

CERCETĂRI PRIVIND INTERFAȚA OM-MAȘINĂ CU APLICAȚII ÎN CONDUCEREA PERSOANELOR CU HANDICAP VIZUAL CU AJUTORUL SUNETELOR

Teză destinată obținerii
titlului științific de doctor inginer
la
Universitatea "Politehnica" din Timișoara
în domeniul
INGINERIE ELECTRONICĂ ȘI TELECOMUNICAȚII
de către

Ing. Zoltán Haraszy

Conducător științific:
Referenți științifici:

prof.univ.dr.ing. VIRGIL TIPONUȚ
prof.univ.dr.ing. ADRIAN GRAUR
prof.univ.dr.ing. GAVRIL TODEREAN
prof.univ.dr.ing. IVAN BOGDANOV

Ziua susținerii tezei: 11.11.2011

Seriile Teze de doctorat ale UPT sunt:

- | | |
|------------------------|---|
| 1. Automatică | 7. Inginerie Electronică și Telecomunicații |
| 2. Chimie | 8. Inginerie Industrială |
| 3. Energetică | 9. Inginerie Mecanică |
| 4. Ingineria Chimică | 10. Știința Calculatoarelor |
| 5. Inginerie Civilă | 11. Știința și Ingineria Materialelor |
| 6. Inginerie Electrică | |

Universitatea „Politehnica” din Timișoara a inițiat seriile de mai sus în scopul diseminării expertizei, cunoștințelor și rezultatelor cercetărilor întreprinse în cadrul școlii doctorale a universității. Seriile conțin, potrivit H.B.Ex.S Nr. 14 / 14.07.2006, tezele de doctorat susținute în universitate începând cu 1 octombrie 2006.

Copyright © Editura Politehnica – Timișoara, 2006

Această publicație este supusă prevederilor legii dreptului de autor. Multiplicarea acestei publicații, în mod integral sau în parte, traducerea, tipărirea, reutilizarea ilustrațiilor, expunerea, radiodifuzarea, reproducerea pe microfilme sau în orice altă formă este permisă numai cu respectarea prevederilor Legii române a dreptului de autor în vigoare și permisiunea pentru utilizare obținută în scris din partea Universității „Politehnica” din Timișoara. Toate încălcările acestor drepturi vor fi penalizate potrivit Legii române a drepturilor de autor.

România, 300159 Timișoara, Bd. Republicii 9,
tel. 0256 403823, fax. 0256 403221
e-mail: editura@edipol.upt.ro

Cuvânt înainte

Teza de doctorat a fost elaborată pe parcursul activității mele în cadrul Departamentului de Electronică Aplicată al Facultății de Electronică și Telecomunicații, Universitatea „Politehnica” din Timișoara. Lucrarea cumulează rezultatele studiilor și cercetărilor desfășurate în ultimii patru ani.

Această teză nu ar fi fost posibil fără sfaturile semnificative, stimularea și încurajarea mai multor persoane. În primul rând, mulțumiri deosebite i se cuvin conducătorului de doctorat Prof.Dr.Ing. Virgil Tiponuț pentru îndrumarea și coordonarea continuă, atență la realizarea acestei lucrări, dar și pentru însuflarea interesului cu privire la domeniul de aplicabilitate al interfețelor om-mașină destinate conducerii persoanelor cu handicap vizual în medii necunoscute. Mulțumiri se cuvin domnului Conf.Dr.Ing. Cătălin Căleanu pentru sfaturile deosebite și sprijinul acordat pe parcursul acestor ani.

Doresc să mulțumesc tuturor colegilor din Departamentul de Electronică Aplicată, care au contribuit într-un fel la elaborarea tezei. Mulțumesc colegilor, împreună cu care am elaborat articolele științifice rezultate ca urmare a preocupărilor în domeniu studiat.

Mulțumesc domnului Director al Departamentului de Electronică Aplicată, Prof.Dr.Ing. Ivan Bogdanov, domnului Decan al Facultății de Electronică și Telecomunicații, Prof.Dr.Ing. Marius Oteșteanu și doamnei Prof.Dr.Ing. Dorina Isar pentru îndemnul continuu de a finaliza cât mai repede teza de doctorat.

Doresc să-mi exprim considerația și recunoștința pentru referenții științifici, Prof.Dr.Ing. Adrian Gaur (Rectorul Universității Ștefan cel Mare din Suceava), Prof.Dr.Ing. Gavril Todorean (Universitatea Tehnică din Cluj-Napoca) și Prof.Dr.Ing. Ivan Bogdanov (Universitatea Politehnica din Timișoara) pentru atenția acordată tezei, dar și pentru critici, aprecieri și sfaturi.

Mulțumiri i se cuvin doctorandului Paul Harfaș (UPT), doctorandei Mengqiu (Karan) Zhang din cadrul Universității Naționale din Australia (ANU) și firmei S.C. Polipol S.A. Foieni pentru sprijinul acordat în ultimul an, cel de prelungire din cadrul studiilor de doctorat, care au contribuit decisiv în eforturile mele de a finaliza teza în timp util.

De asemenea, nu în ultimul rând, doresc să mulțumesc familiei mele, în special mamei mele, Edita, tatălui mele, Andrei-Carol, bunicii mele, Rozália și mătușii mele, Ildikó pentru suportul de excepție, răbdarea, înțelegerea și încurajările nelimitate acordate, de care am beneficiat pe întreaga perioadă a elaborării acestei teze de doctorat. Doresc să transmit mulțumiri speciale familiei Steiger care a fost mereu aproape de mine. Totodată, mulțumesc prietenilor mei, Cezar Ilea, Alex Giuroiu, și Vlad Ilie pentru nenumăratele festivaluri de salsa la care am participat împreună în acești patru ani și care au contribuit decisiv la dezvoltarea mea ca persoană.

Timișoara, Septembrie 2011

Ing. Zoltán Haraszy, M.Sc.

Părinților, Andrei-Carol și Edita

Zoltán Haraszy

Cercetări privind interfața om-mașină cu aplicații în conducerea persoanelor cu handicap vizual cu ajutorul sunetelor

Teze de doctorat ale UPT, Seria 7, Nr. 39, Editura Politehnica, 2011, 147 pagini, 71 figuri, 8 tabele.

ISSN: 1842-7014

ISBN: 978-606-554-366-9

Cuvinte cheie: Realitate Acustică Virtuală, persoane cu handicap vizual, Head Related Transfer Functions, sistem de conducere asistată a deplasării persoanelor cu handicap vizual

Rezumat:

Această teză își propune să aducă contribuții noi în domeniul interfețelor om-mașină destinate conducerii persoanelor cu handicap vizual cu ajutorul sunetelor. Autorul prezintă un sistem al cărui scop este de a asista deplasarea persoanelor cu handicap vizual, compus dintr-un centru de monitorizare și echipament portabil. Interfața om-mașină, dintre echipamentul portabil și persoana în cauză, se bazează pe noul concept de Realitate Acustică Virtuală (RAV), care poate asigura asistarea deplasării persoanelor cu deficiențe de vedere în medii necunoscute. La baza cercetărilor conduse stau funcțiile HRTF (Head Related Transfer Functions), care asigură legătura dintre o sursă de sunet, ce urmează a fi poziționată în spațiu, respectiv semnalele acustice ce ajung la urechile persoanei respective prin intermediul unor căști. În cadrul acestei teze a fost propusă un algoritm de reducere a dimensionalității funcțiilor HRTF. Este descrisă o metodă de măsurare a funcțiilor HRTF în planul orizontal și în fața individului, constând într-un număr de 37 de puncte de măsurare, în condiții apropiate de o cameră fără ecouri. De asemenea, teza propune două metode (una pentru o singură persoană, respectiv una pentru orice persoană) pentru interpolarea funcțiilor HRTF bazate pe rețele neuronale artificiale (RNA). Sunt prezentate interfețele software, ce asigură crearea unei RAV simple folosite în cadrul experimentelor de localizare, care validează structura RNA valabilă unei singure persoane, respectiv algoritmul de reducere a dimensionalității funcțiilor HRTF, și nu numai. De asemenea, este descrisă implementarea hardware și software a unui echipament pentru asistarea deplasării persoanelor cu handicap vizual pe rute precizate folosind RAV, care se bazează pe o placă de dezvoltare puternică la care se atașează o placă experimentală. Acest dispozitiv poate asigura, în momentul de față, conducerea unei persoane cu handicap vizual pe o rută prestabilită fără însă să indice poziționarea posibilelor obstacole, așa cum este descris în conceptul de RAV. La final, sunt enumerate atât concluziile și contribuțiile autorului, cât și posibile direcții viitoare de continuare a cercetării.

CUPRINS

CUPRINS	5
LISTĂ DE FIGURI	7
LISTĂ DE TABELE	10
1. INTRODUCERE	11
1.1. Lumea vizuală și lumea acustică	11
1.2. Motivația cercetării.....	12
1.3. Scopul tezei	13
1.4. Organizarea tezei.....	14
2. SISTEME DE ASISTARE A DEPLASĂRII PENTRU PERSOANE CU HANDICAP VIZUAL	16
2.1. Sisteme de asistare a deplasării pentru persoane cu handicap vizual întâlnite în literatură	16
2.2. Sistemul de asistare a deplasării pentru persoane cu handicap vizual propus ..	18
2.3. Conceptul de Realitate Acustică Virtuală	22
3. AUZ SPATIAL. SUGESTII DE LOCALIZARE. FUNCȚII HRTF	24
3.1. Auz spațial	24
3.1.1. Sistemul auditiv uman	25
3.1.2. Definiția termenilor utilizați.....	27
3.1.3. Sisteme de coordonate.....	29
3.2. Funcții HRTF.....	32
3.2.1. Prezentarea funcțiilor HRTF.....	32
3.2.1.1. Funcții HRTF individuale sau funcții HRTF generale	34
3.2.1.2. Modele ale funcțiilor HRTF	35
3.2.1.2.1. Modele structurale	36
3.2.1.2.2. Modele bazate metode numerice.....	37
3.2.2. Baze de date publice.....	37
3.2.2.1. Baza de date LISTEN HRTF	38
3.2.2.2. Baza de date CIPIC HRTF	39
3.2.2.3. Baza de date KEMAR HRTF	39
3.2.2.4. Alte baze de date	41
3.2.2.5. Alegerea bazei de date	41
3.2.3. Reducerea dimensionalității funcțiilor HRTF.....	41
3.2.4. Sinteza binaurală.....	43
4. MĂSURAREA FUNCȚIILOR HRTF	45
4.1. Aspecte legate de calitatea măsurării funcțiilor HRTF	46
4.2. Scenariul de măsurare.....	47
4.2.1. Locul de desfășurare și echipamentele folosite	47
4.2.2. Procedura experimentală	51
4.2.3. Egalizarea scenariului de măsurare	52
4.2.4. Proiectarea semnalului de test	53
4.2.5. Extragerea semnalului util din semnalul înregistrat	56
4.3. Valori obținute pentru Diferența Interaurală de Timp	64
4.4. Concluzii. Probleme întâlnite și posibile îmbunătățiri	65

5. CONTRIBUȚII LA INTERPOLAREA FUNCȚIILOR HRTF CU AJUTORUL REȚELOR NEURONALE ARTIFICIALE.....	67
5.1. Interpolarea funcțiilor HRTF	67
5.2. Interpolarea funcțiilor HRTF folosind RNA.....	67
5.2.1. Rețele neuronale artificiale	68
5.2.1.1. RNA de tip perceptron cu un singur neuron	69
5.2.1.2. RNA de tip perceptron cu mai multe straturi	70
5.2.2. Structura de RNA pentru interpolarea funcțiilor HRTF pentru o singură persoană	71
5.2.2.1. Evaluare prin criterii numerice	72
5.2.3. Structura de RNA pentru interpolarea funcțiilor HRTF pentru orice persoană	75
5.2.3.1. Evaluare prin criterii numerice	77
5.3. Concluzii. Probleme întâlnite și posibile îmbunătățiri	80
6. EXPERIMENTE DE LOCALIZARE. UNELTE FOLOSITE	81
6.1. Necesitatea egalizării căștilor utilizate.....	81
6.2. Interfață grafică utilizator pentru generarea unei Realități Acustice Virtuale	82
6.3. Experimente de localizare pentru validarea structurii de RNA pentru o singură persoană	86
6.3.1. Scenariul de măsurare	88
6.3.2. Rezultate obținute	91
6.4. Experimente de localizare pentru validarea algoritmului de reducere a dimensionalității funcțiilor HRTF	92
6.4.1. Scenariul de măsurare și rezultatele obținute	94
6.5. Concluzii. Probleme întâlnite și posibile îmbunătățiri	97
7. IMPLEMENTAREA HARDWARE ȘI SOFTWARE A UNUI ECHIPAMENT PENTRU CONDUCEREA PERSOANELOR CU HANDICAP VIZUAL PE RUTE PRECIZATE FOLOSIND REALITATEA ACUSTICĂ VIRTUALĂ.....	99
7.1. Placa de dezvoltare EK-LM3S8962.....	99
7.2. Placa experimentală	101
7.2.1. Implementarea hardware	101
7.2.2. Implementarea software	105
7.3. Experimente.....	107
7.4. Concluzii. Probleme întâlnite și posibile îmbunătățiri	109
8. CONCLUZII. CONTRIBUȚII PERSONALE. DIRECȚII VIITOARE DE CERCETARE	110
8.1. Concluzii	110
8.2. Contribuții personale	110
8.3. Direcții viitoare de cercetare.....	112
8.4. Lista publicațiilor proprii.....	113
BIBLIOGRAFIE	115
ANEXE.....	140
A1. Proiectarea echipamentului prototip.....	140

LISTĂ DE FIGURI

Fig. 2.1. Diagrama bloc a sistemului de supraveghere propus în [69].	19
Fig. 3.1. Sursa de sunet auzită de ascultător.	24
Fig. 3.2. Secțiunea transversală a sistemului auditiv uman, preluat din [86].	26
Fig. 3.3. Traseul unei transmisii acustice dintr-un punct din câmpul liber spre ureche [88].	27
Fig. 3.4. Definirea celor trei planuri de interes.	30
Fig. 3.5. Sistemul de coordonate vertical-polare (partea stângă), respectiv sistemul de coordonate interaural-polare (partea dreaptă) [90].	30
Fig. 3.6. Sistemul de coordonate vertical polare și terminologia folosită frecvent în literatură.	31
Fig. 3.7. Funcție HRTF tipică măsurată la (0°,0°) aparținând subiectului IRC_1031 din baza de date Listen HRTF.	33
Fig. 3.8. Funcție HRIR tipică măsurată la (0°,0°) aparținând subiectului IRC_1031 din baza de date Listen HRTF.	34
Fig. 3.9. Componentele direcționale și nedirecționale ale funcțiilor HRTF [4], [15].	36
Fig. 3.10. Structura sistemului de măsurare LISTEN HRTF. Difuzorul este montat pe o macara controlată de un calculator. (a) Elevația este ajustată la cea dorită și pentru fiecare elevație, subiectul este rotit la azimutul corespunzător și măsurătoarea este realizată. (b) O poză a mediului de măsurare.	38
Fig. 3.11. Funcție HRTF (de fapt, funcție HRIR) nemodificată din baza de date LISTEN HRTF.	42
Fig. 3.12. Funcție HRTF (de fapt, funcție HRIR) modificată prin aplicarea algoritmului propus asupra funcției HRTF (HRIR) din Fig. 3.11.	43
Fig. 3.13. Diagrama bloc al unui sistem simplu de sinteză a sunetului spațial bazat pe funcții HRTF folosind căști.	44
Fig. 4.1. (a) Măsurarea funcțiilor HRTF, (b) sinteză binaurală folosind funcții HRTF.	46
Fig. 4.2. Aranjarea locului de măsurare. Abrevierea H este folosită pentru înălțimea diferitelor perechi, respectiv pentru înălțimea mesei.	47
Fig. 4.3. (a) Plasarea difuzorului, și (b) plasarea scaunului rotativ.	48
Fig. 4.4. Configurația sistemului de măsurare.	49
Fig. 4.5. (a) Plasarea microfonului în urechea dreaptă, respectiv (b) placa de sunet externă și modulul de baterie.	50
Fig. 4.6. Panoul frontal al instrumentului virtual pentru controlul total al procesului de măsurare implementat în LabVIEW.	51
Fig. 4.7. Măsurarea funcțiilor HRTF cu subiectul așezat pe scaunul rotativ.	52
Fig. 4.8. Semnalul sweep în domeniul timp.	54
Fig. 4.9. Semnalul sweep în domeniul frecvență.	55
Fig. 4.10. Semnalul de test folosit în procesul de măsurare, compus dintr-un număr total de 10 secvențe de semnal sweep și pauză.	55
Fig. 4.11. Panoul frontal al aplicației de procesare a datelor măsurate.	56
Fig. 4.12. Determinarea mediei semnalelor sweep din semnalul de măsurare înregistrat, respectiv din semnalul de test.	57
Fig. 4.13. Semnal sweep mediat aparținând direcției (90°,0°), canalul stâng.	58
Fig. 4.14. Semnal sweep mediat aparținând direcției (90°,0°), canalul drept.	58

Fig. 4.15. Determinarea funcțiilor HRTF necompensate din media semnalelor sweep, respectiv din semnalul sweep original.....	59
Fig. 4.16. Calculul răspunsului invers la impuls al întregului sistem de măsurare...	60
Fig. 4.17. Plasarea microfonului față de difuzor în cadrul determinării funcției de transfer.....	61
Fig. 4.18. Răspunsul la impuls măsurat pentru canalul stâng (domeniul timp).	61
Fig. 4.19. Funcția de transfer măsurată pentru canalul stâng (domeniul frecvență).	62
Fig. 4.20. Răspunsul invers la impuls al canalului stâng.	62
Fig. 4.21. Calculul funcțiilor HRTF calibrate folosind funcțiile HRTF neegalizate și răspunsul invers la impuls al sistemului de măsurare.	63
Fig. 4.22. Perechea de funcții HRIR obținute pentru direcția (90°,0°).	63
Fig. 4.23. (a) Umbră acustică produsă de capul ascultătorului, și (b) indicații similare produse de către sursele S, respectiv S'.....	64
Fig. 5.1. Arhitectura perceptronului simplu.....	69
Fig. 5.2. RNA de tip perceptron cu două straturi ascunse.	70
Fig. 5.3. Structura RNA propuse pentru interpolarea funcțiilor HRTF pentru o singură persoană.	71
Fig. 5.4. Un exemplu de grafic reprezentând eroarea medie pătratică în raport cu numărul de epoci de antrenament.....	73
Fig. 5.5. Performanțe obținute în faza de testare a 10 rețele neuronale cu structura propusă.....	74
Fig. 5.6. Arhitectura RNA propuse pentru interpolarea funcțiilor HRTF pentru orice persoană.	77
Fig. 5.7. Un exemplu al dependenței erorii de antrenament de numărul de epoci. .	78
Fig. 5.8. Performanțele obținute după procedura de testare ale celor 10 RNA cu structura propusă.	79
Fig. 6.1. Panoul frontal al aplicației LabVIEW.	84
Fig. 6.2. Diagrama bloc a aplicației propuse.	84
Fig. 6.3. Schema logică a aplicației propuse.	85
Fig. 6.4. Arhitectura rețelei neuronale este aceeași cu cea propusă în [340], singura diferență apare la numărul neuronilor din stratul ascuns.	86
Fig. 6.5. Scenariul experimental folosit în experimentele de localizare.	88
Fig. 6.6. Interfața implementată în timpul rulării.....	89
Fig. 6.7. Un exemplu al erorii de antrenament obținute pentru 250000 de epoci. ..	90
Fig. 6.8. Rezultatele de localizare obținute.	91
Fig. 6.9. Structura RNA utilizată pentru ambele urechi.....	93
Fig. 6.10. Interfața grafică folosită în cadrul experimentelor.	95
Fig. 6.11. Rezultatele de localizare obținute.	96
Fig. 6.12. Clasificarea erorii absolute de azimut raportate la numărul total de teste de localizare conduse.	96
Fig. 7.1. Placa de dezvoltare Stellaris LM3S8962, imagine preluată din [359].	100
Fig. 7.2. Placa experimentală – privire de sus.....	102
Fig. 7.3. Placa experimentală – privire de jos.	102
Fig. 7.4. Diagrama bloc a componentei audio a plăcii experimentale.....	103
Fig. 7.5. Diagrama bloc a componentei GPS a plăcii experimentale.	104
Fig. 7.6. Prototipul dispozitivului de asistare a deplasării persoanelor cu handicap vizual – partea hardware.	104
Fig. 7.7. (a) Schema logică a părții software, și (b) ciclul de generare a semnalului acustic.....	106
Fig. 7.8. Aplicație de comunicație pe port serial implementată în LabVIEW.	107

Fig. 7.9. Ruta predefinită cu cele 10 puncte de trecere reprezentate cu ajutorul interfeței MapQuest (www.mapquest.com).....	108
Fig. A1. Componenta audio a plăcii experimentale în mediul LTspice IV.	140
Fig. A2. Schema electronică a plăcii experimentale în Altium Designer Summer 09.	141
Fig. A3. Cablajul imprimat al plăcii experimentale – vedere sus.	142
Fig. A4. Cablajul imprimat al plăcii experimentale – vedere jos.....	143
Fig. A5. Cablajul imprimat al plăcii experimentale – vedere sus-jos.	144

LISTĂ DE TABELE

Tab. 2.1. Cele mai importante sisteme de asistare a deplasării persoanelor cu handicap vizual.....	17
Tab. 3.1. Rezoluția măsurătorilor LISTEN HRTF. Fiecare linie conține numărul de valori ale azimutului, respectiv pasul de incrementare al azimutului la elevația respectivă.....	39
Tab. 3.2. Rezoluția măsurătorilor KEMAR HRTF. Fiecare linie conține numărul de valori ale azimutului, respectiv pasul de incrementare al azimutului la elevația respectivă.....	40
Tab. 4.1. Specificația tehnică a microfoanelor binaurale folosite.....	49
Tab. 4.2. Câteva modele ale DIT găsite în literatură.....	65
Tab. 5.1. Parametrii antropometrici disponibili pentru fiecare subiect folosit în procesul de antrenament/test (datele marcate nu au fost disponibile pentru fiecare subiect din baza de date folosită).....	76
Tab. 6.1. Parametrii antropometrici selectați în [158]. Valorile specificate sunt exprimate în milimetri.....	87
Tab. 6.2. O privire de ansamblu a performanțelor sistemului.....	92
Tab. 8.1. Granturi de cercetare.....	114

1. INTRODUCERE

1.1. Lumea vizuală și lumea acustică

Ființa umană este înzestrată cu cel puțin cinci simțuri: simțul vizual, simțul auditiv, simțul olfactiv, simțul gustului și simțul tactil. Cu toate că este posibil să folosim toate aceste simțuri pentru localizarea unui obiect, de cele mai multe ori folosim doar simțul vizual, simțul auditiv sau o combinație a celor două.

În lumea vizuală, mai multe mecanisme permit localizarea unui obiect. În primul rând, deplasamentul unghiular al globului ocular la focalizarea obiectului permite stabilirea poziției lui laterale. Poziția verticală este determinată în același mod. Creierul uman înregistrează ambele deplasamente, cel lateral și cel vertical, prin informația tactilă a mușchilor corespunzători din ochi. Acest lucru este valabil pentru cei doi ochi, atât pentru cel drept, cât și pentru cel stâng.

Există și o altă indicație, furnizată prin combinație ochiului stâng și al celui drept. Din cauza distanței dintre cei doi ochi (deplasament lateral), apare un unghi între axele lor optice când ochii focalizează un obiect. Acest unghi (deplasament) ajută evaluarea poziției obiectului. Acest unghi este o indicație, în special, pentru percepția distanței.

Încă o indicație, mult mai decisivă, pentru distanță este claritatea imaginii obiectului vizual – cauzată de curbura lentilelor ochilor. Când ochii focalizează un obiect vizual, imaginea lui este proiectată clar pe retină.

O altă indicație ia naștere din relația dintre obiect și mediul înconjurător. Mărimea cunoscută și dimensiunile unui obiect pot duce la o informație distinctă despre poziția lui în mediul respectiv. De exemplu, obiecte de mărime și formă identică diferă vizual în funcție de distanța lor de la observator. Cu cât mai departe este un obiect plasat, cu atât mai mică este suprafața afectată pe retină (fovea), indicând astfel distanța mai mare.

Câteodată poate fi de ajutor rotirea capului în direcția obiectului pentru a-l avea direct în față. Rotirea capului induce indicații vizuale dinamice, prin modificarea unghiului dintre axele optice. De exemplu, la rotirea capului, deplasamentul unghiular al axelor optice este mai mare în cazul unui obiect mai apropiat decât pentru unul mai la distanță.

Cu toate acestea, dacă ne bazăm în totalitate pe simțul acustic pentru localizarea surselor de sunete, situația este complet diferită și nu la fel de evidentă ca în cazul optic, sau cel puțin mult mai greu de înțeles.

Bregman [1] ilustrează complexitatea procesului de localizare în lumea acustică:

„Imaginează-ți că ești pe marginea unui lac și un prieten te provoacă la un joc.

Jocul este acesta: prietenul tău sapă două canale înguste de la marginea lacului. Fiecare este de 1-2 metri lungime și câțiva centimetri lățime, și sunt plasate la câțiva metri distanță. Pe fiecare canal la jumătatea lungimii, prietenul tău întinde o batistă și o leagă de cele două margini ale canalului. Când valurile ajung la marginea lacului, ele se propagă pe canale și aduc în mișcare cele două batiste. Ți este permis să te uiți doar la batiste și pornind de la mișcările lor să răspunzi la o serie de întrebări:

Câte vapoare există pe lac și unde sunt poziționate ele? Care este cel mai puternic? Care este mai aproape? Bate vântul? A fost vreun obiect de dimensiune mare brusc aruncat în lac? ...

Rezolvarea acestei probleme pare imposibilă, dar este o analogie strictă la problemă întâlnită de sistemul nostru auditiv. [2]"

În contrast cu cazul optic, unde ochii se pot mișca și focaliza un obiect fără rotirea capului, este imposibilă orientarea urechilor externe ale omului spre sursa de sunet. Prin urmare, nu este posibilă folosirea unei corelații de unghiuri pentru a sprijini localizarea (mai ales, în percepția distanței).

Similar cu cazul optic, un deplasament lateral dintre cele două urechi cauzează diferențe interaurale. Pentru surse de sunet poziționate în planul median (planul vertical de simetrie al capului), aceste diferențe interaurale contribuie puternic la localizarea surselor de sunet. Cu toate acestea, în planul median nu există diferențe interaurale (cel puțin teoretic), în consecință, alte indicații trebuie să fie responsabile pentru distingerea surselor de sunet, care se situează în față, pe deasupra sau în spate [3]-[5]. Variațiile spectrale cauzate de forma urechii externe, a capului și a corpului sunt indicații posibile.

În procesul de localizare vizuală, mișcările ochilor și ale capului sunt necesare pentru a permite proiecția pe retină. În cazul acustic, mișcările capului pot avea un impact destul de puternic asupra localizării.

1.2. Motivația cercetării

Cercetările în domeniul conducerii persoanelor cu handicap vizual sunt de importanță deosebită. Acest lucru se datorează faptului că numărul persoanelor, care suferă de diferite deficiențe de vedere, este foarte mare. Primul raport, care estima numărul persoanelor cu handicap vizual a apărut în 1995 [6]. De atunci, odată cu creșterea importanței acestui domeniu, forțată și de numărul mare de persoane care suferă de acest tip de dizabilitate, numărul rapoartelor a crescut [6]-[13]. Totodată, din cauza reducerii bolilor contagioase prin acțiuni publice de sănătate, numărul persoanelor cu handicap vizual s-a redus de la începutul anilor 1990, în ciuda îmbătrânirii globale a populației [12].

Global, în 2011 mai mult de 285 milioane de persoane au dizabilități de vedere, dintre care 246 milioane au vedere slabă și restul de 39 milioane sunt orbi [12], [13]. În întreaga lume pentru fiecare persoană oarbă, o medie de 3,4 persoane au vedere slabă, cu o variație pe țară și regiune în intervalul 2,4-5,5 [9].

Aceste numere sunt estimările științifice cele mai bune ale numărului de persoane cu deficiențe de vedere și constituie rezultatul unor studii efectuate de OMS, Organizația Mondială a Sănătății (World Health Organization, WHO) în fiecare regiune.

Distribuția persoanelor cu probleme de vedere:

- În funcție de vârstă: Persoanele cu deficiențe de vedere sunt distribuite în mod inegal în diferitele grupe de vârstă. Mai mult de 82% dintre toate persoanele fără vedere au o vârstă mai mare de 50 de ani, deși ele reprezintă doar 20% din populația lumii. Cu numărul persoanelor cu vârstă mai mare de 50 ani în creștere din cauza îmbătrânirii populației în multe țări, mai mulți oameni vor fi expuși riscului de a avea handicap vizual din cauza vârstei asociate. De asemenea, din cauza numărului presupus de ani trăiți în orbire, orbirea din copilărie rămâne o problemă semnificativă, cu un număr estimat de 1,4 milioane de copii cu o vârstă de sub 15 ani, dintr-un număr total de 19 milioane de copii cu handicap vizual.

- În funcție de sex: Studiile disponibile indică faptul că în fiecare regiune a lumii, și la orice vârstă, femeile sunt expuse unui risc semnificativ mai mare de a deveni persoane cu probleme de vedere decât bărbații.

- În funcție de așezare geografică: Persoanele cu dizabilități în vedere nu sunt distribuite uniform în lume. Mai mult de 90% dintre ele trăiesc în țări în curs de dezvoltare.

În ultimul deceniu au apărut numeroase sisteme și dispozitive pentru asistarea/conducerea persoanelor cu handicap vizual în medii nestructurate. Momentan nici unul dintre aceste sisteme nu a reușit să devină un standard în comunitatea acestor persoane.

1.3. Scopul tezei

În cadrul acestei teze, autorul propune un sistem de conducere dedicat persoanelor cu handicap vizual cu scopul bine determinat de a asista deplasarea lor în lumea reală. Acest sistem, prezentat în cadrul Cap. 2, are două componente principale: centrul de monitorizare, respectiv echipamentul portabil. Echipamentul portabil se bazează pe conceptul denumit Realitate Acustică Virtuală (RAV) pentru implementarea interfeței om-mașină. Ideea unei astfel de interfețe a fost sugerată de sensibilitatea și acuratețea a sistemului auditiv uman de a identifica poziția sursei de sunet în spațiul 3D, deși acest proces este, de obicei, asistat de percepția vizuală [1], [3]-[5].

În absența percepției vizuale, acest concept al RAV propune înlocuirea realității vizuale cu o realitate acustică virtuală potrivită, în cadrul căruia obstacolele și direcția de deplasare corectă sunt sugerate cu ajutorul sunetelor, care prin poziția sursei lor virtuale să sugereze posibilele obstacole de pe ruta de deplasare și direcția urmată spre țintă. O sursă virtuală de sunet indică un sunet, care este perceput într-o locație/direcție unde nu există nici o sursă de sunet fizic existentă. Posibilitatea plasării surselor virtuale de sunet în anumite direcții dorite este realizată prin folosirea funcțiilor Head-Related Transfer Functions (HRTF, vezi Cap. 3.2.1), procesul fiind denumit sinteză sau tehnică binaurală (vezi Cap. 3.2.4) [14]. Cuvântul 'binaural' ilustrează faptul că ambele urechi sunt implicate în percepția unui eveniment de sunet. Aceste funcții HRTF sunt reprezentări matematice ale influenței sistemului acustic uman, compus din ureche, cap și partea superioară a corpului, asupra schimbării spectrului semnalelor acustice, care ajung urechile ascultătorului, fiind transmise dintr-un anumit punct sau anumită direcție din spațiul 3D. Autorul

aduce contribuții în ceea ce privește interpolarea acestor funcții HRTF, care au rol esențial poziționarea imaginilor acustice vizate să înlocuiască imaginilor vizuale. Metoda de interpolare propusă se bazează pe unelte ale inteligenței artificiale, și anume pe folosirea rețelelor neuronale artificiale (RNA).

Partea aplicativă a tezei de doctorat constă în elaborarea și implementarea unei metode pentru măsurarea acestor funcții HRTF în medii cu ecouri reduse. Sunt implementate două metode de interpolare a funcțiilor HRTF cu ajutorul RNA: una valabilă unei singure persoane, respectiv cealaltă metodă valabilă pentru orice persoană. Metoda de interpolare pentru o singură persoană este validată prin experimente de localizare. Autorul propune un algoritm de extragere a informațiilor esențiale din funcțiile HRTF. Validitatea acestui algoritm este dovedită prin experimente de ascultare. Autorul a implementat o interfață grafică utilizator, care permite simularea conceptului de RAV, fiind posibilă pentru generarea semnalelor acustice cu ajutorul funcțiilor HRTF, care sunt obținute prin una din următoarele posibilități: folosirea unei baze de date de funcții HRTF, folosirea metodelor de interpolare ale funcțiilor HRTF bazate pe RNA, sau folosirea unei funcții obținute prin alte metode de generare ale funcțiilor HRTF. De asemenea, este proiectat și realizat un echipament pentru asistarea deplasării persoanelor cu handicap vizual, format dintr-o placă de dezvoltare cu nucleu ARM-Cortex M3 la care se atașează o placă experimentală pentru asigurarea generării sunetelor specifice ale RAV, reprezentând interfața om-mașină. În momentul de față, dispozitivul poate asigura asistarea deplasării unei persoane cu deficiențe de vedere pe un traseu, prestabilit prin coordonate GPS, fără semnalizarea obstacolelor de pe traseul parcurs.

1.4. Organizarea tezei

Teza de doctorat este structurată pe un număr total de 8 capitole.

Capitolul 1, cel de introducere, prezintă motivația tezei elaborate și scopul cercetărilor conduse.

Capitolul 2 are ca scop principal introducerea sistemului de asistare, respectiv a conceptului de Realitate Acustică Virtuală, reprezentând interfața om-mașină dintre sistem și persoană. Capitolul începe cu prezentarea caracteristicilor generale legate sistemele de asistare/conducere a persoanelor cu handicap vizual, se continuă prin enumerarea a câtorva dintre dispozitivele de conducere din literatură. Urmează o scurtă trecere în revistă a celor mai importante trăsături ale unui astfel de dispozitiv, după care se prezintă întregul concept propus al sistemului de conducere pentru deplasarea asistată a persoanelor cu handicap vizual, respectiv detaliile conceptului de RAV, care stă la baza acestei teze.

Capitolul 3 este compus din două părți: prima prezintă aspecte generale legate de auzul spațial uman, iar al doilea descrie aspecte legate de funcțiile HRTF. În prima parte, are loc prezentarea urechii umane și a sistemului nervos auditiv. Sunt definiți termenii de bază pentru localizarea sunetelor. În partea a doua, sunt prezentate funcțiile HRTF, idei legate de folosirea funcțiilor HRTF individuale sau generale, după care sunt enumerate cele mai importante modele ale funcțiilor HRTF întâlnite în literatură. Se prezintă cele trei sisteme de coordonate folosite pentru definirea direcțiilor sunetelor: sistemul de coordonate interaural-polare, sistemul de coordonate vertical polare, respectiv sistemul de coordonate sus/jos, stânga/dreapta și față/spate. Tot în acest capitol sunt prezentate cele mai importante baze de date publice, care conțin funcții HRTF. Urmează prezentarea algoritmului propus pentru reducerea dimensionalității funcțiilor HRTF, respectiv prezentarea conceptului de

sinteză binaurală, care este de importanță maximă în plasarea surselor de sunet virtuale în direcțiile dorite.

Capitolul 4 descrie detaliile procesului de măsurare a funcțiilor HRTF. Începe cu prezentarea scenariului de măsurare: locul de desfășurare, echipamentele folosite, procedura experimentală. În continuare se prezintă considerații legate de proiectarea semnalului de test, după care urmează procesarea semnalului înregistrat, compus din mai mulți pași: extragerea semnalului util din semnalul înregistrat, egalizarea sistemului de măsurare folosit, respectiv calculul funcțiilor HRTF măsurate. Sunt prezentate valorile obținute pentru Diferența Interaurală de Timp, respectiv cât de bine sunt aproximate aceste valori de către funcții simple de calcul din literatură. Ultimul subcapitol descrie problemele întâlnite și posibilele îmbunătățiri, care pot fi aduse întregului sistem de măsurare a funcțiilor HRTF.

Capitolul 5 conține contribuțiile legate de interpolarea funcțiilor HRTF cu ajutorul inteligenței artificiale. Prima dată sunt enumerate diferitele metode de interpolare din literatură, după care atenția se mută la interpolarea funcțiilor HRTF folosind rețelele neuronale artificiale. Se continuă cu o prezentare generală a rețelelor neuronale, după care este prezentat perceptronul simplu, respectiv rețeaua de tip perceptron cu mai multe straturi. În continuare, sunt prezentate structurile rețelelor neuronale propuse pentru interpolarea funcțiilor HRTF pentru o singură persoană, respectiv pentru orice persoană. Fiecare dintre aceste structuri de rețea sunt evaluate folosind metode numerice.

Capitolul 6 include prezentarea interfeței grafice utilizator pentru crearea unei RAV prin descrierea panoului frontal, respectiv a diagramei bloc a aplicației implementate, cât și descrierea celor două experimente de localizare conduse în cadrul tezei. Prima a fost executată pentru validarea structurii de RNA propusă pentru interpolarea funcțiilor HRTF pentru o singură persoană și conține prezentarea scenariului de măsurare, respectiv a rezultatelor obținute. Experimentul al doilea este destinat validării algoritmului de reducere a dimensionalității funcțiilor HRTF, cu prezentarea scenariului de măsurare, respectiv a rezultatelor experimentului. Sunt prezentate, și în acest capitol, problemele întâlnite și posibilele îmbunătățiri.

Capitolul 7 este destinat implementării hardware și software a unui echipament pentru asistarea deplasării persoanelor cu handicap vizual pe rute precizate folosind RAV. Pentru început este descrisă placa de dezvoltare EK-LM3S8962, care stă la baza dispozitivului de asistare a deplasării. Urmează prezentarea proiectării și implementării hardware, respectiv software a părții experimentabile, care se atașează la placa de dezvoltare LM3S8962. În continuare, sunt descrise experimentele conduse cu dispozitivul proiectat, respectiv problemele, limitările întâlnite și posibilele îmbunătățiri.

În final, capitolul 8 este dedicat prezentării concluziilor finale, contribuțiilor personale aduse de către autor, direcțiilor viitoare de cercetare.

2. SISTEME DE ASISTARE A DEPLASĂRII PENTRU PERSOANE CU HANDICAP VIZUAL

2.1. Sisteme de asistare a deplasării pentru persoane cu handicap vizual întâlnite în literatură

Ultimul deceniu reprezintă un pas mare înainte în dezvoltarea diferitelor tipuri de dispozitive dedicate persoanelor nevăzătoare, care asistă deplasarea lor în medii nestructurate sau care se schimbă dinamic. Termenii folosiți în literatura pentru aceste dispozitive sunt Electronic Travel Aids (ETA) sau Electronic Orientation Aids (EOA). Avantajele dezvoltării uluitoare a tehnologiei sunt evidente când vine vorba despre asistarea deplasării persoanelor cu handicap vizual. Aceste dispozitive îmbunătățesc mobilitatea utilizatorului, în termeni de siguranță și viteză de deplasare, în medii necunoscute.

Cu toate aceste realizări, până în momentul de față, nu există nici un astfel de dispozitiv de asistare/conducere a deplasării destinat acestor persoane, care să fie acceptat de comunitate persoanelor cu deficiențe de vedere. Succesul unui astfel de dispozitiv depinde puternic de interfața om-mașină folosită între utilizatorul orb și sistemul de asistare/conducere, mai exact de felul în care se transmite informația de sistem către persoană. Pentru ca aceste sisteme să fie acceptate de către potențialii utilizatori, cercetări profunde cu privire la factorii umani trebuie să fie conduse. În [16], autorii prezintă o privire de ansamblu a acestor aspecte de factor uman, incluzând eficiența umană în realități virtuale, aspecte legate de sănătate și siguranță, respectiv impactul social al tehnologiei. Totodată, în literatura de specialitate există mai multe lucrări științifice, care prezintă modalități/interfețe prin care informația este transmisă persoanei nevăzătoare [17]-[22]. Ele oferă o privire de ansamblu asupra posibilităților de reprezentare a informației utile, ce se dorește să fie transmisă persoanei respective. Conform unui studiu [23] din 2005, care și-a propus să compare diferite tehnologii folosite în conducerea persoanelor cu handicap vizual, cele mai eficiente două modalități de a sugera acestor persoane direcția de deplasare sunt vocea, respectiv sunetul virtual, ambele fiind furnizate prin căști.

În continuare, în Tab. 2.1 sunt prezentate câteva dintre cele mai importante sisteme de asistare a deplasării persoanelor nevăzătoare întâlnite în literatura de specialitate. Publicații de sinteză sunt, de asemenea, întâlnite în literatură [24]-[28]. Cele mai importante criterii de proiectare ale acestor sisteme/dispozitive dedicate persoanelor cu handicap vizual sunt date de Bourbakis în [27]:

- Mâini libere – utilizatorii să nu fie obligații să le țină în mâini;
- Portabile – să ofere flexibilitate utilizatorilor și să profite de avantajele tehnologiilor portabile;
- Simple – să fie ușor de folosit (funcționarea și interfața să nu fie încărcate cu caracteristici și funcții inutile) și să nu fie necesare perioade lungi de antrenament.

Câteva dezavantaje ale folosirii sunetelor în navigarea persoanelor nevăzătoare sunt prezentate în [29]:

- Utilizatorul trebuie să aibă auz direcțional normal;
- Se necesită purtarea de căști pentru furnizarea sunetelor la urechi;

Sursă	Detectie obstacole	Planificare globală/locală	Mod de conducere	Asistat de un operator
NavBelt [30], [31]	Ultrasonic	Locală	Sunete	Nu
GuideCane [30], [31]	Ultrasonic	Locală	Urmărirea unui robot	Nu
Drishti [32]	Nu	Globală	Voce	Nu
[33]	Cameră video	Locală	Sunete	Nu
[34]	Ultrasonic	Locală	Sunete	Nu
[35]	Cameră video	Globală și locală	Voce	Da
[36]	RFID	Locală	Urmărirea unui robot	Nu
[37]	Nu	Locală	Voce	Nu
Tom Pouce [38], [39]	Infraroșu	Locală	Vibro-tactil	Nu
Teletact [38], [39]	Laser	Locală	Sunete	Nu
[40]	Ultrasonic	Locală	Vibro-tactil	Nu
SVETA [41]	Cameră video	Locală	Sunete	Nu
BlindAid [42]	RFID	Locală	Voce	Nu
TANIA [43]-[45]	Nu	Globală	Vibro-tactil	Nu
Naviton [46]-[54]	Cameră video	Globală și locală	Sunete	Da
Tyflos [55]	Cameră video	Globală	Vibro-tactil și voce	Nu
[56]	Senzori wireless	Locală	Senzor smart	Nu
[57]	Ultrasonic	Locală	Vibro-tactil și sunete	Nu
[58]	Ultrasonic	Locală	Vibro-tactil	Nu
[59]	Ultrasonic	Locală	Vibro-tactil	Nu
CASBlIP [60]	Cameră video	Globală	Sunete	Nu
[61]	RFID	Locală	Voce	Nu
[62]	Ultrasonic, cameră video	Locală	Sunete	Nu
[63]	Ultrasonic	Locală	Vibro-tactil, sunete sau voce	Nu
[64]	Cameră video	Globală și locală	Sunete	Nu
Sistemul propus	Ultrasonic, cameră video 3D	Globală și locală	Sunete și/sau voce	Da

Tab. 2.1. Cele mai importante sisteme de asistare a deplasării persoanelor cu handicap vizual.

- Auzul spațial normal al sunetelor din mediul înconjurător poate să fie blocat prin folosirea căștilor sau să fie mascat [65] de sunetele generate în căști;
- Există un cost adițional al complexității sistemului asociat cu folosirea sunetelor virtuale.

În subcapitolul următor este prezentat sistemul de asistare a deplasării persoanelor cu deficiențe de vedere propus de către autor.

2.2. Sistemul de asistare a deplasării pentru persoane cu handicap vizual propus

Din cauza numărului mare de persoane cu diferite grade de orbire un sistem de conducere dedicată acestora cu scopul de a ajuta deplasarea lor în lumea reală este un echipament foarte util. Un sistem care folosește un modul de Realitate Virtuală Acustică (RAV), concept descris în Cap. 2.3, poate asista deplasarea persoanelor cu handicap vizual. Sistemul auditiv uman deține abilitatea remarcabilă de a identifica poziția sursei de sunet în spațiul 3D, deși acest proces este, de obicei, asistat de percepția vizuală. Această abilitate de a localiza surse de sunet este un aspect crucial al percepției auditive, el creând un simț al spațiului pentru ascultător, furnizând informații despre locația spațială a obiectelor din mediul înconjurător. Motivația folosirii sunetelor pentru asistarea deplasării persoanelor nevăzătoare a fost accentuată în [66]. În cadrul acestei publicații au fost măsurate și comparate performanțele de localizare a sunetelor la persoane cu și fără deficiențe de vedere în planul orizontal. Concluzia trasă de autori este faptul că performanțele de localizare ale persoanelor cu deficiențe de vedere sunt semnificativ mai ridicate comparativ cu persoanele fără deficiențe de vedere, atât pentru localizare de sunete, cât și pentru percepția distanței sursei de sunet la orice azimut din planul orizontal (vezi sisteme de coordonate, Cap. 3.1.3). Blauert trage două concluzii importante în [2]:

- semnalele acustice ascultate prin ureche sunt cele mai importante semnale de intrare pentru auz spațial;
- chiar și cele mai mici modificări ale acestor semnale pot duce la modificări semnificative în auzul spațial.

Din cauza acestor considerente semnalele acustice sunt furnizate prin căști, și pot da celui care ascultă o imagine auditivă foarte realistă a surselor virtuale ale sunetelor, fiind localizate în spațiul din afara capului [3], [4], [67], [68].

Sistemul de asistare a deplasării persoanelor cu handicap vizual, propus de către autor în [69], folosește un mediu auditiv 3D sau imagine acustică pentru a înlocui imaginile, care nu sunt furnizate în lipsa percepției vizuale. Conceptul de imagine acustică este o imagine vizuală descrisă acustic, menită să ajute persoane cu handicap vizual să se orienteze în spațiu.

Arhitectura sistemului propus în [69] se prezintă în Fig. 2.1. Ea include blocurile necesare pentru conducerea (Echipamentul Portabil) și pentru supervizarea (Centrul de Monitorizare) deplasării persoanelor cu deficiențe de vedere pentru atingerea destinației.

Există cel puțin două cauze pentru a urmări persoanele în timpul deplasării lor în aria supervizată:

- Prima este aceea de a se asigura că mișcarea persoanelor este în progres și sunt capabile să ajungă la destinația dorită;
- A doua, este foarte important să fie cunoscută, în orice moment, poziția actuală a indivizilor, pentru a-i putea ajuta în cazul mediilor care se schimbă dinamic sau în cazul unei urgențe.

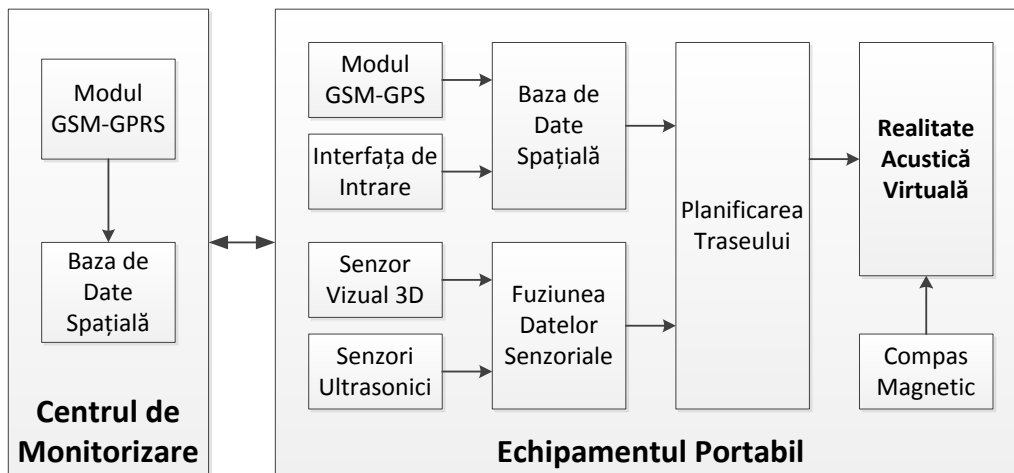


Fig. 2.1. Diagrama bloc a sistemului de supraveghere propus în [69].

Necesitățile menționate mai sus sunt furnizate de prima componentă inclusă în mediul integrat, numit Centrul de Monitorizare (CM). El include un modul GSM/GPRS (Global System for Mobile Communications/General Packet Radio Services), conectat la un calculator personal (PC). Un modul similar este folosit și în cazul Echipamentului Portabil (EP), plasat pe capul fiecărei persoane care navighează în aria supervizată.

Din când în când, CM interoghează EP pentru a primi poziția actuală a subiectului. Locația geografică, în forma unor coordonate carteziene X și Y, este furnizată de modulul GPS (Global Positioning System), inclusă în aceeași unitate cu modulul GSM/GPRS. Coordonatele primite sunt plasate pe o hartă, pe care este desenată traseul, care trebuie să fie urmărit, pentru atingerea destinației [70]. Importanța tehnologiei GPS în asistarea persoanelor nevăzătoare este subliniată în [71]. Deși în forma actuală a acestui sistem de asistare, tehnologia RFID nu este folosită, în [72] se arată că această tehnologie poate constitui o soluție viabilă în viitor, odată ce este rezolvată problema standardizării în domeniul respectiv.

Traseele dorite la destinații diferite sunt stocate în Baza de Date Spațială (BDS), rezidentă pe PC. Orice abatere semnificativă de la traseul propus este detectată și un mesaj de avertizare este trimis subiectului, pentru a corecta în mod corespunzător direcția de deplasare. De observat este faptul că modulele GSM, incluse pe CM și, respectiv, EP, pot fi folosite ca un telefon mobil, pentru comunicație de voce full-duplex. Astfel, subiectul poate stabili o comunicație în orice moment pentru informații adiționale.

EP este un dispozitiv destul de complex, care a fost inspirat din tehnicile bine definite specifice roboților autonomi. El este responsabil pentru următoarele sarcini:

- Generarea traseului la destinația dorită cu evitarea obstacolelor, și
- Furnizarea în timpul deplasării a acestor informații subiectului.

Generarea traiectoriei este efectuată de blocul Planificare Traseu (PT), bazat pe următoarele informații:

- Traseul la destinație, obținut de la BDS stocată pe EP (care este același BDS rezidentă pe PC), în robotică, numit Planificare Globală a Traseului (PGT). Această

BDS de pe EP destinată îndeplinirii funcției PGT este implementată în cadrul prototipului descris în Cap. 7;

- Datele obținute de la sistemul senzorial implementat pe EP. Acesta reprezintă Planificarea Locală a Traseului (PLT).

Componentele hardware, care implementează metodele de planificare menționate mai sus, sunt Controlerul Ierarhic (CI), respectiv Controlerul Reactive (CR). CI include BDS și Interfața de Intrare. Ultima poate fi utilizată de individ pentru selecția, din BDS, destinației dorite (BDS conține trasee la diferite, bine definite destinații). Pe de altă parte, CR previne coliziunea cu obiecte neașteptate din mediul dinamic. El include sistemul senzorial, format din senzorii ultrasonici și senzorul vizual 3D, urmat de blocul de Fuziune a Datelor Senzoriale (FDS). Senzorii ultrasonici sunt o soluție frecvent utilizată în lumea roboților mobili, în timp ce senzorii vizuali 3D constituie o soluție alternativă de valoare pentru viitorul apropiat.

Fuziunea datelor este o procedură obișnuită aplicată datelor senzoriale, pentru a îmbunătăți consistența și fiabilitatea informației achiziționate. O fuziune convenabilă a datelor cumulează avantajele și elimină dezavantajele diferitelor tipuri de senzori.

Blocul PT este implementat în concordanță cu următoarele reguli simple:

- Cât timp nu există obstacole în fața subiectului, traiectoria este generată în acord cu PGT, având prioritate mai mare ca PLT, de exemplu, folosind date obținute de la BDS;
- Altfel, traseul spre destinația dorită este generat folosind PLT, bazată pe date senzoriale (prioritate mai mare ca PGT);
- Odată ce obstacolul a fost depășit, din nou procedura de PGT este luată în considerare.

Sarcina cea mai dificilă în dezvoltarea echipamentului propus este comunicația dintre utilizator, persoana cu handicap vizual, și EP. Trebuie să fie luate în considerare nu numai problemele tehnice ci și alte aspecte specifice la această categorie de persoane. Altfel, o soluție tehnică bună poate fi rejectată de comunitatea persoanelor cu deficiențe de vedere.

Sistemul prezentat nu a fost încă implementat în totalitate. Partea de comunicație GSM/GPRS a EP cu CM a fost realizată cu succes. Coordonatele X și Y, obținute prin sistemul GPS, ale subiecților a fost trimisă cu succes dinspre EP spre CM la o cerere exclusivă a CM, unde aceste coordonate au fost reprezentate pe o hartă utilizând un PC, așa cum este arătat de către autor în [70], [73].

Interfața om-mașină urmează să fie implementată folosind principiile realității virtuale. Acest concept nou, introdus de către autor în [74], este numit Realitate Acustică Virtuală (RAV), și este descris în Cap. 2.3. În acord cu acest concept, poziția obstacolelor și direcția traiectoriei, care se dorește să fie urmărită, sunt sugerate subiectului prin generarea diferitelor tipuri de sunete (imagini acustice).

Există foarte multe aplicații ale sunetului spațial. De exemplu, cele mai uluitoare și fascinante caracteristici ale realității virtuale se bazează pe reproducerea a cât mai multor aspecte ale mediului real și nu bazându-ne doar pe reprezentări vizuale extraordinare. Aceasta înseamnă că orice stimul care aprinde senzație și percepție în ființa umană trebuie cumva reprodusă în lumea virtuală.

În cazul auzului, un sistem audio spațial poate da ascultătorului abilitatea de a percepe poziția surselor ca în viața reală. În general, scopul de bază a sunetului spațial este trimiterea semnalelor ascultătorului care sunt echivalente, perceptual, cu semnalele ce ar fi fost recepționate în mediul care urmează să fie simulat.

Orice sistem de reproducere a sunetelor 3D încearcă să dea ascultătorului un simț al spațiului, și deci trebuie cumva să-l facă pe ascultător să creadă, să aibă impresia că sunetul vine dintr-o poziție unde, de fapt, nu există sursă de sunet reală. Această apropiere este frecvent numită ca reprezentarea surselor virtuale (virtual source imaging).

O parte considerabilă a cercetărilor curente în reprezentarea surselor virtuale se bazează în mare parte pe tehnologia binaurală (binaural technology) [14]. Această tehnică poate fi considerată ca arta "păcălirii" mecanismului auditiv uman pentru localizarea sunetelor. Ea se bazează pe principiul ingineriei sensibile, și anume dacă un sistem de reproducere a sunetelor este capabil să genereze aceeași presiune a sunetelor la membranele timpanice ale ascultătorului ca în cazul în care presiunile ar fi fost reproduse acolo de o sursă de sunet reală, atunci ascultătorul nu ar trebui să fie capabil să facă diferența între o imagine virtuală și o sursă de sunet reală. Pentru determinarea acestor semnale binaurale (binaural signals), sau semnale "țintă", este necesar să cunoaștem cum anume modifică partea superioară a corpului (torso, upper body), capul și pavilionul urechii (pinnae, outer ear) ascultătorului, unde de sunet incidente în acord cu o anumită poziție a sursei de sunet. Această informație poate fi obținută prin măsurători pe capuri manechin (dummy heads) sau pe ființe umane, însă de dorit este ca ea să fie obținută pe ființe umane deoarece rezultatele din [75] arată că modelele existente de capuri artificiale încă nu sunt la fel de utile în înregistrări ca cele ale ființelor umane. Rezultatele unor astfel de măsurători sunt numite uzual funcții HRTF (Head-Related Transfer Functions). Aceste funcții sunt prezentate, mai în detaliu, în Cap. 3.2.1. Orice semnal binaural sintetic poate fi creat prin convoluția (filtrarea) unui semnal de sunet monofonic cu o pereche corespunzătoare de funcții HRTF, procedură numită sinteză binaurală (binaural synthesis), descrisă în Cap. 3.2.4.

Pentru furnizarea corectă a semnalului binaural unui ascultător folosind traductoare, semnalul trebuie egalizat pentru compensarea traseului de transmisie de la traductoare la membranele timpanice (vezi Cap. 4.2.3). Căști sunt folosite frecvent pentru semnalele binaurale, deoarece ele asigură separarea excelentă a canalelor, ele pot izola ascultătorul de sunetele externe și reverberațiile camerei, și traseele de transmisie de la traductoare spre urechi sunt egalizate ușor. Cercetări interesante au fost conduse de Shinn-Cunningham în [76]. Conform acestor cercetări, în ciuda faptului că reverberațiile reprezintă o sugestie importantă pentru distanța sursei de sunet percepute și, totodată, pentru generarea unor sunete externe capului (percepute din exteriorul capului ascultătorului) [77], rezultatele obținute demonstrează faptul că performanțele de localizare direcțională se degradează odată cu includerea chiar și a unor reverberații minore (necesare codării distanței percepute). Aceste concluzii subliniază importanța înțelegerii scopurilor unei anumite aplicații la proiectarea unei RAV. Conform acestor concluzii, reverberațiile nu au fost luate în considerare în cercetările conduse, incluse în această teză.

În cadrul acestei teze, implementarea conceptului de RAV a fost realizată exclusiv cu ajutorul căștilor binaurale. În ciuda beneficiilor evident oferite de folosirea unor căști, există mai multe dezavantaje asociate. Cel mai important, sunetele furnizate prin căști pot veni aparent din interiorul capului ascultătorului (termenii în engleză sunt "in-head localization" sau "inside-the-head localization"), care poate duce la un număr ridicat de confuzii față-spate. Mai mult, confortul datorită purtării unor căști poate deveni o problemă, în special după folosirea lor îndelungată [78].

2.3. Conceptul de Realitate Acustică Virtuală

Succesul unui sistem de asistare a deplasării persoanelor cu handicap vizual depinde puternic de interfața om-mașină folosită între persoana cu handicap vizual și sistemul respectiv. Proiectantul interfeței om-mașină trebuie să fie sigur că informația care se dorește să fie transmisă utilizatorului prin sunete, chiar este informația înțeleasă de către utilizator din acele sunete. Concluzia trasă în [23] și [79] subliniază că sunetele pot fi folosite în mod efectiv în interfețele om-mașină. A fost demonstrat în [74] faptul că Realitatea Acustică Virtuală (RAV), în limba engleză Acoustic Virtual Reality (AVR), poate fi folosită cu succes pentru implementarea unei astfel de interfețe om-mașină. Potrivit acestui concept al RAV, prezența obstacolelor din mediul înconjurător și drumul spre ținta dorită sunt indicate persoanelor în cauză prin sunete ascultate în căști, ale căror poziție virtuală sugerează poziția reală a obstacolelor și direcția de deplasare spre țintă. Folosirea căștilor introduce următoarele erori de localizare [80]: sunete localizate în interiorul capului, confuzii față-spate, surse prea apropiate, elevație deplasată, neclaritatea mișcărilor simetrice în planul median (vezi Fig. 3.4 și Fig. 3.6).

