

NÄYTETAJUUDEN VAIHTO NOPEALLA FOURIER-MUUNNOKSELLA

Vesa Välimäki¹, Stefan Bilbao²

¹ Akustiikan laboratorio, Informaatio- ja tietoliikennetekniikan laitos, Aalto-yliopisto, Otakaari 5, 02150 Espoo
vesa.valimaki@aalto.fi

² Akustiikan ja audion ryhmä, Edinburghin yliopisto, Alison House, 12 Nicolson Square, Edinburgh EH8 9DF, Yhdistynyt kuningaskunta
s.bilbao@ed.ac.uk

Tiivistelmä

Esittelemme uuden menetelmän äänitteen näytetaajuuden vaihtamiseen nopean Fourier-muunnoksen eli FFT:n avulla. Yleensä äänisignaalin näytetaajuus muutetaan digitaalisella suodatuksella, joka perustuu näytearvojen interpolointiin aika-alueessa. Bi ja Mitra ovat osoittaneet, että signaalin näytetaajuus voidaan muuttaa taajuusalueessa, kun signaali jaetaan lyhyisiin ikkunoihin. Tämä artikkeli näyttää kuinka hyvin pitkiä äänitteitä, kuten kokonainen musiikkikappale tai -albumi, voidaan käsitellä yhdellä jättimäisellä FFT:llä ja käänteis-FFT:llä. Näytetaajuutta nostettaessa signaalin FFT-spektrin keskelle lisätään sopiva määrä nollija ja lasketaan käänteis-FFT. Näytetaajuutta laskettaessa FFT-spektrin keskeltä poistetaan sopiva määrä taajuuspisteitä ja tehdään käänteismuunnos, jolloin näytetaajuus laskee ja samalla liian korkeat taajuudet jäävät pois. Äänenlaatu ei kärsi tässä operaatiossa, mutta Nyquist-rajataajuudelle syntyy häiriö, joka kannattaa vaimentaa.

1 JOHDANTO

Äänitekniikassa käytetään useita näytetaajuuksia, ja siksi näytetaajuusmuunnosta tarvitaan jatkuvasti [1, 2]. Yleensä äänisignaalin näytetaajuus vaihdetaan digitaalisella suodatuksella, joka perustuu näytearvojen interpolointiin aika-alueessa [3, 4, 5, 6].

Kuitenkin vuosikymmenten ajan on ollut tiedossa, että Fourier-muunnos interpoloi kaistarajoitettua signaalin arvoja hyvin tarkasti [7, 8, 9]. Viime vuosikymmenellä Bi ja Mitra esittivät kuinka signaalin näytetaajuus voidaan muuttaa taajuusalueessa, kun signaali jaetaan lyhyisiin ikkunoihin [10].

Tässä artikkelissa näytämme kuinka erittäin pitkiä äänitteitä, kuten kokonainen musiikkikappale tai -albumi, voidaan muuntaa yhdellä hyvin pitkällä Fourier-muunnoksella ja Fourier-käänteismuunnoksella. Käytännössä Fourier-muunnokset toteutetaan FFT-

Copyright ©2023 Vesa Välimäki ja Stefan Bilbao. 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 johdannaisteoksia, kunhan tekijän nimi ja lähde mainitaan asianmukaisesti.

algoritmin (engl. Fast Fourier Transform) avulla. Tekniikkaa kutsutaan jättiläis-FFT-menetelmäksi (engl. giant FFT) ja se on julkaistu aiemmin Journal of the Audio Engineering Society -lehdessä vuoden 2023 keväällä [11].

2 NÄYTETAAJUUDEN VAIHTO TAAJUUSALUEESSA

Seuraavaksi esitellään FFT-spektrin rakenne, jättiläis-FFT-menetelmä näytetaajuuden nostamiseen ja laskemiseen sekä uuden algoritmin toiminta vuokaavion avulla.

