



M Ű E G Y E T E M 1 7 8 2

Budapesti Műszaki és Gazdaságtudományi Egyetem
Villamosmérnöki és Informatikai Kar
Hálózati Rendszerek és Szolgáltatások Tanszék

Herczeg Ferenc

**DOLBY SURROUND
RENDSZEREK ÉS WAVE FIELD
SYNTHESIS KOMPATIBILITÁS-
VIZSGÁLATA**

KONZULENS

Firtha Gergely

BUDAPEST, 2018

Tartalomjegyzék

Összefoglaló	4
Abstract.....	5
1 Bevezetés	6
2 Hangtér-reprodukciós rendszerek	8
2.1 Bevezetés	8
2.2 Surround rendszerek	8
2.3 Wave Field Synthesis.....	8
3 Independent Component Analysis	9
3.1 Elméleti áttekintés.....	9
3.2 Skálázási probléma	11
3.3 Permutációs probléma.....	11
4 Spherical Harmonic Transform	12
5 Plane Wave Expansion	13
6 Eredmények.....	15
7 Összefoglalás.....	16
Irodalomjegyzék.....	17

HALLGATÓI NYILATKOZAT

Alulírott **Herczeg Ferenc**, szigorló hallgató kijelentem, hogy ezt a diplomatervet meg nem engedett segítség nélkül, saját magam készítettem, csak a megadott forrásokat (szakirodalom, eszközök stb.) használtam fel. Minden olyan részt, melyet szó szerint, vagy azonos értelemben, de átfogalmazva más forrásból átvettem, egyértelműen, a forrás megadásával megjelöltem.

Hozzájárulok, hogy a jelen munkám alapadatait (szerző(k), cím, angol és magyar nyelvű tartalmi kivonat, készítés éve, konzulens(ek) neve) a BME VIK nyilvánosan hozzáférhető elektronikus formában, a munka teljes szövegét pedig az egyetem belső hálózatán keresztül (vagy hitelesített felhasználók számára) közzétegye. Kijelentem, hogy a benyújtott munka és annak elektronikus verziója megegyezik. Dékáni engedéllyel titkosított diplomatervek esetén a dolgozat szövege csak 3 év eltelte után válik hozzáférhetővé.

Kelt: Budapest, 2018. 05. 25.

.....
Herczeg Ferenc

Összefoglaló

A dolgozatom témája a csatorna és modell alapú hangtér-reprodukciós módszerek közötti kompatibilitás vizsgálata. A két rendszertechnika közötti átjárást a kevert források azonosítása és szétválasztása jelenti, amely a digitális jelfeldolgozásnak jelenleg is kutatás alatt álló tématerülete. A bevezetés után bemutatom a manapság használatos sokcsatornás hangrendszereket és a nehézségeket a források szeparálására nézve. Továbbá ismertetek néhány eljárást, amely alkalmas a feladat végrehajtására.

Részletesen kifejtem az Independent Component Analysis algoritmust, annak előnyeit és hátrányait, valamint megvalósíthatóságát és alkalmazhatóságának feltételeit különféle források használata esetén. Megvizsgálom, hogy milyen előfeldolgozási módszerekkel növelhető a hatékonysága, valamint kifejtem az egyes esetekben a vele kapcsolatos várakozásaimat.

Ezt követően a már implementált algoritmus kimeneteleit hasonlítom össze a különböző jelfeldolgozási alkalmazások esetén, és értékelem az eredményeket az előzetes remények figyelembevételével. A dolgozatom összegzés zárja, amelyben összefoglalom az eredményeket és kitekintést nyújtok a továbbfejlesztési lehetőségek, jövőbeli célok felé.

Abstract

The topic of my thesis is the investigation of the compatibility of channel based and model based sound field reproduction systems. The connection between the two systems technology is the identification and separation of mixed sources, which is still an active research field of digital signal processing. After the introduction I present today's multi-channel sound systems and the difficulties of the source separation. Additionally I present some methods, which are able to solute the problem.

