

Harri Sahavirta

Eero Sormunen

Kuvahaun perusongelmat

Polku Panofskysta digitaaliseen kuva-arkistoon

Harri Sahavirta & Eero Sormunen, Kuvahaun perusongelma. Polku Panofskysta digitaaliseen kuva-arkistoon [Fundamental problems in image retrieval. A track from Panofsky to digital image archives] Informaatiotutkimus 20 (3), 60-70, 2001.

Traditional information retrieval is extremely text-oriented. The mainstream of IR research has focused on matching one text (a query) to another text (a document). Image retrieval is a less studied field of IR research, and the special characteristic of images as retrieval objects are not yet well understood. This article is an attempt to illustrate the special nature of image retrieval using the theories of art history as a starting point. Three major questions are highlighted: the problems of expressing image needs, of representing the contents of images as searchable expressions, and of improving the applicability of feature-based indexing methods. Images are created for visual observation, and it is obvious that browsing has a greater role in image retrieval than in text retrieval.

Address: University of Tampere, Department of Information Studies, FIN-33014 University of Tampere, Finland. Email: harri.sahavirta@pp.inet.fi (Harri Sahavirta), eero.sormunen@uta.fi (Eero Sormunen)

1. Johdanto

Kuvahaun teoreettisten lähtökohtien pohdinnassa on usein haettu taustatukea taidehistorian teorioista. Taidehistoriasta on erityisesti etsitty ratkaisua kuville annettavien asia- ja avainsanojen valinnan vaikeuteen. (Markey 1986, Shatford 1986, Karhula 1989, Ørnager 1995) Tarkastelu onkin rajoittunut kuvien *kuvailun* ongelmiin ja siten *kyseilyyn* perustuvan kuvahaun lähtökohtiin. Tässä yhteydessä monet tutkijat ovat viittaneet erityisesti *Erwin Panofskyn* teoriaan kuvataiteiden merkityksistä (Panofsky 1987).

Taidehistorian teorioita voidaan kuitenkin käyttää myös laajemman tarkastelun lähtökohtana. Kuvien esittävyttä ja merkityksiä koskeva taidehistoriallinen keskustelu tarjoaa hedelmällisen maaperän kuvien asiasanoittamisen ongelmien lisäksi myös *selailun* ja

hahmopohjaisten kuvahaun menetelmien tarkastelulle. Tällöin huomio keskitetään itse kuviin ja siihen, miten katsoja antaa kuville merkityksiä. Tämän merkityksenantoprosessin selvittäminen auttaa ymmärtämään sitä, miksi kuvahaussa kohdataan ongelmia riippumatta siitä, perustuuko haku manuaalisesti suoritettuun kuvailuun vai automaattiseen hahmopohjaiseen indeksointiin. Samalla voidaan myös selkeyttää sitä, millaisia seikkoja kuvahaun teoreettisia lähtökohtia pohdittaessa tulisi huomioida.

Kuvahaun teoreettisiin perusteisiin liittyy oleellisesti kysymys kuvien esittävydestä ja merkityksistä. Vastauksia ei kuitenkaan voida hakea yksistään Panofskyn teoriasta, sillä se keskittyy taideteosten ”korkeamman kertaluvun” merkityksiin kuten teemoihin, motiiveihin ja allegorioihin (jotka ovat taidehistoriallisesti kaikkein mielenkiintoisempia). Kuvahaussa ongelmia tuottavat jo ”alemmman kertaluvun” merkitykset ts. se,

miten kuva tunnistetaan tai identifioidaan jonkin reaali maailman objektin *representaatioksi* eli *esitykseksi* (ks. esim. Shatford 1986). Taidehistoriassa tätä ongelmaa on pohtinut mm. E.H. Gombrich (1996).

Jatkossa tulemme käsittelemään taideesteon esittävyyttä ja merkityksiä Panofskyn ja Gombrichin pohjalta. Lisäksi pohdimme sitä, miten esitystavat ja merkitykset varioivat ja millaisia vaikutuksia tällä varioinnilla on kuvahaun. Tätä teoreettista kehystä täydennämme kuvahaun käytäntöjä valaisevien empiiristen tutkimusten tuloksilla, mutta myös joillakin havaintopsykologian näkemyksillä. Keskeinen väittämä on se, että *koska taideoksen (tai yleensä kuvan) representaatio ja merkitykset varioivat, jokainen kuvahakujärjestelmä on epätäydellinen* – riippumatta siitä käytetäänkö siinä *kielessä muotoiltuja kyselyjä* vai *visuaalisia hahmoja*. Kuvien representaatio ja merkitykset eivät ole invariantteja, eivätkä ne ole tyhjentyvästi määriteltävissä kielen tai tiettyjen hahmojen avulla.

Lähdemme liikkeelle kuitenkin kuvatusteista väittämällä, että *sumeat kuvatusteet* ovat kyselyyn perustuvassa tiedonhaussa periaatteessa ratkaisematon ongelma. Tämäkin näkemys perustuu kuvallisten merkitysten varianssiin.

2. Kuvatusteiden määrittelyä

Kyselyihin perustuva tiedonhaku edellyttää, että tiedontarve on selvärajainen ja määriteltävissä oleva. Taustaoletuksena on, että tiedontarve on pohjimmiltaan *määriteltä* ja *artikuloitu*, vaikka tiedontarvitsija itse ei tiedostaisikaan tarvetta näin selkeästi - sillä vasta määriteltä tiedontarve voidaan muotoilla kyselykielessä kyselylausekkeeksi. Esimerkiksi Ingwersen (1992) on kuitenkin korostanut hakutehtävän määriteltävyysongelmaa kyselykeskeisen tiedonhaun rajoituksena.

Aiemmassa tutkimuksessa myös kuvatusteiden käyttäjien tarpeet on useimmiten oletettu melko selvärajaisiksi. Tyypillisesti kuvia oletetaan kysyttävän nimetystä henkilöstä tai kohteesta, tietystä tapahtumasta tai aiheesta (Ørnager 1995, Keister 1994, Armitage & Enser 1997). Tuoreempi tutkimus, jossa on havainnointi toimittajien työssä ilmeneviä kuvatusteita, niiden määrittelyä sekä hakujen suoritusta, osoittaa tällaisen oletuksen ongelmalliseksi. Kuvatusteideat ovat usein vain intuitiivisesti ymmärrettyjä, jolloin toimittaja pystyy sanallisesti kuvaamaan vain joitakin idean reunaehtoja (Markkula & Sormunen 2000). Lisäksi kuvan käyttökelpoisuuteen vaikuttavat olennaisesti ne vaikutelmat, joita kuva synnyttää tai joita kuvan odotetaan synnyttävän katselijassa.

Toimittajat ilmoittivat konkreettisesti kuvitusteiden valintakriteereikseen kuvan ilmentämän dramaattisuuden, yllätyksellisuuden, epätyypillisyyden, vaikuttavuuden, järkyttävyyden, huvittavuuden tai uhkaavuuden (Markkula & Sormunen 2000).

Tiedonhaun lähtökohtana ei ole mielekäästä puhua täysin määrittelemättömistä tiedontarpeista, jolla ei ole mitään kohdetta. *Sumeat* tarpeet ovat sen sijaan täysin mielekkäitä. Toimittajien kuvatusteet ovat usein sumeita ja tällainen tiedontarve on kyseessä myös silloin, kun etsitään esimerkiksi *analogiaa*, *vertailukohtaa* tai *metaforaa*. Tämä on kuvahaulle tyypillistä, sillä kuvia etsitään usein liitettäväksi johonkin tekstiin tai toisiin kuviin, jolloin niiden ajatellaan toisaalta toimivan konkreettisesti kuvitusteena, mutta muodostavan myös *analogian*, *kontrastin* tai *metaforan*. Kuvitusteesta tekstiä riittää usein se, että kuva *viittaa epäsuorasti* tiettyyn aiheeseen. Samoin kuvan tarkoitus on usein muodostaa *uusia merkityksiä* liittyessään johonkin jo olemassa olevaan tekstiin tai toisten kuvien yhteyteen (ks. esim. Markkula & Sormunen 2000). Tällaiset piirteet tekevät kuvatusteesta sumean.

