

MISURE DI HRTF SU UNA TESTA KEMAR CON PADIGLIONI AURICOLARI INTERCAMBIABILI

Simone Spagnol

Dip. di Ingegneria dell'Informazione
Università degli Studi di Padova
Padova, Italia
spagnols@dei.unipd.it

Kristján Bjarki Purkhús, Sverrir Karl Björnsson, Rúnar Unnthórsson

School of Engineering and Natural Sciences
University of Iceland
Reykjavík, Islanda
runson@hi.is

SOMMARIO

Questo articolo ha come oggetto l'acquisizione di un insieme di misure di Head-Related Transfer Function (HRTF) sviluppata presso la University of Iceland. Le HRTF, campionate su una griglia sferica di 1513 posizioni, sono state misurate con una testa KEMAR dotata di volta in volta di uno tra 20 diversi padiglioni auricolari sinistri artificiali, ottenuti artigianalmente da accurati calchi di teste umane. Le valutazioni preliminari qui contenute suggeriscono la robustezza e la riproducibilità delle misure acustiche, così come la loro fedeltà rispetto alle HRTF KEMAR di riferimento. L'obiettivo della presente raccolta dati è di fornire un input affidabile per metodologie di determinazione di pattern spettrali individuali a partire da rappresentazioni 3D dell'orecchio e, più in generale, di incoraggiare studi sulla relazione tra HRTF e dati antropometrici tramite tecniche di apprendimento automatico.

1. INTRODUZIONE

La maggior parte delle tecniche per la resa binaurale del suono fa affidamento sull'utilizzo di Head-Related Transfer Function (HRTF), ovvero filtri che catturano gli effetti acustici del corpo umano [1]. L'uso di HRTF permette una simulazione realistica del segnale che giunge all'ingresso del canale uditivo in funzione della posizione della sorgente sonora nello spazio (azimut, elevazione e distanza). La resa ideale, ovvero quella che meglio approssima l'ascolto spaziale reale, richiede l'utilizzo di HRTF individuali misurate acusticamente sull'utente. Elaborando un segnale arbitrario tramite una coppia di HRTF individuali, una per canale, l'utente può fruire di autentiche esperienze sonore tridimensionali [2].

Sfortunatamente, ottenere HRTF individuali accurate comporta tutt'oggi avere accesso a strumentazione di laboratorio generalmente costosa e applicare procedure di registrazione invasive e/o faticose per l'utente [3]. Tale è la ragione per cui in pratica viene spesso favorito l'utilizzo di HRTF non individuali, misurate su manichini antropomorfi o su individui terzi. Diversi set di HRTF sono disponibili online; di questi, alcuni sono misurati su teste artificiali come

la KEMAR¹ mentre altri sono inclusi in database di misure individuali [4]. Molti database sono stati recentemente unificati in un formato comune di HRTF conosciuto come *Spatially Oriented Format for Acoustics* (SOFA).²

L'ovvio svantaggio delle HRTF non individuali è che difficilmente sono relazionabili all'antropometria dell'utente, in particolare quella del padiglione auricolare (detto anche *pinna*, termine che utilizzeremo in questo articolo). Nel momento in cui vengono usate per il rendering binaurale del suono, ciò si traduce in probabili e frequenti errori di localizzazione, quali inversioni (*front/back confusion*), errata percezione dell'elevazione, e localizzazione all'interno della testa [5]. Nonostante adattarsi a una copia di HRTF non individuali sia possibile, questo processo richiede un tempo di apprendimento significativo [6].

Per tali ragioni, lo studio della relazione tra caratteristiche della pinna (forma, dimensioni, orientazione) ed HRTF rappresenta un passo fondamentale verso la comprensione dei meccanismi acustici sottintesi. Ciononostante, la precedente letteratura sulla relazione tra effetti acustici ed antropometria - tra cui applicazioni di metodi di regressione su database di HRTF [7, 8, 9, 10] - hanno prodotto risultati controversi, evidenziando come molte di queste relazioni non siano tuttora comprese appieno. Una delle possibili concause di tali ambiguità può essere individuata nel fattore umano, intendendo con questo problematiche quali il difficile controllo del posizionamento relativo tra microfono e orecchio reale, possibili movimenti del soggetto durante le misurazioni, e condizioni di misura non ideali. Tali problematiche possono dare luogo a misure poco accurate e non riproducibili, come nel caso di un precedente lavoro del primo autore riguardante la misurazione di HRTF su un campione di 25 soggetti umani [11]. D'altro canto, anche l'utilizzo di teste artificiali può dare adito ad errori: misure effettuate sulla stessa testa artificiale, infatti, possono addirittura risultare in set di HRTF molto diversi tra loro, a seconda del sistema di misurazione utilizzato [12].

