

9. Visuaalinen tiedonhaku

Informaation perinteinen esitys: teksti

- tiedonhakumenetelmät ensin tekstille, vasta myöhemmin kuville, äänelle, multimedialle

Kuva ja teksti haun kohteena?

- teksti säännöllisempää
- kuvan kohdalla havainto ja sen tulkinta (käsitys) tärkeä
- käsitys kuvan sisällöstä vaikeasti määriteltävissä samanlaisuus? relevanssi?
- indeksointi: 1- ja 2-ulotteisen tiedon ero
- videokuva (videon jakso): lisää ulottuvuuksia kuvien suhteet, liike, ...

Visuaalisuuden muita IR-yhteyksiä:

- kyselyn ilmaisu graafisesti (esim. ehdot, yhdistäminen)
- hakutuloksen havainnollistaminen (tekstidokumenttia, sen osia jne voidaan esittää graafisesti: esim. TileBars)
- dokumenttijoukon suhteiden kuvaus

Lähteitä:

- Gupta & Jain, Visual information retrieval. CACM 40,5 (1997), 71-79.
- Lew, M.S., Next-generation Web searches for visual content. Computer 33, 11 (2000), 46-53.
- Del Bimbo, Visual Information Retrieval. Morgan Kaufmann, 1999.

Visuaalisen tiedonhaun sovelluskohteita:

- taidekokoelmat, museoiden kokoelmat
- valokuva-arkistot
- mallikokoelmat (kankaat, vaatteet, muodot)
- logot yms. symbolikuvat yksilöllisyys
- lääketieteelliset kuvat säännönmukaisuudet, yksilön tilan muutos
- kasvat (tunnistustehtävät)
- suunnittelutehtävät (engineering) eri tasot, kokonaisuus
- maantieteelliset kuvat (kartat, satelliittikuvat) kuvien yhteydet, suhteet, ympäristöt
- videoarkistot (haku, koostaminen)

Esim. hae *'noin kahden sekunnin videopätkä, jossa punainen auto ajaa rantatietä kirkkaana päivänä ja katoaa mutkan taakse'*

- kysely tekstimuodossa (Boolean kyselynä?) monta tekijää
- likimääräiset arvot
- kuvan tuottama vaikutelma vs. määritellyt arvot

Visuaalinen informaatio

- varsinainen kuva: visuaaliset piirteet
- kuvan kuvaus metatietona kuvauspaikka, -aika triviaaleja aihe vaikeampi

Yksinkertaisia visuaalisia piirteitä: pikselitaso

Esim. *'hae kuvat, joissa alue x..y on oranssia'*
(oranssi määriteltynä väriarvojen perusteella:
red = 255, green = 130, blue = 0)

'hae kuvat, joiden keskialue on suunnilleen samaa väriä kuin mallikuvassa'
keskialue?
värien etäisyys?

'hae kuvat, jotka saadaan mallikuvasta enintään k pikselin sivuttaissiirrolla'

- sopivat huonosti käyttäjälle
- pikselimalli on herkkä epätarkkuudelle (noise)
- rotaatiot tms. muunnokset vaikuttavat
- valoisuusaste yms. yleiset tekijät vaikeita hallita

Mitä pikselikuvista voidaan havaita ?

- kuvien väliset selvät erot, esim. peräkkäisten videokuvien ero leikkauskohdassa
- jäsentyneen kuvan (objekteja, alueita jne) piirteitä

Esim. satelliittikuvat:
määritellään kohdeobjektien muodot (ihminen) jaetaan kuva-alue osiin (esim. ruudukkoon)

'hae kohdat, joissa on vähintään 5 ajoneuvoa tietyntyyppisen kohteen lähellä'

'hae kohteet, jotka ovat edellisiin kuviin verrattuna liikkuneet vähintään ... metriä'

- ihmisen osuus näissä tulkinnoissa keskeinen

Ihmisen rooli on kuvainformaation tulkinnessa yleensä vielä paljon vahvempi:

- samaa kuvaa eri ihmiset tulkitsevat eri tavoin (samoin kylläkin tekstin sisältöä?)
- selautustyyppinen 'haku' luonnollista, nopeaa (kuvien esittämisestä voi tulla pullonkaula)

Kuvan sisällön kuvaaminen

- väri, muoto, teksturi, osat ja niiden suhteet

1° Väri

- värisävy (hue) ja kylläisyys (saturation)

voidaan kuvata esim. histogrammoina,
 - voidaan laskea dominoivat värit
 - voidaan laskea kuvien värietäisyys
 - värin lisäksi voidaan johtaa muotoja

- psykologisia mittoja: kirkkaus, lämpimyyys, vaaleus
 - värien nimet
 + monimutkaisempia järjestelmiä (geometria)

Esim. *'vähintään 30 % taivaansinistä ja samoin ruohonvihreää'*
 (yhtenäisinä kenttinä?)

- kuvan jako väriösiin (ositus ja kullekin osalle omat histogrammansa → quad tree)

'vähintään 20 % oranssia tai punaista toisessa yläneljänneksessä, sama määrä keltaista toisessa, vähintään 40 % tummaa ruskeaa alapuoliskossa'
 = auringonlasku ?