Afterwards I expand particularly the Independent Component Analysis algorithm, its advantages and disadvantages, as well as practicability and conditions of applicability using different sources. I will examine how pre-processing methods can increase efficiency, and explicate my expectations with it.

Then I compare the issues of the implemented algorithms for different signal processing applications and evaluate the results attending to the previous hopes. The paper is closed by summary, where I summarize the results and describe further development opportunities.

1 Bevezetés

Napjainkban folyamatos fejlesztéseket érünk el a különféle hangtechnikai eszközök és eljárások területén, amelyek célja a még teljesebb és még több élvezetet nyújtó hangélmény létrehozása. A monó hanghoz képest hatalmas előrelépést jelentett a sztereó bevezetése, amely még mindig a legjobban elterjedt technika. Mára már otthonainkban is elérhetővé váltak a sokcsatornás hangrendszerek, amelyek sokáig csak a filmszínházak különlegességei voltak. A digitális technikának köszönhetően a stúdiókban egyre komplexebb keverések és hanghatások érhetők el, és már házilag is elfogadható minőségű mixeket állíthatunk elő. Ezek célja a minél élet hűbb hangélmény elérése, a fizikai valóság pontosabb leírása, illetve olyan hanghatás keltése, amely lenyűgözi a hallgatót, függetlenül attól, hogy ez a világunkban megtalálható vagy sem.

Azonban ezekben a technológiákban közös, hogy a térhatást csak egyetlen pontban, az úgy nevezett sweet spot-ban képesek elérni. Ezzel szemben más technológiák, például a Wave Field Synthesis (WFS) ezt a pontot kiterjeszti egy nagyobb térrészre, így képes a hangtér teljes, fizikai reprodukálására, persze bizonyos korlátozásokkal. A WFS technológia a Huygens-Fresnel elven alapul, amely kimondja, hogy valamennyi hullámfront előállítható gömbhullámok szuperpozíciójaként. Ez a gyakorlatban egy vonalmentén elhelyezett hangszóró sokaságot jelent, ahol a hangszórók szolgálnak a gömbhullámok forrásaiként. A két rendszertechnológia között az is eltérés, hogy még egy sztereofon rendszer esetén a csatornák jeleit tároljuk, addig a WFS esetén a virtuális források jelei kerülnek tárolásra. Így WFS alkalmazása során az adott hangszóró elrendezésre valós időben tudunk renderelni. Ekkor megjelenik a két hangrendszer közötti kompatibilitás vizsgálata iránti igény. Egy már meglévő kevert hanganyagot, csak akkor tudunk egy WFS rendszeren kielégítően megszólaltatni, ha a kevert audio anyagból ki tudjuk nyerni a virtuális forrásokat, sztereofon és WFS között nem elegendő egy lineáris leképzés. Tehát a valódi problémánk a kompatibilitás vizsgálat során a források szétválasztása, azaz a Sound Source Separation (SSS). Több SSS metódus leírása és véleményezése található meg a szakirodalomban, például Independent Component Analysis (ICA), Panning Index Window (PIW), Multi-Level Thresholding Separation (MuLeTS), illetve Azimuth Discrimination and Resynthesis (ADRes).

A forrás szeparálás nehézségeket rejt magában, hiszen egy kész audiomix számos forrást tartalmazhat és számos előállítási módon készülhetett, így a nem egységes

hanganyagokat különbözőképpen kell feldolgozni és a szétválasztást elvégezni. A konkrét eljárások pedig csak bizonyos feltételek esetén működnek elfogadható minőségben, ezért korlátozottak a lehetőségek.

