

TILAIMPULSSIVASTEIDEN ANALYYSI JA SYNTEESI HUONEAKUSTIIKASSA

Sakari Tervo, Jukka Pätynen ja Tapio Lokki

Aalto-yliopisto, Perustieteiden korkeakoulu, Mediatekniikan laitos
PL 15500, FIN-00076 AALTO
sakari.tervo@aalto.fi

Tiivistelmä

Huoneen impulssivaste on perinteisesti mitattu yhdellä pallo-suuntakuvion omaavalla mikrofonilla, jolloin äänen tulosuuntaa ei pysty mitatusta vasteesta analysoimaan. Jos impulssivaste mitataan monella mikrofonilla yhtäaikaan voidaan mikrofonien välisten erojen perusteella analysoida äänen tulosuunta ajan ja taajuuden funktiona. Näin ollen saadaan tilaimpulssivasteen jokaiselle näytteelle laskehtua suuntainformaatio, jota voidaan käyttää äänen tilallisen käyttäytymisen analysoinnissa sekä monikanavaäänentoistossa. Tässä artikkelissa esitellään monimikrofonitekniikat, joita olemme käyttäneet konserttisalien tilaimpulssivasteiden analyysiin. Lisäksi esitetään tilaimpulssivasteen käyttöä monikanavakuuntelussa, jossa impulssivasteen energia jaetaan toistokaiuttimien kesken ja konvoloidaan kaiuttomien äänitysten kanssa.

1 JOHDANTO

Tilaaänen äänitys ja toistotekniikat ovat tärkeitä välineitä huoneakustiikan tutkimuksessa [1, 2]. Tilojen akustiikan arviointia varten mitataan ensin huoneen tilaimpulssivaste joka koodataan jollekin monikanavaiselle tilakaiutinjärjestelmälle, ja konvoloidaan kaiuttoman musiikin kanssa. Viimeistä osaa prosessissa kutsutaan tyypillisesti konvoluutiokaiuksi.

Edelliset tutkimukset ovat esittäneet useita tilaaänen koodausmenetelmiä, joita voidaan soveltaa tilaimpulssivasteisiin. Koodausmenetelmät voidaan jakaa kolmeen ryhmään niiden tavoitteiden mukaan. Ensimmäisen ryhmän tavoitteena on tuottaa mitattu äänikenttä tietyllä alueella. Näitä menetelmiä ovat ensimmäisen asteen Ambisonics (EAA), korkeamman asteen Ambisonics (KAA) [3] ja aaltokenttäsynteesi (AKS) [4, 5]. Toisessa ryhmässä, binauraalisissa tilaaänen toistomenetelmissä, on tarkoitus toistaa äänenpaine oikein kuuntelijan tärykalvoilla mittaamalla äänenpaine lähellä tärykalvoja tai tärykalvoilla [6]. Kolmannessa ryhmässä lähtökohtana on analysoida ja toistaa ihmisen havaintojen kannalta oleelliset vihjeet oikein [7]. Esimerkki tällaisesta menetelmästä on Spatial Room Impulse Response Rendering (SIRR)-menetelmä [8]. Kaksi ensimmäistä ryhmää vaativat erikoismikrofoneja tai mikrofoniasetelmia, kun taas viimeisen ryhmän menetelmiä voidaan soveltaa useammalle mikrofoniasetelmälle. Kaksi ensimmäistä ryhmää ovat aitoja tilaaänitysmenetelmiä, mutta kolmannen ryhmän menetelmillä tilaaänitys joudutaan tekemään tilaimpulssivasteiden kautta. Näin ollen ne eivät so-

vellu jatkuvan äänen äänittämiseen. Tässä artikkelissa keskitytään tilaimpulssivasteiden koodaamiseen.

Useimpien alalla työskentelevien ammattilaisten mielestä mikä tahansa yllämainituista menetelmistä voi tuottaa uskottavan tilavaikutelman mitatusta akustiikasta. Kuitenkin erityismikrofonien käyttäminen rajoittaa menetelmiä. Esimerkiksi EAA-menetelmässä usein sovelletun ensimmäisen asteen B-formaattimikrofonien tiedetään olevan epätarkkoja suunnan arvioinnissa, varsinkin korkeammilla taajuuksilla, mikä luonnollisesti vaikuttaa analyysin ja tiläänen toiston tarkkuuteen. Lisäksi mittaus, erityisesti AKS-menetelmässä, vaatii joko lukuisia mikrofoneja tai on aikaa vievää, joka voi olla kallista. Lopuksi jotkin menetelmät ovat rajoittuneet vain yhteen mikrofoniasetelmaan tai erityismikrofoniin.

