

HABILITÁCIÓS TÉZISEK

**Szoftveres eljárások akusztikai mérőrendszer és virtuális valóság
szimulátorok optimalizálására látó és vak felhasználók
összehasonlításában**

szerző

DR. WERSÉNYI GYÖRGY

egyetemi docens

Győr

2013

Előzmények

Az akusztikai mérés technika és a térhallás-vizsgálatokhoz szükséges mérőrendszerek fejlődése a digitális technika megjelenésével és a számítástechnika fejlődésével új lehetőségekhez jutott. A korábbi lassú analóg berendezéseket először a hardveres számítástechnika (célhardverek, külső kártyák és érzékelők), majd a tisztán szoftveres megoldások váltották fel. A fejlődés az elmúlt évtizedben lehetővé tette olyan szimulációk valós-idejű megvalósítását, amelyek korábban nem voltak lehetségesek. Ez vonatkozik a mérés technikára, a mért adatok kiértékelésére és azok szimulációban történő felhasználására, valamint végtermékek, alkalmazások implementálására az ember-gép kommunikáció javításának céljából. Utóbbinál különösen érdekesek és fontosak a vakok és látássérültek segítségét célzó alkalmazások, amelyek könnyebbé, gyorsabbá tehetik számukra az oly nehéz grafikus felületeken történő eligazodást.

Vizsgálataim és kutatásom legfontosabb iránya az emberi tér- és irányhallás lehetőségeinek felderítése, beleértve a virtuális valóság szimulátorokat és a szabadtéri lokalizációs vizsgálatokat is. Az emberi irányhallás legfontosabb paramétere a két füljel között fellépő ún. interaurális idő- és szintkülönbségek, amelyek alapján határozzák meg a vízszintes síkban történő lokalizációt. Továbbá, a függőleges síkban – ahol ezek a különbségek zérussá válnak – a külső fül komplex átviteli függvényéi (HRTF – Head-Related Transfer Functions) szolgáltatják az irányinformációt. A binaurális technika egyik legfontosabb területe a HRTF-függvények mérés technikája és a szimulációkban történő felhasználhatósága. A modern mérőrendszerek számítógép-vezérelt, automatizált, sokszor műfejes megoldásokat tartalmaznak. A létrejött adatbázisok feldolgozása szintén komoly erőforrásokat igényel, a valós-idejű implementálás (FIR- vagy IIR-szűrők segítségével, esetleg időtartományi realizációval) lehetőségei pedig gyakran kérdésesek. Munkám egyik része arra irányult, hogy miként lehet egy nagy pontosságú számítógépes műfejes HRTF-mérőrendszert kialakítani, nagy mennyiségű pontos mérést végrehajtani, a létrejött adatbázist kiértékelni, majd virtuális valóság szimulátorok számára elérhetővé és implementálhatóvá tenni.

A virtuális valóság szimulátorok esetében tisztán hangalapúak, amelyeket VAD-nak nevezünk (Virtual Audio Display). Általános jellemzője, hogy a tesztalanyok fejhallgatón keresztül kapják a szimulált hangteret, a hangforrások irányát a HRTF-szűrők implementálásával és/vagy az interaurális eltérések szabályozásával érik el. Egyik rendszer sem mentes a hibáktól (fejhallgatóhibák, kiegyenlítés vagy a fejmozgás hiánya stb.), így az optimalizálás és az újabb vagy egyszerűbb eljárások keresése, bevetése indokolt. A vizsgálatok eredményei később felhasználhatók valós alkalmazások esetén bármely célcsoport számára.

