

# Virtuális akusztikai valóság és auralizáció

HUSZTY CSABA<sup>1</sup>, AUGUSZTINOVICZ FÜLÖP<sup>2</sup>

Rezgésakusztikai és Hangtechnikai Laboratórium, Híradástechnikai Tanszék, Budapesti Műszaki és Gazdaságtudományi Egyetem

<sup>1</sup> [huszty@hit.bme.hu](mailto:huszty@hit.bme.hu)

<sup>2</sup> [fulop@hit.bme.hu](mailto:fulop@hit.bme.hu)

**Kivonat.** Az előadás a virtuális technológiák egy új változatát mutatja be, amely alkalmas arra, hogy a hallgatói környezetben – akár valós időben is – előállítsunk egy olyan hangteret, amelyben a hallgató úgy érzi, mintha egy másik teremben tartózkodna. A konvencionális zengetőktől eltérően ez a technológia valódi termék mérési eredményein alapul, lehetővé teszi a hangmérnök számára a kreatív beavatkozást és támogat minden jelenlegi sokcsatornás hallgatói elrendezést. Alkalmazási lehetőségei az elérhető árú nagy számítási kapacitású processzorok megjelenésével és elterjedésével szélesednek ki igazán, mert így a technológiát használó eszközök a hangstúdiókból és az ipari felhasználóktól az otthoni felhasználókhoz is eljuthatnak.

**Kulcsszavak:** teremakusztikai mérések, konvolúciós zengetés, auralizáció, impulzusválasz, B-format

## 1 Bevezetés

A hangtér reprodukciójának célja már az 1950-es években felmerült. Azóta számos rendszer terjedt el, amely – egyebek mellett az emberi hallás sajátosságait kihasználva – képes a hallgató számára elfogadható, természetű hangzást produkálni. Ezen rendszerek áttekintése és rendszerezése után bemutatunk egy rendszert, amely valós idejű konvolúciót használva utólag hoz létre valódi helyszíneken mért akusztikai környezetet a hallgató számára, így lehetővé téve például azt, hogy egy hangszeret utólag, virtuálisan a térbe helyezzünk. Ennek a technikának számos gyakorlati alkalmazási lehetősége van a hangmérnökök számára kifejlesztett zengetőn keresztül a házimozsi-rendszereken át az autópárig, az alkalmazások pedig csak mostanában kezdenek elterjedni a rendszer jelentős számítási igénye miatt.

## 2 Surround rendszerek

### 2.1 Tér- és irányérzékelés<sup>1</sup>

Az emberi hangérzékelés két hallószervvel működik, amelynek két fontos következménye van. Egyrészt

- a hang által a hallószervekig megtett út eltérő, feltéve, hogy valamelyik fülünkhöz közelebb van a hangforrás, másrészt
- a távolabbi fülünkhöz érkező hang előtt a fejünk akadályt képez, takarja a hangforrást.

E két hatás eredménye a hangra nézve egyrészt időbeli különbség (Interaural Time Difference, ITD), másrészt szintbeli (amplitúdóbeli) különbség (Interaural Level Difference, ILD) lesz. Az időkülönbséget periodikus jeleknél fáziskülönbségnek is értelmezhetjük. A

---

<sup>1</sup> [24],[15]

fülünk abszolút fázist egyáltalán nem érzékel, fáziseltérést azonban igen. A fáziseltérésből adódó irányérzékelés a nagyobb frekvenciák felé egyre bizonytalanabb.

Ha feltesszük, hogy a hangsebesség 342 m/s, az emberi fej átlagos átmérője 18 cm, a fülek pedig 90 illetve 270 foknál helyezkednek el egy gömb alakú fejmodellen, akkor a fülek távolsága (a kerület fele)  $d = \pi \cdot r = 0,2827$  m, ebből pedig a legnagyobb lehetséges időbeli eltérés a fenti hangsebességgel számolva 0,83 ms.

A szint eltéréseket (ILD) nem a távolságból adódó csillapítás befolyásolja elsősorban, hanem a fej takaró hatása. Az ILD frekvenciafüggő mennyiség: ha a hang hullámhossza nagyobb, mint az akadály - a fej - átmérője, a diffrakció jelensége lép fel, ha pedig kisebb, akkor aluláteresztő szűrő hatással lehet számolni. A két jelenség határfrekvenciája a fej átmérőjével számolva 1,89 kHz, azonban e két hatás egyszerre lép fel folytonos átmenettel, és szögfüggő is: a csillapítás egyre meredekebben nő a növekvő frekvenciák felé a beérkezési szög függvényében.

A fentiekén kívül a fülkagyló, a haj és a fej más egyedi tulajdonságai okozta frekvenciafüggő szűrő hatást is figyelembe kell venni.

Összefoglalva tehát a hangforrással kapcsolatban a következő tényezők befolyásolják az érzékelést:

- Szintkülönbség (ILD)
- Fázis és időkülönbség (ITD)
- Fej, fülkagyló, test irányfüggő szűrőhatása (HRTF<sup>2</sup>)

A előre és hátra irányokból érkező hangoknál nem jelentkezik ILD és ITD, így itt csak a fej, a test és a külső fül (fülkagyló) szög- és frekvenciafüggő szűrőhatása jelenthet támpontot az irányérzékeléshez. Az ebből adódó amplitúdó-eltérések akár 30 dB nagyságrendűek is lehetnek. Ezen hatások, valamint a fej mozgása, illetve az egyéb érzékszerveinkből érkező információk együttesen alakítják ki a hangforrásról alkotott érzetünket, mint például az irányérzékelést, a forrás kiterjedtségére vonatkozó ismereteinket, vagy a hangforrás távolságát.

Képesek vagyunk az akusztikai környezet érzékelésére, és ez alapján a fizikai környezetre való következtetésre is. Az akusztikai környezetben többutas terjedés lép fel, az objektumokról való visszaverődések sokasága révén szerzünk tudomást a környezetről. A hallásunk a közvetlen, legkorábban és legnagyobb szinten érkező hangutat részesíti előnyben. Amennyiben a visszaverődések kb. 35 ms nagyságrendjén belül érkeznek, nem hallunk visszaverődést (Haas-hatás), ezek tehát a lokalizációt nem befolyásolják. A távolságra a forrásra vonatkozó információkról és az érzékelt hangszínből tudunk következtetni, mert a távolsággal egyre növekvő mértékű aluláteresztő hatás jelentkezik, valamint zengő térben a direkt-zengő energiaarányból, hiszen minél közelebb van hozzánk a forrás, annál nagyobb arányú a direkt-hang energiája a zengőéhez képest.

## 2.2 Surround rendszerek

A surround rendszerek kialakításának legfőbb célja az volt, hogy a hallgató körül olyan hangteret hozzanak létre, amely számára elfogadható, élethű élmény ad. A surround rendszerek egyaránt szolgálhatnak önálló rendszerként vagy kiegészítő rendszerként vizuális tartalomhoz (pl. film). A hangtér előállítása úgy lehetséges legjobban, ha sikerül előállítani az eredeti forrásból kiinduló hullámfrontot és fenntartani a gerjesztett állapotot akként, ahogy az az eredeti környezetben megvalósult. Ha sikerülne mindez, az ún. sweet spot – azaz az optimális térrész: az a hely, ahol a kívánt hatás az általunk kívánt módon következik be, vagyis a hangtér eredeti formájában áll elő a hallgatói környezetben – tetszőlegesen nagy lehetne,

---

<sup>2</sup> Head-Related Transfer Function vagy Anatomical Transfer Function

azonban a gyakorlatban a hangsugárzók száma, valamint a rendelkezésre álló jelfeldolgozási kapacitás szűkössége okán a sweet spot mérete korlátozott<sup>3</sup>.

A surround reprodukciós rendszerek fogalmának értelmezésekor a következőkre gondolunk:

- felvételi technikák
  - mikrofonrendszerek és -elrendezések
- hangsugárzó-elrendezések
  - hangsugárzók száma, helye, fajtája (fejhallgató, hangsugárzó-rendszer, stb.)
- tárolási módszerek
  - adathordozók, kódolások (Dolby Digital AC-3, DTS, stb.)

A jelenkorban használt surround rendszereket pedig a következőképpen csoportosíthatjuk [24]:

- csak a hangsugárzórendszer elrendezését vagy a hordozó médiumot és formátumot meghatározó rendszerek
  - ITU 5.1, 7.1, stb.
  - Dolby Digital, DTS
- csak a rögzítés (felvétel) és visszajátszás (panorámázás) módját meghatározó rendszerek
  - Wavefield szintézis
  - Ambisonics

Az előbbi rendszerek során a rögzítés és visszajátszás módját és eszközeit nem határozzuk meg, kizárólag a hangsugárzó-rendszer elrendezését vesszük figyelembe, míg az utóbbi esetben nincs definiálva az alkalmazandó hangsugárzó-rendszer.

## 2.3 A hangsugárzórendszer elrendezését meghatározó surround rendszerek

### 2.3.1 Fejhallgató

Bár a fejhallgatóval történő visszahallgatás nem kötődik szorosan e munka témaköréhez, mégis, a teljesség érdekében hadd említsük meg röviden a rendszer sajátosságait!

Amikor fejhallgatóval hallgatunk vissza egy felvételt, két csatorna áll rendelkezésünkre, amelyek közvetlenül a fülünkbe jutnak. Áthallás nem lép fel, ezért a felvételen levő fáziskülönbségek jól érvényesülhetnek, és így az ITD alapú lokalizáció jól működik [15]. Sőt, az amplitúdó különbségek - amennyiben a hangfelvételt hangsugárzóra optimalizálták -, még akár eltűnő is felléphetnek, ami különösen erős oldalra lokalizálódást eredményezhet. A fejhallgatóból ezenkívül közvetlenül jut a fülbe a hang, ezért az irányfüggő szűrő hatás

---

<sup>3</sup> Erre az egyik első példa az 1950-es években a Bell Labs által kifejlesztett kísérleti rendszer, ami a ma wavefield szintézis néven ismert technikához hasonló alapokra épült. Az volt a céljuk, hogy a hallgató pozíciójától függetlenül (nagy kiterjedésű sweet spot) reprodukálhassanak egy hangforrásból kiinduló hullámfrontot. Egy-máshoz nagyon közel elhelyezett mikrofonokat használtak a hangfelvételhez, és szintén közeli hangsugárzókat a reprodukcióhoz, majd azzal kísérleteztek, hogy miként lehet csökkenteni az elemszámot annak érdekében, hogy könnyen használható rendszert kapjanak. Ahogy csökkentették a hangsugárzók számát, a hullámfront úgy lett egyre pontatlanabb (térbeli alulmintavételezés). Valamivel korábban, 1931-ben nyújtott be Blumlein szabadalmi kérelmet (Binaural Reproduction), amelyben két, a fülek távolságának megfelelően elhelyezett gömbi mikrofont használt, amelyek közé elnyelő anyagot tett. Felvételi technikája jó eredményeket adott fejhallgatók meghallgatás esetén, azonban a hangsugárzóknál az áthallás (crosstalk) miatt a rögzített fáziseltérések nem reprodukálódtak a hallgató fülénél. Rájött, hogy kisfrekvencián szintkülönbségeket is produkálni kell annak érdekében, hogy a lokalizáció létrejöhessen. Később, amikor megjelentek a nyolcas amplitúdó-iránykarakterisztikájú szalagmikrofonok, a Blumlein-féle sztereó technika már sokkal jobb eredményeket adott. Ez vezetett oda, hogy koincidens mikrofontechnikát alkalmazzon, ahogyan azt ma is ismerjük.

