Pienten joustavien molekyylien geometrinen hajautus

Juho Muhonen
juho.muhonen@helsinki.fi

Tiedon louhinta biomolekyyliaineistoista
Helsingin yliopisto, tietojenkäsittelytieteen laitos
Raportti C–2003–52, s. 62–68, marraskuu 2003

Tiivistelmä

Raportissa käsitellään pienten joustavien molekyylien tunnistamista geometrisen hajautuksen avulla. Geometrinen hajautus on mallipohjainen hahmontunnistusmenetelmä, jossa hajautus ja etsintä perustuu lokaalin koordinaatiston valintaan. Lokaalilla koordinaatistolla taataan, että mallin esitys ei muutu, vaikka sen orientaatio muuttuisi globaalissa koordinaatistossa. Lokaalissa koordinaatistossa esitetystä mallista valitaan invariantteja, joiden avulla hajautus tehdään. Koska lokaalin koordinaatiston kanta voidaan esittää usealla eri tavalla, mallit joudutaan hajauttamaan jokaisessa mahdollisessa kannassa. Hahmon tunnistus tapahtuu hajauttamalla hahmo samaan tapaan kuin mallit, ja vertaamalla hajautuksen tulosta mallitietokannasta muodostettuun hajautustauluun. Pienillä joustavilla molekyyleillä tarkoitetaan molekyylejä, jotka sisältävät pyöriviä sidoksia. Pyörivien sidosten takia molekyylit joudutaan jakamaan osiin, joissa lokaali koordinaatisto säilyy eheänä, vaikka molekyylin jokin pyörivä sidos muuttaisi asentoaan. Yksittäisten osien etsinnän jälkeen osat yhdistetään mahdollisimman suureksi molekyyliksi vertaamalla osista lähtevien pyörivien sidosten globaaleja asentoja.

1 Johdanto

Geometrinen hajautus on hahmontunnistusmenetelmä, joka on alun perin kehitetty tietokonenäön avuksi. Se perustuu siihen ajatukseen, että mikä tahansa kolmiulotteinen malli säilyttää tiettyjä geometrisia ominaisuuksia omassa paikallisessa koordinaatistossaan, vaikka sen orientaatio muuttuisikin havainnoitsijaan nähden. Näiden ominaisuuksien perusteella voidaan geometrisesti hajautetusta tietokannasta hakea malleja tehokkaasti, eikä tietokannan jokaista alkiota tarvitse erikseen käsitellä. Se, miten hajautuksessa käytettävät geometriset ominaisuudet valitaan, riippuu hajautettavasta aineistosta.

Vaikka geometrinen hajautus alun perin kehiteltiin tietokonenäön avuksi esimerkiksi tilanteeseen, jossa robotin tulee tunnistaa työkalu saamansa kuvan perusteella. Se on tehokas hahmontunnistusmenetelmä missä tahansa tilanteessa, jossa olemassa olevasta tietokannasta halutaan löytää malli geometristen ominaisuuksien perusteella. Tehokkuus saavutetaan massiivisen esikäsittelyvaiheen avulla. Esikäsittelyssä geometrinen hajautus esittää tietokannan mallit useassa eri kannassa. Tämä tarkoittaa käytännössä sitä, että geometrisesti hajautettu tietokanta sisältää redundanttia tietoa, jonka avulla etsintä tunnistaa missä tahansa asennossa olevan hahmon. Redundantti esitys mahdollistaa myös hahmontunnistuksen vaikka osa hahmosta olisi peitossa.

Proteiinien toiminnan perustuminen suurelta osin niiden kolmiulotteiseen rakenteeseen motivoi vertailemaan lääkemolekyylejä geometrisista lähtökohdista. Geometrinen hajautus löytää suuristakin lääketietokannoista annetun lääkemolekyylin kaltaisia molekyylejä suhteellisen nopeasti.