2.1 FFT-spektrin rakenne

Taajuusalueen näytetaajuusmuunnosmenetelmän hahmottamiseksi täytyy ymmärtää reaaliarvoisen äänisignaalin diskreetin spektrin rakenne. Kuvassa 1(a) näytetään yksinkertaistettu magnitudispektri, joka sisältää N spektridatan itseisarvoa ($0, 1, 2, \dots, N - 1$), missä N on sama kuin FFT:n pituus. Taajuuspisteessä 0 on tasavirtakomponentti eli DC. Piste $N/2$ vastaa puolta näytetaajuudesta ja on Nyquistin raja eli korkein taajuus, joka on mahdollista tallentaa. Näissä molemmissa erikoispisteissä spektridata on reaaliarvoista, kun taas muut spektridatan arvot ovat kompleksisia eli luvuilla on reaali- ja imaginaariosa.

Kuvan 1(a) oikea puoli on peilikuvapektri, jossa esiintyvät itseisarvoltaan samat spektridatan arvot kuin alkuperäisessä spektrissä mutta päinvastaisessa järjestyksessä. Usein peilikuvapektriä kutsutaan negatiivisiksi taajuuksiksi. Kuvan 1(a) esimerkissä äänienergian määrä (magnitudi) on suurimmillaan pienillä taajuuksilla ja pienenee kohti suuria taajuuksia. Negatiivisilla taajuuksilla sama tapahtuu toiseen suuntaan. Peilikuvapektrin vaihe on käänteinen signaalin spektriin verrattuna, mikä ei näy tässä kuvassa.

2.2 Näytetaajuuden nosto

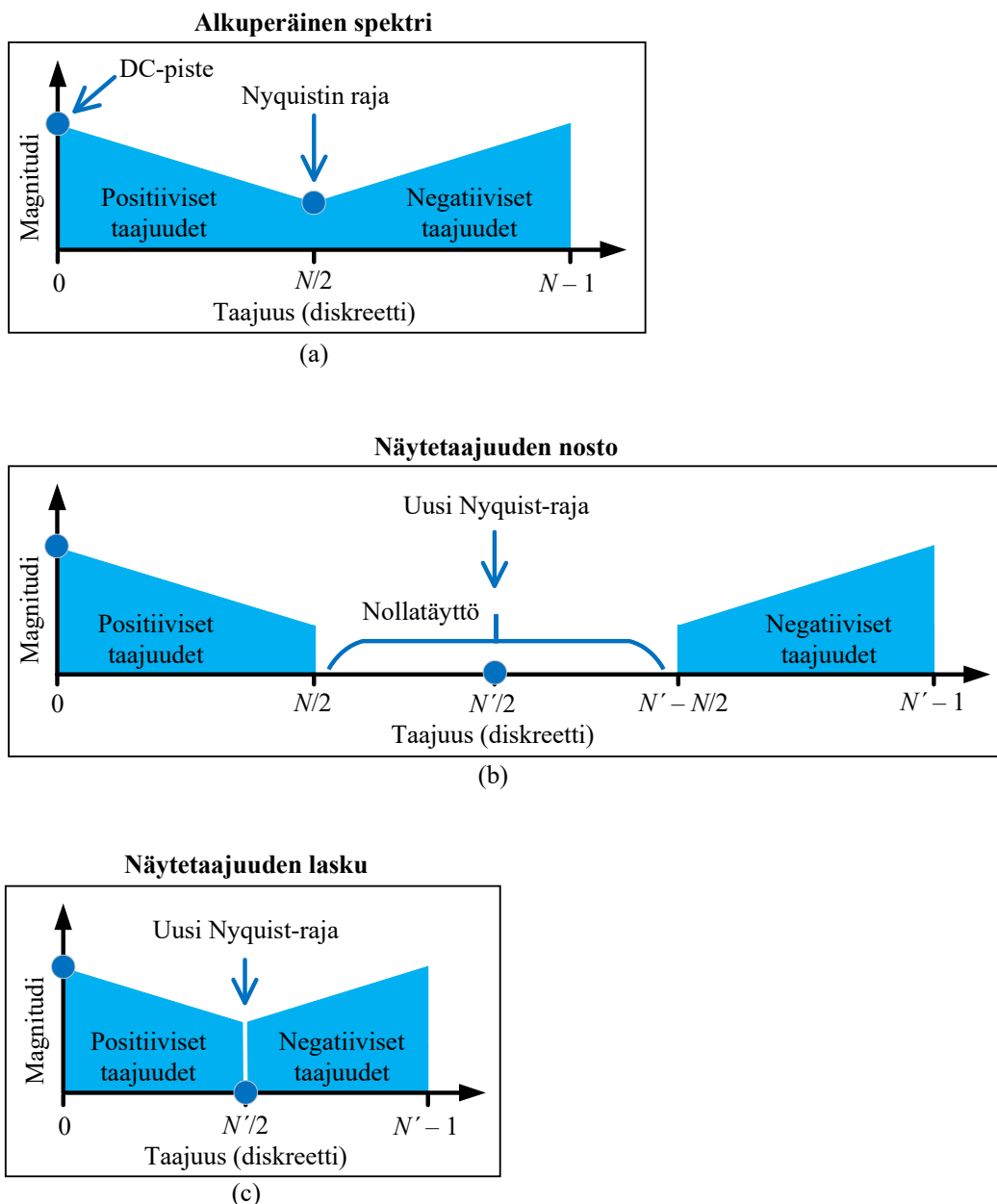
Näytetaajuutta nostettaessa signaalin spektrisisällön ei pidä muuttua. Käytännössä ainakin osa uuden signaalin spektristä jää tyhjäksi eli signaali on ylinäytteistetty. Näytetaajuuden nostamista kutsutaan usein interpoloinniksi, koska näytteiden määrää on lisättävä eli uusia näytearvoja on estimoitava alkuperäisten näytteiden välille (tai tilalle).

