
try
    [xChosen,yChosen] = ginput(1);
catch ME
    uiwait(msgbox('Tracing aborted.', ...
        'Abort','none'));
    return;
end
earCanal = [xChosen yChosen];
hold on;
plot(xChosen, yChosen, '*');

% Costruzione dei vettori Cx e Cy
pixel_to_meters = 0.00026458333; % fattore di conversione pixel-metri
dist = zeros(3,7);
F0 = zeros(3,7);
for i = 1:3 % ciclo sui contorni
    phiIndex = 1;
    for phi = -45:15:45 % ciclo sull'elevazione
        minDiff = 360;
        j = 1;
        while ( Cx(j,i) ~= 0 ) % ciclo sui punti
            angle=radtodeg( atan( (Cy(j,i)-earCanal(2)) / (Cx(j,i)-earCanal(1))));
                if ( angle >= -45 && angle <= 45 )
                    angleDiff = abs(phi - angle);
                    if ( angleDiff < 5 )
                        if ( angleDiff < minDiff )
                            xChosen = Cx(j,i);
                            yChosen = Cy(j,i);
                            minDiff = angleDiff;
                        end
                    end
                end
            j = j+1;
        end
    end
    if ( minDiff ~= 360 )
        dist(i,phiIndex)=pdist([earCanal(1), earCanal(2); ...
            xChosen,yChosen])*pixel_to_meters;
        F0(i,phiIndex) = 343.2 / (2*dist(i,phiIndex));
    else
        dist(i,phiIndex) = 0;
        F0(i,phiIndex) = 0;
    end
end

```

```

        phiIndex = phiIndex + 1;
    end
end
[path, name, ~] = fileparts(imageName);
sep = filesep;
save([path sep name '.mat'],'Cx','Cy', 'earCanal', 'F0');
uiwait(msgbox('Traces saved successfully.', ...
              'Saved','none'));
close();
end

```

## C.8 Funzione *stimulus3*

```

function y=stimulus3(len)
% IN:
% len: lunghezza stimolo in secondi
% OUT:
% y: stimolo

fs=48000;
p=0.25;
len_s=round(fs*len);
ramp = ones(1,len_s);
hann = hanning(.05*fs);
ramp(1:round(0.025*fs)) = hann(1:round(0.025*fs));
ramp(end-round(0.025*fs):end) = hann(round(0.025*fs):end);
sigout = 2*rand(1,len_s)-1;

y = sigout.*ramp;
y = [y zeros(1,round(p*fs)) y zeros(1,round(p*fs)) y];

end

```

## C.9 Funzione *subjects*

```

function varargout = subjects(varargin)
% SUBJECTS M-file for subjects.fig
% SUBJECTS, by itself, creates a new SUBJECTS or raises the existing
% singleton*.
%
% H = SUBJECTS returns the handle to a new SUBJECTS or the handle to
% the existing singleton*.
%
% SUBJECTS('CALLBACK',hObject,eventData,handles,...) calls the local
% function named CALLBACK in SUBJECTS.M with the given input arguments.
%
% SUBJECTS('Property','Value',...) creates a new SUBJECTS or raises the
% existing singleton*. Starting from the left, property value pairs are

```

```

% applied to the GUI before subjects_OpeningFcn gets called. An
% unrecognized property name or invalid value makes property application
% stop. All inputs are passed to subjects_OpeningFcn via varargin.

% Begin initialization code - DO NOT EDIT
gui_Singleton = 1;
gui_State = struct('gui_Name',       mfilename, ...
                  'gui_Singleton',   gui_Singleton, ...
                  'gui_OpeningFcn',  @subjects_OpeningFcn, ...
                  'gui_OutputFcn',  @subjects_OutputFcn, ...
                  'gui_LayoutFcn',   [] , ...
                  'gui_Callback',    []);
if nargin && ischar(varargin{1})
    gui_State.gui_Callback = str2func(varargin{1});
end

if nargout
    [varargout{1:nargout}] = gui_mainfcn(gui_State, varargin{:});
else
    gui_mainfcn(gui_State, varargin{:});
end
% End initialization code - DO NOT EDIT

% --- Executes just before subjects is made visible.
function subjects_OpeningFcn(hObject, eventdata, handles, varargin)
% This function has no output args, see OutputFcn.
% hObject    handle to figure
% eventdata  reserved - to be defined in a future version of MATLAB
% handles    structure with handles and user data (see GUIDATA)
% varargin   command line arguments to subjects (see VARARGIN)

% Choose default command line output for subjects
handles.output = hObject;

% Update handles structure
guidata(hObject, handles);

% Importa dati da subjects.mat files
subjectsStruct = load('subjects');
subjects = subjectsStruct.subjects;
subjectsIdStruct = load('subjects_id');
id = subjectsIdStruct.id;

% Salva le variabili per quest'applicazione
handles.subjects = subjects;
handles.id = id;
handles.dirty = 0;
if ( ~isempty(subjects(:,1)) ),
    handles.gender = subjects{1, 3};

```

```
else
    handles.gender = 'M';
end

guidata(hObject,handles);

subjectsListBox_Callback(hObject, eventdata, handles)
drawListBox(hObject, handles);

% UIWAIT makes subjects wait for user response (see UIRESUME)
% uiwait(handles.figure1);

% --- Outputs from this function are returned to the command line.
function varargout = subjects_OutputFcn(~, eventdata, handles)
% varargout cell array for returning output args (see VARARGOUT);
% hObject handle to figure
% eventdata reserved - to be defined in a future version of MATLAB
% handles structure with handles and user data (see GUIDATA)

% Get default command line output from handles structure
varargout{1} = handles.output;

% --- Executes during object creation, after setting all properties.
function subjectsListBox_CreateFcn(hObject, eventdata, handles)
% hObject handle to subjectsListBox (see GCBO)
% eventdata reserved - to be defined in a future version of MATLAB
% handles empty - handles not created until after all CreateFcns called

% Hint: listBox controls usually have a white background on Windows.
% See ISPC and COMPUTER.
if ispc && isequal(get(hObject,'BackgroundColor'),...
    get(0,'defaultUicontrolBackgroundColor'))
    set(hObject,'BackgroundColor','white');
end

% --- Executes on button press in deleteButton.
function deleteButton_Callback(hObject, eventdata, handles)
% hObject handle to deleteButton (see GCBO)
% eventdata reserved - to be defined in a future version of MATLAB
% handles structure with handles and user data (see GUIDATA)
selected = get(handles.subjectsListBox,'Value');
subjects = handles.subjects;
if ( isempty(subjects(:,1)) ),
    return;
else
    numOfRecords = length(subjects(:,1));
    if ( selected > numOfRecords )
        return;
    end
end
```

```

        end
        subjects(selected,:) = [];
end

handles.subjects = subjects;
handles.dirty = 1;
guidata(hObject,handles);
drawListBox(hObject, handles);

