

Konvolúciós megoldások hangtér-szimulációs feladatokhoz

Convolution: methods for soundfield rendering

WERSÉNYI GYÖRGY

Széchenyi István Egyetem, Távközlési Tanszék, 9026, Győr, Egyetem tér 1.

Absztrakt:

A hangtér-szimulációs eljárások egyik legfontosabb része, amikor a nyers hanganyagot a szükséges térinformációval „ellátjuk”. Ennek során manapság digitális jelfeldolgozási eljárásokkal hozunk létre irányinformációt, térérzetet, utószó (auralizáció). A konvolúciót megvalósíthatjuk a frekvenciatartományban és a mai számítástechnika segítségével az időtartományban is. Ebben a cikkben ismeretterjesztő jelleggel néhány lehetőséget mutatunk be: ismert hangfeldolgozó- és teremakusztikai programokat, másrészt MATLAB alatt megvalósítható és futtatható környezetben is létező megoldást. Továbbá bemutatunk egy binaurális felvevőrendszert, amely a fenti algoritmusokat kiszolgálhatja a hangtér-szimulációs feladatok során.

Abstract:

The most important part of the sound field reproduction is the simulation of spatial and directional information (auralization). Digital signal processing methods help us to add this directional information, the feeling of room size or reverberation. This can be realized in the frequency domain and using state-of-the-art computer technique, also in the time domain. This paper presents some methods and software applications to render fast convolution of input data and impulse response. Some commercially available software solutions and a MATLAB application will be introduced, as well as a binaural recording system that can assist such sound field simulations.

(A cikk folytatása)

4. Alkalmazási példák

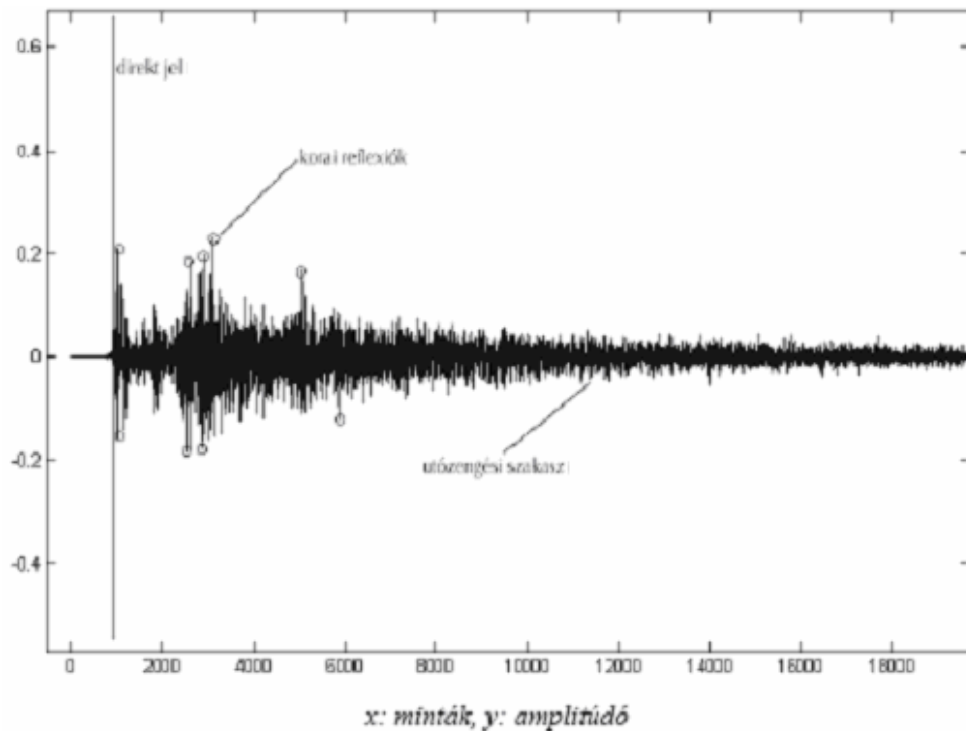
4.1 Teremakusztika

Az alábbiakban példát mutatunk a fentiek gyakorlati megvalósítására. A győri Széchenyi István Egyetem hallgatói [27-29] diplomamunka keretében foglalkoztak a témával. Rögzítettük a felújított, 300 fős, új D1-jelű előadó impulzusválaszait egy laptop és az Adobe Audition segítségével. Majd bevetve az Audition beépített konvolúcióját, sikeresen előállítottuk nyers hangfelvételek „térbeliségét”.