Sumeiden kuvatusteiden ilmaiseminen on vaikeaa, koska hyvin erilaiset kuvan piirteet tai yksityiskohdat voivat tehdä kuvasta ajattelussa käyttöyhteydessään "yllätyksellisen", "epätyypillisen" tai "vaikuttavan". Tällaisten ominaisuuksien ennakoiminen indeksoinnissa on myös mahdotonta, sillä kuva saa erilaisia merkityksiä tai rooleja erilaisten tekstien yhteydessä ja liittyessään toisiin kuviin. Kuvien käytölle ominaista on tässä mielessä *avoimuus*. Kuvien käytön avoimuutta kuvaa myös se, että toimittaja voi ratkaista kuvitusongelman rakentamalla "aasinsillan" kuvan ja tekstin välille sopivalla kuvatekstillä (Markkula & Sormunen 2000).

Kuvatusteita luonnehtiikin hyvin pitkälle *paradoksi*: kuvittajalla on ratkaistavana kuvitustehtävä - siis selkeä tarve löytää kuva - mutta hän ei kykene artikuloimaan, mitä etsii ennen kuin on jo löytänyt etsimänsä. Toimiva kysely voidaan siten muotoilla vasta jälkikäteen. Käytännössä esimerkiksi toimittajat näyttävätkin arvalleval, minkä tyypillisissä konkreettisesti kuvissa (esimerkiksi tietyn henkilön kuvat) haluttuja piirteitä voisi löytyä ja muotoilevat yksinkertaisia kyselyjä tältä pohjalta. Jos ensimmäisen kyselyn tuloksesta ei selauksessa löydy toivottua kuvaa, yritys jatkuu usein uusilla kyselyillä. (Markkula & Sormunen 2000).

Kuvahaut sumeilla perusteilla saattavat olla harvinaisempia kuin selvärajaiset kuvatusteet, mutta samalla tällaiset hakutehtävät ovat huomattavasti vaikeampia. Ainoa mahdollisuus tyydyttää sumea kuvatuste on usein sattumasta kiinni – käytännössä

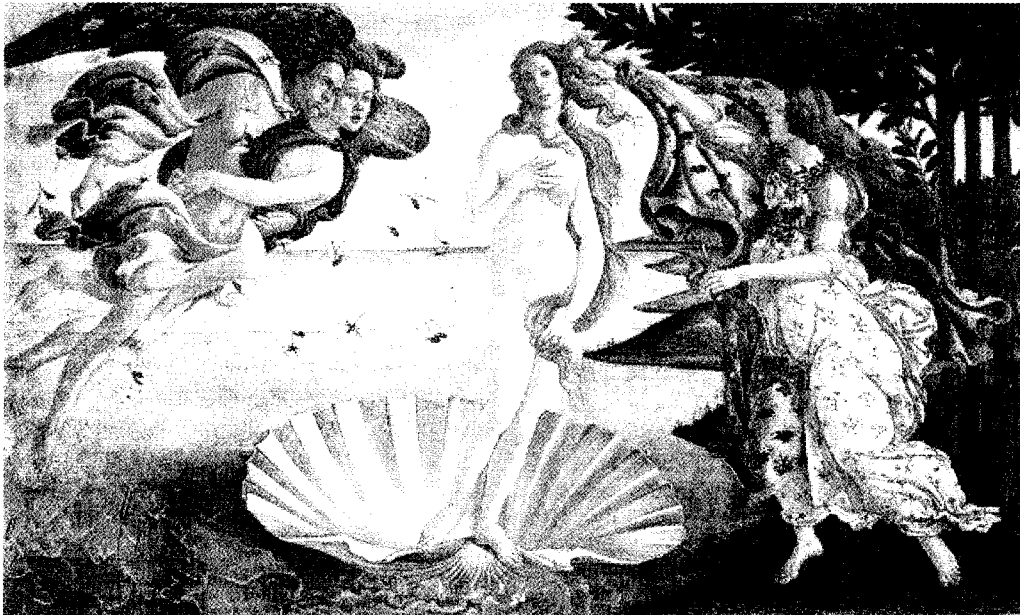
tiedontarvitsija joutuu *selaillemaan* kuvatietokantaa tai väljästi määrittelyyn kyselyn tulosjoukkoa, kunnes sopiva kuva osuu silmiin (Markkula & Sormunen 2000).

Seuraavassa pohdimme, mistä nämä kuvatarpeiden määrittelymisen vaikeudet johtuvat. Haluamme väittää, että pohjimiltaan ne ovat seurausta siitä, miten kuvien merkitykset muotoutuvat ja varioivat.

3. Kuvan merkityksellisistä piirteistä

Perinteinen kyselyorientoitunut tietokantojen kehittäminen nojaa tiedontarpeen määriteltävyyden lisäksi oletukseen *dokumenttien kuvailtavuudesta*. Kuvailtavuuden ongelmana on nähty pääasiassa se, millä menetelmällä kyselyn edellyttämät sanalliset kuvailut liitetään kuvailtaviin dokumentteihin. Tämä huoli näkyy selvästi myös kuvien sisällön kuvailun vaihtoehtoja käsitelmissä julkaisuissa (Markey 1986, Shatford 1986, Rasmussen 1997). Toisaalta esimerkiksi Svenonius (1994) on esittänyt perustellun epäilyn, ettei kuvan esittävyys (representaatio) ja siihen sisältyvät merkitykset ole aukottomasti käännettävissä millekään luonnolliselle tai dokumentaatiokielelle.

Panofskyn teoriakehystä on yleisesti käytetty niiden kuvan piirteiden tyypittelyssä, joita kuvailussa tulisi huomioida. Panofsky tarjoaakin hyvän lähtökohdan pohtia kuvan kuvailtavuuden ongelmaa. Panofskya seuraten voimme sanoa, että taideteoksella on *aihe* (idea) ja *muoto* (form). Luodessaan teoksen taitelija pyrkii löytämään valitsemaansa aiheeseen sopivan muodon: näin *Botticelli* maalatessaan *Venuksen syntymää* (1485) etsi sopivaa muotoa aiheeseen "Venus nousee meren aalloista" (ks. kuva). Muoto koostui sellaisista asioista kuin mallin valinta, asennon valinta, kokonaissommitelma, mutta myös maalausalan valinnasta, pohjamaaleista yms. Nämä kaikki vaikuttavat kuvasta saatavaan kokonaisvaikutelmaan. Aiheen ja muodon yhdistäminen paljastaa (onnistuessaan) teoksen *sisällön* (content). Sisällön Panofsky määrittelee sanomalla, että se on *teokseen kondensoitunut, kansakunnan, aikakauden, luokan, uskonnollisen tai filosofisen opin perusasenne, jonka yksittäinen henkilö on tiedostamattaan kvalifioinut* (Panofsky 1987, 37). Yksinkertaistaen voisimme sanoa, että teoksen sisältöön kuuluvat ne kulttuuriset aatteet, ihanteet ja konventiot, joita teos ilmentää.



