

ÄLYPUHELIMILLA TALLENNETUN STEREOÄÄNEN LAATU

Antti Kelloniemi, Kalle Koivuniemi, Jarkko Punnonen, Sari Suomela

Nokia Oyj
Smart Devices
PL 226
00045 Nokia Group
antti.kelloniemi@nokia.com, kalle.koivuniemi@nokia.com,
ext-jarkko.punnonen@nokia.com, sari.suomela@nokia.com

Tiivistelmä

Uusimmilla älypuhelimilla videoitaessa on mahdollista tallentaa stereofonista ääntä. Kuuntelukokeiden tuloksista nähdään, että parhailla älypuhelimilla kyetään niiden pienestä koosta huolimatta tallentamaan varsin hyvälaatuista ääntä, mutta kokemukset äänensävyn ja suunnan virheistä vaikuttavat kokonaislaatuun.

1 JOHDANTO

Älypuhelimia käytetään nykyään yhä enemmän myös videokameraoina. Edistyneimpien laitteiden ominaisuutena stereofonisen äänen tallentaminen kahta mikrofonia käyttäen on yleistymässä. Laitteet poikkeavat kuitenkin muotoilultaan perinteisistä videokameraista, eivätkä mikrofonit usein ole tätä käyttötarkoitusta varten optimaalisissa paikoissa, mistä johtuen niiden taajuusvasteissa ja suuntakuvioissa voi olla merkittäviä eroja.

Stereofonisen signaalin äänittämiseksi on kehitetty lukuisia erilaisia mikrofoni tekniikoita, joissa mikrofonien suuntakuviot ja keskinäinen sijoittelu vaihtelevat. Eero Aro luokittelee yleisimmät stereofoniset mikrofoni tekniikat mikrofonien keskinäisen etäisyyden perusteella kolmeen ryhmään [1]: erillismikrofoni tekniikat (mikrofonien etäisyydet ≥ 30 cm), koinsidenssitekniikat (mikrofonit mahdollisimman lähellä toisiaan) ja lähes koinsidentti mikrofonisijoittelu (mikrofonien etäisyydet noin 3-35 cm).

Nykyaikaiset älypuhelimet voidaan kokonsa puolesta sijoittaa tämän jaottelun kahteen jälkimmäiseen ryhmään. Mikrofonien lähes tai kokonaan koinsidentista sijoittelusta johtuen stereosignaalien välille ei varsinkaan matalilla äänillä synny äänitysvaiheessa suuria aika- eikä vaihe-eroja. Koska älypuhelimissa käytetään lähes ainoastaan pallokuvioisia mikrofoneja, ei äänitysjärjestelmää voida optimoida myöskään erilaisilla mikrofonien suuntakuvioilla ja niiden suuntaamisella. Käytännössä jäljelle jää signaalinkäsittelyn tarjoamat mahdollisuudet ja algoritmien avulla paranneltu stereokuva.

Stereoäänitteen toisto voidaan yleisesti jakaa kaiutin- ja kuuloketoistoon. Näistä kuuloketoisto hyötyy stereokuvan muodostumisessa suoraan siitä, että ihmisen korvien välinen etäisyys on likimain sama kuin älypuhelimien pituus. Tilantuntua kasvattavat tässä tapauksessa kanavien väliset pienet aika- ja vaihe-erot siitä huolimatta, että tallennettu ääni on väritynnyt erilaiseksi kuin miltä se olisi kuulostanut paikan päällä kuulutuna [2]. Voidaan siis saavuttaa etua sijoittamalla mikrofonit puhelimen päihin, mikä

onkin hyvin tyypillistä nykyisissä puhelimissa. Kuulokekuuntelussa vasemmalle ja oikealle korvalle toistettavat signaalit pysyvät erotettuina toisistaan toisin kuin kaiutintoistossa, johon keskityimme tässä tutkimuksessamme. Kaiuttimilla kuunneltaessa käsittelemättömän stereoäänen lähde paikallistuu kuuloketoistosta poiketen kuuntelijan pään ulkopuolelle, tyypillisesti kaiuttimien väliin.