Kuvassa 1(b) näytetään kuinka näytetaajuutta nostetaan lisäämällä signaalin FFT-spektrin keskelle sopiva määrä nolliä, jolloin puskurin uusi pituus on N' [11]. Sen jälkeen lasketaan käänteis-FFT, jolloin saadaan tuloksena koko signaali uudella näytetaajuudella. Oikea nollatäytön (engl. zero-padding) määrä tulee siitä, että uusi näytetaajuus on N'/N -kertainen alkuperäiseen näytetaajuuteen nähden [11]. Signaalin näytetaajuus kasvaa eikä signaalin taajuussisältö oleellisesti muutu, kun se toistetaan uudella näytenopeudella.

Nollatäytön jälkeen signaalin spektrissä on N' pistettä ja uusi Nyquist-raja on $N'/2$, mikä on merkitty kuvaan 1(b). Pituudet N ja N' voivat olla suuruusluokaltaan satoja miljoonia, joten näytetaajuutta voidaan muuttaa hyvin tarkasti. Äänenlaatu ei kärsi tässä operaatiossa lainkaan, koska muunnoksessa syntyvä virhe on tietokoneen laskentatarkkuuden luokkaa. Jos alkuperäisessä signaalissa on energiaa Nyquist-rajana tuntumassa, muunnoksessa syntyy kuitenkin korkeataajuinen häiriö, joka kannattaa vaimentaa [11].

2.3 Näytetaajuuden lasku

Kuvasta 1(c) nähdään kuinka näytetaajuuden laskussa FFT-spektrin keskeltä poistetaan taajuuspisteitä [11]. Tämän jälkeen voidaan ajaa käänteis-FFT, jolloin saadaan signaali, jonka näytetaajuus on pienentynyt ja samalla liian korkeat taajuudet jäävät pois [11]. Spektrin katkaisusta voi syntyä korkeataajuinen häiriö, joka kannattaa poistaa [11].