% --- Updates ListBox content.
function drawListBox(hObject, handles)
% hObject      handle to deleteButton (see GCBO)
% handles      structure with handles and user data (see GUIDATA)
subjects = handles.subjects;
selected = get(handles.subjectsListBox,'Value');
if ( isempty(subjects(:,1))),
    recordsAsStrings = {'Database empty'};
    set(handles.lastNameText, 'String', '');
    set(handles.firstNameText, 'String', '');
    set(handles.ageText, 'String', '');
    set(handles.maleRB,'Value',1);
    handles.gender = 'M';
    set(handles.chooseImgButton,'Enable','off');
    set(handles.traceEarButton,'Enable','off');
    set(handles.deleteButton,'Enable','off');
    set(handles.updateButton,'Enable','off');
    set(handles.startExperimentButton,'Enable','off');
else
    subjects = sortrows(subjects, [1, 2]);
    numOfRecords = length(subjects(:,1));
    numOfFields = length(subjects(1,:));
    recordsAsStrings = cell(numOfRecords, numOfFields);
    for i = 1:numOfRecords,
        recordsAsStrings{i} = horzcat(recordsAsStrings{i},...
            subjects{i,1}, ' ', subjects{i,2}, ' ', subjects{i,3}, ' ',...
            subjects{i,4}, ' (id: ', int2str(subjects{i,6}) ,')');
    end
end
set(handles.subjectsListBox,'String',recordsAsStrings(:), 'Value',1);
if ( ~isempty(subjects(:,1)) && selected > numOfRecords )
    selected = numOfRecords;
end
set(handles.subjectsListBox,'Value', selected);

if ( ~isempty(subjects(:,1)) ),
    drawEarImage(handles, 0);
end

% Save variables for this application
handles.subjects = subjects;

```

```

guidata(hObject,handles);

% Debug purpose
assignin('base','handles',handles)

function lastNameText_Callback(hObject, eventdata, handles)
% hObject    handle to lastNameText (see GCBO)
% eventdata  reserved - to be defined in a future version of MATLAB
% handles    structure with handles and user data (see GUIDATA)

% Hints: get(hObject,'String') returns contents of lastNameText as text
%        str2double(get(hObject,'String')) returns contents
%        of lastNameText as a double

% --- Executes during object creation, after setting all properties.
function lastNameText_CreateFcn(hObject, eventdata, handles)
% hObject    handle to lastNameText (see GCBO)
% eventdata  reserved - to be defined in a future version of MATLAB
% handles    empty - handles not created until after all CreateFcns called

% Hint: edit controls usually have a white background on Windows.
%        See ISPC and COMPUTER.
if ispc && isequal(get(hObject,'BackgroundColor'),...
    get(0,'defaultUiControlBackgroundColor'))
    set(hObject,'BackgroundColor','white');
end

% --- Executes on button press in addButton.
function addButton_Callback(hObject, eventdata, handles)
% hObject    handle to addButton (see GCBO)
% eventdata  reserved - to be defined in a future version of MATLAB
% handles    structure with handles and user data (see GUIDATA)
subjects = handles.subjects;
lastName = get(handles.lastNameText, 'String');
firstName = get(handles.firstNameText, 'String');
age = get(handles.ageText, 'String');

if ( isempty(lastName) || isempty(firstName) || isempty(age) ),
msgbox('In order to add a new subject, please insert all required data.',...
'Empty field','warn');
return;
end

[~, status] = str2num(age);
if ( ~status ),
    msgbox('Age must be a numeric value.', ...
        'Bad input','warn');
end

```

```

        return;
end

numOfRecords = length(subjects(:,1));
subjects{numOfRecords+1,1} = lastName;
subjects{numOfRecords+1,2} = firstName;
subjects{numOfRecords+1,3} = handles.gender;
subjects{numOfRecords+1,4} = age;
subjects{numOfRecords+1,5} = '';
subjects{numOfRecords+1,6} = handles.id + 1 ;

% Save variables for this application
handles.id = handles.id + 1 ;
handles.subjects = subjects;
handles.dirty = 1;
guidata(hObject,handles);
drawListBox(hObject, handles);

function firstNameText_Callback(hObject, eventdata, handles)
% hObject      handle to firstNameText (see GCBO)
% eventdata    reserved - to be defined in a future version of MATLAB
% handles      structure with handles and user data (see GUIDATA)

% Hints: get(hObject,'String') returns contents of firstNameText as text
%         str2double(get(hObject,'String')) returns contents
%         of firstNameText as a double

% --- Executes during object creation, after setting all properties.
function firstNameText_CreateFcn(hObject, eventdata, handles)
% hObject      handle to firstNameText (see GCBO)
% eventdata    reserved - to be defined in a future version of MATLAB
% handles      empty - handles not created until after all CreateFcns called

% Hint: edit controls usually have a white background on Windows.
%         See ISPC and COMPUTER.
if ispc && isequal(get(hObject,'BackgroundColor'),...
    get(0,'defaultUiControlBackgroundColor'))
    set(hObject,'BackgroundColor','white');
end

function ageText_Callback(hObject, eventdata, handles)
% hObject      handle to ageText (see GCBO)
% eventdata    reserved - to be defined in a future version of MATLAB
% handles      structure with handles and user data (see GUIDATA)

% Hints: get(hObject,'String') returns contents of ageText as text

```

```

%         str2double(get(hObject,'String')) returns contents
%         of ageText as a double

% --- Executes during object creation, after setting all properties.
function ageText_CreateFcn(hObject, eventdata, handles)
% hObject    handle to ageText (see GCBO)
% eventdata  reserved - to be defined in a future version of MATLAB
% handles    empty - handles not created until after all CreateFcns called

% Hint: edit controls usually have a white background on Windows.
%         See ISPC and COMPUTER.
if ispc && isequal(get(hObject,'BackgroundColor'),...
    get(0,'defaultUicontrolBackgroundColor'))
    set(hObject,'BackgroundColor','white');
end

% --- Executes when selected object is changed in genderButtonGroup.
function genderButtonGroup_SelectionChangeFcn(hObject, eventdata, handles)
% hObject    handle to the selected object in genderButtonGroup
% eventdata  structure with the following fields (see UIBUTTONGROUP)
%         EventName: string 'SelectionChanged' (read only)
%         OldValue: handle of the previously selected object
%         or empty if none was selected
%         NewValue: handle of the currently selected object
% handles    structure with handles and user data (see GUIDATA)
guidata(hObject, handles);
if ( hObject == handles.maleRB )
    gender = 'M';
else
    gender = 'F';
end
handles.gender = gender;
guidata(hObject,handles);

% --- Executes on button press in chooseImgButton.
function chooseImgButton_Callback(hObject, eventdata, handles)
% hObject    handle to chooseImgButton (see GCBO)
% eventdata  reserved - to be defined in a future version of MATLAB
% handles    structure with handles and user data (see GUIDATA)
[fn, pn] = uigetfile({'*.jpg;*.tif;*.png;*.gif;*.bmp','All Image Files';...
    '*.*','All Files' },'Choose ear image');
if ( (isnumeric(fn) && fn == 0) || (isnumeric(pn) && pn == 0) )
    return;
end;
cd img;
if ( ~exist(fn, 'file') ),
    msgbox('Please place the ear image in the img folder.', ...

```

```

                                'Invalid image path','warn');
    cd ..;
    return;
end
cd ..;
selected = get(handles.subjectsListBox,'Value');
subjects = handles.subjects;
if ( strcmpi(fn,'blank.png') )
    subjects{selected,5} = '';
    set(handles.traceEarButton,'Enable','off')
else
    subjects{selected,5} = fn;
    set(handles.traceEarButton,'Enable','on')
end
handles.subjects = subjects;
handles.dirty = 1;
guidata(hObject,handles);

drawEarImage(handles, 0);