2 TUTKIMUKSEN KOHDE

Tutkimme neljän älypuhelimien (laitteet A-D) videotallenteiden stereoäänen laatua vertaamalla niitä kuuntelukokeissa ammattilaitaisen kannettavan stereotallentimen ääneen (laite E) sekä alkuperäisiin ääninäytteisiin. Selvitimme, kuinka hyvin tallennetuissa signaaleissa säilyy äänen koettu tulosuunta ja millainen vaikutus taajuusvasteen ja äänen suunnan virheillä on äänen koettuun kokonaislaatuun. Tutkittujen laitteiden mikrofonien suuntakuviot ja taajuusvasteet etusektorilla on esitetty kuvassa 1.

3 TUTKIMUSMENETELMÄ

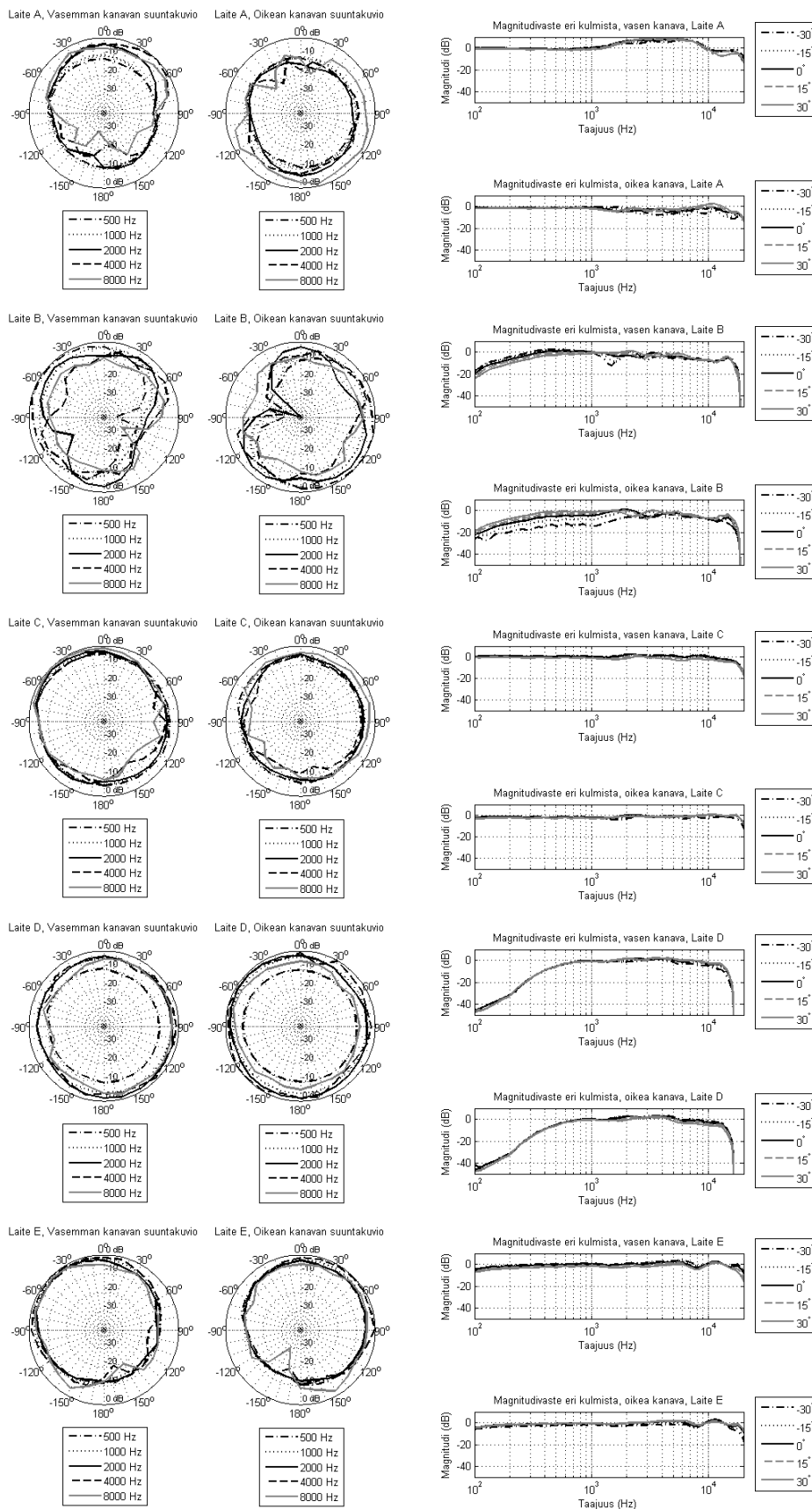
Kuuntelukokeessa soitettiin ääniä, jotka on vastaanotettu tallentimen etupuolelta. Näytteet olivat pinkkiä kohinaa ja suomenkielistä miehen puhetta. Kuuntelut suoritettiin soveltuvilta osin ITU-suosituksen mukaisesti [3].

