

Akusztikus-hangforrás lokalizátor digitális jelfeldolgozással

Dobos Kende János, Braun Ferenc, Kackstädter Péter, Volk János, Molnár Zsolt

Dobos Kende János, hallgató [Óbudai Egyetem, Tavaszmező utca 15-17., 1084, Budapest, Magyarország] doboskende@gmail.com

Braun Ferenc, HUN-REN Energiatudományi Kutatóközpont, Konkoly-Thege Miklós út 29-33, 1121, Budapest, Magyarország, braun.ferenc@ek.hun-ren.hu

Kackstädter Péter, Szent György Hang- és Filmművészeti Technikum, Lenhossék u. 24-28, 1096, Budapest, Magyarország, kackstadter.peter@broadcast.hu

Volk János, HUN-REN Energiatudományi Kutatóközpont, Konkoly-Thege Miklós út 29-33, 1121, Budapest, Magyarország, volk.janos@ek.hun-ren.hu

Molnár Zsolt, Óbudai Egyetem, Bécsi út 96/B, 1034, Budapest, Magyarország, molnar.zsolt@kvk.uni-obuda.hu

Absztrakt: Cikkünk azt a témát járja körül, hogy hogyan lehet egy tetszőleges hangforrás irányát meghatározni, mindezt digitális jelfeldolgozással. Modellt alkottunk a jelenségről, szimulációkat végeztünk, majd létrehoztunk egy rendszert, amely valós időben képes a hangforrás irányát kiszámolni és kimutatni. Bemutatjuk egy blokkdiagramon keresztül, hogy hogyan áll össze ez a rendszer, milyen részegységekből áll és milyen szempontok alapján raktuk össze ezeket az egységeket. Szó esik még egy kiértékelésről is, amely egy mérés során készített hanganyagból számított eredményeket vet össze a hangforrás pozícióját mért eszköz eredményeivel. A jövőbeli fejlesztési lehetőségekről is ejtünk még pár mondatot

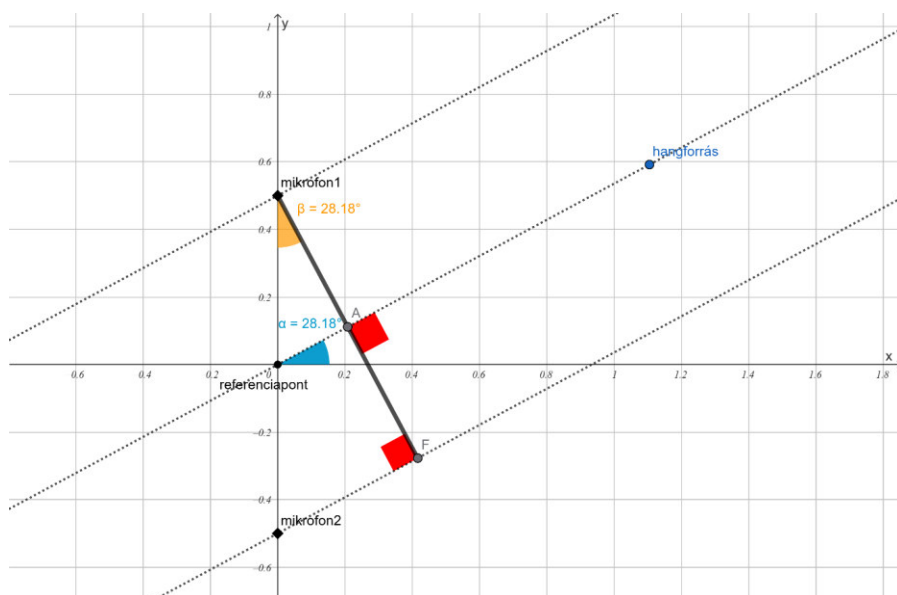
Kulcsszavak: digitális jelfeldolgozás, hangforrás lokalizáció, keresztkorreláció, fft, tdoa

1 Módszerek

1.1 Matematikai módszerek

1.1.1 Hang terjedése síkhullámként

Az ideális hangforrásunk hangjának hullámfrontja mindenképp gömbfelületet alkot, azonban a távolság növekedésével ennek a gömbnek a sugara egyre nagyobb lesz, egyre kevésbé érzékelhető a görbülete és síkfelületnek tekinthetjük. Elsősorban síkban vizsgáljuk a jelenséget.



