Musiikin klusterointimenetelmät

Anna Pienimäki

Tiivistelmä

Musiikkikorpuksien klusteroinnilla tarkoitetaan korpuksen kappaleiden ryhmittelyä näiden keskinäisten samankaltaisuuksien perusteella. Klusterointia voidaan hyödyntää niin soittolistojen generointiin kuin tyylilajitunnistukseenkin. Näiden lisäksi klusterointia voidaan käyttää musiikkianalyysin osana musiikkikorpuksen esiprosessoinnissa ja kappaleiden muotorakenteiden löytämisessä. Tässä esitelmässä esitellään musiikkidatan klusteroinnin vaatimuksia, menetelmiä sekä uusi musiikkikorpuksen klusterointimenetelmä, joka pohjaa musiikkidatan sisältämän rakenteen ja harmonian hyödyntämiseen kappaleiden samankaltaisuutta määritettäessä.

1 Johdanto

Yhteiskunnan digitalisoituminen on vaikuttanut voimakkaasti musiikin tallennusmuotoihin ja saatavuuteen. Digitaaliset tallenteet ovat 2000-luvulle tultaessa vallanneet länsimaiden markkina-alueen analogisilta tallenteilta. Koska digitaalisia äänen talletusmuotoja voidaan levittää myös Internetin välityksellä, on yleisesti saatavilla olevien laillisten ja laittomien digitaalisten musiikkitiedostojen määrä kasvanut räjähdysmäisesti. Internet tarjoaakin nykyisellään suunnattoman joukon myös laillista analysoimatonta musiikkiaineistoa.

Digitaaliset musiikkitiedostot voidaan luokitella karkeasti kahteen ryhmään sisällön tallennustavan perusteella: digitaalisiin suoriin ja symbolisiin representaatioihin. Digitaaliset suorat representaatiot sisältävät digitalisoitua signaalia joko pakatussa, kuten esimerkiksi mp3-tiedostot, tai pakkaamattomassa muodossa. Symbolisella representaatiolla puolestaan tarkoitetaan symbolisessa muodossa tallennettua tietoa musiikkikappaleesta. Digitaalisissa symbolisissa representaatioissa esitetään musiikkikappaleet yleensä joko numero- ja/tai kirjainsarjoina, jotka kuvastavat nuottien piirteitä, kuten sävelkorkeutta ja -kestoa, tai nuotteja kuvaavina vektoreina moniulotteisessa avaruudessa. Symbolista representaatiota voidaan hyödyntää analysoitaessa tarkemmin musiikkikappaleiden rakennetta, harmoniaa tai melodisia seikkoja. Symbolisen musiikkidatan käsittelyssä on tyypillisesti käytetty merkkijonomenetelmiä sekä nyttemmin myös geometrisia menetelmiä [8].

Etenkin suurten datajoukkojen ollessa kyseessä olisi hyödyllistä, jos datasta kyettäisiin erottelemaan nopeasti sellaiset osaset, joissa on jotain uutta, poikkeavaa ja mielenkiintoista. Tällaisen materiaalin etsimistä ja erottelua datajoukosta kutsutaan tietämyksen muodostamiseksi (data mining, knowledge discovery). Tietämyksen muodostamista on sovellettu pääasiassa tekstimuotoisen ja numeerisen datan analyysiin. Viime aikoina on kuitenkin herännyt myös pohdintoja sen menetelmien, kuten assosiaatiosääntöjen, hahmontunnistuksen ja klusteroinnin, musiikkianalyyttisista mahdollisuuksista.

Musiikkianalyysitehtävä koostuu tyypillisesti aineiston keruusta, sen ryhmittelystä tutkimuksen kannalta olennaisten piirteiden nojalla, tutkittavan ryhmän tai ryhmien tarkemmasta analyysista ja johtopäätösten muodostamisesta. Näistä tehtävistä sekä aineiston ryhmittelyyn että osittain sen tarkempaan analyysiin voidaan soveltaa automatisoitua analyysia. Musiikkianalyysin lisäksi musiikkidatan sisällöllistä analyysia on tehty myös musiikkitiedonhauksi kutsutun alan puitteissa.

