



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



Adaptív nyalábformálás dinamikus akusztikai környezetben

Diplomaterv

Banc Roland

konzulens: Dr. Fiala Péter

2023

HALLGATÓI NYILATKOZAT

Alulírott Banc Roland, szigorló hallgató kijelentem, hogy 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, cím, angol és magyar nyelvű tartalmi kivonat, készítés éve, konzulens neve) a BME VIK nyilvánosan hozzáférhető elektronikus formában, a munka teljes szövegét pedig az egyetem belső hálózatán keresztül (vagy autentikált felhasználók számára) közzétegye. Kijelentem, hogy a benyújtott munka és annak elektronikus verziója megegyezik.

Budapest, 2022. 12. 07.

Banc Roland

Tartalom

1	Összefoglaló	5
2	Abstract.....	5
3	Bevezetés	6
4	Mikrofonrendszerek	11
4.1	Iránykarakterisztika	11
4.2	Mikrofonrendszer fogalma	12
4.3	Jelöléstechnika	13
5	Jelterjedés modellezése.....	15
5.1	Több mikrofon egyetlen forrással	17
5.2	Több forrás, több mikrofon.....	19
5.3	Kovarianciamátrix	21
5.4	Nyalábformáló algoritmusok	22
5.5	Akusztikus vászon.....	22
6	Bartlett nyalábformálás	24
7	Capon algoritmus.....	27
7.1	Capon algoritmus frekvenciatartományból analitikusan.....	29
7.2	Capon algoritmus időtartományból becsülve.....	31
8	Robust Capon algoritmus.....	35
9	Jelfeldolgozási folyamat.....	41
9.1	Kovarianciamátrix egyetlen frekvencián.....	41
9.2	Kovarianciamátrix szélessávban.....	42
9.3	Mikrofontömb kimenetének számítása	45
9.4	Teljes rendszer felépítése	49
10	Algoritmus tesztelése modellezett környezetben	50
10.1	Modellezett mérési elrendezés	50
10.2	Generált hangminta ellenőrzése	53
10.3	Bartlett nyalábformálás demonstrálása	54
10.4	Hangminta feldolgozása	57
10.4.1	Egyes forrásokra kvázi-állandósult karakterisztikák	57
10.4.2	Karakterisztika alakulása akusztikus környezet változását követően.....	59
10.4.3	Karakterisztika pontatlan fókuszpozíció esetén	60
10.4.4	Erősítési értékek ábrázolása	64

10.4.5	Átlagolási idő hatása	67
10.4.6	Fókuszált hangforrás hangteljesítménye	68
10.4.7	Szélessávú viselkedés impulzusválasszal.....	69
11	Algoritmus tesztelése valós mikrofontömbön	78
11.1	Mérési helyszín	78
11.2	Mérési elrendezés.....	79
11.2.1	Mikrofontömb	79
11.2.2	hangforrások	81
11.3	Algoritmus vizsgálata valós dinamikus környezetben	84
11.4	Szélessávú működés tesztelése zenei hangfájlokra.....	87
11.5	További mérési tapasztalatok	89
12	Összefoglalás	89
13	Köszönetnyilvánítás.....	90
14	Hivatkozások	91

1 Összefoglaló

A Budapesti Műszaki és Gazdaságtudományi Egyetem Villamosmérnök és Informatikai Karának MSc villamosmérnök képzése során ezen diplomaterv dokumentumom készítését a Hálózati Rendszerek és Szolgáltatások Tanszék által vezetett Hang- és stúdiótechnika mellékspecializáción végeztem.

Dr. Fiala Péter vezetésével vettem részt egy tanszéki projekt fejlesztésében, melynek során mikrofontömbök esetén alkalmazott nyalábformáló eljárásokkal foglalkoztam. Ezek közül is kifejezetten nagy hangsúlyt fektettem az MVDR, vagyis Capon nyalábformáló algoritmusra, mely potenciálisan alkalmas a mikrofontömb karakterisztikájának adaptív formálására dinamikus akusztikus környezetben.

Dolgozatomban bemutatom a mikrofontömbök működésének megértéséhez szükséges fogalmakat és leírásukhoz használt modelleket, bemutatok ezen kívül még nyalábformálási eljárásokat, illetve ezek implementációját mikrofontömbökön. Beszámolok továbbá az implementált algoritmusok szimulált és valós környezetben történő tesztelése során gyűjtött tapasztalatokról is.

2 Abstract

During my MSc electrical engineering studies at the Faculty of Electrical Engineering and Informatics of the Budapest University of Technology and Economics, I prepared this diploma project document in the Sound and Studio Technologies minor specialization led by the Department of Networked Systems and Services.

I participated in the development of a departmental project under the guidance of Dr. Péter Fiala, during which I dealt with beamforming procedures used in the case of microphone arrays. Among these, I particularly focused on MVDR, i.e. Capon's beamforming algorithm, which is potentially suitable for adaptive shaping of the characteristics of the microphone array in a dynamic acoustic environment.

In my thesis, I present the concepts necessary to understand the operation of microphone arrays and the models used to describe them, I also present beamforming procedures and their implementation on microphone arrays. I also report on the experience gained during testing of the implemented algorithms in simulated and real environments.

3 Bevezetés

A Földi élet többmilliárd éves törzsfejlődése során kialakuló állatfajok jelentős része birtokába került az őket körülvevő közeg által kifejtett mechanikai nyomás detektálási képességének. Az ilyen nyomásváltozások fontos információkat hordoznak a közegben zajló mechanikai jellegű eseményekről, zavarokról. A közeg alapvető tulajdonságainak köszönhetően ezek a mechanikai zavarok hullámként képesek abban haladni, és így az említett hatás megnyilvánulásáról, annak pozíciójától nagy távolságra is tanúskodni.

Specializálódott érzékszervek segítségével egyes élőlények az ilyen jellegű nyomáshullámoknak nem csak a detektálására, hanem azon hullámoknak nyomásérték nagysága-, időbeli periodicitása-, és beérkezési iránya szerinti elkülönítésére is képessé váltak, ami a túlélésük során egy rendkívül hasznos képességnek bizonyult.

Az említett érzékelési forma természetesen a *hallás*, melynek fontosságát, és széleskörű alkalmazhatóságát nehéz úgy tárgyalni, hogy közben ne keltsük valamiféle közhely hatását. A hallás képességének állatvilágban betöltött szerepe egészen sokszínű. Segítségével megállapítható, hogy az adott hangforrás milyen hangos, milyen irányban található, illetve hogy konkrétan micsoda. Egyes egyedek képesek arra, hogy a fülükkel csak bizonyos irányból érkező hanghullámokat érzékeljenek, esetleges zavaró hatásokat pedig hallásuk során elnyomjanak, vagy teljesen kizárjanak. Ez nagy előny jelent a számukra relevánsnak tekintett hangforrás pozíciójának pontos megállapításában, illetve a hangforrásról történő információ gyűjtésében. Példa lehet erre egy zsákmány becserkészésének, vagy egy potenciális ragadozó elkerülésének szándéka. A hallásuk ilyen irányérzékenységének megvalósítása érdekében az egyes élőlények nem csak kifinomult alakú hallószervvel rendelkezhetnek, mint például a főemlősök, hanem annak a többi érzékszervüktől független mozgására is képesek lehetnek, ami a hallás irányérzékenységének változtatását is lehetővé teszi, mint például a macskafélék esetében.

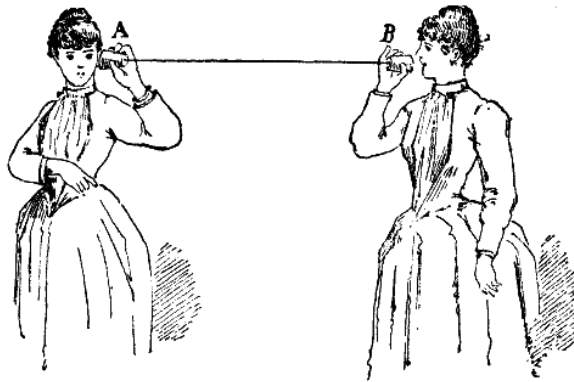
Azonban nem kizárólag a hallószerv változtatásán keresztül van lehetőség az irányérzékenység befolyásolására. Az élővilágban jellemző szimmetriáknak köszönhetően a hallószervekből is jellemzően több, legáltalánosabb esetben kettő áll rendelkezésre egy egyed számára. Az ezekből egyidejűleg származó jelek alapján lehetősége van az élőlénynek a központi idegrendszere segítségével feldolgoznia a környezetében zajló akusztikus eseményeket (akusztikus környezetet), és így a hallószerv akaratlagos mozgása nélkül is megállapítani egy hangforrás irányát.

Erre képes az ember is, sőt ennek tárgyalása zajlik a *Koktélparty effektus* [1] vizsgálata során. Az említett jelenség lényege, hogy egy zsúfolt, zajos eseményen is, ahol a fennálló háttérzaj bár jelentős, képesek vagyunk beszélgetést folytatni, sőt akár mások beszélgetésébe is belehallgatni. Akkor válik ez a jelenség egészen feltűnővé, ha a környezetről egyetlen mikrofonnal felvételt készítünk, akkor ugyanis a beszélgetések

érthetősége nagyban leromlik. A jelenségre adott magyarázatok egyike, hogy bár a fülünkhöz a zavarforrások zaja elér, agyunk automatikusan képes zajok szűrésére a hallószervünk által biztosított jelek együttese alapján.

A hallás élővilágban történő alkalmazásának sokszínűsége műszaki megoldások sorozatát motiválta. Diplomatervezésem során a hangérzékelés irányérzékenységének olyan tudatos alakítása volt a célom, mely zavarforrások esetén is kedvező eredményt ad.

Az egyes távoli hangforrások érzékelésére, és a különféle jelenlevő zavarforrások kiszűrésére megalkotott műszaki megoldások között említhető például a beszélőcső, az akusztikus telefon, vagy a sztetoszkóp. Ezekben közös, hogy már a Bell-fele villamos hangtovábbítást megelőzően is lehetőséget biztosítottak egy kiválasztott hangforrás vizsgálatára úgy, hogy a hangforrás és az érzékelő között olyan közvetlen jelút került kialakításra, amin keresztül az adó jelei a vevőig közel csillapítatlanul juthattak el, az egyes zavarforrások jelei pedig azon keresztül nem haladhattak. Ez kisebb távolságok, illetve együttműködő hangforrások esetén kedvező, és egyszerű megoldásnak bizonyultak, azonban ezek a körülmények sokszor nem állnak fenn.



1. ábra Korai akusztikus távbeszélő készülék

A hallócső [2] segítségével lehetőség van a közvetlen összeköttetésben nem lévő hangforrás jelének zavarszűrt érzékelésére. Ezek az eszközök a 19. században terjedtek el, használatuk során a tölcsér alakú eszköz az adott irányból adott felületre érkező hangteljesítményt összegyűjtve juttatja azt a felhasználó füléhez, így ő az adott hangot hangosabbnak hallja.



2. ábra Hallás javítására szolgáló hallócső [3]

A világháborúk kitörését, és a - viszonylag gyors mozgásra képes, viszont jelentős mennyiségű hangot kibocsájtó - légi járművek megjelenését követően felmerült az igény az ilyen jellegű hallásjavító eszközök fejlesztésére. Ennek során a korábban is tárgyalt módon, viszont sokkal hatékonyabban volt lehetőség egy kiválasztott irányból érkező zaj érzékelésére, majd ezt követően pedig a zajforrás beazonosítására, így biztosítva az esetlegesen bombázó repülőgépek, vagy harci járművek érkezésének korai érzékelését a radar megjelenése előtt.

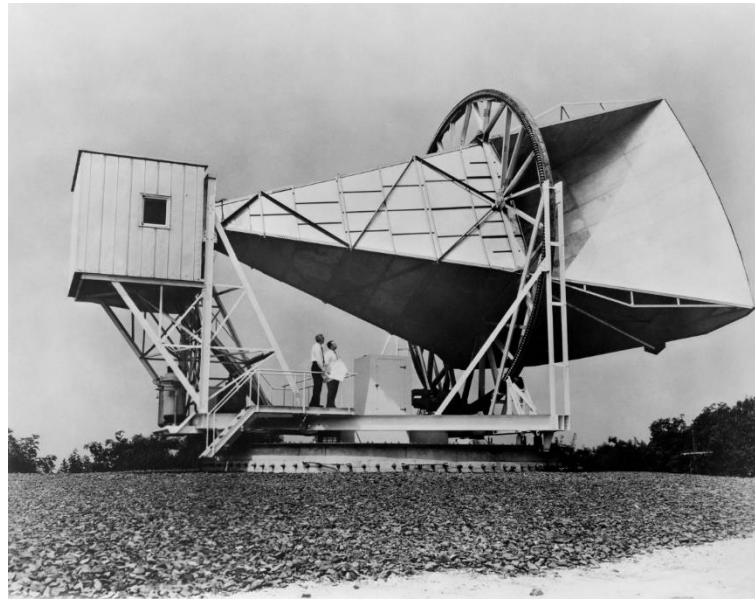
Ezek rendszerint vagy nagyméretű, mozgatható csőszerkezetek, vagy ívelt betonfelületek voltak, melyek esetén a megfigyelőnek az elhelyezkedésének a változtatásával volt lehetősége kiválasztani a hangérzékelés irányát.



3. ábra Légitámadások kiszűréséhez használt fülelőkészülék [4]

A radarok megjelenését követően az elektromágneses hullámok érzékelése során is rendkívül nagy jelentőségű kérdéssé vált az azokat detektáló antennák irányérzékenysége. Ennek céljára rendelkezésre álltak összetett, kis irányérzékenységgű antennamegoldások, azonban a radarok üzemelési sebességéhez képest az ilyen

nagyméretű antennák csupán lassan voltak képesek biztonságos körülmények között üzemelni.



4. ábra A kozmikus mikrohullámú háttérsugárzást felfedezése során alkalmazott, irányított Holmdel kürtantenna [5]

Az irányérzékenység egyszerűbb kialakítására nyújtott megoldást az antennarendszerek bevezetése. Ilyen antennarendszerek csupán egyszerű antennákból épültek fel, viszont azok elhelyezkedése, száma, és viszonya jól megtervezett. Az egyes antennákból érkező jelek utólagos feldolgozásával arra is lehetőség van, hogy az antennarendszer eredő irányérzékenysége módosítható legyen.



5. ábra Ukrán területen található ICBM detekciót szolgáló Duga radar antennarendszere [6]

Az akusztikai jelenségek villamos jelekké alakításra biztosított lehetőséget a mikrofon megjelenése, mely a fejlődő villamos jelfeldolgozás által lehetővé tette hangjelek összetett feldolgozását is. Több egyszerűbb mikrofon alkalmazásával lehetőség van az antennarendszerekhez hasonló irányérzékenység alakítására, az úgynevezett nyalábformálásra is. A digitális jelfeldolgozás elterjedésével ennek az eljárásnak a megvalósíthatósága is nagyban megnőtt.

Ennek jelentősége rendkívül nagy, ugyanis a tengeralattjárókban alkalmazott szonárok, valamint orvosi alkalmazású ultrahangos képalkotó eszközök csak ilyen nyalábformálási eljárások segítségével képesek megvalósítani az eléjük állított követelményeket. Napjainkban elérhetőek ezen kívül olyan mikrofonrendszerek, melyek vagyoni védelmi, felderítési, vagy tudományos ismertszerzési, például geológiai feladatokban kapnak szerepet. Ezek jelenek feldolgozását különféle nyalábformáló algoritmusok végzik.



6. ábra Gömb alakban elhelyezett mikrofonokból kialakított mikrofontömb

Diplomatervezési munkám során olyan mikrofontömbökre alkalmazható nyalábformálási eljárást implementálok, melynek segítségével olyan irányérzékenység érhető el, mely egy kedvező irányból érkező hangokat torzítás nélkül enged át, miközben esetlegesen megjelenő, vagy akár mozgó zavarforrások jelét teljesen ki tudjuk oltani.

Tulajdonképpen a korábban említett *koktélparty effektusban* bemutatott jelenséget van lehetőség így reprodukálni, tehát tetszés szerint kiválaszthatok egy éppen beszélő személyt egy nagyobb tömegben, vagy zaj közelében, és ideális esetben eredményül torzítatlanul, és zavarmentesen hozzájuthatok ahhoz, amit egy közvetlen mellette álló személy hallhatott.

4 Mikrofonrendszerek

4.1 Iránykarakterisztika

Mikrofonok esetén fontos paraméter azok iránykarakterisztikája, mely azt határozza meg, hogy az adott mikrofon a térben adott irányokban, adott frekvencián milyen érzékenységgel rendelkezik, vagyis az az adott irányokból érkező hangteljesítmény milyen mértékben érzékeli.

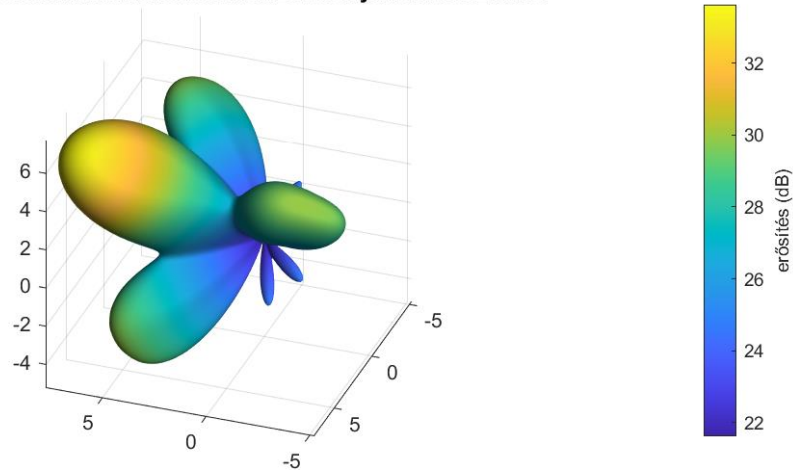
Érzékenység alatt a mikrofonnak azt a tulajdonságát értjük, hogy adott frekvenciájú, adott effektív értékű hangnyomásjel mekkora effektív értékű villamos jelet hoz létre a kimenetén, mértékegysége [V/Pa].

Ez az $S(\vartheta, \varphi)$ iránykarakterisztika az érzékenységnek az adott irány horizontális koordinátarendszerbeli ϑ azimut-, és φ elevációs szögértékének kétváltozós függvénye. Felfogható úgy is, mint egy olyan beiktatott erősítő, melynek erősítési értéke a vizsgált beérkezési irány függvénye.

Ennek ábrázolására történhet például gömbi polár-koordinátarendszerben térben, vagy egyszerűen valamilyen gömbi vetületen (például henger-, vagy síkvetületen) szintvonalakkal, esetleg az érzékenység mértékének színskála segítségével történő érzékeltetéssel. Sokszor az iránykarakterisztika csak egy bizonyos nevezetes síkmetszete kerül polár-koordinátarendszerben ábrázolásra, főleg abban az esetben, ha az tengelyesen szimmetrikus.

A 7. ábra egy véletlenszerűen kiválasztott mikrofon gömbi polár-koordinátarendszerbeli iránykarakterisztikáját ábrázolja dB-ben. Könnyen megfigyelhető rajta főnyalábnak nevezett fő érzékenységi irány, illetve melléknyaláboknak nevezett kevésbé érzékeny, de lokális maximumokat jelentő irányok.

Példa gömbi
polár-koordináta-rendszerben ábrázolt iránykarakteristikára



7. ábra iránykarakterisztika szemléltetése

Izotrópnak nevezzük azt a mikrofont, mely a tér bármely irányából ugyanazzal az érzékenységgel képes a hanghullámok villamos jellé alakítására, tehát az ilyen mikrofon egy középpontosan szimmetrikus gömbi karakterisztikával rendelkezik, a hozzá érkező hullámok beérkezési iránya szempontjából nem szelektív.

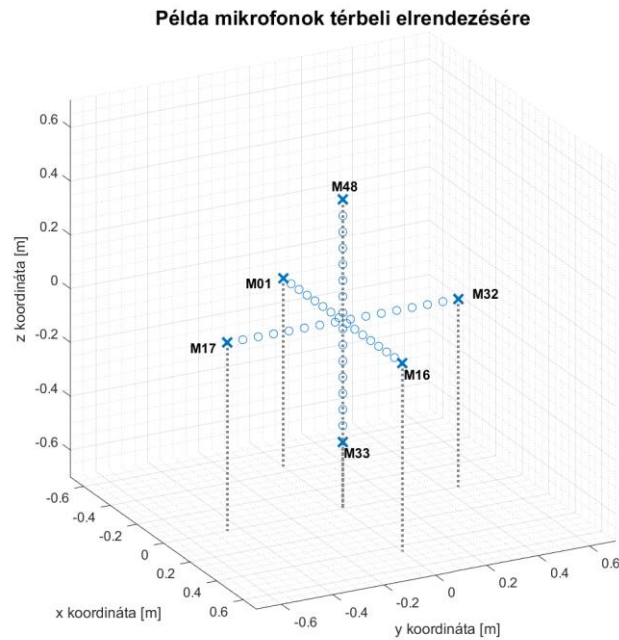
4.2 Mikrofonrendszer fogalma

Adott számú, térben valamilyen módon elhelyezett izotróp mikrofon együttes alkalmazása, majd a mikrofonok által szolgáltatott jelek összegzése esetén azt tapasztalhatjuk, hogy ismét megjelenik az irányszelektivitás a beérkező jelek szempontjából, tehát értelmezhetővé válik a mikrofonokra egy közös eredő iránykarakterisztika.

Ez annak köszönhető, hogy a térbeli elhelyezésemből adódóan az egyes mikrofonokon egy adott forrásból érkező hangjelet a jelterjedésből származó késleltetésnek megfelelően, adott frekvencián eltérő fázissal érzékelik. Ezek alapján a mikrofonok kimenetén megjelenő jelek összegzését követően olyan irányokban tapasztalható erősítés-, illetve kioltás, melyekben a mikrofonok kimenetén a jelfazorok eredője maximális, illetve nulla.

A források jeleinek ezt a térbeli terjedésből adódó, azok késésében megnyilvánuló lineáris torzulását úgy lehet elképzelni, mintha az egyes hangforrásokból származó jelek szűrőkön haladnának keresztül, majd az egyes mikrofonokban összegződnének. Amennyiben az egyes mikrofonokból származó jeleket az összegzés előtt magunk által beiktatott szűrőkön keresztül vezetjük át, melyek azok adott frekvenciákon vett fázisát, illetve amplitúdóját módosítják, úgy a mikrofontömbnek ez az eredő

iránykarakterisztikája is módosul. Az ilyen együttesen alkalmazott, kimeneteiket szűrőkön keresztül összegzett mikrofonokat mikrofonrendszereknek nevezzük.



8. ábra Példa térbeli mikrofonelrendezésre

A mikrofonrendszerek analitikus vizsgálata érdekében elengedhetetlen a mikrofonok térbeli pozíciójának ismerete. Ez azonban tulajdonképpen tetszőleges, vagy akár igényekre optimalizált lehet.

4.3 Jelöléstechnika

A későbbiek során alkalmazott matematikai műveletekkel kapcsolatban, a félreértések elkerülése érdekében az alábbi jelöléseket rögzítem:

Kis félkövér betű minden esetben oszlopvektort jelöl, amit ábrázolás során kézenfekvőbb transzponált sorvektorral megadni. A transzponált jelölése a $[\cdot]^T$. A vektor elemei azonos normál betűvel kerülnek jelölésre. Példa egy N hosszú vektorra:

$$\mathbf{v} = \begin{bmatrix} v_1 \\ \dots \\ v_n \end{bmatrix} = [v_1, \dots, v_n]^T$$

Nagy, félkövér betűvel minden esetben mátrixot jelölünk, annak elemeit pedig azonos normál kisbetűvel. Példa egy $N \times N$ -es mátrixra:

$$\mathbf{M} = [\mathbf{m}_1, \dots, \mathbf{m}_N] = \begin{pmatrix} m_{11} & \cdots & m_{1N} \\ \vdots & \ddots & \vdots \\ m_{N1} & \cdots & m_{NN} \end{pmatrix}$$

Időfüggő fizikai változók esetében minden esetben jelölésre kerül azok időfüggése, illetve időfüggő változók Fourier-transzformáltját az alábbi módon jelölöm:

$$F(\omega) = \mathcal{F}\{f(t)\} = \int_{-\infty}^{\infty} f(t) \cdot e^{-j\omega t} dt$$

Időfüggő fizikai változók konvolúcióját a $*$ műveleti jellel jelölöm:

$$f(t) * g(t) = \int_{-\infty}^{\infty} f(\tau) \cdot g(t - \tau) d\tau$$

A frekvenciatartománybeli változókból alkotott vektorok, illetve mátrixok esetén azok körfrekvenciától való függésének jelölése az összetettebb kifejezések jobb olvashatósága érdekében sokszor elmarad, viszont ezek körfrekvenciafüggése a szövegkörnyezet és a jelölés alapján minden esetben egyértelmű.

$$\mathbf{M} := \mathbf{M}(\omega)$$

$$[\mathbf{m}_1, \dots, \mathbf{m}_N] := [\mathbf{m}_1(\omega), \dots, \mathbf{m}_N(\omega)]$$

$$\begin{pmatrix} m_{11} & \cdots & m_{1N} \\ \vdots & \ddots & \vdots \\ m_{N1} & \cdots & m_{NN} \end{pmatrix} := \begin{pmatrix} m_{11}(\omega) & \cdots & m_{1N}(\omega) \\ \vdots & \ddots & \vdots \\ m_{N1}(\omega) & \cdots & m_{NN}(\omega) \end{pmatrix}$$

A vektorokra, illetve mátrixokra alkalmazott \mathbf{v}^* operátor azok elemenkénti konjugálását, míg a \mathbf{v}^H operátor a vektor adjungáltját, vagyis konjugált transzponáltját, úgynevezett Hermite-féle transzponáltját jelenti, tehát:

$$\begin{aligned} \mathbf{v} &= [v_1 \ v_2 \ \dots \ v_n] \\ \mathbf{v}^* &= [v_1^* \ v_2^* \ \dots \ v_n^*] \\ \mathbf{v}^H &= [v_1^* \ v_2^* \ \dots \ v_n^*]^T \end{aligned}$$

$$\mathbf{M} = \begin{pmatrix} m_{11} & \cdots & m_{1N} \\ \vdots & \ddots & \vdots \\ m_{N1} & \cdots & m_{NN} \end{pmatrix}$$

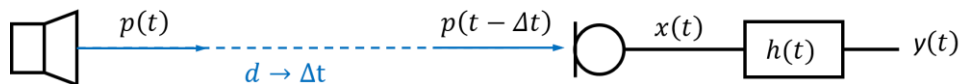
$$\mathbf{M}^* = \begin{pmatrix} m_{11}^* & \cdots & m_{1N}^* \\ \vdots & \ddots & \vdots \\ m_{N1}^* & \cdots & m_{NN}^* \end{pmatrix}$$

$$\mathbf{M}^H = (\mathbf{M}^*)^T = \begin{pmatrix} m_{11}^* & \cdots & m_{1N}^* \\ \vdots & \ddots & \vdots \\ m_{N1}^* & \cdots & m_{NN}^* \end{pmatrix}^T$$

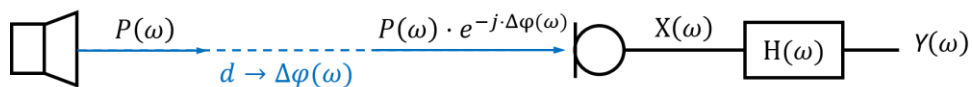
5 Jelterjedés modellezése

Egy adott hangforrás jelének, egy tőle d távolságban lévő mikrofon-, majd a mikrofont követő szűrő kimenetén megjelenő válaszanak leírása történhet idő-, illetve frekvenciatartományban a 9. ábra szerint.

Jelterjedés időtartományban



Jelterjedés frekvenciatartományban



9. ábra Jelterjedés modellezése idő- és frekvenciatartományban

Időtartományban a forrás jelét $p(t)$ frekvenciatartományban pedig $P(\omega)$ írja le. A mikrofon közvetlen kimenetén megjelenő jelet $x(t)$, frekvenciatartományban pedig $X(\omega)$ jelöli. A szűrőt időtartományban a $h(t)$ impulzusválaszával, míg frekvenciatartományban $H(\omega)$ átviteli karakterisztikájával jellemezhetünk. A szűrő kimenetén megjelenő választ pedig időtartományban $y(t)$, frekvenciatartományban pedig $Y(\omega)$ -val jelöljük.

$$P(\omega) := \mathcal{F}\{p(t)\}$$

$$X(\omega) := \mathcal{F}\{x(t)\}$$

$$H(\omega) := \mathcal{F}\{h(t)\}$$

$$Y(\omega) := \mathcal{F}\{y(t)\}$$

Fontos megjegyezni, hogy a modellünkben szereplő izotróp mikrofonok érzékenysége minden frekvencián, minden térirányban azonosan 1 [V/m], így a mikrofon közvetlen közelében pascalban mérhető hangnyomásértékek a mikrofon kimenetén azonos számértékkel jelennek meg Volt mértékben. Ez a modellünk szempontjából egy egyszerűsítés, valós körülmények között a mikrofonok nem izotrópok, illetve átvitelük

nem egyenletes, precízebb alkalmazások esetén ezeket is figyelembe kell venni, viszont ennek kompenzálása viszonylag egyszerű.

A hangforrás jelamplitúdójának a térben való terjedése során a forrástól mért távolság inverzével kellene változnia, így távolabbi mikrofonok ugyanannak a forrásnak a jelét gyengébben kellene érzékeljék, azonban jelenlegi modellünkben síkhullámú terjedést feltételezek, ami abban az esetben lehet indokolt, amennyiben a mikrofontömb térbeli kiterjedése jóval kevesebb, mint az egyes mikrofonok, és a forrás közötti távolság. Így az egyes mikrofonokon megjelenő jelamplitúdó nem, csupán a jelfázis számít.

Időtartományban vizsgálva a hangforrás létrehoz egy $p(t)$ jelet, mely a forrás és a mikrofon közötti d távolságon keresztülhaladva jut el a mikrofonig a közegbeli c terjedési sebességgel. Ezen a d távolságon a hangforrás jelének $\Delta t = \frac{d}{c}$ időt vesz igénybe a mikrofonhoz eljutni, tehát a mikrofon közvetlen közelében, valamint a mikrofon közvetlen kimenetén a hangforrás Δt értlékkel késleltetett függvénye mérhető.

$$x(t) = p(t - \Delta t) = p\left(t - \frac{d}{c}\right)$$

Frekvenciatartományban a forrás és a mikrofon közötti d távolságból, valamint az adott frekvenciához az adott közegben tartozó $k(\omega) = \frac{2\pi}{\lambda}$ hullámszámból következő jelkésleltetés a forrás $P(\omega)$ jelének $\Delta\varphi(\omega) = d \cdot k(\omega) = d \cdot \frac{2\pi}{\lambda(\omega)} = d \cdot \frac{\omega}{c}$ fáziskésésében nyilvánul meg, a mikrofon közvetlen kimenetén is ez mérhető.

$$\begin{aligned} X(\omega) &= \mathcal{F}\{x(t)\} = \mathcal{F}\{p(t - \Delta t)\} = \\ &= P(\omega) \cdot e^{-j\Delta\varphi(\omega)} = P(\omega) \cdot e^{-jd k(\omega)} = P(\omega) \cdot e^{-j \cdot d \cdot \frac{\omega}{c}} \end{aligned}$$

A $h(t)$ impulzusválaszával jellemzett szűrő kimenetén a mikrofon által szolgáltatott jelnek, valamint a szűrő impulzusválaszának a konvolúciója mérhető.

$$y(t) = h(t) * x(t) = h(t) * p\left(t - \frac{d}{c}\right)$$

A $H(\omega)$ átviteli függvényével jellemzett szűrő kimenetén a mikrofonból származó jel spektrumának, illetve az átviteli függvénynek a szorzata jelenik meg.

$$\begin{aligned} Y(\omega) &= \mathcal{F}\{y(t)\} = \mathcal{F}\{h(t) * x(t)\} = \\ &= H(\omega) \cdot X(\omega) = H(\omega) \cdot e^{-j \cdot d \cdot \frac{\omega}{c}} \cdot P(\omega) \end{aligned}$$

A korábban is említettek szerint a térbeli terjedés egy olyan lineáris torzulást eredményez, mely könnyedén modellezhető egy, a mikrofon és forrás közötti szűrő segítségével.

A fent tárgyalt esetekben a jelterjedést modellező szűrő impulzusválasza:

$$d(t) = \delta(t - \Delta t) = \delta\left(t - \frac{d}{c}\right)$$

$$x(t) = p(t) * d(t) = p(t - \Delta t) = p\left(t - \frac{d}{c}\right)$$

$$y(t) = h(t) * x(t) = h(t) * d(t) * p(t)$$

Frekvenciatartományban pedig a jelterjedést modellező szűrő átviteli függvénye:

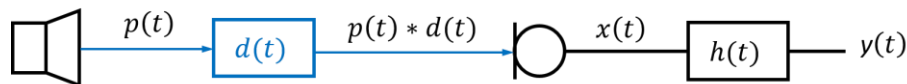
$$\mathcal{F}\{d(t)\} := a(\omega) = e^{-j\Delta k(\omega)}$$

$$X(\omega) = P(\omega) \cdot a(\omega) = P(\omega) \cdot e^{-j\Delta\phi(\omega)}$$

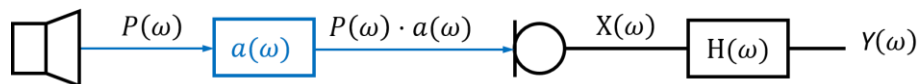
$$Y(\omega) = H(\omega) \cdot X(\omega) = H(\omega) \cdot a(\omega) \cdot P(\omega)$$

Ez a felírasmód a mikrofontömb analitikus kezelése szempontjából nagyon praktikus. A modellezett rendszert a 10. ábra szemlélteti.