% --- Executes on button press in updateButton.
function updateButton_Callback(hObject, eventdata, handles)
% hObject    handle to updateButton (see GCBO)
% eventdata  reserved - to be defined in a future version of MATLAB
% handles    structure with handles and user data (see GUIDATA)

subjects = handles.subjects;
selected = get(handles.subjectsListBox,'Value');
numOfRecords = length(subjects(:,1));
if ( selected > numOfRecords ),
    return;
end

lastName = get(handles.lastNameText, 'String');
firstName = get(handles.firstNameText, 'String');
age = get(handles.ageText, 'String');

if ( isempty(lastName) || isempty(firstName) || isempty(age) ),
    msgbox('In order to edit a subject, please insert all required data.', ...
        'Empty field','warn');
    return;
end
[~, status] = str2num(age);
if ( ~status ),
    msgbox('Age must be a numeric value.', ...
        'Bad input','warn');
    return;
end

subjects{selected,1} = lastName;

```

```

subjects{selected,2} = firstName;
subjects{selected,3} = handles.gender;
subjects{selected,4} = age;

% Save variables for this application
handles.subjects = subjects;
handles.dirty = 1;
guidata(hObject,handles);
drawListBox(hObject, handles);

% -----
function fileMenu_Callback(hObject, eventdata, handles)
% hObject    handle to fileMenu (see GCBO)
% eventdata  reserved - to be defined in a future version of MATLAB
% handles    structure with handles and user data (see GUIDATA)

% -----
function saveMenu_Callback(hObject, eventdata, handles)
% hObject    handle to saveMenu (see GCBO)
% eventdata  reserved - to be defined in a future version of MATLAB
% handles    structure with handles and user data (see GUIDATA)
saveToFileSystem(hObject, handles, 1);
handles.dirty = 0;
guidata(hObject, handles);

% -----
function closeMenu_Callback(hObject, eventdata, handles)
% hObject    handle to closeMenu (see GCBO)
% eventdata  reserved - to be defined in a future version of MATLAB
% handles    structure with handles and user data (see GUIDATA)
close(handles.figure1);

% --- Executes on selection change in subjectsListBox.
function subjectsListBox_Callback(hObject, eventdata, handles)
% hObject    handle to subjectsListBox (see GCBO)
% eventdata  reserved - to be defined in a future version of MATLAB
% handles    structure with handles and user data (see GUIDATA)

% Hints: contents = cellstr(get(hObject,'String')) returns
%        subjectsListBox contents as cell array
%        contents{get(hObject,'Value')} returns selected item from
%        subjectsListBox

subjects = handles.subjects;
selected = get(handles.subjectsListBox,'Value');
numOfRecords = length(subjects(:,1));
if ( selected > numOfRecords ),

```

```

    set(handles.lastNameText, 'String', '');
    set(handles.firstNameText, 'String', '');
    set(handles.ageText, 'String', '');
    set(handles.maleRB, 'Value', 1);
    handles.gender = 'M';
    set(handles.chooseImgButton, 'Enable', 'off');
    set(handles.traceEarButton, 'Enable', 'off');
    set(handles.deleteButton, 'Enable', 'off');
    set(handles.updateButton, 'Enable', 'off');
    set(handles.startExperimentButton, 'Enable', 'off');
    drawEarImage(handles, 1);
    return;
end

set(handles.lastNameText, 'String', subjects{selected,1});
set(handles.firstNameText, 'String', subjects{selected,2});
set(handles.ageText, 'String', subjects{selected,4});
if ( strcmpi( subjects{selected,3}, 'M' ) )
    set(handles.maleRB, 'Value', 1);
    handles.gender = 'M';
else
    set(handles.femaleRB, 'Value', 1);
    handles.gender = 'F';
end

drawEarImage(handles, 0);

set(handles.chooseImgButton, 'Enable', 'on');
set(handles.deleteButton, 'Enable', 'on');
set(handles.updateButton, 'Enable', 'on');
set(handles.startExperimentButton, 'Enable', 'on');
if ( isempty(subjects{selected,5}) ),
    set(handles.traceEarButton, 'Enable', 'off');
    set(handles.startExperimentButton, 'Enable', 'off');
else
    set(handles.traceEarButton, 'Enable', 'on')
end

cd img;
if ( exist(subjects{selected,5}, 'file') )
    [~, earDataFile, ~] = fileparts(subjects{selected,5});
    if ( ~exist( [earDataFile '.mat'], 'file') )
        set(handles.startExperimentButton, 'Enable', 'off');
    end
end

cd ..;

handles.subjects = subjects;
guidata(hObject, handles);

```

```

% --- Executes during object creation, after setting all properties.
function subjectListBox_CreateFcn(hObject, eventdata, handles)
% hObject    handle to subjectListBox (see GCBO)
% eventdata  reserved - to be defined in a future version of MATLAB
% handles    empty - handles not created until after all CreateFcns called

% Hint: listbox controls usually have a white background on Windows.
%         See ISPC and COMPUTER.
if ispc && isequal(get(hObject,'BackgroundColor'),...
    get(0,'defaultUicontrolBackgroundColor'))
    set(hObject,'BackgroundColor','white');
end

function saveToFileSystem(hObject, handles, displayMsg)

guidata(hObject, handles);
subjects = handles.subjects;
id = handles.id;
save('subjects.mat' , 'subjects');
save('subjects_id.mat', 'id');
if (displayMsg)
    msgbox('Saved successfully.', ...
        'Save','none');
end

% --- Executes when user attempts to close figure1.
function figure1_CloseRequestFcn(hObject, eventdata, handles)
% hObject    handle to figure1 (see GCBO)
% eventdata  reserved - to be defined in a future version of MATLAB
% handles    structure with handles and user data (see GUIDATA)

% Hint: delete(hObject) closes the figure

if ( handles.dirty == 0 )
    delete(hObject);
    return;
end
choice=questdlg('Do you want to save changes?', ...
    'Confirm changes', ...
    'Yes', 'No', 'Cancel', 'Cancel');
switch choice
case 'Yes'
    guidata(hObject, handles);
    saveToFileSystem(hObject, handles, 0);
    delete(hObject)
case 'No'
    guidata(hObject, handles);
    delete(hObject)
case 'Cancel'

```

```

        return;
end

function drawEarImage(handles, blank)

if ( blank ),
    cd img;
    absolutePath = which('blank.png');
    imshow(absolutePath,'Parent',gca);
    cd ..;
    return;
end

selected = get(handles.subjectsListBox,'Value');
cd img;
earImageName = handles.subjects{selected,5};
if ( ~exist(earImageName, 'file') ),
    absolutePath = which('blank.png');
else
    absolutePath = which(earImageName);
end
cd ..;
imshow(absolutePath,'Parent',gca);

% --- Executes on button press in traceEarButton.
function traceEarButton_Callback(hObject, eventdata, handles)
% hObject    handle to traceEarButton (see GCBO)
% eventdata  reserved - to be defined in a future version of MATLAB
% handles    structure with handles and user data (see GUIDATA)

selected = get(handles.subjectsListBox,'Value');
cd img
earImage = which(handles.subjects{selected,5});
cd ..;
contour_trace(earImage, 3);

set(handles.startExperimentButton,'Enable','on');
cd img;
if ( exist(handles.subjects{selected,5}, 'file') )
    [~, earDataFile, ~] = fileparts(handles.subjects{selected,5});
    if ( ~exist( [earDataFile '.mat'], 'file') )
        set(handles.startExperimentButton,'Enable','off');
    end
end
cd ..;

