

Irányérzékelésre alkalmas jelfeldolgozó algoritmusok összegyűjtése és értékelése

Azonosító: Rez-1701-DronDet-M.2.2_Jelentes_190930

2019.09.30.

Tartalomjegyzék

1. Bevezetés	2
2. Irányérzékelő algoritmusok	3
2.1. Bevezetés	3
2.2. Az irányérzékelés alapfeladata	3
2.3. Alapfogalmak	4
2.3.1. A terjedés időtartománybeli modellje	4
2.3.2. Frekvenciatartománybeli leírás	4
2.3.3. Virtuális forráspozíciók: az akusztikus vászon	4
2.3.4. Végtelen távolságú vászon	5
2.3.5. Inverz terjedési vektorok	5
2.4. Hagyományos nyalábformálás	5
2.4.1. A mikrofontömb kovarianciamátrixa	6
2.4.2. A pontszórásfüggvény	6
2.4.3. Hagyományos nyalábformálás korrelálatlan pontforrások esetén	6
2.5. A MUSIC algoritmus	6
2.6. A CLEAN algoritmus	7
2.6.1. CLEAN-PSF	7
2.6.2. CLEAN-SC	9
2.7. A DAMAS algoritmus	10
2.8. Compressive Beamforming algoritmusok	11
2.8.1. Compressive Beamforming zajmentes esetben	11
2.8.2. Compressive Beamforming zajos esetben	12
3. Szimulációs esettanulmányok	13
3.1. Kétdimenziós végtelen távolságú nyalábformálás	13
3.1.1. Hagyományos nyalábformálás	13
3.1.2. MUSIC algoritmus	15
3.1.3. CLEAN algoritmus	16
3.1.4. A Compressive Beamforming algoritmus	16
3.2. Háromdimenziós végtelen távolságú nyalábformálás	20
3.2.1. Hagyományos nyalábformálás	20
3.2.2. A MUSIC algoritmus	21
3.2.3. A CLEAN algoritmus	21
3.2.4. A Compressive Beamforming algoritmus	23
3.3. A MUSIC algoritmus zajtűrésének vizsgálata	23
4. Összefoglalás	25

1. fejezet

Bevezetés

Ez a tanulmány sokmikrofonos mérőrendszerek akusztikai forrásdetektálásra való alkalmazásával foglalkozik. A tanulmány célja olyan algoritmusok bemutatása és elvezetése, melyek segítségével sokcsatornás akusztikai mérésekből automatizált módon beazonosítható zajforrások pozíciója, a mikrofontömbhöz képest mért iránya illetve teljesítménye.

A sokcsatornás mikrofonrendszerek akusztikai források lokalizálására és kiemelésére történő használata alapvetően kétféle célt szolgál: ezek az akusztikai fókuszálás és a forráslokalizáció. Az akusztikai fókuszálás során a mikrofonrendszer egyes mikrofonjaira beérkező hangjeleket erősítjük illetve késleltetjük annak érdekében, hogy az adott irányból beérkező hangjeleket összeadva azok konstruktívan interferáljanak egymással. A hullámterjedés fázisviszonyainak köszönhetően ebben az esetben elvárhatjuk, hogy más irányokból érkező hangjelek „véletlenszerű” fázisban összegződnek, így gyengítik egymást. Ezen elv alapján a mikrofonrendszer alkalmas a fókuszálásra, vagyis egy előre meghatározott irányból érkező hangjel kiemelésére és más irányok elnyomására. Az akusztikai fókuszálás témakörével foglalkozott „Forgó légcsavaros eszközök észlelésére alkalmas akusztikai antennák működési elvei és műszaki paraméterei” tanulmányunk [4].

A mikrofonrendszerek alkalmazásának másik fő feladata a forráslokalizáció. Forráslokalizáció során a cél a térben jelen levő, ismeretlen pozíciójú és teljesítményű hangforrások irányának és teljesítményének megbecslése. Ez általában egy virtuális pontháló, az akusztikai vázson felvételével történik, és a forráslokalizáció algoritmus a vázson pontjaiban elhelyezkedő virtuális források közül választja ki azokat, melyek a leginkább összeegyeztethetőek a vett hangtérrel. A forráslokalizáció és a fókuszálás egymástól függetlenül végezhetőek, de tipikus automatizált alkalmazásokban a forráslokalizáció eredményeként forrásirányokat kapunk, majd ezen irányokra fókuszálva lehetőségünk nyílik azok jeleinek háttérzajból való kiemelésére, illetve az egyes források interferenciájának csökkentésére.

Jelen dokumentum forráslokalizációs algoritmusokat ismertet. A 2. fejezet gyűjti össze a leggyakrabban használt nyálábformáló algoritmusokat, és ismerteti azok működési elvét. Bemutatjuk a hagyományos nyálábformálást, a MUSIC algoritmust, a CLEAN és DAMAS képtisztító algoritmusokat, majd a *Compressive Sensing* elven alapuló Compressive Beamforming módszert. A 3. fejezet szimulációs példákkal elemzi az egyes algoritmusok alkalmazásának előnyeit és hátrányait, külön fókuszálva kétdimenziós, mikrofonos alkalmazásokra, illetve háromdimenziós nyálábformáló módszerekre. Az egyes algoritmusok zajmentes esetre elvégzett értékelése után kitérünk a leghatékonyabbnak talált algoritmus zajtűrésének vizsgálatára. A szimulációs vizsgálatok eredményeinek értékelését a 4. fejezet foglalja össze.

2. fejezet

Irányérzékelő algoritmusok

2.1. Bevezetés

Ez a fejezet az irányérzékelő algoritmusok elméleti alapjait foglalja össze. A 2.2. szakaszban megfogalmazzuk az irányérzékelés alapfeladatát, majd a 2.3. szakaszban olyan fogalmakat vezetünk be, melyek mindegyik, később tárgyalt algoritmus leírásának alapjául szolgálnak. Végül, a 2.4–2.8. szakaszok egy-egy irányérzékelésre alkalmas algoritmust mutatnak be.

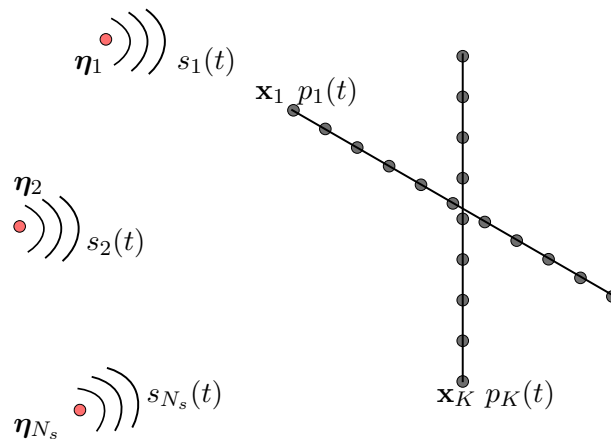
2.2. Az irányérzékelés alapfeladata

Az irányérzékelés alapfeladata az alábbi alakban fogalmazható meg:

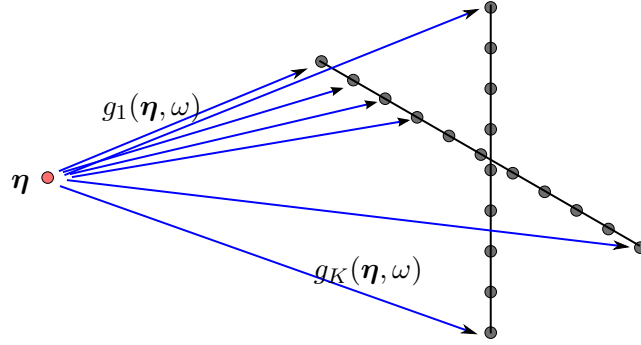
Adott N_s számú hangforrás, melyek az η_j pozíciókban helyezkednek el, ahol $j = 1, 2, \dots, N_s$, és az $s_j(t)$ jelet sugározzák, ahogy azt a 2.1. ábra mutatja. Adott továbbá egy K elemű mikrofontömb, melynek mikrofonjai az $\mathbf{x}_i$ pozíciókban helyezkednek el, ahol $i = 1, 2, \dots, K$.

Feladtunk, hogy becslést adjunk a források N_s számára, azok η_j elhelyezkedésére és S_j teljesítményére. Ehhez lehetőségünk van a mikrofonok $p_i(t)$ jeleinek mérésére.

Amennyiben a források számát és elhelyezkedését jól becsüljük, a mikrofontömb segítségével lehetőségünk nyílik az egyes források $s_j(t)$ jeleinek kiemelésére akusztikai fókuszálás útján.



2.1. ábra. Az irányérzékelés alapfeladata



2.2. ábra. Az η pozícióhoz és a mikrofontömbhöz tartozó $\mathbf{g}(\eta)$ terjedési vektor szemléltetése

2.3. Alapfogalmak

2.3.1. A terjedés időtartománybeli modellje

Az η pozícióban elhelyezkedő és $s(t)$ jelet sugárzó forrás hatására a mikrofontömbön az alábbi $\mathbf{p}(t)$ jel mérhető:

$$\mathbf{p}(t) = \{p_i(t)\} = \int_{-\infty}^{+\infty} \mathbf{g}(\eta, t - \tau) s(\tau) d\tau, \quad (2.1)$$

ahol $\mathbf{g}(\eta, t)$ az η forráspozíció és a mikrofontömb elemi közötti terjedést leíró terjedési impulzusválaszvektor. Homogén, végtelen, reflexiómentes hangtér esetén

$$\mathbf{g}(\eta, t) = \{g_i(\eta, t)\} = \left\{ \frac{\delta(t - r_i/c)}{4\pi r_i} \right\}, \quad (2.2)$$

ahol $r_i = |\mathbf{x}_i - \eta|$, c a hang terjedési sebessége, δ pedig a Dirac-delta eloszlás. A Dirac-delta argumentuma a terjedési késleltetést, a nevező pedig a terjedés során fellépő geometriai csillapítást írja le.

2.3.2. Frekvenciatartománybeli leírás

Frekvenciatartományban ($e^{j\omega t}$ időfüggést feltételezve), a vett mikrofonjelek komplex csúcserőértéke

$$\mathbf{p}(\omega) = \{p_i(\omega)\} = \mathbf{g}(\eta, \omega) s(\omega), \quad (2.3)$$

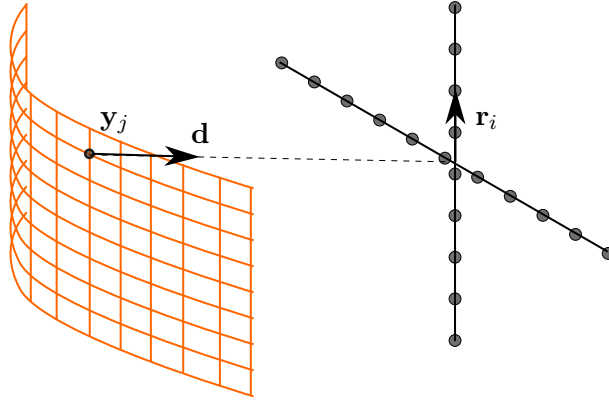
ahol homogén hangtérben a terjedési vektor alakja

$$\mathbf{g}(\eta_j, \omega) = \frac{e^{-jkr_i}}{4\pi r_i}, \quad (2.4)$$

$k = \omega/c$ pedig az akusztikai hullámszám. Itt az exponensben szereplő negatív fázistolás a késleltetést írja le, míg a nevező a geometriai csillapításért felelős. A terjedési vektort szemléletesen ábrázolja a 2.2. ábra.

2.3.3. Virtuális forráspozíciók: az akusztikus vászon

Az irányérzékelő algoritmusok nem képesek általában az η forráspozíciók pontos meghatározására, hanem egy előre megadott, N elemű $\mathbf{y}_j$ forráshalmaz elemei közül választják ki azokat a pozíciókat, melyek a leginkább összeegyeztethetőek a mikrofontömbös méréssel. Az $\mathbf{y}_j$ virtuális pozíciók összességét akusztikus vászonnak nevezzük. A vásznat a 2.3. ábra ábrázolja. A vászon j -edik eleméhez és a mikrofontömbhöz értelemszerűen a $\mathbf{g}_j = \mathbf{g}(\mathbf{y}_j, t)$ vagy $\mathbf{g}(\mathbf{y}_j, \omega)$ terjedési vektor tartozik.