Ideea folosirii sunetelor diferite pentru conducerea persoanelor cu handicap vizual spre ținta dorită, cu evitarea obstacolelor, a fost exploatată în literatura de specialitate, așa cum rezultă din Cap. 2.1. Sistemul auditiv uman dispune de abilități remarcabile în identificarea pozițiilor surselor de sunet în spațiul 3D și permite poziționarea direcțională în spațiu [3]-[5]. Frecvent, acest proces este ajutat de simțul vizual, cunoștințe prealabile și ale intrării senzoriale. În absența simțului vizual, sistemul auditiv poate fi folosit pentru localizarea obstacolelor cu condiția ca aceste obstacole să emită semnale acustice în banda de frecvență a urechii. Ideea de bază a interfeței om-mașină propuse este, deci, substituirea realității vizuale cu o realitate acustică virtuală, conform următoarelor reguli:

- Prezența diferitelor obstacole în mediul înconjurător este semnalizată subiectului prin poziția sursei virtuale ale semnalelor acustice, care sugerează poziția obstacolelor reale;
 - Diferite obstacole se reprezintă prin sunete diferite;
 - Intensitatea și frecvența de repetiție a acestor sunete depind de distanța dintre subiect și obstacol: intensitatea și frecvența de repetiție crește dacă distanța descrește;
 - Un semnal pilot este folosit pentru a indica direcția de deplasare spre destinație, iar subiectul trebuie să urmărească pas cu pas poziția acestei surse virtuale.

Implementarea practică a regulilor de mai sus întâmpină câteva dificultăți tehnice. Cel mai dificil este dezvoltarea unui algoritm simplu și eficient pentru generarea sunetelor potrivite, ale căror surse virtuale să fie percepute de un individ într-un anumit punct din spațiul 3D.

Soluția la problema mai sus menționată este și mai complexă. Când undele acustice se propagă de la o sursă vibrantă spre ascultător, forma de undă a presiunii acustice este modificată de fenomenul de difracție cauzat de partea superioară a corpului (torso), umeri, cap și urechea externă (pinnae). În termeni ingineresti, aceste efecte de propagare pot fi exprimate prin două funcții de transfer, una pentru urechea stângă și încă una pentru urechea dreaptă, care specifică relația dintre presiunea acustică a sursei de sunet și presiunea acustică la intrarea canalului auditiv al urechii din stânga, respectiv dreapta. Ca rezultat, există o pereche de

filtre pentru fiecare poziție a sursei de sunet în spațiu [3]-[5]. Aceste funcții se numesc funcții HRTF, prezentate în Cap. 3.2.1.

Mai mult, comportamentul persoanelor nevăzătoare, în timpul navigării în lumea reală, este foarte similară cu comportamentul altor persoane sănătoase (neavând deficiențe de vedere). Acest comportament adaugă acestui concept al RAV un caracter bioinspirat.

De exemplu, dacă subiectul își rotește capul în planul orizontal, pozițiile surselor virtuale se schimbă corespunzător. Pozițiile acestor surse virtuale depind de parametrii de atitudine ai corpului, capului și urechii externe ai subiectului. Deoarece corelația dintre parametrii de atitudine ai capului subiectului și poziția surselor virtuale, care simulează prezența obstacolelor reale și direcția de deplasare, persoana nevăzătoare își mișcă capul într-un mod asemănător cu un individ cu simțul vizual fără deficiențe în vederea detectării obstacolelor din fața lui. O soluție pentru măsurarea parametrilor de atitudine (direcție, rotire și înclinare) poate fi găsită în [81].

Capitolul următor este dedicat conceptelor legate de auzul spațial uman. Atenția specială este acordată funcțiilor HRTF.

3. AUZ SPATIAL. SUGESTII DE LOCALIZARE. FUNCȚII HRTF

3.1. Auz spațial

Dintre cele minim cinci simțuri cele două simțuri, vederea și auzul, sunt folosite de către ființa umană pentru localizarea obiectelor. În cazul în care nu este disponibil simțul vizual, din cele mai diferite motive, suntem nevoiți să folosim simțul auditiv pentru localizarea obiectelor prin generarea unor sunete ale căror surse virtuale să coincidă cu poziția obiectelor din spațiu. Situația este complet diferită și nu la fel de evidentă ca în cazul simțului vizual, sau cel puțin este mult mai greu de înțeles. Putem spune citându-l pe cercetătorul în domeniul acustic, Georg von Békésy, că "scopul urechilor este să dirijeze ochii".

Capacitatea noastră de a percepe locația sunetelor în spațiu este, cu adevărat, o abilitate remarcabilă. Nu numai că sistemul nostru auditiv poate localiza sunete cu o precizie extraordinară, dar mecanismele legate de auzul spațial pot fi excepțional de robuste, chiar și în condiții de ascultare extreme [3]. Mai mult, sunetele originale emise de diferite tipuri de surse sunt perturbate de reflecții înainte să ajungă la urechea noastră. Pot fi prezente obstacole fizice și mediul poate fi foarte dinamic, cu surse și obiecte mobile. Efortul depus de sistemul auditiv uman este una remarcabilă. Performanțele sistemului auditiv ale persoanelor cu deficiențe de vedere sunt semnificativ mai bune comparativ cu persoanele fără deficiențe de vedere [66].

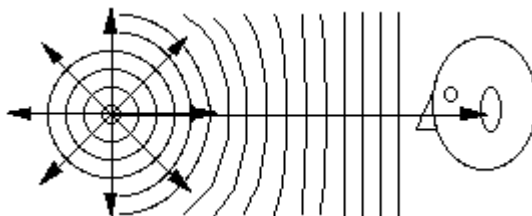


Fig. 3.1. Sursa de sunet auzită de ascultător.

Dacă vederea este un simț focusat într-o anumită direcție (vedem numai acolo unde ne uităm), sunetele, ne înconjoară constant și ne informează de multe dintre obiecte din lumea noastră [82]. Un exemplu este reprezentat în Fig. 3.1. Determinarea locației sunetelor este una dintre cele mai importante abilități biologice ale oamenilor. Pentru a determina sursele sunetelor trebuie să știm că obiectul există, ce este, unde este, dacă se mișcă, etc. Recepționăm sunete simultan de la mai multe surse din mediul înconjurător. Sistemul nervos uman procesează câmpul complex de sunete prin translatarea diferitelor caracteristici ale sunetelor într-un cod neuronal pentru reprezentarea câmpului complex de sunete. Codul neuronal este procesat pentru furnizarea unor informații care ne ajută în determinarea surselor de sunete. În final, informația neuronală auditivă este combinată cu cea furnizată de experiența noastră pentru a se comporta

corespunzător ca răspuns la prezența surselor de sunete. Abilitatea noastră de a localiza ne avertizează de pericol și ne ajută să sortăm sunetele individuale din lumea acustică atât de bogată. Citându-l pe William M. Hartmann, putem enunța: "Bazându-se o varietate de indicații, incluzând intensitatea, sincronizarea, și spectrul, creierul nostru recreează o imagine 3D a unui peisaj acustic din sunetele pe care le auzim" [83]. De asemenea, și alții ilustrează complexitatea procesului de localizare în lumea acustică în [1]-[5].

Putem auzi un sunet în întuneric, unde nu neapărat putem folosi vederea și, în contrast la câmpul vizual limitat, sistemul auditiv este omnidirecțional, permițându-ne să auzim sunete venite din orice poziție a spațiului 3D. Auzul servește, de asemenea, la ghidarea atenției vizuale reglate fin, ușurând sarcina sistemului vizual [4], [84].

În cadrul acestui subcapitol sunt prezentate aspecte legate de auzul spațial, definiția termenilor de interes folosiți, respectiv sistemele de coordonate folosite în literatura de specialitate. Alte aspecte legate de localizarea sunetelor sunt prezentate de către autor în [85].

3.1.1. Sistemul auditiv uman

În acest subcapitol se prezintă sistemul auditiv uman, reprezentat în Fig. 3.2, care este împărțit, în general, în patru părți: urechea externă, urechea mijlocie, urechea internă, și sistemul nervos auditiv central [86], [87]. Sistemul nervos auditiv central, incluzând toate interconexiunile complexe în sistemul auditiv, nu este reprezentat în această figură.

Presiunea acustică variabilă, care produce sunete, ne lovește constant și este colectată de urechea externă. Urechea externă constă din pavilionul urechii (auricula) și canalul care ne duce la membrana timpanică (timpan). Urechea mijlocie este situată între membrana timpanică și o perete osoasă, care conține melcul urechii interne. Urechea internă poate fi împărțită în trei părți: canalele semicirculare, vestibulul, și melcul. Vestibulul afectează mai mult echilibrul decât auzul. Melcul este organul primar auditiv al urechii interne. El seamănă cu un tub de diametru descrescător, care este înfășurat din ce în ce mai mult brusc peste sine însuși. Vederea transversală al melcului arată trei canale pline cu lichid. Canalul din centru este limitat de membrana basilară, o porțiune a pereții externe a melcului. Membrana basilară umană este de o lungime de 34 mm de la bază la vârf. Lățimea ei variază de la 0.16 mm la bază la 0.52 mm la vârf. Capătul dinspre bază este mai îngust și mai solid decât capătul dinspre vârf. Aceste fapte devin importante în considerarea schemei de vibrație a membranei basilare, ca și răspuns pentru stimularea acustică. Răspunsul membranei basilare la stimulări acustice definește așa-numitul model al melcului.

Presiunea acustică este transmisă fluidului urechii interne prin urechea mijlocie și externă. Urechea externă și cea mijlocie ajută la depășirea impedanței urechii medii și a celei interne și, deci, permite o transmisie foarte eficientă a stimulului acustic spre urechea internă. Urechea externă și cea mijlocie furnizează protecția urechii interne împotriva variațiilor excesive ale mediului.

Schema de vibrație a membranei basilare, reprezentând mesajele acustice, ajung la urechea internă prin fereastra ovală, care dirijează melcul umplut cu fluid. Deoarece membrana basilară devine mai lată dacă distanța de la bază spre vârf crește, și deoarece aciditatea ei crește și ea, perioada naturală a vibrației

membranei basilare descrește. Deci, sunete de frecvențe diferite determină ca punctele de vibrație maximă să fie la locații diferite de-a lungul membranei.

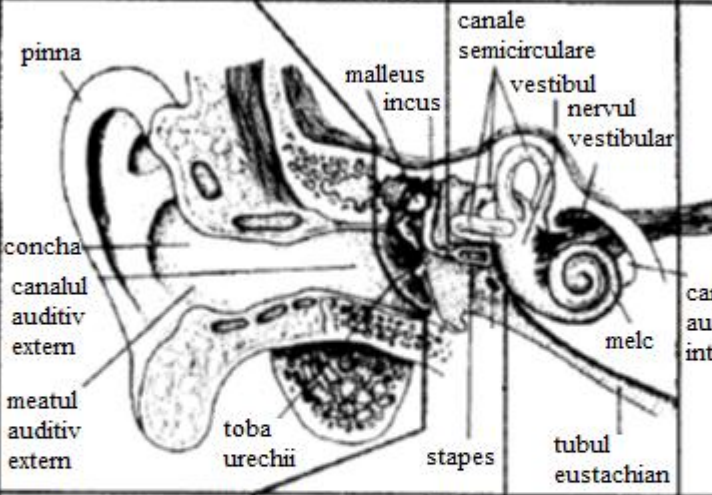

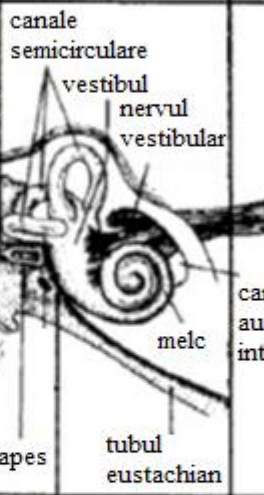
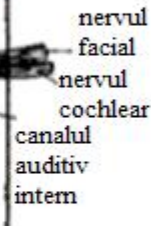
Împărțire	Urechea externă	Urechea mijlocie	Urechea internă	Sistemul nervos auditiv central
Anatomie	 <p>pinna concha canalul auditiv extern meatul auditiv extern toba urechii</p>	 <p>malleus incus stapes</p>	 <p>canale semicirculare vestibul nervul vestibular melc tubul eustachian</p>	 <p>nervul facial nervul cochlear canalul auditiv intern</p>
Mod de operare	Prin vibrația aerului	Prin vibrație mecanică	Mecanic, hidro-dinamic, electrochimic	Electrochimic
Funcție	Protecție, amplificare, localizare	Echivalarea impecănței, stimularea ferestrei ovale selective, egalizarea presiunii	Distribuției filtrării, transmitere	Procesarea informației

Fig. 3.2. Secțiunea transversală a sistemului auditiv uman, preluat din [86].

Dacă două sunete de frecvențe diferite sunt recepționate de melc simultan, ele creează fiecare o deplasare maximă în puncte diferite de-a lungul membranei basilare. Această separare a unui semnal complex în puncte maxime diferite, corespunzătoare sinusoidelor din care semnalul complex este compus, înseamnă că membrana basilară poate fi construită ca o serie de filtre trece-bandă.

Când undele de sunet ajung la urechea externă, ele interacționează cu diferitele părți ale sistemului auditiv uman, așa cum se arată în Fig. 3.3. Părțile sistemului auditiv suportă moduri de rezonanță diferite dependente de direcție de sosire a undei. Aceste rezonanțe pot amplifica sau atenua semnalele acustice, care sosesc, la anumite frecvențe asociate. De exemplu, cavitatea centrală a auriculei, cunoscută ca și melc, are o contribuție majoră în jurul frecvenței de 5.1 kHz la

rezonanța efectivă totală a urechii externe, măbind nivelul semnalului sosit cu 10-12 dB la această frecvență. Consecvent, capul și auricula (și canalul auditiv) pot fi considerate codificatoare spațiale ale semnalelor acustice incidente (Fig. 3.3), care sunt în cele din urmă decodate spațial de către creierul uman.

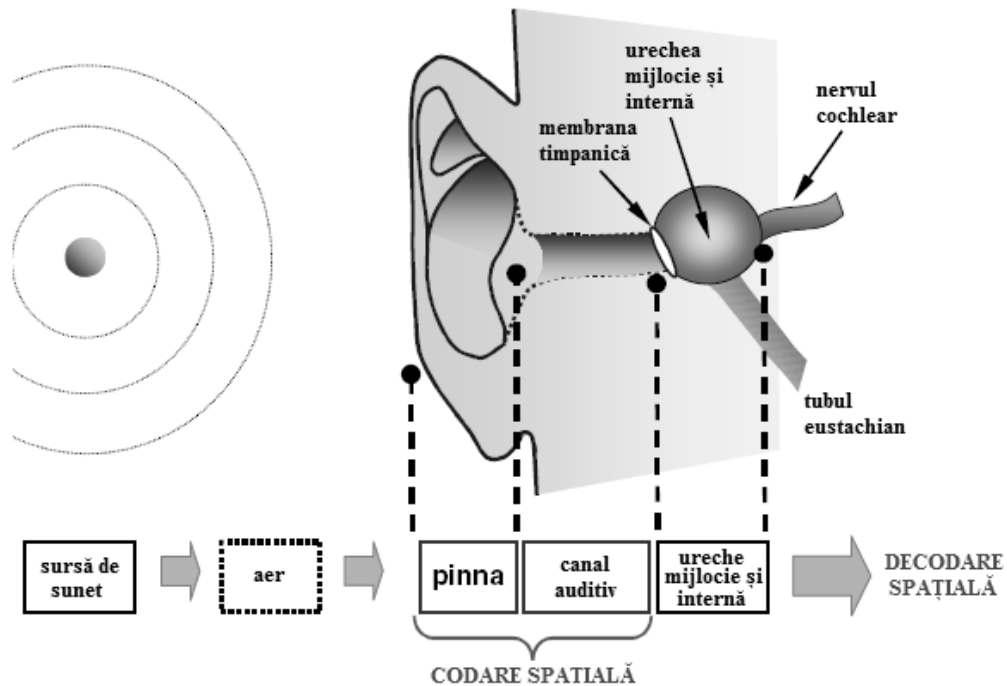


Fig. 3.3. Traseul unei transmisii acustice dintr-un punct din câmpul liber spre ureche [88].

3.1.2. Definiția termenilor utilizați

La prima vedere, semnificația verbului "a localiza" poate fi evidentă: găsirea, localizarea, determinarea poziției. Putem afirma acest lucru pentru că termenul este partea limbajului obișnuit, de exemplu în "localizarea unei greșeli". Cu toate acestea, nu este neapărat clar ce înseamnă în contextul localizării auditive. O definiție exactă este dată la sfârșitul acestui subcapitol. Înainte însă, câțiva termeni de bază necesită explicații.

a) Evenimentul de sunet (sound event)

Un eveniment de sunet este un eveniment fizic, care rezultă în emisia unor unde de sunet. Standardul german DIN 1320 28 definește termenul sunet ca "vibrații și unde mecanice ale unui mediu elastic, în particular în banda de frecvență al auzului uman (16 Hz și 20 kHz)". De exemplu, un sunet ia naștere prin împingerea sau tragerea unei sfori, scufundarea unei pietre în apă, sau mișcarea membranei unui difuzor. Toate aceste evenimente au în comun faptul că o mărime fizică măsurabilă, ca densitatea aerului, variază în timp sau în alte aspecte. Din cauza caracterului fizic și a măsurabilității, un eveniment de sunet apare în lumea reală și este produs de o sursă de sunet.

b) Eveniment auditiv (auditory event)

În contrast cu evenimentul de sunet care există în realitate, evenimentul auditiv are loc doar în imaginația ascultătorului, și nu poate fi demonstrat prin tehnici de măsurare. Mai degrabă este perceput sau simțit, în mod conștient sau inconștient. Câteodată un eveniment auditiv are loc chiar fără vibrația mecanică corespunzătoare sau fără evenimentul de sunet. De exemplu, un simptom, al așa-numitei boli, tinnitus este percepția unui sunet sau al unui zgomot în urechi fără ca orice unde de sunet mecanice să fie prezente.

c) Spațiul sunetelor și spațiul auditiv (sound space and auditory space)

Mediul înconjurător este un spațiu fizic, tridimensional. În acest spațiu, ființa umană poate descrie în mod echivoc fiecare punct prin referință la trei coordonate. Lungimea, lățimea și înălțimea constituie sistemul de coordonate carteziene.

Un alt sistem folosit în mod obișnuit este un sistem orientat după ascultător, originea fiind mijlocul capului ascultătorului. Acest sistem este denumit sistemul cu coordonate sferice sau legate de cap. Cele trei variabile sunt distanța de la origine, azimutul sau unghiul azimutal și înălțarea (elevația) sau unghiul de înălțare (unghi de elevație). Acest sistem este utilizat des pentru experimente de localizare.

Deoarece toate evenimentele de sunet au loc în spațiul fizic real, acest spațiu poate fi denumit spațiul de sunete sau spațiul sunetelor. Cu toate acestea, evenimentele auditive apar într-un spațiu, care există doar în imaginația ascultătorului. Acest spațiu imaginar este denumit, în mod corespunzător, spațiu auditiv.

d) Poziția evenimentului de sunet și poziția evenimentului auditiv (sound event position and auditory event position)

Locația unui eveniment de sunet este denumit poziția evenimentului de sunet, care este un punct în spațiul sunetelor și poate fi descrisă prin trei coordonate. Similar, locul, unde este perceput un eveniment auditiv, este denumit poziția evenimentului auditiv, și poate fi descrisă, în mod echivoc, de trei coordonate.

Termenul de localizare înseamnă legătura dintre un eveniment de sunet, respectiv poziția evenimentului de sunet; și un eveniment auditiv, respectiv poziția evenimentului auditiv.

e) Localizare (localization)

În principiu, prin localizare poate fi înțeleasă ca relația dintre pozițiile corespunzătoare în spațiul sunetelor și în spațiul auditiv. Următoarele exemple pot ilustra această afirmație.

- Un ascultător poate alocă o anumită poziție a unui eveniment auditiv la anumite poziții ale unor evenimente de sunet. Această situație este denumită localizare bună.

- În cazul așa-numitei localizări interne capului subiectului (in-head), ascultătorul percepe poziția unui eveniment auditiv ca fiind localizat în interiorul capului, deși sursele de sunet se situează în exteriorul capului, de exemplu, la ascultarea prin căști.

- Câteodată ascultătorul nu percepe nici o poziție specifică, bine-definită al evenimentului auditiv. Mai degrabă este o impresie neclară, o regiune al evenimentului auditiv – nu doar o singură poziție. Această situație poartă denumirea de neclaritate de localizare (localization blur).

- Percepția auditivă poate să nu aibă nici o poziție specifică, nici măcar o regiune neclară sau o localizare internă în capul subiectului.

Cităm două definiții ale termenului de localizare din literatură [2].

- Definiția lui Blauert.

Blauert [3] definește localizarea ca "o lege sau regulă prin care locația unui eveniment auditiv (de exemplu, direcția ei sau distanța) este legată de una sau mai multe proprietăți specifice ale unui eveniment de sunet, sau ale unui alt eveniment care este corelată într-o anumită manieră cu un eveniment auditiv".

Definiția leagă, nu numai, poziția evenimentului auditiv la locația sursei de sunet, dar afirmă posibilitatea influențării poziției evenimentului auditiv prin anumite proprietăți ale sursei/surselor de sunet, ca de exemplu distribuția spectrală.

- Definiția lui Theile.

Theile restrânge termenul de eveniment de sunet la un sunet ce ia naștere dintr-o singură sursă de sunet, și care determină sau influențează poziția sau forma evenimentului(elor) auditiv(e) corespunzător(oare). Presupunând această definiție restrânsă a evenimentului de sunet, el definește localizarea ca o "lege dintre poziția evenimentului auditiv în afara capului și indicațiile specifice ale unor sau mai multe evenimente de sunet".

Această definiție este una mai strictă decât definiția lui Blauert pentru localizare, deoarece include evenimentele auditive care sunt provocate de mai multe surse de sunet, ca așa-numitele surse fantomă în stereofonie folosind difuzoare.

Determinarea poziției evenimentului auditiv, de exemplu, pentru scopuri de localizare, sunt necesare mai multe indicații de localizare. Unele dintre ele pot fi atribuite doar sursei de sunet.

- Definiția localizării.

În cadrul acestei teze, termenul localizare se referă la legea sau regula dintre poziția evenimentului de sunet și indicațiile specifice ale evenimentului(elor) de sunet, indicațiile specifice ale mediului înconjurător și, de asemenea, indicații specifice ale ascultătorului [2]. Această definiție de localizare completează definiția lui Theile, care, de asemenea, ia în considerare indicațiile mediului și ale ascultătorului.

S-a spus că "scopul urechii este de a dirija ochii". În timp ce abilitatea sistemului auditiv de localizare a surselor de sunet este doar o componentă a sistemului nostru de percepție, ea are o valoare înaltă de supraviețuire, și organismele vii au găsit multe căi pentru extragerea informației direcționale din sunete. Deși mistere perceptuale rămân, sugestiile majore sunt cunoscute de mult timp, și studii psihologice atente au stabilit cât de precis putem lua decizii de localizare. Oricine cine dorește să genereze sunet spațial pentru o interfață om-mașină trebuie să cunoască ce factori influențează sistemul auditiv uman.

3.1.3. Sisteme de coordonate

Pentru a specifica locația unei surse de sunet relativ la ascultător, avem nevoie un sistem de coordonate. O alegere naturală este sistemul de coordonate rectangulare centrată în cap, reprezentat în Fig. 3.4. Aici axa x trece aproximativ prin urechea din dreapta, axa y este dirijat drept înainte, și axa z este verticală. Aceste axe definesc trei planuri standard, xy sau planul orizontal, xz sau planul frontal și yz sau planul median. Este clar că planul orizontal definește separație sus/jos, planul frontal definește separația față/spate, și planul median definește separația stânga/dreapta.

Deoarece capul are o formă aproape sferică, un sistem de coordonate sferice este folosit uzual. Aici, coordonatele standard sunt azimutul, înălțarea (elevație) și distanța. Din nefericire, există mai mult de o posibilă definire a acestor coordonate, și persoane diferite pot să le definească în mod diferit:

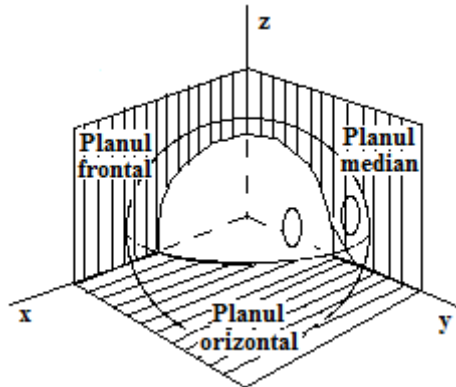


Fig. 3.4. Definirea celor trei planuri de interes.

- **Sistemul de coordonate vertical-polare** (Fig. 3.5, partea stângă) este cel mai des folosit în literatură, mai ales în lucrări ce discută aspecte de psihoacustică în planul orizontal. În acest sistem azimutul este constrâns la valori între $\pm 180^\circ$, iar elevația la valori între $\pm 90^\circ$. În cazul acestui sistem de coordonate, prima dată se măsoară azimutul θ ca unghiul de la planul median la planul vertical conținând sursa și axa z , și pe urmă este măsurat înălțarea ϕ ca unghiul sus (sau jos) de la planul orizontal. Cu această alegere, suprafețe de azimut constant sunt planuri prin axa z , și suprafețe de înălțare constantă sunt conuri concentrice în jurul axei z .

- O alternativă importantă este **sistemul de coordonate interaural-polare**, arătată în Fig. 3.5 în partea dreapta. În acest sistem, azimutul este strâns legat de Diferența Interaurală de Timp (ITD), valorile de azimut definind conuri de confuzii, care au valori ale ITD identice pentru surse la distanță, independente de elevație. Acest sistem este mai des folosit în tehnici de măsurare 'single hoop'. În acest caz, prima dată se măsoară înălțarea ca unghiul de la planul orizontal la planul prin sursă și axa x , care este axa interaurală; azimutul este pe urmă măsurat ca unghiul pe deasupra (sau sub) planului median. Cu această alegere, suprafețe de înălțare constantă sunt planurile prin axa interaurală, și suprafețe de azimut constant sunt conuri concentrice în jurul axei interaurale.

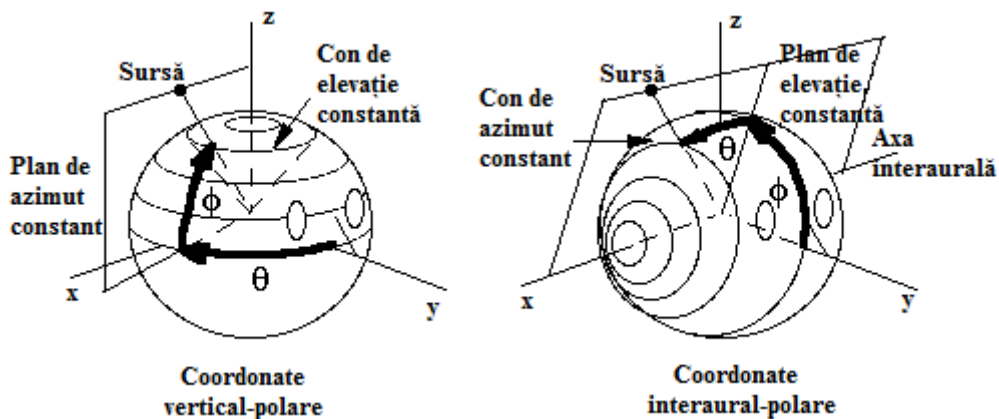


Fig. 3.5. Sistemul de coordonate vertical-polare (partea stângă), respectiv sistemul de coordonate interaural-polare (partea dreaptă) [90].

• **Sistemul de coordonate sus/jos, stânga/dreapta și față/spate** (Up/Down Left/Right and Front/Back Coordinate System, ULF) este al treilea sistem folosit. Un avantaj al acestuia este faptul că este mai simplu să separăm sugestiile: ITD este conținut exclusiv în unghiul stânga/dreapta, confuziile față/spate sunt conținute exclusiv în unghiul față/spate. Unghiul sus/jos este echivalentul elevației din sistemul vertical-polar, în timp ce unghiul stânga/dreapta este echivalentul azimutului din sistemul interaural-polar. Unghiul stânga/dreapta este format între direcția sursei și planul orizontal. Unghiul față/spate este format între direcția sursei și planul median. Unghiul sus/jos este format între direcția sursei și planul frontal [89].

Fig. 3.6 conține câteva terminologii relevante, și arată sistemul de coordonate vertical-polare, cel mai folosit din literatura HRTF.

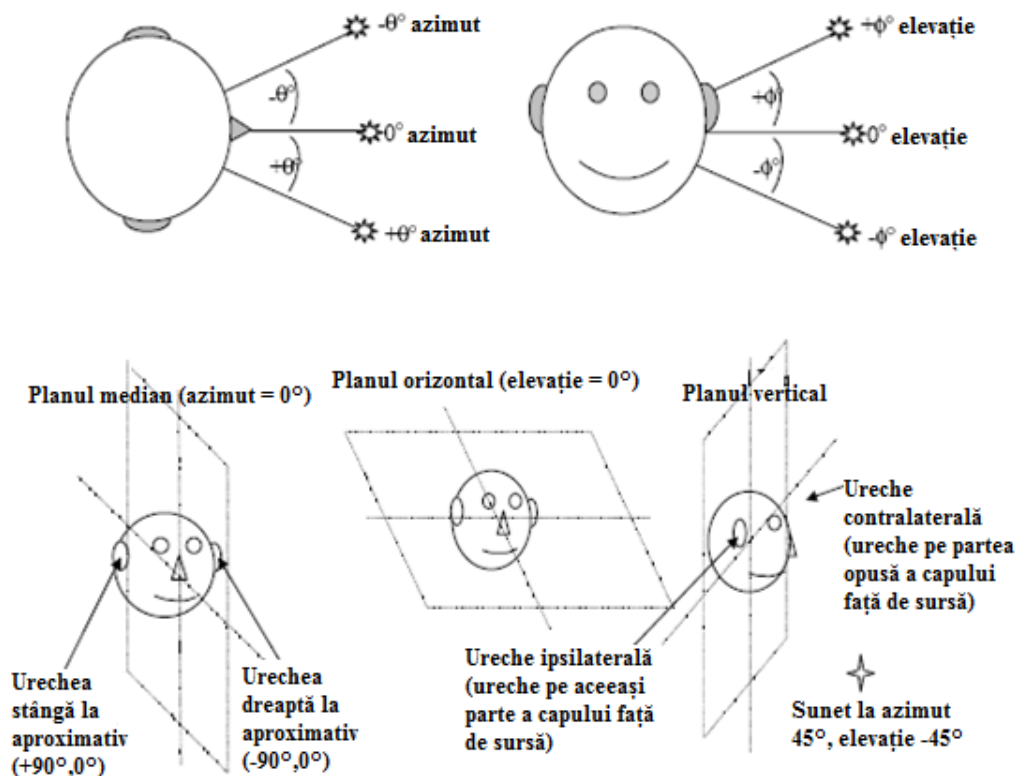


Fig. 3.6. Sistemul de coordonate vertical polare și terminologia folosită frecvent în literatură.

Sistemul vertical-polar este cel mai convenabil pentru descrierea surselor care sunt limitate la planul orizontal, deoarece trebuie specificat doar azimutul ca un unghi între -180° și $+180^\circ$. Cu sistemul interaural-polar, azimutul se află mereu între -90° și $+90^\circ$; în mod surprinzător, distincția față/spate trebuie specificată de înălțare, care este 0° pentru sursele din planul orizontal, și 180° (sau -180°) pentru surse din spate. În cadrul acestei teze, toate valorile de azimut, respectiv elevație sunt date raportând la sistemul vertical-polar, deoarece acesta este sistemul de coordonate folosit în baza de date LISTEN HRTF (Cap. 3.2.2.1), folosit în diversele etape ale elaborării tezei.

3.2. Funcții HRTF

În acest subcapitol se prezintă diferite aspecte legate de funcțiile HRTF, printre care: definiția lor, aspecte legate de avantajele și dezavantajele folosirii funcțiilor HRTF individualizate-generalizate, idei legate de modelarea lor, cele mai importante baze de date HRTF. Urmează prezentarea unui algoritm propus de către autorul tezei pentru reducerea dimensionalității acestor funcții HRTF. În final este descris procesul denumit sinteză binaurală, care stă la baza generării unor sunete spațializate folosind funcțiile HRTF, astfel reprezentând baza conceptului de RAV, descris în Cap. 2.3.

3.2.1. Prezentarea funcțiilor HRTF

Implementarea interfeței om-mașină a sistemului de asistare a deplasării persoanelor cu handicap vizual cu ajutorul conceptului de Realitate Acustică Virtuală se bazează pe funcțiile HRTF. Acest lucru conferă acestor funcții o importanță deosebită în cadrul acestei teze.

Filtrarea spectrală a unei surse de sunet este cauzată în primul rând de către urechea externă. Această filtrare este denumită în literatură Head Related Transfer Function (HRTF), sau funcție HRTF, așa cum este denumită în această teză. Această funcție se rezumă la filtrarea acustică dependentă de direcție a sunetului din spațiul liber, care suferă variații datorită capului, părții superioare ale corpului (torso), și a urechii externe. Folosirea acestei filtrări spectrale reprezintă în mod tipic componenta de bază a sistemelor de sunet 3D. Ele sunt determinate prin măsurări directe, prin modelare și/sau prin interpolare. Acest lucru este bazat pe ipoteza următoare: cel mai precis mod de a produce un sunet spațial este transformarea spectrului sursei de sunet la intrarea canalului urechii cât mai aproape posibil la felul în care ar fi transformat în condiții de auz spațial normal. În literatură, mai există și alți termeni folosiți pentru funcțiile HRTF: Free-Field Transfer Function (FFTF), External Ear Transfer Function (EATF), Ear-Canal Transfer Function (ECTF), Head Transfer Function (HTF), pinnae transform, Outer Ear Transfer Function (OETF), Directional Transfer Function (DTF).

Funcțiile HRTF sunt funcții de patru variabile: azimut, înălțare, distanță și frecvență. Ele sunt definite (vezi ecuația 3.1) ca raportul dintre transformata Fourier a semnalului acustic $P_{L,R}(\omega, \theta, \varphi, r)$ la un punct dat din canalul urechii ascultătorului generat de o sursă plasată la un azimut θ , elevație φ și distanță r , și transformata Fourier a semnalului acustic în câmp liber $P_{ff}(\omega, \theta, \varphi, r)$ care ar fi creată de aceeași sursă de sunet plasată în mijlocul capului ascultătorului, dar în lipsa ascultătorului [14].

$$\begin{aligned} HRTF_L(\omega, \theta, \varphi, r) &= \frac{P_L(\omega, \theta, \varphi, r)}{P_{ff}(\omega, \theta, \varphi, r)} \\ HRTF_R(\omega, \theta, \varphi, r) &= \frac{P_R(\omega, \theta, \varphi, r)}{P_{ff}(\omega, \theta, \varphi, r)} \end{aligned} \quad (3.1)$$

Cu toate că în definiția de mai sus funcția HRTF depinde de distanță, de obicei, ea este neglijată prin considerarea faptului dacă distanța este mare [91] (practica uzuală este presupunerea ca sursa de sunet este situată la o distanță

suficient de mare de la ascultător astfel încât unda de sunet ce ajunge la ascultător poate fi considerată undă plană), putem spune că sursa este în câmp liber. Deoarece majoritatea măsurărilor HRTF se petrec în medii fără ecouri (similare spațiilor libere), se poate neglija efectul distanței. Deci, funcțiile HRTF sunt considerate în continuare că depind doar de trei variabile: azimut, înălțare și frecvență. În acest caz dependența de distanță se simplifică la o scalare după legea pătratului invers. Deci, în majoritatea, dacă nu în toate bazele de date HRTF publice, distanța sursei este constantă (tipic, în intervalul 1-2 m), și nu apare ca variabilă. În Cu toate aceste considerații, ecuațiile precedente devin:

$$\begin{aligned} HRTF_L(\omega, \theta, \varphi) &= \frac{P_L(\omega, \theta, \varphi)}{P_{ff}(\omega, \theta, \varphi)} \\ HRTF_R(\omega, \theta, \varphi) &= \frac{P_R(\omega, \theta, \varphi)}{P_{ff}(\omega, \theta, \varphi)} \end{aligned} \quad (3.2)$$

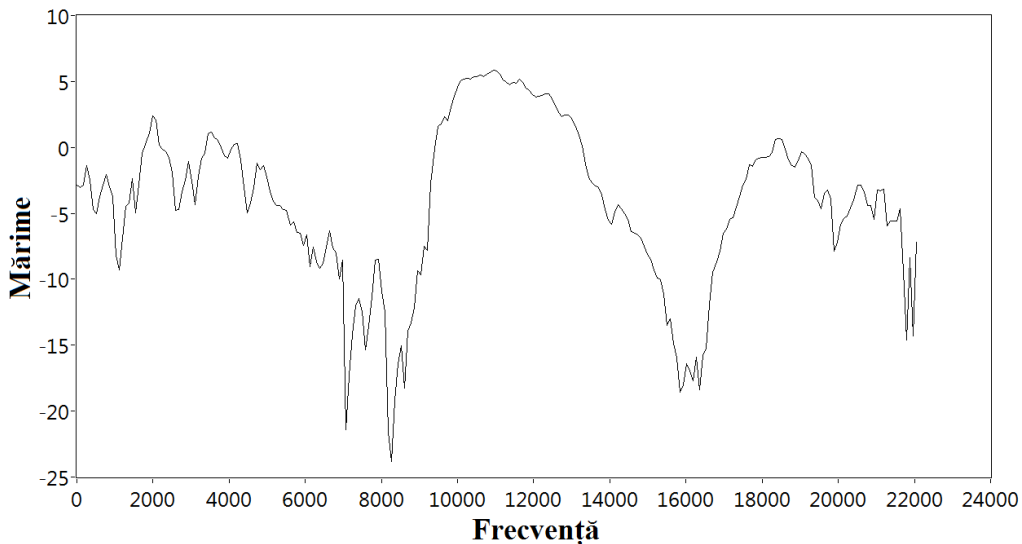


Fig. 3.7. Funcție HRTF tipică măsurată la $(0^\circ, 0^\circ)$ aparținând subiectului IRC_1031 din baza de date Listen HRTF.

Cu toate că funcțiile HRTF sunt funcții în domeniul frecvență, în literatura de specialitate când este folosită denumirea de funcție HRTF (în limba engleză, simplu, HRTFs), ea poate face referire și la funcția HRIR (Head Related Impulse Response), care reprezintă varianta sau perechea în domeniul timp al acestor funcții HRTF.

Formal, o simplă funcție HRTF este definită ca fiind răspunsul în frecvență în spațiul liber al urechii din stânga sau dreapta a unui anumit subiect, măsurat dintr-un punct specific în spațiul liber la un anumit punct în canalul urechii. Funcțiile HRTF sunt frecvent modelate sisteme liniare invariante în timp. Tipic, ambele funcții HRTF, corespunzătoare urechii din stânga sau din dreapta, sunt măsurate empiric de la oameni sau manechini pentru numeroase locații spațiale la o rază fixă de la cap. Locația spațială este alcătuită dintr-o pereche de unghiuri (azimut θ , înălțare φ) și folosind sistemul de coordonate vertical polare (Cap. 3.1.3), direcțiile specificate prin $(-90^\circ, 0^\circ)$ și $(+90^\circ, 0^\circ)$ corespund locațiilor spațiale direct opuse celor două

urechi ale subiectului; în timp ce $(0^\circ, -45^\circ)$ și $(0^\circ, +45^\circ)$ corespund locațiilor spațiale puțin în fața/sub și în fața/deasupra subiectului, etc. Pentru ușurință, definim un set întreg specific unui subiect a acestor funcții HRTF măsurate pentru cazul urechii din stânga și dreapta, ca fiind eșantioane ale funcțiilor HRTF 3D, $HRTF_L(\theta, \varphi, \omega)$ și $HRTF_R(\theta, \varphi, \omega)$. Aceste două funcții reprezintă amplitudinea răspunsului în frecvență, în dB, ale celor două urechi pentru o frecvență dată, ω . Cele două funcții pot fi interpretate ca cantitățile relative de energie recepționate de cele două urechi când un ton pur de frecvență ω este emis de o sursă având coordonatele (θ, φ) .

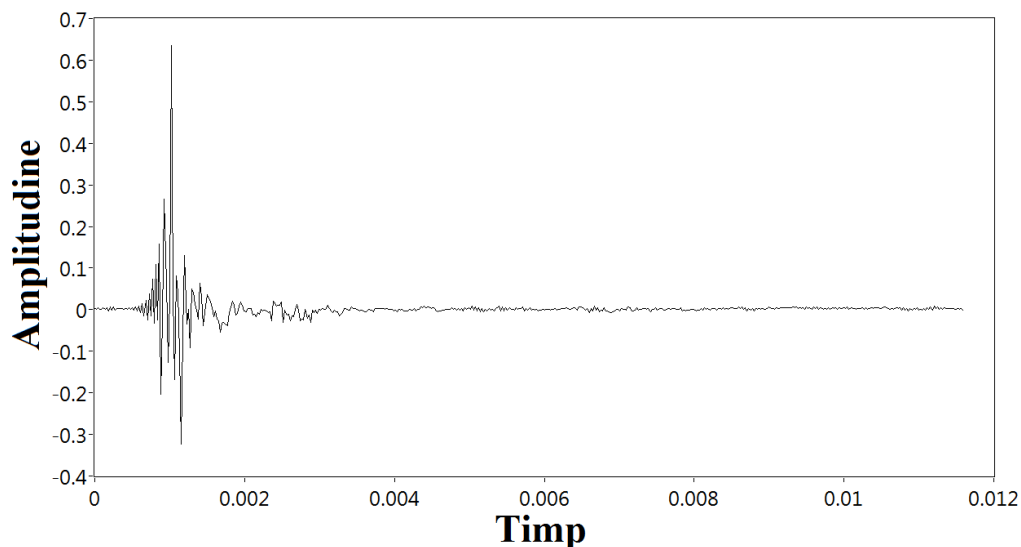


Fig. 3.8. Funcție HRIR tipică măsurată la $(0^\circ, 0^\circ)$ aparținând subiectului IRC_1031 din baza de date Listen HRTF.

Uzual, funcțiile HRTF sunt măsurate într-o anumită locație la un moment dat pentru toate frecvențele de interes folosind tehnici de identificare bazate pe FFT, astfel că $HRTF_{L,R}(\theta, \varphi, \omega)$ sunt frecvent eșantionate neregulat în locații spațiale, dar eșantionate uniform în frecvență.

3.2.1.1. Funcții HRTF individuale sau funcții HRTF generale

Performanța localizării suferă, în general, dacă un ascultător ascultă indicații direcționale sintetizate de la funcții HRTF măsurate de la un cap diferit [92], numite funcții HRTF neindividualizate sau generale. Fiecare cap uman are formă și dimensiune diferită, și există o variație mare în dimensiunea și forma urechii fiecărui individ. Aceasta înseamnă că fiecare individ are un set diferit de indicații direcționale. Cele mai mari diferențe sunt în ceea ce privește transformările tonale la frecvențe înalte cauzate de ureche. Este clar că ne obișnuim cu localizarea sunetelor cu urechile noastre, și deci abilitățile noastre de localizare sunt diminuate dacă ascultăm sunete prin urechile altor persoane. Această idee vine în contrast cu concluziile trase de Hofman în [93], unde sugestiile spectrale a fost modificate prin schimbarea formei urechii externe. Hofman raportează faptul că, deși localizarea sunetelor a fost degradată dramatic imediat după modificarea urechii externe, performanțele de localizare au fost redobândite.

Unicitatea fiecărei ființe umane este sursa celei mai mari limitări ale tehnologiei 3D [94]. Utilizarea funcțiilor HRTF neindividualizate rezultă în două tipuri de erori de localizare care pot fi observate în sisteme audio 3D: confuziile față-spate și erori de înălțare [92]. Confuziile față-spate rezultă dacă ascultătorul percepe sunetul că vine din față în loc de locația actuală din spatele lui, sau invers. Dacă sunetele 3D sunt reproduse de difuzoare frontale, confuziile față-spate tind să fie comune, care înseamnă că unii ascultători pot să nu fie capabili să perceapă sunetele care vin din spate. În practică, acest lucru înseamnă că la mișcarea sursei de sunet din față, în jurul capului, și la spate, rezultatul este perceput ca un sunet translatat pe o parte și înapoi la față. Concluziile din [95]-[98] sugerează că mișcările involuntare ale capului contribuie la reducerea numărului de confuzii față-spate. În literatura de specialitate, în vederea reducerii acestor confuzii față-spate sunt raportate diverse modalități care implică modificarea spectrală a funcțiilor HRTF în [99]-[102], pe lângă existența unor metode raportate în [103]-[106].

Erorile de elevație sunt comune pentru sistemele audio 3D. În practică, dacă un sunet este deplasat din dreapta direct pe deasupra capului, acest lucru poate fi perceput ca un sunet care se mișcă din dreapta direct în față. Aceasta este o manifestare tipică a erorilor de înălțare, observat frecvent la folosirea difuzoarelor. Performanța în ceea ce privește determinarea elevației unei surse de sunet este mult mai bună dacă folosim căști, deoarece indicațiile de frecvență sunt reproduse cu o fidelitate mult mai mare [94]. Dacă comparăm ascultarea sunetelor generate în căști cu cele ascultate în medii obișnuite, numărul confuziilor față-spate crește, în timp ce elevația sursei de sunet este ușor mai slab definit [107]-[108].

O observație foarte interesantă este formulată în [109], și anume faptul că funcțiile HRTF generalizate obținute prin metode deosebit de diferite pot fi interschimbate între ele. Această observație reprezintă o concluzie importantă pentru proiectantul unei interfeței audio spațiale. Când vine vorba de localizarea cu acuratețe a surselor virtuală față de cele reale, ele pot fi localizate cu performanțe mult mai scăzute decât cele reale, fapt care reprezintă un dezavantaj important al utilizării funcțiilor HRTF generalizate [110].

3.2.1.2. Modele ale funcțiilor HRTF

Datorită complexității funcțiilor HRTF, multe sisteme audio spațiale depind de date măsurate experimental. Cel mai important scop deoarece folosim funcțiile HRTF este de a captura corect efectele de azimut și elevație. Dintre aceste efecte sugestiile de elevație sunt deosebit de sensibile la diferențele individuale. O privire de ansamblu a rezultatelor obținute în literatură în localizarea sunetelor din planul orizontal, respectiv cel median, este dată în [80]. O altă privire de ansamblu, furnizat de Salvendy în [111], conține aspecte legate de necesitatea individualizării funcțiilor HRTF, reprezentând un sumar al aproximativ 70 de articole științifice apărute între anii 1977 și 2006.

Din această cauză câteva abordări posibile sunt următoarele:

- Folosirea unui compromis dat de **funcții HRTF generalizate**. Această abordare dă rezultate slabe în localizarea elevației pentru un anumit procentaj din populație, dar cu toate acestea este practică în cazul sistemelor ieftine.
- Folosirea unui **set de funcții HRTF generalizate**. Această abordare necesită măsurarea funcțiilor HRTF pentru un număr mic de persoane reprezentând diferite persoane din populație, și furnizarea unei modalități simple de selecție

pentru alegerea funcțiilor HRTF adecvate [112]. Cu toate că această abordare a fost propusă, nu există nici un astfel de set de funcții.

- Folosirea **funcțiilor HRTF adaptate** la persoana în cauză bazate pe dimensiunile sale antropometrice. Această abordare se bazează pe concluzia din [113], de unde rezultă că diferențele individuale în percepție și dimensiuni anatomice pot fi luate în considerare prin folosirea funcțiilor HRTF adaptate, care sunt selectate pe baza validării psihoacustice;

- Folosirea **funcțiilor HRTF individuale**. Această abordare necesită măsurarea funcțiilor HRTF proprii ascultătorului, care este o procedură inconvenientă și consumatoare de timp. Cu toate acestea, ea dă rezultate excelente, iar folosirea funcțiilor HRTF individuale reprezintă sugestia primară localizării auditive [77].

- Folosirea unui **model HRTF**, care conține parametrii adaptabili fiecărei persoane. Aspecte legate de această opțiune sunt prezentate în acest subcapitol.

3.2.1.2.1. Modele structurale

Criteriile de proiectare a unei aplicații pot determina care dintre caracteristicile corpului uman suplimentare față de urechi (pinnae) trebuie să fie luate în considerare pentru definirea funcțiilor de transfer HRTF. Funcțiile de transfer pot lua în considerare numai urechea externă și capul uman, în timp ce altele înglobează efecte ale corpului uman, ca și canalul urechii, umeri, partea superioară a corpului (torso), care contribuie și ei la rândul lor într-o anumită măsură la caracteristica de filtrare globală. Aceste influențe pot fi împărțite în componente, arătate în Fig. 3.9, care ori depind, ori sunt constante în funcție de direcție.

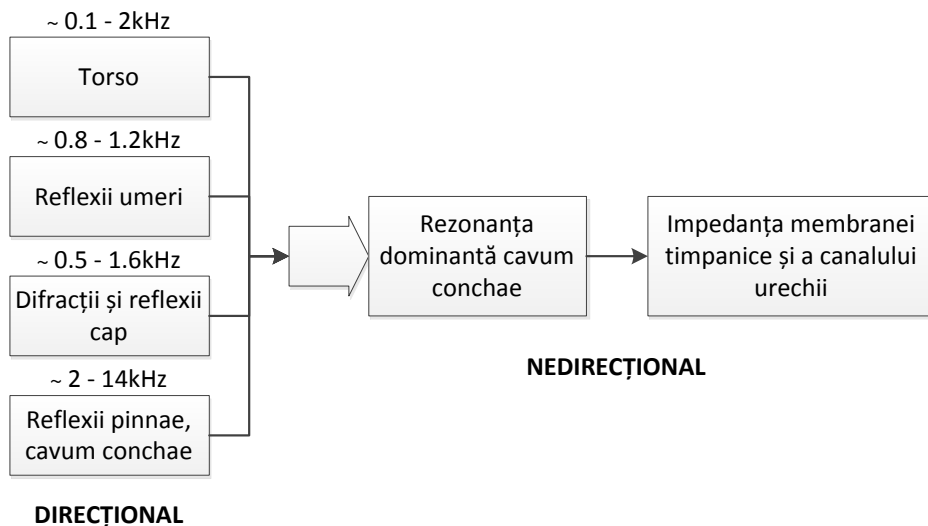


Fig. 3.9. Componentele direcționale și nedirecționale ale funcțiilor HRTF [4], [15].

Pentru simplificarea analizei relației dintre anatomia umană și funcțiile HRTF, este posibilă despărțirea funcțiilor HRTF în componente, care corespund diferitelor contribuții ale părților corpului. În domeniul descompunerii funcțiilor HRTF, cele mai importante contribuții au fost aduse de Duda în [114]-[126]. Fig. 3.9 sugerează că

funcțiile HRTF pot fi aproximate cu o cascadă de filtre, care reprezintă diferitele părți ale corpului uman. Alți cercetători care au avut contribuții la acest domeniu sunt: Chen [127]-[128], Lopez-Poveda [129], Cheng [130], Duraiswami [131], Juskiewicz [132], și Spagnol [133]-[135].

3.2.1.2.2. Modele bazate metode numerice

Domeniul de modelare funcțiilor HRTF a fost și este intens cercetat din cauza neexistenței unui model care poate fi folosit universal pentru eliminarea necesității de măsurare a funcțiilor HRTF în vederea obținerii unei plasări corecte ale surselor de sunet virtuale pentru orice persoană. Diferite tehnici (sau diferite combinații între ele) au fost și sunt folosite pentru cercetarea acestui domeniu, ca:

- analiza componentelor principale (Principal Component Analysis) – Wightman [136], Middlebrooks [137], Park [138]-[144], Hugeng [145]-[147], Salvendy [148]-[150], Wang [151]-[152], Sodnik [153]-[155], Hu [156]-[158], Nishino [159], Matsui [160], Spagnol [161], Katz [162], Kennedy [163], și alții;
- metoda elementului de frontieră (Boundary Element Method) – Otani [164]-[169], Gumerov [170]-[172], Kahana [173]-[175], Katz [176]-[179], Kreuzer [180], Terai [181], Fiala [182], și alții;
- transformata undișor (Wavelet Transform) – Torres [183]-[189], Zhang [190], Xie [191];
- seriile Fourier-Bessel (Fourier-Bessel Series) – Kennedy [192]-[195];
- decompoziția valorii singulare (Singular Value Decomposition) – Bai [196], Rothbucher ;
- modele pol-zero (Pole-Zero Models) – Haneda [198], Durant [199], Zhang [200], Park [201];
- metoda Prony (Prony's signal modeling method) și metoda Steiglitz-McBride (Steiglitz-McBride iteration method) – Barretto [202]-[208];
- spațiul stărilor (State Space) – Adams [209]-[211];
- metoda diferenței finite în domeniul timp (Finite Difference Time Domain method) – Liu [212], Mokhtari [213]-[219];
- rețelele neuronale: Wu [220], Lemaire [221], Goodman [222];
- algoritmi genetici: Cheung [223].

3.2.2. Baze de date publice

Luând în considerare faptul că măsurarea funcțiilor HRTF este o procedură dificilă și consumatoare de timp, există puține baze de date HRTF disponibile pentru comunitatea științifică. Mai mult, din cauza costurilor asociate cu colectarea seturilor de funcții HRTF, cercetătorii și instituțiile care investesc în colectarea datelor, ezită să împărtășească datele obținute cu alți cercetători interesați. Din fericire, însă, există câteva baze de date HRTF disponibile care au fost făcute publice pentru comunitatea științifică interesată. O metodă bazată pe selecție prin competiție a funcțiilor HRTF din mai multe seturi cunoscute este descrisă în [224].

Următoarele secțiuni vor detalia trei astfel de baze de date HRTF:

- Baza de date LISTEN HRTF, măsurată în cadrul proiectului LISTEN [225], [226];
- Baza de date CIPIC HRTF, măsurată de Duda [227];

- Baza de date KEMAR HRTF de la MIT Media Laboratory of Perceptual Computing, măsurată de Gardner [228];
- Alte baze de date.

3.2.2.1. Baza de date LISTEN HRTF

Această bază de date [226] este alcătuită din seturile de funcții HRTF a 49 de persoane și a devenit disponibil spre sfârșitul anului 2002 prin cooperarea dintre Ircam (Institute de Recherche et Coordination Acoustique/Musique) și AKG Acoustics (producător de echipamente acustice), ca parte a proiectului LISTEN [225].

Măsurătorile au fost efectuate într-o cameră fără ecouri. O ilustrare grafică a camerei și al echipamentului folosit se prezintă în Fig. 3.10. Subiectul a fost așezat pe un scaun obișnuit de birou, care a fost montat pe o masă rotativă controlată de un calculator, capabilă de o rotire de 360° (Fig. 3.10b). Un singur difuzor (Tannoy System 600 comandat de un amplificator Yamaha), montat pe o macara, a fost folosit pentru generarea sunetelor (Fig. 3.10a). Elevația difuzorului a fost stabilită (cu ajutorul calculatorului) prin ajustarea elevației de la macara. Pentru fiecare elevație a difuzorului, scaunul a fost rotit la azimutul corespunzător și o pereche (urechea stângă și dreaptă) de măsurători au fost realizată. Asemănător cu baza de date KEMAR HRTF, numărul de măsurători pentru fiecare elevație nu a fost constant, așa cum se rezumă în Tab. 3.1.

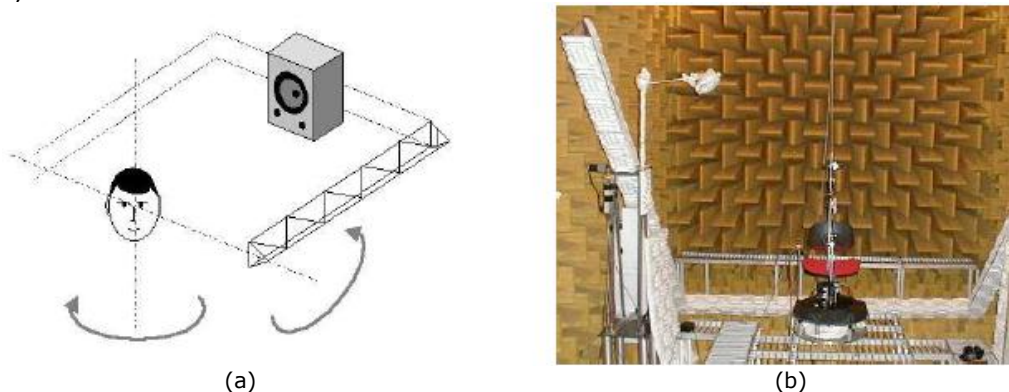


Fig. 3.10. Structura sistemului de măsurare LISTEN HRTF. Difuzorul este montat pe o macara controlată de un calculator. (a) Elevația este ajustată la cea dorită și pentru fiecare elevație, subiectul este rotit la azimutul corespunzător și măsurătoarea este realizată. (b) O poză a mediului de măsurare.

Pentru fiecare subiect au fost realizate 187 de măsurători la locații diferite. Răspunsul a fost măsurat folosind o pereche de microfoane foarte mici (Knowles FG3329). Pentru fiecare subiect, atât măsurătorile neprelucrate (raw), cât și cele egalizate, sunt disponibile în format Microsoft WAV, respectiv Matlab MAT. Măsurătorile neegalizate conțin 8192 de eșantioane (cu durată de 0.186 s) și includ răspunsul întregului sistem de măsurare. Setul de date egalizat a fost obținut prin eliminarea unei porțiuni de la începutul/sfârșitul măsurătorilor pentru evitarea întâzierilor de propagare introduse de reverberații și egalizarea măsurătorilor folosind egalizarea de tip câmp difuz (diffuse field equalization) [229]. Datele egalizate au o lungime de 512 eșantioane (cu durată de 0.012 s). De asemenea,

împreună cu setul de date măsurate, sunt incluse și informații legate de persoane, care au servit ca subiecte de test, informații legate de mediul în care s-au realizat măsurătorile, respectiv informații legate de echipamentele cu care au fost realizate măsurătorile (vârsta persoanelor, dimensiunile camerei fără ecouri, distanța difuzor-subiect, etc.).

<i>Elevație</i>	<i>Număr total de măsurări</i>	<i>Increment unghiular la azimut</i>
-45	24	15
-30	24	15
-15	24	15
0	24	15
15	24	15
30	24	15
45	24	15
60	12	30
75	6	60
90	1	360

Tab. 3.1. Rezoluția măsurătorilor LISTEN HRTF. Fiecare linie conține numărul de valori ale azimutului, respectiv pasul de incrementare al azimutului la elevația respectivă.

3.2.2.2. Baza de date CIPIC HRTF

Baza de date CIPIC HRTF publică [227] constă din 45 seturi de funcții HRTF individuale obținute de la 43 de persoane diferite (27 bărbați și 16 femei) și un manechin KEMAR (cu urechi externe diferite). Pentru fiecare subiect au fost efectuate un total de 1250 de măsurători la fiecare ureche (25 de valori diferite pentru azimut și 50 de valori diferite pentru elevație). Pentru această bază de date a fost folosit un sistem de coordonate interaural-polare. În acest set de date, unghiurile de azimut considerate au fost -80° , -65° , -55° , de la -45° la 45° cu pas de 5° , 55° , 65° , 80° . Elevația a fost măsurată de la -45° la 230.625° , cu pas de 6.625° . Unghiurile sunt raportate la sistemul de coordonate interaural-polare. Fiecare dintre măsurători conține 200 de eșantioane, cu rata de eșantionare de 44100 Hz. Mai mult, măsurătorile au fost egalizate pentru compensarea efectelor întregului sistem de măsurare (microfoane, difuzor, etc.).