2 Hangtér-reprodukciós rendszerek

2.1 Bevezetés

2.2 Surround rendszerek

2.3 Wave Field Synthesis

3 Independent Component Analysis

3.1 Elméleti áttekintés

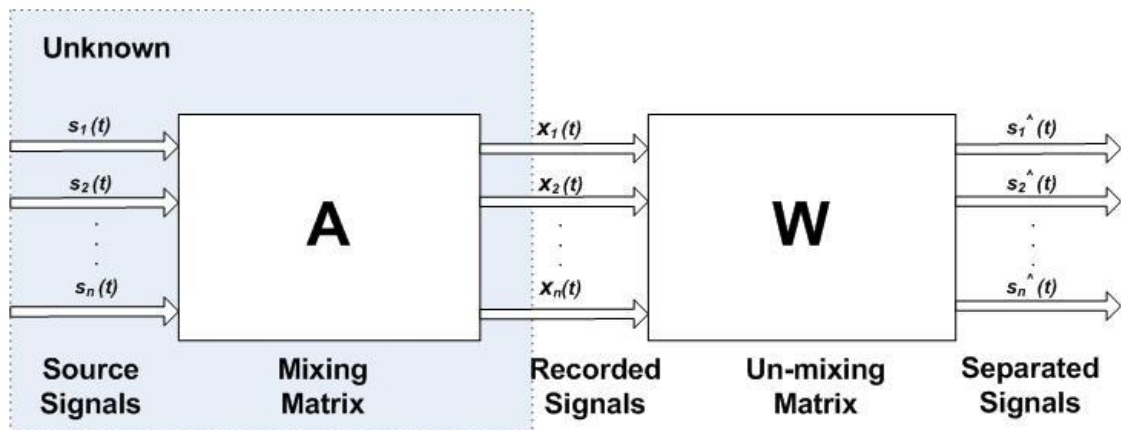
Az Independent Component Analysis (ICA), azaz a független komponens analízis általánosan olyan eljárás, amely statisztikai mintafolyamokban elrejtve lévő adat komponenseket nyer ki. Az algoritmus véletlen jeleken, időfüggvényeken történő feldolgozást hajt végre, amely eredményként annyi független komponenset talál, ahány bemeneti adatvektor áll rendelkezésre. Tehát ez a feldolgozás a pontosan-meghatározott SSS probléma esetén tud eredményesen működni.

Az algoritmus a hangtechnikán kívül többfajta jelfeldolgozási területen is használható, például képfeldolgozás (orvosi diagnosztika) során, gazdasági folyamatok adatsorainak elemzése stb.

A probléma a következőképpen fogalmazható meg:

$$s' = W \cdot x = W \cdot A \cdot s \tag{3.1}$$

Az A úgynevezett keverési mátrix ismerete nélkül W meghatározása pusztán az x jelekből, majd W segítségével s' előállítására.



3.1. ábra: A független források lineáris kombinációja, illetve a szétválasztás sematikus ábrája

Az analízis alapja a centrális határeloszlás-tétel (CHT), azaz nagyszámú független változó összegének eloszlása a normális eloszláshoz tart, tehát egy-egy csatorna jelének, a mixnek, eloszlása tehát Gauss közelibb, mint az egyes forrásoké külön-külön. A CHT azt jelenti, hogy a mix mindenképp "gauss-ibb", mint az egyes források időfüggvényei.

Így a cél egy olyan eljárás készítése, amely során az adatsorok normális eloszlás mértéke csökkenthető, azaz távolodás a normális eloszlástól.

Gauss-eloszlás mértékét kifejező matematikai fogalmak a kurtózis:

$$kurt(x) = E\{x^4\} - 3(E\{x^2\})^2, \quad (3.2)$$

illetve a negentrópia:

$$J(y) = H(y_G) - H(y) \quad (3.3)$$

Mindkettő értéke 0 Gauss eloszlás esetén. Valós, kevert adatsor esetén természetesen ezek az értékek 0-hoz tartanak. Így a feladat a negentrópia maximalizálása.

Az ICA lépései:

jel előfeldolgozás:

centralizálás

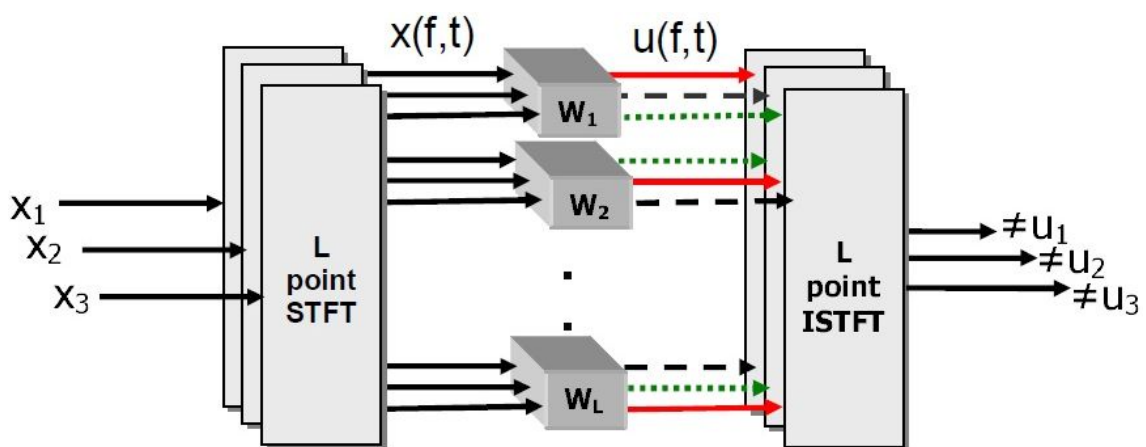
fehérítés: Karhunen-Loeve transzformációval dekorrelálás

negentrópia:

negentrópia kifejezésének közelítése (mivel nem ismerjük a tényleges sűrűségfüggvényeket ezért közelíteni kell) nem-lineáris függvénnyel, ezután ennek a közelítésnek numerikus minimalizálása Newton-Raphson iterációval

eredményként belekonvergálunk a kívánt W mátrixba.

ICA végrehajtása a jelek spektrumán, frekvenciánként külön-külön elvégezve.