In questo articolo presentiamo il design, l'implementazione e l'analisi preliminare di un nuovo set di misure di HRTF acquisite con una testa KEMAR dotata di pinne intercambiabili, con l'obiettivo di fornire un input affidabile per metodologie di determinazione di pattern spettrali individuali a partire da rappresentazioni 3D dell'utente [13]. In particolare, descriviamo la messa a punto un procedimento

Copyright: ©2018 Simone Spagnol et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution License 3.0 Unported](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

¹ <https://www.gras.dk/>

² <http://www.sofaconventions.org/>



Figura 1. Il sistema di misurazione di HRTF utilizzato.

artigianale per creare pinne artificiali per la testa KEMAR da calchi di teste umane, poi applicato a un campione di 20 pinne sinistre. Queste sono quindi state utilizzate come unica variabile antropometrica di una serie di misure di HRTF. Pur limitati dalle risorse disponibili, abbiamo progettato il setup di misurazione per essere quanto più accurato e stabile possibile, al fine di garantire affidabilità e riproducibilità delle relative misure acustiche.

2. METODOLOGIA

2.1 Sistema di misurazione

Le HRTF presentate in questo lavoro sono state acquisite automaticamente tramite il sistema di misurazione rappresentato in Figura 1. Esso consiste di un manichino KEMAR fissato a un supporto cilindrico rotante e di un altoparlante Genelec 8020CPM-6 fissato a un braccio rotante a forma di L. La configurazione KEMAR utilizzata è la 45BB-4 con pinne antropometriche standard (durezza 35 Shore-OO, GRAS KB5000/KB5001) e microfoni a pressione da mezzo pollice (GRAS 40AO) situati all'ingresso dei canali. I due canali audio sono collegati a una scheda RME Fireface 802 collegata a sua volta a un PC che esegue il software MATLAB.

La distanza tra il tweeter dell'altoparlante e il centro della testa è stata fissata alla costante di 1 m, indipendentemente dall'orientazione relativa di manichino e braccio rotante. In fase di misurazione, tale valore di distanza garantisce con sufficiente accuratezza la rappresentazione degli indicatori spettrali di localizzazione nel campo lontano [14]. Grazie all'utilizzo di un sistema di raggi laser indirizzati ai microfoni e alla punta del naso della KEMAR, è stato possibile ottenere un corretto allineamento del sistema di misurazione sui tre assi. Tali scelte sperimentali permettono di ottenere una collezione di HRTF in un range che copre tutti gli azimut ed elevazioni che vanno da -50° a 90° , secondo un sistema di coordinate polari verticali.

La rotazione del manichino e del braccio è gestita indipendentemente da due motori passo-passo (JVL Industri Elektronik MST001A) ad alta coppia (1.2 Nm), controllati attraverso due azionatori digitali (Geckodrive G213V) in modalità full-step (200 passi per rivoluzione, singolo passo di 1.8°). Al fine di aumentare la coppia e la risoluzione angolare del sistema, due riduttori 100 : 1 sono stati installati tra il primo motore e il braccio rotante e tra il secondo motore e il manichino. Di conseguenza, il passo minimo di entrambi i sistemi rotanti è di 0.018° . Inoltre, per ridurre la coppia necessaria sul braccio, un contrappeso di 22 kg è stato applicato all'estremità più corta in modo tale da bilanciare il peso dell'altoparlante e dell'altro segmento del braccio. La comunicazione tra gli azionatori e il PC è anch'essa gestita tramite MATLAB.

Per avere dei riferimenti assoluti di azimut ed elevazione durante le misurazioni, sono stati apposti rispettivamente un marcatore fisso nella parte inferiore del manichino (azimut 0°) e una livella a bolla sul segmento più lungo del braccio (elevazione 0°). Come accorgimento per evitare l'aggrovigliarsi di cavi attorno al supporto del manichino e all'asse del braccio, ogni rivoluzione viene fatta seguire da una controrivoluzione in senso opposto.