A célcsoportok között kiemelkedő a vak és látássérült emberek halmaza, akik a mindennapi tájékozódásban és a számítógépes munkában is igénylik a hangokat (elsősorban a beszédet). A beszéd hátránya, hogy nyelvfüggő és gyakran monoton, unalmas. Csak az utóbbi időben kerültek elő azok a minőségi felolvasók, amelyek hosszútávon is élvezhető minőséget szolgáltatnak. A szokványos felolvasóprogramok és beszédszintézis (Text-To-Speech, TTS) mellett a vakok egyéb hasznos, szórakoztató és a géppel való kommunikációt megkönnyítő alkalmazásokat is keresik. Ezek az informatikai megoldások felgyorsíthatják, egyszerűbbé és könnyebbé tehetik a szoftveres alkalmazásokat, vagy éppen elősegíthetik hardveres eszközök fejlesztését (navigációs és tájékozódás-segítő berendezések). Az ipari alkalmazások az ún. O&M (Orientation and Mobility) témakörben fejlesztenek elektronikus navigációs berendezéseket (ETA – Electronic Travel Aids), amelyek akár irányinformációval ellátott hangokat is használhatnak.

Végül, de nem utolsósorban, közelebb kerülhetünk az alap kutatás témakörébe tartozó kérdéseknek is az akusztikai információ átvitele és feldolgozása témájában. A hallásmodellezés foglalkozik a hallórendszer leírásával a fülektől az agyig terjedő pályán, és minden új eredmény hozzájárulhat a modellek kiegészítéséhez, a perifériális és centrális információfeldolgozási folyamatok megértéséhez és az agy ill. a HRTF-függvények szintjén történő feldolgozás kapcsolatához. Úton a kutatás során arra is keressük a választ, vajon a vakok jobban hallanak-e bármilyen értelemben is látó társaiknál akár virtuális, akár szabadtéri lokalizációs feladatok során. Sokszor feltételezzük ennek automatikus megvalósulását és igaz voltát, holott a különbségek – ha léteznek – a tanult agyi folyamatokban rejlenek. Ez a témakör azonban már távolabb vezet a mérnöki és informatikai alkalmazásoknál.

A következőkben bemutatott és publikációkkal alátámasztott téziscsoportok olyan informatikai megoldásokat foglalnak össze, melyeket az elmúlt évtizedben a fenti kutatási téma körében dolgoztam ki.

Téziscsoport 1.

Informatikai és programozástechnikai megoldásokat dolgoztam ki virtuális valóság szimulátorok számára, amelyek lehetővé teszik a szükséges mérések nagy pontosságú elvégzését és kiértékelését, a lehallgatási tesztek hatékony végrehajtását és a hibák csökkentését, valamint az egyszerű implementációt.

- 1.1 Kidolgoztam egy olyan algoritmust, amely rövid futási idő alatt képes a mérőrendszer zaja és átvitele alapján olyan álvéletlen, szélessávú gerjesztő jel előállítására, amely frekvenciafüggetlen és maximális jel-zaj-viszonyt tesz lehetővé az akusztikai mérőrendszerek számára. Az így előálló pszeudo-random jel az MLS-módszerekkel vagy a Golay-kódokkal egyenértékű szélessávú gerjesztésnek felel meg, egyhez közeli csúcstértékű az időtartományban, ugyanakkor a rendszer saját zajának spektrumát közelíti a frekvenciatartományban. Így optimálisnak tekinthető mindkét tartományban, frekvenciafüggetlen és maximális jel-zaj viszonyú mérést tesz lehetővé az egész (akusztikai) sávban. Hatékonyságát az álvéletlen tulajdonsága miatt a periodikusan ismételt mérések során bekövetkező átlagolás során fejti ki legjobban. Ezzel a módszerrel előállított gerjesztéssel a szokásos kb. 60 dB-es jel-zaj-viszonyú akusztikai mérési környezetek helyett egy annál jóval nagyobb, akár 89 dB feletti eredményt sikerült elérni. Méréssel az is belátható, hogy ennél nagyobbra nincs is szükség (az elérhető maximum, amit a jelfeldolgozás korlátoz, 102 dB). A bemutatott algoritmus a jel előállítása során felváltva dolgozik az idő- és a frekvenciatartományban, hogy a fenti két feltételnek megfelelő kimeneti jelet hozzon létre. [1]
- 1.2 A kiépített műfejes mérőrendszer segítségével és az általam bevezetett HRTFD-számítási eljárással megállapítottam, hogy a fejhez közeli hangtérben bekövetkező apró változások a HRTF-ekre akár 20 dB-es módosító hatással is lehetnek az iránytól és a frekvenciától függően. Ilyen mérésekhez a szükséges mérőrendszer felbontása a szokásosnál jóval finomabb kell legyen: 1 fokos horizontális és 5 fokos vertikális forrásirány-beállítását valósítottam meg, valamint 1 dB-es mérési bizonytalanságot a