Tässä tutkimuksessa esitetellään tiläänen koodausmenetelmä tilaimpulssivasteille nimeltään kuvalähdehajotelma-menetelmä (KLH). Toisin kuin aiemmin kehitetyt menetelmät, KLH-menetelmää voidaan soveltaa mielivaltaiseen mikrofoniasetelmaan, joka koostuu neljästä tai useammasta mikrofonista. Lisäksi koodattu tilainformaatio voidaan toistaa millä tahansa äänentoistotekniikalla. Esitetty menetelmä nojautuu yksinkertaisen oletukseen, että äänen etenemissuunta on kullakin ajanhetkellä t kaikkien mikrofoniasetelmaan saapuvien aaltojen keskiarvo, ja äänenpaine mikrofoniasetelman keskellä on siihen liittyvä paine. Menetelmä analysoi tilaimpulssivasteen tällä oletuksella. Lopputuloksena tilaimpulssivaste on koodattu paine ja- paikkatietona.

Koodattua painetta ja paikkatietoa voidaan hyödyntää tiläänen toistossa sekä akustiikan visualisoinnissa.

2 MENETELMÄ

KLH olettaa, että impulssivaste voidaan esittää joukkona kuvalähteitä. KLH analysoi tilaimpulssivasteen jokaisella diskreetillä ajanhetkellä $\Delta t = 1/f_s$, missä f_s on näytteenottotaajuus. Äänen keskimääräinen tulosuunta arvioidaan paikannusmenetelmällä jokaiselle ajanhetkelle. Tämän tuloksena tilaimpulssivaste voidaan esittää painearvoina ja niiden sijaintina, eli kuvalähteinä. Kuvalähdehajotelma kuvaa tätä prosessia ja siitä menetelmän nimi.

Mikrofoniasetelma voi olla mielivaltainen, kunhan se täyttää seuraavat ehdot:

- 3-D-tiläänen koodausta varten mikrofoneja täytyy olla vähintään 4 ja niiden täytyy virittää kolmiulotteinen avaruus,
- yhden mikrofonin suuntakuviot on "pallo"ja
- asetelman geometria on pieni verrattuna etäisyyteen lähteestä.

Tilaimpulssivasteen $\mathbf{h}(t) = \{h_n(t)\}_{n=1}^N$ analyysi KLHlla N lle mikrofonille etenee seuraavasti.

2.1 Vaihe 1: Paikannus

KLH ratkaisee ensin lähteen ja kuvalähteiden paikat tilaimpulssivasteesta. Mitä tahansa paikannusmentelmää voidaan käyttää. Tässä artikkelissa käytetään aikaeroihin perustuva paikannusmenetelmää.

Aikaerot estimoidaan korrelaatiomenetelmällä [9] jokaiselle diskreetille ajanhetkelle k lyhyessä analyysi-ikkunassa. Lisäksi jokainen aikaero interpoloidaan eksponentiaalisella sovituksella [10]. Aikeroestimaattit kirjoitetaan matriisimuodossa

$$\hat{\boldsymbol{\tau}}_k = [\hat{\tau}_{1,2}^{(k)}, \hat{\tau}_{1,3}^{(k)}, \dots, \hat{\tau}_{N-1,N}^{(k)}]^T,$$

ja vastaava matriisi mikrofonien paikkojen erotusvektoreille kirjoitetaan

$$\mathbf{V} = [\mathbf{r}_1 - \mathbf{r}_2, \mathbf{r}_1 - \mathbf{r}_3, \dots, \mathbf{r}_{N-1} - \mathbf{r}_N]^T.$$

Pienimmän neliösumman ratkaisu hitausvektorille on annettu [11, s. 75]:

$$\hat{\mathbf{m}}_k = \mathbf{V}^+ \hat{\boldsymbol{\tau}}_k, \quad (1)$$

missä $(\cdot)^+$ on Moore-Penrosen pseudo-käänteismatriisi, ja ääniaallon saapumissuunta mikrofoniasetelmaan nähden on $\hat{\mathbf{n}}_k = -\hat{\mathbf{m}}_k / \|\hat{\mathbf{m}}_k\|$. Etäisyys kuvautuu suoraan aikaindeksin k ja äänen nopeuden avulla $d_k = ck\Delta t$.