(HRTF) nem érvényesülhet. A később ismertetendő Ambisonics rendszerrel azonban lehetőség van arra, hogy a hallgató körül virtuális hangszugárzó-elrendezést hozzunk létre a hangszugárzók irányába mutató HRTF függvényekkel, sőt, a fejmozgás is követhetővé válik, ha elegendő jelfeldolgozási kapacitás áll rendelkezésre. (Ebben az esetben arra lenne szükség, hogy a pillanatnyi háromdimenziós orientációnak megfelelő HRTF szűrőt valós időben kikeressük és alkalmazzuk a hangjelen.)

A fejhallgatóéhoz igen hasonló hatást lehet elérni hangszugárzók alkalmazásával is, az ún. sztereó dipól elrendezés segítségével [15], de ezzel a továbbiakban itt nem foglalkozunk.

### **2.3.2 Hangszugárzórendszerek**

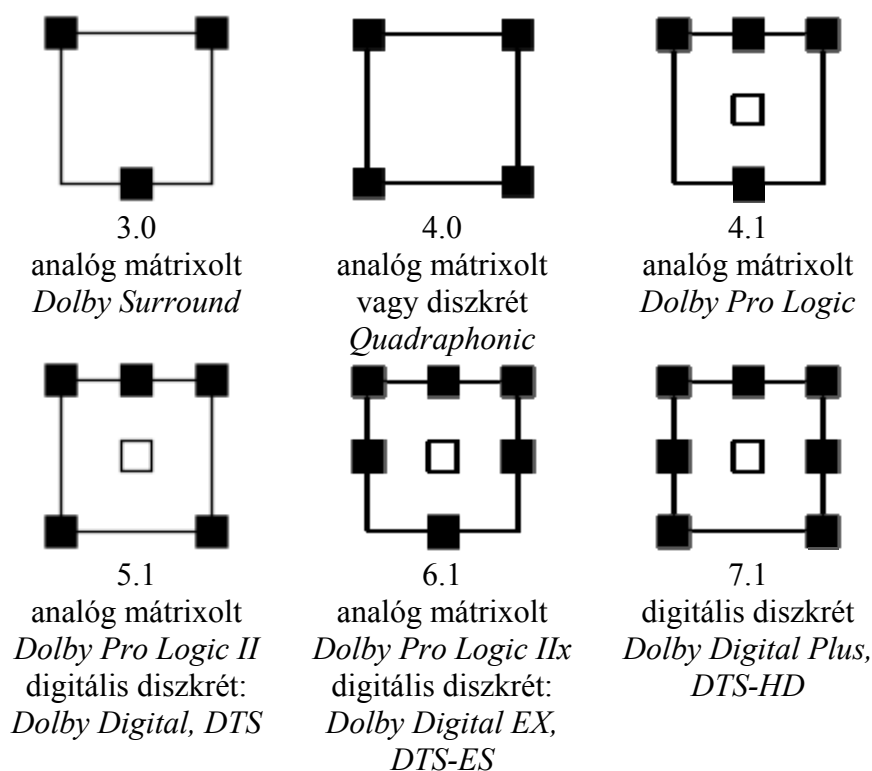
A hangszugárzórendszer elrendezését meghatározó rendszerek többnyire egy korlátozott méretű területet vesznek körül néhány hangszugárzóval, többnyire síkban. A filmszínházakban térbeli elrendezést is alkalmaznak (hiszen sokszor a hallgatók sem a vászonra merőleges síkban ülnek), illetve újabban megjelentek a nem síkbeli elrendezésű kísérleti rendszerek is (pl. 22.2 rendszer), azonban ezek elterjedtsége egyelőre csekély mértékű. A síkbeli elrendezést kör, vagy más alakzat mentén képzeljük el. Ezek helyes késleltetési értékek és szintek beállításával a lehallgatói környezet helyfüggő teremakusztikai viszonyaitól eltekintve ekvivalensnek tekinthetők. A hangszugárzórendszerek és a fejhallgató alkalmazása között két jelentős különbség van.

- Egyrészt áthallás (crosstalk) lép fel a hangszugárzók között, mivel a hangutak mindkét fülünkbe eljutnak
- Másrészt a hangszugárzók által keltett hangot a lehallgató helyiség saját teremakusztikai adottságai is befolyásolják (pl. hangstúdió, filmszínház, vagy szoba).

Az utóbbira a gondos teremakusztikai tervezés és a helyesen beállított hangszugárzók (frekvenciamenet, hangszugárzott egymáshoz viszonyított szintjei, és késleltetései, stb.) együtt adhatnak megoldást.

Ilyen rendszerek alkalmazásakor döntően az amplitúdó-különbségek adnak támpontot az irányérzékeléshez, de kisfrekvencián itt is inkább a fáziskülönbség (ITD) dominál [24],[15]. A hallgató a saját füléhez és fejéhez tartozó HRTF szűrés segítségével jól tudja lokalizálni az előre-hátra irányokat, azonban ha a hangszugárzók közé eső forrásirányt tartalmaz a felvétel, a lokalizáció pontatlanabbá válhat.

A következő ábrákon a szokványos hangszugárzó-elrendezések láthatók megjelölve azt is, hogy milyen technológia tartozik a hangszugárzók bemeneteinek előállításához. A hangszugárzó-elrendezést úgy szokás megjelölni, hogy a szélessávú sugárzók darabszáma után helyezett pontot a keskenysávú kisfrekvenciás kiegészítő csatorna darabszáma követi (pl. 5.1, 10.2, stb.).



1. ábra – Szokványos hangszugárzó-elrendezések

Forrás:[25]

Megjegyezzük, hogy 10.2 és legújabban 22.2 rendszerek is léteznek, de még nem terjedtek el.

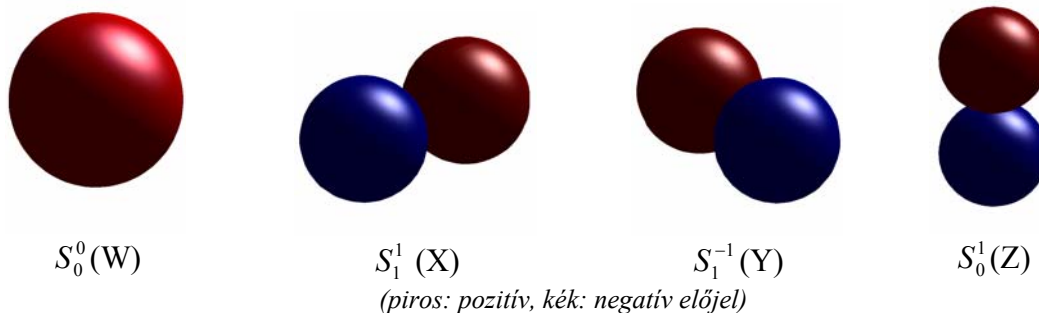
## 2.4 A hangszugárzórendszer elrendezésétől független surround rendszerek

### 2.4.1 Ambisonics

Az Ambisonics [4] koincidens rendszert Michael Gerzon fejlesztette ki 1974-ben [6], módszer elméleti alapja a hangtér gömbfüggvényekkel történő dekompozíciója. E függvények közül a 0. rendű gömbfüggvény által leírt iránykarakterisztika a gömbi (nyomásmikrofon) az 1. rendű gömbfüggvények pedig különböző főirányú nyolcas karakterisztikákat írnak le (nyomásgradiens mikrofon) [15]:

$$\begin{aligned}
 S_0^0 &= \frac{1}{2} \sqrt{\frac{1}{\pi}}; \\
 S_1^0 &= \frac{1}{2} \sqrt{\frac{3}{\pi}} \cos \vartheta; \\
 S_1^1 &= \frac{1}{2} \sqrt{\frac{3}{\pi}} \sin \vartheta \cdot \cos \varphi; \\
 S_1^{-1} &= \frac{1}{2} \sqrt{\frac{3}{\pi}} \sin \vartheta \cdot \sin \varphi;
 \end{aligned}
 \tag{1}$$

A gömbi és a három nyolcas karakterisztikájú jelet megfelelő sorrendben W,X,Y,Z jelöléssel illetik, együttesen pedig B-formát (B-formátumú) jeleknek nevezik. A fenti négy függvény abszolútértéke gömbi koordináta-rendszerben ábrázolva a következő ábrán látható.



2. ábra – B-formátumú csatornák iránykarakterisztikája

Az elsőrendű Ambisonics a 0. és 1. rendű gömbfüggvényeknek megfelelő iránykarakterisztikákat használja, azaz 4 csatornán a következők szerint rögzít:

| csatorna | karakterisztika típus | irány        | gömbfüggvény |
|----------|-----------------------|--------------|--------------|
| W        | nyomás                | -            | $S_0^0$      |
| X        | nyomásgradiens        | előre-hátra  | $S_1^1$      |
| Y        | nyomásgradiens        | balra-jobbra | $S_1^{-1}$   |
| Z        | nyomásgradiens        | fel-le       | $S_1^0$      |

1. táblázat – a B-format és a gömbfüggvények kapcsolata

A B-formátumban rögzített jelek lehetőséget kínálnak a hangtér virtuális reprodukciójára. A rögzített 4 csatornából történő dekódolásra, más szóval a hangszugárzó jeleinek előállítására többféle eljárás is létezik. A gömbi és a három nyolcas karakterisztika – mintha térbeli MS-stereo elrendezés lenne – megfelelő súlyozásával új, tetszőleges virtuális mikrofon-iránykarakterisztika hozható létre. A tetszőleges főirányú virtuális mikrofon jele az alábbi formulával állítható elő [2], [15], [19]:

$$V(\varphi, \vartheta) = \frac{1}{2} \left[ (2-D)\sqrt{2} \cdot W + D \cdot (\cos(\varphi) \cos(\vartheta) X + \sin(\varphi) \cos(\vartheta) Y + \sin(\vartheta) Z) \right] \quad (2)$$

ahol  $D$  a virtuális mikrofon irányítási tényezője. (gömbi: 0; szub-kardiod: 0,5; kardiod: 1; hiperkardiod: 1,5; nyolcas: 2).  $D$  növelésével javítható az irányfelbontás, ám ekkor az ellentétes oldalon ellenfázisú jel jelenik meg, ami csökkenti a *sweet spot* méretét.  $\frac{D}{2-D} = 1$  esetben még nincs ellenfázisú jel (ekkor a virtuális mikrofonjaink kardiod karakterisztikájúak). Amennyiben olyan hangszugárzó-elrendezést használunk, amely nem egy kör mentén helyezkedik el, a  $2-D$  és  $D$  helyén szereplő mennyiségeket a távolság szerint is súlyozni kell, illetve az elrendezésből adódó addicionális késleltetések kompenzálásáról is gondoskodni kell. Összetettebb dekódolási eljárások a már említett konstansok frekvenciafüggését is figyelembe veszik.