Jelterjedés időtartományban



Jelterjedés frekvenciatartományban

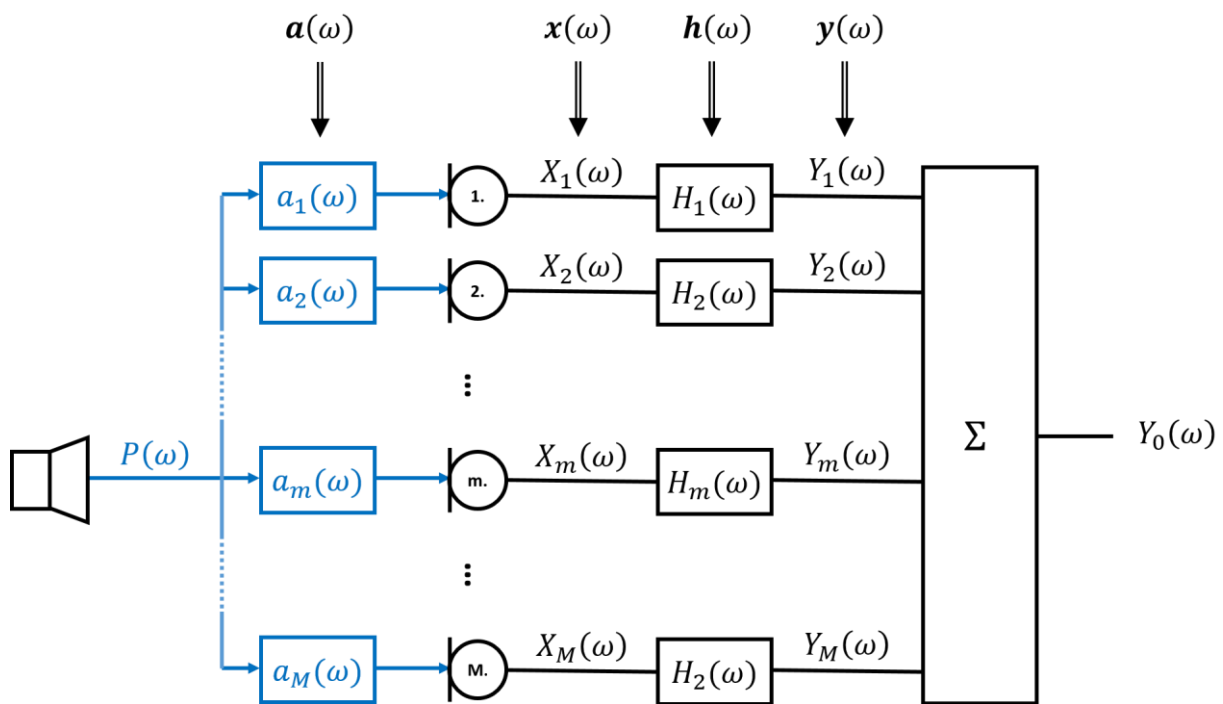


10. ábra Jelterjedési rendszer modellezése idő- és frekvenciatartományban

Amennyiben több mikrofonunk, és több forrásunk van, úgy a helyzet bonyolultabb, hiszen minden mikrofon-jelforrás pár rendelkezik egy saját terjedési szűrővel, valamint több mikrofon terjedést követő jele egyszerre jut egy mikrofonba.

5.1 Több mikrofon egyetlen forrással

Abban az esetben, ha M mikrofonunk van, és azok együttesen veszik egyetlen hangforrás jelét, a 11. ábra által bemutatott rendszermodell állítható fel.



11. ábra M mikrofon, egyetlen forrás

Amennyiben az egyes mikrofonoktól mért távolság d_m , a mikrofonok közvetlenkimenetein megjelenő jelek:

$$X_m(\omega) = e^{-j \cdot d_m \frac{\omega}{c}} \cdot P(\omega) = a_m(\omega) \cdot P(\omega)$$

Ezek szűrését, majd összegzését követően pedig a kimenet:

$$Y(\omega) = \sum_{m=1}^M Y_m(\omega) = \sum_{m=1}^M H_m(\omega) \cdot X_m(\omega) = \sum_{m=1}^M H_m(\omega) \cdot a_m(\omega) \cdot P(\omega)$$

Ez vektoros formában is megfogalmazhatjuk, amennyiben definiáljuk az \mathbf{a} jelterjedési-, valamint \mathbf{h} súlyvektort a következő módon:

$$\mathbf{a}^T(\omega) = [a_1(\omega), \quad a_2(\omega), \quad \dots a_m(\omega), \quad \dots a_M(\omega)]$$

$$\mathbf{x}^T(\omega) = [X_1(\omega), \quad X_2(\omega), \quad \dots X_m(\omega), \quad \dots X_M(\omega)]$$

$$\mathbf{h}^T(\omega) = [h_1(\omega), \quad h_2(\omega), \quad \dots h_m(\omega), \quad \dots h_M(\omega)]$$

$$\mathbf{y}^T(\omega) = [y_1(\omega), \quad y_2(\omega), \quad \dots y_m(\omega), \quad \dots y_M(\omega)]$$

ezek segítségével:

$$\mathbf{x}(\omega) = \mathbf{a}(\omega) \cdot P(\omega)$$

A mikrofontömb kimenetén megjelenő jel:

$$Y_0(\omega) = \mathbf{h}^T(\omega) \cdot \mathbf{x}(\omega) = \mathbf{h}^T(\omega) \cdot \mathbf{a}(\omega) \cdot P(\omega)$$

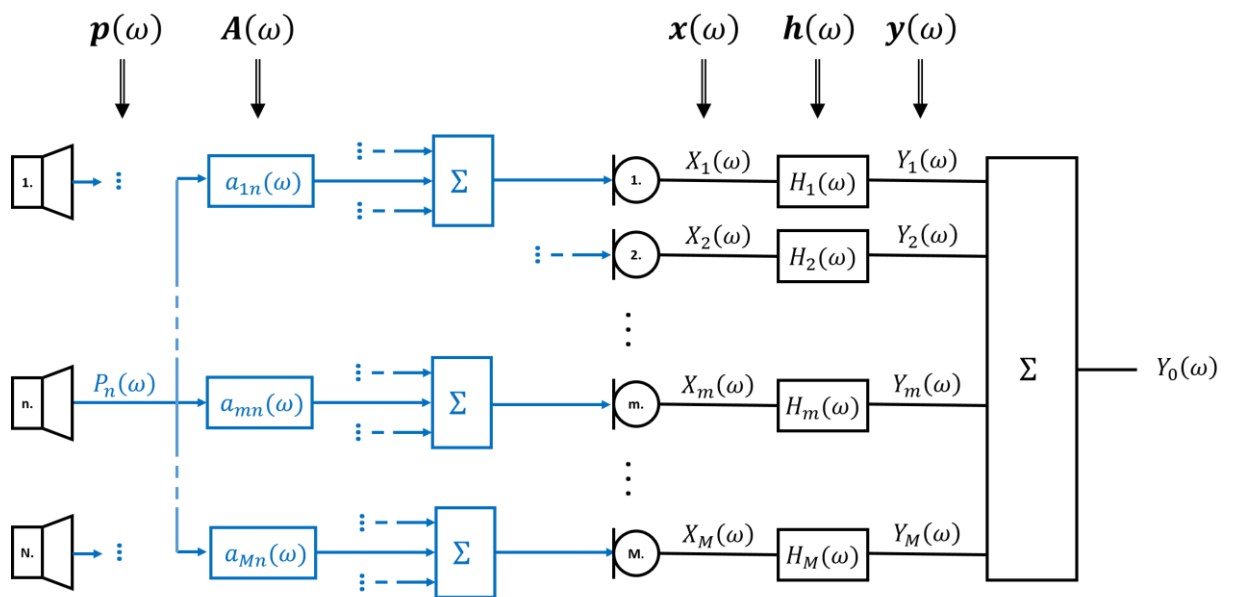
Annak érdekében, hogy az összefüggések áttekinthetőbbek legyenek, ilyen esetben elhagyom az omegától való függés jelölését, a vektoros felírások pedig ezt követően mindig frekvenciatartományban értendők:

$$Y_0 = \mathbf{h}^T \cdot \mathbf{x} = \mathbf{h}^T \cdot \mathbf{P} \cdot \mathbf{a}$$

5.2 Több forrás, több mikrofon

Amennyiben egy M elemű mikrofonrendszerünk, illetve N db hangforrásunk van, úgy az egyes források mindegyikének a jele eljut a rendszer minden mikrofonjához a feljebb tárgyalt módon úgy, hogy az n-edik forrás és az m-edik mikrofon közötti távolságot d_{mn} , az általa megvalósított frekvenciatartománybeli szűrőt pedig $a_{mn}(\omega)$ jelöli.

A rendszer modelljét 12. ábra mutatja.



12. ábra Több forrást, és több mikrofont modellező rendszer

Ez esetben egyetlen mikrofon közvetlen kimenetén megjelenő jel:

$$x_m(t) = \sum_{n=1}^N p_n \left(t - \frac{d_{mn}}{c} \right)$$

ami frekvenciatartományban

$$X_m(\omega) = \sum_{n=1}^N P_n(\omega) \cdot a_{mn}(\omega)$$

Ezek szűrt kimeneteinek az összege az alábbiak szerint írható fel:

$$y_0(t) = \sum_{m=1}^M h_m(t) * x_m(t) = \sum_{m=1}^M \sum_{n=1}^N h_m(t) * p_n\left(t - \frac{d_{mn}}{c}\right) = \sum_{m=1}^M \sum_{n=1}^N h_m(t) * d_n(t)$$

illetve

$$Y_0(\omega) = \sum_{m=1}^M H_m(\omega) \cdot X_m(\omega) = \sum_{m=1}^M H_m(\omega) \cdot \sum_{n=1}^N P_n(\omega) \cdot a_{mn}(\omega)$$

Az egyes források jeleit $\mathbf{p}(\omega)$ forrásvektorként, az egyes forrásokhoz tartozó \mathbf{a}_n jelterjedési vektorokat pedig az $\mathbf{A}(\omega)$ jelterjedési mátrixba fűzve írhatjuk fel.

$$\begin{aligned} \mathbf{p}^T(\omega) &= [P_1(\omega), \quad P_2(\omega), \quad \dots P_m(\omega), \quad \dots P_N(\omega)] \\ \mathbf{A}(\omega) &= [\mathbf{a}_1, \mathbf{a}_2, \dots \mathbf{a}_n, \dots \mathbf{a}_N] = \begin{bmatrix} a_{11} & \dots & a_{1N} \\ \vdots & \ddots & \vdots \\ a_{M1} & \dots & a_{MN} \end{bmatrix} = \\ &= \begin{bmatrix} e^{-jk(\omega) \cdot d_{11}} & \dots & e^{-jk(\omega) \cdot d_{1N}} \\ \vdots & \ddots & \vdots \\ e^{-jk(\omega) \cdot d_{M1}} & \dots & e^{-jk(\omega) \cdot d_{MN}} \end{bmatrix} \end{aligned}$$

Így a korábbi összefüggéseink felírása egyszerűbb alakot ölt:

$$\mathbf{x}(\omega) = \mathbf{A}(\omega) \cdot \mathbf{p}(\omega)$$

$$\mathbf{x} = \mathbf{A} \cdot \mathbf{p}$$

$$Y_0(\omega) = \sum_{m=1}^M h_m \cdot x_m = \mathbf{h}^T \cdot \mathbf{A} \cdot \mathbf{p}$$

Ezen kívül bevezetésre kerül a későbbiek során alkalmazott \mathbf{w} vektor:

$$\mathbf{w}(\omega) = \mathbf{h}^*(\omega) = \text{conj}(\mathbf{h}(\omega))$$

$$Y_0(\omega) = \mathbf{h}^T \cdot \mathbf{x} = \mathbf{w}^H \cdot \mathbf{x} = \mathbf{w}^H \cdot \mathbf{A} \cdot \mathbf{p}$$

5.3 Kovarianciamátrix

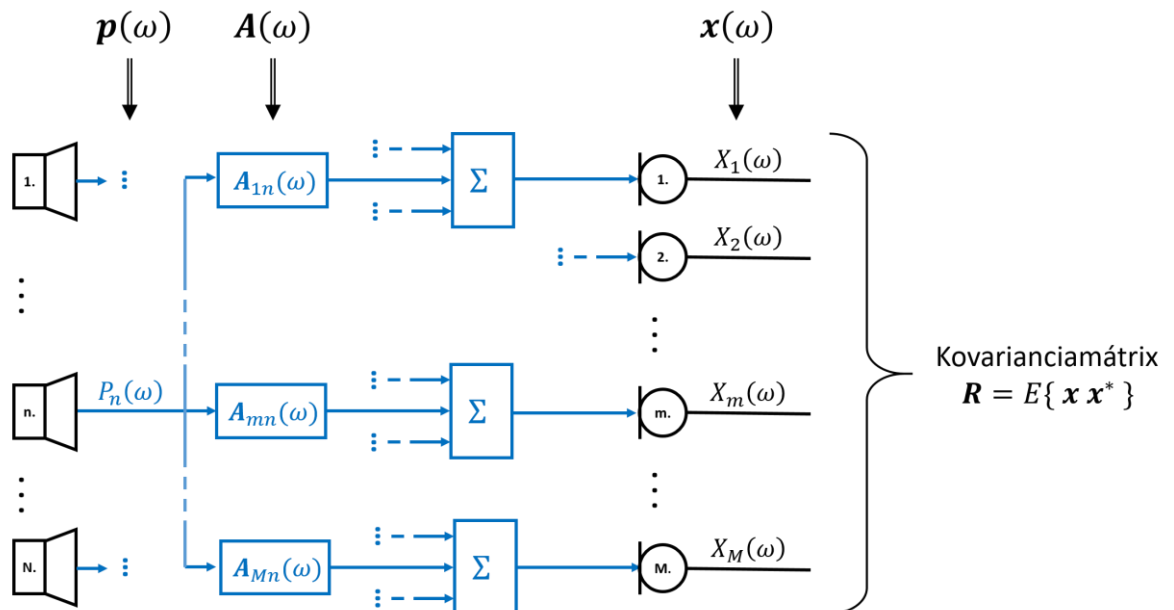
A kimeneten megjelenő teljesítmény a korábbiakban bevezetett jelölések használatával az alábbi módon adható meg:

$$S = |Y_0|^2 = (\mathbf{w}^H \cdot \mathbf{x}) \cdot (\mathbf{w}^H \cdot \mathbf{x})^H = (\mathbf{w}^H \cdot \mathbf{x}) \cdot (\mathbf{x}^H \cdot \mathbf{w}) = \mathbf{w}^H \mathbf{x} \mathbf{x}^H \mathbf{w}$$

Az egyenlet végén megjelenő, vett jel vektor önmaga komplex konjugált transzponáltjával vett szorzata a mikrofonrendszer *Kovarianciamátrixát* eredményezi, melynek definíciója tehát az alábbi:

$$\mathbf{R} = E\{\mathbf{x} \cdot \mathbf{x}^H\} = E\left\{\sum_{m=1}^M \mathbf{x}_m \cdot \mathbf{x}_m^H\right\}$$

A kimenet kompakt leírása során a mikrofonrendszert érő jelek viszonyáról ez az $M \times M$ -es komplex mátrix hordoz információt, a \mathbf{w} súlyvektort kizárólag a mikrofonokat követő szűrők határozzák meg, az a bemenettől független. \mathbf{R} meghatározása történhet analitikusan a fent leírt módon a mikrofonok, illetve hangforrások pozíciói alapján frekvenciatartományban, vagy az időtartománybeli jelek alapján közelítőleg, azok várható értékének számításával.



13. ábra Kovarianciamátrix származtatása

A mikrofonjelek \mathbf{R} kovarianciamátrixának felírása során azt tovább bontva, felhasználva, hogy $\mathbf{x} = \mathbf{A} \cdot \mathbf{p}$, megkaphatjuk a forrásjelek \mathbf{P} kovarianciamátrixát az alábbi módon:

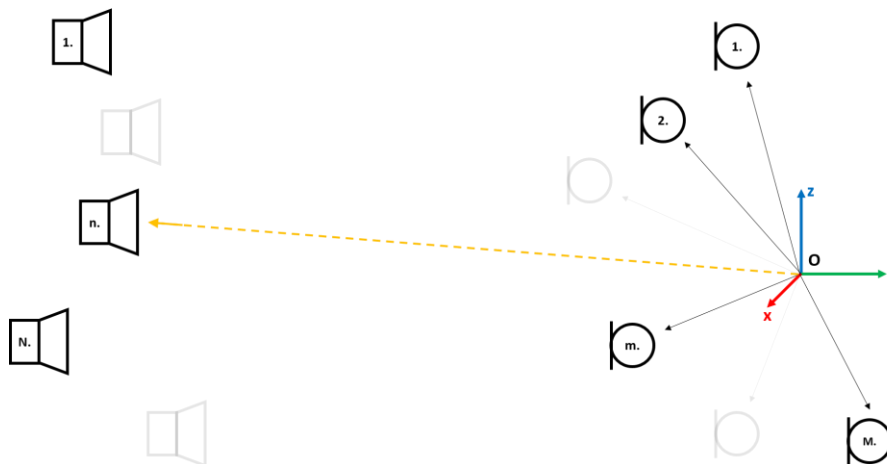
$$\mathbf{R} = \mathbf{x} \cdot \mathbf{x}^H = (\mathbf{A} \cdot \mathbf{p}) \cdot (\mathbf{A} \cdot \mathbf{p})^H = (\mathbf{A} \cdot \mathbf{p}) \cdot (\mathbf{p}^H \cdot \mathbf{A}^H) = \mathbf{A} \cdot \mathbf{p} \mathbf{p}^H \cdot \mathbf{A}^H = \mathbf{A} \mathbf{P} \mathbf{A}^H$$

$$\mathbf{P} = E\{\mathbf{p} \cdot \mathbf{p}^H\} = E\left\{\sum_{n=1}^N \mathbf{p}_n \cdot \mathbf{p}_n^H\right\}$$

A későbbiekben abból indulunk ki, hogy az egyes források jelei egymástól függetlenek, teljesen korrelálatlanok. Ez alapján a forrásjelek kovarianciamátrixának egy olyan egységmátrix adódik, melynek átlójában az egyes források fázorjának hosszánegyzete áll.

5.4 Nyalábformáló algoritmusok

Az olyan eljárásokat, melynek során egy mikrofonrendszerhez tartozó iránykarakterisztika alakja a mikrofontömbhöz tartozó súlyvektor értékeinek hangolásán keresztül állítható, nyalábformálásnak nevezzük. A nyalábformálás megvalósítására, annak érdekében, hogy az iránykarakterisztika a számunkra lehetőleg kedvező legyen, különféle elvárások fogalmazhatók meg, és ezek kielégítésére különféle matematikai eljárásokon alapuló algoritmusok találhatók.



14. ábra Mikrofontömb, valamint annak környezetében elhelyezkedő hangforrások

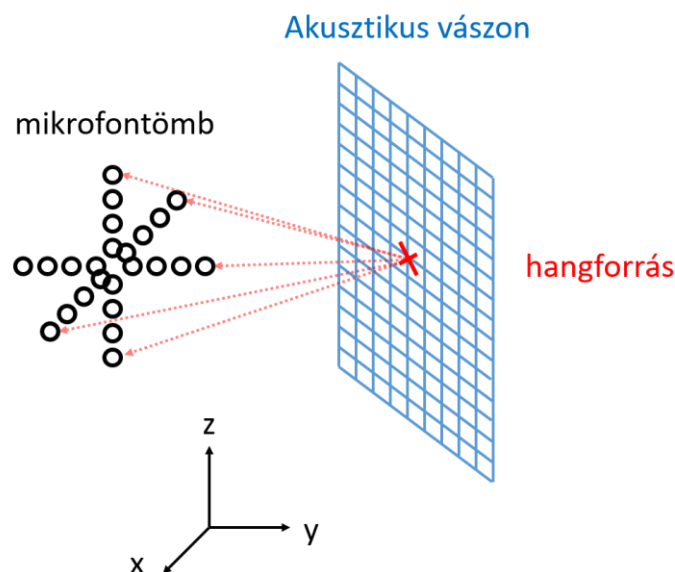
5.5 Akusztikus vászon

Az eddigiek során a mikrofontömb rendszerleírásakor, valamint a szimulációk során abból a feltételezésből indultunk ki, hogy az egyes források a mikrofontömb kiterjedéséhez képest nagyságrendekkel távolabb vannak, így az azokból származó jelek

tulajdonképpen síkhullámokként terjednek. Az iránykarakterisztika is ennek megfelelően lett definiálva, azonban az a terjedési vektor értelmezhető olyan forrásokra is, melyek távolsága összemérhető a mikrofontömb térbeli kiterjedésével, a nyalábformáló algoritmusok számára ez nem jelent nagyobb problémát ahhoz képest, mintha az végtelen távol lenne.

Azonban az így végzett nyalábformálás eredményeképpen kialakuló (végtelenben vizsgált) iránykarakterisztikán általában nem lehet a főnyaláb határozott erősítését, illetve a nullhelyek elnyomását jól megfigyelni, hiszen a főnyaláb, valamint nullhely ilyen esetben nem irányt, hanem térbeli ponthalmazokat, illetve azok környezetét jelöli, ahol a mikrofontömb erősítése maximális, illetve nulla. Tehát a mikrofontömb által végzett nyalábformálás eredményeképpen az iránykarakterisztika helyett egy olyan térbeli erősíteloszlás-függvényt kapunk, mely megadja, hogy egy adott térbeli pontban az erősítés mekkora értékű, és ez változik az origótól vett távolság függvényében, hiába azonos azimut, illetve eleváció szerint haladunk.

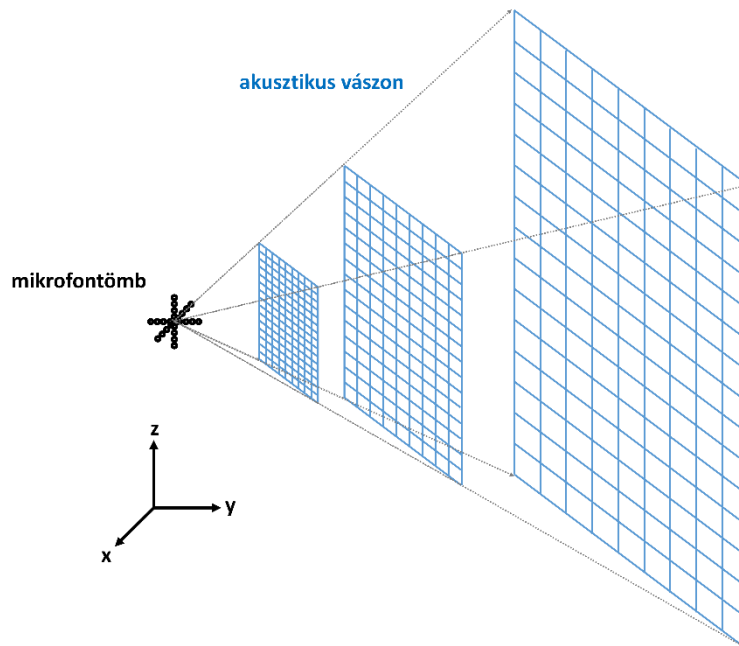
Ennek ábrázolása céljából definiáljuk az *akusztikus vászon* fogalmát, ami azt jelenti, hogy a térbeli pontoknak kiválasztjuk egy olyan halmazát, amiket egyértelmű módon meg tudunk feleltetni síkbeli pontok halmazának, így lehetőség van azok ábrázolására síkbeli felületen.



15. ábra Akusztikus vászon

Az akusztikus vászon az egyszerűség, valamint könnyebb értelmezhetőség kedvéért rendszerint maga is síkbeli, azonban amennyiben a hangforrások nem esnek egy síkba, esetleg a mozgásuk során abból elmozdulnak, úgy a síkbeli akusztikusra vászonra képződő iránykarakterisztika már nem feltétlen tükrözi azt az erősítési értéket, amit az adott jelforrás valójában tapasztal.

Az akusztikus vászon az origó középpontú nagyításával, különféle vászontávolságok mellett, az adott azimut-eleváció szélsőértékeknek megfelelő irányokba vett erősítésértékeket kaphatjuk meg. Amennyiben egy térben elhelyezett pontforrásra végeztünk nyalábformálást, úgy tulajdonképpen ennek a formált nyalábnak a metszetét láthatjuk különféle vászontávolságok mellett. Ez a kép akkor rendelkezik a legjobb erősítésekkel, mikor a hangforrás éppen a vászonra esik. Ezekon kívül a karakterisztika nem rendelkezik határozott körvonalakkal, elmosódott hatást kelt.



16. ábra Akusztikus vászon nagyobb távolságokban

Tulajdonképpen teljesen analóg a helyzet azzal, mint mikor például egy projektorral vetített képet próbálunk vászonra fókuszálni.

Természetesen kvázi végtelen távolra helyezett vászon esetén megkapjuk az eredetileg is definiált iránykarakterisztikát, így az akusztikus vászon alkalmazása egy általánosítása az iránykarakterisztika fogalmának, viszont elengedhetetlen abban az esetben, ha hiteles képet akarunk kapni az algoritmusunk működéséről.

6 Bartlett nyalábformálás

Amennyiben a mikrofonrendszer iránykarakteristikájával szemben támasztott elvárás csupán annyi, hogy az egy adott pontból érkező jeleket a lehető legnagyobb erősítéssel kerüljenek a kimenetre, úgy a legegyszerűbb módszer, ha az ebből a pontból az egyes mikrofonokhoz érkező jeleket épp annyival késleltessük, amennyivel az az egyes

mikrofonokhoz egy bizonyos referenciamikrofonnál korábban érkezik meg. Így elérhető, hogy a forrásirányból az egyes mikrofonra érkező jelek fázishelyesen összegződjenek, így az adott pontból érkeve a konstruktívan erősítsék egymást. Mindeközben más pontokból esetlegesen beérkező jelekre nem feltétlenül fog teljesülni a fázishelyes összegzés, így azok esetében valamilyen mértékű elnyomásra lehet számítani a kimeneten.

A korábbi fejezetben bemutatott fogalmak segítségével a Bartlett nyálábformálás jól leírható. Tegyük fel, hogy egy hangforrás $P(\omega)$ mikrofontömbre eső jelét szeretnénk a nyálábformálás segítségével konstruktívan összegezni. Ennek a jelnek a mikrofontömb kimenetén mérhető hatása az alábbi:

$$Y(\omega) = \mathbf{w}^H(\omega) \cdot \mathbf{x}(\omega) = \mathbf{w}^H(\omega) \cdot P(\omega) \cdot \mathbf{a}(\omega)$$

Ennek értelmében mivel a forgatóvektor elemei kizárólag különböző fázisú, egységnyi hosszú komplex forgató értékek, így az ezekkel való szorzás hatását úgy lehet kiküszöbölni, ha a súlyvektorunk minden eleme a forgatóvektor neki megfelelő elemének komplex konjugáltja. Általánosabban azt lehet mondani, hogy akkor kapható meg a mikrofontömb kimenetén $P(\omega)$ jel, ha

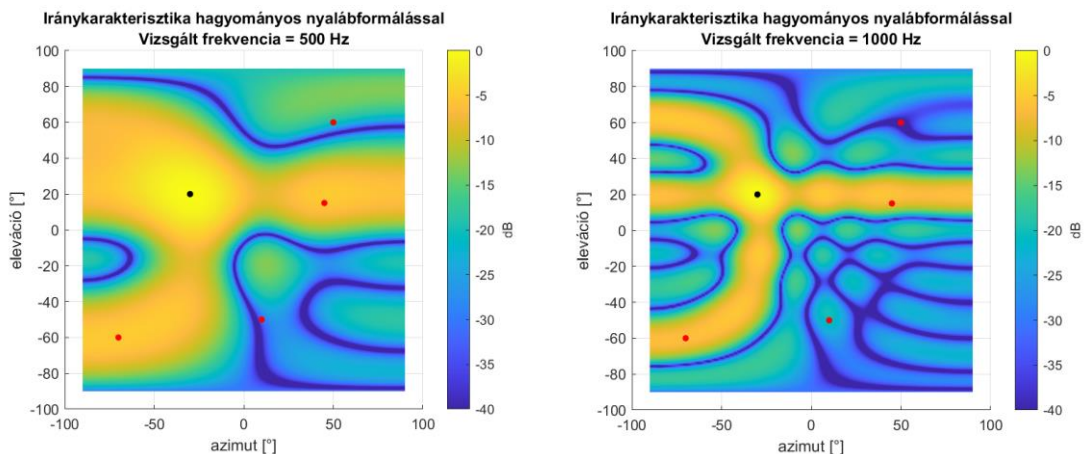
$$\mathbf{h}_{opt} = \frac{1}{M} \cdot [a_1^{-1}, \quad a_2^{-1}, \quad \dots a_m^{-1}, \quad \dots a_M^{-1}]^T = \frac{1}{M} \cdot \mathbf{a}^*$$

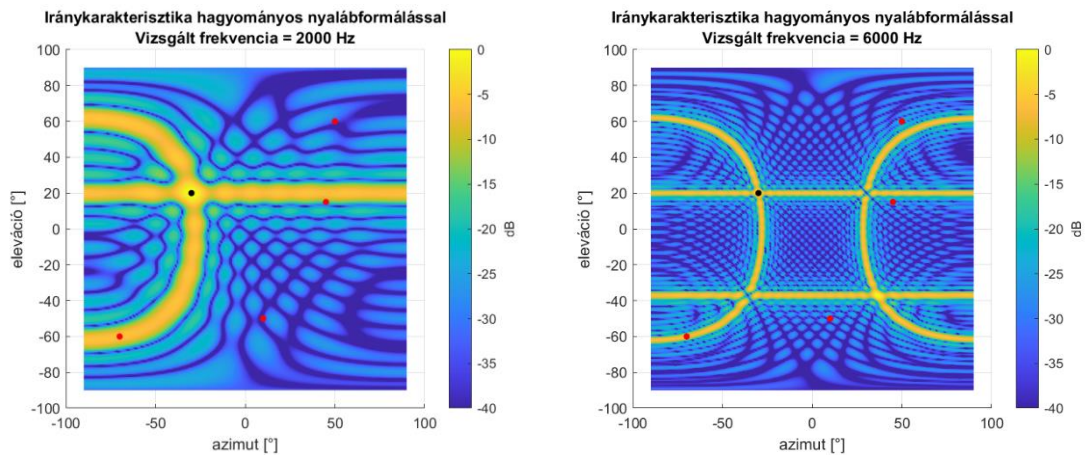
$$\mathbf{w}_{opt} = \mathbf{h}_{opt}^* = \frac{1}{M} \mathbf{a}$$

így

$$\mathbf{w}_{opt}^H \cdot P(\omega) \cdot \mathbf{a} = P(\omega)$$

Tehát a mikrofontömb kimenetén az adott hangforrás irányában az erősítés egységnyi, más irányban viszont ez nem feltétlen teljesül, hiszen azokhoz ennek érdekében más optimális súlyvektor tartozna. A nyálábformált mikrofontömb egyes irányokra vonatkozó erősítési tényezőjéről számot adó iránykarakterisztikákat a 17. ábra mutatja különféle frekvenciákon.





17. ábra Kialakuló iránykarakterisztikák

Látható, hogy a Bartlett nyalábformálás eredményeképpen a fekete koronggal jelölt forrás irányába helyezkedik el az úgynevezett főnyaláb, amerre a mikrofontömb erősítése egységnyi. A karakterisztikát görbék mentén átszelő keskeny nullhelyek figyelhetők meg, ahol az adott frekvencián a késleltetésekből következően teljes kioltás tapasztalható. Az egyes nullhelyek felosztják a teljes térszögtartományt, közöttük lévő térrészben találhatóak nem nulla, de főnyalábnál jelentősen kisebb erősítésű irányok, melyeket melléknyaláboknak nevezünk.

Megfigyelhető, hogy bár a főnyaláb a kijelölt forrás irányába néz, az ilyen módon elért jel-zaj viszony a melléknyalábokon keresztül beszivárgó zavaró jelek miatt nem a lehető legjobb.

A nullhelyek irányába eső forrásokból az adott frekvencián természetesen semmi nem jut a kimenetre, viszont megfigyelhető, hogy a vizsgálati frekvencia függvényében az iránykarakterisztika is változik, a nullhelyek átrendeződnek, a melléknyalábokat tartalmazó sávok keskenyebbekké, a nullhelyvonalak pedig sűrűbbé válnak. Ennek eredménye, hogy bár különböző frekvenciákon vizsgálva a főnyaláb mindig a megadott irányba néz, a nullhelyek, és melléknyalábok frekvenciafüggő átrendeződéséből adódóan zavaró források jele mindig megjelenik a kimeneten, csupán keresztülhaladva a mikrofontömb frekvenciafüggő iránykarakterisztikájából adódó szűrőjén.

Feltűnő jelenség tapasztalható 6000 Hz esetén, ahol az egészen keskenyre szűkült főnyaláb mellett megjelenik egy hasonlóan nagy erősítésű irányosáv, melynek zavaró forrás irányába esése súlyos mértékben rontja a kimeneti jel-zaj viszonyt. Ez a térbeli mintavételi tétel megsérüléséből következik, az adott vizsgálati frekvenciához tartozó hullámhossz legalább kétszerese kell legyen a mikrofonok közötti legkisebb távolságnak, különben az időtartományból ismert átlapolódási jelenség térszögtartományi megfelelője következik be, tehát a főnyaláb ismételt megjelenik az iránykarakterisztikán, ami semmiképpen sem tekinthető kedvező jelenségnek.

Bár a nyalábformálás főiránya folyamatosan változtatható, vele együtt az adott frekvencián kialakuló melléknyalábok, és nullhelyek is fordulnak, aminek eredménye,

hogy zavaró források belekerülhetnek melléknyalábokba, illetve kikerülhetnek nullhelyekről.

A Bartlett nyalábformálás legnagyobb előnye az egyszerűsége, viszont a fent említett előnytelen tulajdonságok miatt használata kritikus esetekben, mikor a releváns forrás jelszintje alacsony, a további forrásokból származó zajszint magas, vagy a vételt esetleg valamilyen nagy teljesítményű jel zavarja. Helyette inkább tehát további nyalábformálási algoritmusok alkalmazását érdemes számba venni. Ilyen például az úgynevezett *Capon* módszer, melynek részletes tárgyalása a következő fejezetben olvasható.

Fontos megjegyezni, hogy bár az iránykarakterisztika frekvenciafüggő, magához a Bartlett nyalábformáló algoritmushoz nem feltétlen szükséges frekvenciatartományban dolgozni, hiszen csupán az egyes mikrofonértékeket kell olyan módon késleltetni, hogy az a kívánt hatást elérje. Az algoritmus egyszerűségéhez ez is hozzájárul.

7 Capon algoritmus

A Capon algoritmus [7] alap gondolata a mikrofontömb kimeneti teljesítményére, és az adott főirányban mérhető erősítés nagyságára tesz megkötéseket, melyek alapján bevezetésre kerül egy lineáris szélsőértékkeresési probléma, melynek megoldásával egy olyan súlyvektorhoz lehet jutni, mely a kezdeti feltételeket teljesíti. Ennek hatására a Bartlett nyalábformálás esetén vázolt problémák kiküszöbölhetőek, ugyanis a Capon algoritmus megfelelő alkalmazásával elérhető, hogy a főirányból érkező jelek torzítatlanul, az azon kívül elhelyezkedő zavaró források jelei pedig teljes kioltással kerüljenek a kimenetre minden frekvencián. További előnyös tulajdonság, hogy az algoritmus ezt az ideális iránykarakterisztikához vezető súlyvektorokat olyan módon képes meghatározni, hogy számára az egyes mikrofonokon megjelenő jeleken kívül kizárólag a főirányt határozzuk meg. A zavarforrások irányába kerülő nullhelyek automatikusan következnek az algoritmus működéséből, tehát azok esetleges elmozdulással sem tudnak hosszútávon kikerülni a teljes kioltást jelentő nullhelyekről.

A standard Capon algoritmusra vonatkozó konkrét matematikai leírás alapján a \mathbf{w}_{opt} súlyvektorral szemben támasztott elvárás úgy fogalmazható meg, hogy az általa alakított kimeneten megjelenő teljesítmény a lehető legkevesebb legyen, miközben a főirány erősítése egységnyi.