Klusteroinnin tarkoituksena on ryhmitellä ohjastamattomasti syötteenä annettu data sen jakautumista mahdollisimman hyvin kuvastaviin ryhmiin. Ryhmittely pohjaa data-alkioiden välisiin similariteetti- eli samankaltaisuusarvoihin. Tavoitteena ryhmittelyssä on, että toisaalta ryhmien sisäisten data-alkioiden keskinäinen samankaltaisuus on mahdollisimman suuri ja toisaalta ryhmien väliset samankaltaisuudet ovat mahdollisimman pieniä. Samankaltaisuuksien laskemiseen on kehitetty runsas joukko erilaisia etäisyysmittoja ja metriikoita.

2 Musiikkidatan klusteroinnista

Musiikkidatan ensimmäiset klusterointimenetelmät kehitettiin 1960- ja 1970-luvuilla Bartokin kansanlaulujen luokittelukriteereiden perusteella [13, s. 57]. Sittemmin klusterointia on kohdistettu niin klassiseen [2] kuin kansanmusiikkiin [5], kokonaisten korpusten luokitteluun [14] ja yksittäisten kappaleiden paradigmaattiseen [2] analyysiin, symboliseen musiikkidataan [5], digitaaliseen musiikkisignaaliin [6, 11] sekä musiikin metatietoihin [10].

Klusteroitaessa dataa, on ensin pohdittava datan klusteroituvuutta: miten klusteroitavat data-alkiot esitetään vertailukelpoisesti ja millainen metriikka soveltuu kuvaamaan niiden välisiä etäisyyksiä. Signaalimuotoisesta musiikkidatasta on melko yksinkertaista muodostaa klusteroitavia data-alkioita ja laatia niille soveltuvia metriikoita. Suurin osa olemassa olevista musiikkidatan klusterointimenetelmistä kohdistetaankin signaalimuotoiseen dataan [6, 11, 14, 4]. Soittolistojen generointiin on lisäksi esitelty myös menetelmiä, jotka eivät ota kantaa musiikkikappaleiden sisältöön vaan hyödyntävät klusteroinnissa musiikkikappaleiden metatietoja [10], jotka luonteeltaan soveltuvat hyvin klusteroitavien data-alkioiden ulottuvuuksiksi.

Symbolinen musiikkidata on luonteeltaan hankala kummankin kysymyksen kannalta. Musiikkikappaleet saattavat olla hyvinkin eripituisia ja niiden samankaltaisuus ei ole todettavissa kappaleen nuotteja lineaarisesti tarkastelemalla. Symbolista musiikkidataa käsittelevät klusterointimenetelmät rajoittavat klusteroinnin tyypillisesti koskemaan vain yksiäänistä melodiaa tai sen osasia [5, 2, 1]. Menetelmissä, joissa muutkin äänet kuin melodia polyfonisesta kudoksesta otetaan huomioon, hajoitetaan kudos erillisiksi ääniksi, jotka analysoidaan erikseen [3].

Melodioiden muokkaukseen klusteroituviksi data-alkioiksi on käytetty pääasiassa kolmea lähestymistapaa: melodian häivyttämistä, melodian kuvausta ja melodian syväanalyysia. Melodian häivytykseen pyrkivissä menetelmissä joko melodiasta laaditaan tilastollinen malli, jonka arvot muodostavat melodiaa kuvaavan data-alkion ulottuvuudet [5], tai melodia tiivistetään [3]. Melodian kuvauksessa [1] melodia tallennetaan melko yksinkertaistetussa esitysmuodossa, kuten sävelkestoina ja korkeuksina, jolloin data-alkion ulottuvuuksina ovat melodian nuotit. Melodian syväanalyysissa puolestaan melodian kulku analysoidaan niin intervalli- kuin rytmiselläkin tasolla, jolloin kukin analyysitaso muodostaa melodiaa kuvailevan data-alkion yhden ulottuvuuden [2].