#### 2.4.2 Hullámtér (Wavefield) szintézis

Az egyik legpontosabb és legnagyobb optimális térrészt (sweep spot-ot) előállító eljárást hullámtérszintézisnek (Wavefield Synthesis, WFS) nevezik. A WFS egymáshoz közel elhelyezett hangszugárzókat alkalmaz, amelyek a Huygens-elv alapján állítják vissza a hallgató körül az eredeti (vagy a modellezett) hangteret. A módszer érvényességi határát a térbeli mintavételezés állította határfeltétel szabja meg: a hangszugárzó vagy hangfalak közti távolság felének megfelelő hullámhossz alatt már nem működik a rekonstrukció (pl. 1 kHz-es műkö-

dési határfrekvenciához 17 cm-es távolság szükséges.) A gyakorlatban nem feltétlenül szükséges a mintavételezési korlátot ilyen szigorúan betartani, mert az egyéb lokalizációt segítő hatások révén a matematikailag nem teljesen korrekt rendszer is kielégítő eredményeket szolgáltat [15], [24]. A módszerrel a virtuális források elhelyezkedése a hallgató pozíciójától nem függ, ám jelentős a jelfeldolgozási (nagy csatornaszám) és a hangszóró igény, ezért széleskörű elterjedtségéről még nem beszélhetünk.

## 2.5 Felvételi technikák

Az itt említett surround rendszereknek szinte mind eltérő felvételi technikákat igényelnek. Azoknál a rendszereknél, ahol a hangszóró-rendszert definiálják és a felvételi vagy visszajátszási technikát nem, különféle felvételi technikák alakultak ki, amelyek közül a hangmérnök választja ki a megfelelőt szubjektív értékítélete, vagy más szempontok alapján. A felvételi technikák a következő megközelítésben is értelmezhetők:

- egy pontban kívánjuk rögzíteni a hangteret (koincidens), vagy
- több pontban mintavételezzük (nem coincidens), esetleg
- a kettő közti átmenetet képző technikát használunk, például a fülek távolsága, vagy valamilyen más paraméter szerint úgy, hogy e pontok viszonylag közel helyezkedjenek el egymáshoz (közel-koincidens).

A mikrofonok térbeli szűrőként viselkednek, a szűrést pedig az iránykarakteristikával adjuk meg, ami komplex, 3 térdimenzió mentén szög- és frekvenciafüggő mennyiség. A gyakorlatban az amplitúdó iránykarakterisztika síkmetszetét használjuk polárdiagramon ábrázolva, kihasználva, hogy az iránykarakterisztika abszolútértéke nagyfokú szimmetriát mutat, így elegendő a síkmetszettel történő reprezentáció. (A fázis-iránykarakteristikát nem szoktuk figyelembe venni.) Az amplitúdó iránykarakteristikája megadja a mikrofon érzékenységének irányfüggését, amit általában relatív értékekkel, a legérzékenyebb irányhoz képest (főirány) viszonyítunk.

A teljesség igénye nélkül megemlítenek néhány a fenti rendszerbe sorolt mikrofonozási technikát [15].

| Koincidens | Közel-koincidens (átmeneti) | Nem coincidens  |
|------------|-----------------------------|-----------------|
| X/Y        | ORTF                        | AB              |
| M/S        | OCT                         | Hamasaki-square |
| Ambisonics | műfej                       |                 |

2. táblázat – Mikrofonozási technikák

Az Ambisonics technikához használhatunk három nyolcas és egy gömbi mikrofont egy pontba helyezve, azonban a valóságban bizonyos távolság mindenképpen lesz a kapszulák között, ami felső határfrekvenciát ad a mérés matematikai pontosságára.

Az ún. SoundField mikrofon közvetett módon B-formátumú jelet rögzít 4 szubkardioid kapszulával, amelyek egy tetraédernek megfelelően vannak elhelyezve. Az így rögzített jelet A-formátumúnak nevezik.

| Kapszula | Oldalszög | Emelkedési szög |
|----------|-----------|-----------------|
| A        | 45°       | 35,3°           |
| B        | 135°      | -35,3°          |
| C        | -45°      | -35,3°          |
| D        | 135°      | 35,3°           |

3. táblázat – a SoundField mikrofon kapszuláinak elhelyezkedése

Az A-formátumú négycsatornás jelből a következő egyenletek segítségével lehet előállítani a B-formátumot (feltéve, hogy az A,B,C,D kapszulák kardioid karakterisztikájúak).

$$\begin{aligned} W &= \frac{1}{2}(A + B + C + D) \\ X &= (A + C) - (B + D) \\ Y &= (A + B) - (C + D) \\ Z &= (A + D) - (B + C) \end{aligned} \quad (3)$$



3. ábra – A SoundField mikrofonkapszula

Forrás: <http://www.soundfieldusa.com>

Az elhelyezett kapszulák azonban nem egy pontban vannak, így némi fáziseltérés is fel fog lépni. Ez azt jelenti, hogy a nulladrendű eredmények egyféleképpen, az elsőrendű karakterisztikák pedig egy másik módon fognak megváltozni. A gyakorlatban úgy adódott, hogy kardioid kapszulákból számított jelek túlságosan ingadozó frekvenciamenetű B-formátumú jeleket eredményeztek, így a gyártó szubkardioid kapszulákat szerelt a mikrofonba. A frekvenciamenet a forrás beesési szögétől valamennyire így is függeni fog, és ezt tovább változtatják a kapszulák egyedi sajátosságai, tökéletlenségei, azonban ennél a rögzítési technikánál a szükséges korrekciók segítségével kb. 10 kHz-ig ko incidens vételt kaphatunk, szemben a stúdiótechnikában használatos más sztereó és ko incidens mikrofonozási technikákkal, amelyek a kapszulák távolsága miatt nagyságrendileg 1,5 kHz tartományig működnek helyesen [24].

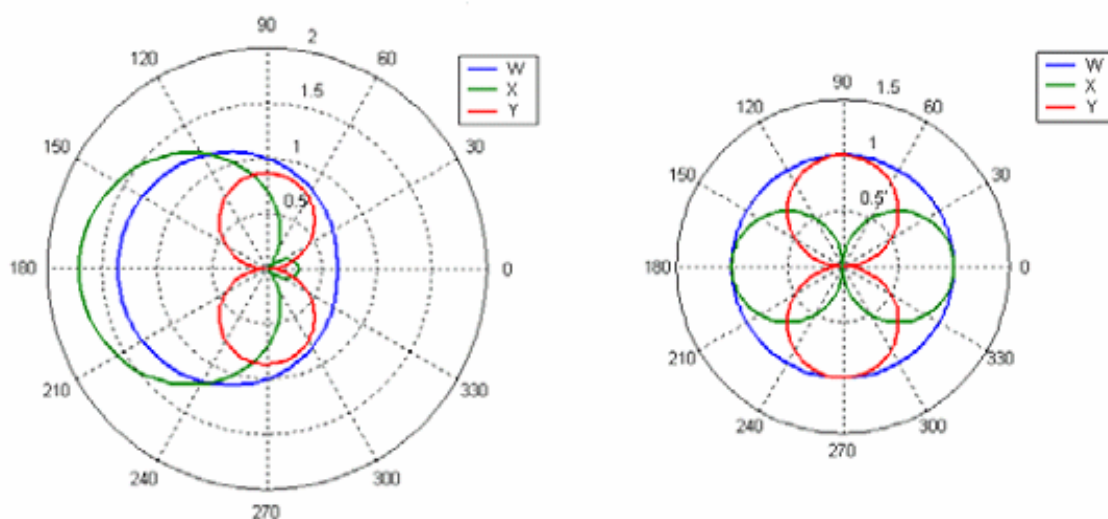
A felvételi technika néhány speciális alkalmazást is lehetővé tesz, amelyek közül a legfontosabb, hogy a B-formátumú jelből bármilyen elsőrendű iránykarakterisztika előállítható. Ebből következik, hogy a B-formátumú jel felhasználásával virtuális mikrofonjelek szintetizálhatók, és olyan utólagos manipulációk válnak lehetségessé, mint a főirány háromdimenziós forgatása, vagy a hangforrásra való ráközelítés. A következőkben megadjuk ezeknek az egyenleteit is [24],[15].

| X-irányú közelítés (zoom)   | Forgatás az X-tengely mentén   | Forgatás a Z-tengely mentén   |
|---|--|---|
| $W' = W + \frac{1}{\sqrt{2}} \cdot d \cdot X$ $X' = X + \sqrt{2} \cdot d \cdot W$ $Y' = Y \sqrt{1 - d^2}$ $Z' = Z \sqrt{1 - d^2}$ <p>ahol</p> <p><math>-1 \leq d \leq 1</math>, <math>d \in \Re</math> ún. dominancia-paraméter</p> | $W' = W$ $X' = X$ $Y' = Y \cdot \cos(\theta) - Z \cdot \sin(\theta)$ $Z' = Z \cdot \cos(\theta) + Y \cdot \sin(\theta)$ <p>ahol</p> <p><math>\theta</math> az elfordítás szöge</p> | $W' = W$ $X' = X \cdot \cos(\theta) + Y \cdot \sin(\theta)$ $Y' = Y \cdot \cos(\theta) - X \cdot \sin(\theta)$ $Z' = Z$ |

4. táblázat – Közelítés és forgatás B-formátumú jelekkel



A közelítés hatása a következő ábrákon látható:



4. ábra – Közelítés hatása az amplitúdó-iránykarakterisztikára horizontális B-formátumú csatornáknál

Forrás: [24], p. 87.

### 3 Konvolúciós zengetés<sup>4</sup>

Az egyes hangesemények megítélésében vitathatatlan szerepe van az akusztikai környezetnek, így egy hangfelvétel is akkor hat természetesen, ha a hangforrás közvetlen (direkt) hangja mellett az akusztikai környezetről is tartalmaz információt. A zenei célú hangfelvétél-készítési technológiák alapvetően produkciós (kreatív) és reprodukciós (Hi-Fi) filozófia mentén alakultak ki (utóbbi esetén „a hangtér keltette hangképnek kell olyannak lennie, hogy az az emberi szubjektumban olyan érzetet keltsen, amelyet a szubjektum természetesnek elfogad.” [11]). Mindkét filozófia követői számára használható eszköz a konvolúciós zengető.