Ensimmäisessä kokeessa soitettiin alkuperäiset äänet 15° välein sijoitetuista, näkyviltä piilotetuista viidestä kaiuttimesta $\pm 30^\circ$ sektorin sisältä. Vastaavat tallenteet sekä vastaaviin suuntiin amplitudipanoroinnilla asetetut äänet toistettiin stereokaiuttimista. Puhenäytteet ja kohinäytteet soitettiin omina sarjoinaan arvotussa järjestyksessä. Kuuntelijoita pyydettiin piirtämään havaitsemansa äänen tulosuunta lomakkeelle, johon oli merkitty kaarelle numerot 1-48. Samat numerot olivat kaarella paperilapuilla kuuntelijan ympärillä. Kukin numero vastasi noin 4° levyistä sektoria. Piirroksista tulkittiin kullakin laitteella tallennetun äänilähteen suunta ja leveys, joita verrattiin vastaavalle alkuperäiselle äänelle annettuihin tuloksiin. Kokeeseen osallistui 14 naista ja 16 miestä, iältään 21-47 vuotta. Kaksi kuuntelijoista työskentelee äänen parissa ja 15 harrastaa jotain musiikkiin tai ääneen liittyvää.

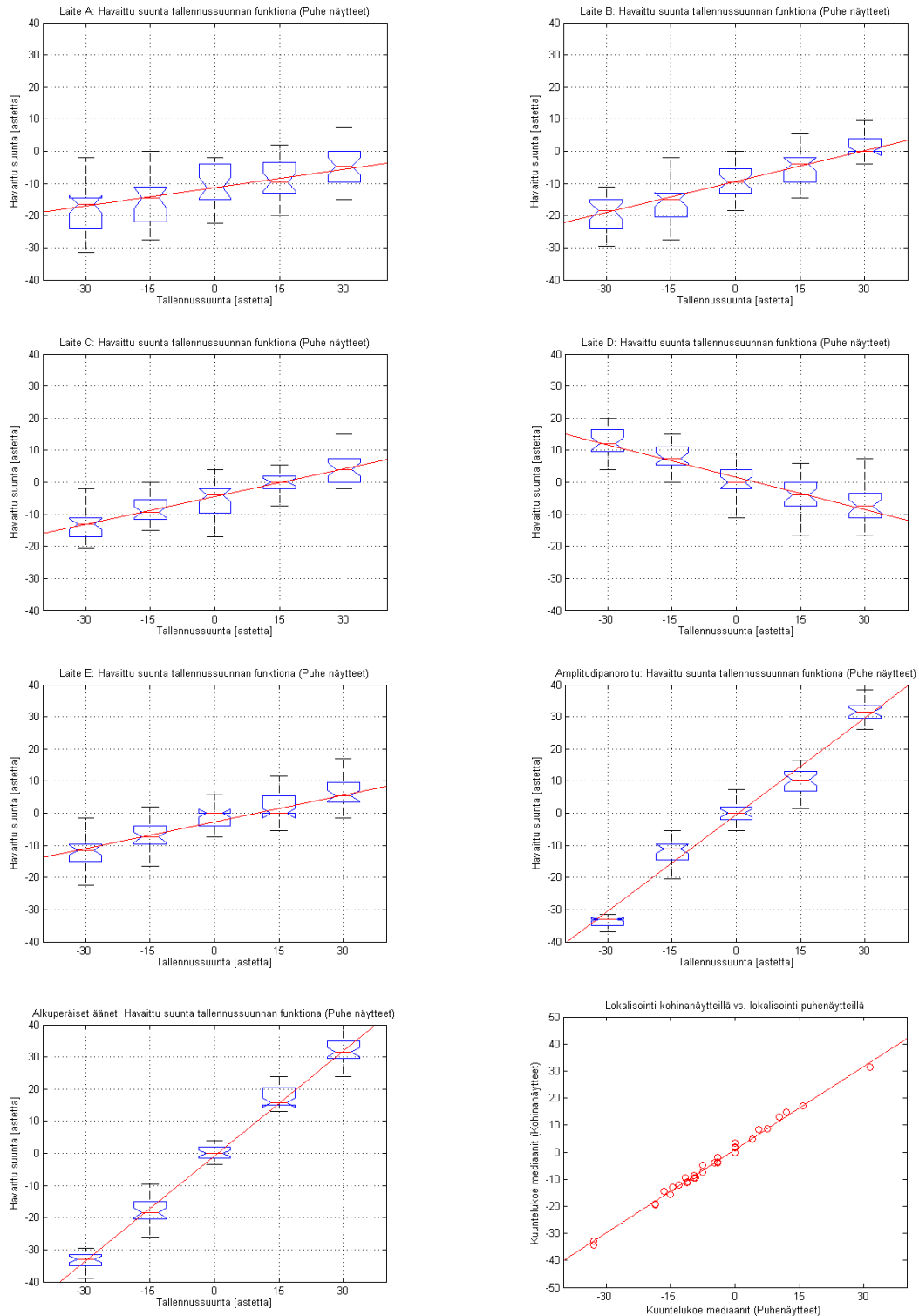
Toisessa kuuntelukokeessa soitettiin satunnaisessa järjestyksessä yksittäisten kanavien ääniä suoraan kuuntelijan edestä ja pyydettiin arvioimaan vain äänenväriin eroavuuksia verrattuna referenssiääneen, joka oli vastaava alkuperäinen ääni.

