

ÄÄNI-INTENSITEETIN ESTIMOINTI LAAJAKAISTAISESTI PAIN- JA ENERGIA-GRADIENTTIEN AVULLA SEKÄ SOVELTAMINEN ÄÄNEN TULOSUUNNAN ANALYSOINNISSA

Jukka Ahonen, Ville Pulkki

Signaalinkäsittelyn ja akustiikan laitos
Aalto-yliopisto Otakaari 5 A, 02015 Espoo
jukka.ahonen@acoustics.hut.fi

1 JOHDANTO

Useissa tilaäänen sovelluksissa hyödynnetään monimikrofonitekniikoita. Sovelluskohteita ovat mm. monikanavainen äänentoisto sekä suuntamikrofonitekniikat, joissa pyritään vahvistamaan yhtä (tai useampaa) tietystä suunnasta saapuvaa ääntä ja toisaalta vaimentamaan ääntä ei-toivotuista suunnista. Perinteisesti em. sovelluksissa mikrofonisignaaleita käytetään joko sellaisenaan tai niistä lasketaan lineaarikombinaatio. Kuitenkin viime aikoina erilaiset parametriset tilaäänen prosessointimenetelmät ovat yleistyneet. Yksi näistä on Directional Audio Coding (DirAC) [1]. Siinä tilaääni parametrisoidaan aika-taajuustasossa laskemalla äänen tulosuunta ja äänikentän diffuusisuus, joita hyödynnetään tilaäänentoistossa eri tavoin. Suunnan ja diffuusiisuuden laskenta perustuu ääni-intensiteettiin, jota estimoidaan äänipaineesta ja hiukkasnopeudesta äänikentän yhdessä pisteessä. Hiukkasnopeus voidaan approksimoida kahden painemikronin välisestä painegradientista. DirAC-prosessoinnissa tarvittavat paine- ja hiukkasnopeussignaalit tuotetaan useissa sovelluksissa neljän painemikrofonin muodostamalla mikrofonihilalla.

Painegradienttiin perustuva ääni-intensiteetin laskenta on kuitenkin vahvasti kaistarajoitettua. Suurilla taajuuksilla mikrofonien välinen painegradientti johtaa virheelliseen intensiteetin ja äänen tulosuunnan estimointiin. Ongelman ratkaisuksi tässä artikkelissa esitellään painemikrofonien välisen energiagradientin käyttö intensiteetin laskennassa, jolloin saavutetaan luotettava suuntaestimaatti myös suurilla taajuuksilla. Energiagradientin laskenta vaatii mikrofonien välisiä tasoeroja äänen tulosuunnasta riippuen. Nämä tasoerot voidaan saavuttaa käyttämällä esim. kardiodisuuntakuvioisia mikrofoneja tai sijoittamalla painemikrofonien väliin akustisesti varjostava esine. Artikkelissa käsitellään suunnan laskentaa neljän, jäykän sylinterin pinnalle sijaitetun painemikrofonin signaalista. Lisäksi artikkelissa evaluoidaan suuntaestimaatteja paine- ja energiagradienteista kaiuttomien mittausten avulla.

2 SUUNNAN ESTIMOINTI ENERGIA-ANALYYSISSA

Energia-analyysi perustuu ääni-intensiteettiin, joka kuvaa virtaavan äänienergian suuntaa ja magnitudia. Kuten edellä mainittiin, intensiteetti saadaan laskettu äänipaineesta ja hiukkasnopeudesta. Nämä voidaan mitata B-formaattimikrofonilla, jonka pallosignaali W vastaa äänipainetta ja dipolisignaalit X , Y ja Z approksimoivat hiukkasnopeutta vastaavalle karteesisen koordinaatiston akselille. B-formaattisignaalit harmonisoidaan STFT-muunnoksella (short-time Fourier transform), jonka aikaikkunan koko on usein 20 ms DirAC prosessoinnissa. Aktiivinen intensiteettivektori [2] voidaan laskea aika-taajuustasossa kaavalla

$$\mathbf{I}(t, f) = \frac{1}{\sqrt{2}Z_0} \text{Re}\{W^*(t, f)\mathbf{X}(t, f)\}, \quad (1)$$

