

POIZVEDOVANJE PO SLIKAH NA PODLAGI VSEBINE*

Tomaž Kranjc
Jasna Maver

Oddano: 20. 9. 2007 – Sprejeto 19. 10. 2007

Pregledni znanstveni članek

UDK 025.4.036:004.65

Izvleček

Ob pojavu znatnega števila velikih slikovnih podatkovnih zbirk, v katerih so bile slike večinoma opisane s tekstovnimi deskriptorji, se je zaradi nepopolnosti tovrstnega načina opisovanja sčasoma pojavila potreba po drugačnem pristopu do indeksiranja slik in poizvedovanja po njih. Alternativa tekstovnim sistemom poizvedovanja po slikah je poizvedovanje po slikah na podlagi vsebine. Slike so indeksirane na osnovi svoje dejanske vsebine in značilnice, kot so barva, tekstura in oblika so izražene s številčnimi vrednostmi. V procesu poizvedovanja sistem izračuna vrednosti značilnic za podano sliko na isti način, kot so izračunane vrednosti za slike, shranjene v zbirki, in opravi primerjavo. Takšni sistemi so znani pod imenom CBIR (content-based image retrieval) sistemi. V delu predstavimo osnovne karakteristike CBIR sistemov, prosto dostopnih prek svetovnega spleta, in jih primerjamo gleda na razvojno okolje, v katerih so nastali, značilnicah, ki jih uporabljajo za opisovanje vsebine slik, in zmožnostjo uporabniškega vmesnika, ki ga sistem ponuja.

Ključne besede: sistemi za poizvedovanje, poizvedovanje, slikah, CBIR, značilnice, ekstrakcija značilnic, podatkovne zbirke, vizualne informacije, informacijski sistemi, informacijska tehnologija, Internet

Review article

UDC 025.4.036:004.65

Abstract

With development of technology, large image databases have become reality. Pictures in these databases were mainly described with keywords, but due to insufficiency of such description, the need for different approach to image indexing and retrieval has

* Članek je nastal na osnovi diplomske naloge, za katero je Tomaž Kranjc leta 2007 prejel Berčičevo nagrado.

arisen. Content-based image retrieval is an alternative to textual based systems. Images are indexed on the basis of their actual content, features like color, texture and shape are presented numerically. In retrieval process, features for input image are extracted with the same algorithm as used for pictures already stored in the database, and values are compared. Such systems are known as CBIR (content-based image retrieval) systems. Basic characteristics of CBIR systems, which are freely available on the Internet, are presented. The systems are compared with respect to the environment in which they were built, applied features for content description and ability of user interface.

Keywords: retrieval systems, image retrieval, CBIR, features, feature extraction, image databases, visual information, information systems, information technology, Internet

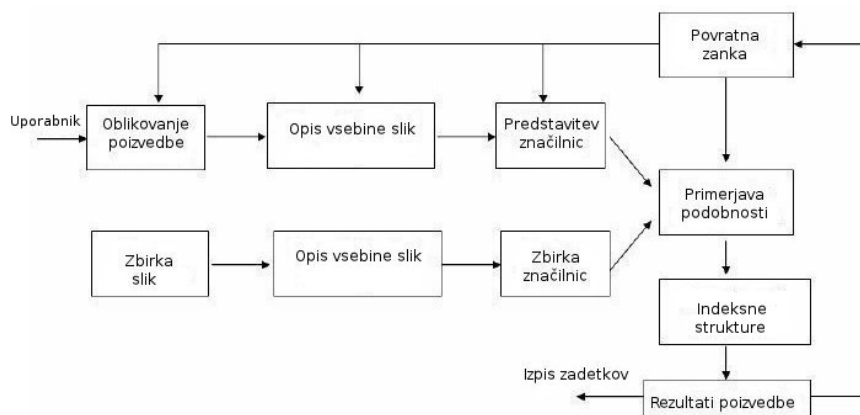
1 Uvod

Danes poznamo številne oblike, v katerih se pojavljajo informacije. Daleč najbolj je raziskano upravljanje s tekstovnimi informacijami, manj pa upravljanje z vizualnimi in zvočnimi. Slike so bile od nekdaj prisotne v človeški interakciji. Že praljudje so se izražali s pomočjo slikovnih informacij, ki so krasile stene njihovih jamskih domovanj. Od nekdaj so bile slike prisotne v človeškem vsakdanu, v modernem svetu pa so se slike uveljavile kot nepogrešljiv pripomoček na številnih strokovnih področjih. Fotografija in televizija sta močno povečali uporabo slik, računalnik pa je pomenil dokončen razmah njihove uporabe, še posebej, ko so bile v devetdesetih razvite tehnike za digitalno zajemanje slik, njihovo obdelavo, shranjevanje, prikazovanje in posredovanje. Ob tem se je pojavilo znatno število velikih slikovnih podatkovnih zbirk, v katerih so bile slike večinoma opisane s tekstovnimi deskriptorji. Zaradi nepopolnosti tovrstnega načina opisovanja (Bird, Elliot in Hayward, 1999) se je pojavila potreba po drugačnem pristopu do indeksiranja slik in poizvedovanja po njih. Cawkell (2000) povzema zgodovino soočanja z velikimi težavami pri tekstovnem opisovanju zelo velikih slikovnih zbirk.

