

SUUNTAKUULON TOIMINNALLISUUDEN MALLINTAMINEN NEUROFYSIOLOGISELLA TASOLLA

Marko Takanen, Olli Santala, Ville Pulkki

Aalto-yliopisto, Sähkötekniikan korkeakoulu, Signaalinkäsittelyn ja akustiikan laitos
PL 13000, 00076 Aalto
marko.takanen@aalto.fi

1 JOHDANTO

Auditorisia malleja on kehitetty 1980-luvulta lähtien mallintamaan ihmiskuulon toimintaa laskennallisten operaatioiden avulla [1]. Tämän kehitystyön tavoitteena on luoda malleja, joiden antamat arviot jostakin äänitapahtumasta vastaavat ihmisen tekemää havaintoa. Tavoitteen saavuttamisen jälkeen voitaisiin käyttää auditorista mallia arvioimaan tutkittavia ääniä ilman aikaa vievien kuuntelukokeiden järjestämistä ja/tai vaarantamatta koehenkilön kuuloa.

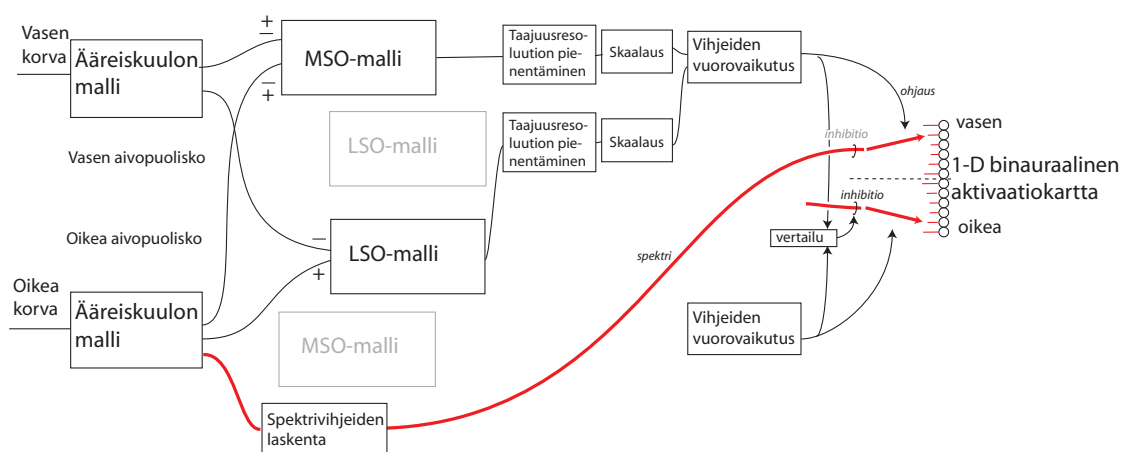
Koko kuuloaistimuksen taustalla olevan prosessoinnin mallintaminen on kuitenkin osoittautunut hyvin haastavaksi tehtäväksi, minkä johdosta auditoristen mallien kehittämisessä on yleensä keskitytty johonkin yksittäiseen osa-alueeseen kuten äänekkyuden havainnoinnin mallintamiseen [2]. Yksi eniten tutkituista osa-alueista on ollut suuntakuulo, jota on perinteisesti mallinnettu käyttäen korviin saapuvien signaalien väliseen ristikorrelaatioon pohjautuvaa laskentaa [3]. Viimeisimmät auditoriset mallit pyrkivät kuitenkin kuulojärjestelmän tarkempaan kuvaukseen mallintaen myös kuulojärjestelmän neurofysiologisia osia toiminnallisella tasolla.