Tässä kirjoituksessa esitellään ensin luvussa 2 geometrisen hajautuksen periaate ja toteutus yleisellä tasolla. Tämän jälkeen luvussa 3 selitetään pienen joustavan molekyylin käsite ja näytetään, miten geometrista hajautusta voidaan käyttää avuksi etsittäessä annetun molekyylin kaltaisia molekyylejä.

2 Geometrinen hajautus

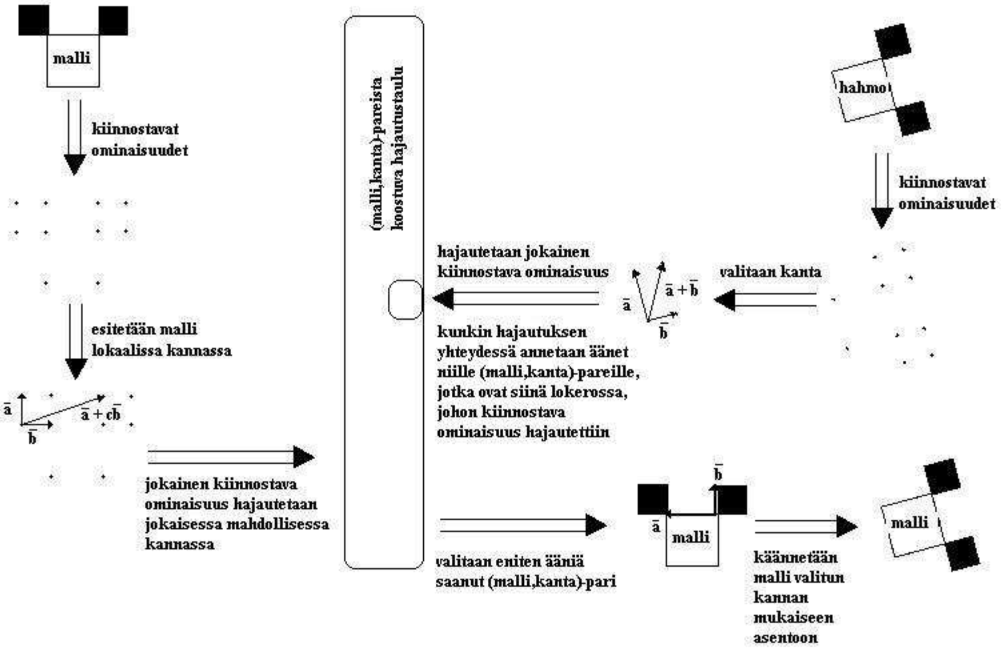
Mallipohjaisen (Model-Based) hahmontunnistuksen perusongelma voidaan formuloida seuraavasti: On annettu *mallien* d_i joukko D sekä *hahmo* c. Ongelmana on löytää joukosta D se malli d_c, joka parhaiten vastaa hahmoa c. Geometrinen hajautus on mallipohjainen hahmontunnistusmenetelmä. Malleina ja hahmona ovat geometriset objektit, joiden orientaatio voi vaihdella globaalissa koordinaatistossa, mutta jotka säilyttävät omassa koordinaatistossaan joitakin ominaisuuksia globaalista orientaatiosta riippumatta.

Geometrinen hajautus [LaW88, WoR97] jakautuu oleellisesti kahteen vaiheeseen. Ensimmäisessä vaiheessa algoritmi hajauttaa kaikki mallit invarienttien ominaisuuksiensa perusteella suureksi hajautustauluksi, ja toisessa vaiheessa algoritmi etsii hahmoa muistuttavia malleja hajauttamalla vuorostaan hahmon. Kuva 1 esittää yksinkertaistettua kaaviota menetelmästä. Tässä luvussa käsitellään ensin geometrisen hajautuksen esikäsittelyvaihetta, minkä jälkeen esitetään, miten algoritmi löytää esikäsiteltyä aineistoa käyttäen annettua hahmoa lähellä olevan mallin tai malleja.