Hangforrásnak nagy termék esetén riasztópisztolyt használhatunk, de a jobb, olcsóbb és egyszerűbb megoldás a léggömb kidurranása. Ezt teremakusztikai mérésekhez gyakran megkérdőjelezzik, noha pld. a Brüel & Kjaer cég a mai napig a lufit javasolja ill. a saját maguk által forgalmazott pisztolyt utóhangsúlyozási idők mérésére. Tapasztalatunk szerint a léggömb egy viszonylag nagy méretű egyetemi előadóban is megfelelő hangenergiát biztosít. Minél nagyobbra fújuk, annál nagyobb a létrehozott hangnyomás. Hátránya ennek a módszernek, hogy nehezen reprodukálható és kis frekvenciákon nem garantált a megfelelő mennyiségű hangenergia kiadása. A másik probléma a mérésekkel, hogy a jelszintek beállítása nehézkes. Gyakran van túlvezérlés és torzítás az időfüggvényekben, máskor pedig túl halk és rossz a felvétel jel-zaj viszonya. A süketszobai teszteknel a lufi durranása a tőle 3 méterre lévő műfej hallójáratának bementén 105 dB (lineáris, impulzus üzemmód) hangnyomás szintet eredményezett. A beállítások jórészt próbálkozással zajlanak és gyakran becslésekre alkalmas a módszer.

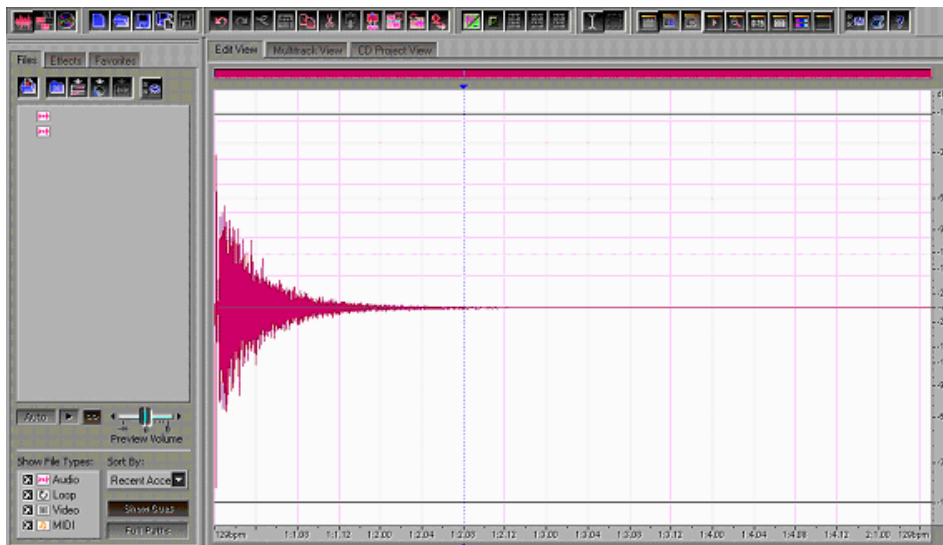
A mérés során az üres teremben elhelyeztük a Sennheiser BSN-ME62 jelű, gömbkarakterisztikájú mikrofont a terem impulzus átvitelének rögzítése céljából. Ezzel egyidőben a 4.2 szakaszban bemutatott binaurális rendszerrel is mértünk két élő emberen kétcsatornás binaurális teremátviteli függvényeket. Gerjesztő jelnek léggömböket és egy riasztópisztoly lövést is kipróbáltunk, utóbbi azonban használhatatlanul hangos volt és túlvezérlést okozott. Több mérést is elvégeztünk, különböző mikrofon- és ülőpozíciókkal, melyek közül a legjobban sikerült (legjobb dinamikájú és jel-zaj viszonyú, torzításmentes) felvételeket tartottuk meg. A terem így rögzített impulzusválaszát a fenti MATLAB-os program is sikerrel használta fel. A mérés során egy Brüel Kjaer 2260-as zajszint analízátorral az utóhangsúlyozási időt és a jelszinteket vizsgáltuk. A műszer tercsávós analízise során kijelzi, mely sávokban nem volt elég a hangenergia, ez alapján a mérések alsó határfrekvenciáját kb. 160-300 Hz között találhatók. A reprodukálhatóság a léggömbös méréseknél nehezen biztosítható, de az impulzusválaszokban „hallható” különbséget nem tapasztaltunk (az auralizáció után sem). Az utóhangsúlyozási idő tercsávós analízise az előbb említett határfrekvencia felett ms-os nagyságrendben mutatott csak eltéréseket, így a mérési módszer a különbségeket az aktuális alsó határfrekvencia változásában mutatta meg. Itt kompromisszumot kell találni a minél hangosabb (egyre jobban felfújott léggömb) gerjesztés és a túlvezérlés elkerülése érdekében.

