

Univerza v Ljubljani
Fakulteta za elektrotehniko

Rudolf Sušnik

**Postopki kodiranja elevacije izvorov zvoka
v akustični sliki prostora**

DOKTORSKA DISERTACIJA

Mentor: prof. dr. Sašo Tomažič

Ljubljana, 2006

ZAHVALA

Zahvaljujem se vsem, ki so v kakršnikoli obliki pripomogli k nastanku tega dela.

Mentorju prof. dr. Sašu Tomažiču se zahvaljujem za strokovno vodstvo, koristne nasvete, konstruktivne predloge in komentarje. Kolegu Jaku Sodniku se zahvaljujem za konstruktivne strokovne debate in predloge ter nenadomestljivo pomoč pri izvajanju poskusov. Za strokovna mnenja se zahvaljujem tudi dr. Antonu Umeku in mag. Goranu Bobojeviću. Prav tako se zahvaljujem ostalim sodelavcem Laboratorija za komunikacijske naprave za vsestransko pomoč.

Za sodelovanje pri poskusih se zahvaljujem dijakom Gimnazije Kranj, dijakom Tehniškega šolskega centra Kranj, učiteljicam Osnovne šole Orehek, študentom Fakultete za elektrotehniko in sodelavcem Laboratorija za komunikacijske naprave.

Posebna zahvala pa gre predvsem domačim – Špeli, staršema in bratu za vse vzpodbudne, optimistične in kritične nasvete ter dejanja v preteklih letih.

Vsem imenovanim in neimenovanim še enkrat hvala.

POVZETEK

Akustična slika je pojem, ki se nanaša na podajanje informacij z neartikuliranimi zvoki, tj. z zvoki, ki jih ne moremo opredeliti kot govor. Tak način komunikacije zajema široko paleto naprav, vse od preprostih zvočnih alarmov do predstavitve prostora z zvokom. Na predstavitev prostora z zvokom se nanaša tudi akustična slika prostora s katero se ukvarjamo v tem delu, kjer nas zanima predvsem vertikalna dimenzija, tj. elevacija v akustični sliki prostora.

Prostor v akustični sliki predstavimo s prostorskim zvokom oz. z navideznimi izvori zvoka v prostoru. Predmete v okolici poslušalca nadomestimo z navideznimi izvori zvoka, ki jih s pomočjo prostorskega zvoka premikamo po prostoru. Prostorski zvok je »tri-dimenzionalen«, tj. poslušalec zmore za vsak zvok določiti položaj njegovega izvora (azimut, elevacija, oddaljenost). Za reprodukcijo zvoka se uporabijo običajne slušalke, kar glede na dejstvo, da ima človek dva sprejemnika za zvok, povsem zadostuje.

Prostorskost zvoka je posledica dveh poglavitnih vzrokov, ki nastaneta pri poslušanju v prostoru: medušesne razlike (razlika v amplitudi valovanja in časovni zamik med trenutkoma, ko zvočno valovanje doseže levo oz. desno uho) in vplivi poslušalčevega telesa (predvsem uhljev) na spekter zvočnega valovanja. Za doseg prostorskih učinkov pri reprodukciji zvoka s slušalkami je omenjene vplive potrebno ustrezno upoštevati. Vplivi so najtočneje opisani s t.i. HRIR (Head Related Impulse Response), ki predstavljajo odziv prenosne poti med izvorom zvoka in ušesom na enotin impulz. Fourierov transform HRIR se imenuje HRTF (Head Related Transfer Functions). Filtriranje poljubnega zvoka s filtrom na osnovi HRIR povzroči, da poslušalec s slušalkami sliši zvok v smeri, ki jo opisuje HRIR. HRIR oz. HRTF delimo na individualne (pripadajo točno določeni osebi) in splošne (izmerjene na lutki).

Z akustično sliko prostora se v tem delu ukvarjamo v smislu orientacijskega in navigacijskega pripomočka, ki naj bi slepim osebam služil namesto bele palice. V literaturi najdemo podatke o že obstoječih izvedbah akustične slike prostora. V teh izvedbah se uporabljajo princip ehelokacije, individualne HRTF in umetni načini kodiranja prostora. Uporaba individualnih HRTF je sicer najboljši način, vendar je zaradi kompleksnosti merjenja individualnih HRTF le-to velika ovira. Eholokacija in umetni načini kodiranja zahtevajo bistveno več učenja, poleg tega pa izvedbe, ki smo jih preizkusili, proizvajajo silno moteče zvoke.

Naša rešitev predvideva kombinacijo splošnih HRTF in umetnega kodiranja ter uporabo zvokov, ki so prijetni za poslušanje. Ker z uporabo splošnih HRTF točnost lokalizacije izvorov zvoka pade le v elevaciji, v naši akustični sliki uporabimo splošne HRTF za kodiranje azimuta, za kodiranje elevacije pa uvedemo umetno kodiranje. Razvoj kodirnih postopkov za kodiranje elevacije izvorov zvoka predstavlja jedro tega dela.

Postopek kodiranja elevacije izvorov zvoka v akustični sliki prostora mora zadostiti več zahtevam. Predvsem je pomembno kolikšno točnost lokalizacije lahko dosežemo – kot orientacijsko vrednost vzamemo točnost lokalizacije izvorov zvoka v prostoru. Nadalje je pomembno, da zvok, ki ga proizvaja akustična slika ni pretirano moteč in da je poslušalec sposoben čim hitreje osvojiti t.i. umetno kodiranje. V vseh akustičnih slikah, ki uporabljajo umetno kodiranje elevacije, se mora poslušalec naučiti katero elevacijo predstavlja posamezen zvok. Lokalizacija po elevaciji je brez predhodnega učenja možna le pri kodiranju z individualnimi HRTF, saj ljudje to kodiranje uporabljamo od rojstva.

Za določitev praktičnih omejitev pri lokalizaciji izvorov zvoka po elevaciji smo izvedli poskus z zvočniki v gluhi sobi. Rezultati kažejo, da je človek sposoben ločiti dva izvora zvoka, ki sta po elevaciji razmaknjena za vsaj 6° . Poskus ugotavljanja ločljivosti po elevaciji smo izvedli tudi s slušalkami – z uporabo splošnih HRTF in z uporabo individualnih HRTF. Splošne HRTF dajejo silno slabe rezultate, za ločevanje dveh izvorov zvoka morata biti izvora kar 23° narazen, z individualnimi HRTF pa se približamo ločljivosti, ki jo dosežemo v prostoru.

Pri razvoju postopka kodiranja elevacije smo izhajali iz predpostavke, da bo čas učenja oz. privajanja uporabnika na umetno kodiranje tem krajši, čim natančneje nam bo uspelo posnemati lastnosti, ki omogočajo človeku lokalizacijo izvorov zvoka v naravnem okolju. Uhlji in ostali deli poslušalčevega telesa delujejo kot filter katerega učinke v odvisnosti od elevacije oz. položaja izvora zvoka lahko opazujemo v HRTF. Učinke filtriranja smo ugotavljali z analizo poteka frekvenc oslabitve uhlja (angl. pinna notch) in porazdelitve energije HRTF v odvisnosti od elevacije. Na podlagi rezultatov analiz in poznanih dejstev o povezavi elevacije in frekvence smo razvili tri modele kodiranja elevacije. Zvoke smo izbrali izmed zvokov, ki se najpogosteje uporabljajo v poskusih povezanih z akustiko: beli šum, rožnati šum, sinus oz. ton.

S kombinacijami zvokov in modelov kodiranja elevacije smo dobili več naborov zvokov (nabor dobimo, ko modelu priredimo zvok). S testom zaznavnih razlik med zvoki istega nabora smo ugotavljali najboljše kombinacije zvoka in modela, ki smo jih nato uporabili za kodiranje elevacije.

Razvili smo šest akustičnih slik prostora z različnimi načini umetnega kodiranja elevacije in preizkusili kako natančno so se testne osebe sposobne naučiti lokalizacije z umetno kodirano elevacijo izvorov zvoka. Pri najboljših izvedbah akustične slike so testne osebe po desetih, približno 5 minut trajajočih serijah učenja, dosegle povprečno

ločljivost 7.5° , kar je rezultat, ki daje dobre obete za nadaljnji razvoj akustične slike prostora z umetnim kodiranjem elevacije.

ABSTRACT

Auditory display or auditory image is the use of nonspeech sound to communicate information to a listener. The field of auditory display encompasses everything from alarms to complex sonification like auditory display of space. Latter is primary concern of this thesis particularly concentrated on elevation in auditory display of space.

The key attribute which enables representation of space with auditory display is spatial sound. Spatial sound enables one to hear in three dimensions which mean that he or she is able to localize sound source with three coordinates - azimuth, elevation and distance. In auditory display, the objects around the listener are therefore replaced with virtual sound sources which can be virtually moved according to the position of listener and object. Sounds of auditory display of space are played through headphones.

There are two main causes which enable spatial sound: interaural (level and time) differences and head related influences which affect sound spectrum. When played through headphones, sound must be correctly pre-processed to invoke spatial sensations. All influences are due to sound propagation from sound source to listener's eardrum and are therefore most accurately described with head related impulse response - HRIR. The Fourier transform of HRIR is called HRTF which stands for head related transfer function. When arbitrary sound is convolved with particular HRIR and played through headphones listener would hear the sound originating from the direction which corresponds to the HRIR used in convolution. HRIRs (HRTFs) are divided into two major groups: individual (measured with particular persons) and non-individual (measured with manikin, e.g. KEMAR).

The aim of our auditory display of space is to become an orientation and navigation tool for visually impaired persons. In literature one can found few existing auditory displays of space. They use different principles - echolocation, artificial coding techniques and sound filtering with individual HRTFs. Individual HRTF filtering would be the best choice since it does not require learning and therefore ensures high localization fidelity. Method's main drawback is necessity of individual HRTFs which requires extensive measurements. The use of echolocation and artificial coding requires time to learn the concept of auditory display and produces unpleasant sounds (at least to our feelings). Learning is indispensable for all auditory displays with artificial elevation coding, since absolute localization in elevation is impossible with any kind of artificial elevation

coding. Without pre-learning only individual HRTFs enables absolute elevation localization. Individual HRTFs represents coding which humans learn since their birth.

Our approach is combination of non-individual HRTFs, artificial coding techniques and non-disturbing sounds. Non-individual HRTFs give good results in azimuth localization but very poor in elevation localization. Therefore our auditory display uses artificial coding only in elevation. Artificial coding of sound source elevation is essence of this thesis.

An artificial elevation coding technique should satisfy a few requirements: reasonable localization accuracy (nearly as good as in natural environment), fast learning and pleasantness of sound.

Localization accuracy in natural environment was established with experiment in anechoic room. Test subjects were able to distinguish between two sound sources located at least 6° apart in elevation. The nearly same result was obtained with headphones and individual HRTFs, while with headphones and non-individual HRTFs (virtual) sound sources should be at least 23° apart.

One of the main hypotheses in auditory display of space development was that similarity of synthesized and natural sounds which represents the same elevation would enable faster learning. Localization cues are introduced by the body, pinna and head filtering and they reflect in HRTFs. Localization cues was therefore studied by analyzing HRTFs – extracting pinna notch frequencies and observing the frequency dependent distribution of HRTFs' energy. According to HRTFs' analyses three elevation coding techniques were developed. Beside coding techniques, sounds must be also defined. We used combinations of sounds which are regularly used in acoustics: white noise, pink noise, pitch.

Using particular coding technique and sound we obtained a few sound sets. The most promising sound sets were identified by the experiment in which we searched for number of distinguishable sounds within one sound set. This experiment was a sort of search for just noticeable differences in spectrum.

We developed six auditory displays of space for the final experiment. Test subjects tried to learn artificial elevation coding. After ten sessions (each lasted approximately 5 minutes) of learning and testing, the resolution of 7.5° was achieved within the best auditory displays. The result encourages us for further researches in field of auditory display of space.

PRISPEVKI K ZNANOSTI

- Razširitev poznavanja mehanizmov zaznavanja elevacije izvorov zvoka na podlagi analize energije HRTF, ki se v splošnem sklada z zakonitostjo povezave elevacije in frekvenc tonov oz. centralnih frekvenc spektra pasovno omejenih zvokovnih signalov. S poskusom smo pokazali, da odvisnost lahko razširimo tudi na pasovno omejene zvokovne signale katerih spektri imajo fiksno spodnjo mejno frekvenco 16 Hz (ali manj) in različno zgornjo mejno frekvenco.
- Predlog učinkovitih postopkov umetnega kodiranja elevacije izvorov zvoka v akustični sliki prostora, ki smo jih določili na podlagi opravljenih poskusov. V poskusih smo uporabili kombinacije postopkov kodiranja in zvokov, ki jih v literaturi še nismo zasledili. Kot je razvidno iz rezultatov, je učinkovitost kodiranja poleg kodirnega postopka odvisna tudi od uporabljenega zvoka.
- Predlog zvokov, ki so primerni za umetno kodiranje elevacije s stališča učinkovite lokalizacije in prijetnosti za poslušalca. Zahtevi sta si v nekaterih primerih (npr. rožnati šum) nasprotni. Z uporabo zvokov, ki so standardni v akustičnih poskusih (beli šum, rožnati šum, pasovni šum, sinusni signal oz. ton) in v tem delu predlaganimi zvoki, ki posnemajo naravne pojave (npr. udarec na boben), smo s poskusom pokazali, da optimum predstavljajo zvoki, ki posnemajo naravne pojave.

KAZALO

1	UVOD.....	3
2	PROSTORSKI ZVOK.....	9
2.1	Zvočno valovanje.....	9
2.2	Človekovo uho.....	12
2.3	Opis prostora.....	14
2.4	Akustika prostora.....	17
3	PSIHOAKUSTIKA IN PROCES UČENJA.....	19
3.1	Glasnost zvoka.....	19
3.2	Maskiranje in najmanjše zaznavne razlike glasnosti ter frekvence.....	22
3.3	Kritični pasovi.....	23
3.4	Prenosne funkcije vpliva glave in telesa HRTF (Head Related Transfer Functions).....	25
3.5	Učenje.....	26
4	LOKALIZACIJA IZVOROV ZVOKA PRI LJUDEH.....	31
4.1	Lokalizacija izvorov zvoka v prostoru in pri predvajanju s slušalkami.....	31
4.2	Opredeleitev pojma »ločljivost«.....	32
4.3	Merjenje individualnih HRTF.....	34
4.4	Točnost lokalizacije (absolutna ločljivost).....	34
4.5	Merjenje diferenčne ločljivosti s slušalkami in z zvočniki.....	35
4.5.1	Merjenje diferenčne ločljivosti po elevaciji z zvočniki.....	35
4.5.2	Merjenje diferenčne ločljivosti po elevaciji s slušalkami in z uporabo splošnih HRTF.....	38
4.5.3	Merjenje diferenčne ločljivosti po elevaciji s slušalkami in uporabo individualnih funkcij.....	40
4.6	Vpliv spektralne vsebine zvoka na lokalizacijo.....	42
4.7	Vpliv glasnosti in trajanja zvočnega signala na lokalizacijo.....	43
4.8	Komentar.....	45
5	ELEVACIJA – ANALIZA HRTF IN AKUSTIČNE ILUZIJE.....	47
5.1	Elevacija v akustični sliki prostora.....	47
5.2	Pregled nekaterih preizkušenih metod kodiranja elevacije v akustični sliki prostora.....	48
5.3	Analiza HRTF.....	49
5.3.1	Oslabitev uhlja (pinna notch).....	50
5.3.2	Energija HRTF.....	53
5.4	Akustične iluzije v elevaciji.....	56
5.4.1	Akustične iluzije v prostoru.....	57
5.5	Komentar.....	62
6	RAZVOJ IN OVREDNOTENJE MODELA ZA KODIRANJE ELEVACIJE.....	63
6.1	Ideja umetnega kodiranja elevacije.....	63
6.2	Postopki obdelave signalov.....	64
6.2.1	Model na osnovi »oslabitve uhlja«.....	64
6.2.2	Model na osnovi porazdelitve energije HRTF.....	65

6.2.3	Modeli na osnovi povezave elevacije in frekvence.....	65
6.3	Zaznavne razlike med zvoki, ki pripadajo istemu naboru.....	65
6.3.1	Nabori zvokov	66
6.3.2	Opis poskusa	67
6.3.3	Testne osebe in oprema	69
6.3.4	Rezultati	69
6.4	Akustična slika v vertikalni smeri.....	71
6.4.1	Zasnova akustične slike in nabori zvokov.....	72
6.4.2	Opis poskusa	77
6.4.2.1	Učenje elevacije v akustični sliki	77
6.4.2.2	Ocenjevanje prijetnosti zvokov v akustični sliki.....	78
6.4.3	Testne osebe in oprema	79
6.4.3.1	Učenje elevacije v akustični sliki	79
6.4.3.2	Ocenjevanje prijetnosti zvokov v akustični sliki.....	79
6.4.4	Rezultati	79
6.4.4.1	Učenje akustične slike.....	79
6.4.4.2	Ocenjevanje prijetnosti zvokov v akustični sliki.....	82
6.4.5	Komentar	83
6.5	Razširitev akustične slike po azimutu	84
7	ZAKLJUČEK.....	85
	LITERATURA.....	87
	SEZNAM KRATIC.....	93

1 UVOD

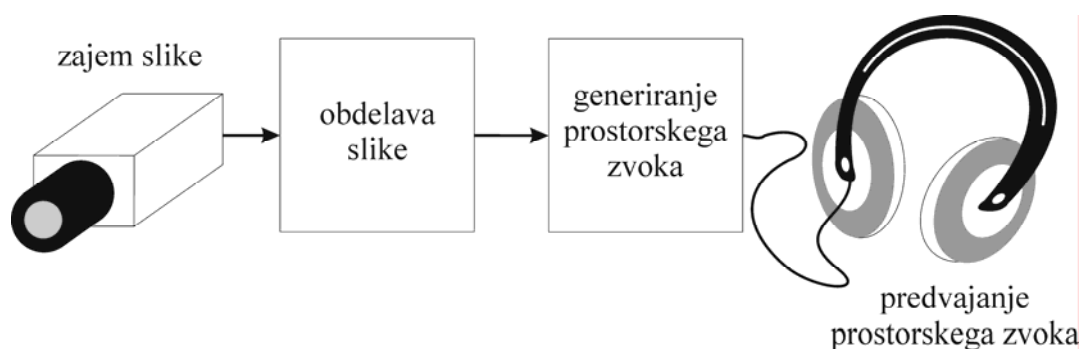
Prostor in dogajanje v prostoru lahko namesto z običajno sliko predstavimo s prostorskim zvokom. Zvok je pomemben pripomoček pri orientaciji v prostoru in preprosto se je moč prepričati, da zvok slišimo v treh dimenzijah. Z natančnostjo nekaj stopinj ali celo manj so možgani sočasno sposobni določiti položaj več izvorov zvoka v prostoru. Določanju položaja izvora zvoka rečemo tudi lokalizacija (angl. localisation) [4, 5, 16, 21, 46].

Za lokalizacijo izvorov zvoka ljudje izrabljamo številne lastnosti, ki so prisotne v zvoku. Te lastnosti so posledica razširjanja zvočnih valov od izvora zvoka do bobniča poslušalca. Vpliv okolice na zvočno valovanje je izražen z odboji, lomi in interferenco na predmetih, ki se nahajajo na poti med izvorom zvoka in poslušalcem, slabljenjem amplitude valovanja zaradi razširjanja, absorpcije ipd. [4, 5, 15, 35, 81]. Ti vplivi so odvisni od trenutne situacije, tj. lokacije poslušalca in lokacije izvora zvoka. Vplivi okolice so pomembni predvsem zato, ker v zvok vnašajo informacijo o oddaljenosti izvora zvoka od poslušalca. Zaznavanje oddaljenosti je namreč posledica odbojev, tako od predmetov v okolici poslušalca kot odbojev od poslušalčevega telesa. Za določanje smeri izvora zvoka (azimut in elevacija) so ključni predvsem vplivi, ki so posledica odbojev, loma in interference zvočnega valovanja zaradi poslušalčevega telesa, glave in uhljev. Vpliv glave, telesa in uhljev na širjenje zvoka je opisan v zbirki odzivov prenosne poti med izvorom zvoka in poslušalčevim bobničem na enotin impulz. Odziv se imenuje HRIR (Head Related Impulse Response), njegov Fourierov transform pa HRTF (HRTF – Head Related Transfer Function). HRIR vsebujejo vse potrebne podatke za določitev smeri izvora zvoka – medušesno časovno zakasnitev (angl. Interaural Time Difference – ITD), medušesno amplitudno razliko (angl. Interaural Level Difference – ILD) in amplitudno karakteristiko [5, 11, 27, 56, 63, 67].

Medušesne časovne zakasnitve in amplitudne razlike so ključnega pomena za lokalizacijo izvora zvoka v horizontalni dimenziji (azimut) [4, 16, 67], medtem ko so za lokalizacijo v vertikalni smeri (elevacija) pomembne razlike med amplitudnimi poteki HRTF [1, 33, 39, 76, 79] pri različnih elevacijah. Razlike v amplitudnih potekih HRTF postanejo očitne pri frekvencah nad 4 kHz [1, 5], ko valovna dolžina zvoka postane primerljiva z velikostjo uhlja. Uhlj vpliva na spektralne lastnosti zvoka predvsem s t.i. oslavitvijo uhlja [1, 30, 79], ki se, odvisno od elevacije izvora zvoka, pojavlja med 5 kHz in 17 kHz [57]. Ta ugotovitev se sklada s trditvijo [1, 39], da mora zvok za natančno

lokalizacijo po elevaciji vsebovati komponente višjih frekvenc. Kljub temu pa je človek sposoben lokalizirati tudi zvoke, ki vsebujejo le frekvence do 3 kHz [1]. Prisotnost višjih frekvence izboljša natančnost lokalizacije, vendar so tudi vplivi odbojev od telesa, ki se pojavljajo pod 3 kHz, dovolj za uspešno lokalizacijo po elevaciji [1]. Frekvenčna vsebina zvoka ima tudi minimalni vpliv na lokalizacijo po azimutu [33], kjer se prav tako izkaže, da širši frekvenčni spekter zvoka zagotavlja natančnejšo lokalizacijo. Pojavi v zvezi z elevacijo in oddaljenostjo so zaenkrat še relativno slabo raziskani, medtem ko je azimut bolje obdelan.

Akustična slika prostora je pripomoček, ki s pomočjo prostorskega zvoka poskuša pomagati slepim osebam pri orientaciji v prostoru, uporablja se tudi kot navigacijski pripomoček posebnim skupinam ljudi kot so to npr. rudarji, tankisti, piloti ipd. [22, 34, 44, 68]. To pomeni, da želimo s pomočjo zvoka čim natančneje opisati vidno sliko, ki jo zajamemo s kamero, radarjem ali sonarjem, zvoke akustične slike pa predvajamo preko slušalk (slika 1.1). Opravka imamo z dokaj različnimi uporabniki, ki imajo to skupno lastnost, da tudi sicer aktivno uporabljajo sluh. Akustična slika jim predstavlja dodaten vir podatkov, ki naj bi čim manj vplival na obstoječe funkcije, zato je potrebno da zvoki niso moteči za poslušalca. Bistvena razlika med uporabniki je tudi pogostost uporabe pripomočka.



Slika 1.1: Akustična slika je transformacija vizualne slike v prostorski zvok, ki ga predvajamo preko slušalk. Iz procesa obdelave slik dobimo podatke o lokacijah objektov, ki so pomembni za orientacijo ali navigacijo (slepe) osebe v prostoru, nato te objekte predstavimo kot virtualne izvore zvoka v prostoru. V tem delu se ukvarjamo izključno z generiranjem prostorskega zvoka.

Najbolj razširjen pripomoček slepim ljudem za orientacijo v prostoru je bela palica, v zadnjem času pa se pojavljajo novi, tehnično naprednejši pripomočki. Med te štejemo akustično sliko prostora, elektronsko belo palico ipd. Pri vsaki metodi gre za transformacijo iz vizualne oblike v neko drugo obliko – najpogosteje akustično ali

taktilno (otip). Z akustično sliko želimo slepim ljudem približati vsaj nekaj tega, kar ljudje z normalno razvitim vidom poznamo kot »vid«.

Poznanih je več načinov ustvarjanja akustične slike prostora. V tej disertaciji zagovarjamo princip razdeljevanja vidnega prostora v več podprostorov. Če se v podprostoru nahaja ovira, poslušalec zasliši zvok, ki prihaja iz ustreznega podprostora. Velikost podprostorov je omejena s človekovo sposobnostjo lokalizacije izvorov zvoka [4, 48, 64].

Generiranje prostorskega zvoka s slušalkami je poseben problem kjer želimo generirati tak zvok, da se poslušalcu zdi, da zvoki resnično prihajajo iz prostora okoli njega [4, 5, 21]. V najboljšem primeru to pomeni, da poslušalec iz zvoka pridobi enako informacijo, kot bi jo, če bi se resnično nahajal v tem prostoru. Pri tem se srečamo s posebnostmi snemanja zvoka oz. z obdelave poznanega zvoka, ko želimo izvor nekega zvoka (navidezno) postaviti v drugo točko prostora. Pri navideznem postavljanju izvorov zvoka v poljubne točke prostora nam povzroča težave odvisnost spektra zvoka od elevacije izvora, odboji od ovir ter časovne in amplitudne razlike. Pri ustvarjanju akustične slike prostora privzamemo prostor brez ovir in zato ne upoštevamo vplivov okoliških predmetov na zvok. Tak prostor simuliramo z uporabo gluhe sobe [15, 16]. V takem primeru načeloma zadostuje obdelava s HRTF, pri čemer se najbolje obnesejo individualne HRTF, to so funkcije, ki so merjene za vsakega poslušalca posebej [11]. Merjenje HRTF [66] je relativno zapleten in dolgotrajen postopek in se mu zato želimo izogniti. Za kodiranje v horizontalni smeri se zadovoljivo obnesejo tudi splošne HRTF, ki pa za kodiranje elevacije niso neposredno uporabne.

Kot smo že omenili, je najboljša rešitev za ustvarjanje akustične slike prostora uporaba individualnih HRTF [77, 78]. S tem se lahko približamo ločljivosti, ki jo doseže človek v prostoru (2° po azimut in 6° po elevaciji) [64, 65]. Pri uporabi splošnih HRTF se ločljivost po elevaciji bistveno poslabša. S splošnimi HRTF knjižnice CIPIC [2] dosežemo po azimutu ločljivost 4° , po elevaciji pa le 23° [65]. Za kodiranje v horizontalni smeri zato lahko uporabimo splošne HRTF, za elevacijo pa moramo uporabiti specifičen način kodiranja.

Kodiranje elevacije, ki ga predstavljamo v tem delu, smo si zamislili na osnovi povezave frekvence zvoka in zaznavanja elevacije. Iz osnovnega zvoka dobimo s pomočjo tehnik obdelave signalov množico zvokov, nato pa vsaki elevaciji priredimo svoj zvok. Da bi dobili 2D akustično sliko, zvok vsake elevacije filtriramo s filtrom za ustrezni azimut. Uporabimo lahko filtre horizontalne ravnine katere od splošnih knjižnic HRTF [2, 20, 54] ali pa zgolj upoštevamo medušesne časovne in amplitudne razlike. V preteklih raziskavah o lokalizaciji izvorov zvoka po elevaciji so uporabljali različne oblike zvokovnih signalov, npr. šumi različnih pasovnih širin, naključni signali ipd. [5, 7, 30, 39, 42, 43, 60, 74, 86]. V poskusih z ozkopasovnimi signali (sinusni signali ali ozkopasovni

šum) se je pokazala povezanost centralne frekvence zvoka in elevacije izvora zvoka [3, 59]. Težava pri poslušanju sintetiziranih zvokov s slušalkami je, da poslušalci ne morejo določiti absolutne elevacije položaja izvora [30]. Absolutni položaj izvora zvoka je možno določiti le, če zvok vsebuje individualne lastnosti, ki jih pri poslušanju v prostoru vnesejo vplivi telesa, glave in uhljev.

V tem delu torej poskušamo določiti primeren postopek za kodiranje elevacije v akustični sliki prostora in izbrati ustrezen zvok, ki ne bo preveč moteč in bo nudil zadovoljivo ločljivost tako v vertikalni kot horizontalni ravnini.

Poglavje ki sledi uvodu, tj. drugo poglavje, je namenjeno predstavitvi prostorskega zvoka s fizikalnega in fiziološkega stališča. Najprej so predstavljene osnovne značilnosti zvočnega valovanja in zgradba ter delovanje človekovega organa za sluh, nato sledi opis prostora in postavitev koordinatnega sistema ter na koncu še nekaj osnovnih pojmov o akustiki prostora, tj. vplivu prostora na človekovo zaznavanje zvoka.

V tretjem poglavju je pozornost namenjena psihoakustičnim pojavom in procesu učenja. Psihoakustika se ukvarja s človekovim dojemanjem in ovrednotenjem zvokov, tu razpravljamo o glasnosti zvoka, maskiranju, minimalnih spektralnih razlikah med dvema zvokoma, ki so potrebne za njuno razločevanje, kritičnih pasovih, prenosnih funkcijah glave in telesa (Head Related Transfer Functions – HRTF) ter o procesu učenja, ki npr. omogoča prilagajanje na spremenjene razmere za orientacijo.

V četrtem poglavju se ukvarjamo z lokalizacijo prostorskih zvokov pri ljudeh. Uvedemo pojma natančnosti (absolutne) lokalizacije in diferenčne ločljivosti. Podrobneje razčlenjena sta primera lokalizacije resničnih izvorov zvoka v prostoru in lokalizacije navideznih izvorov zvoka, predvajanih preko slušalk, ki sta za nas zanimiva zlasti v elevaciji. Poleg predstavitve rezultatov, ki pomenijo mejo človekove zmožnosti lokalizacije izvorov zvoka so v tem poglavju obdelani še trije dejavniki, ki vplivajo na lokalizacijo izvorov zvoka. To so spekter, glasnost in trajanje zvočnega signala. Četrto poglavje je zaključeno s komentarjem.

V petem poglavju se podrobneje seznanimo s pogoji, ki omogočajo lokalizacijo elevacije izvorov zvoka. Za ugotavljanje mehanizmov lokalizacije po elevaciji je ključna analiza HRTF, ki smo jih analizirali z iskanjem frekvenc oslabitev uhlja (angl. pinna notch) in z analizo porazdelitve energije HRTF v odvisnosti od elevacije. V tem poglavju je opisan tudi poskus z ustvarjanjem akustičnih iluzij z zvočniki v prostoru. Sklepni del poglavja je komentar k glavnim ugotovitvam.

V šestem poglavju modelom oz. postopkom kodiranja elevacije izvorov zvoka dodamo zvok in tako dobimo nabore (množice) zvokov, ki so potencialno uporabni za konstruiranje akustične slike prostora. Ker so si zvoki, ki sestavljajo določen nabor podobni, z eksperimentom določimo nabore z najbolj ločljivimi zvoki. Sledi opis konstruiranja akustične slike, ki je zreducirana na eno dimenzijo (elevacija) in poskus

ugotavljanja možnosti lokalizacije navideznih izvorov zvoka v tej sliki. Rezultati vključujejo točnost lokalizacije in prijetnost zvokov. Razdelek o konstruiranju in preizkušanju akustične slike prostora zaključuje komentar k uporabljenim metodam in rezultatom. Šesto poglavje zaključuje opis načrtovane razširitve enodimenzionalne akustične slike v dvodimenzionalno (elevacija in azimut).

Delo zaključujemo s poglavjem Zaključek, kjer povzemamo in komentiramo glavne ugotovitve ter hkrati razmišljamo o novih idejah in izboljšavah, ki bi jih lahko uporabili v nadaljnjem razvoju akustične slike prostora.

2 PROSTORSKI ZVOK

2.1 Zvočno valovanje

Zvok je mehansko valovanje, ki se širi v snovi (trdnini, kapljevini ali plinu), ki je vsaj malo stisljiva. V kapljevinah in plinih je zvok vedno longitudinalno valovanje, v trdninah pa je mogoče izzvati tudi transverzno zvočno valovanje. V ožjem pomenu definiramo kot zvok le zvočno valovanje s frekvencami v slišnem območju človeškega ušesa, to je med 16 Hz in 20 kHz. Zvok s frekvencami nižjimi od 16 Hz imenujemo infrazvok, z višjimi od 20 kHz pa ultrazvok. Zvok in z njim povezane fizikalne pojave preučuje akustika [15, 16, 28, 35], subjektivno človeško zaznavanje zvoka pa psihoakustika [5, 28].

Zvočno valovanje opišemo s hitrostjo delcev, s silo, ki jo delci povzročajo in lastnostjo medija v katerem se valovanje širi. S pomočjo drugega in tretjega Newtonovega zakona klasične mehanike ugotovimo [28], da je odvod produkta hitrosti delcev (v) in gostote medija (ρ) enak gradientu tlaka $p(r,t)$:

$$\text{grad } p(\vec{r}, t) = -\frac{\partial}{\partial t}(\rho \vec{v}) \quad (2.1)$$

V snovi tako nastanejo spremembe tlaka in gostote oz. zgoščine in razredčine, ki so krajevno (r) in časovno (t) odvisne. V zgoščinah se gostota snovi in tlak povečata, v razredčinah pa zmanjšata. Spremembam tlaka v snovi zaradi longitudinalnega valovanja pravimo zvočni tlak. Človeško uho lahko zazna zvok, če je amplituda zvočnega tlaka v zraku ob bobniču najmanj 20 μPa . Najmočnejši zvok, ki ga uho prenese še brez škode, ustreza zvočnemu tlaku okrog 20 Pa.

Zgoščine in razredčine potujejo skozi snov s hitrostjo zvoka, ki je odvisna od lastnosti snovi. Za trdnine in kapljevine velja, da se zvok širi tem hitreje, čim lažja in manj stisljiva je snov, pri plinih pa ugotovimo, da je hitrost zvoka odvisna neposredno od temperature [35]. Longitudinalno valovanje povzroča neprestano stiskanje in raztezanje snovi ter s tem spreminjanje temperature, vendar se to dogaja dovolj hitro, da se toplota ne utegne znatno pretakati in valovanje lahko obravnavamo kot adiabatno spremembo. Z upoštevanjem razmerja specifičnih toplot plina ($\gamma = c_p / c_v$) [35] izračunamo hitrost zvočnega valovanja c :

$$c = \sqrt{\frac{\gamma p_0}{\rho_0}} \quad (2.2)$$

Od tod sledi valovna enačba za širjenje zvočnega valovanja v plinu:

$$\frac{\partial^2 p}{\partial t^2} = c^2 \Delta p \quad (2.3)$$

oz. zapisana v kartezičnem koordinatnem sistemu (x, y, z) :

$$\frac{\partial^2 p}{\partial t^2} = c^2 \left(\frac{\partial^2 p}{\partial x^2} + \frac{\partial^2 p}{\partial y^2} + \frac{\partial^2 p}{\partial z^2} \right)$$

Poleg splošne rešitve valovne enačbe sta zanimivi še rešitev za ravninski val (tudi t.i. enodimenzionalna rešitev) in sferična rešitev [28], kjer so zvočne fronte koncentrične krogle. Večina izvorov zvoka proizvaja valovanje, ki se ga da bolje opisati s sferično rešitvijo kot z ravninsko, vendar pa pri velikih razdaljah ali če opazujemo valovanje na površini z majhnim presekom valovanje lažje opišemo kar z ravninsko rešitvijo. Z ravninskim modelom obravnavamo tudi zvočno valovanje v človekovem zunanjem sluhovodu in vokalnem traktu.

S tem, ko izvor zvoka vzbuja nihanje delcev v snovi opravlja delo, ki se kot energija razširja skozi snov. Tako s hitrostjo zvoka skozi snov potuje energija, ki jo lahko opišemo kot energijski tok P . Energijski tok P je definiran kot kvocient pretečene energije (W) in energijskega intervala (t) v katerem energija preteče skozi presek pravokotno na smer pretakanja [35]:

$$P = \frac{dW}{dt} \quad (2.4)$$

Energijski tok P je zvočna moč, ki jo oddaja izvor zvoka, učinek energijskega toka pa merimo z gostoto zvočnega toka j [15, 35]. Gostoto zvočnega toka imenujemo tudi jakost zvoka:

$$j = \frac{dP}{dS} = c \frac{dW}{dV} \quad \left[\frac{\text{W}}{\text{m}^2} \right] \quad (2.5)$$

V enačbi 2.5, pomeni S presek prečne ploskve, c hitrost zvoka in V prostornino. Gostoto zvočnega toka v poljubni točki prostora ob poljubnem času lahko izrazimo s pomočjo zvočnega tlaka in hitrosti zvoka v snovi [28]:

$$j(x, t) = p(x, t)c(x, t) \quad (2.6)$$

Kot smo že napisali, ima najšibkejši še slišni zvok amplitudo zvočnega tlaka $20 \mu\text{Pa}$. Od tod dobimo gostoto zvočnega toka $0.5 \cdot 10^{-12} \text{ W/m}^2$. Običajno se kot spodnja meja slišnosti vzame kar 10^{-12} W/m^2 , za zgornjo mejo pa 1 W/m^2 .