Za razliko od tekstovnih sistemov poizvedovanja po slikah, deluje poizvedovanje po slikah na podlagi njihove vsebine po bistveno drugačnem principu. Slike so indeksirane na osnovi njihove dejanske vsebine in značilnice, kot so barva, tekstura in oblika so izražene s številčnimi vrednostmi. Tipični CBIR (angl. *Content-Based Image Retrieval*) sistem omogoča uporabnikom, da formulirajo poizvedbo s podajanjem primera podobne slike, ki jo iščejo. Sistem izračuna vrednosti značilnic za podano sliko na isti način, kot so izračunane vrednosti za slike, shranjene v zbirki, in opravi primerjavo. Rezultat poizvedbe so slike, kjer so značilnice po vrednosti najbližje značilnicam podane slike.

Preglednica 1: Osnove delovanja CBIR sistema.

(Predelano po: <http://lmb.informatik.uni-freiburg.de/research/isearch/cbir.h1.gif>)



Najzgodnejša uporaba termina CBIR sega v leto 1992, ko je Kato s tem izrazom opisal svoje eksperimente poizvedovanja po slikah na osnovi avtomatično pridobljenih značilnic barve in oblike. Izraz se je kmalu uveljavil in doživel široko uporabo. Termin se lahko nanaša tako na enostavne kot na semantične značilnice, toda proces ekstrakcije značilnic iz slik mora biti pretežno avtomatski (Eakins in Graham, 1999). CBIR je hitro razvijajoča se tehnologija z visokim potencialom in naraščujočim številom objav. Težave poizvedovanja po slikah so bile devetdesetih deležne široke pozornosti in iskanje rešitev je bilo zelo aktivno.

CBIR se razlikuje od klasičnega poizvedovanja v tem, da so slike v svojem bistvu nestrukturirane. Digitalne slike sestojijo izključno iz slikovnih elementov-pikslov različnih intenzitet, ki niso vnaprej povezani v pomenske skupine, npr. za sliko obraza oko, lica, nos, usta, brada, ... oziroma segmente. Tako za sliko, na kateri je obraz, ni vnaprej poznano, da se na njej nahaja obraz, niti kateri slikovni elementi v sliki pripadajo obrazu. Preden je sploh možno razglabljanje o pomenu slike, je potrebna ekstrakcija uporabnih informacij-značilnic iz sicer brezpomenskih podatkov. Slika se torej bistveno razlikuje od teksta, kjer so črke že strukturirane v besede, ki nosijo pomen, in besede naprej v stavke.

Podobnost slik je subjektivno področje, saj ima lahko ista slika za dva človeka precej različen pomen. Ko nekdo gleda sliko, na kateri so prikazane redke vrste živali, mu bo slika predstavljala »neznane živali«, strokovnjak pa bo natančno vedel, za katere živali gre in bo drugače interpretiral sliko – z natančnim poimevanjem na njej prikazanih živali.

Značilno je, da ljudje radi povezujemo predmete z njihovo funkcijo. Ta fenomen se odraža v tekstovnih sistemih. Če vzamemo za primer kozarec, ga glede na

njegovo funkcijo (pitje) vedno enako poimenujemo, ne glede na veliko potencialno raznolikost kozarcev, ki se lahko odraža v različnih oblikah, barvah, materialih. Ali drugače povedano, če opredelimo kozarec kot semantično kategorijo, lahko znotraj te kategorije beležimo nešteto njenih vizualnih predstavitev. Poleg tega lahko isti kozarec po svoji vizualni vrednosti bistveno variira glede na vidni kot, geometrijsko rotacijo ali osvetlitev. Resnično uporaben CBIR sistem bi moral biti sposoben izvajanja tovrstnih povezav, kar pa je v praksi težko doseči (Johansson, 2000). Brez dvoma obstaja razkorak med človeškim in računalniškim dojetjem podobnosti.

Pogosto se zgodi, da je rezultat poizvedbe v CBIR sistemu bistveno neskladen z uporabnikovimi pričakovanji. Za slike je določitev informacijske vsebine bolj zahtevna kot je to v primeru teksta. Pri tekstu ima vsaka beseda končno število pomenov in njena semantična vrednost je takoj jasna ali pa jo razjasnimo z izbiro ene od končnega števila možnosti, z analizo na nivoju stavka ali odstavka. Kompleksnosti vizualnih informacij se zavemo, ko skušamo samo z besedami opisati sliko. Slika ima lahko več interpretacij. Tipičen primer so iluzije, slike, ki zavajajo človeško zaznavanje, v smislu da možgani očem ob pogledu na sliko sporočajo zaznavo, ki ni skladna z dejanskim stanjem. Poznamo več vrst iluzij, optične iluzije, psihološke iluzije, iluzije kontrasta in druge. Znan primer iluzije je ženski obraz, ki je sočasno tudi podoba saksofonista. Nekateri ljudje tudi po dolgem strmenju v sliko ne bodo ugledali druge podobe.



Slika 1: Primer iluzije, ženski obraz, ki je sočasno podoba saksofonista.
(Vir: http://www.sapdesignguild.org/resources/optical_illusions/images/sax.gif)

Problematiko lahko razvijemo dalje z drugim primerom. Sliko preoranega polja bo sistem morda zaznal kot zelo podobno sliki tablice čokolade, a njuni semantični vrednosti sta povsem različni. Do navidezne enakosti dveh vsebin slik pride

zaradi podobne vrednosti izgleda, ki jo zaznamuje njuna specifična kombinacija barv in tekstur. Gre za problem t. i. sorodnih slik (angl. *rhyming images*), ki po svojem semantičnem pomenu nimajo dosti skupnega, a so si gledano skozi algoritem, po katerem jih primerja sistem, zelo podobne (Gupta, Santini in Jain, 1997).