Kuva. Sandro Botticelli (Italia, 1445-1510): Venuksen synty, maalattu noin v. 1485

Aihe, muoto ja sisältö ovat siis teoksesta löydettävissä olevia merkityksiä. Panofskyn mukaan näihin merkityksiin vaikuttavat kolmenlaiset tekijät. *Taiteilijan intentiot* määrittelevät teoksen merkityksiä ja teoksen tulkinta pyrkii paljastamaan nämä intentiot koko historiallisessa viitekehityksessään. Taiteilijan intentioita korostavia näkemyksiä on tosin kritisoitu, sillä intentiot ovat psykologisia tiloja, joita taitelija ei välttämättä kykene välittämään (ks. mm. Beardsley 1981, s. 17-29). Panofsky myöntääkin, etteivät intentiot välttämättä ole kovin tarkasti määriteltyjä – ja tulkintaan vaikuttaa myös *tulkitsijan asenne*, joka riippuu hänen kokemuksistaan ja historiallisesta tilanteesta. Taideteoksen merkitys on aina jossain määrin *subjektiivinen*, sillä siihen liittyy *uudelleen luomista* (re-enaction) ja *koodausta* (coding and decoding). Kolmanneksi, teoksen merkitykseen vaikuttavat aikakauden, kulttuurin, tyylisuunnan tai koulukunnan sisällä määritellyt *konventiot* – mikä tuo tulkintaan objektiivisen elementin. (Panofsky 1987, s. 36-37.)

Panofskyn teoria keskittyy taideteosten ”korkeamman kertaluvun” merkityksiin, eikä käsittele yksinkertaisempia esittävyiden ongelmia. Nämä kaksi ongelmakenttää ovat kuitenkin tietyllä tavalla analogisia. *Gombrichin* (1996) mukaan nimittäin ei ole olemassa ”viatonta silmää” ts. havainto ja tunnistaminen muotoutuvat aina suhteessa katsojan laajempaan tietämykseen sekä kulttuurin omaksumiin esitystapoihin. Näin teoksen esittävyys määräytyy suhteessa havainnon psykologisiin ehtoihin, joita määrittelee laajempi kulttuuri-konteksti. Gombrichin ajatuksen voisi kiteyttää sanomalla, että *esittävyys* – kuvien ja niiden kohteiden *samankaltaisuus* – on *opittava näkemään*.

Tältä teoreettiselta pohjalta voidaan todeta, että kuvat ovat usein kuvatarpeiden tapaan vaikeasti kuvailtavia ja merkitykseltään *sumeita*. Jatkossa analysoimme tätä kuvien sumeutta ja väitämme, että kuvallinen ilmaisu varioi kolmessa merkityksessä – *merkityksen taso*, mutta myös *representaation taso ja muoto* muuntelevat. Tämä merkityksen variointi on kuvahaun kannalta keskeistä. Luvussa 4 tarkastelemme *merkityksen ja representaation tasoja* ja päädyimme johtopäätöksiin *sanallisen sisällönkuvailun rajoituksista*. Luvussa 5 tarkastelemme *representaation muodon* varioivuutta ja päädyimme johtopäätöksiin *hahmopohjaisen* kuvailun ja haun rajoituksista.

4. Merkityksen ja representaation tasojen vaikutus sisällönkuvailuun

Representaatiosta eli esittävydestä puhuessamme olemme tarkoittaneet pitkälti samaa, mitä Panofsky merkitysten teoriassaan tarkoittaa *ensisijaisella* tai *luonnollisella merkityksellä* (primary or natural subject matter), joka muodostaa merkityksen *esi-ikonografisen tason* (Panofsky 1987, 53). Tällä tasolla kuvasta yksinkertaisesti identifioidaan joitakin luonnollisten objektien representaatioita. Panofsky ei problematisoi tätä aspektia; hän olettaa identifioimisen olevan jotenkin ”luontainen” prosessi, jota voidaan verrata siihen että tunnistamme reaali maailman objekteja. Analyysin jättäminen tälle tasolle saattaa riittää taidehistorian näkökulmasta. Sisällönkuvailun kannalta kuvan ”luonnollisten” kohteiden yhtenäinen nimeäminen ei kuitenkaan ole ongelmatonta, sillä näitä saattaa olla varsin suuri määrä. Kun otetaan tarkasteluun lisäksi teoksen korkeamman tason *merkitysten variointi*, sisällönkuvailijalle alkaakin kasautua yhä vaikeampia ongelmia.

Panofsky jaottelee korkeamman tason merkitykset *ikonografisiin* ja *ikonologisiin* merkityksiin. *Ikonografisella* tasolla merkitys muodostuu konventionaalaisesta *aiheesta* (conventional subject matter), jonka tunnistaminen edellyttää tapojen ja kulttuuristen traditioiden tuntemista. Tällöin kuvasta tunnistetaan *motiivit, käsitteet ja teemat* sekä identifioidaan *mielikuvat* (images), *tarinat* ja *allegoriat* (Panofsky 1987, 54.). Ikonografisella tasolla *Botticellin Venus* tunnistetaan (kreikkalais-) roomalaisen mytologian rakkauden jumalattareksi, joka tarinan mukaan syntyi meren vaahdosta. Ikonografisella tasolla tuleekin lisää kuvailtavia ominaisuuksia, kuten ”kreikkalais-roomalainen mytologia”, ”jumalatar” ja ”rakkaus” ja niiden tulkinta edellyttää laajaa taustatietoa.

Ikonologisella tasolla merkitykset muodostuvat puolestaan *sisällöstä* (content). Panofsky määrittelee tämän sisällön hyvin laajasti kulttuurisiksi perusasteiksi, jotka taitelija on kondensoinut teokseen ja jotka selittävät sekä sommitelmallista metodia että ikonografista merkitystä (Panofsky 1987, 55-56). Ikonologisella tasolla Botticellin Venusta tarkastellaan dokumenttina *Botticellin persoonallisuudesta, italialaisesta renessanssista* ja ajalle tyypillisestä *antiikin aiheiden käytöstä* ja *antiikin ihallusta*, mutta myös tietyn *naisihanteen* ja *eroottisuuden* ilmentymänä.

Panofskya seuraten voidaan sanoa, että taide-teoksella on kolmentasoisia merkityksiä, jotka edellyttävät erilaisia tulkinnan muotoja. Esi-ikono-

grafisella tasolla teoksen luonnollisia merkityksiä *kuvaillaan*; ikonografisella tasolla teoksen konventionaalisia merkityksiä *analysoidaan* ja ikonologisella tasolla teoksen sisältö *tulkitaan*. Nämä tulkinnan muodot ”lepäävät toistensa päällä”. Analyysi perustuu kuvailun oikeellisuudelle ja tulkinta edellyttää korrektein kuvailun ja analyysin (Panofsky 1987, 58.).

Panofskyn näkemykseen täytyy kuitenkin lisätä se, että sama pätee myös käänteisesti. Teoksen esi-ikonografinen kuvailu saattaa edellyttää jo olemassa olevaa ikonografista analyysia ja tämä analyysi edellyttää varsin usein ikonologista tulkintaa. Näin maalaukselle naisesta, jolla on *miekka* kädessä ja verinen miehen pää *tarjottimella*, ei voida määrittää (ikonografisella tasolla) käsitteitä, motiivia tai tarinaa ilman (ikonologisen tason) tulkintaa, sillä ei ole selvää mihin tarinaan maalaus viittaa: Judit leikkasi assyrialisten sotapäälliköltä Holofernekseltä pään *miekalla*, mutta Salome halusi Johannes Kastajan pään *tarjottimella*. Aiheen määrittäminen edellyttää, että tiedämme joitakin tyypillisesti ikonologiselle tasolle kuuluvia periaatteita: Salomea ei ole ollut tapana kuvata oikeutta symboloivan miekan kanssa, mutta Judit on usein kuvattu kantamassa Holoferneksen päätä *tarjottimella* (vaikkei se kuulukaan kertomukseen).