2.1 Esikäsittely

Geometrisen hajautuksen tehokkuus perustuu etukäteen tehtävään laskentaan, jolla vähennetään varsinaisessa hahmontunnistuksessa kuluvaa aikaa muistinkäytön kustannuksella [WoR97]. Käytännössä tämä tarkoittaa sitä, että algortimi hajauttaa mallit etukäteen hajautustauluun, joka sisältää malleista redundanttia tietoa. Esikäsittelyvaiheessa algoritmi etsii jokaisesta mallijoukon D mallista d_i *kiinnostavat ominaisuudet* P_{d_i}, jotka voivat olla esimerkiksi pisteitä, viivoja tai muita vastaavia geometrisia ominaisuuksia [LaW88]. Algoritmilla on etukäteen tiedossa, kuinka monta ominaisuutta $p_k^{d_i} \in B$ tarvitaan, että mallin muut kiinnostavat ominaisuudet $p_l^{d_i} \notin B$ voidaan esittää ominaisuuksien $p_k^{d_i}$ avulla: $p_l^{d_i} = \sum \alpha_k * p_k^{d_i}$. Kiinnostavien ominaisuuksien joukon P_{d_i} osajoukkoa B kutsutaan *kannaksi*. Jokaista kantaa kohden algoritmi lisää hajautustauluun *(malli,kanta)-alkion* (d_i, B) kutakin kantaan kuulumatonta kiinnostavaa ominaisuutta $p_l^{d_i}$ vastaamaan. Algoritmi laskee alkion hajautusluvun kertoimista α_k.

Algoritmi lisää hajautustauluun (malli,kanta)-pareja, jotka sijoitetaan tauluun kiinnostavan ominaisuuden parin kantaesityksen mukaiseen paikkaan. Järjestelystä seuraa, että ku-



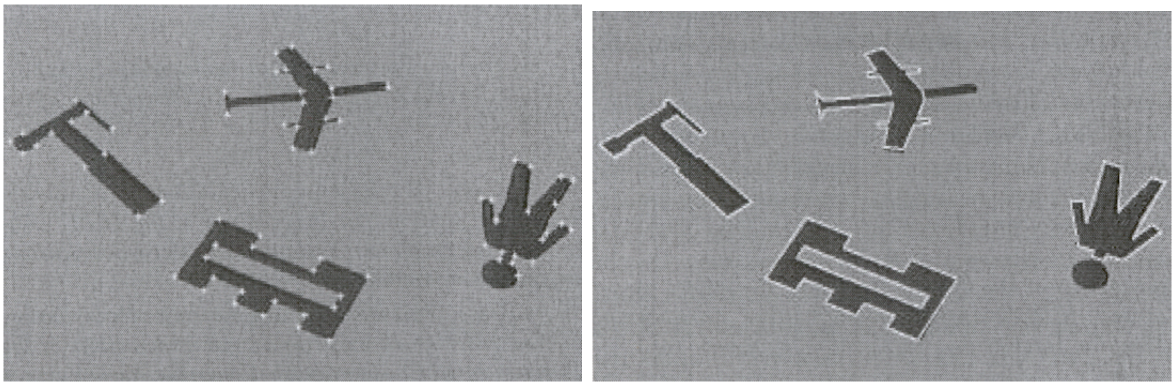
Kuva 1: Kuvassa on kaavio geometrisen hajautuksen eri vaiheista. Se antaa yleiskuvan geometrisen hajautuksen toiminnasta. Hajautustaulun vasemmalla puolella on lueteltu esikäsittelyn eri vaiheet, ja oikealla puolen etsinnän vaiheet. Kaavio on yksinkertaistus menetelmästä, eikä kaaviossa oleva hahmo koostu kuin yhdestä mallista. Hahmo ei myöskään ole lainkaan peitossa, eikä kaavioon ole merkitty tilannetta, jossa kanta jouduttaisiin valitsemaan uudelleen.

takin kiinnostavaa ominaisuutta vastaa hajautustaulussa useita pareja, riippuen siitä, mitkä kiinnostavat ominaisuudet kulloinkin toimivat kantana. Tämä redundantti esitys mahdollistaa hahmon etsimisen, vaikka hahmo olisikin osittain peitossa tai sen orientaatiota ei tunnettaisi.