Velik razkorak med zmožnostjo CBIR sistemov in zmožnostjo človekove interpretacije vizualnih podatkov je posledica dejstva, da človeški možgani daleč presegajo nivo izločanja značilnic. Podobe in oblike interpretirajo kot predmete. Poleg tega imajo možgani sposobnost razpoznave specifičnega predmeta znotraj kategorije (npr. oseb, ki so na sliki). Razkorak med uporabljenimi značilnicami za matematično ponazoritev slike in interpretacijo slike s strani človeških možganov, imenujemo **semantični prepad**. Eden glavnih raziskovalnih problemov je, kako zmanjšati ta prepad (Geradts, 2002).

2 Značilnice

CBIR sistemi operirajo s pomočjo značilnic. Enostavne značilnice, ki predstavljajo vsebino slike, so izračunane za vsako sliko v zbirki in shranjene. Za sliko ali skico, uporabljeno v poizvedbi, se izračunajo značilnice na enak način in izvede se matematična primerjava z vrednostmi, shranjenimi za slike v zbirki.

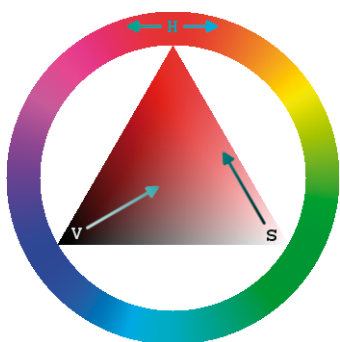
Značilnica je lastnost, pridobljena s pretvorbo originalnega vizualnega predmeta skozi algoritem analize slike in odraža specifične značilnosti slike. Značilnica je tipično predstavljena kot niz številčk oziroma kot vektor. Dve pomembni in pogosti operaciji nad vektorji značilnic sta projekcija in razdalja. V primeru slik se razdalja kot mera podobnosti med slikama pogosteje uporablja. Večja kot je razdalja med značilnicami, manjša je podobnost med slikama.

Uporaba značilnic pri CBIR poveča možnost, da bo poizvedovanje znotraj sistema bolj podobno človeškemu pogledu na slike. Lep primer so barve, ki v resnici sploh ne obstajajo, so samo spekter svetlobe, ampak so del človeškega pogleda na slike. Barva zato predstavlja uporabno značilnico (za razliko od infrardeče svetlobe, ki je ne vidimo). Poleg barv sta dve tipični enostavni značilnici tekstura in oblika. Zdi se, da so enostavne značilnice skupne vidnim sistemom večine ljudi. S tem pa ni izključena uporaba kompleksnih značilnic. Obstaja mnogo načinov uporabe, kjer so uporabne specifične, kompleksne značilnice. Primer so MR slike v medicini ali infrardeča svetloba v primeru satelitskih slik. Vsaka uporaba zahteva rešitev zase. Kompleksne značilnice so ponavadi izračunane iz enostavnih značilnic. Če je za enostavne tipična splošnost, so kompleksne značilnice mnogo bolj področno specifične, v smislu da so izdelane za uporabo v posebnem specifičnem razredu slik ali predmetov. S točno določenimi kompleksnimi

značilnicami se ponavadi ukvarjajo znanstveniki s točno določenega področja. Sistem CBIR, ki v svojem delovanju zajema procesiranje specifičnih kompleksnih značilnic, je pogosto zelo drag in preden se relevantnim slikam prilagodi kompleksne značilnice, se prej odstrani večina nerelevantnih slik iz zbirke.

Barva je najpogosteje uporabljena in pogosto najpomembnejša značilnica v CBIR sistemih. V računalniškem svetu poznamo dva standardna barvna modela:

- RGB (angl. *red, green, blue*) in
- CMY (angl. *cyan, magenta, yellow*).



Slika 2: Ponazoritev barvnega modela HSV.

(Vir: http://upload.wikimedia.org/wikipedia/en/2/21/Hsv_sample.png)

Ta dva barvna modela pa ne ustrezata dobro človeškemu načinu dojemanja slik (Müller, Michoux, Bandon in Geissbuhler, 2004; Geradts, 2002). Vsaka točka je sestavljena iz kombinacije treh barv različne intenzitete, vendar so te tako majhne, da človeško oko vidi točko kot enobarvno. Alvy Ray Smith je leta 1978 ustvaril barvni model HSV (angl. *hue, saturation, value*), ki najbolj ustreza človeškemu dojetanju barv. Gre za uporabniško prilagojen model, ki bolje odraža lastnosti barve. *Hue* predstavlja barvo (npr. rdeča). *Saturation* predstavlja nasičenost barve in se opredeljuje v razponu 0 do 100 odstotkov. Nižja kot je nasičenost barve, manj je barva intenzivna. *Value* predstavlja svetlost barve in ločuje temno barvo od svetle (glej Sliko 2).