2.3. ábra. Az akusztikus vászon

2.3.4. Végtelen távolságú vászon

Sok esetben – amennyiben a valós források mikrofontömbtől mért távolsága elegendően nagy –, a vásznat a mikrofontömbtől végtelen távolságra helyezzük el. Ez esetekben a vásznon levő pontok mikrofontömb felől érzékelt irányyszöge határozza meg a terjedési vektorokat, és azokban a távolságinformációt elhagyjuk. Végtelen távolságú vászon esetén az egyes mikrofonokba beérkező hanghullámjelek párhuzamos utak mentén, síkhullámként terjednek, és ez megkönnyíti a nyálábformáló algoritmusok implementálását. A frekvenciatartományban a terjedési vektor alakja ebben az esetben

$$\{g_i\} = e^{-j\mathbf{k}\mathbf{d}\cdot\mathbf{r}_i}, \quad (2.5)$$

ahol $\mathbf{d}$ a forrásból a mikrofontömb közepe (origó) felé irányuló egységvektor, $\mathbf{r}_i$ pedig az i -edik mikrofon középponttól mért helyvektora, ahogy azt a 2.3. ábra mutatja. Ebben az esetben a mikrofontömb közepéhez tartozik a zérus fáziskésleltetés.

2.3.5. Inverz terjedési vektorok

A későbbi algoritmusok leírásának megkönnyítésére bevezetjük az inverz terjedési vektor fogalmát. Az $\mathbf{y}_j$ virtuális forráspozícióhoz tartozó $\mathbf{h}_j$ inverz terjedési vektor elemei

$$\mathbf{h}_j = \{h_{ij}\} = \{g_{ij}^{-1}\} = \{4\pi r_{ij} e^{jkr_{ij}}\}, \quad (2.6)$$

vagy végtelen távolságú nyálábformálás esetén

$$\mathbf{h}_j = \{h_{ij}\} = \{g_{ij}^{-1}\} = e^{j\mathbf{k}\mathbf{d}\cdot\mathbf{r}_i}. \quad (2.7)$$

Ez az inverz terjedési vektor kompenzál a terjedési fázistolással és – véges távolságú terjedés esetén – az amplitúdócsökkenéssel is. Könnyen látszik, hogy

$$\frac{\mathbf{h}_j \cdot \mathbf{g}_j}{K} = 1. \quad (2.8)$$

2.4. Hagyományos nyálábformálás

A hagyományos nyálábformálás (*Conventional Beamforming*, CBF) a legegyszerűbb forrás-lokalizációs algoritmus.

Hagyományos nyálábformálás során a vászon minden egyes eleméhez előállítjuk a $\mathbf{h}_j$ inverz terjedési vektort, majd azokkal külön-külön kompenzáljuk a mikrofontömb vett $\mathbf{p}$ jelét. Az így adódó $\mathbf{h}_j \cdot \mathbf{p}$ skalárok abszolút négyzeteként adódik az S_j forrásteljesítmény-eloszlás:

$$S_j = \frac{|\mathbf{h}_j \cdot \mathbf{p}|^2}{K^2} = \frac{(\mathbf{h}_j^T \mathbf{p})(\mathbf{p}^H \mathbf{h}_j)}{K^2} \quad (2.9)$$

2.4.1. A mikrofontömb kovarianciamátrixa

A hagyományos nyálábformálás algoritmusának alternatív megfogalmazásához érdemes bevezetnünk a mikrofontömb $\mathbf{C}$ kovarianciamátrixának fogalmát:

$$\mathbf{C} = \mathbf{E} \mathbf{p} \mathbf{p}^H, \quad (2.10)$$

ahol $\mathbf{E}$ az időbeli várhatóértékképzés operátora. A kovarianciamátrix diagonálisában a mikrofontömb jeleinek teljesítményei szerepelnek. A kovarianciamátrix segítségével a becsült forráseloszlás alakja

$$S_j = \frac{\mathbf{h}_j^T \mathbf{C} \mathbf{h}_j}{K^2}. \quad (2.11)$$

2.4.2. A pontszórásfüggvény

A hagyományos nyálábformálás algoritmus segítségével könnyen definiálhatjuk a mikrofontömb PSF pontszórásfüggvényét (*Point Spread Function*). A pontszórásfüggvény a mikrofontömb által becsült S_j forrásteljesítményeloszlást adja meg abban az esetben, mikor a mikrofontömb bejövő jele az $\mathbf{y}_i$ pozícióban elhelyezkedő egységnyi teljesítményű pontforrás. Ebben az esetben a mikrofontömb jele

$$\mathbf{p} = \mathbf{g}_i s, \quad (2.12)$$

a kovarianciamátrix

$$\mathbf{C} = \mathbf{E} \mathbf{p} \mathbf{p}^H = \mathbf{g}_i \mathbf{E} |s|^2 \mathbf{g}_i^H = \mathbf{g}_i \mathbf{g}_i^H, \quad (2.13)$$

a pontszórásfüggvény pedig

$$\text{PSF}_{ji} = \frac{\mathbf{h}_j^T \mathbf{g}_i \mathbf{g}_i^H \mathbf{h}_j}{K^2}. \quad (2.14)$$

Könnyen belátható, hogy a pontszórásfüggvény értéke egységnyi, amennyiben $i = j$, vagyis a hagyományos nyálábformálás a vászon egy pontjában elhelyezkedő valós pontforrás teljesítményét helyesen adja vissza. Ugyanakkor az is látszik, hogy a hagyományos nyálábformálás a valódi pontforráspozícióhoz közeli vászonpozíciókra is az egységnyihez közeli értéket ad vissza.

2.4.3. Hagományos nyálábformálás korrelálatlan pontforrások esetén

Tételezzük fel, hogy a valódi forrásaink az $\mathbf{y}_j$ forráspozíciókban helyezkednek el, ahol $j = 1, 2, \dots, N_s$, és korrelálatlanok, vagyis $\mathbf{E} s_i s_k = \delta_{ik} \sigma_i^2$, ahol δ_{ik} a Kronecker-delta.

Ebben az esetben a mikrofontömb kovarianciamátrixa

$$\mathbf{C} = \mathbf{E} \left(\sum_i \mathbf{g}_i s_i \right) \left(\sum_i \mathbf{g}_i s_i \right)^H = \sum_i \sigma_i^2 \mathbf{g}_i \mathbf{g}_i^H, \quad (2.15)$$

a becsült forráseloszlás pedig

$$S_j = \sum_i \sigma_i^2 \text{PSF}_{ji}. \quad (2.16)$$

Vagyis korrelálatlan források esetén a hagyományos nyálábformálás a pontszórásfüggvények teljesítményekkel súlyozott összegét adja virtuális forráseloszlásként.

2.5. A MUSIC algoritmus

A MUSIC (*Multiple Signal Classification*) algoritmus lineáris algebrai eszköztárat alkalmaz a forrásbecslési probléma megoldásához.

Tételezzük fel, hogy N_s számú valós forrás jelét veszi mikrofontömbünk, vagyis a mikrofontömb nyomásjelének alakja

$$\mathbf{p}(\omega) = \sum_{j=1}^{N_s} \mathbf{g}(\boldsymbol{\eta}_j) s_j(\omega). \quad (2.17)$$

Ez azt jelenti, hogy a K elemű mikrofontömb vett jele a K -dimenziós tér egy N_s dimenziós alterét feszíti ki, melynek bázisát a $\mathbf{g}(\boldsymbol{\eta}_j)$ terjedési vektorok határozzák meg.

Képezzük a mikrofontömb kovarianciamátrixát:

$$\mathbf{C} = \mathbf{E} \mathbf{p} \mathbf{p}^H. \quad (2.18)$$

Tételezzük fel, hogy az s_j források korrelálatlanok, ekkor

$$\mathbf{C} = \sum_{j=1}^{N_s} \sigma_j^2 \mathbf{g}(\boldsymbol{\eta}_j) \mathbf{g}^H(\boldsymbol{\eta}_j), \quad (2.19)$$

aminek értelmében a kovarianciamátrix N_s diád összege. Ekkor rangja legfeljebb N_s lehet.

Mivel a kovarianciamátrix önadjungált, sajátértékei valósak, sajátvektorai pedig merőlegesek egymásra. Mivel a mátrix rangja legfeljebb N_s , K sajátértéke közül $K - N_s$ zérus kell, hogy legyen. Határozzuk meg a kovarianciamátrix $K - N_s$ legkisebb sajátértékhez tartozó sajátvektorait, jelölje őket ϕ_n , $n = 1, 2, \dots, K - N_s$. Ezek a sajátvektorok a $\mathbf{g}(\boldsymbol{\eta}_j)$ vektorokra merőleges $K - N_s$ dimenziós alteret feszítik ki, ezt nevezzük a kovarianciamátrix zajterének. Foglaljuk ezeket a sajátvektorokat egy $K \times (K - N_s)$ típusú mátrixba:

$$\boldsymbol{\Phi} = [\phi_n]. \quad (2.20)$$

Ezek után az iránybecslés a következő módon történik. Meghatározzuk minden egyes $\mathbf{y}_j$ virtuális forráspozícióhoz az onnan ható $\mathbf{g}_j$ terjedési vektor zajterre eső vetületét, és ezen vetület hossz négyzetének reciprokával jellemezzük az adott forráspozíciót. A vászon képének definíciója tehát

$$S_j = \frac{\|\mathbf{g}_j\|^2}{\|\mathbf{g}_j^T \boldsymbol{\Phi}\|^2}. \quad (2.21)$$

A MUSIC algoritmussal kapcsolatban érdemes megjegyezni, hogy ennek bemenő paramétere a források becsült N_s száma. Ugyanakkor, a kovarianciamátrix sajátértékei alapján az algoritmus képes a források számának megbecslésére is. Ehhez azt kell megvizsgálni, hogy hány sajátérték esik adott küszöbszint fölé. Ehhez kihasználhatjuk azt a tételt, miszerint a mátrix nyoma (diagonalizálásának összege) megegyezik sajátértékeinek összegével. A források számához használhatjuk tehát az alábbi szabályt:

$$\frac{\sum_{n=1}^{K-N_s} \lambda_n}{\sum_{i=1}^K |p_i|^2} < \varepsilon, \quad (2.22)$$

ahol ε egy előre meghatározott relatív küszöbszint. Így a küszöbszint alatti teljesítményekhez tartozó sajátvektorok fogják kifeszíteni a zajteret.

2.6. A CLEAN algoritmus

2.6.1. CLEAN-PSF

A CLEAN-PSF (*CLEAN based on Point Spread Function*) egy iteratív képtisztító algoritmus, mely az ismert pontszórás-függvényekkel történő dekonvolúciót közelíti [3]. Az eljárás a hagyományos nyalábformálásból adódó, a (2.11) összefüggéssel számított „elmosódott” képből (*dirty image*) indul ki. Emellett az eljárás létrehoz egy kezdetben üres „tisztított” képet (*clean image*) is. Minden iteráció során a kép a következő módon frissül:

1. A *dirty image*-en megkeressük az amplitúdótérkép maximumához tartozó $i_{\max}$ virtuális forráspozíciót.
2. Kiszámítjuk az ehhez a pozícióhoz tartozó PSF_{ji} pontszórás-függvényt.
3. A *dirty image*-ből kivonjuk a kiszámított pontszórás-függvényt, a *clean image*-en pedig elhelyezünk az $i_{\max}$ pozícióban egy olyan nyalábot, melynek melléknyalábjai már nincsenek. Szélsőséges esetben az elhelyezett nyaláb akár egyetlen pont is lehet.
4. Elölről folytatjuk a ciklust, amíg nem érjük el valamelyik konvergenciakritériumot. Konvergenciakritérium lehet az iterációk maximális számának elérése, a *dirty image*-en a teljes teljesítmény adott referenciaszint alá csökkenése, a pontszórásfüggvény kivonására az összes teljesítmény adott szint alá csökkenése vagy e kritériumok valamilyen kombinációja.

Az algoritmus matematikai megfogalmazása az alábbi:

1. Legyen a $\mathbf{D}$ mátrix kezdetben egyenlő a mikrofontömb $\mathbf{C}$ kovarianciamátrixával. Jelölje $\mathbf{Q}$ a becsült teljesítményeloszlást, melyet zérus értékről indítunk. Jelölje továbbá i az iterációs számot.

$$i \leftarrow 0 \quad (2.23)$$

$$\mathbf{D}^{(i)} \leftarrow \mathbf{C} \quad (2.24)$$

$$\mathbf{Q} \leftarrow \mathbf{0} \quad (2.25)$$

2. Opcionálisan vonjuk ki a $\mathbf{D}$ mátrixból annak diagonálisát. Ezzel a mikrofonok autokorrelációját eldobjuk, mely nem hordoz információt a vett jelek összefüggéseiről. A redukált mátrixot jelöljük $\bar{\mathbf{D}}$ -vel:

$$\bar{\mathbf{D}}^{(i)} \leftarrow \mathbf{D}^{(i)} - \text{diag} \mathbf{D}^{(i)} \quad (2.26)$$

3. Állítsuk elő a becsült forráseloszlást a (2.11) formula alkalmazásával:

$$S_j^{(i)} = \frac{\mathbf{h}_j^T \bar{\mathbf{D}} \mathbf{h}_j}{K^2} \quad (2.27)$$

4. Keressük meg a képen a $P_{\max}$ maximális amplitúdót és ennek indexét:

$$P_{\max} = \max_j S_j \quad (2.28)$$

$$j_{\max} = \arg \max_j S_j \quad (2.29)$$

5. Frissítsük a $\mathbf{D}$ mátrixot $P_{\max}$, $j_{\max}$ és a $j_{\max}$ indexhez tartozó $\mathbf{g}_{\max}$ terjedési vektor felhasználásával:

$$\mathbf{D}^{(i+1)} \leftarrow \mathbf{D}^{(i)} - \alpha P_{\max} \mathbf{g}_{\max} \mathbf{g}_{\max}^H \quad (2.30)$$

Az α együttható egy relaxációs együttható, mellyel az szabályozható, hogy a megtalált maximális teljesítmény mekkora részét távolítsuk el a képről egy lépésben. Amennyiben $\alpha = 1$, úgy a teljes teljesítményt egy lépésben kivonjuk. Utóbbi azonban nem feltétlenül szerencsés választás, mivel ekkor hibás eredményt kapunk, ha a megtalált maximum egy fő- és egy vagy több oldalnyaláb összegződésével állt elő. Ezért a gyakorlatban az $\alpha \approx 0.6 \dots 0.8$ közötti választás a célszerű. Ha α -t ennél kisebbre választjuk, akkor túl sok iteráció szükséges a képtisztításhoz. A fentihez hasonlóan a $\mathbf{D}$ mátrix diagonálisának kinullázása opcionális lépés.

