Automatisoitu musiikin analyysi ja tunnistus

Jarkko Senne

13.11.2003

Tiivistelmä

Musiikin automaattisen analyysin ja tunnistuksen pyrkimyksenä on jäljittää ja kuvata duuri-molli-tonaalisessa musiikissa esiintyviä piirteitä. Suuri osa alueen tutkimuksesta on tehty nimenomaan melodian eri muunnelmien vertailuun ja samankaltaisten mallien, patternien etsimiseen. Tässä esitelmässä käydään läpi muutamia esimerkkejä tonaalisen musiikin analysoimisesta kuten temaattinen indeksi, assosiaatioiden eli yhdistymien suuruusluokan etsiminen, melodian mikroanalyysiä, melodioiden väliseen ”etäisyyteen” perustuvaa analyysiä.

1 Johdanto

Useat automatisoidun musiikin analyysin tutkimukset olettavat ihmisen jakamien musiikillisten elementtien vastaavan musiikin tärkeitä harmonisia tai melodisia osia. Yksi musiikin analyysin perusedellytys on musiikin jakaminen sopiviin segmentteihin, jotka tukevat analyysia myöhemmin. Nämä sopivat segmentit eivät välttämättä ole samoja mitä ihminen jakaa kuunnellessaan musiikki kappaletta. Se millä tavalla ja minkä perusteella tämä jako tapahtuu, riippuu paljon tutkimuksen tarkoituksesta.

2. Melodian varianttien tunnistaminen

Tunnistettaessa samanlaisia melodioita, voidaan käyttää temaattista indeksiä **[DiH],** jossa melodiaa käytetään hakualgoritmissä kuten tietosanakirjan sanoja. Melodioiden tulee olla järjestyksessä, jotta ne voidaan löytää nopeasti. Yksi tapa esittää melodiat on kirjaimilla(C,D,E,F,G,A,B), mutta jos melodiat ovat eri sävellajissa ne pitää transponoida samaan sävellajiin. Toinen tapa on esittää melodiat sävelten välisillä intervalleilla; ”C E G C” esitettäisiin siten: ”+3 +3 -5”. Tämä analysointitapa ei huomioi kuitenkaan melodian muunnelmia; jos vain muutamaa ensimmäistä säveltä muutetaan, temaattinen indeksointi määrittelee ne eri melodioiksi.
Toinen tapa on analysoida assosiaatioiden eli yhdistymien suuruusluokkaa **[DiH]**, jossa melodian tärkeät piirteet kuten ensimmäinen ja viimeinen nuotti sekä korkein ja matalin nuotti, erotellaan ja voidaan täten laskea assosiaatioiden suuruus. Eri ominaisuuksia voidaan painottaa, jos ne havaitaan tärkeiksi musiikin variaation suhteen; lasketaan esimerkiksi identtisten nuottien lukumäärä,

variaatio melodian alussa ei ole sen tärkeämpi kuin missään muuallakaan, painotettuja nuotteja korostetaan enemmän analyysissä. Tämä tapa ei anna täydellistä ratkaisua ongelmaan. Analyysillä saadut tulokset ovat mielenkiintoisia, mutta vaikeasti tulkittavia. Tämä onkin kyseisen lähestymistavan suurin ongelma; tärkeiden piirteiden määrittämiseen ja painottamiseen ei ole olemassa yleistä käytäntöä. Lisäksi assosiaatioiden laskeminen tuhansille eri kappaleille on raskas toimenpide nykyaikaisienkin koneiden kanssa.
Dillon ja Hunter käyttivät seuraavia menetelmiä melodian muunnelmien automatisoituun tunnistamiseen **[DiH].** Helpon toteutuksen ja yleistettävyyden kannalta käytetään muutamia perusmääritelmiä; vain niitä musiikillisia ominaisuuksia käytetään analyysissä, jotka voidaan johtaa DARMS koodatusta musiikista. Ryhmittely operaatiot suunnitellaan samanlaisiksi kuten sanakirjojen prosessointi järjestelmät. Osien hakuun käytetään boolean pohjaista hakukieltä.
Melodian nuotit kirjattiin ylös sävelasteikon asteita kuvaavilla numeroilla; ”C D E F” kirjattaisiin ”1 2 3 4”. Oktaavin alapuolelle menevien nuottien eteen laitetaan L ja yläpuolelle U.
Melodian tunnistaminen jaetaan viiteen eri osa-alueeseen; kaksi osa-aluetta keskittyy melodioiden vertailuun sävelkorkeuksien mukaan. Toisessa näistä käytetään tahtierottelua ja toisessa ei. Kahdessa seuraavassa osa-alueessa verrataan melodioita ainoastaan painottuneiden sävelkorkeuksien vertailuun, joista toisessa ei käytetä tahtierottelua. Viimeinen vertailu perustuu melodian fraasien vertailuun.