Tässä artikkelissa esitellään viimeksi mainittua periaatetta noudattava malli, jossa on mallinnettu kuulojärjestelmän osien toimintaa ulkokorvalta alemmalle nelikukkulatmakkeelle (Inferior Colliculus, IC) asti. Esiteltävän mallin rakenteen ja toiminnallisuuden taustalla ovat tiedot psykoakustiikasta ja ihmisen kuulojärjestelmän neurofysiologiasta sekä eläinkokeista saadut tulokset. Artikkelissa esitellään, kuinka malli muodostaa binauraalisesta signaalista prosessoinnin lopputuloksena binauraalisen aktivaatiokartan, joka osoittaa äänten tulosuunnat taajuuksittain kullakin ajanhetkellä. Lisäksi artikkelissa osoitetaan vastaavuudet mallin ulostulon ja ihmisten tekemien kuulohavaintojen välillä.

2 BINAURAALINEN AUDITORINEN MALLI

Mallin yleisrakenne on esitetty kuvassa 1. Mallin sisääntulosignaalina oleva binauraalinen signaali prosessoidaan ensin vasemman ja oikean korvan ääreiskuulon mallien läpi, jonka jälkeen näiden kahden mallin ulostulot johdetaan keskimmäisen ylemmän oliivitumakkeen (Medial Superior Olive, MSO) ja ja laitimmaisen ylemmän oliivitumakkeen (Lateral Superior Olive, LSO) malleihin. Nämä MSO:n ja LSO:n mallit suorittavat binauraalisten vihjeiden laskennan ääreiskuulon mallien ulostuloista. Saatujen

binauraalisten vihjeiden ja kahden ääreiskuulon mallin ulostuloista laskettujen spektrivihjeiden avulla muodostetaan sitten binauraalinen aktivaatiokartta, joka havainnollistaa spektrivihjeet yksiulotteiselle akselille vasemmalta oikealle. Huomioitavaa on se, että ääreiskuulon, MSO:n ja LSO:n mallit pohjautuvat neuroanatomiseen ja neurofysiologiseen tietoon [4], mutta binauraalisen kartan muodostuksen taustalla on psykoakustinen informaatio.



Kuva 1: Mallin rakenteen kuvaus, jossa on selvytyden vuoksi esitettyä vain vasemmalle aivopuoliskolle saapuvien signaalien muodostumiset.

Binauraalisten vihjeiden laskenta mallissa pohjautuu niin sanottuun count-comparison-periaatteeseen [3] ja on toteutettu suurilta osin artikkelissa [5] esittyjä toteutuksia noudattaen. Seuraavat kappaleet esittelevät mallin eri osien toimintaa yleisellä tasolla.

Ääreiskuulon mallin ensimmäisessä vaiheessa sisääntulevaa signaalia suodatetaan 141-kaistaisella gammatone-suodinpankilla, joka mallintaa basilaarikalvon taajuusselektiivisyyttä. Seuravassa vaiheessa suodinpankin ulostuloista muodostetaan impulssijonoja, joissa yksittäiset impulssit ovat voimakkuudeltaan signaalin tasosta riippuvia ja hiukan levinneitä. Tämä vaihe mallintaa simpukkatumakkeessa tapahtuvaa vaihelukittautumista. Viimeisessä vaiheessa kompensoidaan suodinpankin viiveet ja edellisen vaiheen vaikutus signaalin tasoon.

MSO:n mallin tehtävänä on mallintaa kyseisen aivopuoliskon MSO:n neuronien yhteistä aktiivisuutta molempien aivopuoliskojojen simpukkatumakkeista saapuville herätteille. Yleisesti ottaen voidaan todeta MSO:n olevan herkkä korvienvälisen aikaeron (ITD:n) suhteen [4]. MSO:n mallintaminen on toteutettu viisivaiheisena prosessina, jonka ensimmäisessä vaiheessa tulosignaalien arvoista vähennetään lyhyen aikavälin (10 ms) keskiarvo yli 1500 Hz:n keskitaajuutta vastaavilla kaistoilla impulssien terävöittämiseksi näillä kaistoilla. Toisessa vaiheessa ipsi- ja kontralateraaleista signaaleista lasketaan yhteisaktiivisuuden mallintamiseksi geometriset keskiarvot vierekkäisten kaistojen yli keskiarvoistaen. Kolmannessa vaiheessa kontralateraalilta puolelta saapuva signaali konvoloidaan taajuusriippuvalla signaalilla, jotta MSO:n mallin vaste on suurimmillaan 45° vaihe-eroa vastaavilla viiveillä. Samanaikaisesti ipsilateraalien puolen signaali normalisoidaan välille [0, 1]. Näin saadut signaalit kerrotaan sitten keskenään neljännessä vaiheessa binauraalisen vuorovaikutuksen mallintamiseksi ja lopulta saatu tulos normalisoidaan välille [0, 1] kontralateraalin sisääntulon avulla.