2.2 Pinne artificiali

In questa sezione descriviamo la procedura artigianale utilizzata per creare le riproduzioni in silicone di pinne per la KEMAR. Al fine di avere una varietà sufficiente di forme per le misurazioni di HRTF, abbiamo applicato questa procedura alle pinne sinistre di 20 diversi soggetti. Tra questi, oltre alla KEMAR con pinne antropometriche standard, abbiamo considerato 19 riproduzioni in gesso di teste umane, concesse dal Saga Museum di Reykjavík. Tali teste, costruite a mano tra il 2001 e il 2003,³ riproducono ad alta fedeltà i tratti antropometrici di 19 individui islandesi (7 femmine), di età compresa tra 7 e 77 anni al momento della costruzione. Si riportano di seguito le fasi principali del processo di riproduzione delle pinne.

Step 1: Forma preliminare. Applicare due strati di miscela di silicone all'orecchio della testa artificiale, il primo sottile con un pennello (Figura 2a) e il secondo più spesso (mezza tazza) con una spatola, lasciando asciugare ogni strato per un giorno prima dell'applicazione successiva. Onde ridurre la colatura su superfici verticali, aggiungere un paio di gocce di additivo tissotropico al silicone

³ https://www.youtube.com/watch?v=1102tc_bxKI

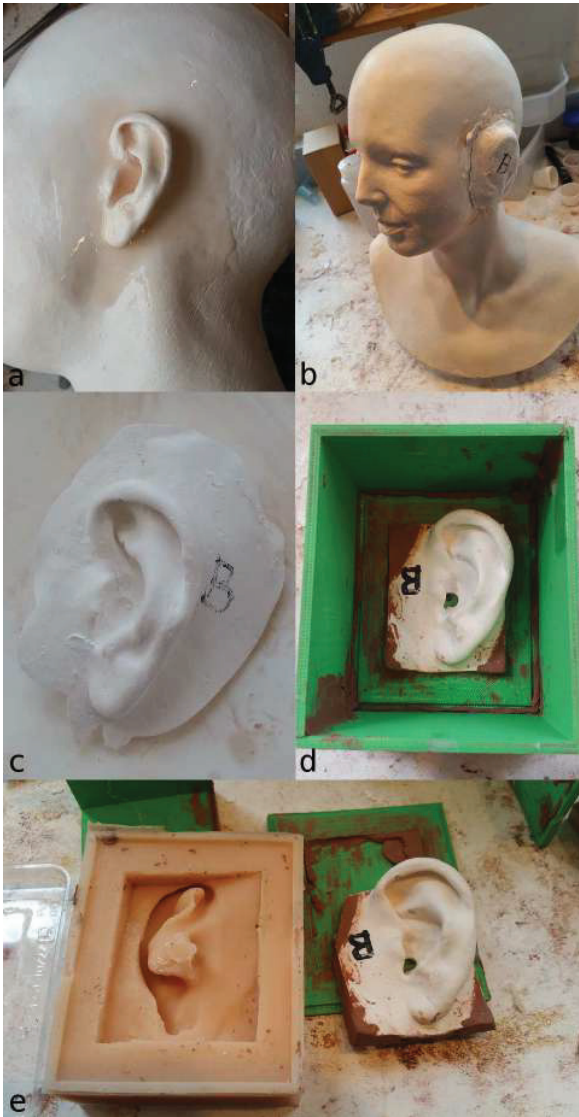


Figura 2. Alcune fasi del processo di costruzione di una forma di pinna per KEMAR.

dello strato spesso. Successivamente, applicare delle strisce di gesso bagnate sulla forma (Figura 2b) per agevolare la rimozione della forma stessa dalla testa artificiale e per evitare che si deformi nello step successivo.

Step 2: Riproduzione in Jesmonite®. Versare una miscela istantanea di mezza tazza d'acqua e due cucchiaini di Jesmonite® nella forma ottenuta nello Step 1, precedentemente posta sopra un tavolo vibrante (onde evitare la formazione di bolle d'aria), e lasciare indurire per 1–2 giorni. Successivamente, rimuovere la pinna in Jesmonite® (Figura 2c) e riempire eventuali fori o irregolarità con dell'argilla. Infine, per adattare la pinna alla guida della KEMAR, forare con un trapano la parte più interna della conca (in corrispondenza del canale uditivo) e smussare la base della pinna con una levigatrice a nastro.

