

ÄÄNEN TULOSUUNNALLE HERKKIEN TUMAKKEIDEN TOIMINNALLISTEN MALLIEN ULOSTULOJEN YHDISTÄMINEN

Olli Santala, Marko Takanen ja Ville Pulkki

Aalto-yliopiston Sähkötekniikan korkeakoulu
Signaalinkäsittelyn ja akustiikan laitos
Otakaari 5, 02150 Espoo
olli.santala@aalto.fi

Tiivistelmä

Ihmisen kuulojärjestelmä käyttää korvienvälisiä aika- ja tasoeroja sekä spektrivihjeitä äänten paikallistamiseen. Neurofysiologiset tutkimukset ovat osoittaneet, että nämä vihjeet muutetaan suuntainformaatioksi kuulohermoradalla sijaitsevissa ylemmän oliivitulomakekompleksin tumakkeissa. Näitä tumakkeita on yhteensä neljä, kaksi samanlaista kullakin aivopuoliskolla, mikä voisi mahdollistaa neljän samanaikaisen suuntahavainnon muodostumisen yhdellä taajuuskaistalla. Lisäksi kuulojärjestelmämme kykenee pääättelemään tulosuuntaa myös verhoikäyrien aikaerojen perusteella, joten teoriassa olisi mahdollista havaita kuusi eri suuntaa samanaikaisesti kullakin taajuuskaistalla. Psykoakustisten kokeiden perusteella näin montaa suuntaa ei kuitenkaan havaita.

Tässä artikkelissa esitellään menetelmä tumakkeiden toiminnallisten mallien tuottaman suuntainformaation yhdistämiseen kuulon toimintaa jäljittelevällä tavalla. Tämä tehdään kummallekin aivopuoliskolle erikseen, jolloin saadaan kaksi samanaikaista vihjetä kullekin taajuuskaistalle. Tällöin voitaisiin havaita kaksi suuntaa yhdellä taajuuskaistalla, mikä vastaa paremmin psykoakustisten kokeiden tuloksia. Yhdistämisessä muodostettuja vihjeitä käytetään auditorisessa mallissa tuottamaan sille syötetystä signaalista aktivaatiokartta, jossa esitetään signaalin aiheuttama aktivaatio eri aivopuoliskoilla ajan funktiona. Mallin toimintaa sekä aktivaatiokarttojen sisältöä esitellään muutamien erilaisten psykoakustisten ilmiöiden avulla.

1 JOHDANTO

Ihmisen kuulojärjestelmän mallinnukseen on panostettu runsaasti muutaman viimeisimmän vuosikymmenen aikana. Tavoitteena on toteuttaa malli, joka pystyisi jäljittelemään ihmisen kykyä havaita äänitapahtumia eli analysoidaan ääniympäristöä samalla tarkkuudella kuin mihin ihminen kykenee. Tässä artikkelissa keskitytään suuntakuulon mallinnukseen neurofysiologian pohjalta.

Suuntakuulon toiminta perustuu erilaisten suuntavihjeiden koodaamiseen: korvienvälinen aikaero (interaural time difference, ITD), korvienvälinen tasoero (interaural level difference, ILD) ja äänen verhoikäyrästä analysoitavat aikaerot ovat merkittävimpiä suuntavihjeitä, ja näiden lisäksi korvalehden suuntariippuvan akustisen suodatuksen aiheuttamat spektraaliset vihjeet auttavat äänten lokalisaatiossa [1]. Neurofysiologiset tutkimukset ovat osoittaneet, että suunta- ja tilakuulemisen kannalta keskeisiä osia kuulohermoradalla

ovat keskimäinen ylempi oliiviumake (medial superior olive, MSO) sekä laitimmainen ylempi oliiviumake (lateral superior olive, LSO). Näistä MSO:n on havaittu olevan herkkä ITD:lle [2], kun taas LSO:n on havaittu olevan herkkä ILD:lle kaikilla taajuuksilla ja lisäksi ITD:lle pienillä taajuuksilla [3].