Matematikailag felírva:

$$\min_{\mathbf{w}} \mathbf{w}^H \mathbf{R} \mathbf{w} \text{ úgy, hogy } \mathbf{w}^H \mathbf{a}_0 = 1$$

ennek eredménye [8] [9] alapján:

$$\mathbf{w}_{opt} = \frac{\mathbf{R}^{-1} \mathbf{a}_0}{\mathbf{a}_0^H \mathbf{R}^{-1} \mathbf{a}_0}$$

ezek alapján pedig felírható, hogy a releváns irányból származó teljesítmény ilyen módon:

$$P(a) = \mathbf{w}^H \mathbf{R} \mathbf{w} = \frac{1}{\mathbf{a}_0^H \mathbf{R}^{-1} \mathbf{a}_0}$$

A korábban már említett \mathbf{R} kovarianciamátrix fontos szerepet játszik a súlyvektor kiszámítása során, hiszen ezen keresztül van információ a mikrofontömböt érő jelteljesítményekről, ami alapján a nullhelykiosztás elvégezhető.

Első közelítésben a kovarianciamátrixot frekvenciatartományban, az egyes hangforrások pozíciójának ismeretében határoztam meg. Így a következő adódott:

$$\mathbf{x} = \sum_{n=1}^N \mathbf{a}_n p_n = \mathbf{A} \mathbf{p}$$

$$\mathbf{R} = \mathbf{x} \mathbf{x}^H$$

Látható, hogy az \mathbf{A} terjedési mátrix N db diád összegeképp adódik, tehát ennek értelmében a kovarianciamátrix rangja N , mérete viszont $M \times M$ -es. Az algoritmus azonban a kovarianciamátrix inverzével dolgozik, ami viszont nem valósítható meg abban az esetben, ha a mátrix rangja nem teljes. Ez azt jelenti, hogy a Capon módszer nem működik olyan esetben, mikor a különböző zavaró forrásirányok száma nem éri el az alkalmazott mikrofonok számát, vagy valamilyen egyéb módon nem bővül teljes rangúvá a kovarianciamátrix.

Érdemes megjegyezni, hogy a források által adódó N db diád összege csak akkor alkot N -ed rangú mátrixot, ha a diádképzéshez felhasznált vektorok, tehát a terjedési vektorok lineárisan függetlenek. Ez persze azt jelenti, hogy az egyes források iránya eltérő, tehát nincs két olyan forrás, mely ugyanabból az irányból szólna, hiszen ekkor ezeket össze lehetne vonni egyetlen forrássá, vagy fordítva: nem lehet \mathbf{R} rangját olyan módon bővíteni, hogy egyetlen forrás jelét két jel szuperpozíciójának tekintem.

Szintén említésre méltó, hogy a KM rangja által szabott korlát a Capon algoritmus működésére azért jogos, mivel az M elemű mikrofontömb felhasználásával az iránykarakterisztikán M db nullsáv helyezhető el. Abban az esetben, ha nekünk M -nél kevesebb zajforrásunk van, úgy több olyan megoldás is lehetséges, hogy a kimenő teljesítmény minimalizálása érdekében az egyes forrásokra nullhelyek kerüljenek, így az

algoritmusnak nincs egyetlen egzakt megoldása, ami tehát úgy nyilvánul meg, hogy a KM nem invertálható.

Gyakorlati körülmények között az említett probléma nem tapasztalható, ugyanis az egyes mikrofonokon kiküszöbölhetetlenül megjelenő termikus, vagy egyéb forrásból származó független, adott szigma szórású zajok a kovarianciamátrixhoz egy $M \times M$ méretű diagonális mátrix formájában járulnak hozzá, mely a kovarianciamátrix rangját azonnal teljessé teszi, így az invertálhatóvá, az algoritmus pedig alkalmazhatóvá válik.

Az algoritmus egyik súlyos alkalmazási problémája, hogy bár az általunk meghatározott irányba mindenképpen kialakul az egységnyi erősítés, azonban amennyiben valamilyen hiba következtében a hangforrásunk nem pontosan abban az irányban található, úgy arra az algoritmus automatikusan nullhelyet illeszt, hiszen így képes a kimeneten megjelenő teljesítményt minimalizálni, és a releváns egységnyi erősítését is megtartani. Ennek kiküszöbölése érdekében lett kifejlesztve a *Robust Capon algoritmus*, mely a forrás irányának hibájával kapcsolatban megengedőbb, tulajdonképpen egy adott térszögtartományon belül tolerálja az iránymeghatározás hibáját, és azon belül nem végez elnyomást. A Robust módszerre a későbbiekben még visszatérek.

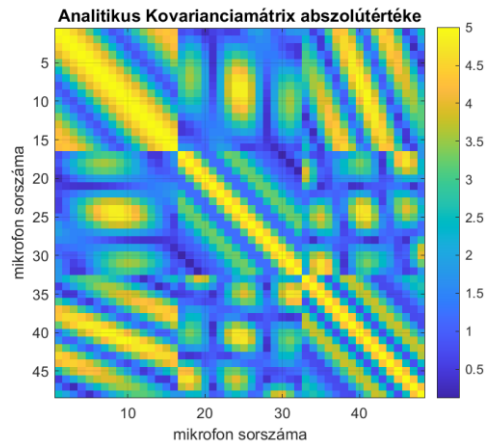
7.1 Capon algoritmus frekvenciatartományból analitikusan

A Capon algoritmus alkalmazásához szükség van a mikrofonjelek R kovarianciamátrixára, mely az alábbi összefüggés alapján megszerezhető analitikusan az A terjedési mátrix, P forrás kovarianciamátrix, illetve σ effektív értékű, mikrofonokon megjelenő független zaj ismeretében.

$$\mathbf{R} = \mathbf{A} \mathbf{P} \mathbf{A}^H + \sigma^2 \mathbf{I}$$

Az így adódó kovarianciamátrix, egy olyan négyzetes, szimmetrikus, komplex elemeket tartalmazó mátrix, melynek sorainak, illetve oszlopainak száma megegyezik a mikrofontömböt alkotó mikrofonok számával.

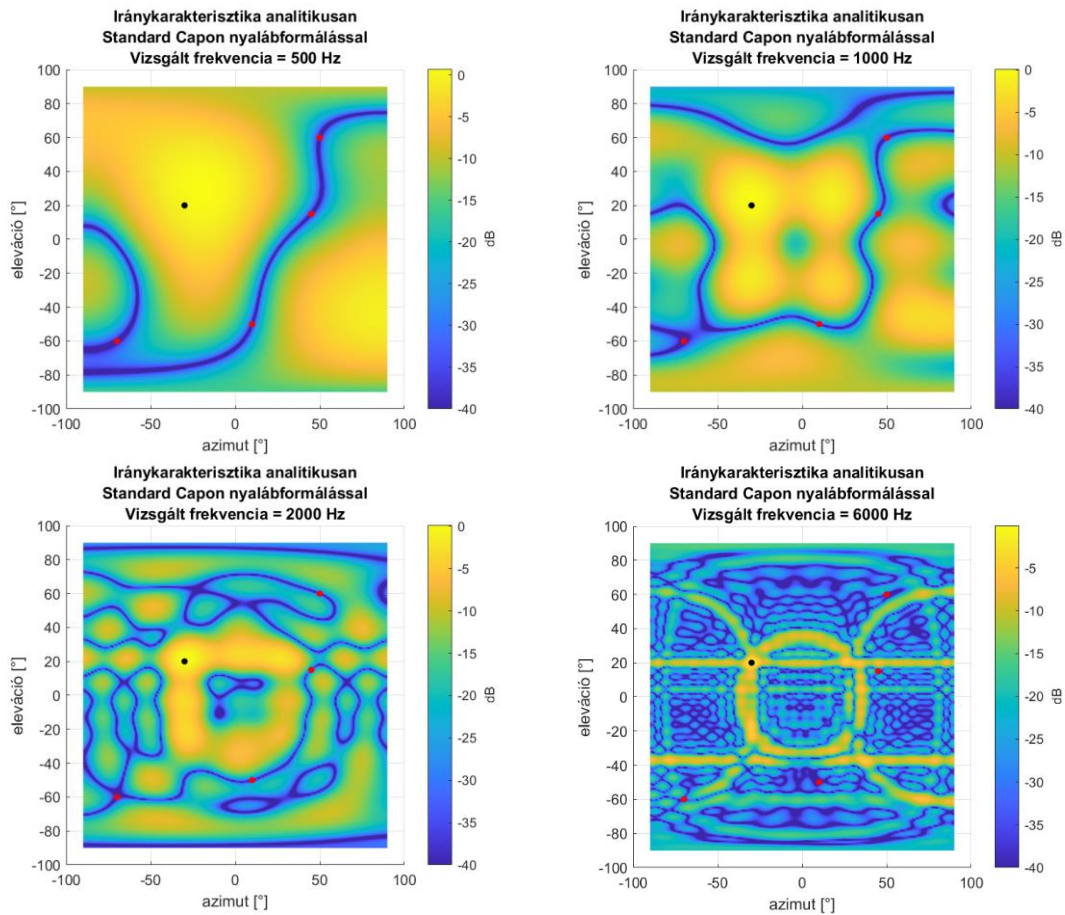
Példaképpen a 18. ábra egy analitikus módon megalkotott kovarianciamátrix abszolútértékét mutatja.



18. ábra Analitikus kovarianciamátrix abszolútértéke

A kovarianciamátrix birtokában lehetőségünk van a Capon algoritmus alkalmazására. A kovarianciamátrix azonban függ a vizsgált frekvenciától, hiszen az közvetlen kapcsolatban áll a terjedési vektorral. Tehát eltérő frekvenciákra eltérő kovarianciamátrixok adódnak ugyanazon rendszer, és akusztikus környezet mellett.

Különböző vizsgálati frekvenciák esetén, analitikus módon szerzett kovarianciamátrixok alapján, Capon algoritmus segítségével kalkulált iránykarakteristikákat a 19. ábra prezentálja.



19. ábra Iránykarakteristikák Capon algoritmussal

Látható, hogy az algoritmus eredményeképpen olyan iránykarakterisztikához juthattunk, melyeken az egyszeres erősítérű főnyaláb az általunk kijelölt irányban található, míg a nullsávok úgy húzódnak, hogy azok az összes zavarforrás irányát érintik, így azok tökéletesen kioltva jelennek meg a kimeneten. Ez a nullhelykiosztás bármilyen frekvencián elérhető, így szélessávú gerjesztés esetére is alkalmas algoritmust kapunk.

Az analitikus megközelítés alapján tehát ez a nyalábformálási módszer rendkívül előnyösnek tűnik, azonban észben kell tartani, hogy jelen esetben a kovarianciamátrixot a releváns, és zavaró forrásirányok ismeretében hoztuk létre. Még így is egészen figyelemreméltó az algoritmus eredménye, viszont az eredeti kikötések alapján a kedvező iránykarakterisztikának a mikrofontömb kovarianciamátrixának felhasználásával abban az esetben is létre kell jönnie, ha kizárólag a releváns forrásirányt határozom meg számára, eredményeképpen pedig a zavaró források irányába nullhelyek kerülnek. Ez persze azt a veszélyt rejti magában, hogy amennyiben a releváns forrásirányt nem kellő pontossággal határozom meg, úgy az algoritmus működéséből adódóan tekintheti azt is zavaró forrásnak, és a kimeneti teljesítmény minimalizálása érdekében létrehoz egy olyan iránykarakterisztikát, ahol a releváns forrásra is nullhely kerül.

7.2 Capon algoritmus időtartományból becsülve

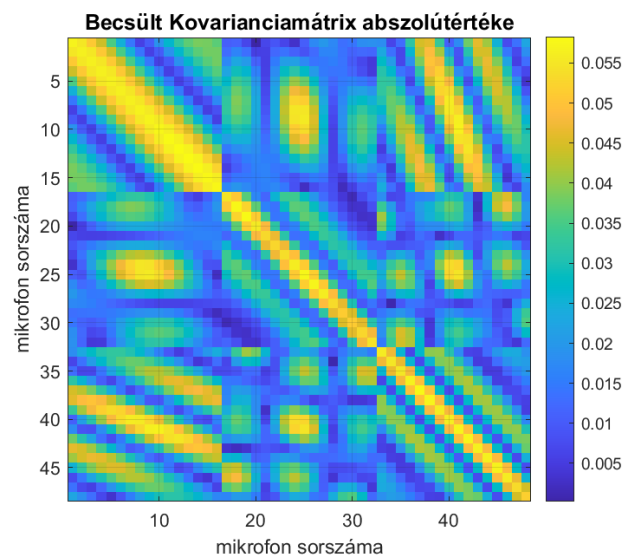
Gyakorlati alkalmazás során, a mikrofontömbön megjelenő jelekkel eredendően időtartományban dolgozunk, és az egyes források mikrofontömbön létrehozott önálló kovarianciamátrixának a meghatározására nincsen lehetőség, csupán a hangtérben összegződött, mikrofonokon megjelenő időtartománybeli jelekkel tudunk operálni.

Az időtartománybeli kovarianciamátrix becsléséhez generáltam minden hangforrásra egy adott hosszúságú Gaussi eloszlású fehérzaj jelet, amiből a jelterjedési mátrix alapján előállítottam a minden mikrofonon megjelenő, egyes forrásokból származó késleltetett jeleket, majd az adott mikrofonokra összegeztem ezeket egy olyan zajjellel együtt, mely a minden mikrofonra egyedi, saját zajjelet modellezi, ennek értéke annyi, hogy a kimeneti jel mellette 30 dB SNR- el jelenik meg. Eredményül minden mikrofonra megkaptam azt a jelet, amit az adott mikrofon az adott elrendezésben tapasztalna.

A mikrofonok közvetlen kimenetén, az időtartománybeli jeleken elvégeztem egy keskenysávú komplex szűrést. Ennek során a mikrofonok kimeneti jeleit lekevertem a vizsgálati frekvenciához tartozó komplex forgó vektorral, majd a lekevert jelet Butterworth aluláteresztő szűrővel kellően keskenysávúvá tettem. Eredményül az adott frekvencián, az egyes mikrofonokon megjelenő komplex alapsávi jelértékeket kaptam, mely az eredeti jel zajjellegéből adódóan időben folytonosan változik. Az egyes források ilyen módon szűrt jelei esetén azt tapasztalhatjuk, hogy a jelek teljesen korrelálatlanok, hiszen a források jelei is teljesen függetlenek. Azonban az egyes források által a

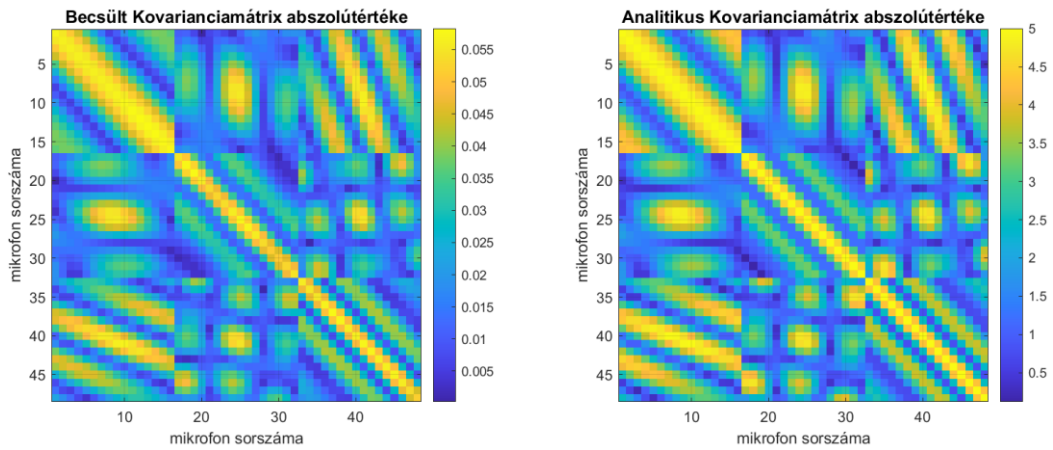
mikrofonokon előállított adott frekvenciájú komplex jelfolyamok nem függetlenek, sőt, kizárólag az adott irányra, és adott frekvenciára vonatkozó terjedési vektoroktól függenek.

Az ilyen módon szűrt, mikrofonok közvetlen kimenetein időtartományban, egyetlen pillanatban adódó komplex értékekből számolt kovarianciamátrix nem fog emlékeztetni az analitikusan kapható eredményre, helyette ezeknek az egy-egy időtartománybeli mintasorból adódó kovarianciamátrixoknak kell venni a várható értékét, amit lehet tenni egyszerűen egy adott intervallumra vonatkozó átlagolással, vagy folyamatos exponenciális átlagolással. Eredményül azt kaptam, hogy 1000 Hz-en, az adott mikrofonelrendezésben vizsgálva néhány másodpercen belül kialakul az analitikus eredményhez hasonló mintázat. A 20. ábra 5 másodpercig tartó exponenciális átlagolás eredményeképpen adódó kovarianciamátrix abszolútértékét mutatja.



20. ábra Becsült kovarianciamátrix abszolútértéke

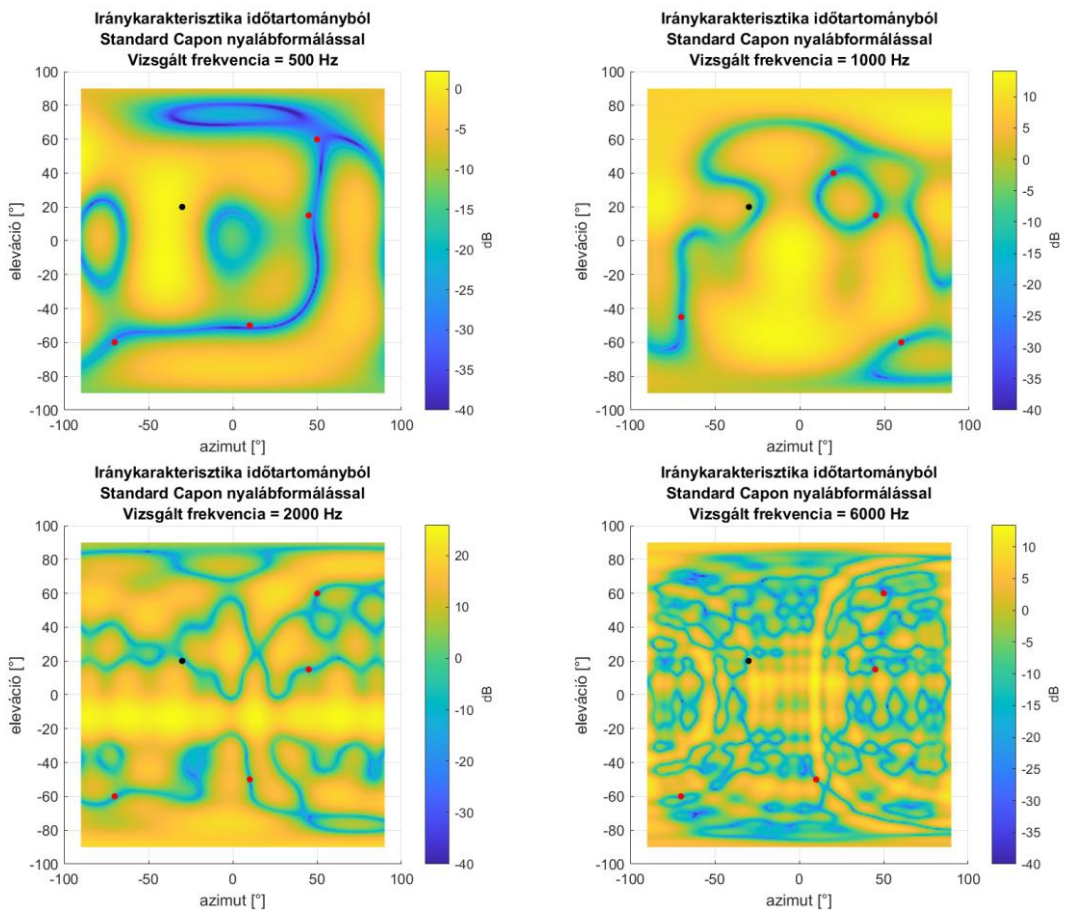
Összehasonlításképpen egymás mellé helyezve az analitikusan, illetve a szimulációs segítségével adódó eredményt mutatja a 21. ábra.



21. ábra Analitikus és Becsült kovarianciamátrixok összehasonlítása

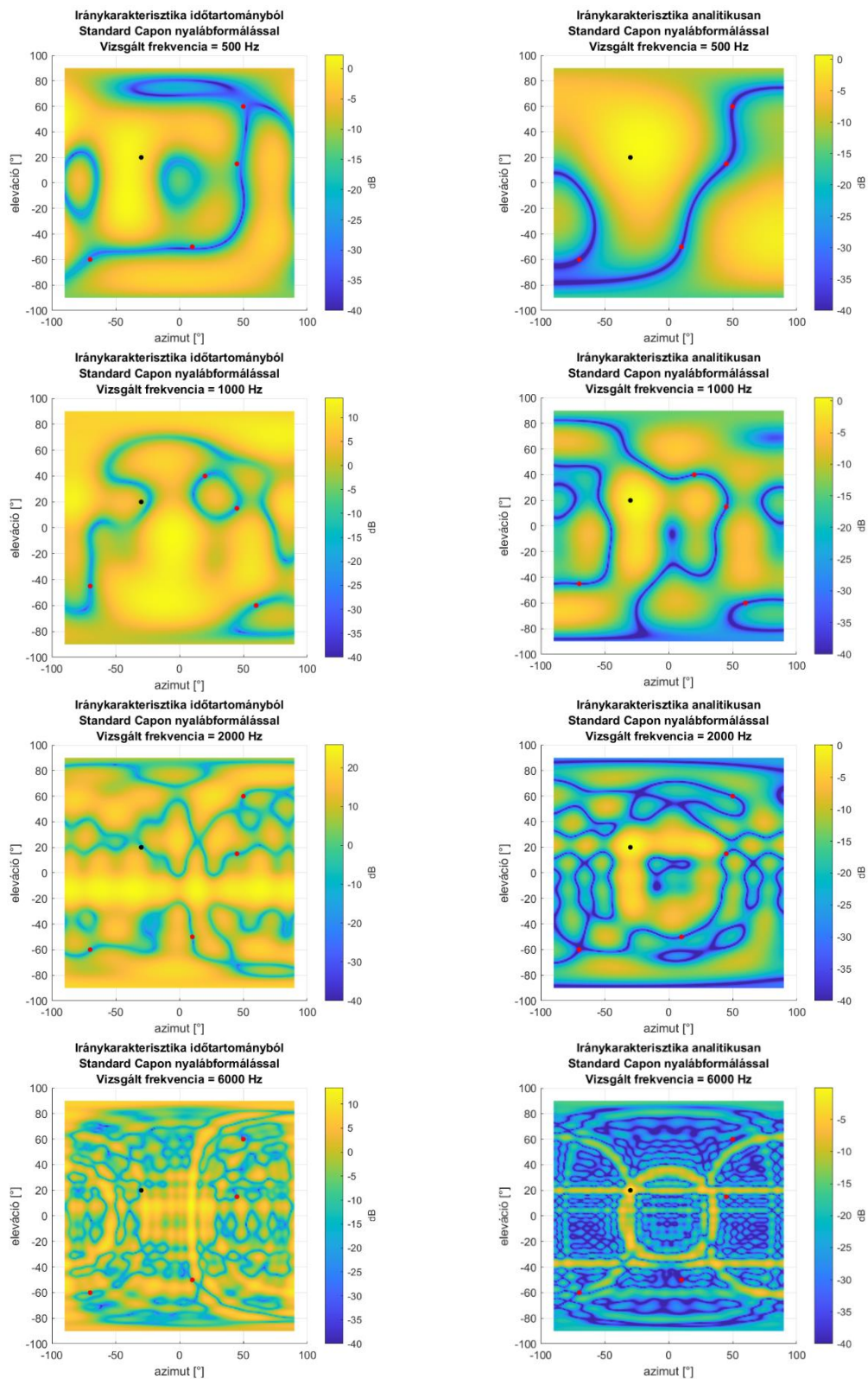
Látható, hogy a két mátrix rendkívül hasonló, és ez igaz a képzetes, illetve valós értékekre is. Az időtartománybeli átlagolás természetesen csak egy közelítő eredményt ad, hiszen teljesen véletlenszerűen változó források jeleiből indulunk ki, viszont az eredményül kapott mátrixok egymáshoz egészen hasonlóak.

Az ezekből számított iránykarakteristikákat pedig a 22. ábra mutatja.



22. ábra Iránykarakteristikák Capon nyálábformálással időtartományból

Ugyanezen a frekvencián alább látható az analitikus, és a szimulált karakterisztika egymás mellett, a jobb összehasonlíthatóság érdekében megtekinthető a 23. ábra segítségével.



23. ábra Analitikus és Becsült karakterisztikák több frekvencián

Látható, hogy az időtartományban modellezett iránykarakterisztika tulajdonképpen egyáltalán nem hasonlít az analitikus párára, főleg a nagyobb vizsgálati frekvenciák esetén. Bár a zavaró forrásokra jellemző, hogy nullhelyre esnek, a megadott irányban pedig egyszeres erősítés tapasztalható, viszont a főnyaláb nem ebbe az irányba esik, sőt, hatalmas erősítések alakulnak ki tulajdonképpen teljesen indokolatlan irányokban.

Ez annak köszönhető, hogy az időtartománybeli vizsgálat eredményeképpen kialakuló kovarianciamátrix nem teljesen azonos az analitikussal, az invertálás eredményeképpen a kis eltérések is nagyon ki hangsúlyozódnak. A kialakult karakterisztikák viszont a specifikációnak megfelelnek, tehát adott irányban egységnyi átvitel, zajforrások irányában pedig komoly elnyomás mérhető. Azonban látható, hogy az algoritmus a releváns irányhoz közel is elvezet egy nullsávot, ami viszont nagy kiemeléseket produkál a karakterisztika más részein. Bár matematikailag ez egy helyes megoldás gyakorlati elrendezésben alkalmazhatatlan, hiszen a nagy erősítésű területeken keresztül könnyedén bezavarhat valami hirtelen felbukkanó zajforrás, illetve a közeli nullsáv miatt, amennyiben a releváns forrás hirtelen odébb mozdul, úgy simán bekerülhet egy nullsávba ameddig a kovarianciamátrix újra nem alakul.

Ez az előnytelen működés korrigálható úgy, ha a mikrofonok saját zaját rendkívüli mértékben növeljük, ekkor ugyanis az algoritmus számára kifejezetten előnytelen lenne a megkövetelt egyszeres erősítésnél nagyobbakat létrehozni, így tehát rákényszerül, hogy a főnyalábot a főirány közelébe helyezze. Ez azonban nem tekinthető egy szofisztikált megoldásnak. Annak érdekében, hogy a releváns irány környékéről az odakerülő nullhelyeket kiszorítsam a Robust Capon eljárást implementáltam, melynek eredeti feladata pontosan ennek a problémának az orvoslása.

8 Robust Capon algoritmus

A korábban már említett módon, a Capon algoritmus egyik legnagyobb gyakorlati hátránya, hogy abban az esetben, ha valóban beérkező jelekből dolgozunk, és nem meglévő ismert zavarirányokra helyezünk analitikusan nullhelyeket, úgy ha a releváns forrás iránya nem pontosan ismert, az algoritmus automatikusan nullhelyet helyezhet tévesen becsült, egységnyi erősítésű vélt forrásirány mellett lévő valódi forrásirányra annak érdekében, hogy a kimeneten megjelenő teljesítményt minimalizálja.

Ez egy nagyon előnytelen viselkedés, ami ellen kifejlesztették az úgynevezett Robust Capon algoritmust [10], melynek a vélt releváns forrásirány mellett bemenete még egy ε bizonytalansági érték, ami azt határozza meg, hogy mennyire vagyunk biztosak a releváns forrás irányában, vagyis mekkora hibát engedjen meg az algoritmus az iránykarakterisztika során. Ennek értelmében egy adott bizonytalansági értékhez egy

olyan ellipszis tartozik a karakterisztikán, melyen belül feltételezzük a valódi forrásirányt. Az algoritmus működése során a rendelkezésére bocsájtott kovarianciamátrix alapján igyekszik olyan irányt találni az adott bizonytalansági ellipszisen belül, hogy az arra alkalmazott Standard Capon módszer a lehető legnagyobb teljesítményt adja, ami ideális esetben a releváns forrás irányának megtalálását jelenti.

Ez a viselkedés az esetünkben is rendkívül hasznos lehet, ugyanis azáltal, hogy az adott ellipszisekből nullhelyek kizáródnak, elérhető, hogy a főnyaláb iránya is arrafele legyen, ami az időtartományból becsült kovarianciamátrix alapú iránykarakterisztikaszámítás során jelentett komoly problémát.

Robust Capon matematikailag:

$$\min_{\mathbf{a}} \mathbf{a}^H \mathbf{R}^{-1} \mathbf{a} \quad \text{úgy, hogy} \quad \|\mathbf{a} - \bar{\mathbf{a}}\|^2 \leq \varepsilon$$

ahol \mathbf{a} az algoritmus során a későbbiekben alkalmazott forrásirányhoz tartozó terjedési vektor, $\bar{\mathbf{a}}$ az algoritmus számára bemenetként meghatározott feltételezett releváns forrásirányból adódó terjedési vektor, \mathbf{R} a mérés során kialakult kovarianciamátrix, ε pedig a bizonytalansági érték, mely értelmezhető úgy, mint a feltételezett terjedési vektor, és a bizonytalansági ellipszis széléhez tartozó terjedési vektorok különbségének az Euklideszi normája.

A Standard Capon algoritmus eredményeképpen adódó w súlyvektor alkalmazásával a kimeneten megjelenő teljesítmény az alábbi:

$$P(\mathbf{a}_0) = \mathbf{w}^H \mathbf{R} \mathbf{w} = \frac{1}{\mathbf{a}_0^H \mathbf{R}^{-1} \mathbf{a}_0}$$

A Robust Capon eljárásban szereplő $\min_{\mathbf{a}} \mathbf{a}^H \mathbf{R}^{-1} \mathbf{a}$ tehát a kimeneti teljesítmény maximalizálásának célját írja le az \mathbf{a} terjedési vektor hangolásán keresztül. Annak érdekében, hogy ne egyszerűen a legnagyobb jelszintű forrás kerüljön a főnyalábra, az algoritmus kiköti, hogy $\|\mathbf{a} - \bar{\mathbf{a}}\|^2 \leq \varepsilon$ tehát a teljesítménymaximumot produkáló \mathbf{a} terjedési vektor nem térhet el az általunk megadott feltételezett forrásirányhoz tartozó \mathbf{a} vektortól egy bizonyos ε értéknél jobban.

Annak érdekében, hogy a triviális $\mathbf{a} = \mathbf{0}$ megoldást kizárjuk, abból a feltételezésből indulunk ki, hogy

$$\|\bar{\mathbf{a}}\|^2 > \varepsilon .$$

Mivel a kezdeti összefüggésre a kényszerkészlet határán mindenképpen található megoldást, így a feladat a következőképpen fogalmazható át:

$$\min_{\mathbf{a}} \mathbf{a}^H \mathbf{R}^{-1} \mathbf{a} \quad \text{úgy, hogy} \quad \|\mathbf{a} - \bar{\mathbf{a}}\|^2 = \varepsilon$$

Melynek megoldására a Lagrange-féle multiplikatormódszer segítségével van lehetőség, ami az alábbi függvényen alapul:

$$f = \mathbf{a}^H \mathbf{R}^{-1} \mathbf{a} + \lambda (\|\mathbf{a} - \bar{\mathbf{a}}\|^2 - \varepsilon)$$

ahol λ az úgynevezett Lagrange multiplikátor. Az előző egyenlet \mathbf{a} szerinti deriválását követően megkapható az optimális $\hat{\mathbf{a}}_0$ megoldás:

$$\mathbf{R}^{-1} \hat{\mathbf{a}}_0 + \lambda (\hat{\mathbf{a}}_0 - \bar{\mathbf{a}}) = 0$$

Ami alapján, illetve a mátrixinverziós lemma felhasználásával:

$$\hat{\mathbf{a}}_0 = \left(\frac{\mathbf{R}^{-1}}{\lambda} + \mathbf{I} \right)^{-1} \bar{\mathbf{a}} = \bar{\mathbf{a}} - (\mathbf{I} + \lambda \mathbf{R})^{-1} \bar{\mathbf{a}}$$

A Lagrange multiplikátor megállapításához az alábbi összefüggés megoldására van szükség:

$$g(\lambda) \triangleq \|(I + \lambda R)^{-1} \bar{\mathbf{a}}\|^2 = \varepsilon$$

Legyen:

$$\mathbf{R} = \mathbf{U} \mathbf{\Gamma} \mathbf{U}^H$$

Ahol \mathbf{U} mátrix oszlopai az \mathbf{R} mátrix sajátvektorai, a $\mathbf{\Gamma}$ mátrix főátlóbeli elemei pedig az egy es sajátvektorokhoz tartozó sajátértékek úgy, hogy $\Gamma_{11} \geq \Gamma_{22} \geq \dots \geq \Gamma_{MM}$.

Legyen $z = \mathbf{U}^H \bar{\mathbf{a}}$, és jelentse z_m a z vektor m -edik elemét, így a fenti egyenlet az alábbi alakban írható le:

$$g(\lambda) = \sum_{m=1}^M \frac{|z_m|^2}{(1 + \lambda \gamma_m)^2} = \varepsilon$$

Mivel g egy folytonos, szigorúan monoton csökkenő, pozitív függvény pozitív lambda értékekre, így az optimális Lagrange multiplikátor érték egyszerű felező módszerrel jól közelíthető.