Panofsky painottaa tämän ”käänteisen järjestyksen” sijasta sitä, että kuvailussa, analyysissa ja tulkinnassa voidaan (ja täytyy) nojata yleisempään tietoon kulttuurista. Tällöin teosten merkitysten etsintää säätelevänä periaatteena ei ole merkityksen ylempi taso, vaan yleinen (ko. tasoa koskeva) historiallinen tietämys. Tämä tuo objektiivisen elementin tulkintaprosessiin. Näin voimme ajatella, että esi-ikonografisella tasolla Botticellin Venuksen luonnollisen aiheen kuvaus (alastomaksi naiseksi) perustuu arkikokemukseemme, jota kuitenkin säätelee tylin historia eli tietomme siitä, miten eri objekteja ja tapahtumia on kuvattu eri aikoina. Ikonografisella tasolla analysoimme maalauksen konventionaalisen merkityksen rakkauden jumalatar Venukseksi kirjallisen ja käsitteellisen tietomme avulla, mutta tätä analyysia säätelee myös tyyppien historia eli tietomme miten eri teemoja ja käsitteitä on kuvattu eri aikoina. Ja lopulta, ikonologisella tasolla teoksen merkitys tulkitaan käyttäen intuitiota, jota säätelee kulttuuristen symbolien historia eli tietomme siitä, miten ihmismielelle ominaisia tendenssejä on kuvattu eri aikoina (Panofsky 1987, 66).

Kuvahaun kannalta keskeistä on se, että taide-teoksen *merkitys* varioi eri (abstraktio)tasoilla, eikä merkitysten määrä ole invariantti; uusia merkityksiä voidaan jatkuvasti löytää. Merkityksiä luonnehtii lisäksi se, etteivät ne usein ole suoraan havaittavissa teoksesta ts. ne eivät perustu mihinkään invarianttiin hahmoon,

vaan siihen, että teoksen eri elementeille annetaan jokin merkitys laajemman historiallisen tiedon kontekstissa. Merkityksiä ei ehkä ole ääretöntä määrää, mutta mielekkäiden merkitysten tyhjentävä kuvailu edellyttäisi niin syvää ja monialaista asiantuntemusta, ettei tällainen sisällönkuvailu ole käytännössä mahdollista.

Tämän lisäksi on huomattava, että myös teoksesta suoraan havaittavat merkitykset varioivat: *representaation taso* muuntelee. Representaation tason varioiminen tarkoittaa yksinkertaisesti sitä, että kuva esittää monia eri asioita riippuen siitä, mihin kiinnitämme huomiomme. Ajatelkaamme jälleen *Botticellin* maalausta *Venuksen syntymä*, jossa alaston Venus nousee merestä simpukankuoren päällä seisten. Maalaus esittää naisvartaloa kokonaisuudessaan, mutta samalla myös eri ihmisruumiin *osia*: päätä, olkapäitä, rintoja, vatsaa ja vaikkapa napaa. Maalaus on kokonaisuus, joka esittää sekä kokonaisuutta että kaikkia osiaan. Maalauksen osista muodostuu varsin suuri määrä ominaisuuksia, joita kuva esittää ja jotka ovat kaikki potentiaalisia hakuominaisuuksia.

Botticellin maalaus esittää kuitenkin enemmän kuin vain osiaan; käsitetasolla se esittää ”naista”, mutta erityisesti ”pitkähiuksista naista” tai ”häpynsä hiuksillaan peittävää naista” – mutta myös ”alastonta”, ”ihmistä” ja ”homo sapiens -lajin edustajaa” tai haluttaessa vaikkapa ”kaksijalkaista höyhenetöntä oliota”. Maalaus esittää tietysti myös ”jumalataria” eikä siis tässä mielessä lainkaan ihmistä. Maalauksen esittävyys voidaan siis tulkita eri käsitetasoilla ja nämä käsitteet ovat potentiaalisia hakuominaisuuksia.

Maalauksen osien ja kokonaisuuden sekä esitettyjen käsitteiden hierarkkisen tason valinnaisuus johtaa päätelmään *kuvailun subjektiivisuudesta*. Teoksen merkitys ja esittävyys ovat aina tulkintoja ja riippuvat tulkitsijan mielikuvituksesta ja mielenkiinnon kohteista – sekä siitä yhteydestä, jossa teosta tarkastellaan. Toisaalta voidaan päätyä johtopäätökseen, että kuvatietokannassa kuvan *sisällönkuvailu on aina epätyhjentävää*. Kuvasta voidaan aina löytää uusi osa tai käsitteellinen ominaisuus, joka kuvailusta puuttuu ja on näin myös kyselyn muotoilijan ulottumattomissa.

Sisällönkuvailun menetelmien kehittäjät ovat olleet varsin hyvin tietoisia edellä esitettyjen epätyhjentävyyden ja epäyhdenmukaisuusongelmien luonteesta (Markey 1986, Shafford 1986, Svenonius 1995). Melko yhtenäinen johtopäätös on ollut se, että sisällönkuvailussa on mielekästä keskittyä sellaisiin piirteisiin, jotka ovat mahdollisimman konkreettisia ja yksiselitteisesti tulkittavia. Tämä on johtaa todennäköisesti kohtuullisen yhtenäiseen ja käyttäjän näkökulmasta ennakoitavaan kuvailukäytäntöön. Kyselyn muotoilija

pystyy hyödyntämään kuvailua silloin, kun hakutarve on määriteltävissä kuvan konkreettisen sisällön perustella. Toisaalta em. suunnannäyttäjät ovat sivuuttaneet kysymyksen miten kuvailulla voitaisiin tukea käyttäjää sumeissa ja abstraktille tulkinnan tasolle nousevissa kuvahakutarpeissa. Ideointia on myös rajoittanut perinteinen viitetietokanta-ajattelu; kuva-arkistojen digitalisoitumisen mahdollistaman kuvien silmäilyn merkitystä ei ole otettu huomioon.

Digitaalisessa ympäristössä on mahdollista ja näyttäisi tärkeältä kehittää kuvatietokantojen sisällönkuvailua siten, että se tukisi tasavertaisesti niin kyselyjä kuin selailuakin (Markkula & Sormunen 2000). Kyselyä tukevassa kuvailussa on ensisijaista luetella kuvan kaikki piirteet, joihin kyselyssä on tarve viitata. Digitaalisessa ympäristössä olennaisia kehittämis-kohteita ovat sellaiset kuvailumenetelmät, jotka mahdollistavat vaihtoehtoisten *selailunäkymien* luomisen kuva-aineistoon. Ne merkitykset ja representaation tasot, jotka ovat kuvailussa ja kyselyissä ongelmallisia, ovat silmäiltäessä usein helposti tulkittavissa. Selailijan näkökulma on ainakin intuitiivisella tasolla kuvatarpeen rajaama ja näin olennaiset kuvan merkitykset ovat käyttäjän tulkittavissa suoraan kuvasta sopivalla representaation tasolla havainnoiden. Selailunäkymiä luova kuvien sisällönkuvailu on toistaiseksi kuitenkin vain löyhästi määritelty idea, joka vaatisi lisäselvityksiä muodostuakseen käyttäntöjä muokkaavaksi menetelmäksi.

