

RESONAATTORIPOHJAINEN DEKORRELAATTORI

Jon Fagerström¹, Vesa Välimäki¹

¹ Aalto-yliopiston sähkötekniikan korkeakoulu
Akustiikan laboratorio, Informaatio- ja tietoliikennetekniikan laitos
Otakaari 5, 02150 Espoo
fagerstrom.jon@gmail.com, vesa.valimaki@aalto.fi

Tiivistelmä

Dekorrelointisuodattimilla voidaan muuntaa monoääni useiksi keskenään dekorreloituiksi kopioiksi. Tässä artikkelissa esitellään uusi resonaattoreihin perustuva dekorrelointisuodin. Kuulokkeilla toteutetun kuuntelutestin avulla määritettiin havaintopohjaiset kanavien väliset vähimmäisviiveet, joilla ERB-suodatettu koherentti kohina saatiin kuulostamaan epäkoherentilta kohinalta. Testin tulosten perusteella määritettiin sopivat resonaattoreiden ryhmäviiveet. Lisäksi työssä esitellään dekorrelaattoriparin vaihevasteen optimointimenetelmä, jonka avulla saavutetaan huomattavasti satunnaista vaihetta pienempi koherenssi.

1 JOHDANTO

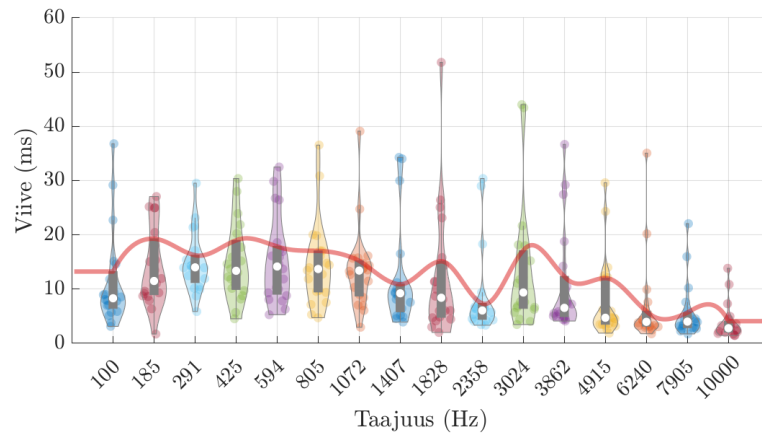
Dekorrelointi tarkoittaa signaalin autokorrelaation tai kahden tai useamman signaalin välisen ristikorrelaation pienentämistä [1, 2]. Äänentoistossa dekorrelointi viittaa usein kanavien välisen korrelaation vähentämiseen, jotta äänen tilantuntu tai stereovaikutelma kasvaisi ilman muutoksia sointiväriin tai ajallisiin ominaisuuksiin.

Varhaiset dekorrelaattorit kehitettiin muuntamaan yksikanavainen ääni stereoääneksi, aluksi sähkömekaanisin keinoin hyödyntämällä mm. viivelinjoja ja kaikukammiota [3, 4, 5]. Modernit digitaaliset menetelmät perustuvat esimerkiksi IIR-kokopäästösuodatukseen [1, 6, 7, 2], samettikohinapohjaiseen harvaan FIR-suodattimeen [8, 9], lyhytaikaiseen Fourier-muunnokseen (STFT) [10], suodinpankkiin [11, 12, 13] ja neuroverkkoihin [14].

Useimmat näistä menetelmistä lisäävät kanavien välistä taajuusriippuvaa viivettä. Tässä työssä määritetään kuuntelukokeella havaintopohjaiset viiveiden ylärajat, joiden perusteella saadaan käytännön ryhmäviiveet binauraalisille dekorrelaattoreille ERB-kaistoittain [15].

Artikkelissa esitellään uusi resonaattoripohjainen dekorrelaattori, joka koostuu resonaattoripankista, mukaillen moodisynteesiin pohjautuvia jälkikaiuntamalleja [16, 17]. Kukin resonaattori tuottaa vaimenemisnopeudesta riippuvan ryhmäviivehuipun, ja niiden vaiheet optimoidaan minimoimaan dekorrelaattoriparin koherenssi. Ehdotettu

Copyright ©2025 Jon Fagerström ja Vesa Välimäki. Tämä on avoimesti julkaistu teos, joka noudattaa Creative Commons NIMEÄ 4.0 Kansainvälinen –lisenssiä (CC BY 4.0). Teosta saa kopioida, levittää, näyttää ja esittää julkisesti ja siitä saa luoda johdannaisiteoksia, kunhan tekijän nimi ja lähde mainitaan asianmukaisesti.