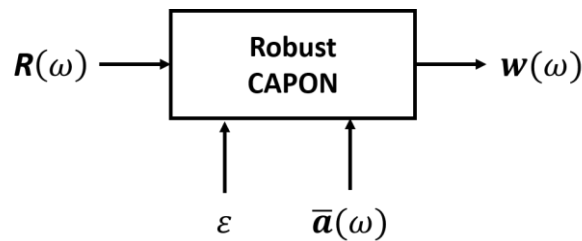
Ezek alapján tehát megkapható az ideális terjedési vektor:

$$\hat{\mathbf{a}}_0 = \bar{\mathbf{a}} - \mathbf{U} (\mathbf{I} + \lambda \mathbf{\Gamma})^{-1} \mathbf{U}^H \bar{\mathbf{a}} = \bar{\mathbf{a}} - (\mathbf{I} + \lambda \mathbf{R})^{-1} \bar{\mathbf{a}}$$

Ennek felhasználásával pedig ismét alkalmazható a Standard Capon módszer, ami az optimális súlyvektorokat adja majd eredményül:

$$\mathbf{w}_0 = \frac{\mathbf{R}^{-1} \hat{\mathbf{a}}_0}{\hat{\mathbf{a}}_0^H \mathbf{R}^{-1} \hat{\mathbf{a}}_0} = \frac{\left(\mathbf{R} + \frac{1}{\lambda} \mathbf{I}\right)^{-1} \bar{\mathbf{a}}}{\bar{\mathbf{a}}^H \left(\mathbf{R} + \frac{1}{\lambda} \mathbf{I}\right)^{-1} \mathbf{R} \left(\mathbf{R} + \frac{1}{\lambda} \mathbf{I}\right)^{-1} \bar{\mathbf{a}}}$$

Ezek alapján tehát rendelkezésünkre áll a súlyvektor, mellyel a korábban mért kovarianciamátrix, a feltételezett releváns forrásirány, valamint a bizonytalansági tényező alapján ideális iránykarakterisztika kapható a 24. ábra szerint:



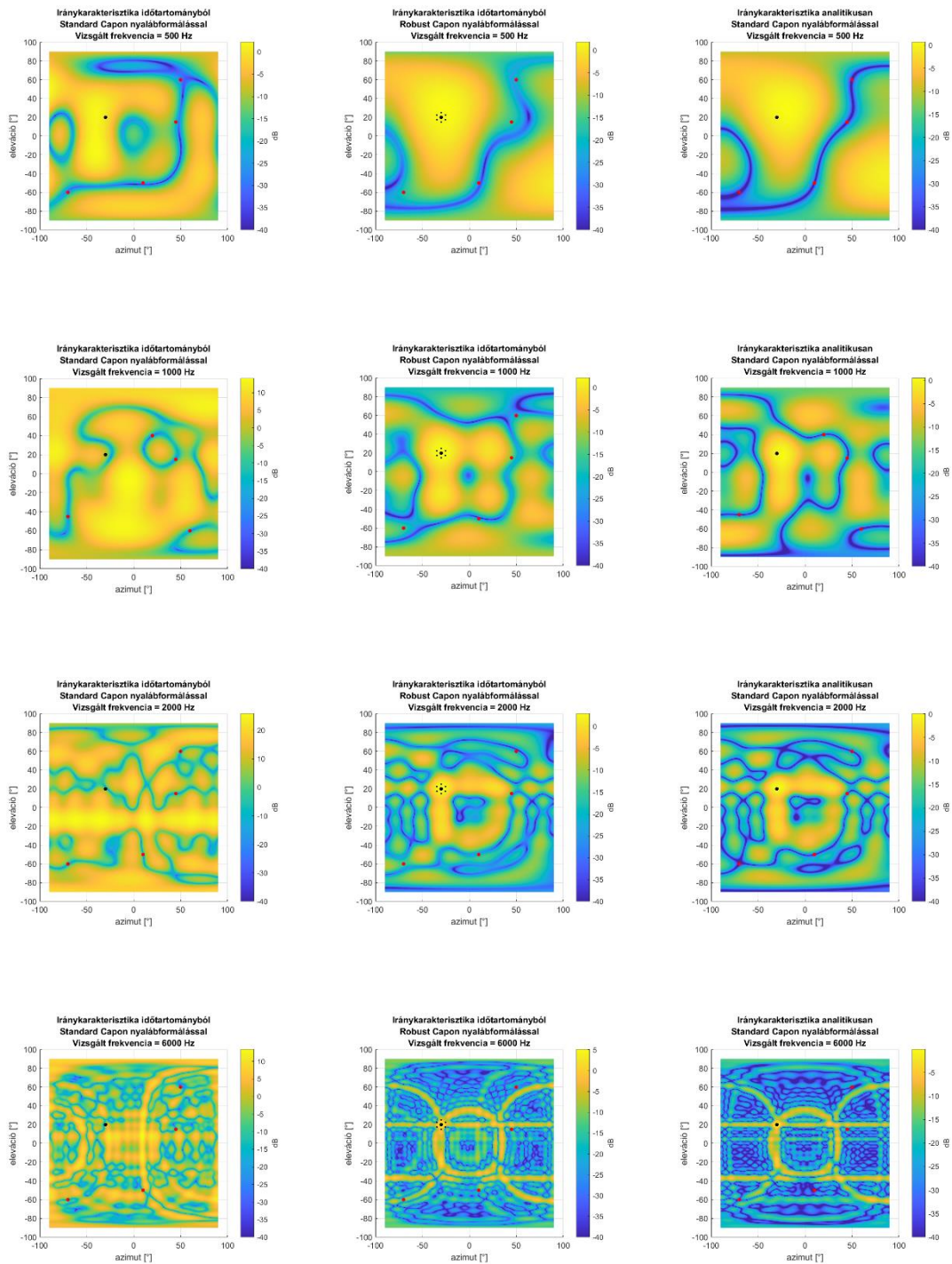
24. ábra Robust Capon algoritmus bemenetei, illetve kimenete

Annak érdekében, hogy a bizonytalansági tényezőhöz szemléletes jelentés kössünk, meghatározhatunk a főirány mellett további olyan irányokat vagy térbeli pontokat, amelyekkel bizonyos értelemben kijelölhetjük a tartományt, amin belül az egységnyi erősítést biztosítani kívánjuk. Az algoritmusához szükséges ε értéket pedig a főirányhoz, illetve a bizonytalansági irányhoz tartozó terjedési vektorok különbségének euklideszi normájaképp kaphatjuk meg.

$$\varepsilon = \|\bar{\mathbf{a}} - \mathbf{a}_{err}\|$$

A bizonytalansági tényező így ezáltal irányok esetén kifejezhető az egyes irányok által bezárt szöggel, vagy térbeli esetben az egyes pontok közötti távolsággal. Több ilyen érték esetén természetesen célszerű az adódó legnagyobb ilyen tényezőértéket választani.

A jobb összehasonlíthatóság érdekében egymás mellett ábrázoltam az analitikusan, időtartományból Standard, illetve Robust Capon-nal kapott karakterisztikáit, ezeket a 27. ábra mutatja.



27. ábra Capon és Robust Capon összehasonlítása

Látható, hogy a Robust Capon által biztosított iránykarakterisztika a Standard Caponhoz hasonló eredményt produkál, nem illeszt nullhelyet a releváns irány közelébe, és nem hoz létre hatalmas erősítésű irányoka. Bár a zavarforrások irányába alacsony frekvenciákon nem tökéletesen illeszt nullhelyeket, ez a viselkedés a későbbiekben talán orvosolhatóvá válik hosszabb idejű átlagolás, vagy az epszilon finomhangolásán keresztül.

Elképzelhető esetleg az is, hogy alacsony frekvenciákon Standard, nagyobbakon pedig Robust Capon kerül alkalmazásra, bár ennek eldöntéséhez további tesztek szükségesek.

9 Jelfeldolgozási folyamat

A korábbi fejezetben részletezett Robust Capon algoritmus bár kedvező eredményeket mutat, annak valós mikrofontömbön történő megvalósítása nem triviális. Maga az algoritmus közvetlenül kizárólag frekvenciatartományban, egy adott vizsgálati frekvencián alkalmazható. A későbbiek során számunkra ennek az algoritmusnak a szélessávú alkalmazása a célunk.

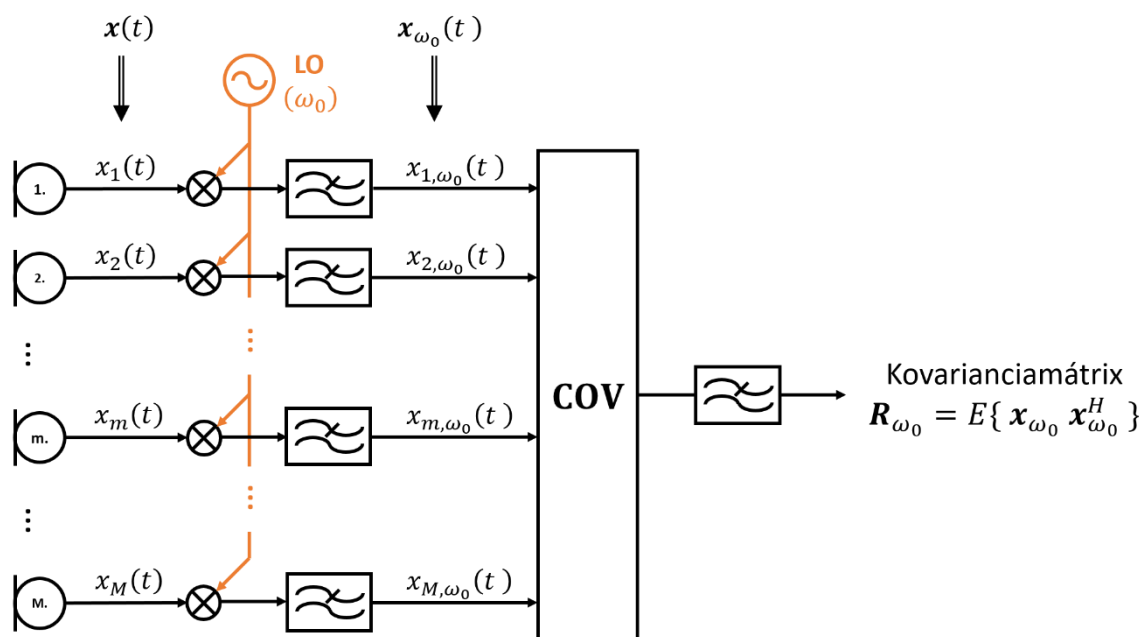
9.1 Kovarianciamátrix egyetlen frekvencián

Az algoritmus kulcsponja az akusztikai környezetet, adott vizsgálati frekvencián, mikrofontömb szempontjából leíró kovarianciamátrixok meghatározása. Ennek folyamatát egyetlen frekvencián történő alkalmazás során a 28. ábra szemlélteti:

Az adott frekvenciájú jelfázisok meghatározásához egyes mikrofonokból származó feszültség-idő függvények adott frekvenciájú alapsávi komponenseinek komplex értékeire van szükségünk, melyek információt hordoznak a mikrofonokon megjelenő adott frekvenciájú jel amplitúdójáról, illetve fázisáról.

Ezeknek meghatározása történhet a mikrofonokból származó $x(t)$ jelek adott ω_0 frekvenciájú lokáloszcillátorból származó komplex jellel történő lekeverésével, majd aluláteresztő-szűrésével. Ennek segítségével rendelkezésünkre állnak a komplex $x_{\omega_0}(t)$ időben változó fázisértékek. Az ezekből képzett mátrix várhatóértéke adja az ω_0 frekvenciához tartozó R_{ω_0} kovarianciamátrixot.

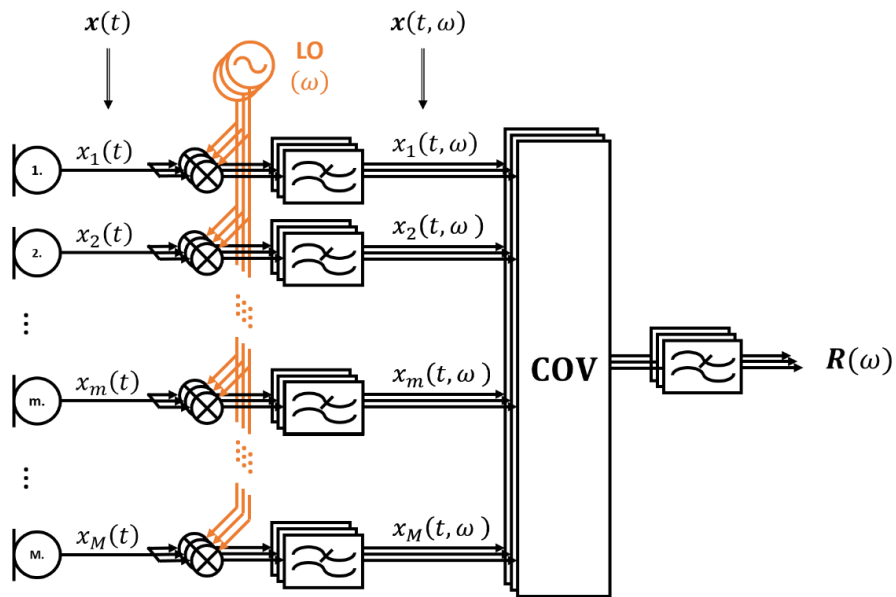
$$R_{\omega_0} = E\{x_{\omega_0} \cdot x_{\omega_0}^H\}$$



28. ábra Kovarianciamátrix becslése egyetlen frekvencián

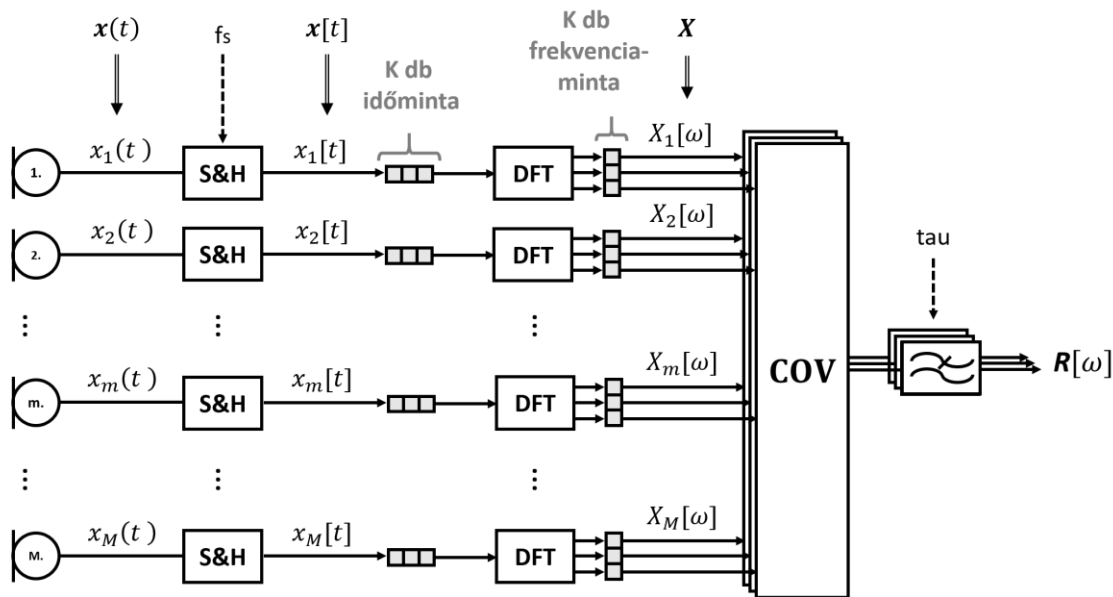
9.2 Kovarianciamátrix szélessávban

Az előző alfejezetben tárgyalt módon lehetőségünk van a kovarianciamátrix meghatározására egyetlen frekvencián, azonban számunkra az algoritmus szélessávú alkalmazása célszerű. Ehhez a kézenfekvő megoldás természetesen az, hogy több frekvenciára is elvégezzük a kovarianciamátrix számítását a korábban részletezett módon, a 29. ábra által szemléltetett módon.



29. ábra Kovarianciamátrix becslése több frekvencián

Ezzel a módszerrel tulajdonképpen egy szűrőbankot hozunk létre, mellyel adott frekvenciaértékekkel közelítjük a mikrofonokon kialakuló jel frekvenciatartománybeli leírását. Ilyen módon minden egyes vizsgált frekvenciához tartozik egy megfelelő keverő, szűrő, illetve kovarianciamátrixot generáló hálózat. A vázolt eljárás során tulajdonképpen a mikrofonokon kialakuló jelek spektrumait mintavételezzük a frekvenciatartományban. Analóg megvalósítás során ez egy elfogadható megoldás lehet, azonban a digitális technika jelenlegi szintje mellett kézenfekvőbb megoldást nyújt a jeleknek az időbeli mintavételezése, hiszen azoknak digitalizálását követően a jelfeldolgozás nagyban leegyszerűsödik. Ennek az eljárásnak a sematikus vázlatát mutatja a 30. ábra.



30. ábra Kovarianciamátrix becslése DFT segítségével

Az egyes mikrofonokból származó folytonos idejű $x(t)$ jelek közvetlenül az S&H-val jelölt időbeli mintavevő- és tartó egység bemenetére kerülnek, aminek kimenetein így már csak a diszkrét idejű $x[t]$ jelek jelennek meg. Ennek a mintavevőnek a mintavételi frekvenciája f_s , ami a Shannon-Nyquist-féle mintavételi tétel értelmében az általunk vizsgált valós, sávkorlátos jelek spektrumában szereplő legnagyobb frekvenciájú komponens frekvenciájának legalább kétszerese kell legyen, a mintavételből következő esetleges spektrumátlapolódási jelenségek elkerülése érdekében.

A jelspektrum összetevőinek meghatározásához a mintavevőből származó diszkrét minták minden F hosszú csoportján Diszkrét Fourier-transzformációt hajtunk végre, melynek eredménye egy ugyancsak K hosszú mintasor, mely az adott mikrofonhoz tartozó komplex spektrumösszetevőket mutatja adott $\frac{f_s}{K}$ frekvenciánként. Az $[\omega]$ -val jelölt értékek olyan frekvenciák jelentenek, amik megkaphatók az $\omega = q \cdot \frac{f_s}{K}$ összefüggéssel, ahol q N -nél kisebb természetes szám.

Így tehát megkaphatjuk bármely m jel mikrofonhoz tartozó $x_m[t]$ diszkrét időbeli jel adott K hosszú mintájának DFT-jét, ami az alábbi összefüggés.

$$X_m[\omega] = \sum_{k=0}^{K-1} x_m[k] \cdot e^{-j \cdot \frac{2\pi \omega k}{K}}$$

ahol

$$\omega = 0, 1, 2, \dots, F - 1$$

Csak egy adott ω diszkrét frekvenciához tartozó spektrumértéket az alábbi módon vektorba foglalva megkaphatjuk a ω diszkrét körfrekvenciához tartozó jelfázisvektor diszkrét megfelelőjét.

$$\mathbf{x}_{\omega}^T = [X_1[\omega], X_2[\omega], \dots X_m[\omega], \dots X_M[\omega]]$$

Ezekből képezhető az alábbi \mathbf{X} jelspektrum-mátrix:

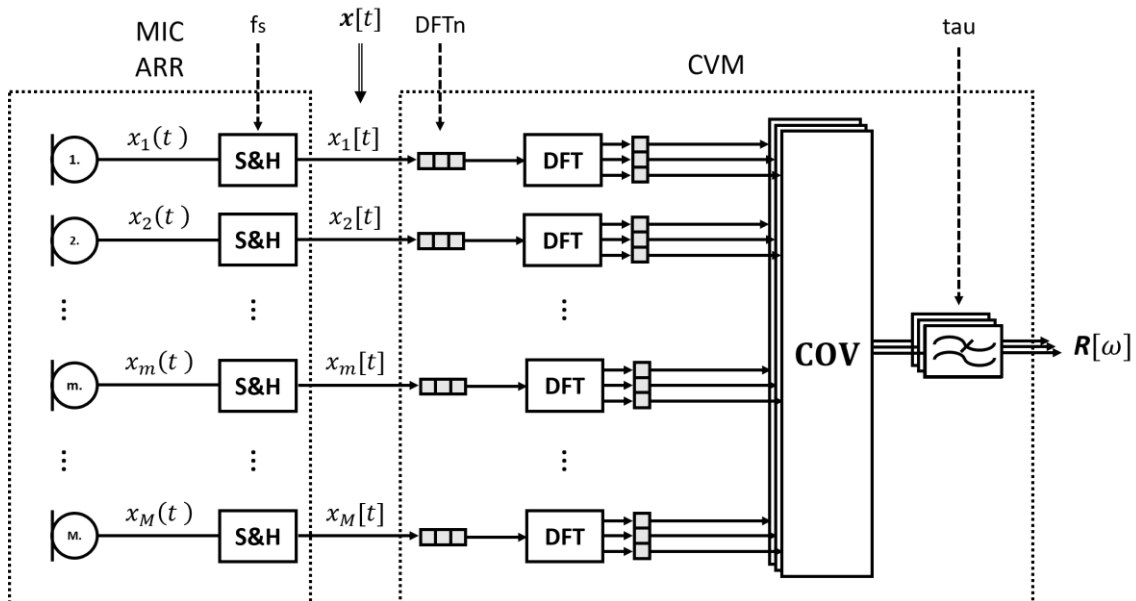
$$\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \dots \mathbf{x}_k, \dots \mathbf{x}_K] = \begin{bmatrix} x_{11} & \dots & x_{K1} \\ \vdots & \ddots & \vdots \\ x_{1M} & \dots & x_{KM} \end{bmatrix}$$

Az egyes mikrofonokból származó, megegyező frekvenciájú spektrumösszetevőkön történik meg a kovarianciamátrix képzése.

$$\mathbf{R}[\omega] = E\{ \mathbf{X}_{:\omega} \cdot \mathbf{X}_{:\omega}^* \}$$

Ez az eljárás tehát a mintavételi frekvencia, illetve a diszkrét Fourier-transzformációhoz felhasznált minták száma által meghatározott frekvenciaértékek mindegyikére egyszerre határozza meg a kovarianciamátrixokat. Az eljárás előnye, hogy a vázolt folyamat számítógép felhasználásával könnyen végezhető, nincsen további hardverigénye, illetve a diszkrét Fourier-transzformáció hatékonyan megvalósítható olyan eljárásokkal, mint például az FFT.

Annak érdekében, hogy a későbbiekben ezekről átláthatóbb képet kapjunk, célszerű az egyes elemeket logikai blokkokba összefoglalni a 31. ábra által szemléltetett módon.

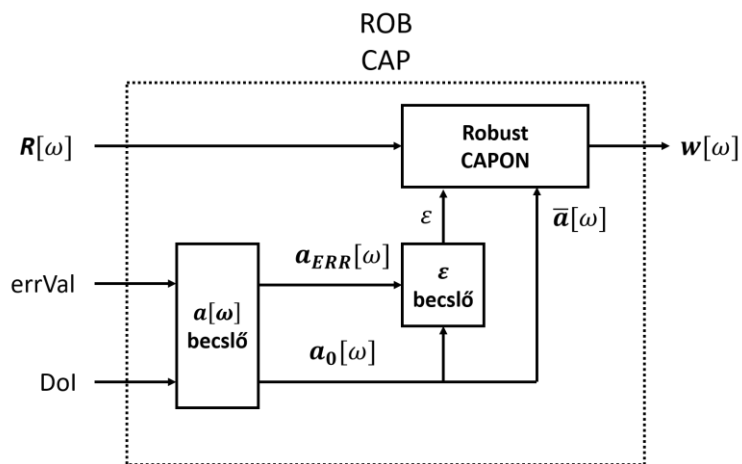
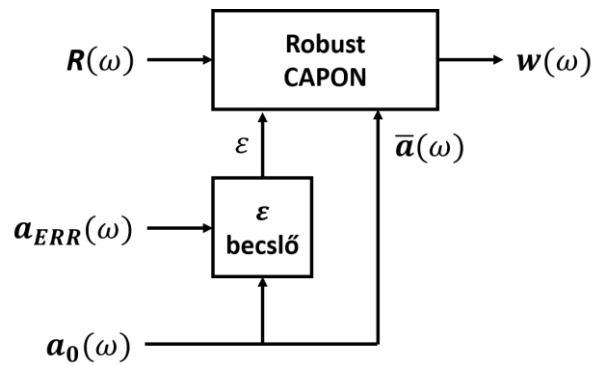


31. ábra Jelfeldolgozási rendszer logikai blokkokban

9.3 Mikrofontömb kimenetének számítása

Az egyes frekvenciához tartozó kovarianciamátrixok, azokhoz adott célirány esetén tartozó terjedési vektorok, illetve ε bizonytalansági értékek felhasználásával, a korábban

részletezettek szerint a Robust Capon algoritmus alkalmazásával megkaphatjuk a mikrofontömbhöz adott frekvenciákon tartozó súlyvektort.



Ezekkel a súlyvektorokkal a korábbiakban ismertettek szerint, F darab $x[t]$ időmintából diszkrét Fourier-transzformáció segítségével adódó $x[\omega]$ diszkrét spektrum esetén megkapható az egyes frekvenciákhoz tartozó kimenet az alábbi módon:

$$Y_0[\omega] = \mathbf{w}^H[\omega] \cdot \mathbf{x}[\omega] = \sum_{m=1}^M w_m^*[\omega] \cdot x_m[\omega]$$

Annak érdekében tehát, hogy rendelkezésünkre álljon az időtartománybeli végeredmény, szükségünk van ennek az inverz diszkrét Fourier-transzformáltjára, ami az alábbi eredményt adja:

$$y[t] = \sum_{k=0}^{K-1} Y[k] \cdot e^{j \frac{2\pi t k}{K}}$$

ahol

$$t = 0, 1, 2, \dots, N - 1$$

A mikrofontömb alkalmazása során azonban kedvezőbb futást tudunk biztosítani, ha szűrést nem frekvenciatartományban, hanem időtartományban végezzük, felhasználva, hogy $W_m^*[\omega] = H_m[\omega]$

$$\begin{aligned} y_0[t] &= \sum_{n=0}^{N-1} Y_0[n] \cdot e^{j \frac{2\pi t n}{N}} = \sum_{n=0}^{N-1} \mathbf{w}^H[n] \cdot \mathbf{x}[n] \cdot e^{j \frac{2\pi t n}{N}} = \\ &= \sum_{n=0}^{N-1} \sum_{m=1}^M \mathbf{w}_m^*[n] \cdot \mathbf{x}_m[n] \cdot e^{j \frac{2\pi t n}{N}} = \sum_{m=1}^M \sum_{n=0}^{N-1} H_m[n] \cdot X_m[n] \cdot e^{j \frac{2\pi t n}{N}} \end{aligned}$$

ahol

$$t = 0, 1, 2, \dots, N - 1$$

a konvolúciós-tétel [11] alapján:

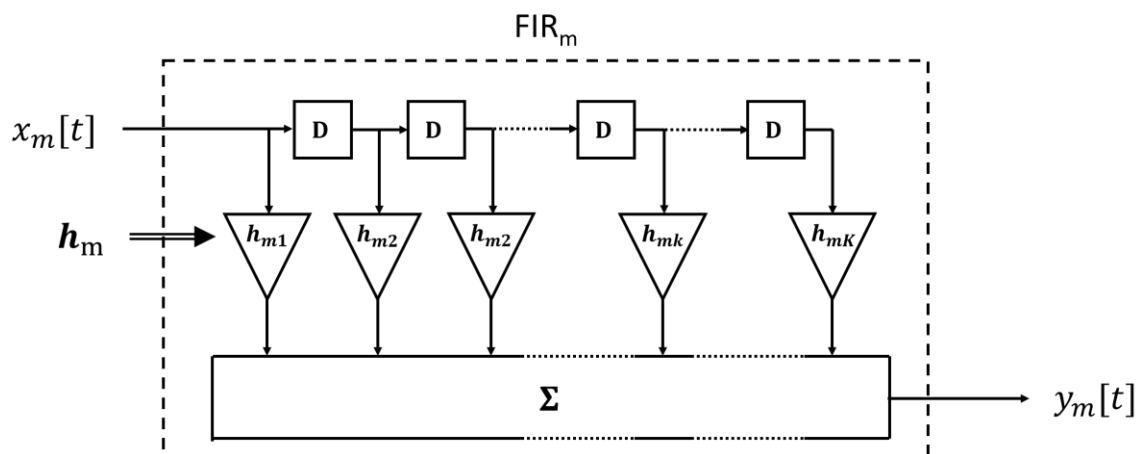
$$H_m[n] \cdot X_m[n] = h_m[n] * x_m[n]$$

így

$$y[t] = \sum_{b=1}^B h_m[t] * x_m[t]$$

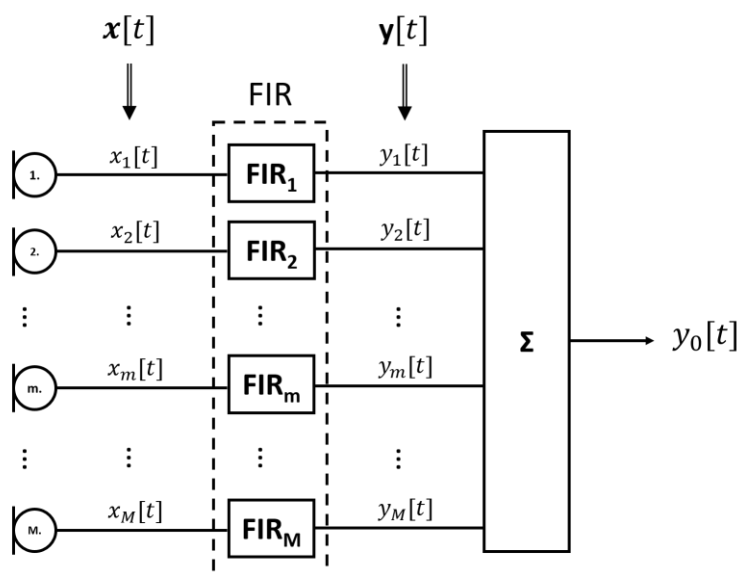
Ennek ismeretében tehát amennyiben szélessávon kellően pontosan ismert Kovarianciamátrixokra elvégzem a Robust Capon algoritmust, akkor mikrofononként birtokába kerülök egy-egy átviteli függvények, amelyből azt inverz diszkrét Fourier transzformálva megkaphatom az egyes mikrofonokhoz tartozó szűrők impulzusválaszát, melyeket összegezve adódik az ideális kimenet.

Az impulzusválasz ismeretében tehát megalkotható minden mikrofonra egy-egy véges impulzusválaszú (FIR) szűrő, mely az egyes frekvenciákon a Robust Capon algoritmus által számított átviteli értékkel rendelkezik, és futásidejű alkalmazást tesz lehetővé. A FIR szűrő felépítését a 32. ábra mutatja, az ebben található erősítőtagoknak az erősítési értéke megegyezik a diszkrét idejű impulzusválasz egyes értékeivel.



32. ábra FIR szűrő blokkvázlata

Minden egyes mikrofon közvetlen kimenetére kerül tehát egy FIR szűrő, melyek eredményeit összegezve megkaphatjuk a mikrofontömb $y_0[t]$ diszkrét kimenetét.

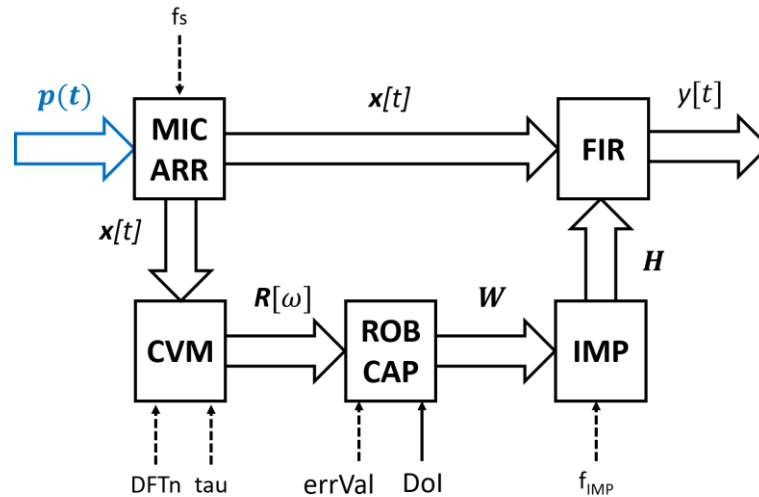


33. ábra FIR blokk alkalmazása

A 33. ábra által bemutatott módon az egyes FIR szűrőket összefoglalhatjuk egy M bemenetű, M kimenetű FIR jelzésű blokkban, hogy a későbbiekben könnyebben tudjunk rá hivatkozni.

9.4 Teljes rendszer felépítése

A teljes digitális adatfeldolgozási rendszert, az egyes paramterekkel az 34. ábra mutatja.



34. ábra Teljes rendszer felépítése

A rendszernek bemenete tehát az egyes mikrofonokra érkező $p(t)$ hangnyomásjelek, valamint a Dol -vel jelölt fókuszálási irány a bemenete. Emellett további paraméterek is megjelennek:

Dol: A fókuszálási irány, vagy fókuszálási pozíció, mellyel meghatározzuk az algoritmus számára, hogy az akusztikus tér mely pontjában, vagy irányában található az általunk relevánsnak vélt hangforrás.

fs: A mikrofonok mintavételi frekvenciája. Ez az érték a vizsgálni kívánt frekvenciatartomány ismeretében fontos. Legtöbb esetben ez az emberi hallástartományhoz igazított, jellemzően 44100 Hz, vagy 48000 Hz értékű.

DFT_n: A DFT alkalmazása során felhasznált időminták száma. Ezt az értéket minél nagyobbra választjuk, annál nagyobb frekvenciafelbontással végezhető a szűrők átviteli függvényének becslése, ami így pontosabb átvitelt tud eredményezni, azonban ennek ára az, hogy ezzel azonos számú $M \times M$ -es kovarianciamátrix tárolására és feldolgozására, van szükségünk, ezek mindegyikére pedig elvégzésre kerül a Robust Capon algoritmus, tehát az FFT minták számával a szükséges számítási kapacitás is lineárisan növekszik. Ezen kívül pedig még természetesen amiatt, hogy adott K hosszú mintasort kell megvárnom, a mikrofontömb kimenetének a késleltetése is növekszik.

tau: A kovarianciamátrix exponenciális átlagolása során alkalmazott időállandó, mely a kovarianciamátrix képzése során igényelt várhatóérték-képzésben játszik szerepet. Természetesen statikus környezet esetén ezt minél hosszabbnak célszerű választani,

hogy az elméleti kovarianciamátrixot minél jobban közelítse, azonban dinamikus esetben a nagy időállandó kedvezőtlen hatást jelenthet egy következő stabil akusztikus környezethez való alkalmazkodás során.

errVal: A *DoI*-hez szorosan tartozó paraméter, mely által kerül meghatározásra a Robust Capon algoritmus bemeneti ε értéke.

f_IMP: A súlyvektorok frissítésének frekvenciája, ami azt határozza meg, hogy a rendelkezésünkre álló kovarianciamátrixokon egységnyi idő alatt hányszor történik meg a Capon algoritmus lefutása. Ennek a kialakuló karakterisztikára természetesen nincsen hatása, hiszen a kovarianciamátrixok az algoritmus elvégzésétől függetlenül ugyanúgy alakulnak, viszont dinamikus esetben problémát jelenthet az, hogy egy már módosult környezet esetén annak egy korábbi állapotára megfelelő szűrőrendszert alkalmazunk.

10 Algoritmus tesztelése modellezett környezetben

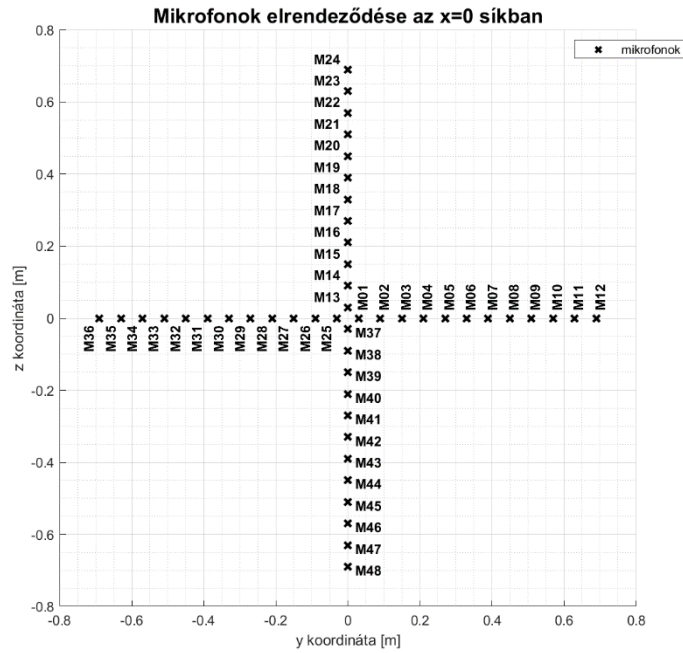
Annak érdekében, hogy a helyes működésről megbizonyosodjak, a korábban részletezett nyálábformáló algoritmust megvalósító jelfeldolgozó rendszert Matlab programkörnyezet felhasználásával teszteltem.