5. Representaation muodon vaihtelun vaikutukset hahmopohjaisiin hakumenetelmiin

Kuvien subjektiivisen tulkinnan aiheuttamaan epäyhdenmukaisuuteen, kuvailun epätyhjentävyyteen sekä kuvallisten ja kielellisten ilmausten käännettävyyden ongelmaan on tarjottu lääkkeeksi *hahmopohjaisia* (content-based) *indeksointi-* ja *hakumenetelmiä* (ks. Gong 1998, del Bimbo 1999). Hahmopohjaisessa indeksoinnissa kuvasta tuotetaan digitaalisen kuvankäsittelyn menetelmin automaattisesti kuvan visuaalista sisältöä edustavia *piirrevektoreita*. Nykyisten hahmoindeksointimenetelmien heikkoutena on kuitenkin se, että ne perustuvat hyvin alhaisen abstraktiotason piirteisiin kuten kuvan väriin, muotoihin ja tekstuurin (Eakins 1996, Rasmussen 1997). Menetelmän etu on siinä, että sekä kuvat että kysely ovat kuvallisia representaatioita eikä käännösongelmia kuvan ja kielen välillä muodostu (Enser 1995).

Hahmopohjaisessa kuvahaussa kysely voidaan periaatteessa suorittaa niidenkin kuvallisten ominaisuuksien varassa joita ei voida kääntää sanalliseksi kuvailuksi. Toisaalta käsitteellisiä, abstrakteja ominaisuuksia joudutaan hakemaan indeksoitujen hakusanojen avulla. Hahmopohjaiset tunnistusmenetelmät eivät myöskään kykene yksilöiden tunnistamiseen kuin hyvin rajoitetuissa aineistoissa. Hahmopohjaisten hakualgoritmien tehokkuudesta on käytettävissä hyvin vähän uskottavia evaluointituloksia (Sormunen ym. 1999).

Hahmopohjaisten kuvahakumenetelmien ongelmia voidaan jälleen tarkastella Panofskyn pohjalta. *Ikonografisella* ja *ikonologisella* tasolla analyysi ja tulkinta eivät rajoitu vain kuvallisesta teoksesta suoraan havaittavissa oleviin piirteisiin. Maalaus viittaa itsensä ulkopuolelle ja kuvastaa jotain itsensä ulkopuolista. Jotta kaikki nämä merkitykset tavoitettaisiin, tulkinnan tulee, Panofskyn mukaan, edetä ”alhaalta ylös” kolmessa vaiheessa ja tukeutua jokaisessa yleisempään (kulttuuriseen) tietoon. Lisäksi ainakin joissakin tapauksissa myös ylempät tasot vaikuttavat alempien tasojen merkityksiin. Tulkinta on siten varsin kompleksinen tapahtuma. Tämän vuoksi merkityksiä ei voida tavoittaa pelkän hahmontunnistuksen avulla ilman hyvin kehittynyttä mallia tulkinnan kohteena olevasta ilmiöstä. Esimerkkejä onnistuneesta arkitietämyksen (common sense) mallintamisesta – saati sitten yleisen kulttuurisen tiedon mallintamisesta – ei kuitenkaan ole tekoälytutkimuksenkaan puolella.

Jos tarkastelu rajoitetaan pelkästään kuvassa konkreettisesti *näkyviin* ominaisuuksiin, *representaation muotojen* variointi muodostaa ongelman, joka hahmopohjaisten menetelmien tulisi ratkaista. Muodon varioiminen tarkoittaa yksinkertaisesti sitä, että tietty kohde voidaan esittää monella eri tavalla. Hakuprosessi aloitetaan kuitenkin määrittelemällä niiden kuvan visuaalisten ominaisuuksien joukko, joka haettavilla kuvilla tulee olla. Käsitteellistä tai visuaalista tiedontarvetta edustaa *yksi* kuvahahmo, vaikka mahdollisia representaatioita (hahmoja) on *useita* – ja näin hakutuloksesta putoaa joitakin relevantteja kuvia pois. Triviaali esimerkiksi olisi se, että jos haluamme löytää kuvatietokannasta kaikki Venusta esittävät kuvat ja valitsemme hakuahmoksi Botticellin Venuksen, hakutuloksesta putoavat mm. kaikki makaavat Venukset.

Ongelma ei kuitenkaan ole triviaali. Kun hakuahmon ominaisuuksia verrataan aineistoon, jossa representaation muodot varioivat, tuloksesta putoaa aina joitakin kuvia, jotka ihmissilmä arvioisi relevanteiksi. Näin Botticellin Venus hakuahmona ei kykenisi erottamaan edes kaikkia klassisessa asennossa seisovia Venuksia

– sillä tämän yhden määritellyn hahmon ominaisuudet eivät kata kaikkia mahdollisia klassisessa asennossa seisovan Venuksen representaation muotoja.

Representaation muotojen variointia on hedelmällisintä tarkastella ottaen lähtökohdaksi representaation muotojen kehittyminen taiteessa (ks. Panofsky 1987, luku 2). *Egyptiläiselle taiteelle* oli luonteenomaista varioimattomuus; maalauksissa ei pyritty syvyyden vaikutelmaan tai luonnon muunnelmien jäljentämiseen, vaan ennemminkin konstruoimaan hahmo tietyistä moduuleista ja näin luomaan eräänlainen *skeema*. Egyptiläisissä maalauksissa esitettävien ihmisten asennot eivät muuntele paljoakaan: seisova ihmishahmo seisoo aina samassa asennossa – ja jalat esitetään sivulta, vartalo edestä ja kasvot profiilissa. Hahmojen variointi rajoittuu seisoviin ja istuviin hahmoihin, sekä joihinkin eleisiin kuten käden nostoon jne. Tämä on kuitenkin vain pääsääntö, johon on poikkeuksia. Egyptiläisessä taiteessa myös ruumiin suhteet ovat vakiot. Käytännössä kuvapinta-ala jaettiin 18 (tai 22) samankokoiseen ruutuun, jolloin yksi ruutu muodosti yhden yksikön ja sen koon avulla voitiin määritellä koko kuvan sekä kaikkien sen osien koko (Panofsky 1987, 87). Hahmojen muoto ja mittasuhteet olivat niin vakioituneet, että kaikki esitykset seisovasta miehestä muistuttavat toisiaan – ne ovat tavallaan kopioita toisistaan ihonväriä myöten (joka määräytyi sukupuolen mukaan).

Egyptiläisen taiteen kukoistuskauden jälkeen representaation muodot ovat kuitenkin alkaneet varioida huomattavasti enemmän. *Kreikkalaiset* noudattivat yhä ihmisvartalon ideaalisia mittasuhteita ja tiettyjä konventionaalisia asentoja, mutta pyrkivät suurempaan totuudenmukaisuuteen: taide oli kreikkalaisille mimeettistä; *todellisuutta jäljittelevää*. Näin kreikkalaiset ottivat käyttöön *elastiset asennot*, mutta pyrkivät myös luomaan vaikutelman *kolmiulotteisuudesta* ja käyttivät mm. *lyhennyksiä* ja todellisten mittasuhteiden *vääristelyä* (Panofsky 1987, 90). Näin taaksepäin ojentuva käsi on lyhennetty, sillä se näyttää lyhyemmältä kuin vartaloa vasten lepäävä käsi. Korkealle sijoitettavan patsaan pää on suhteettoman suuri, jotta se näyttäisi alhaalta katsoen sopusointuiselta vartalon kanssa.