2.2 Hahmon etsiminen

Kun algoritmi on tehnyt mallien esikäsittelyn, etsintäalgoritmi etsii hahmon c kaltaisia malleja hajautustaulusta hajauttamalla hahmon c samaan tyyliin kuin yksittäiset mallit d_i. Hahmolla tarkoitetaan tässä yhdestä tai useammasta mallista koostuvaa kokoelmaa. Osa hahmon malleista voi myös olla osittain näkymättömissä [LaW88]. Ensin algoritmi etsii hahmosta c kiinnostavat ominaisuudet P_c. Tämän jälkeen algoritmi valitsee satunnaisesti $|B|$ kiinnostavaa piirrettä kannaksi B, minkä jälkeen se laskee kantaan kuulumattomattomien ominaisuuksien $P_c - B$ esityksen valitussa kannassa. Kutakin kantaan kuulumatonta piirrettä kohti p_l^c algoritmi etsii hajautustaulun lokeron, johon piirre kantaa B käyttäen kuuluisi. Kun algoritmi on löytänyt lokeron, jokaiselle lokerossa olevalle (malli,kanta)-parille annetaan *ääni* (vote). Lopulta algoritmi etsii sen (malli,kanta)-parin, joka on saanut eniten ääniä, tai vaihtoehtoisesti ne parit, joiden äänimäärä ylittää jonkin tietyn kynnysarvon. Mikäli algoritmi ei löydä paria, jolla olisi riittävän korkea äänimäärä, se valitsee piirteistä P_c tähän mennessä käyttämättömän kannan ja suorittaa hajautuksen uudestaan tämän kannan perusteella. Kun etsintäalgoritmi on löytänyt äänimäärävaatimukset tyydyttävän (malli,kanta)-parin (d_c, B_c), se laskee transformaation, jolla parin kanta B_c asettuu etsinnässä käytetyn kannan B mukaisesti. Tämä transformaatio ilmaisee mallin d_c orientaation annetussa hahmossa c.

Mikäli tiedetään, että hahmo voi sisältää useita malleja, voidaan algoritmi asettaa jatkamaan mallien etsintää antamalla sen poistaa kiinnostavien ominaisuuksien joukosta P_c



(a) ***Hahmo*** (b) ***Löydetyt mallit***

Kuva 2: Kuvassa (a) on kaksiulotteisen avaruuden hahmo. Kuvaan (a) on merkitty geometrisen hajautuksen etsintäalgoritmin löytämät kiinnostavat ominaisuudet. Kuvassa (b) puolestaan on asetettu hahmoon algoritmin löytämät mallit. Lentokoneen kärki ja henkilön pää eivät kuulu malleihin [Ols94].

ne ominaisuudet, jotka kuuluvat löydettyyn malliin, ja käyttämällä jäljelle jääviä piirteitä uutena hahmona etsinnässä [LaW88]. Algoritmi laskee tässä tapauksessa mallin d_c ominaisuuksien sijainnit globaalissa koordinaatistossa ja poistaa vastaavat ominaisuudet joukosta P_c. Kuvassa 2 on annettu esimerkki hahmosta, ja algoritmin löytämistä malleista kaksiulotteisessa koordinaatistossa.