Ennek során modelleztem egy bizonyos akusztikai környezetet, amit a térben adott módon elhelyezett mikrofontömbbel mintavételeztem, majd azt egy hangfájlként tároltam. Ezt a hangfájlt, valamint egy térbeli koordinátát, mint fókuszálási irányt adtam meg az implementált feldolgozóprogram bemenetének, mely a Robust Capon algoritmus alkalmazásával olyan kimenetet igyekezett biztosítani, mely a célkoordinátán lévő hangforrás jel-zaj viszonyát tekintve a lehető legkedvezőbb eredményt adja.

A modellezés során olyan környezetet, illetve térbeli mikrofonelrendezést alkalmaztam, melyen keresztül szemléletesen bemutatható az algoritmus alkalmazásának hatása a kimeneten, valamint amely elrendezés a későbbiekben rendelkezésemre álló eszközök segítségével a valóságban is könnyen megvalósítható.

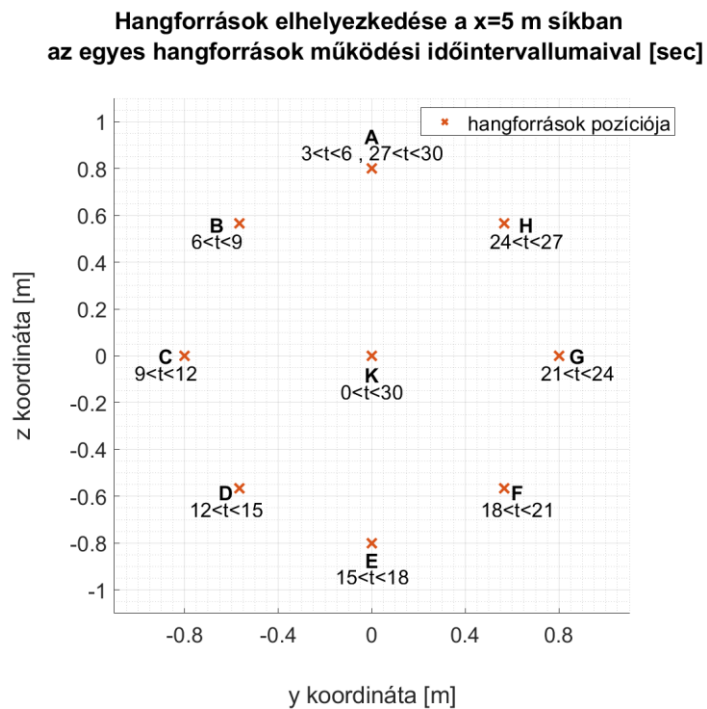
10.1 Modellezett mérési elrendezés

Az általam modellezett mikrofontömb 48 db egyesével sorszámozott izotróp mikrofonból épül fel az $x=0$ síkban a 35. ábra által bemutatott módon. Az egyes mikrofonok által érzékelt hangnyomásjelek a kimeneti .wav kiterjesztésű hangfájl megfelelő sorszámú csatornáján kerülnek tárolásra.



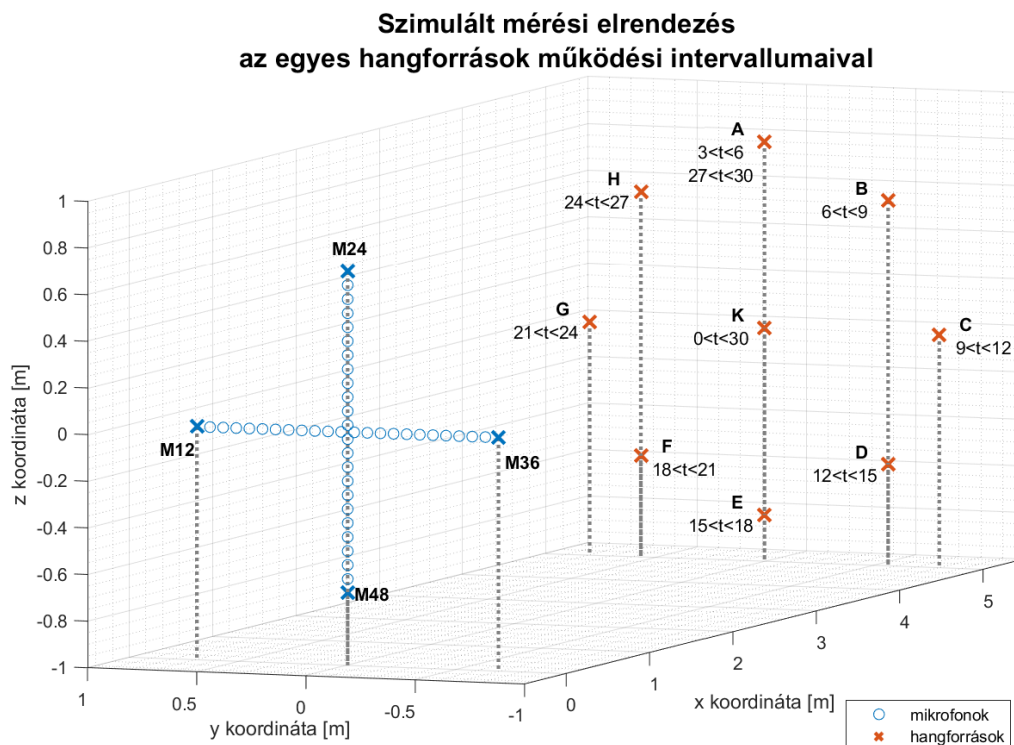
35. ábra Mikrofonok elhelyezkedése

A mikrofontömbhöz képest a szimulált hangforrások az x=5 m síkban, egy 0.8 m sugarú $O=(5; 0; 0)$ m középpontú kör mentén 45° -onként helyezkednek úgy, hogy az az A jelű hangforrás az $A=(5; 0; 0.8)$ pontban helyezkedik el a 36. ábra által bemutatott módon:



36. ábra Hangforrások elhelyezkedése

Az egyes hangforrások a mikrofontömböt alkotó mikrofonokhoz képest az 37. ábra által szemléltetett módon helyezkednek el a térben.



37. ábra Szimulált mérési elrendezés

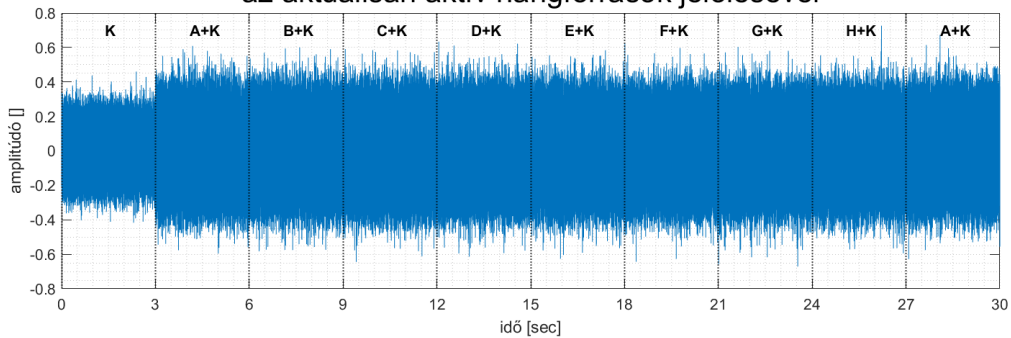
A modellezett mérésünk során alapvetően a K jelű hangforrásból származó akusztikus jeleket szeretnénk rögzíteni, mialatt a célforrásunk mellett további vele inkohérens hangforrások is megjelennek, és kisugárzott teljesítményükkel rontják a K jelű hangforrás vétele során mérhető jel-zaj viszonyt.

A hangforrások közül az K jelű célforrás az egész szimulált minta során, $t=0$ sec-tól $t=30$ sec-ig folyamatosan bocsájt ki magából egy adott effektív értékű Gaussi fehér zajt. Eközben $t=3$ sec-től kezdődően a többi hangforrás egymást követően, 3 másodpercenként felváltva, az ABCDEFGH sorrendben bocsájt ki ugyanilyen effektív értékkel fehér zajt.

Az így rendelkezésemre álló, mikrofonfelvételekből létrejött egy .wav kiterjesztésű, 48 csatornás, 30 másodperc hosszúságú hangminta, mely a feldolgozóprogram bemeneteképp szolgál.

A hangfájl első számú csatornájának a jelét, tehát az M01-es mikrofon által vett jelet a 38. ábra mutatja.

Hangnyomás-idő függvény a szimulált minta M01-es mikrofonján az aktuálisan aktív hangforrások jelölésével



38. ábra Hangnyomás-idő függvény a szimulált minta M01-es mikrofonján

Az időfüggvényen látható, hogy a mikrofon által rögzítésre került hangnyomásértékek valóban zajjellegűek, és ezeknek a nagysága a minta első 3 másodpercében láthatóan kevesebb, mint a fennmaradó részben, hiszen ekkor csak a K jelű hangforrás jele volt mérhető, míg később felváltva egy-egy másiké is. Az éppen aktuálisan megszólaló hangforrásokat a grafikon tetején felirattal, az azok közötti váltás pillanatát pedig a szaggatott függőleges vonalakkal jelöltem.

Erre az említett hangmintára a későbbiek „*dinamikus hangminta*” képpen fogok hivatkozni. Emellett létrehoztam ugyanis egy másik hangmintát, mely esetében a folyamatosan aktív K mellett a zavarforrások közül csak az 'A' jelű aktív $3 < t < 30$ sec intervallumon. Erre fogok később „*statikus hangmintaképp*” hivatkozni.

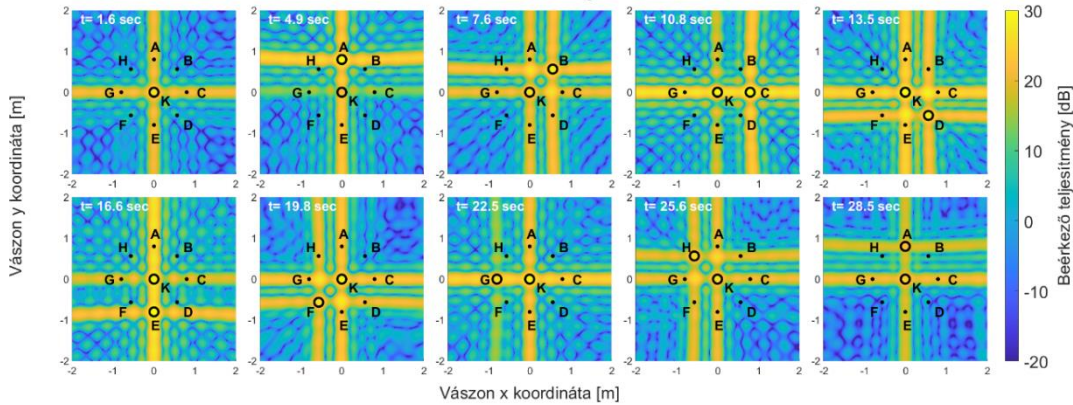
10.2 Generált hangminta ellenőrzése

Annak érdekében, hogy a jelforrás pozíciójának megváltozásáról, illetve a szimuláció helyességéről megbizonyosodjak, a rendelkezésre álló *dinamikus hangminta*n Bartlett-iránymérő algoritmust futtattam, melynek eredményét 6000 Hz vizsgálati frekvencián a 39. ábra mutatja. Az ábrák egy $x=5$ m síkban kifeszített akusztikus vászon egyes pontjairól beérkező hangteljesítményeket mutatják az iránymérő algoritmus futtatása során.

Az egyes ábrák fölött olvasható, hogy azok a vizsgált minta melyik időpontjában adódtak az iránymérő algoritmus eredményeképpen. Az ábrákon láthatók továbbá az egyes hangforrások modellezés során meghatározott pozíciói. Ezek közül karikával jelöltem az adott pillanatban éppen aktív, és ponttal az éppen néma hangforrásokat.

**Szimuláció során megszólaló zajforrások irányának meghatározása
Bartlet iránymérő algoritmus alkalmazásával 6000Hz frekvencián**

FFT mintaszám= 480; vászontávolság= 5.00 m; SNR= 50 dB;

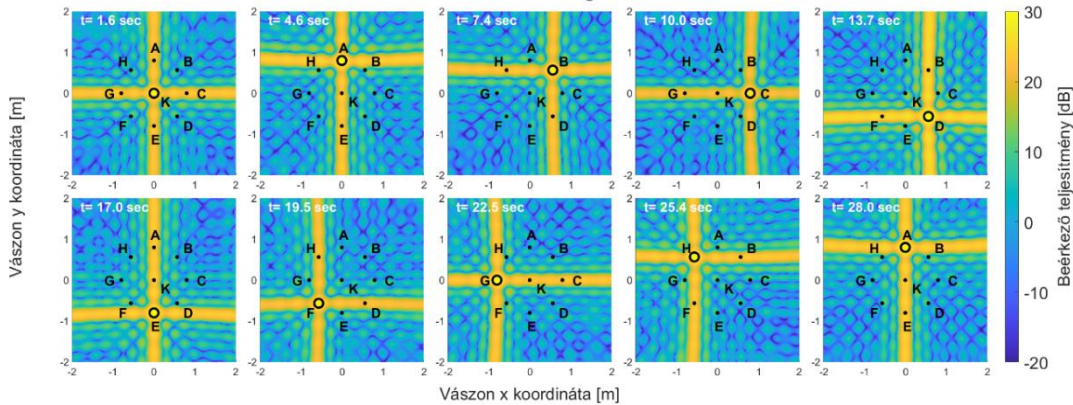


39. ábra Szimulációra alkalmazott iránymérés

Az alkalmazott iránymérési eljárás egyik hátránya, a mikrofontömb térbeli elrendeződéséből származó melléknyalábok szignifikáns jelenléte, ami így megnehezíti az iránymérés eredményének leolvasását, ezért létrehoztam a dinamikus hangminta egy olyan verzióját, melyen a centrális „K” jelű hangforrás csupán az első 3 másodpercben aktív, utána elnémul, viszont az őt körülvevő további hangforrások viselkedése nem változik. Erre a segédhangfájltra elvégezve a szimulációt a 40. ábra adódott, melyen már jól megfigyelhető, hogy az egyes hangforrások pozícióját az iránymérő algoritmus az általunk elvárt eredménnyel határozza meg.

**Szimuláció során megszólaló zajforrások irányának meghatározása
Bartlet iránymérő algoritmus alkalmazásával 6000Hz frekvencián**

FFT mintaszám= 480; vászontávolság= 5.00 m; SNR= 50 dB;



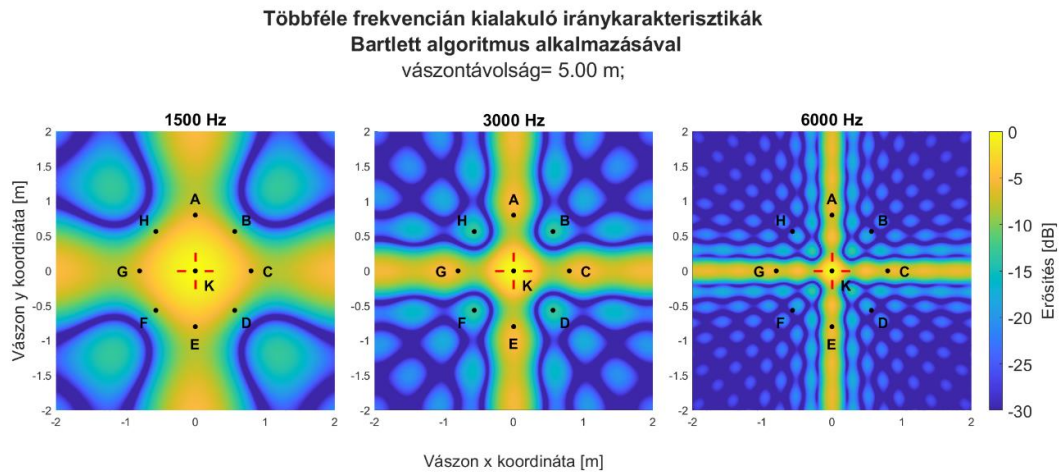
40. ábra Módosított szimulációra alkalmazott iránymérés

10.3 Bartlett nyalábformálás demonstrálása

Az implementált Robust Capon algoritmus hatékonyságának vizsgálatához az annál sokkal kézenfekvőbb, jóval kisebb számítási kapacitást igénylő, és igen népszerű Bartlett

nyalábformáló algoritmust választottam összehasonlítási alapnak. A Bartlett algoritmusnak egyetlen bemenete az esetünkben „K” ponthoz, mint fókuszálási irányhoz adott frekvencián tartozó jelterjedés vektor, aminek eredményeképpen kialakuló karakterisztika tulajdonsága, hogy a fókuszálási pontban a karakterisztikának egyik maximális, egységnyi erősítésű pontja található.

A modellünkben rendelkezésre álló mikrofontömbre, valamint a K fókuszálási pontra alkalmazva a Bartlett nyalábformáló algoritmust, a különféle frekvenciákon, a $x=5$ akusztikus vásznon kialakuló karakterisztikákat a 41. ábra mutatja.



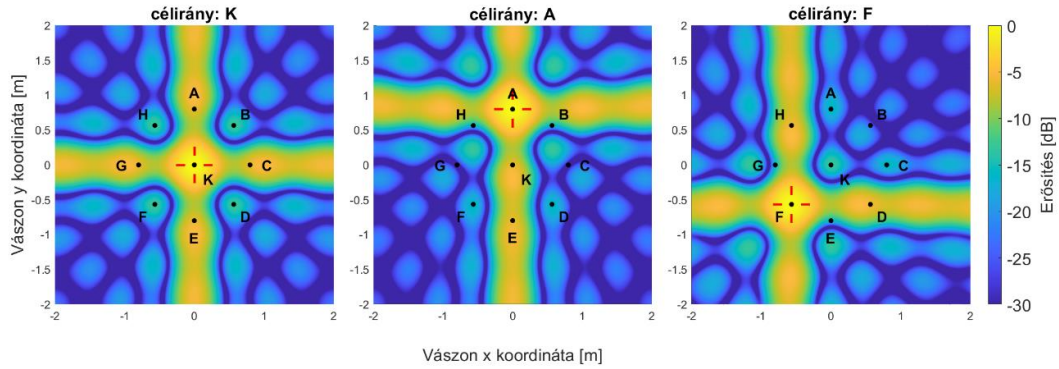
41. ábra Iránykarakterisztikák Bartlett algoritmus alkalmazásával

Látható hogy a jelenlegi elrendezésünk során is teljesül, hogy a Bartlett algoritmussal adódó karakterisztika bár egységnyi erősítést eredményez a releváns „K” pont irányába, a mikrofontömb térbeli elrendeződéséből származó melléknyalábok komoly erősítést biztosítanak a további zajforrások, főleg az A, C, E, G irányába, így az az azok által sugárzott jelek a kimeneten is megjelennek.

A nyalábformáló algoritmusok segítségével természetesen bármelyik hangforrás, vagy általános térbeli pont kiválasztható fókuszálási pontnak, ezt a fókuszálási pontot az ábrákon piros kereszttel jelölöm, ahogy azt a 42. ábra is mutatja.

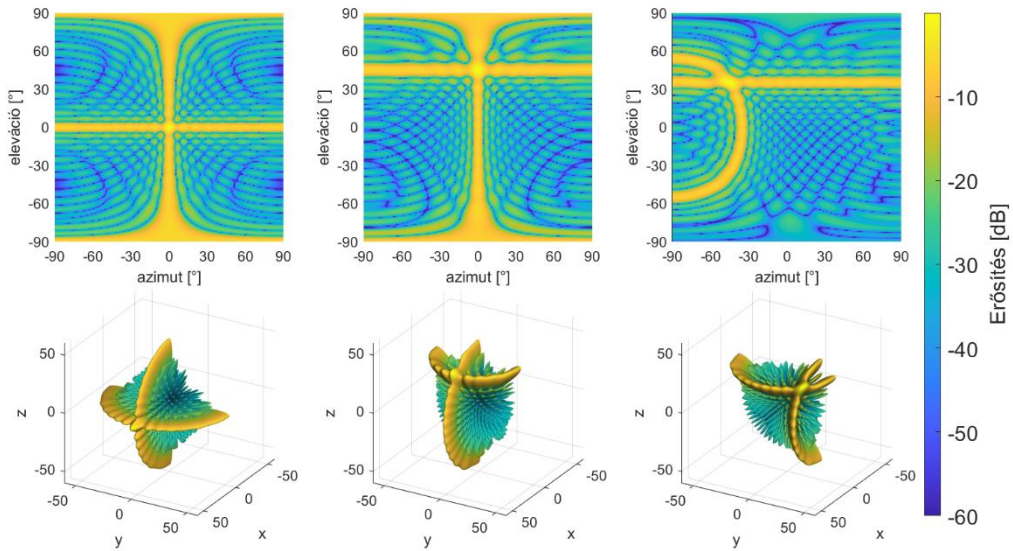
A kialakuló karakterisztikákról szemléletesebb képet ad azoknak a térbeli koordinátarendszerben történő ábrázolása, melyre példát a 43. ábra mutat.

Különbféle irányokba kialakított iránykarakterisztika
Bartlett algoritmus alkalmazásával
vászontávolság= 5.00 m;



42. ábra Különbféle irányokba kialakított karakterisztika Bartlett algoritmussal

Térbeli iránykarakterisztika 3000 Hz frekvencián
Bartlett irányforámoló algoritmussal



43. ábra Bartlett nyálábformálás térbeli polárkoordináta-rendszerben

A későbbiek során tehát annak a bemutatása a célom, hogy a Bartlett nyálábformáló algoritmushoz képest jelentősen nagyobb számítási igényű Robust Capon algoritmus olyan karakterisztikákat eredményez, melyek a Bartlettel összehasonlítva kedvezőbb eredményeket adnak a kimeneti jel-zaj viszony szempontjából.

10.4 Hangminta feldolgozása

Valós idejű feldolgozás esetén az algoritmus beágyazott rendszereken történő implementációja során azt valamilyen gépközeli, gyors futás idejű programnyelven célszerű megvalósítani, azonban a könnyebb kezelhetőség, jobb átláthatóság, és ízléses megjeleníthetőség érdekében a korábban részletezett algoritmust Matlab programkörnyezetben implementáltam. A hangfájl feldolgozásának pontos menetét a Jelfeldolgozási folyamat fejezetben tárgyalt módon végeztem.

A feldolgozás során tulajdonképpen minden impulzusválaszfrissítési lépést követően rendelkezésünkre áll a Robust Capon algoritmusból közvetlenül származó súlyvektor. Ezen a súlyvektoroknak az alapján a korábbiakban tárgyalt módon lehetőség van a mikrofontömb iránykarakterisztikáinak megjelenítésére, amivel ellenőrizni lehet az algoritmus helyes működését, illetve lehetőség van megfigyelni a különböző bemenetekre való viselkedését. Az adódó karakterisztikának természetesen leginkább az $x=5$ m síkban lévő akusztikus vászonra vett vetülete az, ami számunkra a legrelevánsabb erősítésértékeket mutatja a korábban tárgyaltak alapján.

Az iránykarakterisztikák ábrázolása során az aktuálisan rendelkezésre álló akusztikus környezet szemléltetése céljából jelölöm az egyes hangforrásokat, illetve az algoritmus bemeneteképp adódó fókuszálási irányt az modellezett környezet, valamint a vizsgálati időpont ismeretében. Az ábrákon az éppen aktív hangforrások karikával, míg az inaktívak ponttal kerülnek jelölésre, az általuk jelölt hangforrás betűjele mellett. Az algoritmus fókuszálási irányát piros célkereszttel jelöltem, ami a modellünk során általában a K hangforrás pozíciója.

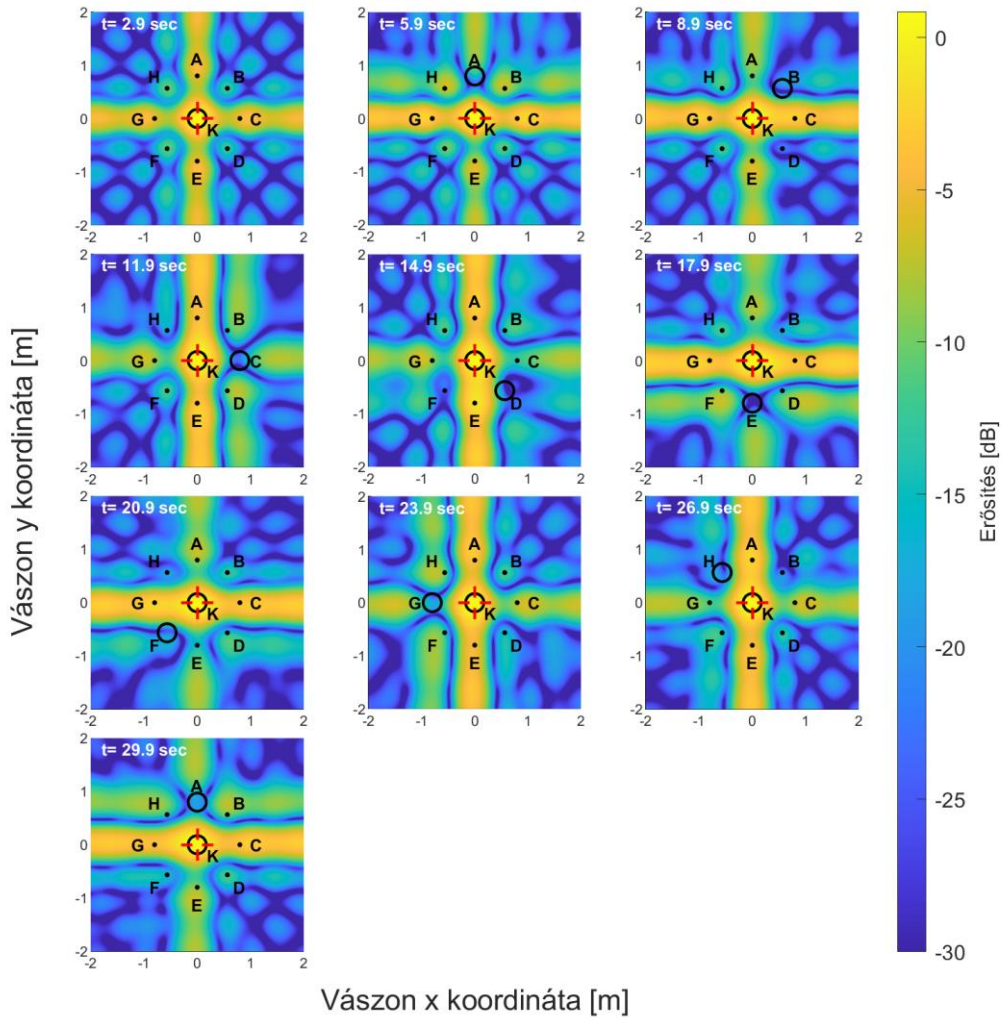
10.4.1 Egyes forrásokra kvázi-állandósult karakterisztikák

Az *dinamikus hangfájlr*a alkalmazott Robust Capon algoritmus során az akusztikus környezet változása előtti közvetlen pillanatokban adódó karakterisztikák a 44. ábra segítségével tekinthetők meg. Az egyéb alkalmazott paraméterek értékei az ábra címsorában szerepelnek. A karakterisztikák időrendben balról jobbra, illetve fentről lefele elhelyezve szerepelnek, de az egyes ábrák bal felső sarkában látható időpillanatokban tapasztalható állapotot jelenítik meg. Ezek az időpontok azért lényegesek, ugyanis ezek közvetlen megelőzik a sorban következő hangforrás aktiválódását, így helyesen beállított paraméterek esetén a karakterisztikáknak ez már kvázi állandósult állapota.

Az egyes karakterisztikákon látható, hogy a körökkel jelölt aktív hangforrások irányába a fókuszirányhoz képest minden esetben jellemzően legalább 10-15 dB elnyomás adódik, ami általában kisebb, mint abban az esetben, mint mikor az adott hangforrás éppen nem aktív. Ennek hatása leginkább a melléknyalábokba eső A, C, E, G jelű

hangforrások esetén nagyon feltűnő, ami felettébb biztató eredmény, hiszen a Bartlett alkalmazása esetén ezekben az irányokban volt leginkább probléma a zaj beszívargása.

Szimulált hangmintán kialakuló karakterisztika az egyes források megszólalását megelőzően, 3000 Hz frekvencián, Robust Capon algoritmus alkalmazásával
 átl. időállandó = 2.00 sec; bizonytalanság= 0.01 m;
 DFT mintaszám= 480 Hz; vászontávolság= 5.00 m;
 SNR= 50 dB;



44. ábra Kialakuló iránykarakterisztikák az akusztikus környezet változása során

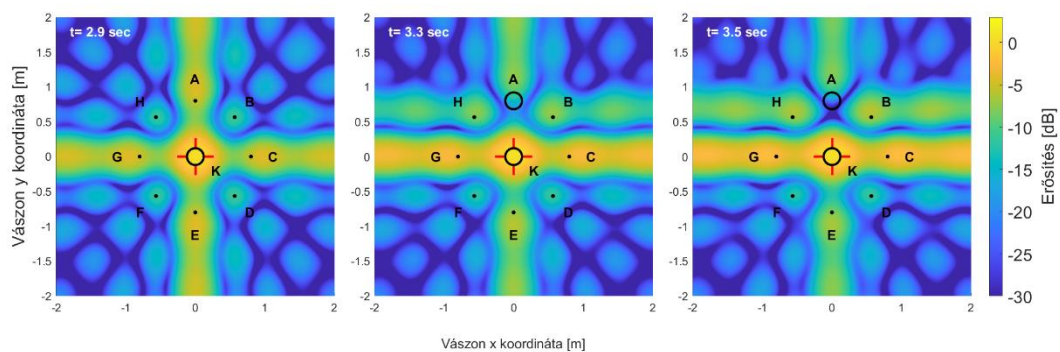
10.4.2 Karakterisztika alakulása akusztikus környezet változását követően

Az algoritmus eddigiekben is hangsúlyozott legnagyobb előnye, hogy a karakterisztika akusztikus környezetnek megfelelő dinamikus alakítására képes. Modellezett környezetben megvizsgáltam a karakterisztika alakulását az akusztikus környezet változásának hatására.

A dinamikus hangminta feldolgozása során a $t=2.9$ sec pillanatban kialakult, helyes paraméterek esetén kvázi állandósult iránykarakterisztika a $t=3$ sec időpont után az 'A' jelű hangforrás megszólalásának következtében az algoritmus elvárt működésének megfelelően módosulni kényszerül, hogy a kimenetre minimalizálja az 'A' hangforrásból származó teljesítményt. Ez a változás a Jelfeldolgozási folyamatban részletezettek szerint, az exponenciális átlagolással alakuló Kovarianciamátrix miatt nem pillanatszerű, hanem folyamatos. Ennek hatása tekinthető meg a 45. ábra által szemléltetett módon, a megszólalást követő néhány tizedmásodpercben.

A 45. ábra alapján az első, 'A' forrás megszólalását megelőző felvételen a karakterisztika tulajdonképpen nagyon hasonló a Bartlett algoritmus alkalmazásával nyert karakterisztikához. Ennek azonban drasztikus változása látható a következő kettő felvételen, ahol a melléknyalásban elhelyezkedő 'A' jelű hangforrásból származó jelteljesítmény az algoritmus hatására a karakterisztikát olyan mértékben torzítja, hogy az irányába eső függőleges melléknyaláb folytonossága tulajdonképpen egy nullhely beiktatásával megszűnik.

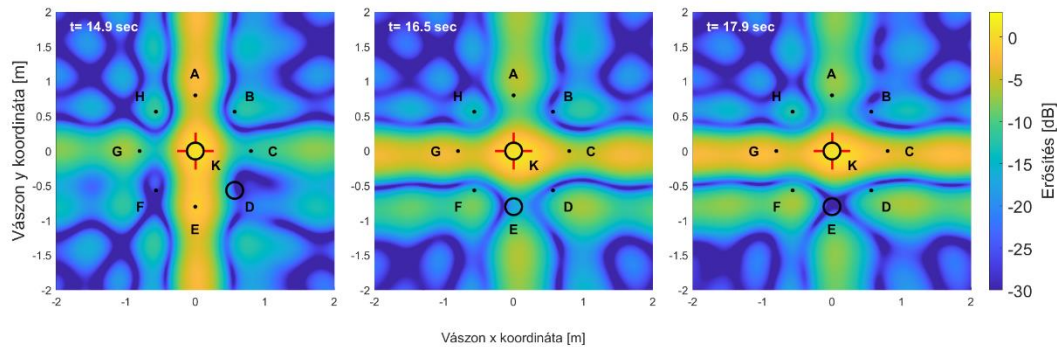
Szimulált hangminta kialakuló karakterisztika 'A' forrás megszólalását követően,
3000 Hz frekvencián, Robust Capon algoritmus alkalmazásával
átl. időállandó = 10.00 sec; bizonytalanság= 0.01 m; DFT mintaszám= 480 Hz; vászontávolság= 5.00 m; SNR= 50 dB;



45. ábra Karakterisztika változása „A” hangforrás megszólalását követően

A 46. ábra a korábban vázolt mérési elrendezést mutatja egy másik időpillanatban. Ekkor az 'E' jelű forrás aktiválódását követő másodpercekben figyelhető meg az iránykarakterisztika alakulása. Külön látványos, ahogy a dominánsan függőleges irányú melléknyaláb vízszintesre cserélődik fel, ami hangforrás legutóbbi pozíciójának hatását mutatja az exponenciálisan átlagoló kovarianciamátrixon.