A konvolúciós zengetők elve a hálózateméletben ismert konvolúciós összefüggésen alapul. Az impulzusválasz és a gerjesztés ismeretében a válasz – a zengetett kimenet – meghatározásának módszere diszkrét időre (mintavételezett eset):

$$y[k] = \sum_{i=-\infty}^k e[i]w[k-i], \quad (4)$$

ahol  $w$  a rendszer impulzusválasza (súlyfüggvény, weight function) és  $e$  a gerjesztés (excitation). Mivel a rendszer kauzális, az impulzusválasza belépő:

$$y[k] = \sum_{i=0}^k e[i]w[k-i]. \quad (5)$$

Szemléletesen úgy lehet a konvolúciót értelmezni, hogy felbontjuk a gerjesztést súlyozott és időben eltolts diszkrét Dirac-delták összegére:

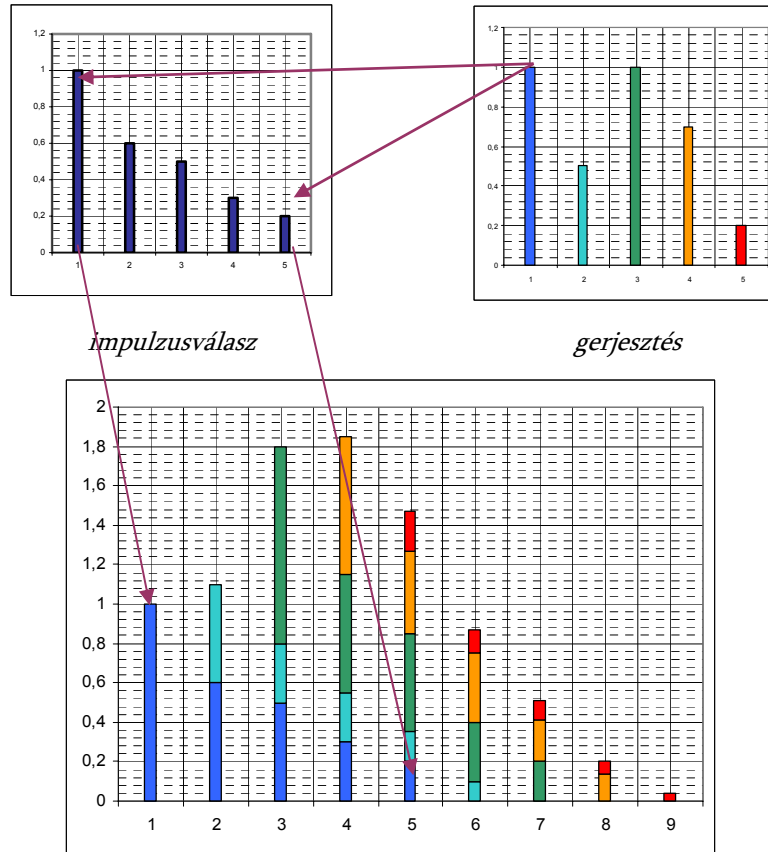
$$e[k] = \sum_{i=0}^{\infty} e_k \delta[k-i], \quad (6)$$

és úgy tekintjük, hogy a rendszer ilyen Dirac-deltákra adott válasza éppen az impulzusválasz  $e_k$ -szorozosa, hiszen a rendszer lineáris. A  $k$ -ik ütemre (a nulladik ütemet is beleszámít-

<sup>4</sup> [7], [8], [9], [10], [20], [23]

va)  $k+1$  darab Dirac-delta válasza érkezik meg, hiszen ugyanennyi Dirac-deltát indított el a gerjesztés.

A rendszer linearitása és időinvarianciája miatt ezek a válaszok a gerjesztések megfelelő eltolásúival szorozódnak és időben eltolódnak, majd összegeződnek.



5. ábra – A konvolúció szemléltetése

Ha felhasználjuk a diszkrét Fourier-transzformáció képletét a konvolúcióra, azt kapjuk, hogy a konvolúció a Fourier-transzformáltak szorzatának inverz Fourier-transzformáltja, így a konvolúció elvégezhető Fourier-transzformálással is – szorzással.

Ez a módszer nemperiodikus jelek esetén is működik, de a periodikus, véges méretű adatblokkokon dolgozó Fourier-transzformációk (DFT, FFT) cirkuláris konvolúciót hajtanak végre [13], azaz egy mintablokk konvolúciója Fourier-transzformációval a blokk végén levő minták konvolváltját behozza a blokk elejére. A jelenség azzal magyarázható, hogy ezek a transzformációk a bemeneti vektort úgy értelmezik, mintha az egy periodikus jel egyetlen mintavételezett periódusa lenne. Ennek megfelelően a válasz is periodikus lesz, és a válaszban azok a minták, amik az eredeti vektor hosszán túlra kerülnének a konvolúció miatt, a periodicitás miatt belapolódnak az előző periódusba is (vagyis abba, amit mi transzformáltunk). Ez nemkívánatos jelenség, de a blokk méretének megduplázásával és a fennmaradó helyek 0-kkal való kitöltésével (*zero padding*) egyszerűen kiküszöbölhető, ami természetesen a számításigény növekedésével jár [7].

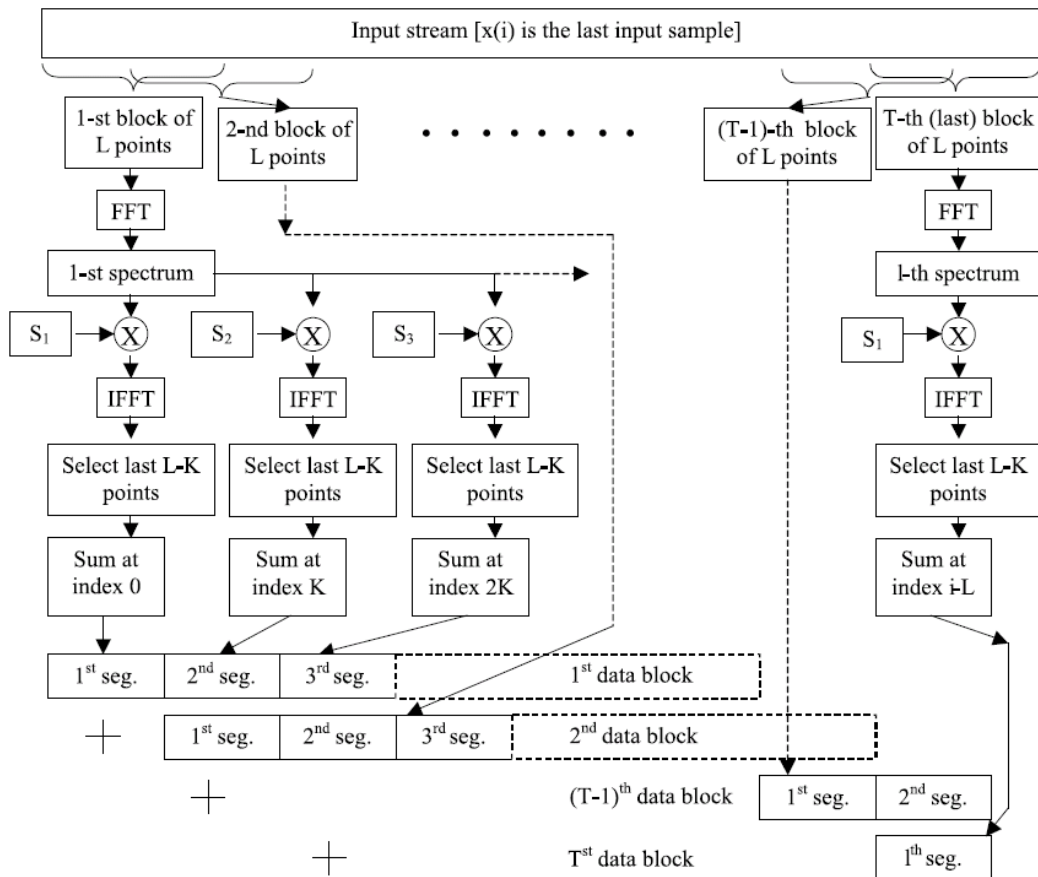
A konvolúciós zengetés gyakorlati megvalósítása azonban rendkívül számításigényes feladat. A számításigény csökkentése és a valósidejűség – vagy kis késleltetés – követelménye egyszerre jelentkeznek. A konvolúciós zengetők alapvető megvalósítási módjai a következők:

| Módszer                       | Előny                                  | Hátrány   |
|-------------------------------|--|---|
| Időtartománybeli (FIR szűrős) | 1 minta késleltetés is megvalósítható  | Rendkívül nagy számításigény (jelenlegi jelfeldolgozó processzorokkal nem valószínűsíthető meg) és felhalmozódó számítási hibák |
| Frekvenciatartománybeli (FFT) | Kisebb számításigény (blokkos kezelés) | A blokkméretnek megfelelő késleltetésű  |

5. táblázat – Konvolúciós zengetők lehetséges megvalósításai

A fentiekből következik, hogy célszerű, ha az impulzusválaszt blokkokra osztjuk, és a blokkokon hajtjuk végre a konvolúciót, majd a részeredményeket felhasználva állítjuk elő a kimenetet.

Az első ilyen felhasználást Stockham javasolta 1966-ban [21]. Módszere egyenlő méretű blokkokra osztotta az impulzusválaszt, amellyel - megfelelően megválasztott zero padding és késleltetés segítségével - az eredeti konvolúcióval ekvivalens átvitelt kapta. A partícionált konvolúció blokkvázlata a következő ábrán látható.  $S_n$  az  $n$ -ik impulzusválasz-blokk frekvenciatartománybeli reprezentációja:



6. ábra – Partícionált konvolúciós zengető blokkvázlata

Forrás: [23]

A minimális késleltetés érdekében célszerű a partíciókat növekvő nagyságúra választani, így az első kisméretű bloknál kis késleltetés érhető el.

|  |         |         |         |         |
|--|---------|---------|---------|---------|
| IR hossza [minta]                              | 114 688 | 122 880 | 126 976 | 130 048 |
| Késleltetés [minta]                            | 16384   | 8192    | 4096    | 1024    |
| Végrehajtás ideje [s] (egyenlő blokkméret)     | 0,45    | 0,71    | 1,33    | 5,25    |
| Végrehajtás ideje [s] (nem egyenlő blokkméret) | 1,06    | 1,19    | 1,27    | 1,45    |
| Partíciók száma (egyenlő blokkméret)           | 14      | 30      | 62      | 254     |
| Partíciók száma (nem egyenlő blokkméret)       | 6       | 8       | 10      | 14      |

6. táblázat – Particionált konvolúciós algoritmusok teljesítménye

Forrás: [23]

## 4 Teremakusztikai mérések alkalmazásai zenetésre

### 4.1 Teremakusztikai mérések célja

A teremakusztikai vizsgálatok kezdetben szubjektív, érzeti és kísérleti úton, később az objektív megítélésre és minősítésre törekedve valósultak meg, a technológiai fejlettség azonban csak a huszadik század közepére tette lehetővé az elfogadható objektív méréseket. Az objektív vizsgálatok háttérét a már korábban említett hálózat- és rendszerelmélet alkalmazása jelentette azzal a feltevéssel, hogy az akusztikai tér mint vizsgálandó objektum lineáris, időinvariáns, kauzális fizikai rendszernek tekinthető – elhanyagolva az esetleges nemlineáris tulajdonságait.

Az akusztikai tér két kijelölt pontja közti átvitelt az impulzusválasszal mint rendszerleíró függvénnyel lehet jellemezni, és az impulzusválasz ismeretében az akusztikai paramétereket számítani, de az impulzusválasz az, amit a konvolúciós zenetők bemenetére kell adni. Az impulzusválasz közvetlen mérése azonban kielégítő minőségben nem lehetséges, mert az ideális impulzust kellő mértékben megközelítő impulzusgerjesztés kibocsátása nem lehetséges: közvetett mérésre van szükség speciális vizsgálójelekkel.

A kezdeti impulzusszerű gerjesztések alkalmazása (pl. taps, vagy szikrakisülés) ma már elavultnak számít és nehezen reprodukálható eredményeket ad. Az 1960-as években kezdődött el az a fejlődési folyamat – a sztochasztikus jelek vizsgálatában, illetve a technológiai fejlődésben –, amelyek eredményeképpen a ma használatos módszerek széles körben elérhetővé váltak. Így született meg az MLS módszer, illetve a számítási kapacitás növekedtével így alkalmazhatták újra a szinuszos gerjesztőjeleket.