Tässä artikkelissa esitellään MSO:n ja LSO:n ulostulojen yhdistämismenetelmä, jonka avulla on mahdollista havainnollistaa äänten aiheuttamaa aktivaatiota eri aivopuoliskoilla. Menetelmä sisältyy binauraaliseen kuulojärjestelmän malliin [4], jonka toiminta perustuu ääreiskuulon ja tumakkeiden osalta neurofysiologiaan ja siitä eteenpäin psykoakustisten kokeiden tuloksiin. Menetelmän toiminnallisuutta esitellään muutamien esimerkitapausten avulla, joissa vertaillaan mallinnuksen tuloksena saatuja aktivaatiokarttoja psykoakustisiin tuloksiin.

2 MENETELMÄT

Artikkelissa käytetyn binauraalisen kuulojärjestelmän mallin aikaisempaa kehitystasetta sekä sen toimintaa on esitelty kirjoittajien artikkelissa [5]. Sittemmin mallia on kehitetty erityisesti MSO:n ja LSO:n jälkeen tapahtuvan analyysin osalta. Tässä artikkelissa keskitytään esittelemään MSO:n ja LSO:n mallien ulostulojen eli suuntavihjeiden yhdistämistä. Aluksi kuitenkin esitellään varsinaisen mallin osiot lyhyesti. Malli esitellään tarkemmin viitteessä [4].

Mallin herätteenä toimii binauraalinen signaali, joka prosessoidaan ensimmäiseksi vasemman ja oikean aivopuoliskon ääreiskuulon mallien läpi. Ne koostuvat epälinearisesta aika-alueen simpukan mallista [6] sekä simpukkatumakkeen mallista. Näiden ulostulot syötetään kullakin puolella kolmeen tumakemalliin, eli MSO:n, LSO:n ja laajakaistaisen MSO:n malleihin, jotka vastaavat eri suuntavihjeiden analysoinnista. Vihjeet yhdistetään tässä artikkelissa esitetyllä menetelmällä. Tämän jälkeen prosessoinnissa on äänten alukkeiden korostunutta vaikutusta simuloiva funktio sekä puoltenvälistä vertailua simuloiva funktio. Lopulta molempien aivopuoliskojen osalta on saatu erityiset *missä* ja *mitä* -vihjeet. *Mitä*-vihjeen sisältö sijoitetaan binauraaliselle aktivaatiokartalle *missä*-vihjeen määräämään paikkaan. Lopputuloksena saatu kartta esittää signaalin aiheuttaman aktivaation eri aivopuoliskoilla ja eri taajuuksilla.

2.1 Ulostulojen yhdistämismenetelmä

Psykoakustisissa kokeissa on saatu tuloksia, joiden mukaan yhdellä kriittisellä kaistalla voidaan havaita vain yksi äänen tulosuunta yhdellä ajanhetkellä [1]. Lisäksi joissain erikoistapauksissa kapeakaistainen heräte voi johtaa kahden samanaikaisen, eri puolilla sijaitsevien suuntien havaitsemiseen. Koska tumakemalleja on tässä mallissa yhteensä kuusi, kolme kummallakin aivopuoliskolla, se johtaisi kuuteen yhtäaikaiseen äänihavaintoon, mikä ei ole psykoakustisesti perusteltavissa. Siksi näihin tietoihin pohjautuen vihjeet yhdistetään tässä menetelmässä kullakin taajuuskaistalla ja kummallakin aivopuoliskolla erikseen.

Jotta eri tumakemallien ulostulot olisivat vertailtavissa ja yhdistettävissä, arvot muunnetaan suuntakulmiksi vertailemalla niitä referenssiarvoihin, jotka laskettiin käyttäen

keinopäällä eri kulmista mitattuja vapaan kentän siirtofunktioita pistemäisiltä lähteiltä tärykalvolle (head-related transfer function, HRTF). Ulostulot muutettiin kullakin aivopuoliskolla diskreeteiksi arvoiksi $-20^\circ, -10^\circ, \dots, 90^\circ$.