2.2 Vaihe 2: Äänenpaine

Toisessa vaiheessa analyysiä, valitaan yksi käytössä olevista pallo-suuntakuvion omaavista mikrofoneista painesignaaliksi h_p . Ihannetapauksessa painesignaalin mikrofoni sijaitsee asetelman geometrisessä keskipisteessä. Tässä tapauksessa kunkin näytteen paine impulssivasteessa $h_p(\Delta tk)$ assosioidaan 3-D-paikkaan $\hat{\mathbf{x}}_k$, joka on saatu analyysin 1. vaiheesta. Siten menetelmä on koodannut tilaimpulssivasteen neljällä arvolla per ajanhetki, paine ja 3-D paikka-arvoina.

Jos mikrofoni ei ole asetelman geometrisessä keskipisteessä, voidaan paine ennustaa kuvalähteiden paikoista. Ensin lasketaan kuvalähteiden etäisyys painemikrofonin sijaintiin nähden \mathbf{r}_p

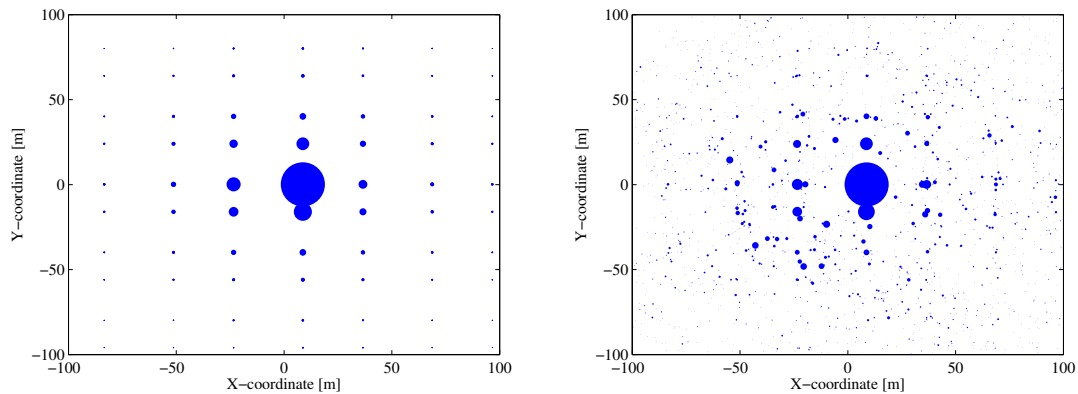
$$d_k = \|\mathbf{r}_p - \mathbf{x}_k\|, \quad (2)$$

ja sitten päivitetään jokaisen kuvalähteen painearvo $h_p(f_s d_k / C)$. Käytettäessä tasoaltomallia, etäisyys lasketaan seuraavasti:

$$d_k = |\mathbf{n}_k(\mathbf{r}_p - \mathbf{x}_{k,0})|, \quad (3)$$

missä \mathbf{n}_k ja $\mathbf{x}_{k,0}$ ovat tason normaali ja piste tasolla.

Sen sijaan, että pyrittäisiin ennustamaan paine asetelman keskellä, kuten edellä, voidaan ennustaa kuvalähteiden sijainti asetelman keskikohtaan nähden. Tämä on laskennallisesti helpompaa, koska signaalia ei tarvitse uudelleennäytteistää. Tässä artikkelissa kumpaakaan menetelmistä ei käytetä, koska painemikrofonin oletetaan olevan geometrisessä keskipisteessä.



(a) Kvalähteiden sijainti ja amplitudit.

(b) Analysoidut kvalähteiden sijainnit ja amplitudit

Kuva 1: Esimerkki kvalähteistä (a) ja analysoiduista kvalähteistä (b) esitellyllä menetelmällä. Lähinnä asetelmaa, eli origoa, olevat kvalähteet analysoidaan oikein.

2.3 Esimerkki

Tässä kappaleessa esitellään KLHn toimintaa tilaimpulssivastesimulaation avulla. Tilaimpulssivaste on simuloitu suorakulmion muotoisessa huoneessa ($20 \times 30 \times 12$) m³. Lähteen paikka on [16, 04, 8, 06, 3, 58] m ja asetelman geometrisen keskipisteen [7, 35, 7.92, 3.22] m. Tilaimpulssivaste on jaettu 1,33 ms pituisiin Hanning-ikkunoihin ja kahden peräkkäisen ikkunan päällekkäisyys on 99 %. Äänen nopeus on $c = 345$ m/s, näytteenottotaajuus $f_s = 48$ kHz, heijastuskerroin 0,85 kaikille pinnoille. Simuloinnissa laskettiin kvalähtet ja 45 kertalukuun asti.