A 5. ábra általánosságban mutatja (MATLAB-bal megjelenítve) egy tetszőlegesen kiválasztott impulzusválasz jellegzetes részleteit: a közvetlen hangutat, a körökkel jelölt, jól beazonosítható korai reflexiókat, valamint a sokszori visszaverődések által okozott diffúz hangtér (utóhangsúlyozás) kialakulását. A 6. ábra az egyetem D1-es termében felvett impulzusválaszt mutatja (Room Impulse Response: RIR), jelen esetben az Adobe Audition programmal megjelenítve.



5. ábra. Az impulzusválasz jellegzetes szakaszai egy teremben felvéve, MATLAB alatt ábrázolva.

Fig. 5. Typical impulse response of a room plotted with MATLAB.



6. ábra. A D1 terem impulzusválasza Adobe Audition-el ábrázolva.

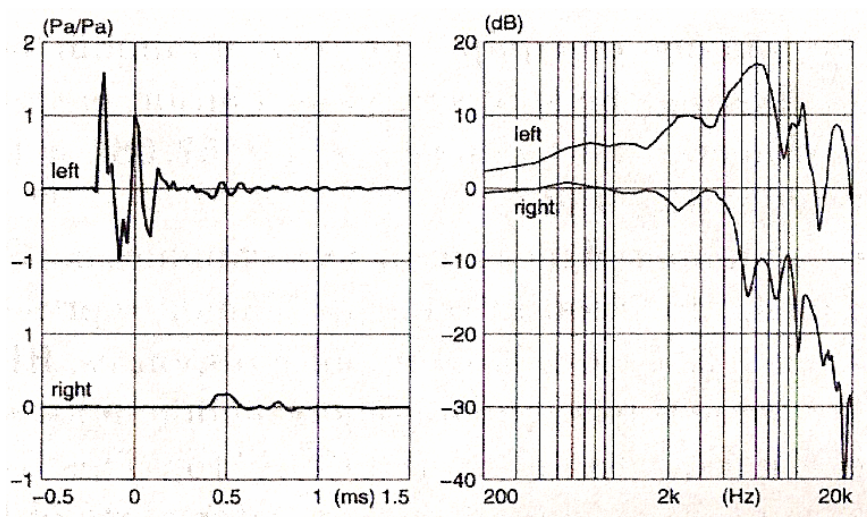
Fig. 6. The impulse response of the D1 lecture room at the university plotted with Adobe Audition.

Amire a mérésnél továbbá ügyelni kell, hogy a felvétel indítása nem fog egybe esni az impulzus kiadásával. Csak a legritkább esetben van a felvétel indítása szinkronban az impulzussal (amikor az impulzus triggereli a mérés indítást). Ez a gyakorlatban azt jelenti, hogy a felvétel előbb indul, és a végén nekünk kell a wave fájlt méretre vágni. Egyrészt levágjuk a felvétel végén (több másodperc után) lévő részeket, hisz azok már túl vannak a lecsengési időközön. Másrészt, a felvétel elejéről is le kell vágni a felesleges részt. Ha nem tesszük meg, a távolságok leképezése hamis lesz. A legjobb módszer, ha tisztában vagyunk a

hangforrás (a durranás) és a felvevő mikrofon távolságával: ekkor a hangsebességből kiszámíthatjuk mennyi idő telik el az első hullámfront beérkezéséig. Ez tipikusan nagyon kis időtartam, de a wave editorral akár mintapontossággal is levághatjuk a felesleges részeket.