% --- Executes on button press in startExperimentButton.
function startExperimentButton_Callback(hObject, eventdata, handles)
% hObject    handle to startExperimentButton (see GCBO)
% eventdata  reserved - to be defined in a future version of MATLAB

```

```

% handles      structure with handles and user data (see GUIDATA)
if ( handles.dirty == 1 )
    choice=questdlg(['In order to proceed with this experiment, you need', ...
        ' to save current changes. Do you want to continue? '], ...
        'Confirm changes', ...
        'Yes', 'Cancel', 'Cancel');
    switch choice
        case 'Yes'
            guidata(hObject, handles);
            saveToFileSystem(hObject, handles, 0);
            handles.dirty = 0;
            guidata(hObject, handles);
        case 'Cancel'
            return;
    end
end

%Find current selected ID
selected = get(handles.subjectsListBox,'Value');
id = handles.subjects{selected,6};

%Check repetitions number
repetitions = get(handles.repetitionsText, 'String');
repetitions = str2num(repetitions);
if ( isempty(repetitions) || ~isnumeric(repetitions) || repetitions <= 0 )
    repetitions = 1;
    box = msgbox('Invalid repetitions number. Using default 1 value.', ...
        'Warning','warn');
    uiwait(box);
end

%Check pauseInterval number
pauseInterval = get(handles.pauseIntervalText, 'String');
pauseInterval = str2num(pauseInterval);
if ( isempty(pauseInterval) || ~isnumeric(pauseInterval) ||...
    pauseInterval <= 0 )
    pauseInterval = 50;
    box = msgbox('Invalid Pause Interval value. Using default 50 value.', ...
        'Warning','warn');
    uiwait(box);
end

%Shared data between GUIs
setappdata(0,'subjectsHandler', hObject);
subjectsHandler = getappdata(0,'subjectsHandler');
setappdata(subjectsHandler,'sharedID', id);
setappdata(subjectsHandler,'repetitions', repetitions);
setappdata(subjectsHandler,'pauseInterval', pauseInterval);

experiment1;

```

```

% -----
function editMenu_Callback(hObject, eventdata, handles)
% hObject    handle to editMenu (see GCBO)
% eventdata  reserved - to be defined in a future version of MATLAB
% handles    structure with handles and user data (see GUIDATA)

% -----
function Untitled_2_Callback(hObject, eventdata, handles)
% hObject    handle to Untitled_2 (see GCBO)
% eventdata  reserved - to be defined in a future version of MATLAB
% handles    structure with handles and user data (see GUIDATA)

% -----
function resultsMenu_Callback(hObject, eventdata, handles)
% hObject    handle to resultsMenu (see GCBO)
% eventdata  reserved - to be defined in a future version of MATLAB
% handles    structure with handles and user data (see GUIDATA)
selected = get(handles.subjectsListBox, 'Value');
cd results;
numOfRecords = length(handles.subjects(:,1));
if ( selected > numOfRecords ||...
    ~exist([int2str(handles.subjects{selected,6}) '.mat'] , 'file') )
    msgbox('There are no results to be displayed for this subject.', ...
        'No results', 'warn');
    cd ..;
    return;
end
results = load([int2str(handles.subjects{selected,6}) '.mat']);
disp(results.results);
cd ..;

% -----
function clearDBMenu_Callback(hObject, eventdata, handles)
% hObject    handle to clearDBMenu (see GCBO)
% eventdata  reserved - to be defined in a future version of MATLAB
% handles    structure with handles and user data (see GUIDATA)
choice = questdlg(['All subject database entries will be removed. ', ...
    'This action cannot be undone. Experiments results and ', ...
    'ear tracing data will not be removed. ', ...
    'Are you sure that you want to continue? '], ...
    'Confirm deletion', ...
    'Yes', 'No', 'No');
switch choice
    case 'Yes'

```

```

        guidata(hObject, handles);
        handles.subjects = cell(0,6);
        saveToFileSystem(hObject, handles, 0);
        handles.dirty = 0;
        handles.selected = 1;
        set(handles.subjectsListBox,'Value', 1);
        guidata(hObject, handles);
        drawListBox(hObject, handles);
        drawEarImage(handles, 1)
    case 'No'
        return;
end

% --- Executes during object creation, after setting all properties.
function genderButtonGroup_CreateFcn(hObject, eventdata, handles)
% hObject    handle to genderButtonGroup (see GCBO)
% eventdata  reserved - to be defined in a future version of MATLAB
% handles    empty - handles not created until after all CreateFcns called

function repetitionsText_Callback(hObject, eventdata, handles)
% hObject    handle to repetitionsText (see GCBO)
% eventdata  reserved - to be defined in a future version of MATLAB
% handles    structure with handles and user data (see GUIDATA)

% Hints: get(hObject,'String') returns contents of repetitionsText as text
%        str2double(get(hObject,'String')) returns contents of
%        repetitionsText as a double

% --- Executes during object creation, after setting all properties.
function repetitionsText_CreateFcn(hObject, eventdata, handles)
% hObject    handle to repetitionsText (see GCBO)
% eventdata  reserved - to be defined in a future version of MATLAB
% handles    empty - handles not created until after all CreateFcns called

% Hint: edit controls usually have a white background on Windows.
%       See ISPC and COMPUTER.
if ispc && isequal(get(hObject,'BackgroundColor'),...
    get(0,'defaultUiControlBackgroundColor'))
    set(hObject,'BackgroundColor','white');
end

function pauseIntervalText_Callback(hObject, eventdata, handles)
% hObject    handle to pauseIntervalText (see GCBO)
% eventdata  reserved - to be defined in a future version of MATLAB
% handles    structure with handles and user data (see GUIDATA)

```

---

```

% Hints:get(hObject,'String') returns contents of pauseIntervalText as text
%     str2double(get(hObject,'String')) returns contents of
%     pauseIntervalText as a double

% --- Executes during object creation, after setting all properties.
function pauseIntervalText_CreateFcn(hObject, eventdata, handles)
% hObject    handle to pauseIntervalText (see GCBO)
% eventdata  reserved - to be defined in a future version of MATLAB
% handles    empty - handles not created until after all CreateFcns called

% Hint: edit controls usually have a white background on Windows.
%     See ISPC and COMPUTER.
if ispc && isequal(get(hObject,'BackgroundColor'),...
    get(0,'defaultUicontrolBackgroundColor'))
    set(hObject,'BackgroundColor','white');
end

% -----
function tutorial_Callback(hObject, eventdata, handles)
% hObject    handle to tutorial (see GCBO)
% eventdata  reserved - to be defined in a future version of MATLAB
% handles    structure with handles and user data (see GUIDATA)

    tutorial;

% --- Executes during object creation, after setting all properties.
function earAxes_CreateFcn(hObject, eventdata, handles)
% hObject    handle to earAxes (see GCBO)
% eventdata  reserved - to be defined in a future version of MATLAB
% handles    empty - handles not created until after all CreateFcns called

% Hint: place code in OpeningFcn to populate earAxes

% --- Executes during object creation, after setting all properties.
function traceEarButton_CreateFcn(hObject, eventdata, handles)
% hObject    handle to traceEarButton (see GCBO)
% eventdata  reserved - to be defined in a future version of MATLAB
% handles    empty - handles not created until after all CreateFcns called