LSO:n on vastaavasti todettu olevan herkkä korvienvälisen tasoeron (ILD:n) suhteen, mutta myös ITD:n suhteen pienillä taajuuksilla [4]. LSO:n malli mallintaa LSO:n neurotien yhteisaktiivisuutta neljävaiheisen prosessin avulla. Ensimmäiseksi eri kaistojen yhteisvaikutusta mallinnetaan, kuten MSO:n mallissakin, geometrisen keskiarvon laskennan avulla. Toisessa vaiheessa mallinnetaan ajallista integraatiota suodattamalla kontraja ipsilateraalien puolen signaaleja ensimmäisen asteen IIR-suodattimella. Kolmannessa vaiheessa suodatettu ipsilateraali signaali jaetaan kontralateraalilla tasoeron laskemiseksi, jonka jälkeen tasoero normalisoidaan välille [0,1] siten, että se saturoituu arvoon 1 yli 18 dB:n eroa vastaavilla arvoilla. Näin saadun signaalin lyhytkestoisia piikkejä korostetaan sitten viimeisessä vaiheessa painotetun keskiarvon laskennan avulla.

Binauraalinen aktivaatiokartta muodostetaan MSO:n ja LSO:n mallien ulostuloista saatavien suuntavihjeiden ja ääreiskuulon mallien ulostuloista laskettavien spektrivihjeiden avulla. Edellä mainitut vihjeet kuvaavat IC-tasolla käytettävissä olevaa informaatiota. Kartan tavoitteena on havainnollistaa akustinen ympäristö samalla tavalla kuin ihmisen kuulohavainnossa hahmotetaan äänikentässä olevia objekteja sisäisellä esitystavalla.

Varsinaisen kartan muodostamista ennen suoritetaan kuitenkin muutamia valmistelevia operaatioita, joista ensimmäisessä lasketaan spektrivihjeet ääreiskuulon mallien ulostuloista kymmenkantaisen logaritmin, liukuvan keskiarvon (10 ms) laskennan ja normalisoinnin avulla. Normalisoinnissa käytetään referenssinä 60 dB äänenpainetasolla toistetun pinkin kohinan avulla laskettuja arvoja. Seuraavana operaationa on aineiston taajuusresoluution pienentäminen kartan selkeyden vuoksi. Tämä operaatio toteutetaan jakamalla suunta- ja spektrivihjeet kuuteen eri taajuusalueeseen niiden kaistojen keski- ja taajuuksien perusteella taulukon 1 rajataajuuksien mukaan ja laskemalla vihjeiden keskiarvot kyseisillä taajuusalueilla. Tämän jälkeen suuntavihjeet skaalataan käyttäen vertaisarvoja, jotka laskettiin käyttäen vapaan kentän siirtofunktioita pistemäisiltä äänilähteiltä tärykalvolle (head-related transfer function, HRTF). Skaalaus tehdään, jotta eri taajuusalueiden vihjeet ovat samankaltaisia laajakaistaisen äänilähteen tapauksessa. Viimeisessä operaatiossa MSO:n ja LSO:n välinen vuorovaikutus mallinnetaan yhdistämällä niiden skaalatut suuntavihjeet yhdeksi suuntavihjeeksi kummallakin aivopuoliskolla. Yhdistäminen tapahtuu valitsemalla kullakin ajanhetkellä ja taajuusalueella edellämmainituista arvoista suurempi.