Najenostavnejši način določanja barvne podobnosti med dvema slikama je primerjava njunih histogramov. Za vsako sliko, dodano v zbirko, se izračuna barvni histogram, ki prikazuje deleže pikslov posameznih barv v sliki. Ti podatki se shranijo in uporabijo pri procesiranju uporabniških vnosov pri poizvedovanju. Ko uporabnik zastavlja poizvedbo, določi želeni delež poljubnih barv v sliki. Po izvedeni primerjavi sistem uporabniku prikaže slike z najbolj podobnimi barvnimi histogrami. To tehniko sta v začetku devetdesetih razvila Swain in Bal-

lard (1991), a histogrami so merili samo globalne vrednosti in v najslabšem primeru sta lahko dve povsem različni sliki imeli identična histograma. Kasneje je bila tehnika nadgrajena, in danes večinoma vsebuje kombiniranje barvnega z elementi prostorskega ujemanja. Čeprav gre za enostaven postopek, so rezultati takih primerjav dobri (Forsyth, 1999).

Tekstura se nanaša na ureditev osnovnih sestavin slike. Na digitalni sliki se tekstura odraža kot prostorska ureditev pikslov na sliki – ta ureditev pikslov se vidi kot sprememba v intenziteti vzorca (Grosky in Stanchev, 2000). Podobnost tekstur je lahko uspešno uporabljena pri razločevanju med področji slik s podobnimi barvami (npr. morje in nebo). Poizvedbe z vnosom tekstur so lahko oblikovane na podoben način kot poizvedbe z vnosom barv – z izbiranjem zelene teksture iz ponujenega nabora (palette). Tipična lastnost teksture je, da jo je skorajda nemogoče opisati z besedami.

Statistično preračunane mere tekstur se pogosto urejajo v vektor, v katerem postanejo medsebojno objektivno primerljive. Tekstura slike lahko vsebuje različne statistično predstavljljive attribute, kot so (Gupta, Santini in Jain, 1997):

- naključnost (pove koliko je naključnosti v sliki, npr. šahovnica jo ima malo, vzorec naključnih točk pa zelo veliko);
- periodičnost (ponavljanje vzorca);
- usmerjenost (npr. črte v zastavi so usmerjene, orientirane).

Zamislimo si sliko z desetimi različnimi regijami in vsaka regija ima drugačno teksturo. Vrednosti tekstur za naključnost, periodičnost in usmerjenost določajo deset točk v koordinatnem sistemu, kjer so koordinate naključnost, periodičnost, usmerjenost. Kako podobna je ta slika neki drugi, ki ima 10 drugačnih teksturnih regij? Funkcije razdalje so določene med točkami in lahko podajo matematično vrednost podobnosti dveh slik, ključno vprašanje pa ostaja, kako dobro odražajo človeški občutek za razlike v videzu.

Oblika je kvaliteta objekta, ki je odvisna od notranje lege točk, ki sestavljajo njegov obris, ali zunanjih površin. Vizualni predmeti so v naravi primarno prepoznani po svoji obliki. Če predstavimo obliko z značilnicami, pridobljenimi iz slike, lahko primerjavo izvajamo podobno kot v primeru drugih enostavnih značilnic, s funkcijo razdalje, kjer večja razdalja pomeni manjšo podobnost. Značilnice, ki ponazarjajo obliko, delimo na globalne in lokalne, pri čemer se globalne nanašajo na prostor celotne slike, lokalne pa na okolico točk zanimanja (angl. *interest points*). Točke zanimanja so lahko npr. mesta v sliki, ki predstavljajo ekstreme, pri filtriranju slike z gaussovimi filtri. Velikost filtra oz. σ , pri kateri je bila dosežena ekstremna vrednost, pa določa velikost okolice. Obstajajo tudi drugi načini predstavitve oblike, npr. s transformacijo, s katero lahko iz osnovnih pridobimo nove informacije.

Prostorske entitete v sliki so točke, črte, regije in predmeti. Prostorske odnose med temi entitetami lahko razvrstimo v dve skupini, v smerne in topološke odnose.

Smerni odnosi (angl. *directional relationship*) upoštevajo položaje predmetov na sliki. Izraženi so kot »levo od«, »desno od«, »nad«, »pod«. Merski odnosi so pogosto povezani s smernimi odnosi in lahko izražajo, kako blizu so si predmeti, ki tvorijo odnos (razdalja) in/ali pod kakšnim kotom ga tvorijo. Smerni odnosi zahtevajo neko globalno orientiranost slike, saj so odvisni od spreminjanja rotacije in skale slike. Topološki odnosi ne vključujejo koncepta razdalje, in so neobčutljivi za tovrstne variacije. Zajemajo odnose med bližnjimi entitetami, kot so stikanje, vsebovanje in prekrivanje (Del Bimbo, 1999).

2.2 Globalni in lokalni pristop

Globalni pristop pomeni ekstrakcijo značilnic iz področja celotne slike, z neupoštevanjem lokalnih posebnosti. Globalne vrednosti značilnic same pogosto niso zadostne. Enostaven primer je s področja medicine, kjer se zdravniki osredotočijo na obolelo tkivo na radiološkem posnetku, in bi bile globalne vrednosti slike neuporabne. Kot drug primer si zamislimo dve sliki z zelo podobno globalno barvno strukturo ali strukturo tekstur, ki sta si v končnem videzu lahko zelo različni, npr. slika japonske zastave in slika otroka, ki je oblečen v rdeč kombinizon in se igra na zasneženem dvorišču.