### 4.2 Az impulzusválasz mérésének módszerei<sup>5</sup>

Egy terem impulzusválaszának mérésére többféle technikát is alkalmazhatunk, amelyeket közvetlen és közvetett módszerekre osztunk fel. A közvetlen – mára már elavult – módszernek az impulzushoz hasonló közvetlen hang kibocsátását hívjuk (pl. taps, szikrakisülés, kidurranó léggömböt vagy pisztolylövés), ahol a rögzített válasz közvetlenül az impulzusválasz közelítését adja. A közvetett módszer ezzel szemben valamilyen inverz szűrőn (vagy dekonvolúción), illetve más számítási elven alapul, azaz nem szolgáltatja közvetlenül az eredményt. Önkényesen további két csoportra osztjuk a közvetett módszereket a gerjesztőjel szerint: konstans pseudo-véletlen jel (Maximum Length Sequence, MLS), illetve pásztázó szinuszos jellel (sweeping sine vagy time stretched pulse, TSP) történő gerjesztés csoportokra. Egyéb gerjesztésekkel külön nem foglalkozunk, de megjegyezzük, hogy elvileg bármilyen ismert gerjesztőjelet használhatnánk – akár rögzített zenét is. Ezek azonban spektrumuk teljesítménybeli kiegyenlítettensége miatt nem szolgáltatnak kielégítő és egyenletes jel-zaj arányt

<sup>5</sup> [8], [10], [15]

az egyes frekvenciasávokban. A lehetséges mérési módszereket az [1], [2], [3], [5], [17], [18] szakirodalom alapján önkényesen csoportosíthatjuk a következők szerint:

- Közvetlen mérési módszerek (impulzus-szerű gerjesztéssel)
  - taps
  - kidurranó léggömb
  - pisztolylövés
  - elektromos szikra, stb.
- Közvetett mérési módszerek (speciális vizsgálójel)
  - MLS (szélessávú, pszeudo-véletlen zaj)
  - TSP (keskenysávú, time stretched pulse; szinuszos pásztázó jel)
    - Időben lineárisan pásztázó jel (felfele/lefele)
    - Időben exponenciálisan pásztázó jel (felfele/lefele)
    - Egyéb módon pásztázó jel
  - Tetszőleges egyéb ismert gerjesztés (változó sáv szélességű, például zene)

A mérési módszerek összehasonlítását az alábbi táblázatban foglaltuk össze.

| Módszer                | Előnye   | Hátránya  |
|------------------------|--|---|
| Közvetlen mérés        | könnyen elvégezhető, kis eszköz-igényű   | közelítő eredményt ad, reprodukálhatatlan, kis energiájú, egyenetlen spektrumú, elavult                 |
| MLS                    | kiejtheti a konstans véletlen háttérzajt   | torzításra és tranziens zajra érzékeny  |
| Lineáris TSP           | jó SNR hosszú mérés alatt, érzéketlen a torzításra és külön eszközben is működhet  | hosszú mérési idő, nem maximális hangnyomásszint a hangsugárzókból lineáris TSP a fehér spektruma miatt |
| Exponenciális TSP      | még jobb SNR   | körülményesen található hozzá hibátlan inverz szűrő   |
| Tetszőleges jel (zene) | kritikus helyeken (ahol zavaró lenne más gerjesztőjel) mérési lehetőséget biztosít | nehezen található megfelelő spektrális tartalmú zene  |

7. táblázat – Teremakusztikai mérési módszerek összehasonlítása

### 4.3 A mérőjel előállítás

Miután a seprt szinuszos jel a legkedvezőbb a felsoroltak közül, a következőkben levezetjük egy általános  $\nu$  szubmodulációs függvényű seprt szinuszos jel (szinuszos sweep jel) időbeli előállításának formuláit. A keresett jel általános alakja:

$$y = A \sin(\Phi(t)) \quad (7)$$

ahol  $A$  az amplitúdó és

$$\frac{d}{dt} \Phi(t) = \omega(t) \quad (8)$$

az időben változó pillanatnyi körfrekvencia. Legyen a felfutás sebessége a  $\nu(b \cdot t + c)$  szubmodulációs függvényvel adott úgy, hogy

$$\omega(t) = a \cdot \nu(b \cdot t + c) + d \quad (9)$$

ahol  $a$  és  $b$  konstansok egyelőre ismeretlenek,  $c$  és  $d$  pedig egy-egy szabadon választott konstans. A pillanatnyi fázis így felírható a következőképpen:

$$\Phi(t) = \int \omega(t) dt = a \int v(bt+c) dt + \int d dt = \frac{a}{b} \left[ F(bt+c) + bC + \frac{b}{a} d \cdot t + \frac{b}{a} D \right] \quad (10)$$

ahol  $F$  a  $v$  primitív függvénye,  $C$  és  $D$  pedig egy-egy az integrálásból adódó egyelőre ismeretlen konstans. Felhasználva, hogy  $T$  a sweep jel időbeli hossza és  $\omega(t=0) = \omega_1$  kezdeti, valamint  $\omega(t=T) = \omega_2$  végső körfrekvencia, valamint, hogy  $c$  és  $d$  is adott, az eddig ismeretlen  $a$  és  $b$  konstansok kifejezhetők az alábbiak szerint:

$$a = \frac{\omega_1 - d}{v(c)} \quad (11)$$

$$b = \frac{1}{T} v^{-1} \left\{ v(c) \frac{\omega_2 - d}{\omega_1 - d} \right\} \quad (12)$$

ahol  $v^{-1}$  a  $v$  függvény inverzét jelöli. Célszerű a jel kezdőfázisát úgy megválasztani, hogy  $\Phi(t=0) = 0$  legyen, mert így a  $C$  és  $D$  konstansok fent szereplő súlyozott összege is kifejezhető lesz, mégpedig úgy, hogy:

$$b \cdot C + \frac{b}{a} D = -F(c) \quad (13)$$

Ez valójában csak azzal a feltétellel igaz, hogy  $a \neq 0$ , vagyis  $v(c) \neq \infty$  és  $\omega_1 \neq d$ . A pillanatnyi fázis mindezek alapján az alábbi formulával írható fel:

$$\Phi(t) = \frac{a}{b} [F(bt+c) - F(c)] + d \cdot t \quad (14)$$

és ezzel fejezhető ki az általános formula, amely szerint:

$$y(t) = A \sin \left( \frac{(\omega_1 - d)T}{v(c)v^{-1} \left\{ v(c) \frac{\omega_2 - d}{\omega_1 - d} \right\}} \left[ F \left( \frac{1}{T} v^{-1} \left\{ v(c) \frac{\omega_2 - d}{\omega_1 - d} \right\} \cdot t \right) - F(c) \right] + d \cdot t \right) \quad (15)$$

Exponenciális sweep esetén a következőket tudjuk:

$$v(x) = \exp(x) \quad (16)$$

és így

$$v^{inv}(x) = \ln(x) \quad (17)$$

ahonnan

$$F(x) = \int \exp(x) dx = \exp(x) \quad (18)$$

Ezt felhasználva legyen

$$\omega(t) = a \cdot v(b \cdot t + c) + d = a \cdot \exp(b \cdot t) \quad (19)$$

A fenti definíciók szerint a segédváltozók értékei:

$$a = \frac{\omega_1 - d}{v(c)} = \omega_1 \quad (20)$$

$$b = \frac{1}{T} v^{-1} \left\{ v(c) \frac{\omega_2 - d}{\omega_1 - d} \right\} = \frac{1}{T} \ln \left( \frac{\omega_2}{\omega_1} \right) \quad (21)$$

És így a fázis időfüggvénye:

$$\Phi(t) = \frac{a}{b} [F(bt + c) - F(c)] + d \cdot t = \frac{\omega_1 T}{\ln \left( \frac{\omega_2}{\omega_1} \right)} \left[ \exp \left( \frac{1}{T} \ln \left( \frac{\omega_2}{\omega_1} \right) \cdot t \right) - \exp(0) \right] \quad (22)$$

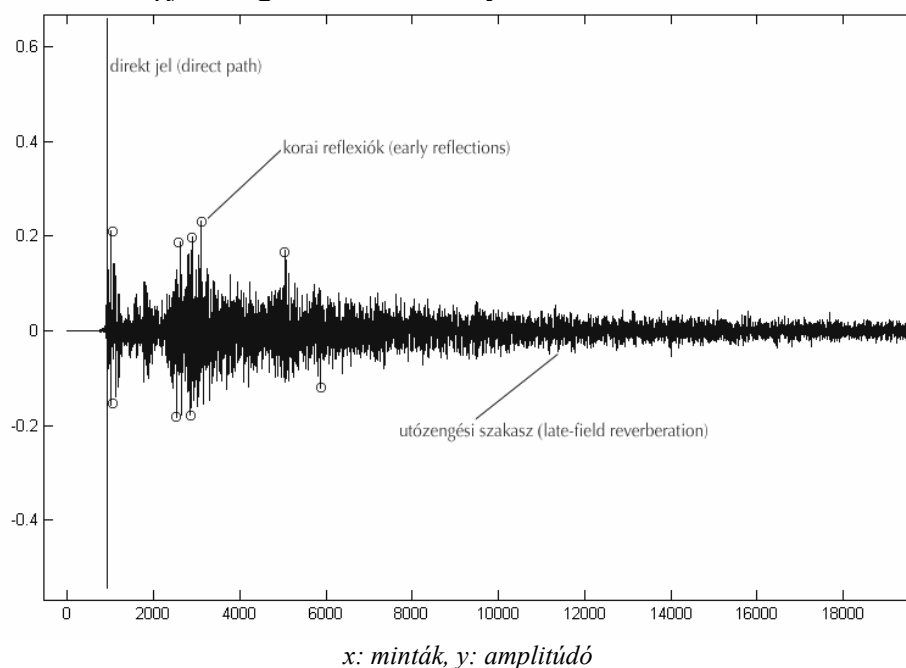
Amiből az exponenciális seprt szinuszos jel időtartománybeli előállítás:

$$y(t) = A \sin \left( \frac{\omega_1 T}{\ln \left( \frac{\omega_2}{\omega_1} \right)} \left[ \exp \left( \frac{1}{T} \ln \left( \frac{\omega_2}{\omega_1} \right) \cdot t \right) - 1 \right] \right). \quad (23)$$

Az illesztett szűrőt, amellyel konvolúciót végezve az ideális impulzust kapjuk nem analitikus formában, hanem dekonvolúcióval célszerű előállítani a frekvenciatartományban, az ideális impulzus és a módosított gerjesztőjel alapján. A fent előállított gerjesztőjelet a felfutás és a végpont körül hangosítani és halkítani célszerű a véges sáv szélesség biztosítása érdekében.

#### 4.4 Az impulzusválasz<sup>6</sup>

Ha jól mértük az impulzusválaszt, akkor a mérés eredményeképpen a rendszer adott két pontja közti átvitelt kapjuk meg. Ez időtartományban az alábbiak szerint alakul.