3. Melodian mikroanalyysi

Stech**[Ste]** määrittelee sovelluksen, jolla voidaan abstrahoida tietyn tyyppisiä melodiamalleja lineaarisesta musiikista. Hän esittää, että sävellys koostuu pienistä musiikillisista ideoista, jotka esiintyvät useamminkin ja myös varioituina, mutta joiden toistuminen on todennäköisesti säveltäjän tiedostamatonta työskentelyä isomman kokonaisuuden suhteen.
Aikaisemmat sovellukset ovat rajoittuneet mm. erittäin spesifisen tiedon hankintaan tietyn säveltäjän tai läheisesti saman tyylisten säveltäjien musiikista. Suurin puute aikaisemmissa tutkimuksissa on niiden peruste laatia tietyn pituinen musiikillinen malli koko analyysin perustaksi. Mallin äänenkorkeuden ja rytmi lukitseminen sekä tulosten epäselvä esittäminen havaittiin myös puutteiksi. Tämän perusteella laadittiin sovellus, jossa voitiin analysoida musiikkia valitsemalla ennalta määritellyistä musiikillisista testeistä tietyt testit analyysiin sen perusteella minkä tyylistä musiikkia analysoidaan ja mikä on analyysin tarkoitus.
Sovellus koodaa äänenkorkeuden ja rytmin. Analyysiä varten tehtiin 26 erilaista melodiatestiä. Jokainen testi suunniteltiin löytämään tietyn tyyppisen melodisen relaation, jotka jaettiin kolmeen eri luokkaan; melodioiden identtinen tai samanlainen äänenkorkeus, melodioiden rytmin vertailu ja viimeisessä luokassa testit, joilla verrataan rytmiä ja äänenkorkeutta.
Äänenkorkeuden vertailuun käytettiin tuttuja melodian relaatioita; alkuperäinen(O, original), inversio(I, inversion) eli käänteinen melodiakulku, lopusta alkuun(R, retrograde), lopusta alkuun yhdistettynä käänteiseen(RI). Rytmiset relaatiot ovat; alkuperäisen rytmikuvion toisto, alkuperäisen rytmikuvion vertailu sen täsmälliseen lopusta alkuun vastineella.
Jotta toistuvien äänenkorkeuksien analyysille saatiin valinnanmahdollisuuksia O, I, R ja RI testeille annettiin kaksi muotoa; ”exact contour”, joilla mallit ovat samanlaisia vain jos kaikilla intervalli relaatioilla on samanlainen sekvenssi; ”tonal contour”, joilla mallit ovat samanlaisia vain melodisen ääriviivan mukaan, esim. fuuga teema tunnistetaan jos se esitetään tonaalisena vastauksena, mutta edellinen muoto havaitsisi vain teeman alkuperäisen muodon.