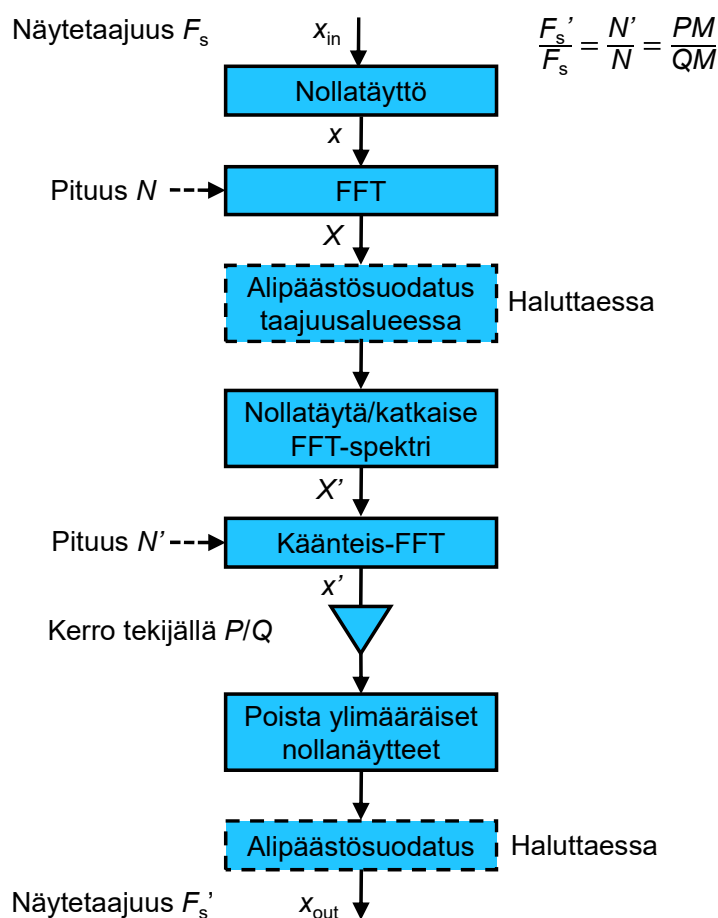
Kuva 1: (a) Yksinkertaistettu äänen magnitudispektri, jossa on mukana myös peilikuvapektri (negatiiviset taajuudet), (b) spektri näytetaajuuden noston jälkeen, kun Nyquist-ajan ympärille on lisätty nollia ja spektrin puolet on siirretty kauemmas toisistaan, ja (c) spektri näytetaajuuden laskun jälkeen, kun spektri on katkaistu symmetrisesti Nyquist-ajan ympäriltä ja puolet liitetty yhteen sekä Nyquist-taajuudelle asetettu yksi nollanäyte.

2.4 Jättiläis-FFT-algoritmi

FFT ja käänteis-FFT voidaan nykyisin laskea minkä pituiselle lukujonolle tahansa, kunhan pysytään alle tietyn maksimipituuden, joka on satoja miljoonia. Pienistä tekijöistä koostuvat pituudet FFT laskee nopeammin kuin suurista tekijöistä koostuvat pituudet. Käänteis-FFT:n laskentanopeus on sama kuin FFT:n.

Esimerkiksi $2^{27} = 134\,217\,728$ -pituisen FFT:n laskeminen Matlabilla kannettavassa tietokoneessa vie keskimäärin 3,4 s. Kun näytetaajuus on 44,1 kHz, tämä pituus vastaa lähes 51 min pitkää äänitettä, johon mahtuu kokonaisen CD-levyn toinen kanava (esim. vasen). Jos signaalin pituutta kasvatetaan hieman alkulukuun 134 217 757, FFT:n laskenta kestää 62 s eli lähes 20 kertaa pitempään. Tämäkään ei välttämättä ole liian pitkä odotusaika. Laskentaesimerkit ajettiin Lenovo P50 -tietokoneella, jossa on Intel Xeon v5 -prosessori [11].