Etäisyysmittoina musiikkidatan klusteroinnissa on käytetty niin euklidista [1], Manhattan- [5] kuin Hamming-etäisyyttäkin [2, 3] joko sellaisenaan tai muokattuina versioina. Klusterointimenetelminä puolestaan on käytetty niin hierarkkisia [5, 3] kuin osittaviakin klusterointialgoritmeja [2] tai kokoelmaa erilaisia algoritmeja, joiden toimintaa on vertailtu [1] saman musiikkidatan puitteissa.

3 Musiikkidatan kaksitasoinen klusterointi

Musiikkidatan kaksitasoinen klusterointimenetelmä pohjaa generatiivisen ja semiologisen musiikintutkimuksen käsityksiin musiikin rakenteellisuudesta [9, 7]. Klusterointimenetelmä koostuu kahdesta osasta: musiikkitiedostojen analyysista, jonka perusteella laaditaan kullekin musiikkitiedostolle vektorimuotoinen esitys, ja näiden vektoriesitysten klusteroimisesta hierarkkista kokoavaa klusterointia käyttäen. Menetelmän ydin voidaan esittää algoritmisessa muodossa seuraavasti:

Algoritmi 4 Rakenteellisen klusteroinnin pääalgoritmi

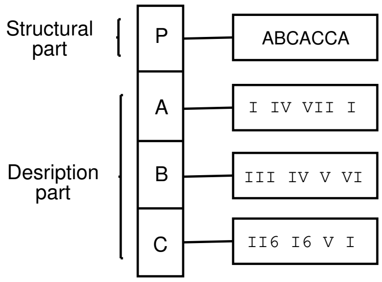
```
Syöte: hakemisto D sisältäen MIDI-tiedostoja
Tuloste: Datajoukon hierarkkista mallia kuvastava puurakenne

(1)   for i = 1,...,n tiedostoa hakemistossa D do
(2)     tee tiedostosta i yksiääninen esitys i'
(3)     ryhmittele i' melodiafraaseiksi
(4)     tee i:n nuoteille arvio niiden painollisuuksista
(5)     jaa i i':n perusteella osiksi g_j, j = 1,...,m
(6)     foreach g_j do
(7)        muodosta g_j:stä harmoninen analyysi h_j
(8)     end
(9)     klusteroi harmoniset analyysit h_j kolmeen klusteriin C_1,...,C_3
(10)    koosta kappaleesta i vektori klusteroinnin perusteella
(11)  end
(12)  klusteroi tiedostoja i kuvailevat vektorit
(13)  return datajoukon hierarkkinen malli ja kappalevektorilista
```

Menetelmän ensimmäisessä vaiheessa (rivit 1–11) kukin klusteroitavan hakemiston kappale muutetaan sisäiseen esitysmuotoon, josta puolestaan muodostetaan kappaleelle yksiääninen variantti poimimalla samanaikaisesti esiintyvistä nuoteista ne, joiden sävelkorkeus on korkein. Yksiääninen esitys analysoidaan ja ryhmitellään melodiafraaseiksi. Moniääniselle versiolle tehdään analyysi sen nuottien painollisuuksista ja jaetaan saatu tulos melodiafraasien rajojen perusteella osasiksi, joille kullekin tehdään harmoninen analyysi.

Fraasien harmonista analyysia kuvastavat vektorit klusteroidaan k:n keskusalkion menetelmällä kolmeen klusteriin, jotka nimetään kirjaimilla A, B ja C. Fraasien etäisyys klusteroinnissa lasketaan muokattuna editointietäisyytenä:

$$d_H(h_1, h_2) = min(D + I + R),$$



Kuva 1: Musiikkikappaletta kuvaava vektori.