3.2. ábra: ICA végrehajtása frekvenciatartományban

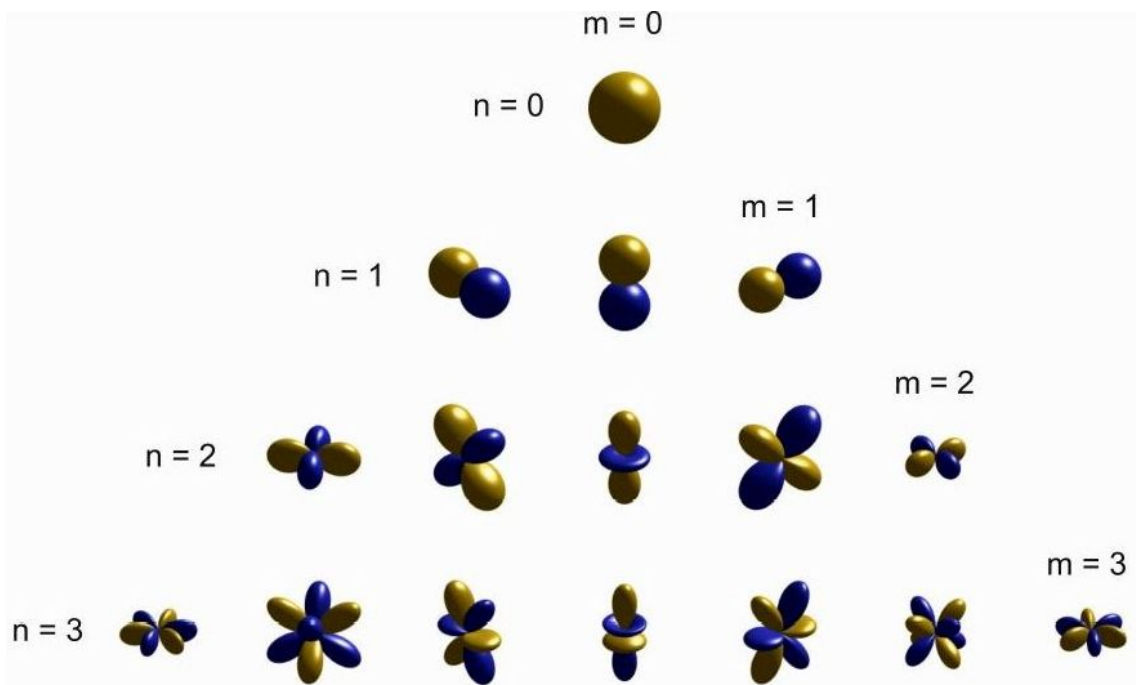
3.2 Skálázási probléma

3.3 Permutációs probléma

4 Spherical Harmonic Transform

A Spherical Harmonic Transform (SHT), azaz a gömbi harmonikus transzformáció, amely egy hangteret adott környezetben a 4.1. ábrán látható SHT bázisok szerint oszt fel komponensekre. Ez gyakorlatilag egy térbeli irányítottságot, iránykarakterisztikát jelent, mely azt mutatja meg, hogy bizonyos térbeli irányokból milyen hanghullámok érkeznek. Nyalábformálás, amely térbeli szelektivitást valósít meg.

Függőleges irányú függést állandónak tekintve 2D-ben ez cirkuláris Fourier-transzformálást jelent a bementi jelen.



4.1. ábra: SHT bázisok

5 Plane Wave Expansion

A Plane Wave Expansion (PWE), azaz a síkhullám sorfejtés során egy adott pont körül a hangteret síkhullámok összegére lehet bontani. Általános esetben az alábbi formula írja le:

$$\psi_{in}(k, \mathbf{r}) = \frac{1}{4\pi} \int_S e^{i\mathbf{k}\cdot\mathbf{r}} \mu_{in}(k, \mathbf{s}) dS(\mathbf{s}), \quad (5.1)$$

azaz a síkhullámok különböző súlyozó tényezőkkel összegezve megadja egy adott pont \mathbf{r} sugarú környezetében a hangteret. Az együtthatók a

$$\mu_{in}(\mathbf{s}) = \sum_{n=0}^{\infty} \sum_{m=-n}^n i^{-n} A_n^m Y_n^m(\mathbf{s}) \quad (5.2)$$

kifejezés írja le, ahol A amplitúdó értékeket, Y pedig a gömbi harmonikus bázisfüggvényeket jelöli.

A források által létrehozott összegtér lokálisan, egy adott pontban síkhullámnak tekinthetjük, így esetünkben is van ráció a sweet spotban előálló tér sorfejtésében. Az 5.1. ábrán egy szimulált pannelt sztereó forrás látható. A 0,0 koordinátájú pontban ez tényleg egy adott irányú síkhullámnak tekinthető.

Mivel a hangforrások egy síkban helyezkednek el a lehallgatási pozícióval, ezért elegendő 2D-ban folytatni a vizsgálatot és konstansnak vehetjük a harmadik koordinátát. 2D-ben a fenti formulák az alábbival közelíthető:

$$P(\mathbf{x}, \omega) \approx \frac{1}{2\pi} \int_{-\pi}^{\pi} \bar{P}(\phi_{pw}, \mathbf{x}_c, \omega) e^{-j\frac{\omega}{c}(\mathbf{x}'|\mathbf{n}_{pw})} d\phi_{pw}, \quad (5.3)$$

ahol a síkhullám sorfejtés együtthatói

$$\bar{P}(\phi_{pw}, \mathbf{x}_c, \omega) \approx \sum_{\mu=M}^M j^{\mu} \check{P}_{\mu}(\mathbf{x}_c, \omega) e^{j\mu\phi_{pw}}, \quad (5.4)$$