3° Teksturi

= hahmontunnistuksen perussuure, joka perustuu pinnan kirkkauden vaihteluun ja kuvaa pinnan 'luonnetta':

sileä (ei varsinainen tekstuurilaji)
 toistava (esim. 'kopanpohja')
 suuntautuva ('pyyhkäisty')
 hienojakoinen (toistava, ilman erityistä muotoa)

- teksturi voidaan esittää numeerisena vektorina, samanlaisuusvertailu etäisyysmitoilla
 - tunnistaminen riippuu tarkkuustasosta

4° osien väliset suhteet

- topologisia ominaisuuksia:
 lähekkäisyys, vierekkäisyys, sisältyminen, päällekkäisyys
 - suunnat: oikea/vasen, ylhäällä/alhaalla

- ihmisen käsitys monipuolisempi: tausta/edusta, pysty- ja vaakasuunnan painotus

2° Muoto (shape)

- geometriset muodot helppoja kuvata metatietona
 - yleisessä tapauksessa:

kuvankäsittelytekniikoilla tunnistetaan kuvasta väritään riittävän yhtenäisiä segmenttejä, ja niille joukko piirteitä:
 väri, koko, pitkulaisuus, suuntaisuus, keskeisyys
 piirteet = vektori, joka on segmentin edustaja

→ kyselyissä voidaan viitata kuvan osiin segmentteinä:

'hae kuvat, joiden keskellä on dominoiva valkoinen neliö'

'hae kuvat, joissa on kaksi sinistä ympyrää ja punainen soikio suunnassa 135° lähellä keskustaa'

Huom.

Piirteet muuttuvat hallitusti useissa operaatioissa, esim. skaalaus, rotaatio, peilaus

Moniin todellisiin muotoihin tarvitaan semanttista tulkintaa; esim. kävelevä hevonen (eläinten muut asennot vielä vaikeampia ...)

Erikoistapauksia:

- sovelluskohtaiset tiedot helpottavat kuvaamista, erityisesti muotojen tulkintaa
 - piirteiden automaattinen tunnistus onnistuu hyvin vain rajoitetulla sovellusalueella
 lääketieteelliset kuvat, eläimet, kasvit ...?

- kasvonpiirteiden tunnistus

- järjestelmälle on voitu opettaa noin 20 piirteiden tunnistaminen kasvokuvasta
 - haku näihin perustuen
 - piirteet eivät välttämättä 'ihmillisiä' eli tyyppiä silmät, nenä, suu, ...

- toinen lähtökohta on juuri mainittujen peruspiirteiden mukainen käsittely (silmänseudun, leuan jne tyyppilliset kuvalliset muodot)

Visuaaliset kyselyt

- selaaminen (browsing) ja sisältöön perustuva kysely usein vuorotellen (ihmiskyvyt käyttöön sopivissa vaiheissa)
- sisältöön perustuvia kyselytyyppejä:
 - avainsanapohjaiset
 - mikäli avainsanoille on määritelty vastineita kuvaominaisuuksissa ('objekttiluokkia')
 - (metatiedon kautta tavanomainen tapa)
 - Esim. *'autot'*
 - esimerkkipohjaiset
 - käyttäjä antaa esimerkkikuvan (tai osan) järjestelmä analysoi esimerkin ja hakee samantyyppisillä ominaisuuksilla varustetut kuvat (järjestyksessä parhaat)
 - relevanssipalautetekniikka sopii hyvin: käyttäjä valitsee seuraavaa vaihetta varten lisäesimerkkejä
- 1° esimerkki kuvakejoukosta
 - ominaisuuskuvaukset ehkä vahvoja
 - esim. vaaleatukkainen ihminen
 - osien väliset suhteet sopivat hyvin
 - 'auton vieressä seisova ihminen'*

2° luonnos tai mallikuva

- luonnos voi olla vähäinen muototekijä
 - esim. vapaamuotoinen perusgeometria
 - viiva, kaari; kirjainmuodot
- tarkempi piirros, jossa esim. värit tärkeitä (vaatii piirtämistaitoa)
 - käyttöliittymä voi tukea:
 - väripaletti, interaktiivinen sekoitus esim. liukusäätimin (myös muotoja, 'clip art')
- mallikuva esim. jostain kirjastosta, jatkokyselyssä edellisestä tuloksesta värien, tekstuurin ym. säätely ja painotus

Samanlaisuuden määrittäminen

- ihmisellä monimutkainen toiminta; erotetaan välitön ja tulkintaan perustuva
 - esim. nurinpäinen ja oikein päin oleva kuva (kasvot tms.)
- tulkintaan perustuva tärkeä useilla sovellusalueilla, käyttäjällä odotuksia tai kokemusta
- välitön: osien suhteet, värit, muoto, tekstuuri vaikuttavat ('objektiiviset' kokeet näitä?)
- samanlaisuus on harvoin täsmällistä
 - mittaluvut, järjestys
- malleja:
 - metriset piirrevektoreihin, etäisyyksiin tms. perustuvat
 - muunnosetäisyydet: elastinen muodon muutos, vähittäiset rajoitteet muodon kehittämisessä