Szimulált hangmintán kialakuló karakterisztika 'E' forrás megszólalását követően,
 3000 Hz frekvencián, Robust Capon algoritmus alkalmazásával
 átl. időállandó = 2.00 sec; bizonytalanság= 0.01 m; DFT mintaszám= 480 Hz; vászontávolság= 5.00 m; SNR= 50 dB;



46. ábra Karakterisztika változása „E” hangforrás megszólalását követően

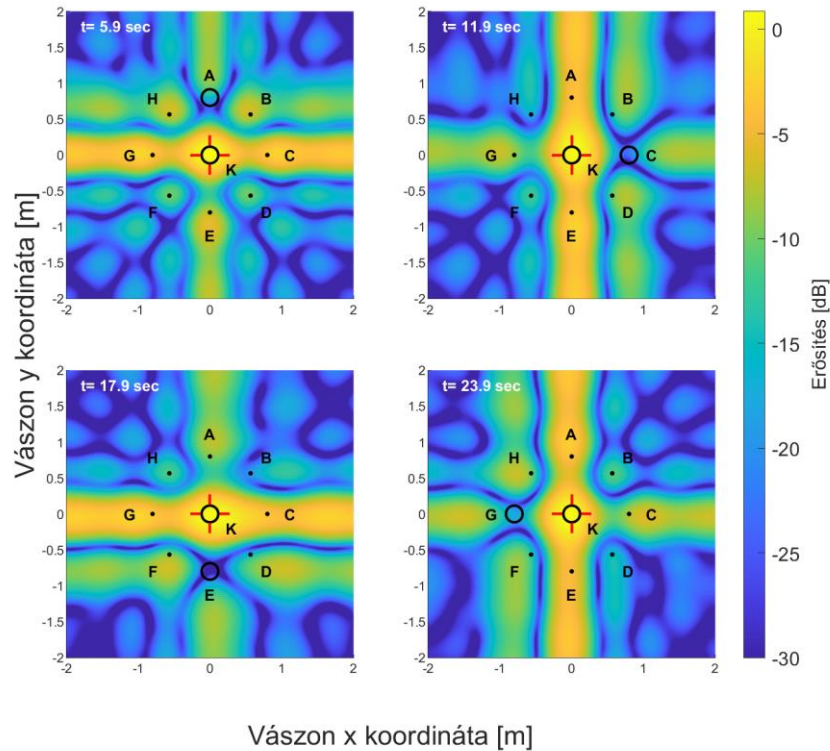
10.4.3 Karakterisztika pontatlan fókuszpozíció esetén

A modellezett akusztikai környezet legnagyobb előnye, hogy biztosak lehetünk abban, hogy egyes hangforrásoknak mi az adott pillanathoz tartozó térbeli elhelyezkedése. Az alkalmazott algoritmussal szemben természetesen alapvető követelmény, hogy a zavarforrások szempontjából ne kelljen rendelkezni ezzel az információval, és azok kioltása automatikusan történjen, viszont a relevánsnak tekintett forrás irányát mindenképpen pontosan kell ismernem, különben az algoritmus nem rendelkezhet információval arról, hogy a relevánsnak tekintett forrás által sugárzott jel nem-e szintén valami külső zavarforrás.

Ez a definit szimulált környezet az algoritmus tesztelése szempontjából rendkívül előnyös, azonban valós alkalmazás során a releváns hangforrás pozícióját nemcsak hogy nem feltétlenül ismerhetjük pontosan, annak hagyományos módszerekkel történő megmérése esetén sem teljesen triviális, hogy az objektumon belül pontosan hova tehető a keltett hanghullámok forrásponjtja.

A 47. ábra mutatja a helyes fókuszirányzás esetén kialakuló kvázi állandósult karakterisztikát az egyes melléknyalábokra eső hangforrások jelenléte mellett a *dinamikus hangminta* esetén. Ennek lényeges tulajdonságait a korábbi fejezetekben már tárgyaltam. Ezen, illetve a következő ábrákon a címsorban feltűntetésre került a fókuszálási pont koordinátája is a könnyebb értelmezhetőség kedvéért.

**Szimulált hangmintán kialakuló karakterisztika
az egyes források megszólalását megelőzően,
3000 Hz frekvencián, Robust Capon algoritmus alkalmazásával**
 átl. időállandó = 2.00 sec; bizonytalanság= 0.01 m;
 DFT mintaszám= 480 Hz; SNR= 50 dB;
 vászontávolság= 5.00 m;
 fókuszpont(x,y,z)= [5.00 , 0.00, 0.00] m;



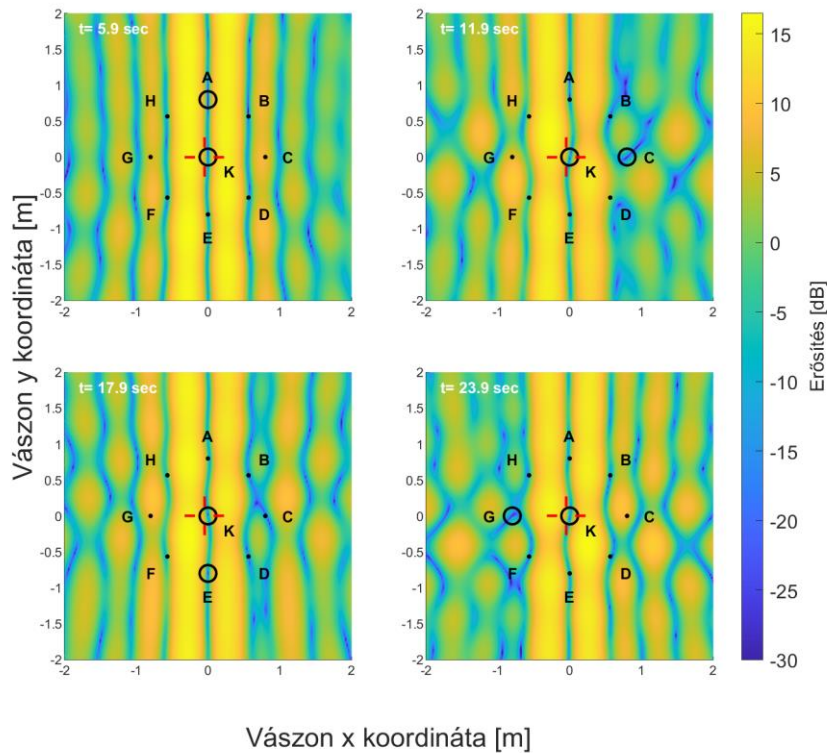
47. ábra Kialakuló karakterisztika pontos fókuszirányzással

Abban az esetben, ha valami oknál fogva az ismeretünk a hangforrás pozíciójáról nem kellőképpen pontos, akkor a kialakuló karakterisztika a fent jelzett ideálistól jelentős mértékben eltérhet. Erre példa 48. ábra, ahol az 5 m re lévő forrás pozíciójához képest 0.05 m hibával került a fókuszálási pont elhelyezésre y irányba 0.01 m-es bizonytalansági érték mellett. Ennek hatására a mikrofontömbünk középpontját is jelentő origóból nézve a feltételezett forrás iránya közel 0.57° -kal tér el a forrás valódi irányától. Az így kialakuló karakterisztika a korábbi helyes fókuszálás során alkalmazottól jelentősen eltér. Ennek leglátványosabb hatása a K jelű forrás közel nullhelyre kerülése, illetve a 15 dB-nél is nagyobb erősítésértékek megjelenése. Ez kétségtelenül kedvezőtlen karakterisztika a K jelű forrás megfigyelésének céljára, viszont említésre méltó, hogy az éppen aktív források irányába ez esetben is nullhely került, és a tévesen feltételezett fókuszálási irányban is közel egységnyi erősítés található. A drasztikus karakterisztikatorzulás oka tehát, hogy a K forrást egy fókuszálási irányhoz nagyon közeli zavarforrásnak tekinti, aminek kioltásához az egységnyi fókuszirányú erősítés mellett a

nagy térbeli erősítésváltozás szükséges, ami a hatalmas erősítésű régiók kialakulásának oka.

Szimulált hangmintán kialakuló karakterisztika az egyes források megszólalását megelőzően, 3000 Hz frekvencián, Robust Capon algoritmus alkalmazásával

átl. időállandó = 2.00 sec; bizonytalanság= 0.01 m;
 DFT mintaszám= 480 Hz; SNR= 50 dB;
 vászontávolság= 5.00 m;
 fókuszpont(x,y,z)= [5.00 , 0.05, 0.00] m;

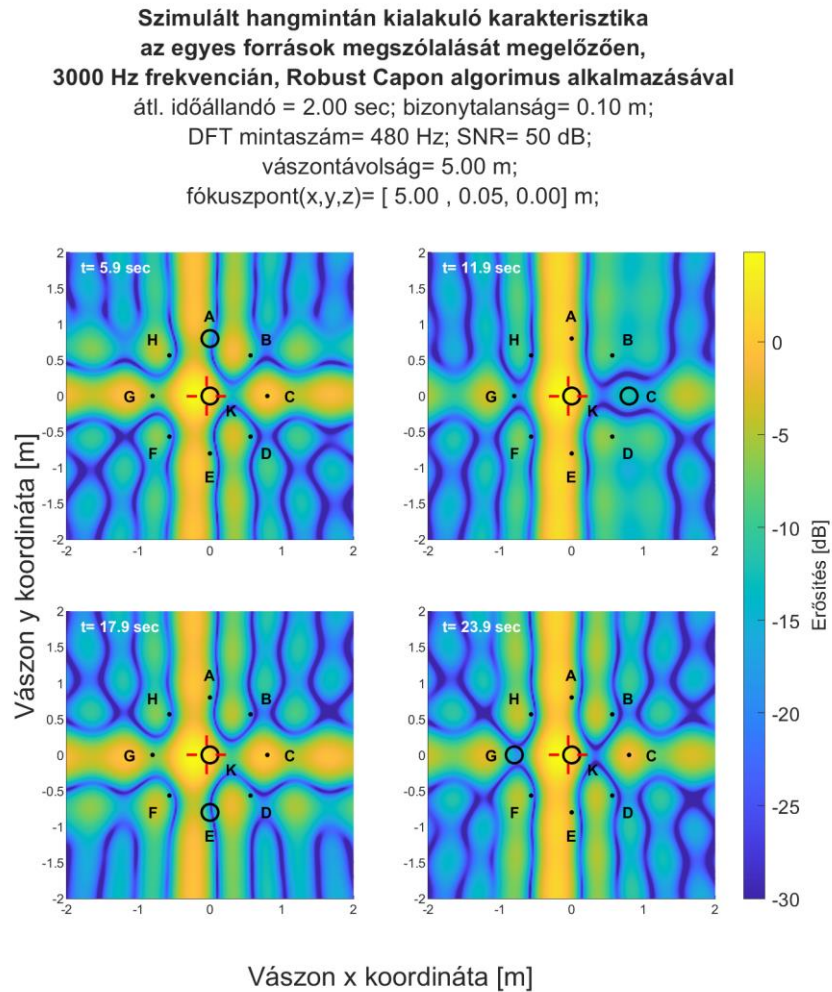


48. ábra Kialakuló karakterisztika pontatlan fókuszirányzással

Tulajdonképpen a kialakult probléma pontosan ugyanaz, mint amit a Standart Capon algoritmus alkalmazása során tapasztalhattunk. Ennek megszüntetésére került bevezetésre a Robust Capon algoritmus által a bizonytalansági érték fogalma, ami tulajdonképpen a jel kontrollált mértékű zajosítása mellett a Capon algoritmus fókuszálási irány szempontjából engedékenyebb alkalmazását biztosítja.

A bizonytalansági érték 0.1 m-re állításának hatására 49. ábra által bemutatott karakterisztika alakul ki a korábbival megegyező, 0.05-m fókuszálási hiba esetén. Ez az a gyakorlati alkalmazás szempontjából sokkal kedvezőbb eredmény, nem láthatóak hatalmas erősítésű melléknyalábok, illetve a releváns K forrás jele is közel egységnyi erősítéssel kerül a kimenetre. Ami természetesen az ára ennek a pontatlanságnak, illetve az azt kiküszöbölő eljárásnak, hogy a valódi zajforrások immár csupán jóval kisebb elnyomással kerülnek a kimenetre a kisebb bizonytalansági értékek alkalmazása során tapasztaltakhoz képest. A 'C' jelű forrás esetén például egyértelműen megfigyelhető,

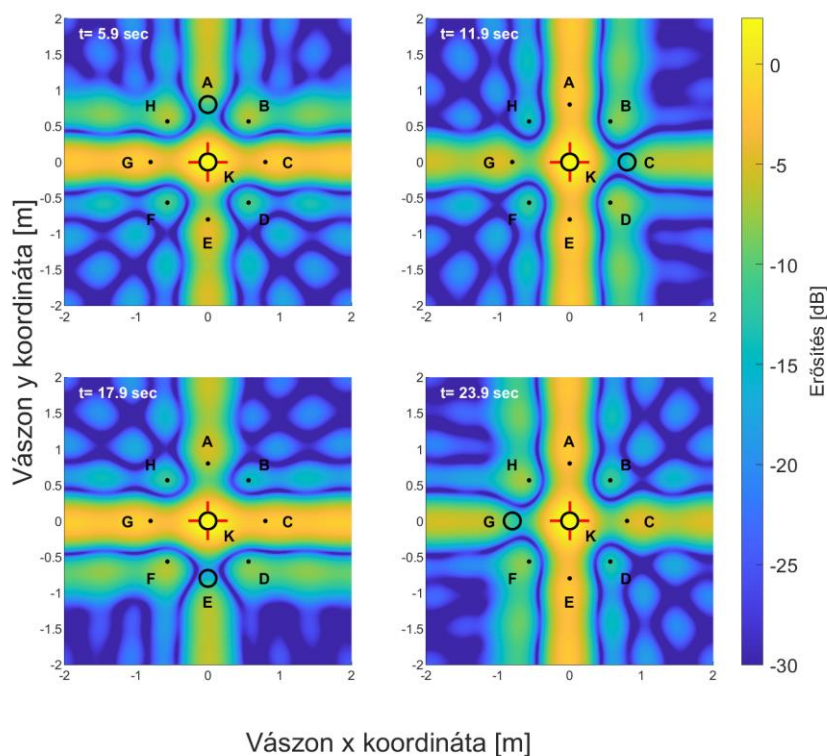
hogy míg kisebb bizonytalanság esetén konkrét nullhely haladt rajta át, addig a korrigált esetben csupán egy nagyobb, közel 15 dB elnyomású melléknyaláiban foglal helyet.



49. ábra Kialakuló karakterisztika pontatlan fókuszirányzás, de nagyobb bizonytalansági érték esetén

Abban az esetben is, ha helyesen történik a releváns forrás pozíciójának meghatározása, a nagyobb bizonytalansági érték kisebb elnyomásokat eredményez a zajforrások irányába, ahogy ezt az 50. ábra is mutatja. Az algoritmus alkalmazása ilyen szempontból a lehető legkisebb bizonytalansági érték mellett célszerű.

Szimulált hangmintán kialakuló karakterisztika az egyes források megszólalását megelőzően, 3000 Hz frekvencián, Robust Capon algoritmus alkalmazásával
 átl. időállandó = 2.00 sec; bizonytalanság= 0.10 m;
 DFT mintaszám= 480 Hz; SNR= 50 dB;
 vászontávolság= 5.00 m;
 fókuszpont(x,y,z)= [5.00 , 0.00, 0.00] m;



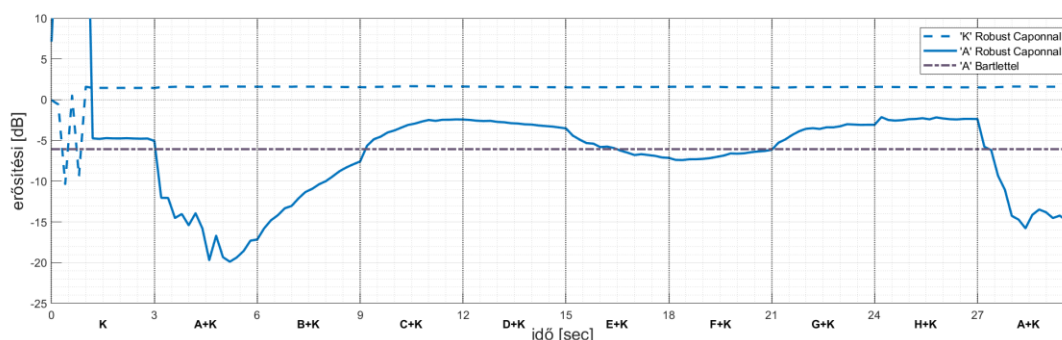
50. ábra Kialakuló karakterisztika pontos fókuszirányzás, de nagyobb bizonytalansági érték esetén

10.4.4 Erősítési értékek ábrázolása

Annak érdekében, hogy a karakterisztika változásáról pontosabb képet kaphassunk, ábrázoltam az adott pontban, adott frekvenciákon mérhető erősítési értéket az idő függvényében a *dinamikus hangmintára* alkalmazott nyalábformálás során.

Az 51. ábra az 'A' pontbeli erősítés értékét mutatja a nyalábformálás során 3000 Hz frekvencián. Ugyanezen a grafikonon színén ábrázoltam 'K' fókuszpontban mérhető erősítés értékét, valamint az 'A' pontban mérhető erősítési értéket Bartlett nyalábformálás alkalmazása esetén.

'A' hangforrás irányába mérhető erősítési érték
 Robust Capon nyalábformálás során 3000 Hz frekvencián
 átl. időállandó = 2.00 sec; bizonytalanság= 0.05 m;
 DFT mintaszám= 480 db; SNR= 50 dB;



51. ábra 'A' hangforrás erősítés-idő függvénye

A grafikonon több említésre méltó jelenség is megfigyelhető.

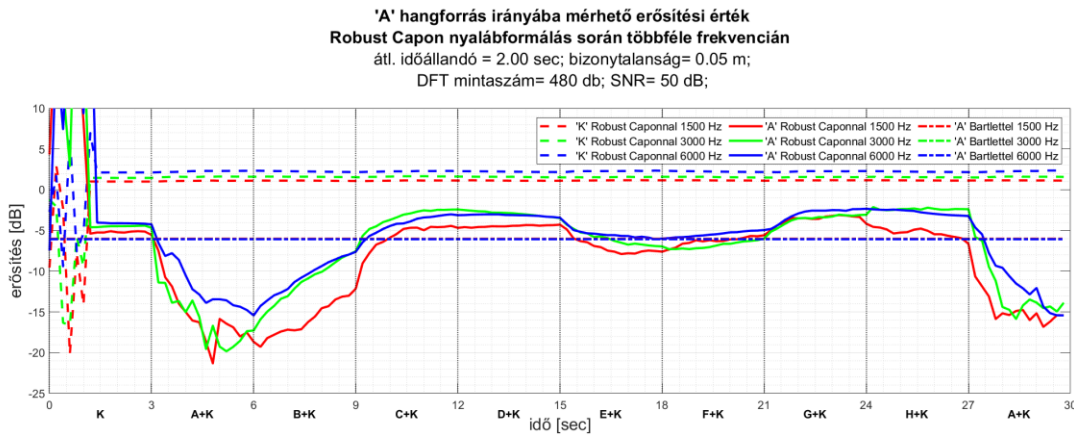
A Bartlett eredményeképpen kapott erősítési érték az egész folyamat során állandó, hiszen az független az akusztikus környezettől. Ehhez képest viszont a Robust Capon algoritmus csupán nagyjából az első másodperc után nyújt értékelhető eredményt, hiszen a kovarianciamátrix kialakítása során alkalmazott várhatóérték képzés miatt szükség van egy adott mennyiségű mintára annak jó közelítéséhez.

Az iránykarakterisztika stabilizálódását követően 0 dB közeli erősítést kaphatunk fókuszálási irányba, és -5 dB közelit az akkor még néma 'A' forrás irányába. Ezek mindegyike nagyobb, mint a Bartlett algoritmus alkalmazása esetén adódó értékek, azonban a $t=3$ sec pillanatban bekövetkező 'A' megszólalását követően az irányában mérhető erősítés értéke drasztikusan csökken, egészen a -20 dB értéket is eléri. A $t=6$ sec pillanatban bekövetkező 'A' forrás elnémulását, és 'B' megszólalását követően 'A' erősítése folyamatos emelkedik, a későbbiekben eléri, sőt bizonyos esetekben meg is haladja a $0 < t < 3$ sec intervallumban tapasztalható értéket. Miután 'A' ismét aktiválódik $t=27$ sec pillanatot követően, az irányába mérhető erősítés ismételen csökkenni kezd, ekkor -15 dB körüli értékeket ér el.

Többféle frekvencián is lehetőség van az analízis elvégzésére, így az algoritmus frekvenciafüggő viselkedésével kapcsolatban szerezhethetünk mélyebb ismereteket. A nyalábformálás szempontjából tulajdonképpen lényegtelen kellene legyen a vizsgálati frekvencia, azonban a jelfeldolgozás során alkalmazott paraméterek, és a mikrofontömb térbeli szerkezete szempontjából a vizsgált frekvencia eltérő viselkedéseket eredményez.

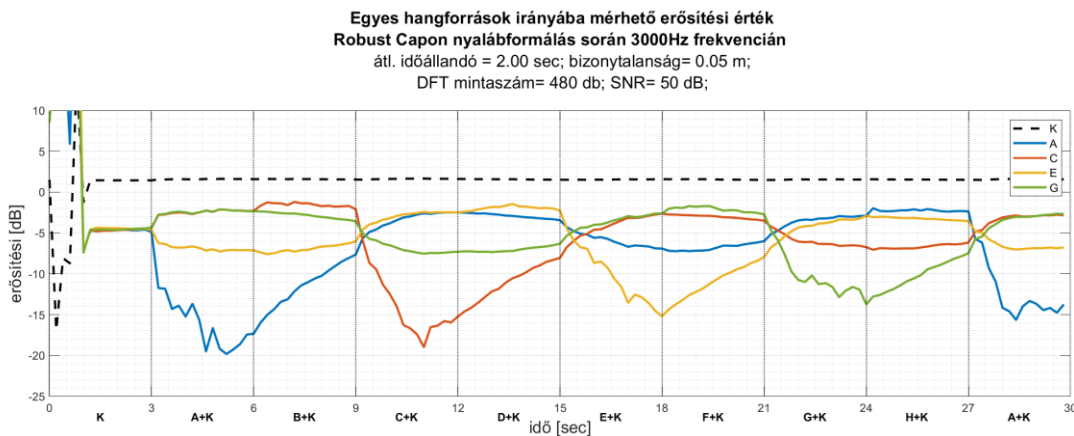
Az 52. ábra által szemléltetett módon 1500 Hz, 3000 Hz illetve 6000 Hz frekvencián is elvégeztem az algoritmus vizsgálatát, majd azok eredményét közös grafikonokon ábrázoltam. A vizsgált frekvenciák mindegyikén megfigyelhetőek a korábbi esetben is tapasztalt viselkedések, ami előnyös az algoritmus szélessávú alkalmazása során.

Ami megállapítható, hogy 6000 Hz esetén az elnyomások nem olyan határozottak, mint alacsonyabb frekvenciákon, illetve hogy 1500 Hz-en tapasztalhatók a legjobb eredmények.



52. ábra „A” hangforrás erősítési értékek több frekvencián

Annak érdekében, hogy pontos képet kaphassunk a nyalábformálás során az egyes források irányába mérhető erősítések értékéről, miközben azok felváltva szólalnak meg, a korábban bemutatott analíziseket elvégeztem úgy is, hogy egyetlen frekvencián, viszont egyszerre több forrás irányába mérhető erősítési érték is szerepel. A források közül a melléknyalábokba eső A, C, E és G jelű forrásokat ábrázolom, ugyanis az alkalmazott algoritmus legnagyobb előnye éppen az, hogy a melléknyalábok struktúráját képes alakítani az akusztikai környezetnek megfelelően.



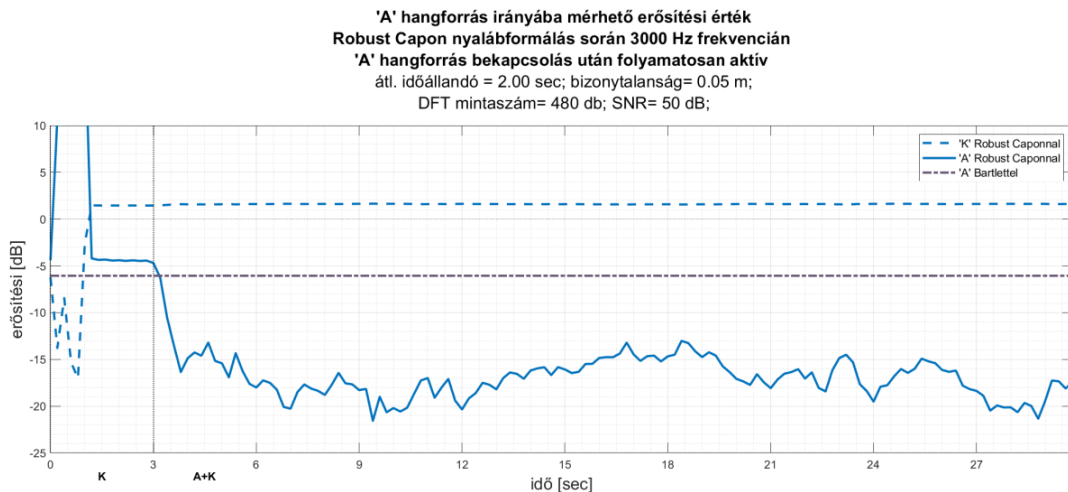
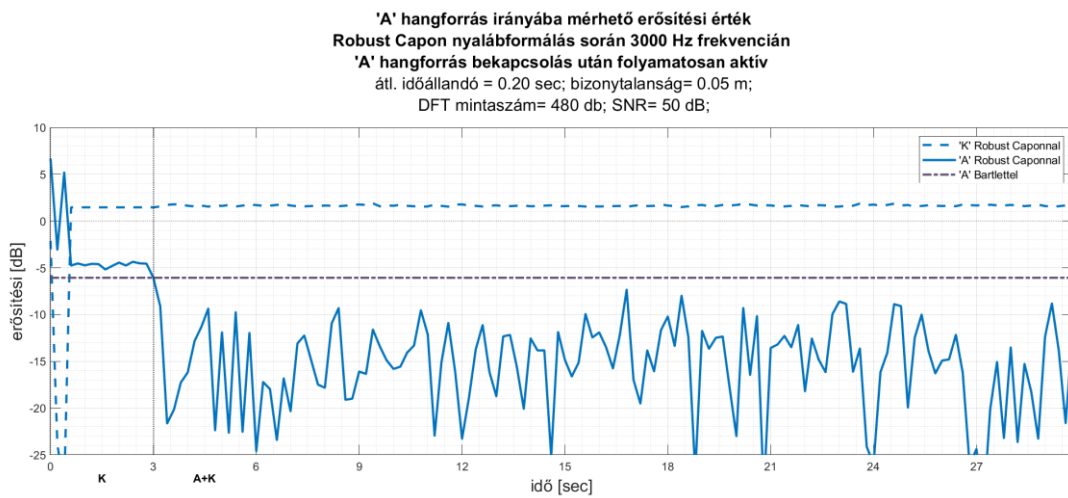
53. ábra Több forrás erősítésértékei

Látható, ahogy az egyes forráspozíciókhoz tartozó erősítésértékek az adott forrás megszólalásával összhangban csökkennek le, egymást követően, sorban. A csillapítások mértéke valami oknál fogva nem azonos minden esetben, egyértelmű csökkenés tapasztalható 'A'-tól 'G' fele haladva, viszont mikor ismét 'A' forrás következik, az erősítésérték ismét megugrik. Mivel az elrendezés teljesen szimmetrikus, ezért a jelenséget mindenképpen a kovarianciamátrix időbeli alakulásához kell kötni, viszont

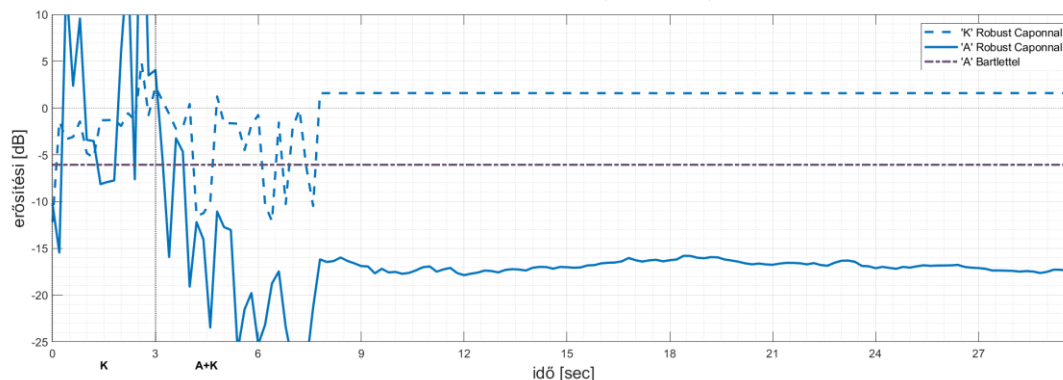
további vizsgálat lenne szükséges általános megállapítások tételéhez. Az algoritmus viszont így is 5-10 dB-lel nagyobb elnyomásokat eredményez a Bartletthez képest.

10.4.5 Átlagolási idő hatása

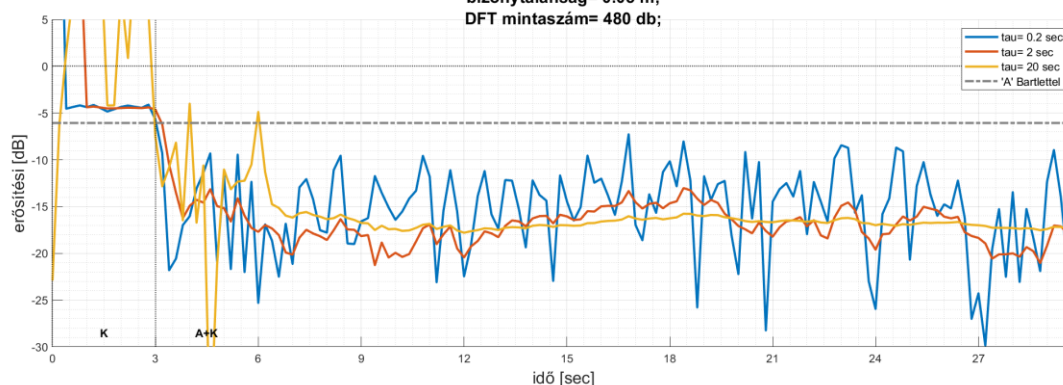
Az 54. ábra a $\tau=0.2, 2$ sec, illetve 20 sec kovarianciamátrix átlagolási idő esetén 'A' irányban kialakuló erősítésértéket mutatja a *statikus hangmintán*. Jól látszik, ahogy minél nagyobb az átlagolási idő, annál kevesebb a már kialakult karakterisztika erősítésértékének ingadozása, viszont cserébe annál lassabban alakul ki ez a stabilizálódott, fókuszálási irányban nagyjából egységnyi erősítéssel rendelkező karakterisztika, illetve ilyen esetben az az akusztikai környezet váltására is lassabban reagál.



'A' hangforrás irányába mérhető erősítési érték
 Robust Capon nyálábformálás során 3000 Hz frekvencián
 'A' hangforrás bekapcsolás után folyamatosan aktív
 átl. időállandó = 20.00 sec; bizonytalanság= 0.05 m;
 DFT mintaszám= 480 db; SNR= 50 dB;



'A' hangforrás irányába mérhető erősítési érték
 Robust Capon nyálábformálás során 3000 Hz frekvencián
 Különféle átlagolási időállandók mellett
 'A' hangforrás bekapcsolás után folyamatosan aktív
 bizonytalanság= 0.05 m;
 DFT mintaszám= 480 db;



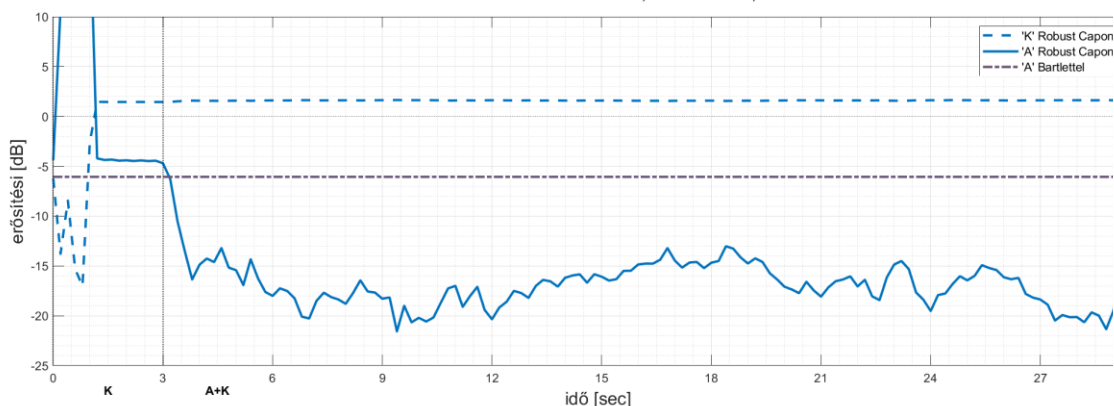
54. ábra Átlagolási idő hatása a módosított hangmintán

10.4.6 Fókuszált hangforrás hangteljesítménye

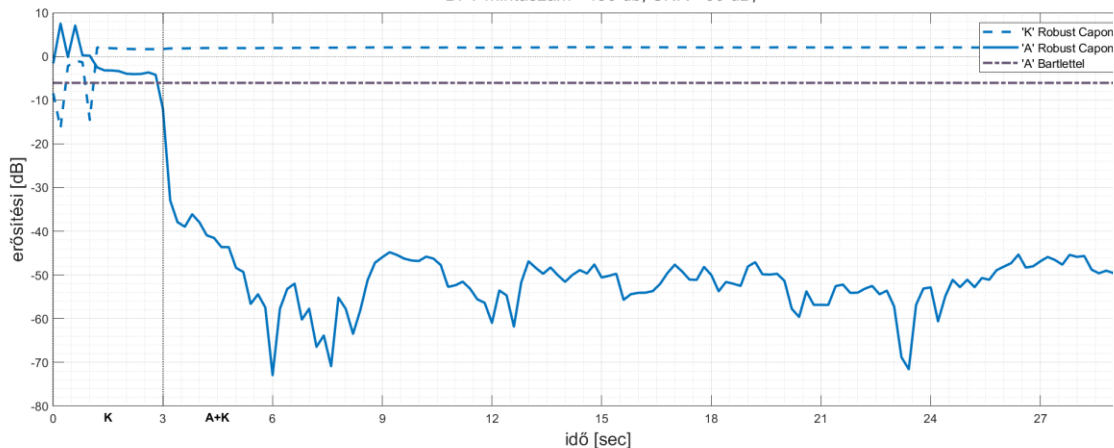
Eddigiek során az egyszerűség kedvéért olyan eseteket figyeltünk meg, ahol a vizsgált forrás és a zavaró források megegyező effektív értékű független Gaussi fehér zajt bocsájtottak ki. Annak vizsgálata érdekében, hogy a milyen hatása van annak, ha a zavaró források jóval nagyobb teljesítményűek, mint a vizsgált hangforrás létrehoztam a statikus hangminta egy olyan módosítását, melyen az 'A' jelű hangforrás 20 dB-lel nagyobb teljesítményű jelet sugároz ki, mint a 'K' jelű.

Az 55. ábra összehasonlítja azt a két esetet, amikor a zajforrás teljesítménye megegyező, illetve 20 dB-el nagyobb értékű, mint a számunkra releváns jelforrás.