6. Vizsgáljuk meg a leállási kritériumot. Egy lehetséges leállási feltétel a $\mathbf{D}$ mátrix elemeinek vizsgálata. Ha a $\mathbf{D}$ mátrix elemeinek abszolútérték-összege nagyobb, mint az előző iterációban, vagyis

$$\sum_{m,n} |D_{mn}^{(i+1)}| > \sum_{m,n} |D_{mn}^{(i)}| \quad (2.31)$$

úgy az 5. lépésben nem vettünk ki teljesítményt a képből. Ebben az esetben nem folytatjuk tovább az iterációt.

7. Ellenkező esetben a $\mathbf{Q}$ kimeneti képen elhelyezünk egy $\alpha P_{\max}$ erősségű nyalábot, majd a következő $i + 1$ -edik iterációt a 3. lépéstől folytatjuk. Így egyetlen pontba sűrítettük az eredetileg véges szélességű nyalábot. Más stratégia is elképzelhető, például véges szélességű, oldalnyaláboktól mentes főnyalábot helyezünk el a $j_{\max}$ pozíció környékén. A teljesítmény korrekt értéken tartása érdekében az elhelyezett új nyaláb összteljesítményének a kivont $\alpha P_{\max}$ teljesítménnyel kell megegyeznie.

2.6.2. CLEAN-SC

A CLEAN egy változata a CLEAN-SC-nek (*CLEAN based on spatial source coherence*) nevezett algoritmus. A fent bemutatott CLEAN-PSF-hez hasonlóan ez az algoritmus is iteratív módszerrel közelíti a dekonvolúciós probléma megoldását. Ahogy arra a módszer neve is utal, az előbbi változathoz képest a különbséget az jelenti, hogy míg a CLEAN-PSF esetén minden iterációban egy virtuális pontforrás által létrehozott nyalábot távolítunk el az amplitúdótérképről, addig a CLEAN-SC esetében a forrásmodellünket nem feltétlenül tekintjük többé pontszerűnek [5]. Ehelyett a (2.30) lépésben egy olyan virtuális forrás által előállított spektrális kereszt-korrelációs mátrixot vonunk ki a $\bar{\mathbf{D}}^{(i)}$ mátrixból, amelyet kiterjedt forrás is létrehozhat. Ezzel relaxáljuk a pontforrás-feltételezést, mellyel javíthatjuk a dekonvolúciós képtisztítás eredményét térben kiterjedt források esetén.

Az iterációs lépések a fentiekkel egyeznek meg, kivéve a (2.30) lépést, mely így módosul:

$$\mathbf{D}^{(i+1)} \leftarrow \mathbf{D}^{(i)} - \alpha P_{\max} \mathbf{h} \mathbf{h}^H, \quad (2.32)$$

ahol a $\mathbf{h}$ vektor (módosított terjedési vektor, mely akár térben kiterjedt forráshoz is tartozhat) iteratív módszerrel határozható meg a következő módon.

1. Legyen kezdetben $\mathbf{h} = \mathbf{g}_{\max}$. Jelölje $\mathbf{w}_{\max}$ a $\mathbf{g}_{\max} / |\mathbf{g}_{\max}|$ normalizált terjedési vektort.
2. Frissítsük a $\mathbf{h}$ vektort a következő módon:

$$\mathbf{h} \leftarrow \frac{1}{\sqrt{1 + \mathbf{w}_{\max} \mathbf{H} \mathbf{w}_{\max}^H}} \left(\frac{\mathbf{w}_{\max} \bar{\mathbf{D}}^{(i)}}{P_{\max}} + \mathbf{w}_{\max} \mathbf{H} \right), \quad (2.33)$$

ahol a $\mathbf{H} = \langle \text{diag}(\mathbf{h} \mathbf{h}^H) \rangle$, vagyis a $\mathbf{h} \mathbf{h}^H$ szorzattal kapott mátrixból a főátlón kívül elemek elhagyásával kapott mátrix.

3. A 2. pontban szereplő iteráció gyorsan konvergál, tipikusan 2–3 lépéssel már elérjük a jó megoldást.

A MUSIC algoritmushoz hasonlóan a CLEAN algoritmusok is akkor működnek jól, ha a különböző források nem koherensek (a CLEAN-SC változat egy térben kiterjedt forrást enged meg, nem a térben elkülönülő egymással koherens forrásokat). Ez a viselkedés abból is látható, hogy a MUSIC-hoz hasonlóan a spektrális kereszt-korrelációs mátrixot dolgozzák fel ezek az algoritmusok is.

2.7. A DAMAS algoritmus

A DAMAS (*Deconvolution approach for the mapping of acoustic sources*) algoritmus egy dekonvolúciós problémát old meg. A (2.16 formula értelmében a hagyományos nyalábformálás kameraképe a források teljesítményeinek és a mikrofotómb pontszórásfüggvényének konvolúciójaként alakul. A dekonvolúciós feladat a kamerakép és a pontszórásfüggvények ismeretében a forráseloszlás meghatározása. A DAMAS algoritmus a dekonvolúciós problémát az előzőektől eltérő módon, de a CLEAN algoritmusokhoz hasonlóan szintén iteratív módon oldja meg [1]. A dekonvolúció közvetlenül nem végezhető el, mivel a rendszer erősen alulhatározott, rosszul kondicionált, így a DAMAS plusz kényszerfeltételek bevezetésével végzi el a dekonvolúciót. Az újonnan bevezetett kényszerfeltétel szerint a $\mathbf{P}$ forráseloszlás csak nemnegatív értékű lehet a térben, mivel a források nem sugározhatnak negatív teljesítményt. Így a matematikailag megfogalmazható kényszerhez jól megfogható fizikai jelentés is társul.

Megoldandó tehát az

$$\mathbf{S} = \mathbf{A}\mathbf{P} \quad (2.34)$$

egyenletrendszer, melyben az egyszerűség kedvéért a korábbi $\mathbf{A}_{\text{PSF}}$ mátrixot az alsó index elhagyásával $\mathbf{A}$ -val jelöltük. Ismert az $\mathbf{S}$ vektor, és meghatározandó a $\mathbf{P}$ vektor, a $\mathbf{P} = \mathbf{A}^{-1}\mathbf{S}$ művelet viszont nem végezhető el, mivel $\mathbf{A}$ rosszul kondicionált, vagyis $\text{rank } \mathbf{A} \ll N$. A (2.34) összefüggés mellett még feltételezzük, hogy $\mathbf{P}$ csak nemnegatív értékeket tartalmazhat, vagyis

$$P_j \geq 0, \quad j = 1, 2, \dots, N. \quad (2.35)$$

A (2.34) egyenlet n -edik egyenletét kibontva írhatjuk:

$$A_{n1}P_1 + A_{n2}P_2 + \dots + A_{nn}P_n + \dots + A_{nN}P_N = S_n \quad (2.36)$$

Az $\mathbf{A}$ mátrixot alkotó pontszórásfüggvények tulajdonságaiból következően feltételezhetjük, hogy $A_{nn} = 1$, így az egyenletet a következő alakba rendezhetjük:

$$P_n = S_n - \sum_{n'=1}^{n-1} A_{nn'}P_{n'} - \sum_{n'=n+1}^N A_{nn'}P_{n'}. \quad (2.37)$$

A (2.37) összefüggést használjuk a dekonvolúcióhoz, bevezetve a $\mathbf{P}^{(i)}$ i -edik iterációban adódó megoldást:

$$P_n^{(i)} = S_n - \sum_{n'=1}^{n-1} A_{nn'}P_{n'}^{(i)} - \sum_{n'=n+1}^N A_{nn'}P_{n'}^{(i-1)}, \quad n = 1, 2, \dots, N \quad (2.38)$$

Minden lépésben alkalmazzuk még a nem-negatív teljesítmény feltételét is, vagyis, amennyiben a fenti formulával negatív $P_n^{(i)}$ érték adódik, azt zérusra állítjuk. Látható, hogy az $n = 2, 3, \dots, N-1$ esetekben az addig kiszámított $P_n^{(i)}$ értékeket is felhasználjuk már a számítás közben. Az iteratív megoldás konvergenciája gyorsítható, ha a (2.38) iterációt úgy valósítjuk meg, hogy minden iteráció során megfordítjuk a léptetés irányát. Eszerint az első iterációban $n = 1, 2, \dots, N$ sorrendben haladunk, a második iterációban viszont az $n = N, N-1, \dots, 1$ sorrendet választjuk és így tovább.

A DAMAS algoritmus előnye robusztussága: ha kellően sok iterációt alkalmazunk, a kép szépen kitisztul. Az algoritmus hátránya viszont, hogy általában nehezen becsülhető meg előre, hogy mennyi iterációra lesz szükség. Emellett az iteratív megoldás meglehetősen számításigényes, így valós időben korlátozottan alkalmazható. Érdekes még megjegyezni, hogy az $\mathbf{A} \in \mathbb{R}^{N \times N}$ mátrix mérete (és így a tárolásához szükséges kapacitás) a virtuális forráspozíciók N számával négyzetesen arányosan nő, ami az algoritmus memóriaigényét is növeli. Alternatívaként megtehető, hogy mivel a mátrixnak egyszerre csak egy sorára van szükségünk, így a mátrixot soronként (vagy részenként állítjuk elő). Ez megtehető, mivel a pontszórásfüggvények számítása nem túlságosan számításigényes feladat.

2.8. Compressive Beamforming algoritmusok

2.8.1. Compressive Beamforming zajmentes esetben

A *Compressive Beamforming* a *Compressive Sensing* (CS) eljárásan alapul. Ez a módszer a forráslokalizációt egy konvex optimalizálási problémára visszavezetve oldja meg, azonban a forrásnak meg kell felelnie egy ritkasági kényszerfeltételnek annak érdekében, hogy a módszer pontos becslést adhasson [2, 6].

A módszer egyszerű matematikai megfogalmazásához bevezetjük az alábbi jelöléseket: Jelölje $\mathbf{x}$ a forráseloszlásvektort, mely az egyes források komplex amplitúdóit tartalmazza (a frekvenciatarományban). Jelölje továbbá $\mathbf{y}$ a mikroföntömb egyes mikrofonjain megjelenő komplex amplitúdókat. Jelölje végül az $\mathbf{A}$ mátrix a két vektor közti lineáris leképezést, vagyis

$$\mathbf{y} = \mathbf{A}\mathbf{x}. \quad (2.39)$$

Jelen esetben az $\mathbf{A}$ mátrix oszlopai az egyes vászonpontokhoz tartozó terjedési vektorok.