Kuva 2 näyttää jättiläis-FFT-algoritmin periaatteen [11]. Muuntosuhte on P/Q , missä P ja Q ovat kokonaislukuja ja voivat olla hyvin suuria. Myös luvut F_s , F'_s , N , N' ja M kuvassa 2 ovat kokonaislukuja. Menetelmän tehokkuus riippuu signaalin pituudesta ja



Kuva 2: Jättiläis-FFT-algoritmin vuokaavio, jossa näytetaajuutta muutetaan tekijällä P/Q , missä sekä P että Q voivat olla suuria kokonaislukuja. Kaavion keskikohdassa FFT-spektri joko nollatäytetään, kuten kuvassa 1(b), tai katkaistaan, kuten kuvassa 1(c).

erityisesti siitä kuinka suuria alkulukuja pituuksissa N ja N' on tekijöinä. Siksi algoritmin ensimmäisessä lohossa alkuperäisen signaalin loppuun lisätään nollia (aika-alueessa), jotta saavutetaan jokin pituus N , jonka suurin alkulukutekijä on pieni, esimerkiksi 7 [11].

Kuvan 2 toisessa lohossa lasketaan signaalin spektri N -pisteisellä FFT:llä. Tämän jälkeen spektrin ylimpiä jäljellejääviä taajuuspisteitä voidaan vaimentaa, jotta spektrin muokkauksesta aiheutuva häiriö saadaan vaimennettua [11]. Nollatäyttö taajuusalueessa voi aiheuttaa spektriin epäjatkuvuuskohdan riippuen signaalin spektristä ja silloin siitä syntyy samankaltainen häiriö kuin spektrin katkaisusta [11]. Sen jälkeen tehdään kuvan 1(b) spektrin täyttö nolilla tai kuvan 1(c) mukainen katkaisu. Muihin spektridatapisteisiin ei tarvitse koskea.

Kun spektriä on muokattu, lasketaan käänteis-FFT, jonka pituus on N' . Myös tällä kertaa suuren luvun N' tulee koostua pienistä tekijöistä, jotta laskenta sujuu nopeasti. Lopputuloksena saatavan signaalin kaikki näytteet on kerrottava näytetaajuussuhteella P/Q . Muuten signaali on vaimentunut, koska interpoloinnissa signaalin keskimääräinen teho näytettä kohti vähenee.

Kuvan 2 toiseksi viimeisessä lohossa aika-alueen signaalin lopusta voidaan poistaa ylimääräiset näytteet, jotta tuloksena saatavan signaalin pituus olisi tarkalleen oikea. Lopuksi signaali voidaan alipäästösuodattaa siten, että uuden näytetaajuuden lähellä olevia hyvin korkeita taajuuksia vaimennetaan. Näin voidaan vielä vähentää spektrin katkaisusta aiheutuvaa häiriötä, joka ei tosin yleensä ole kuultavissa [11].

Jättiläis-FFT-menetelmällä pitkänkin äänitteen näytetaajuus voidaan vaihtaa nopeasti. Käytännössä kokonaista CD-levyä vastaavan musiikkiäänitteen molempien kanavien näytetaajuuden muuntaminen kestää nykyisin alle 20 s, kun pituus N valitaan sopivasti eli lisätään sopiva määrä nollia alkuperäisen signaalin perään.