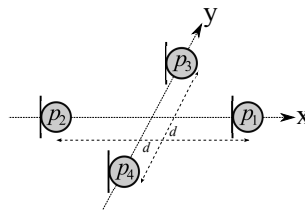
jossa t on aika, f on taajuus, $*$ tarkoittaa kompleksikonjugaattia, $\text{Re}\{\cdot\}$ on reaaliosa kompleksisityksestä, Z_0 on ilman karakteristinen impedanssi ja \mathbf{X} esittää dipolisignaalien STFT-speaktrejä vektorimuodossa, $\mathbf{X}(t, f) = [X(t, f) Y(t, f) Z(t, f)]^T$. Äänen tulosuunta saadaan yksinkertaisesti intensiteettivektorin vastakkaisesta suunnasta.

3 B-FORMAATTI MIKROFONIHILA

Verrattain kalliiden kaupallisten B-formaattimikrofonien sijaan tarvittavat pallo- ja dipolisignaali voidaan myös tuottaa lineaarikombinaatiolla lähelle toisiaan sijoitettujen pallomikrofonien signaaleista. Kuvassa 1 on esitetty neliöhila, jossa on neljä pallomikrofonia vaakatason B-formaattisignaalien tuottamiseen. Dipolisignaali x - ja y -akseleille muodostetaan laskemalla painegradientti vastakkaisten mikrofonien signaaleista

$$\begin{aligned} X(t, f) &= \sqrt{2} \cdot A(f) \cdot [P_1(t, f) - P_2(t, f)] \\ Y(t, f) &= \sqrt{2} \cdot A(f) \cdot [P_3(t, f) - P_4(t, f)], \end{aligned} \quad (2)$$

jossa $P(t, f)_n$ vastaa mikrofonisignaalin STFT-spektriä, kun $n = \{1, 2, 3, 4\}$. Normalisointikerroin $A(f)$ kompensoi dipolisignaalien taajuusriippuvaista vaimenemista. Pallosignaali W muodostetaan laskemalla keskiarvo kaikista neljästä mikrofonisignaalista.



Kuva 1: Neliömikrofonihila, jossa neljä pallomikrofonia on sijoitettu neliön kulmiin. Neliöhilalla voidaan luoda B-formaattisignaaleista pallosignaali W sekä dipolisignaali X ja Y .

Mikrofonien välisen painegradientin laskeminen (2) ei kuitenkaan tuota dipolisignaaleja suurilla taajuuksilla. Äänen puolikkaan aallonpituuden ollessa pienempi kuin mikrofonien välinen etäisyys esiintyy spatiaalista laskostumista. Näin ollen teoreettisen laskostumistaajuuden,