Kuva 1: Kuuntelukokeen tulokset on esitetty viuludiagrammina [20]. Valkoinen piste osoittaa mediaanin ja laatikon reunat 25 %:n ja 75 %:n pisteet. Viulukuvion äärioviiva kuvaa tiheystimaattia, ja punainen käyrä näyttää 75 %:n pisteisiin perustuvan interpoloidun ryhmäviiveprofiilin, jota käytetään dekorrelaattorin suunnittelussa.

suodin voidaan toteuttaa rinnakkaisina toisen asteen suodattimina tai muodostamalla impulssivasteen näytteistä FIR-suodatin. Tämän tutkimuksen tuloksia on esitelty aiemmin DAFx-25-konferenssissa [18].

2 KUUNTELUKOE

Korvien välinen suurin luonnollinen aikaviive on noin $630 \mu\text{s}$ [19], joten aiemmin oletettiin etteivät tätä suuremmat viiveet vaikuta tilavaikutelmaan. Dekorrelaattiosuodattimen impulssivaste voi kuitenkin olla jopa 30 ms pitkä [8], mikä viittaa siihen, että myös pidemmät viiveet vaikuttavat äänen tilantuntuun.

Tämän havainnon perusteella suunniteltiin kuulokekuuntelukoe, jossa määritettiin pienimmät ERB-kaistoittaiset viiveet, joilla viivästetyn kohinan havaittu leveys vastasi korreloimattoman kohinan leveyttä. Referenssiääninä käytettiin gammatone-suodatettua epäkorreloitua binauraalista kohinaa (ts. eri satunnainen kohina kummassakin stereokanavassa) ja testiääninä korreloitua kohinaa. Testin käyttöliittymässä oli liukusäädin, jolla säädettiin kuulokkeiden kanavien välistä viivettä 0–60 ms. Osallistujat kuuntelivat testiäänä ja asettivat liukusäätimen alimpaan mahdolliseen kohtaan, jossa testiäänin kuulosti yhä yhtä leveältä kuin referenssiäänin. Liukusäätimen asteikkoa tai viiveen määrää ei näytetty koehenkilöille.

Testi koostui 16 kokeesta, joissa kohinat oli suodatettu eri ERB-kaistoittain välillä 100 Hz–10 kHz. Kymmenen osallistujaa suoritti testin kahdesti, jolloin kertyi yhteensä 320 havaintopistettä. Keskitaajuuksien järjestys ja se, kumpaan kanavaan viive lisättiin, satunnaistettiin kullekin osallistujalle ja molemmille toistoille.

Kuuntelukokeen tulokset on esitetty viuludiagrammeina kuvassa 1. Päädyimme konservatiiviseen vaihtoehtoon määrittää tarvittava dekorrelaattorin ryhmäviive 75 %:n pisteen mukaan. Intepoloitu ryhmäviiveprofiili näytetään kuvassa 1 punaisena käyränä. Kuuntelukokeen tulokset osoittavat yleisen trendin: matalammat taajuuskaistat vaativat isomman

viiveen leveysvaikutelman syntymiseen kuin korkeammat kaistat. Lisäksi vaadittu viive ei seuraa aiemmin ehdotettua taajuuden käänteislukua [11], vaan osoittaa matalimpia taajuuksia lukuun ottamatta suhteellisesti paljon suurempia viiveitä.

3 MENETELMÄ

Tässä osiossa esitellään ehdotettu dekorrelaattori, joka perustuu resonaattoripankkiin. Lisäksi käsitellään optimointimenetelmää, jolla minimoidaan kahden dekorrelaattorin välinen koherenssi.

3.1 Suodinrakenne

Resonaattoripohjainen dekorrelaattori perustuu eksponentiaalisesti vaimeneviin siniaaltoihin, jotka on lineaarisesti sijoitettu ERB-asteikolle eri taajuuksille koko äänitaajuusalueella. Dekorrelaattorin impulssivaste muodostuu K vaimenevan siniaallon summasta [16]:

$$h(n) = \sum_{k=0}^{K-1} A_k R_k^n \cos(\omega_k n / f_s + \phi_k), \quad (1)$$

missä A_k on alkuamplitudi, $R_k = e^{-\alpha_k / f_s}$ on napasäde, joka määrittää eksponentiaalisen vaimenemistermin, ja ω_k sekä ϕ_k ovat k :nnen resonaattorin taajuus ja vaihe radiaaneina. Väritymän minimoimiseksi impulssivaste valkaistaan ekvalisaattorilla, joka suunnitellaan lineaariennustuksen (engl. linear prediction) avulla [18].