% --- Executes on key press with focus on traceEarButton
%     and none of its controls.
function traceEarButton_KeyPressFcn(hObject, eventdata, handles)
% hObject    handle to traceEarButton (see GCBO)
% eventdata  structure with the following fields (see UICONTROL)
%     Key: name of the key that was pressed, in lower case

```

```

% Character: character interpretation of the key(s) that was pressed
% Modifier: name(s) of the modifier key(s) (i.e., control, shift)
% handles structure with handles and user data (see GUIDATA)

% --- Executes on mouse press over axes background.
function earAxes_ButtonDownFcn(hObject, eventdata, handles)
% hObject handle to earAxes (see GCBO)
% eventdata reserved - to be defined in a future version of MATLAB
% handles structure with handles and user data (see GUIDATA)

```

## C.10 Funzione *experiment1*

```

function varargout = experiment1(varargin)

% EXPERIMENT1 MATLAB code for experiment1.fig
% EXPERIMENT1, by itself, creates a new EXPERIMENT1 or raises the existing
% singleton*.
%
% H = EXPERIMENT1 returns the handle to a new EXPERIMENT1 or the handle to
% the existing singleton*.
%
% EXPERIMENT1('CALLBACK',hObject,eventData,handles,...) calls the local
% function named CALLBACK in EXPERIMENT1.M with the given input arguments.
%
% EXPERIMENT1('Property','Value',...) creates a new EXPERIMENT1 or raises
% existing singleton*. Starting from the left, property value pairs are
% applied to the GUI before experiment1_OpeningFcn gets called. An
% unrecognized property name or invalid value makes property application
% stop. All inputs are passed to experiment1_OpeningFcn via varargin.

% Begin initialization code - DO NOT EDIT
gui_Singleton = 1;
gui_State = struct('gui_Name',       mfilename, ...
                  'gui_Singleton',  gui_Singleton, ...
                  'gui_OpeningFcn', @experiment1_OpeningFcn, ...
                  'gui_OutputFcn',  @experiment1_OutputFcn, ...
                  'gui_LayoutFcn',  [] , ...
                  'gui_Callback',   []);
if nargin && ischar(varargin{1})
    gui_State.gui_Callback = str2func(varargin{1});
end

if nargout
    [varargout{1:nargout}] = gui_mainfcn(gui_State, varargin{:});
else
    gui_mainfcn(gui_State, varargin{:});
end
% End initialization code - DO NOT EDIT

```

```

% --- Executes just before experiment1 is made visible.
function experiment1_OpeningFcn(hObject, eventdata, handles, varargin)
% This function has no output args, see OutputFcn.
% hObject    handle to figure
% eventdata  reserved - to be defined in a future version of MATLAB
% handles     structure with handles and user data (see GUIDATA)
% varargin    command line arguments to experiment1 (see VARARGIN)

% Choose default command line output for experiment1
handles.output = hObject;

% Update handles structure
guidata(hObject, handles);

% Importa dati condivisi (soggetto ID)
subjectsHandler = getappdata(0,'subjectsHandler');
id = getappdata(subjectsHandler,'sharedID');
repetitions = getappdata(subjectsHandler,'repetitions');
handles.repetitions = repetitions;
pauseInterval = getappdata(subjectsHandler,'pauseInterval');
handles.pauseInterval = pauseInterval;
handles.nr_pause=1;
handles.counter=0;
handles.pauseDone=0;

% Carica la lista dei soggetti ed estrai i dati
subjectsStruct = load('subjects');
subjects = subjectsStruct.subjects;
handles.subjects = subjects;

if ( isempty(subjects(:,1)) ),
    disp('Database empty. Abort. ');
    close();
end

numOfRecords = length(subjects(:,1));
for i = 1:numOfRecords,
    if ( subjects{i,6} == id )
        lastName = subjects{i,1};
        firstName = subjects{i,2};
        gender = subjects{i,3};
        age = subjects{i,4};
        [~, earDataFile, ~] = fileparts(subjects{i,5});
        earDataFile = [earDataFile '.mat'];
        cd img;
        earDataFile = load(earDataFile);
        cd ..;
        break;
    end
end

```

```

end

% Estrai il valore F0 dai dati caricati.
F0 = earDataFile.F0;
clear earDataFile.F0;

% Prepara arrays con i dati del test, calcola il numero del test e
% prepara la presentazione random degli stimoli.
load('positions.mat');
load('conditions.mat');
[numPositions, ~] = size(positions);
handles.numOfTests = size(conditions,1) * numPositions * repetitions;
handles.condTest = zeros(2,size(conditions,1) * numPositions);
handles.blockTest=zeros(repetitions*2,size(conditions,1) * numPositions);
for k=1:repetitions
    pos = 1;
    for i = 1:size(conditions,1)
        for j = 1:numPositions
            handles.condTest(1,pos) = i;
            handles.condTest(2,pos)=positions(j,1);
            pos = pos + 1;
        end
    end
    randomOrder = randperm(size(conditions,1) * numPositions);
    handles.condTest = handles.condTest(:,randomOrder);
    handles.blockTest([k*2-1,k*2],:)=handles.condTest([1,2],:);
end
blockTest=zeros(repetitions*2,handles.numOfTests);
blockTest = handles.blockTest(:,:);

handles.currentTest = 0;
handles.id = id;
handles.experimentResults = cell(handles.numOfTests, 3);

% Visualizza informazioni sul soggetto
handles.subjectData = [lastName ' ' firstName ' ' gender ' ' age ];
set(handles.subjectDataText, 'String', handles.subjectData);
set(handles.currentTestText, 'String', [int2str(handles.currentTest)...
    ' of ' int2str(handles.numOfTests)]);
axes(handles.elevationAxes);
img = imread('img/newhead_green.jpg');
imagesc([0, 1], [0, 1], flipdim(img,1));
set(gca,'YDir','normal');
axis image off;
set(hObject, 'Visible', 'on');
create_subject(F0,id);
save(['subjects\' ['subject_' int2str(id) '\'] ['blockTest' '.mat'] ],...
'blockTest');
set(handles.listenButton,'Enable','on');
guidata(hObject, handles);

```

```

% UIWAIT makes experiment1 wait for user response (see UIRESUME)
% uiwait(handles.figure1);

% --- Outputs from this function are returned to the command line.
function varargout = experiment1_OutputFcn(hObject, eventdata, handles)
% varargout cell array for returning output args (see VARARGOUT);
% hObject handle to figure
% eventdata reserved - to be defined in a future version of MATLAB
% handles structure with handles and user data (see GUIDATA)

% Get default command line output from handles structure
varargout{1} = handles.output;

% --- Executes on button press in listenButton.
function listenButton_Callback(hObject, eventdata, handles)
% hObject handle to listenButton (see GCBO)
% eventdata reserved - to be defined in a future version of MATLAB
% handles structure with handles and user data (see GUIDATA)

% Aggiornamento risultati a fine di ogni stimolo
if ( handles.currentTest ~= 0 )
    handles.endTime = round(etime(clock,handles.startTime) * 1000);
    handles.experimentResults{handles.currentTest, 1} = ...
        handles.chosenElevation;
    handles.experimentResults{handles.currentTest, 2}= ...
        handles.blockTest(2*handles.nr_pause,handles.counter);
    handles.experimentResults{handles.currentTest, 3}= ...
        handles.blockTest(2*handles.nr_pause-1,handles.counter);
end

handles.currentTest = handles.currentTest + 1;
handles.counter=handles.counter+1;
set(handles.currentTestText, 'String', [int2str(handles.currentTest)
    ' of ' int2str(handles.numOfTests)]);
guidata(hObject, handles);