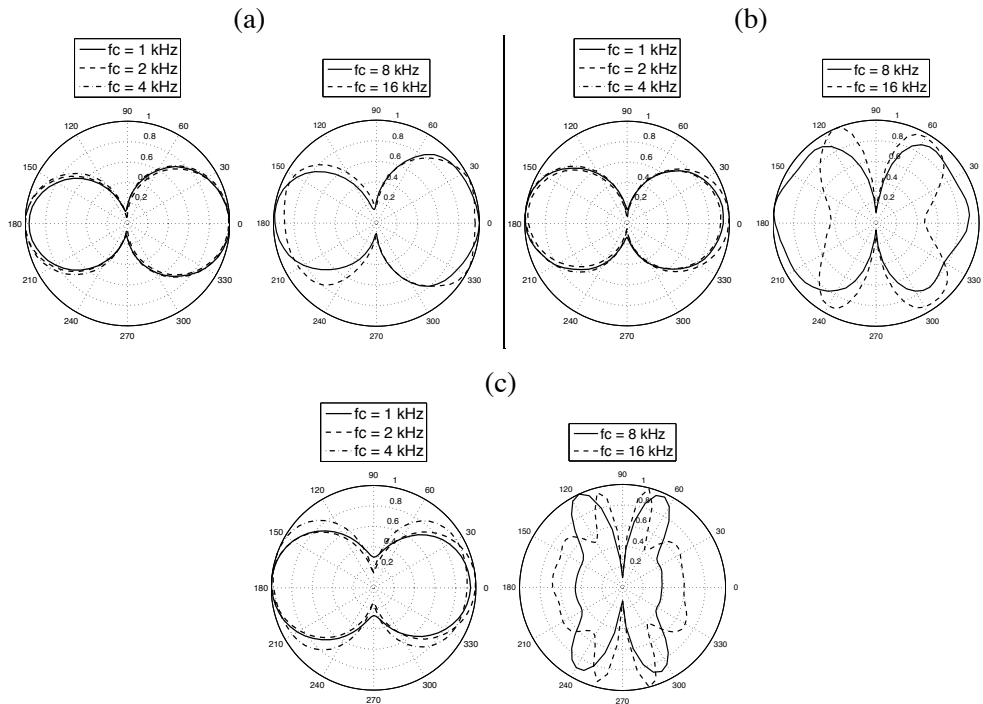
$$f_{sa} = \frac{c}{2d}, \quad (3)$$

yläpuolella dipolin suuntakuviot vääristyy ja suuntaestimaatti on virheellinen. Tämän lisäksi painegradientin tuottamissa dipolisignaaleissa on pienillä taajuuksilla verrattain suuri kohina riippuen mikrofonien välisestä etäisyydestä, mikä johtaa myös virheelliseen suuntaestimaattiin.

3.1 Neliöhilan evaluointi suunta-analyysissä

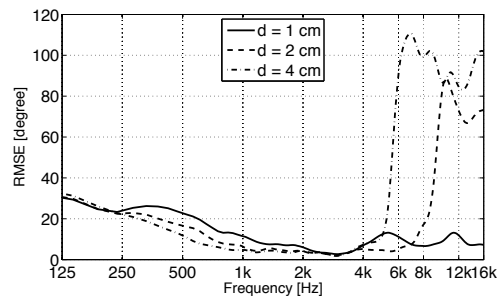
Edellä kuvatun neliöhilan käyttöä evaluointiin suunnan estimoinnissa. Kaiuttomat mittaukset suoritettiin kolmella eri hilalla, joissa vastakkaisten mikrofonien väliset etäisyydet olivat 1 cm, 2 cm ja 4 cm. Mikrofonien (Audio Technica MT350) impulssivasteet mitattiin kaiuttomassa huoneessa viiden asteen välein käyttäen logaritmista taajuuspyyhkäisyä, joka toistettiin kaiuttimella (Genelec 8030A) 1,6 m etäisyydellä neliöhilasta. Kuvassa 2 on esitetty suuntakuviot oktaavikaistoittain painegradientteille, jotka on laskettu mitatuista impulssivasteista kaavan (2) mukaan. Spatiaalisesta laskostumisesta aiheutuva suuntakuvioiden vääristymä näkyy selkeästi suurilla taajuuksilla, kun suuntakuviot 2 cm ja 4 cm neliöhiloilla vääristyvät 8 kHz ja 4 kHz oktaavikaistoilta alkaen. Tämä oli odotettavissa, kun vastaavat teoreettiset laskostumistaajuudet ovat 8,5 kHz ja 4,3 kHz. Neliöhilalla, jossa vastakkaisten mikrofonien välinen etäisyys on 1 cm, painegradientti tuottaa verrattain validin dipolisuuntakuvioiden kaikilla esitetyillä oktaavikaistoilla ($f_{sa}=17$ kHz).