Step 3: Forma definitiva. Riporre la riproduzione in Jesmonite® in una scatola di plastica aperta superiormente e fissarla alla base della scatola con dell'argilla (Figura

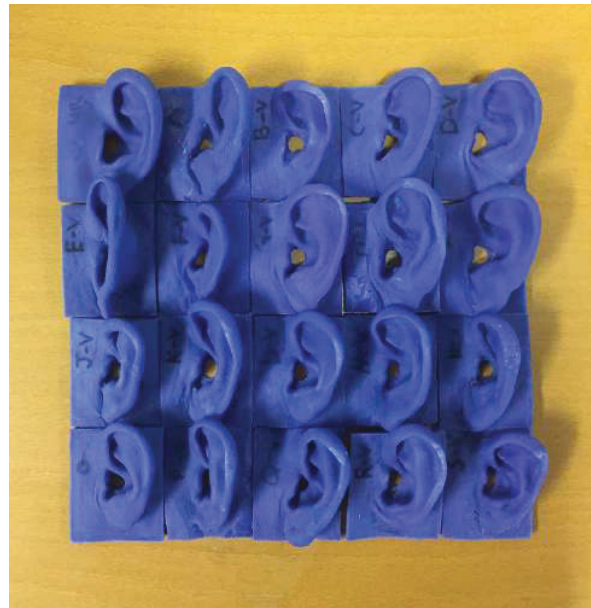


Figura 3. Le 20 riproduzioni di pinne sinistre ottenute con la procedura artigianale.

ra 2d). Successivamente, versare 2 tazze di miscela di silicone all'interno della scatola, precedentemente posta sul tavolo vibrante, e lasciare asciugare per 1 giorno. Infine, rimuovere la scatola e la pinna in Jesmonite® (Figura 2e).

Step 4: Riproduzione in silicone. Versare della cera per auto, in qualità di agente distaccante, nelle parti interne della forma ottenuta nello Step 3, e lasciare asciugare per mezza giornata. Successivamente, versare 1/3 di tazza di una miscela di silicone di durezza 25 Shore-A nella forma, e lasciare asciugare per 1 giorno. Infine, rimuovere la pinna in silicone e tagliare le parti eccedenti sulla base con un coltello. In Figura 3 riportiamo le 20 riproduzioni di pinne sinistre ottenute attraverso la procedura di cui sopra.

2.3 Procedura di misurazione

Le misurazioni sono state condotte durante diversi giorni, al di fuori dell'orario di lavoro (16.00 – 24.00) al fine di ridurre l'inquinamento acustico ambientale, all'interno di una stanza-ufficio dell'edificio Tæknigarður della University of Iceland (vedi Figura 1). Poiché non abbiamo misurato la risposta a campo aperto del sistema, le misurazioni grezze includono la risposta in frequenza dell'altoparlante e dei microfoni così come alcuni effetti della stanza (ad esempio le riflessioni sui muri e sul pavimento, comunque minimizzate in fase di post-processing come spiegato successivamente).

Abbiamo utilizzato il metodo dello sweep logaritmico per registrare le singole risposte acustiche [15]. Il segnale di ingresso, il cui livello è stato tenuto costante durante l'intero programma di misurazione, riproduce tutte le frequenze tra 20 Hz e 20 kHz in 1 s, con una frequenza di campionamento $f_s = 48$ kHz. Il livello SPL medio a 1 kHz per uno stimolo frontale come da misurazione trami-

Tabella 1. Posizioni della sorgente sonora.

Elevazioni [deg]	[-45,45]	[50,70]	[75,85]	90
Passo [deg]	5	15	45	360
No. di azimut	72	24	8	1

te un fonometro di Classe 1 posto al centro del sistema di riferimento è di 82 dB.

La posizione della sorgente sonora è univocamente determinata dall'azimut θ e dall'elevazione ϕ in coordinate polari verticali. Le elevazioni sono campionate uniformemente in passi da 5° tra -45° e 90° , mentre gli azimut sono campionati diversamente a seconda dell'elevazione, come mostrato in Tabella 1. Questa scelta permette di evitare una concentrazione non necessaria di punti circostanti il polo superiore della sfera. Il numero totale di posizioni per misurazione è 1513.