Lokalni pristop pomeni ekstrakcijo značilnic iz določenih delov slike. CBIR sistemi pogosto uporabljajo segmentacijo slik, ki funkcionira najbolje, kadar slika sestavlja eden ali nekaj predmetov, ki se jasno razlikujejo od ozadja. Pri večjem številu manjših predmetov pogosto prihaja do prekrivanja in neželene barvne podobnosti z ozadjem, kar oteži razlikovanje. **Lokalne značilnice** se nanašajo na parametre, pridobljene iz posameznih segmentov slike in razmerja med njim. Vsaka lokalna značilnica je opremljena s pozicijo. **Predmetne značilnice** so značilnice, izračunane za vsak predmet v sliki posebej. Predmeti so lahko ločeni intelektualno ali na nek avtomatski oz. polavtomatski način. V primeru **regionalnih značilnic** je slika razdeljena na regije, ki imajo kompaktne lastnosti glede na vsebino določene slike, lahko pa gre tudi za fiksne, vnaprej določene regije, neodvisne od vsebine slike. Meja med predmeti in regijami je včasih nejasna. Homogene regije se včasih interpretirajo kot predmeti (Johansson, 2000).

Če v CBIR sistemu iščemo slike, ki imajo več kot 40 odstotkov modrine na vrhu in več kot 30 odstotkov oranžne barve na dnu, morda iščemo sliko prizora s plaže. Primer kaže, kako utegne uporabnik opisati sestavne elemente zelene slike in njihovo postavitev v prostoru slike. Sistemi, ki uporabljajo fiksne regije ločuje-

jo prostor slik v določeno število preddefiniranih regij. Lahko gre npr. za preproste bloke osem krat osem pikslov. Uporabniki določijo, katere regije so za njih pomembne. Sistem izračuna značilnice za vsako izbrano regijo in oceni podobnost slik glede na regionalno ujemanje vrednosti značilnic.

Sticker in Dimaijev (1996) sistem uporablja funkcijo, ki daje več teže na osrednjo ovalno regijo – pomembnost se progresivno zmanjšuje, ko gremo stran od centra. Sistem je neobčutljiv za devetdeset-stopinjske rotacije. Izpostavljeno je pomembno vprašanje, kako naj sistem, ki omogoča poizvedovanje v prostoru značilnic, upošteva vizualne predmete, ki so geometrijsko rotirani, različno osvetljeni ali vidni pod različnim kotom, v smislu, da so značilnice, ki jih uporablja, invariantne za tovrstne spremembe.

3 Shranjevanje in poizvedovanje

Že vrsto let se soočamo z eksplozijo rasti količine vizualnih informacij. Vsak dan nastane veliko število medicinskih, satelitskih in drugih vrst slik. Te slike so pretežno v digitalni obliki, kar omogoča enostavno obdelavo, shranjevanje, vzdrževanje in prenos slik. Za dobro izrabo teh slik pa je pomembno, da so organizirane tako, da omogočajo hitro poizvedovanje na zahtevo.

Vključevanje računalnikov v upravljanje s slikami datiramo v leto 1965, ko je Ivan Sutherland v odmevnem projektu *Sketch Pad* demonstriral izvedljivost računalniške kreacije, spreminjanja in shranjevanja slik¹. A vse do srede osemdesetih let je draga strojna oprema omejevala uporabo računalnikov v tovrstne namene. Takrat je hitro rastoči trg računalniških iger povzročil skokovit padec cen računalniške opreme, in področja, tradicionalno odvisna od slik, so kmalu uvedla računalnike v svojo dejavnost. Sledilo je obdobje elektronskih zbirk z omejenim dostopom, v zgodnjih devetdesetih pa je internet omogočil uporabnikom enostaven dostop do velike količine vizualnih podatkov.

Potrebo po učinkovitem shranjevanju slik in poizvedovanju po njih so med prvimi prepoznali upravljavci velikih slikovnih zbirk. Leta 1992 so na delavnici, sponzorirani s strani ameriške nacionalne znanstvene fundacije, izpostavili nekaj področij, kjer je bilo raziskovanje najbolj potrebno. Med drugim so izpostavili področja ekstrakcije značilnic iz slik, indeksiranja in izgradnje učinkovitega uporabniškega vmesnika. Eden glavnih izpostavljenih problemov je bila težavnost lociranja zelene slike v veliki in raznoliki zbirki. Medtem ko je dokaj enostavno identificirati zeleno sliko v majhni zbirki, s preprostim brskljanjem,

¹ Obširneje na <http://en.wikipedia.org/wiki/Sketchpad>

ali pa identificirati želeno sliko v homogeni zbirki, kjer so lahko predmeti iskanja vnaprej določeni, pa se še vedno išče učinkovite tehnike identificiranja slik v velikih raznolikih zbirkah.

3.1 Indeksiranje slik

Tradicionalno so bile slike shranjene v analogni obliki, večinoma v mapah, urejenih po policah. Stopnja indeksiranja v takih zbirkah je bila močno povezana s pomembnostjo zbirke, načinom uporabe in s številom človeških kadrov, ki so bili na voljo. Iskanje posameznih slik v takih zbirkah je bilo neizogibno naporno opravilo in znanje o zbirki je bilo pogosto neločljivo povezano z osebo, ki jo je vzdrževala, mnogo manj pa z dejanskimi uporabniki. Danes je v slikovnih podatkovnih zbirkah še vedno najbolj običajno opisovanje slik z uporabo ključnih besed, pri čemer upravitelji za pomoč pri klasifikaciji pogosto uporabljajo indeksne sheme, ki so jih razvili sami ali v sodelovanju z uporabniki in odražajo specifično naravo teh zbirk. Tehnike indeksiranja slik z uporabo ključnih besed imajo številne prednosti. Tako indeksiranje ima visoko izrazno moč, s ključnimi besedami lahko opišemo skorajda vse aspekte še tako kompleksne vsebine slike, enostavno je dodajati nove koncepte in procesa poizvedovanja ni težko avtomatizirati, saj obstajajo za ta namen številni že razviti programi.