Impulssivastemittausten lisäksi neliöhiloilla äänitettiin valkoista kohinaa yksittäisistä suunnista 0° , 15° , 30° ja 45° . Kohina tuotettiin kaiuttomassa huoneessa kaiuttimella 2 m etäisyydeltä neliöhilasta. A-taajuus- painotettu äänipainetaso oli 70 dB äänituspisteessä. Kuvassa 3 on esitetty suuntaestimaatit keskineliövirheen neliöjuuren avulla (RMSE, root-mean-square error). Suunnat laskettiin DirAC-analyysissä hiloissa olevien pallomikrofonien signaaleista luvussa 2 esitetyllä



Kuva 2: Suuntakuviot oktaavikaistoittain painegradientille, kun painemikrofonit ovat (a) 1 cm, (b) 2 cm ja (c) 4 cm etäisyydellä toisistaan.

tavalla. Suurilla taajuuksilla suuntaestimaatin suuri virhe on seurausta vääristyneestä suuntakuviosta (kuva 2). Dipolien verrattain suuren sisäisen kohinan vaikutus näkyy pienillä taajuuksilla, kun virhe on sitä suurempi mitä pienempi on mikrofonien välinen etäisyys. Suuri sisäinen kohina ja spatiaalinen laskostuminen yhdessä aiheuttavat sen, että suunta voidaan analysoida audiokaistalla ainoastaan rajoitetulla taajuusalueella.



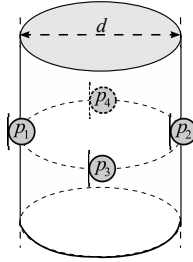
Kuva 3: Suuntaestimaatti neliömikrofonihilalla painegradienttia käyttäen. Suuntaestimaatin virhe on ilmoitettu keskineliövirheen neliöjuuren avulla (RMSE) taajuuden funktiona.

4 SUUNNAN ESTIMOINTI SYLINTERIMIKROFONIHILALLA

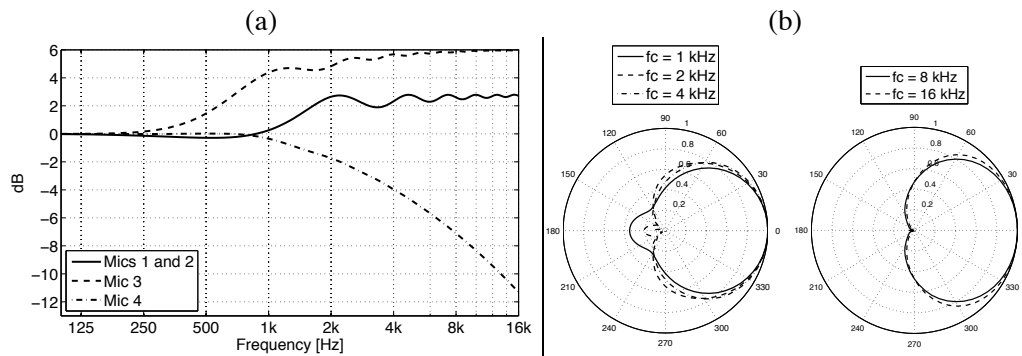
Kuten edellä osoitettiin, B-formaattisignaalien luominen neliöhilasta ja näiden käyttö suunta-analyysissä mahdollistaa kaistarajoitetun suuntaestimaatin. Taajuusalueutta luotettavalle suuntaestimaatille voidaan kuitenkin laajentaa käyttämällä kuvan 4 mukaista *sylinterihilaa* sekä mikrofonisignaalien energiaan perustuvaa laskentaa [3].