Per ogni sessione di misurazione, le HRTF sono state registrate in ordine decrescente di elevazione, partendo da $\phi = 90^\circ$. Per ogni elevazione, il manichino è stato ruotato del corrispondente passo angolare fino a compiere un giro completo e le risposte allo sweep registrate ad ogni azimut. Alla fine del giro, il manichino è stato ruotato di 360° nel senso opposto e il braccio spostato all'elevazione successiva. Onde evitare la registrazione di rumori dovuti alla rotazione del sistema, delle pause di 0.5 s sono state inserite tra ogni operazione dei motori e la registrazione precedente/successiva. Date tali scelte sperimentali, la durata di una sessione di misurazione è di circa 105 minuti.

Ventitré sessioni di misurazione (3 misurazioni di controllo + 20 di test) sono state programmate in totale. Una pinna antropometrica destra standard (35 Shore-OO) è stata installata sul canale destro della KEMAR durante l'intero programma di misurazione, mentre la pinna sinistra cambiava ad ogni sessione. Nelle prime due sessioni di controllo (*CS1* and *CS2*) abbiamo installato rispettivamente la corrispondente pinna antropometrica sinistra e la sua variante di diversa durezza (55 Shore-OO). Queste sessioni sono state introdotte per controllare eventuali differenze tra materiali e per certificare la fedeltà della nostra riproduzione di pinna KEMAR alle originali. Le successive 20 sessioni sono state invece dedicate alle nostre 20 riproduzioni rappresentate in Figura 3, partendo dalla riproduzione KEMAR e procedendo con le 19 pinne umane (etichettate in ordine alfabetico da *A* a *S*). Nella terza e ultima sessione di controllo (*CS3*) abbiamo misurato di nuovo la KEMAR con entrambe le pinne standard, al fine di verificare la riproducibilità delle nostre misure. Tra ogni sessione di misurazione e la successiva, il sistema di misurazione è stato ricalibrato manualmente alla corretta posizione di partenza grazie ai riferimenti angolari assoluti.

2.4 Post-processing

L'estrazione delle HRTF dalle corrispondenti risposte allo sweep grezze è gestita da uno script di post-processing in MATLAB basato sui seguenti passi. Innanzitutto, identifichiamo il campione di partenza della risposta allo sweep tramite il calcolo della funzione di cross-correlazione $\Psi[n]$

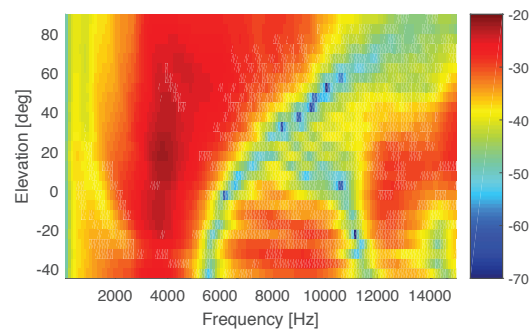


Figura 4. Risposte in ampiezza del set di HRTF *A*, piano mediano (pinna sx: riproduzione, pinna dx: KEMAR 35 Shore-OO).

tra la risposta allo sweep $y[m]$ e il segnale d'ingresso $s[m]$. In particolare, troviamo il valore massimo di correlazione Ψ_M in $\Psi[n]$ ed estraiamo il ritardo n_i del primo campione in $\Psi[n]$ tale che $\Psi[n] > 0.5\Psi_M$. Quindi, il campione di partenza di $y[m]$ corrisponde al ritardo n_j con valore massimo di correlazione all'interno di un intervallo di 21 campioni centrato in $\Psi(n_i)$. Questo escamotage è stato introdotto onde evitare di estrarre direttamente Ψ_M come campione di partenza che, a causa delle condizioni di misura ecoiche, potrebbe corrispondere a una riflessione sulla parete, specialmente nel caso di HRTF controlaterali.

Il passo successivo, coerentemente con il metodo dello sweep logaritmico, è un filtraggio inverso della risposta allo sweep al fine di ottenere la corrispondente risposta all'impulso. Calcoliamo dunque l'inverso $S^{-1}[n]$ dello spettro di riferimento del segnale d'ingresso e applichiamo dei filtri di Butterworth passa-basso e passa-alto del secondo ordine per compensare l'originale assenza di contenuto sotto i 20 Hz e sopra i 20 kHz nel segnale d'ingresso. Quindi, la risposta all'impulso è ottenuta come

$$h[n] = \Re(\mathcal{F}^{-1}(\mathcal{F}(y[n]) * S^{-1}[n])), \quad (1)$$

dove \mathcal{F} e \mathcal{F}^{-1} sono la DFT e la DFT inversa.