Postopek intelektualnega indeksiranja slik ima tudi pomembne pomanjkljivosti. Gre za delovno-intenziven proces, in natančen opis ene slike lahko traja tudi do 40 minut (Eakins in Graham, 1999). V okolju zbirke z milijon in pol slikami to pomeni **milijon ur dela**. Naslednji problem je subjektivnost. Obstajajo široka razhajanja v pogledu različnih posameznikov na to, katere ključne besede ustrezajo opisu neke slike. Celo pri indeksnih jezikih, ki so bili izdelani z namenom točno določene zbirke, so raziskovalci (Enser in McGregor, povzeto po Eakins in Graham, 1999) ugotovili slabo ujemanje med uporabniškimi poizvedbami in indeksnim jezikom. Možne so tudi slovnične napake pri zapisovanju ključnih besed, katerim pa se lahko indeksir izogne, če besede izbira s seznama, uporabnik pa, če mu je sistem sposoben predlagati pravilno besedo. Problem je večplasten. Katalogizatorjev opis iste slike lahko variira tudi glede na različna časovna obdobja, v katerih opisuje isto sliko. Poleg tega se skozi čas besednjak nekega področja razvija in spreminja skupaj z razvojem področja samega. Še en problem je nedvoumnost – nekatere strukturne lastnosti slike so le težko opisljive z besedami, tako ima npr. vsaka barva širok razpon intenzitet, kar je težko opisati z besedami, še težje pa je semantično označiti oblike in teksture. Vsebinsko sliko je praktično nemogoče opisati tako, da bi zajeli poglede vseh različnih uporabnikov in vse različne načine možne uporabe slike.

Avtomatsko indeksiranje po vsebini slik ima potencialno številne prednosti pred intelektualnim indeksiranjem. Je neprimerno hitrejše, cenejše in povsem obje-

ktivno. A najbolj pomembna ostaja učinkovitost poizvedovanja. Obeh tehnik se ne da smiselno primerjati, ker sta izgrajeni, da odgovarjata na različne tipe poizvedb, je pa možno v nekaterih primerih področno specializirane uporabe zaključiti, da se tehnike CBIR bolje obnesejo kot tehnike indeksiranja in poizvedovanja po ključnih besedah. Tipičen primer so zbirke logotipov blagovnih znamk, saj so logotipi sami po sebi pretežno brezpomenski in ne morejo biti ustrezno opisani s tekstovnimi oznakami.

3.2 Hibridni pristop

Ker imata oba pristopa svoje prednosti in slabosti, je zanimivo vprašanje, kako ju integrirati v enem sistemu tako, da bi docela izkoristili prednosti obeh. Eksperimentalni rezultati so pokazali, da ima integriran pristop boljše zmogljivosti poizvedovanja kot katerakoli od obeh tehnik uporabljena sama zase (Hove, 2004). Tekstovno poizvedovanje in poizvedovanje na osnovi vsebine slik se lahko dobro dopolnjujeta. Tekstovne tehnike lahko zajamejo visoko stopnjo abstrakcije, enostavno je izvesti poizvedbo, so pa tekstovni opisi subjektivni in nepopolni in niso zmožni poizvedovanja na osnovi vsebinske podobnosti slik. CBIR sistemi lahko zajamejo enostavne značilnice slik in sprejemajo slikovne poizvedbe, toda ne zmorejo zajeti kompleksnih konceptov. Pri CBIR je za uporabnika ponavadi zahtevno podati začetni vnos, navesti mora namreč ustrezne vrednosti posameznih značilnic, ali pa podati sliko za primer oziroma narisati skico.

Hibridni pristop načeloma ne rešuje problemov, ki izhajajo iz določanja ključnih besed, zato je idealno okolje za implementacijo hibridne tehnike svetovni splet, kjer je možen avtomatski tekstovni opis slik na podlagi teksta, ki se nahaja v HTML dokumentu, ki vsebuje sliko (Lu, Williams in You, 2001). Za natančnejše opisovanje so na voljo algoritmi, ki obtežijo posamezne elemente HTML dokumenta, kar je relativno enostaven postopek, saj so v spletnih dokumentih posamezni elementi označeni s točno določenimi oznakami (angl. *tags*). Tako značilnice kot tekstovni opis slik sta torej pridobljena z avtomatskim indeksiranjem. Uporabniki lahko poizvedbo enostavno zastavijo z vnosom ključnih besed in iz rezultatov izberejo slike, s katerimi nato operirajo skladno s CBIR zmožnostmi sistema. Seznam zadetkov temelji na kombiniranemu ujemanju značilnic in teksta. Implementacija hibridne tehnike bi bila verjetno primerna za okolje digitalne knjižnice, mnogi raziskovalci jo vidijo kot možnost obvladovanja ogromne količine slik na svetovnem spletu (Chang, Smith, Beigi in Benitez, 1997).