V praksi se srečamo z različnimi izvori zvoka, ki jih razporedimo v tri skupine, pri čemer je potrebno upoštevati tudi položaj opazovalca oz. oddaljenost opazovalca od izvora:

- točkasti izvori (angl. point sources): npr. zvočnik, človek, ki govori; gostota zvočnega energijskega toka sledi inverznemu kvadratnemu zakonu, tj. pri podvojitvi oddaljenosti jakost pade za 6 dB. S povečevanjem oddaljenosti od

izvora se povečuje površina valovne fronte, to pa neposredno vpliva na gostoto zvočnega energijskega toka j in s tem na energijski tok P (2.4, 2.5).

- linijski izvori (angl. line sources): npr. prometna cesta; jakost pade za 3 dB pri podvojitvi razdalje. Linijski izvor oddaja osno simetrično valovanje, kjer so valovne fronte koaksialni valjasti plašči. Površina valovnih front narašča premo sorazmerno z oddaljenostjo od linijskega izvora.
- ravninski izvori (angl. area sources): npr. množica zvočnikov v isti ravnini; če zanemarimo absorpcijo valovanja v prostoru, je jakost energijskega toka konstantna. Energijski tok, ki ga izvor oddaja z valovanjem, se razporeja po valovnih frontah, ki se jim z oddaljevanjem površina ne spreminja.

Pri obravnavi izvorov in ponorov zvoka se srečamo še s pojmom smernosti oz. smerne odvisnosti, ki opisuje odvisnost oddane zvočne energije ali občutljivost sprejemnika od smeri.

Ko zvočno valovanje zadane ob ovojo oz. mejo kjer se spremenijo lastnosti snovi, se v splošnem deloma odbije, deloma pa preide skozi mejo in se obenem lomi. Poleg omenjenih dveh pojavov se v primeru, ko valovanje naleti na ovojo srečujemo še z uklonom in interferenco. V primeru zvočnega valovanja so zakonitosti enake kot veljajo tudi za ostale oblike valovanja [28, 35, 49].

Odbojnost oz. absorpcijske lastnosti na meji dveh snovi opisujemo z impedanco. Zvočno valovanje se na meji dveh snovi tem močnejše odbije, čim bolj se akustični impedanci snovi na obeh straneh meje razlikujeta. Karakteristična impedanca snovi je [28]:

$$z = \rho c \quad (2.7)$$

V splošnem pa lahko definiramo specifično impedanco v poljubni točki prostora kot kvocient tlaka in hitrosti:

$$Z(x) = \frac{\mathcal{F}\{p(x)\}}{\mathcal{F}\{v(x)\}} \quad (2.8)$$

pri čemer \mathcal{F} označuje Fourierovo transformacijo.

Poznavanje osnovnih fizikalnih pojavov in oznak v zvezi z zvočnim valovanjem nam bo v nadaljevanju v pomoč pri razumevanju akustike prostora in delovanja človekovega organa za sluh. Slednje je še posebej pomembno v zvezi z uhljem, ki ga lahko primerjamo z anteno, saj filtrira in usmerja zvočno valovanje v sluhovod ter s tem vpliva na človekovo sposobnost lokalizacije izvorov zvoka. Zvočno valovanje in modifikacije, ki so posledica fizikalnih zakonitosti lahko opišemo z linearnim modelom, tj. filtrom, kar se uporablja npr. za sintezo prostorskih zvokov [27, 30, 37, 56].

2.2 Človekovo uho

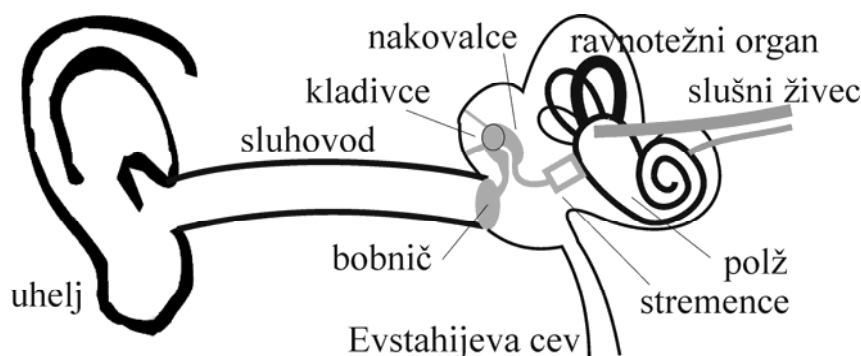
Človeški organ sluha je uho, ki ga sestavljajo zunanje (uhelj, zunanji sluhovod in bobnič), srednje (ušesne koščice – kladivce, nakovalce in stremence) in notranje uho (polkrožni kanali, polž in slušni živec). Zgradbo ušesa prikazuje slika 2.1.

Uhelj prestreže zvočno valovanje, ki se širi iz izvora zvoka in ga usmeri v zunanji sluhovod. Uhelj zaradi svoje značilne oblike deluje kot smerno odvisen filter [4, 5, 21]. Omenjena smernost je še posebej očitna npr. pri zvokih, ki prihajajo izza glave poslušalca. Zaradi uhlja se to valovanje ukloni, to pa povzroči slabljenje visokih frekvenc. Količina zvočne energije, ki jo prestreže uhelj je odvisna od velikosti uhlja. Glede na velikost in obliko so faktorji dobitka med različnimi uhlji vse do 5 dB [5, 28]. Večina sesalcev lahko s posebnimi mišicami giblje uhelj na vse strani in s tem lovi zvočne valove. Človek ima te mišice zakrnele.

Zunanji sluhovod poenostavljeno obravnavamo kot resonator, ki je na enem koncu odprt (vstop v sluhovod), na drugem pa zaprt oz. zaključen z membrano (bobnič). Dolžina zunanjega sluhovoda je približno $L = 3$ cm, v njem se vzbudi četrto-valovna resonanca pri približno 3 kHz ($f_r = c/4L \approx 3000$ Hz). Zunanji sluhovod deluje ojačevalno med prvo in drugo resonančno frekvenco, tj. v območju 3 kHz – 6 kHz. Absolutno vrednost prenosne karakteristike $H(f)$ v frekvenčnem pasu med 3 kHz in 6 kHz med zunanjim svetom in bobničem lahko opišemo s preprostim modelom [28]:

$$|H(f)|^2 \approx \frac{1}{\alpha^2 + (f - f_r)^2} \quad (2.9)$$

Za konstanto α v enačbi 2.9 se uporabi polovična vrednost resonančne frekvence sluhovoda. Zaradi opisanih resonančnih lastnosti sluhovoda je občutljivost človekovega sluha najboljša prav v frekvenčnem pasu med 3 kHz in 4 kHz.



Slika 2.1: Shema ušesa – uhelj prestreže zvočno valovanje, to nadaljuje pot po sluhovodu do bobniča, nato se preko slušnih koščic (kladivce, nakovalce) prenese v polža (membrana polža deluje kot spektralni analizator) in od tod po slušnih živcih v možgane.

Bobnič oz. membrana zapira vhod v srednje uho, ki je majhna, komaj za grah velika votlina. V srednjem ušesu so tri drobne, med seboj gibljivo zvezane slušne koščice - kladivce (lat. *incus*), nakovalce (lat. *malleus*) in stremence (lat. *stapes*). Kladivce je pripeto na bobnič, stremence pa na odprtino, ki je v koščeni steni med srednjim in notranjim ušesom. Zvočni valovi, ki pripotujejo po zunanjem sluhovodu zadenejo ob bobnič in ga zatresejo. Pri tem zanihajo slušne koščice. Stremence vzvalovi tekočino v koščeni votlinah v katerih je polž, ki ga prištevamo k notranjemu ušesu. Koščice pravzaprav predstavljajo prilagoditev impedance med zunanjim in notranjim ušesom. Zaradi velike razlike med akustičnima impedancama zunanjega sluhovoda (medij širjenja je zrak) in s tekočino (perilimfa in endolimfa) napolnjenih votlinic polža, bi se brez posredovanja koščic zvočno valovanje skoraj popolnoma odbilo.

Notranje uho je najbolj zapleteno zgrajeni del ušesa, zato se imenuje tudi blodišče ali labirint. Čutilo za sluh je zavita cevka, imenovana polž, ostali deli notranjega ušesa so sprejemniki za dražljaje ravnotežja in gibanja v prostoru. V notranjosti polža so trije zaviti kanali. Bazilarna membrana razdeli polža na dva ločena kanala in sicer sta to scala vestibuli in scala tympani. Scala vestibuli je z Reissnerjevo membrano še enkrat razdeljena na dva kanala in tako imamo poleg scale vestibuli še kohlearni kanal (*scala media*). Kanali so napolnjeni z dvema različnima tekočinama – perilimfa (*scala vestibuli* in *scala tympani*) in endolimfa (*scala media*). Polžji kanal je zavit okoli živca, ki se imenuje modiolus, v njem pa se nahajajo sistemi za kanale, čez katere gredo vlakna slušnega živca. Polž oz. notranji slušni kanal je v povprečju dolg 35 mm, je okroglega preseka in se proti vrhu zožuje. Na spodnji steni polža so razporejene slušne čutnice, ki imajo na vrhu šop migetalk. Stremence je neposredno povezano s scalo vestibuli preko membrane, ki niha ob spremembi pritiska na membrani bobniča. Ob tem udarjajo slušne čutnice ob mrenico, ki je nad njimi. Ob teh udarcih se vzdražijo čutnice. Čutnice so povezane z živčnimi vlakni, ki vodijo do možganov.

S stališča obdelave signalov notranje uho deluje kot spektralni analizator. Vlogo spektralnega analizatorja opravlja bazilarna membrana, ki je na spodnjem koncu polža ozka in toga, na vrhu polža pa velika in mehka. Njena skupna dolžina je približno 35 mm, njena širina pa je precej spremenljiva od osebe do osebe, zaradi tega v literaturi velikokrat zasledimo različne podatke. Zaradi spreminjajoče dolžine in širine bazilarne membrane se spreminja tudi njena frekvenčna karakteristika oz. odziv. Vsaka točka vzdolž membrane ima neko svojo frekvenco, pri kateri najbolj zaniha kot da bi imela svojo resonančno frekvenco na vsakem odseku. Posamezne dele bazilarne membrane modeliramo kot delce mase m , ki nihajo vpeti na vzmet s konstanto k . Hitrost posameznega dela membrane je sorazmerna s produktom zvočnega tlaka in prenosne funkcije H tega dela membrane.

Frekvenca, na katero se ujame del membrane se imenuje karakteristična frekvenca [28]:

$$f_c(x) = \frac{1}{2\pi} \sqrt{\frac{k(x)}{m(x)}} \quad (2.10)$$

V zgornji enačbi x pomeni lokacijo dela membrane na katerega se nanaša karakteristična frekvenca. Ko val potuje od spodnjega dela polža proti vrhu, se postopoma absorbirajo vse frekvenčne komponente vala, tako da odbojev valovanja in stoječih valov ni. Zvočni tlak frekvenčne komponente $p(x, f)$ pada eksponentno glede na oddaljenost x od začetka polža [28]:

$$p(x, f) = p_0(f) \cdot \Re \left\{ e^{-j \frac{2\pi f x}{c}} \right\} \quad (2.11)$$

V enačbi (2.11) pomeni $p_0(f)$ zvočni tlak na začetku polža in c je hitrost širjenja valovanja, ki v polževi tekočini znaša približno 1600 m/s.

Glavni del bazilarne membrane so lasne celice, ki so oživčene z živčnimi končiči. Živčni končiči sestavljajo slušni ali avditorni živec, ki prenaša podatke o vzburjanju notranjega ušesa v možgane. Proženje končičev je odvisno od gibanja lasnih celic, na katere so pritrjeni končiči. Ko se bazilarna membrana upogne v smeri kanala scala vestibuli, se dlačice na različne načine deformirajo, kar ima za posledico generiranje električnih signalov ali elektricitov, ki dražijo slušni živec. Ko se bazilarna membrana upogne v smeri kanala scala tympani, pa se dlačice ne deformirajo in ne generirajo električnih signalov. V mirovnem stanju je električni potencial lasne celice -50 mV, ko pa se zaradi upogibanja membrane generira električna napetost se potencial lasne celice približa 0 V. Vidimo, da v primeru tišine lasne celice pošiljajo enosmerno napetost, ko se pojavi zvoka pa signal pulzirajoče oblike.

2.3 Opis prostora

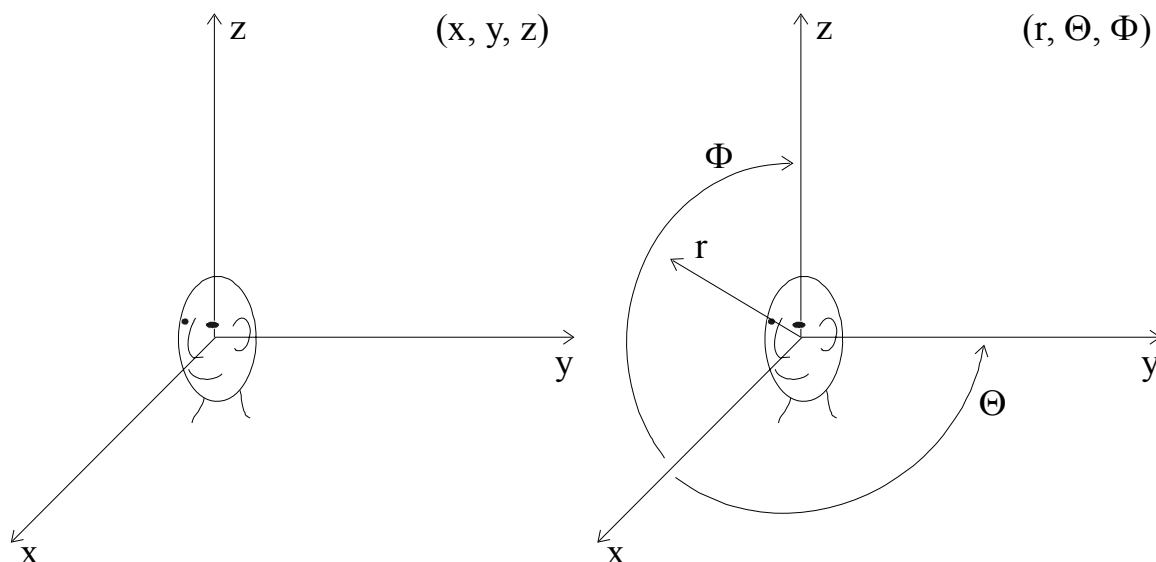
Za opis prostorskega zvoka potrebujemo tudi opis prostora. Prostor opišemo s pomočjo koordinatnega sistema. Najsplošnejši je kartezični koordinatni sistem, glede na obravnavani problem pa uporabljamo tudi druge oblike koordinatnih sistemov (cilindrični, krogelni ipd.). V našem primeru se najbolje obnese krogelni koordinatni sistem, ker glavo najlažje aproksimiramo kot kroglo (slika 2.2). V nadaljevanju sledi natančnejša definicija krogelnega koordinatnega sistema in pretvorbe med kartezičnim in krogelnim koordinatnim sistemom.

V kartezičnem koordinatnem sistemu imamo tri med seboj ortogonalne smeri: x , y in z . Razpon osi v kartezičnem koordinatnem sistemu:

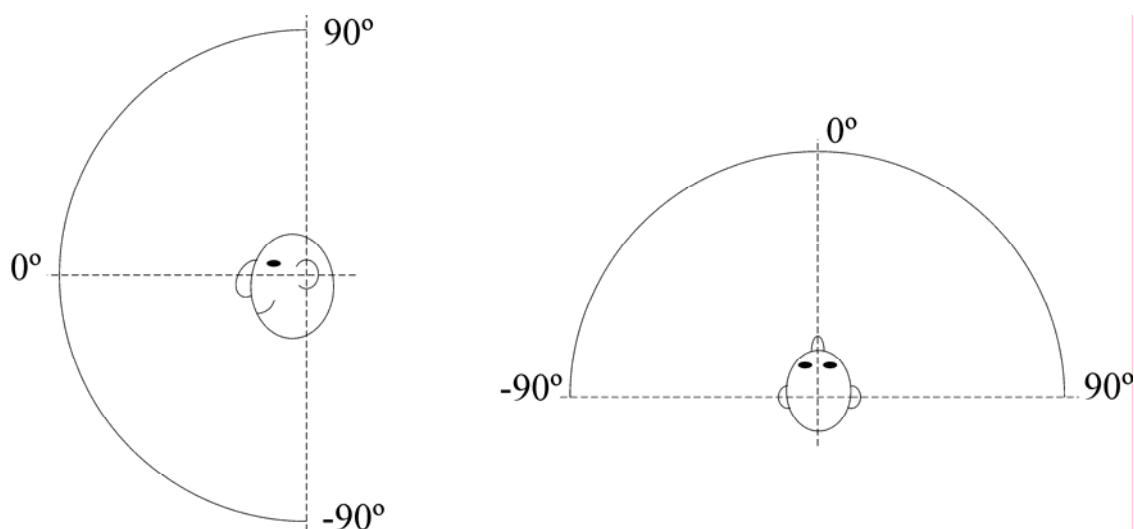
$$-\infty < x < \infty; \quad -\infty < y < \infty; \quad -\infty < z < \infty \quad (2.12)$$

V krogelnem koordinatnem sistemu določimo točko v prostoru z razdaljo od koordinatnega izhodišča (r) ter smerjo, ki jo določata kota Θ in Φ . Razpon osi v krogelnem koordinatnem sistemu:

$$0 \leq r < \infty; \quad 0^\circ \leq \Theta < 360^\circ; \quad -90^\circ \leq \Phi \leq 90^\circ \quad (2.13)$$



Slika 2.2: Oznaka smeri v kartezičnem (levo) in krogelnem (desno) koordinatnem sistemu. Smeri v krogelnem koordinatnem sistemu poimenujemo razdalja (r), azimut (Θ) in elevacija (Φ).



Slika 2.3: Elevacija (levo) in azimut (desno) v krogelnem koordinatnem sistemu glede na poslušalca oz. njegovo glavo. Središče koordinatnega sistema je v razpolovišču daljice, ki povezuje levo in desno uho poslušalca. Ravnino kjer je azimut 0° poimenujemo sredinska (sagitalna) ravnina, ravnino z elevacijo 0° pa horizontalna ravnina.

Pretvorba med kartezičnim in krogelnim koordinatnim sistemom:

$$\begin{aligned}x &= r \cdot \cos \Theta \cdot \cos \Phi \\y &= r \cdot \sin \Theta \cdot \cos \Phi \\z &= r \cdot \sin \Phi\end{aligned}\tag{2.14}$$

$$\begin{aligned}r &= \sqrt{x^2 + y^2 + z^2} \\ \Theta &= \operatorname{arctg} \frac{y}{x} + n\pi; n \in \mathbb{Z} \\ \Phi &= \arcsin \frac{z}{r} = \operatorname{arctg} \frac{z}{\sqrt{x^2 + y^2}}\end{aligned}\tag{2.15}$$

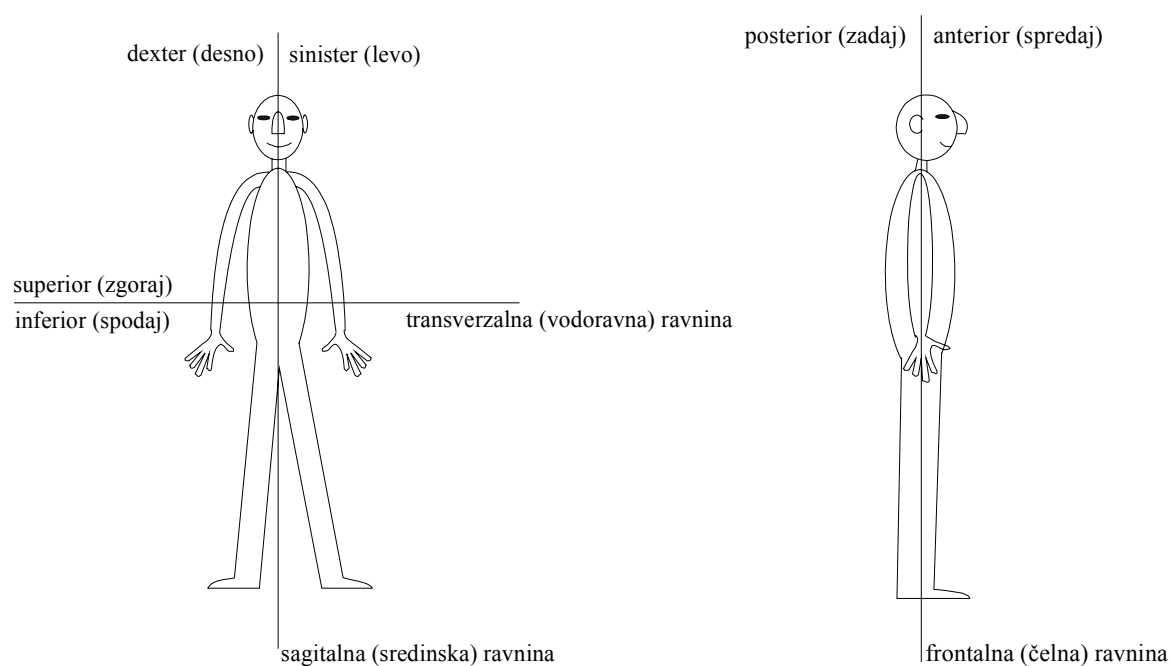
Za potrebe akustike (veja fizike, ki se ukvarja z zvokom in z njim povezanimi fizikalnimi pojavi) oz. psihoakustike (disciplina, ki preučuje subjektivno človeško zaznavanje zvoka) poimenujemo smer oz. kot, ki ga označujemo s Θ azimut in kot Φ elevacija. Slika 2.3 natančneje prikazuje pomen pojmov azimut in elevacija.

V psihoakustiki je ključnega pomena človek, zato potrebujemo še opis njegovega telesa. Človeško telo opisujemo v pokončni drži. Glede na aksialni skelet, ki ga predstavlja hrbtenica uporabljamo pri opisovanju človeškega telesa orientacijske ravnine oz. prereze, ki si jih zamišljamo položene skozi telo. Orientacijske ravnine (slika 2.4):

- a) transverzalna (vodoravna, horizontalna) ravnina (angl. transverse plane): prereže telo počez in ga razdeli na dva dela:
 - superior – zgornji,
 - inferior – spodnji;
- b) frontalna (čelna) ravnina (angl. frontal, coronal plane): poteka vzporedno s čelom, opredelitve glede na frontalno ravnino so:
 - anterior - sprednji, kar leži spredaj,
 - posterior - zadnji, kar leži zadaj,
 ali:
 - ventralis – trebušni, na trebušni strani,
 - dorsalis – hrbtni, na hrbtni strani;
- c) sagitalna (sredinska) ravnina (angl. saggital, midsaggital, median plane) poteka vzporedno s središčnico ali telesno osjo, opredelitve glede na sagitalno ravnino so:
 - dexter – desni,
 - sinister – levi;
- d) za potrebe psihoakustike dodatno definiramo še horizontalno ravnino (angl. horizontal plane), ki ustreza ravnini x - y v koordinatnem sistemu na sliki 2.2. Horizontalna ravnina je torej vzporedna transverzalni ravnini in je položena skozi središče človekovih uhljev.

Smeri pa poimenujemo (slika 2.4):

- cranialis - kar leži v smeri proti glavi,
- caudalis - kar leži v smeri proti trtici;
- medialis – kar leži bliže sredini,
- lateralis – kar leži dlje od sredine oz. bolj ob strani,
- ipsilateralis – kar leži na isti strani,
- contralateralis – kar leži na drugi strani,
- bilateralis – kar leži na obeh straneh.



Slika 2.4: Orientacijske ravnine in podprostori za opisovanje človekovega telesa. Izrazi so zapisani v latinščini, v oklepajih so slovenski izrazi.

2.4 Akustika prostora

Pomemben vpliv na to kar poslušalec sliši ima tudi okolica poslušalca. Zvok, ki ga oddaja zvočni izvor se odbija od mejnih ploskev prostora in od predmetov v okolici poslušalca in izvora zvoka. Pri večjih razdaljah med izvorom zvoka in poslušalcem ima vpliv tudi gibanje zračnih tokov, porazdelitev temperature ipd. [15, 35].

Poslušalca najprej doseže direktni zvočni val, nato pa še odbiti valovi. Če odbiti valovi dosežejo poslušalca z manj kot 35 ms zakasnitve glede na direktni val, jih poslušalec zazna skupaj z direktnim zvokom kot enoten zvok. Pri zakasnitvah večjih od 35 ms pa uho zazna direktni in odbiti val ločeno. Zaznavanje direktnega in odbitih valov kot enotnega zvoka se človek nauči po približno 16 tednih življenja [4]. Pri lokalizaciji

zvočnih izvorov imajo odboji od okoliških ovir pomembno nalogo, saj omogočajo poslušalcu ugotoviti oddaljenost zvočnega izvora [4, 5, 81].

V prostoru z odbojnimi površinami se vzpostavi stojno valovanje, ki ima lahko zelo zapleteno porazdelitev. Pomembna lastnost prostora so resonančne frekvence prostora, ki povzročajo ojačanje zvoka resonančnih frekvenc in čas odmeva oz. reverberacije RT60 (Reverberation Time 60 dB, tj. čas v katerem pade prvotna zvočna energija na eno milijoninko začetne vrednosti) [15, 16]. Ta dva parametra sta pomembna predvsem pri reprodukciji in snemanju zvoka v prostoru. Najbolj kritično je seveda predvajanje posnetega zvoka, saj posnetki vsebujejo lastnosti prostora v katerem so nastali, nanje pa vpliva tudi prostor v katerem se predvajajo.

Poseben izziv predstavlja generiranje tri dimenzionalnega zvoka za potrebe t.i. navidezne resničnosti (angl. virtual reality) in obogatene resničnosti (angl. augmented reality), kjer mora biti poleg 3-D slike tudi zvok tri dimenzionalen. Če želimo zvok v navidezni ali obogateni resničnosti predvajati preko slušalk, mora vsebovati značilnosti, ki so sicer v realnih razmerah posledica lastnosti prostora. Lastnosti prostora lahko ugotavljamo s pomočjo odziva na enotin impulz (Room Impulse Response – RIR), ki pa je poleg prostora odvisen tudi od položajev generatorja enotnega impulza in sprejemnika [52].

Zaradi nepredvidljivosti porazdelitve zvočnega valovanja v prostorih z odbojnimi površinami za meritve v akustiki in psihoakustiki uporabljamo prostore s stabilnimi razmerami, ki zadostujejo temeljnima pogojema meritev, tj. ponovljivosti in splošni veljavnosti. Pogojema zadostimo z idealnim odprtim prostorom, ki omogoča zvočnemu valovanju svobodno razširjanje. V praksi takega prostora ne moremo najti, saj na zvočno valovanje vplivajo vremenske razmere, spremembe temperature, vlage ipd. Prostor, ki se po lastnostih najbolj približa lastnostim idealnega odprtega prostora je t.i. gluha soba. Najpomembnejši lastnosti gluhe sobe sta nizek nivo zunanlega hrupa (tj. dobra zvočna izolacija) in zadušitev odbojev od sten sobe. Z ustrezno zadušitvijo odbojev je izpolnjena zahteva za širjenje valovanja točkastega izvora po inverznem kvadratnem zakonu.

Meritve, ki smo jih izvajali v gluhi sobi, so nastale v gluhi sobi Fakultete za elektrotehniko. Da se preprečijo stojni valovi, oblika te sobe ni pravokotna. Višina konstrukcije je od 3.25 m do 3.75 m, tla imajo dimenzije 6 m x 6,5 m. Soba je podprta z gumijastimi podporami, katerih lastna frekvenca znaša okoli 7 Hz, nosilna konstrukcija je iz lahkega profiliranega železa premera 10 cm, na tleh pa je nosilna jeklena mreža. Debelina absorpcijskega materiala je 75 cm [15].

3 PSIHOAKUSTIKA IN PROCES UČENJA

Psihoakustika je veda, ki preučuje subjektivno človeško zaznavanje zvoka. Je podzvrst psihofizike, ki preučuje odzive človeških čutil (zaznavanje zvoka, svetlobe, dotika in ostalih pojavov). Psihoakustika se torej ukvarja z vprašanji, ki so povezana z glasnostjo zvoka, maskiranjem, kritičnimi pasovi, zaznavnimi razlikami med zvoki ipd.

Osnova človeškega dojetanja oz. interpretacije občutkov, ki jih dobimo prek čutil iz zunanjega sveta je Weber-Fechnerjev zakon [28], ki pravi, da pri spremembi fizikalne veličine x za vrednost dx človek zazna spremembo dojetanja občutka y za vrednost dy :

$$dy \propto \frac{dx}{x} \quad (3.1)$$

Od tod sledi, da je človekova občutljivost na zunanje dražljaje logaritemska:

$$y(x) = \log(x) + C \quad (3.2)$$

Našega dojetanja sveta, interpretacije signalov, ki jih možgani dobijo s pomočjo čutil in končno tudi vsega kar vemo in znamo pa smo se naučili. To sklepamo na podlagi dejstva, da so možgani novorojenega otroka tako rekoč prazni, nato pa se postopoma naučimo vsega kar potrebujemo za življenje [6, 19]. Človeški možgani so torej izjemen procesor, ki se je sposoben prilagoditi skoraj vsaki zahtevi [12, 14, 29, 32]. To velja tudi za dojetanje prostorskega zvoka, možgani novorojenega otroka npr. še ne znajo ugotoviti, da direktni in zakasneni zvočni val pravzaprav spadata skupaj, tega se naučijo šele z nadaljnjim razvojem [4].

V tem poglavju se ukvarjamo s človekovim dojetanjem zvoka, delovanjem človekovih možganov in možnostmi kako prilagoditi možgane na spremenjene koncepte predstavitve sveta.

3.1 Glasnost zvoka

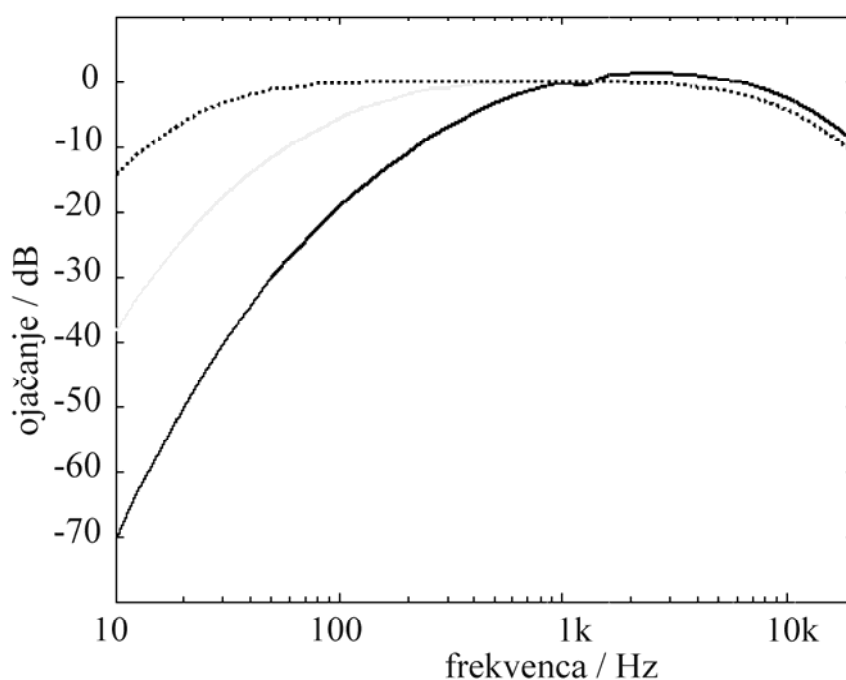
V prejšnjih poglavjih smo napisali, da človeško uho zazna zvok, če je gostota zvočnega toka oz. jakost zvoka ob ušesu večja od 10^{-12} W/m^2 in manjša od 1 W/m^2 . Razpon gostote zvočnega toka je zelo velik in je možen zaradi logaritemske občutljivosti ušesa. Občutljivost ušesa pa ni odvisna le od jakosti zvoka temveč predvsem od frekvence zvoka. Človeško uho je najbolj občutljivo za zvoke s frekvencami 3 kHz – 4 kHz, kar pomeni, da je za dan fiziološki učinek pri tej frekvenci potrebna najmanjša jakost zvoka. Omenjeni fiziološki učinek imenujemo glasnost. Jakost zvoka je fizikalna lastnost,

medtem ko je glasnost razumska (psihična) interpretacija, ki je zaradi svoje subjektivne narave ne moremo natančno meriti [15].

V fiziki je glasnost (J) normirana glede na najmanjšo zaznavno gostoto zvočnega toka (j_0) ali tlaka (p_0) in se meri v decibelih (dB) ter ne upošteva frekvenčno odvisne občutljivosti ušesa [35]. To veličino (J) imenujemo tudi nivo zvočnega tlaka (angl. sound pressure level – *SPL*) [15]:

$$J = 10 \log \frac{J}{j_0} = 20 \log \frac{P}{p_0} \quad [\text{dB}] \quad (3.3)$$

Frekvenčna odvisnost se pri meritvah nivoja zvočnega tlaka (*SPL*) upošteva z uporabo različnih utežnih koeficientov oz. krivulj *A*, *B*, *C* in »Flat«. Vhodni signal je pri meritvi z utežjo *A* (*A*-weighted *SPL*, IEC 179) pred izvedbo meritve filtriran s filtrom, ki posnema občutljivost človeškega ušesa (več o tem v nadaljevanju). Pri merjenju z utežjo »Flat« gre za neposredno meritev fizikalne veličine. Karakteristiki »Flat« je zelo podobna utež *C*, utež *B* pa je nekakšna vmesna varianta med *A* in *B* kot prikazuje slika 3.1.

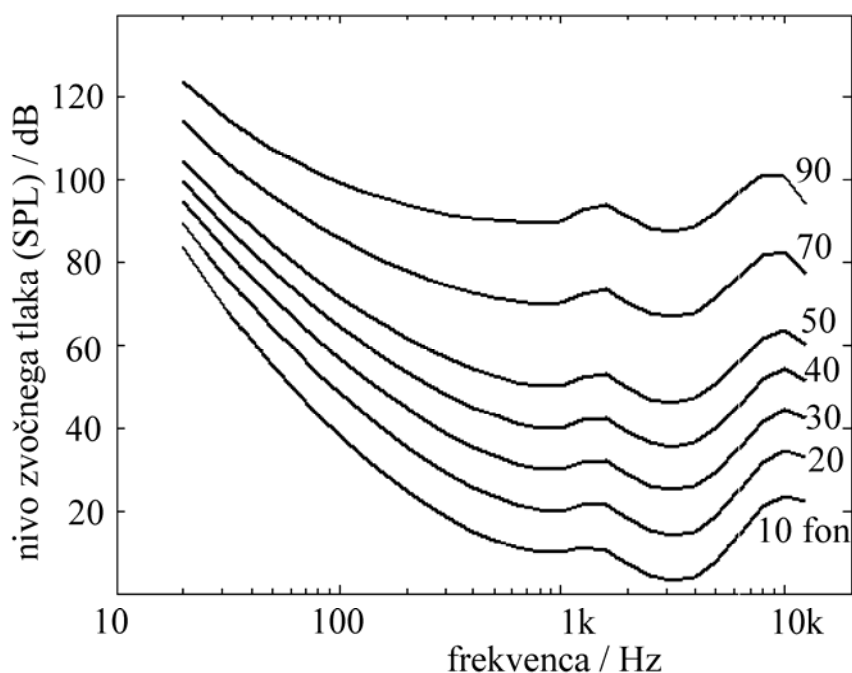


Slika 3.1: Potek uteži *A* (črna), *B* (siva) in *C* (pikčasta) za merjenje nivoja zvočnega tlaka (*SPL*) oz. glasnosti. Utež *A* posnema odvisnost občutljivosti človekovega ušesa od frekvence, utež *C* ima raven potek preko več dekad, zato je uporabna za merjenje glasnosti zvokov velikih jakosti, utež *B*, ki predstavlja nekakšno srednjo varianto med *A* in *C* se redko uporablja. Krivulje smo narisali po podatkih iz standarda IEC 179 [80].