A ritkasági kényszerfeltétel azt jelenti, hogy kevés forrásunk van a vászon pontjainak számához képest (amik elég sűrűn helyezkednek el ahhoz, hogy ne adódjon probléma a források rossz illeszthetősége miatt). Ezt matematikailag úgy is megfogalmazhatjuk, hogy a forrást leíró vektornak (vagy mátrixnak) kevés nem-nulla eleme van. Ha ez a feltétel teljesül, akkor a célunk az, hogy a forrásra végzett becslésünknek minél kevesebb zérustól különböző értéke legyen, miközben $\mathbf{A}\mathbf{x} = \mathbf{y}$ teljesül:

$$\min_{\mathbf{x} \in \mathbb{C}^N} \|\mathbf{x}\|_0 \quad \text{úgy, hogy} \quad \mathbf{A}\mathbf{x} - \mathbf{y} = \mathbf{0}. \quad (2.40)$$

ahol $\mathbf{x}$ a becslésünk a forráseloszlásra, $\mathbf{y}$ a vett jelek mátrixa, $\mathbf{A}$ az előre meghatározott és fix ($\mathbf{x}$ -től és $\mathbf{y}$ -től nem függő) érzékelési mátrix, $\|\mathbf{x}\|_0$ pedig az $\mathbf{x}$ vektor $\mathbb{L}_0$ -normája. Egy adott vektor $\mathbb{L}_p$ -normája általánosan

$$\|\mathbf{x}\|_p = \left(\sum_{i=1}^n |x_i|^p \right)^{\frac{1}{p}}. \quad (2.41)$$

ami a $p = 0$ esetben éppen a nemzérus elemek számát adja meg:

$$\|\mathbf{x}\|_0 = \sum_{i=1}^n 1_{x_i \neq 0}. \quad (2.42)$$

Az $\mathbf{A}$ érzékelési mátrix a vászon pontjai és a mikroföntömb elemei közti kapcsolatot írja le, oszlopai tehát az egyes vászonpontokhoz tartozó terjedési vektorokat tartalmazzák:

$$\mathbf{A} = [\mathbf{g}_1 \quad \mathbf{g}_2 \quad \dots \quad \mathbf{g}_N]. \quad (2.43)$$

A (2.40) probléma megoldására hatékony algoritmus nem létezik. Azonban, ha az $\mathbb{L}_0$ -norma szerinti feltételt $\mathbb{L}_1$ normára relaxáljuk, akkor jóval könnyebben megoldható problémához jutunk:

$$\min_{\mathbf{x} \in \mathbb{C}^N} \|\mathbf{x}\|_1 \quad \text{úgy, hogy} \quad \mathbf{A}\mathbf{x} - \mathbf{y} = \mathbf{0}. \quad (2.44)$$

A *Compressive Sensing* módszerek területén nagy áttörés volt annak bizonyítása, hogy megfelelően ritka forrás esetén (jóval kevesebb valós forrásunk van, mint ahány virtuális forráspozíció), (2.44) ekvivalens a (2.40) problémával, és felhasználható nyálábformáló algoritmusokban.

2.8.2. Compressive Beamforming zajos esetben

A gyakorlatban a mikrofonok által vet jelre zaj is rakódik, amit a nyálábformáló algoritmusnak figyelembe kell vennie. Additív zajt feltételezve:

$$\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{n} \quad (2.45)$$

a konvex optimalizálási problémának a következő alakot kell öltetnie:

$$\min_{\mathbf{x} \in \mathbb{C}^N} \|\mathbf{x}\|_1 \quad \text{úgy, hogy} \quad \|\mathbf{A}\mathbf{x} - \mathbf{y}\|_2 < \varepsilon, \quad (2.46)$$

ahol ε a zaj miatt bevezetett hibahatár. A hibahatár értékét úgy kell megválasztani, hogy teljesüljön az $\|\mathbf{n}\|_2 \leq \varepsilon$ feltétel.

Ha a zajszint ismeretlen, nekünk kell becslést adnunk rá, azonban felmerül a kérdés, hogy az alul- vagy a felülbecslés közül melyik a kedvezőbb. A zajszint alulbecslése esetén előfordulhat, hogy a megoldás kevésbé lesz ritka a valóságosnál, azaz olyan helyeken is forrást látunk, ahol valójában nincsenek. Felülbecsléskor ennek az ellenkezője történik, azaz lehet, hogy a megoldás túl ritka lesz, és a kisebb erősségű forrásokat nem találjuk meg. A kettő közül az alulbecslés a kedvezőbb, ugyanis a ritkaság csökkenése kiküszöbölhető újrásúlyozott $\mathbb{L}_1$ -minimalizációval és szintkorrekcióval.

Az újrásúlyozott $\mathbb{L}_1$ -minimalizálás alkalmazásánál az alábbi módon járunk el:

Először egy kezdeti becslést hajtunk végre a 2.46 formula segítségével. Ezután iteratíván megoldjuk az alábbi súlyozott problémát:

$$\min \|\mathbf{W}\mathbf{x}\|_1 \quad \text{úgy, hogy} \quad \|\mathbf{A}\mathbf{x} - \mathbf{y}\|_2 < \varepsilon, \quad (2.47)$$

ahol $\mathbf{W}$ egy diagonális mátrix, mely elemeinek értéke a kezdeti becslés előtt 1, utána pedig a legutóbbi becslés alapján adhatóak meg.

$$w_i = \frac{1}{\hat{x}_i + \xi} \quad (2.48)$$

Itt ξ egy nullánál nagyobb szám a legkisebb várható amplitúdó nagyságrendjében, amely biztosítja, hogy a súlyozás egyik eleme se legyen végtelen, $\hat{x}_i$ a forrásra végzett legutóbbi becslés i -edik eleme. Az iterációt akkor hagyjuk abba, amikor az utolsó két becslés közel megegyezik. Ez általában két iterációval megvalósul:

$$|w_i x_i|^{k+1} = \begin{cases} \frac{\hat{x}_i}{\hat{x}_i + \xi} \approx 1, & \text{ha } |x_i| \geq 0 \\ 0 & \text{egyébként} \end{cases} \quad (2.49)$$

Az újrásúlyozott megoldás tovább javítható szintkorrekció alkalmazásával. Ha az érzékelési mátrixnak csak azokat az oszlopait hagyjuk meg, amelyek a megoldásban csúcserősségnek felelnek meg, a többinek az elemeit pedig nullával helyettesítjük, akkor a kapott $\mathbf{A}_a$ mátrixot felhasználva oldjuk meg a következő egyenletrendszert:

$$\hat{\mathbf{x}}_a = \mathbf{A}_a^+ \mathbf{y} \quad (2.50)$$

ahol $\mathbf{A}_a^+$ a mátrix pszeudoinverze.

3. fejezet

Szimulációs esettanulmányok

Ez a fejezet szimulációs esettanulmányokat mutat be, melyek segítségével az egyes nyalábformáló algoritmusokat összehasonlítjuk és értékeljük.

3.1. Kétdimenziós végtelen távolságú nyalábformálás

Az első szimulációs esetben kétdimenziós nyalábformálással, vagyis mikrofonosor alkalmazásával foglalkozunk. A valós jelek hangterjedését háromdimenziós, gömbi terjedési modell írja le, a nyalábformáláshoz pedig végtelen távolságú, síkhullámú terjedést tételezünk fel.

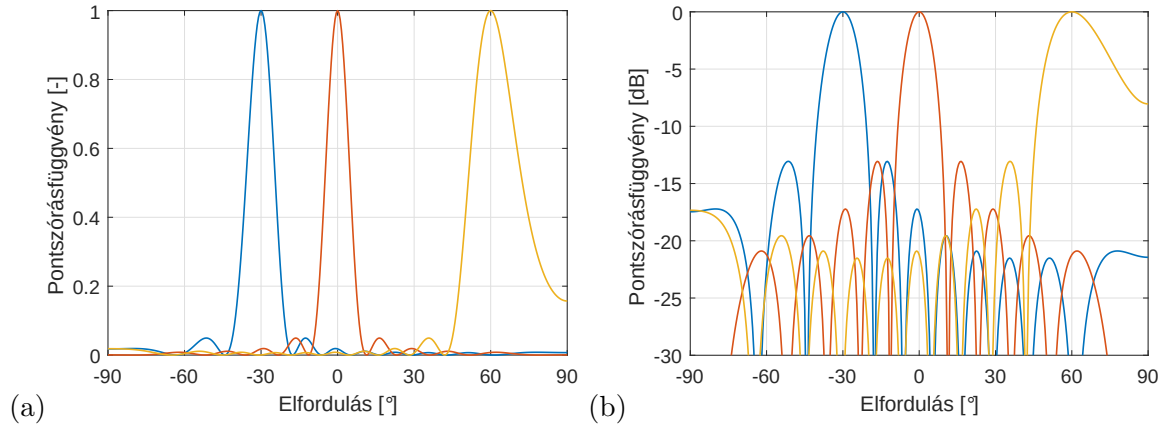
Szimulációs rendszerünk elemeit az alábbi blokséma mutatja be:

- A térben pontforrások helyezkednek el, melyek fehérzajjelet sugároznak. A szimulátorban a fehérzajjeleket $f_s = 48$ kHz mintavételi frekvenciával generáltuk.
- A hangterjedést leíró anyagi paraméterek az alábbiak:
 - A levegő hőmérséklete: $T = 20$ °C
 - A relatív páratartalom: $h = 0$ %
 - A statikus légköri nyomás: $P_0 = 101.325$ kPa
 - A levegő sűrűsége $\rho = 1.20$ kg/m³
 - A hangsebesség $c = 343.45$ m/s
- A mikrofonosor $K = 12$ elemű, ekvidisztáns, szélessége $D = 80$ cm. A mikrofonok távolsága $d = 7.27$ cm. A térbeli átlapolódás határfrekvenciája eszerint $f_{\text{lim}} = \frac{c}{2d} = 2361$ Hz.
- A vászon pontjait – kétdimenziós, síkhullámú terjedést feltételezve – irányszögek definiálják. A lefedett szögtartomány $\phi \in [-90^\circ, +90^\circ]$, a szögfelbontás $\Delta\phi = 0.5^\circ$.

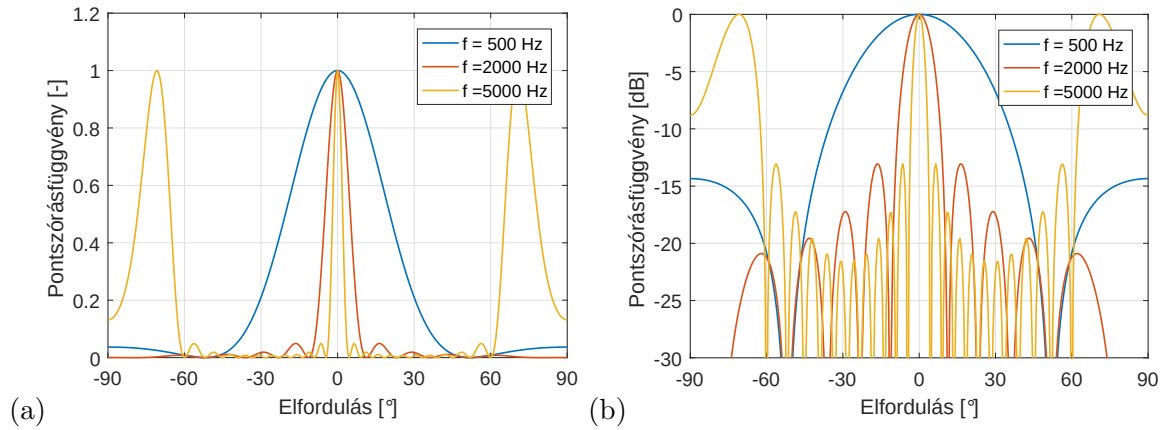
3.1.1. Hagyományos nyalábformálás

A mikrofontömb 2000 Hz frekvenciához tartozó pontszórásfüggvényét a 3.1. ábra mutatja. Az (a) ábra lineáris skálán, a (b) ábra logaritmikus, decibelskálán ábrázolja ugyanazokat a pontszórásfüggvényeket. Megfigyelhető, hogy a mikrofontömb pontszórásfüggvényének főnyalábszélessége hozzávetőlegesen 10° , és ezen a frekvencián számos melléknyaláb jelenik meg a főnyaláb mellett. A legerősebb melléknyaláb szintje a főnyalábét 13 dB-lel múlja alul. Az ábrán az is megfigyelhető, hogy amennyiben a mikrofonosor oldalirányra fókuszál, a főnyaláb szélessége megnövekszik. Ez természetesen az oldalirányú fókuszálás felbontásának csökkenésével jár.

A 3.2. ábra a főirányra való fókuszáláskor adódó pontszórásfüggvényeket hasonlítja össze különböző frekvenciákon. Az 500 Hz-es frekvencia alacsonynak számít, ebben az esetben a főnyaláb szélessége meghaladja a 30° -ot. Megfigyelhető továbbá, hogy mivel az 5000 Hz-es frekvencia az



3.1. ábra. 12 elemű, 80 cm széles mikrofonosor pontszórásfüggvénye 2000 Hz frekvencián, három különböző irányra fókuszálva. Az (a) ábra lineáris, a (b) ábra decibelskálán ábrázolja ugyanazt a pontszórásfüggvényt.