36. ábra: síkhullám terjedési irányszögének visszafejtése

Ez az ábra, amelyet GeoGebrában készítettem, azt szemlélteti, hogy a síkfelületként kezelt hullám a mikrofonokba érkezési idő különbségéből, hogyan tudunk következtetni a hangforrás irányára, azaz az α szögre. A szaggatott vonalak a hullám terjedési irányát mutatják és mivel síkhullámról beszélünk, ezek a vonalak párhuzamosak. Ilyenkor a terjedési irányra merőleges hullámfront a fekete vastag szakasz lesz. Ez mutatja azon pontok halmazát a síkban, amelyek a síkhullám hullámfrontját képezik abban az időpillanatban, amikor a hang az 1.

mikrofonba érkezik. Ilyenkor létrejön egy útkülönbség, amely az $sd=d(F; \text{mikrofon2})$ szakaszként írható le.

Matematikailag adódik, hogy az α és β szögek merőleges szárú szögek, így azok egyenlők. Ilyenkor a keresett irány, mint főiránytól való szögeltérés, a következőképp számítható:

$$\alpha = \sin^{-1} \frac{sd}{d_{\text{mikrofon}}} \quad (2)$$

ahol d_{mikrofon} a két mikrofon közti távolság.

Mivel azonban az sd útkülönbséget nem tudjuk közvetlenül mérni, át kell számolni a mért td időkülönbséget a hangsebességgel való osztással, így:

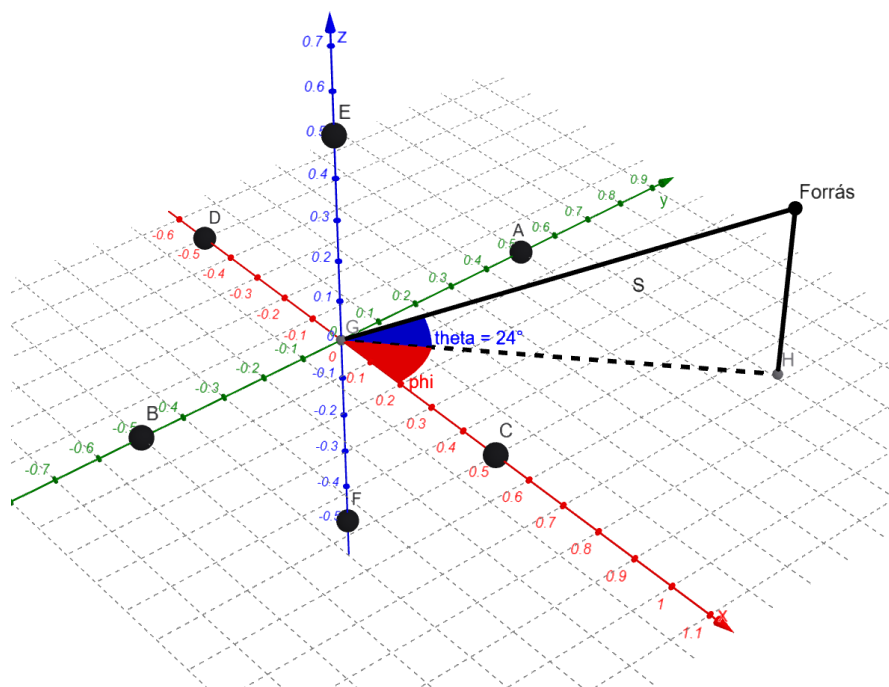
$$\alpha = \sin^{-1} \frac{\frac{td}{v_{\text{hang}}}}{d_{\text{mikrofon}}} \quad (2)$$

Ez az ún- TDOA (Time Difference Of Arrival) módszer [5], amelynek lényege, hogy a beérkezési időkülönbségekből következtetünk a hangforrás irányára, csak úgy, mint az emberi irányhallás esetén [2].