Successivamente, ogni risposta all'impulso $h_w[n]$ viene moltiplicata per una finestra di Hann da 128 campioni, allo scopo di rimuovere i contributi di tutte le onde riflesse sul sistema di misurazione e all'interno della stanza registrate approssimativamente oltre i 2.5 ms dal campione di partenza, ottenendo il segnale $h_w[n]$. La HRTF viene quindi infine semplicemente calcolata come risposta in ampiezza della DFT di $h_w[n]$. Come ultimo passo, la HRTF viene sottocampionata di un fattore 100 per ridurre la notevole dimensione dei dati (da 24000 a 240 campioni per HRTF).

3. RISULTATI

In questa sezione riportiamo alcuni risultati preliminari come validazione della fedeltà e riproducibilità del set di HRTF raccolto. Nella Figura 4 viene riportato un esempio di risposta in ampiezza delle HRTF sul piano mediano per una delle pinne artificiali (set *A*). Possiamo riconoscere degli effetti notevoli, quali l'avvallamento tra 1 e 2 kHz dovuto

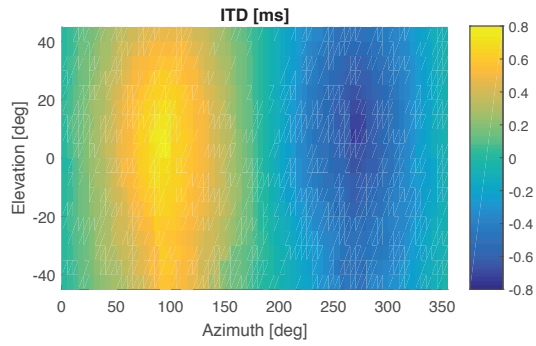


Figura 5. ITD del set di HRTF *A* (pinna sx: riproduzione, pinna dx: KEMAR 35 Shore-OO).

alla riflessione sulla spalla, la risonanza omnidirezionale tra 3 e 4 kHz e il pattern di picchi e notch dipendente dall'elevazione [1], con il più alto numero di notch registrato alle elevazioni inferiori [11]. La Figura 5 mostra le *interaural time differences* (ITD) per la stessa sessione di misurazione, calcolate come differenze temporali tra i campioni di partenza dei canali sinistro e destro di ogni risposta allo sweep. Coerentemente con la precedente letteratura [16], i valori di ITD variano tra zero (piano mediano) e ± 0.8 ms per le direzioni laterali. Si rileva inoltre l'asimmetria tra canale sinistro e destro, che risulta in massimi e minimi della ITD registrati a elevazioni differenti, dovuta alla presenza di due diverse pinne nella misurazione (pinna KEMAR e riproduzione da testa umana).

Per valutare la fedeltà delle nostre misure KEMAR, in Figura 6 riportiamo le misurazioni con la pinna di riferimento (35 Shore-OO, misurata nella sessione *CS1*) sul piano orizzontale e le paragoniamo con precedenti misurazioni della KEMAR alla distanza di 1 m ad opera di Brungart e Rabinowitz (Figura 5 in [17]). Se escludiamo la regione a più bassa frequenza, dove le differenze sono maggiori a causa della diversa configurazione del canale uditivo della KEMAR nei due setup sperimentali, possiamo riconoscere le stesse feature spettrali salienti sopra i 2 kHz, corrispondenti ai cinque punti di riferimento A-E (picchi e notch). Tale risultato qualitativo conferma l'efficacia del nostro setup di misurazione nel riprodurre ad alta fedeltà le misure KEMAR di riferimento.

Al fine di valutare la robustezza e la riproducibilità delle nostre misure, abbiamo calcolato la distorsione spettrale media nell'intervallo 3 – 10 kHz tra ogni set di HRTF e il set di riferimento KEMAR 35 Shore-OO. I risultati sono riportati in Tabella 2. Mentre per il canale destro si evidenzia soltanto rumore di misurazione, come dimostrato dalla bassa distorsione spettrale (< 0.5 dB), i risultati del canale sinistro mostrano differenze notevoli tra gli effetti spettrali delle varie pinne. La distorsione spettrale media tra la pinna KEMAR e ognuna delle 19 riproduzioni di pinne umane è infatti sempre superiore ai 4.5 dB. Per contro, la riproduzione di pinna KEMAR fa registrare una distorsione media sensibilmente più bassa (inferiore a 1 dB), sorprendentemente inferiore anche a quella della pinna KEMAR originale di diversa durezza (55 Shore-OO)