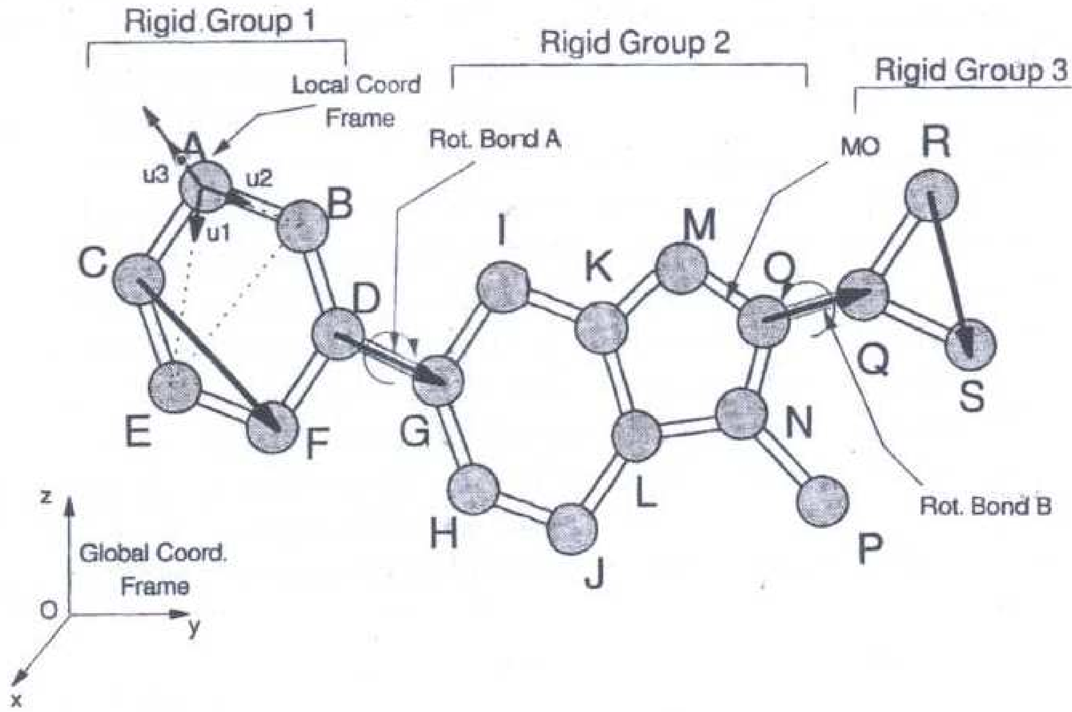
3 Pienet joustavat molekyylit

Pienillä joustavilla molekyyleillä tarkoitetaan molekyylejä, jotka koostuvat yhdestä tai useammasta *jäykästä ryhmästä* (rigid group), jotka liittyvät toisiinsa akselinsa ympäri *pyörivillä sidoksilla*. Jäykkä ryhmä on ryhmä atomeja, joiden sijainti toistensa suhteen pysyy muuttumattomana. Akselinsa ympäri pyörivä sidos puolestaan tarkoittaa sidosta, joka yhdistää kaksi jäykkää ryhmää toisiinsa, mutta sallii näiden pyörimisen toisiinsa nähden sidoksen suuntaisen akselin ympäri. Kuvassa 3 on esitetty esimerkki pienestä joustavasta molekyylistä. Molekyylin mallinnuksesta kerrotaan lisää luvussa 3.1.

Yksi lääketieteen tutkimusaloista on uusien lääkkeiden kehitys ja etsintä. Etsintää voidaan tehdä olemassa olevien tietokantojen avulla esimerkiksi tutkimalla lääkkeen vaikutusten ja sen kolmiulotteisen mallin välistä suhdetta. Tässä luvussa kerrotaan, kuinka pieniä joustavia molekyylejä voidaan etsiä tietokannasta niiden kolmiulotteisen rakenteen avulla käyttäen geometrista hajautusta. Esikäsittely ja etsintä eroavat hieman aiemmin kuvaillusta geometrisen hajautukseksen esityksestä, mutta perusidea säilyy.

3.1 Mallinnus

Ennen kuin molekyylitietokanta voidaan esikäsitellä geometrisen hajautuksen vaatimalla tavalla, yksittäiset molekyylit täytyy mallintaa siten, että molekyylin kiinnostavat ominaisuudet ja geometriset invariantit voidaan määrätä. Ilmeinen valinta molekyylin kiinnostaviksi ominaisuuksiksi ovat yksittäisten atomien sijainnit. Invarianttien valinnassa ei kuitenkaan riitä molekyylin sisäisen koordinaatiston kiinnitys kolmen atomin sijainnin avulla, koska molekyylissä mahdollisesti olevat pyörivät sidokset rikkovat sisäisen koordinaatiston [WSS99]. Pyörivien sidosten vuoksi lokaalin koordinaatiston kanta-atomien tulee sijaita samassa jäy-