Microfoanele au fost plasate puțin în afara canalului urechii, deci răspunsul canalului urechii nu este inclus. Măsurători antropometrice (dimensiuni ale capului, ale corpului, ale urechii) detaliate sunt, de asemenea, incluse în baza de date pentru fiecare subiect.

3.2.2.3. Baza de date KEMAR HRTF

Această bază de date HRTF [228], care conține funcții HRTF neprelucrate, a fost obținută prin măsurarea funcțiilor de interes folosind un manechin artificial antropometric KEMAR (Knowles Electronic Manikin for Acoustic Research). Manechinul KEMAR a fost echipat cu două urechi externe (pinnae) diferite, care permiteau măsurători simultane pentru obținerea a două seturi de funcții HRTF, unul pentru fiecare model de ureche externă corespunzătoare.

Manechinul KEMAR a fost montat pe o masă rotativă controlată electronic, capabilă să fie rotită 360° și plasată într-o cameră fără ecouri, la o distanță de 1.4 metri de la sursa de sunet (un difuzor Realistic Optimus Pro 7), fiind posibilă orientarea manechinului la orice azimut dorit. La rândul său, și difuzorul a fost plasat pe un stand controlat electronic, permițându-se astfel poziționarea lui la orice elevație relativă la manechinul KEMAR. Prin plasarea difuzorului la un anumit azimut și elevație relativ la manechinul KEMAR, măsurarea funcțiilor HRTF corespunzătoare la acea locație particulară a fost realizată prin emisia unui sunet prin difuzor și înregistrarea sunetului cu câte un microfon plasat în fiecare ureche al manechinului KEMAR. În total, 710 măsurători au fost obținute, o elevație la un anumit moment în timp, prin deplasarea difuzorului la elevații particulare, de la -40° la 90° (cu pas de 10°) și rotirea manechinului un total de 360°, cu pas constant pentru fiecare elevație. Tab. 3.2 ilustrează pasul cu care azimutul a fost incrementat pentru fiecare elevație. Răspunsul la impuls măsurat la fiecare ureche conține un număr de 16383 de eșantioane, cu o rată de eșantionare de 44100 Hz și salvat cu o rezoluție pe 16 biți.

Elevație	Număr total de măsurări	Increment unghiular la azimut
-40	56	6.43
-30	60	6
-20	72	5
-10	72	5
0	72	5
10	72	5
20	72	5
30	60	6
40	56	6.43
50	45	8
60	36	10
70	24	15
80	12	30
90	1	360

Tab. 3.2. Rezoluția măsurătorilor KEMAR HRTF. Fiecare linie conține numărul de valori ale azimutului, respectiv pasul de incrementare al azimutului la elevația respectivă.

Deși pentru fiecare măsurătoare, răspunsul la impuls măsurat avea 16383 de eșantioane, nu toate eșantioanele au fost incluse în baza de date. Fiecare răspuns a fost redus la 512 de eșantioane. Deoarece sunetul nu se propagă instantaneu, durează un timp foarte scurt ca sunetul să ajungă de la difuzor la ureche. De asemenea, există o întârziere suplimentară de 50 de eșantioane introduse de sistemul de măsurare. În concluzie, primele 200 de eșantioane nu au fost luate în considerare (presupunând viteza de propagare a sunetului în aer de 345 m/s, timpul de propagare exprimat în număr de eșantioane este de aproximativ 180 de eșantioane). Similar, ultimele 15671 de eșantioane nu au fost nici ele luate în considerare pentru evitarea coruperii procedurii de măsurare din cauza reverberațiilor cauzate de reflexiile undelor acustice cu alte obiecte din camera fără ecouri, ca manechinul KEMAR, standul și camera rotativă.

În plus față de interacțiunea dintre undele acustice și manechinul KEMAR, răspunsul la impuls mai conține și răspunsul întregului sistem de măsurare (difuzor,

microfon, manechin, camera, etc.) și poate produce rezultate slabe în anumite situații. Funcțiile HRTF pot fi egalizate pentru compensarea acestor efecte nedorite.

Manechinul KEMAR a fost echipat cu câte un model al urechii externe și răspunsul a fost măsurat simultan la fiecare ureche. În consecință, răspunsurile nu conțin informații despre DIT (ITD). Dacă este necesar, informațiile despre DIT trebuie să fie adăugate de sistem. Răspunsul difuzorului, al căștilor și al sistemului de măsurare sunt, de asemenea, incluse separat și pot fi folosite pentru egalizarea funcțiilor HRTF măsurate.

3.2.2.4. Alte baze de date

În afară de aceste trei baze de date, care sunt cel mai des folosite în literatură, mai există câteva baze de date elaborate în ultimii ani și au fost puse parțial sau în totalitate la disponibilitatea comunității științifice. Ele sunt următoarele:

- Baza de date HRTF [230], ca rezultat al colaborării dintre Universitatea Tohoku, Universitatea Prefecturală Akita, respectiv Shimizu Corporation, Japonia;
- Baza de date HRTF [231] de la Universitatea Maryland, SUA;
- Baza de date HRTF [232] de la Universitatea de Tehnologie China de Sud, China;
- Baza de date HRTF [233], [234] de la Universitatea Nagoya, Japonia;
- Baza de date HRTF [235] de la Universitatea din Oldenburg, Germania;
- Baza de date HRTF [236] ca rezultat al colaborării dintre Universitatea din Bridgeport și Universitatea Internațională din Florida, SUA;
- Baza de date HRTF [237] de la Universitatea Tehnică din Berlin, Germania;
- Baza de date Romeo-HRTF [238] de la Institutul de Tehnologie din Paris, Franța;
- Baza de date ARI HRTF [239] de la Institutul de Cercetări în Acustică, Austria.

3.2.2.5. Alegerea bazei de date

În cadrul acestei teze, baza de date LISTEN HRTF a fost cea selectată dintre cele publice disponibile din următoarele cauze: numărul mare de persoane pentru care funcțiile HRTF au fost măsurate (în total, 49), numărul mare de măsurări antropometrice disponibile pentru fiecare persoană (în medie, 24), și numărul mare de eșantioane care formează fiecare dintre cele 187 de funcții HRIR egalizate pentru fiecare persoană măsurată (în total, 512).

3.2.3. Reducerea dimensionalității funcțiilor HRTF

În acest subcapitol este descris un algoritm de reducere a dimensionalității funcțiilor HRTF propus de către autorul tezei, prezentat în [240]. Algoritmul își propune să extragă părțile importante ale unei funcții HRTF, specificate prin funcția răspuns la impuls, HRIR. Acest algoritm poate fi aplicat, deci, variantei din domeniul timp al funcțiilor HRTF, funcțiilor HRIR. Pentru validarea acestui algoritm, au fost conduse experimente de localizare. Rezultatele obținute au fost comparate cu rezultatele obținute după experimentele de localizare din [241]-[243], reprezentând

cazul când algoritmul de reducere a dimensionalității nu a fost aplicat funcțiilor HRTF. Aceste experimente sunt prezentate în Cap. 6.

În literatura mai există și alte câteva metode de reducere a dimensionalității funcțiilor HRTF, ca cele raportate în [151], [244]-[250].

În cadrul experimentelor, funcțiile HRTF folosite au fost preluate din baza de date LISTEN HRTF (Cap. 3.2.2.1). A fost obținută reducerea numărului de eșantioane de la 512 la un număr între 60-150, depinzând de valorile eșantioanelor, asupra cărora urmează să fie aplicat algoritmul, și de parametrii algoritmului.

Algoritmul are doi parametri principali: *GetInRule* și *GetOutRule*. Acești doi parametri determină lungimea funcțiilor HRTF filtrate. Funcțiile HRTF filtrate reprezintă funcții HRTF asupra cărora a fost aplicat algoritmul. Fig. 3.11 și Fig. 3.12 furnizează un exemplu despre cum anume o funcție HRTF originală (variantă din domeniul timp, arătată în Fig. 3.11) este filtrată (Fig. 3.12) de algoritmul propus. Este vizibil că algoritmul, practic, reține valorile funcției HRTF dintre două puncte de tăiere determinate de cei doi parametri.

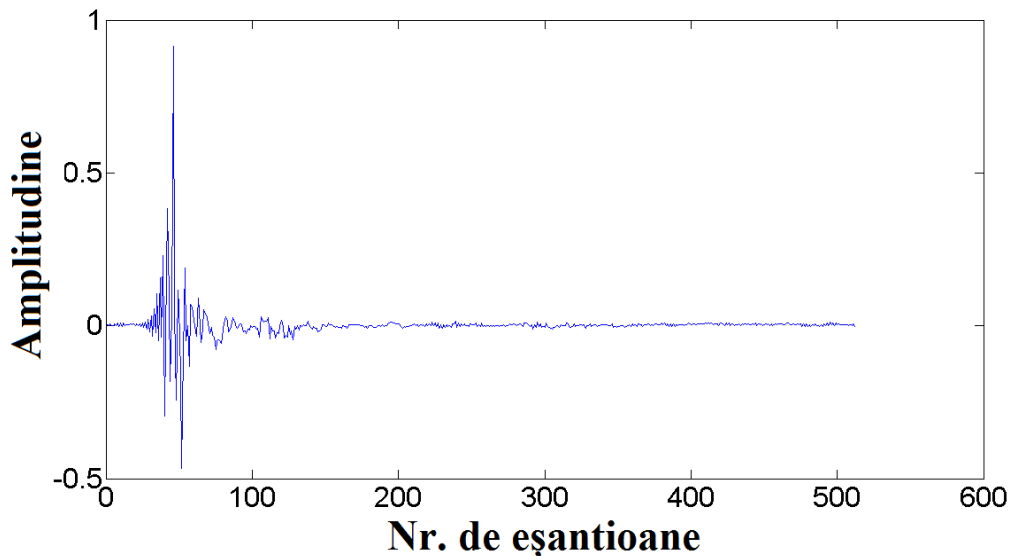


Fig. 3.11. Funcție HRTF (de fapt, funcție HRIR) nemodificată din baza de date LISTEN HRTF.

Algoritm de reducere a dimensionalității funcțiilor HRTF pentru o persoană este bazat pe determinarea punctului de început (de unde să taie), respectiv a punctului de sfârșit (până unde să taie) al părții relevante funcțiilor HRTF pentru extragerea ei din funcții HRTF măsurate, spre exemplu. Pentru determinarea acestor puncte, trei funcții HRTF sunt considerate ca referință:

- Prima referință este perechea de funcții HRTF din stânga subiectului (azimut -90 sau 270 de grade, elevație 0 grade). Această pereche furnizează punctul de început pentru funcțiile HRTF referitoare la urechea stânga, respectiv punctul de sfârșit de la funcțiile HRTF referitoare la urechea dreaptă;
- A doua referință este perechea de funcții HRTF din față subiectului (azimut 0 grade, elevație 0 grade). Această pereche a fost considerată pentru depistarea posibilelor erori;
- A treia referință este perechea de funcții HRTF din dreapta subiectului (azimut 90 de grade, elevație 0 grade). Această pereche furnizează punctul de

început pentru funcțiile HRTF referitoare la urechea stângă, respectiv punctul de sfârșit de la funcțiile HRTF referitoare la urechea stângă.

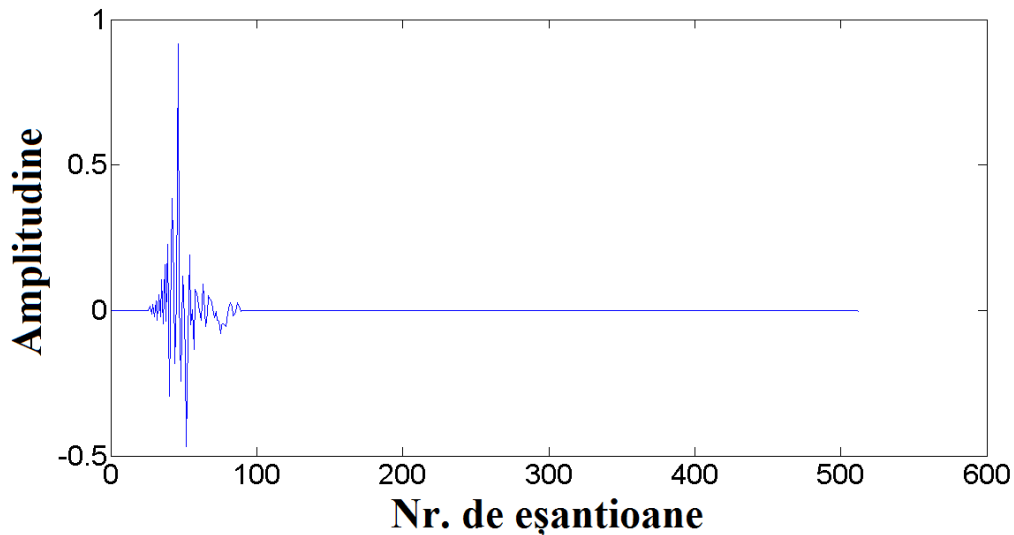


Fig. 3.12. Funcție HRTF (de fapt, funcție HRIR) modificată prin aplicarea algoritmului propus asupra funcției HRTF (HRIR) din Fig. 3.11.

Toate aceste unghiuri sunt date în raport cu sistemul de coordonate vertical-polare, conform alegerii făcute în Cap. 3.2.2.5.

După determinarea celor două puncte de interes (început, respectiv sfârșit) pentru cele trei referințe prezentate, este determinat:

- Punctul de început pentru urechea stângă prin luarea valorii minime dintre cele trei puncte de început, aflate în prealabil;
- Punctul de sfârșit pentru urechea dreaptă prin luarea valorii maxime dintre cele trei puncte de sfârșit, aflate în prealabil.

Această metodologie este aplicată și este identică pentru ambele părți. Deci, odată ce știm care dintre eșantioane trebuie să fie păstrate pentru ambele părți, filtrarea este posibilă pentru orice funcție HRTF a unei persoane pentru obținerea funcțiilor HRTF esențiale.

Aceste funcții HRTF esențiale (filtrate) sunt folosite în vederea antrenării ambelor RNA, care au fost folosite în experimentele de localizare descrise în Cap. 6.4 pentru validarea algoritmului. Următorul capitol prezintă sinteza binaurală, proces care stă la baza conceptului de RAV.

3.2.4. Sinteza binaurală

Procesul de sinteza binaurală se bazează pe funcțiile HRTF, și stă la baza plasării surselor de sunet virtuale (obstacole), ce urmează să fie folosite în vederea implementării conceptului de RAV pentru asistarea deplasării persoanelor cu handicap vizual.

Acest proces important pornește de la un semnal acustic $s(t)$ disponibil și care conține toate informațiile despre un eveniment de sunet dat. Semnalul $s(t)$ mai poate fi întâlnit și sub denumirea de sursă de sunet monaural. În cazul cel mai tipic,

el este o înregistrare a unui sunet ce urmează să fie plasat într-un punct dat din spațiul 3D, specificat de perechea de funcții HRTF ce urmează să fie folosită. Semnalele acustice, $s_L(t)$ și $s_R(t)$, ce ar fi fost generate la urechile unui ascultătorului, sunt date de produsul dintre perechea de funcții HRTF (stânga $HRTF_L(\theta, \phi, \omega)$, respectiv dreapta $HRTF_R(\theta, \phi, \omega)$) și spectrul semnalului acustic ce va fi plasat într-un punct dat (θ, ϕ) . Acest proces este mult mai des implementat în domeniul timp ca și convoluția dintre semnalul acustic respectiv și perechea de funcții HRIR, formând baza sintezei binaurale. În concluzie, ținând cont de linearitatea procesului de sinteză binaurală, el poate fi extins pentru obținerea semnalelor binaurale corespunzând la un număr de surse plasate la un număr de locații diferite. În acest caz, cele două semnale binaurale sunt obținute prin însumarea semnalelor binaurale corespunzătoare fiecărei surse singulare [251].

Procesul de sinteză binaurală este ilustrată în ecuațiile (3.3):

$$\begin{aligned} S_L(\theta, \phi, \omega) &= S(\omega) \cdot HRTF_L(\theta, \phi, \omega) \Leftrightarrow s_L(\theta, \phi, t) = s(t) * HRIR_L(\theta, \phi, t) \\ S_R(\theta, \phi, \omega) &= S(\omega) \cdot HRTF_R(\theta, \phi, \omega) \Leftrightarrow s_R(\theta, \phi, t) = s(t) * HRIR_R(\theta, \phi, t) \end{aligned} \quad (3.3)$$

unde $S(\omega)$ este transformata Fourier a semnalului acustic $s(t)$, $S_L(\theta, \phi, \omega)$ și $S_R(\theta, \phi, \omega)$ reprezintă transformatele Fourier a semnalelor binaurale $s_L(\theta, \phi, t)$ și $s_R(\theta, \phi, t)$, operatorul \cdot reprezintă înmulțirea în domeniul frecvență, în timp ce operatorul $*$ reprezintă convoluție în domeniul timp.

Sinteza binaurală funcționează extrem de bine când sunt folosite pentru sintetizarea indicațiilor de localizare propriile funcții HRTF ale ascultătorului [107]-[108]. Măsurarea funcțiilor HRTF este o procedură complicată, și din această cauză, sistemele 3D audio utilizează un singur set de funcții HRTF măsurat de la un individ particular sau de la un manechin. Fig. 3.13 arată un sistem simplu bazat pe sinteza binaurală a semnalelor acustice folosind căști.

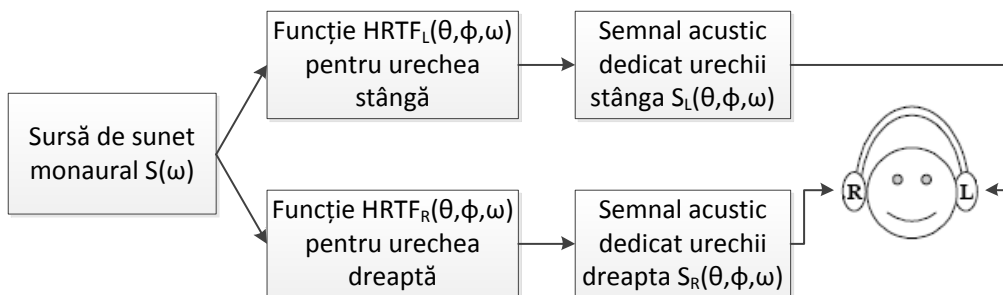


Fig. 3.13. Diagrama bloc al unui sistem simplu de sinteză a sunetului spațial bazat pe funcții HRTF folosind căști.

4. MĂSURAREA FUNCȚIILOR HRTF

Procesul de măsurare a funcțiilor HRTF este executat, de obicei, în medii fără ecouri similare spațiilor libere. Cu toate acestea sunt raportate câteva situații de măsurare și în medii obișnuite [252]-[258]. Scenariul de bază este următoarea: se transmite un semnal acustic cunoscut la un difuzor plasat la azimut θ , elevație φ și distanță r față de capul ascultătorului, și se înregistrează semnalele acustice filtrate folosind microfoane binaurale plasate în canalul sau la intrarea canalului urechii deschise sau blocate. Acest proces este arătat în Fig. 4.1a. Fiecare funcție HRTF, constând tipic din mai multe sute de numere, descrie întârzierea de timp, amplitudinea, și transformarea tonală pentru locația particulară a sursei de sunet la stânga sau la dreapta de subiect. Procedura de măsurare se repetă pentru multe locații ale sursei de sunet relativ la cap, rezultând o bază de date a mai multor sute de funcții HRTF, care descrie caracteristicile de transformare ale sunetului pentru un cap particular [94].

Din cauza numărului mare de componente software și hardware diferite implicate cu un sistem HRTF, măsurarea HRTF poate fi un proces complicat, intensiv pentru subiecți și consumator de timp, un proces care este vulnerabil la erori. Putem enunța, că acuratețea cu care funcțiile HRTF pot fi măsurate este de interes major. O metodă extrapolare în vederea obținerii funcțiilor HRTF măsurate la alte distanțe decât cea care a fost folosită în procesul de măsurare este dată în [259].

Un sistem 3D audio lucrează prin imitarea procesului de auz natural, reproducând indicațiile localizării de sunet la urechile ascultătorului. Acest lucru se realizează cel mai simplu prin folosirea unei perechi de funcții HRTF măsurate. Dacă un semnal de sunet este filtrat de aceste funcții și ascultat prin căști, indicațiile de localizare a sunetului sunt reproduse pentru fiecare ureche, și ascultătorul ar trebui să perceapă sunetul la locația specificată de funcțiile HRTF. Acest proces se numește sinteză binaurală, prezentată în Cap. 3.2.4. Procesul de sinteză binaurală este reprezentat în Fig. 4.1b [94].

În acest capitol este descris un scenariu de măsurare viabil într-un mediu aproape fără ecouri pentru cazuri practice, unde nu este disponibilă o cameră fără ecouri. În comparație cu majoritatea bazelor de date publice, intenția noastră este obținerea funcțiilor HRTF pentru direcții în planul orizontal pentru unghiuri de azimut cuprinse între -90 și $+90$ de grade. De asemenea, măsurarea este efectuată la o rezoluție mai mare, ca de obicei, care reprezintă un avantaj pentru interpolarea funcțiilor HRTF în dezvoltarea RAV (vezi Cap. 5).

În vederea măsurării funcțiilor HRTF, există două metode: prima este denumită metoda directă (exemple în [226]-[228]), în timp ce a doua este denumită metoda reciprocă. Metoda directă [231], [260]-[272] constă în măsurarea funcțiilor HRTF prin difuzarea semnalului de test la diferite locații din spațiul 3D și înregistrarea semnalului acustic recepționat în punctul de măsurare din urechea respectivă. Cealaltă metodă [273]-[277], cea reciprocă se bazează pe principiul reciprocității lui Helmholtz, în care locul difuzării, respectiv a înregistrării sunt interschimbate. Scenariul de măsurare, descris în acest capitol, este bazat pe metoda directă de măsurare a funcțiilor HRTF.

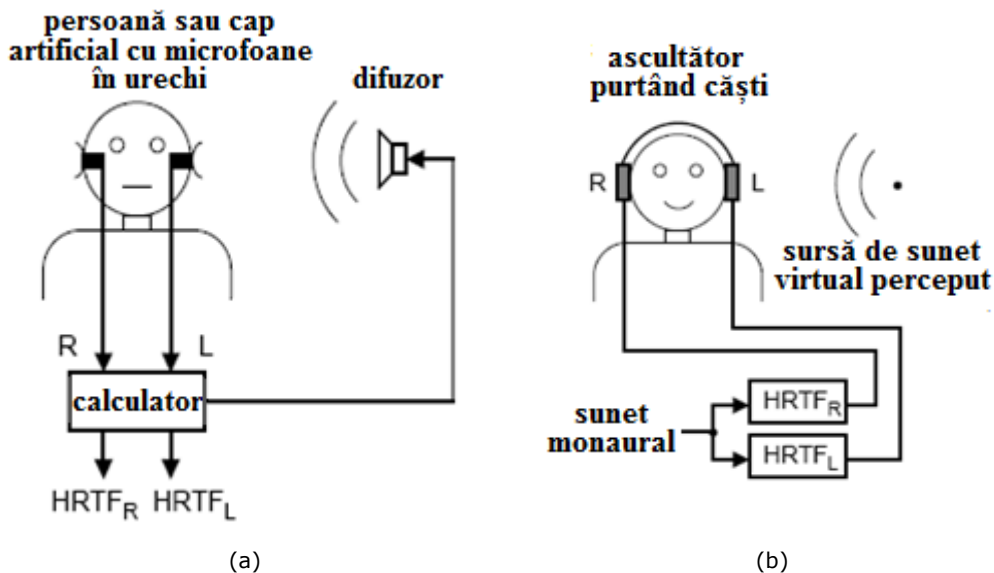


Fig. 4.1. (a) Măsurarea funcțiilor HRTF, (b) sinteză binaurală folosind funcții HRTF.

4.1. Aspecte legate de calitatea măsurării funcțiilor HRTF

Ca toate celelalte experimente practice, procesul de măsurare a funcțiilor HRTF este susceptibilă la diverse erori de natură electrică, acustică și fizică, care pot afecta calitatea rezultatelor obținute [278].

Există două aspecte de natură fizică, ce nu pot fi eliminate, din procesul de măsurare a funcțiilor HRTF:

- Umbra capului contribuie la scăderea raportului semnal-zgomot în cazul semnalului măsurat la urechea contralaterală. Acest lucru reprezintă o sugestie de localizare importantă prin DIN;
- Asimetria minoră a capului, respectiv a corpului subiectului face analiză rezultatelor mai dificilă.

Alte posibile influențe nedorite, care pot fi prezente în măsurările funcțiilor HRTF, reprezintă orice fel de combinație a celor ce urmează a fi enumerate:

- Erori provocate de sistemul de măsurare și/sau mediul de măsurare;
- Erori umane datorate experimentatorului;
- Erori rezultate din poziționarea subiectului de test sau al diferitelor părți ale corpului subiectului;
 - Erori datorate mișcării capului subiectului;
 - Efecte nedorite din cauza materialelor folosite pentru poziționarea corectă a microfoanelor folosite;
 - Erori provocate de poziționarea microfoanelor binaurale.

Pe lângă cele enumerate, există alte dificultăți/limitări dacă măsurările sunt executate/repetate la intervale de pauză îndelungate:

- Instabilitatea sau variația în timp al sistemului de măsurare, de exemplu prin schimbarea componentelor hardware, efecte provocate de folosirea altor microfoane, etc.;

- Schimbări în aparența subiectului, de exemplu prin modificarea aranjării părului, a vestimentației, etc.;
- Schimbări în anatomia subiectului de test datorate, de exemplu, creșterii dimensiunilor urechii, ale capului la subiecți tineri, sau unor operații după o traumă suferită, etc.

Această listă a posibilelor erori, ce pot apărea într-un proces de măsurare, subliniază complexitatea de a efectua măsurarea cu acuratețe a funcțiilor HRTF. Subiectul repetabilității acestor măsurări este îndelung tratat de Riederer în [279]. Tot Riederer a propus o metodă de evaluare calitativă a funcțiilor HRTF bazată pe prelucrarea digitală a semnalelor în [280]. Un alt criteriu de evaluare din perspectiva raportului spectral de putere în [281].

4.2. Scenariul de măsurare

4.2.1. Locul de desfășurare și echipamentele folosite

Amenajarea locului de măsurare este prezentată în Fig. 4.2.

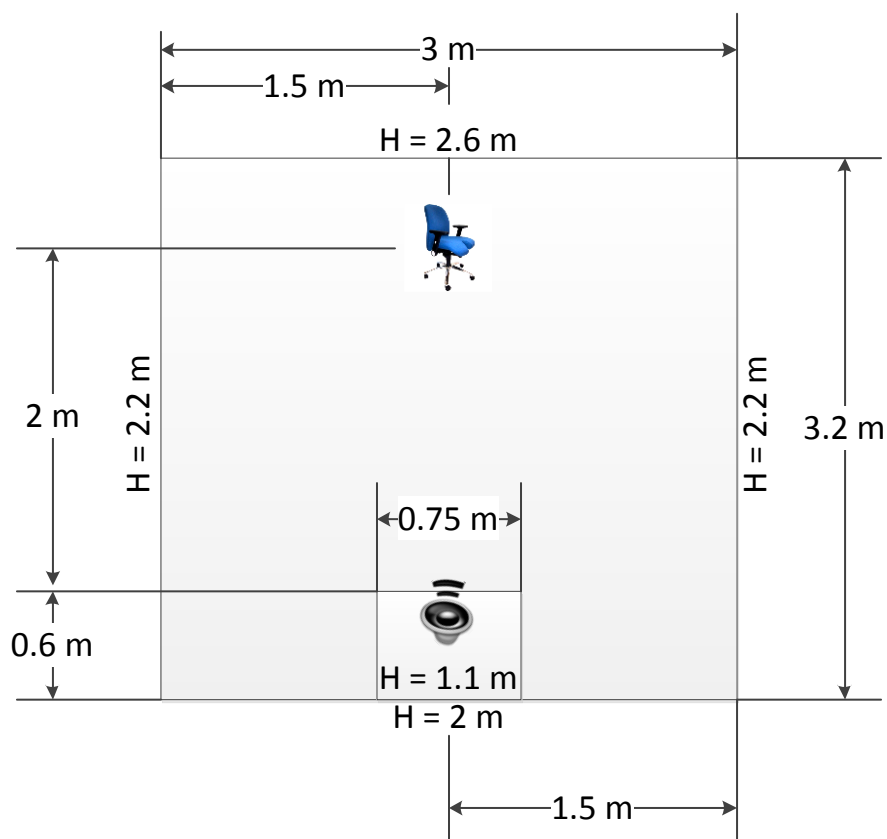


Fig. 4.2. Aranjarea locului de măsurare. Abrevierea H este folosită pentru înălțimea diferitelor perechi, respectiv pentru înălțimea mesei.

Camera de măsurare a fost construită la o fabrică locală de mobilier din cauza cantității mari de spumă disponibilă, utilizată în fabricarea mobilei. Camera construită avea următorii parametri: 3.2 metri lungime, 3 metri lățime, cel puțin 2 metri înălțime și a fost construită din spumă cu o grosime de 0.6 metri. Podeaua a fost acoperită cu spumă de o grosime minimă de 0.2 metri. Scopul acestei camerei experimentale a fost crearea unor condiții de măsurare apropiate de o cameră fără ecouri prin reducerea reflexiilor interioare și izolarea zgomotului exterior.

Difuzorul (vezi Fig. 4.3a) a fost plasat pe o masă construită, tot din spumă, (0.75 metri x 0.6 metri) la o înălțime egală (1.1 metri de la podea) cu capul ascultătorului pentru asigurarea măsurărilor funcțiilor HRTF în planul orizontal (la elevație 0 grade). Masa a fost plasată aproape de unul dintre pereții camerei, așa cum este arătat în Fig. 4.2. Scaunul rotativ (vezi Fig. 4.3b) a fost poziționat la o distanță de aproximativ 2 metri de difuzor pentru asigurarea faptului că distanța de la sursa de sunet, respectiv centrul capului este egală cu 2 metri, pentru o bună aproximare a unei plane în punctele de măsurare din planul orizontal [91].



(a)

(b)

Fig. 4.3. (a) Plasarea difuzorului, și (b) plasarea scaunului rotativ.

Configurația sistemului de măsurare este arătată în Fig. 4.4, și conține următoarele componente:

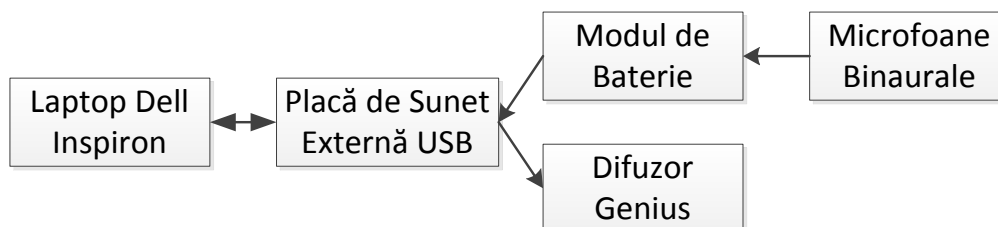


Fig. 4.4. Configurația sistemului de măsurare.

- Laptop Dell Inspiron 1520, pe care rulează un sistem de operare Windows 7 Ultimate pe 32 biți cu mediul de programare grafică LabVIEW controlând tot procesul de măsurare;
- Placă de sunet externă de fidelitate ridicată Andrea Pureaudio USB-SA (vezi Fig. 4.5b) alimentată prin conexiune USB de la laptop. Comunicația cu laptopul se realizează, de asemenea, prin USB. Placa de sunet asigură două intrări stereo pentru microfoanele binaurale și două ieșiri stereo pentru difuzoare, prin conectori jack 3.5 mm atât pentru microfoane, cât și pentru difuzoare [282];
- Modul de baterie Sound Professionals SP-SPSB-8 cu baterie standard de 9V (vezi Fig. 4.5b), pentru performanțe îmbunătățite la nivelele ale presiunii de sunet ridicate. Modulul de baterie este conectat la placa de sunet USB externă, respectiv la microfoanele binaurale folosind cablul de extensie stereo dublu ecranat de calitate ridicată Sound Professionals SP-SPSC-1;
- Microfoane binaurale de zgomot scăzut cu sensibilitate standard Sound Professionals SP-TFB-2-80015, folosite pentru înregistrarea semnalelor acustice binaurale, care au montate suporturi foarte confortabile pentru plasarea corespunzătoare în urechi. Specificațiile tehnice ale microfoanelor folosite sunt date în Tab. 4.1.

Răspuns în frecvență	20-20,000 Hz
Raport semnal-zgomot	Senzitivitate standard 60dB
Senzitivitate la circuit deschis	Senzitivitate standard -42dB
Nivel maxim al sunetului de intrare	105dB
Gamă dinamică	81dB
Directivitate	Omnidirecțională
Lungime cablu	48"
Material corp microfon	Plastic
Alimentare	2-12VDC, alimentare de la modulul de baterie

Tab. 4.1. Specificația tehnică a microfoanelor binaurale folosite.

Spume de forma sferică (foam windscreens) montate pe microfoanele binaurale (Sound Professionals SP-MINI-WS) au fost folosite pentru asigurarea poziționării stabile a microfoanelor relativ la urechi de-a lungul întregului proces de măsurare. Poziționarea stabilă este de o importanță deosebită, deoarece chiar și mișcări mici ale microfoanelor produc modificări semnificative în caracteristica funcțiilor HRTF [278]. În cadrul măsurărilor efectuate, microfoanele binaurale au fost plasate la intrarea canalului deschis a urechii, unde rezultate obținute în

literatură indică faptul că pentru majoritatea ascultătorilor măsurarea funcțiilor HRTF este de o calitate foarte ridicată [283], deși variația între persoanele pentru care funcțiile HRTF sunt măsurate la intrarea canalului blocat a urechii este mai mică decât în același punct dar cu intrarea canalului deschis a urechii [284]. Spuma folosită pentru fixarea microfoanelor este transparentă din punct de vedere acustic, permițând trecerea sunetelor la urechi complet neschimbată. Fig. 4.5a arată plasarea microfoanelor în urechea dreaptă a subiectului.

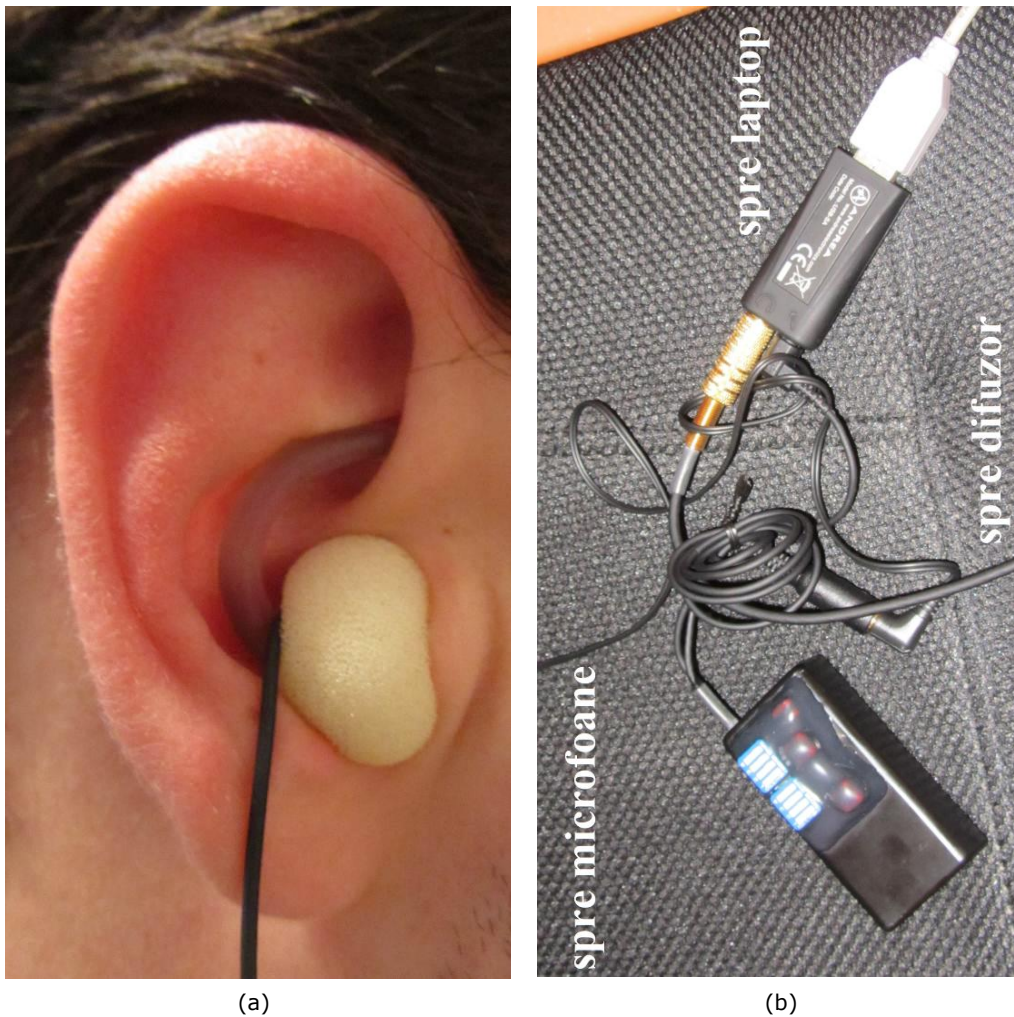


Fig. 4.5. (a) Plasarea microfonului în urechea dreaptă, respectiv (b) placa de sunet externă și modulul de baterie.

- Difuzoare stereo Genius SP-G16, cu un răspuns în frecvență între 20 Hz și 20 kHz, destinat difuzării semnalului de test pentru măsurări.

4.2.2. Procedura experimentală

Așa cum am menționat deja, experimentele au condus la funcții HRTF numai pentru direcții specificate de unghiuri de azimut între -90 și $+90$ de grade pentru un unghi de elevație de 0 grade. Rezoluția măsurărilor efectuate este de 5 grade, rezoluția necesară pentru funcțiile HRTF în planul orizontal conform [285], deși autorii din [286] pretind că rezoluția spațială a măsurărilor HRTF poate fi redusă de la 5 la aproximativ 10 grade fără introducerea unor diferențe perceptuale puternice. În cazul acestor măsurări rezoluția de 5 grade înseamnă un total de 37 de măsurări pentru subiectul de test. Subiectul a folosit un compas de câmp plasat orizontal între picioare pentru ajustarea simplă a unghiului de azimut.

Înainte de începerea procesului de măsurare, asistentul de măsurare trebuie să configureze corect interfața grafică utilizator, implementată în mediul de programare grafică LabVIEW de la firma National Instruments, prin setarea parametrilor semnalului de măsurare. Acești parametri sunt descriși în Cap. 4.2.3. Programul de control al măsurării, prezentat în Fig. 4.6, este, de fapt, un instrument virtual (VI).

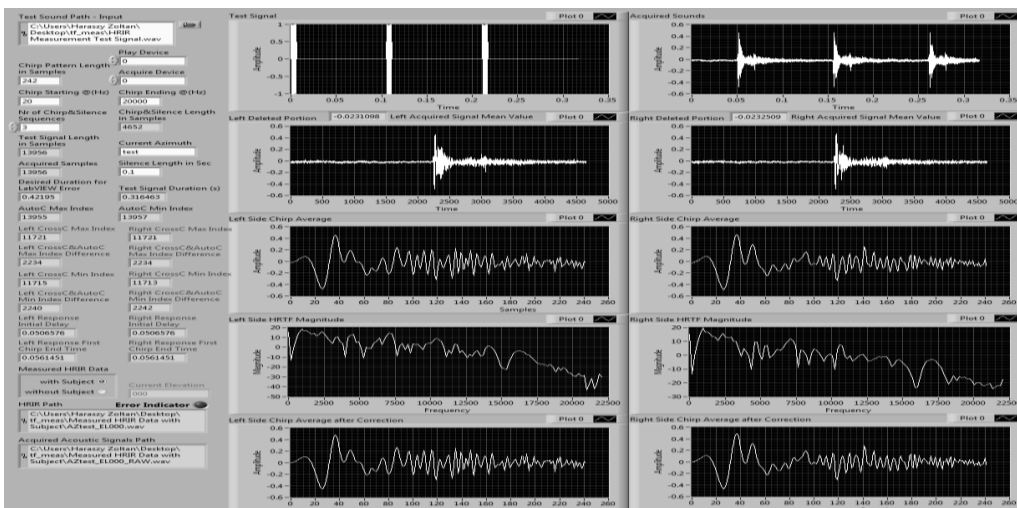


Fig. 4.6. Panoul frontal al instrumentului virtual pentru controlul total al procesului de măsurare implementat în LabVIEW.

Fig. 4.7 reprezintă un moment în cadrul măsurării funcțiilor HRTF cu subiectul așezat pe scaunul rotativ, având microfoanele binaurale montate în cele două urechi.

În cele ce urmează este descris cum anume a decurs măsurarea pentru o singură direcție. În primul pas, subiectul de test ajustează poziția scaunului rotativ la direcție dorită, după care în al doilea pas, persoana care asistă la întregul proces de măsurare execută instrumentul virtual. Cele două semnale binaurale înregistrate, folosind microfoanele binaurale, sunt salvate în formatul Microsoft Wave (.wav) pe hard discul laptopului pentru extragerea ulterioară a funcțiilor HRTF din semnalele de măsurare înregistrate. Această procedură este repetată pentru fiecare din cele 37 direcții de măsurare.



Fig. 4.7. Măsurarea funcțiilor HRTF cu subiectul așezat pe scaunul rotativ.

De remarcat este faptul că măsurările efectuate, descrise în cadrul acestui capitol, nu iau considerare efectul mediului acustic apropiat de capul ascultătorului, ca de exemplu, părul uman. Aspecte legate de influența acestui mediu apropiat pot fi consultate în [80], [287]-[288].

4.2.3. Egalizarea scenariului de măsurare

Un astfel de sistem de măsurare introduce propria filtrare asupra semnalului transmis de către difuzorul folosit, deci asupra semnalului înregistrat de microfoanele binaurale. Pentru eliminarea acestui dezavantaj, semnalele binaurale înregistrate trebuie să fie egalizate după efectuarea măsurărilor. Egalizarea unui sistem de măsurare înseamnă, de fapt, eliminarea sau compensarea influențelor introduse de părțile hardware, care alcătuiesc sistemul de măsurare. Acest lucru asigură reproducerea corectă în cazul folosirii funcțiilor HRTF măsurate într-un sistem de spațializare audio virtual.

Egalizarea sistemului de măsurare este efectuată, de obicei, în aceleași condiții în care au fost măsurate funcțiile HRTF. Conform celor raportate în [289], este posibil ca la nevoie această egalizare să fie făcută în alt mediu decât cel de măsurare a funcțiilor HRTF. În cadrul acestor măsurări, egalizarea s-a efectuat în aceleași condiții ca întregul proces de măsurare.

În vederea obținerii compensării (egalizării) corecte a semnalelor binaurale, există două metode care pot fi folosite.

Prima metodă este măsurarea răspunsurilor de referință pentru toate direcțiile de măsurare. Acest tip de compensare a fost folosită în [271]-[272]. În cadrul acestei proceduri, întregul sistem (difuzor, microfoanele binaurale, etc.) funcționează identic ca în cadrul oricărei măsurări normale, singura diferență fiind că nu există ascultător poziționat pe scaunul rotativ. Prin folosirea acestei metode pot fi limitate numeroase efecte negative, care scad acuratețea măsurărilor:

- Influența caracteristicilor de directivitate și frecvență a componentelor hardware folosite în măsurare;
- Influența caracteristicii acustice a camerei de test, respectiv reflexiile de la diferite aparate folosite, sunt semnificativ reduse.

A doua metodă este măsurarea funcției de transfer (răspuns la impuls) a fiecărui dispozitiv a întregului sistem de măsurare sau măsurarea răspunsului la impuls al întregului sistem, și determinarea răspunsului la impuls invers pentru folosirea acestuia în vederea egalizării funcțiilor HRTF necompensate.

În momentul de față, sistemul de măsurare prezentat nu permite măsurarea precisă a răspunsului la impuls al tuturor direcțiilor, deoarece nu există o metodă de poziționare precisă a microfoanelor binaurale în aceeași locații pentru fiecare direcție de măsurare, dar cu ascultătorul pe scaunul rotativ.

În concluzie, prima metodă nu poate fi aplicată în acest caz. Consecvent, a doua metodă este folosită pentru compensarea influențelor configurației hardware de măsurare prin măsurarea răspunsului la impuls al întregului sistem. Detaliile acestei calibrări sunt prezentate în Cap. 4.2.5, etapa a treia (C) de prelucrare.

4.2.4. Proiectarea semnalului de test

În ultimele decenii au fost dezvoltate numeroase semnale de excitație pentru măsurarea funcțiilor de transfer cu diferite forme de variație în timp: impulsuri, secvențe de lungime maximă, semnale periodice de lungime 2^N , sinusoide baleiate¹ (termenul în engleză este 'swept sine', 'sine sweep', sau simplu 'sweep'). Dintre aceste semnale de excitație, alegerea optimă este folosirea sinusoidei baleiate din mai multe cauze. Ele exclud toate produsele de distorsiune armonică din răspunsurile la impuls, lăsând doar zgomotul de fond ca limită pentru raportul semnal-zgomot realizabil. Semnalele sweep pot, deci, excita difuzorul cu putere semnificativ mai mare fără introducerea unor deformații în răspunsul la impuls măsurat. Măsurările bazate pe semnalele sweep sunt, de asemenea, semnificativ mai puțin vulnerabile la efecte nedorite a dispersiei temporale [290]. Spre deosebire de alte metode, utilizarea semnalelor sweep nu necesită o calibrare dificilă în vederea obținerii unor rezultate foarte bune, reprezentând cea mai bună tehnică de măsurare a răspunsurilor la impuls în medii silențioase și neocupate [291]. Cele mai recente progrese în domeniul măsurărilor răspunsurilor la impuls, folosind semnale sweep, sunt date în [292]-[295].

În experimentele noastre, semnalul de test este generat din LabVIEW. Semnalul sweep standard, implementat în LabVIEW este un semnal sweep liniar. În consecință, acest semnal sweep liniar este folosit.

Matematic, semnalul sweep cu frecvența variabilă liniar poate fi descris ca în ecuația (4.1):

$$x(t) = A \sin(f(t)) \quad (4.1)$$

Frecvența instantanee este dată de derivata argumentului sinusoidei. Dacă presupunem că frecvența variază liniar, începând de la f_1 și până la f_2 în timpul total de T , obținem (4.2), (4.3):

¹ În cadrul acestei teze, ne vom referi la semnalul sinusoidă baleiată ca semnal sweep.

$$\frac{d(f(t))}{dt} = \omega_1 + \frac{\omega_2 - \omega_1}{T} t \quad (4.2)$$

$$\frac{d(f(t))}{dt} = 2\pi \left(f_1 + \frac{f_2 - f_1}{T} t \right) \quad (4.3)$$

care este adevărat, dacă (4.4), (4.5):

$$f(t) = \omega_1 t + \frac{\omega_2 - \omega_1}{T} \cdot \frac{t^2}{2} \quad (4.4)$$

$$f(t) = 2\pi \left(f_1 t + \frac{f_2 - f_1}{T} \cdot \frac{t^2}{2} \right) \quad (4.5)$$

În cazul nostru, frecvența semnalului sweep variază liniar în timp de la 20 Hz până la 20 kHz. Durata T a semnalului sweep depinde de distanța dintre difuzor și capul subiectului. Așa cum rezultă din Fig. 4.2, în aceste măsurări, distanța este de 2 metri. Acest lucru înseamnă că timpul necesar semnalului sweep, ca să parcurgă această distanță, este de aproximativ 5.83 ms (care înseamnă 257 de eșantioane la rata de eșantionare de 44.1 kHz, considerând viteza sunetului la 343.2 m/s. Pentru a evita suprapunerea dintre reflexiile cauzate de subiectul de test și reflexiile cauzate de pereții înconjurători sau alte echipamente, am ales să folosim un număr total de 242 de eșantioane pentru semnalul sweep, care este echivalent cu o durată de aproximativ 5.49 ms. Deci, există un număr total de 242 de eșantioane la rata de eșantionare de 44.1 kHz. O tehnică similară este folosită în [271], unde semnalul sweep este filtrat cu un filtru cu răspuns finit la impuls în vederea creșterii componentelor de frecvență joasă pentru îmbunătățirea raportului semnal-zgomot [290], dar lăsând celelalte frecvențe nemodificate.

Fig. 4.8, respectiv Fig. 4.9, arată semnalul sweep în timp, respectiv în frecvență.

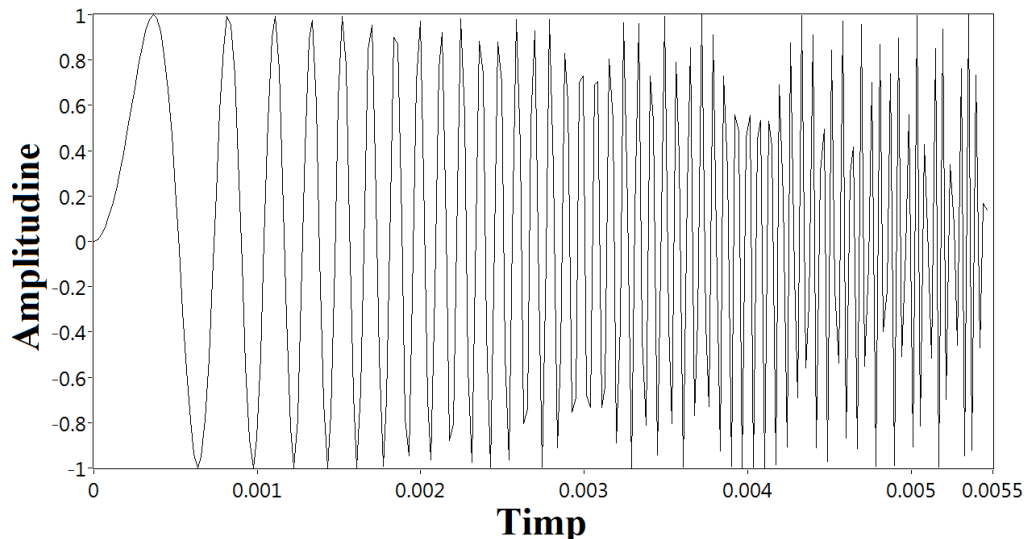


Fig. 4.8. Semnalul sweep în domeniul timp.

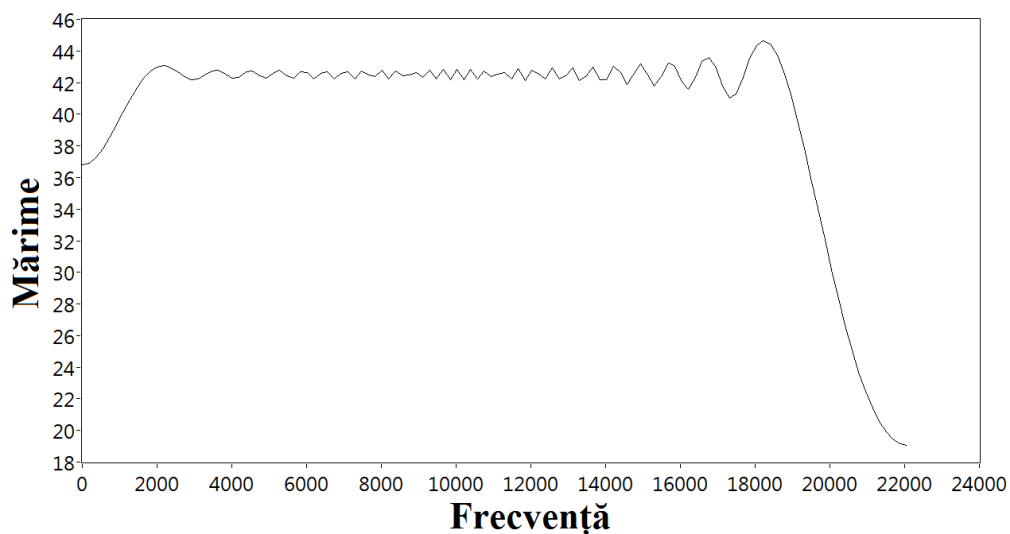


Fig. 4.9. Semnalul sweep în domeniul frecvență.

În cadrul măsurărilor curente, semnalul de test a fost compus din 10 secvențe compuse din semnal sweep și pauză, de lungime constantă, concatenate una după cealaltă. Semnalul de test folosit în cadrul măsurărilor este arătat în Fig. 4.10.

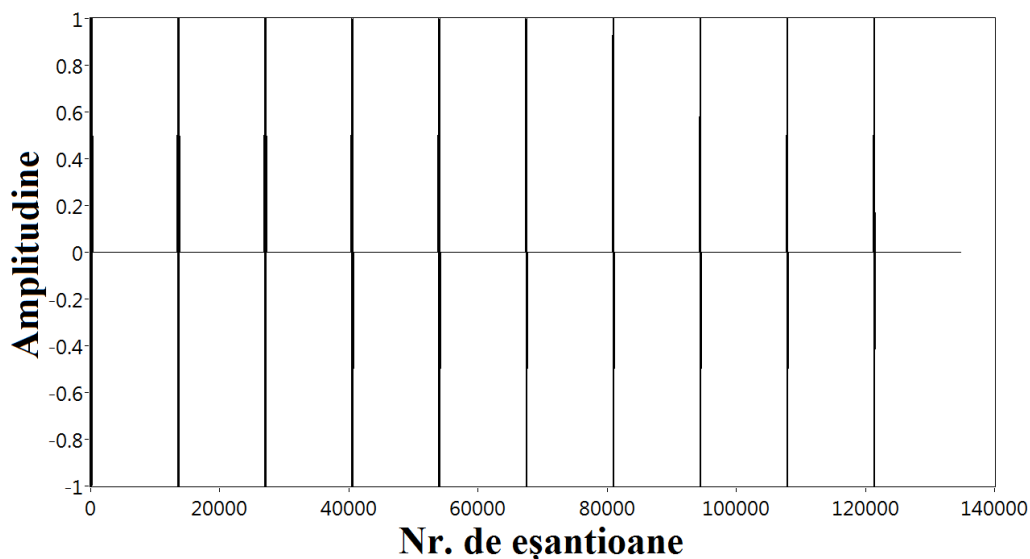


Fig. 4.10. Semnalul de test folosit în procesul de măsurare, compus dintr-un număr total de 10 secvențe de semnal sweep și pauză.

Semnalul sweep este repetat de mai multe ori în vederea medierii răspunsului sistemului în domeniul timp. Astfel, este îmbunătățit raportul semnal-zgomot al semnalelor măsurate. Deoarece răspunsul sistemului este convoluționat cu semnale reale, semnalele sweep nu pot fi oricât de lungi. Mai mult, dacă ele ar fi prea lungi, timpul necesar întregului proces de măsurare ar fi și mai mare. De

asemenea, prin folosirea semnalului de test, canalele de achiziții de date trebuie să fie declanșate doar o singură dată pentru fiecare direcție de măsurare, care evită posibila pierdere a sincronizării cauzate de pornirea și oprirea diferitelor dispozitive de mai multe ori.

Așa cum am menționat mai sus, semnalul sweep au avut o lungime de 242 de eșantioane. Liniștea între semnalele sweep a fost de o durată de 0.3 secunde pentru a se asigura că posibile reverberații se atenuează suficient de mult, care înseamnă un număr de 13230 de eșantioane la rata de eșantionare de 44.1 kHz. Deci, semnalul de test, arătat în Fig. 4.10, care include 10 secvențe de semnale sweep și liniște, are o lungime de 134720 de eșantioane, care înseamnă aproximativ 3.05 secunde la rata de eșantionare de 44.1 kHz.

Toți parametrii semnalului de test folosit în procesul de măsurare, ca frecvența de start, frecvența de stop, lungimea liniștei, numărul secvențelor compuse dintr-un semnal sweep și pauză, sunt reglate din interfața LabVIEW.

4.2.5. Extragerea semnalului util din semnalul înregistrat

Acest capitol conține pașii de procesare efectuați pentru fiecare direcție de măsurare. Toți acești pași sunt implementați în forma unei interfețe în LabVIEW, prezentată în Fig. 4.11.

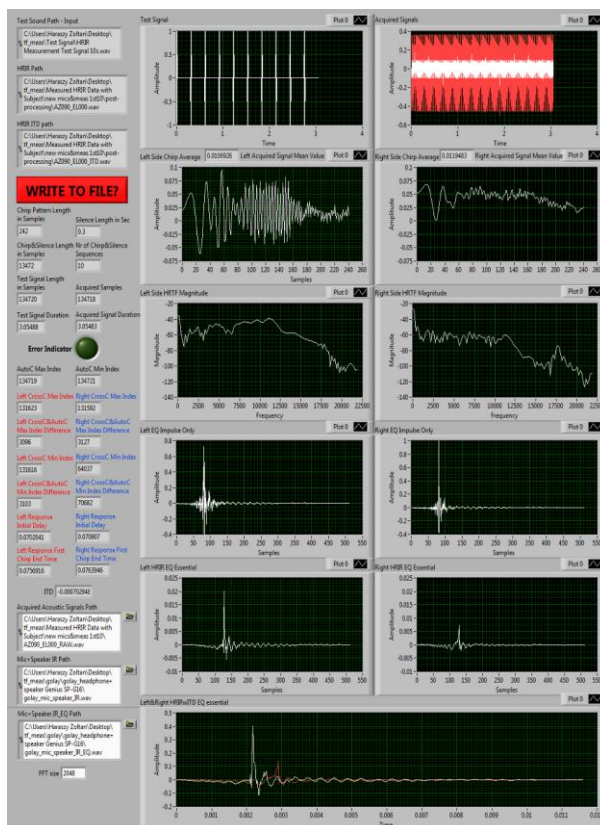


Fig. 4.11. Panoul frontal al aplicației de procesare a datelor măsurate.

Pentru fiecare dintre cele 37 de direcții de măsurare, au fost înregistrate două semnale acustice neprelucrate, care reprezintă versiunile filtrate ale semnalului de test emis de către difuzor (vezi Fig. 4.10), și includ funcțiile HRTF pentru fiecare direcție. Lungimea ambelor semnale înregistrate a fost de 134718 eșantioane.

Acest capitol conține o descriere detaliată a modalității de determinare a funcțiilor HRTF compensate. Procedura, ce urmează să fie descrisă, este valabilă pentru ambele semnale înregistrate.

În primul pas, trebuie să fie calculată media semnalelor sweep a tuturor celor 10 semnale sweep înregistrate (separate prin liniște) după extragerea lor din semnalele înregistrate. Al doilea pas conține determinarea funcțiilor HRTF necompensate din media semnalelor sweep calculată. În al treilea pas, răspunsul invers la impuls al sistemului de măsurare este obținut în vederea calibrării funcțiilor HRTF neegalizate. În final, urmează determinarea funcțiilor HRTF egalizate, folosind rezultatele obținute în al doilea, respectiv în al treilea pas.

Toți acești pași de procesare post-măsurare sunt implementați într-un instrument virtual dezvoltat în LabVIEW.

A. Determinarea mediei semnalelor sweep înregistrate, separate de liniște

Întreaga procedura de determinare a mediei aritmetice a semnalelor sweep înregistrate, separate prin liniște, este arătată în Fig. 4.12.