Egyptiläiset maalaukset on konstruoitu osista, jotka ovat invariantteja ”hahmoja”, mutta kreikkalaisessa taiteessa ”hahmot” varioivat riippuen todellisuudessa tapahtuvasta muuntelusta, asennosta ja aiotusta katselukulmasta. Näin voidaan ajatella, että egyptiläisessä kuva-aineistossa hahmopohjaisen kuvahakuun riittäisi melko yksinkertainen ohjelma. Haluttaessa hakea kaikki ihmishahmoa etsittävät kuvat jouduttaisiin määrittelemään ihmisvartalon osien hahmot: ihmisellä tulee olla kaksi (sivusta kuvattua) jalkaa, (edestä

kuvattu) vartalo, kaksi kättä sekä pää (jonka muoto saattaa varioida ihmispäästä koiran tai linnun päähän). Tämän lisäksi tulisi sallia tiettyjen nivelien liike: hahmo voi taipua mm. lonkka- ja polvinivelien kohdalta (istuva hahmo) tai olka- ja kyynärnivelistään (kättä nostava hahmo). Tällainen ihmisen hahmo voitaisiin tietysti määritellä kuvatesaurukseen ja sitä voitaisiin sen jälkeen käyttää aina tarvittaessa. Kreikkalaisessa kuva-aineistossa kuvatesaurukseen sisältyvien hahmojen tulisi olla jo huomattavasti kompleksisempia: ihmishahmoa esittäviä kuvia haettaessa tarvittaisiin jonkinlaisia kolmiulotteinen virtuaalimalli ihmisestä: hahmo, jonka mittasuhteet ja muodot saattavat varioida, mutta jota voidaan ennen kaikkea käännettä ja liikuttaa kolmiulotteisessa tilassa perspektiivin lait huomioiden.

Kolmiulotteisuuden ja perspektiivin lisääminen hakuhaamoon ei kuitenkaan vielä tyhjennä representaatioiden variaatioita. Taiteessa tosin kuvattiin pitkään ideaalia ihmisruumista, mutta samalla totuudenmukaisuuteen ja luonteenomaisuuteen pyrkiminen ovat olleet vahvoja tendenssejä. Todellisuudessa ihmisvartaloiden mittasuhteet vaihtelevat hyvinkin paljon ja taiteilija saattaa vääristää niitä vielä lisää saavuttaakseen jollakin tavalla ”luonteenomaisen hahmon”: manierismin aikana (1500-luvulla) jäseniä venytettiin maalauksissa luonnottomiin mittasuhteisiin ja nykytaiteilijoista vaikkapa *Salvador Dali* on oiva esimerkki mittasuhteiden vääristelystä ilman että kuvan esittävyys heikkenisi. Hahmopohjaisen kuvahakuohjelman tulisi siten pyöritellä hahmoja tilassa, mutta myös sallia hahmojen *venyttäminen, paisuttaminen ja sekoittaminen*.

Kysymykseen, kuinka paljon vääristelyä hahmo kestää ilman että sen identiteetti muuttuu, ei ole olemassa mitään yksiselitteistä vastausta. *Dalin norsut* muistuttavat kirahveja pitkine ohuine jalkoineen, mutta ne tunnustetaan oitis norsuiksi ja *Botticellin Venuksen* kaltaista hyvin tunnettua kuvaa voidaan vääristellä melkein loputtomasti ilman että sen tunnustettavuus kärsisi. Venus voidaan lihottaa, laihduttaa, venyttää tai pukea farkkuihin. Venukselle voidaan piirtää viikset (kuten *Mona Lisalle*) tai lines *Ankka* voidaan asettaa samaan asentoon – mutta Venus voidaan myös korvata alastomalla miesfiguurilla (kuten *Gallen-Kallelan Aino* mainoksessa) tai konstruoida vihanneksista. *Botticellin Venuksen* tunteva henkilö tunnistaa piilevän yhteyden, koska kaikissa tapauksissa hahmossa säilyy jotakin olennaista. Me emme vain tiedä, mitä se on – tai edes sitä, mikä erottaa lopulta Venuksen hahmon amforan hahmosta. Hahmoja voidaankin monissa tapauksissa modifioida loputtomiin ilman että niiden identiteetti katoasi – ja tämä sama pätee myös hahmon osittamiseen. *Botticellin Venus* tunnustetaan pelkästään

päästään tai vartalostaan – tai ääriviivoistaan tai vasemmasta puoliskostaan jne. Modifioinnin lisäksi erilaiset rajaukset kuuluvat representaation muodon varioiviin piirteisiin.

Representaatio siis varioi sekä tasonsa että muotonsa puolesta, eikä tälle varioinnille voida asettaa mitään täsmällisiä rajoja. Yhtä representaation muotoa vastaa useampi käsite ja yhtä käsitettä vastaa useampi mahdollinen representaation muoto. Tätä ongelmaa voidaan tarkastella myös havaintopsykologiselta pohjalta. Gombrich tuo tämän aspektin esiin taidehistoriassa ja huomauttaa, että esittävyiden tunnistaminen perustuu *suhteiden* havaitsemiseen, eikä suinkaan yksittäisten elementtien tunnistamiseen. Kuva muotoutuu toisaalta hahmon suhteesta taustaan ja toisaalta kirkkauden vaihtelusta hahmon sisällä. Näin hahmon absoluuttinen muoto, väri tai kirkkaus eivät ole ainoa olennainen asia – kontrastit ja suhteet ovat ainakin yhtä merkittäviä. Tätä havainnollistaa se, että hahmo voi olla vaalea kuvio tummalla pohjalla tai tumma kuvio vaalealla pohjalla – ja toisaalta kuvan vaalein kohta mielletään valoisaksi riippumatta sen todellisesta kirkkaudesta asteesta. (Gombrich 1996, 37, 46-49.) Kuvan sisäisistä kontrasteista rakentuu hahmon muoto ja syvyysvaikutelma. Havaintopsykologia tukee tätä näkemystä, sillä nyky näkemysten mukaan visuaalinen järjestelmä muoaa aktiivisesti monin tavoin optista ärsykettä ja kontrastit ovat oleellinen osa tätä prosessia. Silmä korostaa kuvassa esiintyviä kontrasteja niin että mm. värin (havaittu) kirkkaus riippuu taustan väristä ja ns. Machin kuviossa kahden toisiinsa rajoittuvan alueen välissä nähdään tummempi rajaviiva (jota ei todellisuudessa ole). Visuaalinen järjestelmä tarkentaa (ja jopa liioittelee) verkkokalvolle muodostuvan hahmon ääriviivoja. (Gleitman 1991, 180-183.)

Havainto on siten konstruoitu jo fysiologisella tasolla. Gombrich huomauttaa kuitenkin lisäksi, että nämä suhteiden havainnointi tapahtuu aina jossakin viitekehyksessä; "odotusten horisontissa", joka saa havainnoitsijan "liioitellun herkäksi" eroille ja vaihteluille. (Gombrich 1996, 53.) Katsomme kuvia koko kulttuurisen tietomme vaikuttaessa taustalla ja näin mm. kuvan todenmukaisuus tai illusorisuus riippuvat viime kädessä siitä millaisia kuvia olemme totuneet näkemään. Tätä ilmiötä on ehkä hieman vaikea mieltää, ellei ajattele vaikkapa sitä, että varhaisrenessanssin suuren mestarin *Giotton* maalaukset vaikuttivat aikalaisista hätkädyttävän totuudenmukaisilta, mutta meistä hieman kankeilta ja "skemaattisilta"; meistä on omiutusta ajatella että rakennus, joka on esitetty samanaikaisesti ulko- ja sisäpuolelta ja jota ei ole piirretty perspektiiviin, "pettäisi silmää" niin ettei sitä erota mallistaan. Vastaavasti (meidän mielestämme) "luonnollisten" värien tullessa

maisemamaalaukseen 1800-luvun alkupuolella aikalaiset kokivat uudet maalaukset räikeinä ja lähes rivoina – reaktio, jota on vaikea ymmärtää katsoessa taitetuilla ruskean ja vihreän sävyillä maalattuja maisemia.