4.2 Binaurális mérőrendszer

A térhallásunk alapja az emberi fülkagyló ún. külső fül átviteli függvényei, melyet az angol Head-Related Transfer Function elnevezésből magyarul is HRTF-knek nevezünk [30-34]. Ezek az iránytól függően szűrik a hangforrás hangját. Ha ezeket rögzítjük és ezekkel szűrjük a nyers felvételeket, irányszimulációt hajtunk végre. A lejátszáshoz szükséges fajhallgatót ki kell egyenlíteni [5]. Ez a feladat pontosan ugyanaz, mint a fentiek. Az átviteli rendszer ebben az esetben az emberi test, fej, fülkagyló kétsatornás hatása. E rendszert leírhatjuk az impulzusválasszal: egy adott hangforrásirányhoz tartozik egy bal fül és egy jobb fül impulzusválasz (7. ábra). Ezek rögzítése egyszerűbb és gyorsabb, mint a zajmérés. Ezek neve Head-Related Impulse Response (HRIR). A mérés jel-zaj viszonya azonban alacsonyabb lesz.



7. ábra. A külső fül rendszerleíró függvényei. Balra az impulzusválasz (HRIR), jobbra az abból számított átviteli függvény (HRTF) a bal és a jobb fülhöz, élő ember blokkolt hallójáratának bejáratán mérve „bal” irányból [35].

Fig. 7. The HRTFs and the HRIRs for the left and the right ear measured at the blocked ear canal entrance from the direction „left” [35].

Lejátszáskor, a szimuláció során, a bemenő időfüggvényt egyidőben a bal és a jobb fül impulzusválaszával is konvolváljuk, így állítva elő az irányinformációt. Elvben ez után szükséges még egy művelet: a fejhallgató átvitelének kiegyenlítése. Ez nem csupán a fejhallgatót, hanem az egész lejátszó lánc, az erősítők, hangkártya, kábelek stb. együttes átvitelét jelenti. A pontos hallásvizsgálatokhoz ez a lépés elengedhetetlen. A jó hír, hogy ez is ugyanaz a feladat, mint a korábbiak: egy átviteli lánc átvitelét kell meghatározni méréssel, majd számítással annak inverzét előállítani. Ezt a frekvenciatartományban egyszerűbb megvalósítani, reciprokképzéssel. Az így előálló kiegyenlítő fokozatot aztán beillesztjük a lejátszási láncba, méghozzá annak impulzusválaszával reprezentálva. Ezt előre előállíthatjuk, elmenthetjük és a konvolúciónak alávetethetjük. Érezhető, hogy a feladat adott esetben gyorsítható azzal, hogy ezt a két konvolúciót egybeépítjük. A gyakorlatban azonban ezeket egymás után kell elvégezni, hiszen amíg a kiegyenlítő fokozat átvitele állandó, addig a HRIR függvények a szimuláció során változhatnak.

A MATLAB-ban lehetőségünk van teljes szimulációra is, melynek során dinamikusan cserélgetve a HRIR függvényeket különböző forrásirányokat, mozgásokat szimulálhatunk. Jelenlegi kutatásunk egyik célja éppen ez: az emberi HRTF függvények behatóbb vizsgálata egy ilyen lejátszó rendszer segítségével [36-38].

A HRTF függvények mérése történhet műfejen vagy élő embereken. Előbbi „türelmesebb” és lehetőség van nagy jel-zaj viszonyú, hosszantartó zajgerjesztéses mérésekre is. A Békésy György Akusztikai Laborban a kilencvenes évek végén sok ilyen mérést végeztünk műfejjel [39-41]. Az emberek órákon keresztül nem képesek mozdulatlanul ülni, az ilyen mérések térbeli pontossága is kérdéses. Ezért itt csak impulzusválaszos HRIR mérések jönnek szóba. A további probléma, hogy két miniatűr, jó minőségű mikrofont kell elhelyeznünk a fülékben, melyek drágák és viselésük is kényelmetlen. Miért tesszük ezt mégis meg? Mert az így nyert HRTF készlet individuális lesz: a későbbi szimuláció során olyan érzésünk lesz, mint mindig: a saját fülünkkel halljuk a szimulációt is. Ez lényeges és döntő szempont, kutatások igazolták, hogy a lokalizációs feladatok megoldásában az individuális HRTF készlet használata a legjobb [42-44]. Ez műfej esetén nincs így, ott egy idegen, átlagos méretekkel rendelkező bábu fülein keresztül hallunk. Ennek eredménye pedig rosszabb felbontóképesség, nehezebb lokalizáció és romló eredmények a feladatok során.