Pentru extragerea corectă a semnalelor sweep recepționate din semnalul de măsurare înregistrat, este necesară determinarea precisă a poziției fiecărui semnal sweep aflat în interiorul semnalului înregistrat. Pentru obținerea poziției primului eșantion din primul semnal sweep (din semnalul înregistrat), trebuie să calculăm valoarea minimă dintre următoarele două valori:

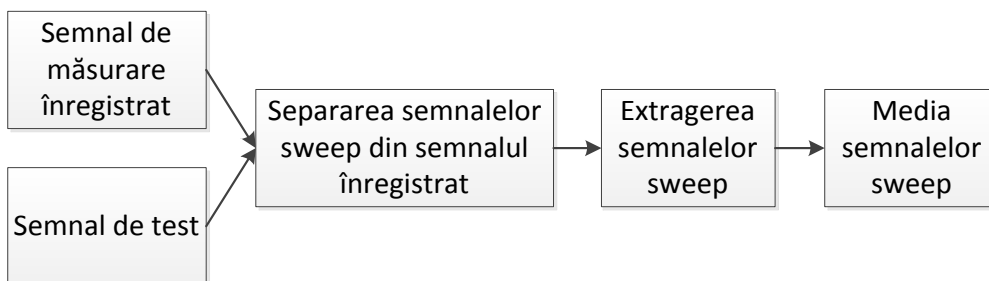


Fig. 4.12. Determinarea mediei semnalelor sweep din semnalul de măsurare înregistrat, respectiv din semnalul de test.

- Diferența dintre pozițiile unde este situată
 - valoarea maximă a autocorelației semnalului de test și
 - valoarea maximă a intercorelației dintre semnalul de test (cele 10 secvențe de semnal sweep și pauză) și semnalul înregistrat;
- Diferența dintre pozițiile unde este situată
 - valoarea maximă a autocorelației semnalului de test și
 - valoarea minimă a intercorelației dintre semnalul de test (cele 10 secvențe de semnal sweep și pauză) și semnalul înregistrat;

În vederea obținerii poziției ultimului eșantion din primul semnal sweep, trebuie să adăugăm lungimea unui singur semnal sweep la valoarea minimă obținută în prealabil, care ne va furniza poziția primului eșantion al primului semnal sweep recepționat. Odată ce este cunoscută poziția primului semnal sweep, putem

afla pozițiile primelor și ale ultimelor eșantioane ale fiecărui semnal sweep în parte, prin adăugarea la aceste valori lungimea unei singure secvențe de semnal sweep și pauză (în eșantioane), care este de 13230 eșantioane în cazul nostru. Următoarea etapă este determinarea poziției fiecărui semnal sweep recepționat dintre cele 10, și extragerea lor în vederea calculului mediei lor, denumită media semnalului sweep.

După efectuarea acestor operații se obțin semnale mediate arătate în Fig. 4.13, respectiv Fig. 4.14.

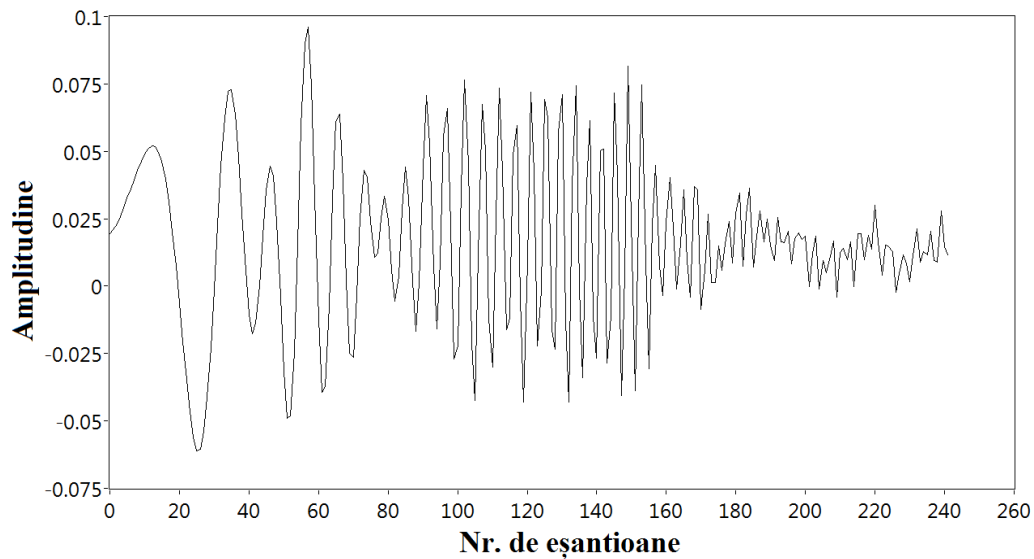


Fig. 4.13. Semnal sweep mediat aparținând direcției $(90^\circ, 0^\circ)$, canalul stâng.

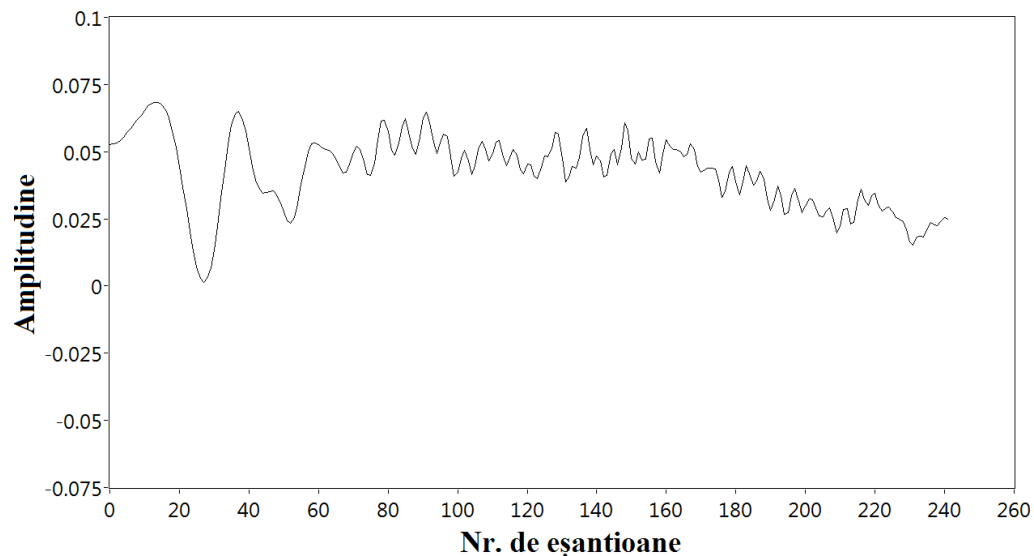


Fig. 4.14. Semnal sweep mediat aparținând direcției $(90^\circ, 0^\circ)$, canalul drept.

Aceste două semnale conțin perechea de funcții HRTF pentru direcția specificată de azimut +90 de grade, elevație 0 grade (dreapta subiectului).

De observat este faptul că:

- pozițiile relative ale eșantioanelor de început ale primelor semnale sweep pentru ambele părți, stânga și dreapta, pot fi folosite pentru determinarea timpilor de sosire ale semnalelor sweep recepționate. Diferența dintre acești doi timpi permit obținerea Diferenței Interaurale de Timp (DIT) dintre semnalele înregistrate, care este o sugestie de localizare majoră la frecvențe joase [296];
- media semnalelor sweep, calculată după extragerea semnalelor sweep recepționate, nu conține informații despre întârzierea temporală a semnalelor înregistrate cu microfoanele binaurale. Diferența acestor întârzieri temporale furnizează DIT. Această informație esențială pentru localizarea sunetelor trebuie să fie reconstruită după determinarea funcțiilor HRIR pentru a obține perechile de funcții HRIR corecte.

B. Determinarea funcțiilor HRTF necompensate

Întregul proces de calcul al funcțiilor HRTF necompensate, folosind un singur semnal sweep original și media semnalelor sweep calculate la pasul precedent, este prezentat în Fig. 4.15.

Pentru a calcula funcțiile HRTF neegalizate:

- Aplicăm transformata FFT pentru media semnalelor sweep, și respectiv pentru semnalul sweep original;
- Efectuăm împărțirea celor două transformate FFT rezultate la punctul precedent. Rezultatul acestei împărțiri reprezintă chiar o funcție HRTF neegalizată.

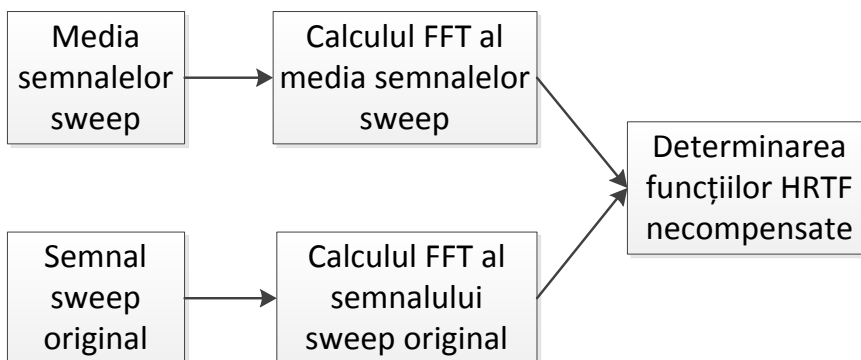


Fig. 4.15. Determinarea funcțiilor HRTF necompensate din media semnalelor sweep, respectiv din semnalul sweep original.

C. Determinarea răspunsului la impuls invers al sistemului de măsurare

Pașii pentru calculul răspunsului la impuls invers al sistemului de măsurare sunt reprezentați în Fig. 4.16.

Prima dată, pentru calculul răspunsului la impuls invers al sistemului, trebuie să aflăm răspunsul la impuls al lui (vezi Fig. 4.16). Pentru a afla răspunsul la impuls au fost folosite unelte de măsurare a funcției de transfer dezvoltate în cadrul proiectului RealSimple [297]. Dintre cele două metode prezentate, autorul a folosit metoda bazată pe semnale sweep datorită avantajelor [290]-[295] acestor tipuri de semnale față de metoda bazată pe codurile Golay [298].

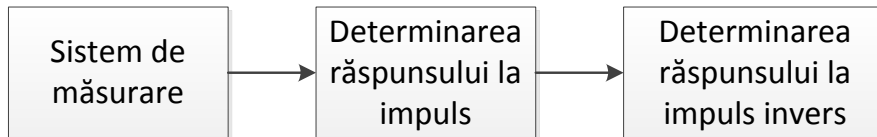


Fig. 4.16. Calculul răspunsului invers la impuls al întregului sistem de măsurare..

În cele ce urmează se prezintă detaliile determinării răspunsului la impuls al întregului sistem, folosind semnale sweep urmând procedura descrisă în [297]. Semnalul folosit este un semnal sweep cu frecvența crescătoare în mod exponențial de la ω_1 până la ω_2 de-a lungul a T secunde.

$$s(t) = \sin \left[K \left(e^{-\frac{n}{L f_s}} - 1 \right) \right] \quad (4.6)$$

unde $K = \frac{\omega_1 T}{\ln \frac{\omega_2}{\omega_1}}$ și $L = \frac{T}{\ln \frac{\omega_2}{\omega_1}}$.

Codul MATLAB `generate_sinesweeps.m` generează semnalul sweep corespunzător.

Proprietatea importantă a $s(t)$ este faptul că întârzierea de timp Δt_N dintre oricare eșantion t_0 și unul mai întârziat cu o frecvență instantanee de N ori mai mare decât frecvența instantanee la $s(t_0)$ este constantă:

$$\Delta t_N = T \frac{\ln(N)}{\ln \frac{\omega_2}{\omega_1}} \quad (4.7)$$

Această caracteristică implică faptul că după filtrarea inversă a răspunsului măsurat, componentele neliniare sunt plasate la locuri specifice în semnalul final. Deci, contribuția liniară a răspunsului poate fi separată de componentele neliniare.

Procedura de măsurare folosind semnalul sweep este prezentată în continuare:

1. Se generează semnalul sweep, folosind `generate_sinesweeps.m` cu parametrii $f_1 = 20$ Hz, $f_2 = 20$ kHz, $N = 16$.

2. Se deschide programul `sinesweeps.pd` în PureData, care este un limbaj de programare vizual dezvoltat de Miller Puckette în anii 1996.

3. Se asigură că programul nu este în mod de editare, și se bifează căsuța "compute audio" în fereastra principală.

4. Se ajustează volumul "Output Volume" astfel încât dacă se apasă butonul "Record Response To The Sine Sweeps", sistemul se comportă linear și semnalul de intrare nu este prea zgomotos.

5. Se asigură faptul că volumul de la intrare este cât mai aproape de nivelul maxim fără să se ajungă la limitare când se apasă butonul "Record Response To The Sine Sweeps".

6. Odată ce rezultatele sunt acceptabile, răspunsul este salvat pe disc, apăsând "Write Responses to Disk".

7. PureData creează un fișier pe disc, care poate fi redenumită cu o denumire sugestivă.

8. Se rulează programul MATLAB *sinesweeps_response.m* pentru analiza răspunsului măsurat. Ca rezultat, răspunsul la impuls al sistemului de măsurare este dat într-un fișier pe disc în format .wav.

A fost măsurat răspunsul la impuls al întregului sistem folosind semnale sweep prin plasarea fiecărui microfon la o distanță de 2 cm de la difuzor, așa cum este arătat în Fig. 4.17.



Fig. 4.17. Plasarea microfonului față de difuzor în cadrul determinării funcției de transfer.

Configurația hardware al întregului lanț de măsurare a rămas neschimbată, fiind identică cu cea prezentată în Cap. 4.2.1, respectiv arătată Fig. 4.4. Răspunsul la impuls obținut pentru canalul stâng este prezentat în domeniul timp în Fig. 4.18, respectiv în domeniul frecvență în Fig. 4.19. Rezultatele obținute pentru canalul drept sunt foarte similare.

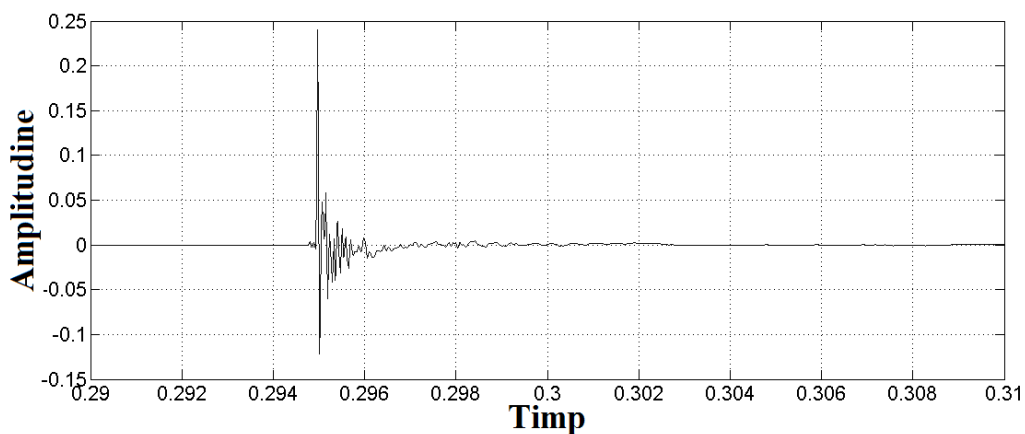


Fig. 4.18. Răspunsul la impuls măsurat pentru canalul stâng (domeniul timp).

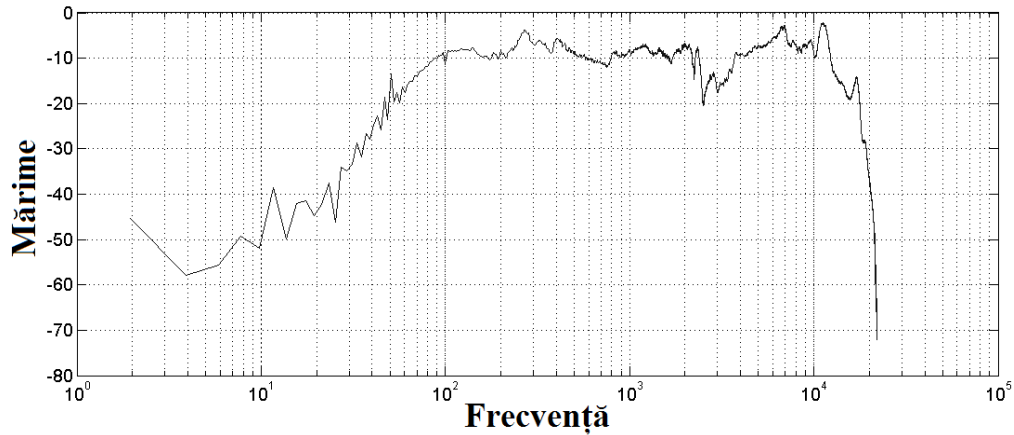


Fig. 4.19. Funcția de transfer măsurată pentru canalul stâng (domeniul frecvență).

După ce ambele răspunsuri la impuls sunt cunoscute, putem trece la determinarea răspunsurilor la impuls inverse, folosind codurile MATLAB furnizate de B. Gardner în [228], mai exact funcția MATLAB *invert.m*. O metodă utilă de egalizare este prezentată în [298], care se bazează pe date extrase din semnalul acustic ce urmează să fie egalizat. Răspunsul invers la impuls al canalului stâng este dat în Fig. 4.20.

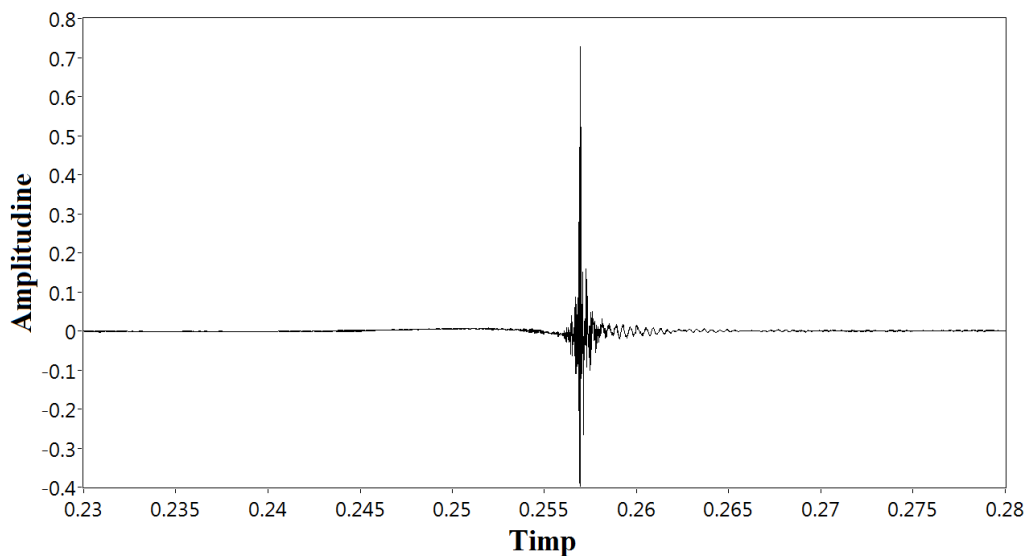


Fig. 4.20. Răspunsul invers la impuls al canalului stâng.

D. Determinarea funcțiilor HRTF compensate

Metoda folosită în cadrul acestor procesări pentru obținerea funcțiilor HRIR (funcții HRTF în domeniul timp) din funcțiile HRTF necompensate și din răspunsul invers la impuls, obținut în etapa precedentă, este prezentată în Fig. 4.21.

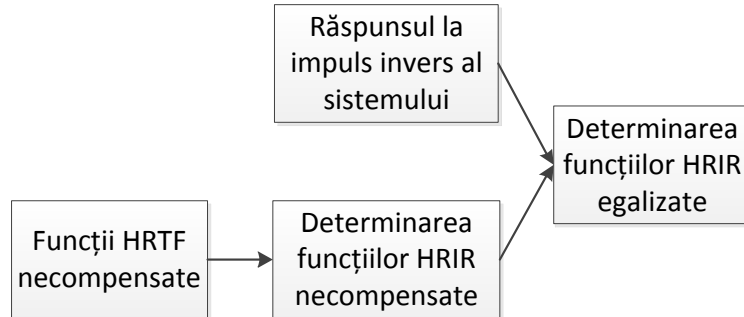


Fig. 4.21. Calculul funcțiilor HRTF calibrate folosind funcțiile HRTF neegalizate și răspunsul invers la impuls al sistemului de măsurare.

Funcțiile HRTF necompensate sunt obținute ca rezultat al etapei a doua de prelucrare. Prin calculul transformatei FFT inverse asupra lor, se obțin funcțiile HRIR necompensate. Următorul pas este folosirea rezultatului etapei a treia, răspunsul invers la impuls al întregului sistem de măsurare, împreună cu funcțiile HRIR compensate pentru determinarea funcțiilor HRIR compensate, care reprezintă de fapt variantele în domeniul timp ale funcțiilor HRTF. Ele reprezintă rezultatul final al procesării efectuate după încheierea procesului de măsurare.

După procesarea datelor măsurate, așa cum s-a explicat în acest capitol, am obținut 37 de funcții HRIR pentru ambele canale (37x2 funcții HRIR), corespunzător celor 37 de direcții de măsurare din planul orizontal (la unghiul de elevație 0 grade) cu valori ale unghiului azimut între -90 și +90 de grade cu o rezoluție de 5 grade între măsurări. Fig. 4.22 arată funcțiile HRIR obținute după procesarea descrisă pentru direcție specificată de perechea (90°, 0°).

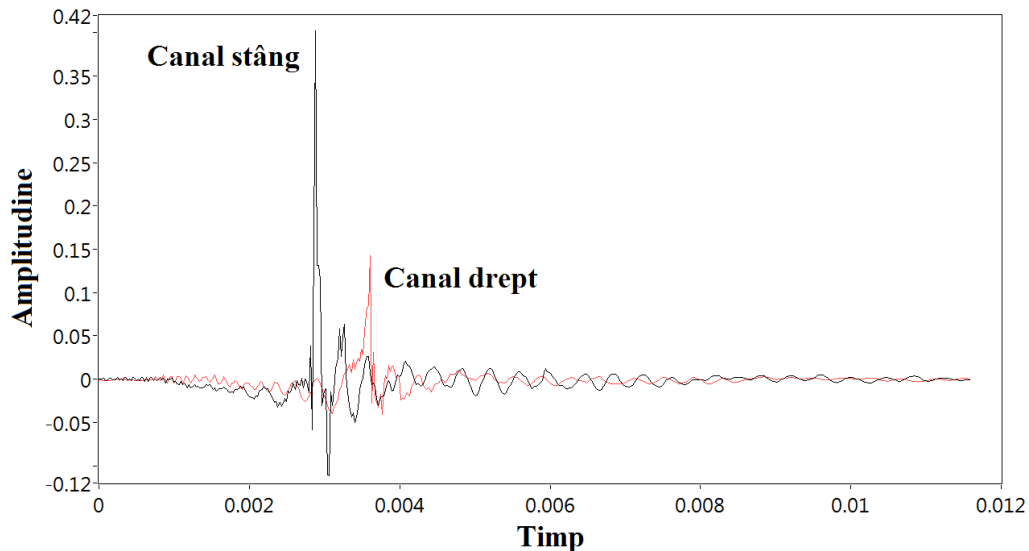


Fig. 4.22. Perechea de funcții HRIR obținute pentru direcția (90°, 0°).

4.3. Valori obținute pentru Diferența Interaurală de Timp

Teoria duplex (numită câteodată și sugestie azimutală), formulată de Lord Rayleigh este o teorie a localizării sunetelor la oameni bazată pe două sugestii binaurale, Diferența Interaurală de Timp, DIT (ITD, Interaural Time Difference) și Diferența Interaurală de Nivel, DIN (ILD, Interaural Level Difference) și presupunerea că capul are formă sferică fără urechi externe (pinnae). Mărimea DIT mai poate fi numită și Întârziere Interaurală de Timp (Interaural Time Delay). În cadrul acestei teze se va folosi prima denumire, adică DIT. Mărimea DIN mai este numită în unele lucrări și materiale ca Diferența Interaurală de Intensitate (Interaural Intensity Difference). Aceste două indicații sau sugestii, numite și sugestii interaurale (Interaural Cues), se bazează pe faptul că cele două urechi nu au aceeași poziție în spațiu, dar mai mult, sunt separate de un cap de dimensiune relativ mare [130], [300].

Cunoscând această separare, dacă sursa de sunet nu se află într-un punct din planul median (în acest caz distanța către cele două urechi este egală), distanța parcursă de undele de sunet, emise de sursa de sunet, până la urechea stângă, respectiv dreaptă a ascultătorului este diferită. Acest lucru are efectul că sunetul ajunge mai repede la urechea ipsilaterală (urechea mai aproape de sursa de sunet, urechea care se află pe o parte a capului cu sursa de sunet), și cu o anumită întârziere ajunge la urechea contralaterală (urechea mai îndepărtată de sursa de sunet, urechea care se află pe partea opusă cu sursa de sunet). Diferența dintre momentele de timp la care sosesc sunetele la cele două urechi este cunoscută ca DIT. În timp ce, cunoscând separația celor două urechi, când lungimea de undă a sunetului ce ajunge la urechi sunt relativ mici în comparație cu capul, el se comportă ca o umbră acustică, arătată în Fig. 4.23a, atenuând (în mod special la frecvențe înalte) nivelul presiunii de sunet (termenul corespunzător în limba engleză este Sound Pressure Level, SPL), al undelor care ajung la urechea contralaterală. Această diferență de nivele, între undele care se recepționează la urechea ipsilaterală, respectiv cea contralaterală, este denumită ca DIN.

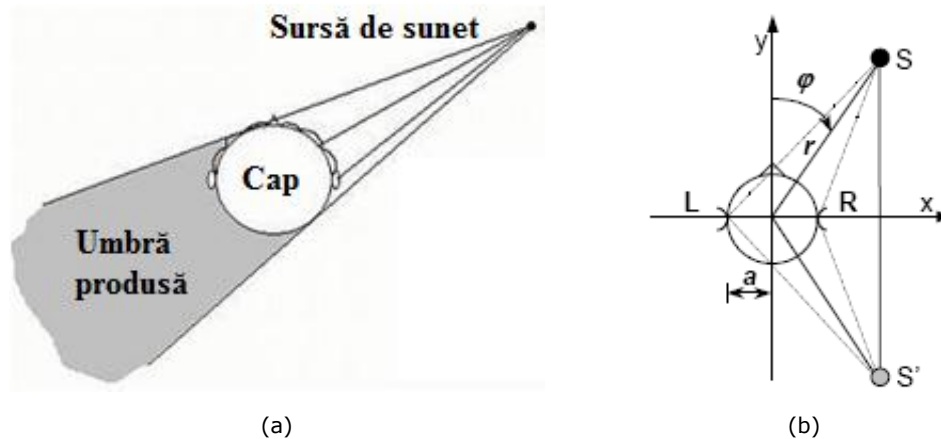


Fig. 4.23. (a) Umbră acustică produsă de capul ascultătorului, și (b) indicații similare produse de către sursele S , respectiv S' .

Dacă sursa de sunet se află în planul median, distanța de la sursa de sunet la cele două urechi este egală, cauzând ca sunetul să ajungă la cele două urechi

deodată. În plus, nivelul presiunii de sunet la ambele urechi este egală. Ca rezultat, atât DIT, cât și DIN, sunt zero sau foarte apropiate de zero. Dacă sursa se deplasează spre dreapta sau spre stânga, cele două mărimi cresc până când sursa ajunge exact la dreapta sau la stânga ascultătorului, respectiv la un unghi azimut de $\pm 90^\circ$. În acest caz, ambele ating valoarea lor maximă. Pentru un azimut de $\pm 90^\circ$, DIT se situează în jurul a 0.7 ms pentru un cap uman cu dimensiuni tipice [3]. Se poate observa, conform Fig. 4.23b, faptul că sursa de sunet S și imaginea ei S' față de axa interaurală introduc indicații similare.

Există numeroase formule de aproximare a DIT în literatură. În continuare sunt prezentate câteva dintre ele. În Tab. 4.2 sunt calculate erorile medii date de aceste formule față de valorile DIT obținute în cadrul măsurărilor. Cele mai mici erori au fost obținute pentru formula lui Woodworth, respectiv pentru legea sinusoidei, erorile fiind de 66.9227 μs , respectiv de 54.4349 μs .

Denumire	Formulă	Eroare medie
Woodworth [301]	$DIT = \frac{r}{c}(\theta + \sin \theta)$	66.9227 μs
-	$DIT = \frac{r}{c}(\theta + \cos \theta)$	229.52 μs
Legea sinusoidei [301]	$DIT = \frac{2r}{c} \sin \theta$	54.4349 μs
-	$DIT = \frac{r}{c}(1 - \sin \theta)$	345.706 μs

Tab. 4.2. Câteva modele ale DIT găsite în literatură.

Literatura furnizează și alte modele pentru reconstrucția DIT prin diverse procedee mai complicate în [119], [302]-[306]. Însă în majoritatea cazurilor cele două formule (Woodworth, legea sinusoidei) sunt folosite datorită simplității lor, respectiv erorilor mici obținute.

4.4. Concluzii. Probleme întâlnite și posibile îmbunătățiri

În acest capitol, a fost descris, în detaliu, procesul de măsurare a funcțiilor HRTF implementat și efectuat de către autor, începând cu locul de desfășurare, scenariul hardware, procedura experimentală, controlul prin software, proiectarea semnalului de test, procesarea datelor după măsurare, respectiv rezultatele obținute.

Putem concluziona faptul că scenariul de măsurare implementat în medii aproape fără ecouri este posibilă. Acest lucru asigură că măsurările sunt mult mai efective din punctul de vedere al costurilor. Funcțiile HRIR obținute vor fi folosite pentru reproducerea unor sunete virtuale spațiale ca parte a unei versiuni viitoare al prototipului prezentat în Cap. 7, proiectat pentru asistarea deplasării persoanelor cu handicap vizual.

Posibilele îmbunătățiri ale scenariului descris sunt următoarele:

- Se dorește modificarea scenariului de măsurare pentru a avea posibilitatea de a compensa funcțiile HRTF și prin prima metodă descrisă în Cap. 4.2.3;

- O modificare minoră a semnalului de test descris în Cap. 4.2.4, și anume intenția este înlocuirea semnalului sweep cu frecvența liniar variabilă cu un semnal sweep al cărei frecvență variază logaritm. Acest lucru este de dorit deoarece o rată de creștere exponențială are avantajul (printre altele) de a furniza mai mult timp și energie la frecvențe joase decât metoda folosită în cadrul acestor măsurări [307]. De asemenea, schimbarea frecvențelor de început, respectiv sfârșit este posibilă în funcție de răspunsul la impuls al sistemului de măsurare;
- Este dorit includerea unui suport pentru capul ascultătorului pe scaunul rotativ pentru a asigura o poziție stabilă a capului de-a lungul procesului de măsurare, deoarece funcțiile HRTF măsurate fără folosirea unui astfel de suport sunt foarte probabil afectate de o rată mare a erorii cauzate de posibilele mișcări involuntare ai capului ascultătorului [308]-[310];
- Se dorește, de asemenea, modificarea scenariului de măsurare pentru eliminarea dezavantajului de a putea măsura funcții HRTF doar în planul orizontal. Deci, este de dorit măsurarea funcțiilor HRTF în orice direcție dorită;
- Autorul dorește repetarea întregului proces de măsurare pentru același subiect, deoarece măsurarea curentă a fost efectuată cu microfoanele plasate la intrarea canalului deschis al urechii. Autorii din [311] raportează faptul că cel mai potrivit punct pentru măsurarea funcțiilor HRTF și pentru înregistrări binaurale este intrarea canalului blocat al urechii, deoarece sunetul în acest punct include toată informația spațială, cât și o cantitate minimă de informații individuale. Măsurările la intrarea canalului blocat al urechii sunt validate în [312]. Această îmbunătățire permite compararea celor două modalități de măsurare;
- În cazul selecției unor persoane vizate pentru diferite experimente este utilă evaluarea persoanelor vizate prin chestionare specifice proiectate în funcție de aplicația dorită. O astfel de selecție este descrisă în [313]. În cadrul experimentelor conduse de-a lungul elaborării acestei teze, nu s-a folosit nici un criteriu de selecție pentru alegerea persoanelor de test.

5. CONTRIBUȚII LA INTERPOLAREA FUNCȚIILOR HRTF CU AJUTORUL REȚELELOR NEURONALE ARTIFICIALE

Interfețe om-mașină, bazate pe informații audio de fidelitate spațială ridicată, pot fi create prin obținerea semnalelor acustice la urechile ascultătorului, care ar fi produse de sunete naturale emise de surse fizice. Acest lucru poate fi atins prin măsurarea proprietăților de filtrare dependentă de direcție a corpului ascultătorului (torso, cap și urechi). Următorul pas este filtrarea sunetelor cu perechea de filtre corespunzătoare (câte un filtru pentru fiecare ureche), după care semnalele acustice obținute sunt prezentate ascultătorului prin intermediul unor căști [14]-[15].

Filtrarea dependentă de direcție a corpului uman este descrisă de funcțiile HRTF ale ascultătorului (vezi Cap. 3.2). Ele sunt măsurate, de obicei (vezi Cap. 4), prin plasarea unor microfoane de dimensiune mică la intrarea urechii ascultătorului, și înregistrarea semnalului acustic emis de la diferite locații din jurul ascultătorului. Este imperativ ca ascultătorii să rămână nemișcați de-a lungul procesului de măsurare a funcțiilor HRTF. Rezoluția spațială a locațiilor, unde sunt executate aceste măsurări, este în mod rar mai mare de aproximativ 10 grade în azimut sau elevație. Datorită acestui neajuns în vederea obținerii unei RAV cu rezoluție acceptabilă, apare necesitatea interpolării funcțiilor HRTF măsurate.

5.1. Interpolarea funcțiilor HRTF

Tehnicile de interpolare ale funcțiilor HRTF pot fi evaluate în două moduri diferite. Funcțiile HRTF măsurate și interpolate pot fi comparate numeric (vezi Cap. 5.2.2, respectiv Cap. 5.2.3) sau psihoacustic (vezi Cap. 6.2), când este comparată percepția sunetelor filtrate cu funcțiile HRTF măsurate și interpolate. Dintre cele două metode enumerate, a doua este cea preferată, deoarece metoda psihoacustică poate valida, în mod echivoc, o metodă de interpolare propusă [314].

În literatura de specialitate sunt raportate numeroase metode de interpolare a funcțiilor HRTF, scopul acestora fiind reducerea sau eliminarea procesului complex de măsurare. Câteva dintre ele sunt raportate de: Freeland [315]-[319], Ajdler [285], [320]-[321], Takane [322]-[323], Duraiswami [324]-[325], Scarpaci [326], Xie [327], Park [328], McKanally [314], Diepold [329]-[330], Wang [331]-[332], Chen [333], Torres [334], Nishimura [335], Qinghua [336], Sousa [337], Otani [338], Breebaart [286].

5.2. Interpolarea funcțiilor HRTF folosind RNA

Pentru generarea sunetelor, ale căror surse virtuale sugerează, prin poziția lor, locația obstacolelor reale, pentru puncte ale spațiului 3D, sunt necesare funcțiile HRTF, prezentate în prealabil în Cap. 3.2.1. O pereche de funcții HRTF, formată din funcția HRTF pentru urechea stânga, respectiv una pentru urechea dreapta, pentru

aceeași direcție, reprezintă legătura dintre poziția în spațiul 3D a unei surse de sunet (plasate virtual) și presiunea acustică prezentă la urechile subiectului. Aceste funcții, care corespund fiecărui punct al spațiului 3D, pot fi determinate printr-o procedură destul de complicată de măsurare (vezi Cap. 4), care necesită mulți măsurători. Mai mult, această procedură trebuie să fie repetată pentru fiecare individ implicat în generarea surselor de sunet virtuale, datorită existenței diferențelor individuale indiscutabile dintre indivizi.

După cele mai recente informații ale autorului, folosirea rețelelor neuronale artificiale reprezintă o idee nouă în domeniul interpolării funcțiilor HRTF. Acest capitol conține câteva dintre ideile autorului referitoare la obținerea funcțiilor HRTF prin interpolare folosind rețele neuronale. În momentul de față interpolarea funcțiilor HRTF poate fi aplicată atât pentru cazul unei singure persoane, cât și pentru orice persoană. Această separare se manifestă în diferența structurii rețelelor neuronale folosite în vederea interpolării, și anume prin includerea parametrilor antropometrici, ca date de intrare în rețele (vezi Cap. 5.2.3). Dintre cele două variante menționate, până în momentul de față, ambele variante au fost validate prin criterii numerice (calculul unor erori dintre funcțiile HRTF interpolate, respectiv cele măsurate), în timp ce, doar prima a fost validată prin experimente psihoacustice (experimente de localizare).

În cele ce urmează sunt prezentate foarte concis aspecte legate de rețelele neuronale artificiale, dintre care în particular rețeaua de tip perceptron cu mai multe straturi în Cap. 5.2.1, metoda de interpolare propusă pentru o singură persoană în Cap. 5.2.2, respectiv metoda de interpolare propusă pentru orice persoană în Cap. 5.2.3.

5.2.1. Rețele neuronale artificiale

Rețele neuronale artificiale (RNA), denumite uneori procesoare paralele distribuite, neurocomputere sau modele conexiuniste, reprezintă modele matematice sau modele computaționale inspirate de structura și/sau aspectele funcționale ale rețelelor neuronale biologice caracteristice organismelor vii. Ca o definiție generală, se poate spune că RNA reprezintă un sistem de procesare a semnalelor, compus dintr-un număr mare de procesoare elementare interconectate, denumite neuroni artificiali sau noduri, și care cooperează pentru rezolvarea unor sarcini specifice [339]. Modalitatea de adaptare la condițiile specifice mediului constă în modificarea ponderilor asociate conexiunilor dintre neuroni și eventual a structurii RNA pe durata fazei de învățare sau antrenament. Rețelele neuronale moderne reprezintă unelte de modelare statistice neliniare ale datelor. Ele sunt folosite în vederea modelării legăturilor complexe dintre intrări și ieșiri sau în vederea găsirii tiparelor din date.

Astfel de modele conexiuniste oferă anumite avantaje, caracteristice sistemelor neuronale biologice și care nu sunt întâlnite în cazul sistemelor de calcul tradiționale, secvențiale [339]:

- O proprietate deosebit de importantă a RNA este aceea de a învăța și de a se adapta;
- Posibilitatea de a opera cu date imprecise;
- Capacitatea de generalizare, în sensul în care RNA operează corect și cu date de intrare care nu au fost prezentate în timpul procesului de antrenament;

- Datorită gradului ridicat de paralelism, funcționarea defectuoasă sau chiar pierderea unui număr de neuroni nu afectează semnificativ performanța sistemului global. RNA reprezintă, deci, sisteme tolerante la erori;
- Capacitatea de a aproxima orice funcție continuă neliniară cu gradul de acuratețe dorit. Astfel RNA pot fi folosite cu succes în modelarea sistemelor neliniare;
- Datorită numărului mare de intrări și ieșiri, RNA modelează cu ușurință sisteme multivariabile;
- Implementările hardware ale RNA, de exemplu prin intermediul circuitelor integrate pe scară largă (VLSI), fac posibilă utilizarea RNA pentru cazul aplicațiilor în timp real.

În cele ce urmează sunt prezentate RNA de tip perceptron cu un singur neuron, respectiv RNA de tip perceptron multistrat.

5.2.1.1. RNA de tip perceptron cu un singur neuron

Perceptronul simplu are o aplicabilitate practică limitată datorită valorii binare a ieșirii sau datorită imposibilității de clasificare a tiparelor neliniari. El a fost inventat de către Frank Rosenblatt în anul 1957 în Cornell Aeronautical Laboratory și reprezintă cea mai simplă rețea neuronală cu propagarea înainte a semnalului (feedforward). Perceptronul simplu constituie un punct de plecare în studiul perceptronului multistrat.

Arhitectura unui astfel de RNA este prezentată în Fig. 5.1.

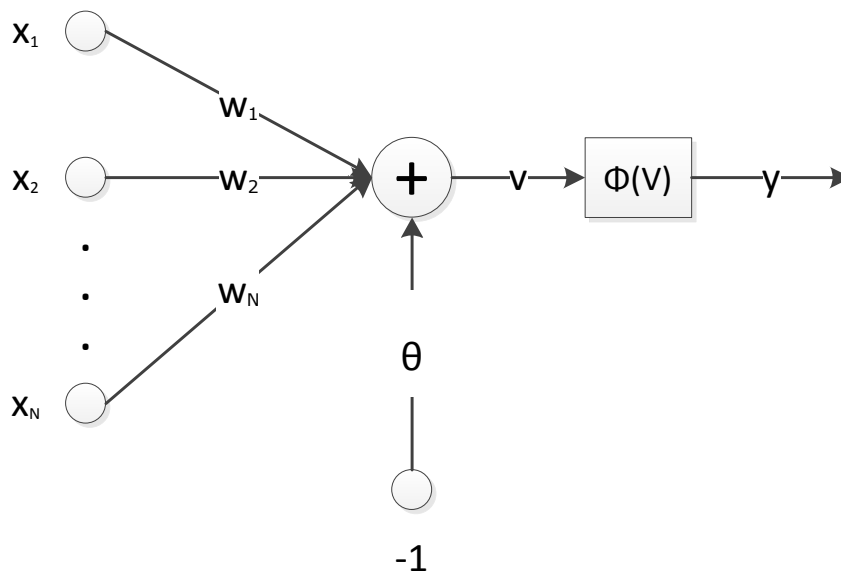


Fig. 5.1. Arhitectura perceptronului simplu.

Scopul perceptronului simplu este de a clasifica un set de stimuli exteriori. Funcționarea poate fi descrisă prin ecuațiile (5.1):

$$\left. \begin{array}{l} y = \Phi(v) \\ v = \sum_{i=1}^N w_i x_i - \theta \end{array} \right\} \Rightarrow y = \Phi \left(\sum_{i=1}^N w_i x_i - \theta \right) \quad (5.1)$$

unde x_i reprezintă intrarea i , w_i arată ponderea intrării i , v este rezultatul însumării intrărilor ponderate, Φ reprezintă funcția de activare, respectiv y constituie valoarea dată de către perceptron prin propagarea înainte a semnalului x .

5.2.1.2. RNA de tip perceptron cu mai multe straturi

Perceptronul multistrat (RNA-MLP, Multilayer Perceptron) reprezintă o generalizare a perceptronului simplu, prezentat în prealabil. El este o RNA de tip feedforward, arătată în Fig. 5.2, compusă din [339]:

- Un strat de intrare;
- Unul sau mai multe straturi ascunse;
- Un strat de ieșire.

Perceptronul multistrat este o RNA cu propagarea înainte a semnalului, care mapează seturi de date de intrare în seturi de date de ieșire corespunzătoare. O RNA-MLP constă din straturi multiple de noduri, unde fiecare strat este total conectat la următorul. Cu excepția nodurilor de la intrare, fiecare nod este un neuron (adică element de procesare) cu o funcție de activare. Această RNA folosește tehnica de învățare supervizată denumită antrenament cu propagarea înapoi a erorii.

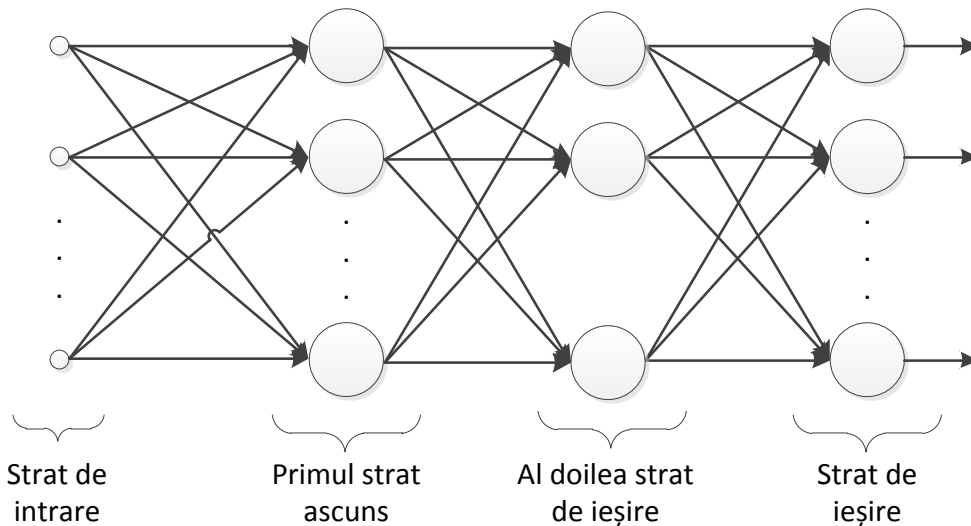


Fig. 5.2. RNA de tip perceptron cu două straturi ascunse.

Se deosebesc două etape în realizarea unei aplicații cu RNA. Prima reprezintă etapa de învățare sau etapa de antrenament, în care sunt aplicate perechi de tipare intrare-ieșire corect asociate, iar RNA își modifică parametrii liberi (ponderi și praguri) pentru a învăța aceste asociații. A doua etapă presupune utilizarea propriu-zisă a RNA, fază în care se pot aplica vectori diferiți de cei din

etapa de antrenament, urmând ca RNA, pe baza capacității sale de generalizare, să furnizeze un răspuns adecvat.

5.2.2. Structura de RNA pentru interpolarea funcțiilor HRTF pentru o singură persoană

O metodă bazată pe rețele neuronale artificiale (RNA) este propusă de către autorul tezei în [340] pentru determinarea funcțiilor HRTF pentru orice puncte din spațiul 3D. Metoda este valabilă pentru o singură persoană. Două RNA sunt necesare pentru generarea funcțiilor HRTF corespunzătoare fiecărei urechi în parte. Azimutul și înălțarea, mărimi care definesc poziția unui anumit punct al spațiului 3D, sunt aplicate la intrările RNA, în timp ce valorile, ce definesc funcțiile HRTF, sunt obținute la ieșirea RNA.

Implementarea unei RNA include dezvoltarea structurii ei și faza de testare. Pentru antrenarea rețelei sunt necesare perechi de coordonate ale punctelor (azimut θ , înălțare φ) și valorile funcțiilor HRTF, pentru un număr limitat de puncte din spațiul 3D. O astfel de bază de date poate fi obținută numai ca rezultat al unor măsurători experimentale. Există câteva baze de date disponibile pentru comunitatea științifică, care au fost prezentate în Cap. 3.2.2.

Structura RNA propuse este prezentată în Fig. 5.3.

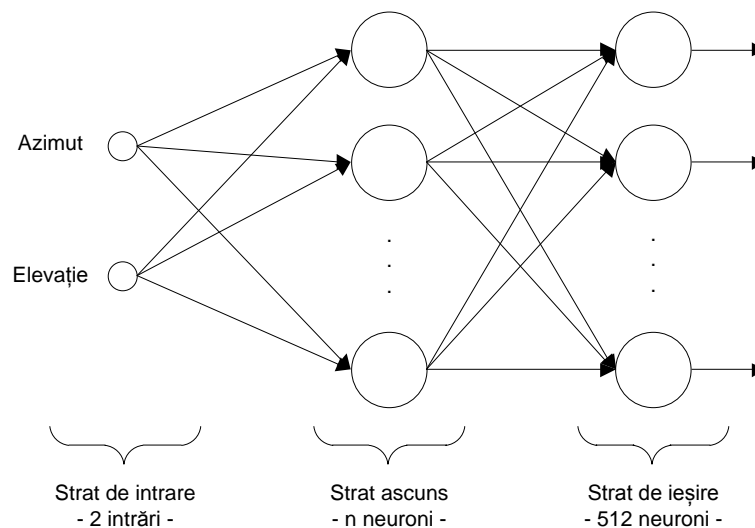


Fig. 5.3. Structura RNA propuse pentru interpolarea funcțiilor HRTF pentru o singură persoană.

Rețeaua constă din trei părți: un strat de intrare pentru nodurile sursă (2 intrări), un strat ascuns (n neuroni) și un strat de ieșire (512 neuroni). Considerând faptul că o pereche de astfel de rețele sunt necesare pentru o singură persoană corespunzătoare perechii de funcții HRTF, cele două intrări necesare, fiecărei RNA dintre acestea două, sunt azimutul și înălțarea sursei de sunet virtuale dorite. Fiecare dintre aceste două RNA furnizează, la ieșirile lor, un set de 512 valori, care reprezintă funcțiile HRIR pentru sursa de sunet virtuală dorită. Aceste funcții HRIR sunt perechi Fourier cu funcțiile HRTF (vezi Cap. 3.2.1). Numărul optim al neuronilor

din stratul ascuns este dificil de estimat. Acest număr poate fi stabilit prin evaluarea performanțelor rețelelor, sau prin folosirea algoritmilor genetici. În această fază, autorul a folosit un număr de n neuroni pentru stratul ascuns. Acest număr n de neuroni afectează performanțele RNA și este ales după experimentări atente. Rezultatele obținute sugerează ca n trebuie să fie undeva între 40 și 60.

Autorul a folosit baza de date LISTEN HRTF, descrisă în Cap. 3.2.2.1, care este o bază de date publică, pentru antrenarea și testarea RNA.

Rețeaua neuronală propusă este o rețea perceptron multistrat cu propagarea înainte a semnalului și cu propagarea înapoi a erorii (multilayer perceptron feedforward backpropagation network). Această rețea a fost selectată pentru simplitatea ei și, de asemenea, este considerată potrivită pentru începutul experimentelor autorului. În viitor, se va încerca luarea în considerare ale altor rețele neuronale.

După considerare atentă, setul de date existent pentru un singur subiect a fost împărțit în două seturi mai mici, așa cum urmează. Fiecare a patra pereche de funcții HRTF a fost selectată și setul de date obținute a fost folosită în faza de testare a RNA. Setul rămas, după extragerea funcțiilor HRIR selectate, a fost folosit ca set de date de antrenament pentru RNA. Preprocesarea datelor de intrare sau postprocesarea datelor de ieșire nu a fost folosită. Necesitatea procesării datelor de intrare și/sau ieșire ale rețelei rămâne o întrebare de viitor.

Funcția de transfer selectată pentru experimentele efectuate a fost: funcția de transfer 'hiperbolic tangent sigmoid' pentru stratul ascuns și funcția de transfer liniară pentru stratul de ieșire.

Întreaga rețea a fost antrenată utilizând funcția de antrenament *traindx*. Această funcție actualizează valorile ponderilor și ale pragurilor conform metodei gradientului descendent cu moment și rată de învățare adaptivă.

5.2.2.1. Evaluare prin criterii numerice

Experimentele au fost efectuate folosind uneltele dedicate rețelelor neuronale (Neural Network Toolbox), incluse în MATLAB, care este un limbaj de calcul tehnic de nivel înalt și, totodată, un mediu interactiv pentru dezvoltare de algoritmi, vizualizare de date, analiză de date, și calcul numeric. Toolbox-ul Neural Network extinde mediul MATLAB cu unelte pentru proiectarea, implementarea, vizualizarea, și simularea rețelelor neuronale. Rețelele neuronale sunt de neprețuit pentru aplicații, unde analiza formală este greu sau imposibil de folosit, ca recunoașterea tiparelor, controlul și identificarea sistemelor neliniare. De asemenea, Toolbox-ul Neural Network furnizează suport larg pentru modele de rețele demonstrate, de asemenea, pentru interfețe grafice de utilizator (GUI), care permit proiectarea și administrarea rețelelor. Proiectarea modulară, deschisă și elastică al acestui toolbox simplifică crearea funcțiilor și rețelelor speciale, care urmează să fie dezvoltate pentru un scop bine determinat.

Așa cum a fost menționat anterior, setul de date (funcții HRTF extrase din baza de date LISTEN HRTF) existent a fost împărțit în date de antrenament și date de test. Folosind acest mod de abordare, a fost posibilă evaluarea performanțelor rețelei neuronale propuse.

Inițial, cele două RNA sunt create. Fiecare dintre ele este antrenată, iar antrenamentul durează un număr de epoci folosind setul de date de antrenament obținute după împărțirea datelor unei persoane din baza de date LISTEN HRTF. Datele de antrenament constă din perechi de intrări (perechi azimut-înălțare) și

ieșiri (valorile corespunzătoare ale funcțiilor HRIR) cunoscute. Numărul optim de epoci de antrenament este greu de determinat și depinde de valorile inițiale ale valorilor ponderilor și pragurilor după crearea RNA. Experimentele au fost conduse pentru un număr de 56 de neuroni în stratul ascuns. După faza de antrenament, următoarea țintă este evaluarea performanțelor celor două rețele. Ca un criteriu de performanță, a fost aleasă eroarea medie pătratică (mean squared error, MSE). Eroare medie pătratică a fost calculată între valorile obținute la ieșirile rețelelor și valorile corespunzătoare ale datelor de test, alese din baza de date LISTEN HRTF, în mod aleator. Pentru a da măsură mărimii acestei erori, eroarea medie pătratică obținută a fost comparată cu valoarea absolută maximă ale ieșirilor rețelei. Concluzia autorului este, că această eroare este, de obicei, foarte mică, undeva în jur de 1-2%. Valorile erorii se schimbă în funcție de valorile inițiale ale ponderilor și pragurilor. De asemenea, depinzând de aceste valori inițiale, sunt necesare mai multe sau mai puține epoci de antrenament pentru obținerea valorilor mici ale erorii. Numărul maxim de epoci utilizate pentru antrenament a fost 1000, dar, în mod obișnuit, un număr de 500 de epoci sunt mai mult decât suficiente.

Un exemplu al erorii de antrenament obținute este prezentat în Fig. 5.4. Se poate observa, că eroarea descrește odată cu creșterea numărului de epoci de antrenament.

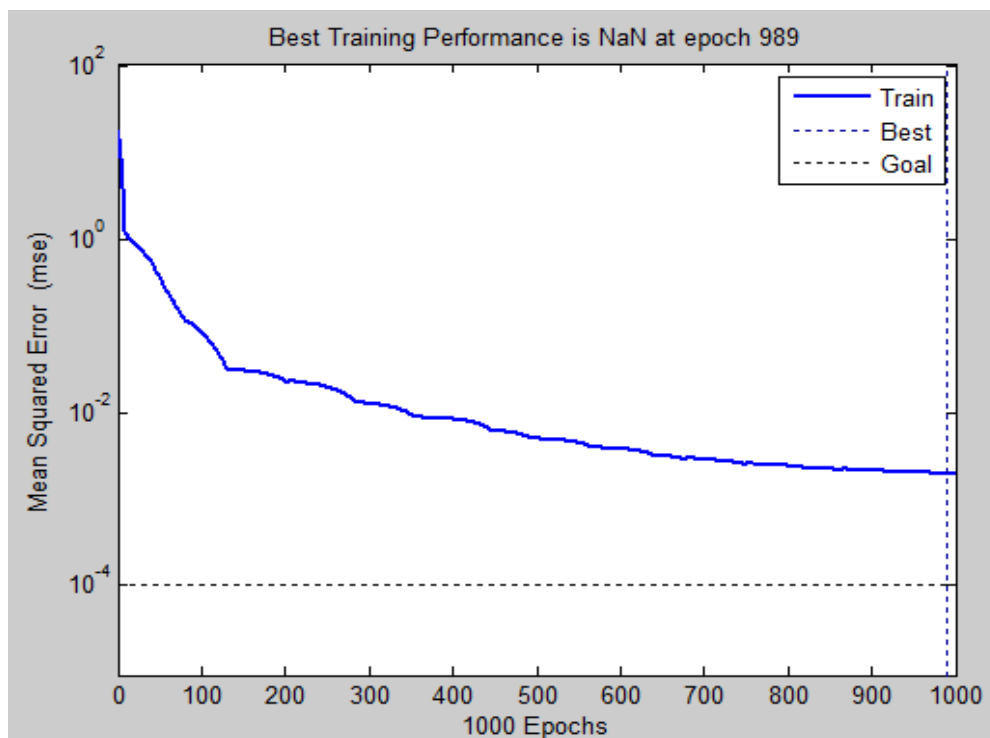


Fig. 5.4. Un exemplu de grafic reprezentând eroarea medie pătratică în raport cu numărul de epoci de antrenament.

Autorul dorește să prezinte, pe scurt, rezultate obținute de la 10 rețele neuronale, care toate au același structură, cea propusă, din Fig. 5.3. Valorile obținute confirmă afirmația de mai înainte, că un număr de 500 de epoci este

suficient pentru antrenamentul RNA. Fig. 5.5 prezintă eroarea medie pătratică obținută pentru cele 10 RNA corespunzătoare urechii din stânga pentru un anumit subiect de test. Antrenamentul fiecărei RNA a fost oprit după 100, 200, 300, 500, respectiv 1000 de epoci și eroarea medie pătratică a fost calculată, folosind setul de date de test ales. Se poate observa, că eroarea scade odată cu creșterea numărului de epoci de antrenament.

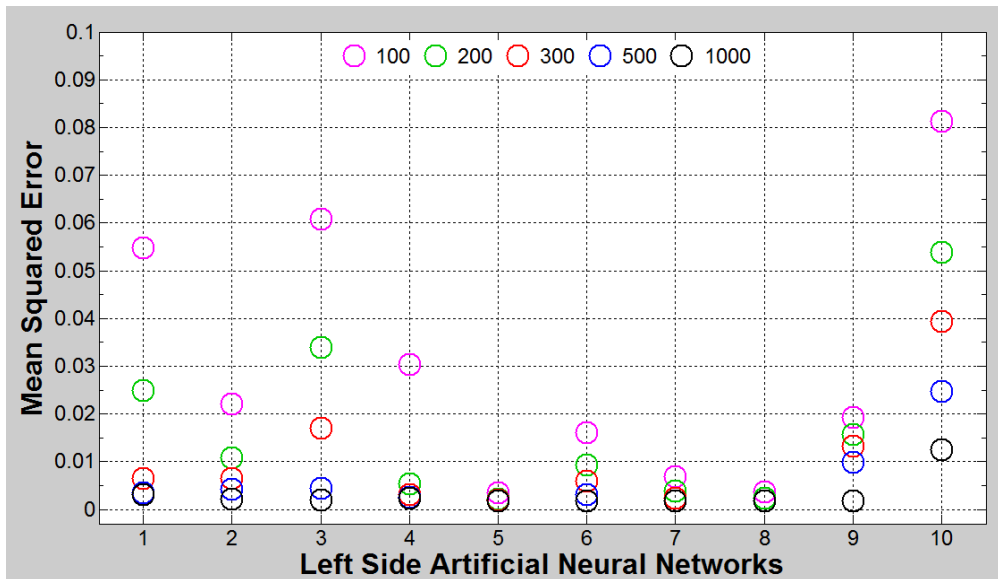


Fig. 5.5. Performanțe obținute în faza de testare a 10 rețele neuronale cu structura propusă.

Bazând pe rezultatele obținute în cercetarea descrisă în [340], se poate afirma că generarea funcțiilor HRTF pentru o anumită persoană este posibilă cu ajutorul rețelelor neuronale artificiale. Această metodă reduce semnificativ, durata procesului complex de măsurare necesar în momentul actual pentru obținerea funcțiilor HRTF pentru o anumită persoană. Dacă numărul epocilor de antrenament folosite pentru cele două RNA este adecvat, erorile obținute sunt mici. În concluzie, utilizarea RNA în generarea funcțiilor HRTF este o unealtă complementară procesului complex de măsurare, folosit de obicei.

Etapa următoare în dezvoltarea acestei soluții de interpolare este verificarea ei practică prin experimente de localizare conduse (vezi Cap. 6.3), în vederea evaluării funcțiilor HRTF obținute pe această cale. Acest aspect reprezintă principala, cea mai importantă metodă de evaluare a funcțiilor HRTF, deoarece dacă o pereche de funcții HRTF obținute prin soluția propusă chiar plasează sursa de sunet virtuală, obținută prin procesul de sinteză binaurală, la locația dorită, specificată prin perechea folosită pentru o anumită persoană, înseamnă că metoda de interpolare este validă și poate fi folosită cu succes în crearea unei Realități Acustice Virtuale (vezi Cap. 2.3) dedicată asistării deplasării persoanelor cu handicap vizual.