Yhdistämisessä eri tumakemallien ulostuloja painotetaan eri tavalla taajuusalueesta riippuen, sillä psykoakustisten kokeiden perusteella voidaan sanoa ITD:n, ILD:n ja äänen verhoikäyrän aikaerojen merkittävyyden vaihtelevan taajuudesta riippuen [1]. Painotus toteutettiin liittämällä jokaiseen tumakemallin ulostuloon energiasignaali ja kontrolloimalla tämän voimakkuutta taajuudesta riippuen. Lisäksi yhdistämisessä huomioidaan se, että kuuntelukokeet ovat osoittaneet jo yhden suuntavihjeen muuttamisen johtavan suuntahavainnon muuttumiseen [9]. Esimerkiksi ääni, jossa pelkästään ITD osoittaa äänen tulosuunnan olevan vasemmalla, havaitaan tulevan vasemmalta, vaikka ILD osoittaa suoraan eteenpäin. Tämän takia vihjeiden yhdistämismenetelmän toteutuksessa suositetaan laidoille osoittavia vihjeitä.

Vihjeet yhdistetään erikseen kullakin taajuuskaistalla ja aivopuoliskolla laskemalla energiapainotettu keskiarvo:

$$\hat{\gamma}_m = \sqrt[3]{\frac{e_{i,m}\gamma_{i,m}^3 + e_{j,m}\gamma_{j,m}^3 + e_{k,m}\gamma_{k,m}^3}{e_{i,m} + e_{j,m} + e_{k,m}}} - 20^\circ, \quad (1)$$

missä e on energia ja γ on vihje ja joiden alaindeksien m tarkoittaa joko vasenta tai oikeaa aivopuoliskoa ja $[i, j, k]$ ITD-, ILD- ja laajakaistaista ITD-vihjettä. Kukin vihje korotettiin kolmanteen potenssiin, jotta laidoille osoittavat vihjeet korostuisivat. Laskennassa lisättiin kaikkiin γ :n arvoihin väliaikaisesti 20° , jotta vältettiin negatiiviset arvot kuutiojuuren sisällä. Tämä kompensoitiin lopuksi vähentämällä 20° saadusta tuloksesta.

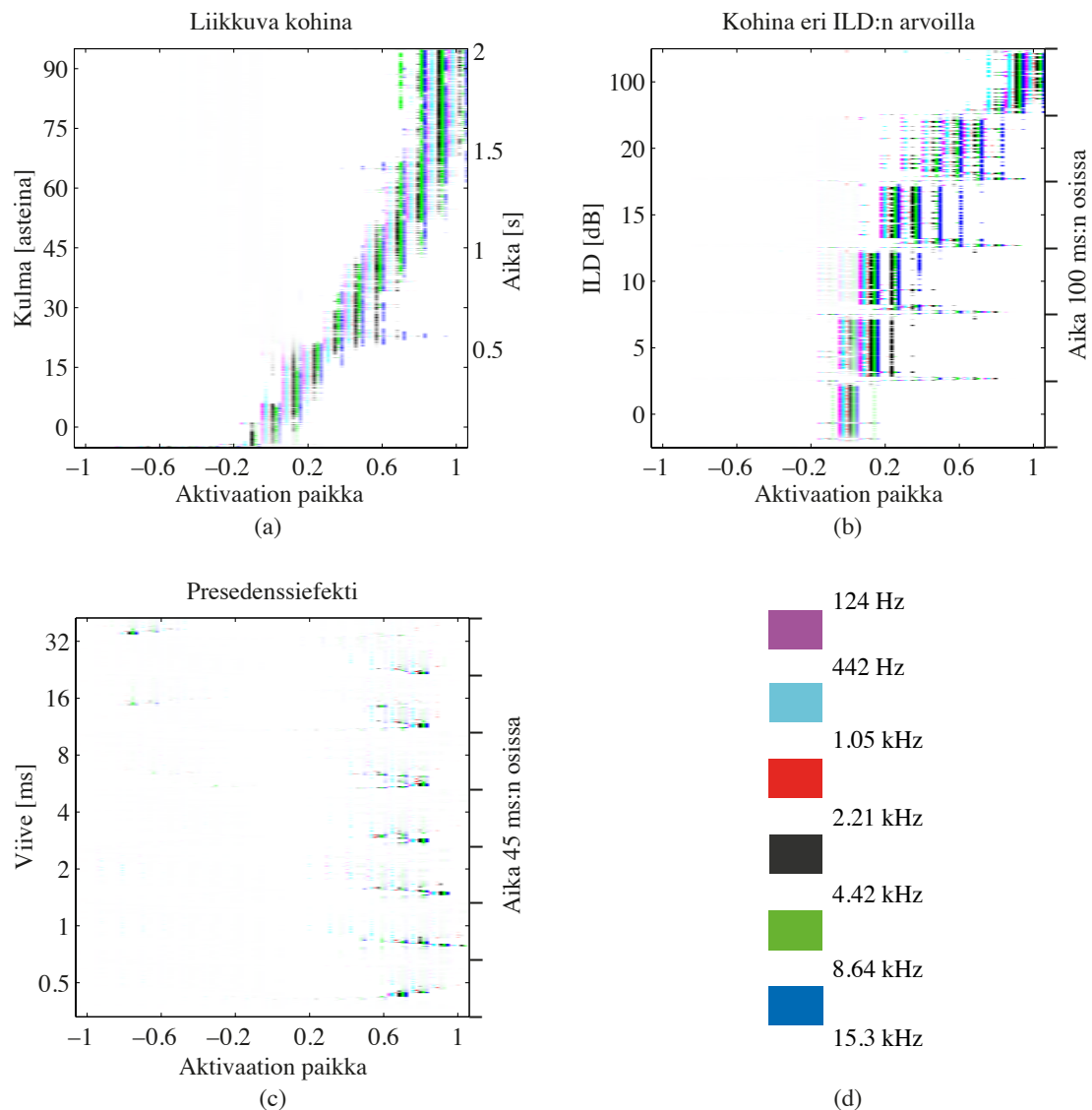
3 MENETELMÄN TOIMINNALLISUUDEN ARVIOINTI