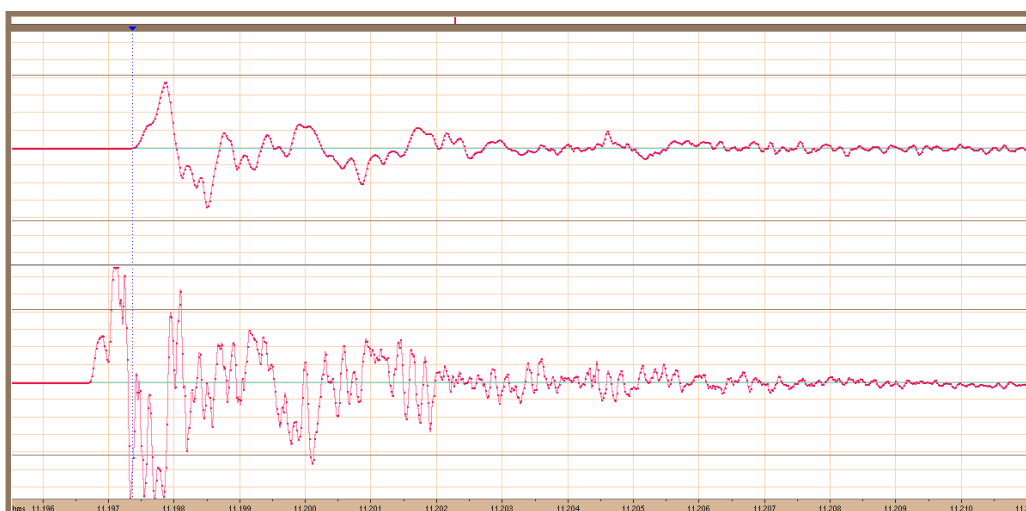
A győri egyetem rendelkezik egy kiépített, telepített ún. binaurális felvevő rendszerrel is, mellyel ezek a felvételek, mérések elvégezhetők. Ez áll egy közönséges laptopból és az azon futó Audition szoftverből. Nem a saját hangkártyáját használjuk, hanem a Digigram VXPocket PCMCi-foglalatba illeszthető kártyát a jó minőség érdekében. A mikrofon a Brüel Kjaer 4101 jelű binaurális mikrofon, mely egy kis illesztő elektronikán át kap tápfeszültséget és csatlakozik a hangkártyához (8. ábra) [45].



8. ábra. A BK 4101 binaurális mikrofon, a tápillesztő áramkör és a PCMCi hangkártya. A mikrofon a blokkolt hallójárat bemenetén helyezkedik el, cserélhető szivacsokkal.

Fig. 8. The BK 4101 binaural microphone, the power supply and the PCMCi sound card for the laptop to record individual HRTFs or for binaural recordings.

A rendszerrel próbaméréseket végeztünk a süketszobában, felvéve két ember néhány individuális HRTF függvényét, valamint a 4128-as műfej saját beépített és a külsőleg elhelyezett mikrofonnal történő HRTF eredményeit vizsgáltuk. Gerjesztésnek – tanulva a pisztoly okozta túlvezérlésből – ismét léggömböket választottunk. A hangforrás távolsága 2 méter volt, a mérendő személy ill. a műfej egy széken foglalt helyet. Analizáláshoz, editáláshoz ismét az Adobe Audition-t használtuk. Nem okozott meglepetést, hogy ezek hasonlítanak egymásra, ugyanakkor nem sikerült kétszer sem azonos mérési eredményeket produkálni. A 9. ábra élő emberi fejen mért impulzusválaszokat mutat a bal és a jobb fül esetén. A hangforrás 90 fokban, azaz az egyik fül tengelyében található. Jól látszik az amplitúdó- és futásidő különbség a két csatorna között. Ez az idő kb. megfelel a fej fél kerületének megfelelő távolságnak (30-40 cm).



9. ábra. Binaurális mikrofonnal történő impulzusválasz mérés kinagyított eredménye emberi fejen a bal ill. a jobb fül esetén, ha a hangforrás 90 fokos irányban helyezkedik el a horizontális síkban.

Fig. 9. Zoomed figures of impulse responses measured on a real human head (left and right ear). Sound source direction is 90 degrees.