Taulukko 1: *Eri taajuusalueiden alarajataajuudet ja taajuusalueita vastaavat värit.*

Taajuus	124 Hz	337 Hz	811 Hz	1,56 kHz	2,84 kHz	5,04 kHz
Väri	sininen	violetti	punainen	keltainen	vihreä	syaani

Binauraalisen aktivaatiokartan muodostamisessa suuntavihjeet ohjaavat spektrivihjeet vasemmalta oikealle ulottuvalle akselille. Näin ollen kartta osoittaa äänilähteen sijainnin lisäksi myös sen taajuussisällön, sillä värit kartalla ovat sitä kirkkaampia, mitä enemmän energiaa kullakin taajuusalueella on. Puoltenvälisen inhibition mallintamisen seurauksena toiselle aivopuoliskolle osoittavat suuntavihjeet kuitenkin heikentävät toiselle aivopuoliskolle meneviä spektrivihjeitä. Inhibition mallintaminen on toteutettu vertaamalla kyseisen puolen suuntavihjetä toisen puolen suuntavihjeeseen. Mikäli näiden vihjeiden

väläinen erotus on negatiivinen, kerrotaan kyseisen aivopuoliskon spektrivihjeitä kertoimella, joka on sitä lähempänä nollaa, mitä pienempi erotuksen tulos on.

3 TOIMINNALLISUUDEN ARVIOINTI

Kehitetyn mallin toiminnallisuuden arvioimiseksi luotiin kolme erilaista testisignaalia, jotka prosessoitiin mallin avulla. Testisignaalien sisältämät äänilähteet simuloitiin piste-mäisinä äänilähteinä prosessoimalla niiden signaalit mitatuilla HRTF:illä (Cortex MK II -keinopää). Tämän jälkeen mallin toiminnallisuutta arvioitiin tarkastelemalla visuaalisesti mallin ulostulonaan antamia binauraalisia aktivaatiokarttoja.

Testisignaalit valittiin sen perusteella, että niiden havaitsemista on tutkittu kuuntelukokeissa (esim. [6, 7]) ja että ne edustavat erilaisia tilanteita suuntakuulon kannalta. Ensimmäisenä testisignaalina käytettiin signaalia, jossa yksittäinen äänilähde liikkuu simuloituun kuuntelijaan nähden kymmenen asteen askelin kulmien 0° ja 90° välillä. Äänilähteen tuottamana signaalina käytettiin pinkkiä kohinaa.

Toisen testisignaalin simuloidussa tilanteessa on kaksi samanaikaista puhujaa, jotka sijaitsevat simuloituun kuuntelijaan nähden $\pm 30^\circ$ kulmissa. Toinen puhujista on nainen, joka lausuu lauseen "*in arithmetic, infinitely many numbers*", ja toinen puhujista on mies, joka lausuu lauseen "*three four five*".