Za izražanje frekvenčne odvisnosti občutljivosti ušesa uporabljamo enoto glasnosti zvoka – fon (angl. phon), ki je definirana pri frekvenci 1 kHz. Ker je občutljivost ušesa odvisna od frekvence, se dB in foni ujemajo le pri 1 kHz, sicer pa je glasnost zvoka podana z množico krivulj, kot jih podaja slika 3.2. Dojemanje glasnosti (LL_x) lahko analitično zapišemo kot produkt enačbe 3.3 in utežne krivulje $\alpha(f)$, ki jo najdemo npr. v [28]. Povezava med SPL in foni je definirana v standardu ISO 226 (normal equal-loudness-level contours).

Žal pa Webrov zakon ne ustreza povsem človeškemu dojemanju glasnosti zvoka, to pomeni, da npr. zvoka glasnosti 40 fonov ne dojema dvakrat glasneje od zvoka glasnosti 20 fonov. Izkazalo se je, da dojemanje glasnosti natančneje opišemo s kubičnim korenem energijskega toka zvoka. Tako dobimo novo mero glasnosti zvoka – son (angl. sone). Najboljša rešitev za ugotavljanje glasnosti v sonih je uporaba krivulj za ugotavljanje glasnosti v fonih (LL_x) in nato pretvorba v sone (L_x) s pomočjo naslednje enačbe:

$$L_x = 0.04610 \frac{LL_x}{30} \quad [\text{son}] \quad (3.4)$$



Slika 3.2: Krivulje enake glasnosti (equal loudness curves) prikazujejo frekv. odvisnost občutljivosti človekovega organa za sluh, ki je najbolj občutljiv med 3 kHz in 4 kHz. To pomeni, da pri nizkih ali visokih frekvencah potrebujemo večjo jakost zvoka, da izzovemo enak občutek glasnosti, tj. enako vrednost v fonih. Krivulje enake glasnosti sta prva izmerila Fletcher in Munson, kasneje pa še Robinson in Dadson [4, 28]. Na osnovi teh meritev je nastal standard ISO 226 [4, 80], od koder smo tudi črpali podatke za sliko.

Tabela 3.1 natančneje podaja krivuljo enake glasnosti za 0 fon-ov. Krivulja je vzeta iz standarda ISO 226, ki podaja povezavo med glasnostjo v fonih (L) in nivojem SPL (J) za frekvence od 20 Hz do 12.5 kHz. Standard ISO 226 [4] izhaja iz krivulj oz. meritev, ki so jih izvajali Fletcher in Munson ter Robinson in Dadson [4, 15, 28].

Kot smo videli, za glasnost uporabljamo več različnih enot – dB, fon in son. Če je glasnost podana v dB to pomeni, da gre za neposredno meritev jakosti zvoka (SPL), kadar je uporabljena katera od utežnih krivulj pa je to v enoti označeno, npr. dB(A). Kadar je glasnost izražena v fonih ali sonih pa vemo, da je bila frekvenčna odvisnost občutljivosti človeškega ušesa že upoštevana..

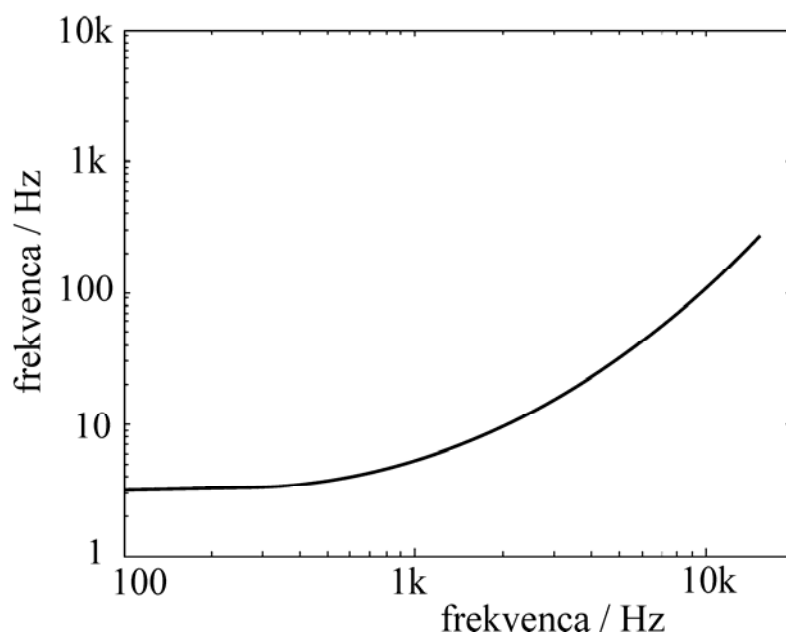
f / Hz	SPL / dB	f / Hz	SPL / dB
20	76.6	630	0.3
25	65.6	800	-0.3
32	55.1	1000	0.0
40	45.5	1250	1.0
50	37.6	1600	-1.2
63	30.9	2000	-4.1
80	25.0	2500	-7.0
100	20.5	3150	-9.0
125	16.6	4000	-8.5
160	13.1	5000	-4.4
200	10.1	6300	3.3
250	7.5	8000	9.8
315	5.1	10000	10.5
400	3.1	12500	8.4
500	1.5		

Tabela 3.1: Krivulje enake glasnosti (equal loudness curves), vrednosti SPL potrebne za glasnost 0 fon (vir: ISO 226 [4, 80]).

3.2 Maskiranje in najmanjše zaznavne razlike glasnosti ter frekvence

Maskiranje je pojav, ki se zgodi kadar hkrati ali zelo hitro eden za drugim nastopita dva zvoka. Maskirani zvok postane neslišen zaradi maskirnega zvoka ali maskerja. Maskiranje nastopi tudi, če je en izmed dveh zvokov bistveno glasnejši kot drugi ali kadar se dva zvoka podobne glasnosti le malo razlikujeta po frekvenci. Maskiranje lahko opišemo s pogojem, da mora biti razlika glasnosti vsote maskerja in maskiranega zvoka ter glasnosti maskerja samega manjša od 0.41 dB (0.41 dB predstavlja mejo zaznavanja glasnosti, tj. če se glasnost zvoka spremeni za 0.41 dB ali več, človek zazna spremembo glasnosti; angl. *JND* – Just Noticeable Difference in loudness) za vse frekvence v slišnem področju [47].

Kadar je razlika frekvenc dveh podobno glasnih tonov, ki ju predvajamo enega za drugim premajhna, je slišati le en ton. Minimalno frekvenčno razliko, ki še zagotavlja da slišimo dva tona poimenujemo najmanjša razločljiva frekvenčna razlika (angl. *JND* – Just Noticeable Difference in frequency). Najmanjša razločljiva frekvenčna razlika je odvisna od frekvence tonov in se z naraščanjem frekvence tonov povečuje, prav tako pa je natančnost človekovega ušesa slabša tudi pri nizkih frekvencah (<200 Hz). Vzroki za pojav maskiranja so povezani z bazilarno membrano, natančneje s pojavi, ki so posledica vzbujanja in nihanja membrane oz. delovanja tega »človekovega spektralnega analizatorja«. Slika 3.3 prikazuje najmanjšo razločljivo frekvenčno razliko v odvisnosti od frekvence tona [83]. Velja, da je najmanjša razločljiva frekvenčna razlika približno 30 krat manjša od pasovne širine ustreznega kritičnega pasu [83] o katerem bomo razpravljali v naslednjem razdelku.



Slika 3.3: Minimalna zaznavna razlika v frekvenci (*JND* in frequency) pomeni minimalno razliko frekvenc dveh tonov, da ju (povprečno) človeško uho še zazna kot dva tona. Odvisnost je prvi izmeril Zwicker s sodelavci [83], kar je tudi vir podatkov za sliko.

3.3 Kritični pasovi

Obstoj kritičnih pasov je posledica delovanja bazilarne membrane, ki deluje kot spektralni analizator. Zvočno valovanje v polžu vzbuja posamezne dele bazilarne membrane, to pa lahko modeliramo kot nabor pasovno prepustnih filtrov s pasovno širino enako kritičnim pasovom [83, 84].

Če zvok vsebuje dva tona s frekvencama f_1 in f_2 , ki pripadata istemu kritičnemu pasu, oba tona vzbujata isti del bazilarne membrane. Glasnost, ki jo zazna poslušalec je v tem primeru manjša kot če bi tona vzbujala različna dela bazilarne membrane, oz. bi bila tona iz dveh različnih kritičnih pasov. Podobno velja tudi, če je zvok pasovno omejen šum. Dokler vse frekvenčne komponente šuma ležijo znotraj enega kritičnega pasu in je pasovna širina šuma manjša ali enaka pasovni širini tega kritičnega pasu ter gostota zvočnega toka šuma vedno enaka, poslušalec ne zazna razlike v glasnosti zvoka ne glede na dejansko pasovno širino šuma. V skrajnem primeru poslušalec ne zazna razlike v glasnosti med pasovno omejenim šumom, ki sovpada z določenim kritičnim pasom in tonom, katerega frekvenca leži v istem kritičnem pasu, pri čemer sta gostoti zvočnih tokov obeh signalov enaki.

zap. številka kritičnega pasu	f_{sp} / Hz	f_{zg} / Hz	B / Hz
1	20	110	90
2	110	200	90
3	200	295	95
4	295	395	100
5	395	503	108
6	503	625	120
7	625	755	130
8	755	900	145
9	900	1060	160
10	1060	1250	190
11	1250	1460	210
12	1460	1700	240
13	1700	1970	270
14	1970	2290	320
15	2290	2670	380
16	2670	3120	450
17	3120	3680	560
18	3680	4360	680
19	4360	5200	840
20	5200	6200	1000
21	6200	7500	1300
22	7500	9300	1800
23	9300	11700	2400
24	11700	15000	3300
25	15000	19600	4600

Tabela 3.2: Pasovna širina ter spodnja in zgornja meja kritičnih pasov (vir: [83]). Kritični pasovi so posledica delovanja bazilarne membrane, ki deluje kot spektralni analizator. Velja, da dva tona katerih frekvenci sovpadata z istim kritičnim pasom, poslušalec zazna manjšo glasnost kot bi jo, če bi frekvenci tonov pripadali različnima kritičnima pasovoma.

Tabela 3.2 podaja kritične pasove (spodnjo in zgornjo mejno frekvenco ter pasovno širino). Velja, da je širina kritičnega pasu približno 30 krat večja kot najmanjša razločljiva frekvenčna razlika o kateri smo razpravljali v prejšnjem razdelku [83].

3.4 Prenosne funkcije vpliva glave in telesa HRTF (Head Related Transfer Functions)

Pri preučevanju mehanizmov človekovega zaznavanja prostorskega zvoka imata impulzni odziv, oz. prenosna funkcija poti med izvorom zvoka in začetkom poslušalčevega sluhovoda ali prenosna funkcija poti med izvorom zvoka in poslušalčevim bobničem, poseben pomen. Ko zaradi ponovljivosti meritve izvajamo v gluhi sobi, na karakteristiko prenosne poti pomembno vplivata le še poslušalčeva glava in telo. Zato govorimo o t.i. prenosnih funkcijah vpliva glave in telesa oz. HRTF (Head Related Transfer Functions). Impulzni odziv vpliva glave in telesa imenujemo HRIR (Head Related Impulse Response), HRTF pa je Fourierov transform impulznega odziva [11].

HRTF so zanimive tudi s stališča reprodukcije zvoka s slušalkami. Če s pomočjo HRTF konstruiramo filter in nato s tem filtrom filtriramo poljuben zvok, ki ga preko slušalk predvajamo poslušalcu, se poslušalcu zdi kot da zvok prihaja iz smeri, ki ji ustreza HRTF s pomočjo katere smo konstruirali filter. HRTF, ki so izmerjene v dobri gluhi sobi, ne vsebujejo odbojev od sten, kar pomeni pomanjkanje informacije o oddaljenosti izvora zvoka. Preprosto filtriranje zvoka s filtrom HRTF daje poslušalcu občutek, da se izvor zvoka nahaja v njegovi glavi, kar pa za poslušalca ni najbolj prijetno. Reševanje tega problema, tj. prestavljanje navideznega izvora zvoka iz poslušalčeve glave, se imenuje eksternalizacija (angl. externalization) [26]. Eksternalizacijo je mogoče izvajati matematično, na osnovi fizikalnih zakonov o širjenju zvoka [4, 5].

HRTF so specifične za vsakega posameznika, nanje vpliva predvsem človekova konstitucija – oblika in dimenzije telesa, uhljev, glave ipd. Poleg telesnih karakteristik človeka je potek HRTF odvisen od položaja izvora zvoka glede na poslušalca. Načeloma bi bilo torej potrebno izmeriti odziv prenosne poti za vsak položaj posebej, vendar ima človek omejeno zmožnost ločevanja zvočnih izvorov, ki po azimutu znaša približno 2° , po elevaciji pa 6° [64, 65]. V gluhi sobi, kjer ni odbojev od sten, tudi razdalja med izvorom zvoka in poslušalcem delno izgubi pomen (odboji so večinoma izločeni, ostane le slabljenje zaradi razširjanja valovanja). Pri razdaljah med izvorom zvoka in poslušalcem manjših od 3 m se HRTF v isti smeri na različnih razdaljah razlikujejo v amplitudnem in faznem poteku, pri razdaljah nad 3 m pa je razlika v isti smeri zreducirana zgolj na razliko v jakosti amplitudnega dela spektra (skladno z inverznim kvadratnim zakonom), pri razdaljah nad 15 m pa se pojavi frekvenčno odvisno slabljenje, ki vpliva na višje frekvence [4].

V splošnem delimo HRTF na individualne in splošne (generalne). Individualne HRTF so funkcije, izmerjene na točno določeni osebi, s pojmom splošne funkcije pa mislimo na funkcije, izmerjene na doprsni lutki (npr. KEMAR – Knowles Electronic Manikin for Auditory Research [20]) ali funkcije, ki jih dobimo s tehnikami obdelave signalov [17, 24, 27, 37, 56, 58]. Oblika lutke je lahko povzeta po točno določeni osebi, običajno pa gre za lutko z dimenzijami, ki bi ustrezale neki povprečni osebi ali pa je lutka narejena celo kot zelo poenostavljen človeški model, ki ga je možno tudi matematično opisati. Kot bomo videli v nadaljevanju, je v lokalizaciji zvočnih izvorov pri reprodukciji zvoka s pomočjo slušalk bistvena razlika med splošnimi in individualnimi HRTF. Ta je v primeru uporabe individualnih HRTF bistveno boljša kot v primeru splošnih. Poleg tega omogoča uporaba individualnih HRTF absolutno lokalizacijo izvorov zvoka (tj. poslušalec lokalizira izvor zvoka), medtem ko je uporaba splošnih funkcij ne omogoča (poslušalec sicer lahko določi relativno pozicijo več izvorov, ni pa sposoben ugotoviti dejanske lokacije izvorov).

Kot smo že omenili, prideta pri merjenju HRTF v gluhi sobi do izraza predvsem horizontalna in vertikalna dimenzija (azimut in elevacija), zato se meritve izvajajo pri konstantni razdalji med izvorom zvoka in testno osebo [2, 20, 54, 66]. Smeri v katerih merimo HRTF lahko smiselno določimo z upoštevanjem človekovih sposobnosti ločevanja izvorov zvoka.

Merjenje HRTF je relativno dolgotrajen in zamuden postopek, še posebej, če želimo meriti individualne funkcije. Še zahtevnejša od natančne postavitve zvočnega izvora je naloga testne osebe, ki mora v času meritve držati glavo v čim bolj nespremenjenem položaju kar navkljub raznim pripomočkom za stabilizacijo ni preprosto.

Meritve individualnih HRTF smo izvedli tudi sami, vendar smo zaradi zahtevnosti merili le s tremi osebami [66]. Poleg lastnih HRTF smo v analizah uporabljali predvsem javno dostopne HRTF knjižnice CIPIC [2] ter funkcije knjižnic MIT Media Lab [20] in Itakura [66].

3.5 Učenje

Človeški možgani so izjemen procesor, ki se je sposoben prilagoditi skoraj vsaki zahtevi. To sklepamo na podlagi dejstva, da so možgani novorojenega otroka tako rekoč prazni in se pravzaprav vsega česar se zavedamo, kar znamo, vemo, uporabljamo naučimo [6, 19]. Kot pravi Johannes Muller (1801 – 1858) v »Law of Specific Nerve Energies«, se ljudje sveta zavedamo na podlagi impulzov v možganih za katere velja, da enaki impulzi, ki vzbudijo različne dele možganov povzročajo različne občutke, različni impulzi, ki vzbudijo iste dele možganov pa povzročijo enake občutke. Zavedanje občutkov in učenje

je posledica povezav med nevroni. V človekovih možganih je približno 100 milijard nevronov. Povezavam med njimi rečemo sinapse.

Učenje oz. vzpostavljanje povezav v možganih je najintenzivnejše takoj po rojstvu. Ker je učenje zahteven proces, je po mnenju razvojnih psihologov to vzrok, da dojenčki večino dneva prespijo. Po drugi strani pa velja tudi, da s stalno aktivnostjo možganov in telesno aktivnostjo povečamo zmožnost učenja [6].

Tako kot vsega ostalega, se ljudje naučimo tudi poslušanja oz. zaznavanja prostorskega zvoka [4, 6, 19]. Kljub temu, da senzorji (v primeru zvoka so to ušesa) že od vsega začetka dostavljajo možganom signale o zvokih, ti podatki človeku v zgodnji fazi življenja še ne pomenijo kaj dosti, ker še nima ustrezno razvitega sistema za obdelavo zvokov. Pri normalnem razvoju sluha in vida, se hitreje razvija del za obdelavo slike. Pri približno enem mesecu starosti naj bi otrok že ločil kateri predmeti so bližje in kateri dlje [19]. Povezovanje oz. medsebojno sodelovanje čutil igra pomembno vlogo pri dojetju sveta in razvoju različnih, še posebej motoričnih sposobnosti [29]. Razvoj sposobnosti zaznavanja smeri izvora zvoka je pri normalnem razvoju sluha in vida povezan tudi s sposobnostjo vidnega zaznavanja o čemer govori t.i. vizualna teorija [4]. Otrok naj bi postal popolnoma sposoben določiti izvor zvoka med drugim in četrtem mesecem starosti [19]. Pri lokalizaciji izvora zvoka je morda najtežji problem zaznavanje oddaljenosti izvora, ki poleg relativno preprostega zaznavanja smeri (časovne in amplitudne razlike med levim in desnim bobničem) vključuje še vrsto zakasnjenih zvočnih valov. Očitno je prav zaznavanje zakasnjenih odbojev zvoka kot nečesa, kar je del prvega vpadnega vala tega zvoka, ena težjih nalog – v povprečju naj bi se ljudje tega naučili po 16 tednih življenja [4].

Slika sveta, ki si jo ustvarimo v možganih se, tako kot v prvih mesecih življenja, lahko spreminja tudi kasneje. Vsak, ki nosi očala, se je srečal z izkušnjo, ko jih je prvič uporabil – pogled skozi očala prvi hip deluje nekoliko nenavadno in moteče, vendar običajno že nekaj dni zadostuje, da se možgani prilagodijo novi »optiki«. S fenomenom sposobnosti prilagajanja možganov na spremembe se je že v 19. stoletju začel ukvarjati Hermann von Helmholtz. S področja prilagajanja možganov na spremembe v vidnem zaznavanju so najbolj znani eksperimenti z optičnimi prizmi, ki so bili do danes v različnih oblikah izvedeni vsaj dva tisočkrat [12, 29]. V teh poskusih so testnim osebam nataknili očala, ki so npr. sliko obrnila na glavo, jo ukrivljala ali kako drugače popačila. Manj poznani so poskusi s področja akustike, kjer so uporabljali podobne metode kot v optiki, npr. zamenjevanje smeri izvora zvoka [29].

Za učenje je bistvena Helmholtzeva ugotovitev, ki pravi, da še število napačnih interpretacij neke konkretne situacije, ki se pojavi zaradi sprememb v zaznavanju (tj. ko si npr. natakne očala, ki obračajo sliko) zmanjšuje s številom ponovitev oz. soočenj s to

situacijo, kar je bilo kasneje potrjeno tudi v [18]. To z drugimi besedami pomeni, da lahko z učenjem možgane prilagodimo spremenjenim razmeram.

Hitrost učenja oz. hitrost prilagajanja možganov novim razmeram je odvisna zlasti od aktivnosti učenca. V literaturi [29] najdemo ugotovitve, da se možgani novim razmeram najhitreje prilagodijo, če v procesu učenja aktivno sodelujemo, kar pomeni, da poleg poslušanja ali gledanja tudi sami poskušamo določiti smer, obliko ali kako drugo karakteristiko v neznanih (spremenjenih) razmerah. Proces prilagajanja oz. stopnja prilagoditve možganov novim razmeram poimenujemo korelacija. Prilagoditev možganov novim razmeram pa pomeni, da je v primeru ponovne spremembe na stare razmere zopet potrebno učenje. Slikovit prikaz tega je primer, ko so testni osebi natakneli očala s takšnimi stekli, ki zamenjajo levo in desno stran. V povprečju so se možgani testne osebe prilagodili novim razmeram v 4 do 8 tednih, ko pa so sneli očala, je bilo potrebno ponovno prilagajanje na nekoč že znane razmere. Prilagajanje na prejšnje stanje imenujemo dekorelacija. Hitrost dekorelacije je poleg intenzivnosti učenja odvisna tudi od poprejšnje stopnje korelacije, tj. v kolikšni meri so se možgani že prilagodili spremenjenim razmeram [29].

Podoben primer iz akustike je poskus v katerem so testnim osebam spremenili obliko uhljev [32, 61, 62], nato pa so ugotavljali sposobnost lokalizacije zvočnih izvorov. V horizontalni smeri ni bilo očitnih odstopanj, v vertikalni smeri pa so ljudje z modificiranimi uhlji bistveno slabše ločili zvočne izvore kot ljudje z »normalnimi« uhlji.

Primer spremenjenih oz. na poseben način prirejenih razmer v akustiki je kodiranje elevacije v akustični sliki prostora projekta VoICE – Seeing with sound [44, 45], kjer je elevacija kodirana neposredno s frekvenco. Rezultati tega projekta kažejo, da se je uporabnikom uspelo naučiti prepoznavanja slike, kljub temu, da v prvem poskusu zagotovo ni moč slišati nič drugega kot zgolj nekakšno piskanje. Podoben primer kodiranja slike in barv z zvokom je projekt slepega podiplomskega študenta univerze Cornell [14], ZDA, ki s pomočjo zvoka bere oz. kodira vremenske karte.

Na lokalizacijo zvokov ima poseben vpliv vizualna izkušnja. Poskus, katerega opis najdemo v [4], govori o možnosti prevlade vizualne percepcije nad slušno. V tem poskusu so levo ali desno od poslušalca namestili zvočnik, zvok pa so poslušalcu predvajali preko slušalk iz nasprotne smeri kot je bil postavljen zvočnik. Poslušalec je poročal, da sliši zvok iz smeri, v kateri se je nahajal zvočnik. Ko je imel zavezane oči, pa je slišal zvok iz nasprotne smeri. Ko so poslušalcu zavezali oči šele po daljšem intervalu poslušanja z nezavezanimi očmi, je poslušalec še nekaj časa poročal, da sliši zvok v smeri zvočnika, šele nato je postopoma začel poročati, da sliši zvok iz smeri kot bi pričakovali glede na zvok iz slušalk. Z vplivom vizualne izkušnje na lokalizacijo zvoka se ukvarja t.i. vizualna teorija (angl. visual theory) [4].

Na osnovi predstavljenih dejstev je očitno, da se je človek sposoben naučiti poljubnega načina kodiranja elevacije v akustični sliki prostora. Ker pa je naš namen čim hitrejša prilagoditev razmeram, sklepamo, da je to možno uresničiti ob upoštevanju naslednjih predpostavk:

- glede na razmerja korelacija – dekorelacija je hitrost učenja možno povečati s primernim izborom zvokovnih signalov, tj. posnemanjem razmer, kot smo jih vajeni iz vsakdanjega življenja;
- ker želimo poleg akustične slike uporabniku omogočiti tudi spremljanje zvokov iz okolice, možganom postavljamo dodatno zahtevo, da hkrati upoštevajo in se odzivajo na zahteve obeh modelov zaznavanja zvoka – naravni zvoki iz okolice in zvoki akustične slike;
- z izborom zvokov, ki so bliže naravni predstavitvi, poslušalca dodatno motiviramo za učenje s tem, da bi si že ob prvem poslušanju poskušal predstavljati prostorsko zasnovo akustične slike.

4 LOKALIZACIJA IZVOROV ZVOKA PRI LJUDEH

4.1 Lokalizacija izvorov zvoka v prostoru in pri predvajanju s slušalkami

Lokalizacija zvoka ima v vsakdanjem življenju velik pomen, saj nam pomaga pri orientaciji v prostoru, omogoča izogibanje nevarnostim ipd. Nasploh nas v vsakdanjem življenju zvok v različnih oblikah spremlja na vsakem koraku. Kljub pomembni funkciji, ki jo opravlja organ sluha, pa vendarle lahko opazimo, da je sluh pri človeku v primerjavi z vidom manj razvit in ima torej manj pomembno vlogo v življenju. Opazimo tudi, da je evolucija nekoliko prikrajšala človekov organ sluha v primerjavi z nekaterimi živalmi, ki imajo možnost premikanja uhljev in s tem precej boljše možnosti lokalizacije izvorov zvoka.

Z natančnostjo lokalizacije zvoka opišemo, kako natančno je človek sposoben določiti smer iz katere prihaja zvočno valovanje. V splošnem je nekoliko nenatančno govoriti o lokalizaciji izvorov zvoka, saj v prostoru na lokalizacijo vplivajo dejavniki, ki jim je zvočno valovanje podvrženo na poti med izvorom zvoka in človekovim uhljem, ti dejavniki pa lahko povzročijo, da zvok slišimo prihajati iz popolnoma druge smeri, kot se izvor zvoka dejansko nahaja. Slednje poimenujemo akustična iluzija (angl. acoustic illusion). Dejavnike, ki vplivajo na razširjanje zvočnega valovanja in s tem na akustične iluzije lahko tudi koristno uporabimo. Za merjenje oz. določanje natančnosti lokalizacije izvorov zvoka pri človeku pa jih želimo izločiti, zato meritve opravljamo v gluhi sobi.

Točnost lokalizacije je pomemben podatek za načrtovanje akustične slike prostora, ki jo poslušamo s slušalkami. Akustična slika se nanaša na predvajanje prostorskih zvokov preko slušalk, kar obsega različne aplikacije – od navigacijskih sistemov za slepe in slabovidne [22, 34, 44] do multimedijskih aplikacij [71]. V akustični sliki želimo predvajati prostorske zvoke s čim boljšo natančnostjo lokalizacije, tj. natančnostjo, ki jo je človek sposoben razviti v svojem naravnem okolju in ki jo lahko izmerimo v gluhi sobi. Natančnost lokalizacije pri poslušanju s slušalkami je v prvi vrsti odvisna od načina generiranja prostorskega zvoka, za kar uporabljamo filtre, izpeljane iz prenosnih karakteristik HRTF ali pa kako drugo vrsto kodiranja. Z uporabo tehnik za generiranje prostorskega zvoka navidezno premikamo izvore zvoka po prostoru.

Pri lokalizaciji izvorov zvoka imajo posebej pomembno vlogo psihološki dejavniki. Če se torej omejimo zgolj na lokalizacijo izvorov zvoka, ugotovimo, da poleg neposredno prestreženega zvočnega valovanja pri lokalizaciji sodelujejo tudi organ vida, organ tipa,

ravnotežni organ in zvok, ki se širi po drugih poteh (npr. po lobanjskih kosteh) [4]. Pogosto se zgodi, da poslušalec, ki ne vidi izvora zvoka ni povsem prepričan o lokaciji izvora zvoka. Ta pojav je še posebej opazen pri predvajanju prostorskih zvokov preko slušalk [78], kjer poslušalci najpogosteje zamenjujejo smeri spredaj in zadaj (angl. »front-back confusion«). Čeprav so ta pojav opazili mnogi raziskovalci, (še) nimamo niti splošno sprejete razlage o mehanizmu, ki omogoča ločevanje izvorov v prostoru pred in za poslušalcem. Še najbolj verjetna se zdi teorija, da si človek v »dvoimljivih« situacijah pomaga z minimalnimi premiki glave [32], kar je primerljivo s premikanjem uhljev pri nekaterih živalih. Pomemben vpliv na lokalizacijo izvora zvoka ima tudi organ vida, kjer gre za povezavo vidnega in slišane (t.i. vizualna teorija [4]). Pojav, da poslušalec ropot helikopterja lokalizira nad svojo glavo pa je posledica naučenega dejstva, da helikopterji vedno letijo nad nami. Primer psihološkega vpliva opazimo tudi pri predvajanju šumečih impulzov različnega trajanja. Poslušalci krajše impulze običajno zaznavajo kot točkaste izvore, daljše impulze pa kot izvore z večjo površino [70].

Za sistematično raziskovanje fizikalnih in psiholoških pojavov v zvezi z lokalizacijo izvorov zvoka in prostorskimi zvoki uporabimo en izvor zvoka, tj. en zvočnik. S fizikalnega stališča predstavlja uporaba enega izvora zvoka elementarni primer. Če namreč poznamo učinek, ki ga povzroči en izvor zvoka v poslušalčevem sluhovodu ali bobniču in če ta učinek poznamo za vsak zanimiv položaj zvočnika in poslušalca, potem lahko z zakonom superpozicije določimo tudi učinek več izvorov. Ta zakonitost sledi neposredno iz lastnosti linearnosti zvočnega valovanja. Ker nas zanimajo predvsem zakonitosti, ki so lastne človeku, dejavnike prostora izločimo z uporabo gluhe sobe, dejavnike merilne opreme pa s filtriranjem z njihovimi inverznimi karakteristikami.

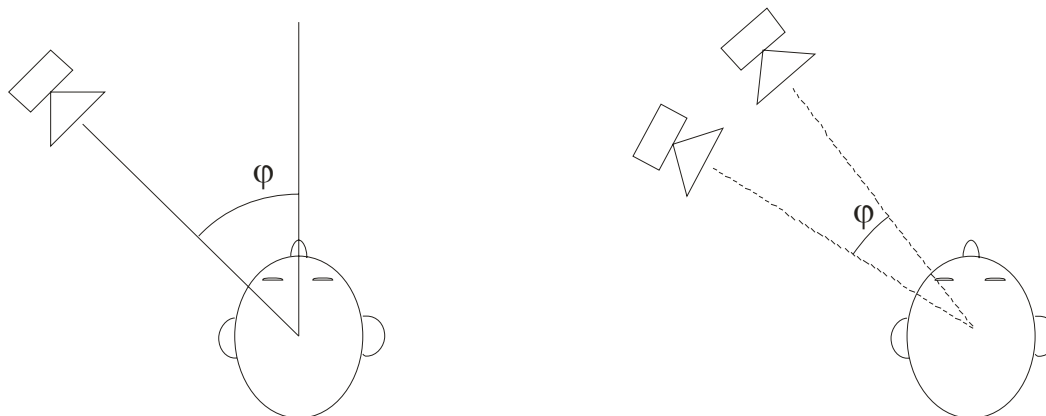
S psihološkega vidika so stvari bolj zapletene, ker človeško dojemanje oz. interpretacija občutkov, ki jih dobimo prek čutil iz zunanjega sveta ni linearna. Kljub temu so rezultati, pridobljeni z enim samim izvorom zvoka pomembni, ker enak zvok v sluhovodu povzroči enake občutke ne glede na način nastanka tega zvoka.

4.2 Opredelitev pojma »ločljivost«

Obsežne raziskave o ločljivosti oz. resoluciji izvorov zvoka pri ljudeh so se začele po letu 1958, ko je bil objavljen članek [48]. Prve raziskave so bile osredotočene na ločljivost statičnih izvorov zvoka v horizontalni smeri (azimut), ki so jim kasneje sledile raziskave o ločljivosti v vertikalni smeri (elevacija) in raziskave ločljivosti gibljivih izvorov zvoka. Kot testni signal se v raziskavah povezanih z ločljivostjo največkrat pojavlja beli šum, omejen s frekvenčnim pasom razpona človekovega sluha (16 Hz – 20 kHz), drugačni

zvokovni signali pa se uporabljajo predvsem v smislu ugotavljanja vpliva spektralne vsebine na ločljivost.

S pojmom absolutna ločljivost ali točnost lokalizacije opisujemo zmožnost poslušalca, da ugotovi smer v kateri se nahaja izvor zvoka. Rezultate podajamo z razmerjem med dejansko lokacijo izvora zvoka in lokacijo, ki jo je ocenil poslušalec.



Slika 4.1: Absolutna ločljivost (levo) – zanima nas smer izvora zvoka, tj. azimut in elevacija. Diferenčna ločljivost (desno) – zanima nas najmanjši kot φ , ko bo poslušalec še ločil, da gre za dva izvora zvoka, pri čemer nas dejanski položaj izvorov zvoka ne zanima.

Diferenčna ločljivost (angl. localization blur [4]) se nanaša na ugotavljanje ločljivosti med dvema izvoroma zvoka, ko je pomembno le ali poslušalec loči dva izvora zvoka in pri tem absolutni položaja izvorov zvoka za poslušalca ni pomemben. Kljub temu, da nas pri ugotavljanju diferencialne ločljivosti absolutni položaj izvora ne zanima, se izkaže, da je diferencialna ločljivost odvisna od absolutnega položaja izvorov zvoka, kar bomo videli v nadaljevanju. Diferenčno ločljivost oz. resolucijo merimo v kotnih stopinjah, vrednost predstavlja najmanjši kot, ko poslušalec še zazna dva ločena zvočna izvora, z nadaljnjim zmanjšanjem kotne razlike pa ju ne loči več. Pojma absolutne in diferencialne ločljivosti prikazuje tudi slika 4.1.

Kadar izvori zvokov mirujejo, govorimo o statični ločljivosti in najmanjši kot poimenujemo MAA (Minimum Audible Angle). Ločljivost je odvisna od več dejavnikov, med njimi velja omeniti spekter zvoka, čas trajanja zvočnega signala in absolutno smer izvora zvoka. Za ločljivost gibajočih se izvorov zvoka uporabljamo kratico MAMA (Minimum Audible Movement Angle). Ločljivost gibajočih se izvorov zvoka je odvisna predvsem od frekvenčne vsebine zvoka in hitrosti gibanja izvora zvoka. V merilnem postopku MAMA se izvor zvoka v želeni smeri giblje po krožnem loku, naloga poslušalca pa je, da ugotovi smer gibanja izvora. Z manjšanjem kota, ki ga opiše izvor zvoka, ugotavljamo najmanjši kot ločljivosti (MAMA).

V nadaljevanju tega poglavja se bomo ukvarjali z ločljivostjo statičnih, tj. nepremičnih izvorov zvoka s poudarkom na ločljivosti v elevaciji.

4.3 Merjenje individualnih HRTF

Individualne HRTF smo merili na treh osebah. Postopek in pogoji preproste meritve so natančneje opisani v [66], tu pa navajamo le najpomembnejše podatke.

Meritve smo izvajali v gluhi sobi Fakultete za elektrotehniko [15], uporabili smo prenosni računalnik Acer TravelMate 4000, programski paket Matlab 6.1, zvočno kartico Digigram VX Pocket 440, zvočnike Genelec 8030 APM in mikrofone SoundMan OKM-II Professional. Izmerjeni signal smo vzorčili s frekvenco 44100 Hz, razmerje signal/šum (*SNR*) v gluhi sobi je bilo 50 dB, izmerjeno z merilnikom Lutron SL-4012. Zvočnik, ki je služil kot izvor zvoka, smo premikali po krožnici s polmerom 1 m glede na središče poslušalčeve glave. Meritve smo opravili za 8 različnih elevacij (0° do 35° s korakom 5°) pri azimutu 0° , razen za elevacijo 0° , kjer smo merili pri azimutih od -90° do 90° s korakom 5° . Pri računanju HRTF smo upoštevali tudi odzive prazne sobe in merilne opreme, ki smo jih izločili z inverznim filtriranjem.

4.4 Točnost lokalizacije (absolutna ločljivost)

Pri ugotavljanju oz. merjenju točnosti lokalizacije izvorov zvoka moramo zagotoviti način s katerim bo testna oseba nedvoumno nakazala položaj izvora zvoka. Pristopi k reševanju tega problema so raznovrstni, od različnih načinov merjenja položaja glave [7], merjenja premikov oči [30], opisnih odgovorov s pomočjo polarnih ali urnih (položaj kazalca kaže smer izvora zvoka) koordinat [77] do pritiskanja na tipke. V literaturi najdemo meritve absolutne ločljivosti v prostoru, tj. z zvočniki [7, 77], meritve točnosti lokalizacije navideznih izvorov zvoka s slušalkami in uporabo individualnih HRTF [7, 77, 39] ali s slušalkami in individualnimi HRTF druge osebe [7].