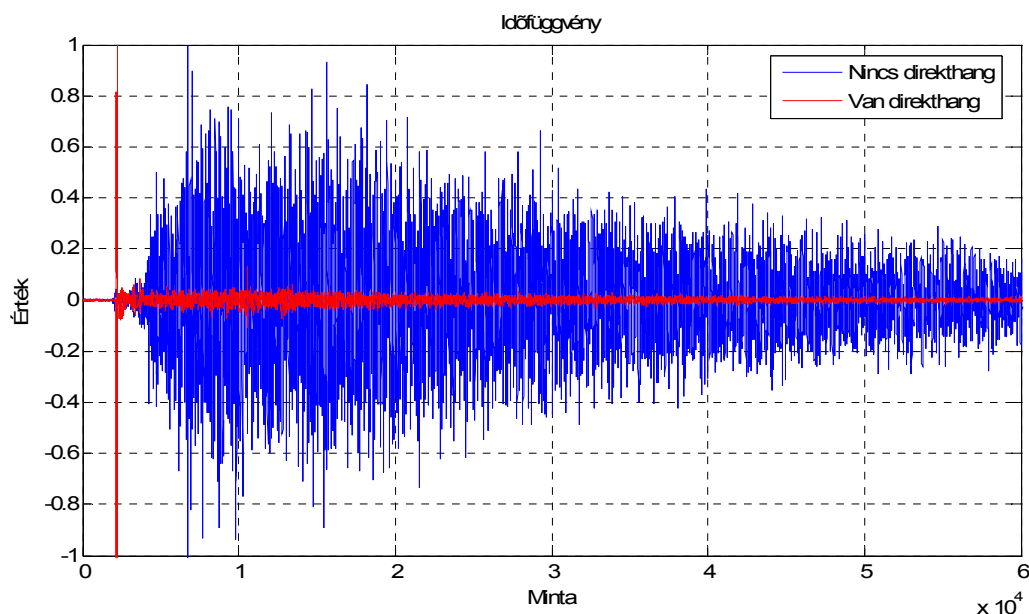


7. ábra – az impulzusválasz felépítése (Mátyás-templom)

<sup>6</sup> [8]

Az ábrán látható, hogy az impulzusválasz egyrészt egy késleltetéssel indul, ami a hangforrás és a mikrofon közti távolságból adódik, másrészt néhány jellegzetes szakaszra bontható, amelyek a későbbiekben a modellezés szempontjából fontos szerepet fognak játszani. A direkt hang egy Dirac-impulzus formájában jelenik meg, amit néhány korai reflexió követ. Ezek a környező falfelületekről (oldalfal, padló, stb.) történt visszaverődések eredményei, többnyire még jól elkülönülő formában, majd az utózengési szakasz következik, amely már nem tartalmaz követhető visszaverődéseket.

Érdekességképpen a következő ábrán két impulzusválaszt rajzoltunk egymásra, az egyiknél jelentős direkthang mérhető, a másikonál viszont a láthatóság nem teljesült a hangforrás és a mikrofon között, így nem tudunk direkthangot mérni.



8. ábra – Direkthang és direkthang nélküli impulzusválasz időfüggvénye

#### 4.5 Előfeldolgozás zengetéshez<sup>7</sup>

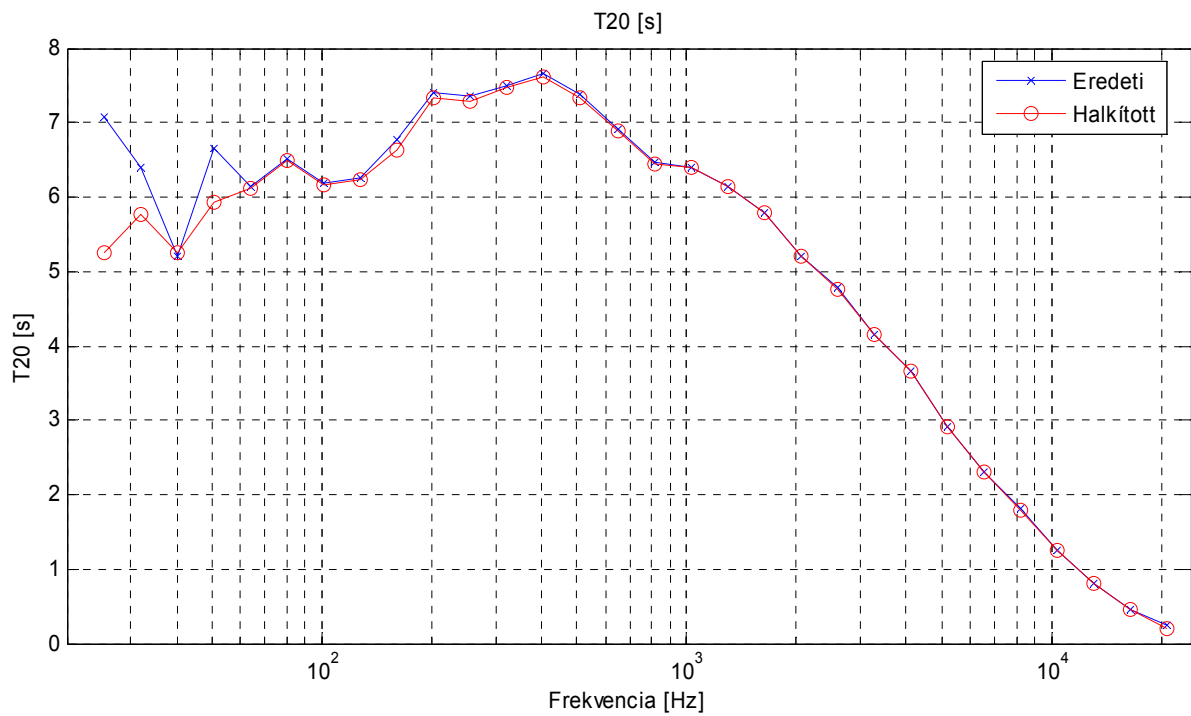
A gyakorlati tapasztalat szerint a teremakusztikai mérések során legfeljebb 40 és 60 dB közötti (szélessávban mért) jel-zaj arányt lehet elérni a kívülről beszűrődő és a teremben tapasztalható alapszaj miatt. Ha az impulzusválaszt akusztikai paraméterek kiértékelésére használjuk, többnyire ritkán van szükség ennél jobb jel-zaj viszonyra, hiszen a lecsengési görbéről kiolvasott értékek extrapolációjával becsülhető a zajszint alatti viselkedés is, a zengetésnél azonban a zajos impulzusválasz kitarított zúgó hangot eredményez, ami elengedhetlenné teszi a korrekciót, de a rögzítési technika és a számábrázolás lehetővé teszi a nagyobb dinamika kihasználását.

A szélessávú zaj, amelyet eszközeink rögzítenek, többnyire nem fehér és nem is biztos, hogy nem tartalmaz determinisztikus komponenseket – pl. ismétlődő részeket. Ennek érdekében célszerű az impulzusválaszok zajszint alatti tartalmát időtartományban beillesztett, frekvenciafüggő módon a lecsengéshez illesztett kontúrvonalak mentén lehalkított fehérzajra cserélni. A beillesztő algoritmus a közepes jel-zaj viszony szintjéhez illesztett zaj-kontúrvonal mint referencia kontúrvonal segítségével automatizálható figyelembe véve, hogy az eredeti jelet az impulzusválasz tényleges hosszának függvényében meghatározott idő alatt, a kismagasságú frekvenciák fele egyre növekvő hosszon keverjük majd át fehérzajba. A halkítás eredményeképp

<sup>7</sup> [8]

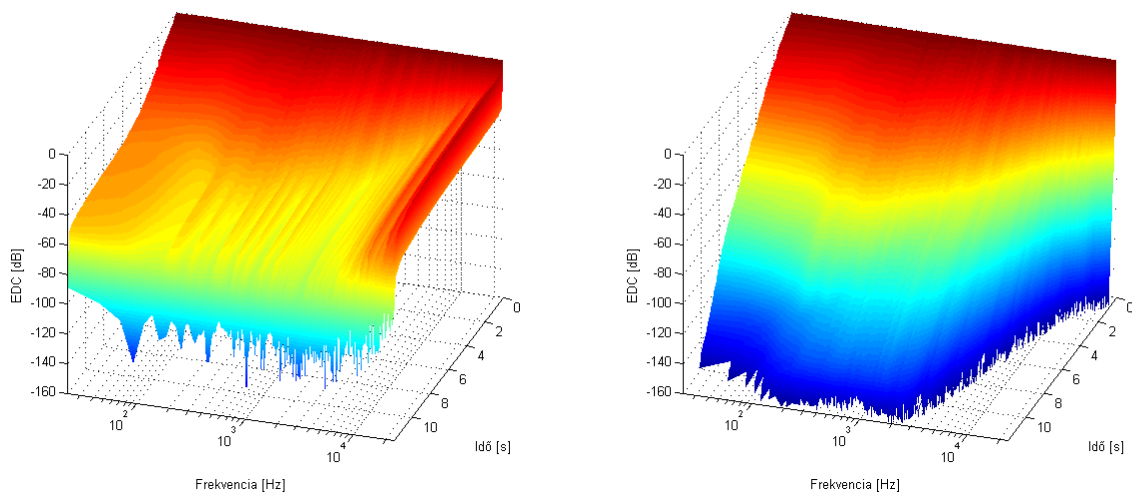


a teremakusztikai paraméterek lényegesen nem változnak, csak az igen rossz jel-zaj viszonyú kisfrekvenciás tartományban mutatkozik némi eltérés.



9. ábra – T20 (utözengési idő) görbe halkítás után és előtt

Az energia-lecsengési görbe (EDC) frekvenciafüggő ábrázolásával (EDC vizesés felület) jól látható az előfeldolgozó algoritmus működésének eredménye.



Szegedi Dóm, eredeti impulzusválasz

Szegedi Dóm, halkított impulzusválasz

1025 lineáris frekvenciasáv logaritmus frekvenciatengelyen (20-24000 Hz)  
a színskála a z-tengelyen látható értékeknek felel meg

10. ábra – EDC vizesés felület az eredeti és az előfeldolgozott impulzusválasz esetében

## 4.6 A teremakusztikai mérések alkalmazásának korlátai<sup>8</sup>

Korábban már említettük, hogy a teremérzet kialakításában nem kizárólagos szerepű a zengés minősége, és a hangtér valóság-hű reprodukciója. Hangsugárzórendszerrel auralizálva a lehallgatói környezetnek minél inkább semlegesnek kell lennie, leginkább a süketszobás visszahallgatás lenne ideális, azonban ilyen helységben hosszan tartózkodni természetellenes érzetet kelt, ami nincs jó hatással az értékítéletre. A szubjektív érzethez a hangzáson túl egyéb érzékszerveink is sokat hozzátesznek, és ezek együtt alakítják ki azt a benyomást, hogy egy templomban, vagy éppen egy hangversenyteremben tartózkodunk.