Sylinterihilassa neljä pallomikrofonia on sijoitettu jäykän sylinterin pinnalle, jolloin sylinteri varjostaa ja heijastaa saapuvaa ääntä. Tämä aiheuttaa mikrofonien välille tasoeroja riippuen äänen tulosuunnasta sekä taajuudesta. Sylinterin vaikutusta saapuvalla äänelle on havainnollistettu

kuvassa 5, kun neljä ideaalista pallomikrofonia on mallinnettu jäykän ja äärettömän pitkän sylinterin ($d = 8$ cm) pinnalle viitteen [4] mukaisesti. Kuvan (a)-kohdassa on esitetty mikrofonien taajuusvasteet tasoallon saapuessa etu- ja takamikrofonien (3 ja 4 kuvassa 4) välisen akselin suuntaisesti. Suurilla taajuuksilla ääni on vaimentunut merkittävästi takamikrofonissa akustisesta varjostuksesta johtuen ja toisaalta voimistunut etumikrofonissa saapuneen ja sylinterista heijastuneen äänen summautuessa keskenään. Kuvan (b)-kohdassa on esitetty suuntakuviot yhdelle ideaaliselle mikrofonille sylinterin pinnalla. Kuvasta nähdään, että suuntakuviot muistuttaa kardioidia suurilla taajuuksilla.



Kuva 4: Sylinterimikrofonihila, jossa neljä pallomikrofonia on sijoitettu jäykän sylinterin (hal-kaisija d) pinnalle.



Kuva 5: (a) Taajuusvasteet neljälle ideaaliselle pallomikrofonille, jotka on mallinnettu äärettömän pitkän sylinterin ($d = 8$ cm) pinnalle. Tasaalta saapuu etu- ja takamikrofonien (3 ja 4) välisen akselin suuntaisesti. (b) Suuntakuviot oktaavikaistoittain ideaaliselle pallomikrofonille sylinterin pinnalla.

Seuraavaksi käsitellään laskentaa, joka hyödyntää mikrofonien välisiä tasoeroja äänen tulouunnan estimoinnissa. Jos oletetaan yksittäinen äänilähde vapaassa kentässä (ilman varjostavaa sylinteriä), niin ääni-intensiteettivektori x -akselille voidaan määrittää kaavalla

$$I_x = -E \cos(\theta), \quad (4)$$

jossa E on tasoallon energia ja θ on sen saapumiskulma. Kun äänikenttään sijoitetaan varjostava sylinteri, niin sen pinnalla mitattu äänienergia eri tasoallon saapumiskulmilla muistuttaa suurilla taajuuksilla kardioidin suuntakuviota, mikä on osoitettu kuvassa 5 (b). Tällöin energia voidaan approksimoida kaavalla

$$|P_m|^2 = E|(1 + \cos(\theta - \theta_m))|^2, \quad (5)$$

jossa θ_m vastaa mikrofonin m pääakselin suuntaa. Nyt x -akselin suuntaista ääni-intensiteettivektoria voidaan approksimoida laskemalla energiagradiendi kahden vastakkaisiin suuntiin (tässä 0° ja 180°) osoittavan mikrofonin väliltä

$$\begin{aligned} I_x &= -[|P_1|^2 - |P_2|^2] \\ &= -[E|(1 + \cos(\theta - 0^\circ))|^2 - E|(1 + \cos(\theta - 180^\circ))|^2] \\ &= -E \cos(\theta). \end{aligned} \quad (6)$$

Edellä kuvattu toteutetaan sylinterihilan tapauksessa vähentämällä vastakkaisten mikrofonien tehospektrit toisistaan,

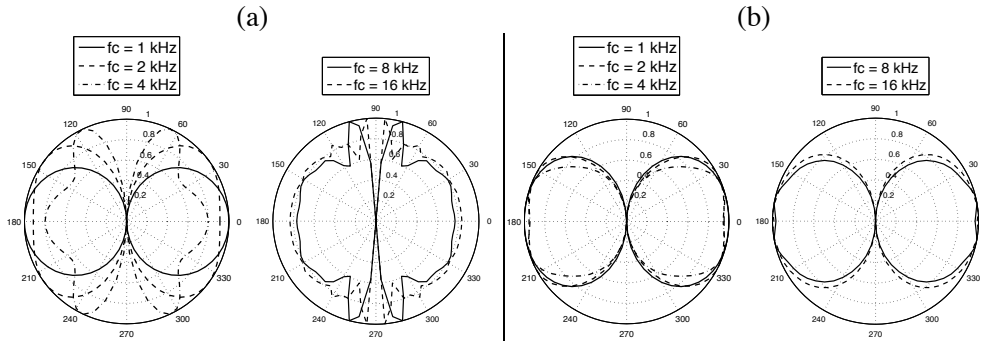
$$\begin{aligned}\tilde{I}_x(t, f) &= |P_1(t, f)|^2 - |P_2(t, f)|^2 \\ \tilde{I}_y(t, f) &= |P_3(t, f)|^2 - |P_4(t, f)|^2,\end{aligned}\quad (7)$$