Izkaže se, da sta točnost lokalizacije v prostoru in točnost lokalizacije pri poslušanju s slušalkami in individualnimi HRTF primerljivi, medtem ko daje uporaba splošnih HRTF nekaj slabše rezultate. V raziskavi [7] je bilo ugotovljeno, da točnost lokalizacije v elevaciji ni odvisna od položaja izvora zvoka, medtem ko v raziskavah [77, 78] opazimo določeno odstopanje, ki je izrazitejše pri nekoliko višjih elevacijah, tj. več kot 36° . Glede na rezultate, objavljene v [7, 78] je odstopanje med dejanskim in »slišanim« položajem izvora zvoka v prostoru med 8° [7] in 25° [78], pri uporabi slušalk in individualnih HRTF med 13° [7] in 25° [78] ter pri uporabi slušalk in splošnih HRTF nad 15° [7]. Pri rezultatu

s slušalkami in splošnimi HRTF je potrebno upoštevati, da je točnost odvisna predvsem od razlik med poslušalcu lastnimi HRTF in uporabljenimi HRTF.

Pri meritvah v prostoru je potrebno testni osebi prekriti oči ali primerno zatemniti prostor, zato tako pri poskusih v prostoru kot pri poskusih s slušalkami, naletimo na pojav zamenjave smeri. To pomeni, da poslušalec npr. sliši izvor zvoka zadaj, čeprav se le-ta nahaja pred njim. Zamenjava smeri je opazna še v smeri gor – dol in po diagonali. Pojav je precej bolj očiten pri poslušanju s slušalkami, v [7] najdemo podatek, da jih je v primeru s slušalkami kar dvakrat toliko kot pri poslušanju v prostoru.

4.5 Merjenje diferenčne ločljivosti s slušalkami in z zvočniki

Mera diferenčne ločljivosti izvorov zvoka pri človeku je podana z ločljivostjo v prostoru, tj. merjeno z zvočniki v gluhi sobi. V literaturi najdemo podatke o ločljivosti po azimutu med 1° [4, 48] in 3° [4, 28] ter ločljivostjo po elevaciji med 4° [4] in 10° [28]. Naše meritve potrjujejo te rezultate. S poskusom, ki ga bomo opisali v nadaljevanju, smo v smeri naravnost pred poslušalcem izmerili ločljivost po elevaciji 6° . Pri tem velja omeniti, da je ločljivost odvisna od smeri (absolutnega položaja izvorov zvoka) in pada z oddaljevanjem od horizontalne ravnine.

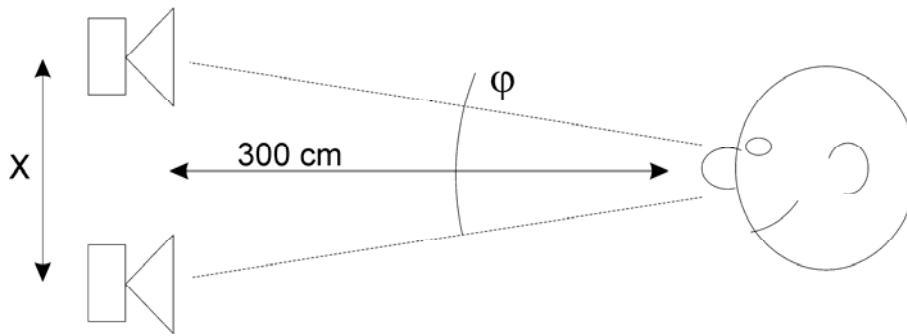
V nadaljevanju si najprej pogledjmo postopek in rezultate merjenja diferenčne ločljivosti resničnih izvorov zvokov po elevaciji, ki smo ga izvedli z zvočniki v gluhi sobi, nato pa še določanje ločljivosti navideznih izvorov zvokov po elevaciji z uporabo slušalk in splošnih HRTF knjižnice CIPIC [2].

4.5.1 Merjenje diferenčne ločljivosti po elevaciji z zvočniki

Meritev smo izvajali v gluhi sobi. Za generiranje zvoka smo uporabili prenosni računalnik IBM ThinkPad 30, zvočno kartico Digigram VX Pocket 440 in zvočnike Genelec 8030 APM. Nivo šuma v gluhi sobi smo izmerili z merilnikom Lutron SL-4012 in je znašal 28 dB, merjeno po karakteristiki *A*. Razmerje signal-šum med potekom meritve je bilo 50 dB.

V poskusu je sodelovalo 32 testnih oseb starih med 15 in 50 let. Nobena od testnih oseb ni poprej sodelovala pri kakršnemkoli podobnem poskusu povezanem z akustiko, prav tako so vse testne osebe potrdile, da nimajo težav z vidom ali sluhom. Ker se je zdel testnim osebam poskus zanimiv, menimo, da je to vplivalo tudi na primerno koncentracijo in s tem korektnost rezultatov.

Za potrebe testa smo v frekvenčnem prostoru s pomočjo Matlaba generirali signal z ravno amplitudno karakteristiko in naključno fazo ter ga z inverzno Fourierovo transformacijo pretvorili v časovni prostor [73]. Signal smo predvajali v 100 ms sekvencah, izmenično na prvem in drugem zvočniku. Ko je bil aktiven prvi zvočnik, na drugem zvočniku ni bilo signala in obratno. Pasovna širina testnega signala je bila 20 kHz.



Slika 4.2: Postavitev zvočnikov za merjenje ločljivosti po elevaciji v gluhi sobi. Povezava med kotom φ in razmikom med zvočnikoma x je navedena v tabeli 4.1.

Testno osebo smo posadili na stol in ji okoli oči zavezali trak, da smo preprečili morebitno »pomoč« vida. Meritev smo izvajali v sagitalni ravnini (azimut 0°), testna oseba je bila od zvočnikov oddaljena 3 m, s spreminjanjem razmika med zvočnikoma smo nastavljali kot elevacije, pri čemer nas je zanimal minimalni kot, ko poslušalec še razloči, da zvok prihaja iz dveh izvorov zvoka. Razmik med zvočnikoma za preizkušene kote podaja tabela 4.1, ki skupaj s sliko 4.2 nazorneje prikazuje razmere pri poskusu.

$\varphi / ^\circ$	56	45	28	23	12	6	4	2
x / cm	273	250	145	120	63	31.5	21	31.5

Tabela 4.1: Koti in razdalje med zvočnikoma (postavljenima v sagitalni ravnini) pri meritvi ločljivosti po elevaciji. Razdalja med poslušalcem in zvočnikoma je bila 3 m (glej tudi sliko 4.2).

Z vsako testno osebo smo izvedli 6 meritev. V treh primerih smo razdaljo med zvočnikoma zmanjševali od največje ($\varphi = 56^\circ$) proti najmanjši, v treh primerih pa je bil postopek obrnjen in smo razdaljo med zvočnikoma povečevali od najmanjše ($\varphi = 4^\circ$) proti največji. Postopek meritve je za vsako osebo trajal približno 15 min.

Vrednosti v grafu na sliki 4.3 so podane kot povprečne vrednosti vzorcev vsakega posameznika (število vzorcev $N = 6$) z označenim intervalom zaupanja.

Interval zaupanja μ definiramo kot:

$$\mu = \bar{x} \pm s \tag{4.1}$$

kjer je \bar{x} srednja vrednost (4.2) in s disperzija (4.3):

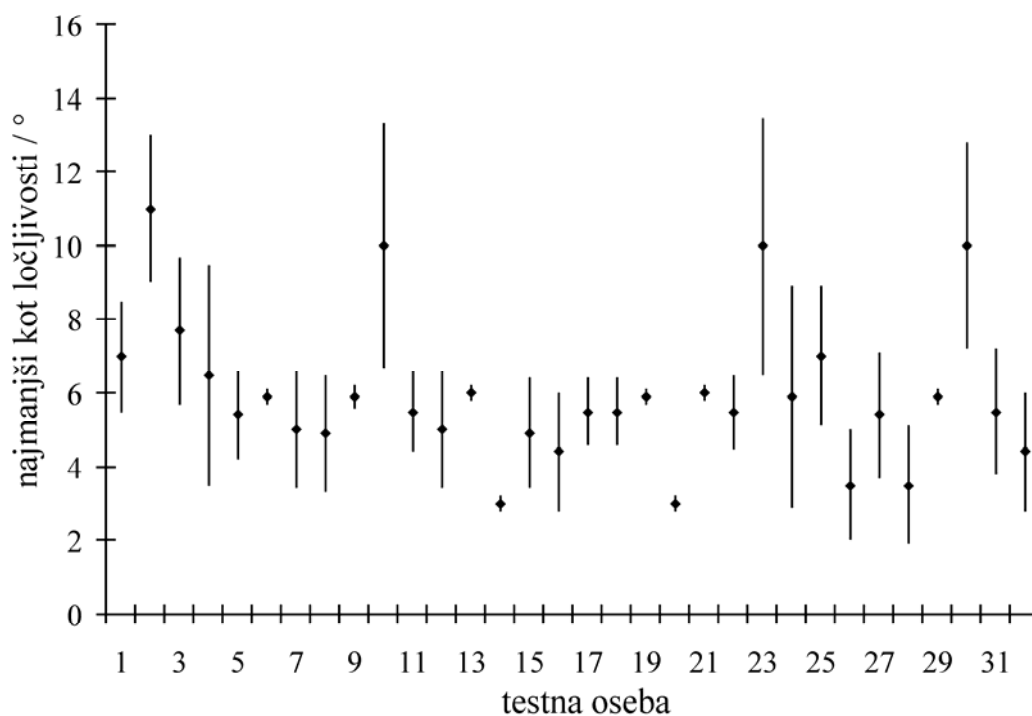
$$\bar{x} = \frac{1}{N} \sum_{i=1}^N x_i \quad (4.2)$$

$$s = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (x_i - \bar{x})^2} \quad (4.3)$$

Tabela 4.2 podaja pregled rezultatov po statističnih razredih, ki so osnovani na podlagi postavitev zvočnikov.

statistični razred	populacija (%)
$\varphi < 4^\circ$	12.5
$4^\circ \leq \varphi < 6^\circ$	56.3
$6^\circ \leq \varphi < 12^\circ$	31.2
$\varphi \geq 12^\circ$	0.0

Tabela 4.2: Rezultati lokalizacije po elevaciji, razdeljeni po statističnih razredih.



Slika 4.3: Diferenčna ločljivost po elevaciji (interval zaupanja z označeno srednjo vrednostjo) merjena z resničnimi zvočnimi izvori (zvočniki) v gluhi sobi. Zvočna izvora sta bila postavljena v sagitalni (azimut = 0°) ravnini, simetrično okoli točke (azimut, elevacija) = (0° , 0°). Povprečna diferenčna ločljivost v elevaciji je približno 6° .

4.5.2 Merjenje diferenčne ločljivosti po elevaciji s slušalkami in z uporabo splošnih HRTF

Meritve s slušalkami smo izvajali v prostorih Laboratorija za komunikacijske naprave. Nivo šuma, izmerjen z merilnikom Lutron SL-4012 je bil 40 dB, merjeno po karakteristiki A. Za generiranje zvoka smo uporabili prenosni računalnik IBM ThinkPad 30, zvočno kartico Digigram VX Pocket 440 in studijske slušalke Sennheiser HD270 s slabljenjem zunanjih šumov 10 dB – 15 dB.

Tudi v tem poskusu je sodelovalo 32 testnih oseb starih med 15 in 50 let, o katerih smo že pisali v prejšnjem razdelku.

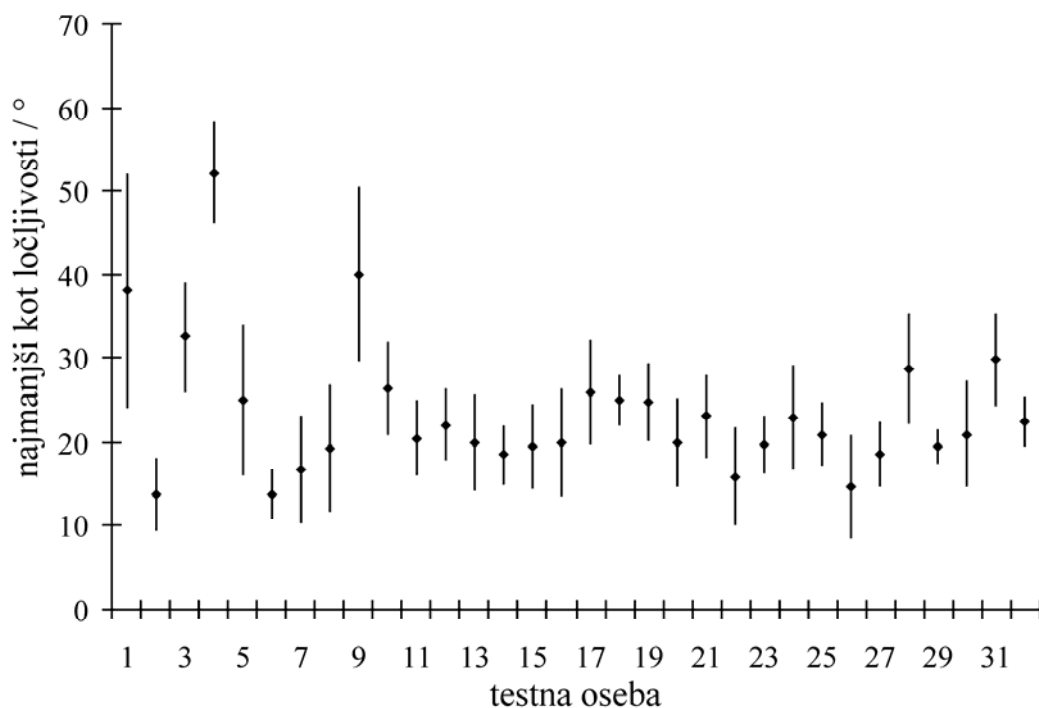
Za potrebe testa smo v frekvenčnem prostoru s pomočjo Matlaba generirali signal z ravno amplitudno karakteristiko in naključno fazo ter ga z inverzno Fourierovo transformacijo pretvorili v časovni prostor. Pasovna širina testnega signala je bila 20 kHz. Za generiranje prostorskega zvoka oz. za postavljanje izvorov zvoka na poljubne virtualne položaje smo uporabili HRTF knjižnice CIPIC [2], kjer elevacijo izvora zvoka lahko spreminjamo med -45° in 230° v koraku po 5.625° . Ugotavljali smo ločljivost za primere, ko je bil elevacijski kot med izvoroma zvoka 5.625° , 11.25° , 22.5° , 28.125° , 45° in 56.25° . Merjenje smo izvajali v sagitalni ravnini (azimut 0°), položaj testnih izvorov zvoka smo izbirali čim bolj simetrično glede na horizontalno ravnino. Dejanske razmere prikazuje tabela 4.3.

diferenca elevacijskega kota / $^\circ$	elevacija izvora A / $^\circ$	elevacija izvora B / $^\circ$
5.625	5.625	0
11.25	5.625	-5.625
22.5	11.25	-11.25
28.125	16.875	-11.25
45	22.5	-22.5
56.25	28.125	-28.125

Tabela 4.3: Virtualne pozicije izvorov zvoka pri merjenju s slušalkami in splošnimi HRTF (knjižnica CIPIC). Virtualni izvori so bili v vseh primerih v sagitalni (azimut = 0°) ravnini. Poskus smo izvajali po dveh scenarijih – z zmanjševanjem in povečevanjem difference elevacijskega kota.

Signal smo predvajali v 100 ms sekvencah, izmenično na prvi in drugi virtualni poziciji. Posamezne 100 ms odseke zvoka smo prekinjali z 200 ms tišine. Z vsako testno osebo smo izvedli 6 meritev. V treh primerih smo elevacijski kot med virtualnima izvoroma zmanjševali od največjega ($\varphi = 56.25^\circ$) proti najmanjšemu, v ostalih treh primerih pa je bil postopek obrnjen in smo elevacijski kot med virtualnima izvoroma

povečevali od najmanjšega ($\varphi = 5.625^\circ$) proti največjemu. Postopek meritve je za vsako osebo trajal približno 4 min.



Slika 4.4: Diferenčna ločljivost po elevaciji (interval zaupanja z označeno srednjo vrednostjo), merjena z virtualnimi zvočnimi izvori (slušalke), generiranimi s pomočjo splošnih HRTF. Povprečna diferenčna ločljivost je približno 23° .

Vrednosti v grafu na sliki 4.4 so podane kot povprečne vrednosti vzorcev vsakega posameznika (število vzorcev $N = 6$) z označenim intervalom zaupanja. Podrobnosti v zvezi z računanjem rezultatov so razložene v prejšnjem razdelku.

Tabela 4.4 podaja pregled rezultatov po statističnih razredih, določenih z uporabljenimi elevacijskimi koti.

statistični razred	populacija (%)
$\varphi < 5.625^\circ$	0
$5.625^\circ \leq \varphi < 11.25^\circ$	0
$11.25^\circ \leq \varphi < 22.5^\circ$	59.4
$22.5^\circ \leq \varphi < 45^\circ$	37.5
$\varphi \geq 45^\circ$	3.1

Tabela 4.4: Rezultati lokalizacije po elevaciji, razdeljeni po statističnih razredih.

4.5.3 Merjenje diferenčne ločljivosti po elevaciji s slušalkami in uporabo individualnih funkcij

Zvočno valovanje, ki ga vzbudimo s pomočjo slušalk in individualnih HRTF je teoretično enako kot bi ga vzbudilo zvočno valovanje zvočnika. Fizikalno gledano sta primera enakovredna, seveda pa pri tem ne upoštevamo vseh psiholoških dejavnikov, ki vplivajo na lokalizacijo. Najbolj izrazit pojav je zamenjevanje smeri spredaj-zadaj (front-back confusion) o čemer smo podrobneje pisali v razdelku 4.1.

V poskusu so sodelovale 3 testne osebe stare med 23 in 28 let za katere smo izmerili individualne HRTF s postopkom, opisanim v razdelku 4.4 in v [66].

Za potrebe ovrednotenja ločljivosti po elevaciji pri uporabi slušalk in individualnih HRTF smo uporabili funkcije sagitalne ravnine – azimut 0° in elevacije med 0° in 35° . Za testni signal smo v frekvenčnem prostoru s pomočjo Matlaba generirali signal z ravno amplitudno karakteristiko in naključno fazo [73] ter ga z inverzno Fourierovo transformacijo pretvorili v časovni prostor. Pasovna širina testnega signala je bila 20 kHz. Signal smo predvajali v 100 ms sekvencah, izmenično na prvi in drugi virtualni poziciji. Posamezne 100 ms odseke zvoka smo prekinjali z 200 ms šuma. Z vsako testno osebo smo izvedli 6 meritev. V treh primerih smo elevacijski kot med virtualnima izvoroma zmanjševali od največjega ($\varphi = 30^\circ$) proti najmanjšemu, v ostalih treh primerih pa je bil postopek obrnjen in smo elevacijski kot med virtualnima izvoroma povečevali od najmanjšega ($\varphi = 5^\circ$) proti največjemu. Postopek meritve je za vsako osebo trajal približno 4 min. Virtualne položaje za posamezne testne primere podaja tabela 4.5.

razlika elevacijskih kotov / $^\circ$	elevacija navideznega izvora A / $^\circ$	elevacija navideznega izvora B / $^\circ$
5	15	20
10	15	25
15	10	25
20	10	30
25	5	30
30	5	35

Tabela 4.5: Virtualne pozicije izvorov zvoka pri merjenju s slušalkami in uporabo individualnih HRTF.

Meritve ločljivosti smo izvajali v Laboratoriju za komunikacijske naprave. Nivo šuma, izmerjen z merilnikom Lutron SL-4012 je bil 40 dB, zmerjeno po karakteristiki A. Za generiranje zvoka smo uporabili prenosni računalnik Acer TravelMate 4000, zvočno kartico Digigram VX Pocket 440 in studijske slušalke Sennheiser HD270 s slabljenjem zunanjih šumov 10 dB – 15 dB.

Kot prikazuje tabela 4.6, smo za vse tri osebe ugotovili ločljivost 10° .

zaporedna št. osebe	ločljivost po elevaciji / $^\circ$
1	10
2	10
3	10

Tabela 4.6: Povprečna diferenčna ločljivost po elevaciji pri merjenju s slušalkami in individualnimi HRTF. Predvidevamo, da je v resnici kot ločljivosti še nekoliko manjši, ki pa nam ga zaradi preprostega postopka merjenja individualnih HRTF ni uspelo ugotoviti.

elevacija središčne točke / $^\circ$	elevacija izvora A / $^\circ$	elevacija izvora B / $^\circ$
5	0	10
10	5	15
15	10	20
20	15	25
25	20	30
30	25	35

Tabela 4.7: Virtualne pozicije izvorov zvoka za ugotavljanje ločljivosti po elevaciji v odvisnosti od absolutnega položaja izvorov zvoka. Virtualne izvore smo generirali z individualnimi HRTF in jih predvajali preko slušalk.

elevacija središčne točke / $^\circ$	ločljivost posamezne testne osebe / $^\circ$			povprečna ločljivost / $^\circ$
	#1	#2	#3	
5	10	10	10	10
10	10	10	10	10
15	10	10	10	10
20	10	10	10	10
25	10	10	10	10
30	10	10	10	10

Tabela 4.8: Diferenčna ločljivost navideznih izvorov zvoka (generiranih z individualnimi HRTF in predvajanih preko slušalk) v omejenem obsegu elevacij ($5^\circ - 30^\circ$) ne kaže odvisnosti od absolutnega položaja izvora zvoka

Nadalje nas je zanimala še odvisnost diferenčne ločljivosti od absolutnega položaja izvorov zvoka. Med navideznima izvoroma zvoka smo nastavili kot 10° in spreminjali njun položaj tako kot podaja tabela 4.7. Rezultati, dobljeni v omejenih pogojih ne kažejo, da bi bila ločljivost odvisna od absolutnega položaja (tabela 4.8), kot je to npr. očitno pri ločljivosti v horizontalni smeri [67]. Majhno odvisnost ločljivosti od absolutnega položaja

središčne točke v za nas zanimivem delu elevacije (-90° do 90°) so potrdili tudi drugi raziskovalci. V [4] najdemo podatek, da ločljivost pada od 9° , pri središčni točki 0° , do ločljivosti 13° pri središčni točki 74° . Situacija pa se precej spremeni, ko izvore zvoka postavimo za poslušalčevo glavo, tj. elevacija je večja od 90° . Ločljivost se poslabša približno za faktor 2, kar npr. pomeni ločljivost 22° pri elevaciji 112° ali ločljivost 15° pri elevaciji 153° [4].

Izmerjena ločljivost kaže, da je dejanska ločljivost pri uporabi individualnih HRTF zelo verjetno še boljša kot 10° . Žal nam naš preprost merilni postopek individualnih HRTF ne omogoča merjenja HRTF po elevaciji z večjo natančnostjo, zato se v tej fazi raziskav zadovoljimo z dobljenimi rezultati.

4.6 Vpliv spektralne vsebine zvoka na lokalizacijo

Za zaznavanje elevacije izvorov zvoka je bistvenega pomena spekter zvoka, ki ga prestreže poslušalčev bobnič [1, 4, 13, 30, 57]. Kot smo že napisali, na spekter vplivajo dimenzije in oblika poslušalčevega telesa, glave in uhljev, ki povzročajo uklone, odboje in interferenco zvočnega valovanja. Posebno vlogo pripisujemo uhlju, ki povzroča t.i. oslabitev uhlja (pinna notch). Centralna frekvenca tega slabljenja je odvisna od elevacije izvora zvoka in se spreminja od 5 kHz za nizke elevacije do 17 kHz [57] za visoke elevacije.

V literaturi najdemo množico opisov poskusov v katerih so poskušali dokazati vpliv posameznih frekvenčnih pasov na lokalizacijo izvorov zvoka po elevaciji in ki se ukvarjajo z ugotavljanjem zakonitosti zaznavanja elevacije [1, 31, 33, 39, 43]. Omenili smo že oslabitev uhlja, ki jo nekateri viri, npr. [5], navajajo kot glavni parameter lokalizacije po elevaciji. Če je prisotnost te oslabitve v amplitudnem poteku HRTF potreben pogoj za zaznavo elevacije, potem je za zaznavo elevacije v zvokovnem signalu potrebna prisotnost višjih frekvenc, kar pa ne drži. Poskusi s signali nizkih frekvenc so namreč pokazali, da komponente višjih frekvenc v zvokovnem signalu niso nujno potrebne [1]. Če uporablja poslušalec za reprodukcijo zvoka individualne HRTF, je zaznavanje elevacije povsem zadovoljivo tudi s signali, ki vsebujejo le nizke frekvence (< 3 kHz) [1]. Vendar pa obstoj spektralne vsebine pri višjih frekvencah zagotavlja boljšo ločljivost [1, 7, 39] in zmanjša verjetnost zamenjave smeri [7].

V raziskavi [39] pa ugotavljajo, da so za točnost lokalizacije širokopasovnih zvokovnih signalov bistvene spektralne značilnosti v pasu 4 kHz – 16 kHz. V omenjeni raziskavi so uporabili 200 ms trajajočo sekvenco šuma z ravnim spektrom med 0 Hz in 22 kHz, ki so jo filtrirali s pasovno zapornimi filtri. Na podlagi tega so ugotovili, da spektralna vsebina zvoka v pasu 5.7 kHz – 11.3 kHz odločilno vpliva na točnost

zaznavanja elevacije, spektralna vsebina v pasu 8 kHz – 16 kHz na zaznavanje ali se izvora zvok nahaja pred ali za poslušalcem, spektralne značilnosti zvoka v pasu pod 4 kHz pa imajo le minimalen vpliv na lokalizacijo po elevaciji. Avtorji raziskave glede na rezultate ugotavljajo, da sta za lokalizacijo po elevaciji ključna dejavnika oslabitev uhlja in ojačitev, ki na frekvenčni osi sledi slabljenju uhlja.

4.7 Vpliv glasnosti in trajanja zvočnega signala na lokalizacijo

Možgani oz. ustrezni nevroni lokalizirajo prostorske zvoke na podlagi akustičnih značilnosti zvoka. Kot smo že omenili, je človekov slušni sistem občutljiv za časovne in amplitudne razlike med levim in desnim ušesom (angl. binaural hearing) ter spekter zvoka (angl. monaural hearing). Slednje je posledica loma in odbojev zvočnega valovanja od telesa, glave in uhljev.

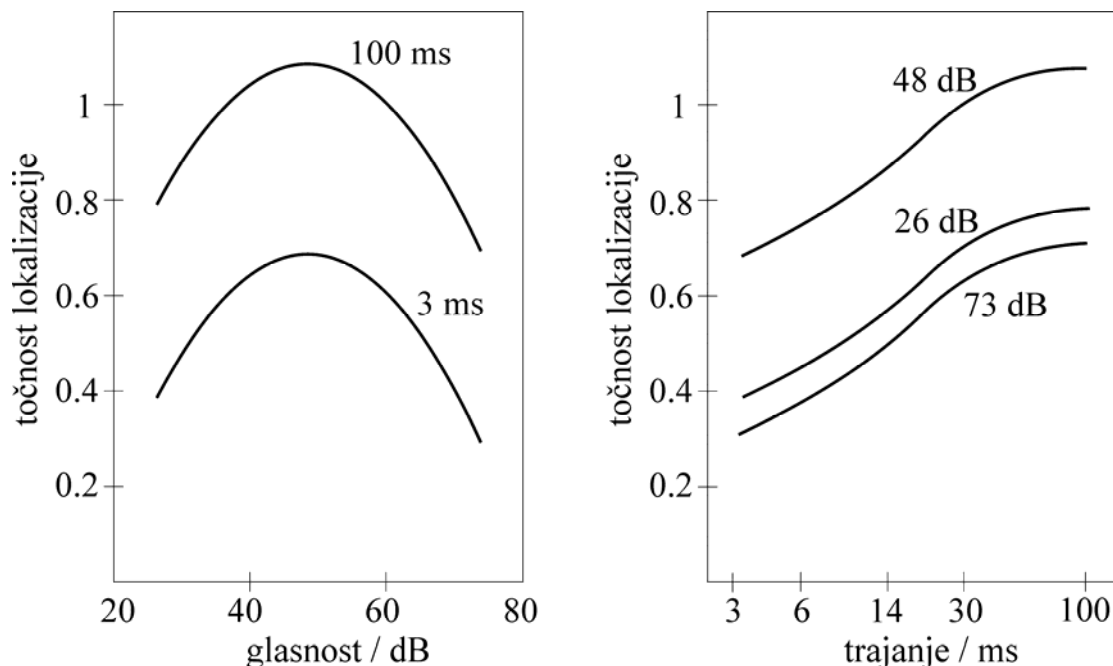
Začetniki raziskav človekovega dojemanja prostorskega zvoka so poskuse izvajali večinoma z relativno dolgo trajajočimi signali s stacionarnimi [72] lastnostmi. Šele kasneje so se pojavile prve raziskave o vplivu glasnosti in dolžine trajanja signalov na lokalizacijo [13, 31, 42, 75] ter s temi raziskavami povezane hipoteze, ki jih bomo na tem mestu zgolj poimensko omenili: »the neural integration hypothesis« (hipoteza nevronske integracije) [31], »negative level effect« (efekt zmanjšanja glasnosti, tj. s povečevanjem glasnosti se točnost lokalizacije zmanjšuje enako kot če bi glasnost zmanjševali) [42] in »adaptation hypothesis« (hipoteza prilagoditve) [25]. V omenjenih raziskavah so raziskovali vpliv glasnosti in trajanja zvočnega signala na lokalizacijo po azimutu in elevaciji, pri čemer so eksperimente izvajali z zvočniki v prostoru. Pregled hipotez in poenoteno teorijo na katero se bomo sklicevali v nadaljevanju najdemo v [75], kjer so za testni signal uporabili beli Gaussov šum trajanja 3 ms – 100 ms in glasnosti 26 dB – 73 dB *SPL*. Rezultati meritev se nanašajo na absolutno lokalizacijo izvorov zvoka.

V splošnem velja, da lokalizacija po azimutu ni omejena ne s trajanjem kot tudi ne z glasnostjo zvokovnega signala, če je le ta v mejah dinamike človekovega sluha [31, 75].

Na lokalizacijo po elevaciji pa vplivata oba dejavnika, ki ju raziskujemo v tem razdelku [75]. Točnost lokalizacije monotono (vendar neenakomerno) narašča s povečevanjem trajanja zvokovnega signala, kar je še posebej izrazito pri časih do 30 ms, nad 30 ms pa je naraščanje točnosti lokalizacije po elevaciji manj izrazito. Sama odvisnost med točnostjo lokalizacije in trajanjem zvokovnega signala je neodvisna od glasnosti zvoka, slednja pa bistveno vpliva na dejansko točnost lokalizacije. Slika 4.5 (povzeta po rezultatih objavljenih v [75]) prikazuje točnost lokalizacije po elevaciji (angl. elevation gain), ki je podana s kvocientom dejanske lokacije in lokacije, ki jo je »slišala« testna oseba, tj. kvocient 1 pomeni najboljše ujemanje. Potek krivulje je sicer določen s

parametri, ki so izračunani na podlagi nelinearne regresije rezultatov merjenja točnosti lokalizacije [31, 75, 50] (najboljše ujemanje, angl. best fit).

Odvisnost točnosti lokalizacije od glasnosti zvoka ni monotona in je pri nizkih (< 30 dB *SPL*) in visokih glasnostih (> 65 dB *SPL*) majhna, največja pa pri srednjih glasnostih kot lahko vidimo tudi na sliki 4.5.



Slika 4.5: Točnost lokalizacije v odvisnosti od glasnosti in trajanja zvokovnega signala. Vidimo, da obstaja optimum izbire glasnosti, medtem ko z daljšim trajanjem zvoka točnost le izboljšamo. Podatke za sliko smo dobili v [75].

Za razliko od omenjene raziskave [75] smo se v naši raziskavi [70] ukvarjali z odvisnostjo diferenčne ločljivosti po azimutu od trajanja zvokovnega signala. Meritve smo izvajali s pomočjo slušalk, za doseg prostorskega učinka smo uporabili HRTF knjižnice MIT [20]. Za testni signal smo uporabili beli šum filtriran s pasovnim filtrom 350 Hz – 8000 Hz, nivo glasnosti je bil približno 65 dB *SPL*, trajanje signala je bilo 5 ms – 300 ms. Poskus smo izvajali v treh različnih smereh (azimuti 0°, -45° in -90°) v horizontalni ravnini (elevacija 0°). Ugotovili smo, da je ločljivost v smeri pred poslušalcem (azimut 0°) neodvisna od trajanja zvokovnega signala, pri ostalih dveh preizkušanih lokacijah (azimut -45° in -90°) pa je ločljivost začela upadati, ko smo trajanje zvoka zmanjšali pod 100 ms.

4.8 Komentar

V tem poglavju smo se ukvarjali z lokalizacijo izvorov zvoka pri ljudeh. Zanimala nas je točnost absolutne lokalizacije izvorov zvoka, diferenčna ločljivost izvorov zvoka ter vplivi spektra, trajanja in glasnosti zvoka na lokalizacijo. Najboljše rezultate lokalizacije smo dosegli v prostoru, kar je pričakovano, saj testne osebe večinoma niso imele predhodnih izkušenj s prostorskimi zvoki predvajanimi preko slušalk, hkrati pa je lokalizacija zvokov v prostoru za njih nekaj vsakdanjega.

Rezultati izmerjeni v prostoru dajejo podlago za oceno mej lokalizacije. Vrednostim, doseženim v prostoru, se pri predvajanju zvoka preko slušalk lahko z uporabo individualnih zvokov približamo, hkrati pa nastopijo dejavniki s psihološkimi značilnostmi. Slednje so povezane predvsem s človekovo izkušnjo (zamenjava smeri spredaj-zadaj, lokalizacija znanih zvokov – npr. ropot helikopterja). Kot smo videli, je pri reprodukciji zvoka na slušalkah in uporabi splošnih HRTF diferenčna ločljivost bistveno slabša, absolutna lokalizacija pa niti ni mogoča.

Ker so v poskusih sodelovale osebe z normalno razvitim vidom in sluhom, se postavlja vprašanje kakšne rezultate bi dosegli slepi ljudje. Čeprav v javnosti prevladuje mnenje, da imajo slepi ljudje bistveno boljše sposobnosti percepcije zvoka kot ljudje z normalno razvitim vidom [40], je ta hipoteza dvomljiva, rezultati v [74, 82] namreč ne potrjujejo tega. Celo nasprotno, avtorji raziskave [82] trdijo, da omenjena »splošna resnica« nima realne osnove.

5 ELEVACIJA – ANALIZA HRTF IN AKUSTIČNE ILUZIJE

5.1 Elevacija v akustični sliki prostora

Akustična slika v splošnem pomeni predvajanje prostorskih zvokov preko slušalk. To zajema različne aplikacije, vse od navigacijskih pripomočkov do sistemov navidezne in obogatene resničnosti. Raziskave, ki jih opisujemo v tem delu se nanašajo na razvoj orodja za pomoč slepim ljudem pri orientaciji v prostoru. Orodje lahko služi tudi kot pripomoček za zagotavljanje dodatnih orientacijskih in navigacijskih podatkov posebnim skupinam ljudi, ki delajo v oteženih okoliščinah (npr. rudarji, piloti, tankisti ipd.).

Generiranje prostorskih zvokov za poslušanje s slušalkami je zanimivo tudi za uporabo v multimedijskih aplikacijah – primere prostorsko bogatih zvokov lahko slišimo npr. v sodobnih računalniških igricah. Problem generiranja prostorskega zvoka za multimedijske aplikacije se od akustične slike, s katero se ukvarjamo tu, razlikuje predvsem v tem, da zvokov v multimedijskih aplikacijah ne moremo poljubno spreminjati, medtem ko je eno od pomembnih izhodišč pri snovanju akustične slike prav izbira ustreznega zvoka, saj smo pokazali, da imajo parametri zvokovnega signala (spekter, glasnost, trajanje ipd.) nezanemarljiv vpliv na lokalizacijo.