Kuva 3: Kuvassa on esimerkkimallinnus pienestä joustavasta molekyylistä. Molekyylin atomit on nimetty aakkosin ja sidoksia vastaavat atomiparit. Kuvassa on kolme jäykkää ryhmää $rg_1 = \{A, B, C, D, E, F\}$, $rg_2 = \{G, H, I, J, K, L, M, N, O, P\}$ ja $rg_3 = \{Q, R, S\}$. Lisäksi kuvaan on merkitty jäykän ryhmän rg_1 kannan $\{A, E, B\}$ muodostama paikallinen koordinaatisto sekä akselinsa ympäri pyörivät sidokset (D, G) ja (O, Q) [WSS99].

kässä ryhmässä. Tällöin jäykkään ryhmään kuuluvat atomit voidaan esittää valitussa lokaalissa koordinaatistossa ja jäykän ryhmän atomien sijainti on invariantti niin globaalissa koordinaatistossa tapahtuvien siirtojen ja pyöritysten kuin molekyylin koordinaatistossa tapahtuvien sidosten pyöritysten suhteen [WSS99].

Edellä esitetty lokaalin koordinaatiston kannan valinta johtaisi käytännössä tilanteeseen, jossa molekyyli olisi hahmo, jonka malleina olisivat jäykät ryhmät. Tavoitteena on kuitenkin kohdella molekyyliä hahmona, joka koostuu yhdestä mallista. Tämän seurauksena jäykkien ryhmien välille täytyy luoda yhteys, jonka avulla ryhmät voidaan jotenkin yhdistää. Otetaan käyttöön *laajennetun jäykän ryhmän* käsite (augmented rigid group). Laajennettu jäykkä ryhmä on jäykkä ryhmä, johon on otettu lisäksi mukaan jäykästä ryhmästä lähtevien pyörivien sidosten päässä olevat atomit. Esimerkiksi kuvan 3 jäykkää ryhmää rg_2 vastaa laajennettu jäykkä ryhmä $rg_2^A = \{D, G, H, I, J, K, L, M, N, O, P, Q\}$. Alkuperäiseen ryhmään on lisätty atomit D sekä Q. Tämä laajennos ei vaikuta jäykän ryhmän geometrisiin invariantteihin. Nyt jokaisesta laajennetusta ryhmästä lähtevä pyörivä sidos voidaan ilmaista laajennetun jäykän ryhmän paikallisessa koordinaatistossa [WSS99].

3.2 Esikäsittely

Geometrisen hajautuksen käyttäminen vaatii molekyylitietokannan hajautusta (malli,kanta)-pareiksi. Esikäsittelyalgoritmi eroaa hieman geometrisen hajautuksen esikäsittelyalgoritmista (malli,kanta)-pareja muodostettaessa. Käyttäen formalismia, joka esiteltiin luvussa 2.1, ja edellisessä luvussa muodostettua molekyylin mallia geometrisen hajautuksen esikäsittelyvaihe toteutetaan pienten joustavien molekyylien tietokannalle seuraavasti. Mallia d_i vastaa molekyylin m laajennettu jäykkä ryhmä $rg^A_{m_j}$. Malli sisältää molekyylin tunnisteen m, jäykän ryhmän tunnisteen j, laajennetun jäykän ryhmän sisältävät pyörivät sidokset sekä atomien tunnisteet ja koordinaatit lokaalissa koordinaatistossa. Malli on siis järjestetty nelikko $(m, j, V, rg^A_{m_j})$, missä $V = \{(X, Y)_i\}$ tarkoittaa pyörivien sidosten joukkoa ilmaistuna atomien X ja Y avulla. Mallin kiinnostavia ominaisuuksia P_{d_i} vastaavat atomien sijainnit. Kiinnostavia ominaisuuksia tarvitaan kolme, jotta jäykän ryhmän lokaali koordinaatisto saadaan yksikäsitteisesti määrätyksi. Kolmen atomin sijainti riittää määräämään tason ja siis koordinaatiston kolmiulotteisessa avaruudessa [LaW88]. Kanta muodostuu siis kolmen jäykkään ryhmään kuuluvan atomin sijainnista. Algoritmi ei kuitenkaan lisää hajautusalkiota erikseen jokaiselle kiinnostavalle ominaisuudelle jokaisessa kannassa vaan ainostaan jokaiselle kannalle, koska hahmon kaikki kiinnostavat ominaisuudet ovat saatavilla, eikä hahmo voi olla osittain peitossa. Algoritmi laskee hajautusalkion hajautusluvun kannan ominaisuuksista. Hajautuslukua laskettaessa voidaan käyttää esimerkiksi kannan atomien välisiä kulmia ja etäisyyksiä [WSS99].