5.2.3. Structura de RNA pentru interpolarea funcțiilor HRTF pentru orice persoană

Ideea de a obține funcțiile HRTF într-un mod cât mai simplu, în orice punct din spațiul 3D și pentru orice persoană a căror parametrii antropometrici se cunosc, stă la baza acestei metode noi, bazate pe rețele neuronale. Primii pași în dezvoltarea acestei metode sunt descrise în [341]. Principala deosebire între această metodă nouă și cea prezentată în [340] și [241]-[243], constă în faptul că ea permite obținerea funcțiilor HRTF doar pentru o singură persoană în orice punct din spațiul 3D, cu condiția ca să fie disponibile un număr de funcții HRTF măsurate în prealabil pentru persoana respectivă pentru a putea antrena cele două RNA în vederea realizării interpolării între aceste puncte cunoscute. În timp ce, așa cum s-a menționat mai sus, metoda descrisă în [341] își propune obținerea lor pentru orice persoană și în orice punct din spațiul 3D, propunându-și reducerea/eliminarea timpului și a costurilor necesare în procesul complex de măsurare ale funcțiilor HRTF. Metoda este similară cu cea propusă de către autor în [340] și [241]-[243], valabilă doar pentru o singură persoană.

Similar cu metoda din [340] și [241]-[243], două RNA sunt necesare pentru obținerea funcțiilor HRTF, câte una pentru fiecare ureche. Aceste RNA au la intrările lor următoarele date:

- Anumiți parametrii antropometrici, prezentați în Tab. 5.1 și mai în detaliu în [227]. Acești parametrii definesc un anumit individ în mod unic;
- Azimut și înălțare, care definesc poziția unui anumit punct (direcție) în spațiul 3D.

Ieșirile celor două RNA furnizează o mulțime de valori, care definesc o pereche de funcții HRTF.

Implementarea RNA include următoarele etape: proiectarea structurii și faza de antrenament. Pentru faza de antrenament, datele necesare sunt alcătuite din:

- Parametrii antropometrici anterior menționați, respectiv direcția (perechea azimut-înălțare) dorită, la intrările RNA;
- Valorile funcțiilor HRIR, la ieșirile RNA.

Procesul de antrenament este realizat pentru un număr limitat de puncte ale spațiului 3D, pentru care funcțiile HRIR sunt disponibile din baza de date folosită. Cum s-a mai menționat, aceste funcții HRIR pot fi obținute numai ca rezultat al unor măsurători experimentale, disponibile publicului interesat. Principalele baze de date sunt prezentate în Cap. 3.2.2.

Așa cum s-a menționat mai sus, două RNA sunt necesare pentru generarea unei perechi de funcții HRIR/HRTF (una pentru fiecare ureche). Structura celor două RNA sunt identice și sunt prezentate în Fig. 5.6. Rețeaua include trei părți: stratul de intrare ($21 + 2 = 23$ intrări), stratul ascuns (n neuroni) și stratul de ieșire (512 neuroni). Datele de intrare necesare sunt formate din 21 de parametrii antropometrici (aplicați la intrări în ordinea din Tab. 5.1), respectiv poziția (formată din azimutul și elevația) sursei de sunet virtuale din spațiul 3D. Ca rezultat, fiecare dintre rețelele neuronale furnizează un set de 512 valori, reprezentând funcțiile HRIR particularizate pentru sursa de sunet virtuală. Numărul optim al neuronilor din stratul ascuns este greu de determinat. Acest număr poate fi estimat experimental prin modificarea numărului de neuroni din stratul ascuns și evaluarea performanțelor rețelelor obținute. Rezultatele obținute sugerează ca acest număr n de neuroni trebuie să fie în intervalul 40 și 60.

Denumire	Parametru
x1	Head width
x2	Head height
x3	Head depth
x4	Pinna offset down
x5	Pinna offset back
x6	Neck width
x7	Neck height
x8	Neck depth
x9	Torso top width
x10	Torso top height
x11	Torso top depth
x12	Shoulder width
x13	Head offset forward
x14	Height
x15	Seated height
x16	Head circumference
x17	Shoulder circumference
d1	Cavum concha height
d2	Cymba concha height
d3	Cavum concha width
d4	Fossa height
d5	Pinna height
d6	Pinna width
d7	Integral incisure width
d8	Cavum concha depth
θ1	Pinna rotation angle
θ2	Pinna flare angle

Tab. 5.1. Parametrii antropometrici disponibili pentru fiecare subiect folosit în procesul de antrenament/test (datele marcate nu au fost disponibile pentru fiecare subiect din baza de date folosită).

Autorul a folosit, din nou, baza de date LISTEN HRTF, care este o bază de date HRTF publică, pentru antrenarea și testarea RNA. Ea este descrisă în Cap. 3.2.2.1.

Rețeaua neuronală propusă este o rețea neuronală de tip perceptron multistrat cu propagarea înainte a semnalului și cu propagarea înapoi a erorii (multilayer perceptron feedforward backpropagation network). Această rețea a fost aleasă pentru simplitatea ei și, de asemenea, este considerată potrivită pentru începutul experimentelor autorului. În viitor, se va încerca folosirea ale altor rețele neuronale pentru scopul propus.

După considerație atentă, baza de date publică folosită a fost împărțită în două părți mai mici, așa cum urmează. Astfel, seturile de funcții HRIR disponibile pentru persoanele cunoscute din baza de date ca IRC_10XX, unde XX ia valori între 08 și 49, rezultând un număr total de 35 de persoane, au fost alese să fie utilizate pentru antrenarea celor două RNA. În timp ce, restul datelor disponibile pentru persoanele cunoscute ca IRC_10XX, unde XX ia valori între 50 și 59, au fost folosite ca date de test pentru cele două RNA (persoanele denumite IRC_10XX, unde XX ia valori între 02 și 07, nu au fost folosite din cauza unor parametrii antropometrici

lipsă). În experimentele realizate, preprocesarea datelor de intrare sau postprocesarea datelor de ieșire nu a fost folosită. Necesitatea procesării datelor de intrare și/sau ieșire ale rețelei rămâne o întrebare de viitor.

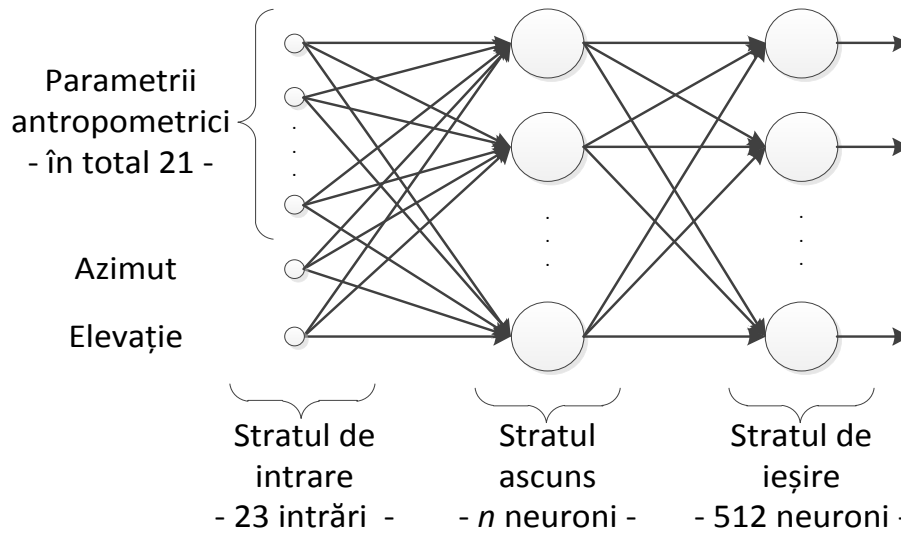


Fig. 5.6. Arhitectura RNA propuse pentru interpolarea funcțiilor HRTF pentru orice persoană.

5.2.3.1. Evaluare prin criterii numerice

Experimentele, la fel ca cel din [340], au fost efectuate în Toolbox-ul Neural Network, inclus în MATLAB.

Mai multe detalii despre rețeaua neuronală sunt prezentate în cele ce urmează. Funcția de transfer aleasă, pentru experimentele conduse, sunt cele din [340], adică funcția de transfer 'hiperbolic tangent sigmoid' pentru stratul ascuns și funcția de transfer liniară pentru stratul de ieșire. Întreaga rețea neuronală a fost antrenată utilizând funcția de antrenament de rețea *trainrp*. Această funcție actualizează valorile pragurilor și ale ponderilor bazându-se pe algoritmul resilient backpropagation (Rprop), implementat în Neural Network Toolbox al MATLAB.

Așa cum a fost menționat anterior, setul de date (funcțiile HRIR) existente a fost împărțit în date de antrenament și date de test. Folosind acest mod de abordare, a fost posibilă evaluarea performanțelor rețelei neuronale propuse.

După crearea celor două RNA, fiecare dintre ele a fost antrenată pentru un număr de epoci folosind datele de antrenament, care constau din perechi de intrări (parametrii antropometrici și perechea azimut-înălțare) și ieșiri (valorile ale eșantioanelor corespunzătoare ale funcțiilor HRIR) cunoscute. Procesul de antrenament a fost executat până când eroarea a coborât sub un anumit prag. Pragul de eroare acceptat în cadrul acestor experimente a fost 3-5%. Această eroare a fost obținută pentru datele de test. Experimentele a fost realizate pentru un număr de 50 de neuroni în stratul ascuns.

Un exemplu al erorii de antrenament obținute este prezentat în Fig. 5.7. Se poate observa că eroarea scade odată cu creșterea numărului de epoci.

După faza de antrenament, următoarea țintă a fost evaluarea performanțelor celor două RNA. Criteriul de performanță ales a fost eroarea medie

pătratică (MSE), la fel ca în [340], care a fost calculată între valorile obținute la ieșirile RNA și valorile corespunzătoare din setul de date folosit pentru test. Pentru a evalua mărimea acestor erori, autorul a comparat această eroare medie pătratică cu valoarea absolută maximă a ieșirilor rețelei. În funcție de valorile inițiale ale ponderilor și pragurilor, numărul de epoci de antrenament necesar pentru obținerea unor erori mici poate varia, dar această variație nu este una semnificativă. Numărul maxim de epoci folosit pentru antrenament a fost 150, dar de obicei un număr de 100 de epoci este mai mult decât suficient.

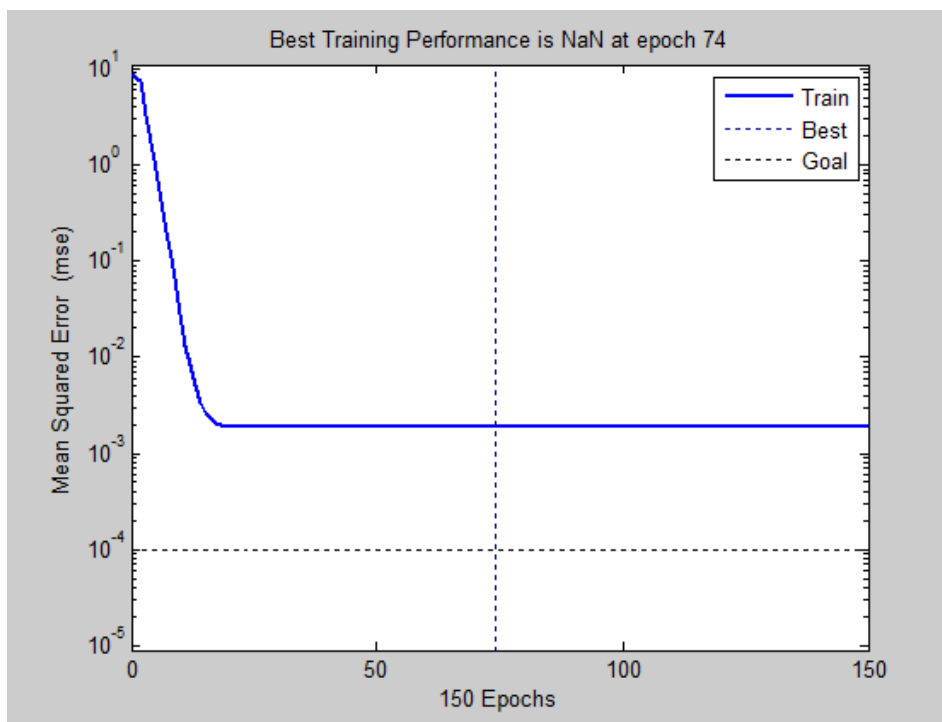


Fig. 5.7. Un exemplu al dependenței erorii de antrenament de numărul de epoci.

Autorul dorește să sublinieze faptul că dependența erorii obținute de valorile inițiale ale pragurilor și ponderilor. Această dependență se diminuează dacă rețelele sunt antrenate pentru un număr mare de epoci (peste 1000).

În cele ce urmează se prezintă, pe scurt, rezultatele obținute pentru 10 rețele neuronale, care toate au aceeași structură, cea propusă. Singurele diferențe dintre ele constau din valorile inițiale ale pragurilor și ponderilor înainte de procesul de antrenament.

Fig. 5.8 prezintă eroarea MSE calculată pentru 10 RNA corespunzătoare urechii din dreapta pentru un anumit subiect. Pe axa orizontală, sunt reprezentate în fiecare din cele 10 coloane erorile obținute pentru una dintre cele 10 RNA evaluate. Procesul de antrenament a fost oprit pentru fiecare dintre aceste RNA după 10, 20, 30, 50, respectiv după 100 de epoci, și eroarea MSE corespunzătoare a fost calculată folosind setul de date de test. Erorile MSE sunt reprezentate pentru un număr diferit de epoci (faze de evaluare) prin cercuri de culoare diferită (roz – 10 epoci, verde – 20 de epoci, roșu – 30 de epoci, albastru – 50 de epoci, negru – 100

de epoci). Pe axa verticală, erorile MSE sunt reprezentate pentru fiecare etapă de evaluare.

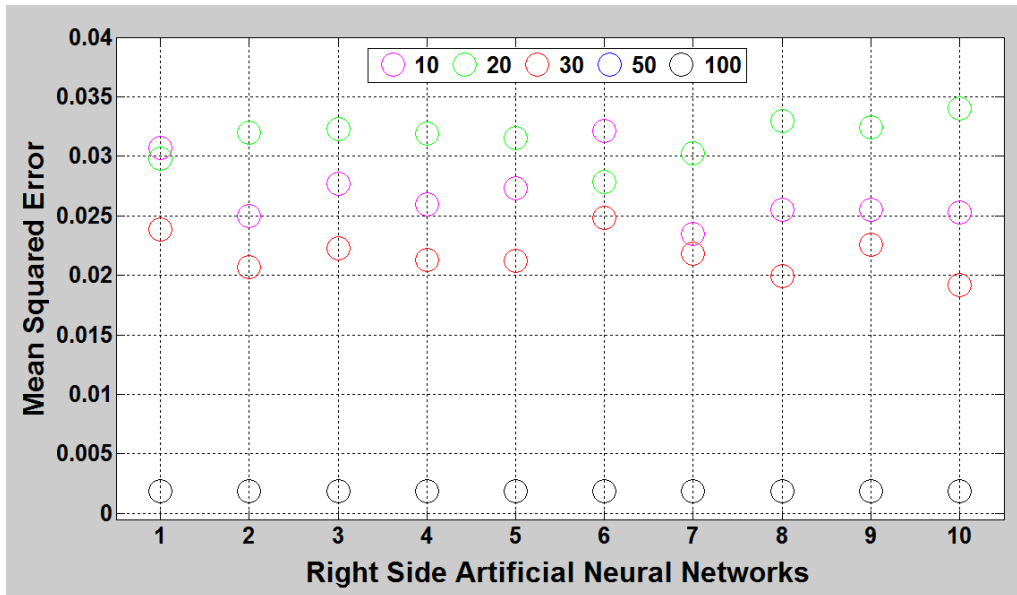


Fig. 5.8. Performanțele obținute după procedura de testare ale celor 10 RNA cu structura propusă.

Se poate observa că eroare MSE pentru RNA investigate variază din cauza valorilor inițiale ale ponderilor și pragurilor. De asemenea, se menționează că în fiecare coloană există doar 4 cercuri (în loc de 5), din cauza faptului că cele două cercuri (albastru – 50 de epoci, respectiv negru – 100 de epoci) se suprapun, însemnând că cele două erori obținute sunt de o mărime foarte apropiată. Se poate observa, că eroarea MSE calculată scade odată cu creșterea numărului de epoci de antrenament. Valorile rezultate confirmă faptul că un număr de 100 de epoci de antrenament sunt suficiente pentru obținerea pragului de eroare acceptată de 3-5%.

Bazând pe rezultatele obținute în cercetare condusă, se poate afirma că generarea funcțiilor HRTF pentru o anumită persoană este posibilă cu ajutorul rețelelor neuronale artificiale. Această metodă elimină sau, cel puțin, reduce semnificativ, durata procesului complex de măsurare necesar actual pentru obținerea funcțiilor HRTF pentru un anumit subiect. Dacă numărul epocilor de antrenament folosite pentru cele două RNA este adecvat, erorile obținute sunt suficiente de mici pentru aplicații practice. În concluzie, utilizarea RNA în generarea funcțiilor HRTF oferă o alternativă bună procesului complex de măsurare.

Rămâne însă de văzut dacă funcțiile HRTF astfel obținute (funcții HRTF cu un prag de eroare acceptată de 3-5%), prin metoda propusă, oferă într-adevăr percepția sursei de sunet virtuale dorite în spațiul auditiv al unei persoane particulare. Cu alte cuvinte, dacă sunetul obținut, prin filtrarea sunetului monaural cu funcțiile HRTF, vine cu adevărat din, sau cel puțin aproape de, direcția specificată de perechea azimut-elevație la intrările RNA. Versiunea simplificată a acestui experiment de localizare, valabilă doar pentru un singur subiect, este descrisă în [241]-[243], iar rezultatele obținute sunt prezentate în Cap. 6.3.

Ca o idee de viitor, autorul își propune proiectarea unor rețele neuronale noi, capabile să furnizeze direct semnalele acustice pentru una sau mai multe persoane, în loc de obținerea funcțiilor HRTF la ieșirile RNA. Astfel, ar fi posibilă obținerea directă, eficientă și rapidă, ale semnalelor acustice pentru orice poziție a sursei de sunet virtuale dorite și pentru orice persoană de interes. Această variantă necesită o RNA cu mai multe intrări și mai mulți neuroni în stratul ascuns.

5.3. Concluzii. Probleme întâlnite și posibile îmbunătățiri

Autorul consideră că față de metodele de interpolare enumerate în Cap. 5.1, metodele propusă bazate pe RNA, descrise în Cap. 5.2.2, respectiv Cap. 5.2.3, sunt metode convenabile de obținere a funcțiilor HRTF. Dificultatea lor constă în alegerea corectă a numărului de neuroni în stratul ascuns, respectiv în antrenarea corespunzătoare ale RNA. Odată ce RNA au fost antrenate corect, ele reprezintă o metodă foarte simplă și convenabilă de a obține, în mod rapid și eficient, funcțiile HRTF ale uneia sau mai multor persoane, depinzând de varianta folosită. Acest lucru permite generarea unei RAV în mod simplu, și în consecință, o semnalizare eficientă a obstacolelor, dacă metoda este folosită într-un dispozitiv de asistare a deplasării persoanelor cu handicap vizual.

Autorul dorește ca pe viitor să verifice aceste funcții HRTF obținute prin interpolare printr-o metodă numerică mai puternică decât MSE. Metoda vizată este cea propusă de Lee într-un studiu foarte recent [344], care despre un criteriu de evaluare a diferențelor dintre funcțiile HRTF măsurate, respectiv cele interpolate, pentru estimarea diferențelor auditive în localizarea sunetelor. Acest criteriu de performanță oferă o posibilitate importantă de evaluare a funcțiilor HRTF obținute prin interpolare.

6. EXPERIMENTE DE LOCALIZARE. UNELTE FOLOSITE

În cadrul acestui capitol sunt descrise experimentele de localizare conduse pentru validarea, evaluarea și compararea unor soluții propuse în cadrul tezei. Experimentele de localizare, așa numitele criterii psihoacustice, reprezintă cele mai importante criterii de validare în domeniul psihoacusticii. Un astfel de experiment poate dovedi dacă soluție propusă poate fi folosită, în mod eficient, în practică uzuală sau nu. Ele sunt o parte esențială a validării oricăror metode noi care vizează obținerea prin diferite modalități ale funcțiilor HRTF.

Cap. 6.1 discută necesitatea egalizării căștilor folosite în cadrul experimentelor de localizare conduse, respectiv motivează decizia de a nu folosi nici o modalitate de compensare a acestora în cadrul experimentelor conduse. Cap. 6.2 descrie interfața grafică elaborată în LabVIEW, folosită în diverse etape de elaborare a tezei curente. În Cap. 6.3 este descris experimentul de ascultare, cât și rezultatele obținute din [241]-[243] pentru validarea metodei de interpolare propus în cadrul Cap. 5.2.2, în timp ce în Cap. 6.4 este detaliat experimentul dedicat validării algoritmului de reducere a dimensionalității funcțiilor HRTF, introdus de către autor în [240] și explicat în Cap. 3.2.3. Cap. 6.5 este dedicat problemelor întâlnite și a posibilelor îmbunătățiri, ce ar putea fi aduse experimentelor prezentate.

6.1. Necesitatea egalizării căștilor utilizate

Funcția de transfer a căștilor utilizate a fost și este un subiect de discuție în literatură. Pot fi găsite atât experimente conduse care iau în considerare acest aspect [107]-[108], [345]-[347], cât și alte experimente care nu iau în considerare egalizarea căștilor [130]. Deoarece funcția de transfer a căștilor este independentă de direcția sursei, contribuția acestor funcții la semnalul recepționat la urechi este echivalentă cu filtrarea semnalului original cu o funcție fixă de transfer fără influențarea sugestiilor de localizare [210]. Diferite metode de egalizare a căștilor și efectele acestora asupra percepției surselor de sunet pot fi găsite în [348]-[352]. În timp ce sugestiile monaurale pot influența percepția locației sursei de sunet, rezultatele obținute în literatură sugerează faptul că influența funcției de transfer a căștilor folosite, asupra locației de sunet percepute, este neglijabilă [353]-[354]. În practică, luarea în considerare exactă a influențelor căștilor este dificilă. Funcția de transfer a căștilor depinde atât de căștile folosite, cât și de ascultător [355], de unde rezultă necesitatea măsurărilor individuale pentru fiecare ascultător cu o pereche particulare de căști [353], [356]. Mai mult, funcția de transfer a căștilor se modifică dacă ascultătorul modifică poziția căștilor chiar și foarte puțin [357].

În mod general, putem spune că egalizarea necesară, datorată influențelor introduse de căștile folosite în ascultare, este o procedură dificilă și pretențioasă. Mai mult, datorită unor rezultate din literatură [353]-[354] are efecte neglijabile asupra performanțelor de localizare. Datorită acestor cauze, autorul a ales să neglijeze compensarea căștilor în cadrul acestei teze.

6.2. Interfață grafică utilizator pentru generarea unei Realități Acustice Virtuale

Diferite versiuni ale interfeței grafice utilizator, ce urmează să fie descrisă în acest capitol, au fost folosite de-a lungul diferitelor etape de elaborare a acestei teze. Ea permite generarea unei RAV simple și poate fi folosită pentru reprezentarea grafică atât în domeniul timp, cât și în domeniul frecvență, a funcțiilor HRTF, respectiv a semnalelor acustice obținute. Interfața elaborată a fost folosită în cadrul experimentelor de psihoacustice din acest capitol pentru generarea semnalelor acustice folosite în aceste experimente. O versiune mai veche a acestei interfețe a fost prezentată de către autor în [342].

Aplicația propusă a fost dezvoltată folosind mediul LabVIEW de la firma National Instruments, care este un mediu de programare grafică foarte utilizat de ingineri și cercetători pentru dezvoltarea unor sisteme sofisticate de măsurare, test și control, folosind icoane grafice și fire, care alcătuiesc o diagramă, denumită instrument virtual (VI). LabVIEW oferă integrare nemaiîntâlnită cu mii de dispozitive hardware și furnizează sute de librării încorporate pentru analiză avansată și vizualizare de date.

În Fig. 6.1 se prezintă panoul frontal al aplicației dezvoltate. Ideea de bază a aplicației este implementarea conceptului de RAV, descris în Cap. 2.3. Prin implementare se înțelege obținerea a două semnale acustice, stânga și dreapta pentru ascultare prin intermediul unor căști, ale căror surse de sunet virtuale este un punct (sau direcție) din spațiul 3D, folosind un sunet monaural, generat local (adică în cadrul aceluiasi VI) și perechea de funcții HRIR corespunzătoare. Există mai multe modalități de obținere ale acestor perechi de funcții HRTF: din baze de date HRTF publice (vezi Cap. 3.2.2), ca rezultat al unor modele propuse în literatură (vezi Cap. 3.2.1.2), sau ca rezultat al unor metode de interpolare existente (vezi Cap. 5.1). Interfața grafică dezvoltată permite, în forma ei actuală, folosirea perechilor de funcții HRTF: din baza de date LISTEN HRTF (bază de date folosită în cadrul diferitelor experimente din cadrul tezei curente), sau ca rezultat al unor metodelor de interpolare propuse de către autor în cadrul Cap. 5.2.

Aplicația propusă, implementată sub forma unui instrument virtual (VI) complex, este prezentată în detaliu. Pentru ca aplicația să fie rulată, persoana, care dorește să o folosească, trebuie să aibă un calculator personal (PC) echipat cu o placă de sunet și mediul LabVIEW, respectiv modulul Mathscript RT de la National Instruments, instalat. Pe lângă cele enumerate, pentru performanțe optime, este obligatorie folosirea unor căști de calitate conectate la PC. O discuție legată de necesitatea compensării căștilor folosite este purtată în Cap. 6.1. În acest mod, fiecare semnal acustic ajunge la destinația lui (la urechea stângă, respectiv dreaptă a subiectului) fără să fie nevoie de diferite modalități de anulare a diafoniei (cazul folosirii difuzoarelor).

Panoul frontal al instrumentului virtual poate fi împărțit în patru părți: generarea sunete monaurale, informații legate de funcțiile HRTF, selecția parametrilor antropometrici (unde se aplică) și vizualizarea diferitelor date de interes.

1. **Partea de generare de sunete** (Sound generation) permite utilizatorului generarea diferitelor tipuri de sunete, ca zgomot alb uniform, zgomot alb gaussian, zgomot gamma, zgomot poisson, zgomot bernoulli, etc. Alți parametri ai semnalului pot fi, de asemenea, modificați, ca durata, frecvența de eșantionare, amplitudinea. Tot aici, mai plasate două butoane: primul, SOUND, care dezactivează generarea

sunetelor la ieșirile audio ale PCului, partea de vizualizare a diferitelor mărimi rămânând funcțională, respectiv al doilea, STOP, care oprește aplicația din rulare.

2. **Partea de informații legate de funcții HRTF** (HRTF related information) este responsabilă pentru selecția subiectului, a azimutului și a înălțării. De asemenea, este posibilă specificarea utilizării funcțiilor HRTF compensate (adică după eliminarea influențelor dispozitivelor de înregistrare) sau a funcțiilor HRTF necompensate (adică fără eliminarea influențelor respective). Totodată este posibilă cu ajutorul unui buton de selecție (MATLAB vs. LabVIEW) alegerea unei perechi de funcții HRIR existente din baza de date LISTEN HRTF sau a unei perechi de funcții HRIR externe, obținute de utilizator printr-o metodă de interpolare implementată în MATLAB, folosind RNA descrise în [340], [241], [240], [341]. Controlul denumit MATLAB permite alegerea folosirii stării RNA dorite (deja gata antrenate). Indicatorul ANN Structure & Mapping afișează informații asupra structurii de RNA folosite, respectiv dacă este utilizată pre- sau post-procesarea datelor de antrenament. După folosirea îndelungată a acestei interfețe, autorul a considerat că este necesară crearea unui combo box, care permite alegerea pasului cu care azimutul se poate modifica, pentru a permite generarea surselor de sunet virtuale cu un pas selectabil (5, 10, sau 15 grade). Această opțiune a fost folosită în experimentele psihoacustice conduse, descrise în Cap. 6. Alte informații, ca numărul total de eșantioane disponibile pe canal, rata de eșantionare, numărul canalelor, numărul de biți pe eșantion, respectiv atenționări în cazul unor erori sunt, de asemenea, disponibile.

3. **Partea dedicată selecției parametrilor antropometrici** (poziționate pe partea stângă a graficelor) permite testarea alegerea parametrilor corespunzători, ce au fost folosiți, în prealabil, pentru antrenarea RNA dedicate interpolării funcțiilor HRTF, varianta pentru orice persoană. Modificarea lor, în procesul de interpolare pentru o singură persoană, nu are nici un efect. Această opțiune de selecție permite folosirea și testarea corespunzătoare a oricăror variante de RNA multi-persoană prin a permite alegerea oricăror parametrii antropometrici specificați în [148]-[150], [158], [343].

4. **Ultima parte este dedicată vizualizării diferitelor date de interes**, cu ajutorul a numeroaselor grafice disponibile. După ce utilizatorul selectează parametrii sunetului monaural, semnalul acustic dorit este generat și reprezentat atât în domeniul de frecvență, cât și în domeniul timp, pe prima linie al panoului frontal. Selecția perechii de funcții HRIR dorite (prin alegerea unui subiect, azimut, înălțare, status) este urmată de convoluția cu sunetul generat. Semnalele acustice obținute, pentru cele două urechi, reprezintă cel mai simplu caz al conceptului RAV. Ultimele două rânduri, din partea inferioară a panoului frontal, conțin grafice, care reprezintă semnalele RAV obținute, atât în domeniul de frecvență, cât și în domeniul timp, pentru cele două urechi.

Fig. 6.2 ilustrează diagrama bloc a instrumentului virtual implementat. VIul conține multe subVIuri, care îndeplinesc diferite funcții, ca: formarea căii utilizată în selecția funcțiilor HRIR, generarea sunetelor, conversia între funcții HRIR și funcții HRTF, convoluția sunetului monaural cu funcțiile HRIR pentru obținerea semnalelor acustice pentru cele două urechi prin intermediul căștilor, realizarea legăturii dintre MATLAB și LabVIEW prin intermediul modulului Mathscript RT în vederea simulării RNA gata antrenate, alegerea și citirea parametrilor antropometrici ale persoanelor (doar pentru varianta RNA multi-persoană), și altele.

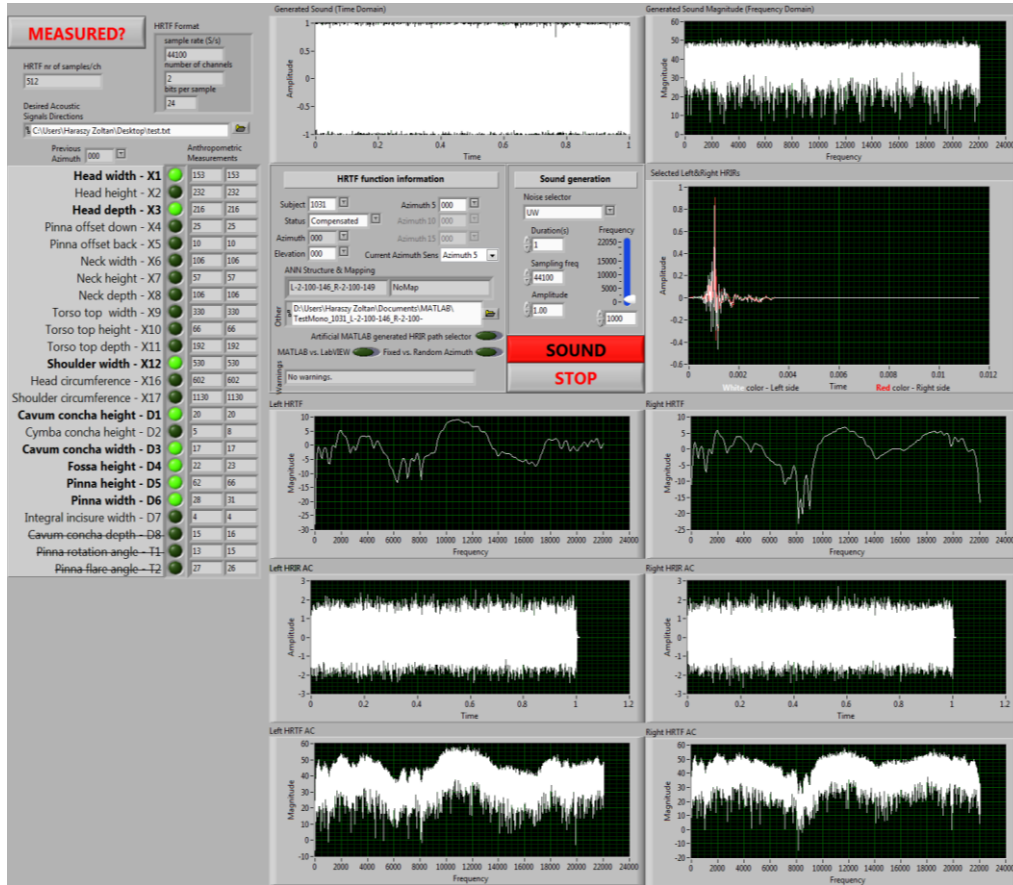


Fig. 6.1. Panoul frontal al aplicației LabVIEW.

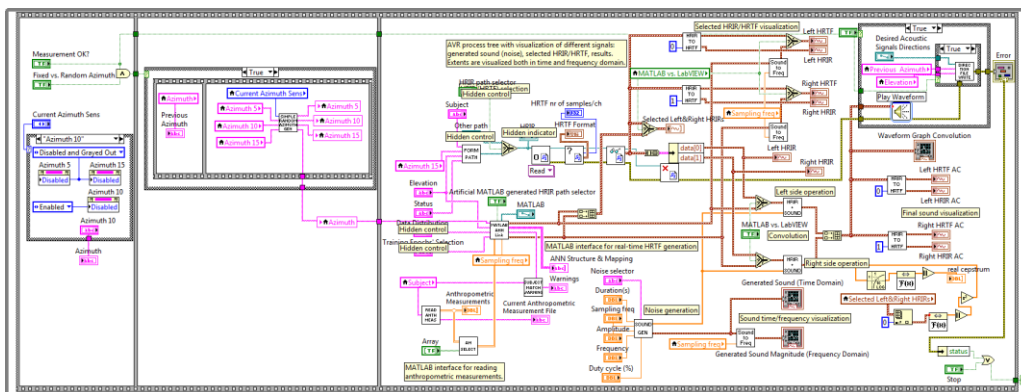


Fig. 6.2. Diagrama bloc a aplicației propuse.

Fig. 6.3 arată schema logică, care descrie funcționarea acestui VI. După ce se lansează în execuție, aplicația verifică starea momentană a unor butoane, de unde sunt citite diferite date de intrare, ca: parametrii sunetului monaural, modalitatea de selecție a funcțiilor HRTF (MATLAB sau LabVIEW), parametrii de

intrare pentru obținerea funcțiilor HRTF. Ca mărimi de intrare ale procesului de sinteză binaurală ajung: sunetul monaural generat local conform parametrilor specificați, respectiv perechea de funcții HRTF conform modalității de selecție și a parametrilor dați. Modalitățile de generare sunt: obținerea funcțiilor prin selecție dintr-o bază de date (LabVIEW) sau obținerea lor prin metode bazate pe rețele neuronale din mediul MATLAB. Rezultatul procesului de sinteză binaurală reprezintă o pereche de semnale acustice, care sunt ascultate prin intermediul unor căști (dacă acest lucru este dorit prin apăsarea unui buton).

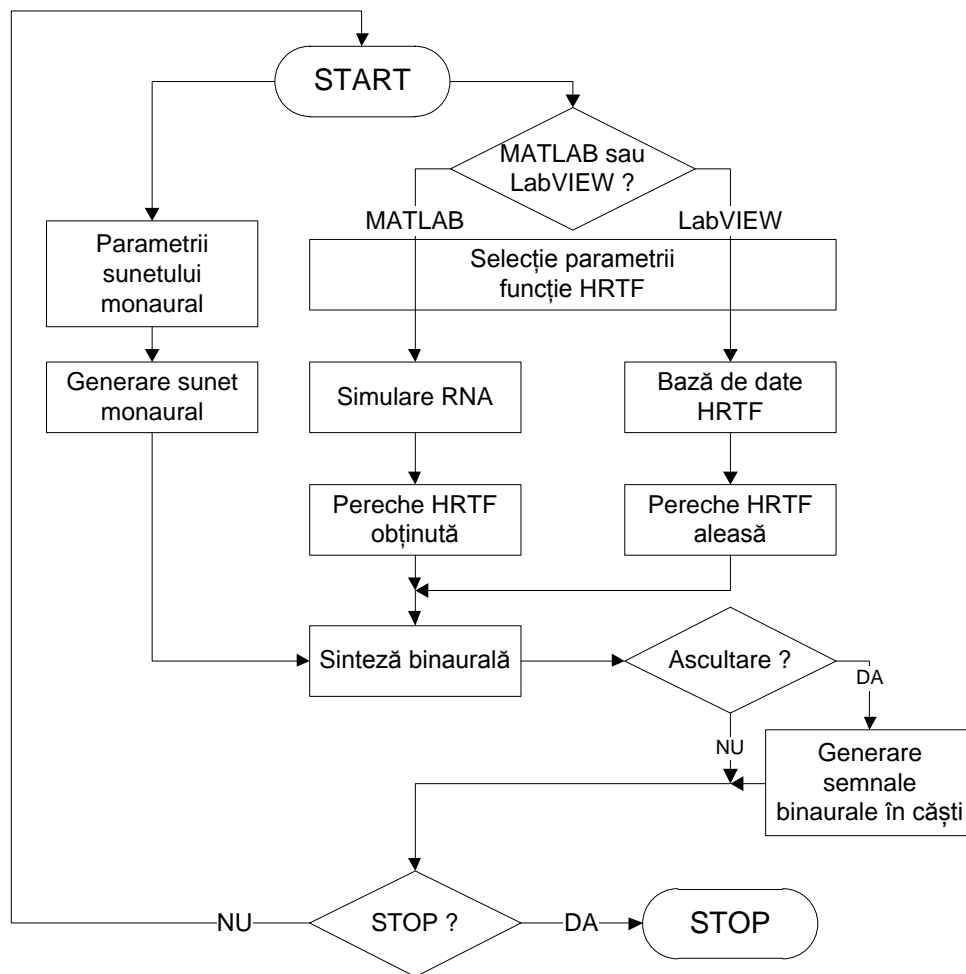


Fig. 6.3. Schema logică a aplicației propuse.

Instrumentul virtual rulează în continuu până când utilizatorul apasă butonul STOP. O facilitate utilă a aplicației dezvoltate este faptul, că ea se oprește dacă apare o eroare în funcționarea ei, pe ecranul calculatorului apărând informații despre natura erorii.

6.3. Experimente de localizare pentru validarea structurii de RNA pentru o singură persoană

Structura RNA propuse este prezentată în Fig. 6.4. Rețeaua constă din trei părți: un strat de intrare pentru nodurile sursă (2 intrări), un strat ascuns (n neuroni) și un strat de ieșire (512 neuroni). Rețeaua folosită în cadrul acestor experimente are aceeași structură cu cea propusă în [340], singură diferență este numărul de neuroni diferiți din stratul ascuns.

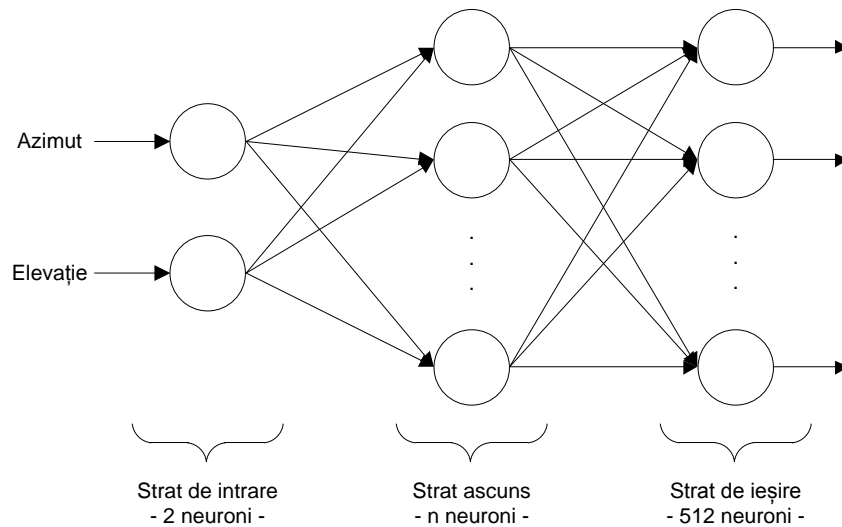


Fig. 6.4. Arhitectura rețelei neuronale este aceeași cu cea propusă în [340], singura diferență apare la numărul neuronilor din stratul ascuns.

Considerând faptul că o pereche de astfel de rețele sunt necesare pentru un singur subiect de test pentru că ființele umane au două urechi, cele două intrări necesare, fiecărei RNA dintre acestea două, sunt azimutul și înălțarea sursei de sunet virtuale dorite. Fiecare dintre aceste două RNA furnizează, la ieșirile lor, un set de 512 valori temporale (eșantioane în domeniul timp), care reprezintă funcțiile HRIR (vezi Cap. 3.2.1) pentru sursa de sunet virtuală dorită. Numărul optim al neuronilor din stratul ascuns este dificil de determinat. Acest număr n de neuroni afectează performanțele RNA și este ales după experimentări atente. Rezultatele obținute sugerează că n trebuie să fie undeva între 40 și 60. În experimentele de localizare conduse, prezentate mai în detaliu în cele ce urmează, stratul ascuns a fost alcătuit din 50 de neuroni.

Rețeaua neuronală propusă este o rețea perceptron multistrat cu propagarea înainte a semnalului și cu propagarea înapoi a erorii (multilayer perceptron feedforward backpropagation network). Baza de date folosită pentru antrenarea acestei rețele neuronale este aceeași, și anume baza de date LISTEN HRTF, descrisă în Cap. 3.2.2.1.

Baza de date LISTEN HRTF conține date măsurate despre un total de 49 de persoane. Dintre toate aceste persoane disponibile, o singură persoană a fost aleasă în funcție de asemănările antropometrice cu subiectul nostru de test (ZH). Persoana aleasă din baza de date LISTEN HRTF este subiectul cunoscut ca IRC_1031.

Parametrii antropometrici comparați au fost prezentați în Tab. 6.1. Ele au fost preluate din [158]. Alte metode de selecție a parametrilor antropometrici sunt raportate în [148]-[150], [343].

Antrenarea RNA a fost realizată folosind setul de date (funcții HRIR) disponibil pentru subiectul de test ales.

<i>Denumire</i>	<i>Parametru</i>	<i>IRC_1031</i>		<i>Subiect</i>	
		<i>L</i>	<i>R</i>	<i>L</i>	<i>R</i>
<i>x1</i>	<i>Head width</i>	153		162	
<i>x3</i>	<i>Head depth</i>	216		225	
<i>x12</i>	<i>Shoulder width</i>	530		550	
<i>d1</i>	<i>Cavum concha height</i>	20	20	18	19.3
<i>d3</i>	<i>Cavum concha width</i>	17	17	15.4	15.4
<i>d4</i>	<i>Fossa height</i>	22	23	21.2	16.7
<i>d5</i>	<i>Pinna height</i>	62	66	66.5	64.7
<i>d6</i>	<i>Pinna width</i>	28	31	26.9	28.8

Tab. 6.1. Parametrii antropometrici selectați în [158]. Valorile specificate sunt exprimate în milimetri.

Scopul propus în cadrul acestor experimente de localizare este evaluarea practică, folosind experimente de localizare, a performanțelor structurii rețelelor propuse. Mediul în care au fost efectuate experimentele de localizare este descrisă în Cap. 7.2.2. Experimentele își propun determinarea faptului dacă sunetul obținut oferă percepția dorită, asigurată prin plasarea corespunzătoare celor două valori (sau cât mai aproape de ele), a sursei de sunet virtuale în spațiul auditiv al subiectului ales. Cu alte cuvinte, se dorește stabilirea prin aceste experimente a faptului dacă sunetul obținut, prin filtrarea sunetului monaural cu funcțiile HRTF(HRIR) obținute, chiar pare să vină din, sau cât mai aproape de, direcția perechii azimut-elevație aduse la intrările RNA.

Considerând setul de date selectat, există două cazuri posibile (și disponibile din interfața grafică implementată în LabVIEW) din punctul de vedere al distribuției datelor de antrenament și ale celor de test:

- În primul caz, setul de date existent, pentru subiectul de test ales, a fost împărțit în două seturi mai mici, așa cum urmează. Fiecare a patra pereche de funcții HRIR a fost selectată și setul de date, astfel obținut, a fost folosit în faza de test pentru cele două rețele neuronale (date de test însemnând faptul că pe lângă experimentele de localizare conduse, performanțele rețelelor sunt evaluate prin criteriul erorii medii pătratice (MSE), la fel ca în [340]). Restul datelor a fost utilizat în procesul de antrenament pentru cele două RNA. Pe scurt, aproximativ 75% din toate datele disponibile sunt folosite pentru antrenarea RNA și restul de 25% pentru testarea RNA.

- În al doilea caz, toate datele existente pentru subiectul selectat (187 de perechi de funcții HRIR) au fost folosite în procesul de antrenament ale celor două RNA (100% antrenament – 0 % test). Autorul se așteaptă la performanțe mai bune de localizare față de primul caz, din cauza numărului mai mare de perechi intrare-ieșire disponibile la antrenamentul rețelelor.

6.3.1. Scenariul de măsurare

Scenariul experimental, folosit în experimentele de localizare conduse, este prezentat în Fig. 6.5.

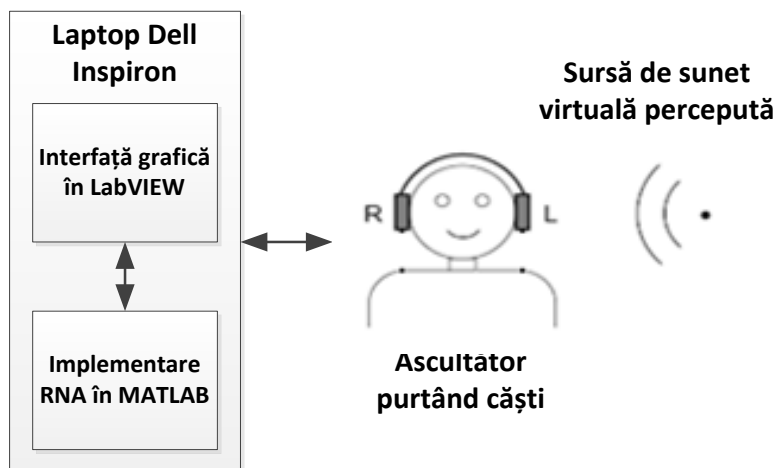


Fig. 6.5. Scenariul experimental folosit în experimentele de localizare.

O scurtă descriere a configurației experimentale este dată în continuare. Experimentele au fost efectuate pe un notebook Dell Inspiron 1520 rulând sistemul de operare Windows 7 Ultimate, versiunea pe 32 de biți, cu următoarea configurație hardware: Intel Core 2 Duo 2.00 GHz CPU, 2GB DDR2 RAM, 160 GB HDD. Căști Sennheiser HD 435 au fost conectate la ieșirea audio a calculatorului și folosite pentru ascultarea semnalelor acustice generate. Ca mediu software, autorul a folosit o versiune mai nouă și îmbunătățită a interfeței grafice de utilizator (GUI), prezentate în [342]. Interfața a fost dezvoltată în mediul de programare grafică LabVIEW al firmei National Instruments.

În cele ce urmează este prezentată, pe scurt, interfața grafică de utilizator, ce a fost dezvoltată. Fig. 6.6 ilustrează interfața elaborată în timpul rulării acestuia. Ideea de bază, care stă la baza dezvoltării ei, este implementarea unei versiuni de bază al conceptului RAV, prezentat în Cap. 2.3. Utilizatorul poate obține o simplă RAV, care înseamnă generarea semnalelor acustice, pentru cele două urechi menite să fie ascultate în căști, ale căror sursă virtuală de sunet este un anumit punct (definit prin perechea azimut-elevație) din spațiul 3D. Semnalul acustic, pentru una dintre urechi, este obținut prin convoluția sunetului generat local (de exemplu: zgomot alb sau sinusoidă) cu funcțiile HRIR corespunzătoare punctului dorit din spațiu. Aceste funcții HRIR pot fi obținute dintr-o bază de date publică sau folosind o metoda bazată pe RNA prezentată (vezi Cap. 5.2).

Interfața oferă posibilitatea de a alege între diferitele stări ale rețelelor neuronale după antrenarea lor pentru un anumit număr de epoci (50000, 100000, 150000, 200000, 250000 sau 300000) și diferitele distribuții ale datelor de antrenament/test (100 % antrenament – 0 % test, respectiv 75 % antrenament – 25 % test). Posibilitatea de alegere dintre stări este posibilă din cauza faptului că rețelele neuronale propuse au fost antrenate înainte de conducerea experimentelor de localizare și starea lor a fost salvată în fișiere .mat pe HDD.

În cele ce urmează atenția se mută pe metoda propusă, bazată pe RNA. Metoda a fost implementată folosind uneltele din Neural Network Toolbox, inclus în mediul MATLAB al firmei MathWorks.

Mai multe detalii despre rețeaua neuronală sunt prezentate în continuare. Funcțiile de transfer alese, pentru experimentele conduse, sunt cele din [340], adică funcția de transfer 'hiperbolic tangent sigmoid' pentru stratul ascuns și funcția de transfer liniară pentru stratul de ieșire. Întreaga rețea neuronală a fost antrenată utilizând funcția de antrenament de rețea *trainrp*, față de *traingdx* folosită în [340]. Această funcție actualizează valorile pragurilor și ale ponderilor bazându-se pe algoritmul resilient backpropagation (Rprop), implementat în Neural Network Toolbox. Funcția *traingdx* a fost înlocuită cu funcția *trainrp* din cauza performanțelor sale mai ridicate obținute în faza de experimentare a cercetărilor conduse.

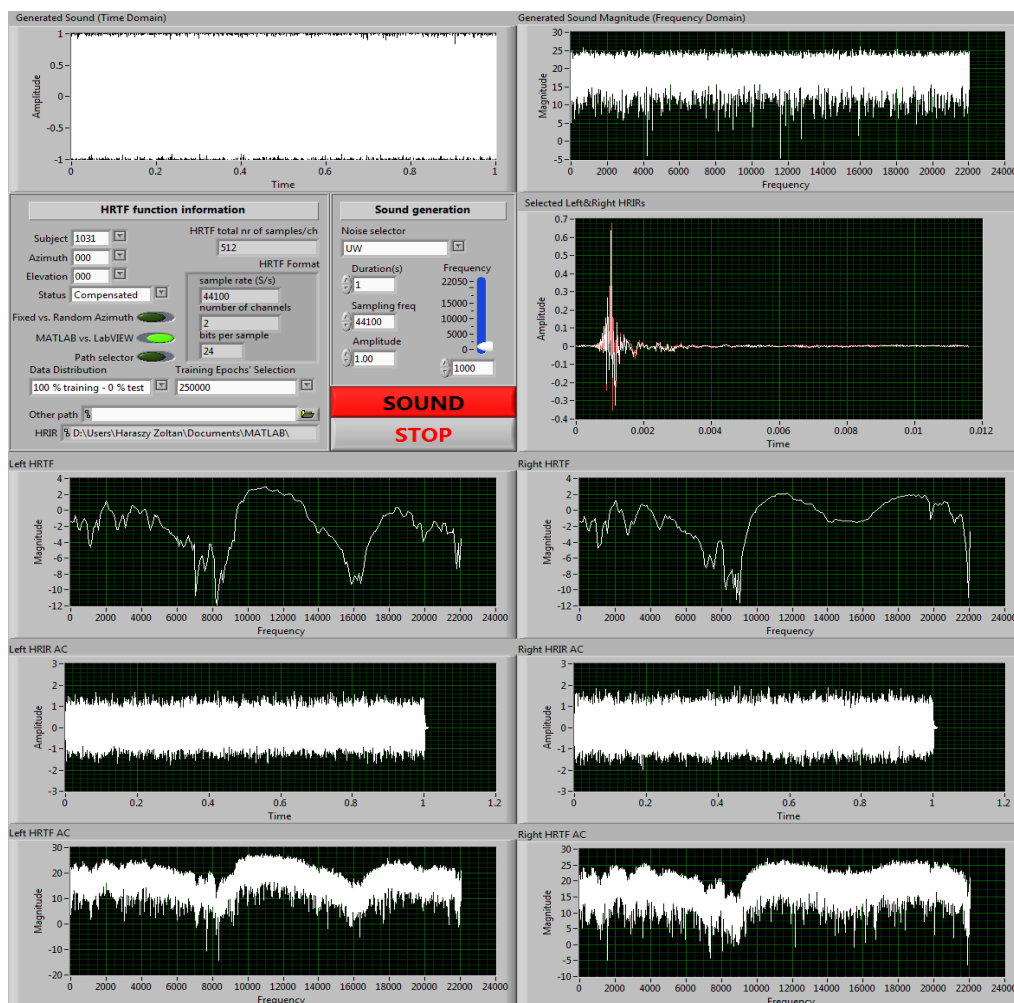


Fig. 6.6. Interfața implementată în timpul rulării.

Un exemplu al erorii de antrenament obținute este prezentat în Fig. 6.7. Se poate observa că eroarea scade odată cu creșterea numărului de epoci.

Experimentele de localizare au fost realizate în Laboratorul de Sisteme Bioinspirate al Departamentului de Electronică Aplicată din cadrul Universității Politehnica din Timișoara. Subiectul de test (ZH) a fost așezat pe un scaun rotativ, cu un compas plasat orizontal între picioare pentru măsurarea simplă a deplasamentului unghiular. Un difuzor a fost plasat drept înaintea subiectului de test la aceeași înălțime cu capul subiectului. Poziția difuzorului, exprimat în coordonate azimut-elevație, este la $(0^\circ, 0^\circ)$, adică drept înaintea persoanei de test orientat cu corpul și cu fața spre el. Distanța dintre difuzor și capul subiectului a fost fixată la 1.5 metri. În aceste experimente, difuzorul a servit ca o sursă de referință de sunet. Semnalul acustic de referință emis de difuzor a fost compus din următoarea secvență (repetată pe durata experimentelor de localizare): 1 s de zgomot alb, urmată de 0,2 s de liniște.

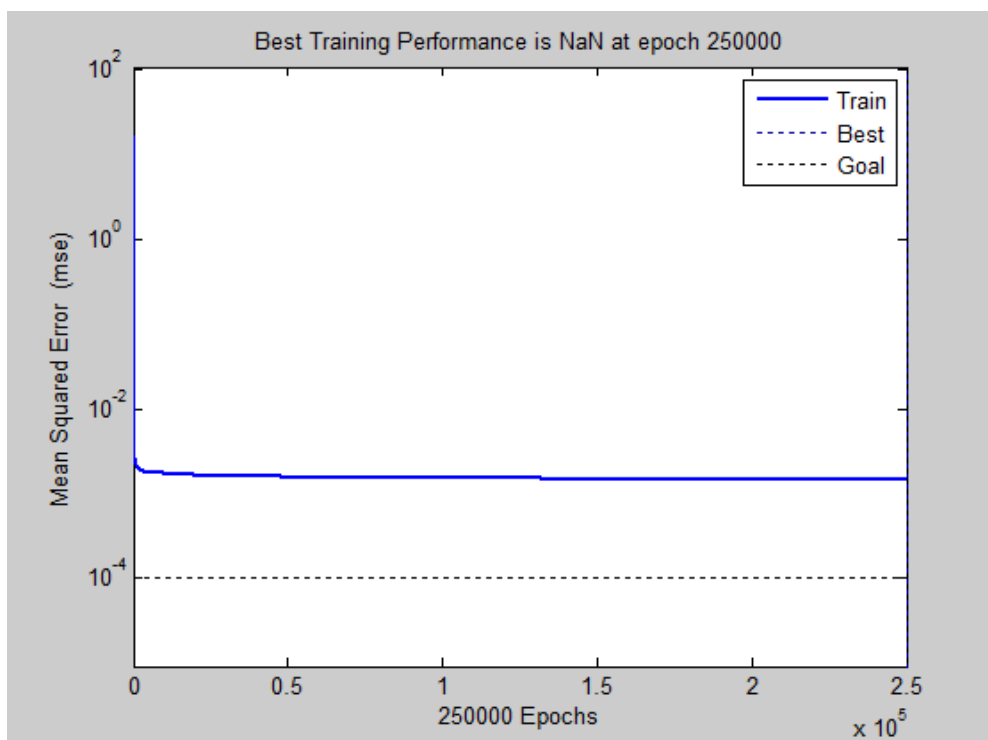


Fig. 6.7. Un exemplu al erorii de antrenament obținute pentru 250000 de epoci.

Semnalele acustice (stimulul) a fost ascultate de subiect prin căști, din interfața dezvoltată în mediul LabVIEW. Subiectul a fost așezat pe scaun, trebuia să se rotească cu scaunul fără să-și rotească capul și fără să modifice înălțimea relativă a capului față de podea, până când poziția sursei de sunet virtuale a semnalelor acustice ascultate în căști coincide cu poziția sursei de sunet existente fizic (cea a semnalului de referință, care se aude din difuzor). În momentul în care cele două surse coincid, deplasamentul unghiular (azimutul), indicat de compas, este citit. Stimulul ales pentru căști a fost compus din următoarea secvență (repetată pe durata experimentelor de localizare): 1 s de zgomot alb uniform, urmată de 0.3 s de liniște. Secvența de stimul a fost repetată până când subiectul de test a dat o estimare a azimutului sursei de sunet virtuale în spațiul său auditiv prin fixarea

poziției scaunului rotativ. În faza următoare, poziția sursei de sunet virtuale este comparată cu perechea azimut-elevație folosită la intrările RNA (prin calculul diferenței absolute dintre valorile de azimut folosite în generarea semnalelor acustice, respectiv cele percepute) pentru verificarea acurateții localizării, și astfel a performanțelor întregului scenariu experimental.

6.3.2. Rezultate obținute

Rezultatele obținute după realizarea experimentelor de localizare sunt prezentate în Fig. 6.8.

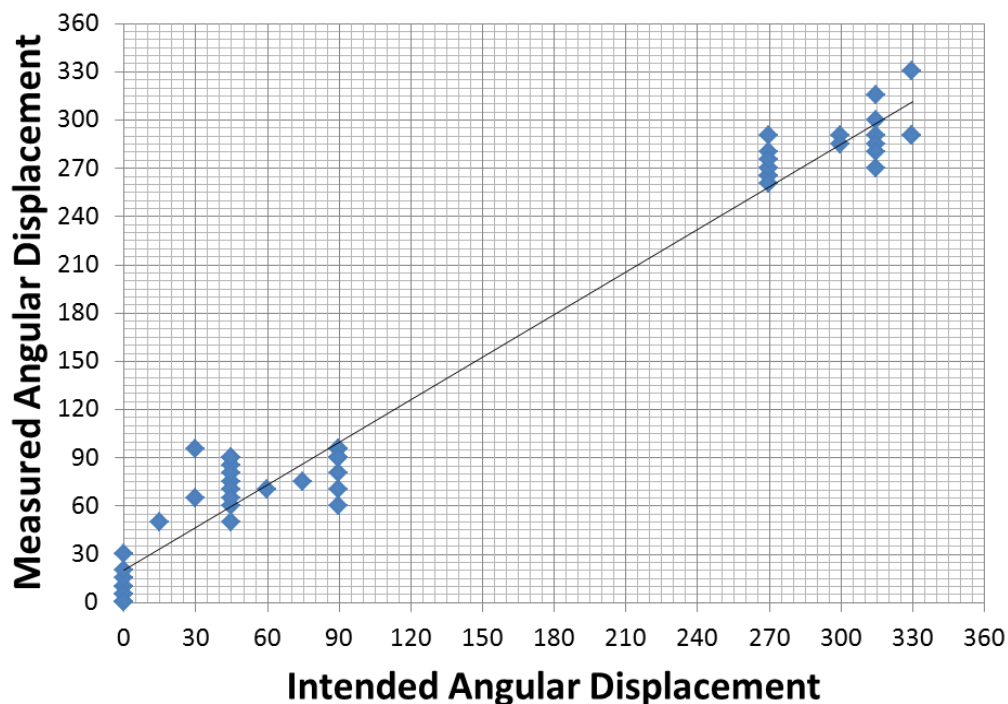


Fig. 6.8. Rezultatele de localizare obținute.