3 YHTEENVETO

Tässä artikkelissa esiteltiin jättiläis-FFT-menetelmä äänisignaalin näytetaajuuden vaihtamiseen FFT:n avulla. Oleellinen uutuus on se, että muunnettavan signaalin pituus voi olla valtava, kuten satoja miljoonia näytteitä. Tämä mahdollistaa pitkienkin äänitteiden näytetaajuuden muuntamisen yhden FFT- ja käänteis-FFT-operaation avulla. Koko muunnettu signaali saadaan valmiiksi kerralla eikä näyte näytteeltä, kuten digitaaliseen suodatukseen perustuvissa aika-alueen ratkaisuisissa. Näytetaajuusmuunnoksen laskentatarkkuus jättiläis-FFT-menetelmällä on myös hämmästyttävän hyvä, lähellä tietokoneen laskentatarkkuutta. Ainoastaan Nyquist-ajan tuntumaan syntyy häiriö, joka on helppo vaimentaa [11].

Yksi merkittävä etu on myös se, että jättiläis-FFT-menetelmä ei rajoitu näytetaajuuden muuttamiseen kokonais- tai rationaalilukutekijällä, vaan mikä tahansa irrationaalinenkin suhde voidaan toteuttaa erinomaisella tarkkuudella. Muuntosuhde määräytyy käänteis-FFT- ja FFT-pituuksien suhteesta eli se on kahden suuren kokonaisluvun suhde. Jättiläis-FFT-menetelmää voidaan käyttää, kun äänisignaalin näytetaajuutta täytyy muuttaa vakiotekijällä.

VIITTEET

- [1] Audio Engineering Society, Inc. AES recommended practice for professional digital audio – Preferred sampling frequencies for applications employing pulse-code modulation. AES Standard AES5-2018, 2018.
- [2] V. R. Melchior. High-resolution audio: A history and perspective. *Journal of the Audio Engineering Society*, 67(5):246–257, 2019. <https://doi.org/10.17743/jaes.2018.0056>.
- [3] J. O. Smith and P. Gossett. A flexible sampling-rate conversion method. In *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, pages 112–115, 1984. <https://doi.org/10.1109/ICASSP.1984.1172555>.
- [4] T. Saramäki and T. Ritoniemi. An efficient approach for conversion between arbitrary sampling frequencies. In *Proceedings of the IEEE International Symposium on Circuits and Systems (ISCAS)*, pages 285–288, 1996. <https://doi.org/10.1109/ISCAS.1996.541702>.
- [5] K. Rajamani, Y.-S. Lai, and C. W. Farrow. An efficient algorithm for sample rate conversion from CD to DAT. *IEEE Signal Processing Letters*, 7(10):288–290, 2000. <https://doi.org/10.1109/97.870683>.
- [6] A. Zeineddine, A. Nafkha, S. Paquelet, C. Moy, and P. Y. Jezequel. Comprehensive survey of FIR-based sample rate conversion. *Journal of Signal Processing Systems*, 93:113–125, 2021. <https://doi.org/10.1007/s11265-020-01575-6>.
- [7] D. Fraser. Interpolation by the FFT revisited—An experimental investigation. *IEEE Transactions on Acoustics, Speech and Signal Processing*, 37(5):665–675, 1989. <https://doi.org/10.1109/29.17559>.
- [8] J. O. Smith. *Mathematics of the Discrete Fourier Transform (DFT) With Audio Applications*. W3K Publishing, <http://www.w3k.org/books/>, 2007.
- [9] R. G. Lyons. *Understanding Digital Signal Processing*. Pearson Education, Boston, MA, 3rd edition, 2011. <https://ptgmedia.pearsoncmg.com/images/9780137027415/samplepages/0137027419.pdf>.
- [10] G. Bi and S. K. Mitra. Sampling rate conversion in the frequency domain [DSP Tips and Tricks]. *IEEE Signal Processing Magazine*, 28(3):140–144, 2011. <https://doi.org/10.1109/MSP.2011.940413>.
- [11] V. Välimäki and S. Bilbao. Giant FFTs for sample-rate conversion. *Journal of the Audio Engineering Society*, 71(3):88–99, 2023. URL <http://www.aes.org/e-lib/browse.cfm?elib=22033>.