teljes frekvenciasávban és térirányokban. A virtuális szimuláció során jelentős, a szabadtéri lokalizáció során azonban elhanyagolható hatásúak lehetnek ezek a változások (pl. a fejhez közeli mindennapi tárgyak, mint a haj, szemüveg, ruházat stb. hatása). Mindez az akusztikai információ agyi feldolgozásának fontosságára utal: míg korábban csupán perifériális (a fülek szintjén végbemenő) folyamatnak tekintették a HRTF-szűrést és feldolgozást, méréseim újabb eredményekkel támasztották alá, hogy a centrális (agyi) információfeldolgozás már ezen a szinten is aktív. Definiáltam a HRTFD fogalmát: a HRTFD (HRTF Difference) azonos forrásirányból, de eltérő környezeti paraméterek mellett mért HRTF-ek spektrális különbsége a frekvenciában (dB-ben). Előnye, hogy nem szükséges az individuális (emberi) mérés, mert a különbségképzés során az individuális jellemzők kiesnek: lehetőség van, sőt, indokolt a nagy pontosságú műfejes mérőrendszer használata. Továbbá meghatároztam és kimértem a fül binaurális és monaurális érzékenységi tartományait, a fejrányék hatását és határfrekvenciájának térbeli ill. frekvenciabeli függését, valamint a fülkagyló még idáig meg nem mért hatását 11 kHz környékén, 70-80 fokos horizontális irányból. E fülkagyló-hatás felderítése csak a mérőrendszer egyfokos térbeli és 1 dB-es felbontásbeli pontosságának, valamint a HRTFD-eljárás robusztusságának volt köszönhető. [2, 5]

- 1.3 Virtuális valóság szimulátorokban a helyes döntéseket előmozdító eljárást dolgoztam ki valós-idejű, egyszerű alul- és felüláteresztő szűréssel a virtuális hangforrásirányok detekciójához. Ezzel az eljárással a vertikális lokalizációhoz szükséges függőleges irányú HRTF-szűrés akár ki is küszöbölhető a szimuláció során. A korábbi mérésekből megállapítható volt, hogy a vertikális forrásirányok a felhasználók többsége számára nem detektálható (lokalizációhiány), valamint, hogy egy virtuális hangfelszínen megvalósított 3x3 ill. 5x2-es kétdimenziós térbeli felbontás használhatósága (a találati arány) a járulékos szűréssel jelentősen megnövelhető a felhasználók 90%-a számára. [3, 4]
- 1.4 Kidolgoztam egy számítógépes szimulációs eljárást a drága és körülményesen használható külső hardveres fejhelyzet-érzékelő részleges kiváltására, amely a fej mozgatása helyett pusztán szoftveres alapon a hangforrás mozgatásával éri el a fejközép-lokalizációs hiba jelentős csökkentését. Az elv az, hogy az életben jelenlévő apró, néhány fokos fejmozdításokat a virtuális szimuláció során a hangforrás néhány fokos

mozgatásával váltam ki. Ekkor a célpontnak tekintett hangforrásirány stabilan tartandó helyzete helyett, azt a HRTF-szűrők gyors, dinamikus változtatásával néhány fokkal véletlenszerűen mozgatom anélkül, hogy a hallgató ezt a mozgást érzékelné. Megállapítottam, hogy a forrás ilyesfajta virtuális, véletlenirányú, kétfokos rezgetése a fejközép-lokalizációs hibát az alanyok egyharmada számára megszüntette, ugyanakkor az elől-hátul döntések hibaszázalékát nem befolyásolta. A négyfokos mozgást az alanyok már mozgásnak érzékelték. Továbbá, hét különféle jellel és rögzített kétfokos mozgatással dolgozva bizonyítottam, hogy a szélessávú jelek esetén az eljárás hatékonyabb, és hogy a mozgítás sebessége ezen a foktartományon belül nem kritikus: egy átlagosnak tekinthető 2 Hz-es frissítés megfelelő sebességű. [6]