Kolmannessa kokeessa soitettiin jälleen tallennetut stereoäänet kahdesta kaiuttimesta ja pyydettiin kuuntelijoita antamaan arvio äänen kokonaislaadusta verrattuna referenssiääneen ja huomioimaan äänten sävyerojen lisäksi myös äänen tulosuunnan. Referenssiääni toistettiin tässä kokeessa amplitudipanoroituna äänen alkuperäistä suuntaa vastaavalla tavalla, minkä tarkoituksena oli edustaa optimaalista stereotallennusta.

Toiseen ja kolmanteen kokeeseen osallistui 20 miestä ja 8 naista, iältään 25-57 vuotta. Kuuntelijoista 6 työskentelee audion parissa ja 11 harrastaa ääneen tai musiikkiin liittyviä asioita. Näissä kokeissa kuuntelijoilla oli käytössään tietokoneen käyttöliittymä, jonka avulla he voivat toistaa ääninäytteitä ja referenssiääntä sekä antaa arvosanansa liukuvalitsimilla. Käyttöliittymässä oli toistettavissa yhtäaikaaisesti aina referenssiääni ja yksi kullakin laitteella tallennettu ääninäyte.



Kuva 1: Tutkittujen laitteiden vasemman ja oikean kanavan suuntakuviot ja magnitudivasteet äänille, jotka on tallennettu 15 asteen välein etusektorilta.