Tämä viittaa siihen, että kuvien hahmottaminen ja esittävyiden tunnistaminen perustuu aina johonkin viitekehykseen: skemaattisiin muotoihin, kategoriioihin ja luokituksiin. Eikä tämä päde ainoastaan kuvien katsomiseen, vaan myös tekemiseen. Gombrich esittääkin monia esimerkkejä siitä, miten taiteilijan ennako-oletukset ovat vaikuttaneet vääristäen teoksiin, joissa on yritetty kopioida jotakin kohdetta luonnonomukaisesti. (Gombrich 1996, 59-63.) Kuviota onkin hyvin vaikea hahmottaa, piirtää tai edes kopioida, ellei kykene mieltämään sitä joksikin, luokittelemaan jotenkin. Gombrichin johtopäätös on se, että kaikki taide on käsitteellistä: taiteen alkuperä on ihmismieleessä ja sen reaktioissa maailmaa kohtaan, eikä suinkaan näkyvässä maailmassa itsessään. Esittämistä ja hahmojen tunnistamista sitovat kyllä jonkinlaiset säännöt, kategoriat ja konventiot; tyyli – mutta tämä tyyli vaihtelee alati, sillä luokituksia voidaan soveltaa ja muunnella aina tarpeen mukaan. (Gombrich 1996, 76.)

Oppiessamme katsomaan kuvia opimme siten yksilöimään, artikuloimaan ja tekemään eroja siellä, missä näyttäisi olevan vain muodotonta massaa. (Gombrich 1996, 86.) Kuvan esittävyys ei perustu jonkin tietyn yleisen hahmon (passiiviseen) erottamiseen, vaan ennemminkin projektioon; aktiiviseen näkemiseen jonakin.¹ Tosin on myönnettävä, että usein näkeminen jonakin kuitenkin edellyttää jonkinasteista *samankaltaisuutta* – mitä tahansa ei voi nähdä minä tahansa. Kuitenkin Picasso onnistui eräessä veistoksessaan luomaan leluautosta simpanssin kasvot. (Gombrich 1996, p. 89.) Lopulta projektiolla on siis rajansa vain ihmisen mielikuvituksessa. Gombrich painottaa taidehistorian keskeisenä elementtinä tätä ihmisen kykyä projisoida – katsojan taipumus nähdä kuvassa "enemmän" kuin siinä objektiivisesti on nousi taiteessa keskeiselle sijalle jo renessanssin aikana. Taitelijat ovatkin tietoisesti luottaneet siihen, että katsoja täydentää kuvan: maalaus saattaa muodostua tuhruisista siveltimenvedoista ja väriäiskistä, jotka kauempaa katsottuna kuitenkin muuttuvat esittäviksi muodoiksi. Monet taitelijat antavat tietoisesti katsojalle "enemmän tehtävää"; katsojan tulee mielessään luoda teos taiteilijan antamien vihjeiden varassa. (Gombrich 1996, 165-169.)

Havaintopsykologia tukee tämän tapaista teoriaa. Nykyisin pidetään tosiasiana, että esimerkiksi visuaalisen kentän syvyysvaikutelma syntyy erilaisten

syvyyshivhjeiden perustella: näköhavainnon syvyys vaikutelma konstruoidaan kahdesta erilaisesta verkkokalvolle syntyvästä kuvasta – ja lisäksi etäisyyttä arvioidaan objektien suhteellisen koon, aseman ja perspektiivin avulla. Kolmiulotteisuus projisoidaan kaksiulotteiseen kuvaan verkkokalvolla. (Gleitman 1991, 200-201.) Vastaavasti hahmottaminen ts. visuaalinen jäsentäminen ei perustu vain siihen, mitä ärsyke sisältää. Hahmo eristetään taustastaan (ja nähdään artikuloitumpana kuin tausta, joka jää epämääräisemmäksi), mutta sen lisäksi havaintoaines järjestetään ja ryhmitellään mm. läheisyyden, samankaltaisuuden ja jatkuvuuden perustella sekä mahdollisesti täydennetään. (Gleitman 1991, 211-214.)

Hahmon tunnistaminen ei näytä etenevän vain osista kokonaisuudeksi: kolmio ei muodostu vain kolmesta kulmasta eivätkä kasvot vain suusta, nenästä ja silmistä. Osien lisäksi oleellista on niiden suhteet – kasvot tunnistetaan vasta jos nenä on silmien välissä ja suun niiden alapuolella. Tällaisia suhteita ei ole onnistuttu analysoimaan yleisellä tasolla. Hahmon tunnistaminen eteneekin toisaalta osista kokonaisuuteen, mutta samalla myös kokonaisuudesta osiin siten, että yleinen tietämys, odotukset ja konteksti vaikuttavat osien tulkintaan. (Gleitman 1991, 218-220.) Näiden piirteiden lisäksi siihen mitä havaitaan vaikuttaa tietysti myös havaintosijan tarkkaavaisuus. Havaintopsykologian näkökulmasta ei olekaan selvää, missä näkeminen päättyy ja tietämys alkaa. (Gleitman 1991, 237.)

Näyttää tosiasialta, että myös taideteoksessa nähtävät hahmot ja esittävyys eivät edes yksinkertaisimmissa muodoissaan perustu vain teoksessa konkreettisesti havaittaviin piirteisiin. Katsoja jäsentää, muokkaa ja täydentää näkemäänsä niin paljon, että esittävyyden voidaan väittää olevan aina katsojan projektiota. Monet taiteilijat ovat tietoisesti käyttäneet ihmisen taipumusta projektiioon hyväkseen. Kuvan esittävyys; ”alemmman tason” merkitykset ovat siten yhtä kompleksisia ja varioivat yhtä paljon kuin kuvan ”korkeamman tason” merkitykset. Lisäksi näille molemmille näyttää olevan luonteenomaista se, että tulkinta tai tunnistaminen etenee toisaalta alhaalta ylös (osista kokonaisuuteen), mutta myös ylhäältä alas (kokonaisuudesta osiin). Nämä rinnakkaiset prosessit tukeutuvat edelleen koko katsojan kulttuuriseen tietämykseen.

Nykyiset hahmopohjaiset indeksointi- ja hakumenetelmät pystyvät parhaimmillaan hyödyntämään segmentointia – toisin sanoen tunnistamaan kuvassa näkyvien kohteiden kaksiulotteisia hahmoja (Gong 1998, del Bimbo 1999). Tällöin ei puhuta yleensä vielä kyvystä antaa jokin semanttinen merkitys tunnistetuille kohteille. Toimivan hahmopohjaisen menetelmän tulisikin pystyä *täydentämään* ja *korostamaan* joitakin

kuvassa näkyviä *kohteita* ja niiden *suhteita* – sekä ideaalitapauksessa *projisoimaantulkintaan kulttuurista tietämystä*.

6. Lopuksi

Kysymykseen ”mitä tämä kuva esittää?” ei ole useinkaan yksiselitteistä vastausta. Kuvien merkitykset vaihtelevat ”konkreettisesta” esittävydestä hyvinkin abstrakteihin käsitteellisiin merkityksiin. Nämä käsitteelliset merkitykset *Panofsky* jakaa ikonografisiin ja ikonologisiin merkityksiin, jotka muotoutuvat konventioiden perusteella, mutta nojaavat viime kädessä laajaan kulttuuriseen tietämykseen. Kuvan käsitteelliset merkitykset ovat monitasoisia ja joskus varsin vaikeasti löydettävissä.

Toisaalta ”konkreettinen” esittävyyskään ei ole kovin konkreettista, eikä se perustu yksiselitteisesti kuvan ja todellisuuden samankaltaisuudelle tai tiettyihin ”hahmoinhin”. *Gombrich* painottaa esittävyyden olevan pohjimmiltaan aina kuvan luomista mielessä; näkemistä jonakin tiettyssä kulttuurisessa kontekstissa. Tämä projektiivinen kyky perustuu yleisempään ilmiöön, jota filosofiassa ja psykologiassa kutsutaan havainnon teoriapitoisuudeksi.