- 1.5 MATLAB alatt programozva összehasonlító kiértékelést végeztem a megmért műfejes HRTF-függvényeken, különös tekintettel a lehetséges 2D és 3D ábrázolásmódok sajátosságaira. Megmutattam, hogy a 2D normál négyszögletes ábrázolás mellett a polárdiagram is alkalmas a nagy mennyiségű HRTF- vagy HRTFD-adatsereg megjelenítésére, míg a szintfelületes kontúrábra nem. Az általam megmért és összeállított műfejes adatbázist korábbi mérésekkel összehasonlítva látványosabb és jobban használható ábrákat hoztam létre, továbbá mozgókép-videószekvenciát is. Létrehoztam egy MATLAB alatt működő virtuális valóság szimulátort, amely egyben képes a fejhallgató-kiegyenlítés, a valós-idejű vagy off-line HRTF-szűrés és a különböző ITD-beállításokon keresztül történő hangforrás-szimulációra a jövőben. [8]

Téziscsoport 2.

Méresi eljárások és pszichoakusztikai vizsgálatok során vakok és látók statisztikai összehasonlítása alapján virtuális és szabadtéri (kültéri) feladatok megoldásában bizonyítottam, hogy a vak alanyok nem teljesítenek lényegesen jobban a virtuális környezetben és a szabadtéri feladatok során sem mindig, így az esetleges jobb lokalizációs képességeik nem perifériális szintűek, hanem tanult agyi folyamatokkal vannak összefüggésben. A „kifinomultabb hallás” nem jelent jobb érzékenységet és nem feltétlenül alakul ki bármiféle emeltszintű képesség a látás elvesztésével automatikusan. Sok vak nem képes az önálló közlekedésre, külső referenciahangforrás nélkül ún. egyenestartásra vagy megtanulni a Braille-írást sem. Ugyanakkor, az echolokáció során képesek lehetnek reflexiók alapján pontosabb akadályérzékelésre (nagy felületű tárgyak esetén).

2.1 Virtuális valóság szimulátor esetén látók és vakok összehasonlításából megmutattam, hogy a fejhallgató és a HRTF-szűrés bevezetése mindkét csoport számára azonos mértékű lokalizációs hibákhoz (fejközép-lokalizáció, térbeli felbontóképesség) vezet. A vakok azonban kevesebb elől-hátul döntés hibát vétének, ami az ezzel összefüggésben lévő vizsgálatok során jobb értékeket eredményeznek. A Z-teszt és T-teszt statisztika megmutatta, hogy a vakok szignifikánsan jobbak a látóknál a 3x3 VAD felszínen és a „szemben” lévő források detektálása esetén. Hasonlóan, a fej körüli „keringő” mozgás irányának megállapításában is. Látók ugyanakkor jobbnak bizonyultak „hátról” irányból szimulált források esetén, de ez betudható a nagyobb számú elől-hátul döntések hibájának. Egyéb, mozgó források közötti irányeltérés megállapításában és a fejközép-lokalizációs eredményekben nem volt szignifikáns az eltérés a két csoport között. További paraméterek, mint életkor, nem, zenei képzettség, abszolút hallás vagy a vakság foka és a megvakulás ideje sem volt befolyásoló tényező. Virtuális vizsgálatokhoz nem szükséges a két csoportot szétválasztani, látókkal nyert eredmények alkalmazhatók vakokra is, ami megkönnyíti a pszichoakusztikai lehallgatási tesztek kivitelezését. A vakok tehát pusztán az „elől” irány megállapításában voltak jobbak, amely feltehetőleg arra vezethető vissza, hogy a valós életben gyakrabban találkozunk ilyen helyzetekkel, különösen számítógép- és fejhallgató-használat során. [10]