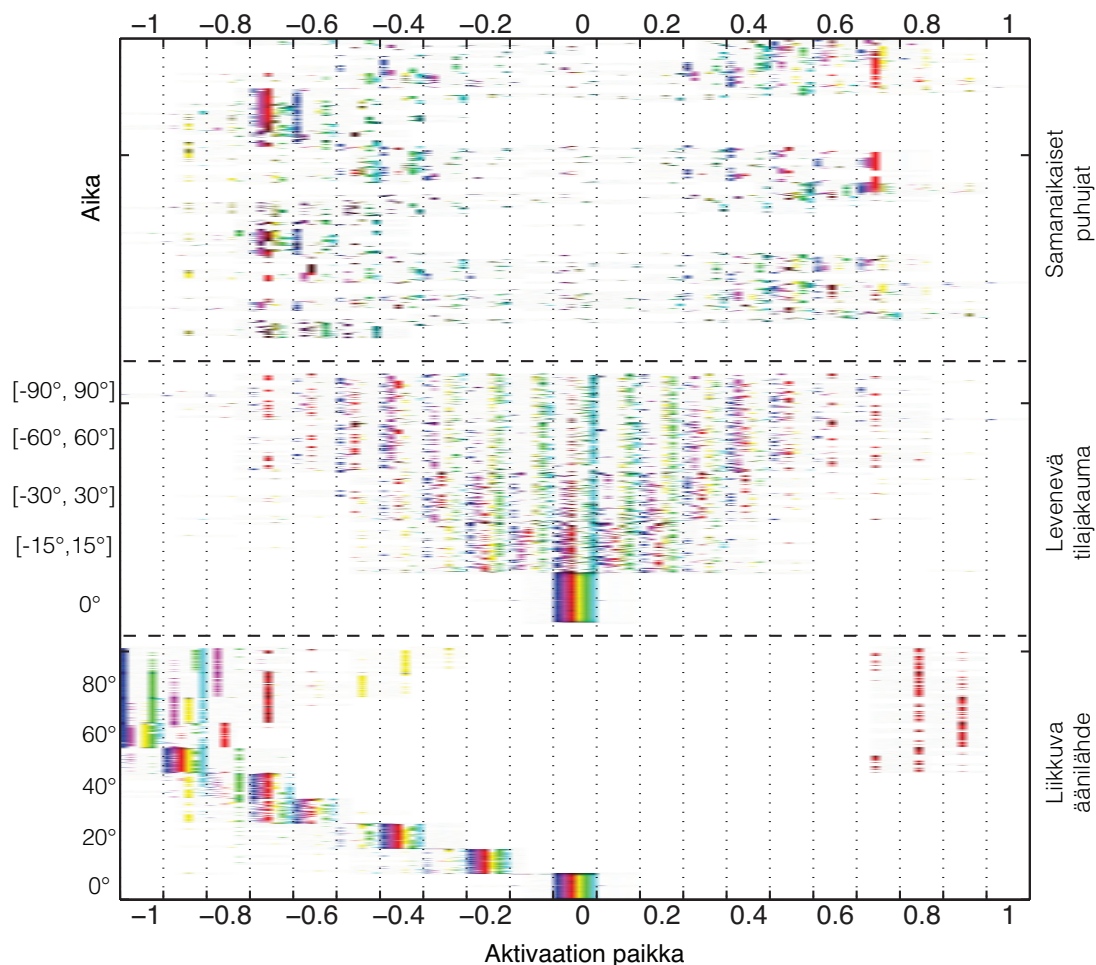
Kolmannen testisignaalin tilanteessa on yhtäaikaisesti läsnä useita samankaltaisia äänilähteitä. Äänilähteiden tuottamina signaaleina käytettiin korreloimattomia pinkin kohinan näytteitä. Äänilähteiden lukumäärä kasvoi asteittain testisignaalin aikana yhdestä kolmeentoista siten, että äänilähteet olivat jakautuneet tasaisesti 15° välein simuloitun kuuntelijan vasemmalle ja oikealle puolelle.

3.1 Tulokset ja analyysi

Mallin ulostulonaan palauttavat binauraaliset aktivaatiokartat käytetyille testisignaalille ovat esitettynä kuvassa 2. Alimpana esitetyn yksittäisen äänilähteen aktivaatiokartasta nähdään, kuinka aktivaation paikka siirtyy vasemmalle seuraten näin äänilähteen vaiheittaista siirtymistä simuloitun kuuntelijan oikealle puolelle. Aktivaatiokartasta huomataan myös, että mallin erottelukyky ja tarkkuus on parhaimmillaan äänilähteen ollessa edessä, alle 30° kulmassa kuuntelijaan nähden. Vastaavasti erottelukyky on huonoimmillaan äänilähteen ollessa kuuntelijan sivulla. Nämä tulokset korreloivat hyvin ihmisille tehtyjen kuuntelukokeiden tulosten kanssa [6].