Ahhoz, hogy pontosabb irányt tudjunk számítani, érdemes további 2-2 mikrofont alkalmazni a tér további 2 dimenziójára.

1.1.2 Irányszögek számítása

A térbeli irányszögek számítási módszerét szemlélteti a következő ábra:



37. ábra: irányszögek szemléltetése

A fekete gömbök helyén a 3 pár mikrofon található, melyek közt időkülönbséget mérünk páronként. Ha az x-tengelyt tekintjük főiránynak, akkor a síkbeli- vagy vízszintes szögeltérést a phi (piros) szöggel ábrázolhatjuk, a magasságbeli- vagy függőleges szögeltérést phi-hez képest pedig a theta (kék) szög ábrázolja. A következőkben sokszor fogok ezen paraméterekre ily néven hivatkozni. A következő számítással kaphatjuk meg a hangforrás irányát a 3 tengelyre vonatkoztatott időkülönbségek ismeretében:

$$\phi = \tan^{-1} \left(\frac{v_{hang} \times \Delta t_y}{v_{hang} \times \Delta t_x} \right) = \tan^{-1} \left(\frac{\Delta t_y}{\Delta t_x} \right), \quad (3)$$

$$\theta = \tan^{-1} \left(\frac{v_{hang} \times \Delta t_z}{\sqrt{(v_{hang} \times \Delta t_x)^2 + (v_{hang} \times \Delta t_y)^2}} \right) = \tan^{-1} \left(\frac{\Delta t_z}{\sqrt{\Delta t_x^2 + \Delta t_y^2}} \right), \quad (4)$$

ahol Δt_x , Δt_y , Δt_z az x, y és z tengelyen mért időkülönbségek, v_{hang} pedig a hang terjedési sebessége, mely kiesik a képletből.

1.1.3 Időkülönbség meghatározás

A két mikrofon kimeneti jelei közt számított keresztkorrelációs függvény számításával meg tudjuk állapítani a mikrofonok közt fellépő időkülönbség mértékét.

A keresztkorrelációs függvény a következőképpen adható meg:

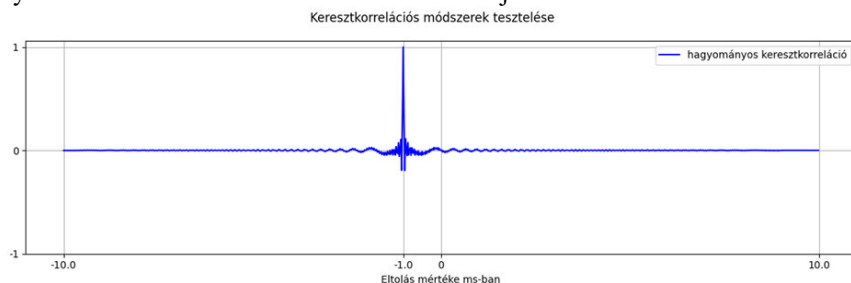
$$W_{xy}(\tau) = \lim_{T_v \rightarrow \infty} \frac{1}{2 \times T_v} \int_{-T_v}^{T_v} x(t - \tau) \times y(t) dt, \quad (5)$$

ahol T_v a vizsgált időtartomány ideje, $x(t)$ és $y(t)$ pedig a két mikrofon bemenete [1].

Ez a függvény -1 és 1 közti értékeket vehet fel, a határokat beleértve. Ha a függvény pozitív értéket vesz fel, akkor a két jel közti hasonlóság -azaz korreláció- is pozitív. Ha negatív a függvény által felvett érték, akkor negatív ez a hasonlóság, azaz az egyik jel polaritásának fordításakor lenne pozitív a korreláció. Mivel azonban, ha a két mikrofonba érkező hang, 0 s időkülönbség esetén azonos polaritású villamos jelet ad a kimenetén, elég a legnagyobb szélsőérték helyett a keresztkorrelációs időfüggvény globális maximumát megkeresni. Ahol ez az időfüggvény a legnagyobb értéket veszi fel, az lesz a két mikrofonhoz tartozó időkülönbség értéke:

$$\Delta t = \operatorname{argmax} (W_{xy}(\tau)). \quad (6)$$

Szimulált, digitális jelekkel is vizsgáltuk az esetet. Az egyik bemeneti, sweep jelünk frekvenciája 20 Hz-től lineárisan nő 20 kHz-ig, mindezt 10 ms időtartam alatt, 48kHz-es mintavétellel és 32 bites lebegőpontos bitmélységgel. Az egyik jel ez, a másik pedig ugyanez, csak -1 ms-mal eltolva időben. Az eredményeket egy Pythonban általam írt kóddal hoztam létre és rajzoltam ki:



38. ábra: sweep jel keresztkorrelációs függvénye

Jól kiadja a függvény a két jel közti időkülönbséget.

Nem lehet a mikrofon távolságnál nagyobb hullámhosszú jel esetén időkülönbséget mérni, ez állítás bizonyításaként 3 kHz-es szinuszos jelet is vizsgáltunk:



39. ábra: két 3 KHz-es jel keresztkorrelációja

Közelebbről:



40. ábra: két 3 KHz-es jel keresztkorrelációja kinagyítva

3 kHz hullámhossza :

$$\lambda = \frac{343 \frac{m}{s}}{3 \text{ kHz}} \approx 11 \text{ cm} . \quad (7)$$

Két egymástól 15 cm-re lévő mikrofon közt a legnagyobb lehetséges időkülönbség:

$$\tau_{max} = \frac{0.15 \text{ m}}{343 \frac{m}{s}} \approx 43 \text{ us} . \quad (8)$$

Ez a bizonyos τ_{max} fel is van tüntetve a vízszintes tengelyen. Periodikus a keresztkorrelációs függvény és ezen a $\pm\tau_{max}$ -os tartományon több helyen is 1-et

vesz fel a függvény. Egy lineáris fáziskarakterisztikájú aluláteresztő szűrő, aminek a törésponti frekvenciájának hullámhossza a két mikrofon távolsága, jól ki tudja küszöbölni a magas frekvenciák átlapolódásából keletkező hibás értékeket.

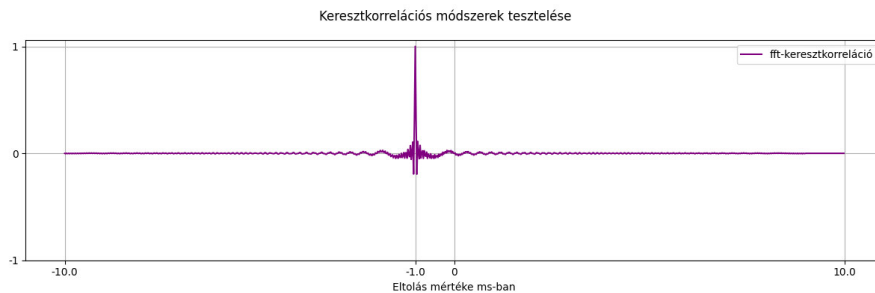
Ha valós időben akarjuk ezt megkapni, esetleg hosszabb jeleken, a számítási erőforrások használatának optimalizálásakor sokkal jobban járunk, ha a Wiener-Hincsin tételt alkalmazzuk [1]:

$$x(f) = \mathcal{F}\{x(t)\} \text{ és } y(f) = \mathcal{F}\{y(t)\} \quad (9)$$

esetén:

$$W_{xy}(\tau) = \mathcal{F}^{-1}\{x(f) \times y^*(f)\}. \quad (10)$$

Ennek tükrében, ha a két bemeneti jelen FFT-t végzünk, majd az első jel spektrumát összeszorozzuk a második jel spektrumának komplex konjugáltjával, akkor megkapjuk a két jel keresztspektrumát, azaz a keresztkorrelációs időfüggvény spektrumát. Ezen a keresztspektrumon IFFT-t végezve megkapjuk a két jelen végzett keresztkorrelációs időfüggvényt, ez a módszer az úgynevezett generalizált keresztkorreláció [6]. A korábban szimulált sweep jeleken elvégezve az FFT-alapú keresztkorrelációt a következő eredményt kapjuk:

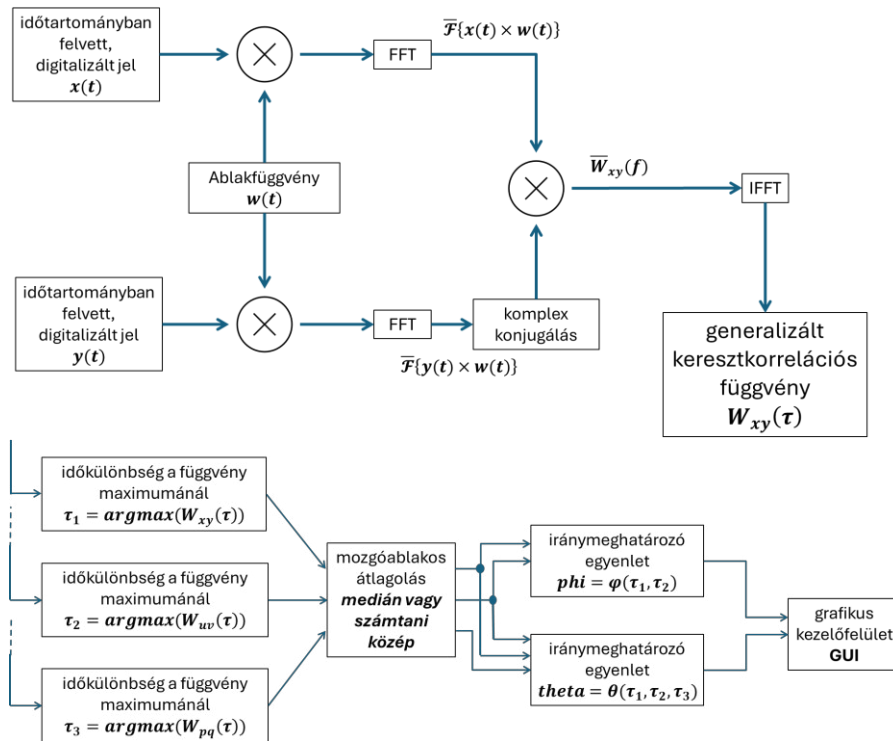


41. ábra: FFT-alapú keresztkorrelációs függvény

Az eredmény tökéletesen ugyanaz, a kulcs azonban a lefutásának idejében van. Az hagyományos keresztkorreláció futásideje **6,6 s**, míg az FFT-vel végzett verzió esetén **16 ms** volt csak.

1.1.4 Az algoritmus blokkdiagramja

A korábban leírtak alapján készítettem egy blokkdiagramot, mely bemutatja az iránymeghatározó algoritmus működését, kiértékelés módszertanát. Az egyetlen újdonság, hogy a mért időkülönbségeket átlagoljuk a mérés során, ezzel csökkentve a véletlen hiba hatását.



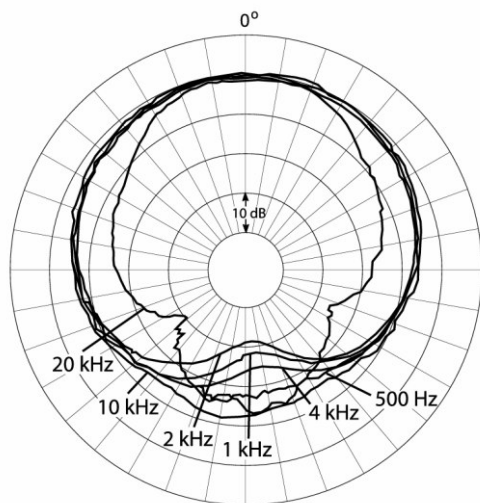
42. ábra: a mérőrendszer blokkdiagramja

1.2 Kísérleti módszerek

1.2.1 Mikrofonkarakterisztika

Minden irányból ugyanakkora érzékenységgel kell vennünk a hangot, ezért a legcélszerűbb megoldásnak gömbkarakterisztikájú mikrofon használata számít.