% Fine esperimento e memorizzazione dei dati finali
if ( handles.currentTest > handles.numOfTests )
    cd results;
    results = handles.experimentResults;
    save([int2str(handles.id) '.mat'],'results');
    cd ..;
    uiwait(msgbox('Experiment completed and saved.', ...
        'Completed','none'));
    guidata(hObject, handles);
    close();
    return;

```

```

end

if ( handles.currentTest - 1 ~= 0 )
    set(handles.listenButton,'Enable','off');
    delete(handles.elevationPlotHandler);
else
    set(handles.listenButton,'Enable','off');
end

% Pausa e memorizzazione dei risultati parziali
if ( handles.currentTest-1 ~= 0 && mod(handles.currentTest-1, ...
    handles.pauseInterval) == 0 )
    cd results;
    results = handles.experimentResults;
    save([int2str(handles.id) '_autosave.mat'],'results');
    cd ..;
    set(handles.statusText,'String',...
    'Take a short break before continuing..');
    pause(2);
    countdown(0,30);
    pause(32);
    box = msgbox('Click ok when ready: the next trial will start.', ...
        'Take a break!','help');
    uiwait(box);
    handles.nr_pause=handles.nr_pause+1;
    handles.counter=1;
end

% Presentazione degli stimoli
handles.startTime = clock;
set(handles.listenButton,'Enable','off');
set(handles.continueButton,'Enable','off');
pause(0.5);
set(handles.statusText, 'BackgroundColor', 'yellow');
set(handles.statusText,'String','Sound playing ...');
pause(0.3);
y=wavread(['subjects\' ['subject_' int2str(handles.id)] '\stimoli_wav\'...
    ['cond_' int2str(handles.blockTest(2*handles.nr_pause-1,...
    handles.counter)) '_elev_'...
    int2str(handles.blockTest(2*handles.nr_pause,handles.counter))]);
wavplay(y,48000);

% Scelta elevazione da parte del soggetto
clear y;
set(handles.statusText, 'BackgroundColor', 'green');
set(handles.statusText,'String','Choose elevation inside the red circle.');
```

```

pause(0.2);
alreadyChosenElevation = false;

while ( true )
    try

```

```

        [xChosen,yChosen] = ginput(1);
    catch ME
        return
    end

    if ( xChosen < 0 || xChosen > 1 || yChosen < 0 || yChosen > 1 )
        uiwait(msgbox('Choose a point inside the image.', 'Error','warn'));
        continue;
    end

    if ( gca == handles.elevationAxes )
        xElevationChosen = xChosen;
        yElevationChosen = yChosen;
        hold on;
        if ( alreadyChosenElevation == true )
            delete(elevationPlotHandler);
            elevationPlotHandler = plot(xElevationChosen, ...
                yElevationChosen, '*');
        else
            elevationPlotHandler = plot(xElevationChosen, ...
                yElevationChosen, '*');
            alreadyChosenElevation= true;
        end
    end

    if ( alreadyChosenElevation )
        break;
    end

end

% Update con le risposte del soggetto
handles.chosenElevation = getAngle(0.5, 0.5, 1, 0.5, xElevationChosen,...
yElevationChosen);
if ( xElevationChosen > 0.5 && yElevationChosen < 0.5 )
    handles.chosenElevation = -handles.chosenElevation;
end
if ( xElevationChosen < 0.5 && yElevationChosen < 0.5 )
    if ( handles.chosenElevation ~= 180 )
        handles.chosenElevation = 360 - handles.chosenElevation;
    end
end

% Presentazione nuovo stimolo
handles.elevationPlotHandler = elevationPlotHandler;
set(handles.statusText, 'BackgroundColor', 'red');
set(handles.statusText,'String', 'Loading next sound');
set(handles.listenButton,'Enable','off');
guidata(hObject, handles);
pause(0.5);
listenButton_Callback(hObject, eventdata, handles);

```

```
function output = getAngle(x1,y1, x2,y2, x3,y3)

% Calcola angolo in gradi
temp = atan2( abs((x2-x1)*(y3-y1)-(y2-y1)*(x3-x1)), ...
    (x2-x1)*(x3-x1)+(y2-y1)*(y3-y1) );
output = radtodeg(temp);

function out = radtodeg(in)

out = 180.*in/pi;

function subjectDataText_Callback(hObject, eventdata, handles)
% hObject    handle to subjectDataText (see GCBO)
% eventdata  reserved - to be defined in a future version of MATLAB
% handles    structure with handles and user data (see GUIDATA)

% Hints: get(hObject,'String') returns contents of subjectDataText as text
%        str2double(get(hObject,'String')) returns contents of
%        subjectDataText as a double

% --- Executes during object creation, after setting all properties.
function subjectDataText_CreateFcn(hObject, eventdata, handles)
% hObject    handle to subjectDataText (see GCBO)
% eventdata  reserved - to be defined in a future version of MATLAB
% handles    empty - handles not created until after all CreateFcns called

% Hint: edit controls usually have a white background on Windows.
%        See ISPC and COMPUTER.
if ispc && isequal(get(hObject,'BackgroundColor'),...
    get(0,'defaultUiControlBackgroundColor'))
    set(hObject,'BackgroundColor','white');
end

% --- Executes on button press in continueButton.
function continueButton_Callback(hObject, eventdata, handles)
% hObject    handle to continueButton (see GCBO)
% eventdata  reserved - to be defined in a future version of MATLAB
% handles    structure with handles and user data (see GUIDATA)

guidata(hObject, handles);

function currentTestText_Callback(hObject, eventdata, handles)
% hObject    handle to currentTestText (see GCBO)
% eventdata  reserved - to be defined in a future version of MATLAB
% handles    structure with handles and user data (see GUIDATA)

% Hints: get(hObject,'String') returns contents of currentTestText as text
%        str2double(get(hObject,'String')) returns contents of
```

```

%         currentTestText as a double

% --- Executes during object creation, after setting all properties.
function currentTestText_CreateFcn(hObject, eventdata, handles)
% hObject     handle to currentTestText (see GCBO)
% eventdata   reserved - to be defined in a future version of MATLAB
% handles     empty - handles not created until after all CreateFcns called

% Hint: edit controls usually have a white background on Windows.
%         See ISPC and COMPUTER.
if ispc && isequal(get(hObject,'BackgroundColor'),...
    get(0,'defaultUicontrolBackgroundColor'))
    set(hObject,'BackgroundColor','white');
end

% --- Executes when user attempts to close figure1.
function figure1_CloseRequestFcn(hObject, eventdata, handles)
% hObject     handle to figure1 (see GCBO)
% eventdata   reserved - to be defined in a future version of MATLAB
% handles     structure with handles and user data (see GUIDATA)

% Hint: delete(hObject) closes the figure
if ( handles.currentTest > 1 && handles.currentTest <= handles.numOfTests )
    choice = questdlg('Save partial results?', ...
        'Save experiment', ...
        'Yes', 'No', 'Cancel', 'Cancel');
    switch choice
        case 'Yes'
            cd results;
            results = handles.experimentResults;
            save([int2str(handles.id) '.mat'],'results');
            cd ..;
        case 'No'

        case 'Cancel'
            return;
    end
end
delete(hObject);

```

## C.11 Funzione MLP