Podrobnejši pregled hibridnih sistemov so opravili Yanai, Shindo in Noshita (2004). Podan je vpogled v njihovo delovanje in predstavljene so možne naprednejše rešitve.

3.3 Indeksne strukture

Od sistemov podatkovnih zbirk se pričakuje, da bodo podpirali učinkovit dostop do podatkov in omogočili hitro iskanje, ne glede na velikost podatkovne zbirke. Za te sisteme je izjemnega pomena kratek odzivni čas. CBIR indeksiranje slik se bistveno razlikuje od tekstovnega načina indeksiranja. Slike so predstavljene z več značilnicami, katerih vrednosti naseljujejo indeksne strukture. Ko dodajamo v sistem nove slike, se poseljenost prostora veča. To je informacijski prostor, v katerem poteka iskanje. V primeru velikih podatkovnih zbirk, z večdimenzionalnimi indeksnimi strukturami, se uporabljajo metode za zmanjševanje prostora značilnic, v katerem se poižveduje, s čimer se izboljša hitrost iskanja podatkov (Müller, Michoux, Bandon in Geissbuhler, 2004). Različne tehnike indeksiranja in zmanjševanja tega prostora skrbijo za učinkovit dostop do slik v zbirki. Pri procesiranju poižvedbe sistem na podlagi teh tehnik izloči nerelevantne slike, brez da bi moral za to prehajati skozi celotno zbirko (Del Bimbo, 1999).

3.4 Iskanje po slikah

Poižvedovanje po slikah zajema širok spekter možnih tipov poižvedb. Lahko poižvedujemo le po enostavnih atributih slike (npr. količina določene barve). Lahko poižvedujemo po prisotnosti ali ureditvi določenih predmetov na sliki (npr. stoli okrog mize). Lahko poižvedujemo po sliki, ki prikazuje tip dogodka (npr. avtomobilska dirka). Lahko poižvedujemo po konkretnih osebah, lokacijah ali dogodkih (npr. papež na obisku v Ljubljani). Lahko poižvedujemo po čustvenih stanjih, s katerimi sliko povezujemo (npr. sreča na obrazu). Lahko pa poižvedujemo po atributih, ki jih ni mogoče pridobiti iz same slike (npr. kje je slika nastala in kdaj). Tako širok razpon tipov poižvedb je eden od razlogov, zakaj je indeksiranje slik tako zahtevno opravilo.

Iskanje slik na spletu poteka večinoma še prek splošnih, prostotekstovnih iskalnikov. Ti imajo sicer vgrajene specifične iskalnike slik, a princip je isti, vnos so ključne besede, ki opisujejo sliko, dodatne možnosti se nanašajo le še na velikost in tip slik.

Pri CBIR različni sistemi uporabljajo različne značilnice, kombinacije značilnic in obtežitve značilnic. Dobro je, če lahko uporabniki določijo pomembnost posameznih značilnic pri iskanju. Problem CBIR tehnologij je, kako učinkovito priklicati iz zbirke nabor slik čim bolj podobnih temu, kar je podano v poižvedbi. Tu se kaže povsem drugačen princip delovanja CBIR v primerjavi s tekstovnimi sistemi, kjer je podani deskriptor bodisi prisoten bodisi pa odsoten v opisu, shranjenem v zbirki. Procesiranje poižvedbe v takem primeru večinoma sestoji

iz prepoznavanja tistih dokumentov, ki vsebujejo podani deskriptor oziroma so povezani z njim.

Santini in Jain (1997) pravita, da je najvažnejši cilj tradicionalnih tekstovnih sistemov pri poizvedovanju razdelitev zbirke na dva dela: na relevantne in nerelevantne dokumente, četudi so pripadniki prve skupine lahko rangirani po večji ali manjši relevantnosti. V nasprotju s tem pa je primarni cilj CBIR sistema, da glede na uporabniški vnos celotno zbirko razvrsti po podobnosti.

CBIR poizvedovanje se danes v praksi odvija predvsem na stopnji 1, ki je najnižja od treh stopenj CBIR poizvedovanja, kot jih navaja Eakins (1996). Tako poizvedovanje obsega uporabo enostavnih značilnic kot so barva, tekstura, oblika, prostorski odnosi in njihovo kombiniranje. Z uporabo teh parametrov je moč najti slike, na katerih npr. prevladuje oranžna barva, v desnem spodnjem kotu pa se nahaja zelen predmet. Poizvedovanje na tem nivoju najpogosteje poteka po principu *»najdi mi več slik podobnih tej«*. Sistemom je skupno zanašanje na avtomatizirano ekstrakcijo značilnic iz slik, brez navezave na zunanjo zbirko znanja pri indeksiranju ali drugih opravilih. Uspešno se tako poizvedovanje uporablja predvsem v zbirkah, kjer slike same po sebi pretežno nimajo pomena (npr. v zbirki logotipov blagovnih znamk).