A kísérletek során használt DPA 4053 gömbkarakterisztikájú mikrofon polárdiagramja [4]:



43. ábra: DPA 4053 mikrofon irányérzékenységi karakterisztikája

1.2.2 mérés a Pannónia Filmstúdióban

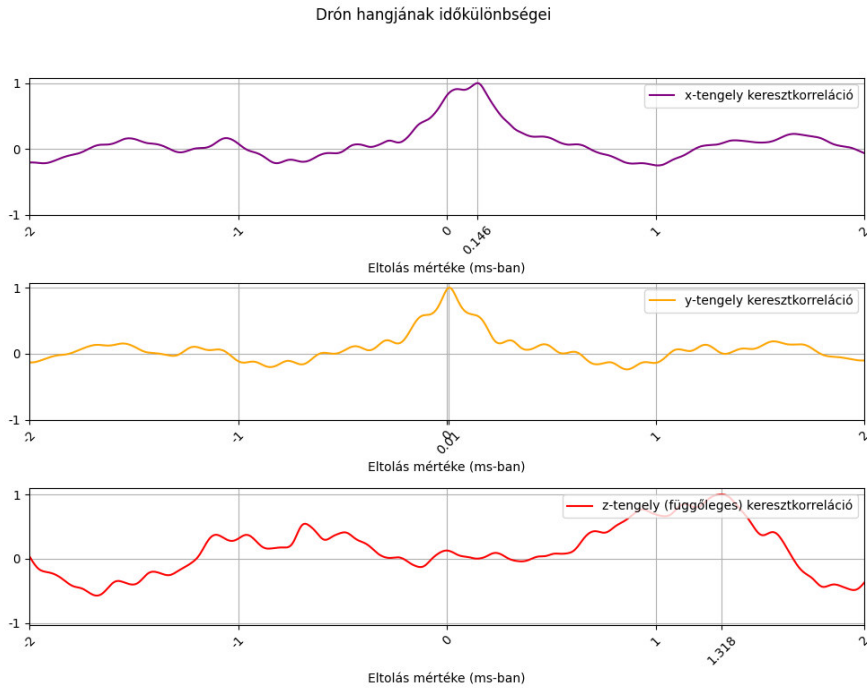
Továbbiakban a hangforrás szerepét egy pilóta nélküli légitáncos töltötte be. Ahhoz, hogy kellő pontossággal tudjuk ezeknek a hangját vizsgálni, szükségünk volt egy olyan kontrollált környezetre, ahol nem fúj a szél és nem volt zaj. Mivel kültéren nem sok ilyen hely van, beltéren kellett elvégezni ezt a mérést, ilyenkor a kellően nagy teremterfogat mellett azonban kulcsfontosságú minimalizálni a falról visszaverődő hangok intenzitását, így a Pannónia Filmstúdió nagy feljátszótermébe szerveztünk ezt.

A csoportmikrofonokon alkalmazott mikrofon típusa DPA 4053 gömbi iránykarakteristikájú, professzionális nyomásérzékeny kondenzátor mikrofon volt. Ezek egy speciális szélzaj ellen védő befogóba kerültek, hogy a propellerek által gerjesztett levegőmozgás ne torzítsa a hasznos hangot, erről a szemléltetés a lejjebb látható A hangfelvételek 192 kHz mintavételi frekvenciával és 24 bites felbontással készültek, miközben a vizsgálathoz különböző típusú pilóta nélküli légitáncosokkal többféle manővert hajtottunk végre.



44. ábra: a mérés során használt mikrofoncsoport

A három tengelyre számolt időkülönbség:



45. ábra: a három tengelyre számított keresztkorrelációs függvény

A függőleges tengelyen a mikrofontávolság 52 cm volt, a többi tengelyen 6 cm, emiatt kaptunk csak nagyságrenddel nagyobb értéket, ezt a különbséget korrigáltuk a számítás során. A theta képletében a nevezőben.