Ehdotettu dekorrelaattori voidaan toteuttaa joko korkea-asteisena FIR-suotimena näytteistämällä yhtälön (1) katkaistu impulssivaste tai vaihtoehtoisesti resonaattoripankkina, jossa jokainen yksittäinen resonaattori toteutetaan toisen asteen IIR-suotimena. IIR-toteutuksen lähtösignaali $y(n)$ lasketaan tulosaalille $x(n)$ seuraavan differenssiyhtälön avulla [18]:

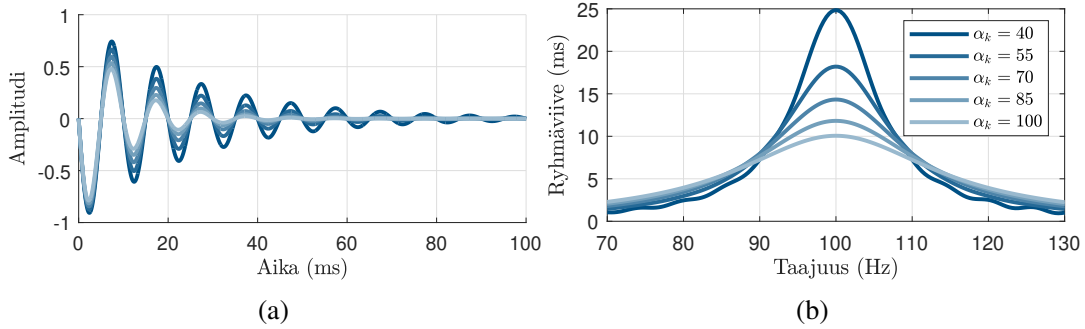
$$y(n) = \sum_{k=0}^{K-1} \left[b_{0k} x(n) + b_{1k} x(n-1) + b_{2k} x(n-2) + a_{1k} y(n-1) + a_{2k} y(n-2) \right], \quad (2)$$

missä $b_{0k} = A_k$, $b_{1k} = -2A_k \cos(\phi_k)$, $b_{2k} = A_k$, $a_{1k} = -2R_k \cos\left(\frac{\omega_k}{f_s}\right)$ ja $a_{2k} = R_k^2$.

Kuvissa 2a ja 2b esitetään 100-Hz resonaattorin impulssivaste ja ryhmäviive eri vaimenemiskertoimen α_k arvoilla. Hitaasti vaimenevien resonaattoreiden (pieni α_k) ryhmäviiveen maksimiarvo on suuri, mikä osoittaa vaimenemiskertoimen ja ryhmäviiveen olevan kääntäen verrannollisia toisiinsa nähden.

3.2 Optimointi

Tehokkaan dekorrelaattorin toteuttamiseksi minimoimme kahden resonaattoripohjaisen dekorrelaattorin välisen koherenssin. Optimointi toteutettiin geneettisellä algoritmilla [21] MATLABilla. Optimoitava muuttuja on kunkin resonaattorin vaihe ϕ_k , ja optimoinnille asetettiin ylä- ja alarajoiksi π ja $-\pi$. Ennen optimointia kunkin resonaattorin vaihe satunnaistettiin tasaisesti rajojen välille ja alkuamplitudit A_k välille -1 ja 1 .



Kuva 2: Yhden 100 Hz resonaattorin (a) impulssivaste ja (b) ryhmäviive vaimenemiskertoimen α_k funktiona.

Häviöfunktiona käytettiin havaintopohjaisen koherenssin neliöllisen virheen keskiarvoa (MSE) desibeleinä ekvalisoidun dekorrelaattorin impulssivasteelle:

$$J = 10 \log_{10} \left(\sqrt{\frac{1}{M} \sum_{m=0}^{M-1} \hat{\Phi}_{\text{per}}^2(\omega_m)} \right), \quad (3)$$

missä havaintopohjainen koherenssi määritellään seuraavasti [18]:

$$\hat{\Phi}_{\text{per}}(\omega) = \sum_{b=0}^{B-1} \frac{\sum_{t_b=0}^{T_b-1} |H_{bL}(t_b, \omega) H_{bR}^*(t_b, \omega)|^2}{\sum_{t_b=0}^{T_b-1} |H_{bL}(t_b, \omega)|^2 \sum_{t_b=0}^{T_b-1} |H_{bR}(t_b, \omega)|^2}, \quad (4)$$

missä H_{bL} ja H_{bR} ovat vasemman ja oikean kanavan b :nnen ERB-kaistan STFT-kertoimet, t_b on b :nnen ERB-kaistan aikakehyksen indeksi ja T_b on STFT:n aikakehysten lukumäärä b :nnen ERB-signaalin osalta. Tässä työssä käytettiin arviointiin $B = 16$ ERB-kaistaa.