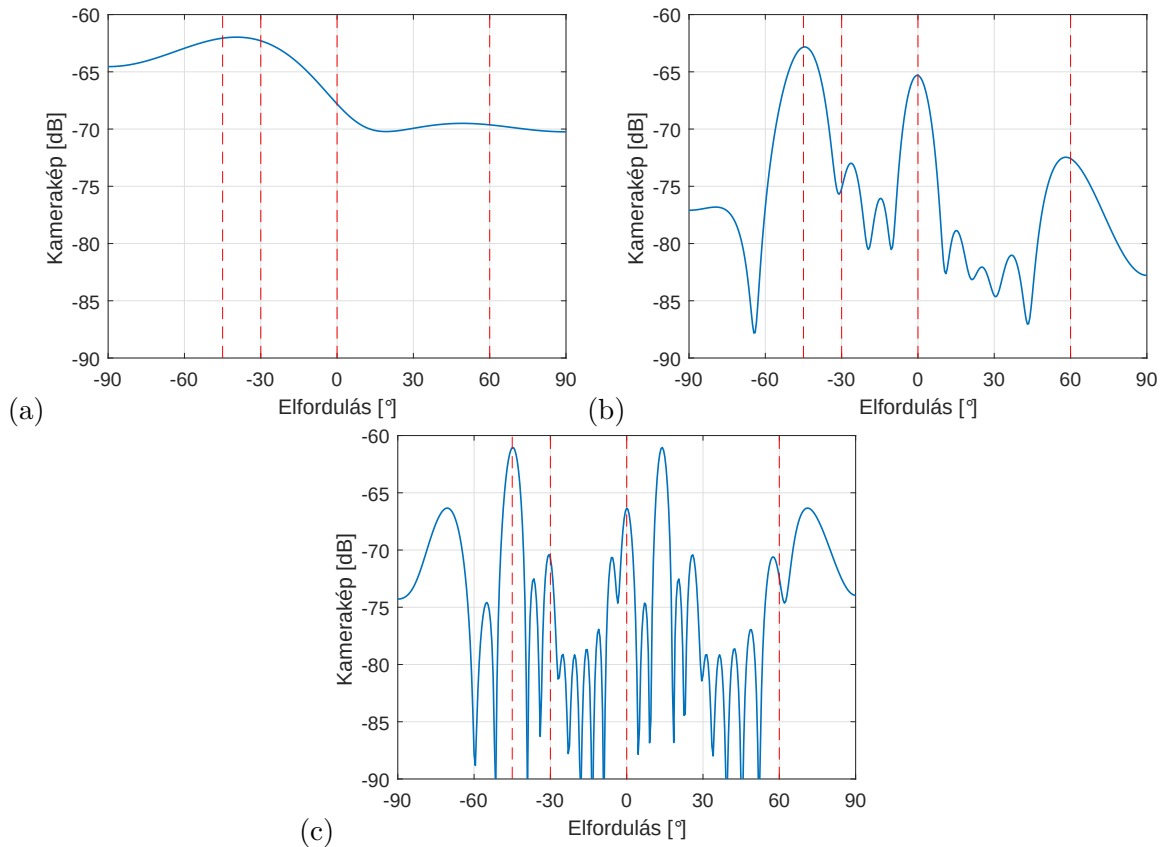
3.2. ábra. 12 elemű, 80 cm széles mikrofonosor pontszórásfüggvénye különböző frekvenciákon, a középipírány ($\phi = 0^\circ$) felé fókuszálva. Az (a) ábra lineáris, a (b) ábra decibelskálán ábrázolja ugyanazt a pontszórásfüggvényt.

átlapolódási határfrekvencia fölött helyezkedik el, itt már oldalirányban belapolódnak a főnyalábal azonos szintű melléknyalábok. Ezek természetesen hamis forrásbecsléshez vezetnek, hiszen elvileg lehetetlen eldönteni, hogy melyik forrás idézte őket elő.

Az alábbiakban a hagyományos nyalábformálás kameraképét vizsgáljuk az alábbi forráselrendezés esetében: A térben négy pontforrás helyezkedik el, ezek mikrofonosor normálisától mért irányszögei $\phi_0 = \{-45^\circ, -30^\circ, 0^\circ, +60^\circ\}$. A források sávkorlátozott fehérzajt sugároznak, teljesítményük rendre $\sigma^2 = \{9, 1, 4, 1\}$. A szimulációknál a források távolságát azonos 75 m-es értékre állítottuk.

A hagyományos nyalábformálás implementációjakor az alábbi módon jártunk el:

- A vett jeleket lekevertük az alapsávba egy f_0 analízisfrekvencián oszcilláló komplex moduláló jellel.
- A modulált jelet egy harmadfokú, $f_c = 20$ Hz vágási frekvenciájú aluláteresztő szűrővel szűrtük.
- A szűrt jelekből állítottuk elő azok kovarianciamátrixát. Itt az időbeli várhatóértékképzést (átlagolást) egy 0.1 s időállandójú exponenciális átlagoló szűrővel végeztük. Így, a jelblokkok feldolgozásának eredményeként olyan kovarianciamátrixokat kapunk, melyek 0.1 s-onként frissülve, mindig az elmúlt 0.1 s hosszú szakaszra átlagolt kovarianciaértékeket közelítik.



3.3. ábra. 12 elemű, 80 cm széles mikrofonsor kameraképe hagyományos nyálábformálással (a) 500 Hz, (b) 2000 Hz és (c) 5000 Hz frekvencián. A valós, az origótól azonos távolságra elhelyezkedő zajforrások pozícióit piros vonalak jelölik. A források teljesítményei elfordulás szerint növekvő sorrendben: $\sigma^2 = [9, 1, 4, 1]$.

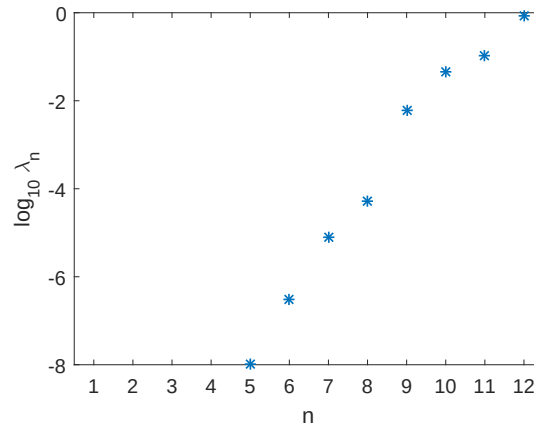
- A fehérzajjelek véletlenszerű volta miatt a 0.1 s időközönként kinyert kovarianciamátrixok, így az algoritmus kimenete is időfüggő, véletlenszerű lesz. A továbbiakban egy-egy véletlenszerűen kiragadott időpillanatban ábrázoljuk az algoritmusok eredményeit.

A 3.3. ábra a hagyományos nyálábformálással kapott kameraképet ábrázolja, mely a különböző irányokhoz tartozó pontszórásfüggvények teljesítményekkel súlyozott összege. Az 500 Hz-hez tartozó (a) ábrán látszik, hogy a széles főnyálábszögű pontszórásfüggvények összegéből nem lehet kinyerni a források irányait. A 2000 Hz-es ábrán mind a négy forrásnál felfedezhetőek csúcsok a kameraképen, de a jelentős teljesítménykülönbség és kis szögtávolság miatt a második forrás „elrejtőzik” a mellényálábok között. Az átalapolódási frekvencia fölötti 5000 Hz-es ábrán a -45° -os forrás fantompárja is megjelenik 10° környékén, és ellehetetleníti a helyes forrásazonosítást.

3.1.2. MUSIC algoritmus

A továbbiakban ugyenezen források, mikrofonsor és vászon alkalmazásával adódó nyálábformálási eredményeket vizsgáljuk a MUSIC algoritmus esetében. A MUSIC algoritmusnak lényeges paramétere a valós források becsült száma. A korábbiakban tárgyaltaknak megfelelően ezt a becsült értéket a mikrofonsor kovarianciamátrixának sajátértékei segítségével határozhatjuk meg.

A 3.4 ábra a kovarianciamátrix sajátértékeit ábrázolja növekvő sorrendben. Az ábrán a sajátértékeket azok összegére normálva ábrázoltuk. Látszik, hogy a legnagyobb sajátérték erősen dominál, értéke majdnem megegyezik az összes sajátérték összegével. A negyedik legerősebb sajátértékhez tartozik az 1%-os érték, az ötödikhez pedig már a 10^{-4} . Ebből sejthető, hogy a valós források száma 4, ami megegyezik a valós szimulált esettel.



3.4. ábra. A 12 mikrofonos mikrofonosor kovarianciamátrixának normalizált sajátértékei növekvő sorrendben, 2000 Hz frekvencián. A normalizálást a sajátértékek összegére való osztással végeztük.

A MUSIC algoritmus kameraképét a 3.5. ábra mutatja a korábbiakhoz hasonlóan három különböző frekvencián. Szembetűnő különbség a hagyományos nyalábformáláshoz képest, hogy a MUSIC algoritmus az alacsony 500 Hz-es frekvencián is szeparálja az egymáshoz közeli, és jelentős amplitúdókülönbségű forrásokat (1. és 2. forrás). A frekvenciát 2000 Hz-re emelve a kamerakép csúcsai kiélesednek, itt a MUSIC algoritmus az irányokat nagy bizonyossággal, „hibátlanul” ismeri föl. Az 5000 Hz-es frekvencián a hamis oldalnyalábok természetesen ismét megjelennek, és a helyes iránybecslés már nem valósítható meg.

3.1.3. CLEAN algoritmus

A CLEAN-PSF algoritmust – a fenti forrásokkal, mikrofonosorral és vászonnal – az alábbi paraméterekkel futtattuk:

- A (2.30) összefüggésben definiált α relaxációs együttható értékét $\alpha = 0.5$ -nek választottuk.

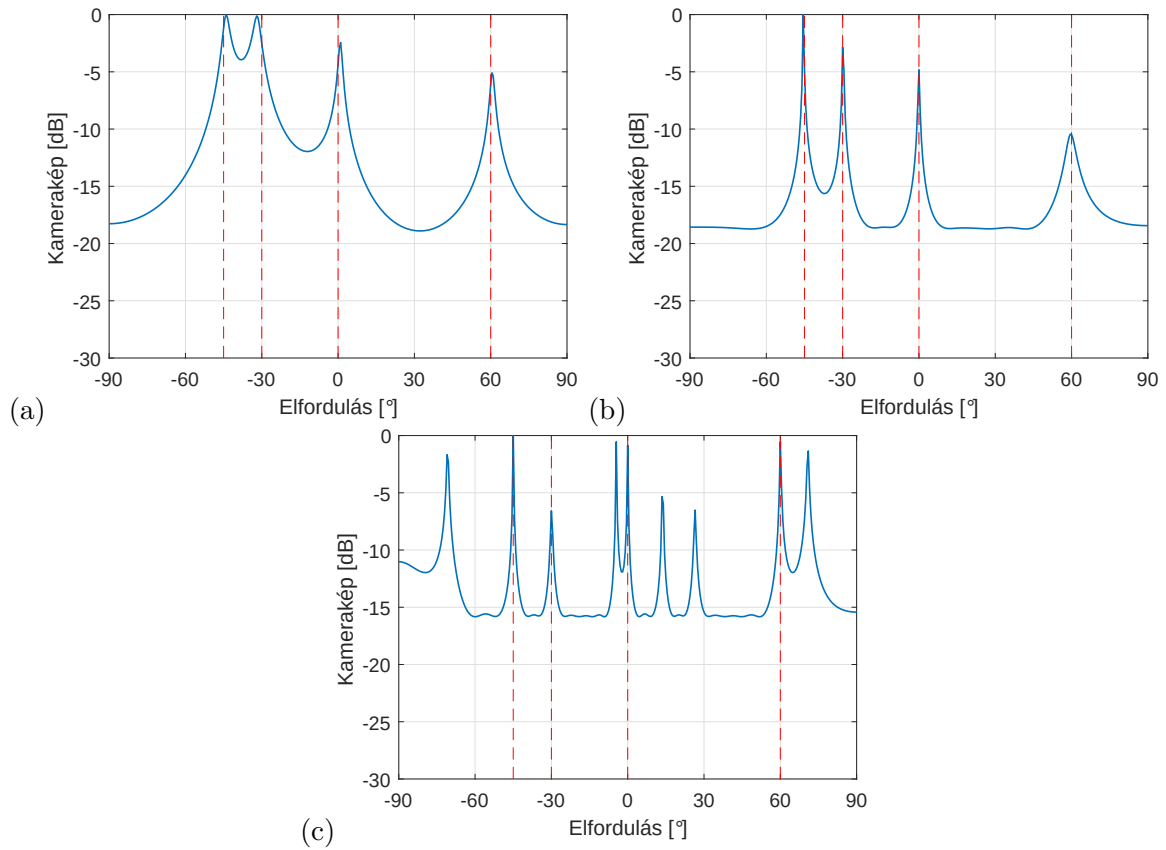
A CLEAN-Psf algoritmus által adott kameraképeket a 3.6. ábra hasonlítja össze különböző frekvenciákon. A korábbiakkal ellentétben – mivel a CLEAN algoritmus sok zérus érték melletti „pontoszerű” eredményeket szolgáltat, a kameraképet nem logaritmikus, hanem lineáris skálán ábrázoltuk. Szemmel láthatólag a CLEAN algoritmus nem képes szétbontani az 500 Hz-es hagyományos nyalábformálás -45° és -30° irányokhoz tartozó pontszórásfüggvényeit, még úgy sem, ha itt a forrásamplitúdókat azonosnak vettük, szemben a korábbi 9 : 1 aránnyal. Az ábrák mutatják az algoritmus iterációs számát is. Az 500 Hz-es eredményt 6 iterációval határozta meg az algoritmus. Mivel a CLEAN algoritmus iterációi nem nagy számítási igényűek, ez a magas iterációs szám is valós idejű végrehajtást eredményez.