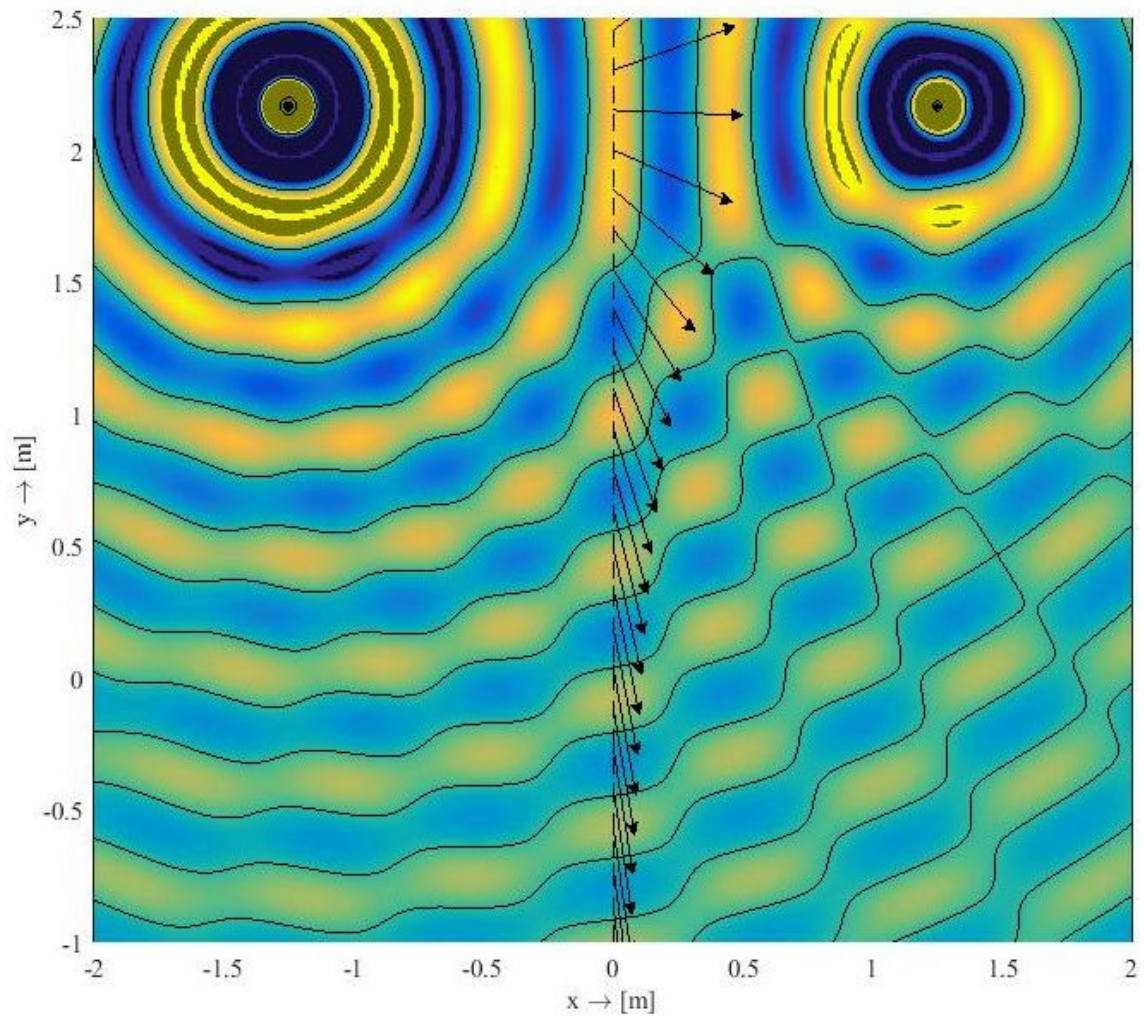
ahol továbbá a $\check{P}_{\mu}(\mathbf{x}_c, \omega)$ az úgynevezett *circular expansion* együtthatók, melyek

$$\check{P}_{\mu}(\mathbf{x}_c, \omega) = e^{-j\frac{\omega}{c}(\mathbf{x}_c|\mathbf{n}_s)} j^{-\mu} e^{-j\mu\phi_s} \quad (5.5)$$

síkhullám források esetében, illetve

$$\check{P}_{\mu}(\mathbf{x}_c, \omega) = \frac{j^{|\mu|-\mu}}{4\pi} \left(-j\frac{\omega}{c}\right) h_{|\mu|}^{(2)}\left(\frac{\omega}{c}\rho'_s\right) e^{-j\mu\phi'_s} \quad (5.6)$$

gömbhullám források esetében.



5.1. ábra: Sztereó forrás tere pannelt forrás esetén

6 Eredmények

7 Összefoglalás

Irodalomjegyzék

[1]