Az 1. táblázat összehasonlítja a műfejes ill. élő emberi fejes mérések, felvételek tulajdonságait.

	Műfej	Élő ember
Mérőjel	Zaj, impulzusválasz, sweep	Impulzusválasz
Kényelmi szempontok	Nincsenek	Vannak
A mérés jel/zaj viszonya	Magas (átlagolással és zajgerjesztéssel)	Általában alacsonyabb
Átlagolás	Lehetséges (zajmérésnél)	Nincs
Pontosság, reprodukálhatóság	Nagy, egyszerű	Kicsi, nehézkes
Individuális HRTF mérés, binaurális felvételek	Nincs	Van
HRTF készlet térbeli felbontása	Nagy (1 fokos)	Általában kicsi (15 fokos vagy rosszabb)
Megítélés	objektív	szubjektív

1. táblázat. Élő emberi és műfejes technika összehasonlítása.

Table 1. Comparison of dummy-head and real human head measurement techniques.

Végezetül az összehasonlítás céljából a D1-es előadóban is végeztünk binaurális impulzusválasz méréseket. Ezeket a függvényeket BRIR-nek hívják (Binaural Room Impulse Response). A 10. ábra mutat mérési eredményeket. Felhasználva a BRIR függvényeket a fenti programokkal (fejhallgató lejátszás során) képesek vagyunk binaurális élmény nyújtására. Az auralizáció során a MATLAB rutin által előállított kimeneti wave fájlokat fejhallgatón át meghallgatva győződünk meg a rendszer használhatóságáról. Első lépésként tehát sikerrel jártunk, pontos lokalizációs kísérletek azonban még hátravannak, így a fejhallgató kiegyenlítés moduljának implementálása. A MATLAB-os alkalmazás népszerűsége és előnye itt is megmutatkozik: a fő program tetszőlegesen bővíthető, átprogramozható, és kiegészíthető pld. esetünkben egy fejhallgató átviteli függvény mérő- és kiegyenlítő modullal, mely szolgáltatást a szokványos konvolváló programok nem biztosítják.

Végezetül jegyezzük meg, hogy a binaurális mikrofon nem csupán individuális HRIR és BRIR felvételekre alkalmas, hanem olyan zajmérésekre és felvételekre, ahol a műfej nem vethető be. Ilyen eset, ahol a tesztalany aktív beavatkozása szükséges a mérés során, pld. autóvezetés.



10. ábra. A D1-es terem impulzusválaszának spektruma (átviteli karakterisztikája) gömb karakterisztikájú mikrofonnal (fent) ill. binaurális felvétel során élő emberen (lent). A fülkagyló 3-4 kHz közötti megszokott, HRTF-kben is jelentkező kiemelése szembevetű. Fig. 10. Spectral representation of the response of the D1 lecture room measured with omnidirectional microphone (above) and using the binaural microphone on a real human head (below). The 3-4 kHz notch of the pinna is visible.

5. Összefoglalás

A digitális jelfeldolgozás és a számítástechnika fejlődése lehetővé tette, hogy gyorsan, egyszerűen, gyakran valós-időben leprogramozzunk olyan feladatokat, melyek korábban nem

voltak lehetségesek. Ennek tipikus példája az átviteli rendszerek vizsgálata, esetünkben egy terem vagy éppen az emberi fej, netán egy fejhallgató inverz működését kell megmérni és szimuláció során wave fájlok manipulációjára felhasználni. Korábban FFT algoritmusok és a frekvenciatartománybeli komplex szorzás állt rendelkezésre, manapság az időtartománybeli konvolúciót több program, célszoftver vagy éppen felhasználóbarát fejlesztői környezet támogatja. Bemutattuk, hogy drágább ill. olcsóbb szoftvermegoldások léteznek, valamint példát láttunk arra, miként lehet MATLAB környezetben is megoldást találni a feladatra. Utóbbi a felhasználóbarát programozói környezettel és a feladat megoldásának egyénre szabásával tűnik ki a többi megoldás közül, gyors alkalmazáshoz azonban a beépített fft és konvolúció használható, a valós idejű alkalmazás kérdéses, az időtartományban pedig kivitelezhetetlen. A MATLAB offline ill. gyors számításokban, wave-fájlok kezelésében és a szokványos funkciók mellett elérhető matematikai apparátusával segíti a munkát, amelyeket más programcsomagok nem biztosítanak. Az alkalmazási területek része a teremakusztikai szimuláció, az impulzusválasz mérések és felhasználás változatai, vagy éppen virtuális valóság szimulátorok bevetése a hallásvizsgálatokban.

