

Signaalinkäsittelymenetelmiä rumpujen nuotintamiseen ja musiikin muotoanalyysiin

Lectio praecursoria

Jouni Paulus

8.1.2010

Sisältö

Tässä esitelmäni sisältö. Ensin sivuan hieman digitaalista signaalinkäsittelyä ja musiikin sisältöanalyysiä yleisellä tasolla. Tämän jälkeen käsitelen väitöskirjan aihepiiriin liittyviä ongelmia rumpujen nuotinnoksessa ja musiikin muotoanalyysissä.

Signaalinkäsittely

Idealisoituna lähes kaikki havainnointi ja toiminta perustuu signaalinkäsittelyyn, myös ihmisten toiminta. Otetaan esimerkiksi väitöskirjan aihetta sivuava äänen havainnointi. Normaalissa ympäristössämme ääni on ilmassa etenevää paineenvaihtelua. Korvassa tämä paineenvaihtelu liikuttaa tärykalvoa. Tärykalvon värähtely välittyy kuuloluiden muodostaman vivuston kautta simpukan sisältämän nesteen edestakaiseksi liikkeeksi. Nesteen liike puolestaan liikuttaa simpukan sisällä olevaa basilaarikalvoa ja kalvolla olevat aistinsolut muuntavat liikkeen hermoimpulsseiksi. Nämä impulssit välittyvät aivoihin, joissa varsinaisen kuulohavainto muodostuu. Tietoteknisesti tapahtuma on samankaltainen. Ilman värähtely liikuttaa mikrofonin kalvoa. Kalvo muuntaa ilman liikkeen jännitemuutokseksi. Tämä jännitesignaali kelpaa jo sellaisenaan tiettyihin sovelluksiin. Digitaalisen signaalinkäsittelyssä jatkuva-aikainen ja -arvoinen jännitesignaali muutetaan digitaalseksi näytteistämällä sitä. Tämä digitaalinen signaali välitetään sitten prosessointiyksikölle, joka muodostaa siitä "havainnon".

Musiikin sisältöanalyysi

Musiikkia on levitetty digitaalisessa muodossa CD:llä jo lähes kolme vuosikymmentä. Viime aikoina tietoverkkojen kehittyessä painopiste on kuitenkin siirtynyt immateriaalisten tuotteiden välittämiseen, eli levyn sijaan kuluttaja ostaakin musiikin vain tiedostona. Maailmanlaajuisesti fyysisten levyjen myynti putosi vuodesta 2007 vuoteen 2008 15 prosenttia, kun taas digitaalinen myynti kasvoi samana aikana 24 prosenttia. Myynnin lisäksi kuluttajilla on pääsy yhä suurempiin musiikkikokoelmiin verkossa.

CD-aikakaudella albumien käsittely oli melko yksinkertaista, koska normaaliin soittimeen mahtuu kerrallaan vain yksi levy. Tilanne on täysin erilainen kun matkapuhelimessa tai tietokoneella onkin pääsy miljooniin kappaleisiin. Miten valikoiman voi esittää käyttäjälle? Miten valikoimasta voi etsiä haluamaansa musiikkia? Voiko musiikilla tehdä jotain muutakin kuuntelun lisäksi? Näihin pyrkii vastaamaan alana musiikin tiedonhaku. Musiikkitiedonhaku itsessään on varsin laaja monitieteellinen kenttä, jossa yhteistyötä tekevät kirjastojen informaatikot, musiikkitieteilijät ja tietotekniikan asiantuntijat. Monitieteellisyydestään johtuen musiikkitiedonhaku kattaa varsin laaja-alaisesti digitaalisten musiikkikokoelmien käsittelyyn liittyviä operaatioita, sisältäen mm. musiikin kulutuksen sosiologisia ja taloudellisia kysymyksiä, käyttöliittymäsuunnittelua, musiikin semanttista käsittelyä (genren ja tunnelman tunnistamista), musikologista tutkimusta (tyypillisiä sointukulkuja, rytmisiä rakenteita), esimerkkipohjaista hakua tietokannoista (query by humming, sormenjälki), musiikin nuotinnos ja erilaiset esitysmuodot.