A 2000 Hz-es esetben a CLEAN algoritmus – némileg magasabb iterációs szám mellett – helyesen határozza meg a mind a négy forrás pozícióját. Az 5000 Hz-es eredményt a korábbiakhoz hasonló jelentős iránytévesztések terhelik.

3.1.4. A Compressive Beamforming algoritmus

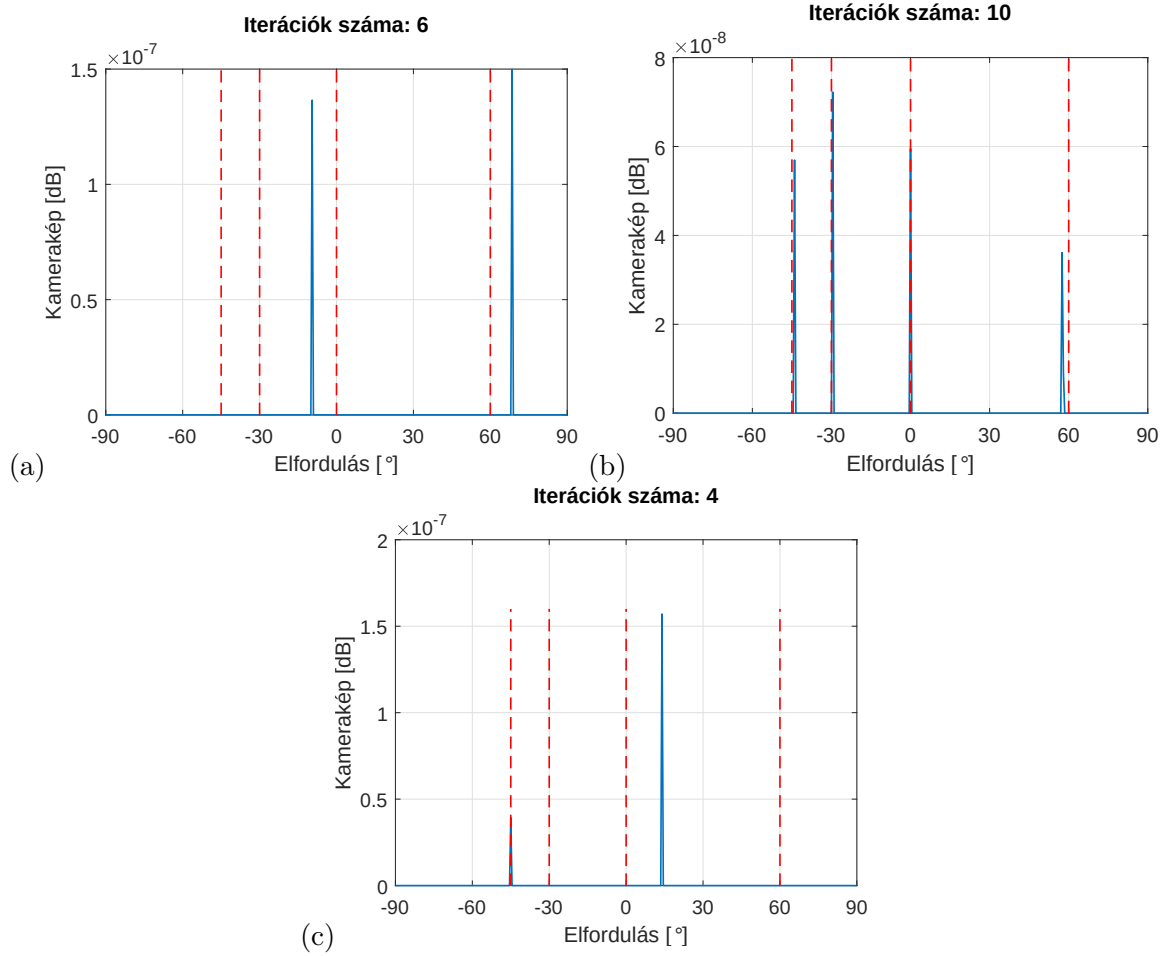
A Compressive Beamforming algoritmus által adott kameraképet a 3.7. ábra mutatja. Ezekről a kameraképekről természetesen azt várjuk el, hogy csupán néhány zérustól különböző értéket tartalmaznak, hiszen az a *Compressive Sensing* algoritmus ritkasági feltétele. A kérdés itt a néhány nemzérus érték pozíciójának pontossága, illetve a konvex optimalizálási algoritmus konvergenciája.

Jelen szimulációk esetén a konvex optimalizálási algoritmus minden időszelvényben és mindhárom vizsgált frekvencián talált megoldást. A 3.7. ábrán ezek közül egy-egy kiragadott esetet

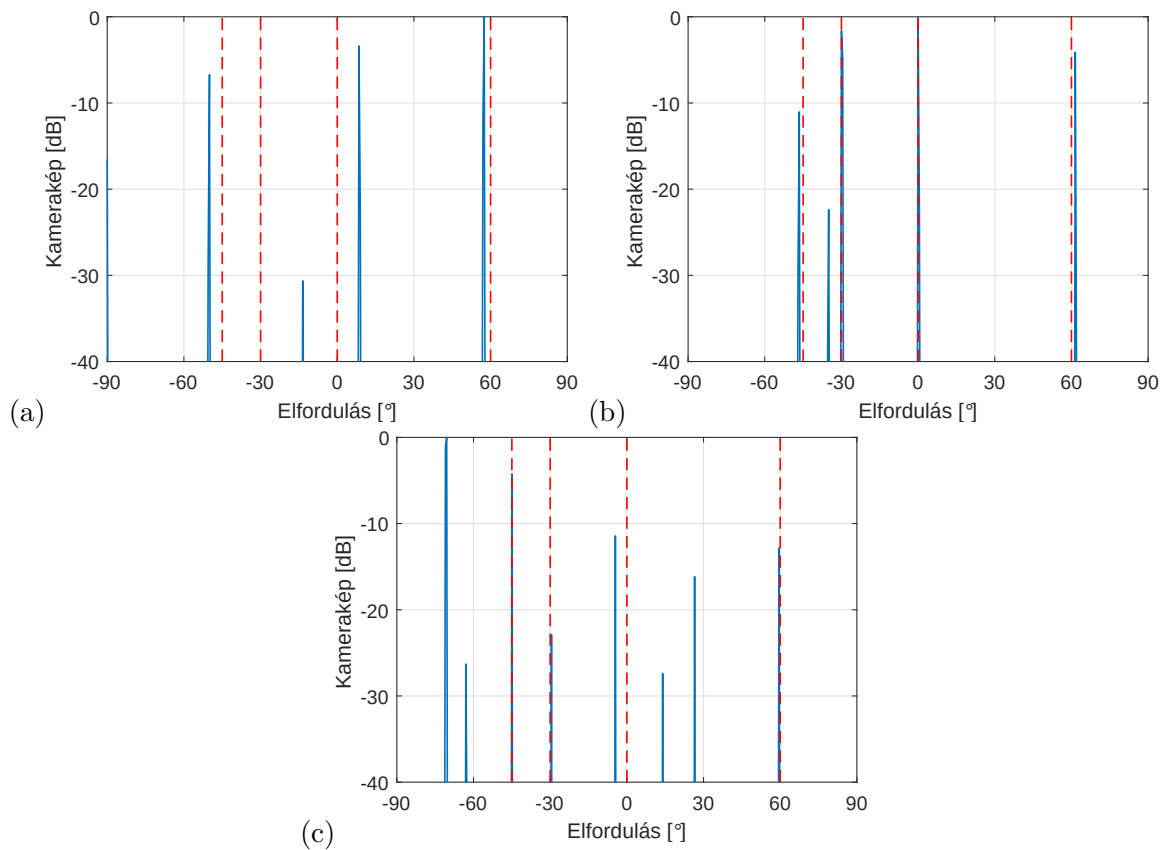


3.5. ábra. 12 elemű, 80 cm széles mikrofonsor kameraképe MUSIC algoritmussal (a) 500 Hz, (b) 2000 Hz és (c) 5000 Hz frekvencián. A valós, az origótól azonos távolságra elhelyezkedő zajforrások pozícióit piros vonalak jelölik. A források teljesítményei elfordulás szerint növekvő sorrendben: $\sigma^2 = [9, 1, 4, 1]$

látunk. A kisfrekvenciás esetben az egymáshoz közeli források elkülönítése problémákba ütközött. A 2000 Hz-es esetben mind a négy forrást helyesen (fél fokos hibával, ami a vászon felbontása) megtalálja az algoritmus. Az ötödik forrás szintje több mint 10 decibellel alulmúlja a negyedik legerősebbet.



3.6. ábra. 12 elemű, 80 cm széles mikrofon sor kameraképe CLEAN-PSF algoritmussal (a) 500 Hz, (b) 2000 Hz és (c) 5000 Hz frekvencián. Az origótól azonos távolságra elhelyezkedő valós zajforrások pozícióit piros vonalak jelölik. A források teljesítményei azonosak.

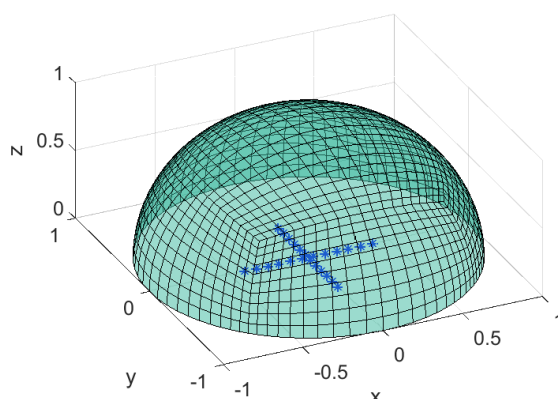


3.7. ábra. 12 elemű, 80 cm széles mikrofon sor kameraképe Compressed Beamforming algoritmussal (a) 500 Hz, (b) 2000 Hz és (c) 5000 Hz frekvencián. Az origótól azonos távolságra elhelyezkedő valós zajforrások pozícióit piros vonalak jelölik. A források teljesítményei azonosak.

Forrás	θ	ϕ
1.	0°	$+90^\circ$
2.	-45°	$+45^\circ$
3.	$+60^\circ$	$+45^\circ$

3.1. táblázat. A háromdimenziós végtelen távolságú nyalábformálási tesztekhez használt forrásirányok. Az irányok értelmezése a hagyományos ϕ, θ gömbi koordinátákat jelentik, ahol θ a z tengellyel bezárt szög, ϕ pedig az xy síkra eső vetület x -tengellyel bezárt szöge.

3.2. Háromdimenziós végtelen távolságú nyalábformálás



3.8. ábra. A háromdimenziós nyalábformálási tesztekhez használt keresztelrendezés és vászon. Az ábrázolt vászon pontjai irányokat jelölnek, a számításokhoz használt valós vászon végtelen távolságú.