```

function MLP()

clear all
clc
load('mpa_hda200.mat'); %carica RETSPLs delle cuffie HDA200
global freq;

```

```

tracking = 'MLP';
s = MLParameters; %obtaining the ML procedure parameters
mkdir('results',s.name);
fr=[1000 500 1000 2000 3000 4000 6000 8000 500 250 125]; % freq. da testare
freq=1000;
count=0;
results=zeros(length(fr),s.nblocks+1,2);
if ~isempty(s)
  for ear=1:2 % ciclo per l'orecchio destro e sinistro
    for u=1:length(fr)
      freq=fr(u);
      if u==4
        t=results(u-1,s.nblocks+1)-results(1,s.nblocks+1);
        if abs(t)>10 %condizione di controllo sulla frequenza di 1000 Hz
          error('Error. It is necessary to stop testing.The first ...
            and second threshold at 1000 Hz do not agree within 5 dB..');
        end
      end
    end
    for i=1:s.nblocks %running the blocks
      clc
      disp('Ear:');
      if ear==1
        disp('Right');
      else
        disp('Left');
      end
      disp(' ');
      disp('Tone Frequency:');
      disp(freq);
      input('Press return to begin the block ', 's');
      clc
      fprintf ('Block number %1.0f\n\n', i);
      pause(1)

      for j = 1:s.ntrials %running the trials
        pause(0.5)
        count = count+1;

        if j>1 && ~s.nafc && rand<s.catchtrial
          s.StimulusLevel(j)=s.standard;
        end

        fprintf('%1.0f] ', j);
        fun = [s.experiment,'(' num2str(s.standard) ...
          ', s.StimulusLevel(j) ', num2str( s.nafc) ')'];

        [CorrectAnswer, Question] = eval(fun);

        s.SubjectAccuracy (j)= EvaluateAnswer(CorrectAnswer, ...
          s.feedback,Question,j); %evaluating the subject answer
      end
    end
  end
end

```

```

while j == 1 && s.SubjectAccuracy (j) == 0 && s.repeatft
    fprintf('%1.0f] ', j);
    fun = [s.experiment,' (' num2str(s.standard)...
        ', s.StimulusLevel(j) ', num2str( s.nafc)  ')'];
    [CorrectAnswer, Question] = eval(fun);
    s.SubjectAccuracy (j)= EvaluateAnswer(CorrectAnswer,...
        s.feedback,Question,j); %evaluating the subject answer
end
%estimating the stimuli intensity for the successive trial
[s.TemporaryThreshold(j),s.FA(j)]=FindThreshold(s.p_target,...
    s.StimulusLevel(1:j), ...
    s.SubjectAccuracy (1:j), s.midpoints,...
    s.beta, s.gamma, s.lambda);

if j < s.ntrials
    s.StimulusLevel(j+1)=s.TemporaryThreshold(j);
elseif j == s.ntrials
    fprintf('Threshold equal to %1.2f\n',...
        s.TemporaryThreshold(j));
    fprintf('Press return to continue\n');
    pause
end

s.MATSAVEDATA(count,:) = [i,j,s.StimulusLevel(j),s.FA(j),...
    s.SubjectAccuracy (j),s.TemporaryThreshold(j)];
end %end trials
fprintf (' \nBLOCK ENDED\n');
results(u,i,ear)=s.TemporaryThreshold(j);

pause(2)
end %end blocks
results(u,s.nblocks+1,ear)=mean(results(u,1:s.nblocks,ear));

end

end

save(['results\' s.name '\\' s.name, '.mat'], 'results');
fprintf (' \nEXPERIMENT ENDED\n\n');
if s.SaveResults
    WriteDataFile(tracking,s.fileout,s.MATSAVEDATA,...
        s.nsub,s.name,s.gender,s.age,s.note);
end

% preparazione, stampa e salvataggio dell'audiogramma
hearing_level_dx=(results((3:end),s.nblocks+1,1)+78.6+10)-mpa_hda200(:,2);
hearing_level_sx=(results((3:end),s.nblocks+1,2)+78.6+10)-mpa_hda200(:,2);
hearing_level_x=[125 250 500 1000 2000 3000 4000 6000 8000];
hearing_level_y_dx=[hearing_level_dx(end) hearing_level_dx(end-1)...
    hearing_level_dx(end-2) hearing_level_dx(1) hearing_level_dx(2)...
    hearing_level_dx(3) hearing_level_dx(4)...

```

```
        hearing_level_dx(5) hearing_level_dx(6)];
hearing_level_y_sx=[hearing_level_sx(end) hearing_level_sx(end-1)...
    hearing_level_sx(end-2) hearing_level_sx(1) hearing_level_sx(2)...
    hearing_level_sx(3) hearing_level_sx(4)...
    hearing_level_sx(5) hearing_level_sx(6)];
hfig=figure;
semilogx(hearing_level_x,hearing_level_y_dx, '--or',hearing_level_x,...
    hearing_level_y_sx, '--ob',hearing_level_x,...
    0*ones(size(hearing_level_x),'-k','linewidth',3);
set(gca,'XTick',hearing_level_x);
set(gca,'XTickLabel',[125;250;500;1000;2000;3000;4000;6000;8000]);
set(gca,'XLim',[125 8000]);
set(gca,'YDir','reverse');
xlabel('Frequency (Hz)');
ylabel('Level (dbHL)');
title('Audiogram','FontSize',18);
hleg1 = legend('ear right','ear left','0 dbHL Reference',...
    'Location','SouthEast');
grid on;