Ker notranje uho samo po sebi nima sposobnosti lokalizacije izvorov zvoka, je, kot smo že ugotovili, za lokalizacijo ključnega pomena vpliv posameznikovega telesa, ki zvok preoblikuje v odvisnosti od smeri iz katere prihaja zvočno valovanje. Za konstruiranje akustične slike oz. nasploh za generiranje prostorskih zvokov za poslušanje s slušalkami bi bilo zato optimalno uporabiti individualne HRTF [57]. Ker je merjenje individualnih HRTF zahteven postopek, tak pristop v splošnem ni mogoč. S postopki obdelave signalov se optimalni obliki prenosne funkcije lahko zgolj približamo. Za lokalizacijo po azimutu to povsem zadostuje, ker sta glavna dejavnika časovna in amplitudna razlika med ušesoma, ki ju lahko natančno določimo [67]. Za lokalizacijo po elevaciji pa to ne zadostuje, ker (še) ne poznamo postopka s katerim bi vnesli vse detajle, ki poslušalcu dajo avtentičen občutek. Zaradi neavtentičnosti oz. ne-individualnosti zvoka, ki ga poslušalcu predvajamo preko slušalk, je okrnjena zlasti točnost absolutne lokalizacije izvorov zvoka [65]. Dodatne možnosti, ki bi omilile izpostavljeni problem iščemo tudi z ustrežno izbiro zvoka, kjer gre v prvi vrsti za povezavo med elevacijo in frekvenco [59]. Povezava pojmov visoko in nizko v smislu »visoka frekvenca – visoka

elevacija izvora zvoka« je nekaj tako naravnega, da obstaja v skoraj vseh jezikih sveta [10].

V nadaljevanju tega dela se bomo ukvarjali s postopki obdelave signalov in izbiro zvokov za kodiranje elevacije v akustični sliki prostora. Posvetili se bomo torej ne-individualnim zvokom, ki sicer ne omogočajo absolutne lokalizacije navideznih izvorov zvoka po elevaciji, vendar pa se lahko poslušalec ob primerni ločljivosti posameznih zvokov nauči kje se posamezen izvor zvoka nahaja. Zato se lahko poleg značilnosti v zvoku, pomembnih za zaznavanje elevacije pri ustvarjanju akustične slike prostora opremo tudi na lastnost človeških možganov, ki so se sposobni prilagajati na različne zahteve. Ob upoštevanju ugotovitev psihologov, ki smo jih navedli v poglavju o učenju in glede na rezultate že delujočih sistemov za navigacijo slepih oseb s pomočjo zvoka [22, 34, 44, 45] sklepamo, da bi bilo po določenem času učenja možgane možno naučiti lokalizacije zvoka na poljuben način. Lokalizacijo izvorov zvoka, ki smo se jo naučili glede na pojave, ki so posledica širjenja zvoka v vsakdanjih razmerah, imenujemo naravna lokalizacija. Pri uporabi ne-individualnih zvokov za kodiranje elevacije lahko vpeljemo še pojem umetne lokalizacije, ki se ga možgani naučijo v skladu z vsiljeno predstavitvijo prostora. Pri učenju je čas ena bistvenih komponent, predpostavljamo pa, da predstavitev, ki je bližja zvokom vsakdanjega življenja skrajša čas učenja in hkrati ne povzroča dodatnih dvomnosti pri ločevanju zaznavanja položajev resničnih in navideznih izvorov zvoka.

5.2 Pregled nekaterih preizkušenih metod kodiranja elevacije v akustični sliki prostora

Za kodiranje elevacije oz. nasploh za ustvarjanje akustične slike se uporabljajo različni pristopi, od katerih nekateri v večji, drugi pa v manjši meri upoštevajo princip lokalizacije, ki ga je človeška vrsta razvila tekom evolucije. Ne glede na način kodiranja je za uporabnike akustične slike pomembno kolikšen je predviden čas učenja in kako prijetni ali moteči so zvoki, ki se uporabljajo v akustični sliki. V nadaljevanju si pogledjmo značilnosti dveh delujočih sistemov akustične slike [34, 44] in enega testnega sistema [22].

Akustična slika, opisana v [34], posnema princip orientacije, ki ga uporabljajo netopirji in torej uporablja popolnoma naraven način kodiranja (čeprav ni lasten človeku). Sistem vsebuje oddajnik in sprejemnik ultrazvoka. Oddajnik oddaja ultrazvok v različne smeri, sprejemnik pa sprejme odbito zvočno valovanje, izvede frekvenčno pretvorbo ter ga prek slušalk predvaja poslušalcu. Ta princip imenujemo princip odmeva ali preprosteje princip ehlokacije (angl. echo location) [4, 5, 34]. Gre za relativno preprost ter dokaj

učinkovit princip, ki pa zahteva dolgotrajno učenje. Zvok, ki ga sliši poslušalec je precej neprijeten, poleg tega pa uporaba tega sistema ne onemogoča poslušanja zvokov iz okolice. Avtorji trdijo, da se je slepa oseba z uporabo tega sistema zmožna npr. naučiti vožnje kolesa.

Akustična slika, ki jo je skonstruiral Meijer [44, 45] je prav tako relativno zahtevna za učenje in proizvaja poslušalcu precej neprijetne zvoke. Zaradi specifične tehnike snemanja prostora obstaja možnost, da akustična slika ne zazna hitro premikajočih se objektov. V tem sistemu je elevacija kodirana neposredno s frekvenco, tj. nizki toni za nizke elevacije in visoki toni za visoke elevacije. Akustična slika je po poročanju uporabnikov zelo uporaben in dobrodošel pripomoček za slepe ljudi.

Metoda, opisana v [22], ima veliko podobnosti z našim načrtom konstrukcije akustične slike prostora [68]. Gre za razdelitev prostora v podprostore, katerih velikost je odvisna od diferenčne ločljivosti v dani smeri. Ko se v nekem podprostoru pojavi ovira, poslušalec zasliši zvok, izhajajoč iz tega podprostora. Avtorji v [22] za kodiranje vseh dimenzij prostora predlagajo HRTF, kar je bistvena razlika glede na naš predlog, ki uvaja posebno kodiranje elevacije. V primerjavi z [22] v vertikalni smeri dolgoročno pričakujemo boljšo točnost lokalizacije in boljšo diferenčno ločljivost kot bi jo dosegli z uporabo splošnih HRTF.

5.3 Analiza HRTF

V tem razdelku opisujemo rezultate, ki smo jih dobili s preučevanjem HRTF. Če namreč velja, da je vsa informacija o položaju izvora zvoka vsebovana v zvoku, ki doseže poslušalčevo notranje uho, potem lahko zakonitosti lokalizacije izvorov zvoka ugotovljamo prav s študijem HRTF. Zvočno valovanje, ki prispe v poslušalčev sluhovod oz. naprej v notranje uho, je v praznem prostoru brez odbojnih površin odvisno izključno od poslušalca oz. njegovega telesa. Tako lahko ugotovimo, da ima človekovo telo oz. njegova oblika posebno vlogo pri lokalizaciji izvorov zvoka. Tu lahko še posebej izpostavimo obliko uhljev [24], ki opravljajo neke vrste »smerno filtriranje«.

Za obravnavo ločljivosti po elevaciji razdelimo spekter v tri pasove [1]:

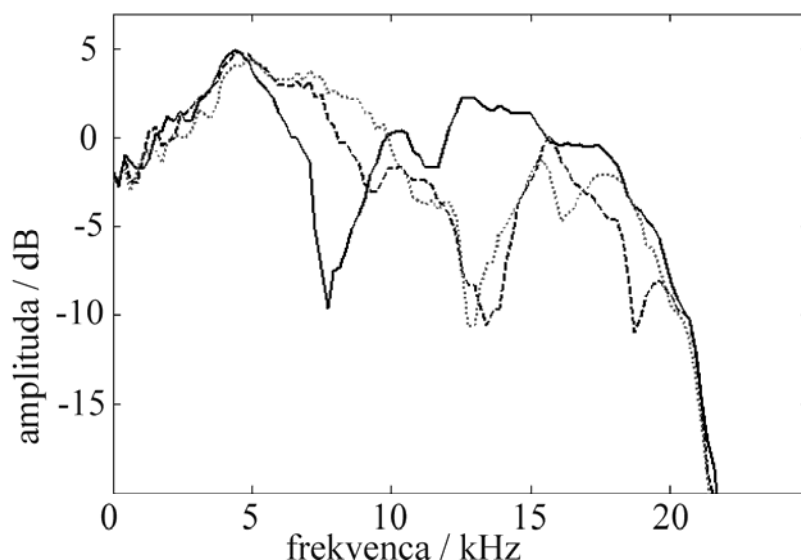
- nizke frekvence, < 4 kHz;
- srednje frekvence, 4 kHz – 16 kHz, pas 8 kHz - 16 kHz poimenujemo pas višjih frekvenc in
- visoke frekvence, > 16 kHz.

Za potrebe analiz HRTF, ki jih opisujemo v nadaljevanju tega poglavja smo uporabili funkcije iz knjižnice CIPIC [2].

5.3.1 Oslabitev uhlja (pinna notch)

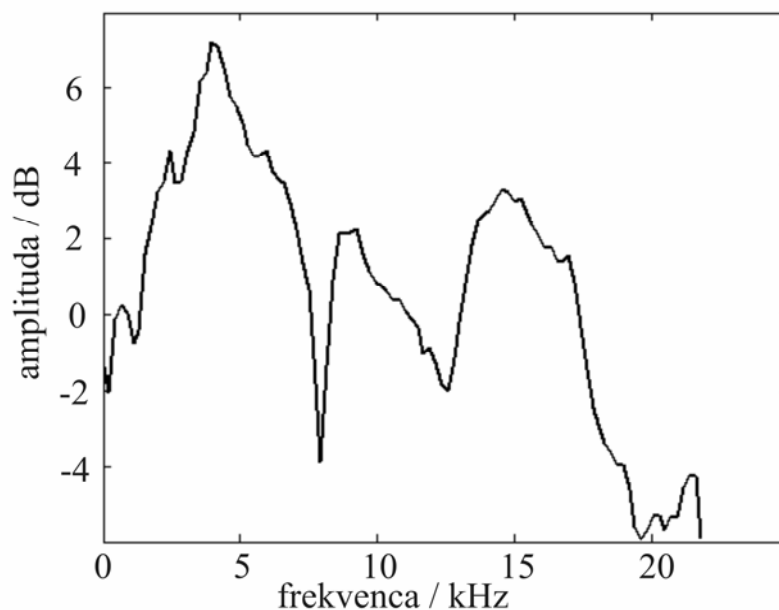
Glavna značilnost, ki jo viri [1, 31, 39] navajajo kot ključni faktor zaznavanja elevacije, je oslabitev v amplitudnem poteku HRTF. Ta oslabitev je posledica vpliva uhlja in jo zato poimenujemo oslabitev uhlja (angl. pinna notch). Oslabitev uhlja je pravzaprav več, odboji zvočnega valovanja v uhlju namreč povzročijo več frekvenčnih lukenj [57].

Karakteristike oslabitev (centralna frekvenca, pasovna širina, nivo oslabitve) so odvisne od vsakega posameznika, centralna frekvenca prve (glavne) oslabitve v amplitudnem poteku HRTF pa je naraščajoča funkcija elevacije (izjemoma je pri nekaterih osebah funkcija v krajših intervalih tudi padajoča). Pri spreminjanju elevacije med -45° in 90° ležijo centralne frekvence oslabitev uhlja med 5 kHz in 17 kHz [57]. Spreminjanje centralne frekvence slabljenja uhlja lahko opazujemo na sliki 5.1, ki prikazuje poteke treh HRTF iste osebe (lutka KEMAR – CIPIC oseba 165) pri azimutu 0° in različnih elevacijah.



Slika 5.1: Potek amplitudnega spektra HRTF za osebo z oznako 165 (tj. lutka KEMAR) iz knjižnice CIPIC pri azimutu 0° in elevacijah 5° (polna črta), 28° (temna pikčasta črta) in 62° (svetla pikčasta črta).

Kot smo omenili, zaradi odbojev v uhlju nastane več oslabitev. Za razliko od slike 5.1, kjer v vsaki HRTF nastopa le ena izrazita oslabitev uhlja, je v funkciji HRTF na sliki 5.2 več oslabitev. Poleg oslabitev uhlja nastopajo še oslabitve, ki so posledica odbojev od drugih delov telesa, vendar so te manj izrazite.

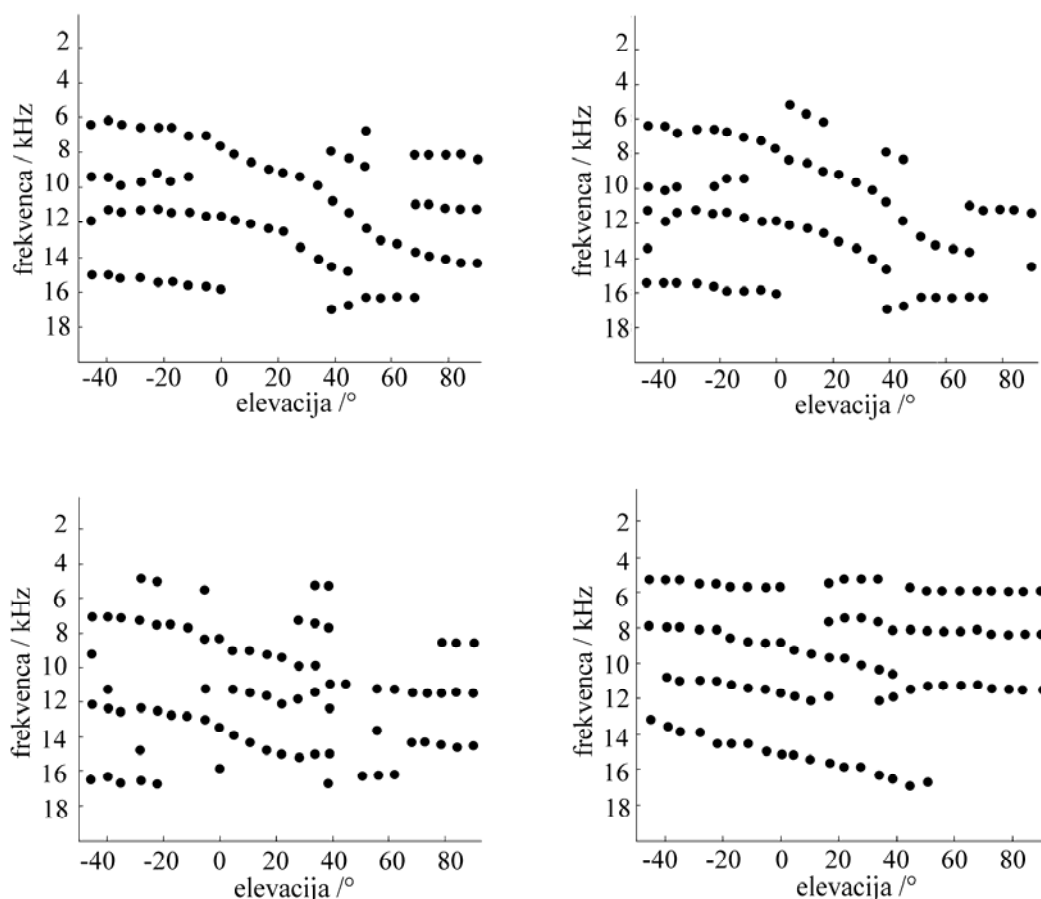


Slika 5.2: Potek amplitudnega spektra HRTF (CIPIC oseba 131) pri azimutu -45° , elevaciji 0° .

Določanje frekvenc oslabitev je mogoče z različnimi postopki obdelave signalov. Nekateri postopki temeljijo na analizi amplitudnega poteka spektra (npr. z linearno predikcijo), v zadnjem času pa je velike pozornosti deležen postopek, ki deluje na osnovi skupinskih zakasnitev [57, 58]. Ker smo ta postopek uporabili tudi v naši analizi, tu na kratko opisujemo algoritem:

- Analizo izvajamo v časovnem prostoru, zato uporabimo impulzne odzive HRIR, ki jih poravnamo tako, da jim odrežemo začetne časovne zakasnitve [64].
- Nato signal filtriramo z linearno prediktivnim filtrom reda 10 – 12 in rezultat filtriranja odštejemo od nefiltriranega signala. Rezultat (razlika oz. napaka) predstavlja tisti del signala, ki ga z metodo linearne predikcije ne moremo napovedati. V naši analizi smo uporabili filter reda 12.
- Za izločitev resonanc in neizrazitih oslabitev, ki niso posledica uhlja, signal oknimo s polovičnim Hannovim [57] ali polovičnim Kaiserjevim [58] oknom dolžine 1 ms (dolžina okna ni kritičen parameter, vendar se najpomembnejše stvari zgodijo približno v 1 ms; polovično okno pomeni, da ne uporabimo celega okna, pač pa le drugo polovico). Oknjenje v splošnem zmanjša frekvenčno ločljivost in vnese popačitve, ki lahko vplivajo na točnost določanja frekvence oslabitve, zato smo te napake kompenzirali z linearno prediktivnim filtriranjem, izvedenim v prejšnji točki algoritma. Uporabili smo Hannovo okno.

- Nato izračunamo avtokorelacijo oknjenege signala.
- Rezultat avtokorelacije ponovno povprečimo oz. oknimo s polovičnim Hannovim [57] oknom dolžine 1 ms ali Kaiserjevim oknom dolžine 0.7 ms [58]. Uporabili smo Hannovo okno.
- Izračunamo Fourierov transform oknjenege signala in odvajamo njegov fazni potek. Odvod faznega poteka je skupinska zakasnitev $\tau(\omega) = -\frac{d\theta(\omega)}{d\omega}$. Iskane frekvence oslabitev uhlja so lokalni minimumi, ki imajo skupinsko zakasnitev manjšo od 0 s. V praksi se je izkazalo, da je bolje vzeti minimume s skupinsko zakasnitvijo manjšo od -1 s, da se izločijo morebitni produkti oknjenja [57], poleg tega moramo izločiti še vse minimume, ki so sicer ustrezni vendar imajo frekvenco manjšo od 3 kHz.



Slika 5.3: Odvisnost frekvenc oslabitve uhlja v odvisnosti od elevacije za lutko KEMAR (oseba 165 v knjižnici CIPIC) – levo uho. Grafi prikazujejo situacijo pri azimutih 0° (levo zgoraj), -15° (desno zgoraj), -45° (levo spodaj) in -65° (desno spodaj).

Namesto neposrednega računanja odvoda faznega poteka, določimo skupinsko zakasnitev s pomočjo Fourierovega transformira funkcij $x[n]$ in $n \cdot x[n]$:

$$\begin{aligned} X(\omega) &= \sum_{n=0}^{N-1} x[n] e^{-j\omega n} = X_R(\omega) + jX_I(\omega) \\ Y(\omega) &= \sum_{n=0}^{N-1} nx[n] e^{-j\omega n} = Y_R(\omega) + jY_I(\omega) \end{aligned} \quad (5.1)$$

Od tod sledi [57], da je skupinska zakasnitev $\tau(\omega)$:

$$\tau(\omega) = -\frac{d\theta(\omega)}{d\omega} = \frac{X_R(\omega)Y_R(\omega) + X_I(\omega)Y_I(\omega)}{X_R^2(\omega) + X_I^2(\omega)} \quad (5.2)$$

Slika 5.3 prikazuje odvisnost frekvence oslabitev uhlja v odvisnosti od elevacije za nezasenčeno uho lutke KEMAR (oseba 165 knjižnice CIPIC) pri različnih azimutih (0° , -15° , -45° , -65°). V nadaljevanju nas bo zanimala predvsem prva oslabitev uhlja, o kateri smo že zapisali, da njena frekvenca (z nekaj izjemami) narašča v odvisnosti od elevacije.

Pomen oslabitve uhlja potrjujejo tudi druge raziskave [46, 78], kjer so ugotovili, da je za zaznavanje elevacije bistvenega pomena potek amplitudnega dela spektra v pasu med 5.7 kHz in 11.3 kHz. Vendar amplitudni potek spektra v omenjenem frekvenčnem pasu ne predstavlja edinega faktorja za ločevanja elevacije izvora zvoka, saj je elevacijo izvora zvoka možno določiti tudi za zvoke, ki ne vsebujejo komponent v tem delu spektra [1]. Na zvoke frekvenc pod 3 kHz vplivajo predvsem odboji od telesa in lom valovanja zaradi glave. Te motnje so izrazitejše izven sagitalne ravnine. Kljub temu, da elevacijo izvora zvoka lahko ugotovimo pri zvokih nižjih frekvenc, pa prisotnost višjih frekvenčnih komponent izboljšuje natančnost lokalizacije [1].

5.3.2 Energija HRTF

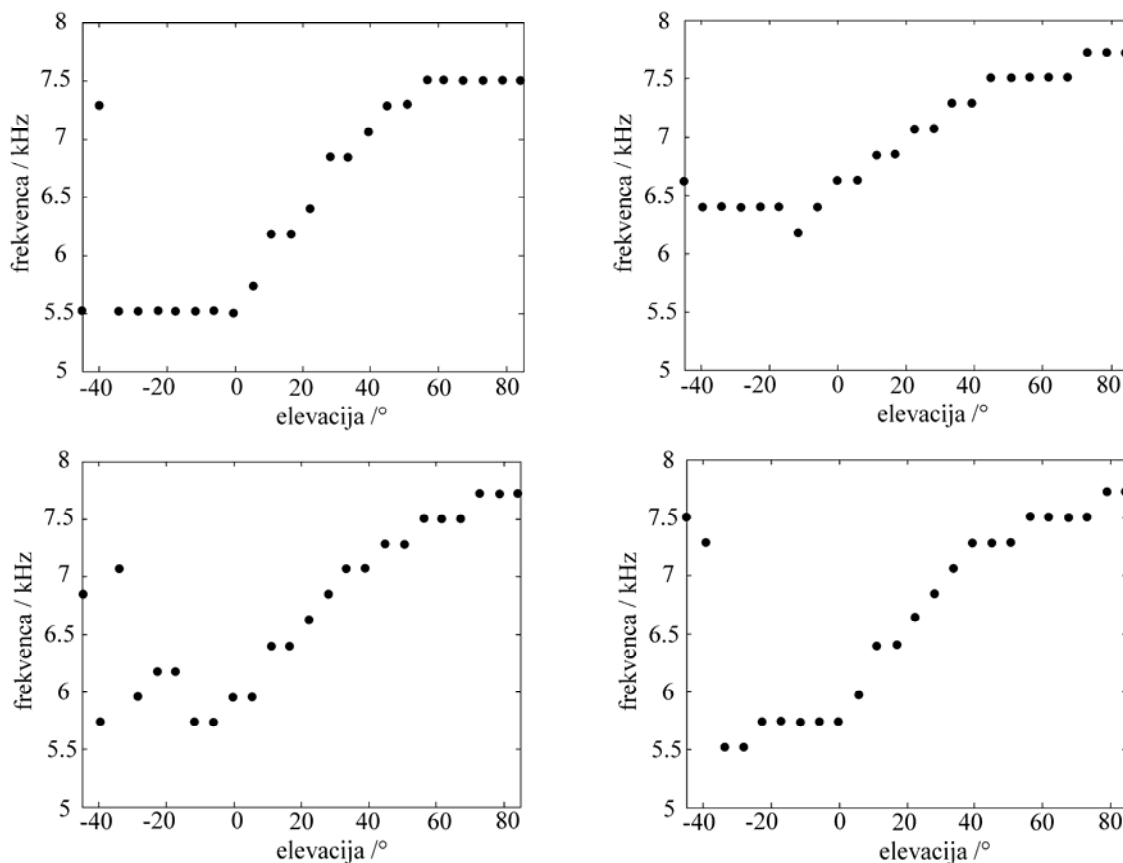
Že s površnim pregledom amplitudnih potekov HRTF pri azimutih med -80° in 80° ter elevacijah med -45° in 90° ugotovimo, da je večina energije koncentrirane pri nizkih (< 4 kHz) in srednjih frekvencah (4 kHz – 16 kHz). V nadaljevanju bomo za posamezne elevacije ugotavljali frekvenčni pas v katerem se nahaja večina energije ter poskusili ugotoviti, če v zvezi s tem obstaja kaka zakonitost.

Zvočno valovanje, ki ga uho prestreže kot posledico oddanega enotnega zvočnega impulza obravnavamo kot širokopasovni (zvočni) signal, kar je razvidno tudi iz amplitudnega poteka pripadajočega HRTF. Minimalna zaznavna razlika v jakosti zvoka (angl. just noticeable difference of intensity) je enaka za ozkopasovne in širokopasovne zvočne signale [47] in je ocenjena na 0.41 dB za zvoke z jakostjo večjo od 30 dB. Slednji pogoj je v naših raziskavah povsod izpolnjen.

Kot smo omenili v poglavju o slabljenju uhlja, se centralna frekvenca oslabitve spreminja v odvisnosti od elevacije. Ker ima ta odvisnost vpliv na razporeditev energije

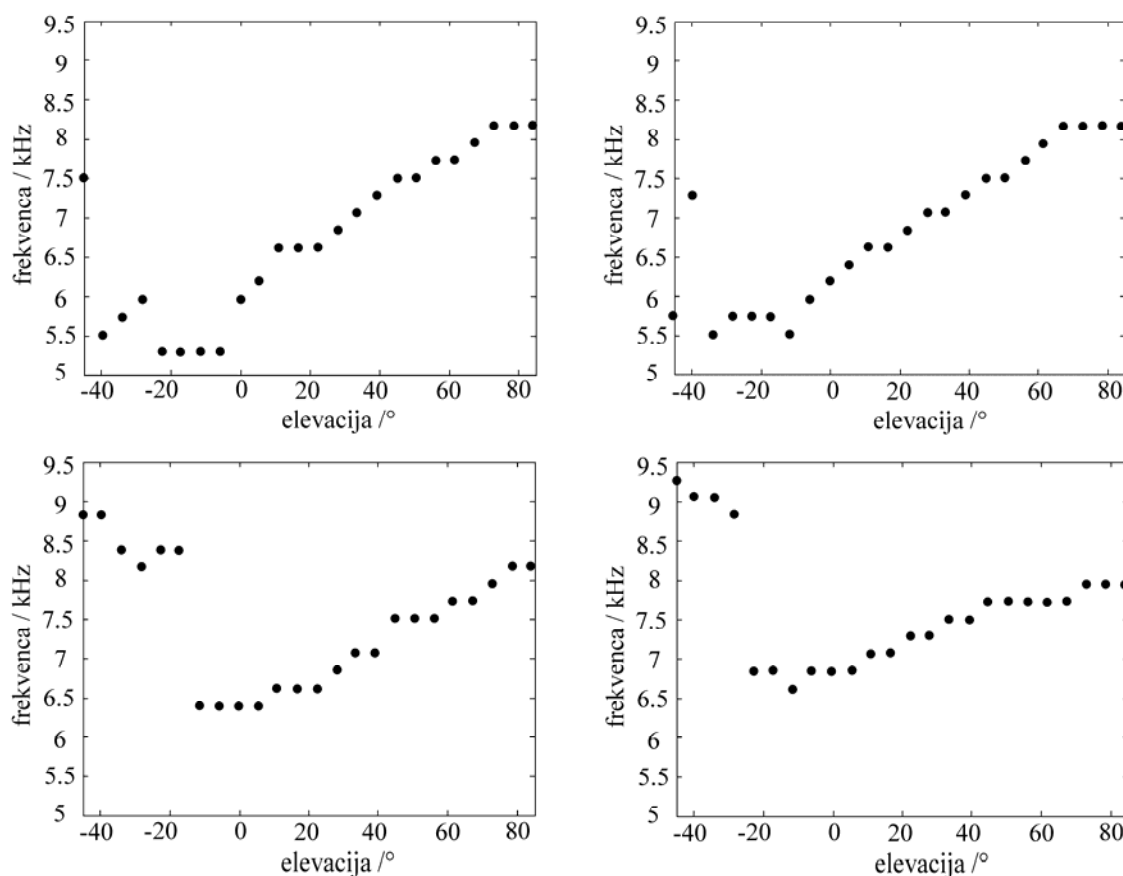
signala v odvisnosti od frekvence, nas bo zanimala zgornja meja frekvenčnega pasu v katerem je skoncentrirana večina energije, pri čemer za računanje energije signal (tj. ustrezen HRTF) obtežimo z utežjo karakteristike A standarda ISO 226 (več o utežnih karakteristikah smo napisali v poglavju 3.1). Zgornjo frekvenčno mejo določimo s pogojem, da je količina energije nad to mejo zanemarljiva, tj. ni zaznavna. V frekv. pasu med 0 in zgornjo frekvenčno mejo je torej skoncentrirano 91 % (-0.41 dB) vse energije odziva na enotin zvočni impulz za posamezno lokacijo, ki je določena z elevacijo in azimutom.

Sliki 5.4 in 5.5 prikazujeta zgornjo frekvenčno mejo v odvisnosti od elevacije pri različnih azimutih za dve osebi – knjižnica CIPIC, lutka KEMAR (oznaka 165) in oseba z oznako 119. Funkcijske odvisnosti so podane za nezasenčeno uho, tj. uho, na katero zvočno valovanje vpada neposredno.



Slika 5.4: Frekvenčne meje do katere je zbrane 91 % vse energije HRTF posamezne elevacije. Odvisnosti frekvenčne meje od elevacije veljajo za lutko KEMAR (oseba 165 knjižnice CIPIC) in so podane za naslednje azimute: 0° (levo zgoraj), -10° (desno zgoraj), -20° (levo spodaj) in -55° (desno spodaj).

Na sliki opazimo naraščajoč trend, zgornja frekvenčna meja se z večanjem elevacije zvišuje, kar pomeni, da je pri nizkih elevacijah energija koncentrirana pretežno pri nizkih frekvencah, z višanjem elevacije pa se razporedi po širšem frekvenčnem pasu. O naraščanju lahko govorimo le v globalnem smislu, na posameznih delih opazimo tudi padanje, ki je še posebej značilno za elevacije manjše od 0° , tj. pod horizontalno ravnino. Primerjava zgornjih frekvenčnih mej v odvisnosti od elevacije za vse funkcije knjižnice CIPIC pokaže, da v splošnem res velja globalni trend naraščanja, lokalna naraščanja in upadanja pa so značilna zlasti za nizke in zelo visoke elevacije. Lokalni poteki so dokaj naključni in ne najdemo skupnih zakonitosti, ki bi veljale za vse osebkke. Zato sklepamo, da je ta naključnost posledica razširjanja valovanja oz. vpliv človekovega telesa. Pri nekaterih funkcijah oz. osebkih, kot npr. oseba z oznako 60, so ti vplivi še posebej izraziti, pri ostalih, kot vidimo tudi na slikah 5.4 in 5.5, pa precej manj.



Slika 5.5: Frekvenčna meja do katere je zbrane 91 % vse energije HRTF posamezne elevacije. Odvisnosti frekvenčne meje od elevacije so podane za osebo 119 iz knjižnice CIPIC za naslednje azimute: 0° (levo zgoraj), -10° (desno zgoraj), -20° (levo spodaj) in -55° (desno spodaj).

Kot smo omenili v uvodu tega razdelka, je energija skoncentrirana pri nizkih frekvencah. V nasprotju s centralno frekvenco oslavitve uhlja, ki se v odvisnosti od elevacije spreminja med 5 kHz in 17 kHz, zgornja frekvenčna meja (91 % energije) leži v vseh primerih bistveno nižje – med 4 kHz in 6.5 kHz. Pri razlagi slednjega si lahko pomagamo s primerjanjem amplitudnih potekov HRTF pri fiksnih azimutih in različnih elevacijah, kjer ugotovimo relativno podoben amplitudni potek med 0 in neko frekvenco, ki je v primeru azimuta 0° približno 4.5 kHz. V okolici te frekvence vsi amplitudni poteki dosežejo maksimum, nato pa začno upadati ter dosežejo minimum pri oslavitvi uhlja. Pri približno 85 % HRTF je to globalni maksimum, v ostalih primerih pa nivo tega maksimum ni dosti nižji kot nivo globalnega maksimuma. V večini ostalih 15 % primerov globalni maksimum leži pri frekvencah nad 10 kHz.

5.4 Akustične iluzije v elevaciji

V raziskavah, ki se ukvarjajo z obdelavo zvokovnih signalov z namenom ustvarjanja akustičnih iluzij oz. navideznih izvorov zvoka (tj. prepričati poslušalca, da sliši izvor zvoka na drugi lokaciji kot je v resnici) omejenih na elevacijo, obdelava zvokovnih signalov temelji na postopkih frekvenčne manipulacije. Gre za filtriranje z resonatorji, s filtri z zarezo in pasovno zapornimi filtri (posnemanje oslavitve uhlja) ali za neposredno preslikavo med elevacijo in frekvenco [4, 44, 59].

Glavna razlika med ustvarjanjem akustičnih iluzij v prostoru ali s slušalkami je, da na zvok, ki ga sliši poslušalec pri poslušanju v prostoru, vpliva prenosna pot, ki »vsiljuje« dejansko in ne navidezno lokacijo izvora. To pomeni, da morajo biti v zvočnem valovanju, ki ga prestreže poslušalec lastnosti navidezne lokacije prevladujoče nad lastnostmi, ki jih prispeva dejanska prenosna pot.

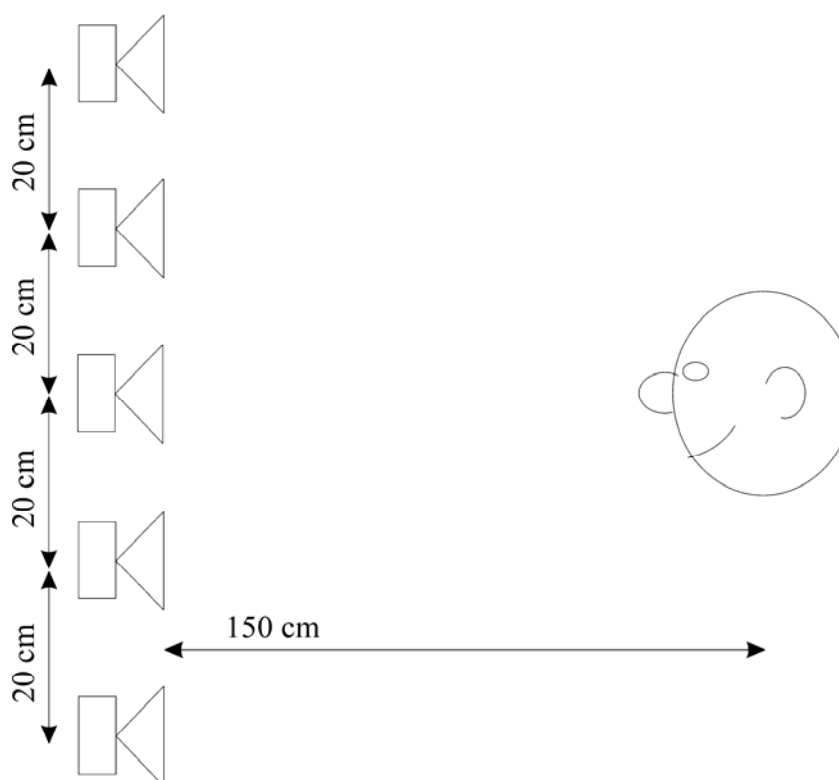
Odgovor na vprašanje ali je s frekvenčno manipulacijo zvokovnega signala možno navidezno premikati izvor zvoka v prostoru, če uporabimo zvočnik, ki je postavljen na fiksni poziciji nam daje npr. raziskava [5]. V tej raziskavi so uporabili šum z ravnim amplitudnim potekom med 6 kHz in 12 kHz, ki so ga pri različnih frekvencah filtrirali s filtrom z zarezo (notch filter) ter ga predvajali po zvočniku, ki je bil nameščen na lokaciji (azimut, elevacija) = (90°, 60°). Po drugem zvočniku so predvajali šum z ravnim amplitudnim potekom med 6 kHz in 12 kHz. Ta zvočnik je bil premičen in je služil za ugotavljanje navideznega položaja zvočnika, ki je predvajal filtriran signal. Z opisanim načinom obdelave šuma so dosegli navidezno premikanje izvora zvoka, pri čemer je bila ločljivost 15°. Podoben poskus, opisan v točki 5.4.1 smo izvedli tudi sami.

Pri ustvarjanju akustičnih iluzij za poslušanje s slušalkami pa je največja težava ta, da so vsi modeli bolj ali manj ne-individualni, zaradi ne-individualnosti sintetiziranih zvokov pa ne moremo ustvariti absolutne lokalizacije. To pomeni, da poslušalec pri poslušanju

dveh zvokov sicer ugotovi, kateri pripada izvoru z višjo in kateri izvoru z nižjo elevacijo, ne more pa določiti absolutnega položaja izvora zvoka. Absolutni položaj elevacije izvora zvoka lahko poslušalec ugotovi le, če je zvok obdelan z individualnimi HRTF ali če se je predhodno že naučil »umetnega« načina kodiranja.