A teremakusztikai mérések során a fentiekén kívül még néhány további körülményt is figyelembe kell venni. Az eszközök tökéletlenségén kívül az egyik – és talán leglényegesebb –, az alkalmazott hangsugárzó iránykarakterisztikája. A zenetnét kívánt hangforrásunk, például egy zongorahang, vagy egy beszélő személy mint hangforrás nem gömbi iránykarakterisztikájú, hanem valamilyen egyedi, jellemzően ismeretlen iránykarakterisztikával rendelkezik. A zengetett felvétel a mérés során használt hangsugárzó – frekvenciafüggő – iránykarakterisztikájával fog a teremben virtuálisan megszólalni, tehát ha egy hangszer hátrafelé jobban sugároz, mint előre, a mérés során pedig előre felé fordított hangsugárzóval mértünk, akkor személetesen és meglehetősen nagyvonalú közelítéssel úgy tekinthetjük, mintha a hangszert háttal megfordítva hallgatnánk. A teremakusztikai méréseink során óhatatlanul irányított hangsugárzót használunk, mert még a mérési célú hangsugárzók sem tekinthetők szélessávon gömbinek. A valódi megoldás a hangsugárzók iránykarakterisztikájának kompenzálására az lehet, ha a mérés során a forrást többféle irányba fordítva irányonként szeparált mérést végzünk. A másik jelentős korlátozás, hogy az impulzusválaszok rögzített pozíciókhoz kötöttek, vagyis nem lehetséges a források folytonos virtuális mozgatása a térben. Valamiféle szimuláció, vagy a mérési eredményekkel kalibrált modellezés, ami a tér geometriáját is figyelembe veszi, megoldhatja ezt a problémát, de ez szintén további erőforrást igényel a zengetéskor.

A harmadik figyelembe veendő körülmény pusztán mérés-technikai jellegű: jogosan mérülhet fel a kétely, hogy az akusztikai tér valóban lineáris időinvariáns rendszernek tekinthető-e (hőingadozás, légmozgás, szél, páratartalom-változás, stb.). A mérési módszer – vagyis a vizsgálójel – megválasztása az idővariancia figyelembevételével kell, hogy történjen.

Mindezek ellenére a mérések során a termék jellegzetes hangzása a tapasztalataink szerint felismerhető.

## 5 Kreatív beavatkozási lehetőségek az impulzusválasz módosításával

### 5.1 Közvetlen konvolúciós zengetés: auralizáció

A vizsgált akusztikai tér meghallgatható és reprodukálható szimulációja az auralizáció [16]. Az auralizáció megvalósításához offline vagy valós idejű konvolúcióra van szükség.

Az auralizáció főbb alkalmazási területei szubjektívek, a teljesség nélkül néhány kiragadott példával élve használatosak a teremakusztikai tervezési fázisban is egyfajta előzetes meghallgatás formájában, valamint a szempontunkból sokkal érdekesebb területen hangforrások a térbe történő utólagos, virtuális elhelyezésekor.

Az impulzusválaszokon alapuló auralizáció során lényeges kérdés, hogy visszahallgatásra milyen akusztikai környezetben kerül sor. Fejhallgatós meghallgatás esetén figyelembe kell venni az emberi fej átviteli függvényeit (HRTF) és az impulzusválaszokat ennek megfelelően korrigálni, vagy rögzíteni (műfejes vevő). A fejhallgatóval történő visszahallgatásnál rontja a szubjektív megítélést, hogy az elrendezés a fej elmozdulására érzéketlen. A hangsug-

---

<sup>8</sup> [8], [15]

gárgós visszahallgatás mindezen hiányosságokat kiküszöböli és több hangsugárzó egyidejű használatával lehetővé válik a pontosabb irány-lokalizáció is, azonban a visszahallgató helyiség saját akusztikai jellemzői módosítják a térérzetet.

Fontos, de nem jól mérhető szubjektív tulajdonság a visszahallgatásra használt terem egyéb tulajdonságai (burkolatok, színek, szagok) és minden más, a személlyel összefüggő érzet (hangulat, kedélyállapot, fizikai tünetek), és bár ezekkel a továbbiakban nem foglalkozunk, megjegyezzük, hogy a kísérleti személy vélhetőleg könnyebben felismer egy templomot egy kőburkolatú, hideg szobában hallgatva, mint egy ezzel nem összhangban levő teljesen szokatlan környezetben.

A saját akusztikai tér hatása mérhető és megfelelő feldolgozással – minden egyes rendszernél külön-külön – kompenzálható, a fejhallgatós kiértékelés pedig egyedi HRTF méréseket igényelne minden egyes kísérleti személyhez. Ennek elkerülésére lehetőség van általánosan alkalmazott HRTF függvényekkel is dolgozni, azonban ennek a kérdésnek, illetve a visszahallgatás kompenzációjának a vizsgálatával itt most nem foglalkozunk, inkább maguknak az impulzusválaszoknak a módosításához térünk vissza.

## **5.2 Konvolúciós és konvencionális zengetés**

Amíg az auralizáció során a valóság-hű reprodukció érdekében nem változtatunk az impulzusválaszon, a hangmérnökök azonban megszokták, hogy bizonyos paramétereket szubjektív értékítéletüknek megfelelően megválasszthassanak, és a zenei produktumhoz igazíthatóknak.

Felmerülhet a kérdés, hogy ennek figyelembevételével milyen lényegi különbség, előny jellemzi a konvolúciós zengetőket a konvencionális zengetőkhöz képest.

Az akusztikai tér a valóságban nem tartalmaz visszacsatolást, a konvencionális zengetők pedig jóformán kivétel nélkül igen, és ez az elvi különbség számos következménnyel jár, amelyek a zengő hang szubjektív érzetét is befolyásolják. A hatás különösen hosszú impulzusválaszoknál érvényesül (templomok), de természetesen a teljes impulzusválaszt érintik, így felfedezhetők strukturális különbségek a korai szakaszban is. A konvencionális zengetők nem arra szolgálnak, hogy valós akusztikai viszonyokat lehessen előállítani a segítségükkel, de még csak nem is ilyenekből indulnak ki, így lényegében bármilyen beállítást is használunk, annak aligha lesz köze a valósághoz. Megszokni szinte bármilyen eszközt vagy hangzást lehetséges, azonban azáltal, hogy a konvolúciós zengetőknél is lehetővé válik bizonyos fokú kreatív beavatkozási lehetőség, de valós mérési adatokból indulunk ki, természetesebb hangzás érhető el, a konvolúciós zengetés viszonylag nagy számítási kapacitása árán.

## **5.3 Beavatkozási lehetőségek**

A hangmérnökök, amikor zengetőkkel operálnak, néhány jellemző paraméterrel találkoznak, amelyek olykor könnyen, olykor nehezebben feleltethetők meg az impulzusválasz különféle paramétereinek. A megfeleltetés sokszor nem egyértelmű, zengetőnként eltérő is lehet. Az itt közzétett megfeleltetés önkényes, a szerzők által is fejlesztett konvolúciós zengetőhöz kötődik.

| Konvencionális zengető paramétere                | Megfeleltetés a konvolúciós zengetőhöz, illetve az impulzusválaszhoz          |
|--|---|
| Teremméret (room size)                           | Utózungési idő (frekvenciafüggetlen), reflexiók sűrűsége, hangszínszabályozás |
| Utózungési idő (reverberation time)              | Utózungési idő (frekvenciafüggetlen)  |
| Csillapítás (damping)                            | Utózungési idő (frekvenciafüggő)  |
| Visszhangosság-simaság, diffuzitás (diffusivity) | Reflexiók sűrűsége az impulzusválaszban, hangszínszabályozás                  |
| Korai visszaverődések aránya (ER)                | Korai-késői szakasz amplitúdóaránya   |
| Sztereó-kép szélessége (Stereo width)            | Nincs megfeleltetés, a mérési módszerből közvetlenül adódik                   |
| Nagyfrekvenciás vágás                            | Hangszínszabályozás   |
| Kisfrekvenciás vágás                             | Hangszínszabályozás   |
| Wet/Dry keverés                                  | Direkt-zengő arány állítása   |

8. táblázat – Beavatkozási lehetőségek az impulzusválasz módosításával

Ha jobban szemügyre vesszük a fenti megfeleltetéseket, rájöhethetünk, hogy a megfeleltetések tulajdonképpen néhány egyszerű elemre vezetnek. A frekvenciafüggő utózungési idő szabályozása az impulzusválasz alkalmas hangszínszabályozásával és nyújtásával elérhető, így végül a következő műveletek maradnak:

- Időbeli nyújtás-zsugorítás
  - interpoláció, decimálás
  - „sűrítés”, „ritkítás”
- Hangszínszabályozás
- Erősítés
- Késleltetés

A konvolúciós zengetőben a direkthang és a zengő szakasz jelét célszerű külön kezelni. Ez sokféle optimalizációs lehetőséget felvet, illetve lehetővé teszi a hangmérnök számára, hogy a zengetőt pusztán pozícionálásra használja egy teremben, majd egyetlen zengő rész lecsengési szakaszával konvolválva korábban sosem létező, mégis valóságghű hangzasképet alakítson ki. A következőkben egy példán szemléltetjük a sűrítő-ritkító algoritmus hatását. Azt vizsgáljuk meg, hogy a hangtisztasági fokot hogyan befolyásolja az impulzusok sűrítése, illetve ritkítása a korai szakaszban.

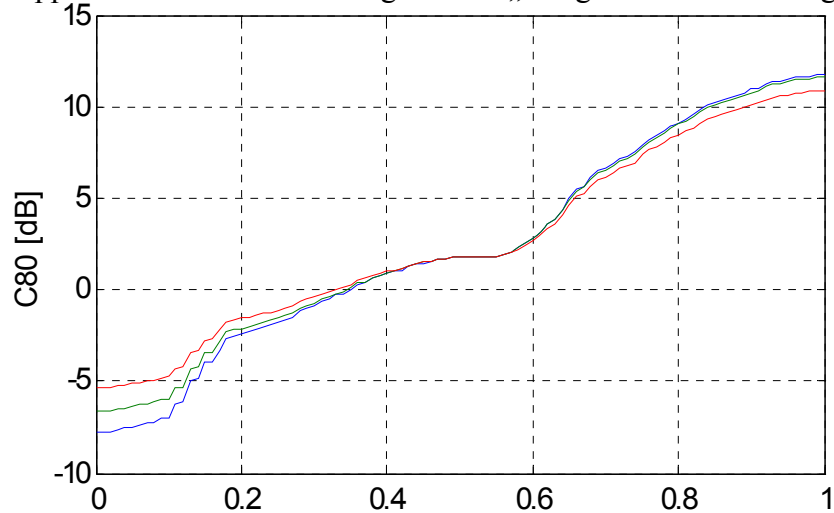
A hangtisztaság az ISO 3382 szabványban [12] definiált teremakusztikai paraméter. Bevezetését az a szubjektív érzet indokolja, hogy minél rövidebb idő alatt érkezik meg az energia nagy része a hallgatóhoz, annál tisztábbnak, érthetőbbnek tűnik számára a hang. Zenei és beszédcélú alkalmazásokban eltérő empirikus értéket határoztak meg a benyomás számszerűsítésére.

Jelölje  $E_{t_0}^{te}$  a korai energiatartalmat a  $t_0$  pillanattól 80 ms-ig (illetve beszéd esetén 50 ms-ig),  $E_{te}^{\infty}$  pedig az ezt követő összes energiát (késői energiatartalom), amit az impulzusválasz mintáinak négyzetösszegéből nyerhetünk. A  $t_0$  paraméter értéke nulla, ha az impulzusválaszban nincs késleltetés, vagyis az első mintája a direkt hang. A hangtisztasági fok (Clarity) zenei célú megítélés esetén az alábbi formulával kapható:

$$C_{80} = 10 \lg \frac{E_{t_0}^{te}}{E_{te}^{\infty}} \text{ [dB]}, \quad (24)$$

A hangtisztasági fok zenei célú termék esetén -4 dB és +2 dB között kedvező.