Tässä kappaleessa esitellään muutamia testitapauksia, joiden avulla binauraalisen kuulojärjestelmän mallin ja vihjeiden yhdistämismenetelmän toimivuutta arvioitiin. Tapaukset simuloitiin vastaamaan kaksikorvaista kuuntelua käyttämällä keinopäällä mitattuja HRTF:iä, joiden avulla äänet voitiin sijoittaa eri tulokulmiin kuulokekuuntelussa. Nämä kaksikanavaiset signaalit syötettiin malliin, ja ulostulona oleva aktivaatiokartta esittää eri aivopuoliskoilla esiintyvän aktivaation ajan funktiona. Saatujen tulosten tarkkuutta arvioidaan vertaamalla niitä psykoakustisista kokeista saatuihin tuloksiin.

Aktivaatiokartalla olevan informaation voidaan ajatella esittävän hermosoluja, jotka ovat järjestäytyneet vasen-oikea -akselille niin, että kussakin kartan aktivaatiopaikassa on useita taajuusselektiivisiä hermosoluja. Visualisoinnin helpottamiseksi kartan tarkkuudeksi on asetettu kymmenen astetta ja taajuudet on jaettu kuuteen eri väri/harmaasävyillä esitettävään taajuusalueeseen, jotka on esitetty kuvassa 1(d).

Ensimmäisessä testitapauksessa testiäänenä oli pinkkiä kohinaa, jota prosessoitiin HRTF:illä niin, että ääni liikkui tasaisesti suoraan edestä vasemmalle eli nolasta 90:een asteeseen. Koko signaalin kesto oli kaksi sekuntia. Kuuntelukokeesta saatujen tulosten perusteella koehenkilöt kykenevät osoittamaan liikkuvan pinkin kohinan suunnan eri ajanhetkillä noin 12° :n tarkkuudella vaakatasossa [10]. Kuvassa 1(a) esitetyssä aktivaati-

tiokartassa nähdään, että aktivaation paikka siirtyy tasaisesti keskeltä laitaan. Huomionarvoista on, että aktivaatio siirtyy oikealle, sillä vasemmalta tulevat äänet aiheuttavat aktivaatiota oikeassa aivopuoliskossa. Edessä olevien tulokulmien osalta kaikkien taajuusalueiden aktivaation on samassa paikassa, ja laidalle päin mentäessä aktivaatio leviää jonkin verran. Suurimmilla tulokulmilla aktivaatio on laidassa useammassa aktivaatiopaikassa, mikä vastaa sitä, että ihmisen suuntahavaintotarkkuus on heikompi sivulla olevien äänten osalta kuin suoraan edessä. Yleisesti ottaen aktivaatio seuraa hyvin äänihavainnon suuntaa, ja tarkkuus vastaa kuuntelukoetuloja.