4. Musiikillisten sekvenssien vertailu

Jotta voidaan välttää sallittujen transformaatioiden ja vääristymien ennalta määrittelyä melodioita tulisi sen sijaan verrata kahden melodian välisenä etäisyytenä, joka on riippuvainen tonaalisesta ääriviivasta (contour) ja rytmisestä rakenteesta **[MoS].** Tällaisen etäisyyteen perustuvan analyysin etuna on sen mahdollisuus jakaa ne hierarkisesti tai esittää melodiat metrisesti kaksi- tai kolmiulotteisessa spatiaalisessa tilassa.
Perinteisesti sekvenssien vertailua on tehty samanmittaisten elementtien suhteen, mutta yleistys tästä olisi etsiä kahden elementin optimaalinen vastaavuus kaikkien niiden vastaavuuksien joukosta jotka toteuttavat jonkin ehdon, esim. elementtien järjestyksen identtisyyden.
Luonnollinen tapa kvantifioida kaksi sekvenssiä on laskea pienin transformaatioiden summa, joilla kaksi elementtiä vastaavat toisiaan. Tyypillisiä transformaatioita ovat; yksikön poistaminen, lisääminen ja vaihtaminen. Näihin muutoksiin voidaan liittää painoarvoja, joita käytetään melodian vertailun laskennassa.
Partituuri esitetään parien jonona, jossa ensimmäinen osa paria on nuotin äänenkorkeus ja toinen osa nuotin pituus sekunteina. Nuotin äänenkorkeus merkitään kuinka monta puolisävelaskeletta se on perussävelestä. Nuotin pituus taas merkitään kuudestoistaosa arvoina.
Transformaatioiden painoarvojen määrittäminen on hienovarainen toimenpide. Pieni painoarvo asetetaan kun korvataan nuotti oktaavia tai viidenneltä sävelaskeleelta, koska nämä ovat kolmisoinnun intervalleja. Jotta voidaan varmistaa, että painoarvot eivät ole riippuvaisia sävellajista, merkitään nuotti puolisävelaskeleina perussävelestä, esim. toinen sävel on G duurissa ja toinen G mollissa. Asetetaan painoarvo nolla nuotin B korvaamiselle alennetulla B:llä, koska molemmat nuotit ovat kolmansia askeleita sävellajiensa skaalalta.
Vertailtaessa sekvenssejä pitää sallia kaksi muutakin transformaatiotyyppiä, fragmentaatio ja konsolidointi. Fragmentoinnissa muutetaan esimerkiksi kokonuotti neljäksi saman äänenkorkeuden neljäsosanuotiksi. Konsolidointi taas on fragmentoinnin käänteinen transformaatio. Jos katsotaan tätä lisäyksien ja poistojen suhteen, painoarvo olisi korkeampi, mutta kun sallimme tällaiset toimenpiteet, fragmentaatio ja konsolidointi voidaan laskea yhdeksi transformaatioksi.
Kun kaikki transformaatio menetelmät kootaan yhteen saadaan kuvan 1 esittämä algoritmi kahden melodian eroavaisuuden laskentaan

$$d_{ij} = \min \begin{cases} d_{i-1,j} + w(a_i, 0) & \text{(poisto)} \\ d_{i,j-1} + w(0, b_j) & \text{(lisääminen)} \\ (d_{i-1,j-k} + w(a_i, b_{j-k+1}, \ldots, b_j),\ 2 \leqslant k \leqslant j) & \text{(fragmentaatio)} \\ (d_{i-k,j-1} + w(a_{i-k+1}, \ldots, a_i, b_j),\ 2 \leqslant k \leqslant j) & \text{(konsolidointi)} \\ d_{i-1,j-1} + w(a_i, b_j) & \text{(vaihto)} \end{cases}$$

Kuva 1.

missä, fragmentaatio ja konsolidointi funktioiden ”w(…)” ovat ennalta määrättyjä painoarvoja. Vaikkakin algoritmillä on monia hyviä puolia, kuten sen sävellaji ja tempo riippumattomuus, on sillä myös paljon heikkouksiakin. Melodia joka esiintyy partituurissa useita kertoja ja toisessa kappaleessa vain kerran, voidaan ”linkittää” vain yhteen edellisen partituurin monista esiintymistä. Jos kaksi säveltä ovat identtisiä, paitsi pienen osan esiintyminen toisen kappaleen alussa ja toisen lopussa, useat samankaltaisuudet jää löytämättä.

Viitteet

[Ste] Stech, A.D., A Computer-Assisted Approach to Micro-Analysis of Melodic Lines, Computers and the Humanities, 15:211-221, 1981

[DiH] Dillon, M., Hunter, M., Automated Identification of Melodic Variants in Folk Music, Computers and the Humanities, 16:107-117, 1982

[MoS] Mongeau, M., Sankoff, D., Comparison of Musical Sequences, Computers and the Humanities, 24:161-175, 1990