5.4.1 Akustične iluzije v prostoru

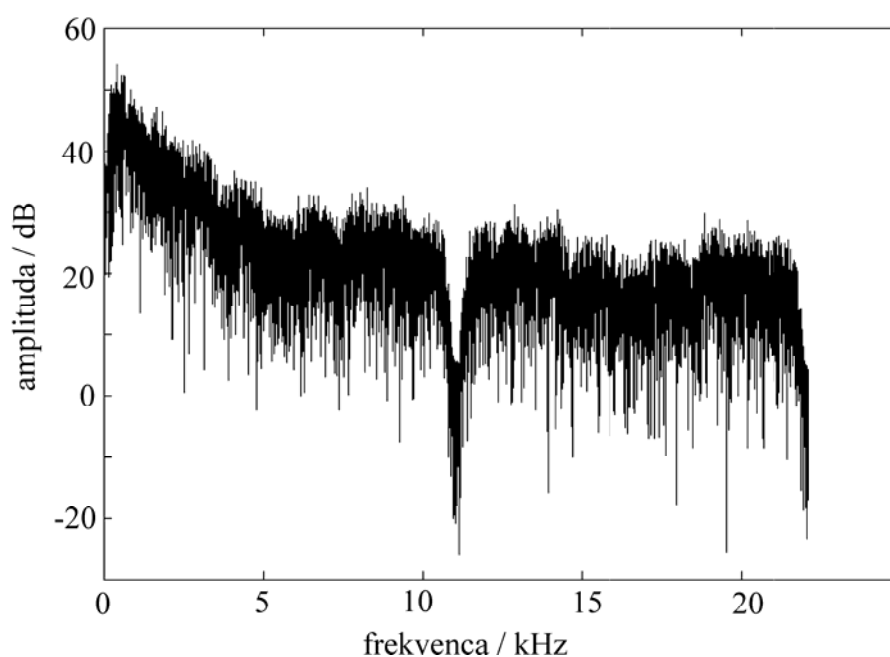
Podobno kot v [5] smo izvedli poskus ustvarjanja akustičnih iluzij z zvočniki v prostoru. Ker nas je zanimalo predvsem ali je akustične iluzije v vertikalni smeri sploh možno ustvariti in če je to možno v nekem običajnem okolju, smo za kraj poskusa izbrali prostore Laboratorija za komunikacijske naprave na Fakulteti za elektrotehniko, kjer je jakost hrupa 40 dB, merjeno po lestvici *A* z merilnikom Lutron SL-4012. Uporabili smo pet navadnih računalniških zvočnikov Genius (model: SP-G06), ki smo jih zložili v vertikalni smeri, 20 cm vsaksebi kot kaže slika 5.6. Poslušalec je bil od zvočnikov oddaljen 150 cm. V poskusu je sodelovalo 8 oseb, ki po njihovih lastnih zagotovilih niso imeli težav s sluhom ali vidom.



Slika 5.6: Postavitev zvočnikov pri poskusu ustvarjanja akustičnih iluzij v vertikalni smeri v prostoru. Zvočniki so postavljeni v sagitalni ravnini, med centri zvočniških membran je razmik 20 cm, testna oseba je od zvočniške konstrukcije oddaljena 150 cm.

Od petih zvočnikov, ki jih je videl poslušalec, smo zvok vedno predvajali le na srednjem zvočniku (lokacija glede na poslušalca: azimut 0° , elevacija 0°). Uporaba več zvočnikov je tako le vizualna iluzija, ki glede na vizualno teorijo (vpliv vizualne percepcije na lokalizacije izvorov) zagotavlja zanesljivejšo ustvarjanje iluzij. Poskus smo izvedli s petimi nabori zvokov od katerih je vsak nabor vseboval tri zvoke. Zvoki posameznega nabora so bili rezultat obdelave istega osnovnega zvoka. Vsak od zvokov posameznega nabora naj bi predstavljal svojo elevacijo. Uporabili smo naslednje zvoke in tehnike obdelave signalov:

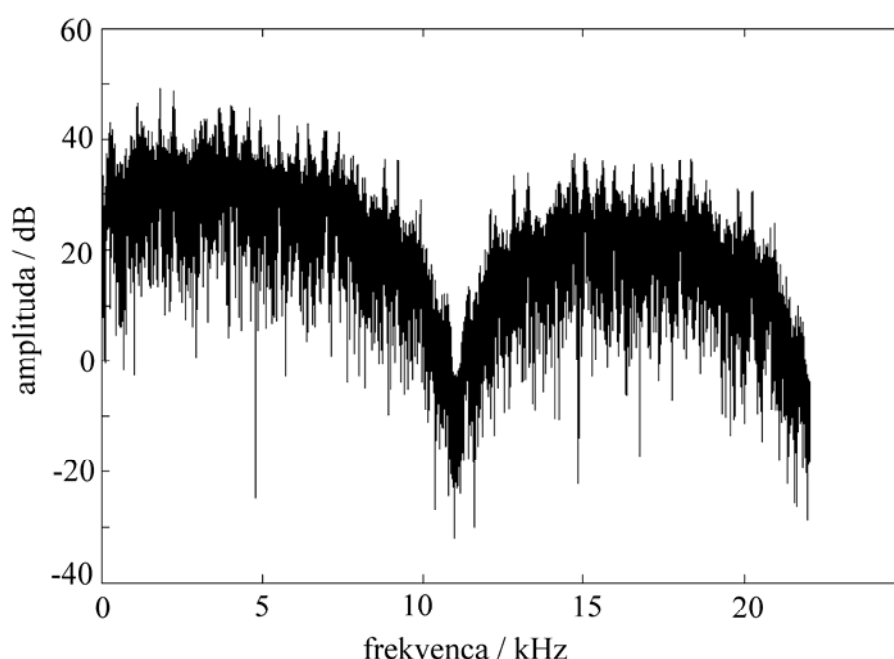
- a) zvokovni signali iz narave (zvok ob puku puške in zvok pri razbitju kozarca, spektra signalov sta prikazana na slikah 5.7 in 5.8): obdelava je bila izvedena z množenjem osnovnega signala s sinusnim signalom različnih frekvenc (za potrebe poskusa smo uporabili sinusa s frekvencama 5 kHz in 10 kHz, tretji zvok nabora je bila osnovna oblika zvoka), ideja o frekvenčni prestavitvi osnovnega spektra izvira iz povezave spektra zvoka in elevacije [44, 59];



Slika 5.7: Spekter zvokovnega signala, ki nastane ob puku puške je eden izmed dveh zvokov iz narave, ki smo jih uporabili pri poskusu ustvarjanja akustičnih iluzij v elevaciji.

- b) šum z ravnim amplitudnim potekom (beli šum): za obdelavo smo uporabili nizko prepustni filter s spremenljivo zgornjo frekvenco (za potrebe poskusa smo uporabili tri filtre z zgornjimi frekvencami 5 kHz, 10 kHz in 20 kHz) in takim ojačanjem, da so bile energije filtriranih signalov vedno enake; ideja postopka izvira iz porazdelitve energije HRTF v odvisnosti od elevacije, ki smo jo opisali v razdelku 5.3.2;

- c) rožnati šum: obdelavo smo izvedli s filtriranjem z nizko prepustnim filtrom s spremenljivo zgornjo frekvenco (za potrebe poskusa smo uporabili tri filtre z zgornjimi frekvencami 5 kHz, 10 kHz in 20 kHz) in takim ojačanjem, da so bile energije filtriranih signalov vedno enake; ideja postopka izvira iz porazdelitve energije HRTF v odvisnosti od elevacije, ki smo jo opisali v razdelku 5.3.2;
- d) rožnati šum: obdelava je bila izvedena z množenjem rožnatega šuma s sinusnim signalom različnih frekvenc (za potrebe poskusa smo uporabili sinusa s frekvencama 5 kHz in 10 kHz, tretji zvokovni signal je bil rožnati šum v osnovni obliki), tudi tu ideja o frekvenčni prestavitvi osnovnega spektra izvira iz povezave frekvence tona in elevacije [44, 59].



Slika 5.8: Spekter zvokovnega signala, ki nastane ob razbitju kozarca je eden izmed dveh zvokov iz narave, ki smo jih uporabili pri poskusu ustvarjanja akustičnih iluzij v elevaciji.

Trajanje posameznega zvokovnega signala je bilo 100 ms, pri predvajanju smo med posamezne zvoke vključili 1 s trajajočo tišino. Preizkusni postopek se je razlikoval glede na uporabljeno tehniko obdelave signalov – vrstni red predvajanja zvokov smo določili glede na predvideno navidezno elevacijo, ki naj bi jo zvok izzval pri poslušalcu, pri čemer smo se oprli na povezavo elevacije in frekvenčne vsebine zvoka [59]. V primeru obdelave signala z množenjem smo predvideli, da najnižjo elevacijo zavzame osnovna oblika zvoka, višjo elevacijo zvok pomnožen s sinusom 5 kHz in najvišjo zvok pomnožen s sinusom 10 kHz. Za zvoke obdelane s filtrom smo ob upoštevanju enakih predpostavk predvideli, da najnižjo elevacijo izzove zvok filtriran s filtrom zgornje frekvence 5 kHz, najvišjo elevacijo zvok filtriran s filtrom zgornje frekvence 20 kHz, vmesno pa zvok

filtriran s filtrom zgornje frekvence 10 kHz. Poslušalcem smo predvajali zvoke v zaporedju od »najnižje« do najvišje »elevacije«.

Zaporedje za zvoke, ki smo jih dobili s pomočjo množenja je bil (ponavljajoči) vrstni red naslednji:

1. osnovna oblika zvokovnega signala;
2. zvokovni signal obdelan (množen) s sinusom frekvence 5 kHz;
3. zvokovni signal obdelan (množen) s sinusom frekvence 10 kHz;
4. zvokovni signal obdelan (množen) s sinusom frekvence 5 kHz;
5. osnovna oblika zvokovnega signala.

V primeru obdelave z nizkoprepustnim filtrom pa je bil (ponavljajoči) vrstni red predvajanja zvokov naslednji:

1. signal obdelan (filtriran) s filtrom zgornje frekvence 5 kHz;
2. signal obdelan (filtriran) s filtrom zgornje frekvence 10 kHz;
3. signal obdelan (filtriran) s filtrom zgornje frekvence 20 kHz;
4. signal obdelan (filtriran) s filtrom zgornje frekvence 10 kHz;
5. signal obdelan (filtriran) s filtrom zgornje frekvence 5 kHz;

Zaporedje, opisano v zgornjih točkah, smo vsakemu poslušalcu ponovili 3 krat. Naloga poslušalcev je bila, da z navzgor, navzdol ali vodoravno usmerjenim palcem nakažejo (navidezno) spremembo lokacije izvora zvoka. Poslušalce, ki so vsaj dvakrat pravilno sledili celotno zaporedje smo uvrstili v »uspešno«, ostale pa v »neuspešno«. Rezultate podajata tabela 5.1 in graf na sliki 5.9.

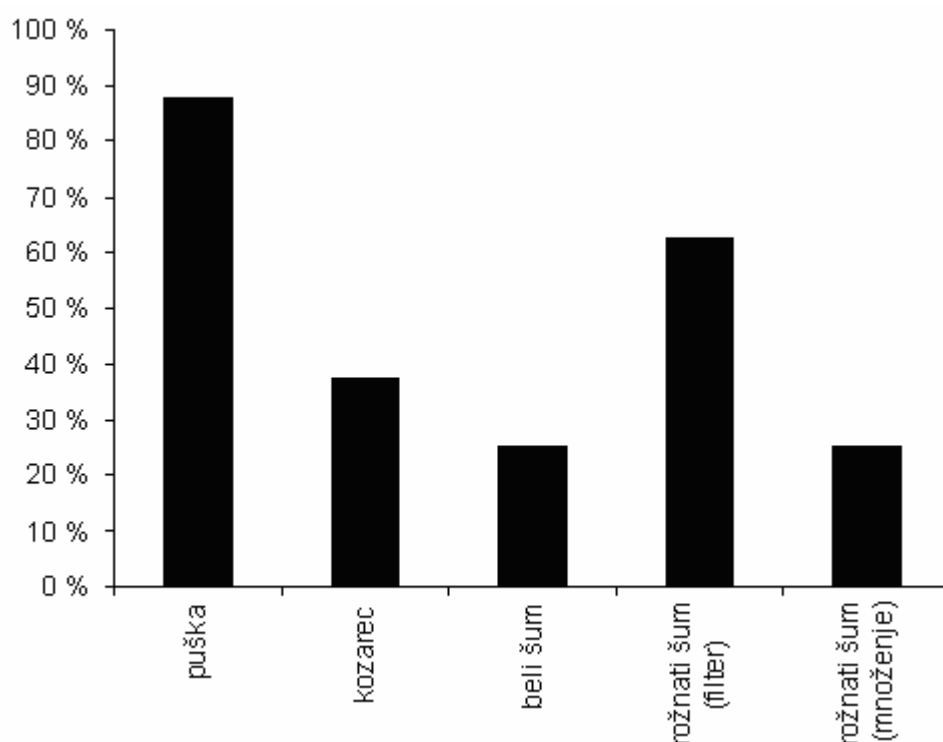
pok puške		razbitje kozarca		beli šum		rožnati šum (filtriran)		rožnati šum (množen)	
usp.	neusp.	usp.	neusp.	usp.	neusp.	usp.	neusp.	usp.	neusp.
7	1	3	5	2	6	5	3	2	6

Tabela 5.1: Število uspešnih oz. neuspešnih sledenj zaporedju navideznih premikov izvora zvoka. Z vsako testno osebo smo izvedli tri ponovitve za vsak zvok. Za uvrstitev v »uspešno« je testna oseba morala vsaj dvakrat pravilno slediti celotno zaporedje navideznega spreminjanja elevacije izvora posameznega zvoka.

Pred izvajanjem poskusa, tj. pred poslušanjem prvega nabora zvokov, smo vsakemu poslušalcu predvajali uvajalno proceduro, kjer smo uporabili osnovne oblike zvokov (oz. šume s pasovno širino 20 kHz), ki smo jih predvajali na posameznih zvočnikih, od tistega z najnižjo elevacijo do najvišjega in nazaj. To zaporedje smo ponovili trikrat. Uvajalna procedura s posameznim zvokom je trajala približno 30 s. Po koncu uvajalne procedure je testna oseba imela 5 min premor. Prav tako je 5 min odmor sledil tudi poslušanju posameznega nabora zvokov, tj. pred poslušanjem naslednjega nabora.

Rezultati kažejo, da je ustvarjanje akustičnih iluzij v vertikalni smeri v prostoru izvedljivo, vendar močno odvisno od izbire zvoka in postopka obdelave signalov. Najboljši rezultat je bil dosežen z zvokom puške, ki smo ga množili s sinusnimi signali različnih frekvenc. V spektru signala (slika 5.7) opazimo, da so izrazitejše komponente pri nižjih frekvencah, ki jih z množenjem premikamo po frekvenčni osi. Mehanizem iluzije je potemtakem povezava dominantnega dela frekvenčnega spektra in elevacije, kot je med drugim ugotovil tudi Blauert, ki je eksperimentiral z ozkopasovnimi zvokovnimi signali [3].

S spektrom, ki smo ga ravnokar omenili, ima nekaj podobnosti spekter rožnatega šuma, vendar množenje rožnatega šuma in sinusnih signalov ni povzročilo prepričljivih akustičnih iluzij. Vzrok bi bil lahko preozek dominantni del frekvenčnega spektra, pa tudi spektralno siromašnejši in manj naraven zvok kot zvok puške. Akustične iluzije, izvedene z zvokom ob razbitju kozarca, so zanemarljivo bolj prepričljive kot v primeru prej omenjenega rožnatega šuma. Sklepamo, da to majhno prednost zagotovi naravnost zvoka.



Slika 5.9: Uspešnost akustičnih iluzij ustvarjenih z zvočniki v vertikalni smeri v odvisnosti od zvokovnega signala in postopka obdelave. Graf podaja delež testnih oseb, ki so sintetizirane (virtualne) izvore zvoka lokalizirali na navidezni in ne na dejanski lokaciji.

Filtriranje šumov z nizko prepustnim filtrom in s tem povezana zastopanost spektralnih komponent je eden izmed mehanizmov, ki smo ga uporabili tudi pri poskusih

s slušalkami. Ustvarjanje akustičnih iluzij v vertikalni smeri v prostoru je bilo prepričljivejše pri uporabi rožnatega šuma kot pri uporabi belega šuma.

5.5 Komentar

Preučevanje vplivov, ki vnašajo spremembe v zvočno valovanje, kar je ključno za lokalizacijo, je silno živahna raziskovalna disciplina. Poleg analize je veliko pozornosti namenjeno tudi sintezi, predvsem gre za iskanje modelov za simulacijo HRTF. Glavni problem, ki zaenkrat ostaja nerešen, je ne-individualnost na tak način pridobljenih smernih funkcij. V smeri reševanja tega problema so zanimiva zlasti modeliranja vplivov glave, telesa in uhljev na zvočno valovanje, saj se fizične karakteristike posameznikovih delov telesa bolj ali manj razlikujejo.

Ker smo v naši raziskavi problem lokalizacije razdelili na dva neodvisna dela, se z zgoraj opisanimi izzivi nismo ukvarjali. Ker nas zanimajo specifični modeli, povezani zgolj z elevacijo, smo se tudi pri analizi HRTF osredotočili zgolj na to dimenzijo. Tako oslabitev uhlja kot povezave med spektrom in elevacijo, ki smo ju preučevali, sta predmet mnogih raziskav, kljub temu pa do sedaj v literaturi nismo zasledili poskusov, ki bi izhajali iz razporeditve energije po frekvenčni osi na način kot smo ga predstavili tu.

6 RAZVOJ IN OVREDNOTENJE MODELA ZA KODIRANJE ELEVACIJE

V skladu z zastavljenim ciljem, tj. uvedbo umetnega kodiranja elevacije v akustično sliko prostora, bomo v tem poglavju podrobneje predstavili konkreten poskus kodiranja elevacije v enodimenzionalni akustični sliki, rezultate, ki so jih dosegle testne osebe in predlog razširitve akustične slike v dve dimenziji. Z analizo rezultatov bomo poskušali natančneje opredeliti nadaljnje možnosti razvoja umetnega kodiranja elevacije in predvsem sposobnosti učenja absolutne lokalizacije navideznih izvorov zvoka, kar je pri uporabi ne-individualnih (tj. splošnih) funkcij ključnega pomena.

6.1 Ideja umetnega kodiranja elevacije

Najučinkovitejši postopek za doseg prostorskega učinka s poljubnim zvokom predstavlja uporaba individualnih HRTF [57], tj. konvolucije zvokovnega signala in odziva prenosne poti med (navideznim) izvorom zvoka in poslušalčevim uhljem na enotin impulz. Zaradi težavnosti merjenja individualnih HRTF želimo uporabiti splošen model HRTF, kot so to npr. HRTF, izmerjene na lutki. Težava splošnih modelov je ne-individualnost, kar pomeni, da poslušalec ni zmožen absolutne lokalizacije navideznega izvora zvoka oz. se mora naučiti njegove prostorske lokacije. V osnovni ideji smo predpostavili, da v akustično sliko uvedemo posebno (umetno) kodiranje le po elevaciji, za kodiranje po azimutu pa uporabimo splošne HRTF ali le ustrezne časovne zakasnitve in amplitudne razlike med obema uhljema [67]. Na tak način se poslušalčevo učenje reducira na učenje lokalizacije navideznih izvorov zvoka po elevaciji. Kot bomo videli v nadaljevanju, pa ustrezen model kodiranja elevacije pripomore k hitrejšemu učenju.

Za reševanje problema lokalizacije po elevaciji in razvoja modelov, smo problem razdelili na dva dela:

- izbira zvoka in
- razvoj postopkov obdelave signalov s katerimi obdelujemo izbrane zvoke.

Zvoke, ki smo jih uporabili v modelih za kodiranje elevacije smo izbrali izmed zvokov, ki se najpogosteje pojavljajo v sorodnih poskusih. Poskušali smo zajeti čim bolj raznolike zvoke glede na njihov frekvenčni spekter. V zvezi s psihoakustiko v literaturi največkrat zasledimo raznobarvne šume, sinusne signale oz. tone, psevdonaključne

signale, kombinacije navedenih signalov ipd. Za potrebe našega poskusa smo izbrali naslednje zvokovne signale:

- šum z ravnim frekvenčnim potekom med 0 Hz in 20 kHz (beli šum),
- rožnati šum in
- sinusni signal (ton).

Pri razvoju postopkov obdelave signalov smo se oprli na analize, ki smo jih opisali v prejšnjem poglavju. Na podlagi analiz smo razvili več modelov, od katerih vsak na svoj način opisuje povezavo med elevacijo in frekvenco.

V vsakem postopku obdelave signalov nastopa frekvenčni parameter, ki ga poimenujemo karakteristična frekvenca in predstavlja odvisnost od elevacije. Število frekvenčnih parametrov posameznega modela je v splošnem poljubno. Ker večino postopkov obdelave, ki smo jih uporabili lahko realiziramo z različnimi filtri, je karakteristična frekvenca pravzaprav parameter filtra, npr. mejna frekvenca, centralna frekvenca ipd. Nabor karakterističnih frekvenc, ki jih uporabimo v posameznem modelu torej pomeni število različnih signalov (zvokov), ki jih dobimo kot rezultat obdelave. Množico zvokov, ki imajo različne karakteristične frekvence in so posledica obdelave skupnega (osnovnega) zvoka, imenujemo zvokovni nabor. Nabor zvokov predstavlja tudi nabor elevacij, ki jih s tem modelom kodiramo – karakteristična frekvenca je neposredno povezana z zvokom, ki mu priredimo točno določeno elevacijo.

Elevacije, ki jih želimo predstaviti v akustični sliki lahko kodiramo z zvoki, ki pripadajo istemu zvokovnemu naboru, lahko pa uporabimo zvoke različnih naborov. V naših eksperimentih v posamezni akustični sliki prostora vedno nastopajo zvoki enega nabora. Z vključitvijo zvokov iz več naborov bi povečali raznolikost in morda posledično pripomogli k večji ločljivosti, vendar pa vključevanje več naborov v isto akustično sliko ni predmet tega dela.

6.2 Postopki obdelave signalov

6.2.1 Model na osnovi »oslabitve uhlja«

Za neposredno modeliranje oslabitve uhlja smo razvili model v katerem uporabimo filter z zarezo oz. pasovno zaporni filter. Uporabimo različne pasovne širine (B) s čimer poudarimo izločitev spektralne vsebine, centralna frekvenca filtra f_c je parameter, ki povezuje model z navidezno elevacijo izvora zvoka.

6.2.2 Model na osnovi porazdelitve energije HRTF

V petem poglavju smo ugotovili, da je v odzivih prenosne poti med izvorom zvoka in poslušalčevim uhljem na enotin impulz energija razporejena tudi v odvisnosti od elevacije izvora zvoka. Kot smo videli, je za nizke elevacije energija koncentrirana v ožjih frekvenčnih pasovih pri nizkih frekvencah, za višje elevacije pa v širših frekvenčnih pasovih s spodnjo frekvenco 0.

Na podlagi napisanega smo model poimenovali nizko pasovni model, ki ga realiziramo z nizko prepustnim filtrom. Zgornja mejna frekvenca filtra, ki je v sorazmerju s pasovno širino filtra, predstavlja povezavo z elevacijo.

6.2.3 Modeli na osnovi povezave elevacije in frekvence

Na osnovi zakonitosti o povezavi elevacije izvorov zvoka in pripadajočih spektrov [3, 59] smo razvili tri modele:

- Model 1 – uporabimo pasovno prepustni filter: eksperimentiramo z različnimi pasovnimi širinami, centralna frekvenca filtra f_c pomeni povezavo z elevacijo izvorov zvoka.
- Model 2 – zvokovni signal množimo s sinusom, kar pomeni, da celoten zvokovni signal premaknemo po frekvenčni osi, frekvenca sinusa oz. tona pomeni povezavo z elevacijo izvora zvoka.
- Model 3 – osnovni zvok sestavimo iz enega ali več tonov s pripadajočimi višjimi harmonskimi komponentami (v razponu 16 Hz – 20 kHz). Za kodiranje elevacije spreminjamo frekvenco osnovnih tonov in v skladu s temi spremembami tudi frekvence harmonskih komponent. Nabor frekvenc osnovnih tonov predstavlja karakteristične frekvence oz. karakteristični nabor. Osnovne tone, višje harmonske komponente in njihove amplitude določimo tako, da si najprej izberemo naraven zvok, ki ga povzroči nek preprost akustični pojav kot npr. nihanje strune kitare, udarec na boben, pisk piščalke ipd. S pomočjo Fourierove transformacije ugotovimo najznačilnejše frekvence, njihove amplitude in pripadajoče višje harmonske komponente. Postopek določanja osnovnih frekvenc naravnega zvoka v tej fazi razvoja akustične slike ni strogo določen.

6.3 Zaznavne razlike med zvoki, ki pripadajo istemu naboru

Prostor, ki bi ga želeli z akustično sliko predstaviti poslušalcu ima vertikalni razpon od -45° do 90° . Z ločljivostjo po elevaciji, kot jo človek izkazuje v prostoru (6°), bi to v

našem modelu zahtevalo 23 zvokov za kodiranje elevacije. Ta vrednost naj zaenkrat služi zgolj za orientacijo, natančneje bomo število potrebnih zvokov določili kasneje.

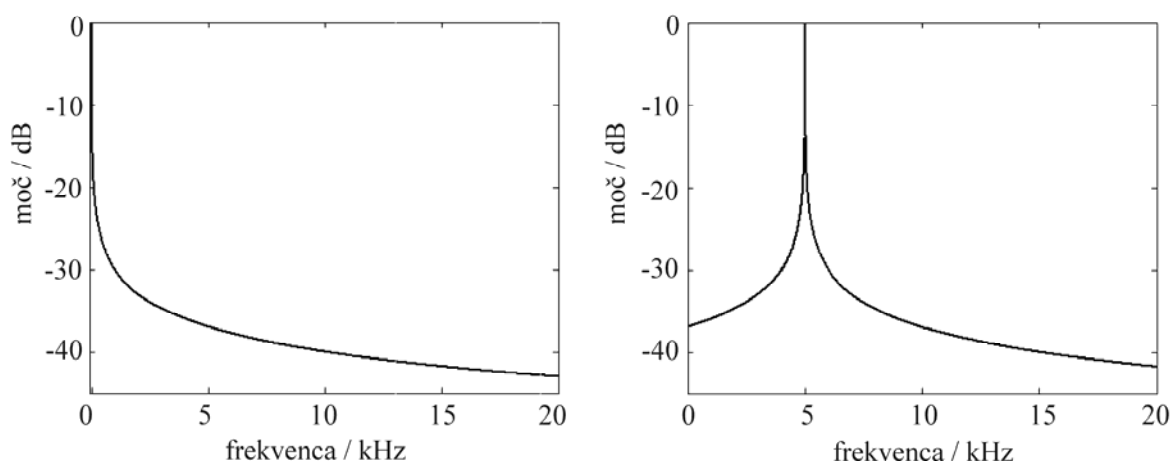
Preden smo se lotili ugotavljanja zaznavnih razlik med zvoki, ki pripadajo istemu naboru, smo predpostavili konstantno razliko (250 Hz) med karakterističnimi frekvencami v celotnem slišnem območju (16 Hz – 20 kHz). Ker so rezultati o potrebni razliki v frekvenci [83] podani le za sinusni signal (ton), je za ostale oblike zvokov potreben preizkus, ki ga opisujemo v nadaljevanju.

6.3.1 Nabori zvokov

Z uporabo nizko prepustnega, pasovno prepustnega in pasovno zapornega filtra ter z uporabo množilnika smo generirali sedem različnih naborov zvokov. Vsak nabor zvoka je vseboval 73 različnih zvokov.

- a) Rožnati šum množen s sinusnim signalom. Rožnati šum se pogosto uporablja v akustiki, njegova glavna značilnost je frekvenčno odvisen amplitudni oz. močnostni spekter. Moč šuma upada obratno sorazmerno s frekvenco – konstantno 3 dB na oktavo:

$$P = \frac{1}{f}. \quad (6.1)$$



Slika 6.1: Močnostni spekter rožnatega šuma (levo) in močnostni spekter rožnatega šuma množenega s sinusnim signalom frekvence 5 kHz (desno). Rožnati šum se pogosto uporablja v akustiki, njegova glavna lastnost je konstantno upadanje moči 3 dB na oktavo.

Produktu rožnatega šuma in sinusa ne moremo več reči rožnati šum, saj ne ustreza zakonitosti enačbe (6.1). Slika 6.1 prikazuje rožnati šum v osnovni obliki in rožnati šum množen s sinusnim signalom frekvence 5 kHz. V

eksperimentu smo uporabili naslednje frekvence sinusnega signala, $f = \{1, 1.25, 1.5, 1.75, 2, 2.25, 2.5, \dots 18.5, 18.75, 19\}$ kHz.

b) Beli šum, filtriran s pasovno prepustnim filtrom. Beli šum se v tem primeru nanaša na signal z ravnim spektrom v področju slišnih frekvenc, tj. med 16 Hz in 20 kHz. Model ima dva parametra – pasovno širino B in centralno frekvenco f_c . Zvoki, ki jih dobimo pri filtriranju s konstantno pasovno širino in s spreminjanjem centralne frekvence f_c tvorijo en nabor. Glede na povezavo med elevacijo in frekvenco [59], zvoki z nižjimi f_c predstavljajo nižje elevacije, zvoki z višjimi f_c pa višje elevacije. V eksperimentu smo uporabili pasovno širino $B = 2$ kHz in naslednje centralne frekvence $f_c = \{1, 1.25, 1.5, 1.75, 2, 2.25, 2.5, \dots 18.5, 18.75, 19\}$ kHz.

c) Beli šum, filtriran s pasovno zapornim filtrom. Beli šum se v tem primeru nanaša na signal z ravnim spektrom v področju slišnih frekvenc, tj. med 16 Hz in 20 kHz. Kot smo že omenili, je eden izmed glavnih faktorjev zaznavanja elevacije izvorov zvoka oslabitev, ki je posledica vpliva uhlja (»pinna notch«). Centralna frekvenca oslabitve f_c je odvisna od elevacije, z naraščanjem elevacije narašča tudi centralna frekvenca oslabitve. Nabor zvokov dobimo s spreminjanjem centralne frekvence oslabitve, z različnimi pasovnimi širinami zapornega pasu pa dobimo več naborov.

V fazi načrtovanja smo predvidevali, da bi s širšo spektralno luknjo izboljšali možnost lokalizacije. Uporabili smo naslednje centralne frekvence $f_c = \{1, 1.25, 1.5, 1.75, 2, 2.25, 2.5, \dots 18.5, 18.75, 19\}$ kHz in pasovne širine B :

- 30 Hz,
- 600 Hz,
- 2 kHz.

d) Beli šum, filtriran z nizko prepustnim filtrom. Beli šum se nanaša na signal z ravnim spektrom v področju slišnih frekvenc, tj. med 16 Hz in 20 kHz. V nizko prepustnem modelu, ki ga realiziramo z nizko prepustnim filtrom je parameter zgornja mejna frekvenca f_{zg} . Uporabili smo naslednje mejne frekvence $f_{zg} = \{1, 1.25, 1.5, 1.75, 2, 2.25, 2.5, \dots 18.5, 18.75, 19\}$ kHz.

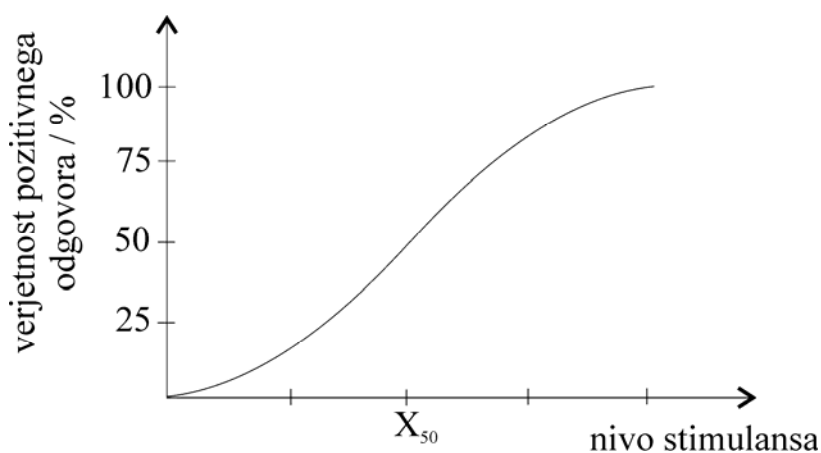
e) Rožnati šum, filtriran z nizko prepustnim filtrom. Za ta primer veljajo enake zakonitosti in predpostavke kot za model, opisan v točki e. Razlika med modeloma je zgolj v osnovnem zvokovnem signalu.

6.3.2 Opis poskusa

S sedmimi nabori zvokov, opisanimi v prejšnjem razdelku, smo izvedli test ločljivosti zvokov, ki pripadajo posameznim naborom (angl. just noticeable differences, JND). Cilj

poskusa je bil ugotoviti koliko zvokov v posameznem naboru je ločljivih oz. kolikšna minimalna razlika karakterističnih frekvenc (zaradi omejenega nabora je v našem primeru ta lahko le večkratnih 250 Hz) je potrebna, da razločimo dva zvoka.

Za test smo uporabili transformirano metodo gor-dol (angl. transformed up-down procedure) [41, 85]. Gre za eno izmed adaptivnih metod, kjer je spreminjanje parametrov stimulatorja (tj. testnega signala) odvisna od prejšnje vrednosti parametrov in od rezultata poskusa. S postopnim spreminjanjem vrednosti parametrov se asimptotično približujemo mejni vrednosti, ki jo poslušalec zazna s 50 % verjetnostjo. Mejna vrednost je zaradi svoje psihofizične narave določena statistično, ponazorimo jo lahko s t.i. psihometrično funkcijo, ki jo prikazuje slika 6.2.



Slika 6.2: Psihometrična funkcija se uporablja v poskusih psihofizične narave. Običajno se kot iskana meja ločljivosti (tj. rezultat poskusa) vzame nivo stimulansa (npr. glasnost zvoka) pri katerem je verjetnost iskanega odgovora 50 %.

Literatura [41] podrobneje opisuje različne načine določanja rezultata na psihometrični funkciji. Med najučinkovitejše metode spadajo tudi metoda gor-dol (angl. simple up-down method) in njene izvedenke – transformirane metode gor-dol.

Za razlago metode gor-dol si pogledjmo primer, ko imamo ton s frekvenco 5 kHz in želimo določiti ton s frekvenco večjo od 5 kHz tako, da bosta tona ločljiva in da bo frekvenčna razlika med njima najmanjša. Na začetku si izberemo začetni odmik, npr. 100 Hz in korak, npr. 10 Hz (iz [83] vemo, da je frekvenčna razlika, ki jo iščemo, približno 30 Hz). Če ugotovimo, da sta preiskovani ton in ton s frekvenco, ki ustreza začetnemu odmiku ločljiva, zmanjšamo frekvenco drugega tona za en korak (10 Hz). To počnemo vse dokler poslušalec ne loči več tonov (prvi tek, angl. first run). V naslednjih korakih nato povečujemo frekvenco drugega tona toliko časa, da poslušalec ponovno loči tona (drugi tek). Mejo med posameznimi teki imenujemo tudi obrat (angl. turnaround). Za določitev frekvenčne razlike, ki jo iščemo, moramo izvesti najmanj 6 tekov, končni

rezultat pa je povprečje srednjih vrednosti vseh sodih tekov. Srednjo vrednost teka določimo iz začetne in končne (v našem primeru) frekvence drugega tona v pripadajočem teku.

Transformirane metode gor-dol so v osnovi enake metodi iz katere izhajajo. Glavna razlika je uvedba skupin gor in dol. To pomeni, da frekvenčno razliko med preizkušeniema tonoma zmanjšamo (dol) oz. povečamo (gor) na podlagi več odgovorov, poleg tega pa lahko uvedemo tudi spremenljiv frekvenčni korak. V našem primeru sta bila za zmanjšanje razlike potrebna dva zaporedna pozitivna odgovora (poslušalec loči zvoka), za povečanje razlike pa je zadostoval en negativen odgovor (poslušalec ne loči zvokov). Za razliko od osnovne metode s tem pristopom dosežemo 70 % verjetnost na psihometrični funkciji. S spremenljivim frekvenčnim korakom pa izboljšamo hitrost konvergence rezultata. Na začetku tako lahko izberemo večji korak, ki ga nato ob vsakem obratu prepolovimo. Tako bi v prej omenjenem primeru dveh tonov izbrali začetni odmik npr. 100 Hz in korak 40 Hz. Če vemo, da bo rezultat 30 Hz, potem bo prvi obrat ko bomo frekvenčno razliko med tonoma zmanjšali na 20 Hz, tedaj bomo prepolovili korak na 20 Hz in naslednji obrat bo pri frekvenčni razliki 40 Hz. Sedaj bomo korak zmanjšali na 10 Hz, kar pomeni, da bomo ob naslednjem poslušanju že dosegli končni rezultat. Več o tej in ostalih transformiranih metodah gor-dol je napisano v [41].