În cadrul acestor experimente, stimulii au fost generați în planul frontal la elevația de 0° . Pentru azimut, acest lucru înseamnă valori între -90° și 90° . Autorul dorește să sublinieze faptul că în experimentele de localizare efectuate a fost folosită una dintre cele 12 stări salvate în prealabil în fișiere, și anume:

- Distribuția datelor – 75 % training and 25 % test;
- Numărul de epoci de antrenament folosit – 300000.

Tab. 6.2 ne dă o privire de ansamblu al erorii absolute obținute după experimente. Prima coloană reprezintă intervalele de eroare considerate pentru azimut. A doua coloană arată numărul de valori de azimut percepute în intervalul de eroare din prima coloană. În sfârșit, a treia coloană furnizează procentajul numărului de valori de azimut percepute, care intră în intervalul de eroare considerat, din numărul total de măsurători efectuate.

După compararea rezultatelor din Tab. 6.2 cu cele obținute de H. Hu și alții în [158], se poate spune că performanțele de localizare prezentate în [241]-[243] sunt mai bune decât cele din [158], dacă considerăm cazul funcțiilor HRTF neindividuale. În schimb, dacă se iau în considerare rezultatele obținute folosind funcții HRTF personalizate, atunci localizarea folosind metoda prezentată de autor este de o acuratețe mai scăzută. De asemenea, este de menționat că experimentele din [158] au avut loc în tot planul orizontal (azimut între 0° - 360°), adică la elevație 0° , în timp ce experimentele conduse de autor în [241]-[243] au avut loc pentru stimuli din planul frontal (azimut între -90° și 90°) și elevație 0° .

Intervale de eroare absolută considerate	Numărul de valori de azimut care cad în intervalul de eroare dat	Procentajul obținut de fiecare interval
[0,15)	32	59,26 %
[15,30)	10	18,52 %
[30,45)	11	20,37 %
[45,60)	0	0 %
[60,∞)	1	1,85 %

Tab. 6.2. O privire de ansamblu a performanțelor sistemului.

Bazându-se pe rezultatele din Fig. 6.8, respectiv Tab. 6.2, autorul consideră că generarea funcțiilor HRTF este posibilă pentru orice persoană cu deficiențe de vedere. Această metodă prescurtează, durata procesului complex de măsurare, necesar în momentul de față pentru obținerea funcțiilor HRTF pentru o anumită persoană, dacă există cel puțin un subiect inclus în baza de date LISTEN HRTF sau altele, al cărui parametrii antropometrici sunt apropiați de cei mai importanți parametrii antropometrici ai persoanei de interes. Dacă numărul de epoci folosit pentru antrenarea celor două RNA este adecvat, erorile de localizare obținute sunt mici (cel puțin pentru pozițiile în care au fost evaluate performanțele de localizare). În concluzie, utilizarea rețelelor neuronale în obținerea funcțiilor HRTF oferă o alternativă bună pentru procesul complex de măsurare necesar.

6.4. Experimente de localizare pentru validarea algoritmului de reducere a dimensionalității funcțiilor HRTF

Implementarea cu succes a unei RAV simple (vezi Cap. 7) pe un sistem de dezvoltare a condus la necesitatea unui astfel de algoritm. În vederea proiectării și implementării unei astfel de RAV, este necesar să luăm toată structura (valorile ponderilor și ale pragurilor) unei rețele neuronale antrenate în prealabil (o RNA pentru fiecare ureche, vezi Cap. 5) și să o salvăm într-un fel pe sistemul de dezvoltare. Pe placa folosită am avut de ales între două opțiuni:

- Prima opțiune este să stocăm cele două structuri în memoria internă a plăcii de dezvoltare. Această opțiune reduce timpul de execuție pentru obținerea semnalelor acustice folosind funcțiile HRTF calculate local pe placă;
- A doua opțiune este să salvăm cele două structuri pe o cartelă de memorie externă. În mod evident, această soluție este mult mai lentă deoarece generarea funcțiilor HRTF locale pe placă este precedată de citirea fiecărei valori de pondere, respectiv prag trebuie executată de pe cartela de memorie (în loc de memoria

internă), care introduce o întârziere semnificativă în execuție, prelungind astfel ciclul de generare a semnalelor acustice.

Dacă dimensiunile celor două RNA, una pentru obținerea funcțiilor HRTF pentru urechea stângă și una pentru urechea dreapta, sunt prea mari, este posibil ca implementarea nu este posibilă folosind prima opțiune (nu există suficientă memoria internă disponibilă pentru salvarea tuturor valorilor de prag și pondere pentru ambele RNA). Dacă acest lucru este adevărat, implementarea este posibilă doar prin a doua opțiune. Dacă funcțiile HRTF sunt prescurtate sau reduse suficient de mult folosind algoritmul propus [240], folosirea primei opțiuni ar putea să fie posibilă cu toate avantajele ei.

Baza de date aleasă este baza de date LISTEN HRTF (vezi Cap. 3.2.2.1), cea folosită de către autor în toate experimentele conduse. Pe baza metodologiei de alegere a unui subiect prezentată în Cap. 6.2, a fost ales același subiect denumit IRC_1031, deoarece persoana care a fost supusă experimentelor de localizare a rămas tot aceeași. Setul de date corespunzător subiectului IRC_1031 a fost filtrat cu algoritmul prezentat în Cap. 3.2.3. Acest set filtrat, corespunzător subiectului IRC_1031, a fost utilizat în vederea antrenării celor două RNA destinate obținerii funcțiilor HRTF. Structura acestor două RNA este prezentată în cele ce urmează.

Este evident că sunt necesare două RNA corespunzătoare celor două urechi pentru obținerea unei perechi de funcții HRIR. Fiecare dintre aceste RNA este o rețea de tip perceptron cu mai multe straturi (vezi Cap. 5.2.1.2). Structura celor două RNA, arătată în Fig. 6.9, este similară cu cea prezentată în [340] și [241], dar în cazul actual este posibil ca numărul neuronilor (L, respectiv R) din stratul de ieșire să fie diferită pentru cele două urechi, datorită folosirii funcțiilor HRTF filtrate în procesul de antrenament.

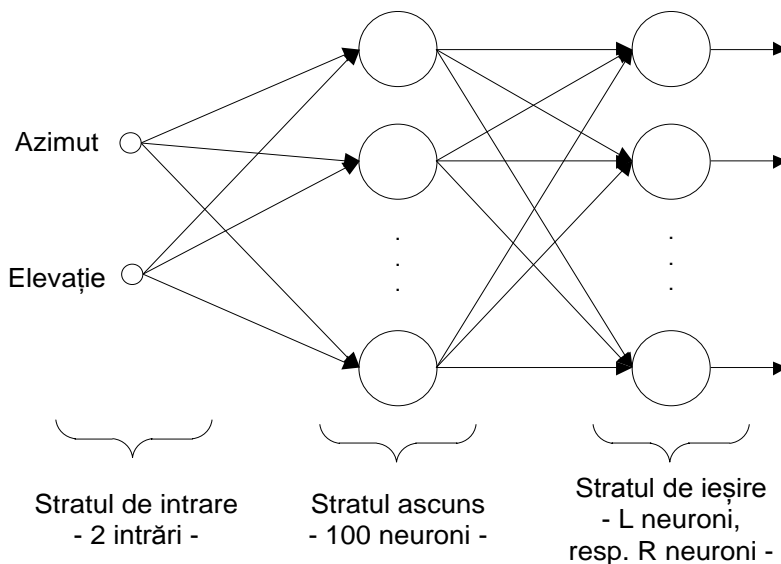


Fig. 6.9. Structura RNA utilizată pentru ambele urechi.

Rețeaua este formată din trei părți: stratul de intrare (2 intrări), un singur strat ascuns (format aici din 100 de neuroni), și stratul de ieșire (format dintr-un număr de L, respectiv R, neuroni pentru RNA stângă, respectiv RNA dreapta). Numărul neuronilor din stratul ascuns a fost determinat experimental și este

considerat optim pentru aplicația noastră. Cele două intrări sunt formate din perechea de azimut-elevație pentru sursa de sunet dorită. Ca rezultat așteptat, fiecare dintre aceste două RNA furnizează un set de L (partea stânga), respectiv R (partea dreapta) valori corespunzătoare funcțiilor HRIR pentru sursa de sunet specificată prin direcția dată la intrarea RNA.

Aplicarea algoritmului de reducere a dimensionalității, celor 187 de perechi de funcții HRIR, are ca rezultat obținerea numărului neuronilor de ieșire pentru cele două RNA, care sunt antrenate cu acest set obținut după filtrare. Numărul de neuroni de ieșire pentru partea stângă este de 146, rezultând structura de RNA de 2-100-146, respectiv numărul de neuroni de ieșire pentru partea dreaptă este de 149, rezultând structura de RNA de 2-100-149.

Ca rezultat al aplicării algoritmului propus lungimea funcției HRIR a fost redusă la 146 de eșantioane temporale pentru partea stângă, și la 149 de eșantioane pentru partea dreaptă, de la un total de 512 de eșantioane. Această reducere a numărului de eșantioane ale funcțiilor HRIR poate fi foarte utilă la implementarea unei RAV pe o placă de dezvoltare, așa cum a fost explicat în detaliu în Cap. 3.2.3.

6.4.1. Scenariul de măsurare și rezultatele obținute

Scopul acestui capitol este evaluarea performanțelor întregului lanț compus din RNA și algoritmul de reducere propus prin experimente de localizare (psihoacustice). Prin folosirea acestui scenariu se dorește să se determine dacă semnalele acustice, obținute prin folosirea sintezei binaurale (vezi Cap. 3.2.4) bazate pe funcțiile HRIR obținute după antrenarea rețelelor cu structura din Fig. 6.9, oferă percepția dorită a surselor virtuale de sunet în spațiul auditiv al persoanei de test supuse experimentelor.

În cazul de față, perechea de funcții HRIR pentru direcția dorită este obținută prin simularea celor două RNA antrenate în prealabil. Pentru ca acest lucru să fie posibil interfața grafică din LabVIEW trebuie să fie interfațată cu mediul de programare MATLAB folosind un modul denumit Mathscript RT al LabVIEW. În esență de fiecare dată perechea azimut-elevație specificată din LabVIEW este transmisă în MATLAB, unde cele două RNA, deja gata antrenate, sunt simulate obținându-se astfel eșantioanele temporale ale funcțiilor HRIR (146 pentru partea stângă, respectiv 149 pentru partea dreaptă, conform aplicării algoritmului propus), care sunt transmise înapoi în LabVIEW în vederea obținerii semnalelor acustice, ce urmează să fie ascultate în căști. Este de menționat faptul că simularea celor două RNA este posibilă, deoarece ele sunt salvate după antrenarea lor pe hard disc în format .mat, de unde pot fi foarte simplu încărcate în mediul de lucru din MATLAB, care este accesat din LabVIEW, folosind modulul Mathscript RT.

Fig. 6.10 prezintă interfața grafică folosită în cadrul experimentelor, care a fost implementată în LabVIEW.

Antrenarea celor două RNA a fost executată în MATLAB cu ajutorul Neural Network Toolbox, așa cum rezultă din Fig. 6.5. Funcțiile de transfer folosite, respectiv funcția de antrenament, sunt cele date în Cap. 6.3.1. Numărul de epoci de antrenament folosit a fost determinat experimental (100000 pentru ambele RNA).

Scenariul experimental folosit este aceeași cu cea prezentată în Cap. 6.3.1, respectiv Fig. 6.5, deci nu este descrisă din nou.

Numărul total de semnale acustice generate (ascultate în căști) a fost de 140, față de 54, din Cap. 6.3.1. Rezultatele obținute în experimentele psihoacustice

sunt arătate în Fig. 6.11. În cadrul acestor experimente au fost generate semnale acustice cu surse virtuale doar în planul orizontal în mod aleator. Valorile de azimut posibile au fost variate între -90° și 90° în fața ascultătorului.

Fig. 6.12 furnizează o privire de ansamblu a erorilor de localizare obținute în cadrul acestor experimente. Rezultă faptul că majoritatea erorilor aparute de-a lungul măsurărilor sunt mai mici decât 15 grade, reprezentând 64.29% din totalul experimentelor. Puțin mai mult de 90% din testele de ascultare au rezultat în erori mai mici de 30 de grade.

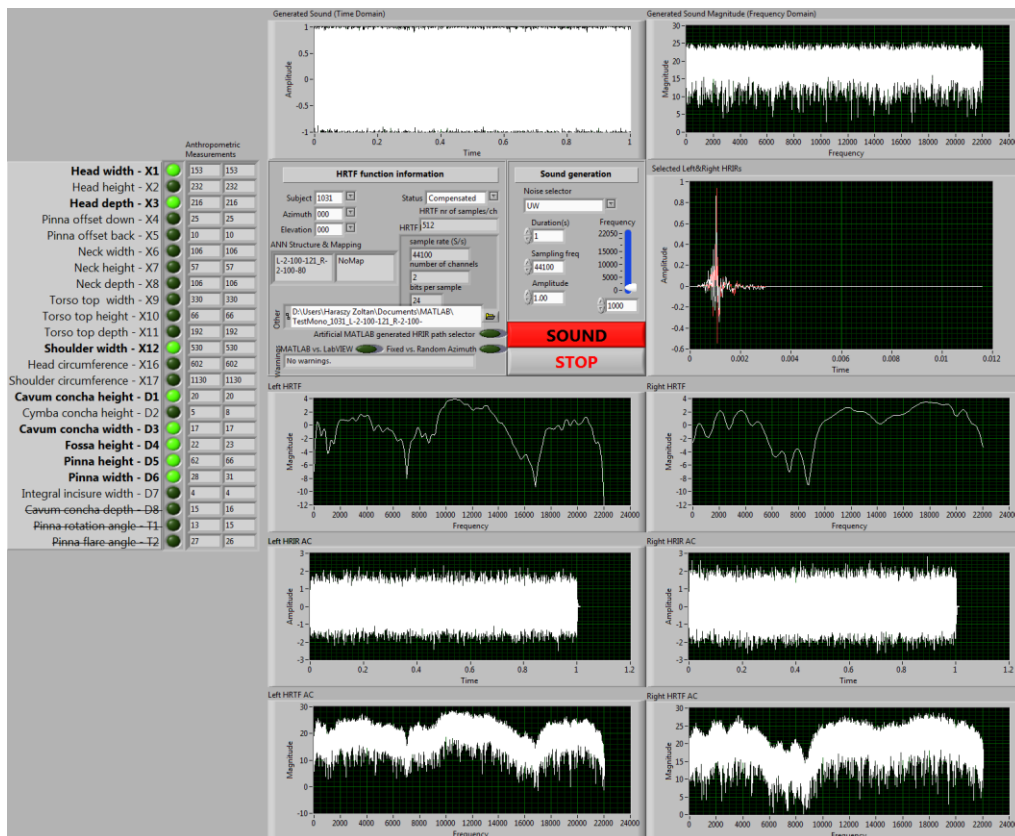


Fig. 6.10. Interfața grafică folosită în cadrul experimentelor.

După compararea rezultatelor din Fig. 6.12 cu cele obținute de Hu în [158], putem spune că performanțele de localizare sunt mai bune, dacă le comparăm cu cele obținute de Hu, folosind funcții HRTF neindividualizate. În timp ce, considerând rezultatele lui Hu obținute prin folosirea funcțiilor HRTF individuale, metoda de localizare este mai puțin precisă. Considerând faptul că antrenamentul rețelelor neuronale nu a plecat de la funcții HRTF personalizate subiectului folosit, ci doar de la funcții HRTF selectate pe baza anumitor parametrii antropometrici dintr-o bază de date publică, aparținând unui subiect care seamănă cu subiectul de test folosit, putem enunța că rezultatele obținute sunt de importanță ridicată. Ele sugerează obținerea unor rezultate și mai bune folosind funcții HRTF personalizate.

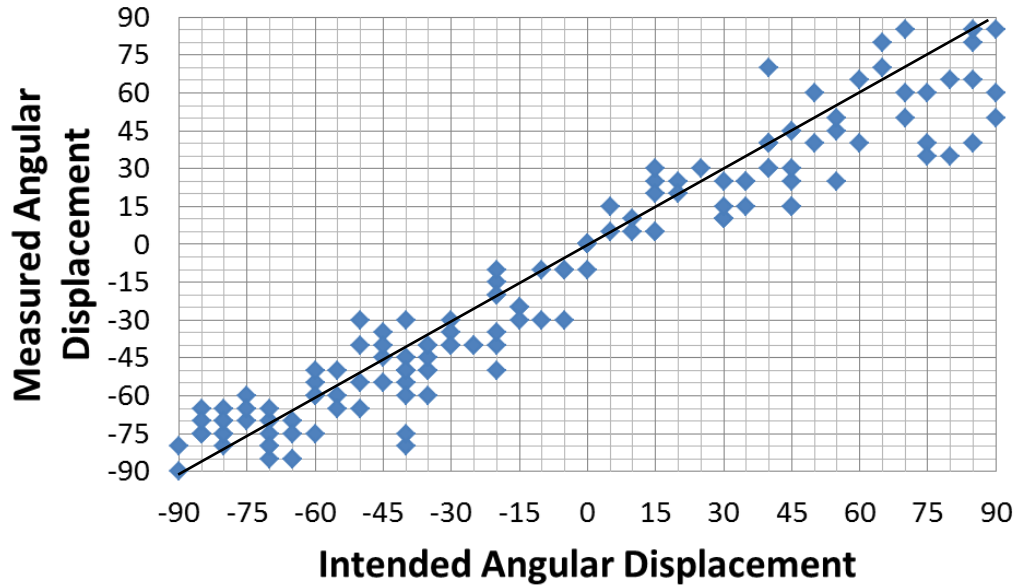


Fig. 6.11. Rezultatele de localizare obținute.

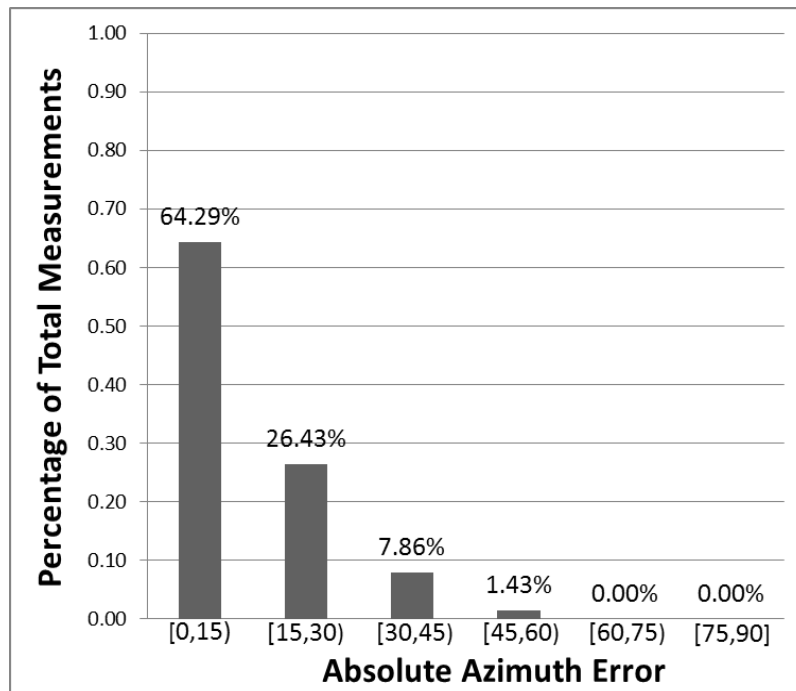


Fig. 6.12. Clasificarea erorii absolute de azimut raportate la numărul total de teste de localizare conduse.

De asemenea, după compararea rezultatelor obținute cu cele din Cap. 6.3.2, se poate observa, că experimentele au fost conduse pentru un număr mai mare de

semnale acustice (140 vs. 54). Comparând rezultatele din Tab. 6.2 cu cele din Fig. 6.12, putem spune că rezultatele din acest capitol sunt mai bune, ceea ce este o observație interesantă, deoarece asupra datelor de antrenament originale a fost aplicată o metodă de reducere a dimensionalității (în forma algoritmului propus în Cap. 3.2.3).

Bazându-se pe rezultatele obținute, se pot enunța că algoritmul de reducere propus poate fi aplicat cu succes în vederea extragerii părții relevante a funcțiilor HRIR unei persoane. Algoritmul poate fi aplicat dacă persoana în cauză sunt cunoscute funcțiile HRIR măcar într-un număr limitat de puncte (printre care să fie cele trei coordonate importante prezentate în Cap. 3.2.3). Pentru cazul în care funcțiile HRIR nu sunt disponibile pentru persoana respectivă, trebuie găsită o persoană a căror parametrii antropometrici selectați seamănă cât mai mult cu persoana în cauză, în vederea obținerii unui set de funcții HRIR cât mai apropiate de cele ale persoanei în cauză.

Algoritmul are avantajele (+) și dezavantajele (-) ei:

- Structura celor două RNA este mai simplă (numărul neuronilor din stratul de ieșire este mai mică după aplicarea algoritmului); (+)
- Timpul necesar antrenării RNA este redus; (+)
- În cazul implementării RAV pe o placă de dezvoltare (vezi Cap. 7), structura RNA este un factor principal, când vine vorba despre luarea unei decizii asupra locului (memoria internă vs. cartelă de memorie externă) de stocare a valorilor ponderilor și ale pragurilor. Acest punct poate fi tradus în timp de execuție mai lung sau mai scurt pentru fiecare ciclu de generare a semnalului acustic; (+)
- Este posibil ca precizia de localizare să fie mai scăzută (implicit precizia de localizare a obstacolelor semnalizate să fie mai scăzută), datorită filtrării funcțiilor HRTF originale, deși rezultatele obținute nu confirmă acest punct de vedere. (-)

6.5. Concluzii. Probleme întâlnite și posibile îmbunătățiri

În cele ce urmează sunt enumerate câteva posibile îmbunătățiri:

- În cadrul acestor experimente de localizare generarea surselor de sunet virtuale s-a produs doar în planul orizontal la valori de azimut între -90 și 90 de grade. O posibilă extensie este generarea surselor de sunet virtuale în tot planul orizontal. Această extensie ar da posibilitatea de evaluare a confuziilor față-spate care apar frecvent datorită folosirii căștilor (vezi Cap. 3.2.1.1). Autorul consideră că această extensie nu este neapărat necesară scopului final al acestor cercetări, care intenționează să dezvolte un dispozitiv de asistare a deplasării persoanelor cu handicap vizual, deoarece obstacolele depășite (care, conform conceptului RAV, sunt plasate deja în spate) pe ruta de deplasare a persoanei nevăzătoare nu mai trebuie să fie semnalate prin surse de sunet virtuale, deplasarea fiind în față. Încă un argument favorabil acestei decizii este dat în [358], unde autorii ajung la concluzia că percepția caracteristicilor surselor de sunet este mai sensibilă pentru surse din față, decât pentru surse aflate în spatele subiectului;
 - Metoda de reglare a deplasamentului unghiular în procesul de localizare din planul orizontal poate fi imprecisă, deoarece valoarea indicată depinde de poziția compasului de câmp, care în mod ideal trebuie ținut perfect orizontal pentru a indica valoarea corectă. O posibilă îmbunătățire ar fi găsirea unei metode precise de măsurare a deplasamentului unghiular;
 - Un alt aspect util ar fi evaluarea unei soluții de interpolare similare, care să permită localizarea obstacolelor nu numai în planul orizontal, ci și la diferite valori

de elevație (peste/sub planul orizontal). Pentru a implementa această soluție în mod eficient este necesară măsurarea funcțiilor HRTF personalizate în emisfera din fața subiectului, ca pe urmă funcțiile HRTF măsurate să fie folosite în antrenarea RNA, în vederea efectuării experimentelor de localizare;

- Altă posibilă îmbunătățire este compensarea influențelor introduse de căștile folosite în experimentele de ascultare;
- Un alt punct de vedere posibil este folosirea unor căști de calitate foarte ridicată (posibil pre-egalizare) de-a lungul experimentelor psihoacustice.

7. IMPLEMENTARE HARDWARE ȘI SOFTWARE A UNUI ECHIPAMENT PENTRU CONDUCEREA PERSOANELOR CU HANDICAP VIZUAL PE RUTE PRECIZATE FOLOSIND REALITATEA ACUSTICĂ VIRTUALĂ

În acest capitol se prezintă un prototip al dispozitivului dedicat persoanelor cu handicap vizual pentru conducerea deplasării lor în medii reale, necunoscute și dinamic schimbătoare.

Prototipul își propune să asiste deplasarea persoanelor nevăzătoare prin implementarea unei Realități Acustice Virtuale simple (vezi Cap. 2.3), care folosește semnale acustice dedicate ascultării prin intermediul unor căști. Poziția surselor virtuale ale acestor semnale acustice urmează să sugereze direcție de deplasare spre destinația dorită.

Conceptul de RAV este implementat folosind rețele neuronale pe o placă de dezvoltare la care se conectează o placă experimentală proiectată și implementată de către autorul tezei. Placa experimentală permite generarea semnalelor acustice dedicate ascultării prin intermediul unor căști folosind amplificatoare în clasa D. Întregul prototip este validat pe o rută prestabilită, unde prototipul este capabil să asiste cu succes o persoană de la începutul până la sfârșitul rutei prestabilite.

Dacă cineva dorește să genereze un sunet care vine dintr-o anumită direcție a spațiului 3D, următorii pași trebuie să fie parcurși. Prima dată, un sunet monaural trebuie ales (zgomot alb, zgomot roz, semnal sweep, etc.) pe baza aplicației. Al doilea pas constă din filtrarea acestui sunet ales cu o pereche de funcții HRTF corespunzătoare direcției alese (vezi Cap. 3.2.1). Ca rezultat al acestui pas sunt obținute două semnale binaurale, care depind de perechea de funcții HRTF, acestea având rolul de a plasa sunetul într-o anumită direcție (vezi Cap. 3.2.4).

Funcțiile HRTF sunt măsurate cu o rezoluție finită, rezultând un număr finit de direcții în spațiul 3D. Pentru eliminarea acestui dezavantaj, autorul a propus o metodă bazată pe RNA, valabilă pentru o singură persoană a căror funcții HRTF sunt cunoscute într-un număr limitat de puncte (obținute prin măsurarea acestor funcții). Ea este descrisă în [340], [241]-[243], respectiv în Cap. 5.2.2 și Cap. 6.2. Metoda folosește RNA pentru obținerea perechilor de funcții HRTF pentru o direcție dorită.

Capitolul este organizat așa cum urmează. Cap. 7.1 prezintă placa de dezvoltare LM3S8962 de la firma Luminary Micro. În Cap. 7.2 este prezentată placa experimentală, care include proiectarea și implementarea practică a părților hardware și software. Cap. 7.3 descrie pe scurt experimentul care vizează conducerea unei persoane pe o rută prestabilită. În final, Cap. 7.4 este dedicat concluziilor, problemelor întâlnite și posibilelor îmbunătățiri.

7.1. Placa de dezvoltare EK-LM3S8962

Ansamblul Stellaris LM3S8962 [359] reprezintă o platformă de evaluare compactă și versatilă bazată pe un microcontroler cu nucleu Stellaris ARM Cortex-

M3. Această placă de dezvoltare, prezentată în Fig. 7.1, este considerată o alegere optimă din punctul de vedere al puterii de calcul și al funcționalității oferite. Acestea fiind motivele principale pentru care autorul a optat să-și dezvolte prototipul bazat pe placa LM3S8962.

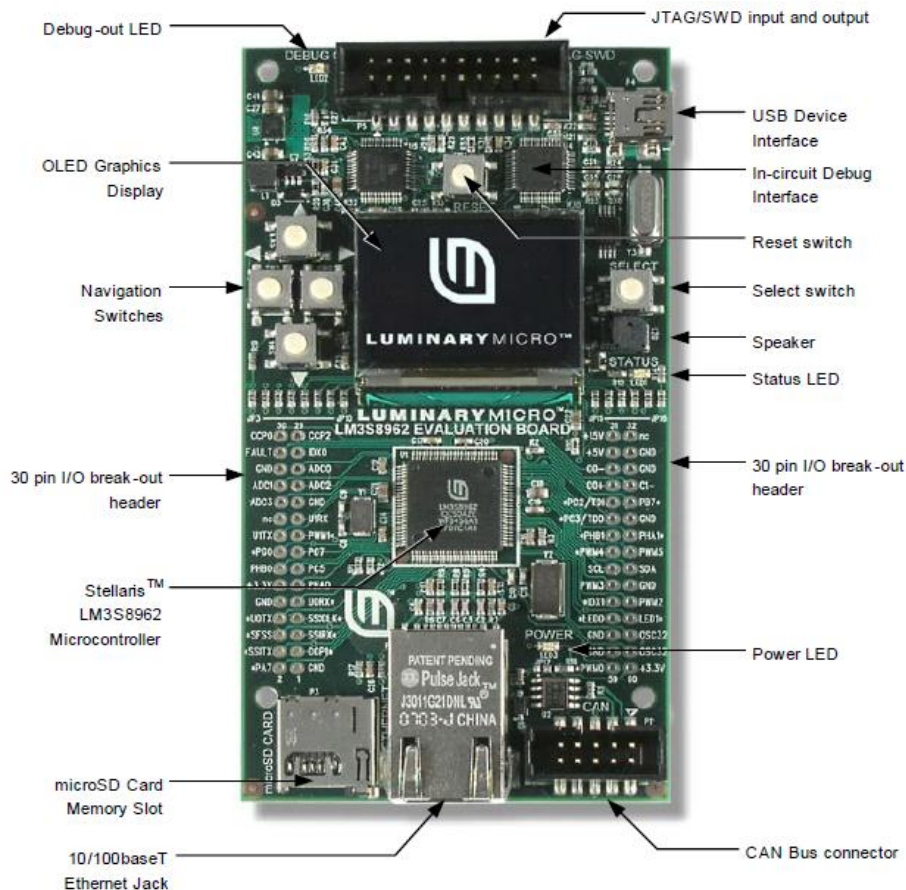


Fig. 7.1. Placă de dezvoltare Stellaris LM3S8962, imagine preluată din [359].

Fiecare placă are o interfață ICDI (In-Circuit Debug Interface) care permite depanarea hardware a dispozitivelor de pe placă prin interfața USB. Placa de dezvoltare este alimentată prin conexiune USB de la un PC. Ansamblul de evaluare conține toate cablurile, partea software, și documentația necesară pentru dezvoltarea și rularea rapidă a aplicațiilor pe microcontrolerul LM3S8962. Ansamblul include:

- Placă de evaluare-dezvoltare LM3S8962;
- Posibilitate de extindere a memoriei prin prezența unui slot micro SD;
- Microcontroler Stellaris LM3S8962 cu controller Ethernet integrat;
- Afișaj grafic OLED cu o rezoluție de 128x96 pixeli;
- LED pentru utilizator, diferite butoane de navigare, respectiv selecție;
- Difuzor magnetic;
- Pini de intrare-ieșire disponibile pentru acces ușor;

- Conector de depanare Stellaris ARM JTAG cu 20 de pini;
- Placă pentru interfața CAN construită în jurul microcontrolerului LM3S2110;
- Cablu ethernet, cablu CAN, cabluri USB și JTAG;
- CD care conține unelte software de evaluare, documentație, manual quick start, software StellarisWare, și exemple de coduri sursă.

7.2. Placa experimentală

Placa experimentală este o placă proiectată și implementată de către autor. Ea se atașează la placa de dezvoltare prin 2 conectori și reprezintă legătura dintre placa de dezvoltare LM3S8962 și căștile plasate pe capul ascultătorului. Practic, placa experimentală este responsabilă de generarea semnalelor acustice în căști pentru asistarea deplasării persoanei cu handicap vizual.

În cele ce urmează se prezintă implementarea hardware, respectiv software. Tehnici de elaborare ale sistemelor dedicate, folosite în proiectarea plăcii experimentale, pot fi consultate în [360].

7.2.1. Implementarea hardware

Partea hardware a prototipului este format din trei părți principale: placa de dezvoltare LM3S8962, placa experimentală, respectiv o pereche de căști. Placa de dezvoltare LM3S8962 reprezintă baza părții hardware, pe ea fiind construită placa experimentală. Placa experimentală a fost proiectată și implementată de către autor și este arătată în Fig. 7.2, respectiv Fig. 7.3. Căștile folosite în cadrul acestui dispozitiv au fost căștile Sennheiser HD448, ele fiind căști închise, circumaurale. Ele oferă o percepție a sunetului stereo excelentă, asigurând totodată o purtare confortabilă în jurul urechilor. Ele vin cu un cablu de extensie, lungimea maximă a cablului fiind de 3 metri. Căștile sunt conectate la placa experimentală prin intermediul unui conector jack de 3.5 mm.

În continuare este descrisă placa experimentală. Ea este construită din două componente principale: componenta audio, respectiv componenta GPS.

Scopul componentei audio este de a genera semnalele acustice corespunzătoare dedicate ascultării prin intermediul căștilor, având la intrare două semnale PWM (Pulse Width Modulation), care conțin informația audio, atât pentru urechea stângă, cât și pentru urechea dreaptă, în factorii de umplere. Pentru fiecare parte, stânga sau dreapta, semnalul PWM este aplicat direct la controlerul de comutare (switching controller), alimentat de la unul dintre regulatorii de tensiune LDO. Controlerul de comutare este urmat de etajul de ieșire, alimentat de la celălalt regulator de tensiune. Etajul de ieșire este urmat de un filtru trece-jos implementat printr-un filtru LC pasiv, al cărui frecvență de tăiere este de aproximativ 22 kHz. De observat este faptul că, toate componentele sunt montate pe partea de sus al cablajului imprimat cu excepția celor doi conectori, care asigură legătura la placa LM3S8962, respectiv a modulului GPS. Componenta audio, arătată în Fig. 7.4, are următoarele părți:

- Doi conectori (Samtec SSQ-115-01-G-D) pentru asigurarea legăturii la placa de dezvoltare LM3S8962;
- Doi conectori (Samtec TLW-110-05-G-S) pentru testare funcțională, depanare și acces ușor la semnale de interes;

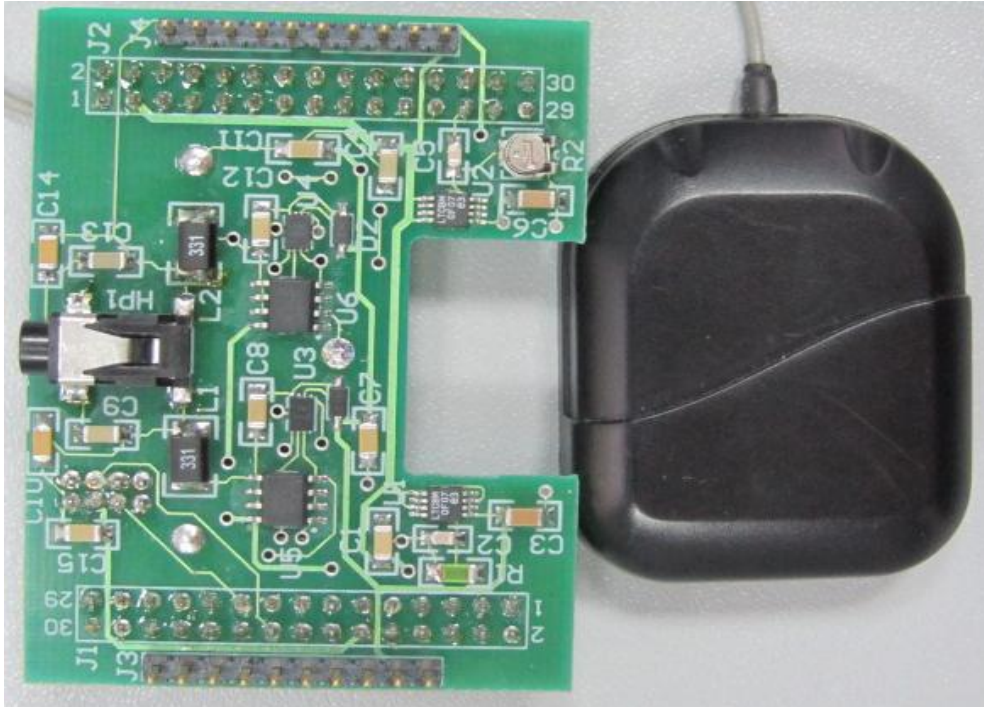


Fig. 7.2. Placa experimentală – privire de sus.



Fig. 7.3. Placa experimentală – privire de jos.

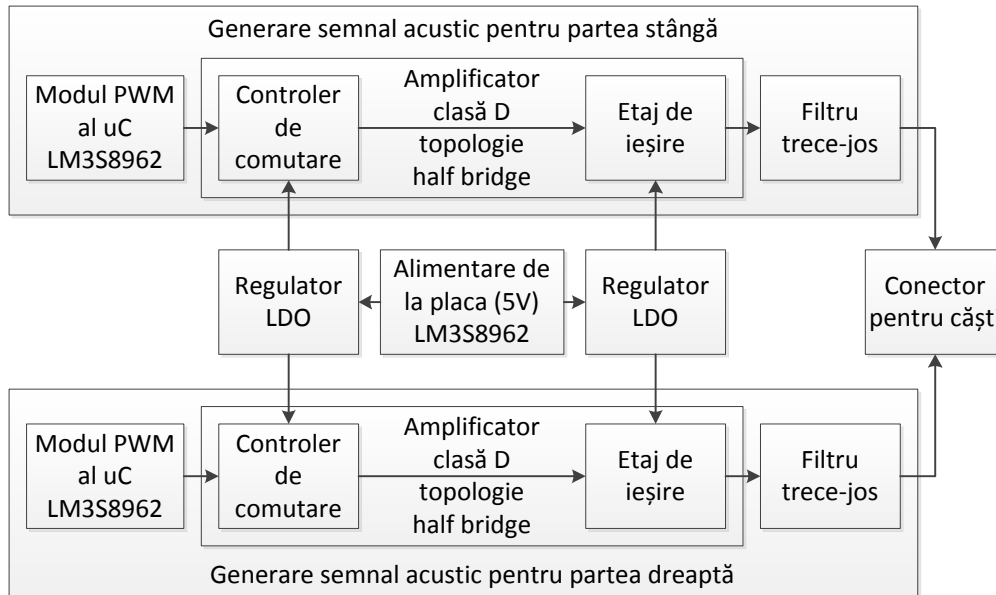


Fig. 7.4. Diagrama bloc a componentei audio a plăcii experimentale.

- Două regulatoare de tensiune LDO (Linear Technology LT3080EMS8E#PBF) pentru asigurarea alimentării celor două circuite integrate (două drivere MOSFET – controlere de comutare, respectiv două etaje de ieșire în clasă D);
- Două drivere MOSFET cu canal N de viteză mare (Linear Technology LTC4449IDCB#PBF) pentru cele două etaje de ieșire în clasă D în configurație half bridge [361]-[365];
- Două etaje de ieșire formate din tranzistoare MOSFET cu canal N cu diode Schottky (Vishay Siliconix SI4816BDY-T1-E3) reprezentând ieșirile amplificatorului în clasa D;
- Două filtre trece-jos necesare pentru rejectia componentelor de frecvență nedorite, proiectate după [365];
- Un conector de căști (Lumberg 1503 02) în vederea ascultării semnalelor acustice prin intermediul căștilor Sennheiser HD448.

De menționat este faptul că marea majoritatea a componentelor, folosite în cadrul proiectării și implementării acestui prototip, au fost obținute de la firme producătoare în forma unor mostre gratuite trimise de acestea. Celelalte componente folosite (conectorul de căști de la firma Lumberg, respectiv diodele Schottky SMD 1N5819HW de la firma Diodes Incorporated) nu au fost obținute în forma unor mostre gratuite, și au fost cumpărate de la firma Farnell (www.ro.farnell.com).

Scopul componentei GPS, arătată în Fig. 7.5, este achiziția coordonatelor GPS curente la intervale specifice, determinate de durata unui Ciclu de Generare a Semnalelor Acustice (CGSA). Aceste coordonate GPS sunt folosite de partea software pentru generarea semnalelor acustice corespunzătoare bazându-se pe coordonatele GPS precedente, respectiv cele curente. Componenta GPS este construită din modulul de recepție GPS (Epson S4E39860001), respectiv antena GPS (Epson S4E39851). Aceste două componente au fost obținute după demontarea lor de la plăci implementate în urmă cu câțiva ani.

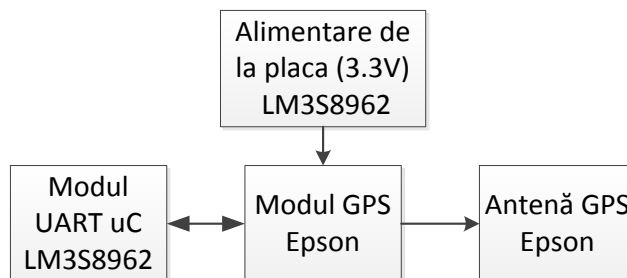


Fig. 7.5. Diagrama bloc a componentei GPS a plăcii experimentale.

Fig. 7.6 prezintă placa experimentală montată pe placa de dezvoltare Stellaris LM3S8962 cu căștile montate.

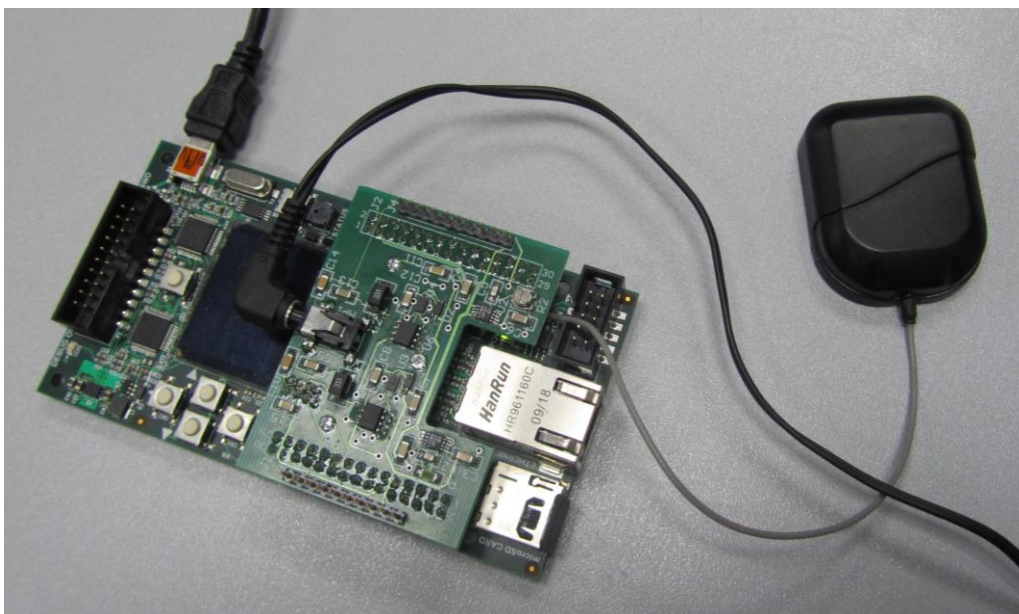


Fig. 7.6. Prototipul dispozitivului de asistare a deplasării persoanelor cu handicap vizual – partea hardware.

Dimensiunile fizice ale plăcii experimentale proiectate sunt 62 mm x 47 mm. Toate componentele au fost lipite manual pe cablajul imprimat de către tehnicianul Ivan Ciortuș din cadrul Departamentului de Electronică Aplicată.

Partea hardware (vezi Anexa 1) a fost dezvoltată folosind mediul Altium Designer Summer 09 (fostul Protel 99), dezvoltat de firma Altium Limited din Australia. Altium Designer este un pachet de software EDA (Electronic Design Automation) dedicat printre altele și proiectării cablajelor imprimate PCB (Printed Circuit Board).

Testarea funcțională a conceptului de funcționare a plăcii experimentale (doar componenta audio), a fost condusă în simulatorul gratuit LTspice IV de la firma Linear Technology (vezi Anexa 1). LTspice IV este un simulator SPICE de performanță înaltă, fiind disponibile modelele tuturor circuitelor integrate folosite în

proiectarea componentei audio. Acesta fiind și factorul decisiv pentru alegerea acestui simulator.

7.2.2. Implementarea software

Așa cum a fost descris la începutul Cap. 7, pentru obținerea funcțiilor HRTF în orice direcție dorită, metoda bazată pe RNA pentru o singură persoană (vezi Cap. 6.2) este cea implementată pe acest prototip. Implementarea metodei este posibilă doar prin stocarea întregii structuri de RNA (praguri și ponderi), antrenate în prealabil în mediul MATLAB, pe placa de dezvoltare LM3S8962.

Așa cum s-a explicat în Cap. 3.2.3, există două opțiuni pentru stocarea acestor valori:

- Prima este stocarea lor în memoria internă a microcontrolerului LM3S8962, care aduce avantajul unui timp de execuție mai scurtă prin reducerea unui CGSA;
- A doua este stocarea lor pe o cartelă de memorie externă (micro SD), care aduce dezavantajul unui timp de execuție mai lung prin prelungirea unui CGSA, din cauza necesității de citire a fiecărei valori, individual de pe cartelă de memorie.

În funcție de mărimea memoriei interne disponibile pentru utilizator, este posibil ca folosirea primei variante să fie imposibilă, datorită numărului mare de praguri și ponderi care trebuie să fie stocate (depinde de structura RNA). O variantă posibilă de reducere a structurii RNA necesare este prezentată în [240], respectiv Cap. 3.2.3, și se bazează pe reducerea eșantioanelor temporale ale funcției HRIR, care înseamnă implicit mai puține valori de stocat. Prin aplicarea acestui algoritm este posibil ca toate valorile necesare să fie stocate în memoria internă (varianta mai rapidă). Toate cele trei versiuni au fost implementate software și sunt funcționale. Schemele logice, descrise în continuare, se aplică fiecărei versiuni dintre cele trei amintite.

Programul principal, arătat în Fig. 7.7, începe cu inițializarea sistemului și a variabilelor folosite. După această fază inițială, prototipul așteaptă să recepționeze un mesaj GLL valid de la modulul GPS pentru a trece la următorul pas, care constă dintr-un CGSA. După acest ciclu, programul verifică dacă butonul STOP a fost apăsat. În cazul în care a fost apăsat prototipul intră în modul sleep și așteaptă un reset pentru a urma aceeași funcționare, cu cea prezentată în Fig. 7.7a.

CGSA (vezi Fig. 7.7b) este detaliat în cele ce urmează. La începutul ciclului, trei mărimi sunt considerate pentru determinarea direcției (specificate prin perechea azimut-elevație) următorului punct de trecere, conținut în baza de date ale coordonatelor GPS stocate, care definesc ruta prestabilită de deplasare a persoanei în cauză. Ele sunt:

- coordonatele GPS precedente valide (extrase din precedentul mesaj GPGLL valid, recepționat),
- coordonatele GPS curente valide (salvate în vederea utilizării lor ca coordonate precedente în următorul CGSA), și,
- în final, punctul de trecere următor din ruta predefinită.

De observat este că direcția următorului punct de trecere definește direcția de deplasare, în care sunt generate semnalele acustice. Odată ce se cunoaște direcția punctului de trecere următor, următorul pas este calculul perechii funcției HRIR corespunzătoare prin simularea celor două RNA stocate în prealabil. După ce a fost obținută perechea de funcții HRIR, urmează o simplă convoluție cu sunetul monaural achiziționat sau generat local, care furnizează semnalele acustice pilot pentru ambele urechi prin intermediul căștilor folosite. Aceste semnale acustice trebuie să indice direcția corectă de deplasare a persoanei în cauză spre următorul

punct de trecere de pe ruta prestabilită. În momentul de față, partea software a prototipului este proiectată să indice doar direcția de deplasare pe o rută predefinită prin coordonatele GPS salvate ca puncte de trecere.

Pentru încheierea acestei părți sunt date niște detalii tehnice.

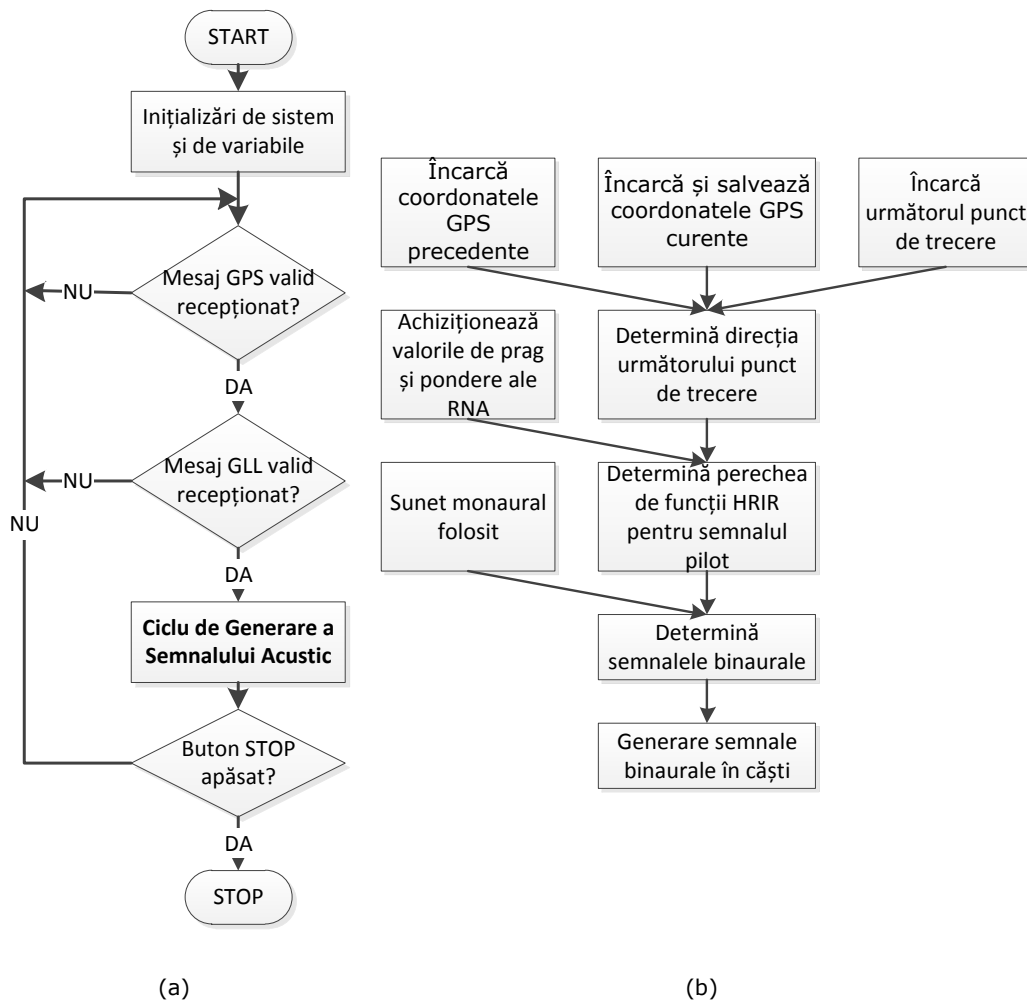


Fig. 7.7. (a) Schema logică a părții software, și (b) ciclul de generare a semnalului acustic.

Întreaga parte software a fost scrisă în limbajul C folosind mediul de dezvoltare integrat Keil μ Vision 4. O descriere detaliată despre procedura de elaborare a unei aplicații software în mediul Keil μ Vision 4 poate fi consultată în [360], scrisă de autor și alții. Biblioteca de funcții Stellaris EK-LM3S8962 Firmware (StellarisWare) a fost folosit pentru controlul eficient asupra microcontrolerului LM3S8962. În afara de această bibliotecă de funcții, autorul a mai folosit modulul FatFs pentru controlul eficient asupra sistemului de fișiere FAT folosită pentru stocarea structurii RNA (doar în versiunea de software, unde valorile pragurilor și a ponderilor celor două RNA, deja antrenate, au fost salvate pe cartela de memorie externă micro SD). Programarea și testarea întregului prototip (părțile hardware și

software) a fost făcută cu ajutorul simulatorului din mediul Keil μ Vision 4. Următorul pas a fost ca întreaga testare să fie efectuată pe direct pe prototip. Prototipul, conectat la PC prin cablul USB, apare ca un port serial virtual în sistemul de operare (Windows 7 Ultimate pe 32 biți). Consecvent, marea majoritate a depanării a fost făcută prin trimiterea mesajelor corespunzătoare la PC la atingerea unei anumite etape din program. Pentru implementarea comunicației seriale cu prototipul proiectat, autorul a dezvoltat o aplicație simplă de comunicație pe portul serial, prezentat în Fig. 7.8, folosind mediul LabVIEW de la firma National Instruments.

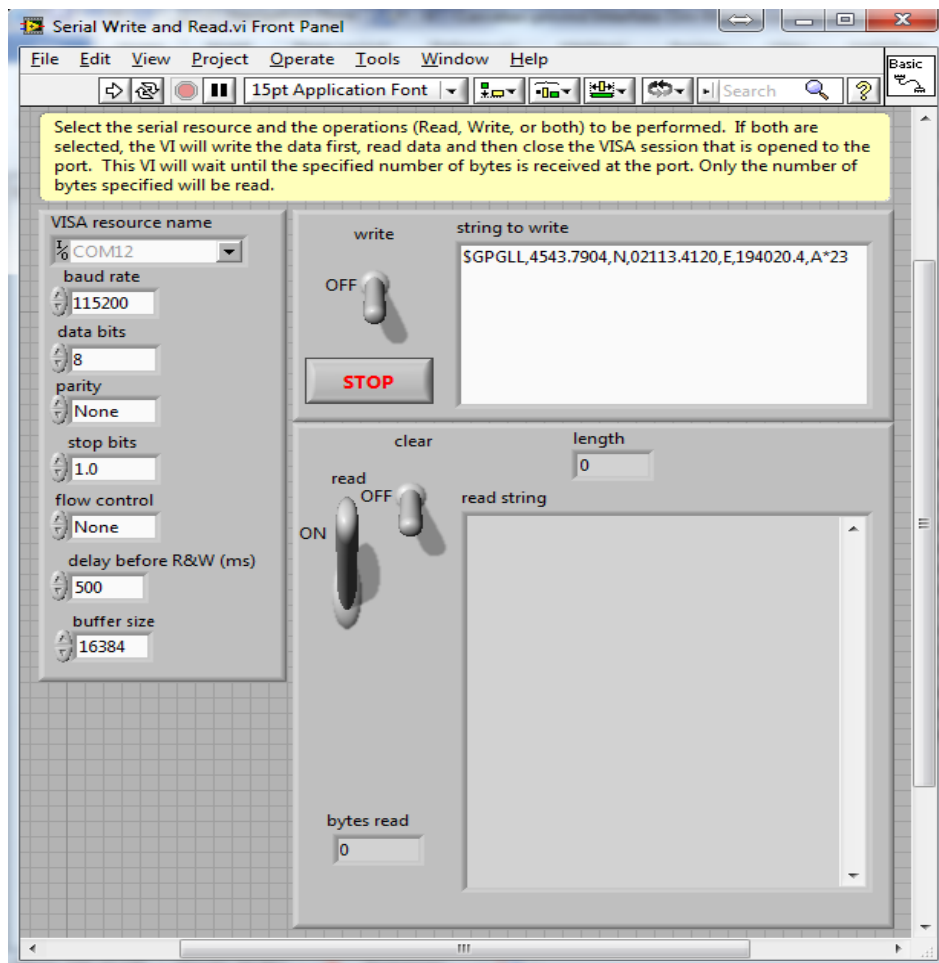


Fig. 7.8. Aplicație de comunicație pe port serial implementată în LabVIEW.

7.3. Experimente

În momentul de față, ruta predefinită este compusă dintr-un număr total de 10 coordonate GPS salvate în prealabil (numite puncte de trecere). Toate punctele de trecere sunt situate în jurul clădirii Electro al Universității Politehnica din

Timișoara. Între două puncte de trecere ruta optimă este linia dreaptă. Consecvent, dacă persoana, care folosește prototipul alege ruta optimă, semnalul pilot este plasat mereu în drept în față, indicând faptul că el/ea trebuie să-și continue deplasarea drept înainte pentru a ajunge la următorul punct de trecere. Punctele de trecere sunt numerotate de la 1 la 10, așa cum este arătat în Fig. 7.9.

De observat este faptul că punctul de trecere nr. 5 este foarte aproape de punctul de trecere nr. 6. Din această cauză punctul de trecere nr. 5 este puțin acoperit în Fig. 7.9. În momentul de față, prototipul este capabil să-și îndeplinească în mod corect obiectivul numai dacă persoana, care ascultă semnalele acustice furnizate prin căști, își ține capul orientat drept înainte. Acest dezavantaj se datorează faptului că nu există nici un dispozitiv montat pe placa experimentală, care să țină cont de orientarea momentană a capului ascultătorului.



Fig. 7.9. Ruta predefinită cu cele 10 puncte de trecere reprezentate cu ajutorul interfeței MapQuest (www.mapquest.com).

În experimentele efectuate, sunetul monaural de test, folosit pentru convoluția cu perechea de funcție HRIR, are o durată de aproximativ 2 secunde. Semnalele acustice sunt generate odată la aproximativ 6 secunde. Această durată este limitată de durată unui ciclu de generare a semnalelor acustice (CGSA). Experimentele sunt bazate pe prima opțiune dintre cele două posibile, enumerate în Cap. 3.2.3, și anume când toate valorile necesare salvării întregii structuri de RNA (ponderi și praguri) sunt stocate în memoria internă a plăcii de dezvoltare LM-3S8962 (vezi Cap. 7.1).

De asemenea, de menționat este faptul că prototipul este alimentat de la un laptop prin conexiune USB. Din această cauză experimentele conduse au fost efectuate cu un laptop plasat într-un rucsac purtat de persoana în cauză.

Subiectul de test a fost capabil să urmeze semnalele acustice generate prin atingerea punctelor de trecere de la 1 la 10 până să ajungă la destinație (punctul de trecere nr. 10). Chiar dacă subiectul a fost nevoit să devieze puțin de la rută optimă, el a reușit să ajunge la următorul punct de trecere pentru continuarea cu succes a călătoriei sale.

7.4. Concluzii. Probleme întâlnite și posibile îmbunătățiri

Autorul a reușit proiectarea și implementarea cu succes a unui prototip dedicat pentru asistarea deplasării persoanelor cu handicap vizual prin ascultarea unor semnale acustice generate prin intermediul unor căști. În momentul de față, prototipul permite doar generarea unui semnal pilot pentru conducerea persoanei la destinația dorită prin urmărirea acestui semnal.

Printre posibilele îmbunătățiri ale acestui prototip se numără extinderea funcționalității lui prin adăugarea unei părți care să sugereze prezența obstacolelor [366]-[367] printr-o realitate virtuală mai complexă, nu doar direcția de deplasare dorită pe o rută prestabilită spre destinație. De asemenea, momentan prototipul nu permite ca persoana în cauză să-și miște capul, deoarece nu există nici un dispozitiv care să țină cont de orientarea capului. În vederea eliminării acestei obligații se dorește adăugarea unui accelerometru, care să țină cont de orientarea capului, pe placa experimentală, respectiv modificarea corespunzătoare a părții software.