4 ARVIOINTI

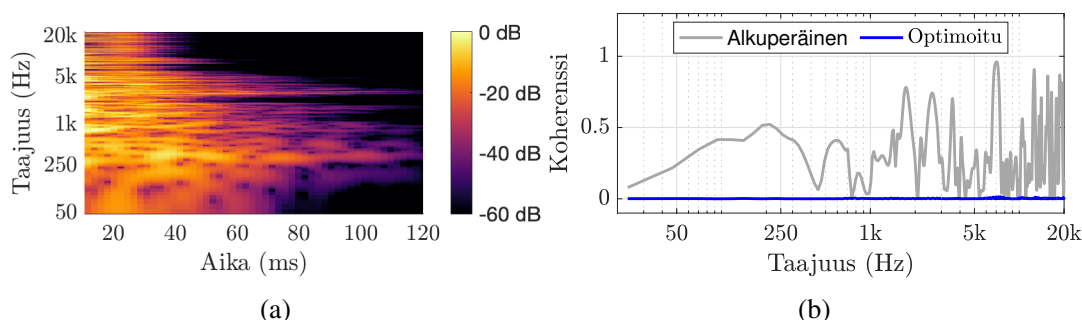
Seuraavaksi arvioidaan lyhyesti resonattoripohjaisen dekorrelaattorin toimivuus. Katavampi vertailu ehdotetun ryhmäviiveprofiilin ja kahden aiemman profiilin välillä on esitetty aiemmassa julkaisussa [18]. Ääniesimerkkejä on saatavilla verkossa. ¹

Kuva 3a esittää vasemman kanavan optimoidun impulssivasteen spektrogrammin. Odotusti impulssivasteen nähdään vaimenevan hitaammin keskitaajuuksilla ja nopeammin matalilla ja korkeilla taajuuksilla, mikä täsmää käytettyyn ryhmäviiveprofiiliin (ks. kuva 1). Kuva 3b näyttää havaintopohjaisen koherenssin ennen optimointia satunnaisilla vaiheilla (harmaa) ja optimoinnin jälkeen (sininen). Kuvan 3 perusteella geneettinen algoritmi onnistui säätämään vaiheet siten, että koherenssi on lähellä nollaa koko taajuusalueella.

5 YHTEENVETO

Tässä artikkelissa esitellään optimoitava resonattoripohjainen menetelmä binauraalisten dekorrelaattiosuodattimien suunnitteluun. Ehdottamamme dekorrelaattori on suodinpank-

¹<http://research.spa.aalto.fi/publications/papers/dafx25-reso-deco/>



Kuva 3: *Optimoidun resonaattoripohjaisen dekorrelaattorin (a) impulssivasteen spektrogrammi ja (b) koherenssi (sininen). Vertailun vuoksi näytetään alkuperäinen koherenssi ennen optimointia (harmaa).*

ki, joka muodostetaan toisen asteen resonaattoreista. Kunkin resonaattorin vaimenemisnopeus määrittää sen keskitaajuuden ryhmäviiveen. Optimoitava muuttuja on resonaattoreiden vaihe, jonka säätö minimoi kahden dekorrelaattorin välisen koherenssin. Vaiheiden asettelu ohjaa vierekkäisten resonaattoreiden rakentavaa ja tuhoavaa interferenssiä, mikä johtaa pienempään koherenssiin. Lisäksi menetelmä vaatii ekvalisaattorin, joka valkaisee resonaattoripankin amplitudivasteen. Uusi dekorrelaattiosuodin voidaan toteuttaa joko IIR-suodattimena, joka koostuu useista rinnakkaisista toisen asteen resonaattoreista ja ekvalisointisuodattimesta, tai yhtenä korkea-asteisena FIR-suodattimena, jonka kertoimet saadaan suoraan ekvalisoidun dekorrelaattorin katkaistusta impulssivasteesta.

Ehdotetun dekorrelaattorin ryhmäviiveprofiili suunniteltiin kuuntelukokeen perusteella, jossa osallistujat määrittivät ERB-kaistoittain minimiviiveen, jolla kuuntelukokemus vastaa korreloimatonta kohinaa. Tulokset osoittavat, että optimointi vähentää koherenssiä merkittävästi verrattuna satunnaisiin vaiheisiin. Menetelmää voidaan hyödyntää monosignaalin muuttamisessa stereo- tai monikanavaisiksi signaaliksi. Tulevaisuudessa voidaan tutkia useiden dekorrelaattorien yhteisoptimointia monikanavaisessa ylösmiksauksessa (engl. upmixing) ja arvioida menetelmän havaittu laatu kuuntelutestillä.