- 2.2 Kültéri, szabad térhez közeli O&M (Orientation and Mobility) vizsgálatokban megmutattam, hogy a vakok külső akusztikai jel követése nélkül ugyanúgy nem képesek egyenesen menni (egyenestartás), ahogy a bekötött szemű látók sem. Szélessávú ill. nagy frekvenciás akusztikus céljel esetén (fehérszaj, click-train) mindkét csoport képes egyenesen menni, így ebben a képességben nincs különbség a két csoport között. Echolokáció esetén a fehér bot okozta visszaverődéseket a vakok jobban hasznosítják „csendes objektumok”, pl. sarokészlelés során. Bár a fehér bot elsősorban nem akusztikus szerepet tölt be a navigáció során, a kopogtatás mint akusztikai információ is hasznosítható és a hasznosítása tanulható. Az épület sarkának megtalálása vakok esetén szignifikánsabb jobb volt a bekötött szemű látókhoz képest. Megismételt kísérlet során javulás a második menetben az elsőhöz képest nem volt tapasztalható egyik csoport esetén sem, így a tanulási folyamat ennél sokkal több időt vesz igénybe ilyen esetekben. [9, 10]
- 2.3 Szoftveres megoldást fejlesztettem ki a szokványos beszédalapú Text-to-speech (TTS) alkalmazások kiegészítésére, amely során nyelvfüggetlen hanghatásokkal a számítógépes környezet felhasználhatóságát javítani, a GUI-hoz való hozzáférést gyorsítani lehet. Ez elsősorban vak felhasználók számára készült, de a látók is profitálhatnak belőle az akusztikus visszacsatolás útján. Ennek során új akusztikus hangjel-osztályt definiáltam és hoztam létre (auditory emoticons). E hangok jellemzője, hogy emberi hangon alapulnak, de nem beszédjelek, így nyelvfüggetlenek, könnyen létrehozhatók felvétellel (akár férfi és női verzióban is), gyorsan tanulhatók, használatuk szórakoztató és képesek érzelmi tartalom továbbítására. [7]
- Rögzítettem (emberi felvételek és TTS-felolvasó alapján) és kiértékeltem a speciálisan tömörített ún. spearcon jeleket magyar és német nyelvre. A spearcon-adatbázisok rövid, beszédalapú, időben tömörített hangjelek, amelyek elsősorban menünavigáció, névfelolvasás esetén robusztusak. Az angol adatbázis mellé elkészítettem a magyart és különböző tájsozlásokkal a német változatokat is. Megállapítottam, hogy felhasználhatóságuk jó és indokolt, rövid tanulási szakasz után bevethetők olyan események leírására, amelyhez a szokványos hangok nem rendelhetők hozzá hatékonyan. Nem volt egyértelmű preferencia a férfi vagy női beszélő ill. az akcentusmentes német változat javára, így feltételezhető, hogy a tömörítés során ez az információ elvész és nem lényeges a beszélő személye. [7]

Kiválasztottam látók és vakok összehasonlító-kérdőíves kikérdezése alapján azt a 30-40 legfontosabb eseményt, programot, alkalmazást stb., amelyeket beszéd helyett vagy mellett hanghatásokkal lehet és érdemes is kiegészíteni. Az ehhez a szükséges hanghatásokat rögzítettem, jelfeldolgozási eljárásokkal optimalizáltam, a felhasználók által minősítettem és adatbázis formájában elérhetővé tettem. Ez a hangadatbázis az auditory emoticonokkal, a spearconokkal és a TTS-felolvasóval együtt alkalmazható, így bárki képes a használat gyors megtanulására. A későbbiekben ez az interfész implementálásra kerül, és térbeli információval is kiegészíthetővé válik. [7]