Idei pentru dezvoltări ulterioare sunt:

- Realizarea nu numai a simulării, ci și a antrenamentului celor două rețele neuronale pe prototip, similar cu implementarea RNA pe sisteme dedicate descrisă de Cotton în [368];
- Implementarea unui astfel de dispozitiv de asistare a deplasării persoanelor cu handicap vizual folosind tehnologia FPGA, similar cu [369].

8. CONCLUZII. CONTRIBUȚII PERSONALE. DIRECȚII VIITOARE DE CERCETARE

8.1. Concluzii

Teza intitulată "Cercetări privind interfața om-mașină cu aplicații în conducerea persoanelor cu handicap vizual cu ajutorul sunetelor", elaborată de către autor, își propune să aducă contribuții în domeniul conducerii persoanelor cu deficient de vedere prin propunerea unui sistem de asistare a deplasării dedicate acestor categorii de persoane. Un astfel de sistem poate reprezenta un ajutor real acordat acestor persoane, dacă el este elaborat în strânsă colaborare cu ele, deoarece nu este de ajuns doar simpla elaborare a lui, ci ea trebuie să fie făcută astfel încât sistemul să fie acceptat de comunitatea persoanelor nevăzătoare.

Elaborarea întregului sistem a fost realizată doar parțial, autorul concentrându-se mai mult asupra aspectelor legate de interfața om-mașină prin conceptul de Realitate Acustică Virtuală. Acest concept își propune folosirea sunetelor pe lângă indicarea direcției de deplasare spre destinație, și semnalizarea obstacolelor prezente pe ruta de deplasare a unei persoane nevăzătoare. Sunetele sunt intenționate să fie ascultate prin intermediul căștilor.

Conceptul de Realitate Acustică Virtuală se bazează pe procesul de sinteză binaurală, care reprezintă obținerea semnalelor acustice, ascultate în căști, prin convoluția unui sunet monaural cu o pereche de funcții HRTF. Datorită diferențelor individuale remarcabile ale corpului uman, funcțiile HRTF diferă de la o persoană la alta. Din această cauză, măsurarea funcțiilor HRTF este o procedură consumatoare de timp, care trebuie efectuată pentru fiecare persoană implicată în parte. Măsurarea funcțiilor HRTF este făcută în câteva direcții din spațiul 3D, de unde apare necesitatea interpolării acestor funcții HRTF. Autorul propune două metode de interpolare bazate pe rețele neuronale artificiale: o metodă pentru obținerea funcțiilor HRTF pentru o singură persoană în orice direcție, respectiv o metodă pentru obținerea acestora pentru orice persoană în orice direcție. Dintre aceste metode doar prima este validată prin experimente psihoacustice în cadrul acestei teze. Pe lângă aceste metode de interpolare, autorul mai propune un algoritm pentru reducere dimensionalității funcțiilor HRTF, care este și el validat cu succes prin metoda psihoacustică. În penultimul capitol se prezintă prototipul unui echipament de asistare a deplasării persoanelor cu handicap vizual, implementarea lui hardware și software, respectiv rezultatele obținute cu acesta. El este deocamdată capabil să conducă o persoană pe o rută prestabilită fără semnalizarea obstacolelor existente pe rută respectivă.

8.2. Contribuții personale

O sinteză a principalelor contribuții personale pe care autorul consideră că le-a adus prin elaborarea acestei teze sunt următoarele:

1. Contribuții cu caracter preponderent teoretic:

- 1.1. Propunerea unui sistem de asistare a deplasării pentru persoane cu handicap vizual, format dintr-un centru de monitorizare și echipament portabil, propus în Cap. 2.2;
- 1.2. Propunerea conceptului de Realitate Acustică Virtuală (vezi Cap. 2.1) ca interfață om-mașină dintre dispozitivul de asistare a deplasării și persoana cu handicap vizual, care asigură semnalizarea obstacolelor de pe o rută precizată, respectiv direcția spre destinație, cu ajutorul unor semnale acustice ascultate prin intermediul unor căști;
- 1.3. Propunerea unui algoritm de reducere a dimensionalității funcțiilor HRTF în MATLAB (vezi Cap. 3.2.3);
- 1.4. Propunerea unei structuri de RNA pentru interpolarea funcțiilor HRTF în orice punct din spațiul 3D pentru o anumită persoană;
- 1.5. Propunerea unei structuri de RNA pentru interpolarea funcțiilor HRTF în orice punct din spațiul 3D pentru orice persoană;
2. Contribuții cu caracter preponderent aplicativ:
 - 2.1. Implementarea unui algoritm de reducere a dimensionalității funcțiilor HRTF în MATLAB (vezi Cap. 3.2.3);
 - 2.2. Propunerea unui scenariu de măsurare pentru măsurarea funcțiilor HRTF în planul orizontal cu o rezoluție de 5 grade;
 - 2.3. Realizarea unui pachet de programe (instrumente virtuale în LabVIEW) pentru măsurarea semnalelor binaurale, respectiv extragerea funcțiilor HRTF din aceste semnale;
 - 2.4. Măsurarea răspunsului la impuls al sistemului de măsurare în vederea eliminării influențelor introduse de acesta prin compensarea funcțiilor HRTF măsurate;
 - 2.5. Realizarea unei comparații între Diferența Interaurală de Timp obținută pentru fiecare direcție de măsurare, respectiv diferite formule pentru aproximarea acesteia din literatură;
 - 2.6. Implementarea tuturor variante de interpolare în mediul MATLAB;
 - 2.7. Proiectarea și implementarea unei interfețe grafice utilizator în LabVIEW pentru generarea unei Realități Acustice Virtuale folosind diferite surse de obținere a funcțiilor HRTF (din baze de date, respectiv prin intermediul metodelor din Cap. 5);
 - 2.8. Propunerea și implementarea unui scenariu experimental de validare, prin experimente psihoacustice, a metodei de interpolare, respectiv a algoritmului de reducere a dimensionalității funcțiilor HRTF;
 - 2.9. Compararea rezultatelor obținute prin interpolarea funcțiilor HRTF pentru o singură persoană, respectiv interpolarea variantelor filtrate ale funcțiilor HRTF pentru o singură persoană;
 - 2.10. Proiectarea, simularea și implementarea părții hardware a unui echipament prototip pentru asistarea deplasării persoanelor cu handicap vizual pe rute precizate folosind Realitatea Acustică Virtuală;
 - 2.11. Proiectarea, simularea și implementarea părții software (trei variante) a unui echipament prototip pentru asistarea deplasării persoanelor cu handicap vizual pe rute precizate folosind Realitatea Acustică Virtuală;
 - 2.12. Realizarea unor experimente simple cu prototipul implementat pe o rută compusă din 10 coordonate GPS;
3. Contribuții aduse prin parcurgerea bibliografiei:

- 3.1. Realizarea unei comparații în forma Tab. 2.1 a celor mai importante sisteme de asistare a deplasării persoanelor cu handicap vizual întâlnite în literatură;
- 3.2. Parcurgerea literaturii în vederea enumerării principalelor modele ale funcțiilor HRTF (vezi Cap. 3.2.1.2);
- 3.3. Parcurgerea literaturii în vederea enumerării principalelor metode de interpolare ale funcțiilor HRTF (vezi Cap. 5.1);

În viziunea autorului cele mai importante contribuții sunt cele următoarele 1.1, 1.2, 1.4, 1.5, 2.2, 2.3, 2.10, 2.11, 3.1.

De-a lungul activității de cercetare, autorul a publicat un număr total de 19 lucrări științifice (5 ca prim autor, toate fiind enumerate în Cap. 8.4), în reviste sau în proceedings-urile unor conferințe științifice internaționale organizate în țară sau în străinătate, dintre care 6 sunt cotate ISI Proceedings, respectiv 3 sunt indexate în Baze de Date Internaționale (BDI), respectiv un număr total de 2 cărți în colaborare (vezi Cap. 8.4);

De-a lungul activității de cercetare, participarea de către autor, ca membru în echipa de cercetare, la realizarea a 2 contracte de cercetare finanțate de către Autoritatea Națională pentru Cercetare Științifică (ANCS), respectiv Unitatea Executivă pentru Finanțarea Învățământului Superior și a Cercetării Științifice Universitare (UEFISCSU).

8.3. Direcții viitoare de cercetare

Multitudinea posibilităților din domeniul de cercetare ales face ca orice încercare de abordare a acestui domeniu atât de divers să fie neexhaustivă. Această afirmație este subliniată și de faptul că până în momentul de față nu există un ghid unanim acceptat pentru asistarea deplasării persoanelor cu handicap vizual. Totodată în ceea ce privește funcțiile HRTF, care stau la baza interfeței om-mașină propuse, prin folosirea acestora pentru plasarea surselor de sunet virtuale în vederea indicării direcției de mers, respectiv a obstacolelor, nu există o metodă ideală ce permite obținerea lor în mod simplu și eficient, care să elimine necesitatea procedurii de măsurare complexe și individuale.

Sub această rezervă, autorul își propune câteva posibile direcții viitoare de cercetare:

- Îmbunătățirea procedurii de măsurare a funcțiilor HRTF descrise și crearea unei baze de date proprii;
- Validarea prin experimente psihoacustice a metodei de interpolare a funcțiilor HRTF pentru orice persoană;
- Îmbunătățirea prototipului propus prin adăugarea unei părți de detectare a obstacolelor pe ruta precizată.

8.4. Lista publicațiilor proprii

A. Articole publicate la diverse manifestări științifice interne și internaționale:

1. V. Tîponut, **Z. Haraszy**, D. Ianchis. Using a Mobile Robot as an Educational Tool, The 2006 International Conference Applied Electronics (AE2006), September 6-7, 2006, Pilsen, Czech Republic.
2. V. Tîponut, **Z. Haraszy**, D. Ianchis. Tools for Student's Project Development Acceleration, The 2006 International Conference Applied Electronics (AE2006), September 6-7, 2006, Pilsen, Czech Republic.
3. V. Tîponut, D. Ianchis, **Z. Haraszy**, S. Popescu. Monitoring System of Assisted Movement of Visually Impaired, The 6th WSEAS International Conference on Cybernetics (CIMMACS'07), December 14-16, 2007, Tenerife, Spain.
4. V. Tîponut, **Z. Haraszy**, D. Ianchis, I. Lie. Acoustic Virtual Reality Performing Man-machine Interfacing of the Blind, The 12th WSEAS International Conference on Systems, July 22-24, 2008, Heraklion, Greece.
5. **Z. Haraszy**, D. Ianchis, V. Tîponut. Generation of the Head Related Transfer Functions Using Artificial Neural Networks, The 13th WSEAS International Conference on Circuits, July 22-24, 2009, Rodos Island, Greece.
6. V. Tîponut, D. Ianchis, **Z. Haraszy**. Assisted Movement of Visually Impaired in Outdoor Environments, The 13th WSEAS International Conference on Circuits, July 22-24, 2009, Rodos Island, Greece.
7. **Z. Haraszy**, D. Ianchis, V. Tîponut. Acoustic Virtual Reality Generation with Head Related Transfer Function Visualization, Sesiunea de comunicări științifice "Doctor ETC 2009", September 24-25, 2009, Timisoara, Romania.
8. D. Ianchis, **Z. Haraszy**, V. Tîponut. Collision Detection Inspired by Locust Neural System, Sesiunea de comunicări științifice "Doctor ETC 2009", September 24-25, 2009, Timisoara, Romania.
9. R. Mirsu, V. Tîponut, L. Petromanjanc, **Z. Haraszy**. Improved p-Delta Learning Algorithm, The 14th WSEAS International Conference on Systems, July 22-24, 2010, Corfu Island, Greece.
10. V. Tîponut, D. Ianchis, M. Basch, **Z. Haraszy**. Work Directions and New Results in Electronic Travel Aids for Blind and Visually Impaired People, The 14th WSEAS International Conference on Systems, July 22-24, 2010, Corfu Island, Greece.
11. **Z. Haraszy**, S. Micut, V. Tîponut, T. Slavici. Multi-Subject Head Related Transfer Function Generation using Artificial Neural Networks, The 14th WSEAS International Conference on Systems, July 22-24, 2010, Corfu Island, Greece.
12. **Z. Haraszy**, D.-G. Cristea, V. Tîponut, T. Slavici. Improved Head Related Transfer Function Generation and Testing for Acoustic Virtual Reality Development, The 14th WSEAS International Conference on Systems, July 22-24, 2010, Corfu Island, Greece.
13. D.-G. Cristea, Y. Shacham-Diamand, V. Tîponut, H. Ben-Yoav, M. Basch, **Z. Haraszy**. VLSI Universal Signal Conditioning Circuit for Electrochemical and Bioluminescent Sensors, The 26th IEEE Convention of Electrical and Electronics Engineers in Israel, November 17-20, 2010, Eilat, Israel.
14. V. Tîponut, D. Ianchis, **Z. Haraszy**, M. Basch, I. Bogdanov. Work Directions and New Results in Electronic Travel Aids for Blind and Visually Impaired People, Journal of WSEAS Transactions on Systems, vol. 9, n. 10, October 2010, pp. 1086-1097.

15. D.-G. Cristea, M. E. Basch, **Z. Haraszy**, N. Fishelson, V. Taponut. A New Electronic Equivalent Circuit for Electrochemical Cells, The 2nd International Conference on Modelling and Simulations Virtual Forum (ICOMOS2011 - VF), July 11-25, 2011.

16. Z. Haraszy, D.-G. Cristea, S. Micut, I. Bogdanov. Efficient Algorithm for Extracting Essential Head Related Impulse Response Data for Acoustic Virtual Reality Development, The 15th WSEAS International Conference on Systems, July 14-16, 2011, Corfu Island, Greece.

17. S. Micut, **Z. Haraszy**, V. Taponut, D. Ianchis, A New QCM Based E-NOSE Model Using Decay Method, The 15th WSEAS International Conference on Circuits, July 14-16, 2011, Corfu Island, Greece.

18. D. Ianchis, V. Taponut, S. Popescu, **Z. Haraszy**, Improved Collision Detection System Inspired from the Neural Network of the Locust, The 9th IEEE International Symposium on Intelligent Systems and Informatics (SISY 2011), September 8-10, 2011, Subotica, Serbia.

19. D.-G. Cristea, M.-E. Basch, **Z. Haraszy**, R.-I. Lorincz, V. Taponut. A New Model For Simulating Electrochemical Sensors Array, The 17th IEEE International Symposium for Design and Technology in Electronic Packaging (SIITME 2011), October 20-23, 2011, Timisoara, Romania.

B. Cărți publicate:

1. D. Isar, **Z. Haraszy**. Echipamente Electronice de Interfațare – Curs pentru învățământ la distanță, Tipografia Universității "Politehnica" din Timișoara, 2009.

2. V. Taponut, D. Ianchis, **Z. Haraszy**, T. Slavici. Tehnici de elaborare ale sistemelor dedicate, Fundația "Ioan Slavici" Timișoara, ISBN: 978-606-92306-7-1, 2010.

C. Granturi de cercetare:

Numele proiectului	Director de grant	Poziție ocupată	Program	Contract (finanțator, nr., valoare, interval)
1. Research on Emotional Facial Expression recognition in Complicated Environment	Conf.Dr.Ing. Cătălin-Daniel CĂLEANU	Membru în echipa de cercetare	PN II, CAPACITATI, Modul III, proiecte de cercetare bilaterale, România-China, 39-5/2008	ANCS, 222/15.04.2009, 22 luni, 29304.45 lei.
2. Noi metode de analiză și recunoaștere a expresiei faciale	Conf.Dr.Ing. Cătălin-Daniel CĂLEANU	Membru în echipa de cercetare	PN II, IDEI, Proiecte de cercetare exploratorie, cod 945/2008	UEFISCSU, 599/19.01.2009, 36 luni, 277416.15 lei.

Tab. 8.1. Granturi de cercetare.

BIBLIOGRAFIE

- [1] A. S. Bregman, Auditory Scene Analysis (The MIT Press, 1990).
- [2] P. Mackensen, Auditive Localization. Head movement, an additional cue in Localization, Ph.D. dissertation, Technical University of Berlin, Berlin, Germany, 2004.
- [3] J. Blauert, Spatial Hearing – The Psychophysics of Human Sound Localization, Revised Edition (The MIT Press, 1997).
- [4] D. R. Begault, 3-D Sound for Virtual Reality and Multimedia (Academic Press Professional, Inc., 2000).
- [5] M. Vorländer, Auralization, Fundamentals of Acoustics, Modelling, Simulation, Algorithms and Acoustic Virtual Reality, First edition (Springer Verlag, Inc. 2008).
- [6] B. Thylefors, A.-D. Negrel, R. Pararajasegaram, K. Y. Dadzie, Global data on blindness, Bulletin of the World Health Organization, vol. 73, n. 1, January 1995, pp. 115-121.
- [7] S. Resnikoff, D. Pascolini, D. Etya'ale, I. Kocur, R. Pararajasegaram, G. P. Pokharel, S. P. Mariotti, Global data on visual impairment in the year 2002, Bulletin of the World Health Organization, vol. 82, n. 11, November 2004, pp. 844-851.
- [8] L. Pizzarello, A. Abiose, T.Ffytche, R. Duerksen, R. Thulasiraj, H. Taylor, H. Faal, G. Rao, I. Kocur, S. Resnikoff, VISION 2020: The Right to Sight – A Global Initiative to Eliminate Avoidable Blindness, Archives of Ophthalmology, vol. 122, n. 4, April 2004, pp. 615-620.
- [9] World Health Organization, Fact sheet 282, VISION 2020: The Right to Sight, the Global Initiative for the Elimination of Avoidable Blindness. Magnitude and causes of visual impairment, December 2004, http://www.vision2020.org/documents/WHO%20Publications/WHO_282_Magnitude_and_causes_of_visual_impairment.pdf.
- [10] L. Dandona, R. Dandona, What is the global burden of visual impairment?, BMC Medicine, vol. 4, n. 1, March 2006.
- [11] World Health Organization, Action plan for the prevention of avoidable blindness and visually impairment 2009-2013, 2010, http://www.who.int/blindness/ACTION_PLAN_WHA62-1-English.pdf.
- [12] World Health Organization, Fact sheet No. 282, Visual impairment and blindness, <http://www.who.int/mediacentre/factsheets/fs282/en/index.html>, April 2011.
- [13] World Health Organization and International Agency for the Prevention of Blindness, VISION 2020 Global Facts, June 2011, <http://www.vision2020.org/documents/Global%20Facts%20June%202011.pdf>.
- [14] H. Moller, Fundamentals of Binaural Technology, Applied Acoustics, vol. 36, n. 3-4, 1992, pp. 171-218.
- [15] H. W. Gierlich, The Application of Binaural Technology, Applied Acoustics, vol. 36, n. 3-4, 1992, pp. 219-243.

-
- [16] K. M. Stanney, R. R. Mourant, R. S. Kennedy, Human Factors Issues in Virtual Environments: A Review of the Literature, *Teleoperators and Virtual Environments - Presence*, vol. 7, n. 4, August 1998, pp. 327-351.
- [17] R. T. Azuma, A Survey of Augmented Reality, *Teleoperators and Virtual Environments - Presence*, vol. 6, n. 4, August 1997, pp. 355-385.
- [18] M. J. Schuemie, P. Van Der Straaten, M. Krijn, C. A. van der Mast, Research on Presence in Virtual Reality: A Survey, *Cyberpsychology and Behavior*, vol. 4, n. 2, 2001, pp. 183-201.
- [19] D. Laurendeau, A. Branzan-Albu, E. Boivin, R. Drouin, H. Martel, D. Ouellet, J.-M. Schwartz, Survey of the State-of-the-Art on Synthetic Environments, Sensori-Motor Activities in Synthetic Environments, Simulation Frameworks and Real-World Abstraction Models, Contract W7701-023981/QCA, Final Report, Computer Vision and Systems Laboratory, Department of Electrical and Computer Engineering, Laval University, Québec City, Canada, 2003.
- [20] T. C. Davies, C. M. Burns, S. D. Pinder, Mobility Interfaces for the Visually Impaired: What's missing?, *Mobility interfaces for the visually impaired: what's missing?*, The 7th ACM SIGCHI New Zealand Chapter's International Conference On Computer-Human Interaction: design centered HCI, July 1-4, 2007, Hamilton, New Zealand.
- [21] Q. Zhao, A survey on virtual reality, *Science in China Series F: Information Sciences*, vol. 52, n. 3, March 2009, pp. 348-400.
- [22] C. Power, H. Jurgensen, Accessible presentation of information for people with visual disabilities, *Universal Access in the Information Society*, vol. 9, n. 2, 2009, pp. 97-119.
- [23] J. M. Loomis, J. R. Marston, R. G. Golledge, R. L. Klatzky, Personal Guidance System for People with Visual Impairment: A Comparison of Spatial Displays for Route Guidance, *Journal of Visual Impairment and Blindness*, vol. 99, n. 4, April 2005, pp. 219-232.
- [24] M. A. Hersh, M. A. Johnson, *Assistive Technology for Visually Impaired and Blind People* (Springer Verlag, Inc. 2008).
- [25] N. A. Giudice, G. E. Legge, Blind navigation and the role of technology, In A. Helal, M. Mokhtari, B. Abdulrazak (Eds.), *The Engineering Handbook of Smart Technology for Aging, Disability and Independence* (John Wiley & Sons, 2008, 479-500).
- [26] R. Velázquez, *Wearable Assistive Devices for the Blind*, In A. Lay-Ekuakille, S. C. Mukhopadhyay (Eds.), *Wearable and Autonomous Biomedical Devices and Systems for Smart Environment: Issues and Characterization* (Springer Verlag, 2010, 331-349).
- [27] D. Dakopoulos, N. G. Bourbakis, *Wearable Obstacle Avoidance Electronic Travel Aids for Blind: A Survey*, *IEEE Transactions on Systems, Man, and Cybernetics - Part C: Applications and Reviews*, vol. 40, n. 1, January 2010, 25-35.
- [28] E. M. Havik, F. J. J. M. Steyvers, H. van der Velde, J. C. Pinkster, A. C. Kooijman, Design and Evaluation of a Protocol to Assess Electronic Travel Aids for Persons Who Are Visually Impaired, *Journal of Visual Impairment and Blindness*, vol. 104, n. 2, February 2010, pp. 84-94.
- [29] J. M. Loomis, R. G. Golledge, R. L. Klatzky, Navigation System for the Blind: Auditory Display Modes and Guidance, *Presence*, vol. 7, n. 2, April 1997, pp. 193-203.
- [30] S. Shoval, I. Ulrich, J. Borenstein, Computerized obstacle avoidance systems for the blind and visually impaired, In H. N. L. Teodorescu and L. C. Jain

- (Eds.), *Intelligent Systems and Technologies in Rehabilitation Engineering* (Boca Raton: CRC Press LLC, 2001, 414-448).
- [31] S. Shoval, I. Ulrich, J. Borenstein, *Robotics-based Obstacle Avoidance Systems for the Blind and Visually Impaired*, Invited article for the *IEEE Robotics and Automation Magazine*, Special Issue on Robotics in Bio-Engineering, vol. 10 n. 1, March 2003, pp. 9-20.
- [32] A. Helal, S. E. Moore, B. Ramachandran, *Drishti: An Integrated Navigation System for Visually Impaired and Disabled*, The 5th IEEE International Symposium on Wearable Computers (ISWC'01), October 7-9, 2011, Zurich, Switzerland.
- [33] F. Fontana, A. Fusiello, M. Gobbi, V. Murino, D. Roccheso, L. Sartor, A. Panuccio, *A Cross-Modal Electronic Travel Aid Device*, In F. Paterno (Ed.), *Mobile HCI* (Berlin: Springer-Verlag, 2002, 393-397).
- [34] H. Shim, J. Lee, E. Lee, *A Study on the Sound-Imaging Algorithm of Obstacles Information for the Visually Impaired*, The 2002 International Conference on Circuits/Systems, Computers and Communications (ITC-CSCC2002), July 16-19, 2002, Phuket Arcadia, Phuket, Thailand.
- [35] V. Garaj, R. Jirawimut, P. Ptasinski, F. Cecelja, W. Balachandran, *A system for remote sighted guidance of visually impaired pedestrians*, *The British Journal of Visual Impairment*, vol. 21, n. 2, 2003, pp. 55-63.
- [36] V. Kulyukin, C. Gharpure, J. Nicholson, S. Pavithran, *RFID in Robot-Assisted Indoor Navigation for the Visually Impaired*, The 2004 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2004), September 28 - October 2, 2004, Sendai, Japan.
- [37] M. Bousbia-Salah, M. Fezari, *An Electronic Travel Aid for Blind People*, The 17th IMACS World Congress (IMACS'2005), July 11-15, 2005, Paris, France.
- [38] R. Damaschini, R. Legras, R. Leroux, R. Farcy, *Electronic Travel Aid for blind people*, The 8th European Conference for the Advancement of Assistive Technology in Europe (AAATE 2005), September 6-9, 2005, Lille, France.
- [39] R. Farcy, R. Leroux, A. Jucha, R. Damaschini, C. Grégoire, A. Zogaghi, *Electronic Travel Aids and Electronic Orientation Aids for blind people: technical, rehabilitation and everyday life points of view*, Conference on Assistive Technology for Vision and Hearing Impairment (CVHI 2006), July 18-21, 2006, Kufstein, Austria.
- [40] C. I. Tan, C. S. Tan, M. Moghavvemi, *Electronic Travel Aid for Visually Impaired*, IEEE International Conference on Electrical and Electronics, 2005.
- [41] G. Balakrishnan, G. Sainarayanan, R. Nagarajan, S. Yaacob, *Wearable Real-Time Stereo Vision for the Visually Impaired*, *Engineering Letters*, vol. 14, n. 2, 2007, pp. 6-14.
- [42] S. Mau, N. A. Melchior, M. Makatchev, A. Steinfeld, *BlindAid: An Electronic Travel Aid for the Blind*, CMU-RI-TR-07-39, Technical Report, The Robotics Institute, Carnegie Mellon University, Pittsburgh, Pennsylvania, 2008.
- [43] A. Hub, *Precise Indoor and Outdoor Navigation for the Blind and Visually Impaired Using Augmented Maps and the TANIA System*, The 9th International Conference on Low Vision (VISION 2008), July 7-11, 2008, Montreal, Quebec, Canada.
- [44] A. Hub, B. Schmitz, *Addition of RFID-Based Initialization and Object Recognition to the Navigation System TANIA*. The 24th Annual International Technology and Persons with Disabilities Conference (CSUN 2009), March 16-21, 2009, San Diego, California, USA.

- [45] B. Schmitz, The ViibraCane - A White Cane for Tactile Navigation Guidance, The 25th Annual International Technology and Persons with Disabilities Conference (CSUN 2010), March 22-27, 2010, San Diego, California, USA.
- [46] M. Bujacz, P. Strumillo, Stereophonic representation of virtual 3D scenes - A simulated mobility aid for the blind, University of Bialystok Science Letters, 134, pp. 157,162, 2006.
- [47] P. Strumiłło, P. Pełczyński, M. Bujacz, M. Pec, Space Perception by means of Acoustic Images: An Electronic Travel Aid for the Blind, ACOUSTICS High Tatras 06 - 33rd International Acoustical Conference - EAA Symposium, October 4-6, 2006, Štrbské Pleso, Slovakia.
- [48] M. Bujacz, P. Strumillo, Synthesizing a 3D auditory scene for use in an electronic travel aid for the blind, Photonics Applications in Astronomy, Communications, Industry, and High-Energy Physics Experiments 2007 (SPIE 2007).
- [49] M. Bujacz, P. Baranski, M. Moranski, P. Strumiłło, A. Materka, Remote Guidance for the Blind - A Proposed Teleassistance System and Navigation Trials, 2008 Conference on Human System Interactions (HSI 2008), May 25-27, 2008, Krakow, Poland.
- [50] M. Bujacz, P. Barański, M. Morański, P. Strumiłło, A. Materka, Remote mobility and navigation aid for the visually disabled, The 7th International Conference on Disability, Virtual Reality and Associated Technologies with Art ArtAbilitation (ICDVRAT 2008), September 8-11, 2008, Maia, Portugal.
- [51] M. Bujacz, P. Skulimowski, G. Wróblewski, A. Wojciechowski, P. Strumiłło, A proposed method for sonification of 3D environments using scene segmentation and personalized spatial audio, The Conference & Workshop on Assistive Technologies for People with Vision & Hearing Impairments Past Successes and Future Challenges (CVHI 2009), April 20-24, 2009, Wrocław, Poland.
- [52] P. Baranski, P. Strumillo, M. Bujacz, A. Materka, A Remote Guidance System Aiding the Blind in Urban Travel, The Conference & Workshop on Assistive Technologies for People with Vision & Hearing Impairments Past Successes and Future Challenges (CVHI 2009), April 20-24, 2009, Wrocław, Poland.
- [53] P. Baranski, M. Polaczyk, P. Strumillo, A Remote Guidance System for the Blind, The 12th IEEE International Conference on e-Health Networking Applications and Services (IEEE HealthCom 2010), July 1-3, 2010, Lyon, France.
- [54] M. Bujacz, Representing 3D scenes through spatial audio in an electronic travel aid for the blind, Ph.D. dissertation, Faculty of Electrical, Electronic, Computer and Control Engineering, Technical University of Lodz, Lodz, Poland, 2010.
- [55] N. Bourbakis, Sensing Surrounding 3-D Space for Navigation of the Blind, IEEE Engineering in Medicine and Biology Magazine, vol. 27, n. 1, January-February 2008, pp. 49-55.
- [56] K. Beydoun, V. Felea, H. Guyennet, Wireless sensor network system helping navigation of the visually impaired, The IEEE International Conference on Information and Communication Technologies: from Theory to Applications (ICTTA'08), April 7-11, 2008, Umayyad Palace, Damascus, Syria.
- [57] L. Kim, S. Park, S. Lee, S. Ha, An electronic traveler aid for the blind using multiple range sensors, IEICE Electronics Express, vol. 6, n. 11, 2009, pp. 794-799.

- [58] C. Gearhart, A. Herold, B. Self, C. Birdsong, L. Slivovsky, Use of Ultrasonic Sensors in the Development of an Electronic Travel Aid, IEEE Sensors Applications Symposium (SAS 2009), February 17-19, 2009, New Orleans, Los Angeles, USA.
- [59] S. Kursun, Wearable Obstacle Detection System Integrated with Conductive Fibers for Blinds, Ph.D. dissertation, Textile Technologies and Design Faculty, Istanbul Technical University, Istanbul, Turkey, 2009.
- [60] G. P. Fajarnes, L. Dunai, V. S. Praderas, I. Dunai, CASBlIP - A New Cognitive Object Detection and Orientation System for Impaired People, The 4th International Conference on Cognitive Systems (CogSys 2010), January 27-28, 2010, ETH Zurich, Switzerland.
- [61] R. Ivanov, Indoor Navigation System for Visually Impaired, The 11th International Conference on Computer Systems and Technologies and Workshop for PhD Students in Computing on International Conference on Computer Systems and Technologies (CompSysTech'10), June 17-18, 2010, Sofia, Bulgaria.
- [62] A. Kumar, R. Patra, M. Manjunatha, J. Mukhopadhyay, A. K. Majumdar, An Electronic Travel Aid for Navigation of Visually Impaired Persons, The 3rd International Conference on Communication Systems and Networks (COMSNETS 2011), January 4-8, 2011, Bangalore, India.
- [63] P. H. Tarange, S. T. Gawas, R. B. Lohani, Design and Implementation of Electronic Aid for Visually Challenged, National Conference on Electronic Technologies (NCET 2011), April 15-16, 2011, Goa, India.
- [64] J. José, M. Farrajota, J. M. F. Rodrigues, J. M. Hans du Buf, The Smart Vision Local Navigation Aid for Blind and Visually Impaired Persons, International Journal of Digital Content Technology and its Applications, vol. 5, n. 5, May 2011, pp. 362-375.
- [65] International Organization for Standardization (ISO), Normal equal-loudness-level contours - ISO 226:2003 Acoustics, 2nd edition, 2003.
- [66] M. Ohuchi, Y. Iwaya, Y. Suzuki, T. Munekata, A comparative study of sound localization acuity of congenital blind and sighted people, Acoustical Science and Technology, vol. 27, n. 5, September 2006, pp. 290-293.
- [67] H. Moller, M. F. Sorensen, C. B. Jensen, D. Hammershoi, Binaural Technique: Do We Need Individual Recordings?, The Journal of Audio Engineering Society, vol. 44, n. 6, June 1996, pp. 451-469.
- [68] T. Liitola, Headphone Sound Externalization, M.Sc. thesis, Laboratory of Acoustics and Audio Signal Processing, Department of Electrical and Communications Engineering, Helsinki University of Technology, Helsinki, Finland, 2006.
- [69] V. Tiponut, D. Ianchis, Z. **Haraszy**, Assisted Movement of Visually Impaired in Outdoor Environments - Work Directions and New Results, The 13th WSEAS International Conference on Systems (CSCC '09), July 22-24, 2009, Rodos Island, Greece.
- [70] Z. **Haraszy**, Sistem de conducere prin GPRS/GPS al persoanelor cu handicap vizual, B.Sc. thesis, Department of Applied Electronics, Faculty of Electronics and Telecommunications, Politehnica University of Timisoara, Timisoara, Romania, 2007.
- [71] T. Coverstone, C. Cronin, S. Kniazeva, GPS Technology to Aid the Blind and Partially Sighted in Copenhagen, Denmark, B.Sc. thesis, Worcester Polytechnic Institute, Worcester, Massachusetts, USA, 2007.

-
- [72] S. Carcieri, S. Morris, B. D. Perry, RFID Technology to Aid in Navigation and Organization for the Blind and Partially Sighted, B.Sc. thesis, Worcester Polytechnic Institute, Worcester, Massachusetts, USA, 2009.
- [73] V. Tiponut, D. Ianchis, Z. **Haraszy**, S. Popescu, Monitoring System of Assisted Movement of Visually Impaired, The 6th WSEAS International Conference on Computational Intelligence, Man-Machine Systems and Cybernetics (CIMMACS '07), December 14-16, 2007, Tenerife, Spain.
- [74] V. Tiponut, Z. **Haraszy**, D. Ianchis, I. Lie, Acoustic Virtual Reality Performing Man-Machine Interface of the Blind, The 12th WSEAS International Conference on Systems (CSCC '08), July 22-24, 2008, Heraklion, Greece.
- [75] P. Minaar, S. K. Olesen, F. Christensen, H. Moller, Localization with Binaural Recordings from Artificial and Human Heads, The Journal of the Audio Engineering Society, vol. 49, n. 5, May 2001, pp. 323-336.
- [76] B. G. Shinn-Cunningham, I-F. Lin, T. Streeter, Trading Directional Accuracy for Realism in a Virtual Auditory Display, The 11th International Conference on Human-Computer Interaction / 1st International Conference on Virtual Reality (HCII 2005), 22-27 July 2005, Las Vegas, Nevada, USA.
- [77] D. R. Begault, E. M. Wenzel, Direct Comparison of the Impact of Head Tracking, Reverberation, and Individualized Head-Related Transfer Functions on the Spatial Perception of a Virtual Speech Source, The Journal of the Audio Engineering Society, vol. 46, n. 10, October 2001, pp. 904-916.
- [78] B. Kapralos, M. Jenkin, E. Milios, Auditory Perception and Spatial (3D) Auditory Systems, Technical Report CS-2003-07, York University, Ontario, Canada.
- [79] S. Brewster, Providing a Structured Method for Integrating Non-Speech Audio into Human-Computer Interfaces, Ph.D. dissertation, Human-Computer Interaction Group, Department of Computer Science, University of York, England, 1994.
- [80] Gy. Wersényi, HRTFs in Human Localization: Measurement, Spectral Evaluation and Practical Use in Virtual Audio Environment, Fakultät für Maschinenbau, Elektrotechnik und Wirtschaftsingenieurwesen der Brandenburgischen Technischen Universität Cottbus, Cottbus, Germany, 2002.
- [81] V. Tiponut, S. Popescu, I. Bogdanov, C. Căleanu, Obstacles Detection System for Visually Impaired Guidance, The 12th WSEAS International Conference on Systems (CSCC '08), July 22-24, 2008, Heraklion, Greece.
- [82] S. C. Peres, V. Best, D. Brock, C. Frauenberger, T. Hermann, J. Neuhoff, L. V. Nickerson, B. Shinn-Cunningham, A. Stockman, Auditory interfaces, In P. Kortum (Ed.), HCI Beyond the GUI (San Francisco: Morgan Kaufman, 2008, 147-195).
- [83] W. M. Hartmann, How we localize sound, Physics Today, vol. 52, n. 11, November 1999, pp. 24-29.
- [84] B. Kapralos, Auditory Perception and Virtual Environments, Ph.D. dissertation, Department of Computer Science, York University, North York, Ontario, Canada, 2003.
- [85] Z. **Haraszy**, Studiu privind circuitele VLSI neuromorfe cu aplicații în interfața om-mașină, M.Sc. thesis, Department of Applied Electronics, Faculty of Electronics and Telecommunications, Politehnica University of Timisoara, Timisoara, Romania, 2009.
- [86] C. J. Pu, A neuromorphic microphone for sound localization, Ph.D. dissertation, University of Florida, Florida, USA, 1998.

- [87] P. Lessing, A Binaural 3D Sound System Applied to Moving Sources, M.Sc. thesis, Institute of Electronic Music and Acoustics, University of Music and Dramatic Arts of Graz, Graz, Austria, 2004.
- [88] A. Sibbald, Virtual audio for headphones, Sensaura, 2000.
- [89] J. W. Scarpaci, Creation of a system for real time virtual auditory space and its application to dynamic sound localization, Ph.D. dissertation, College of Engineering, Boston University, Boston, Massachusetts, 2006.
- [90] R.O. Duda, CIPIC Interface Laboratory - HRTF Tutorial, Department of Electrical Engineering, Department of Electrical and Computer Engineering, University of California Davis, February 2011, <http://interface.cipic.ucdavis.edu/sound/tutorial/psych.html#coord>.
- [91] M. Otani, T. Hirahara, Numerical study on source-distance dependency of head-related transfer functions, The Journal of the Acoustical Society of America, vol. 125, n. 5, May 2009, pp. 3253-3261.
- [92] E. W. Wenzel, M. Arruda, D. J. Kistler, F. L. Wightman, Localization using nonindividualized head-related transfer functions, The Journal of the Acoustical Society of America, vol. 94, n. 1, July 1993, pp. 111-123.
- [93] P. M. Hofman, J. G. A. Van Riswick, A. J. Van Opstal, Relearning sound localization with new ears, Nature Neuroscience, vol. 1, n. 5, September 1998, pp. 417-421.
- [94] W. G. Gardner, 3D Audio and Acoustic Environment Modeling, Wave Arts, Inc., 1999, Arlington, Massachusetts, USA.
- [95] F. L. Wightman, D. J. Kistler, Resolution of front-back ambiguity in spatial hearing by listener and source movement, The Journal of the Acoustical Society of America, vol. 105, n. 5, May 1999, pp. 2841-2853.
- [96] P. A. Hill, P. A. Nelson, O. Kirkeby, H. Hamada, Resolution of front-back confusion in virtual acoustic imaging systems, The Journal of the Acoustical Society of America, vol. 108, n. 6, December 2000, pp. 2901-2910.
- [97] Y. Iwaya, Y. Suzuki, D. Kimura, Effects of head movement on front-back error in sound localization, Acoustical Science and Technology, vol. 24, n. 5, September 2003, pp. 322-324.
- [98] Y. Iwaya, Sound space perception in virtual environments with head movements, The 1st International Workshop on the Principles and Applications of Spatial Hearing (IWPASH 2009), November 11-13, 2009, Miyagi, Japan.
- [99] N. Gupta, A. Barretto, C. Ordonez, Spectral modification of Head-Related Transfer Functions for improved virtual sound spatialization, The 2002 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP2002), May 13-17, 2002, Orlando, Florida, USA.
- [100] K. Watanabe, R. Kodama, S. Sato, S. Takane, K. Abe, Influence of simplifying the spectral form of HRTFs on the contralateral side upon localization, The 1st International Workshop on the Principles and Applications of Spatial Hearing (IWPASH 2009), November 11-13, 2009, Miyagi, Japan.
- [101] R. H. Y. So, B. Ngan, A. Horner, J. Braasch, J. Blauert, K. L. Leung, Toward orthogonal non-individualised head-related transfer functions for forward and backward directional sound: cluster analysis and an experimental study, Ergonomics, vol. 53, n. 6, June 2010, pp. 767-781.
- [102] R. H. Y. So, N. M. Leung, A. Horner, J. Braasch, K. L. Leung, Effects of Spectral Manipulation on Nonindividualized Head-Related Transfer Functions (HRTFs), The Journal of the Human Factors and Ergonomics Society, vol. 53, n. 3, June 2011, pp. 271-283.

- [103] A. Kudo, H. Higuchi, H. Hokari, S. Shimada, Improved method for accurate sound localization, *Acoustical Science and Technology*, vol. 27, n. 3, May 2006, pp. 134-146.
- [104] A. Ovcharenko, Front-Back Confusion Resolution in 3D Sound Localization with Databases Built with Dummy Head, M.Sc. thesis, Computer Engineering and Information Technology, University of Ulsan, Ulsan, South Korea, 2006.
- [105] K. Koo, H. Cha, Enhancement of a 3D Sound Using Psychoacoustics, *World Academy of Science, Engineering and Technology*, n. 37, 2008, pp. 162-166.
- [106] Gy. Wersényi, Effect of Emulated Head-Tracking for Reducing Localization Errors in Virtual Audio Simulation, *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 17, n. 2, February 2009, pp. 247-252.
- [107] F. L. Wightman, D. J. Kistler, Headphone simulation of free-field listening. I. Stimulus synthesis, *The Journal of the Acoustical Society of America*, vol. 85, n. 2, February 1989, pp. 858-867.
- [108] F. L. Wightman, D. J. Kistler, Headphone simulation of free-field listening. II. Psychophysical validation, *The Journal of the Acoustical Society of America*, vol. 85, n. 2, February 1989, pp. 868-878.
- [109] J. A. MacDonald, P. K. Tran, The Effect of Head-Related Transfer Function Measurement Methodology on Localization Performance in Spatial Audio Interfaces, *Human Factors*, vol. 50, n. 2, April 2008, pp. 256-263.
- [110] J. Sodnik, R. Susnik, M. Stular, S. Tomazic, Spatial sound resolution of an interpolated HRIR library, *Applied Acoustics*, vol. 66, n. 11, November 2005, pp. 1219-1234.
- [111] S. Xu, Z. Li, G. Salvendy, Individualization of Head-Related Transfer Function for Three-Dimensional Virtual Auditory Display: A Review, *Lecture Notes in Computer Science*, 2007, vol. 4563/2007, pp. 397-407.
- [112] S. Xu, Z. Li, G. Salvendy, Individualized head-related transfer functions based on population grouping (L), *The Journal of the Acoustical Society of America*, vol. 124, n. 5, November 2008, pp. 2708-2710.
- [113] W. L. Martens, Perceptual evaluation of filters controlling source direction: Customized and generalized HRTFs for binaural synthesis, *Acoustical Science and Technology*, vol. 24, n. 5, September 2003, pp. 220-232.
- [114] R. O. Duda, Modeling Head Related Transfer Functions, *The 27th Asilomar Conference on Signals, Systems and Computers (ACSSC 1993)*, November 1-3, 1993, Pacific Grove, California, USA.
- [115] C. P. Brown, R. O. Duda, An efficient HRTF model for 3-D sound, *IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (ASSP 1997)*, October 19-22, 1997, New Paltz, New York, USA.
- [116] R. O. Duda, W. L. Martens, Range-dependence of the HRTF for a spherical head, *IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (ASSP 1997)*, October 19-22, 1997, New Paltz, New York, USA.
- [117] C. P. Brown, R. O. Duda, A Structural Model for Binaural Sound Synthesis, *IEEE Transactions on Speech and Audio Processing*, vol. 6, n. 5, September 1998, pp. 476-488.
- [118] R. O. Duda, W. L. Martens, Range dependence of the response of a spherical head model, *The Journal of the Acoustical Society of America*, vol. 104, n. 5, November 1998, pp. 3048-3058.
- [119] R. O. Duda, C. Avendano, V. R. Algazi, An adaptable ellipsoidal head model for the interaural time difference, *1999 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP '99)*, March 15-19, 1999, Phoenix, Arizona, USA.

-
- [120] C. Avendano, V. R. Algazi, R. O. Duda, A head-and-torso model for low-frequency binaural elevation effects, 1999 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (ASPAA 1999), October 17-20, 1999, New Paltz, New York, USA.
- [121] V. R. Algazi, R. O. Duda, R. P. Morrison, D. M. Thompson, Structural Composition and Decomposition of HRTFs, IEEE Workshop on the Applications of Signal Processing to Audio and Acoustics, October 21-24, 2001, New Paltz, New York, USA.
- [122] V. R. Algazi, C. Avendano, R. O. Duda, Estimation of a Spherical-Head Model from Anthropometry, *The Journal of the Audio Engineering Society*, vol. 49, n. 6, June 2001, pp. 472-479.
- [123] V. R. Algazi, R. O. Duda, D. M. Thompson, The Use of Head-and-Torso Models for Improved Spatial Sound Synthesis, The 113th Audio Engineering Society Convention, October 5-8, 2002, Los Angeles, California, USA.
- [124] V. R. Algazi, R. O. Duda, R. Duraiswami, N. A. Gumerov, Z. Tang, Approximating the head-related transfer function using simple geometric models of the head and torso, *The Journal of the Acoustical Society of America*, vol. 112, n. 5, November 2002, pp. 2053-2064.
- [125] P. Satarzadeh, A Study of Physical and Circuit Models of the Human Pinnae, M. Sc. thesis, Department of Electrical and Computer Engineering, University of California Davis, Davis, California, USA, 2006.
- [126] P. Satarzadeh, V. R. Algazi, R. O. Duda, Physical and Filter Pinna Models Based on Anthropometry, The 122nd Audio Engineering Society Convention, May 5-8, 2007, Vienna, Austria.
- [127] J. Chen, B. D. Van Veen, K. E. Hecox, External ear transfer function modeling: A beamforming approach, *The Journal of the Acoustical Society of America*, vol. 92, n. 4, October 2002, pp. 1933-1944.
- [128] J. Chen, B. D. Van Veen, K. E. Hecox, A spatial feature extraction and regularization model for the head-related transfer function, *The Journal of the Acoustical Society of America*, vol. 97, n. 1, January 2002, pp. 439-452.
- [129] E. A. Lopez-Poveda, R. Meddis, A physical model of sound diffraction and reflections in the human concha, *The Journal of the Acoustical Society of America*, vol. 100, n. 5, November 1996, pp. 3248-3259.
- [130] C. I. Cheng, Visualization, measurement, and interpolation of Head-Related Transfer Function (HRTF's) with applications in electro-acoustic music, Ph.D. dissertation, Department of Electrical Engineering, University of Michigan, Ann Arbor, Michigan, 2001.
- [131] V. C. Raykar, R. Duraiswami, L. Davis, B. Yegnanarayana, Extracting significant features from the HRTF, The 2003 International Conference on Auditory Display (ICAD2003), July 6-9, 2003, Boston, Massachusetts, USA.
- [132] R. S. Juskiewicz, An improvement to anthropometry-based head and torso HRTF synthesis models for locations near the frontal median plane, M.Sc. thesis, University of Miami, Coral Gables, Florida, USA, 2007.
- [133] M. Geronazzo, S. Spagnol, F. Avanzini, Estimation and modeling of Pinna-Related Transfer Functions, The 13th International Conference on Digital Audio Effects (DAFx-10), September 6-10, 2010, Graz, Austria.
- [134] S. Spagnol, M. Geronazzo, F. Avanzini, Fitting pinna-related transfer functions to anthropometry for binaural sound rendering, IEEE International Workshop on Multimedia Signal Processing (MMSp'10), October 4-6, 2010, Saint-Malo, France.

- [135] S. Spagnol, M. Geronazzo, F. Avanzini, Structural Modeling Of Pinna-Related Transfer Functions, The 7th International Conference on Sound and Music Computing (SMC2010), July 21-24, 2010, Barcelona, Spain.
- [136] D. J. Kistler, F. 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, vol. 91, n. 3, March 1992, pp. 1637-1647.
- [137] J. C. Middlebrooks, D. M. Green, Observations on a principal components analysis of head-related transfer functions, The Journal of the Acoustical Society of America, vol. 92, n. 1, July 1992, pp. 597-599.
- [138] S. Hwang, Y. Park, Median HRIR Customization via Principal Components Analysis, Journal of Korean Society for Noise and Vibration Engineering, vol. 17, n. 7, July 2007, pp. 638-648.
- [139] S. Hwang, Y. Park, HRIR Customization in the Median Plane Via Principal Components Analysis, The Audio Engineering Society 31st International Conference, June 25-27, 2007, London, England.
- [140] S. Hwang, Y. Park, Y. Park, Comparison of head-related transfer function models based on principal components analysis, Journal of Korean Society for Noise and Vibration Engineering, vol. 18, n. 6, June 2007, pp. 642-653.
- [141] K. H. Shin, Y. Park, Enhanced Vertical Perception through Head-Related Impulse Response Customization Based on Pinna Response Tuning in the Median Plane, IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences, vol. E91-A, n. 1, January 2008, pp. 345-356.
- [142] S. Hwang, Y. Park, Interpretations on principal components analysis of head-related impulse responses in the median plane, The Journal of the Acoustical Society of America Express Letters, vol. 123, n. 4, April 2008, pp. 65-71.
- [143] S. Hwang, Y. Park, Y. Park, Modeling and Customization of Head-Related Transfer Functions using Principal Component Analysis, The 2008 International Conference on Control, Automation and Systems (ICCAS'08), October 14-17, 2008, Seoul, South Korea.
- [144] S. Hwang, Modeling, Customization, and Interpolation of Head-Related Impulse Responses based on Principal Components Analysis, Ph.D. dissertation, School of Mechanical, Aerospace and Systems Engineering Division of Mechanical Engineering, Korea Advanced Institute of Science and Technology, 2009.
- [145] Hugeng, W. Wahab, D. Gunawan, Effective Preprocessing in Modeling Head-Related Impulse Responses Based on Principal Components Analysis, Signal Processing: An International Journal (SPIJ), vol. 4, n. 4, October 2010, pp. 175-246.
- [146] Hugeng, W. Wahab, D. Gunawan, Enhanced Individualization of Head-Related Impulse Response Model in Horizontal Plane Based on Multiple Regression Analysis, 2010 Second International Conference on Computer Engineering and Applications (ICCEA 2010), March 19-21, 2010, Bali Island, Indonesia.
- [147] Hugeng, W. Wahab, D. Gunawan, Improved Method for Individualization of Head-Related Transfer Functions on Horizontal Plane Using Reduced Number of Anthropometric Measurements, Journal of Telecommunications, vol. 2, n. 2, May 2010, pp. 31-41.
- [148] S. Xu, Z. Z. Li, L. Zeng, G. Salvendy, A Study of Morphological Influence on Head-Related Transfer Functions, The 2007 IEEE International Conference on

- Industrial Engineering and Engineering Management (IEEM2007), December 2-4, 2007, Singapore.
- [149] S. Xu, Z. Z. Li, G. Salvendy, Improved method to individualize head-related transfer function using anthropometric measurements, *Acoustical Science and Technology*, vol. 29, n. 6, November 2008, pp. 388-390.
- [150] S. Xu, Z. Z. Li, G. Salvendy, Identification of Anthropometric Measurements for Individualization of Head-Related Transfer Functions, *Acta Acustica united with Acustica*, vol. 95, n. 1, 2009, pp. 168-177.
- [151] L. Wang, F. Yin, Z. Chen, HRTF compression via principal components analysis and vector quantization, *IEICE Electronics Express*, vol. 5, n. 9, 2008, pp. 321-325.
- [152] L. Wang, F. Yin, Z. Chen, Head-related transfer function interpolation through multivariate polynomial fitting of principal component weights, *Acoustical Science and Technology*, vol. 30, n. 6, November 2009, pp. 395-403.
- [153] J. Sodnik, S. Tomazic, Directional information in Head Related Transfer Functions, The 2004 IEEE Region 10 Conference (TENCON 2004), November 24, 2004, Chiang Mai, Thailand.
- [154] J. Sodnik, R. Susnik, S. Tomazic, Resolution enhancement of a general HRTF library, *Acoustics 2005*, November 9-11, 2005, Busselton, Western Australia.
- [155] J. Sodnik, R. Susnik, S. Tomazic, Principal Components of Non-Individualized Head Related Transfer Functions Significant for Azimuth Perception, *Acta Acustica united with Acustica*, vol. 92, n. 2, April 2006, pp. 312-319.
- [156] H. Hu, L. Zhou, J. Zhang, H. Ma, Z. Wu, Head Related Transfer Function Personalization Based on Multiple Regression Analysis, The 2006 International Conference on Computational Intelligence and Security (CIS 2006), November 3-6, 2006, Guangzhou, China.
- [157] H. Hu, L. Chen, Z. Wu, The Estimation of Personalized HRTFs in Individual VAS, The 4th International Conference on Natural Computation (ICNC'08), October 18-20, 2008, Jinan, Shandong, China.
- [158] H. Hu, L. Zhou, H. Ma, Z. Wu, HRTF personalization based on artificial neural network in individual virtual auditory space, *Applied Acoustics*, vol. 69, n. 2, February 2008, pp. 163-172.
- [159] T. Nishino, N. Inoue, K. Takeda, F. Itakura, Estimation of HRTFs on the horizontal plane using physical features, *Applied Acoustics*, vol. 68, n. 5, August 2007, pp. 897-908.
- [160] K. Matsui, A. Ando, Estimation of individualized head-related transfer function based on principal component analysis, *Acoustical Science and Technology*, vol. 30, n. 5, September 2009, pp. 338-347.
- [161] S. Spagnol, F. Avanzini, Real-time binaural audio rendering in the near field, The 6th Sound and Music Computing Conference (SMC 2009), July 23-25, 2009, Porto, Portugal.
- [162] D. Schönstein, B. F. G. Katz, HRTF selection for binaural synthesis from a database using morphological parameters, The 20th International Congress on Acoustics (ICA 2010), August 23-27, 2010, Sydney, Australia.
- [163] M. Zhang, R. A. Kennedy, T. D. Abhayapala, W. Zhang, Statistical method to identify key anthropometric parameters in HRTF individualization, 2011 Joint Workshop on Hands-free Speech Communication and Microphone Arrays (HSCMA), May 30-June 1, 2011, Edinburgh, United Kingdom.
- [164] S. Ise, M. Otani, Real time calculation of the Head Related Transfer Function based on the boundary element method, The 2002 International Conference on Auditory Display (ICAD2002), July 2-5, 2002, Kyoto, Japan.

- [165] M. Otani, S. Ise, A fast calculation method of the head-related transfer functions for multiple source points based on the boundary element method, *Acoustical Science and Technology*, vol. 24, n. 5, September 2003, pp. 259-266.
- [166] M. Otani, S. Ise, Fast calculation system specialized for head-related transfer function based on boundary element method, *The Journal of the Acoustical Society of America*, vol. 119, n. 5, May 2006, pp. 2589-2598.
- [167] H. Muraoka, Y. Takahashi, T. Haedar, M. Otani, T. Hirahara, The Head Related Transfer Function simulation by FEM/IEM and reciprocal theorem, *The 14th International Congress on Sound and Vibration (ICSV14)*, July 9-12, 2007, Cairns, Australia.
- [168] M. Otani, T. Hirahara, S. Ise, Numerical study on source-distance dependency of head-related transfer functions, *The Journal of the Acoustical Society of America*, vol. 125, n. 5, May 2009, pp. 3253-3261.
- [169] M. Otani, Y. Iwaya, Y. Suzuki, K. Itoh, Numerical analysis of HRTF spectral characteristics based on sound pressures on a pinna surface, *The 20th International Congress on Acoustics (ICA 2010)*, August 23-27, 2010, Sydney, Australia.
- [170] N. A. Gumerov, R. Duraiswami, Z. Tang, Numerical study of the influence of the torso on the HRTF, *2002 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP '02)*, May 13-17, 2002, Orlando, Florida, USA.
- [171] N. A. Gumerov, R. Duraiswami, A broadband fast multipole accelerated boundary element method for the three dimensional Helmholtz equation, *The Journal of the Acoustical Society of America*, vol. 125, n. 1, January 2009, pp. 191-205.
- [172] N. A. Gumerov, A. E. O'Donovan, R. Duraiswami, D. N. Zotkin, Computation of the head-related transfer function via the fast multipole accelerated boundary element method and its spherical harmonic representation, *The Journal of the Acoustical Society of America*, vol. 127, n. 1, January 2010, pp. 370-386.
- [173] Y. Kahana, Numerical modelling of the Head-Related Transfer Function, Ph.D. dissertation, Institute of Sound and Vibration Research, Faculty of Engineering and Applied Science, University of Southampton, Southampton, United Kingdom, 2000.
- [174] Y. Kahana, P. A. Nelson, Numerical modelling of the spatial acoustic response of the human pinna, *Journal of Sound and Vibration*, vol. 292, n. 1-2, April 2006, pp. 148-178.
- [175] Y. Kahana, P. A. Nelson, Boundary element simulations of the transfer function of human heads and baffled pinnae using accurate geometric models, *Journal of Sound and Vibration*, vol. 300, n. 3-5, March 2007, pp. 552-579.
- [176] B. 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*, vol. 110, n. 5, November 2001, pp. 2440-2448.
- [177] B. F. G. Katz, Boundary element method calculation of individual head-related transfer function. II. Impedance effects and comparisons to real measurements, *The Journal of the Acoustical Society of America*, vol. 110, n. 5, November 2001, pp. 2449-2455.
- [178] B. F. G. Katz, D. R. Begault, Round Robin Comparison of HRTF Measurement Results: Preliminary Results, *The 19th International Conference on Acoustics (ICA)*, September 2-7, 2007, Madrid, Spain.

- [179] R. Greff, B. F. G. Katz, Round Robin Comparison of HRTF Simulation Results: Preliminary Results, The 123rd Audio Engineering Society Convention, October 5-8, 2007, New York, New York, USA.
- [180] W. Kreuzer, P. Majdak, Z. Chen, Fast multipole boundary element method to calculate head-related transfer functions for a wide frequency range, The Journal of the Acoustical Society of America, vol. 126, n. 3, March 2009, pp. 1280-1290.
- [181] K. Terai, I. Kakuhari, HRTF calculation with less influence from 3-D modeling error: Making a physical human head model from geometric 3-D data, Acoustical Science and Technology, vol. 24, n. 5, September 2003, pp. 333-334.
- [182] P. Fiala, J. Huijssen, B. Pluymers, R. Hallez, W. Desmet, Fast Multipole BEM Modeling of Head Related Transfer Functions of a Dummy Head and Torso, The 2010 International Conference on Noise and Vibration Engineering (ISMA2010), September 20-22, 2010, Leuven, Belgium.
- [183] J. C. B. Torres, M. R. Petraglia, R. A. Tenenbaum, Room Acoustic Auralization: HRTF Modeling using Wavelet Transforms, The X International Symposium on Dynamic Problems of Mechanics (DINAME 2003), March 10-14, 2003, Ubatuba, Brazil.
- [184] J. C. B. Torres, M. R. Petraglia, R. A. Tenenbaum, An efficient wavelet-based HRTF model for auralization, Acta Acustica united with Acustica, vol. 90, n. 1, January-February 2004, pp. 108-120.
- [185] J. C. B. Torres, M. R. Petraglia, R. A. Tenenbaum, Low-order modeling and grouping of HRTFs for auralization using wavelet transforms, The IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP '04), May 17-21, 2004, Montreal, Quebec, Canada.
- [186] J. C. B. Torres, M. R. Petraglia, R. A. Tenenbaum, Low-order modeling of head-related transfer functions using wavelet transforms, The 2004 International Symposium on Circuits and Systems (ISCAS 2004), May 23-26, 2004, Vancouver, Canada.
- [187] J. C. B. Torres, M. R. Petraglia, R. A. Tenenbaum, Efficient structure for auralization using a wavelet-based HRTF model, The 4th IASTED International Conference on Circuits, Signals and System, November 20-22, 2006 San Francisco, California, USA.
- [188] J. C. B. Torres, R. A. Tenenbaum, M. R. Petraglia, Efficient Auralization by Grouping Directions and Modeling HRTFs Using Wavelets, The XII International Symposium on Dynamic Problems of Mechanics (DINAME 2007), February 26-March 2, 2007, Ilhabela, Brazil.
- [189] J. C. B. Torres, M. R. Petraglia, HRTF interpolation in the wavelet transform domain, 2009 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (ASPAA'09), October 18-21, 2009, New Paltz, New York, USA.
- [190] J. Zhang, A non-linear spatial hearing model based on bases pursuit algorithm, Frontiers of Electrical and Electronic Engineering in China, vol. 2, n. 4, 2007, pp. 388-394.
- [191] Z. Liang, B. Xie, A head-related transfer function model for fast synthesizing multiple virtual sound images, The 20th International Congress on Acoustics (ICA 2010), August 23-27, 2010, Sydney, Australia.
- [192] W. Zhang, T. D. Abhayapala, R. A. Kennedy, Horizontal plane HRTF reproduction using continuous Fourier-Bessel functions, The Audio Engineering Society 31st International Conference, June 25-27, 2007, London, England.