'A' hangforrás irányába mérhető erősítési érték
 Robust Capon nyálábformálás során 3000 Hz frekvencián
 'A' hangforrás bekapcsolás után folyamatosan aktív
 átl. időállandó = 2.00 sec; bizonytalanság= 0.05 m;
 DFT mintaszám= 480 db; SNR= 50 dB;



'A' hangforrás irányába mérhető erősítési érték
 Robust Capon nyálábformálás során 3000 Hz frekvencián
 'A' hangforrás bekapcsolás után folyamatosan aktív
 és 20 dB-el nagyobb teljesítményű, mint 'K'
 átl. időállandó = 2.00 sec; bizonytalanság= 0.05 m;
 DFT mintaszám= 480 db; SNR= 50 dB;



55. ábra Zavarforrás teljesítményének hatása a nyálábformálás során kialakuló elnyomása szempontjából

Látható, hogy abban az esetben, amennyiben a zavarforrás 20 dB-el nagyobb teljesítményű, mint a megfigyelni kívánt forrás, a zavarforrás irányába 50 dB körüli elnyomásérték a jellemző, míg abban az esetben, ha teljesítményük azonos, úgy csak nagyjából 15 dB adódik erre azonos körülmények között, és azonos paraméterek alkalmazása mellett. Ez alapján tulajdonképpen jobb jel-zaj viszony érhető el az algoritmus alkalmazásán keresztül abban az esetben, ha az eredeti jel-zaj viszony rosszabb.

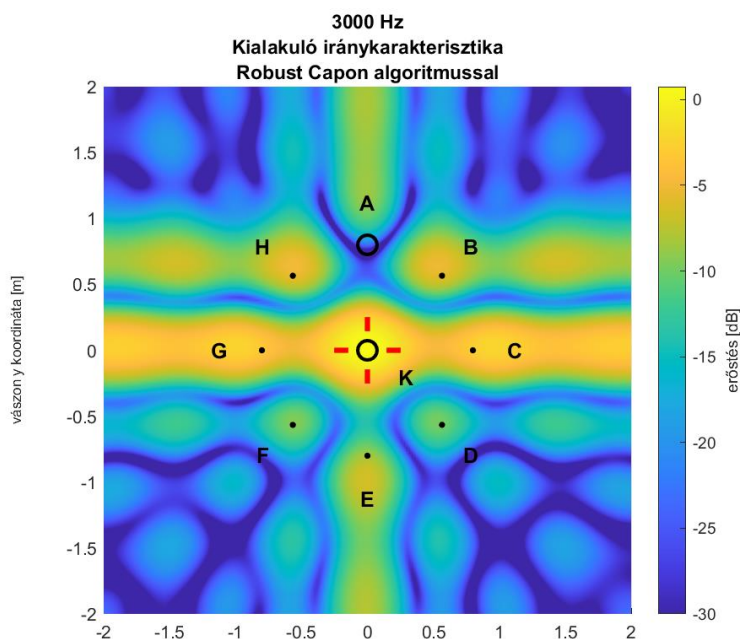
10.4.7 Szélessávú viselkedés impulzusválasszal

Az algoritmus hatékony alkalmazása érdekében a korábban is említetteknek megfelelően a legkézenfekvőbb megoldás egy olyan 33. ábra által is bemutatott FIR szűrő-sor létrehozása, melyeken keresztül a megfelelő mikrofonok jeleit tovább vezetve, majd azokat összegezve az elvárt eredményhez juthatunk.

Az akusztikus környezethez Robust Capon nyalábformálás segítségével folyamatosan alkalmazkodó szűrősort létrehozó algoritmust a korábbi fejezetben már bemutatott szimulációs környezetben teszteltem. Az ilyen szimulált környezetnek a korábban említettekén kívül további előnye, hogy lehetőség van olyan mérési összeállítások létrehozására, amire valós fizikai körülmények között igencsak nehézkes lenne.

A *statikus hangmintát* alkalmazva, mikrofontömbömmel a korábbiakhoz hasonló módon 'K' hangforrás pozícióját jelöltem ki fókuszálási irányképp. A hangminta végére a nyalábformálás eredményeképpen kialakuló iránykarakteristikákon az 56. ábra által is szemléltetett módon megfigyelhető az elnyomás 'A' hangforrás irányába.

**'A' forrás hatására alakuló
Robust Capon nyalábformálás aktív 'K' fókuszirányú hangforrás mellett**
 átl. időállandó = 2.00 sec; bizonytalanság= 0.01 m;
 DFT mintaszám= 480 db; SNR= 50 dB;
 t= 30.0 sec;



56. ábra Nyalábformálás során kialakuló iránykarakterisztika

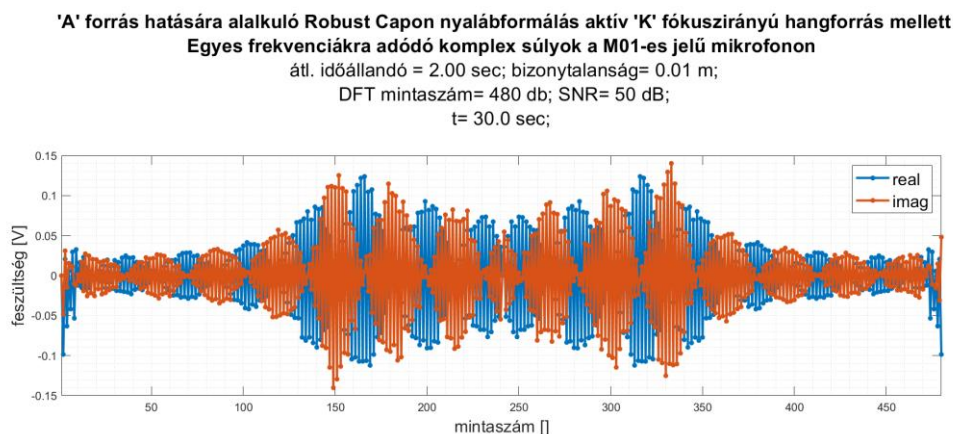
A hangminta feldolgozása során a mikrofonok által vett jelek alapján a 30. ábra által bemutatott módon, szélessávban végeztem a kovarianciamátrixok folyamatos frissítését, minden egyes DFT által adódó frekvenciaértékre.

Adott időközönként az éppen rendelkezésre álló kovarianciamátrixokra a már említett további paraméterek mellett alkalmaztam a Robust Capon algoritmust, melynek eredményeképpen minden egyes mikrofonra, minden egyes DFT által adódó frekvenciaértéken rendelkezésemre állt egy-egy komplex súlyérték. Ez tulajdonképpen felfogható úgy, mint az egyes mikrofonok átviteli függvénye. Ebből Inverz DFT segítségével megkapható az adott mikrofon diszkrét értékű impulzusválasza.

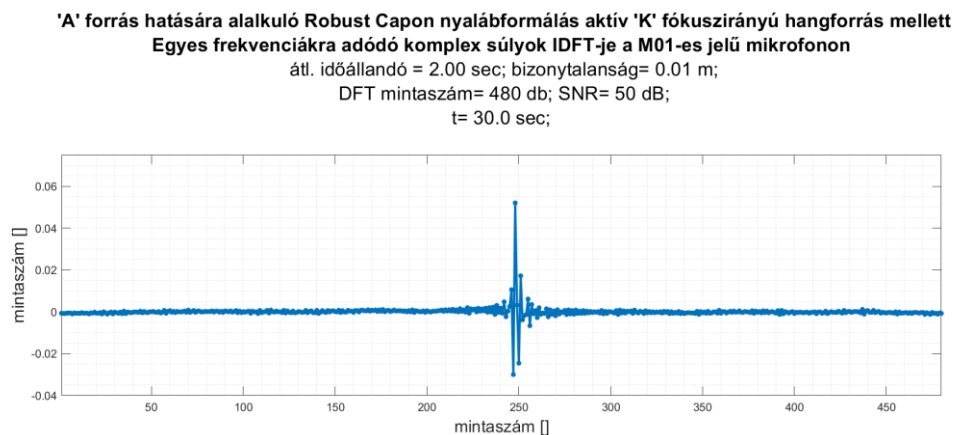
A probléma azonban, hogy az egyszerű Capon algoritmussal ellentétben a Robust Capon esetén az átviteli függvény inverz DFT-je nem tisztán valós. Ez annak köszönhető, hogy az átviteli függvény elemei a második elemtől haladva nem egyeznek meg annak elemeinek a komplex konjugáltjaival a végéről előrefele haladva.

Ez arra vezethető vissza, hogy bár a jelterjedés, és a terjedési vektor számításakor ugyanazt az eredményt kapnánk az $n+1$ -edik komplex súlynak, mint az $N-n$ -edik súly komplex konjugáltjának, azonban a bizonytalansági érték számításakor annak hullámhossztól való függősége miatt ez a szimmetria nem teljesül. A probléma kiküszöbölése érdekében a diszkrét átviteli függvény második felét egyszerűen eldobtam, és az első felének komplex konjugált tükörképével helyettesítettem. Így olyan átviteli függvényhez jutottam, melynek Fourier-transzformáltja tisztán valós, és továbbra is teljesíti az elvártakat. Egy ilyen átviteli függvényre példa az 57. ábra. Ennek IDFT-je pedig a mikrofonhoz tartozó szűrő impulzusválasza, amit az 58. ábra mutat.

A referenciául szolgáló Bartlett algoritmus esetén is megszerezhetők az egyes mikrofonok szűrőihez tartozó impulzusválaszok a bemutatott módszer segítségével. Néhány ilyen mikrofonoszűrőhöz tartozó impulzusválaszt mutat az 59. ábra.



57. ábra Egyes frekvenciaértékekre adódó komplex súlyok a nyalábformálás során egy adott mikrofon esetén

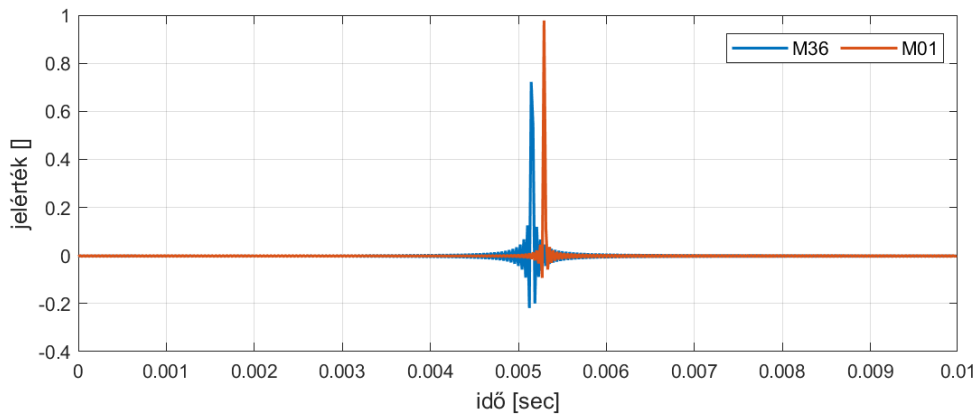


58. ábra M01-es mikrofonhoz tartozó szűrő impulzusválasza az átviteli függvény IDFT-je alapján

Bartlett nyalábformálás impulzusválasza néhány mikrofonon

átl. időállandó = 2.00 sec; bizonytalanság= 0.01 m;

DFT mintaszám= 480 db;



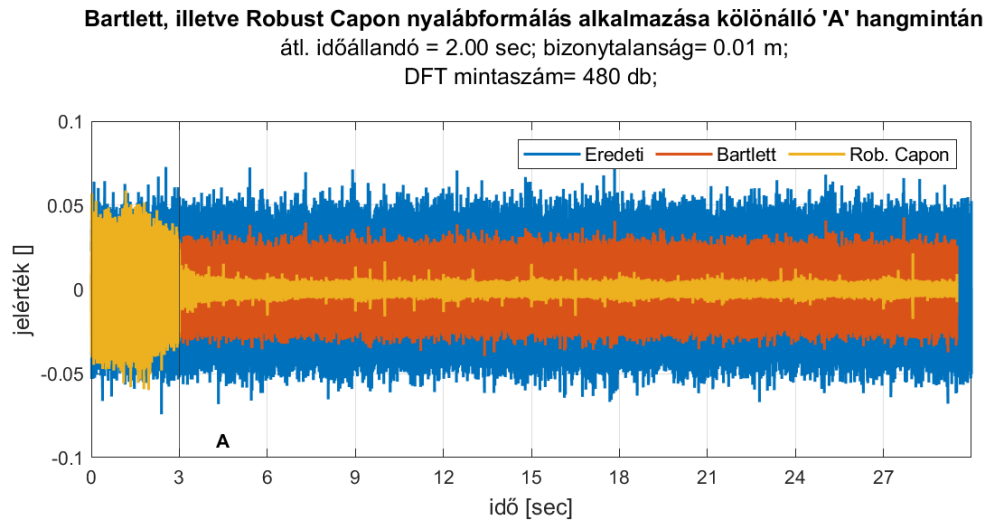
59. ábra Bartlett nyalábformálás során adódó impulzusválaszok

Az adaptív algoritmushoz tartozó impulzusválaszok a folyamatosan frissülő kovarianciamátrixokból képződnek, azonban nem célszerű minden egyes kovarianciamátrix frissülést követően frissíteni az impulzusválaszokat is, ezért azok frissítését csupán jóval ritkábban, 0.5 másodpercenként végeztem. Ezt nevezzük az *impulzusválasz frissítési időnek*. Két frissülés között a legutóbbiak mikrofonkimenettel vett konvolúcióinak összege adja a teljes mikrofontömb kimenetét. Annak érdekében, hogy a kimenet közel folytonos hatást keltsen, figyelni kell arra, hogy az újonnan frissült impulzusválasz hatására megjelenő kimenetekhez még hozzá kell adni az azt megelőző impulzusválasz hatására keletkező, de még le nem zajlott kimeneteket, így az egymást követő szűrők hatását mintegy összefésülve.

A szimulált környezet által lehetőség van arra, hogy az egyes algoritmusok hatását az egyes jelforrásokra egymástól függetlenül vizsgálhassuk meg. A fejezet elején tárgyalt, végig aktív 'K', majd $t=3$ sec pillanatban megszólaló, vele megegyező effektív értékű 'A' zajforrások, valamint 'K' irányú fókuszálás hatására adott időközönként frissülő impulzusválaszokat kaptam az egyes mikrofonokra. Ezeket az impulzusválaszokat nem az eredeti mikrofonjelekre alkalmaztam, amiknek hatására ezek kialakultak, hanem azokra a jelekre, amik a mikrofonok kimenetén külön az 'A', illetve külön a K hangforrás hatására jelennének meg.

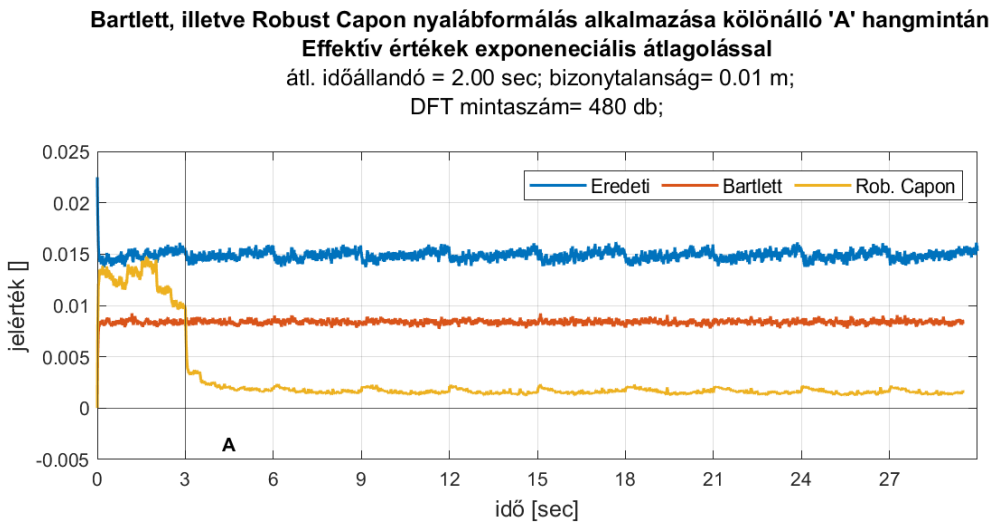
A különálló, folyamatos 'A' hangmintára ilyen módon adódó kimentet mutatja a 60. ábra melyen megfigyelhető az 'A' hangforrás hatása bármiféle szűrés nélkül az M01-es mikrofon közvetlen kimenetén, illetve Bartlett, valamint Robust Capon nyalábformálás mellett az említett korábbi hangmintára. Látható, hogy a $t=3$ sec pillanatban bekövetkező 'A' forrás megszólalásáig az 'A' irányából szóló próbajel Robust Capon esetén minimális csillapítást szenved, azt követően viszont hirtelen lecsillapodik. Ehhez

képezt a Bartlett hatására a próbajel végig azonos szinten mérhető a kimeneten, ami a $t=3$ sec pillanatot követően nagyobb, mint a Robust Capon alkalmazása esetén.



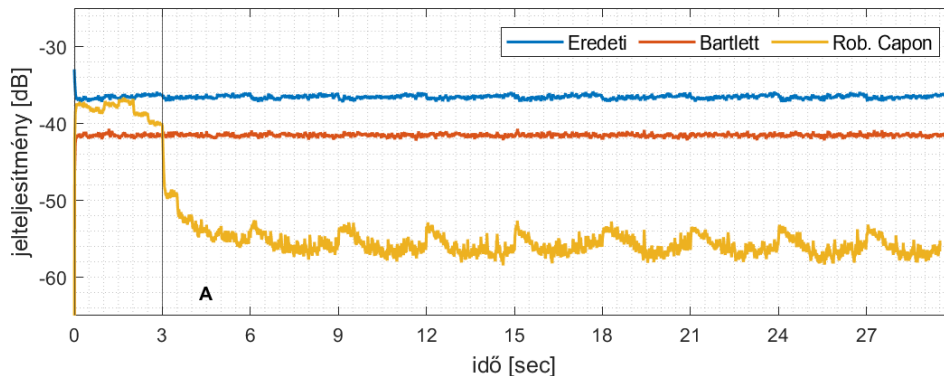
60. ábra Nyalábformálás eredménye különálló 'A' jelforrás hatására

A jelteljesítmények jobb átláthatósága érdekében a 61. ábra mutatja a 60. ábra esetén is ábrázolt jelek effektív értékeinek alakulását idő függvényében, valamint ugyanezt logaritmusos tengelyen az 62. ábra.



61. ábra 'A' próbajel által különféle nyalábformálás során létrehozott kimenetek effektív értékei idő függvényében

Bartlett, illetve Robust Capon nyálábformálás alkalmazása különálló 'A' hangmintán
Effektív értékek exponenciális átlagolással
 átl. időállandó = 2.00 sec; bizonytalanság= 0.01 m;
 DFT mintaszám= 480 db;

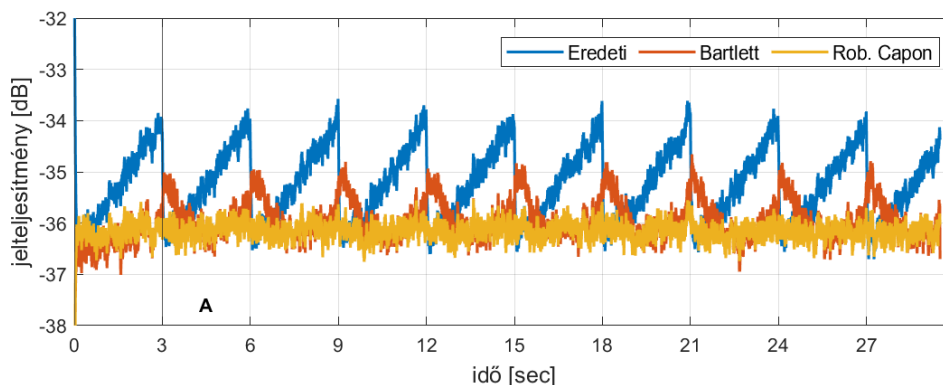


62. ábra 'A' próbajel által különféle nyálábformálás során létrehozott kimenetek effektív értékei idő függvényében logaritmus skálán

Ezek alapján látható, hogy időtartományban vizsgálva, szélessávú jelek esetén, fókuszirányú jel mellett is megvalósul, hogy az 'A' forrás jele 15 dB-el nagyobb csillapítást szenved el Robust Capon alkalmazása mellett, mint ugyanilyen körülmények között Bartlett nyálábformálás esetén. Az is megfigyelhető, ahogy az 'A' irányba mérhető csillapítás mértéke csupán az A hangforrás megszólalását követően kezd el, nagyjából exponenciális jelleggel csökkenni, megközelítve a 28 dB elnyomási értéket. Ez a jelenség utal az alkalmazott nyálábformáló algoritmus adaptív viselkedésére.

Az előzőekben tárgyalt analízist elvégeztem a 'K' fókuszirányú próbajelre is, ennek eredményét a 63. ábra mutatja. Ezen jól megfigyelhető, hogy 'A' próbajellel ellentétben 'K' effektív értéke nem változik meg 'A' hangforrás jelének megjelenésének hatására néhány dB-nél nagyobb mértékben, ami kiválóan mutatja a fókuszirányú erősítés biztosításának tulajdonságát. Érdekes jelenség az eredeti bemeneti jel néhány dB-es periodikus ingadozása. Ez az akusztikus környezet modellező algoritmusnak valamilyen tisztázatlan viselkedése. A periodicitás ahhoz köthető, hogy az adott pontból sugárzott hangminta 3 másodperces blokkonként lett szimulálva, majd azok egymás után fűzve. Úgy tűnik, valamiért a szimuláció során néhány dB-el ingadozik a kibocsájtott jel teljesítménye, bár ez a szempontunkból nem okoz jelentősebb problémát.

Bartlett, illetve Robust Capon nyalábformálás alkalmazása különálló 'K' hangmintán
Effektív értékek exponenciális átlagolással
 átl. időállandó = 2.00 sec; bizonytalanság= 0.01 m;
 DFT mintaszám= 480 db;



63. ábra 'K' próbajel által különféle nyalábformálás során létrehozott kimenetek effektív értékei idő függvényében logaritmus skálán

Az eddig vázolt vizsgálatokat elvégeztem a szimulált *dinamikus* környezetet esetén is. A módszer ezúttal is az, hogy a kombinált K + ABCDEFGH hangforrások által kialakuló impulzusválaszokat alkalmaztam a különálló K, valamint felváltva megjelenő ABCDEFGH (vagyis összefoglalóan S) hangforrásokra.

Az önálló 'S' hangforrásmintára alkalmazva a 64. ábra által bemutatott eredményt kaphatjuk. Ezek effektív értékeit a 65. ábra, azokat pedig logaritmus skálán a 66. ábra mutatja.

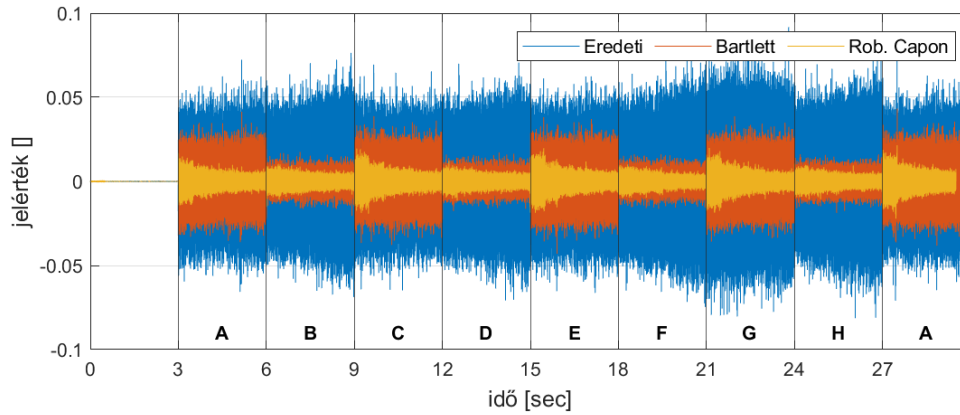
Ezekon megfigyelhető, hogy az 'eredeti', vagyis az M01-es mikrofon közvetlen kimenetén megjelenő, egyszerre mindig csak egyetlen forrásból származó jel teljesítménye bár bizonyos mértékben változik annak függvényében, hogy a próbahangforrás éppen milyen távol van az adott mikrofontól, de ezt leszámítva nem ingadozik. Ehhez képest a Bartlett nyalábformálás eredményeképp adódó kimeneten egyértelműen megfigyelhető, hogy a melléknyalábokban elhelyezkedő A, C, E, G hangforrások jelei 10-15 dB-el nagyobb teljesítménnyel jutnak a kimenetre, mint a B, D, F, H források jelei. A Robust Capon által biztosított kimeneten viszont egyértelműen látható, hogy az egyes hangforrások közötti váltást követően bár megugrik a kimeneten mérhető zajteljesítmény, annak mértéke határozottan csökkenést mutat idő függvényében az algoritmus adaptív működésének köszönhetően. A korábban is említett, melléknyalábokba eső források jelei itt is látványosan nagyobb teljesítménnyel jutnak a kimenetre a források váltását követően közvetlenül, azonban ezután közel 20 dB -es csökkenések is megfigyelhetők. A melléknyalábban, illetve azon kívül lévő források kimeneten megjelenő jelteljesítményei között a következő hangforrás megszólalásáig legfeljebb csupán néhány dB különbség van, a zajforrások mikrofonokra, illetve kimenetre jutó teljesítményei között pedig 50 dB elnyomás is megfigyelhető

Robust Capon alkalmazása során az adott paraméterek mellett, míg ez az érték Bartlett estén 15-30 dB attól függően, hogy hol helyezkedik el a forrás.

Bartlett, illetve Robust Capon nyálábformálás alkalmazása különálló 'S' hangmintán

átl. időállandó = 2.00 sec; bizonytalanság= 0.01 m;

DFT mintaszám= 480 db;



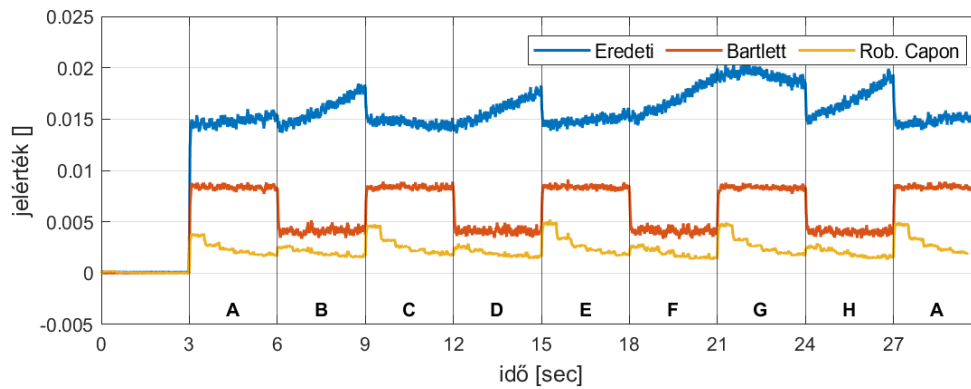
64. ábra Nyálábformálás eredménye különálló 'S' jelforrás hatására

Bartlett, illetve Robust Capon nyálábformálás alkalmazása különálló 'S' hangmintán

Effektív értékek exponenciális átlagolással

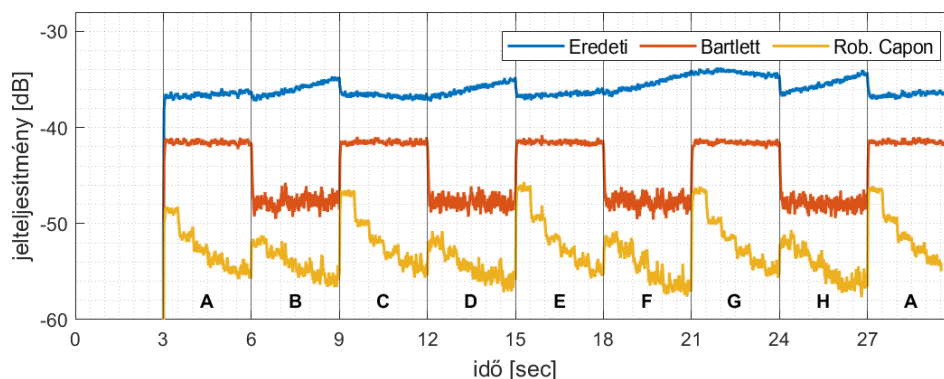
átl. időállandó = 2.00 sec; bizonytalanság= 0.01 m;

DFT mintaszám= 480 db;



65. ábra 'S' próbajel által különféle nyálábformálás során létrehozott kimenetek effektív értékei idő függvényében

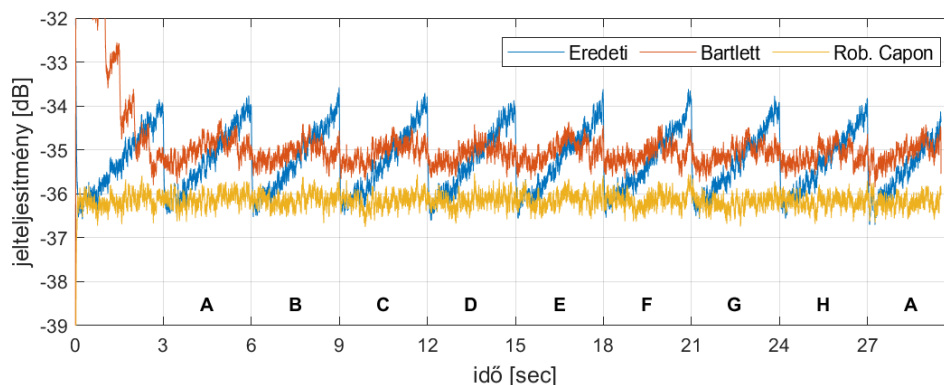
Bartlett, illetve Robust Capon nyalábformálás alkalmazása különálló 'S' hangmintán
Effektív értékek exponenciális átlagolással
 átl. időállandó = 2.00 sec; bizonytalanság= 0.01 m;
 DFT mintaszám= 480 db;



66. ábra 'S' próbajel által különféle nyalábformálás során létrehozott kimenetek effektív értékei idő függvényében logaritmikus skálán

Dinamikus esetben is megvizsgáltam a fókuszirányban lévő önálló próbahangforrás kimeneten megjelenő jelét a folyamatosan pozíciót váltó zavaró források hatására alakuló karakterisztika mellett. Ennek eredményét mutatja a 67. ábra. Az figyelhető meg, hogy szintén nem tapasztalható néhány dB-t meghaladó ingadozás az alkalmazott impulzusválaszok folyamatos változása mellett sem.

Bartlett, illetve Robust Capon nyalábformálás alkalmazása különálló 'K' hangmintán
Effektív értékek exponenciális átlagolással
 átl. időállandó = 2.00 sec; bizonytalanság= 0.01 m;
 DFT mintaszám= 480 db;



67. ábra 'K' próbajel által különféle nyalábformálás során létrehozott kimenetek effektív értékei idő függvényében logaritmikus skálán

Annak érdekében viszont, hogy megállapítsam, hogy Robust Capon alkalmazásával a kimeneten megjelenő jel valóban a fókuszálási irányból származó jel, ennek a két függvénynek vettem a keresztkorrelációját, ennek eredményét mutatja a 68. ábra. Az itt 0 sec körül látható, zajszint fölötti 30 dB-es kiemelés egyértelműen jelzi, hogy a kimeneten megjelenő jel nem csupán közel azonos teljesítményű a bemeneten mérhető

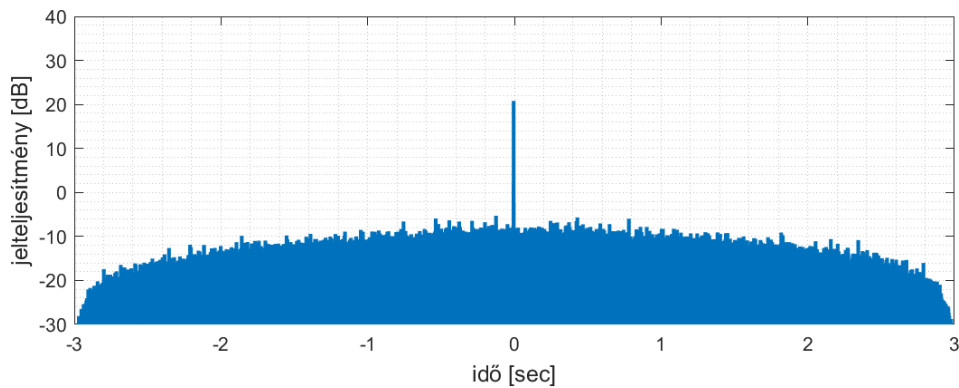
jellel, hanem azok egymáshoz is rendkívül hasonlítanak. Így tehát valóban közel azt a jelformát lehet a kimeneten mérni, amit a megfigyelt forrás produkál.

Keresztkorreláció az eredeti 'K' hangminta, és a Roust Capon kimenet között

átl. időállandó = 2.00 sec; bizonytalanság= 0.01 m;

DFT mintaszám= 480 db;

t= 30.0 sec;



68. ábra Keresztkorreláció az eredeti 'K' hangminta és a mikrofontömb kimenete között

11 Algoritmus tesztelése valós mikrofontömbön

Az eddigiekben részletezett algoritmus alkalmazásával végzett szimulációk során biztató eredményeket tapasztaltunk, így sor kerülhetett annak tesztelésére fizikailag realizált hangforrás, és mikrofontömb segítségével. Ennek során a Budapesti Műszaki és Gazdaságtudományi Egyetem, Hálózati Rendszerek és Szolgáltatások Tanszék Akusztika laborja bocsájtott rendelkezésemre műszereket, illetve eszközöket a mérések elvégzéséhez.

Ennek célja tehát végsősoron az volt, hogy megvizsgáljuk, hogy az algoritmus valóban alkalmazható-e mikrofonok, illetve hozzájuk képest valamilyen módon elhelyezett hangforrások esetében, illetve ennek milyen limitációi vannak.

11.1 Mérési helyszín

Az ilyen jellegű akusztikai vizsgálatok során a legnagyobb problémák jellemzően a hangforrások jeleinek az akusztikai környezetből származó többutas terjedéseiből adódnak. Ezeknek köszönhetően az egyes mikrofonok számára az adott jelkomponensek amplitúdó- illetve fázistorzulása lép fel. Ennek a hatásnak a minimalizálása érdekében a méréseket a tanszék Akusztika laborjának süketszobájában végeztem. A mérőterem

úgynevezett *félszabad hangterű akusztikai mérőhelyiség*, ami egy olyan téglatest alakú termet jelent, melynek az oldalfalai, illetve mennyezete is az azokra beérkező akusztikus jelek jelentős elnyelődését biztosító anyaggal van borítva. Ez azt jelenti, hogy az azokról érkező visszhang minimális, így az oldalfalak, illetve a plafon szempontjából nem kell aggódnunk a többutas terjedés okozta zavarokkal. A padló esetén viszont ennek megvalósítása igen nehézkes, ezért az a falakkal ellentétben úgy lett megalkotva, hogy azt ideális visszaverőnek lehessen tekinteni, így annak bár van hatása a mérési eredményre, de az ismert, és így az utófeldolgozás során kezelhető. Mindezekon kívül a mérőhelyiség tulajdonképpen szoba-a-szobában felépítésű, tehát maga a mérőterem egy olyan szoba, ami megfelelő kialakítású rezgéscsillapítókön keresztül van alátámasztva az épületen belül, így a kívülről érkező zavarok csak jelentősen elnyomva képesek oda bejutni, természetesen nyílászáró is ennek megfelelően lett kialakítva.