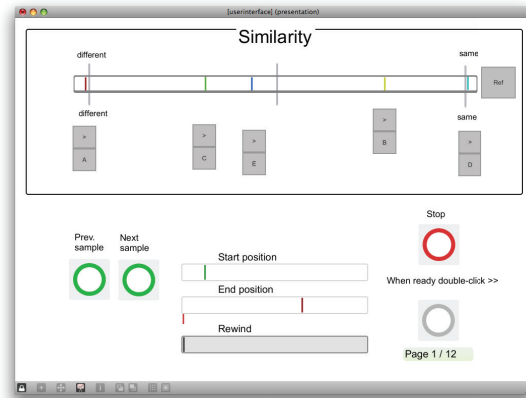
Kuvassa 1 esitetään analyysin tulokset. Ympyrän säde vastaa amplitudia ja paikka vastaa kvalähteen paikkaa. Kuten nähdään kuvasta 1, lähellä asetelmaa olevien kvalähteiden paikka ja amplitudi (a) on hyvin samanlainen kuin analysoitujen kvalähteiden (b). Kauempana olevien kvalähteiden paikkaa ei saada tarkasti, mutta havainnon kannalta tarkkuus oletetaan riittäväksi.

3 KUUNTELUKOKEET

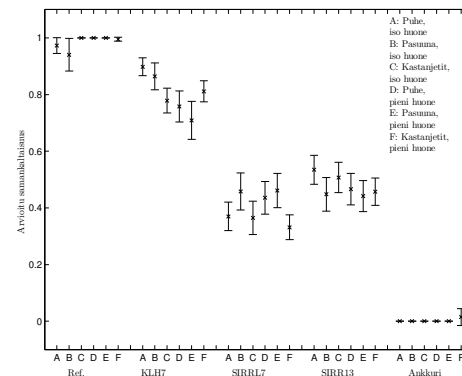
Kuuntelukokeissa verrattiin esitettyä KLH-menetelmää SIRR-menetelmään simuloituissa tapauksissa. Mikrofonien paikat valittiin oktaedrin kulmapisteisiksi, jonka säde oli $d_{\text{spc}}/2$. Simuloidut tilaimpulssivasteet tuotettiin kvalähdemenetelmällä taulukossa 1 esitettyjen parametrien mukaan. Lisäksi impulssivasteet suodatettiin ilman absorptiosuotimella.

Viittä tilaäänien koodausmenetelmää verrattiin kuuntelukokeissa. Menetelmät olivat seuraavat:

- KLH yhdellä mikrofoniaasetelmalla $d_{\text{spc}} = 100$ mm, (KLH7)
- SIRR yhdellä mikrofoniaasetelmalla $d_{\text{spc}} = 100$ mm, (SIRR7) ja



(a) Kuvakaappaus



(b) Kuuntelukoetulokset

Kuva 2: (a) Kuvakaappaus kuuntelukokeen käyttöliittymästä. Koehenkilöt voivat vapaasti liikutella arvioitaan ja kuunnella näytteitä mistä tahansa kohdasta. (b) Arvioitu samankaltaisuus tiläänenkoodausmenetelmien ja referenssin välillä. Tulokset on esitetty keskiarvoilla ja 95 % luotettavuusvälillä.

Taulukko 1: Lähde ja kuuntelupaikat, lähdesignaalit, huonegeometria ja näytteiden nimeämiskäytäntö kuuntelukokeessa.

Näyte (Ääni)	Lähteen paikka			Kuuntelupaikka		
	x [m]	y [m]	z [m]	x [m]	y [m]	z [m]
Iso huone (30 × 20 × 12) m ³						
A (Pu.)	16.04	8.06	3.58	7.35	7.92	3.22
B (Pa.)	17.44	12.81	2.88	2.64	13.48	3.72
C (Ka.)	20.37	11.99	2.52	3.10	12.10	2.86
Pieni huone (5 × 3 × 2.8) m ³						
D (Pu.)	3.44	0.80	1.53	1.02	0.64	1.40
E (Pa.)	3.87	1.45	1.65	0.76	1.39	1.33
F (Ka.)	3.78	0.85	1.81	1.24	0.97	2.07

Pu.: Puhe, Pa.: Pasuuna, and Ka.: Kastanjetit

- SIRR kahdella mikrofoniaasetelmalla $d_{\text{spc}} = 100$ mm ja $d_{\text{spc}} = 25$ mm (SIRR13)

Lisäksi yksi tiläänen koodausmenetelmä oli piilotettu referenssinäyte ja toinen oli referenssinäyte soitettuna yhdestä kaiuttimesta (mono).

Koehenkilöiden tehtävä oli arvioida kuinka samankaltainen kuunneltu näyte on referenssinäytteen kanssa. Esimerkki kuuntelukokeen käyttöliittymästä on esitetty Kuvassa 2(a).