- [193] W. Zhang, R. A. Kennedy, T. D. Abhayapala, Efficient Continuous HRTF Model Using Data Independent Basis Functions: Experimentally Guided Approach, *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 17, n. 4, May 2009, pp. 819-829.
- [194] W. Zhang, T. D. Abhayapala, R. A. Kennedy, R. Duraiswami, Insights into head-related transfer function: Spatial dimensionality and continuous representation, *The Journal of the Acoustical Society of America*, vol. 127, n. 4, April 2010, pp. 2347-2357.
- [195] K. Zhang, R. A. Kennedy, T. D. Abhayapala, Efficiency evaluation and orthogonal basis determination in functional HRTF modeling, May 22-27, 2011, Prague, Czech Republic.
- [196] M. R. Bai, K. Y. Ou, Head-related transfer function (HRTF) synthesis based on a three-dimensional array model and singular value decomposition, *Journal of Sound and Vibration*, vol. 281, n. 3-5, March 2005, pp. 1093-1115.
- [197] M. Rothbucher, M. Durkovic, H. Shen, K. Diepold, HRTF customization using Multiway Array Analysis, *The 18th European Signal Processing Conference (EUSIPCO-2010)*, August 23-27, 2010, Aalborg, Denmark.
- [198] Y. Haneda, S. Makino, Y. Kaneda, N. Kitawaki, Common-Acoustical-Pole and Zero Modeling of Head-Related Transfer Functions, *IEEE Transactions on Speech and Audio Processing*, vol. 7, n. 2, March 1999, pp. 188-196.
- [199] E. A. Durant, G. H. Wakefield, Efficient Model Fitting Using a Genetic Algorithm: Pole-Zero Approximations of HRTFs, *IEEE Transactions on Speech and Audio Processing*, vol. 10, n. 1, January 1999, pp. 18-27.
- [200] J. Zhang, H. Ma, Z. Wu, Optimal approximation of head-related transfer function's pole-zero model based on genetic algorithm, *Frontiers of Electrical and Electronic Engineering in China*, vol. 1, n. 4, 2006, pp. 455-459.
- [201] Ki. Shin, Y. Park, Modeling of non-individualized head-related transfer functions for nearby sources, *The 9th Western Pacific Acoustics Conference (WESPAC IX 2006)*, June 26-28, 2006, Seoul, South Korea.
- [202] A. Barretto, N. Gupta, Dynamic Modeling of the Pinna for Audio Spatialization, *WSEAS Transactions on Acoustics and Music*, vol. 1, January 2004, pp. 77-82.
- [203] K. J. Faller II, A. Barretto, N. Gupta, N. Rishe, Decomposition and Modeling of Head-Related Impulse Responses for Customized Spatial Audio, *WSEAS Transactions on Signal Processing*, vol. 1, 2005, pp. 354-361.
- [204] K. J. Faller II, A. Barreto, N. Gupta, N. Rishe, Enhanced Modeling of Head-Related Impulse Responses Towards the Development of Customizable Sound Spatialization, *The 4th WSEAS International Conference on Computational Intelligence, Man-Machine Systems and Cybernetics (CIMMACS '05)*, November 17-19, 2005, Miami, Florida.
- [205] K. J. Faller II, A. Barreto, N. Gupta, N. Rishe, Accelerated Method for the Reduced-Parameter Modeling of Head-Related Transfer Functions for Customizable Spatial Audio, *The 5th WSEAS International Conference on Circuits, Systems, Electronics, Control and Signal Processing (CSECS '06)*, November 1-3, 2006, Dallas, Texas.
- [206] K. J. Faller II, A. Barreto, N. Gupta, N. Rishe, Time and Frequency Decomposition of Head-Related Impulse Responses for the Development of Customizable Spatial Audio Models, *WSEAS Transactions on Signal Processing*, vol. 2, 2006, pp. 1465-1472.
- [207] K. J. Faller II, A. Barreto, Sampling Rate Impact on the Performance of a Head-Related Impulse Response Decomposition Method, *The 8th WSEAS*

- International Conference on Acoustics & Music: Theory & Applications, June 19-21, 2007, Vancouver, Canada.
- [208] N. Gupta, A. Barretto, A Customizable Model of Head-Related Transfer Functions Based on Pinna Measurements, *International Journal of Scientific and Statistical Computing (IJSSC)*, vol. 1, n. 2, 2011, pp. 20-26.
- [209] N. H. Adams, Low-Order State-Space Models of Head-Related Transfer Function Arrays, Technical Report, University of Michigan, Ann Arbor, Michigan, 2007.
- [210] N. H. Adams, A Model of Head-Related Transfer Functions based on a State-Space Analysis, Ph.D. dissertation, Department of Electrical Engineering, University of Michigan, Ann Arbor, Michigan, USA, 2008.
- [211] N. H. Adams, G. H. Wakefield, State-Space Synthesis of Virtual Auditory Space, *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 16, n. 5, July 2008, pp. 881-890.
- [212] T. Xiao, Q. H. Liu, Finite difference computation of head-related transfer function for human hearing, *The Journal of the Acoustical Society of America*, vol. 113, n. 5, May 2003, pp. 2434-2441.
- [213] P. Mokhtari, H. Takemoto, R. Nishimura, H. Kato, Comparison of simulated and measured HRTFs: FDTD simulation using MRI head data, *The 123rd Audio Engineering Society Convention*, October 5-8, 2007, New York, New York, USA.
- [214] P. Mokhtari, H. Takemoto, R. Nishimura, Toward HRTF personalization: an auditory-perceptual evaluation of simulated and measured HRTFs, *The 14th International Conference on Auditory Display (ICAD2008)*, June 24-27, 2008, Paris, France.
- [215] P. Mokhtari, H. Takemoto, R. Nishimura, H. Kato, Computer Simulation of HRTFs for Personalization of 3D Audio, *The 2008 Second International Symposium on Universal Communication (ISUC'08)*, December 15-16, 2008, Osaka, Japan.
- [216] P. Mokhtari, H. Takemoto, R. Nishimura, H. Kato, Acoustic simulation of KEMAR's HRTFs: verification with measurements and the effects of modifying head shape and pinna concavity, *The 1st International Workshop on the Principles and Applications of Spatial Hearing (IWPASH 2009)*, November 11-13, 2009, Miyagi, Japan.
- [217] H. Takemoto, P. Mokhtari, H. Kato, R. Nishimura, K. Iida, Pressure distribution patterns on the pinna at spectral peak and notch frequencies of head-related transfer functions in the median plane, *The 1st International Workshop on the Principles and Applications of Spatial Hearing (IWPASH 2009)*, November 11-13, 2009, Miyagi, Japan.
- [218] P. Mokhtari, H. Takemoto, R. Nishimura, H. Kato, Acoustic sensitivity to micro-perturbations of KEMAR's pinna surface geometry, *The 20th International Congress on Acoustics (ICA 2010)*, August 23-27, 2010, Sydney, Australia.
- [219] P. Mokhtari, H. Takemoto, R. Nishimura, H. Kato, Pinna sensitivity patterns reveal reflecting and diffracting surfaces that generate the first spectral notch in the front median plane, *The 36th IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP2011)*, May 22-27, 2011, Prague, Czech Republic.
- [220] Z. Wu, T. Weng, W. Wang, T. F. Lo, H. Y. Chan, F. K. Lam, Neural network model of binaural hearing based on spatial feature extraction of the head related transfer function, *The 20th IEEE Engineering in Medicine and Biology*

- Society Conference (EMBS 1998), October 29-November 1, 1998, Hong Kong, China.
- [221] V. Lemaire, F. Clerot, S. Busson, R. Nicol, V. Choqueuse, Individualized HRTFs From Few Measurements: a Statistical Learning Approach, The 2005 IEEE International Joint Conference on Neural Networks (IJCNN '05), July 31-August 4, 2005, Montreal, Quebec, Canada.
- [222] D. F. M. Goodman, R. Brette, Learning to localise sounds with spiking neural networks, Neural Information Processing Systems Conference 2010 (NIPS 2010), December 6-9, 2010, Vancouver, Canada.
- [223] N.-M. Cheung, S. Trautmann, A. Honer, Head-related transfer function modeling in 3-D sound systems with genetic algorithms, The 1998 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 1998), May 12-15, 1998, Seattle, Washington, USA.
- [224] Y. Iwaya, Individualization of head-related transfer functions with tournament-style listening test: Listening with other's ears, *Acoustical Science and Technology*, vol. 27, n. 6, November 2006, pp. 340-343.
- [225] G. Eckel, Immersive Audio-Augmented Environments - The LISTEN Project, The 5th International Conference on Information Visualization (IV2001), July 25-27, 2001, Los Alamitos, California, USA.
- [226] O. Warusfel, LISTEN HRTF Database, 2003, <http://recherche.ircam.fr/equipements/salles/listen/>.
- [227] V. R. Algazi, R. O. Duda, D. M. Thompson, C. Avendano, The CIPIC HRTF database, IEEE Workshop on the Applications of Signal Processing to Audio and Acoustics (WASPAA'01), October 21-24, 2001, Mohonk Mountain House, New Platz, New York, USA.
- [228] W. G. Gardner, K. D. Martin, HRTF measurements of a KEMAR, *Journal of the Acoustical Society of America*, vol. 97, n. 6, June 1995, pp. 3907-3908.
- [229] V. Larcher, J.-M. Jot, G. Vandernoot, Equalization methods in binaural technology, The 105th Audio Engineering Society Convention, September 26-29, 1998, San Francisco, California, USA.
- [230] S. Takane, D. Arai, T. Miyajima, K. Watanabe, Y. Suzuki, T. Sone, A database of Head-Related Transfer Functions in whole directions on upper hemisphere, *Acoustical Science and Technology*, vol. 23, n. 3, May 2002, pp. 160-162.
- [231] E. Grassi, J. Tulsi, S. Shamma, Measurement of head-related transfer functions based on the empirical transfer function estimate, The 2003 International Conference on Auditory Display (ICAD2003), July 6-9, 2003, Boston, Massachusetts, USA.
- [232] B. Xie, X. Zhong, D. Rao, Z. Liang, Head-related transfer function database and its analyses, *Science in China Series G: Physics, Mechanics & Astronomy*, vol. 50, n. 3, June 2007, pp.267-280.
- [233] S. Hosoe, T. Nishino, K. Itou, K. Takeda, Development of micro-dodecahedral loudspeaker for measuring head related transfer functions in the proximal region, IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP2006), May 14-19, 2006, Toulouse, France.
- [234] T. Nakado, T. Nishino, K. Takeda, Head-Related Transfer Function measurement in sagittal and frontal coordinates, *Acoustical Science and Technology*, vol. 29, n. 5, September 2008, pp. 335-337.
- [235] H. Kayser, S. D. Ewert, J. Anemuller, T. Rohdenburg, V. Hohmann, B. Kollmeier, Database of Multichannel In-Ear and Behind-the-Ear Head-Related and Binaural Room Impulse Responses, *EURASIP Journal on Advances in*

- Signal Processing, Special Issue on Digital Signal Processing for Hearing Instruments, vol. 2009, Article ID 298605.
- [236] N. Gupta, A. Barreto, M. Joshi, J. C. Agudelo, HRTF Database at FIU DSP Lab, IEEE International Conference on Acoustics Speech and Signal Processing (ICASSP2010), March 14-19, 2010, Dallas, Texas, USA.
- [237] H. Wierstorf, M. Geier, A. Raake, S. Spors, A Free Database of Head-Related Impulse Response Measurements in the Horizontal Plane with Multiple Distances, The 130th Audio Engineering Society Convention, May 13-16, 2011, London, United Kingdom.
- [238] M. Maazaoui, Y. Grenier, Romeo-HRTF: A Multimicrophone Head Related Transfer Functions Database, March 2011.
- [239] Acoustics Research Institute, ARI HRTF Database, Austrian Academy of Sciences, <http://www.kfs.oeaw.ac.at/content/view/608/606/>, July 2011.
- [240] Z. **Haraszy**, D. G. Cristea, S. Micut, I. Bogdanov, Efficient Algorithm for Extracting Essential Head Related Impulse Response Data for Acoustic Virtual Reality Development, The 15th WSEAS International Conference on Systems, July 14-16, 2011, Corfu Island, Greece.
- [241] Z. **Haraszy**, D. G. Cristea, V. Tiponut, T. Slavici, Improved Head Related Transfer Function Generation and Testing for Acoustic Virtual Reality Development, The 14th WSEAS International Conference on Systems, July 22-24, 2010, Corfu Island, Greece.
- [242] V. Tiponut, D. Ianchis, M. Basch, Z. **Haraszy**. Work Directions and New Results in Electronic Travel Aids for Blind and Visually Impaired People, The 14th WSEAS International Conference on Systems, July 22-24, 2010, Corfu Island, Greece.
- [243] V. Tiponut, D. Ianchis, Z. **Haraszy**, M. Basch, I. Bogdanov, Work Directions and New Results in Electronic Travel Aids for Blind and Visually Impaired People, Journal of WSEAS Transactions on Systems, vol. 9, n. 10, October 2010, pp. 1086-1097.
- [244] F. Keyrouz, Y. Naous, K. Diepold, A New Method for Binaural 3-D Localization Based on HRTFs, The 2006 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2006), May 14-19, 2006, Toulouse, France.
- [245] K. Tsujino, W. Kobayashi, T. Onoye, Y. Nakamura, Automated Design of Digital Filters for 3-D Sound Localization in Embedded Applications, The 2006 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2006), May 14-19, 2006, Toulouse, France.
- [246] R. Nicol, V. Lemaire, A. Bondu, S. Busson, Looking for a relevant similarity criterion for HRTF clustering: a comparative study, The 120th Audio Engineering Society Convention, May 20-23, 2006, Paris, France.
- [247] L. Wang, F. Yin, Z. Chen, A hybrid compression method for head-related transfer functions, Applied Acoustics, vol. 70, n. 9, September 2009, pp. 1212-1218.
- [248] C. Masterson, S. Adams, G. Kearney, F. Boland, A method for Head Related Impulse Response simplification, The 17th European Signal Processing Conference (EUSIPCO 2009), August 24-28, 2009, Glasgow, Scotland.
- [249] M. Rothbucher, H. Shen, K. Diepold, Dimensionality Reduction in HRTF by using Multiway Array Analysis, International Workshop on Human-Centered Robotic Systems (HCRS 2009), November 19-20, 2009, Bielefeld, Germany.

- [250] C. Masterson, G. Kearney, F. Boland, HRIR factorisation: A regularized approach, The 18th European Signal Processing Conference (EUSIPCO 2010), August 23-27, 2010, Aalborg, Denmark.
- [251] T. Papadopoulos, Inverse Filtering for Virtual Acoustic Imaging Systems, Ph.D. dissertation, Institute of Sound and Vibration Research, Faculty of Engineering, Science and Mathematics, University of Southampton, Southampton, United Kingdom, 2006.
- [252] J. S. Abel, S. H. Foster, Measuring HRTFs in a Reflective Environment, 2nd International Conference on Auditory Displays (ICAD'94), November 7-9, 1994, Santa Fe, New Mexico, USA.
- [253] S. Takane, K. Abe, S. Sato, Acquisition of individual HRTFs from measured transfer functions in ordinary sound field, The Japan-China Joint Conference of Acoustics 2007 (JCA2007), June 4-7, 2007, Sendai, Japan.
- [254] S. Takane, M. Nabatame, K. Abe, K. Watanabe, S. Sato, Subjective evaluation of HRIRs linearly predicted from impulse responses measured in ordinary sound field, AES Japan Section Conference in Osaka, July 25-26, 2008, Fukushima, Osaka, Japan.
- [255] S. Takane, Some further investigations on estimation of HRIRs from impulse responses acquired in ordinary sound field, The 127th Audio Engineering Convention, October 9-12, 2009, New York, USA.
- [256] S. Takane, Estimation of HRIRs from impulse responses measured in ordinary sound field, International Workshop on the Principles and Applications of Spatial Hearing 2009 (IWPASH 2009), November 11-13, 2009, Zao, Miyagi, Japan.
- [257] Q. Ye, Q. Dong, Y. Zhang, X. Li, Fast Head-Related Transfer Function Measurement in Complex Environments, 20th International Congress on Acoustics (ICA 2010), August 23-27, 2010, Sydney, Australia.
- [258] M. Durkovic, F. Sagstetter, K. Diepold, HRTF Measurements With Recorded Reference Signal, The 129th Audio Engineering Society Convention, November 4-7, 2010, San Francisco, California, USA.
- [259] S. Spors, H. Wierstorf, J. Ahrens, Interpolation and Range Extrapolation of Head-Related Transfer Functions using Virtual Local Wave Field Synthesis, The 130th Audio Engineering Convention, May 13-16, 2011, London, United Kingdom.
- [260] Gy. Wersényi, Measurement system upgrading for more precise measuring of the Head-Related Transfer Functions, The 29th International Congress and Exhibition on Noise Control Engineering (InterNoise 2000), August 27-30, 2000, Nice, France.
- [261] S. Fontana, A. Farina, Y. Grenier, A System for Rapid Measurement and Direct Customization of Head Related Impulse Responses, The 120th Audio Engineering Society Convention, May 20-23, 2006, Paris, France.
- [262] D. R. Begault, M. Godfroy, J. D. Miller, A. Roginska, M. R. Anderson, E. M. Wenzel, Design and verification of HeadZap, a semi-automated HRIR measurement system, The 120th Audio Engineering Society Convention, May 20-23, 2006, Paris, France.
- [263] P. Plaskota, P. Pruchnicki, Practical aspects of using HRTF measuring device, Archives of Acoustics, vol. 31, n. 4, 2006, pp. 439-444.
- [264] K. Fukudomea, T. Suetsugub, T. Ueshinb, R. Idegamib, K. Takeya, The fast measurement of head related impulse responses for all azimuthal directions using the continuous measurement method with a servo-swiveled chair, Applied Acoustics, vol. 68, n. 8, August 2007, pp. 864-884.

- [265] Gy. Wersényi, Localization in a HRTF-based Minimum Audible Angle Listening Test on a 2D Sound Screen for GUIB Applications, The 115th Audio Engineering Society Convention, October 10-13, 2003, New York, New York, USA.
- [266] M. Pec, M. Bujacz, P. Strumillo, Personalized head related transfer function measurement and verification through sound localization resolution, The 15th European Signal Processing Conference (EUSIPCO 2007), September 3-7, 2007, Poznan, Poland.
- [267] S. Sima, HRTF Measurements and Filter Design for a Headphone-Based 3D-Audio System, Department of Computer Science, Faculty of Engineering and Computer Science, Hamburg University of Applied Sciences, Hamburg, Germany, 2008.
- [268] F. Volk, M. Straubinger, L. Roalter, H. Fastl, Measurement of Head Related Impulse Responses for Psychoacoustic Research, The 35th German Annual Conference on Acoustics (NAG/DAGA 2009), March 23-26, 2009, Rotterdam, Netherlands.
- [269] Y. Lee, Y. Park, Y. Park, Newly Designed HRTF Measuring, International Conference on Control, Automation and Systems 2009. August 18-21, 2009, Fukuoka International Congress Center, Japan.
- [270] D. S. Brungart, G. Romigh, B. D. Simpson, Rapid Collection of HRTFs and Comparison to Free-Field Listening, The 1st International Workshop on the Principles and Applications of Spatial Hearing 2009 (IWPASH 2009), November 11-13, 2009, Zao, Miyagi, Japan.
- [271] M. Zhang, W. Zhang, R. A. Kennedy, T. D. Abhayapala, HRTF measurement on KEMAR manikin, Acoustics 2009, November 23-25, 2009, Adelaide, Australia.
- [272] A. Dobrucki, P. Plaskota, P. Prucknicki, M. Pec, M. Bujacz, P. Strumillo, Measurement System for Personalized Head-Related Transfer Functions and Its Verification by Virtual Source Localization Trials with Visually Impaired and Sighted Individuals, The Journal of the Audio Engineering Society, vol. 58, n. 9, September 2010, pp. 724-738.
- [273] D. N. Zotkin, R. Duraiswami, E. Grassi, N. A. Gumerov. Fast head-related transfer function measurement via reciprocity, The Journal of the Acoustical Society of America, vol. 120, n. 4, October 2006, pp. 2202-2215.
- [274] R. Duraiswami, D. N. Zotkin, A. E. O'Donovan, Capturing and Recreating Auditory Virtual Reality, The 1st International Workshop on the Principles and Applications of Spatial Hearing 2009 (IWPASH 2009), November 11-13, 2009, Zao, Miyagi, Japan.
- [275] N. Matsunaga, T. Hirahara, Re-examination of HRTF measurement system via reciprocity, The 1st International Workshop on the Principles and Applications of Spatial Hearing 2009 (IWPASH 2009), November 11-13, 2009, Zao, Miyagi, Japan.
- [276] N. Matsunaga, T. Hirahara, Fast near-field HRTF measurements using reciprocal method, The 20th International Congress on Acoustics (ICA 2010), August 23-27, 2010, Sydney, Australia.
- [277] N. Matsunaga, T. Hirahara, Reexamination of fast head-related transfer function measurement by reciprocal method, Acoustical Science and Technology, vol. 31, n. 6, November 2010, pp. 414-416.
- [278] K. A. J. Riederer, Part IIIa: Effect of microphone position changes on blocked cavum conchae head-related transfer functions, The 18th International Congress on Acoustics (ICA 2004), April 4-9, 2004, Kyoto, Japan.

- [279] K. A. J. Riederer, Repeatability analysis of head-related transfer function measurements, The 105th Audio Engineering Society Convention, September 26-29, 1998, San Francisco, California, USA.
- [280] K. A. J. Riederer, Computational quality assessment of HRTFs, The 10th European Signal Processing Conference (EUSIPCO2000), September 5-8, 2000, Tampere, Finland.
- [281] J. D. Miller, HRTF Error Analysis using Spectral Power Ratios, Technical Report, Stanford University, 2001.
- [282] Andrea Electronics Corporation, Pureaudio USB-SA Hi-Fidelity External USB Sound Card Spec Sheet, 2011, http://www.andreaelectronics.com/pdf_files/ProductSpecs/USB-SASpec.pdf.
- [283] F. L. Wightman, D. J. Kistler, S. H. Foster, J. Abel, A Comparison of Head-Related Transfer Functions Measured Deep in the Ear Canal and at the Ear Canal Entrance, The 18th Midwinter Meeting of the Association for Research in Otolaryngology, 1995, St. Petersburg Beach, Florida, USA.
- [284] H. Moller, M. F. Sorensen, D. Hammershoi, C. B. Jensen, Head-Related Transfer Functions of Human Subjects, The Journal of the Audio Engineering Society, vol. 43, n. 5, May 1995, pp. 300-321.
- [285] T. Ajdler, C. Faller, L. Sbaiz, M. Vetterli, Interpolation of the Head Related Transfer Functions Considering Acoustics, the 118th Audio Engineering Society Convention, May 28-31, 2005, Barcelona, Spain.
- [286] J. Breebaart, F. Nater, A. Kohlrausch, Spectral and Spatial Parameter Resolution Requirements for Parametric, Filter-Bank-Based HRTF Processing, The Journal of the Audio Engineering Society, vol. 58, n. 3, March 2010, pp. 126-140.
- [287] B. F. G. Katz, Acoustic absorption measurement of human hair and skin within the audible frequency range, The Journal of the Acoustical Society of America, vol. 108, n. 5, November 2000, pp. 2238-2242.
- [288] B. E. Treeby, The Effect of Hair on Human Sound Localization Cues, Ph.D dissertation, Centre for Acoustics, Dynamics and Vibration, School of Mechanical Engineering, The University of Western Australia, Perth, Western Australia, Australia, 2007.
- [289] M. Itokazu, K. Watanabe, K. Ozawa, Individual equalization in binaural reproduction of a reverberant sound field, Acoustical Science and Technology, vol. 30, n. 1, January 2009, pp. 25-32.
- [290] S. Müller, P. Massarani, Transfer-Function Measurement with Sweeps, The Journal of the Audio Engineering Society, vol. 49, n. 6, June 2001, pp. 443-471.
- [291] G.-B. Stan, J.-J. Embrechts, D. Archambeau, Comparison of Different Impulse Response Measurement Techniques, The Journal of the Audio Engineering Society, vol. 50, n. 4, April 2002, pp. 249-262.
- [292] A. Farina, Advancements in impulse response measurements by sine sweeps, The 122nd Audio Engineering Society Convention, May 5-8, 2007, Vienna, Austria.
- [293] P. Majdak, P. Balazs, B. Laback, Multiple Exponential Sweep Method for Fast Measurement of Head-Related Transfer Functions, The Journal of the Audio Engineering Society, vol. 55, n. 7-8, July-August 2007, pp. 623-637.
- [294] A. Farina, Silence Sweep: a novel method for measuring electro-acoustical devices, The 106th Audio Engineering Society Convention, May 7-10, 2009, Munich, Germany.

- [295] Cs. Huszty, S. Sakamoto, Time-domain sweeplets for acoustic measurements, *Applied Acoustics*, vol. 71, n. 10, October 2010, pp. 979-989.
- [296] F. L. Wightman, D. J. Kistler, The dominant role of low-frequency interaural time differences in sound localization, *The Journal of the Acoustical Society of America*, vol. 91, n. 3, March 1992, pp. 1648-1661.
- [297] E. J. Berdahl, J. O. Smith, Transfer function measurement toolbox, June 2007, http://ccrma.stanford.edu/realsimple/imp_meas/.
- [298] P. Zahorik, Limitations in using Golay codes for head-related transfer function measurement, *The Journal of the Acoustical Society of America*, vol. 107, n. 3, March 2000, pp. 1793-1796.
- [299] S. van de Par, J. Breebaart, M. McKinney, An automatic equalization algorithm for audio, *The 8th International Conference on Music Information Retrieval (ISMIR 2007)*, September 23-27, 2007, Vienna, Austria.
- [300] C. I. Cheng, G. H. Wakefield, Introduction to Head-Related Transfer Functions (HRTF's): Representations of HRTF's in Time, Frequency, and Space, *The Journal of the Audio Engineering Society*, vol. 49, n. 4, April 2001, pp. 231-249.
- [301] H. Viste, Binaural Localization and Separation Techniques, Ph.D dissertation, Ecole Polytechnique Fédérale de Lausanne, Lausanne, Switzerland, 2004.
- [302] J. Mouba, S. Marchand, A Source Localization-Separation-Respatialization System Based on Unsupervised Classification of Interaural Cues, *The 9th International Conference on Digital Audio Effects (DAFx06)*, September 18-20, 2006, Montreal, Canada.
- [303] X. L. Zhong, B. Xie, A Novel Method of Interaural Time Difference Based on Spatial Fourier Analysis, *Chinese Physics Letters*, vol. 24, n. 5, May 2007, pp. 1313-1316.
- [304] J. Nam, J. S. Abel, J. O. Smith III, A Method for Estimating Interaural Time Difference for Binaural Synthesis, *The 125th Audio Engineering Society Convention*, October 2-5, 2008, San Francisco, California, USA.
- [305] P. Parhizkari, Binaural Hearing-Human Ability of Sound Source Localization, M.Sc. thesis, Department of Signal Processing, School of Engineering, Blekinge Institute of Technology, Sweden, 2008.
- [306] A. Lindau, J. Estrella, S. Weinzierl, Individualization of dynamic binaural synthesis by real time manipulation of the ITD, *The 128th Audio Engineering Society Convention*, May 22-25, 2010, London, United Kingdom.
- [307] A. Farina, Simultaneous Measurement of Impulse Response and Distortion with a Swept-Sine Technique, *The 108th Audio Engineering Society Convention*, February 19-22, 2000, Paris, France.
- [308] K. A. J. Riederer, Part Va: Effect of Head Movements on Measured Head-Related Transfer Functions, *18th International Congress on Acoustics (ICA 2004)*, April 4-9, 2004, Kyoto, Japan.
- [309] T. Hirahara, D. Morikawa, M. Otani, Effect of Head Movement in HRTF Measurements, *The 1st International Workshop on the Principles and Applications of Spatial Hearing 2009 (IWPASH 2009)*, November 11-13, 2009, Zao, Miyagi, Japan.
- [310] T. Hirahara, H. Sagara, I. Toshima, M. Otani, Head movement during head-related transfer function measurement, *Acoustical Science and Technology*, vol. 31, n. 2, March 2010, pp. 165-171.
- [311] D. Hammershoi, H. Moller, Sound transmission to and within the human ear canal, *The Journal of the Acoustical Society of America*, vol. 100, n. 1, July 1996, pp. 408-427.

- [312] V. R. Algazi, C. Avendano, D. Thompson, Dependence of Subject and Measurement Position in Binaural Signal Acquisition, *The Journal of the Audio Engineering Society*, vol. 47, n. 11, November 1999, pp. 937-947.
- [313] F. Wickelmaier, S. Choisel, Selecting participants for listening tests of multichannel reproduced sound, *The 118th Audio Engineering Society Convention*, May 28-31, 2005, Barcelona, Spain.
- [314] R. Martin, K. McAnally, Interpolation of Head-Related Transfer Functions, Technical Report, DSTO-RR-0323, Air Operations Division, Defence Science and Technology Organisation, Australia, 2007.
- [315] L. W. P. Biscainho, F. P. Freeland, P. S. R. Diniz, Using inter-positional transfer functions in 3D-sound, *The 2002 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP2002)*, May 13-17, 2002, Orlando, Florida, USA.
- [316] F. P. Freeland, L. W. P. Biscainho, P. S. R. Diniz, Efficient HRTF Interpolation in 3D Moving Sound, *The AES 22nd International Conference on Virtual, Synthetic and Entertainment Audio*, June 15-17, 2002, Espoo, Finland.
- [317] F. P. Freeland, L. W. P. Biscainho, P. S. R. Diniz, Interpositional Transfer Function for 3D-Sound Generation, *The Journal of the Audio Engineering Society*, vol. 52, n. 9, September 2004, pp. 915-930.
- [318] F. P. Freeland, L. W. P. Biscainho, P. S. R. Diniz, Interpolation of Head-Related Transfer Functions (HRTFs): A Multi-Source Approach, *The 12th European Signal Processing Conference (EUSIPCO2004)*, September 6-10, 2004, Vienna, Austria.
- [319] F. P. Freeland, L. W. P. Biscainho, P. S. R. Diniz, HRTF Interpolation Through Direct Angular Parameterization, *The 2007 IEEE International Symposium on Circuits and Systems (ISCAS 2007)*, May 27-30, 2007, New Orleans, Los Angeles, USA.
- [320] T. Ajdler, L. Sbaiz, M. Vetterli, Plenacoustic function on the circle with application to HRTF interpolation, *The 2005 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP2005)*, March 19-23, 2005, Philadelphia, Pennsylvania, USA.
- [321] T. Ajdler, C. Faller, L. Sbaiz, M. Vetterli, Sound Field Analysis along a Circle and Its Applications to HRTF Interpolation, *The Journal of the Audio Engineering Society*, vol. 56, n. 3, March 2008, pp. 156-175.
- [322] K. Watanabe, S. Takane, Y. Suzuki, Interpolation of Head-Related Transfer Functions based on the Common-Acoustical-Pole and Residue model, *Acoustical Science and Technology*, vol. 24, n. 5, September 2005, pp. 335-337.
- [323] K. Watanabe, S. Takane, Y. Suzuki, A Novel Interpolation Method of HRTFs Based on the Common-Acoustical-Pole and Zero Model, *Acta Acustica United with Acustica*, vol. 91, n. 6, November-December 2005, pp. 958-966.
- [324] R. Duraiswami, D. N. Zotkin, N. A. Gumerov, Interpolation and range extrapolation of HRTFs, *The IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP '04)*, May 17-21, 2004, Montreal, Quebec, Canada.
- [325] R. Duraiswami, V. C. Raykar, The manifolds of spatial hearing, *The 2005 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP2005)*, March 19-23, 2005, Philadelphia, Pennsylvania, USA.
- [326] J. W. Scarpaci, H. S. Colburn, Principal Components Analysis Interpolation of HRTF's Using Locally Chosen Basis Functions, *The Journal of the Acoustical Society of America*, vol. 117, n. 4, pt. 2, April 2005, pp. 2561-2562.

- [327] B. Xie, Spatial interpolation of HRTFs and signal mixing for multichannel surround sound, *Chinese Journal of Acoustics*, vol. 25, n. 4, 2006, pp. 330-341.
- [328] S. Hwang, Y. Park, Time delay estimation from HRTFs and HRIRs, *The 8th International Conference on Motion and Vibration Control (MOVIC 2006)*, August 27-30, 2006, Daejeon, South Korea.
- [329] F. Keyrouz, K. Diepold, A novel biologically inspired neural network solution for robotic 3D sound source sensing, *Soft Computing*, vol.12, n. 7, May 2008, pp. 721-729.
- [330] F. Keyrouz, K. Diepold, A New HRTF Interpolation Approach for Fast Synthesis of Dynamic Environmental Interaction, *The Journal of the Audio Engineering Society*, vol. 56, n. 1-2, January 2008, pp. 28-35.
- [331] L. Wang, F. Yin, Z. Chen, An indirect interpolation method for Head-Related Transfer Function pole-zero models, *Acoustical Science and Technology*, vol. 29, n. 5, September 2008, pp. 329-331.
- [332] L. Wang, F. Yin, Z. Chen, Head-related transfer function interpolation through multivariate polynomial fitting of principal component weights, *Acoustical Science and Technology*, vol. 30, n. 6, November 2008, pp. 395-403.
- [333] L. Chen, H. Hu, Z. Wu, Head-Related Impulse Response Interpolation in Virtual Sound System, *The 4th International Conference on Natural Computation (ICNC '08)*, October 18-20, 2008, Jinan, Shandong, China.
- [334] J. C. B. Torres, M. R. Petraglia, HRTF Interpolation in the Wavelet Transform Domain, *2009 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (ASPAA'09)*, October 18-21, 2009, New Paltz, New York, USA.
- [335] R. Nishimura, H. Kato, N. Inoue, Interpolation of head-related transfer functions by spatial linear prediction, *The 2009 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2009)*, April 19-24, 2009, Taipei, Taiwan.
- [336] H. Qinghua, F. Yong, Interpolation of head-related transfer functions using spherical fourier expansion, *Journal of Electronics (China)*, vol. 26, n. 4, July 2009, pp. 571-576.
- [337] G. H. M. de Sousa, M. Queiroz, Two approaches for HRTF interpolation, *The 12th Brazilian Symposium on Computer Music (SBCM 2009)*, September 7-9, 2009, Recife, Brazil.
- [338] Y. Iwaya, W. Sato, T. Okamoto, M. Otani, Y. Suzuki, Interpolation method of head-related transfer functions in the z-plane domain using a common-pole and zero model, *The 20th International Congress on Acoustics (ICA 2010)*, August 23-27, 2010, Sydney, Australia.
- [339] V. Tiponut, C.-D. Căleanu, *Rețele neuronale. Arhitecturi și algoritmi* (Editura Politehnica, 2002).
- [340] Z. **Haraszy**, D. Ianchis, V. Tiponut, Generation of the Head Related Transfer Functions Using Artificial Neural Networks, *The 13th WSEAS International Conference on CIRCUITS*, July 22-24, 2009, Rodos Island, Greece.
- [341] Z. **Haraszy**, S. Micut, V. Tiponut, T. Slavici, Multi-Subject Head Related Transfer Function Generation using Artificial Neural Networks, *The 14th WSEAS International Conference on SYSTEMS*, July 22-24, 2010, Corfu Island, Greece.
- [342] Z. **Haraszy**, D. Ianchis, V. Tiponut, Acoustic Virtual Reality Generation with Head Related Transfer Function Visualization, *Sesiunea de comunicări științifice "Doctor ETC 2009"*, September 24-25, 2009, Timisoara, Romania.

- [343] X.-Y. Zeng, S.-G. Wang, L.-P. Gao, A hybrid algorithm for selecting head-related transfer function based on similarity of anthropometric structures, *Journal of Sound and Vibration*, vol. 329, n. 19, September 2010, pp. 4093-4106.
- [344] K.-S. Lee, S.-P. Lee, A Relevant Distance Criterion for Interpolation of Head-Related Transfer Functions, *IEEE Transactions on Audio, Speech, and Singal Processing*, vol. 19, n. 6, August 2011, pp. 1780-1790.
- [345] J.-M. Jot, V. Larcher, O. Warusfel, Digital signal processing issues in the context of binaural and transaural stereophony, *The 98th Audio Engineering Society Convention*, February 25-28, 1995, Paris, France.
- [346] D. Pralong, S. Carlile, Generation and Validation of Virtual Auditory Space, In S. Carlile (Ed.), *Virtual Auditory Space: Generation and Applications*, 4 (Austin: R. G. Landes Company, 1996, 109-151).
- [347] E. H. A. Langendijk, A. W. Bronkhorst, Fidelity of three-dimensional-sound reproduction using a virtual auditory display, *The Journal of the Acoustical Society of America*, vol. 107, n. 1, January 2000, pp. 528-537.
- [348] Z. Schärer, A. Lindau, Evaluation of Equalization Methods for Binaural Signals, *The 126th Audio Engineering Society Convention*, May 7-10, 2009, Munich, Germany.
- [349] R. Nishimura, P. Mokhtari, H. Takemoto, H. Kato, Headphone Calibration for 3D-Audio Listening, *The 3rd International Universal Communication Symposium (IUCS'09)*, December 3-4, 2009, Tokyo, Japan.
- [350] F. Brinkmann, A. Lindau, On the effect of individual headphone compensation in binaural synthesis, *36. Deutsche Jahrestagung für Akustik (DAGA 2010)*, March 15-18, 2010, Berlin, Germany.
- [351] R. Nishimura, H. Kato, P. Mokhtari, H. Takemoto, Effects of headphone calibration functions on subjective impressions of spatial sound, *The 20th International Congress on Acoustics (ICA 2010)*, August 23-27, 2010, Sydney, Australia.
- [352] A. Lindau, F. Brinkmann, Perceptual evaluation of individual headphone compensation in binaural synthesis based on non-individual recordings, *The 3rd ISCA/DEGA Tutorial and Research Workshop on Perceptual Quality of Systems*, September 6-8, 2010, Bautzen, Germany.
- [353] D. Pralong, S. Carlile, The role of individualized headphone calibration for the generation of high fidelity virtual auditory space, *The Journal of the Acoustical Society of America*, vol. 100, n. 6, December 1996, pp. 3785-3793.
- [354] D. Schonstein, L. Ferre, B. F. G. Katz, Comparison of headphones and equalization for virtual auditory source localization, *Acoustics'08 Paris*, June 29-July 4, 2008, Paris, France.
- [355] H. Moller, C. Jensen, D. Hammershoi, M. Sorensen, Transfer Characteristics of Headphones Measured on Human Ears, *The Journal of the Audio Engineering Society*, vol. 43, n. 4, April 1995, pp. 203-217.
- [356] W. L. Martens, Individualized and generalized earphone correction filters for spatial sound reproduction, *The 2003 International Conference on Auditory Display (ICAD2003)*, July 6-9, 2003, Boston, Massachusetts, USA.
- [357] A. Kulkarni, H. S. Colburn, Variability in the characterization of the headphone transfer-function, *The Journal of the Acoustical Society of America*, vol. 107, n. 2, February 2000, pp. 1071-1074.
- [358] S. Hwang, Y. Park, Y. Park, Analysis on perceptual sensitivity to head-related impulse responses in the median plane, *Journal of Mechanical Science and Technology*, vol. 23, n. 12, December 2009, pp. 3340-3348.

-
- [359] Texas Instruments, Stellaris LM3S8962 Evaluation Board, User's Manual.
- [360] V. Tîponut, D. Ianchis, Z. **Haraszy**, T. Slavici, Tehnici de elaborare ale sistemelor dedicate, Fundația "Ioan Slavici" Timișoara, ISBN: 978-606-92306-7-1, 2010.
- [361] W. M. Leach, The Class-D Amplifier, In W. M. Leach (Ed.), Introduction to Electroacoustics and Audio Amplifier Design, Second Edition - Revised Printing (Kendall/Hunt, 2001).
- [362] S. S. Moreno, Class D audio amplifiers: theory and design, Revision A, 2005.
- [363] Maxim, Class D Amplifiers: Fundamentals of Operation and Recent Developments, Application Note 3977.
- [364] J. Honda, J. Adams, Class D Audio Amplifier Basics, International Rectifier Application Note AN-1071, 2008.
- [365] J. Guy, Class D Amplifier FAQ, National Semiconductor Corp, 2009.
- [366] D. Ianchis, Z. **Haraszy**, V. Tîponut, Collision Detection Inspired by Locust Neural System, Sesiunea de comunicări științifice "Doctor ETC 2009", September 24-25, 2009, Timisoara, Romania.
- [367] D. Ianchis, V. Tîponut, S. Popescu, Z. **Haraszy**, Improved Collision Detection System Inspired from the Neural Network of the Locust, The 9th IEEE International Symposium on Intelligent Systems and Informatics (SISY 2011), September 8-10, 2011, Subotica, Serbia.
- [368] N. J. Cotton, A Neural Network Implementation on Embedded Systems, Ph.D. dissertation, Auburn University, Auburn, Alabama, USA, 2010.
- [369] T. Matsumura, N. Iwanaga, T. Onoye, W. Kobayashi, I. Shirakawa, I. Arungsrisangchai, 3D Sound Movement System for Embedded Applications, The 2005 IEEE International Symposium on Circuits and Systems (ISCAS 2005), May 23-26, 2005, Kobe, Japan.

ANEXE

A1. Proiectarea echipamentului prototip

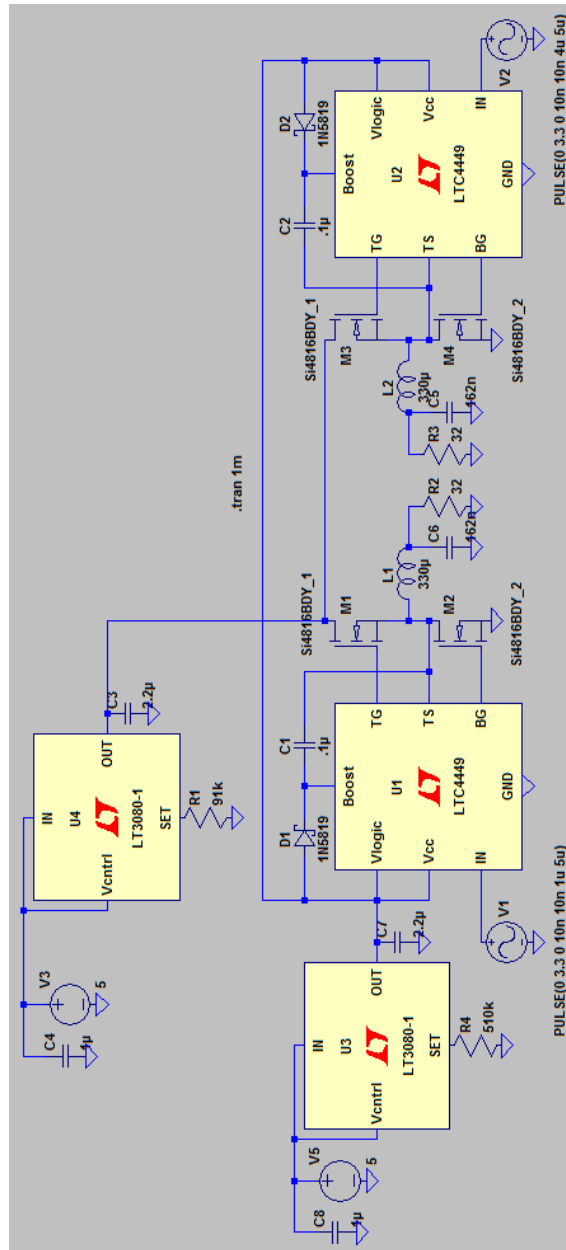


Fig. A1. Componenta audio a plăcii experimentale în mediul LTspice IV.

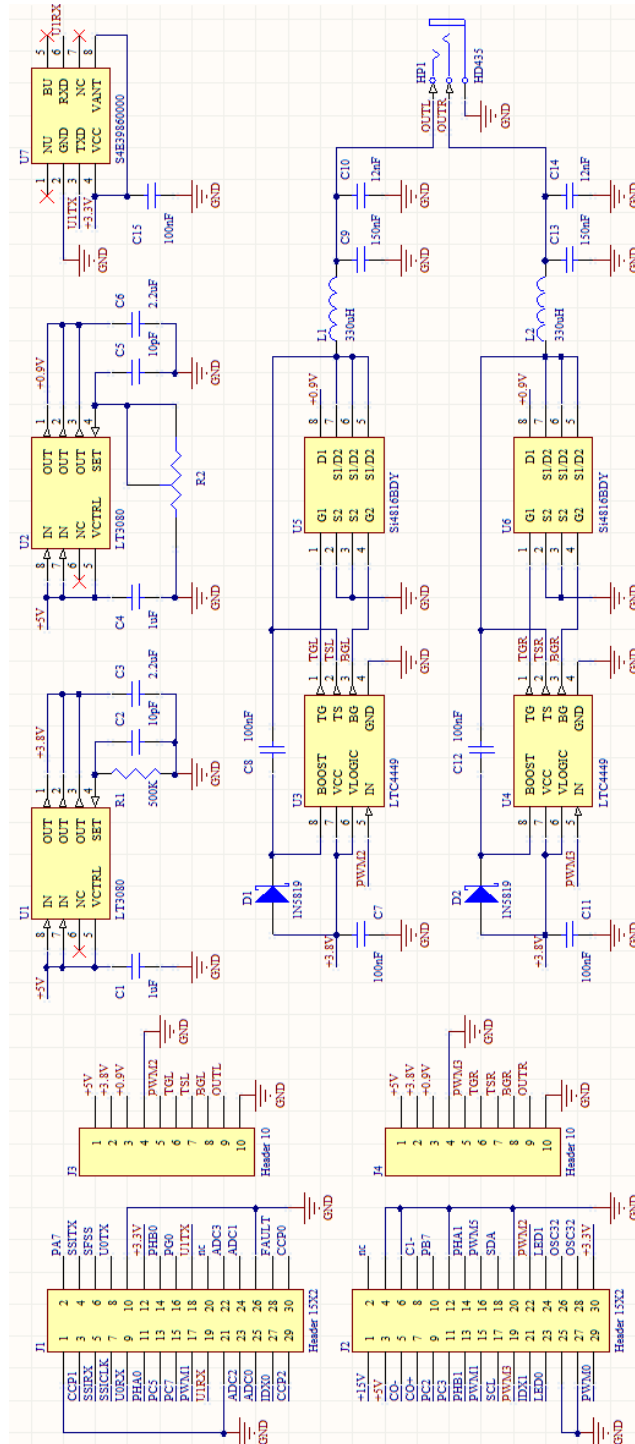


Fig. A2. Schema electronică a plăcii experimentale în Altium Designer Summer 09.

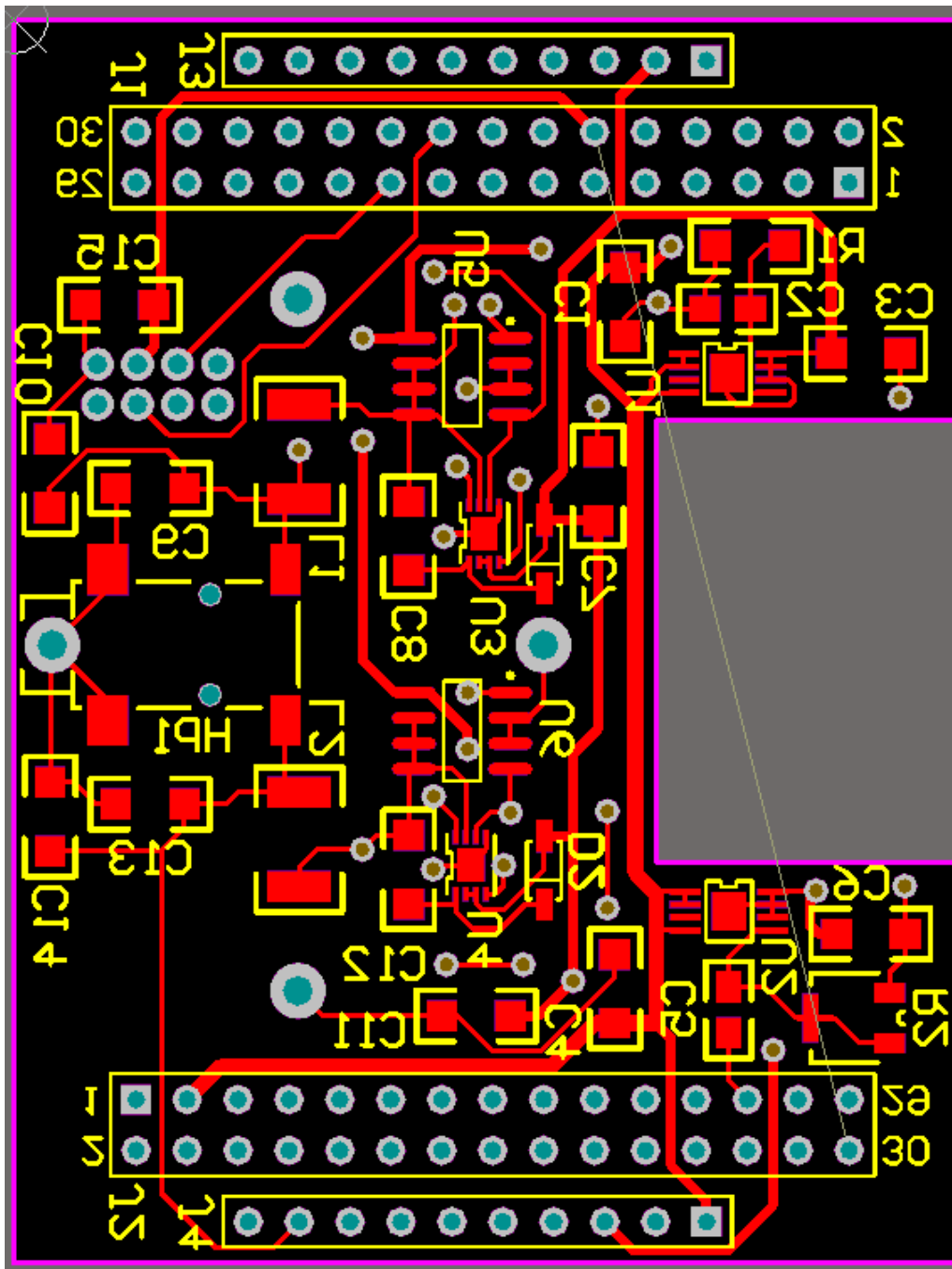


Fig. A3. Cablajul imprimat al plăcii experimentale – vedere sus.

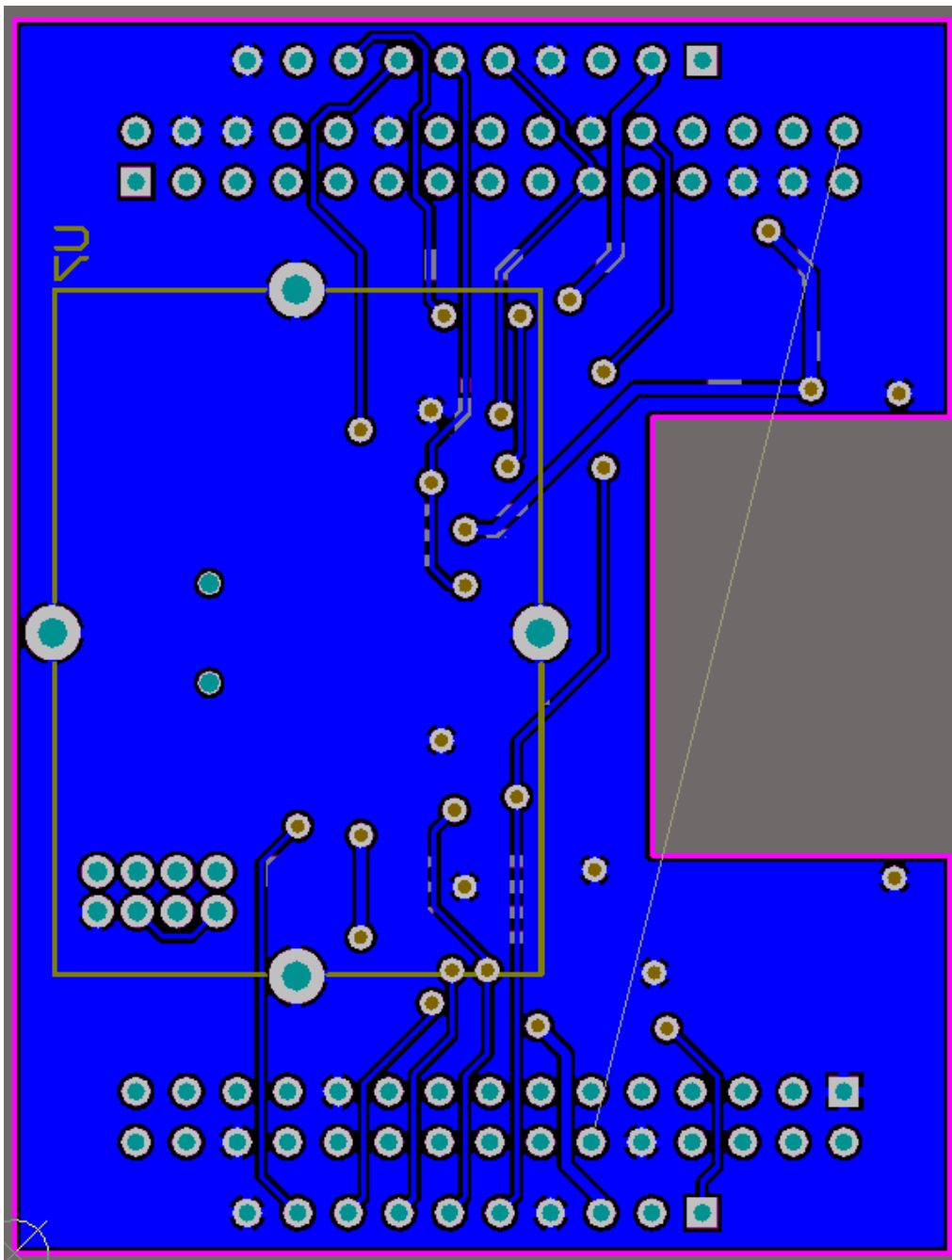


Fig. A4. Cablajul imprimat al plăcii experimentale – vedere jos.

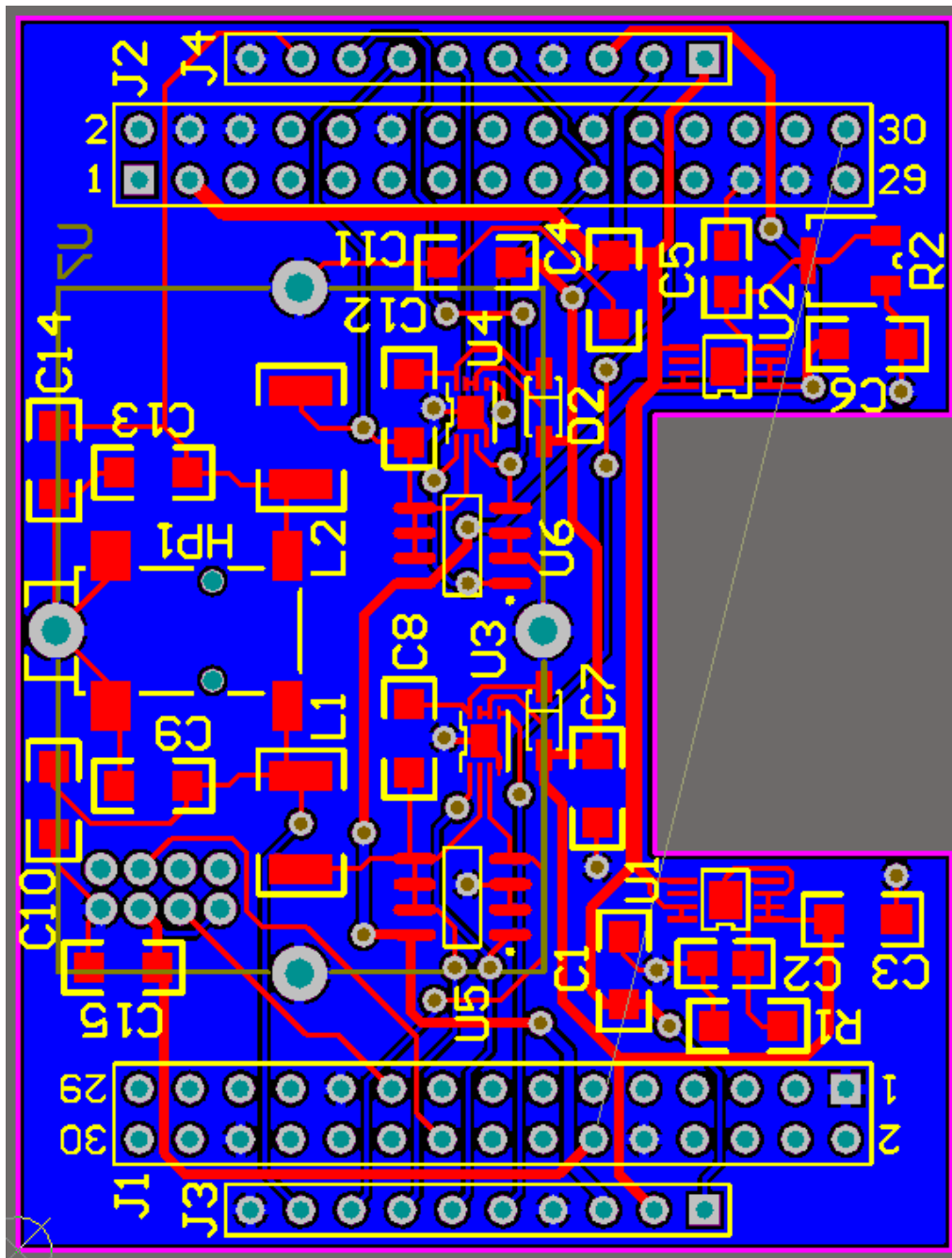


Fig. A5. Cablajul imprimat al plăcii experimentale – vedere sus-jos.