Összefoglalás és jövőbeni munka

Az itt bemutatott két téziscsoportban olyan informatikai eljárásokat mutattam be, amelyek az akusztikai mérés technikában fontos HRTF-mérések, valamint azok implementációját megkövetelő virtuális valóság szimulátorok optimalizációja felé vezetnek. A HRTF-függvények analízise mentén fény derült az emberi agy néhány információfeldolgozási folyamatára is. Sikerült pár egyszerű, valós-idejű, könnyen megvalósítható módszerrel néhány a szimulációk során bekövetkező hibát csökkenteni, vagy éppen drága célhardvert részben kiváltani. A kutatási eredményekből kész termékek előállítás is cél a jövőben, így egy komplett virtuális valóság szimulátor MATLAB alatt, egy a vakokat segítő képfeldolgozással együttműködő navigációs eszköz kifejlesztése, illetve egy szoftver a hangadatbázis felhasználásával. Végül, de nem utolsósorban, az alap kutatási téma is aktuális: vajon milyen következtetések vonhatók le a vakok és a látók hallásképességének, információfeldolgozásának összehasonlításából – avagy jobban hallanak-e a vakok, mint a látók?

Válogatott publikációk

- [1] WERSÉNYI, GY., ILLÉNYI, A., *Test Signal Generation and Accuracy of Turntable Control in a Dummy-Head Measurement System*, Journal of the Audio Engineering Society (JAES), Vol. 51., No.3, March 2003, pp. 150-155.
- [2] WERSÉNYI, GY., ILLÉNYI, A., *Differences in Dummy-Head HRTFs Caused by the Acoustical Environment Near the Head*, Electronic Journal of "Technical Acoustics" (EJTA), Russia, <http://www.ejta.org>, 2005, 1. /15 pages/
- [3] WERSÉNYI, GY., *Localization in a HRTF-based Minimum-Audible-Angle Listening Test for GUIB Applications*, Electronic Journal of "Technical Acoustics" (EJTA), Russia, <http://www.ejta.org>, 2007, 1. /16 pages/
- [4] WERSÉNYI, GY., *Localization in a HRTF-based Virtual Audio Synthesis using additional High-pass and Low-pass filtering of Sound Sources*, Journal of the Acoust. Science and Tech. of Japan, Vol. 28, Nr.4, 2007 July, pp. 244-250.
- [5] WERSÉNYI, GY., *Directional Properties of the Dummy-Head in Measurement Techniques based on Binaural Evaluation*, Journal of Engineering, Computing & Architecture, Vol. 1, No. 2, 2007. www.scientificjournals.org. /15 pages/
- [6] WERSÉNYI, GY., *Effect of Emulated Head-Tracking for Reducing Localization Errors in Virtual Audio Simulation*, IEEE Transactions on Audio, Speech and Language Processing, Vol. 17, No. 2. February 2009, pp. 247-252.
- [7] WERSÉNYI GY., *Auditory Representations of a Graphical User Interface for a Better Human-Computer Interaction*. in S. Ystad et al. (Eds.): Auditory Display. CMMR/ICAD 2009 post proceedings edition, Lecture Notes in Computer Science (LNCS) 5954, Springer Verlag, Berlin, 2010, pp. 80-102.
- [8] WERSÉNYI, GY., *Representations of HRTFs using MATLAB: 2D and 3D plots of accurate dummy-head measurements*. Proceedings of the 20th International Congress on Acoustics (ICA), Sydney, Australia, 2010 Aug. 23-27. /9 pages/
- [9] WERSÉNYI, GY., RÉPÁS, J., *The Influence of Acoustic Stimuli on "Walking Straight" Navigation by Blindfolded Human Subjects*. Acta Technica Jaurinensis, Vol. 5, No. 1, 2012, pp. 1-18.
- [10] WERSÉNYI, GY., RÉPÁS, J., *Comparison of Localization Performance of Blind and Sighted Subjects on a Virtual Audio Display and in Real-life Environments*. Audio Engineering Society (AES) Convention Preprint Paper 8611, AES 132nd Convention, Budapest, Hungary, 2012 April 26-29. /6 pages/

Dr. Wersényi György

Szoftveres eljárások akusztikai mérőrendszer és virtuális valóság szimulátorok
optimalizálására látó és vak felhasználók összehasonlításában