Kuva 1: Kolmen eri testitapauksen binauraaliset aktivaatiokartat: (a) HRTF-prosessoitu liikkuva kohina; (b) kohina eri ILD:n arvoilla; (c) presedenssiefekti havainnollistettuna tapauksella, jossa esitetään eri suunnista kaksi lyhyttä äänipursketta vaihtuvalla viiveen arvolla. Aktivaatio eri aivopuoliskoilla on esitetty x-akselilla ja ajan suhteen muuttuva suure y-akselilla. Eri taajuusalueet on esitetty väri/harmaasävyillä (d):n mukaisesti.

Kuten aikaisemmin mainittiin, jo yhden suuntavihjeen muuttaminen voi aiheuttaa suuntahavainnon siirtymisen. Kuuntelukokeissa on havaittu, että havainto siirtyy täysin yhdelle puolelle, kun ILD on tarpeeksi suuri, n. 10 – 20 dB henkilöstä riippuen [11]. Lisäksi äänihavainto muuttuu pistemäisestä leveäksi, kun ILD lisääntyy, mikä vaikeuttaa täsmällisen suuntahavainnon muodostumista. Toisessa testitapauksessa tutkittiin tätä ilmiötä pitämällä ITD nollassa samalla, kun ILD-vihjeen arvo vaihteli 0 dB:n ja 100 dB:n välillä niin, että kullakin arvolla esitettiin 100 ms:n näyte pinkkiä kohinaa. Kuvassa 1(b) esitetään testitapauksen aktivaatiokartta, jossa on nähtävissä, että aktivaatio siirtyy oikealle aivopuoliskolle ILD:n lisääntyessä. Erityisesti ILD:n arvoilla 15 dB ja 20 dB aktivaatio on levittäytynyt melko laajalle alueelle. 100 dB:llä aktivaatio on selkeästi laidassa. Aktivaatiota ei kuitenkaan ole keskellä kuin pienimmillä ILD-arvoilla, vaikka ITD-vihje osoittaaakin siihen suuntaan. Tämä johtuu yhdistämismenetelmän piirteestä, jonka mukaisesti suositaan laidoille osoittavia vihjeitä. Aktivaation paikka vastaa hyvin kuuntelukokeiden tuloksia lukuunottamatta sitä, että aktivaatiokartassa 20 dB:n ILD:llä aktivaatio on vielä levinnyt melko laajalle.

Kolmas testitapaus keskittyi presedenssiefektiin eli ilmiöön, jonka mukaisesti lyhyen ajan sisällä saapuvat samankaltaiset äänet havaitaan yhtenä kokonaisuutena ensimmäisen äänen suunnasta. Ilmiön ansiosta esimerkiksi luentosalissa puhuvan henkilön puhe paikallistetaan puhujan suuntaan huoneen pinnoista tulevista heijastuksista huolimatta. Testitapauksessa keskityttiin kahden eri suunnasta tulevan äänipurskeen välisen viiveen vaikutukseen jälkimmäisen äänen havaitsemisessa. Äänet prosessoitiin HRTF:illä niin, että ensimmäisen äänen tulokulma oli 45° vasemmalla ja toisen 45° oikealla. Lyhyillä viiveillä, suunnilleen alle 4 ms, havaitaan ääntä vain ensimmäisen äänen suunnasta, mutta kun viive kasvaa suuremmaksi, havaitaan kaksi ääntä [12]. Kuvan 1(c) aktivaatiokartassa on esitetty mallin ulostulo seitsemällä eri viiveen arvolla, jotka mukailevat suoritettua kuuntelukoeetta [12]. Kuvasta nähdään, että viiveillä 0.5 ms – 4 ms aktivaatiota on vain oikealla aivopuoliskolla eli ensimmäisen äänen suunnassa, ja kuuntelukokeiden tuloksia mukaillen viiveen arvolla 8 ms sekä erityisesti arvoilla 16 ms ja 32 ms myös vasemmalla on nähtävissä aktivaatiota.