Kannattaa huomioida, että hajautustauluun talletettava (malli,kanta)-pari määrittelee yksikäsitteisesti mallin kuvaaman jäykän ryhmän sijainnin ja orientaation globaalissa koordinaatistossa, jos kannan sijainti tiedetään. Lisäksi parista voidaan laskea, mihin suuntaan mallin pyörivät sidokset osoittavat. Tätä tietoa käytetään myöhemmin, kun laajennettuja jäykkiä ryhmiä kootaan suuremmiksi kokonaisuuksiksi etsittäessä hahmoa vastaavaa molekyyliä [WSS99].

3.3 Etsintä

Pienten joustavien molekyylien etsintä edellä kuvatun kaltaisesta hajautetusta tietokannasta poikkeaa hieman luvussa 2.2 esitellystä etsinnästä. Tämä johtuu siitä, että vaikka hahmo koostuu esitetyssä formalismissa useammasta mallista eli laajennetusta jäykästä ryhmästä, niin etsintä kohdistuu mallien joukkoon eli molekyyliin. Löydetyt mallit yritetään yhdistää molekyyliksi tarkastelemalla pyörivien sidosten sijoittumista.

Luvun 2.2 formalismissa hahmon tunnistus tapahtuu seuraavasti. Etsittävä hahmo c on annettu molekyyli. Hahmon kiinnostavat piirteet ovat atomien sijainnit $P_c = \cup R_{m_j}$. Molekyylistä ei ole tarpeen erottaa erikseen jäykkiä ryhmiä, koska usean jäykän ryhmän atomeista muodostetut kannat eivät vastaa hajautettuja kantoja kuin sattumalta. Algoritmi valitsee satunnaisesti kolme atomia kannaksi B. Seuraavassa askeleessa algoritmi poikkeaa aikaisemmin esitetystä geometrisen hajautuksen etsintäalgoritmista. Sen sijaan, että se esittäisi kunkin kantaan kuulumattoman atomin valitussa kannassa B, se laskee hajautusarvon suoraan kannan B avulla. Tämän jälkeen algoritmi muodostaa kustakin hajautuslokerossa olevaa (malli,kanta)-paria $(d_i, B^{d_i}_j)$ vastaavan *hypoteesin*. Hypoteesi on parin malli asetettuna kannan B mukaiseen asentoon eli laajennettu jäykkä ryhmä asetettuna tiettyyn orientaatioon globaalissa koordinaatistossa [WSS99]. Algoritmi muodostaa hypoteesin asettamalla parin kannan $B^{d_i}_j$ kannan B mukaiseen asentoon. Jos algoritmi on jo muodostanut saman hypoteesin aikaisemmin, se antaa hypoteesille yhden lisä-äänen. Hypoteesien samanlaisuutta tulee tarkastella tiettyjen parametrien valossa. Parametrit ilmaisevat, kuinka paljon hypoteesien sallitaan olla eri asennossa globaaleissa koordinaateissa, että ne tulkitaan samaksi hypoteesiksi [WSS99].