6.3.3 Testne osebe in oprema

V poskusu je sodelovalo 12 oseb, starih od 17 do 37 let. Osebe niso imele predhodnih izkušenj s poslušanjem prostorskih zvokov na slušalkah, prav tako pa so potrdile, da nimajo težav s sluhom in vidom.

Poskus smo izvajali v različnih prostorih, kjer nivo okoliškega šuma ni presegel 40 dB, merjeno po lestvici *A* z inštrumentom Lutron SL-4012, nivo glasnosti testnih zvokov je bil približno 65 dB. Za generiranje in predvajanje zvokov smo uporabili prenosni računalnik Acer TravelMate 4000, zvočno kartico Digigram VXpocket 440 in slušalke Sennheiser Control HD270.

6.3.4 Rezultati

Tabela 6.1 in slika 6.3 podajata povprečno število zvokov, ki so jih testne osebe razločile iz 73 zvokov posameznega nabora.

S pomočjo testov za ugotavljanje porazdelitev (t-test, Smirnov-Kolmogorov, Jarque-Bera) [8, 50] smo ugotovili, da imajo rezultati normalno porazdelitev. Nadalje smo s

pomočjo t-testa ugotovili, da zvoke lahko razdelimo v 4 skupine s statistično značilnimi povprečnimi vrednostmi.

Zvoki, ki pripadajo posameznim skupinam nimajo statistično značilnih razlik v povprečnih vrednostih:

- Skupina 1: beli in rožnati šum filtrirana z nizko prepustnim filtrom, beli šum filtriran s pasovno prepustnim filtrom.
- Skupina 2: rožnati šum množen s sinusom, beli šum filtriran s pasovno zapornim filtrom $B = 2$ kHz.
- Skupina 3: beli šum filtriran s pasovno zapornim filtrom $B = 600$ Hz.
- Skupina 4: beli šum filtriran s pasovno zapornim filtrom $B = 30$ Hz.

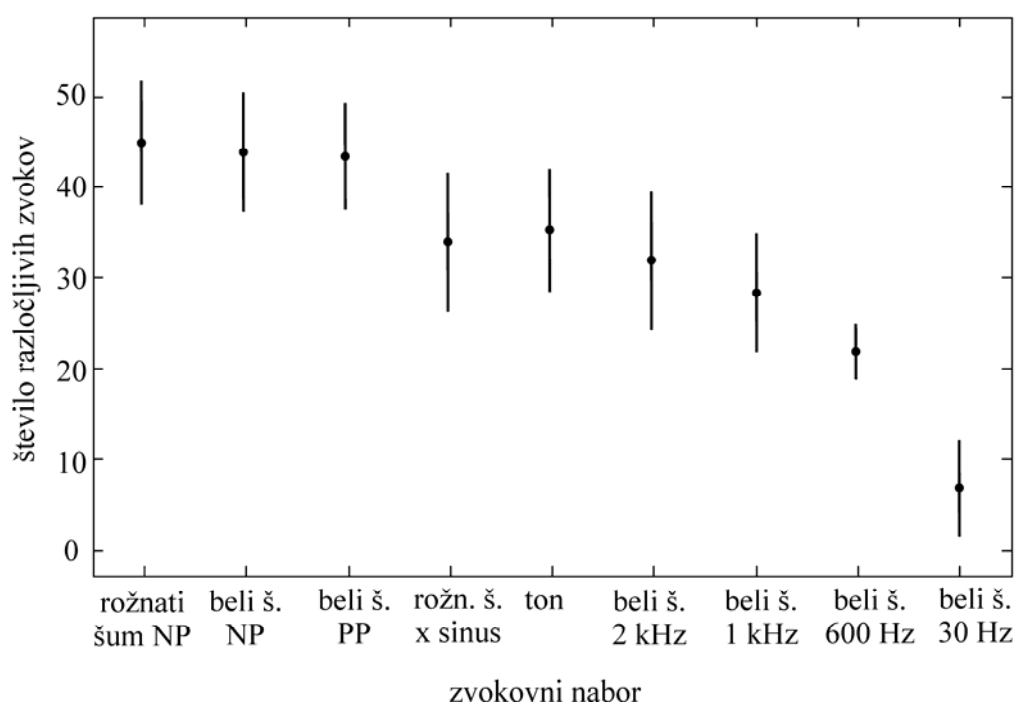
nabor	ločljivost / število zvokov
Rožnati šum množen s sinusom	35±7.87
Beli šum, filtriran s pasovno prepustnim filtrom, $B = 2$ kHz	42±6.11
Beli šum, filtriran s pasovno zapornim filtrom $B = 30$ Hz	8±6.67
Beli šum, filtriran s pasovnozapornim filtrom $B = 600$ Hz	22±3.91
Beli šum, filtriran s pasovno zapornim filtrom $B = 2$ kHz	35±8.23
Beli šum, filtriran z nizko prepustnim filtrom	43±7.15
Rožnati šum, filtriran z nizko prepustni filtrom	45±7.39

Tabela 6.1: Število zvokov (z intervalom zaupanja), ki so jih testne osebe razločile izmed 73 zvokov posameznega nabora. Vsak nabor je bil sestavljen iz zvokov, ki so se med sabo le malo razlikovali (npr. v centralni frekvenci pasovno omejenega šuma). Grafična predstavitev rezultatov je na sliki 6.3.

Glede na prej podano oceno o potrebnem številu razločljivih zvokov za akustično sliko prostora – 23, izpade beli šum filtriran s pasovno zapornim filtrom $B = 30$ Hz, medtem ko je beli šum filtriran s pasovno zapornim filtrom $B = 600$ Hz ravno na meji.

Najboljše rezultate smo dobili s pasovnim in obema nizko pasovnima šumoma. Glede na to, bi za nabor rožnatega šuma množenega s sinusom, ki ga lahko obravnavamo kot ozko pasovni šum pričakovali večje število ločljivih zvokov. Slednje bi pričakovali tudi na podlagi primerjave ločljivosti sinusnega signala [83] in ločljivosti nizko pasovnega in pasovnega šuma [55, 74], ki kaže, da ima sinusni signal najboljšo ločljivost – rožnati šum množen s sinusom je izrazito ozkopasoven signal, kar ga z upoštevanjem delovanja ušesa pri višjih frekvencah uvršča bliže tonu kot pasovnemu šumu. Kot je nakazano v [74] gre vzroke za slabšo ločljivost najverjetneje iskati v povezavi trajanja in nivoja signalov, katere posledice so v našem primeru očitno izrazitejše za rožnati šum množen s sinusom.

Poleg tega nenatančnost vnaša še relativno velik frekvenčni korak uporabljen pri meritvi. Ker pa imajo rezultati pomen predvsem v okviru našega raziskovanja v zvezi z akustično sliko prostora, je zastavljeni cilj dosežen. Rezultati, tj. minimalne zaznavne razlike v spektru za signale različnih oblik zaradi fiksnega koraka karakteristične frekvence (250 Hz) seveda niso splošni, kar v našem primeru, kot že napisano ni potrebno. Vsekakor pa bi bili splošni rezultati nujni v primeru, ko bi se lotili izboljševanja ločljivosti akustične slike.



Slika 6.3: Grafična predstavitev rezultatov ločljivosti v posameznem naboru zvokov (rožnati šum filtriran z nizko prepustnim filtrom, beli šum filtriran z nizko prepustnim filtrom, beli šum filtriran s pasovno prepustnim filtrom, rožnati šum množen s sinusom, beli šum filtriran s pasovno zapornim filtrom pasovne širine 2 kHz, beli šum filtriran s pasovno zapornim filtrom pasovne širine 600 Hz in beli šum filtriran s pasovno zapornim filtrom pasovne širine 30 Hz), v vsakem naboru je bilo 73 zvokov. Rezultati so tabelarično podani v tabeli 6.1.

6.4 Akustična slika v vertikalni smeri

Če želimo ustvariti akustično sliko prostora z umetnim kodiranjem, moramo ugotoviti koliko podatkov lahko sploh posredujemo poslušalcu, ki nima izkušenj z umetnim kodiranjem in na kakšno učinkovitost naj pri tem računamo. Ker nas v tem delu zanima

predvsem elevacija, smo zgradili enodimenzionalno akustično sliko in s pomočjo testnih oseb ugotavljali njeno uporabnost.

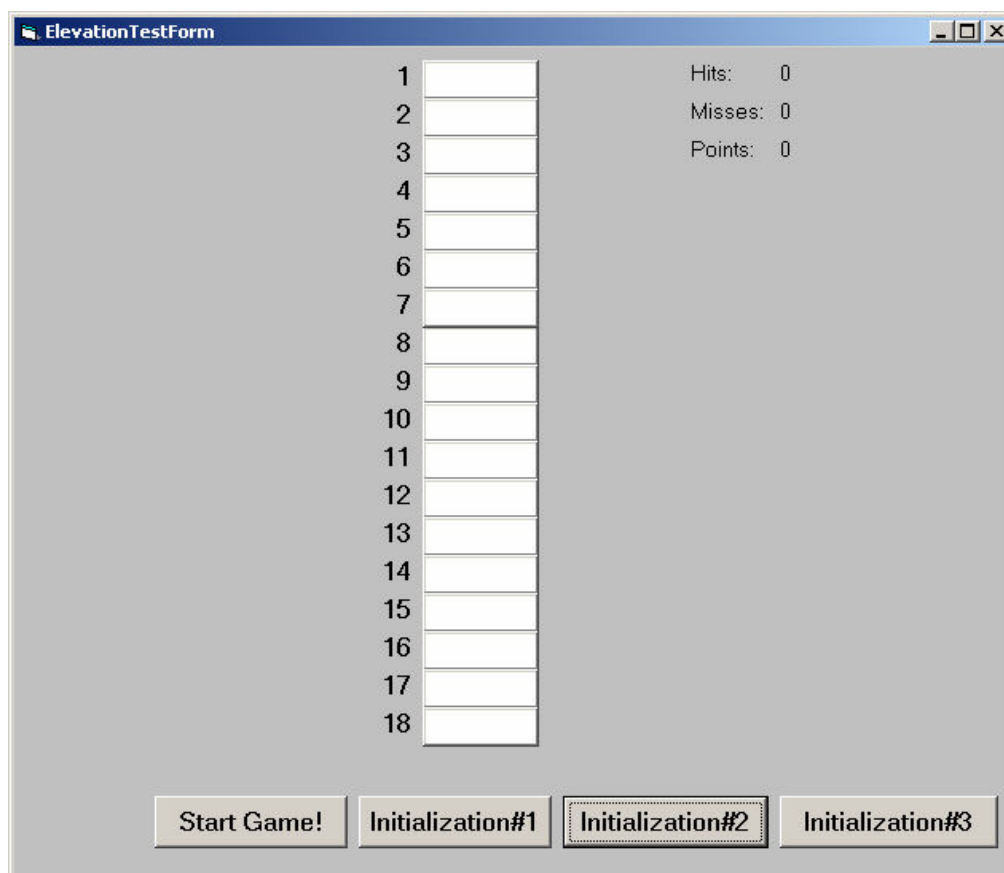
Ker uporaba ne-individualnih prostorskih zvokov ne omogoča absolutne prostorske lokalizacije izvorov zvokov, smo z akustično sliko opisali grafični uporabniški vmesnik na računalniškem zaslonu. Uporabniški vmesnik je bil bistvenega pomena pri učenju – služil je za predstavitev lokacij posameznih izvorov zvoka, ki se jih mora poslušalec naučiti. Ustrezna zasnova in razporeditev zvokov naj bi bila poslušalcu v pomoč pri učenju. Poleg ločljivosti med zvoki, ki predstavljajo različne elevacije, je za učenje v pomoč, da zvoki vsebujejo značilnosti, ki jih poslušalec pozna iz vsakdanjega življenja, tj. predvsem vplivi poslušalčevega telesa na zvočno valovanje, ki se kažejo npr. v HRTF.

6.4.1 Zasnova akustične slike in nabori zvokov

Slika 6.4 prikazuje grafični vmesnik testne aplikacije, ki smo jo razvili v Microsoft Visual Basic-u. Osemnajst gumbov razporejenih v vertikalni smeri predstavlja osemnajst (navideznih) izvorov zvoka. Ker preizkušamo zvoke na splošni osnovi, tj. ne-individualne zvoke, vsakemu gumbu oz. elevaciji priredimo zvok, ki naj bi po pričakovanjih najbolje predstavljal elevacijo. Ker glede na ne-individualnost zvokov težko govorimo o »zvoku, ki najbolje predstavlja elevacijo«, gre pri razporejanju zvokov za zagotavljanje enega ali več faktorjev z monotono odvisnostjo v celotnem obsegu elevacije, ki jo želimo zajeti. Na primer, kadar elevacijo kodiramo neposredno s sinusnim signalom, je ta faktor kar frekvenca.

Če želimo z akustično sliko predstaviti dogajanje pred poslušalcem oz. dogajanje v prostoru, ki ga vidi človek z normalno razvitim vidom, tudi akustično sliko prostora omejimo v skladu s človekovim vidnim kotom. Glede na sposobnosti zaznavanja človekov vid razdelimo v glavni in periferni. Vidni kot glavnega vida je $\pm 15^\circ$ v azimutu in elevaciji, periferni vid pa zajema vse ostalo področje do $\pm 90^\circ$ v horizontalni in približno $\pm 68^\circ$ v vertikalni smeri. Ločljivost v perifernem delu je 15 % – 50 % slabša kot v glavnem vidu, odvisno od kota. To pomeni, da moramo tudi v akustični sliki izvore zvoka najgosteje razporediti v bližini horizontalne ravnine, stran od nje pa razdalje med izvori zvoka povečevati. Število izvorov zvoka, ki jih potrebujemo v akustični sliki dobimo z oceno povprečnega odstopanja po celotnem vidnem polju, ki je glede na predstavljene podatke približno 25 %. To odstopanje prištejemo ločljivosti človeškega sluha v prostoru (6°) in z upoštevanjem človekovega vidnega kota v vertikalni smeri (136°) ugotovimo, da potrebujemo 18 izvorov zvoka. Če zgolj za orientacijo za hip odmislimo neenakost ločljivosti v glavnem in perifernem delu vida, to pomeni, da vsak izvor pokriva elevacijski kot 7.5° . Ker za predstavitev prostora uporabljamo grafični

vmesnik na računalniškem zaslonu, ki ga vidimo pod precej manjšim vidnim kotom kot 136°, pred poslušalca postavljamo še dodatno nalogo – razširitev vidnega kota.



Slika 6.4: Grafični vmesnik enodimenzionalne akustične slike, ki smo jo uporabili v poskusu učenja umetno kodirane elevacije. Testne osebe so se poskušale čim natančneje naučiti povezave med zvokom in lokacijo (gumbi z oznakami 1 – 18). Za učenje so bile na razpolago tri učne serije (»Initalization«), za preverjanje naučenega (gumb »StartGame!«) pa je aplikacija naključno izbirala zvok. Za večjo motivacijo testnih oseb je aplikacija štela tudi točke.

Osemnajst izvorov zvoka oz. elevacij bi morali razvrstiti simetrično glede na horizontalno ravnino, kar pa zaradi povezave elevacije s frekvenco v vseh obravnavanih modelih ni trivialen problem. Kot smo videli pri obravnavi najmanjših zaznavnih razlik tona, obravnavi kritičnih pasov in pri ugotavljanju ločljivosti zvokov istega nabora, za razločevanje potrebna razlika v frekvenci dveh zvokov ni konstantna. Razlike v karakterističnih frekvencah (za osvežitev spomina – s pojmom karakteristična frekvenca mislimo na parameter, ki ga v posameznem modelu spreminjamo, tj. frekvenca sinusa, frekvenca osnovnih tonov, zgornja mejna frekvenca filtra, centralna frekvenca filtra), ki so potrebne za ločevanje zvokov, dobimo iz rezultatov poskusa ugotavljanja ločljivosti

zvokov istega nabora, za oceno psihofizičnega dojetanja razlike med zvoki pa uporabimo Webrov zakon (enačba 3.2), kjer kot neodvisno spremenljivko uporabimo karakteristično frekvenco.

pozicija zvoka	a	b	c	d	e
	f / kHz	f_{cut} / kHz	f_{cut} / kHz	f_c / kHz, $B = 2$ kHz	f_c / kHz, $B = 1$ kHz
1	14	14	14	14	14
2	11	11	11	11	11
3	9	9	9	9	9
4	8	8	8	8	8
5	7	7	7	7	7
6	6	6	6	6	6
7	5	5	5	5	5
8	4	4	4	4	4
9	3.5	3.5	3.5	3.5	3.5
10	3	3	3	3	3
11	2.75	2.75	2.75	2.75	2.75
12	2.5	2.5	2.5	2.5	2.5
13	2.25	2.25	2.25	2.25	2.25
14	2	2	2	2	2
15	1.75	1.75	1.75	1.75	1.75
16	1.5	1.5	1.5	1.5	1.5
17	1.25	1.25	1.25	1.25	1.25
18	1	1	1	1	1

Tabela 6.2: Povezava med elevacijo izvora zvoka in karakteristično frekvenco zvoka za naslednje nabore zvokov: (a) rožnati šum množen s sinusnim signalom, (b) rožnati šum filtriran z nizko prepustnim filtrom, (c) beli šum filtriran z nizko prepustnim filtrom, (d) beli šum filtriran s pasovno prepustnim filtrom in (e) beli šum filtriran s pasovno zapornim filtrom.

Ker smo v akustični sliki uporabili tudi zvoke, ki smo jih generirali že pri poskusu ločevanja zvokov istega nabora, nam ni povsem uspelo slediti konstantnemu upadanju ločljivosti vida pri oddaljevanju od horizontalne ravnine (elevacije 4 – 7 v tabeli 6.2), ki

bi bilo potrebno za pravilno modeliranje perifernega vida – razlike karakterističnih frekvenc pri omenjenih elevacijah bi morale upadati in ne ostati konstantne. Gledano izključno za naš poskus to ne pomeni kake ovire ali pomanjkljivosti, je pa opombo vsekakor smiselno upoštevati pri načrtovanju akustične slike za praktično uporabo. Seznam karakterističnih frekvenc in prireditve zvoka elevacijam z oznakami kot so na sliki 6.4 je v tabelah 6.2 in 6.3. Pri načrtovanju zvoka, sestavljenega iz več osnovnih tonov in njim pripadajočih harmonskih komponent (tabela 6.3) smo uporabili nekoliko drugačen pristop, kar bo opisano v nadaljevanju.

	osn. 1	osn. 2	osn. 3	osn. 4	osn. 5	osn. 6
pozicija zvoka	$f/ \text{ Hz}$	$f/ \text{ Hz}$	$f/ \text{ Hz}$	$f/ \text{ Hz}$	$f/ \text{ Hz}$	$f/ \text{ Hz}$
1	250	260	300	310	330	630
2	280	290	340	350	370	710
3	320	330	390	400	420	800
4	360	370	440	450	480	910
5	410	420	500	510	540	1030
6	460	480	570	580	610	1170
7	520	540	650	660	690	1330
8	590	610	740	750	780	1510
9	670	690	840	850	880	1710
10	760	780	950	960	1000	1940
11	860	880	1080	1090	1130	2200
12	970	1000	1220	1240	1280	2490
13	1100	1130	1380	1410	1450	2820
14	1250	1280	1560	1600	1640	3200
15	1420	1450	1770	1810	1860	3630
16	1610	1640	2010	2050	2110	4110
17	1820	1860	2280	2320	2390	4660
18	2060	2110	2580	2630	2710	5280

Tabela 6.3: Povezava med elevacijo izvora zvoka in karakterističnimi frekvencami za nabor zvoka s harmoniki. Uporabili smo šest osnovnih tonov (osn. 1, osn. 2, ... osn. 3), ki so dodeljeni posameznim elevacijam (1 – 18). Višje harm. komponente posameznega tona si sledijo do frekvence 20 kHz.

V testni aplikaciji smo uporabili šest različnih naborov zvokov za šest izvedb akustične slike. Trajanje posameznega intervala zvoka je bilo 100 ms, med zaporedno predvajanimi zvoki je bilo 250 ms tišine.

- a) Rožnati šum množen s sinusnim signalom oz. tonom. V tabeli 6.2 je seznam frekvenc sinusnega signala, ki smo jih uporabili v akustični sliki in prireditev signalov elevaciji.
- b) Rožnati šum, filtriran z nizko prepustnim filtrom: v nizko prepustnem modelu, ki ga realiziramo z nizko prepustnim filtrom je parameter zgornja mejna frekvenca f_{zg} . Pri filtriranju spreminjamo ojačanje filtra tako, da imajo vsi zvoki enako jakost. Tabela 6.2 prikazuje zgornje mejne frekvence, ki smo jih uporabili v akustični sliki in prireditev signalov elevaciji.
- c) Beli šum, filtriran z nizko prepustnim filtrom: beli šum se nanaša na signal z ravnim spektrom v področju slišnih frekvenc, tj. med 16 Hz in 20 kHz. V nizko prepustnem modelu, ki ga realiziramo z nizko prepustnim filtrom je parameter zgornja mejna frekvenca f_{zg} , s spremenljivim ojačanjem filtra zagotovimo, da imajo vsi zvoki enako jakost. Tabela 6.2 prikazuje zgornje mejne frekvence, ki smo jih uporabili v akustični sliki in prireditev signalov elevaciji.
- d) Beli šum, filtriran s pasovno prepustnim filtrom: za beli šum veljajo enake predpostavke kot v točki c, model ima dva parametra – pasovno širino B in centralno frekvenco f_c . Uporabili smo filter s pasovno širino $B = 2$ kHz, centralne frekvence in povezava z elevacijo je prikazana v tabeli 6.2.
- e) Beli šum, filtriran s pasovno zapornim filtrom. Za beli šum veljajo enake predpostavke kot v točki c. Uporabili smo filter s pasovno širino $B = 1$ kHz, nabor zvokov dobimo s spreminjanjem centralne frekvence oslabitve. V tabeli 6.2 najdemo seznam centralnih frekvenc in povezavo zvokov in elevacije.
- f) Zvok s harmoniki. Za osnovo smo uporabili 100 ms izsek zvoka, ki nastane ob udarcu na boben. Za osnovne tone smo uporabili naslednje frekvence: 250 Hz, 260 Hz, 300 Hz, 310 Hz, 330 Hz in 630 Hz. Glede na prejšnje primere smo v tem primeru izbrali precej manjše razmike med karakterističnimi frekvencami za posamezne elevacije, ker želimo, da ima tudi ton z najvišjo frekvenco pri vsaki elevaciji vsaj 3 ali 4 harmonske komponente. Za zagotovitev ločljivosti zvokov smo se oprli na najmanjšo zaznavno razliko v frekvenci tonov [83], ki je pri tonu 300 Hz enaka 4 Hz. Vrednost smo 10-krat povečali in jo po Weber-Fechnerjevem zakonu (enačba 3.2) uporabili za določitev frekvenc ostalih elevacij za vseh šest tonov. V tabeli 6.3 je seznam frekvenc osnovnih tonov za posamezne elevacije. V tabeli 6.4 so navedene relativne amplitude harmonskih

komponent glede na harmonsko komp. z najvišjo amplitudo v posameznem harmonskem sestavu. Razmerje med amplitudami osnovnih komponent je:

$$A_0(250 \text{ Hz}) : A_0(260 \text{ Hz}) : A_0(300 \text{ Hz}) : A_0(310 \text{ Hz}) : A_0(330 \text{ Hz}) : A_0(630 \text{ Hz}) = 22 : 24 : 16 : 22 : 36 : 100 \quad (6.2)$$

f_0 / Hz	A_0	A_1	A_2	A_3	A_4	A_5	A_6	A_7	A_8	A_9
250	0.67	0.39	0.57	1	0.69	0.23	0.35	0.15	0.25	0.11
260	1	0.36	0.24	0.44	0.20	0.42	0.18	0.36	0.10	0.24
300	1	0.64	0.52	0.46	0.42	0.39	0.37	0.36		
310	0.35	1	0.14	0.28	0.14	0.09	0.05			
330	1	0.32	0.65	0.30	0.53	0.28	0.46	0.25	0.40	0.24
630	1	0.04	0.11	0.03	0.07	0.02				

Tabela 6.4: Razmerja amplitud komponent harmonskih sestavov glede na komponento z največjo amplitudo v harmonskem sestavu. Razmerja med osnovnimi komponentami posameznih harmonskih sestavov podaja enačba 6.2.

6.4.2 Opis poskusa

6.4.2.1 Učenje elevacije v akustični sliki

Kot že omenjeno, je bil cilj poskusa, da se poslušalci naučijo pomenov zvokov v akustični sliki, kar pomeni, da znajo na podlagi slišane zvoka ugotoviti navidezno elevacijo izvora tega zvoka. Poslušalci, tj. testne osebe, so zato s pomočjo testne aplikacije izvajali učenje in preizkušanje naučenega.

Kot vidimo na sliki 6.4, je testna aplikacija omogočala tri različne scenarije učenja (»Initialization#1«, »Initialization#2« in »Initialization#3«) in preizkušanje naučenega (»Start Game!«). V fazi učenja je računalnik izbiral elevacije naključno (»Initialization#1«), od najnižje do najvišje (»Initialization#2«) oz. od najvišje do najnižje (»Initialization#3«). Glede na izbrano elevacijo je poslušalec zaslišal ustrezen zvok, poleg tega pa se je še osvetlil pripadajoči gumb. Faza preizkušanja je bila zasnovana kot nekakšna računalniška igrice, kjer je računalnik naključno izbiral elevacije in predvajal pripadajoči zvok, poslušalec pa je s klikom na gumb sporočil kateri elevaciji po njegovem mnenju ustreza zvok. Če je bil izbrani gumb pravilen, je poslušalec zaslišal kratek pisk potrditvenega tona, če pa je bila izbira napačna pa kratek pisk, ki je signaliziral nepravilno izbiro, poleg tega pa se je osvetlil gumb s pravilno elevacijo. Zgolj za motivacijo so se v

zgornjem desnem kotu grafičnega vmesnika izpisovale dosežene točke (točen zadetek 3 točke; zgrešil za eno elevacijo 2 točki; zgrešil za dve elevaciji 1 točka; v ostalih primerih 0 točk), število točnih zadetkov in število zgrešenih zadetkov (ne glede na razdaljo do pravilne elevacije). V »log« datoteko pa so se zapisovali podatki o pravilni elevaciji in elevaciji, ki jo je izbral poslušalec.

Poskus preizkušanja naučenega se je izvajal na način, ki ga opisujemo v naslednjih vrsticah. Naključna izbira elevacije (porazdelitev naključne spremenljivke je enakomerna) in največ osem ponovitev predvajanja pripadajočega zvoka (100 ms) z vmesno tišino (250 ms), kar skupaj traja 2.8 s. V tem času se je moral poslušalec odločiti kateri elevaciji zvok pripada. Če poslušalec ni podal svoje izbire, je računalnik štel rezultat za napačen z največjim možnim odstopanjem, ki ga v posameznem primeru izračunamo kot maksimum razlik pravilne elevacije in najnižje oz. najvišje elevacije akustične slike. Med pregledom rezultatov smo ugotovili, da so se poslušalci vedno odločili za neko vrednost, torej ta scenarij ni bil nikoli upoštevan. Posamezen poskus je obsegal ugotavljanje elevacij osmih naključno izbranih izvorov zvoka, kar časovno znese slabih 23 s. Zaporedje desetih poskusov imenujemo serija. V povprečju je ena serija trajala 5 minut. Poslušalcem je bilo naročeno, naj pred vsakim prvim poskusom v seriji vsaj 2 minuti poslušajo zvoke učenja (»Initialization#1«, »Initialization#2« in »Initialization#3« na sliki 6.4). Učenje med naslednjimi poskusi v seriji ni bilo dovoljeno.

Testne osebe so poskuse izvajale samostojno. Naročeno jim je bilo, naj jih izvajajo le, ko so se sposobne primerno skoncentrirati. Zahtevali smo, da naj med dvema zaporednima serijama mine vsaj ena ura. Poslušalci, ki so se ukvarjali z več kot eno akustično sliko, so po končanem eksperimentiranju ene slike lahko nadaljevali z naslednjo po štirih dneh.

6.4.2.2 Ocenjevanje prijetnosti zvokov v akustični sliki

Poleg možnosti učenja akustične slike z umetnim kodiranjem elevacije nas je zanimala še prijetnost zvokov v akustični sliki. V ta namen smo izvedli ločen poskus v katerem so poslušalci eksperimentirali z vsemi šestimi akustičnimi slikami in ocenjevali zvoke posamezne slike z ocenami od 1 do 5. Poslušalci so imeli nalogo, da z vsako sliko izvedejo dvakrat po pet poskusov in podajo ocene, ki naj odražajo relativno primerjavo med posameznimi nabori zvokov.

6.4.3 Testne osebe in oprema

6.4.3.1 Učenje elevacije v akustični sliki

V poskusu je sodelovalo 27 oseb, starih od 17 do 36 let. Testne osebe so potrdile, da nimajo težav s sluhom in vidom. Vsaka od oseb je eksperimentirala z največ tremi različnimi akustičnimi slikami, večina le z dvema ali celo z eno samo.

Poskus so testne osebe izvajale samostojno in pri tem uporabljale različne računalnike. Zahtevali smo, naj poskuse izvajajo v prostorih s čim manj hrupa.

6.4.3.2 Ocenjevanje prijetnosti zvokov v akustični sliki

V poskusu je sodelovalo 6 oseb, starih od 27 do 36 let. Osebe so imele predhodne izkušnje s poslušanjem prostorskih zvokov na slušalkah, potrdile so, da nimajo težav s sluhom in vidom.

Poskus so testne osebe izvajale samostojno in pri tem uporabljale različne računalnike. Zahtevali smo, naj poskuse izvajajo v prostorih s čim manj hrupa.

6.4.4 Rezultati

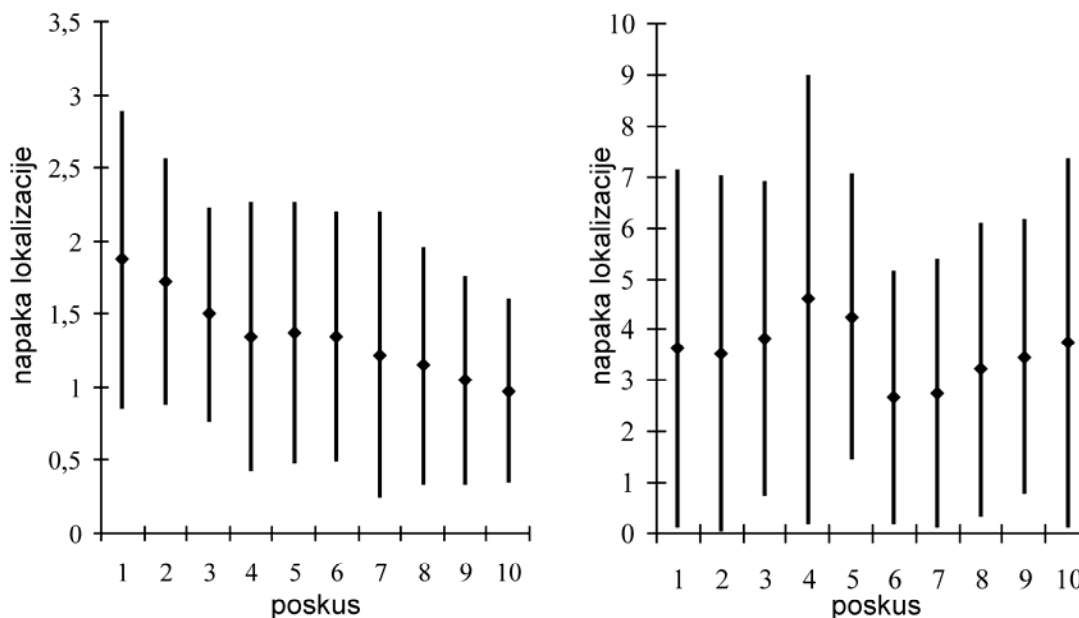
6.4.4.1 Učenje akustične slike

Sprotno preverjanje rezultatov je pokazalo, da razen v sliki, ki uporablja beli šum filtriran s pasovno zapornim filtrom, testne osebe dosegajo napredek pri učenju. Zato smo nadaljnje poskuse z omenjeno akustično sliko ustavili in izključili iz nadaljnjih analiz.

Slika 6.5 prikazuje dva poteka povprečne ločljivosti v naključno izbrani seriji naključne testne osebe. Poteki ločljivosti v serijah drugih testnih oseb so precej podobni, tako da lahko rečemo, da slika 6.5 predstavlja nek tipični potek. Na sliki sta potek ločljivosti v akustični sliki z belim šumom filtriranim z nizko prepustnim filtrom in potek ločljivosti v akustični sliki z belim šumom filtriranim s pasovno zapornim filtrom. Glede na podobnost napredkov v vseh akustičnih slikah, razen v sliki z belim šumom filtriranim z nizko prepustnim filtrom, so poteki ločljivosti za preostale akustične slike (rožnati šum množen s sinusom, rožnati šum filtriran z nizko prepustnim filtrom, beli šum filtriran s pasovno prepustnim filtrom in zvok s harmoniki) podobni poteku v levem delu slike 6.5, z manjšimi razlikami v povprečni vrednosti in disperziji.

Po izločitvi akustične slike z belim šumom filtriranim s pasovno zapornim filtrom, smo pri vseh testnih osebah ugotovili napredek pri učenju. Slika 6.6 prikazuje napredek v ločljivosti za posamezne akustične slike. Ločljivost je izračunana kot povprečna vrednost rezultatov vseh testnih oseb, ki so izvajale eksperiment z določeno akustično sliko. Mera

ločljivosti oz. natančnosti lokalizacije predstavlja povprečno odstopanje od prave vrednosti oz. razliko med elevacijo resničnega in izbranega izvora zvoka. Razlika se meri v številu elevacijskih stopenj (1 – 18) in ne s kotom (glej tudi tabeli 6.2 in 6.3).

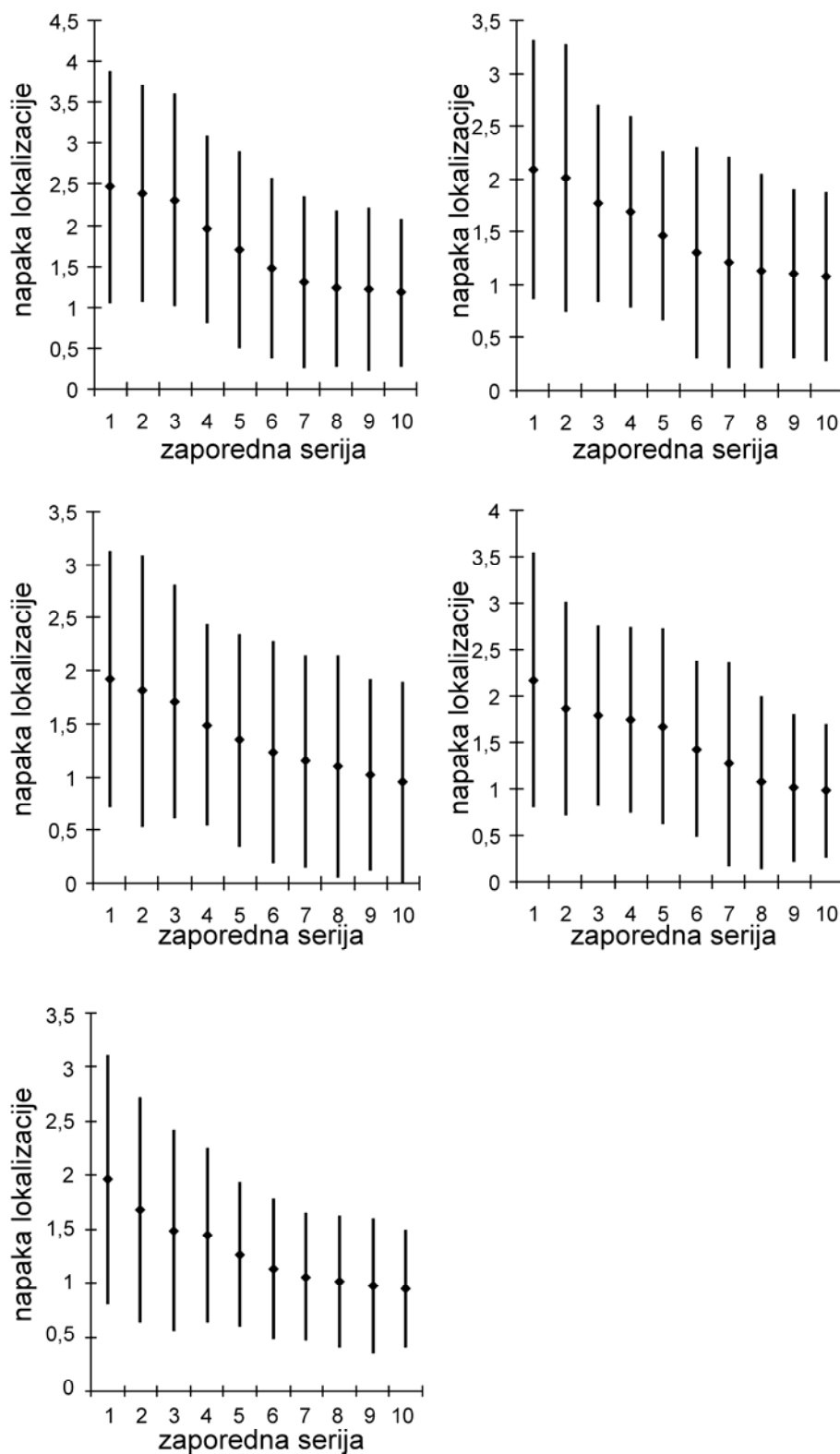


Slika 6.5: Tipični potek ločljivosti v seriji (tj. 10 zaporednih poskusov) z akustično sliko z belim šumom filtriranim z nizko prepustnim filtrom (levo) in belim šumom filtriranim s pasovno zapornim filtrom (desno). Slednji se je izkazal za neprimerne – napaka lokalizacije (razlika med dejansko elevacijo in elevacijo, ki jo je ugotovil poslušalec) je občutna in se s podaljševanjem učenja ne zmanjšuje kot je to opazno v levem primeru.