Irodalom (folytatás)

- [26] <http://www.ludd.luth.se/~torger/brutefir.html>.
- [27] A. Petrovszki, *Konvolúció és teremszimuláció MATLAB program segítségével*. SZE Távközlési Tanszék, diplomaterv dolgozat, 2008. <http://vip.tilb.sze.hu/~wersenyi/PA.pdf>
- [28] Z. Györi, *Binaurális mérőrendszer kiépítése*. SZE Távközlési Tanszék, diplomaterv dolgozat, 2008. <http://vip.tilb.sze.hu/~wersenyi/GYZ.pdf>
- [29] A. Nyikos, *Teremakusztikai és binaurális szimuláció akusztikai programokkal*. SZE Távközlési Tanszék, diplomaterv dolgozat, 2007. <http://vip.tilb.sze.hu/~wersenyi/NYA.pdf>
- [30] Gy. Wersényi, *HRTFs in Human Localization: Measurement, Spectral Evaluation and Practical Use in Virtual Audio Environment*. Ph.D. doctoral thesis, BTU Cottbus, 2002.
- [31] C. B. Jensen, M. F. Sorensen, D. Hammershøi, H. Møller, *Head-Related Transfer Functions: Measurements on 40 human subjects*. Proc. of 6th Int. FASE Conference, Zürich, 1992, pp. 225-228.
- [32] H. Møller, M. F. Sorensen, D. Hammershøi, C. B. Jensen, *Head-Related Transfer Functions of human subjects*. J. Audio Eng. Soc., vol. 43(5), 1995, pp. 300-321.
- [33] V. Mellert, K. F. Siebrasse, S. Mehrgardt, *Determination of the transfer function of the external ear by an impulse response measurement*. J. Acoust. Soc. Am., vol. 56, 1974, pp. 1913-1915.
- [34] J. Kawaura, Y. Suzuki, F. Asano, T. Sone, *Sound localization in headphone reproduction by simulating transfer functions from the sound source to the external ear*. J. Acoust.Soc.Japan, E 12, 1991, pp. 203-215.
- [35] D. Hammershoi, H. Moller, *Binaural Technique – Basic Methods for Recording, Synthesis, and Reproduction*. In J. Blauert (Editor) Communication Acoustics, Springer verlag, 2005, pp 223-254.
- [36] Gy. Wersényi, A. Illényi, *Test Signal Generation and Accuracy of Turntable Control in a Dummy-Head Measurement System*. Journal of the Audio Engineering Society, vol.51 (3), March 2003. pp. 150-155.
- [37] Gy. Wersényi, A. Illényi, *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. pp 1-15.
- [38] A. Illényi, Gy. Wersényi, *Environmental Influence on the fine Structure of Dummy-head HRTFs*. Proceedings of the Forum Acusticum 2005, Budapest. pp. 2529-2534.

- [39] Gy. Wersényi, *On the amplification-damping-effect of the pinnae and of the head*. Akusztikai Szemle, vol.VI., Nr.2., 2005, Budapest. pp 23-28.
- [40] Gy. Wersényi, *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. pp. 1-16.
- [41] Gy. Wersényi, *Directional Properties of the Dummy-Head in Measurement Techniques based on Binaural Evaluation*. Journal of Engineering, Computing & Architecture, Vol.1, Issue 2, 2007. www.scientificjournals.org. pp. 1-15.
- [42] H. Møller, M. F. Sorensen, C. B. Jensen, D. Hammershøi, *Binaural Technique: Do We Need Individual Recordings?* J. Audio Eng. Soc., vol. 44(6), 1996, pp. 451-469.
- [43] P. Majjala, *Better binaural recordings using the real human head*. Proceedings of Inter-Noise 97, Budapest, 1997, pp. 1135-1138.
- [44] P. Minnaar, S. K. Olesen, F. Christensen, H. Møller, *Localization with Binaural Recordings from Artificial and Human Heads*. J. Audio Eng. Soc., vol 49(5), 2001, pp. 323-336.
- [45] <http://www.bksv.com/doc/bp1992.pdf>