4 YHTEENVETO

Artikkelissa esitettiin binauraalisen kuulojärjestelmän malliin sisältyvä keskeinen menetelmä, jossa yhdistettiin eri tumakemallien ulostuloista saatavat suuntavihjeet, jotta kokonaisuuden ulostulo vastaa psykoakustista tietoa kuulojärjestelmän toiminnasta. Menetelmässä muodostetaan kuudesta suuntavihjeestä kaksi vihjetä kullekin taajuusalueelle ja ajanhetkelle. Mallin ja menetelmän toiminnallisuutta esiteltiin aktivaatiokartoilla, joissa havainnollistettiin äänten aiheuttamaa aktivaatiota eri aivopuoliskoilla. Aktivaatiotulosten osoitettiin vastaavan psykoakustisia tuloksia. Mallin avulla voidaan tutkia kuulojärjestelmää monipuolisesti. Psykoakustisissa kokeissa voidaan käyttää herätteitä, jotka vaikuttavat tarkasti tiettyyn kuulojärjestelmän ominaisuuteen ja saatua tulosta voidaan käyttää ominaisuutta vastaavan mallin osan säätöön. Myös erilaisia kuulojärjestelmän häiriöitä voidaan tutkia simuloimalla mallilla jonkin tietyn osan toimintahäiriötä ja vertaamalla saatuja tuloksia potilaan suorituskykyyn. Lisäksi malli voi olla hyödyllinen työkalu parametrusten tilaäänentoistomenetelmien kehittämisessä, sillä mallilla voidaan arvioida menetelmän muutosten vaikutuksia toistetun äänen tilavaikutelmaan.

KIITOKSET

Tätä tutkimusta ovat tukeneet Suomen Akatemia, Walther Ahlströmin säätiö sekä Nokia Säätiö. The research leading to these results has received funding from the European Research Council under the European Community's Seventh Framework Programme (FP7/2007-2013) / ERC grant agreement No. [240453].

VIITTEET

- [1] J. Blauert. *Spatial Hearing*. The MIT Press, Cambridge, MA, USA, revised edition, 1997.
- [2] B. Grothe. Sensory systems: New roles for synaptic inhibition in sound localization. *Nat. Rev. Neurosci.*, 4:540–550, 2003.
- [3] P. Joris. Envelope coding in the lateral superior olive. II. Characteristic delays and comparison with responses in the medial superior olive. *J. Neurophysiol*, 76: 2137–2156, Oct. 1996.
- [4] M. Takanen, O. Santala, and V. Pulkki. Visualization of functional count-comparison-based binaural auditory model output. Julkaisematon käsikirjoitus, 2013.
- [5] M. Takanen, O. Santala, and V. Pulkki. Suuntakuulon toiminnallisuuden mallintaminen. In *Akustiikkapäivät 2011*, pages 69–74, Tampere, 11.-12.5. 2011.
- [6] S. Verhulst, T. Dau, and C. A. Shera. Nonlinear time-domain cochlear model for transient stimulation and human otoacoustic emission. *J. Acoust. Soc. Am.*, 132(6): 3842 – 3848, 2012.
- [7] D. McAlpine, D. Jiang, and A. R. Palmer. A neural code for low-frequency sound localization in mammals. *Nat. Neurosci.*, 4(4):396 – 401, Apr. 2001.
- [8] E. J. Macaulay, W. M. Hartmann, and B. Rakerd. The acoustical bright spot and mislocalization of tones by human listeners. *J. Acoust. Soc. Am.*, 127(3):1440 – 1449, Mar. 2010.
- [9] W. A. Yost. Lateral position of sinusoids presented with interaural intensive and temporal differences. *J. Acoust. Soc. Am.*, 70(2):397–409, Aug. 1981.
- [10] M. Gröhn, T. Lokki, and T. Takala. Static and dynamic sound source localization in virtual room. In *Proc. of the 22nd Intl. Conf. of the Audio Eng. Soc.*, Espoo, Finland, June 2002. Paper No: 213.
- [11] B. McA. Sayers. Acoustic-Image Lateralization Judgments with Binaural Tones. *J. Acoust. Soc. Am.*, 36(5):923 – 926, May 1964.
- [12] R. L. Freyman, R. K. Clifton, and R. Y. Litovsky. Dynamic processes in the precedence effect. *J. Acoust. Soc. Am.*, 90(2):874–883, Aug. 1991.