VIITTEET

- [1] G. S. Kendall. The decorrelation of audio signals and its impact on spatial imagery. *Computer Music J.*, 19(4):71–87, 1995.
- [2] E. K. Canfield-Dafilou and J. S. Abel. A group delay-based method for signal decorrelation. In *Proc. Audio Eng. Soc. Conv. 144*, Milan, Italy, May 2018.
- [3] M. R. Schroeder. An artificial stereophonic effect obtained from a single audio signal. *J. Audio Eng. Soc.*, 6(2):74–79, Apr. 1958.
- [4] B. B. Bauer. Some techniques toward better stereophonic perspective. *J. Audio Eng. Soc.*, 17(4):410–415, Aug. 1969. doi: 10.17743/jaes.17.4.410.
- [5] R. Orban. A rational technique for synthesizing pseudo-stereo from monophonic sources. *J. Audio Eng. Soc.*, 18(2):157–164, Apr. 1970.
- [6] M. A. Gerzon. Signal processing for simulating realistic stereo images. In *Proc. Audio Eng. Soc. Conv. 93*, San Francisco, CA, USA, Oct. 1992. paper 3423.

- [7] C. Uhle and P. Gampp. Mono-to-stereo upmixing. In *Proc. Audio Eng. Soc. Conv. 140*, Paris, France, May 2016.
- [8] B. Alary, A. Politis, and V. Välimäki. Velvet-noise decorrelator. In *Proc. Int. Conf. Digital Audio Effects (DAFx)*, pages 405–411, Edinburgh, UK, Sep. 2017.
- [9] S. J. Schlecht, B. Alary, V. Välimäki, and E. A. P. Habets. Optimized velvet-noise decorrelator. In *Proc. Int. Conf. Digital Audio Effects (DAFx)*, pages 87–94, Aveiro, Portugal, Sep. 2018.
- [10] S. Disch. Decorrelation for immersive audio applications and sound effects. In *Proc. Int. Conf. Digital Audio Effects (DAFx)*, Copenhagen, Denmark, Sep. 2023.
- [11] M. Bouéri and C. Kyriakakis. Audio signal decorrelation based on a critical band approach. In *Proc. 117th Audio Eng. Soc. Conv.*, San Francisco, CA, Oct. 2004. URL <https://aes2.org/publications/elibrary-page/?id=12948>.
- [12] G. Potard and I. S. Burnett. Control and measurement of apparent sound source width and its applications to sonification and virtual auditory displays. In *Proc. Int. Conf. Auditory Display (ICAD)*, Sydney, Australia, Jul. 2004.
- [13] R. Penniman. A general-purpose decorrelator algorithm with transient fidelity. In *Proc. AES Conv. 137*, Los Angeles, CA, USA, Oct. 2014. paper 9170.
- [14] C Anemüller, O. Thiergart, and E. A. P. Habets. Neural audio decorrelation using generative adversarial networks. In *Proc. IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA)*, pages 1–5, Oct. 2023. doi: 10.1109/WASPAA58266.2023.10248167.
- [15] B. C. J. Moore and B. R. Glasberg. Suggested formulae for calculating auditory-filter bandwidths and excitation patterns. *J. Acoust. Soc. Am.*, 152:942–953, Sep. 2022.
- [16] M. Karjalainen and H. Järveläinen. More about this reverberation science: Perceptually good late reverberation. In *Proc. 111th AES Conv.*, New York, USA, Sep. 2001.
- [17] J. S. Abel, S. Coffin, and K. Spratt. A modal architecture for artificial reverberation with application to room acoustics modeling. In *Proc. 137th AES Conv.*, Los Angeles, USA, Oct. 2014. paper 9208.
- [18] J. Fagerström, N. Meyer-Kahlen, S. J. Schlecht, and V. Välimäki. Perceptual decorrelator based on resonators. In *Proc. Int. Conf. Digital Audio Effects (DAFx)*, pages 214–221, Ancona, Italy, Sept. 2025.
- [19] J. Blauert. *Spatial Hearing - Revised Edition: The Psychoacoustics of Human Sound Source Localisation*. MIT Press, Cambridge, MA, 1997.
- [20] B. Bechtold. Violin Plots for Matlab, Github Project. Available at <https://github.com/bastibe/Violinplot-Matlab>, accessed Mar 27, 2024.
- [21] D. E. Goldberg. *Genetic Algorithms in Search, Optimization, and Machine Learning*. Addison-Wesley, Boston, MA, USA, 1989.