jossa etsitään minimiä poistojen D, lisäysten I ja vaihdosten R yhteismäärästä, joilla astejonosta h_1 saadaan muokattua astejono h_2. Kullekin operaatiolle määritellään kustannus, joka lisäyksen ja poiston yhteydessä on 1. Vaihto puolestaan jaotellaan viiteen eri samankaltaisuusasteeseen: täsmäykseen, käännökseen, soinnunasteluokan sisäiseen vaihtoon, duuri-molli-vaihtoon ja totaaliseen vaihtoon. Näiden kustannukset ovat seuraavat:

- Jos $h_{1_i} = h_{2_j}$, kustannus on 0
- Jos h_{1_i} on saman soinnunasteen kuin h_{2_j} eri käännös, kustannus on 0.25
- Jos h_{1_i} kuuluu samaan soinnunasteluokkaan kuin h_{2_j}, kustannus on 0.5
- Jos h_{1_i} on sama soinnunaste mollissa kuin h_{2_j} duurissa tai toisin päin, kustannus on 0.75
- Jos yksikään edellämainituista ehdoista ei toteudu, kustannus on 1

Soinnunasteluokat puolestaan määritellään seuraavasti: toonikaluokkaan kuuluvat asteet I ja VI, subdominanttiluokkaan asteet II ja IV, dominantti-luokkaan asteet V ja VII ja muut-luokkaan aste III (ks. esim [12]).

Kun fraasit on klusteroitu, voidaan kappaleesta muodostaa vektori (kuva 1), jonka ensimmäisenä alkiona on kappaleen muotorakenne-esitys: järjestetty jono kappaleen fraaseja esitettynä sillä kirjaimella, mihin klusteriin fraasi sisältyy. Muut vektorin alkiot puolestaan sisältävät kuvauksen kustakin klusterista, käytännössä tämän keskusalkion.

Klusteroitaessa kappalevektoreita joudutaan etäisyyden laskeminen määrittelemään uudelleen hyödyntäen kuitenkin sekä editointietäisyyden perusmallia, että muokattua editointietäisyyttä. Musiikkikappaleita kuvailevien vektoreiden välinen etäisyys muodostetaan kahdesta tekijästä: kappaleen muotorakennetta kuvaavien merkkijonojen editointietäisyyksistä ja kunkin fraasin harmoniaa kuvaavien vektoreiden muokatuista editointietäisyyksistä:

$$d_S(s_1, s_2) = min_{s_1 s_2, s_2 s_1}(D + E + mE + \sqrt{d_H(P_1)^2 + ... + d_H(P_n)^2}),$$

jossa $min_{s_1 s_2, s_2 s_1}$ valikoi sen operaatiosuunnan, joka tuottaa pienemmän etäisyysarvon, D on muotorakenteiden välinen Levenshteinin editointietäisyys, E on painotettujen muotorakenteiden välinen editointietäisyys, mE on etäisyysmuokattujen painotettujen muotorakenteiden välinen editointietäisyys ja $d_H(P_1)...d_H(P_n)$ ovat toisiaan lähinnä olevien harmoniavektoreiden välisiä etäisyysarvoja.

Muotorakenteiden painotuksessa pidemmän muotorakenteen pituus jaetaan lyhemmällä ja jos jaon kokonaislukuosuus on yhtä suurempi, kerrotaan lyhyempi muoto kokonaislukuosuudella. Nyt editointietäisyys voidaan laskea käyttäen lyhemmän hahmon sijasta sen pidennettyä versiota. Etäisyysmuokatulla muotorakenteella puolestaan tarkoitetaan sellaista muotorakenteen versiota, joka on muodostettu tarkastelemalla verrattavien harmoniavektoreiden läheisyyttä toisiinsa.

Varsinainen klusterointi suoritetaan hierarkkista kokoavaa klusterointimenetelmää käyttäen, jolloin tuloksena on pelkkien klustereiden ja näiden keskusalkioiden sijaan koko musiikkikappaleita sisältävän hakemiston sisällön kuvaus puumuodossa.