mikä approksimoi intensiteettivektorin komponentteja x- ja y-akseleille. Äänen tulosuunta saadaan ratkaistua yksiselitteisesti intensiteettivektorin komponenteista \tilde{I}_x ja \tilde{I}_y , jolloin mikrofonien vaihetta ei oteta huomioon eikä spatiaalinen laskostuminen vaikuta suuntaestimaattiin.

Edellä kuvattu energigradientin laskenta vaatii kuitenkin merkittäviä tasoeroja mikrofonien välille riittävällä signaali-kohinasuhteella. Pienillä taajuuksilla sylinteri ei varjosta saapuvaa ääntä, jolloin energigradienttia ei voida laskea luotettavasti. Tämän takia mikrofonisignaalien välille lasketaan pienillä taajuuksilla painegradientti, ja suunta-analyysi toteutetaan samoin kuin perinteisellä neliöhilalla. Rajataajuus paine- ja energigradienttien käytölle lasketaan kaavalla [3]

$$f_{\text{lim}} = \frac{c}{\pi d}.\quad (8)$$

Kuvassa 6 on esitetty suuntakuviot paine- ja energigraduateille, jotka on laskettu kahden ideaalisen pallomikrofonin signaaleista (mikrofonit 1 ja 2 kuvassa 4). Mikrofonit on mallinnettu jäykän sylinterin ($d = 8$ cm) pinnalle. Kuvasta nähdään, että painegradientti tuottaa validin dipolin 1 kHz oktaavikaistalla. Energigradientti tuottaa dipolimaisen suuntakuviota kaikilla esitetyillä taajuuskaistoilla.

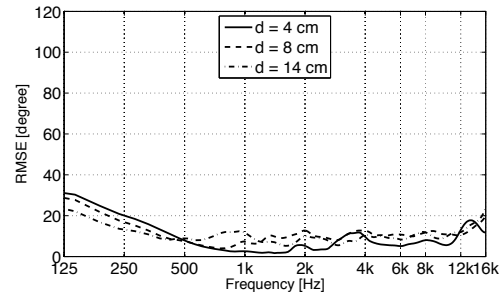


Kuva 6: Suuntakuviot oktaavikaistoittain (a) paine- ja (b) energigraduateille, jotka on laskettu kahden ideaalisen pallomikrofonin signaaleista. Pallomikrofonit on mallinnettu jäykän äärettömän pituisen sylinterin ($d = 8$ cm) pinnalle.