Kuvassa ylimpänä esitetystä samanaikaisten puhujien binauraalisesta aktivaatiokartasta on erotettavissa yksi äänilähde kummallakin aivopuoliskolla. Esimerkiksi miehen lausumat sanat ovat selvästi nähtävissä oikean aivopuoliskon aktivaatiossa. Näin ollen voidaan todeta mallin kykenevän toimimaan myös samanaikaisten äänilähteiden tapauksessa.

Levenevän, useasta samankaltaisesta äänilähteestä koostuvan, tilajakauman aktivaatiokartasta (kuvassa keskimmäisenä) nähdään, että aktivaatio levenee yhä laajemmalle alueelle tilajakauman leventyessä. Aktivaatio ei kuitenkaan ylety aivan niin laitaan kum-



Kuva 2: Binauraalinen aktivaatiokartta, jossa eri värit kuvaavat eri taajuusalueita taulukon 1 mukaisesti. Kartassa positiiviset aktivaation paikan arvot viittaavat aktivaatioon oikealla aivopuoliskolla.

mallakaan aivopuoliskolla tilajakauman ollessa leveimmillään kuin olisi kuuntelukokeiden tulosten [7] ja yksittäisen äänilähteen aiheuttaman aktivaatiokartan perusteella odotettavissa. Näin ollen voidaan todeta mallin toiminnallisuuden olevan lupaava myös useiden samankaltaisten äänilähteiden tapauksessa mutta vaativan myös joitakin muutoksia, jotta saavutetaan vielä parempi vastaavuus kuuntelukokeiden tulosten kanssa.

4 DISKUSSIO

Tässä artikkelissa esiteltiin perinteisistä ristikorrelaatiomalleista poikkeava auditorinen malli, joka pyrkii mallintamaan ihmisen kuulojärjestelmän toiminnallisuutta neurofysiologisella tasolla laskennallisten operaatioiden avulla. Artikkelissa esitellyt tulokset osoittavat mallin toiminnallisuuden sekä yksittäisen että samanaikaisten äänilähteiden tapauksissa. Myös useasta samankaltaisesta äänilähteestä koostuvan tilajakauman ta-

pauksessa mallin tulokset ovat lupaavia. Viimeksi mainittuun tapaukseen liittyvän kehitystyön lisäksi mallin kehitystyössä tullaan jatkossa tutkimaan mallin toiminnallisuutta kaiuntaisissa ympäristöissä.

4.1 Kiitokset

Tätä tutkimusta on rahoittanut Suomen Akatemia. 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] KARJALAINEN M, A new auditory model for the evaluation of quality measurements and spatial hearing studies, in *IEEE on Acoust., Speech and Sig. Proc.*, pages 608 – 611, Tampa, USA, 1985.
- [2] MOORE B, GLASBERG B, & BAER T, A model for the prediction of thresholds, loudness and partial loudness, *J. Audio Eng. Soc.*, **45**(1998) 4, 224 – 237.
- [3] COLBURN H S, Computational models of binaural processing, in H Hawkins & T McMullen, editors, *Springer Handbook on Auditory Research*, Springer, New York, 1996.
- [4] GROETHE B, New roles for synaptic inhibition in sound localization, *Nat. Reviews Neurosci.*, **4**(2003), 540 – 550.
- [5] PULKKI V & HIRVONEN T, Functional count-comparison model for binaural decoding, *Acta Acustica united with Acustica*, **95**(2009), 883 – 900.
- [6] BLAUERT J, *Spatial Hearing*, MIT Press, Cambridge, MA, USA, revised edition, 1997.
- [7] SANTALA O & PULKKI V, Directional perception of distributed sound sources, *J. Acoust. Soc. Am.*, **129**(2011) 3, 1522 – 1530.