saveas(hfig,['results\' s.name '\' s.name'],'jpg');
saveas(hfig,['results\' s.name '\' s.name'],'fig');
else
    clear all
end
```



# Bibliografia

- [1] ISO 389-8. Acoustics-reference zero for the calibration of audiometric equipment- part 8: Reference equivalent threshold sound pressure level for pure tones and circumaural earphones, 2004.
- [2] V. Ralph Algazi, Richard O. Duda, Ramani Duraiswami, Nail A. Gumerov, and Zhihui Tang. Approximating the head-related transfer function using simple geometric models of the head and torso. *The Journal of the Acoustical Society of America*, 112(5):2053–2064, 2002.
- [3] Ana Alves-Pinto and Enrique A. Lopez-Poveda. Detection of high-frequency spectral notches as a function of level. *The Journal of the Acoustical Society of America*, 118:2458–2469, 2005.
- [4] Futoshi Asano, Yoiti Suzuki, and Toshio Sone. Role of spectral cues in median plane localization. *The Journal of the Acoustical Society of America*, 88(1):159–168, 1990.
- [5] Dwight W. Batteau. The role of the pinna in human localization. *Proceedings of the Royal Society of London. Series B. Biological Sciences*, 168(1011):158–180, 1967.
- [6] Alberto Bedin. Head Related Transfer Function selection techniques applied to multimodal environments for spatial cognition. Master’s thesis, Università di Padova, Facoltà di Ingegneria, 2013.
- [7] Durand R. Begault, Elizabeth M. Wenzel, and Mark R. Anderson. Direct comparison of the impact of head tracking, reverberation, and individualized head-related transfer functions on the spatial perception of a virtual speech source. *Journal of the Audio Engineering Society*, 49(10):904–916, 2001.
- [8] Jens Blauert. *Spatial hearing: the psychophysics of human sound localization*. MIT press, 1997.
- [9] C. Phillip Brown and Richard O. Duda. A structural model for binaural sound synthesis. *Speech and Audio Processing, IEEE Transactions on*, 6(5):476–488, 1998.
- [10] R. Buchlein. The audibility of frequency response irregularities. *Journal of the Audio Engineering Society*, 29(3):126–131, March 1981.

- 
- [11] Ana Isabel Alves Pinto Lopes da Silva. *Psychophysical and physiological assessment of the representation of high-frequency spectral notches in the auditory nerve*. PhD thesis, 2007.
- [12] David Doukhan and Anne Sédès. CW\_binaural~: A binaural synthesis external for Pure Data. In *Proc. PdCon09-3rd Puredata Int. Convention*, 2009.
- [13] Michael J. Evans, James A. S. Angus, and Anthony I. Tew. Analyzing head-related transfer function measurements using surface spherical harmonics. *The Journal of the Acoustical Society of America*, 104(4):2400–2411, 1998.
- [14] Armando Faller II, K. J. and Barreto, Navarun Gupta, and Naphtali Rishe. Time and frequency decomposition of head-related impulse responses for the development of customizable spatial audio models. *WSEAS Transactions on Signal Processing*, 2(11):1465–1472, 2006.
- [15] Michele Geronazzo, Simone Spagnol, and Federico Avanzini. Estimation and modeling of pinna-related transfer functions. In *Proc. of the 13th Int. Conference on Digital Audio Effects (DAFx-10)*, pages 6–10, 2010.
- [16] Michele Geronazzo, Simone Spagnol, and Federico Avanzini. A head-related transfer function model for real-time customized 3-d sound rendering. In *Signal-Image Technology and Internet-Based Systems (SITIS), 2011 Seventh International Conference on*, pages 174–179. IEEE, 2011.
- [17] Massimo Grassi and Alessandro Soranzo. MLP: A Matlab toolbox for rapid and reliable auditory threshold estimation. *Behavior Research Methods*, 41(1):20–28, 2009.
- [18] David M. Green. A maximum-likelihood method for estimating thresholds in a yes–no task. *The Journal of the Acoustical Society of America*, 93:2096–2105, 1993.
- [19] Lyn Haber, Ralph Norman Haber, Suzanna Penningroth, Kevin Novak, and Hilary Radgowski. Comparison of nine methods of indicating the direction to objects: Data from blind adults. *Perception*, 22:35–37, 1993.
- [20] Jack Hebrank and D. Wright. Spectral cues used in the localization of sound sources on the median plane. *The Journal of the Acoustical Society of America*, 56:1829–1834, 1974.
- [21] Jack Hebrank and D. Wright. Are two ears necessary for localization of sound sources on the median plane? *The Journal of the Acoustical Society of America*, 56(3):935–938, 2005.
- [22] Kazuhiro Iida, Motokuni Itoh, Atsue Itagaki, and Masayuki Morimoto. Median plane localization using a parametric model of the head-related transfer function based on spectral cues. *Applied Acoustics*, 68(8):835–850, 2007.

- [23] Brian F. G. Katz. Boundary element method calculation of individual head-related transfer function. I. Rigid model calculation. *The Journal of the Acoustical Society of America*, 110(5):2440–2448, 2001.
- [24] Doris J. Kistler and Frederic L. Wightman. A model of head-related transfer functions based on principal components analysis and minimum-phase reconstruction. *The Journal of the Acoustical Society of America*, 91(3):1637–1647, 1992.
- [25] Vincenzo Lombardo e Andrea Valle. *Audio e Multimedia*. 2005.
- [26] Ewan A. Macpherson and John C. Middlebrooks. Vertical-plane sound localization probed with ripple-spectrum noise. *The Journal of the Acoustical Society of America*, 114:430–445, 2003.
- [27] Brian C. J. Moore, Simon R. Oldfield, and Gary J. Dooley. Detection and discrimination of spectral peaks and notches at 1 and 8 kHz. *The Journal of the Acoustical Society of America*, 85:820–836, 1989.
- [28] Masayuki Morimoto. The contribution of two ears to the perception of vertical angle in sagittal planes. *The Journal of the Acoustical Society of America*, 109:1596–1603, 2001.
- [29] P. W. Poon and John F. Brugge. Sensitivity of auditory nerve fibers to spectral notches. *Journal of neurophysiology*, 70(2):655–666, 1993.
- [30] Vikas C. Raykar, Ramani Duraiswami, and B. Yegnanarayana. Extracting the frequencies of the pinna spectral notches in measured head related impulse responses. *The Journal of the Acoustical Society of America*, 118:364–374, 2005.
- [31] Patrick Satarzadeh, V. Ralph Algazi, and Richard O. Duda. Physical and filter pinna models based on anthropometry. In *Proc. 122nd Convention of the Audio Engineering Society*, pages 5–8, 2007.
- [32] Giovanni Schiesaro. Registrazione ed analisi di Headphone Impulse Response per la spazializzazione del suono. Bachelor’s thesis, Università di Padova, Facoltà di Ingegneria, 2013.
- [33] E. A. G. Shaw. External ear response and sound localization. *Localization of sound: Theory and applications*, pages 30–41, 1982.
- [34] E. A. G. Shaw, T. R. Anderson, and R. H. Gilkey. Binaural and spatial hearing in real and virtual environments, chapter acoustical features of human ear. *R. H. Gilkey and T. R. Anderson, Lawrence Erlbaum Associates, Mahwah, NJ, USA*, pages 25–47, 1997.
- [35] E. A. G. Shaw and R. Teranishi. Sound pressure generated in an external-ear replica and real human ears by a nearby point source. *The Journal of the Acoustical Society of America*, 44(1):240–249, 2005.

- 
- [36] Simone Spagnol, Michele Geronazzo, and Federico Avanzini. Fitting pinna-related transfer functions to anthropometry for binaural sound rendering. In *Multimedia Signal Processing (MMSP), 2010 IEEE International Workshop on*, pages 194–199. IEEE, 2010.
- [37] Simone Spagnol, Michele Geronazzo, and Federico Avanzini. On the relation between pinna reflection patterns and head-related transfer function features. *Audio, Speech, and Language Processing, IEEE Transactions on*, 21(3):508–519, 2013.
- [38] Aurelio Uncini. *Audio digitale*. 2006.
- [39] Anthony J. Watkins. Psychoacoustical aspects of synthesized vertical locale cues. *The Journal of the Acoustical Society of America*, 63:1152–1165, 1978.
- [40] Charles S. Watson, John R. Franks, and Donald C. Hood. Detection of tones in the absence of external masking noise. I. Effects of signal intensity and signal frequency. *The Journal of the Acoustical Society of America*, 52:633–643, 1972.
- [41] Elizabeth M. Wenzel, Marianne Arruda, Doris J. Kistler, and Frederic L. Wightman. Localization using nonindividualized head-related transfer functions. *The Journal of the Acoustical Society of America*, 94(1):111–123, 1993.
- [42] Frederic L. Wightman and Doris J. Kistler. Headphone simulation of free-field listening. II: Psychophysical validation. *The Journal of the Acoustical Society of America*, 85:868–878, 1989.