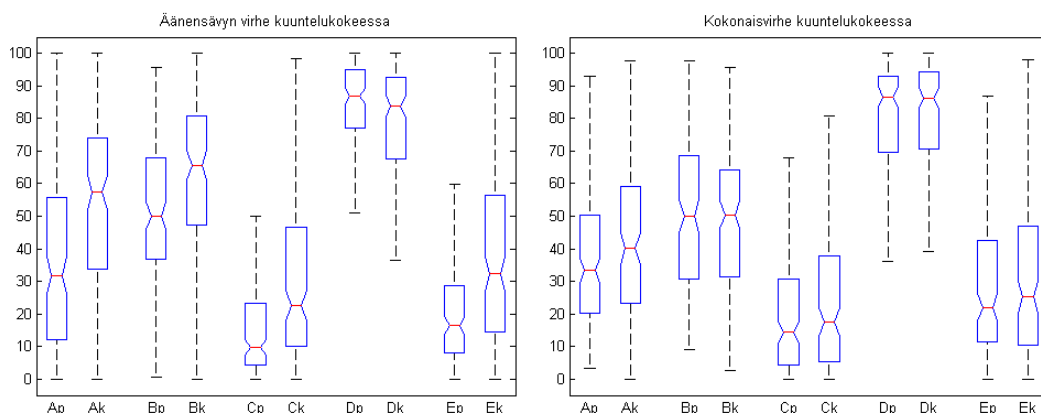
Kuva 2: Äänen tulosuunnan kuuntelukokeen tulokset puhenäytteellä kullekin viidelle laitteelle, amplitudipanoroitulle äänelle ja alkuperäisille ääninäytteille sekä alhaalla oikealla puhe- ja kohinanäytteillä kuultujen äänen tulosuuntien mediaanien korrelaatio. Kunkin kulman tuloksista piirretyn laatikon keskiviiva esittää kuuntelukokeen tulosten mediaania, laatikon ylä- ja alareunat 25% ja 75% fraktiilia ja katkoviivat tulosten hajontaa. Laatikon kapeampi osa mediaani ympärillä osoittaa mediaanin 95% luottamusväliä.

4 KOKEIDEN TULOKSET

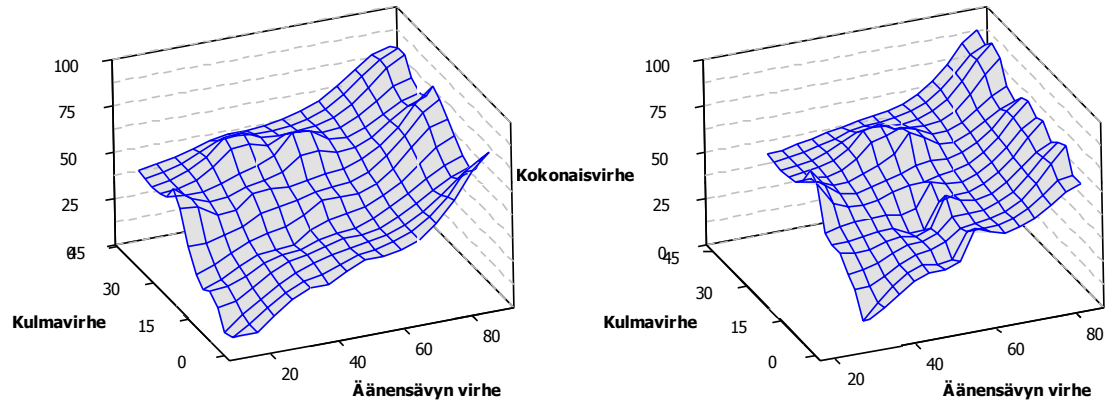
Ensimmäisen kokeen tuloksista kuvassa 2 nähdään, että laitteissa oli merkittäviä eroja äänen suunnan tallennuksen tarkkuudessa. Kanavatasapaino oli useissa laitteissa kallistunut vasemmalle, mikä johtunee mikrofoniin sijoittelusta älypuhelimissa. Oikeanpuoleinen mikrofoni on usein sijoitettu puhelimen näytön puolelle, optimoimaan suuntaavuutta puhekäytössä, mutta tällainen sijoitus heikentää stereotallennuksen laatua. Kuvassa on esitetty puhenäytteellä saadut tulokset ja puhe- ja kohinanäytteiden saamien arvioiden mediaanien suhde, mistä nähdään että näissä ei ollut merkittävää eroa. Puhenäytteillä kuuntelijat olivat äänen tulosuunnasta enemmän yksimielisiä, kohinanäytteillä vastausten hajonta oli hieman suurempaa.