Kun algoritmi on käsitellyt kaikki hahmon kolmen atomin kannat, sillä on on käytettävissään joukko hypoteeseja. Hypoteesien atomien ja pyörivien sidosten koordinaatit ovat ilmaistuna globaalissa koordinaatistossa. Algoritmi järjestää hypoteesit molekyylien, jäykkien ryhmien ja lopulta äänien mukaiseen järjestykseen. Tämän jälkeen algoritmi tarkastelee kunkin molekyylin kohdalla, ovatko vierekkäiset jäykät ryhmät asettuneet globaaliin koordinaatistoon siten, että niiden vastaavat pyörivät sidokset voitaisiin yhdistää. Lopulta, kun jokaisen molekyylin kaikki mahdolliset jäykät ryhmät on yhdistetty, algoritmi palauttaa muodostamansa molekyylit tai osittaiset molekyylit tapauksissa, joissa se ei voinut liittää kaikkia molekyylin jäykkiä ryhmiä. Jäykkien ryhmien saama äänimäärä korreloi suoraan sen kanssa, kuinka hyvä geometrinen peitto hahmon ja palautetun molekyylien välillä on. Äänimäärälle voidaan asettaa jokin kynnysarvo, jonka perusteella päätetään, mitkä molekyylit algoritmi palauttaa [WSS99].

4 Yhteenveto

Geometrinen hajautus on menetelmä, joka sopii mallipohjaiseen hahmonetsintään suurista tietokannoista, jotka sisältävät geometrisia malleja. Se koostuu kahdesta vaiheesta: esikäsittelystä ja hahmon etsinnästä. Esikäsittelyvaiheessa geometrinen hajautus käyttää runsaasti aikaa mallien muokkaamiseen muotoon, jossa ne on esitetty useissa mahdollisissa mallin lokaaleissa koordinaatistoissa. Esikäsittelyvaiheen hyödyt näkyvät etsintävaiheessa, jossa hahmo voidaan tunnistaa riippumatta sen asennosta globaalissa koordinaatistossa. Esikäsittelyn ansiosta geometrisen hajautuksen etsintävaihe on huomattavasti nopeampi kuin etsintäalgoritmeissa, joissa saatua hahmoa verrataan vuorotellen kaikkiin tietokannasta lyötyviin malleihin, koska malli löytyy hajautuksen avulla.

Geometrinen hajautus sopii pienin muutoksin pienten joustavien molekyylien etsintään esimerkiksi lääketietokannoista. Menetelmässä molekyyli jaetaan pienempiin jäykkiin osiin, jotka hajautetaan esikäsittelyssä. Etsinnässä saatavat jäykät osat yhdistetään yhteisten sidosten avulla mahdollisimman suuriksi molekyyleiksi. Menetelmä ei ota huomioon molekyylien mahdollisia muita ominaisuuksia, vaan perustuu ainoastaan niiden kolmiulotteisen rakenteen tarkasteluun.

Viitteet

[LaW88] Lamdan, Y. ja Wolfson, H. J., Geometric hashing: A general and efficient model-based recognition scheme. *Second International Conference on Computer Vision*, 1988, sivut 223–249. `http://ieeexplore.ieee.org/iel2/4606/12966/00589995.pdf`. [13.10.2003]

[Ols94] Olson, C. F., Time and space efficient pose clustering. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Seattle, Washington, 1994, sivut 251–258. `citeseer.nj.nec.com/article/olson93time.html`. [30.9.2003]

[WSS99] Rigoutsos, I., Platt, D., Califano, A. ja Silverman, D., Representation and matching of small flexible molecules in large databases of 3d molecular information. Teoksessa *Pattern Discovery in Biomolecular Data: Tools, Techniques and Applications*, Wang, J. T. L., Shapiro, B. A. ja Shasha, D., toimittajat, Oxford University Press, 1999, sivut 111–129.

[WoR97] Wolfson, H. J. ja Rigoutsos, I., Geometric hashing: An overview. *IEEE Computational Science & Engineering*, 4,4(1997), sivut 10–21. `citeseer.nj.nec.com/wolfson97geometric.html`. [30.9.2003]