Kuuntelukokeet järjestettiin pienessä kuunteluhuoneessa, ja tilääni toistettiin 14 Genelec 8030A kaiuttimella. Tiläänen toistojärjestelmänä toimi vektorikanta-amplitudipanorointi [7].

Kuuntelukokoiden tulokset kuvassa 2(b) osoittavat, että KLH tuottaa simuloitussa tilanteessa tiläänihavainnon, joka on lähempänä referenssiä kuin SIRR. Parhaassa tapauksessa KLH tuottaa yhtä hyvän tuloksen kuin piilotettu referenssinäyte. Viime aikoina esitellyt parannukset Directional Audio Coding-menetelmässä [12] voisivat mah-

dollisesti hyödyttää myös SIRR-menetelmää. Niiden soveltamista tulisi harkita SIRR-menetelmä käytettäessä.

4 KESKUSTELUA JA JOHTOPÄÄTÖKSET

Tässä artikkelissa esiteltiin menetelmä tiläänen koodaukseen. Menetelmää voidaan soveltaa tiläänen analyysiin, kuten on tehty esimerkiksi artikkelissa [13]. Tässä artikkelissa menetelmää testattiin tiläänen toistoon. Parhaassa tapauksessa koodattua tilääntä ei erotettu referenssistä. Menetelmä on toteutettu tässä paperissa laajakaistaisena. Taajuuskaistoittaisen toteutuksen etuna olisi parantunut tarkkuus, sillä reaaliolanteessa heijastukset voivat olla suodattuneita.

Kiitokset

Tätä tutkimusta on rahoittanut Suomen Akatemia, projekti [257099] sekä Euroopan tutkimusneuvosto, ERC grant agreement no. [203636].

VIITTEET

- [1] LOKKI T, VERTANEN H, KUUSINEN A, PÄTYNEN J, & TERVO S, Concert hall acoustics assessment with individually elicited attributes, *J. Acoust. Soc. Am.*, **130**(2011), 835–849.
- [2] LOKKI T, PÄTYNEN J, TERVO S, SILTANEN S, & SAVIOJA L, Engaging concert hall acoustics is made up of temporal envelope preserving reflections, *J. Acoust. Soc. Am.*, **129**(2011), EL223–EL22.
- [3] DANIEL J, NICOL R, & MOREAU S, Further investigations of high order ambisonics and wavefield synthesis for holophonic sound imaging, in *114th Audio Eng. Soc. Convention, Amsterdam, the Netherlands*, 2003, paper 5788.
- [4] BERKHOUT A, DE VRIES D, & VOGEL P, Acoustic control by wave field synthesis, *J. Acoust. Soc. Am.*, **93**(1993) 5, 2764–2778.
- [5] BOONE M, VERHEIJEN E, & VAN TOL P, Spatial sound-field reproduction by wave-field synthesis, *J. Audio Eng. Soc.*, **43**(1995) 12, 1003–1012.
- [6] HAMMERSHØI D & MØLLER H, *Communication acoustics*, chapter 9 Binaural technique–basic methods for recording, synthesis, and reproduction, pages 223–254, Springer-Verlag, New York, NY, USA, 2005.
- [7] PULKKI V, Virtual sound source positioning using vector base amplitude panning, *J. Audio Eng. Soc.*, **45**(1997) 6, 456–466.
- [8] MERIMAA J & PULKKI V, Spatial impulse response rendering I: Analysis and synthesis, *J. Audio Eng. Soc.*, **53**(2005) 12, 1115–1127.
- [9] KNAPP C & CARTER G, The generalized correlation method for estimation of time delay, *IEEE Trans. Acoust., Speech and Signal Proc.*, **24**(1976) 4, 320–327.
- [10] ZHANG L & WU X, On cross correlation based-discrete time delay estimation, in *IEEE International Conference on Acoustics, Speech, and Signal Processing*, volume 4, pages 981–984, 2005.
- [11] PIRINEN T, *Confidence Scoring of Time Delay Based Direction of Arrival Estimates and a Generalization to Difference Quantities*, Ph.D. thesis, Tampere University of Technology, 2009, publication; 854.
- [12] PIHLAJAMÄKI T & PULKKI V, Low-delay directional audio coding for real-time human-computer interaction, in *130th Audio Eng. Soc. Convention, London, UK*, 2011, paper No. 8413.
- [13] PÄTYNEN J, TERVO S, & LOKKI T, Analysis of concert hall acoustics via visualizations of time-frequency and spatiotemporal responses, *J. Acoust. Soc. Am.*, **133**(2013), 17.