Kuulluksi tulosuunnaksi tulkittiin kuuntelijan merkitsemien numeroiden keskiarvo. Kuultujen tulosuuntien mediaanien luottamusvälit menevät päällekkäin laitteen A tuloksissa ja lisäksi laitteilla B, D ja E kahden suunnan välillä, kolmen muun suunnan erottuessa selvästi niistä ja toisistaan. Äänen tulosuunnan erot olivat siis tallenteista pääosin selvästi havaittavissa. Laitteiden A ja B tuloksista nähdään, että kaikki etusektorilta tallennetut äänet kuultiin keskilinjalta tai sen vasemmalta puolelta. Laitteiden C-E tallenteissa äänilähteet kuulostivat sijaitsevan melko tasaisesti keskilinjalla molemmilla puolilla, joskin vain ± 15 asteen välillä. Amplitudipanoroidut ja alkuperäiset äänet sen sijaan leviävät jopa hieman yli ± 30 asteen alueelle. Lisäksi tuloksista havaittiin, että laitteen D kanavat olivat väärin päin, mitä ei voinut päätellä sen käyttöliittymästä. Tuloksia analysoitaessa todettiin lisäksi, että kaikissa stereokaiuttimilla toistetuissa äänissä äänilähteen koettu leveys riippui vain kuuntelijasta, tallentimien välille ei löydetty eroa. Yksittäisillä kaiuttimilla toistetut referenssiäänet kuultiin kapeampina kuin stereokaiuttimista toistetut tallenteet ja amplitudipanoroidut referenssiäänet.

Toisen ja kolmannen kuuntelukokeen tuloksista kuvassa 3 nähdään, että eri laitteiden tallenteissa oli merkittäviä äänensävyllisiä eroja, mikä oli ennakoitavissa kuvan 1 taajuusvasteita vertailemalla. Toisessa kokeessa kohinanäytteet saivat suuremman virhearvon kuin puhenäytteet muilla paitsi laitteella D. Kolmannen kokeen laitekohtaiset tulokset ovat puolestaan samankaltaiset ääninäytteestä riippumatta.



Kuva 3: Äänensävyn ja kokonaislaadun virheet puheella (p) ja kohinalla (k) kullekin laitteelle



Kuva 4: Kolmen kuuntelukokeen tulosten mediaanit puhenäytteellä (vasen) ja kohinalla (oikea)

Kun tulokset esitetään yhdessä toistosuunnan virheen itseisarvon kanssa, joka saatiin ensimmäisestä kuuntelukokeesta, havaitaan kuvasta 4, että toistosuunnan ja taajuusvasteen oikeellisuus ovat molemmat tärkeitä kokonaislaadun tekijöitä. Tätä kuvaa tuloksiin sovitettu yhtälö,

$$E_{\text{kok}} = -23,3 + 1,10 E_{\text{sävy}} + 2,14 E_{\text{kulma}} - 0,0201 E_{\text{sävy}} E_{\text{kulma}}, \quad (1)$$

jossa E_{kok} ja $E_{\text{sävy}}$ ovat kokonaisvirhe ja äänensävy virhe asteikolla 0-100 ja E_{kulma} on kulman virhe asteina.

5 JOHTOPÄÄTÖKSET

Äänensävy virheiden lisäksi äänen tulokulman virheellä näyttää olevan merkittävä vaikutus kokonaisvirheen määrittämisessä. Siksi stereotallennuksen toteutukseen, kuten mikrofonien suuntakuvioihin ja kanavatasapainoon, on syytä kiinnittää huomiota tutkitun kaltaisten laitteiden suunnittelussa. Kuuntelukokeen tuloksista nähdään, että laitteiden kyvyssä tallentaa äänen tulosuunta siten, että se voidaan toistaa stereokaiutinparilla, on merkittäviä eroja. Heikoimmassa tapauksissa kaikki etusektorilta tallennetut äänet paikallistuvat keskilinjän vasemmalle puolelle tai äänet paikallistuvat jopa keskilinjän väärille puolille.

Parhailla testin laitteilla on niiden pienestä koosta huolimatta mahdollista tallentaa korkealaatuista stereoääntä.

VIITTEET

[1] Aro, E. Tilaääni. Idemco Oy, Riffi-julkaisut, Helsinki, 2006.

[2] Blauert, J. Spatial Hearing - The Psychophysics of Human Sound Localization. The MIT Press, Cambridge (MA), 1997, pp. 358- 367.

[3] ITU-R BS 1116-1 Methods for the subjective assessment of small impairments in audio systems including multichannel sound systems, International Telecommunication Union, 2001