4 Yhteenveto

Musiikillisen datan klusterointia on käytetty lähinnä soittolistojen muodostamiseen ja tyylisuuntien arviointiin. Klusterointia voidaan kuitenkin hyödyntää myös automatisoidun musiikkianalyysin apuvälineenä. Symbolisen musiikkidatan klusteroinnissa voidaan hyödyntää tietoutta musiikkidatan sisältämien melodisten ainesten lisäksi sen rakenteellisuudesta, harmonisista ja tilastollisista piirteistä.

Symbolisen musiikkidatan kaksitasoinen klusterointimenetelmä pohjaa musiikin generatiiviseen ja semiologiseen analyysiin. Kutakin klusteroitavaa musiikkikappaletta kohden muodostetaan vektori, joka sisältää kappaleen paradigmaattisen tason lisäksi kunkin paradigman sisällöllisen kuvauksen. Musiikkikappaleita kuvaavien vektoreiden etäisyys lasketaan syvä- ja pintatasojen samankaltaisuuksien avulla ja näiden klusterointiin käytetään hierarkkista klusterointimenetelmää.

Viitteet

[1] C. Agnastopoulou, K. Höthker ja D. Hörnel. Investigating the Influence of Representation and Algorithms in Music Classification. *Computers and the Humanities 35(1)*, s. 65–79, 2001.

[2] E. Cambouropoulos. *Towards a General Computational Theory of Musical Structure.* PhD thesis, University of Edinburgh, Faculty of Music, Department of Artificial Intelligence, 1998.

[3] R. Cilibrasi, P. Vitanyi, and R. de Wolf. *Algorithmic Clustering of Music.* Technical paper, Computer Research Repository (CoRR), 2003.

[4] S. Dixon, E. Pampalk ja G. Widmer. Classification of Dance Music by Periodicity Patterns. Proc. ISMIR'03 Fourth International Conference on Music Information Retrieval, Baltimore, Maryland, October 26–30, s. 159–165, 2003.

[5] T. Eerola, T. Järvinen, J. Louhivuori ja P. Toiviainen. Statistical Features and Perceived Similarity of Folk Melodies. *Music Perception*, 18, s. 275–296, 2001.

[6] M. Grimaldi, A. Kokaram ja P. Cunningham. *Classifying music by genre using the wavelet packet transform and a round-robin ensemle.* Technical report TCD-CS-2002-64, Trinity College Dublin, 2002.

[7] F. Lerdahl ja R. Jackendoff: A Generative Theory of Tonal Music. MIT Press, Cambridge, Mass., 1983.

[8] D. Meredith, G. A. Wiggins ja K. Lemström: Pattern Induction and Matching in Polyphonic Music and Other Multi-Dimensional Datasets. *Proc. the 5th World Multi-Conference on Systemics, Cybernetics and Informatics (SCI2001)*, Volume X, s. 61 – 66, Orlando, Florida, July 22 – 25, 2001.

[9] J.-J. Nattiez: Music and Discourse – Toward a Semiology of Music. Princeton University Press, 1990.

[10] S. Pauws ja B. Eggen. PATS: Realization and User Evaluation of an Automatic Playlist Generator. *Proc. ISMIR'02 Third International Conference on Music Information Retrieval*, Paris, France, October 13-17, s. 222–230, 2002.

[11] A. Rauber, E. Pampalk ja D. Merkl: Using Psycho-Acoustic Models and Self-Organizing Maps to Create a Hierarchical Structuring of Music by Sound Similarity. *Proc. ISMIR'02 Third International Conference on Music Information Retrieval*, Paris, France, October 13-17, s. 71–80, 2002.

[12] E. Salmenhaara: *Sointuanalyysi.* Otava, 1997.

[13] N. Schuler. Methods of Computer-Assisted Music Analysis: History, Classification, and Evaluation. Michigan State University, PhD Thesis.

[14] W.-H. Tsai, H.-M. Wang, D. Rodgers, S.-S. Cheng ja H.-M. Yu. Blind Clustering of Popular Music Recordings Based on Singer Voice Characteristics. *Proc. ISMIR'03 Fourth International Conference on Music Information Retrieval*, Baltimore, Maryland, October 26–30, s. 167–173, 2003.