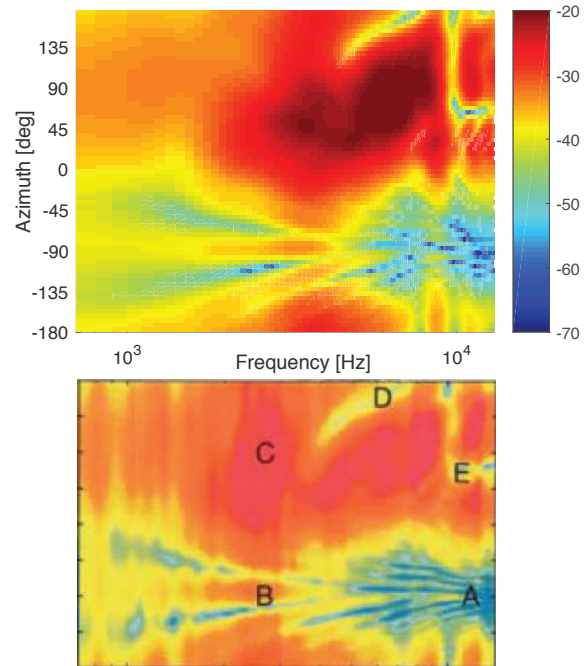


Figura 6. Misurazioni KEMAR sul piano orizzontale nel presente lavoro (riquadro superiore) e in Brungart e Rabinowitz (riquadro inferiore, riprodotto da [17]).

Tabella 2. Distorsione spettrale media [dB] tra il set di HRTF di riferimento (*CS1*, KEMAR 35 Shore-OO) e tutti gli altri set (L: canale sx, R: canale dx).

Set HRTF	L	R
<i>CS2</i>	1.3736	0.1982
<i>KEMAR</i>	0.8407	0.2245
<i>A</i>	4.5978	0.1885
<i>B</i>	5.7394	0.2162
<i>C</i>	6.2859	0.3389
<i>D</i>	4.8688	0.1997
<i>E</i>	5.7817	0.2291
<i>F</i>	5.2341	0.2122
<i>G</i>	5.1931	0.1656
<i>H</i>	4.9951	0.3458
<i>I</i>	5.6352	0.3494
<i>J</i>	4.7819	0.1965
<i>K</i>	5.6527	0.1990
<i>L</i>	4.4466	0.2882
<i>M</i>	4.7295	0.3364
<i>N</i>	5.7147	0.2055
<i>O</i>	5.0146	0.4611
<i>P</i>	5.2498	0.2209
<i>Q</i>	4.7395	0.2648
<i>R</i>	6.5572	0.2806
<i>S</i>	4.5860	0.1877
<i>CS3</i>	0.7902	0.4186

misurata in *CS2*. Questo risultato suggerisce l'accuratezza della nostra procedura di riproduzione di pinne e l'impatto trascurabile della diversa durezza del silicone utilizzato. Infine, i risultati di *CS3* indicano che nonostante il riposizionamento della pinna KEMAR sul canale sinistro introduca della distorsione addizionale, questa non eccede il doppio del valore medio di distorsione dovuta al rumore di misurazione (canale destro).

4. CONCLUSIONI

In questo articolo abbiamo presentato una collezione di HRTF di una testa KEMAR con pinne intercambiabili. Il setup e la procedura di misurazione sono stati descritti nel dettaglio, insieme alle operazioni di post-processing necessarie per estrarre le HRTF finali dalle risposte misurate. I nostri risultati preliminari certificano l'accuratezza e la varietà dei dati collezionati, che auspichiamo possano incoraggiare studi sulla relazione tra HRTF e dati antropometrici tramite tecniche di apprendimento automatico o altre recenti metodologie [18]. La collezione di HRTF, accompagnata da dettagliate scansioni e dati antropometrici delle 20 riproduzioni di pinne, verrà rilasciata online a breve e sarà scaricabile gratuitamente.

5. RINGRAZIAMENTI

Gli autori vorrebbero ringraziare Ernst Backman, proprietario del Saga Museum di Reykjavík, per la sua preziosa assistenza nel processo di realizzazione delle pinne artificiali e per l'accesso al suo laboratorio e strumentazione.