Analizo rezultatov, tj. povprečnih vrednosti smo izvedli z ne-parametričnim testom Kruskal-Wallis [50]. Za vse analizirane akustične slike smo ugotovili, da je napredek v ločljivosti statistično značilen. Analiza je ovrgla hipotezo o statistično neznačilnih razlikah med prvo in zadnjo serijo (verjetnost, ki jo vrne Kruskal-Wallis je bila med 0.001 in 0.005), sicer pa primerjave parov zaporednih serij ne izkazujejo vedno statistično značilnih razlik, kar pomeni, da napredek ločljivosti med serijami ni bil konstanten. Na grafih v sliki 6.6 to lahko opazimo vizualno. Vidimo tudi, da je največji napredek v prvih nekaj testih, nato se napredovanje nekoliko upočasni.

Za ugotavljanje razlik oz. primerjavo naborov zvokov smo primerjali povprečne ločljivosti med akustičnimi slikami v posameznih serijah. Primerjava pokaže, da lahko obravnavanih pet akustičnih slik razdelimo v dve skupini:

- Skupina 1: akustične slike z belim šumom filtriranim z nizko prepustnim filtrom, belim šumom filtriranim s pasovno prepustnim filtrom, rožnatim šumom filtriranim z nizko prepustnim filtrom in zvokom s harmoniki.
- Slika 2: akustična slika z rožnatim šumom množnim s sinusom.



Slika 6.6: Napredek pri učenju v akustični sliki prostora je odvisen od časa učenja in od nabora zvokov: rožnati šum množeni s sinusom (levo zgoraj), rožnati šum filtriran z nizko prepustnim filtrom (desno zgoraj), beli šum filtriran s pasovno prepustnim filtrom (levo v sredini), beli šum filtriran z nizko prepustnim filtrom (desno v sredini) in zvok s harmoniki (levo spodaj).

Razlike v točnosti lokalizacije med akustičnimi slikami iz skupine 1 so statistično manj značilne kot razlika med katerokoli akustično sliko iz skupine 1 primerjano z akustično sliko iz skupine 2. Rezultat analize s Kruskal-Wallis testom v prvem primeru daje rezultate velikosti približno 0.3, v drugem primeru (tj. primerjava akustične slike iz skupine 2 s katerokoli akustično sliko iz skupine 1) pa velikosti približno 0.01. To pomeni, da je, kot lahko sklepamo tudi na podlagi grafične predstavitve rezultatov, povprečna ločljivost v akustični sliki z rožnatim šumom množenim s sinusom nekoliko slabša kot v ostalih štirih akustičnih slikah.

V tabeli 6.5 so zbrani podatki o končni ločljivosti v posameznih akustičnih slikah.

akustična slika oz. nabor zvokov	končna ločljivost / število elevacijskih nivojev
Rožnati šum množen s sinusom	1.18±0.9
Rožnati šum filtriran z nizko prepustnim filtrom	1.08±0.8
Beli šum filtriran z nizko prepustnim filtrom	0.98±0.72
Beli šum filtriran s pasovno prepustnim filtrom	0.95±0.95
Zvok s harmoniki	0.95±0.6

Tabela 6.5: Končna ločljivost, dosežena po desetih ponovitvah učenja in preverjanja naučenega. Ločljivost je podana kot razlika nivojev (s pojmom nivo so mišljene elevacije z oznako 1 – 18) izbrane in prave elevacije.

6.4.4.2 Ocenjevanje prijetnosti zvokov v akustični sliki

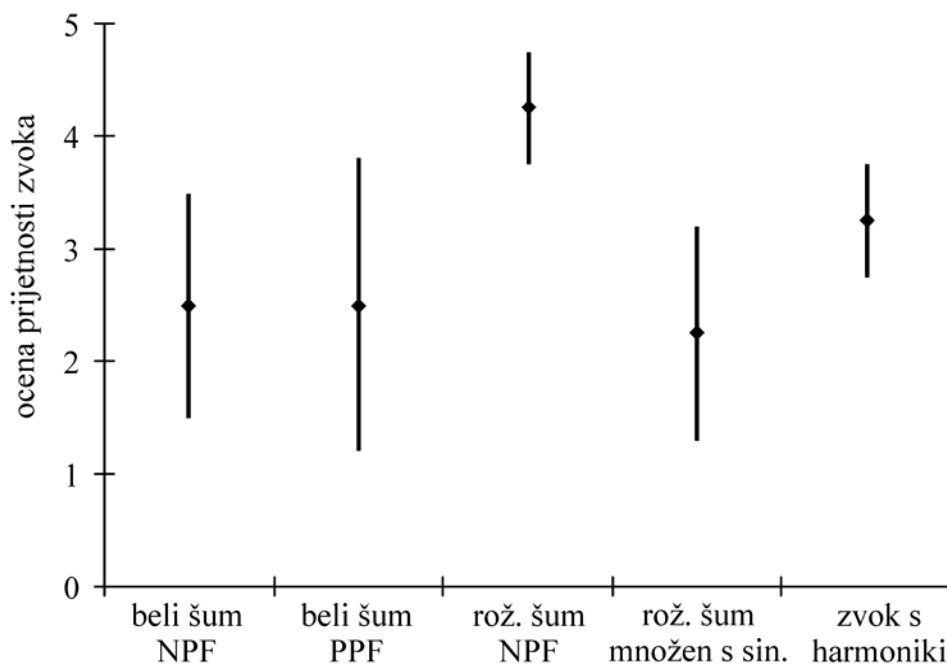
Slika 6.7 prikazuje ocene prijetnosti naborov zvokov, ki so bili uporabljeni v akustični sliki prostora.

Za analizo rezultatov smo uporabili t-test [8, 50]. Za stopnjo značilnosti (parameter α) vzamemo vrednost 0.05 in analiza pokaže, da nabore zvokov lahko razdelimo v dve skupini, ki se statistično razlikujeta:

- Skupina 1: rožnati šum filtriran z nizko prepustnim filtrom.
- Skupina 2: zvok s harmoniki, beli šum filtriran s pasovno prepustnim filtrom, beli šum filtriran z nizko pasovnim filtrom in rožnati šum množen s sinusom.

Pri analizi s t-testom ugotovimo še, da sta zvok s harmoniki in rožnati šum množen s sinusom na meji skupine 2 – zvok s harmoniki na zgornji (po prijetnosti se približuje rožnatemu šumu filtriranem z nizko prepustnim filtrom) in rožnati šum množen s sinusom na spodnji meji. Teza o prijetnosti zvoka z vsebnostjo harmonikov se je izkazala za ustrezno, predvidevamo pa, da bi s še natančnejšim izborom harmonikov rezultat lahko

tudi izboljšali. Hkrati lahko ugotovimo, da prijeten zvok vsebuje relativno velik del nizkih frekvenc, višje frekvence pa so nekoliko zadušene, kar se sklada z ugotovitvami v [5].



Slika 6.7: Ocena prijetnosti naborov zvokov (beli šum filtriran z nizko prepustnim filtrom, beli šum filtriran s pasovno prepustnim filtrom, rožnati šum filtriran z nizko prepustnim filtrom, rožnati šum množen s sinusom in zvok s harmoniki), ki so bili uporabljeni v akustični sliki prostora. Ocene so med 5 (najprijetnejši zvok) in 1 (najbolj moteč zvok). Ocenjevanje je bilo relativno – vsaka testna oseba je podala oceno za vsak zvok.

6.4.5 Komentar

Poskus z učenjem umetno konstruirane akustične slike je dal rezultate za ovrednotenje modelov kodiranja elevacije v akustični sliki prostora. Zaradi uporabe ne-individualnih zvokov je glavno vlogo prevzelo zaznavanje na osnovi t.i. vizualne teorije [4], ki govori o povezavi med videnim in slišanim. Čeprav akustične značilnosti zvokov v tem primeru niso igrale glavne vloge pa so, kot bomo s podrobnejšo razčlenbo ugotovili v nadaljevanju, prispevale pomemben delež k učenju.

Ugotovili smo, da so se bile testne osebe zmožne dokaj uspešno naučiti povezav med elevacijo izvora zvoka in pripadajočimi zvoki z izjemo nabora belega šuma filtriranega s pasovno zapornim filtrom. Pri vseh ostalih izvedbah akustične slike smo zabeležili napredek v ločljivosti.

Hipotezo hitrejšega napredovanja pri učenju v odvisnosti od izbire zvokov lahko potrdimo le s primerjavo neuspešnega in uspešnih naborov zvoka. Med preostalimi štirimi uspešnimi izvedbami, pa glede na omejen čas učenja oz. eksperimentiranja in količino rezultatov, ne moremo jasno potrditi ali ovreči omenjene hipoteze.

Potrjena pa je predpostavka o povezavi frekvenčne vsebine zvoka in elevacije. V učljivih naborih so bile nizke elevacije kodirane z zvoki, ki vsebujejo nizke frekvence, visoke elevacije pa z zvoki, ki vsebujejo višje ali visoke frekvence (rožnati šum množen s sinusom, beli šum filtriran s pasovno prepustnim filtrom in zvok s harmoniki) oz. visoke in nizke frekvence (beli in rožnati šum filtrirana z nizko prepustnim filtrom). Z upoštevanjem povezave elevacije in frekvence si razložimo tudi neuspešnost nabora belega šuma filtriranega s pasovno zapornim filtrom – v tem naboru zvok pri nizkih elevacijah (nizka centralna frekvenca filtra) ni vseboval nizkih frekvenc, pri visokih elevacijah (visoka centralna frekvenca filtra) pa ne visokih frekvenc.

6.5 Razširitev akustične slike po azimutu

V poglavju 6.1, kjer smo predstavili idejo akustične slike prostora z umetnim kodiranjem elevacije, smo napisali, da za kodiranje v horizontalni dimenziji oz. azimutu uporabimo splošne HRTF horizontalne ravnine ali zgolj upoštevamo medušesne časovne in amplitudne razlike.

Ker filtriranje s splošnimi HRTF vpliva na spekter, predvidevamo da bi zgolj upoštevanje medušesnih časovnih in amplitudnih razlik utegnilo biti boljša rešitev. Spekter zvoka na lokalizacijo po azimutu namreč ne vpliva [23] oz. ta vpliv ni primerljiv z vplivom, ki ga ima spekter zvoka na lokalizacijo po elevaciji [33, 64, 65]. Rešitev z upoštevanjem medušesnih razlik lahko podpremo tudi z rezultati poskusa, kjer smo s pomočjo analize PCA (Principal Component Analysis) ugotavljali in primerjali uspešnost diferenčne ločljivosti izvorov zvokov po azimutu pri uporabi splošnih HRTF in pri uporabi le najznačilnejših komponent PCA (medušesne časovne in amplitudne razlike) [67].

7 ZAKLJUČEK

V pričujočem delu smo razvili več postopkov kodiranja elevacije izvorov zvoka za uporabo v akustični sliki prostora, ki jo poslušalcu predvajamo preko slušalk. Postopke, ki dajo dobre rezultate šele v kombinaciji z ustreznim zvokovnim signalom, smo s pomočjo enodimenzionalne akustične slike tudi praktično preizkusili in izvedli statistično analizo rezultatov. Postopke kodiranja smo razvili na osnovi analize HRTF in na podlagi poznavanja ostalih fizikalnih, fizioloških in psiholoških dejavnikov percepcije prostorskega zvoka pri ljudeh. Opisani postopki predstavljajo poskus v smeri uporabe ne-individualnih zvokovnih signalov v akustični sliki prostora. Glede na dejstvo, da najučinkovitejši postopek za ustvarjanje akustične slike temelji na individualnih funkcijah HRTF in da je merjenje individualnih HRTF zapleten postopek, problem s katerim smo se ukvarjali v tem delu, kljub raznovrstnim poskusom še nima končne rešitve. Kot smo ugotovili, je pri lokalizaciji ne-individualnih prostorskih zvokov problem elevacija, ki je »naravno« kodirana z odvisnostjo spektra zvoka od elevacije izvora, medtem ko kodiranje azimuta ni problematično, saj sta glavna faktorja (časovne in amplitudne medušesne razlike) dokaj dobro obvladljiva.

V fazi razvoja smo prostorski zvok najprej obdelali s stališča fizikalnih zakonitosti in človekove fiziologije. Ker je zvok oz. zvočno valovanje pri potovanju od izvora (zvočnik) do ponora (poslušalec) v splošnem izpostavljeno mnogim valovnim dejavnikom, za ponovljivost merilnih rezultatov izvajamo meritve v gluhi sobi, kjer z minimalno odbojnimi stenami simuliramo neskončen, odprt prostor. Ugotovili smo, da tedaj pridejo do veljave valovni pojavi, ki jih povzroča poslušalčevo telo, pri čemer je potrebno še posebej izpostaviti uhlje, ki opravljajo funkcijo smernega filtriranja in zato igrajo ključno vlogo pri percepciji prostorskega zvoka, kar se odraža tudi v pomenu t.i. oslabitve uhlja. Filtriranje vpadnega zvočnega valovanja pri katerem ima poleg uhljev pomemben vpliv tudi zgornji del poslušalčevega telesa (glava, rame), vnaša v zvočno valovanje smerno odvisne značilnosti (amplitudne in časovne razlike med ušesoma, vpliv na spekter), ki služijo možganom za določitev izvora zvoka oz. lokalizacijo. Smerno filtriranje, ki je opisano z odzivom prenosne poti med izvorom zvoka in poslušalcem (HRIR), je ključnega pomena za iskanje značilnosti, ki omogočajo prostorsko zaznavanje zvoka.

Za razvoj postopkov kodiranja v katerih ne uporabljamo individualnih zvokov, so posebnega pomena tudi psihološke zakonitosti človekove percepcije zvoka – psihoakustika. Kot smo ugotovili, te lastnosti vplivajo na izbiro in razporeditev zvokov, ki

jih uporabimo za kodiranje v akustični sliki prostora. V naših postopkih je še posebej pomembna meja minimalne zaznavne razlike v frekvenci, ki podaja minimalno razliko v spektru dveh zvokov, da ju poslušalec še loči (npr. frekvenci dveh tonov).

S področja psihologije je pri uporabi ne-individualnih zvokov v akustični sliki prostora ključnega pomena učenje. Kot smo videli, se človek nauči pomena informacij, ki mu jih posredujejo čutila. S časom se tako iz kognitivnega dožemanja razvije nezavedna (podzavestna) percepcija. V primeru ne-individualnih zvokov je poslušalčeva zmožnost percepcije podobna percepciji ob rojstvu, tj. v kognitivni fazi. Bistvene razlike pri lokalizaciji individualnih in ne-individualnih izvorov zvoka smo ugotovili pri reprodukciji s slušalkami. Z ne-individualnimi zvoki ni možna absolutna lokalizacija, pa tudi diferenčna ločljivost izvorov zvoka je bistveno slabša. Pri reprodukciji zvoka v prostoru pa o ne-individualnih zvokih pravzaprav ne moremo govoriti, saj tudi v primeru ustvarjanja akustičnih iluzij prenosna pot med dejanskim izvorom zvoka in poslušalcem poskrbi za individualizacijo.

Spomnimo še enkrat, da kljub mnogim poskusom individualiziranja zvoka in s tem nepotrebnost prilagajanja oz. učenja pomena zvokov v akustični sliki prostora ostaja nerešen problem, kar je navsezadnje tudi vzrok poskusov iskanja optimalnih postopkov kodiranja elevacije izvorov zvoka v akustični sliki prostora. Primeren postopek kodiranja omogoča čim boljše ločljivost izvorov zvoka, poleg tega pa je hitro učljiv in je zgrajen iz zvokov, ki so prijetni za poslušanje. Iz rezultatov našega poskusa je razvidno, da obstajajo učljivi in ne-učljivi postopki (čeprav slednjih nismo načrtovali namenoma), ni pa bilo med učljivimi postopki statistično pomembne razlike v hitrosti učenja. Na koncu zaključimo z ugotovitvijo, da so rezultati naših poskusov vzpodbudni in dajejo motivacijo za nadaljnje raziskovanje v začrtani smeri.

LITERATURA

- [1] Algazi VR, Avendano C, Duda RO. Elevation localization and head-related transfer function analysis at low frequency. *Journal of Acoustical Soc. of America*, 109(3): str. 1110 – 1122, marec 2001.
- [2] Algazi VR, Duda RO, Thompson DM, Avedano C. The CIPIC HRTF Database. *Proceedings of 2001 IEEE Workshop on Applications of Signal Processing to Audio and Electroacoustics*, Mohonk Mountain House, New Paltz, ZDA, oktober 2001, str. 99 – 102.
- [3] Blauert J. Description of hearing experiments by means of a simple, system-theoretical model. *Kybernetik*, 6(2): 45 – 49, l. 1969.
- [4] Blauert J. *Spatial Hearing, The Psychophysics of Human Sound Localization*. MIT Press, tretji ponatis, 2001.
- [5] Bloom PJ. Creating Source Elevation Illusions by Spectral Manipulation. *Journal of the Audio Engineering Society*, 25(9): str. 560 – 565, september 1977.
- [6] Bransford JD, Brown AL, Cocking RR (uredniki): *How People Learn: Brain, Mind, Experience, and School*. National Academy Press, Washington D.C., 1999.
- [7] Bronkhorst AW. Localization of real and virtual sound sources. *Journal of Acoustical Soc. of America*, 98(5): str. 2542 – 2553, november 1995.
- [8] Bronštejn IN, Semendjajev KA, Musiol G, Muhlig H. *Matematični priručnik*. Tehniška založba Slovenije, 1997.
- [9] Brown CP, Duda RO. A Structural Model for Binaural Sound Synthesis. *IEEE Trans. on Speech and Audio Processing*, 6(5): str. 476 – 488, september 1998.
- [10] Cabrera D, Ferguson S, Tilley S. Recent studies on the effect of signal frequency on auditory vertical localization. *Proceedings of ICAD-05*, Limerick, Ireland, julij 2005.
- [11] Cheng CI, Wakefield GH. Introduction to head-related transfer functions (HRTF's): representations of HRTF's in time, frequency, and space (invited tutorial). *Journal of the Audio Engineering Society*, 49(4): str. 231 – 249, april 2001.
- [12] Cohen MM. Visual feedback, distribution of practice, and intermanual transfer of prism aftereffects. *Perceptual and motor Skills*, 37: str. 599 – 609, l. 1973.
- [13] Dizon RM, Litovsky RY. Localization dominance in the median-sagittal plane: Effect of stimulus duration. *Journal of Acoustical Soc. of America*, 115(6): str. 3142 – 3155, junij 2004.
- [14] Ebert J, Scientists with disabilities: Access all areas. *Nature*, 435: str. 552 – 554, junij 2005.
- [15] Fefer D, Jeglič A. *Osnove akustike*. Akademsko založba, Ljubljana, 1992.
- [16] Fefer D, Jeglič A. *Elektroakustika*. Založba FER, Ljubljana, 1993.
- [17] Ferguson BS, Bogner RE, Warwryk S. A bottle model for Head-Related Transfer Functions. *Proceedings ICASSP98*, Seattle, ZDA, maj 1998, str. 3533 – 3536.

- [18] Foxton JM, Brown ACB, Chambers S, Griffiths TD. Training Improves Acoustic Pattern Perception. *Current Biology*, 14: str. 322 – 325, l. 2004.
- [19] Gabriel G. *Child and Brain: The Stages of Development*. <http://www.brainconnection.com/topics/?main=fa/child-brain>, avgust 2005.
- [20] Gardner B, Martin K. HRTF Measurements of KEMAR Dummy-Head Microphone. *MIT Media Lab Perceptual Computing – Technical Report #280*, l. 1994.
- [21] Gilkey RH, Anderson TR (uredniki). *Binaural and spatial hearing in real and virtual environments*. Lawrence Erlbaum Associates, 1997.
- [22] González Mora JL, Rodríguez-Hernández A, Rodríguez-Ramos LF, Díaz-Saco L, Sosa N. Development of a new space perception system for blind people, based on the creation of a virtual acoustic space. *Proceedings of the International Work Conference on Artificial and Natural networks*, 1999, str. 321 – 330 (zvezek 2).
- [23] Grantham DW, Hornsby BWY, Erpenbeck EA. Auditory spatial resolution in horizontal, vertical, and diagonal planes. *Journal of Acoustical Society of America*; 114(2): str. 1009 – 1022, l. 2003.
- [24] Gupta N, Barreto A, Choudhury M. Modeling Head-Related Transfer Functions Based on Pinna Anthropometry. *Proceedings of LACCET 2004*, Miami, ZDA, junij 2004.
- [25] Hartmann WM, Rakerd B. Auditory spectral discrimination and the localization of clicks in the sagittal plane. *Journal of Acoustical Soc. of America*, 94(4): str. 2083 – 2092, oktober 1993.
- [26] Hartmann WM, Wittenberg A. On the externalization of sound images. *Journal of Acoustical Soc. of America*, 99(6): str. 3678 – 3688, junij 1996.
- [27] Hasegawa H, Kasuga M, Matsumoto S, Koike A. Simply Realization of Sound Localization Using HRTF Approximated by IIR Filter. *IEICE Trans. Fundamentals*, E83-A(6): str. 973 – 978, junij 2000.
- [28] Hasegawa-Johnson M. *Audio Engineering, Lecture Notes for ECE 403*. University of Illinois at Urbana-Champaign, januar 2005.
- [29] Held R. Plasticity in sensory-motor systems. *Scientific American*, 213(5): str. 84 – 94, november 1965.
- [30] Hofman PM, Van Opstal AJ. Bayesian reconstruction of sound localization cues from responses to random spectra. *Biological Cybernetics*, 84(4): str. 305 – 316, l. 1998.
- [31] Hofman PM, Van Opstal AJ. Spectro-temporal factors in two-dimensional human sound localization. *Journal of Acoustical Soc. of America*, 103(5): str. 2634 – 2648, maj 1998.
- [32] Hofman PM, Van Riswick JGA, Van Opstal AJ. Relearning sound localization with new ears. *Nature Neuroscience*, 1 (l. 1998): str. 417 – 421.
- [33] Jin C, Corderoy A, Carlile S, Van Schaik A. Contrasting monaural and interaural spectral cues for human sound localization. *Journal of Acoustical Soc. of America*, 115(6): str. 3124 – 3141, junij 2004.
- [34] *The bat 'K' sonar cane*: <http://www.batforblind.co.nz/>.
- [35] Kladnik R. *Visokošolska fizika 3. del – valovni pojavi*. DZS, Ljubljana, 1989.
- [36] Kuhn GF. Model for the interaural time differences in the azimuthal plane. *Journal of Acoustical Soc. of America*, 62(1): str. 157 – 167, julij 1977.

- [37] Kulkarni A, Colburn HS. Infinite-impulse-response models of the head-related transfer functions. *Journal of Acoustical Soc. of America*, 115(4): str. 1714 – 1728, april 2004.
- [38] Langendijk EHA, Bronkhorst AW. Fidelity of three-dimensional-sound reproduction using virtual auditory display. *Journal of Acoustical Soc. of America*, 107(1): str. 528 – 537, januar 2000.
- [39] Langendijk EHA, Bronkhorst AW. Contribution of spectral cues to human sound localization. *Journal of Acoustical Soc. of America*, 112(4): str. 1583 – 1595, oktober 2002.
- [40] Lessard N, Pare M, Lepore F, Lassonde M. Early-blind human subjects localize sound sources better than sighted subjects. *Nature*, 395(6699): str. 278 – 280, september 1998.
- [41] Levitt H. Transformed Up-Down Methods in Psychoacoustics. *Journal of Acoustical Soc. of America*, 49(2): str. 467 – 477, februar 1971.
- [42] Macpherson EA, Middlebrooks JC. Localization of brief sounds: Effects of level and background noise. *Journal of Acoustical Soc. of America*, 108(4): str. 2834 – 1849, oktober 2000.
- [43] Ewan A. Macpherson, John C. Middlebrooks. Vertical-plane sound localization probed with ripple-spectrum noise. *Journal of Acoustical Soc. of America*, 114(1): str. 430 – 445, julij 2003.
- [44] Meijer PBL. An Experimental System for Auditory Image Representations. *IEEE Trans. on Biomedical Engineering*, 39(2): str. 112 – 121, februar 1992.
- [45] Meijer PBL. Seeing with Sound for the Blind: Is it Vision. *Conference on Consciousness*, Tucson, ZDA, april 2002.
- [46] Middlebrooks JC, Green DM. Sound localization by human listeners. *Annual Rev. Psychology*, 42: 135 – 159, l. 1991.
- [47] Miller GA. Sensitivity to Changes in the Intensity of White Noise and Its Relation to Masking and Loudness. *Journal of Acoustical Soc. of America*, 19(4): str. 609 – 619, julij 1947.
- [48] Mills AW. On the minimum audible angle. *Journal of Acoustical Soc. of America*, 30(4): str. 237 – 246, april 1958.
- [49] Mlakar J. *Elektromagnetno valovanje*. Fakulteta za elektrotehniko, Ljubljana 2002.
- [50] Montgomery DC, Runger GC. *Applied Statistics and Probability for Engineers*. Wiley & Sons, 2003.
- [51] Nandy D, Ben-Arie J. An Auditory localization model based on high frequency spectral cues. *Annals of Biomedical Engineering*, 24(6): str. 621 – 638, l. 1996.
- [52] Neely ST, Allen JB. Invertibility of room impulse response. *Journal of Acoustical Soc. Of America*, 66(1): str. 165 – 169, julij 1979.
- [53] Neuhoff JG (urednik). *Ecological Psychoacoustics*. Elsevier Inc., 2004.
- [54] Nishino T, Hosoe S, Takeda K, Itakura F. Measurement of the Head Related Transfer Function using the Spark Noise. *Proceedings of the 18th International Congress on Acoustics (ICA2004)*, Kyoto, Japonska, 2004, str.1437 – 1438 (zvezek II).
- [55] Pickett JM, Daly RL, Brand SL. Discrimination of spectral cutoff frequency in residual hearing and in normal hearing. *Journal of the Acoustical Society of America*, 38(5): str. 923(A), november 1965.
- [56] Rao KR, Ben-Arie J. Optimal Head Related Transfer Functions for Hearing and Monaural Localization in Elevation: A Signal Processing Design Perspective. *IEEE Transactions on Biomedical Engineering*, 43(11): str. 1093 – 1115, november 1996.

- [57] Raykar VC, Duraiswami R, Yegnanarayana B. Extracting the frequencies of the pinna spectral notches in measured head related impulse responses. *Journal of Acoustical Soc. of America*, 118(1): 364 – 374, julij 2005.
- [58] Rodriguez SG, Ramirez MA. Extracting and modeling approximated pinna-related transfer functions from HRTF data. *Proc. of ICAD 05 – Eleventh Meeting of the International Conference on Auditory Display*, Limerick, Irska, 6. – 9. julij, 2005, str. 269 – 273.
- [59] Rogers ME, Butler RA. The linkage between stimulus frequency and covert peak areas as it relates to monaural localization. *Perception and Psychophysics*, 52(5): 536 – 546, november 1992.
- [60] Van Schaik A, Jin C, Carlile S. Human Localisation of Band-Pass Filtered Noise, *Proceedings of 2nd European Workshop on Neuromorphic Systems (EWNS2)*, Stirling, Škotska, september 1999.
- [61] Shinn-Cunningham BG, Durlach NI, Held RM. Adapting to supernormal auditory localization cues. I. Bias and resolution. *Journal of Acoustical Soc. of America*, 103(6): str. 3656 – 3666, junij 1998.
- [62] Shinn-Cunningham BG, Durlach NI, Held RM. Adapting to supernormal auditory localization cues. II. Constraints on adaptation of mean response. *Journal of Acoustical Soc. of America*, 103(6): str. 3667 – 3676, junij 1998.
- [63] Sodnik J, Sušnik R, Tomažič S. Acoustic signal localization through the use of Head Related Transfer Functions. *Journal of systemics, cybernetics and informatics*, 2(6): str. 1 – 4, l. 2004.
- [64] Sodnik J, Sušnik R, Bobojević G, Tomažič S. Smerna ločljivost navideznih izvorov zvoka pri človeku. *Elektroteh. vestn.*, 71(3): str. 121 – 127, l. 2004.
- [65] Sodnik J, Sušnik R, Štular M, Tomažič S. *Spatial sound resolution of an interpolated HRIR library*, Applied Acoustics, 66: str. 1219 – 1234, l. 2005.
- [66] Sodnik J, Sušnik R, Tomažič S. Merjenje individualnih prenosnih funkcij glave in ušes, *Zbornik štirinajste mednarodne Elektrotehniške in računalniške konference ERK 2005*, Portorož, str. 265 – 268 (zvezek B), september 2005.
- [67] Sodnik J, Sušnik R, Tomažič S. Principal Components of Non-Individualized Head Related Transfer Functions Significant for Azimuth Perception. *Acta Acustica united with Acustica*, 92(2): str. 312 – 319, marec-april 2006.
- [68] Sušnik R, Sodnik J, Tomažič S. Elevation coding in auditory image. Poslano v revijo *Applied Acoustics*, junij 2006.
- [69] Sušnik R, Sodnik J, Tomažič S. Sound source choice in HRTF acoustic imaging. *HCI International 2003: adjunct proceedings*. Heraklion, Grčija, junij 2003, str. 101 – 102.
- [70] Sušnik R, Sodnik J, Tomažič S. Vpliv dolžine impulza na ločljivost v akustični sliki prostora. *Zbornik dvanajste mednarodne Elektrotehniške in računalniške konference ERK 2003*, Ljubljana, Slovenija, september 2003, str. 91 – 94.
- [71] Sincaglia N, Rebaud S. Product design considerations for implementing 3D audio algorithms within existing multichannel formats. *Proceedings of AES 16th Int. conference on Spatial Sound Reproduction*, Rovaniemi, Finska, april 1999.
- [72] Tasič J. *Uvod v postopke digitalne obdelave signalov*. Založba FE in FRI, Ljubljana, 2001.
- [73] Tomažič S, Leonardis S. *Digitalni signali in sistemi*. Založba FE in FRI, Ljubljana, 2004.
- [74] Vickers DA, Faulkner A. The discrimination of the bandwidth of noises by normal-hearing and severe-to-profoundly hearing-impaired listeners. *Speech Hearing and Language: work in progress*, 10, l. 1997 (<http://www.phon.ucl.ac.uk/home/sh110/debi/debishl.htm>).

- [75] Vliegen J, Van Opstal AJ. The influence of duration and level on human sound localization. *Journal of Acoustical Soc. of America*, 115(4): str. 1705 – 1713, april 2004.
- [76] Watkins AJ. Psychoacoustical aspects of synthesized vertical local cues. *Journal of Acoustical Soc. of America*, 63(4): str. 1152 – 1165, april 1978.
- [77] Wightman FL, Kistler DJ. Headphone simulation of free-field listening. I: Stimulus synthesis. *Journal of Acoustical Soc. of America*, 85(2): str. 858 – 867, februar 1989.
- [78] Wightman FL, Kistler DJ. Headphone simulation of free-field listening. II: Psychophysical validation. *Journal of Acoustical Soc. of America*, 85(2): str. 868 – 878, februar 1989.
- [79] Wightman FL. Monaural sound localization revisited. *Journal of Acoustical Soc. of America*, 101(2): str. 1050 – 1063, februar 1997.
- [80] Spletna enciklopedija Wikipedia, www.wikipedia.org.
- [81] Zahorik P. Direct-to-reverberant energy ratio sensitivity. *Journal of Acoustical Soc. of America*, 112(5): str. 2110 – 2117, november 2002.
- [82] Zwiers MP, Van Opstal AJ, Cruysberg JRM. Two-dimensional sound-localization behavior of early-blind humans. *Experimental Brain Research*, 140 (2): str. 206 – 222, september 2001.
- [83] Zwicker E, Flottorp G, Stevens SS. Critical bandwidth in loudness summation. *Journal of Acoustical Soc. Of America*, 29(5): str. 548 – 557, maj 1957.
- [84] Zwicker E. Subdivision of the audible frequency range into critical bands (Frequenzgruppen). *Journal of Acoustical Soc. Of America*, 33(2): str. 248, februar 1961.
- [85] Zwislocki J, Maire F, Feldman AS, Rubin H. On the Effect of Practice and Motivation on the Threshold of Audibility. *Journal of Acoustical Soc. of America*, 30(4): str. 254 – 262, april 1958.
- [86] Žibera G, Zazula D. Računalniška tvorba 3D zvoka v virtualnih prostorih. *Elektroteh. vestn.*, 70(3): str. 96 – 102, 1. 2003.

SEZNAM KRATIC

CIPIC	Center for Image Processing and Integrated Computing	Laboratorij (University of California Davis) katerega banka HRTF je prosto dostopna na spletu (http://www.cipic.ucdavis.edu/).
HRIR	Head Related Impulse Response	Impulzni odziv glave in telesa, tj. impulzni odziv prenosne poti med izvorom zvoka in poslušalcem.
HRTF	Head Related Transfer Function	Fourierov transform impulznega odziva glave in telesa (HRIR).
IAD	Interaural Amplitude Difference	Medušesna amplitudna razlika, tj. razlika v amplitudah zvočnega vala na zasenčenem in nezasenčenem ušesu (enako kot ILD).
IEC	International Electrotechnical Commission	Mednarodno standardizacijsko telo za področje elektrotehnike (http://www.iec.ch/).
ILD	Interaural Level Difference	Medušesna amplitudna razlika, tj. razlika v amplitudah zvočnega vala na zasenčenem in nezasenčenem ušesu (enako kot IAD).
ISO	International Standards Organization	Mednarodno standardizacijsko telo (http://www.iso.org/).
ITD	Interaural Time Difference	Medušesna časovna razlika, tj. zakasnitev med trenutkoma ko zvočno valovanje doseže nezasenčeno in zasenčeno uho.
JND	Just Noticeable Difference	Minimalne zaznavne razlike (v glasnosti, frekvenci, času ...).
KEMAR	Knowles Electronic Manikin for Acoustic Research	Posebna lutka, ki se uporablja predvsem za merjenje splošnih HRTF.

MAA	Minimum Audible Angle	Minimalna kotna ločljivost dveh statičnih izvorov zvoka.
MAMA	Minimum Audible Movement Angle	Minimalna kotna ločljivost dveh premikajočih se izvorov zvoka.
MIT	Massachusetts Institute of Technology	Laboratorij katerega banka HRTF je prosto dostopna na spletu (http://www.mit.edu/).
RIR	Room Impulse Response	Impulzni odziv prostora.
RT60	Reverberation Time 60 dB	Čas v katerem pade prvotna zvočna energija na eno milijoninko začetne vrednosti (kot posledica odbojev – reverberacije v prostoru).
vOICE	Oh, I see.	Projekt akustične slike prostora nizozemskega fizika Petra Meijerja (http://www.seeingwithsound.com/).

IZJAVA

Izjavljam, da sem doktorsko disertacijo samostojno izdelal pod vodstvom mentorja prof. dr. Saša Tomažiča, rednega profesorja na Fakulteti za elektrotehniko Univerze v Ljubljani. Izkazano pomoč drugih sodelavcev sem v celoti navedel v zahvali.

Rudolf Sušnik