A főnti esetre a számított irányok, az *x-tengelyt* választva főirányul:

$$\phi = \tan^{-1} \left(\frac{\Delta t_y}{\Delta t_x} \right) = \tan^{-1} \left(\frac{10,417 \text{ us}}{145,833 \text{ us}} \right) = 4.1^\circ, \quad (11)$$

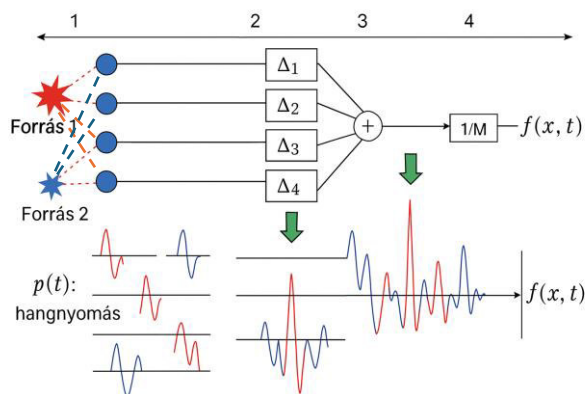
és

$$\theta = \tan^{-1} \left(\frac{\Delta t_z \times \frac{d_{x1,x2}}{d_{z1,z2}}}{\sqrt{(\Delta t_y)^2 + (\Delta t_x)^2}} \right) = \tan^{-1} \left(\frac{1317,708 \text{ us} \times \frac{6 \text{ cm}}{52 \text{ cm}}}{\sqrt{10,417 \text{ us}^2 + 145,833 \text{ us}^2}} \right) = 46.1^\circ. \quad (12)$$

Ezek voltak a pilóta nélküli légi jármű tényleges irányai a térben.

2. Fejlesztési lehetőségek

További fejlesztési lehetőségek között szerepel egy olyan program megvalósítása is, amely valós időben tudja kiszámítani ezeket az irányokat. Érdekes lenne még módosítani a számítási algoritmust, hogy zajos, reflexiókkal telt környezetben is jó közelítést kapjunk az időkülönbségekre. MEMS mikrofonok bevezetésével miniatürizálni lehetne a csoportmikrofonokat, ezáltal például egy mikrokontrollerrel beágyazott rendszerbe tudnánk foglalni az eszközt. A TDOA módszer mellett még a beamforming módszer [3] is bevált a hangforrás lokalizáció területén, ezzel is érdemes lehet dolgozni.



46. ábra: beamforming módszer

3. Összefoglalás

A kezdeti modellalkotás során felvetett módszerek és szempontok a valóságban is megállták a helyüket a rendszer tervezését és üzemeltetését illetően. A mérések kiértékelése során kirajzolt kimutatás kellően pontos eredményt tükröz ahhoz, hogy szemléltetni tudjam, működik a TDOA alapú hangforrás irány- és helymeghatározás módszertana és lehetséges hang alapján lokalizálni egy pilóta nélküli légitármű helyét. A program futása során kiirattam a parancssorba az iránymeghatározó algoritmus futási idejét, ennek nagyságrendje néhány ms volt. A keresztkorrelációs módszer kiválóan implementálható bármilyen mikrokontrolleres vagy más típusú beágyazott rendszerbe.

Hivatkozások

- [1] K. József, Szerző, *Jelfeldolgozás I-II.* [Performance]. Óbudai Egyetem, 2025.
- [2] „Wikipedia,” [Online]. Available: https://en.wikipedia.org/wiki/Sound_localization. [Hozzáférés dátuma: 04. 10. 2025.].
- [3] „www.polytec.com,” Polytec GMBH, [Online]. Available: <https://www.polytec.com/us/acoustics/technology/beamforming>. [Hozzáférés dátuma: 10. 2025.].
- [4] D. M. A/S, „DPA 4051/4052/4053 Compact Omnidirectional Microphones”.
- [5] G. S. R. Y. a. Z. Y. B. Xu, „High-Accuracy TDOA-Based Localization without Time Synchronization,” *IEEE Transactions on Parallel and Distributed Systems*, vol. 24, 2013.
- [6] C. K. a. G. Carter, „The generalized correlation method for estimation of time delay,” *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 1976.