A szerzők által fejlesztett algoritmust egy templomi impulzusválasz korai szakaszán vizsgáltuk. A sűrítés mint bemenő paraméter 0 és 1 között változott, 0.5 érték jelentette a változatlanul hagyást. A következő ábrán látható, hogy a beavatkozás viszonylag széles skálán tudta változtatni az impulzusválasz hangtisztaságát, a hatás pedig jól hallható volt: a ritkítás eredményeképpen a várakozásoknak megfelelően „levegősebb” érzetű hang keletkezett.



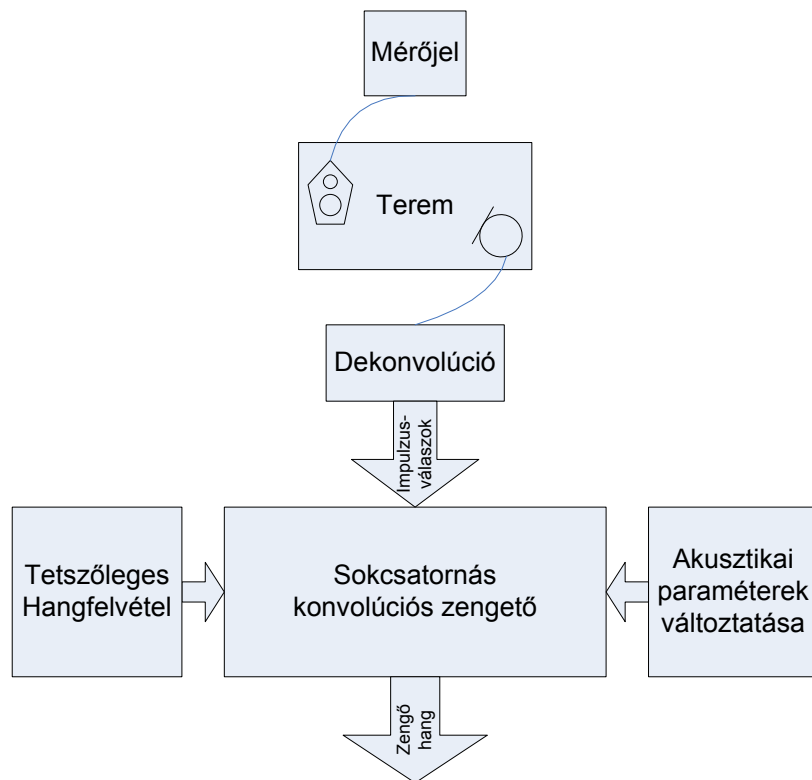
piros: 80%, zöld: 50%, kék: 1% átkeverési idő a módosított és az eredeti korai szakasz között

11. ábra – Sűrítés és ritkítás hatása a hangtisztaságra az impulzusválasz korai szakaszában a sűrítési paraméter függvényében

## 6 Az Ambisonics technikájú teremakusztikai mérésekre támaszkodó konvolúciós zengető rendszer felépítése<sup>9</sup>

A reprodukálendő teremben (pl. egy hangversenyterem) elhelyezett hangsugárzóra adott mérőjel segítségével Ambisonics technikával rögzítjük a mérőjelet. Dekonvolúció segítségével előállítjuk az átvitelre jellemző impulzusválaszokat (pozícióként négyet a B-formátumnak megfelelően), majd az ún. konvolúciós zengetőbe töltjük. A zengető a bemenetére adott zengetendő hangjel, az impulzusválaszok, és a módosítandó teremakusztikai paraméterek függvényében előállítja a zengetett hangot. A zengetőben található az Ambisonics dekóder is, ami egy mátrixművelet segítségével a bemenő négy csatornából tetszőleges hangsugárzó-elrendezést állít elő, így a kimenet lehet akár több csatornás is, mint a bemenet.

<sup>9</sup> [8], [9], [10], [15]



12. ábra – Teremakusztikai mérésre épülő konvolúciós zengető rendszer blokkvázlata

## 7 Gyakorlati alkalmazási lehetőségek

Az Ambisonics rendszerben végrehajtott teremakusztikai mérésekre támaszkodó konvolúciós zengetés alkalmazásai számos területre kiterjedhetnek. Elsőként talán nem is a zengetés hangmérnöki alkalmazásaira célszerű gondolni, hanem inkább annak egy általánosabb megközelítésére, ami az egyik virtuális akusztikai térből a másikba való áthelyezést teszi lehetővé. Amennyiben a hanghordozón a hangjel kis utózengezési idővel, vagy zengés nélkül lett rögzítve, esetleg sávonként is rendelkezésre áll, és adott a szükséges számítási kapacitás is, a rendszer valós időben is működőképes. Néhány példa a lehetséges alkalmazásokra:

- Zeneszerzők: virtuális hangszerek (sound samples, virtual instruments) zengetése
- Hangmérnökök: hangstúdióban rögzített hangsávok virtuális térbe helyezése
- Felhasználók: számítógépes játékok, filmek és videók a megjelenített tartalomhoz igazodó valósidejű hang megjelenítése
- Kutatási alkalmazások: auralizáció, szubjektív tesztek, vizsgálatok
- Egyéb alkalmazások: aktív teremakusztikai beavatkozások

Az alkalmazások közül néhányat az előadást során példákkal is szemléltetünk.

## 8 Összefoglalás

A dolgozat bemutatta egy olyan rendszer elemeinek felépítését, amelynek segítségével ideális lehallgatói környezetben valódi akusztikai terek mérésével alátámasztott valósidejű zengetéssel bizonyos korlátok között reprodukálható az akusztikai környezet az alkalmazott hangsugárzó-rendszer elrendezésétől függetlenül.

Áttekintettük a surround reprodukciós technikákat, és az ezekhez szükséges felvételi technikákat nagy vonalakban, majd bemutattuk az Ambisonics rendszer előnyeit és sajátosságait.

A dolgozatban ismertettük a konvolúciós zengetők működését, levezettük a teremakusztikai méréshez szükséges vizsgálójel teljesen általános alakját időtartománybeli előállításához, az eredményt megoldottuk exponenciális felfutású gerjesztőjelre, majd áttekintettük a teremakusztikai mérések alkalmazhatóságának korlátait. Bemutattuk a zengetéshez szükséges élőfeldolgozást az impulzusválaszokon, majd megmutattunk néhány olyan módszert is, amelyek alkalmasak arra, hogy a hangmérnökök számára megszokott módon befolyásolhassuk a rögzített impulzusválaszokat a zengetés során, így lehetővé téve számukra a valódi mérési eredményeken alapuló kreatív beavatkozási lehetőségeket.

## 9 Hivatkozások

- [1] A. FARINA: Simultaneous Measurement of Impulse Response and Distortion with a Swept-Sine Technique, 108th AES Convention, Paris 18-22 February 2000
- [2] A. FARINA, F. RIGHINI: Software implementation of an MLS analyzer, with tools for convolution, auralization and inverse filtering, Pre-prints of the 103rd AES Convention, New York, 26-29 September 1997
- [3] BORISH J. ANGELL J.B.: An Efficient Algorithm for Measuring the Impulse Response Using Pseudorandom Noise J. Audio Eng. Soc. Vol. 31, No. 7 1983 July/August
- [4] D.G. MALHAM: Spatial Hearing Mechanisms and Sound Reproduction, University of York Music Technology Group
- [5] FUMIAKI S., JIN H., SHINICHI S., HIDEKI T.: Comparison between the MLS and TSP methods for room impulse response measurement under time-varying condition, April 11-13, 2004, International Symposium on Room Acoustics: Design and Science, Kyoto
- [6] GERZON, M. A.: Sound Reproduction Systems. Patent No. 1494751, 1974
- [7] GARDNER W.G.: Efficient convolution without input-output delay, JAES vol. 43, n. 3, 1995 March, pp. 127–136.
- [8] HUSZTY CS.: Korszerű impulzusválasz-mérések alkalmazása a teremakusztikai értékelésben, Diplomaterv dolgozat, 2006
- [9] HUSZTY CS. BALÁZS G.: Valós idejű és offline 7.1 rendszerű sokcsatornás zengetés megvalósítása impulzusválasz-mérések eredményeivel, BME TDK Konferencia, 2004
- [10] HUSZTY CS. MARSCHALL M. JUHÁSZ F. LAKATOS G.: A teremakusztikai hangtérképezés új módszerei és alkalmazásai, BME TDK Konferencia, 2006
- [11] HUSZTY D.: Gondolatok a "Hi-Fi" fogalmáról, Kép és Hangtechnika, XVIII. évf. 1972. 6. füzet, 167-173.  
[http://www.husztty.org/hd\\_hifi.pdf](http://www.husztty.org/hd_hifi.pdf)
- [12] ISO/FDIS 3382, ACOUSTICS: Measurement of the Reverberation Time of rooms with reference to other acoustical parameters, International Organisation for Standardisation, 1997
- [13] J. W. COOLEY, J. W. TUKEY: An Algorithm for the Machine Calculation of Complex Fourier Series – A Butterfly FFT módszer  
<http://www.ph.utexas.edu/~itiq/chiu/cooley/>
- [14] L. L. BERANEK.: Concert and Opera Halls — How They Sound. Acoustical Society of America, Woodbury, NY, USA, 1996.
- [15] MARSCHALL M.: Térbeli impulzusválasz mérése és alkalmazása a teremakusztikában, Budapesti Műszaki és Gazdaságtudományi Egyetem, Diplomaterv dolgozat, 2006
- [16] MENDEL KLEINER: Research Profile (Chalmers University of Technology)  
[www.chalmers.se/HyperText/Prof-E/Kleiner-E.html](http://www.chalmers.se/HyperText/Prof-E/Kleiner-E.html)
- [17] P. D'ANTONIO, J.H. KONNERT: Complex Time Response Measurements Using Time Delay Spectrometry, 83rd AES Convention, New York (October 1987), Preprint No. 2542 [B-1].
- [18] P. FAUSTI, A. FARINA: Acoustic measurements in opera houses: comparison between different techniques and equipment, Journal of Sound and Vibration, vol.232, no. 1, April 2000, ISSN 0022-460X, pp. 213-229  
<http://www.angelfarina.it/Public/Papers/141-JSV00.PDF>
- [19] R. FURSE: First and Second Order Ambisonic Decoding Equations.  
<http://www.muse.demon.co.uk/ref/speakers.html>
- [20] REALLY, D.S. MCGRATH: Convolution processing for realistic reverberation, Pre-Prints of the 98th AES Convention, 1995 February 25–28, Paris
- [21] T. G. STOCKHAM JR.: High-speed convolution and correlation”, AFIPS Proc. 1966 Spring Joint Computer Conf., Vol. 28, Spartan Books, 1966, pp. 229 -233.
- [22] Y. SUZUKI, F. ASANO, H.-Y. KIM, T. SONE: An optimum computer-generated pulse signal for the measurement of very long impulse responses, J. Acoust. Soc. Am. Vol.97 (2), pp.-1119–1123, 1995
- [23] A. TORGER, A. FARINA: Real-Time partitioned convolution for Ambiphonics surround sound, Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA01), Mohonk Mountain Resort, NY, 21-24 October 2001
- [24] B. WIGGINS: An investigation into the real-time manipulation and control of three-dimensional sound fields, PhD thesis, University of Derby, UK, 2004
- [25] WIKIPEDIA: Surround sound  
[http://en.wikipedia.org/wiki/Surround\\_sound](http://en.wikipedia.org/wiki/Surround_sound)