4.1 Sylinterihilan evaluointi suunta-analyysissä

Sylinterihilan käyttöä suunnan analysoinnissa ja edellä kuvattua laskentaa evaluoitiin kaiuttomilla mittauksilla. Neljä pallomikrofonia (Audio Technica MT350) sijoitettiin vuorotellen kolmen halkaisijaltaan eri kokoisen sylinterin pinnalle. Sylinterien halkaisijat olivat 4 cm, 8 cm ja 14 cm, ja pituus 10 cm. Sylinterihilat asetettiin kovan levyn päälle (0,83 m x 0,83 m), sen keskelle. Tämän katsottiin vastaavan tilannetta, jossa hila on sijoitettu pöydälle. Mikrofonit sijaitsivat sylinterin levyn puoleisessa päässä, lähellä levyä. Sylinteri oli kiinnitetty tiiviisti levyyn. Sylinterihiloilla äänitettiin valkoista kohinaa yksittäisistä suunnista: 0° , 15° , 30° ja 45° . Kohina toistettiin kaiuttimelle (Genelec 8030A) 2 m etäisyydellä sylinterihilasta. A-taajuuspainotettu äänipainetaso oli 70 dB äänityspisteessä. Kuvassa 7 on esitetty suunta-analyysin tulokset RMSE-arvoilla eri sylinterihiloille. Tuloksista nähdään, että suuntaestimaatti on luotettavaa suurilla taajuuksilla, mikä on huomattava parannus 2 cm ja 4 cm neliöhiloihin verrattuna (kuva 3). Lisäksi

vastakkaiset mikrofonit voidaan sijoittaa kauemmaksi toisistaan sylinterihilassa, jolloin saadaan luotettavampaa suuntaestimaattia pienillä taajuuksilla verrattuna neliöhiloihin.



Kuva 7: Suuntaestimaatti sylinterimikrofonihilalla paine- ja energiagradienttien kombinaatioita käyttäen. Suuntaestimaatin virhe on ilmoitettu keskineliövirheen neliöjuuren avulla (RMSE) taajuuden funktiona.

5 YHTEENVETO

Artikkelissa käsiteltiin äänen tulosuunnan analysointia laajakaistaisesti neljän toisiaan lähelle sijoitetun painemikrofonin signaaleista. Tulosuunnan analysointi perustuu ääni-intensiteettiin, jota estimoidaan mikrofonisignaaleista lasketuilla paine- ja energiagradienteilla taajuusriippuvaisesti. Perinteisesti intensiteetti lasketaan kaikilla taajuuksilla painegradientin avulla. Painegradientti tuottaa kuitenkin toivotun dipolin suuntakuvion ainoastaan pienillä taajuuksilla. Suurilla taajuuksilla suuntakuviot ovat vääristyneet spatiaalisesta laskostumisesta johtuen, mikä johtaa virheelliseen suuntaestimaattiin. Tämä voidaan kuitenkin välttää estimoimalla intensiteettiä energiagradientilla suurilla taajuuksilla. Energiagradientin laskenta vaatii mikrofonien välisiä tasoeroja, jotka voidaan saavuttaa sijoittamalla painemikrofonien väliin ääntä varjostava sylinteri. Laajakaistaista, paine- ja energiagradienttiin perustuvaa suuntaestimaattia, evaluoitiin tila-äänien prosessointiin tarkoitetussa DirAC-menetelmässä (Directional Audio Coding). Tulokset osoittivat, että paine- ja energiagradienttien taajuusriippuvainen käyttö tuottaa luotettavaa suuntaestimaattia koko audiokaistalla, kun taas perinteisellä menetelmällä (painegradientti kaikilla taajuuksilla) suunta voidaan analysoida luotettavasti ainoastaan rajoitetussa taajuusikkunassa.

6 KIITOKSET

Tutkimusta on rahoittanut Fraunhofer IIS. This work has been supported by the European Research Council under the European Community's Seventh Framework Programme (FP7/2007-2013) / ERC Grant agreement No. 240453.

VIITTEET

- [1] PULKKI V, Spatial sound reproduction with Directional Audio Coding, *J. Audio Eng. Soc.*, **55**(2007), 503–516.
- [2] FAHY F, Sound intensity, *Elsevier Science Publishers Ltd.*, (Essex, England, 1989).
- [3] AHONEN J, GALDO G D, KUECH F, & PULKKI V, Directional analysis for spatial audio using a microphone array mounted on a rigid cylinder, *unpublished manuscript*, (2011).
- [4] TEUTSCH H, Wavefield decomposition using microphone arrays and its application to acoustic scene analysis, *Doctoral Thesis, Friedrich-Alexander University of Erlangen-Nurnberg*, (2005).