Signaalinkäsittely liittyy musiikkitiedonhakuun läheisesti siten, että sen tarjoamat menetelmät ovat taustalla tekemässä monia korkeamman tason tiedonhakuovelluksia mahdollisiksi, sillä esimerkiksi mikäli saatavilla on vain musiikkiäänite on siitä varsin vaikeaa sellaisenaan analysoida esimerkiksi sointukulkuja. Käytännössä kuitenkin tällaista työtä on tehty jo pitkään: musiikillisen koulutuksen saaneet henkilöt kuuntelevat kappaleita ja pyrkivät kuulopohjalta kuvaamaan haluttua ominaisuutta musiikista. Vaikka tähän tapaan työskentely on mahdollista, on se kuitenkin hyvin aikaavievää. Toisin sanoen, tässä ongelmakentässä

ja tämän väitöskirjatyön yhteydessä signaalikäsitteily pyrkii tarjoamaan menetelmiä musiikin sisältöanalyysiin. Termi ``sisältö" tarkoittaa sitä, että pelkän yksinkertaisen suodatuksen sijaan prosessoinnin tulokselle on annettavissa jokin musiikillinen merkitys, kuten äänenkorkeus, tempo, tai lista käytetyistä soittimista.

Rumpujen nuotinnos

Tämä väitöskirjatyö käsittelee kahta erityistä ongelmaa musiikin sisältöanalyysiin liittyen: rumpujen nuotintamista ja musiikin muotoanalyysiä. Rumpujen nuotinnos on näistä kahdesta paremmin määritelty: annetusta äänisignaalista pitää löytää ajanhetket, jolloin kohdejoukkoon kuuluvia rumpuja on lyöty ja tunnistaa mitä rumpuja kyseisellä ajanhetkellä on lyöty. Lisäksi tässä työssä keskitytään pääasiassa länsimaisessa pop ja rock-musiikissa käytettyihin rumpuihin.

Rumpujen nuotinnos mahdollistaa useita erilaisia sovelluksia. Tällainen on esimerkiksi rumpusisällön muokkaaminen. Musiikin kuuntelija ei enää säädäkään basson ja diskantin tasapainoa, vaan säätää kuuntelemaansa kappaleen rumpujen tasoa tai käytettyjä instrumentteja. Musiikin tuotannossa puolestaan tehdyn äänitteen muokkaaminen helpottuu huomattavasti, kun tiedetään missä kohdassa rumpuraitaa eri rumpujen iskut ovat. Opetuksessa nuotinnosta voi käyttää toimimaan kriitikkona soitolle, kun tiedetään mitä piti soittaa ja nuotinnos kertoo mitä soitettiin.

Tässä työssä on lähestytty rumpujen nuotinnosta useilla eri menetelmillä. Ensimmäinen niistä on ns. segmentointi ja tunnistus. Siinä äänisignaali paloitellaan ajallisesti pienempiin yksiköihin esimerkiksi tunnistamalla rummuniskujen alkuajanhetket, ja pyritään tunnistamaan kunkin palan sisältämät rummut. Eräs tämän lähestymistavan ongelmista on tunnistuksen epävarmuus. Väitöskirjatyössä ehdotetaan tässä hyödynnettäväksi korkeamman tason musiikillista tietoa. Rumpujen ominaisuutena musiikissa kun on, että varsin usein ne esiintyvät jollain tapaa säännöllisinä kuvioina, jotka vielä toistuvat useita kertoja peräkkäin. Esimerkiksi kuvan tapauksessa pyrittäisiin tunnistamaan kysymysmerkillä indikoidun palan sisältöä. Jopa tietämättä sen tarkemmin itse äänisisältöä olisi melko helppo arvata pystysuuntaisen nuolen edustaman jaksollisen ennustuksen mukaan, että kyseisessä kohdassa saattaisi esiintyä bassorummun ja hihatın yhdistelmä BH.