Kuvien merkityksiin (kaikilla tasoilla) liittyy toisaalta *säännönmukaisuutta* ja konventioita, mutta toisaalta myös tietoista *säännönrikkomista*. Yllä olemme pyrkineet esittelemään tätä ilmiötä taidehistorian valossa: primitiivinen taide – tai vaikkapa egyptiläinen taide – on hyvin tiukasti konventioiden säätelemää sekä esittämistapojensa että (oletettavasti) abstraktien merkitystensä puolesta. Kreikkalaisen taiteen nousu perustui pitkälti tällaisen konventionaalisuuden hylkäämiseen, vaikka se yhä noudatti kaanonina mm. ihmisruumiin representaatioissa. Renessanssista lähtien uusia konventionaalisia esittämistapoja on luotu useita, mutta samalla on ilmennyt vahva tendenssi konventioiden rikkomiseen ja niiden rajojen kokeilemiseen. Taiteilijat ovat toisaalta pyrkineet konventionaalisiin keinoihin kertomaan teoksillaan enemmän kuin ennen. Toisaalta mielenkiinto näyttää usein suuntautuneen kysymykseen, kuinka pitkälle säännöistä tai samankaltaisuudesta voidaan luopua esittävyyden silti kärsimättä. Samalla katsoja on noussut uudella tavalla keskeiseen asemaan: teoksen esittävyys perustuu viime kädessä katsojan kykyyn nähdä eroja ja projisoida merkityksiä, sillä objektin ja sen kuvallisen esityksen välinen samankaltaisuus on löydettävä tai luotava.

Tässä taidehistoriallisessa viitekehyksessä olemme tarkastelleet kolmea kuvahaun perusongelmaa: *kuvatari-*

peiden määriteltävyyttä, kuvien sisällön kuvailtavuutta ja hahmopohjaisen indeksoinnin rajoittuneisuutta. Kuvien merkitysten varioivuus ja sumeus vaikuttavat näihin kaikkiin oleellisella tavalla. Kuvan kaikkia esittäviä ja käsitteellisiä merkityksiä ei voida indeksoida ja ainakin osa näistä merkityksistä on sellaisia, ettei niitä voida löytää periaatteessakaan hahmojen avulla; ne eivät ole hahmojen ominaisuuksia, vaan tiedostavan ihmisielen antamia merkityksiä. Osa näistä merkityksistä syntyy edelleen vasta asiayhteydessä ts. kun kuva asetetaan johonkin kontekstiin. Lähes mikä tahansa näistä potentiaalisista merkityksistä on kuitenkin mielekäs tiedontarve. Kuvat tarpeet heijastavatkin tätä kuville ominaista merkitysten sumeutta ja siten ko. tiedontarpeetkin ovat usein olennaisesti sumeita. Kuinka pitkälle (ja millä keinoilla) kuvahakujärjestelmän tulisi kyetä tyydyttämään näitä tarpeita on avoin kysymys.

Vaikka tarkastelu onkin tapahtunut taidehistoriallisesta näkökulmasta ja esimerkkinä on käytetty taideteoksia, argumentaation voidaan väittää pätevän myös laajemmin. Lähes kaikki kuvat – valokuvat, piirrookset, merkit, kaaviot – ovat merkityksiltään varsin kompleksia. Niiden merkitys ei muodostu yksi - yhteen suhteesta kohteen kanssa vaan siitä, mitä niiden tulkitaan esittävän.

Viite:

¹Kuten Gombrich huomauttaa tähän perustuu mm. Rorschachin mustetahra testi, joissa koehenkilön tulee sanoa, mitä hän näkee muodottomassa mustetahrassa.

Hyväksytty julkaistavaksi 12.9.2001.

Lähteet:

- Armitage, L.H. & Enser, P.G. (1997). Analysis of user need in image archives. *Journal of Information Science* 23:287-299.
- Beardsley, M.C. (1981): *Aesthetics – Problems in the Philosophy of Criticism*. Hackett, Indianapolis.
- del Bimbo, A. (1999). *Visual Information Retrieval*. San Francisco, Morgan Kaufmann.
- Eakins, J.P. (1996) Automatic image content retrieval - are we getting anywhere? Teoksessa: Proceedings of the Third International Conference on Electronic Library and Visual Information Research, De Montfort University, Milton Keynes, May 1996. s. 123-135.
- Enser, P.G. (1995) Pictorial information retrieval. *Journal of Documentation*, 51:126-170.
- Gleitman, H. (1991): *Psychology*. New York, Norton & company.
- Gombrich, E.H. (1996): *Art and Illusion – A Study in the Psychology of Pictorial Representation*. Phaidon Books, Lontoo. (1 painos 1960)
- Gong, Y. (1998). Intelligent Image Databases. Towards Advanced Image Retrieval. Kluwer, Dordrecht.
- Ingwersen, P. (1992). *Information Retrieval Interaction*. Taylor Graham, Lontoo.
- Karhula, P. (1989). Sanomalehtien sisällönkuvailun kehittäminen. Tampereen yliopisto, Informaatiotutkimuksen laitos, Pro gradu –tutkielma, 216 s. + liitt. 13 s.
- Keister, L.H. (1994). User Types and queries: impact on image access systems. Teoksessa: Fidel R. ym. (toim.), *Challenges in indexing electronic text and images*. Learned Information, Medford, New Jersey, 1994. s. 7-22.
- Markey, K. (1986). Subject Access to Visual Resource Collections: A Model for Computer Construction of Thematic Catalogues. Greenwood Press, Westport.
- Markkula, M. & Sormunen, E. (2000), End-User Searching Challenges Indexing Practices in the Digital Newspaper Photo Archive. *Information Retrieval* 1(4):259-285.
- Ørnager, S. (1995) The newspaper image database: empirical supported analysis of users' typology and word association clusters. Teoksessa: *Proc. of the 18th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*. ACM, New York, 208-218.
- Panofsky, E. (1987). *Meaning in the Visual Arts*. Penguin Books. (1. painos 1955.)
- Rasmussen, E.M. (1997). Indexing images. Teoksessa: Williams M-E, (toim.) *Annual Review of Information Science and Technology* 32. Information Today, Medford, New Jersey, 1997. s. 169-196.
- Shatford, S. (1986) Analyzing the subject of a picture: a theoretical approach. *Cataloguing and Classification Quarterly*, 6:39-62.

Sormunen, E., Markkula, M. & Järvelin, K. (1999). The Perceived Similarity of Photos - Seeking a Solid Basis for the Evaluation of Content-based Retrieval Algorithms. Teoksessa: Draper, S. et al. (toim.): *Mira 99: Evaluating interactive information retrieval*. Glasgow, UK, 4-16 April, 1999. *Electronic Workshops in Computing*. URL: <http://www.ewic.org.uk/ewic/workshop/fetch.cfm/MIRA-99/>.

Svenonius, E. (1994). Access to Nonbook Materials: The Limits of Subject Indexing for Visual and Aural Languages. *Journal of the American Society for Information Science* 45(8):600-606.

Tämän numeron kirjoittajat:

Anttiroiko, Veli-Antti, Professori, Tampereen yliopisto

Saarti, Jarmo, FT, Kuopio

Sahavirta, Harri, FM, Helsinki

Savolainen, Reijo, Professori, Tampereen yliopisto

Sormunen, Eero, YTT, Tampereen yliopisto

Tuominen, Kimmo, YTT, Espoo