11.2 Mérési elrendezés

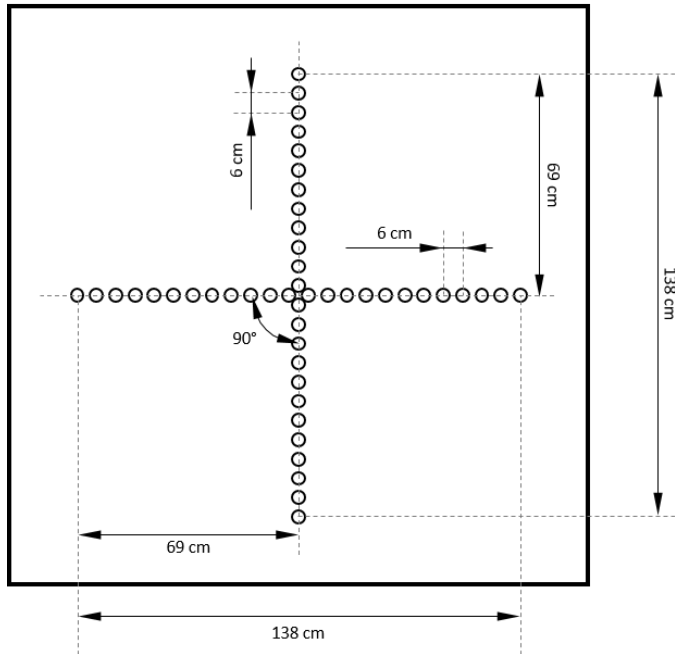
11.2.1 Mikrofontömb

A méréseket a tanszéken rendelkezésre álló mikrofontömb, az ahhoz tartozó feldolgozórendszer, illetve a jeleket biztosító hangforrások segítségével végeztem.

A mikrofontömb 48 kondenzátormikrofonól épül fel, melyek egymáshoz, illetve a mérőhelyiséghez viszonyított térbeli elrendeződését egy faelemekből kialakított szerkezet biztosítja. Ez tulajdonképpen egy sík lap, melyben előre kialakított furatokba helyezhetőek az egyes mikrofonok. Ilyen módon a mikrofonoknak kizárólag síkbeli elrendezését lehet elérni, viszont azok pozíciói jól meghatározhatók. Minden mikrofon rendelkezik egy rajta feltüntetett sorszámval 1-től 48-ig. A sorszámok megtalálhatóak a mikrofonokat tartó szerkezet foglalatai mellett is, így azok elhelyezkedése mérésenként megegyező lehet.

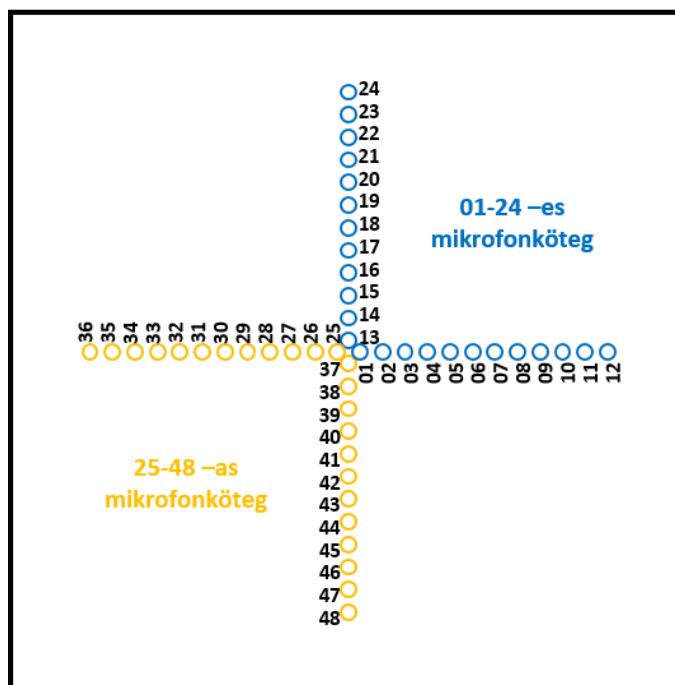
Az alakzat, amiben a mikrofonokat alkalmaztuk, az a már a szimulációs környezetből is ismert, két egymásra merőleges, ekvidisztáns lineáris mikrofonosor, melyek egymás felezőpontjaiban találkoznak. A kereszt elrendezésű geometriai pozícióját a 69. ábra, az egyes mikrofonok elhelyezkedéseit pedig a 70. ábra mutatja.

Mikrofontömb foglalata az abban
KERESZT konstellációban elhelyezett mikrofonok relatív
pozícióival



69. ábra Mikrofontömb foglalata a mikrofonok pontos pozíciójával

Mikrofontömb foglalata az abban
KERESZT konstellációban elhelyezett mikrofonok
sorszámaival



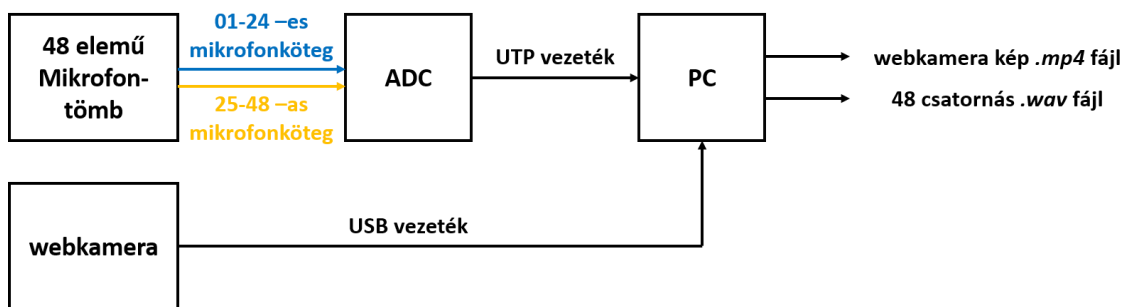
70. ábra Mikrofontömb foglalata a mikrofonok sorszámaival

A 48 mikrofon jelei egy 48 csatornás ADC-be érkeznek be, mely csatornánként 48 kHz mintavételi sebességet képes biztosítani. A mikrofonok 1-től 24-ig, illetve 25-től 48-ig két különálló 24-es kötegben, „ostorban” vannak csoportosítva, ezek számára rendelkezésre áll egy-egy 24 csatornás bemeneten, amin keresztül kapcsolódnak az ADC-hez, onnan pedig a jelek UTP kábelen keresztül jutnak el a feldolgozóegységbe annak hálózati kártyáján keresztül.

A feldolgozóegység tulajdonképpen egy megfelelő hardveres tulajdonságokkal rendelkező személyi számítógép, melyen megtalálható a tanszék által fejlesztett program a mikrofontömb jeleinek feldolgozásához. Itt lehetőség van az egyes mikrofonokból érkező jeleket különféle nyálábformáló algoritmusokkal közvetlenül feldolgozni, és annak eredményét közvetlenül lejátszani, viszont lehetőség van a mikrofonokból érkező nyers jelek közvetlen tárolására is egy 48 csatornás *.wav* kiterjesztésű fájlban.

Ezen kívül a mikrofontömb mellett elhelyezésre került egy webkamera is, melynek segítségével összevethetőek voltak az egyes mérési esetekben a hangforrások pozíciói, illetve, illetve az algoritmus alkalmazásával kapott eredmények.

Mérési elrendezés sematikus vázlata



71. ábra Mérési elrendezés sematikus vázlata

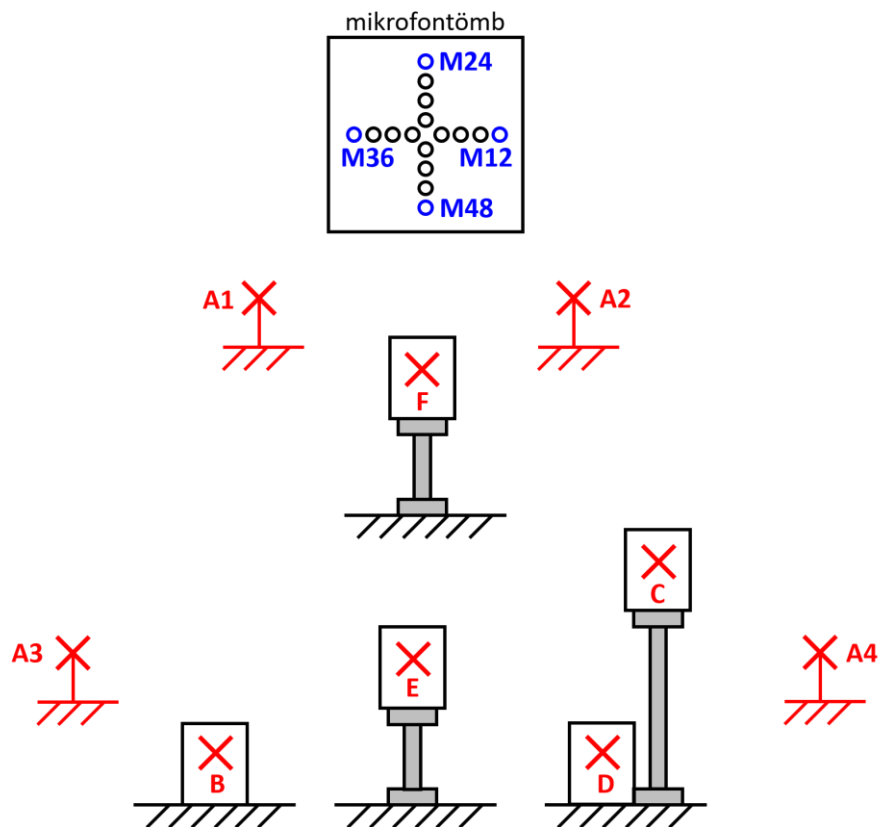
A későbbi feldolgozás során feltűnt, hogy a mentett nyers fájlban az 25-48 -as köteg jelei nem tisztázott oknál fogva közel 11 mintavételyit, tehát 230 μ s-ot késnek az 1-24-es kötegehez képest, ami a karakterisztikák torzításához vezetett. Ezt a hibát a felfedezése után könnyen kompenzálni lehetett az 1-24-es mikrofonköteg jeleinek feldolgozás során ugyanilyen mértékű késleltetésével.

11.2.2 hangforrások

A mérés során a mikrofontömbbel szemben többféle pozícióban helyeztem el hangforrásokat, illetve azokat megadott pályákon mozgattam ezek között. A mérési elrendezés sematikus vázlatát a 72. ábra mutatja.

Ezen látható a mikrofontömb, illetve azon feltüntetve a 12-es, 24-es, 36-os, illetve 48-as mikrofonok rendre **M12**, **M24**, **M36**, illetve **M48** jelöléssel. Ezekkel szembe fordítva helyezkedtek el az egyes hangforrások, melyeknek pozícióit a **B**, **C**, **D**, **E**, **F** pontok jelölik. Ezeknek mérésekor igyekeztem a hangsugárzók homloklületének geometriai középpontját kiválasztani, ami sajnos abból adódóan, hogy a hangsugárzók két egymástól független membránnal rendelkeznek a mély, illetve magas hangok lesugárzásának céljából, mindenképpen csak egy közelítő eredményt jelent. Ezekon kívül kijelöltem a padlón az **A1**, **A2**, **A3** és **A4** jelű pontokat.

Mérési elrendezés sematikus vázlata



72. ábra Mérési elrendezés sematikus vázlata

Ezeket a pontokat a mérés, illetve feldolgozás során egy olyan koordinátarendszerben helyeztem el, aminek az origója a mikrofonosorok metszéspontja, az y tengelye az origóból M12 felé, az z tengelye az origóból M24 fele, az x tengelye pedig a jobbkézszabálynak megfelelő módon a mikrofontömb síkjára merőlegesen, a hangforrások felé mutat.

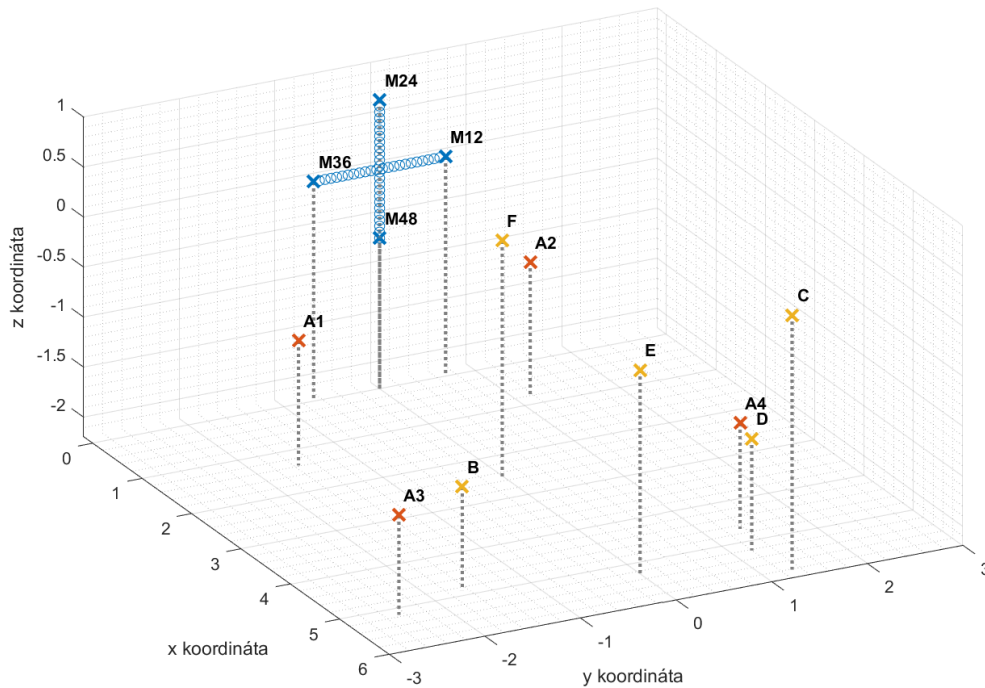


73. ábra Kamerafelvétel a mérési elrendezésről

Az algoritmus alkalmazása során a vizsgálni kívánt hangforrás mikrofontömbhöz viszonyított pozíciójának ismerete rendkívül fontos, hiszen a Capon algoritmus bemeneti értéke, az adott frekvenciához tartozó jelterjedésvektorok ebből származtathatók. Emellett a talajreflexióból adódó esetleges hatások vizsgálata érdekében a talaj elhelyezkedésére is szükség, amit az azon kijelölt 4 pont segítségével tudtam meghatározni.

Annak érdekében, hogy a hangforrások mérés során tapasztalható pontos elhelyezkedését megállapítsam, megmértem a távolságokat a jelzett mikrofonok, és a további pontok, tehát az ábrán minden késsel, és pirossal jelölt pontpár között. Ezt követően a kereszt elrendezésű mikrofontömb geometriáját felhasználva, azokat elhelyeztem a korábban leírt térbeli derékszögű koordinátarendszerben. Azokat ezt követően térben ábrázolva a 74. ábra által bemutatott elrendezést kaptam.

Süketszoba mérési elrendezés 2022.10.13.



74. ábra Süketszoba mérési elrendezés

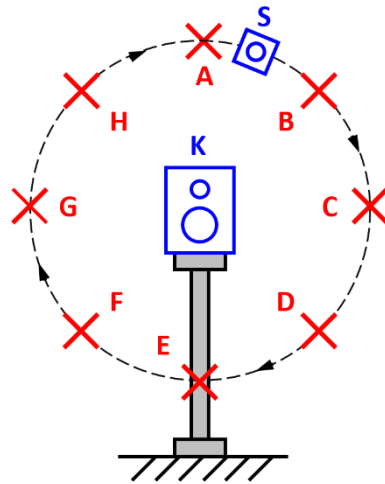
Fontos megjegyezni, hogy a mérés során a mikrofontömb síkja az alátámasztásból adódóan nem volt merőleges a talajra, így a talajon kijelölt pontok a mikrofontömb-középpontú koordináta-rendszerben egy olyan síkot határoznak meg, ami egyik tengelyre sem merőleges, ami tehát nem mérési hiba.

11.3 Algoritmus vizsgálata valós dinamikus környezetben

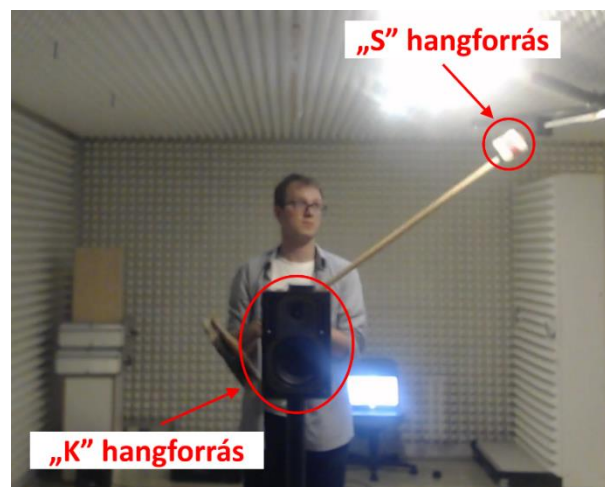
A fizikailag realizált mikrofontömbön alkalmazott adaptív nyálábformáló algoritmus tesztelése érdekében összehasonlítottam annak viselkedését a szimulációk során tapasztaltakkal. Ennek érdekében a szimulációkkal megegyező felépítésű mikrofontömb számára olyan akusztikus környezetet hoztam létre, mely a szimulációk során is alkalmazott *dinamikus* hangmintának megfelelő jelet eredményez az egyes mikrofonok bemenetein.

Ennek megvalósítása érdekében egyik mérési elrendezés során egy későbbiekben 'K'-val jelzett hangforrást helyeztem el a már korábban jelzett mérési pontok egyikében, és akörül függőleges, nagyjából 0.8 m sugarú, koordináta-rendszer origója, tehát

mikrofontömb közepe fele mutató normálisú síkra illeszkedő körpálya mentén 3 másodpercenként, és 45°-onként léptettem körbe egy másik, 'S' jelű hangforrást miközben ezek előzetesen generált fehérzajt bocsájtottak ki. Annak érdekében, hogy az egyes léptetések valóban 3 másodpercenként kövessék egymást, az általuk lejátszott hangfájlbba rövid, 2 kHz-es jelzéseket illesztettem, ami alapján tudtam, mikor szükséges a hangforrás elmozdítanom.



75. ábra Mozgó forrásokat tartalmazó akusztikai környezet megvalósításának sematikus ábrája



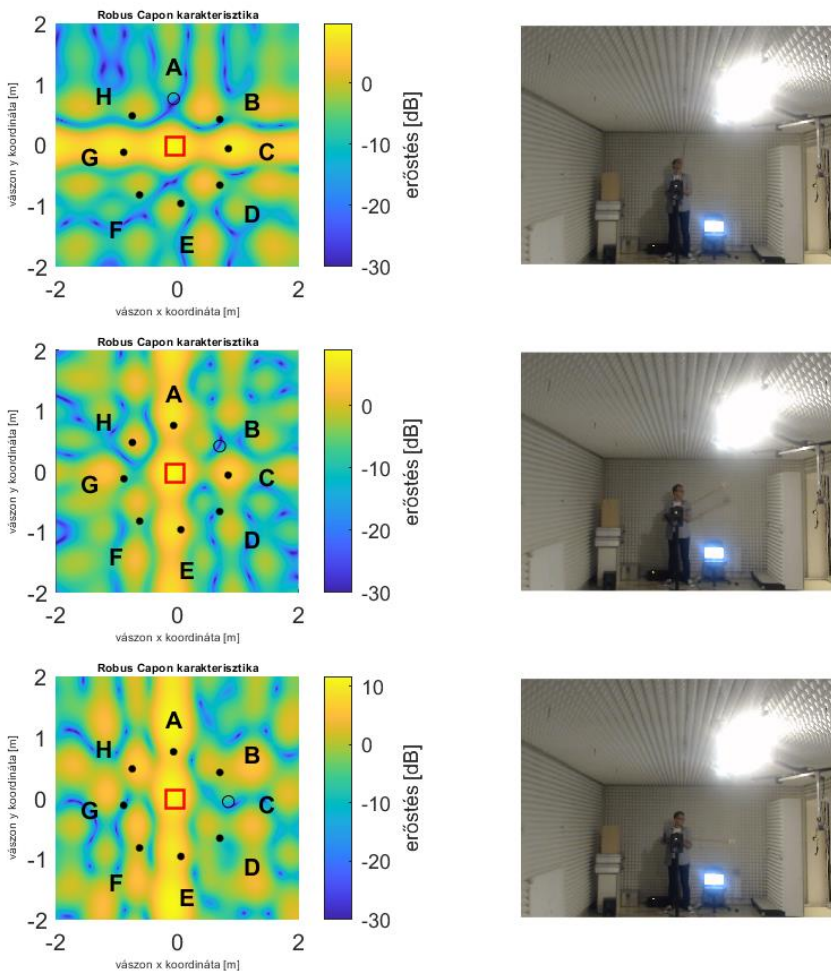
76. ábra Kamerakép a mozgatott hangforrásról

A mikrofontömb, illetve a rajta található webkamera segítségével kép-, illetve hangfelvételt készítettem az ilyen módon mozgatott hangforrásról, melyet később a Robust Capon algoritmus, megfelelően megválasztott paraméterek, illetve a korábban leírt forráspozíciók ismeretében feldolgoztam.

Egy ilyen feldolgozási folyamatot mutat az alábbi ábra, melyen egy olyan hangminta került feldolgozásra, amin a folyamatosan léptetett, aktívan sugárzó 'S' hangforrás mellett a fókuszirányban lévő 'K' hangforrás végig néma volt. Az egyes sorokban egyazon

időpillanathoz tartozó ábrák láthatók. Jobb oldalon a webkamera éppen aktuális felvétele, amin megfigyelhető a kezemben tartott hangforrás pozíciója, a baloldalon pedig az ehhez a pillanathoz tartozó, Robust Capon algoritmus eredményeképpen adódó, kvázi-állandósult iránykarakterisztika, melyen vörös négyzet jelöli a fókuszálási irányt, illetve fekete karika a hangforrás Bartlett Iránymérésből származó aktuális pozícióját, míg a további fekete pontok az 'S' hangforrás által érintett további pontokat jelölik.

Kialakuló iránykarakteristikák, és videófelvétel az algoritmus tesztelése során
 átl. időállandó = 0.10 sec; bizonytalanság= 0.50 m;
 vágási frekvencia= 50 Hz; vászontávolság= 5.27 m;

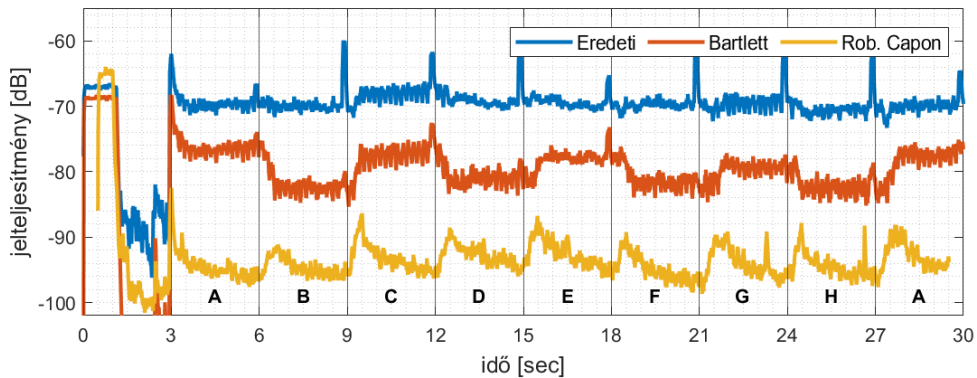


77. ábra Kialakuló iránykarakteristikák összehasonlítása videófelvételekkel

Jól megfigyelhető, ahogy az aktuális mérhető forrásirányban a karakterisztikának a szimulációk esetében is tapasztalt módon nullhelye van, míg a fókuszirányban mérhető erősítés a karakterisztikán tapasztalható maximális.

Ezen mérési összeállítás, és alkalmazott nyálábformálás eredményeképpen a kimeneten mérhető jel effektív értékének az alakulását a 78. ábra mutatja.

Bartlett, illetve Robust Capon nyálábformálás alkalmazása különálló 'S' hangmintán
Effektív értékek exponenciális átlagolással
 átl. időállandó = 2.00 sec; bizonytalanság= 0.01 m;
 DFT mintaszám= 480 db;



78. ábra Bartlett és Robust Capon összehasonlítása különálló 'S' mintán

Ezen jól megfigyelhetők a szimulációk során is tapasztalt jelenségek. Az eredeti zajforrás jeléhez képest 5-10 dB elnyomást biztosít a Bartlett algoritmus, ennek mértékét szintén nagyban befolyásolja a zajforrás pozíciója, tehát a melléknyalábba eső A, C, E, G forrásirányok esetén közel 5-10 dB-el kisebb elnyomásértékek tapasztalhatók, mint a többi forrásirány esetén.

A Robust Capon algoritmus ezzel szemben kvázi-állandósult esetben közel 25 dB elnyomást eredményez a bemeneti jelhez képest. A kimeneten mérhető zaj effektív értéke a hangforrás pozíciójának változtatását követően megugrik, majd határozottan lecsökken, bizonyítva ezzel az algoritmus adaptív viselkedését fizikai környezetben realizált mikrofontömbön is.

A szimulált esetben kapható eredményhez képest eltérést jelentenek a hangforráspozíciók váltásakor megugró bemeneti zajteljesítmények. Ezek magyarázata a hangmintába ágyazott, forráspozíció váltásra figyelmeztető hangjelzések adott pillanatokban történő megszólalása. Ezen kívül feltűnhet még, hogy a Robust Capon kimeneti teljesítmény görbéje a forráspozíció váltás pillanatában nem mutat olyan hirtelen ugrást, mint a szimulációk során. Ennek magyarázata, hogy ebben a valós környezetben a hangforrás tehetetlenségéből adódóan bizonyos időt igényelt annak következő forráspozícióba juttatása, ezzel magyarázható a kimeneti zajteljesítmény kevésbé meredek emelkedése.

11.4 Szélessávú működés tesztelése zenei hangfájlokra

Az algoritmus egyik felhasználási módja lehet a számunkra releváns irányokban elhelyezkedő hangforrások lehallgatása, az esetleges környező zavarforrások kiszűrése mellett. Ennek tesztelése érdekében összeállításra került egy olyan mérési elrendezés,

melyben a süketszoba B pontjában zenei audiofájlt játszó hangforrást adott pillanatban a D jelű mérési pontba mozgattam, miközben a C jelű hangforrás szintén folyamatosan sugárzott egy teljesen független másik zenei audiofájlt.

C jelű forrásra való fókuszálás mellett az volt a célunk, hogy az az általa sugárzott hangjelet minél tisztábban, a zavarforrás minél kisebb zavaró hatása mellett legyen lehetőségünk hallgatni. Az egyes mikrofonok által rögzített hangfelvételek esetében erre lehetőség nincs, C jelű forrás által sugárzott hangok élvezhetőségét, és értelmezhetőségét rendkívüli mértékben zavarta a másik forrás által sugárzott zene. Ez megfigyelhető mellékelt *raw.wav* hangfelvételen, ami az M01-es mikrofon által rögzített hangfelvételt tartalmazza. Ezen ahogy az a 2kHz-es hangjelzés alapján is megállapítható, a hangforrás elmozdítása nagyjából a $t=10$ sec pillanatban kezdődött, és a $t=15$ sec pillanatban fejeződött be, ahogy az a hangforrás mozgatása során bekövetkező hangerőcsökkenés alapján is érzékelhető.

Bár a Bartlett algoritmus alkalmazásának köszönhetően a C forrás értelmezhetősége jelentősen javult a zavarforrás elnyomása által a mérés első szakaszában, ez annak mozgatását követően megváltozott, ugyanis D jelű pozícióba helyezett zavarforrás éppen a C irányú Bartlett fókuszálás melléknyalábjába esett, ahonnan annak jelei már ismét zavaró mértékben voltak képesek a kimenetre jutni. A Bartlett alkalmazásával szerzett kimenetet a *bart.wav* hangfájl rögzíti. Ezen nagyon feltűnő, ahogy $t=15$ sec pillanatban hirtelen tisztán, és zavaró mértékben jelenik meg az addig csak alacsonyabb frekvenciák alapján sejthető zavarforrás hangja.

Robust Capon algoritmus alkalmazása esetén viszont, a kedvező paraméterek, illetve forrásirány megtalálása esetén kifejezetten jó minőségben voltunk képesek C forrás jelét hallgatni. A zavarforrás jelei bizonyos mértékben megjelentek a kimeneten, ennek mértéke hirtelen megugrott annak mozgatása során, azonban ennek hatása idővel egészen gyorsan lecsökkent, a zavarforrás ismét kiszűrésre került. A *RobCap.wav* nevű hangfájl segítségével meghallgatható a Robust Capon algoritmus által eredményezett kimenet, amin bár végig erősebb mértékben jelen van a zavarforrás hangja, mint a Bartlett alkalmazásával, főleg a mélyebb hangok esetében, de a $t=14$ sec pillanatnál megfigyelhető megerősödő zavarjel gyorsan elhalkul, és visszaál a kezdeti mértékre, így határozottan jobb eredményeket produkálva, mint amit a Bartlett esetében tapasztalhattunk.

A Robust Capon algoritmus hatására nem volt annyira kiváló hangminőség elérhető, mint a Bartlett esetében, azonban a zavarjel kioltásának mértéke egészen rendkívüli hozzá képest. Egyedül az egészen mély hangok voltak képesek a kimeneten megjelenni. Ezen problémák természetesen a paraméterek további hangolása által nyilvánvalóan tovább javíthatók.

11.5 További mérési tapasztalatok

Az elvégzett mérések közül leglátványosabb eredményt a fent tárgyalt, néma 'K' hangforrást körüljáró 'S' hangforrás által generált hangminta eredményezte. A többi esetben az a tapasztalat, hogy a fókuszirányba eső hangforrás zajforrásokat jócskán meghaladó teljesítményének köszönhetően azokra a Robust Capon alkalmazása nem eredményezett kifejezetten említésre méltó eredményt. Ezen kívül a forrás pozíciójának rendkívül pontos ismerete szükséges az algoritmus alkalmazásához, amin viszont nagyban ront annak a ténye, hogy a méréseken alkalmazott 'K' célirányba helyezett hangforrás külön hangsugárzó felülettel rendelkezett a magas, illetve mély hangok hatékony lesugárzása céljából. Ez kedvezőtlen a Robust Capon algoritmus, hiszen így nem jól megválasztott paraméterek esetén ezek közül valamelyik kioltását eredményezheti az algoritmus, illetőleg ha a bizonytalanságot kellően nagyra választjuk ennek elkerüléséhez, úgy az algoritmus hatékonyságát rontjuk.

12 Összefoglalás

A diplomatervezés dolgozatom megírása során rengeteg olyan fogalmat, eszközt, és műszert volt szerencsém megismerni, illetve alkalmazni, melyekkel bár előzetes tanulmányaim során találkozhattam már, de a velük kapcsolatos ismeretek elmélyíteni úgy érzem, csak ezen munka végzése által tudtam.

A választott téma is rendkívül szimpatikus volt számomra, komoly motivációt jelentett, hogy azt érezhettem, hogy egy ilyen lenyűgöző eljárással, illetve műszerekkel dolgozhatok. Munkám során megismerhettem a mikrofontömböket leíró matematikai modelleket, elmélyíthettem tudásom a Matlab programkörnyezet alkalmazásában, illetve a tanszéki süketszoba, illetve ahhoz tartozó műszerek kezelésében.

Az algoritmus finomítása, illetve kimeneti eredmények javítása érdekében további mérések összeállítására, módszerek bevezetésére, és feldolgozási paraméterek kipróbálására lenne szükség, viszont az eddigi eredmények alapján is kijelenthető, hogy a bemutatott algoritmus kifejezetten figyelemre méltó lehetőségeket rejt a szélessávú adaptív nyálábformálás szempontjából.

A későbbiekben célszerű lenne az alkalmazása szabadtéren is, a mikrofontömb kiterjedéséhez képest jelentősen nagyobb távolságra elhelyezkedő források által biztosított akusztikai környezet mellett is.

Ezen kívül mivel az algoritmus hatékonysága nagyban múlik a megfigyelni kívánt hangforráspozíció ismeretének pontosságán, ezért az algoritmust célszerű lenne

valamilyen iránymérő algoritmussal együttesen használni, mely automatikusan a kimért zajforráspozícióra korrigálja annak bemenetként megadott becsült értékét.

13 Köszönetnyilvánítás

Szeretném megköszönni Dr. Fiala Péternek a diplomatervezés témám kidolgozása, az említett algoritmusok értelmezése, az alkalmazott programkörnyezet kezelése, a laboratóriumi mérések kivitelezése, valamint ezen dokumentum tartalmi ellenőrzése során nyújtott hatalmas segítséget!

Szeretném megköszönni továbbá Minkó Dávidnak a dokumentum formázásával kapcsolatos tanácsokat, illetve javaslatokat!

14 Irodalomjegyzék

- [1] [Online]. Available: <https://www.oxfordreference.com/view/10.1093/oi/authority.20110803095621172>. [Hozzáférés dátuma: 4. 12. 2022.].
- [2] D. Gyöngyi, „hallasspecialista.hu,” 2014. október 28.. [Online]. Available: http://hallasspecialista.hu/2014/10/28/a_hallasjavito_eszkozok_tortenete_i/. [Hozzáférés dátuma: 29. 11. 2022.].
- [3] Peter. [Online]. Available: <https://fallsofsound.com.au/cow-horns-high-tech-history-hearing-aid-technology/>. [Hozzáférés dátuma: 28. 11. 2022.].
- [4] -, „<https://rarehistoricalphotos.com/aircraft-detection-radar-1917-1940/>,” 2022.11.28. [Online]. [Hozzáférés dátuma: 29. 11. 2022.].
- [5] „www.bell-labs.com,” Bell-Labs, [Online]. Available: <https://www.bell-labs.com/about/history/innovation-stories/confirming-big-bang/#gref>. [Hozzáférés dátuma: 29. 11. 2022.].
- [6] Y. AFANASYENKO, „www.rbth.com,” [Online]. Available: <https://www.rbth.com/science-and-tech/333557-duga-abandoned-air-radar-chernobyl>. [Hozzáférés dátuma: 29. 11. 2022.].
- [7] J. Capon, High-Resolution Frequency Wavenumber Spectrum Analysis, IEEE, 1969..
- [8] O. L. Frost, An algorithm for linearly constrained adaptive array processing, Palo Alto: Proceedings of the IEEE, 1972..
- [9] Y. Narita, „<https://www.frontiersin.org/articles/10.3389/fphy.2019.00008/full>,” 2019. [Online]. [Hozzáférés dátuma: 29. 11. 2022.].
- [10] J. Li, „On Robust Capon Beamforming and Diagonal Loading,” *IEEE TRANSACTIONS ON SIGNAL PROCESSING*, 2003..
- [11] J. O. S. III, „ccrma.stanford.edu,” [Online]. Available: https://ccrma.stanford.edu/~jos/sasp/Convolution_Theorem_DTFT.html. [Hozzáférés dátuma: 29. 11. 2021.].

[12] „rarehistoricalphotos.com,” [Online]. Available:
<https://rarehistoricalphotos.com/aircraft-detection-radar-1917-1940/>. [Hozzáférés
dátuma: 29. 11. 2022.].