Hieman erilainen lähestymistapa nuotinnokseen on unohtaa suora piirteenirrotus ja luokittelu, ja käsitellä alkuperäistä signaalia äänilähteiden erottelumenetelmillä. Siinä signaali pyritään jakamaan niin, että jokaisen kohderummun tuottama signaali on erillään muista. Tällöin onkin mahdollista saavuttaa nuotinnos havainnoimalla iskujen alkuhetket erotelluista signaaleista, eikä erillistä tunnistusta enää tarvita.

Polyfoninen musiikki kaikkine muine instrumentteineen on nuotinnosjärjestelmille vaikea lähtökohta. Erityisesti rummuniskujen ajallinen paikantaminen on vaikeaa. On kuitenkin mahdollista opettaa nuotinnosjärjestelmä huomioimaan miten signaalista irrotetut piirteet käyttäytyvät kun tiettyä kohderumpua on lyöty ja miten taas kun sitä ei ole lyöty. Hyödyntäen tämänkaltaisia malleja on mahdollista suorittaa rummuniskujen ajallinen paikallistaminen ja tunnistaminen samanaikaisesti samassa prosessissa.

Seuraavaksi soitan esimerkin tästä viimeksimainitusta menetelmästä. Ensin järjestelmälle annetaan sisäänmenon ja sitten signaalin, joka saadaan syntetisoimalla nuotinnostulos.

Muotoanalyysi

Kun rumpujen nuotinnoksessa yksittäisten iskujen tunnistaminen tapahtuu kymmenien millisekuntien aikaskaalalla, toinen väitöskirjatyön käsittelemä aihepiiri, muotoanalyysi, pyrkii analysoimaan musiikkisignaalia paljon karkeammalla aikaskaalalla. Etenkin länsimaisen populaarimusiikin kappaleet voidaan jakaa eri osiin, joita ovat mm. säkeistö, kertosäkeistö ja soolo. Varsin tyypillinen esimerkki kappaleen rakenteesta on "intro, säkeistö, kertosäkeistö, säkeistö, kertosäkeistö, soolo, kertosäkeistö". Kun yksittäinen osa kestää yleensä noin 20-30 sekuntia, niin tämä kuvaus kattaa jo ns. normaalimittaisen kolmen minuutin radiosoihtokappaleen. Muotoanalyysissä pyritään löytämään tällainen rakennekuvaus pelkästään akustisen signaalin perusteella.

Muotoanalyysi itsessään mahdollistaa useita musiikkitiedonhakuun liittyviä sovelluksia, seuraavassa muutamia esimerkkejä. Koska saman osan eri esiintymät ovat usein akustisesti hyvin samankaltaisia, tulosta on mahdollista hyödyntää signaalin pakkauksessa: jokaisesta osasta tallennetaan vain yksi esiintymä kokonaan ja osan toistojen kohdissa tallennetaan erot tästä esiintymästä. Näiden lisäksi on mahdollista tarjota muototieto suoraan loppukäyttäjälle ja mahdollistaa uusia viihdesovelluksia. Yksinkertainen esimerkki on tarjota kuvaus kappaleen rakenteesta soitto-ohjelmassa. Tämä mahdollistaa helpomman navigoinnin kappaleen sisällä, ja entuudestaan tuntemattoman kappaleen sisällön hahmottaminen on hieman helpompaa.

Eräs muotoanalyysissä usein käytetty esitysmuoto on ns. samankaltaisuusmatriisi. Se kuvaa jokaisen signaalin ajanhetken samankaltaisuutta kaikkien muiden saman signaalin ajanhetkien kanssa. Kuten voidaan huomata tästä esimerkikuvasta, matriisiin muodostuu erilaisia kuvioita. Päädiagonaalilla on luonnollisesti hyvin suurta samankaltaisuutta, koska kukin ajanhetki on samankaltainen itsensä kanssa. Sen lisäksi matriisiin muodostuu muuallekin vastaavia vinoraitoja tai matalan etäisyyden laatikoita. Nämä viittaavat siihen, että kyseisissä kohdissa signaali toistuu samankaltaisena.