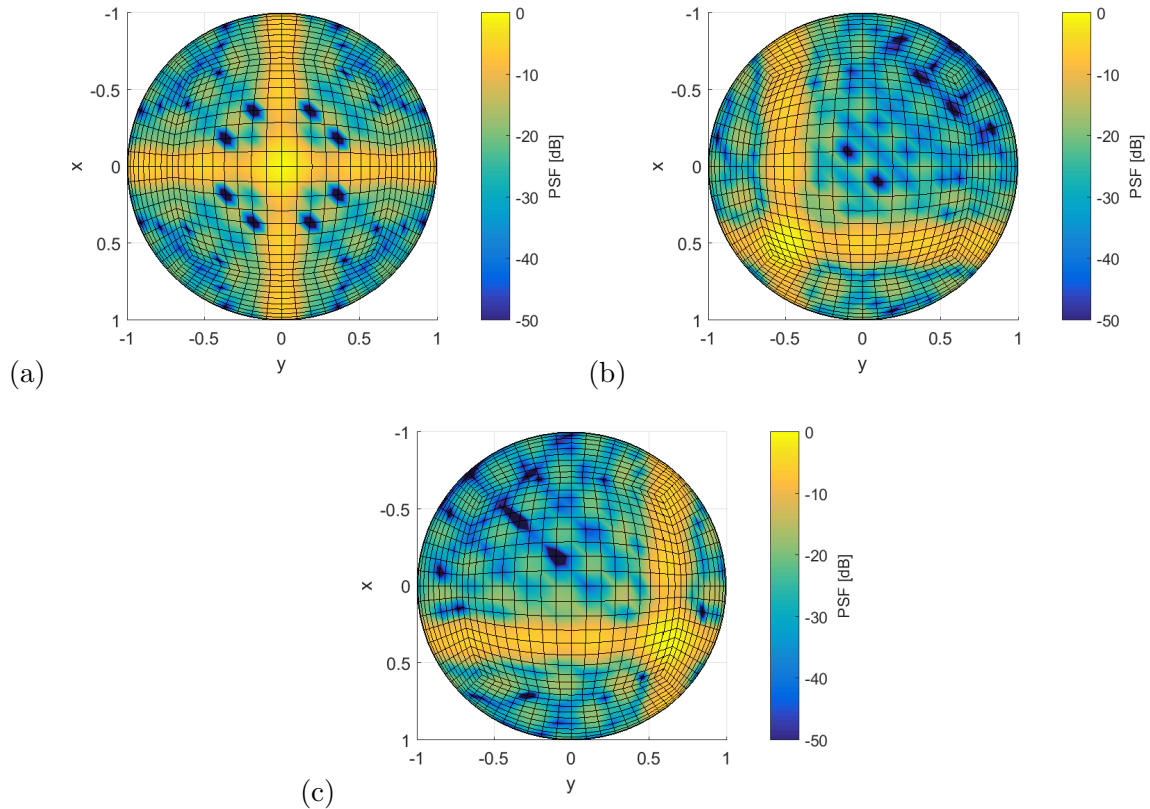
A háromdimenziós nyalábformáláshoz használt elrendezést a 3.8. ábra mutatja. Az ábrán látható a $2 \times 12 = 24$ mikrofonból álló, 80 cm széles keresztelrendezés, mely szomszédos mikrofonjainak távolsága $d = 7.3$ cm, és melynek átlapolódási határfrekvenciája 2360 Hz. Az ábrán látható az 1241 pontból álló vászon, melynek pontjai az origótól vett irányokként értelmezendők. Ennek az az oka, hogy az alábbi számításokban végtelen távolságú nyalábformálást vizsgálunk, melynél az egyes mikrofonokhoz érkező hanghullámok párhuzamos utakon terjednek.

A mikrofonrendszer pontszórásfüggvényét a 3.9. ábra mutatja 2000 Hz frekvencián, a 3.1. táblázatban mutatott irányokra fókuszálva. A kétdimenziós esethez hasonlóan megállapíthatjuk, hogy a főnyaláb szélessége meglehetősen nagy, hozzávetőlegesen 20° körüli, és a melléknyalábelnyomás itt is 13 dB. A kétdimenziós esethez képest előny viszont, hogy a magas melléknyalábok csak a kereszt ágaival párhuzamos irányokban jelennek meg, átlós irányokban jóval nagyobb, 20–25 dB-es melléknyalábelnyomás tapasztalható.

A további szimulációkban három azonos teljesítményű fehérzajforrást tételezünk fel, melyek a mikrofonrendszer középpontjától 75 m távolságban, a 3.1. táblázatban megadott irányokban helyezkednek el.

3.2.1. Hagyományos nyalábformálás

A hagyományos nyalábformálással kapott kameraképet a 3.10. ábra mutatja. Az 500 Hz-es ábrán a mikrofonrendszer pontszórásfüggvényeinek széles főnyalábjai miatt nem lehet elkülöníteni a három forrást. A 2000 Hz-es esetben az egyes források jól elkülöníthetők, és a kamerakép csúcsai közel helyezkednek el a tényleges forrásirányokhoz. Ugyanakkor megfigyelhető, hogy noha a források azonos teljesítményűek, a kamerakép csúcsai eltérő szintűek. Ennek több oka is van.



3.9. ábra. 24 elemű, 80 cm széles keresztelrendezés pontszórásfüggvénye 2000 Hz frekvencián, a 3.1. táblázatban mutatott három irányra fókuszálva.

- Egyrészt a 2. és 3. forrás y -tengellyel párhuzamos melléknyalábjai erősítik egymást, míg az 1. forrás pozíciójára csak egy főnyaláb esik.
- A jelenség másik oka az, hogy a szimulációk fehér zajból vett regisztrátumok véges sáv-szélességű sávszűrővel szűrt mintáiból származnak, melyek természetesen magukban hordozzák a forrásamplitúdók véletlenszerűségét.

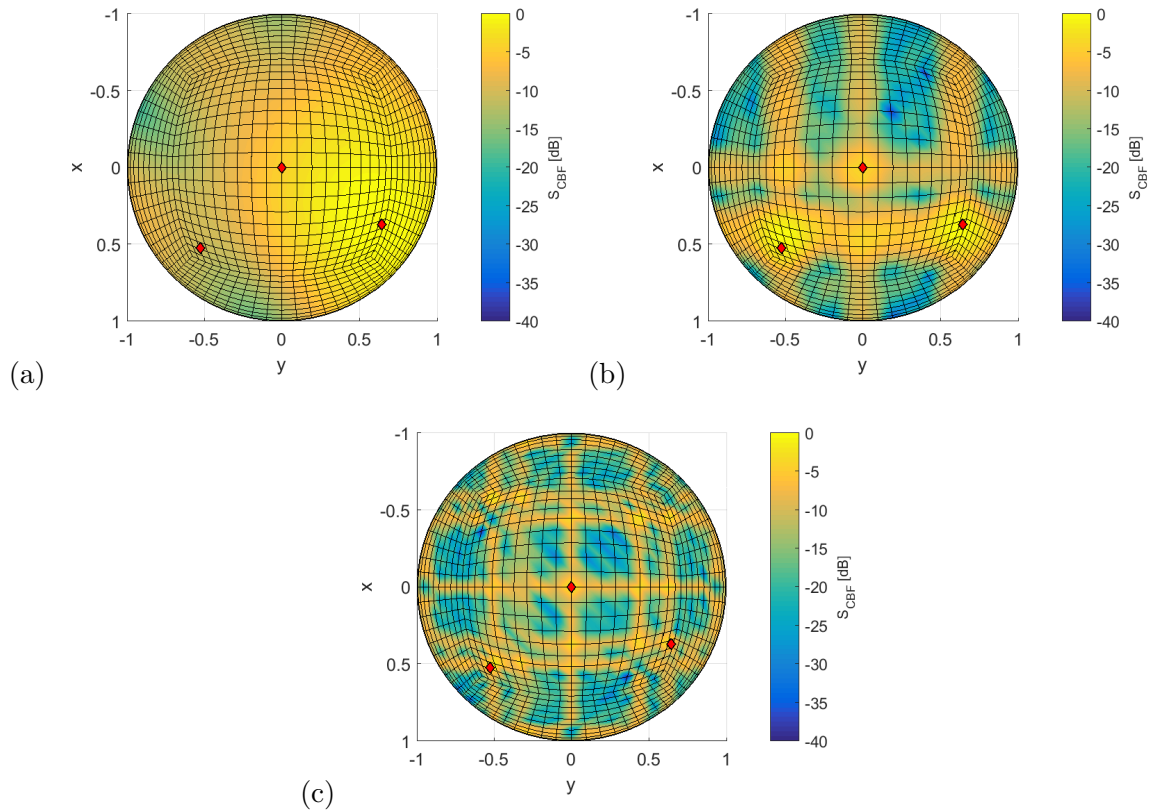
3.2.2. A MUSIC algoritmus

A MUSIC algoritmust – a kétdimenziós esethez hasonlóan – úgy futtattuk, hogy a becsült források száma megegyezett a valós források számával. Ez a jelen háromdimenziós esetben is egyértelműen kiderül a kovarianciamátrix sajátértékeinek nagyságából. Az algoritmus által adott kamerakép a 3.11. ábrán látható kis- és nagyfrekvencián. Itt a térbeli átlapolódási frekvencia fölötti esetet már nem mutatjuk.

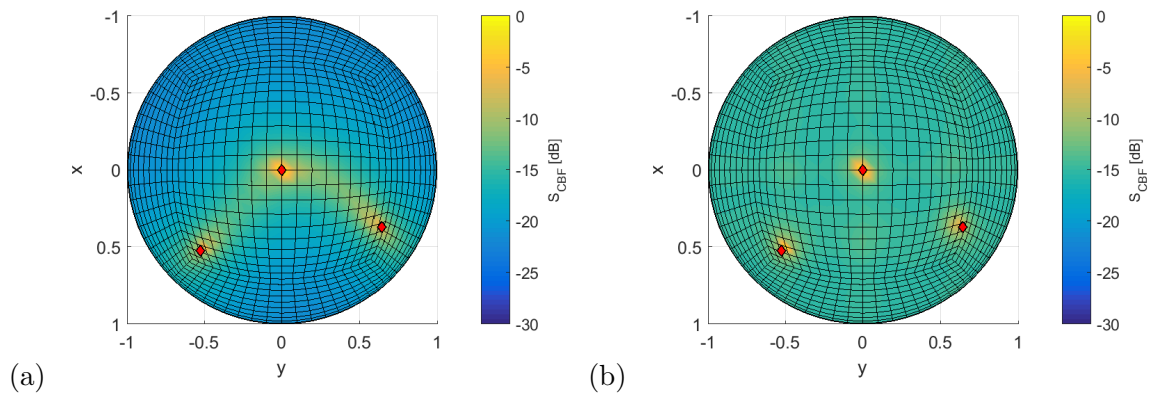
A hagyományos nyálábformálás eredményével összevetve a MUSIC algoritmus sokkal egyértelműbb eredményt mutat. A források pozícióját még a kisfrekvenciás esetben is jól meghatározza, és nincsenek nagy melléknyalábjai, hanem a teljes tértartományban egyenletes zajszintet eredményez. Ez a háttérzajszint a kisfrekvenciás esetben a megtalált forrás szintje alatt 20 dB, a 2000 Hz-es esetben 15 dB.

3.2.3. A CLEAN algoritmus

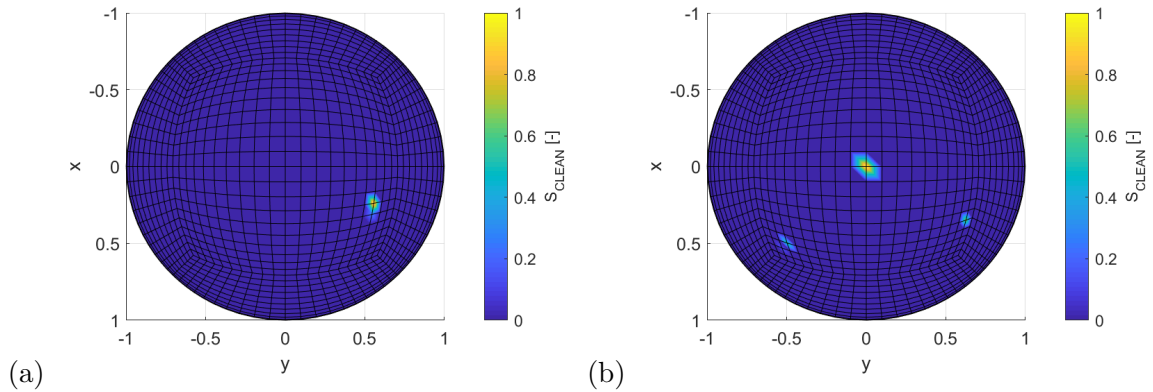
A CLEAN algoritmust a kétdimenziós esethez hasonlóan $\alpha = 0.5$ relaxációs konstans alkalmazásával futtattuk. kameraképe a 3.12. ábrán látható. A CLEAN algoritmus előnye, hogy kisfrekvencián is pontszerű eredményt produkál – az alacsony relaxációs konstans ellenére –, de



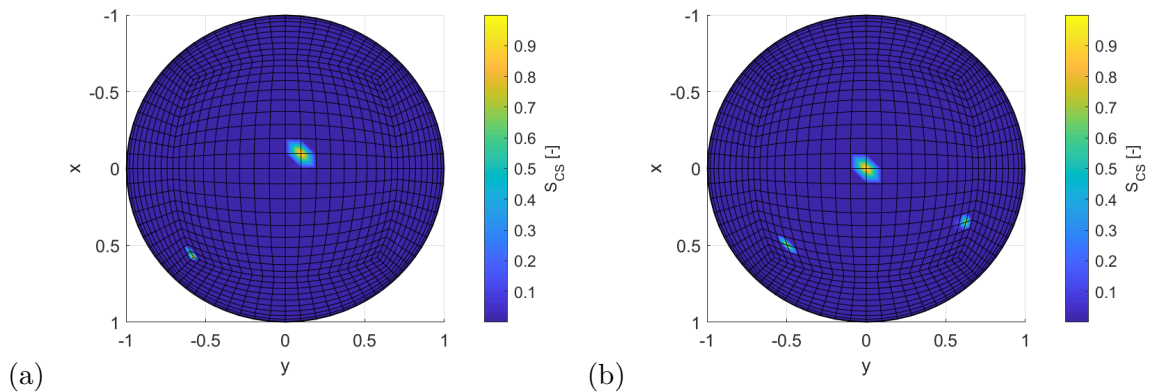
3.10. ábra. 24 elemű, 80 cm széles kereszt elrendezés kameraképe hagyományos nyálábformálással (a) 500 Hz, (b) 2000 Hz és (c) 5000 Hz frekvencián. A piros jelölők az azonos teljesítményű valós források irányait ábrázolják.