6. BIBLIOGRAFIA

- [1] C. I. Cheng and G. H. Wakefield, "Introduction to head-related transfer functions (HRTFs): Representations of HRTFs in time, frequency, and space," *J. Audio Eng. Soc.*, vol. 49, pp. 231–249, April 2001.
- [2] A. W. Bronkhorst, "Localization of real and virtual sound sources," *J. Acoust. Soc. Am.*, vol. 98, pp. 2542–2553, November 1995.
- [3] B. Xie, *Head-Related Transfer Function and Virtual Auditory Display*. Plantation, FL, USA: J.Ross Publishing, 2nd ed., June 2013.
- [4] V. R. Algazi, R. O. Duda, D. M. Thompson, and C. Avendano, "The CIPIC HRTF database," in *Proc. IEEE Work. Appl. Signal Process., Audio, Acoust.*, (New Paltz, New York, USA), pp. 1–4, October 2001.
- [5] H. Møller, M. F. Sørensen, C. B. Jensen, and D. Hammershøi, "Binaural technique: Do we need individual recordings?," *J. Audio Eng. Soc.*, vol. 44, pp. 451–469, June 1996.
- [6] C. Mendonça, "A review on auditory space adaptations to altered head-related cues," *Front. Neurosci.*, vol. 8, pp. 1–14, July 2014.
- [7] P. Bilinski, J. Ahrens, M. R. P. Thomas, I. J. Tashev, and J. C. Platt, "HRTF magnitude synthesis via sparse representation of anthropometric features," in *Proc. 39th IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP 2014)*, (Firenze, Italy), pp. 4501–4505, May 2014.
- [8] F. Grijalva, L. Martini, S. Goldenstein, and D. Florencio, "Anthropometric-based customization of head-related transfer functions using Isomap in the horizontal plane," in *Proc. 39th IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP 2014)*, (Firenze, Italy), pp. 4506–4510, May 2014.
- [9] S. Spagnol and F. Avanzini, "Frequency estimation of the first pinna notch in head-related transfer functions with a linear anthropometric model," in *Proc. 18th Int. Conf. Digital Audio Effects (DAFx-15)*, (Trondheim, Norway), pp. 231–236, December 2015.
- [10] S. Spagnol, S. Galesso, and F. Avanzini, "Stima di feature spettrali di HRTF mediante modelli antropometrici non lineari per la resa di audio 3D," in *Proc. XXI Colloquium on Music Informatics (XXI CIM)*, (Cagliari, Italy), pp. 129–135, October 2016.
- [11] S. Spagnol, M. Hiipakka, and V. Pulkki, "A single-azimuth pinna-related transfer function database," in *Proc. 14th Int. Conf. Digital Audio Effects (DAFx-11)*, (Paris, France), pp. 209–212, September 2011.
- [12] A. Andreopoulou, D. R. Begault, and B. F. G. Katz, "Inter-laboratory round robin HRTF measurement comparison," *IEEE J. Select. Topics Signal Process.*, vol. 9, pp. 895–906, August 2015.
- [13] S. Spagnol, M. Geronazzo, D. Rocchesso, and F. Avanzini, "Synthetic individual binaural audio delivery by pinna image processing," *Int. J. Pervasive Comput. Comm.*, vol. 10, pp. 239–254, July 2014.
- [14] S. Spagnol, E. Tavazzi, and F. Avanzini, "Distance rendering and perception of nearby virtual sound sources with a near-field filter model," *Appl. Acoust.*, vol. 115, pp. 61–73, January 2017.
- [15] S. Müller and P. Massarani, "Transfer-function measurement with sweeps," *J. Audio Eng. Soc.*, vol. 49, pp. 443–471, June 2001.
- [16] G. F. Kuhn, "Model for the interaural time differences in the azimuthal plane," *J. Acoust. Soc. Am.*, vol. 62, pp. 157–167, July 1977.
- [17] D. S. Brungart and W. M. Rabinowitz, "Auditory localization of nearby sources. Head-related transfer functions," *J. Acoust. Soc. Am.*, vol. 106, pp. 1465–1479, September 1999.
- [18] M. Geronazzo, S. Spagnol, and F. Avanzini, "A modular framework for the analysis and synthesis of head-related transfer functions," in *Proc. 134th Conv. Audio Eng. Soc.*, no. 8882, (Rome, Italy), May 2013.