Tässä väitöskirjatyössä muotoanalyysiä lähestytään määrittelemällä mitä ominaisuuksia hyvällä muotokuvauksella on. Ensimmäinen lähestymistapa pyrkii tarjoamaan kuvauksen vain toistuvista osista. Siinä määritellään hyvän kuvauksen omaavan seuraavat ominaisuudet: saman osan esiintymiksi väitettyjen osien tulee olla akustisesti mahdollisimman samankaltaiset, kuvauksen tulee kattaa mahdollisimman paljon kappaleesta, ja kuitenkin olla mahdollisimman yksinkertainen. Painottamalla eri osatekijöitä eri tavoin saadaan useita erilaisia kuvauksia saman kappaleen rakenteelle.

Toisessa lähestymistavassa pyritään muodostamaan koko kappaleen kattava kuvaus. Siinä otetaan vielä yksinkertaisempi määritelmä hyvälle muotokuvaukselle: saman osan esiintymien tulee olla keskenään samankaltaisia ja poiketa muiden osien esiintymistä. Varsinen tässä ongelma muodostuu osien samankaltaisuuden määrittelystä. Menetelmä hyödyntää samankaltaisuusmatriisiin muodostuvia erilaisia kuvioita (raidat indikoivat toistuvaa sekvenssiä, esimerkiksi melodiaa, ja laatikot puolestaan koko osan yli samankaltaisena pysyvää ominaisuutta, esimerkiksi instrumentaatiota). Ehdotettu menetelmä hyödyntää näitä etsiessään hyvää muotokuvausta.

Muotoanalyysijärjestelmät palauttavat tuloksenaan signaalin ajallisen paloittelun ja kuvauksen siitä mitkä palat ovat saman osan esiintymiä. Jos kuvaus tarjotaan suoraan loppukäyttäjälle, olisi kuitenkin hyödyllistä mikäli osille voisi antaa myös jonkinlaisen musiikillisesti merkityksellisen nimen, kuten säkeistö tai kertosaakeistö. Väitöskirjan viimeisenä muotoanalyysiin liittyvänä osana ehdotetaan menetelmää, joka nimeää nämä palaryhmät hyödyntäen tilastollista tietoa siitä, miten eri osat seuraavat toisiaan ajallisesti. Esimerkiksi jos otamme kaikki Beatlesien kappaleet ja katsomme ensimmäisen asteen ajallisia riippuvuuksia osien kesken, saamme tämänkaltaisen tilakoneen. Siirtymiä kuvaavissa nuolissa olevat numeroarvot kuvaavat kyseisen siirtymän yleisyyttä välillä nolasta yhteen.

Tässä on esimerkki siitä miten rakennekuvausta voidaan hyödyntää soitto-ohjelmassa. Nämä värilliset laatikot kuvaavat kappaleen eri osia, sama väri indikoi samaa osaa. Tässä esitysmuodossa on helppo siirtyä kuuntelemaan jotain tiettyä kohtaa kappaleesta tai vaikka kuunnella kaikki tietyn osan esiintymät yhtä aikaa.

Yhteenveto

Signaalinkäsittely on jo arkipäivää ja pyrkii tarjoamaan työkaluja eri ongelmien ratkaisemiseksi. Tässä väitöskirjassa käsitellään työkaluja musiikin sisältöanalyysiin akustisen signaalin perusteella. Ensimmäinen osakokonaisuus käsittelee rumpujen nuotinnosta musiikista, joka voi olla pelkkiä rumpuraitoja, moniraitaäänityksiä, mielivaltaisilla äännähdyksillä tuotettua rytmää, tai mitä tahansa polyfonista musiikkia. Kussakin osaongelmassa toimii hieman erilaiset menetelmät, joissain paremmin kuin toisissa. Toinen osakokonaisuus käsittelee musiikkikappaleen muotoanalyysiä. Siihen liittyy hieman erilaisia ongelmia kuin rumpujen nuotintamiseen, lähinnä ongelmamäärittelyn monikäsitteisyydestä johtuen. Kuitenkin työssä on löydetty menetelmiä, jotka hyödyntävät samankaltaisia ominaisuuksia kuin ihmisetkin analysoidessaan kappaleen rakennetta ja menetelmät pystyvät tuottamaan jo kohtuullisen tuloksen rakenteelle.