3.11. ábra. 24 elemű, 80 cm széles kereszt elrendezés kameraképe MUSIC algoritmmal (a) 500 Hz és (b) 2000 Hz frekvencián. A piros jelölők az azonos teljesítményű valós források irányait ábrázolják.



3.12. ábra. 24 elemű, 80 cm széles kereszt elrendezés kameraképe CLEAN-PSF algoritmussal (a) 500 Hz és (b) 2000 Hz frekvencián.



3.13. ábra. 24 elemű, 80 cm széles kereszt elrendezés kameraképe Compressive Beamforming algoritmussal (a) 500 Hz és (b) 2000 Hz frekvencián.

kisfrekvencián nem képes különválasztani az egymáshoz – a pontszórásfüggvény nyálabszélességéhez mérten – közeli csúcsokat, így csak egyet talál meg a kameraképen.

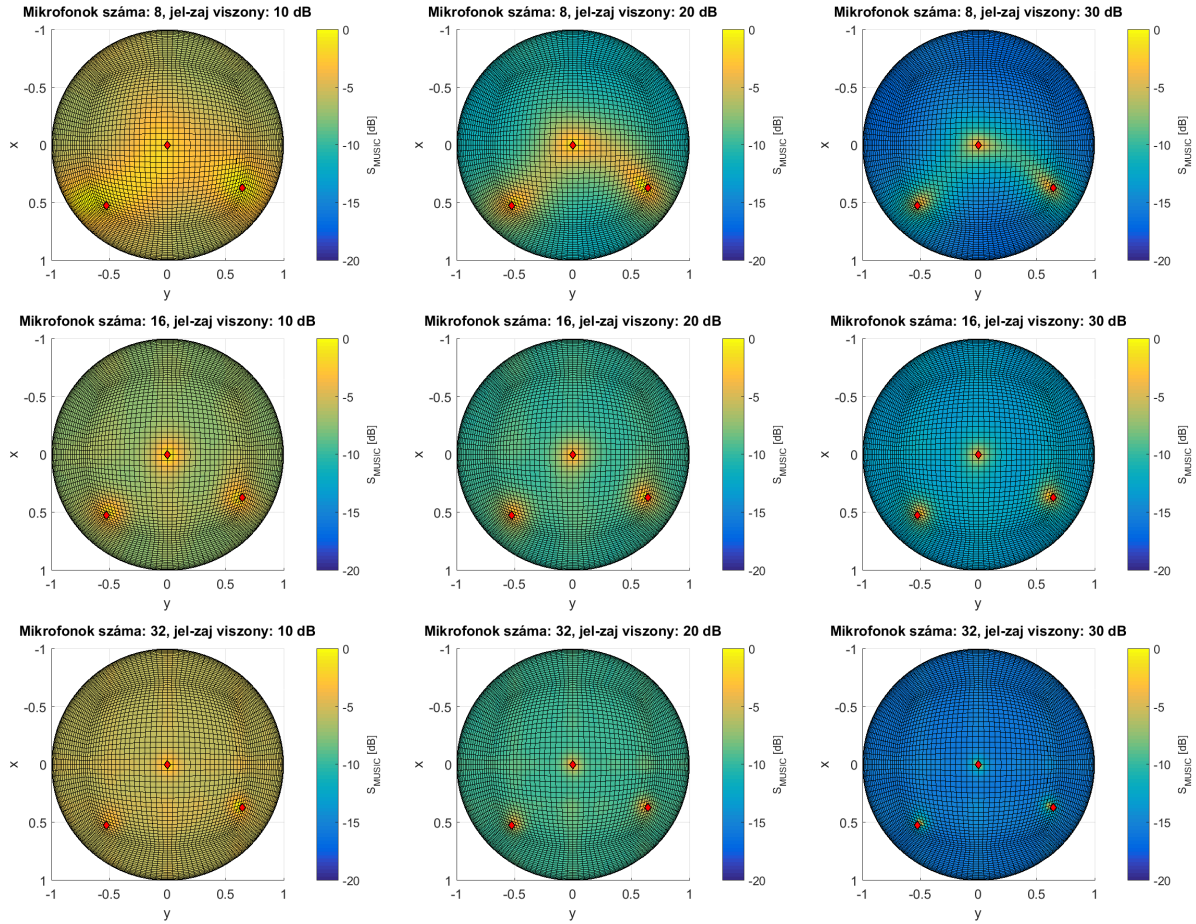
A nagyfrekvenciás esetben a pontszórásfüggvények főnyálabszélességének csökkenése miatt a becslés pontossága növekszik, és az algoritmus megtalálja mindhárom valós forrást. Ezek mellett azonban – kisebb szinttel ugyan, de – egyéb hamis forrásokat is megtalál.

3.2.4. A Compressive Beamforming algoritmus

A *Compressive Sensing*en alapuló Compressive Beamforming algoritmus eredményeit a 3.13. ábra mutatja. Itt megállapíthatjuk, hogy az algoritmus által megkövetelt ritkasági feltétel eredményeként jó pontosságú, pontszerű csúcsokat mutató kameraképet kapunk a kisfrekvenciás esetben is, és a kisfrekvenciás eredmény megbízhatóbb, mint a CLEAN algoritmusé. A nagyfrekvenciás eredményen mindhárom forrás kivehető. Megjegyezzük, hogy itt – mivel a ritksági feltétel majdnem minden képpontban zérus értékű kameraképet eredményez – a kameraképet lineáris színskálával ábrázoltuk.

3.3. A MUSIC algoritmus zajtűrésének vizsgálata

Mivel a fenti szimulációkban a MUSIC algoritmus bizonyult a leghatékonyabbnak, az alábbiakban a MUSIC zajtűrését vizsgáljuk különböző mikrofonelrendezések esetén.



3.14. ábra. A MUSIC algoritmus kameraképe keresztelrendezés mellett, állandó mikrofontávolság mellett, változó mikrofonszámmal és változó jel-zaj viszonytal

Az első esettanulmányban keresztelrendezést vizsgálunk $d = 7.3$ cm állandó mikrofontávolsággal, ami 2358 Hz-es átlapolási frekvenciát jelent, és ez alatt, $f = 2000$ Hz-en vizsgáljuk az algoritmus kameraképét a 3.1. táblázatban mutatott irányú valós források esetében. A zajérzékenység vizsgálatához a mikrofonok vett jeleihez fehérzajt adunk, melynek σ szórását az egyes mikrofonok p_{rms} rms-szintjéhez viszonyítjuk az SNR jel-zaj viszony alábbi definíciója szerint:

$$\text{SNR} = 20 \log_{10} \frac{p_{\text{rms}}}{\sigma}. \quad (3.1)$$

A 3.14. ábra a MUSIC algoritmus kameraképét mutatja változó jel-zaj viszony és mikrofonszám mellett. Kis (4 + 4) mikrofonszám mellett a jel-zaj viszony alapvetően meghatározza a MUSIC algoritmus iránydetekciójának pontosságát: míg 10 dB jel-zaj viszony mellett a források éppen elkülöníthetőek, 30 dB jel-zaj viszonytal éles csúcsokként jelentkeznek a kameraképen. Az ábrán megfigyelhető, hogy kis jel-zaj viszony mellett a mikrofonok számának jelentős növelésével a detekció pontossága lényegesen javítható, de a kamerakép melléknyalábelnyomása nem csökkenthető 5–6 dB alá. A 30 dB jel-zaj viszonyhoz tartozó szimulációk eredménye alapján kimondhatjuk, hogy a mikrofonok számának növelésével a melléknyalábelnyomás a 20 dB szint fölé csökkenthető.

4. fejezet

Összefoglalás

Tanulmányunkban irányérzékelésre alkalmas, sokmikrofonos nyalábformáló algoritmusokat tekintettünk át és értékeltünk. A nyalábformálás alapfeladatának ismertetését követően hat nyalábformáló, forráslokalizációra használható algoritmus működését írtuk le. A hagyományos nyalábformálást, a lineáris algebrai eszköztárral dolgozó MUSIC algoritmust, a hagyományos nyalábformálás eredményét korrelálatlan források esetében képfeldolgozási módszerekkel javító CLEAN-PSF, CLEAN-SC és DAMAS algoritmusokat, illetve a konvex optimalizálás elvén alapuló Compressive Beamformingot. Az algoritmusokat közös keretrendszerben implementáltuk, mely lehetőséget nyújt a módszerek azonos rendszerben történő összehasonlítására.

Az összehasonlítást először kétdimenziós esetben végeztük, melynél az akusztikus kamera egy egyenes vonal mentén elhelyezkedő mikrofonsor. Megmutattuk, hogy a hagyományos nyalábformálás nehezen választja szét az alacsony frekvenciás forrásokat, mivel kisfrekvencián nagyon széles főnyalábbal dolgozik, illetve nagyfrekvencián (az átlapolódási frekvencia alatt) nehezen különít el egymáshoz közel eső vagy lényegesen eltérő teljesítményű forrásokat. A MUSIC algoritmus mindkét szempontból lényegesen javítja a forrásdetekció minőségét. Kisfrekvenciás esetben is hatékonyan szelektál a források között, illetve nagyfrekvencián is képes elkülöníteni a közeli forrásokat; ennek ára a megnövekedett „alapzajszint” a kameraképen. A MUSIC algoritmus további erőssége, hogy alkalmazható a források számának automatikus meghatározására is. A korrelálatlan források feltételezésén alapuló képtisztító algoritmusok előnye, hogy éles csúcsokat tartalmazó kameraképet szolgáltatnak. Hátrányuk, hogy mivel a hagyományos nyalábformálás kameraképből kiindulva képfeldolgozó módszerekkel érik el a végső kameraképet, pontosságukat alapvetően a hagyományos nyalábformálás felbontása határozza meg. Így kisfrekvencián vagy közeli források esetén ugyan pontszerű, de pontatlan pozíciójú eredményt szolgáltatnak. A kameraképet és a mikrofontömb kimenetét lineáris rendszerrel összekötő, majd ritka kamerakép feltételezésén alapuló Compressive Beamforming algoritmus a nagyfrekvenciás esetben hatékony alternatívája lehet a MUSIC algoritmusnak. A szimulációk alapján a forrásokat helyesen találja meg, és a kamerakép felbontását növelni tudja a MUSIC-hoz képest. A MUSIC-kal szemben hátrányosnak bizonyult kisfrekvenciás alkalmazása.

Az algoritmusokat összehasonlítottuk háromdimenziós esetben is, itt keresztelrendezésű mikrofontömböt használtunk. A háromdimenziós összehasonlítás konklúziói teljes mértékben megegyeznek az előző bekezdésben megadott kétdimenziós eredményekkel.

Mivel az összehasonlításokban a MUSIC algoritmus volt az, mely mind kis-, mind nagyfrekvencián kielégítő pontossággal volt képes szétválasztani a forrásokat, külön megvizsgáltuk a MUSIC zajtűrését keresztelrendezés esetén, úgy, hogy a mikrofontömb kimenetére additív gauszi fehérzajt szuperponáltunk. A zajtűrés vizsgálat eredményei azt mutatták, hogy a MUSIC algoritmus kameraképének „háttérzajszintje” a mikrofonok számától függetlenül a mikrofonok jel-zaj viszonyának szintjéhez közeli értéket vesz fel. A mikrofonok számának növelésével a háttérzajszintet lényegesen nem, de a detektált csúcsok élességét javíthatjuk.

Irodalomjegyzék

- [1] T. F. Brooks and W. M. Humphreys. A deconvolution approach for the mapping of acoustic sources (DAMAS) determined from phased microphone arrays. *Journal of Sound and Vibration*, 294:856–879, 2006.
- [2] P. Gerstoft, Mecklenbräuer C.F., W. Seong, and M. Bianco. Introduction to compressive sensing in acoustics. *Journal of the Acoustical Society of America*, 143(6):3731–3736, 2018.
- [3] J.A. Högbom. Aperture synthesis with a non-regular distribution of interferometer baselines. *Astronomy and Astrophysics Supplement*, 15(41):417–426, 1974.
- [4] Reziduum Kft. Forgó légcavaros eszközök észlelésére alkalmas akusztikai antennák működési elvei és műszaki paraméterei. Technical report, 2019.
- [5] P. Sijtsma. CLEAN based on spatial source coherence. Technical Report NLR-TP-2007-345, National Aerospace Laboratory NLR, 2007.
- [6] A. Xenaki, P. Gerstoft, and K. Mosegaard. Compressive beamforming. *Journal of the Acoustical Society of America*, 136(1):260–271, 2014.