Sistemi, ki bi v praksi uporabljali poizvedovanje na eni izmed dveh višjih stopenj, so redki in so predmet raziskovanja. Številni raziskovalci menijo, da bo ravno razvoj teh sistemov razširil uporabo CBIR tehnologije. Stopnja 2 obsega poizvedovanje po logičnih značilnicah, ki vsebujejo določeno stopnjo logičnega sklepa o identiteti predmetov, prikazanih na sliki. Poizvedovanje na tem nivoju zajema prepoznavanje prizorov in predmetov (npr. določenih tipov zgradb). Pogosto je pomembno, da identificiramo skupen tip prizora, ki ga prikazuje slika, saj je to lahko pomemben filter pri iskanju in pomoč pri prepoznavanju predmetov na sliki. Stopnja 3 obsega poizvedovanje po abstraktnih atributih. Ti zajemajo razglabljanje o pomenu in namenu predmetov, ki jih scene na slikah prikazujejo. Poizvedbe znotraj te stopnje lahko delimo na poizvedovanje po imenovanih dogodkih ali tipih aktivnosti (npr. najdi mi slike slovenskih ljudskih običajev) in poizvedovanje po slikah s čustvenim ali religioznim pomenom (npr. najdi mi slike, ki prikazujejo trpljenje).

Medtem ko stopnji 2 in 3 pogosto obravnavamo skupaj, je razkorak med poizvedovanjem na stopnjah 1 in 2 zelo velik. Stopnji 2 in 3 je moč povezati pod pojmom semantično poizvedovanje, razkorak med stopnjama 1 in 2 pa lahko poimenujemo semantični prepad. Ta klasifikacija sicer ignorira nekatere tipe poizvedovanja po slikah, kot je npr. poizvedovanje po metapodatkih, ne ker bi bili ti podatki nepomembni, ampak ker so izključno tekstovne narave. Kombinacija CBIR in tekstovnih oznak lahko izboljša iskanje, vendar ne naslavlja problemov na stopnjah 2 in 3.

4 Uporabniški vmesniki

Sposobnost, da uporabniki enostavno in natančno izrazijo svojo iskalno zahtevo, je v vsakem poizvedovalnem sistemu ključnega pomena. Poizvedovanje po slikah ni izjema, ni pa jasno, kako to v praksi izvesti. Uporaba poizvedovalnih jezikov, kot je SQL in temu podobnih, je bila prisotna v nekaterih zgodnjih CBIR sistemih, a njihova uporaba se že takrat ni zdela optimalen način oblikovanja poizvedb po vizualni vsebini. Danes je v CBIR sistemih najbolj razširjena uporaba slike za primer.

Načini poizvedovanja se lahko med sabo razlikujejo tudi glede na posameznikov pristop k iskanju. Pri **ciljnem iskanju** uporabnik natančno ve, katero sliko išče, in da jo locira, mora samo pravilno podati sistemu ustrezne podatke. Če uporabnik **približno ve, kaj išče**, bo najbrž najlažje iskal znotraj skupine določenih slik. Če ima sliko za primer ali pa mu sistem v naključnem začetnem naboru slik ponudi sliko, ki približno ustreza, mu bo to dobro izhodišče za iskanje. Pri **splošnem brskljanju** je uporabniku zeleni rezultat nejasen ali celo neznan. Uporabnik išče v smislu »vedel bom, ko bom videl«. Uporabnik lahko pregleda veliko količino slik v zbirki, preden bo ugotovil, kaj je tisto, kar išče.

Izbira uporabniškega vmesnika je pomemben korak pri načrtovanju CBIR sistema, saj je uporabniški vmesnik neposredni stik uporabnika s sistemom. V današnjih CBIR sistemih poznamo naslednje oblike vmesniških interakcij med uporabnikom in sistemom.

Podajanje slike za primer

Gre za vrsto vmesniške interakcije, ki omogoča uporabniku relativno enostaven vnos. Slika za primer je vzorec, ki pokaže, kakšen rezultat je zaželen, in naproša sistem, naj najde nadaljnje podobne primere slik. Ta pristop sta že leta 1981 opisala Chang in Fu (povzeto po Eakins in Graham, 1999) s svojim QPE vmesnikom (angl. *query by pictorial example*). Treba je vedeti, da uporabniki nimajo vedno pri roki slike za primer, zato je zaželeno, da lahko uporabnik sliko za primer izbere iz več virov. Možna je izbira slike za primer iz lastnega računalnika, izbira z uporabo spletne povezave in izbira ene izmed slik iz zbirke same. Zadnja možnost je v praksi najpogostejša, sistem uporabniku ponudi naključen ali tipičen nabor slik iz zbirke in uporabnik s klikom miške na eno izmed slik prične poizvedbo.

Izbira (in obteževanje) značilnic

Pri takem vmesniku uporabnik sam izbere značilnice, po katerih poizveduje, in sam določi njihove zelene vrednosti. Gre za vrsto vmesnika, ki je naporen za uporabo, gledano s stališča povprečnega uporabnika. Naprednejša različica omogoča, da uporabnik obteži pomembnost posameznih izbranih značilnic, kar

pa ne olajšuje uporabe vmesnika, prej nasprotno. Seveda uporabnik obtežuje posamezne značilnice tudi pri drugih tipih vmesnikov, a le pri tem tipu to počne z vnašanjem številčnih vrednosti. Tak vmesnik lahko za zelo izkušenega uporabnika pomeni prednost, možnost natančnejšega podajanja poizvedbe. Različica tega tipa vmesnika je podajanje vrednosti v obliki histograma, kjer lahko namesto z vnašanjem številčk uporabnik poda poizvedbo z grafično ponazoritvijo npr. količine barve, kar pa je že zelo podobno vnašanju lastne skice ali slike za primer.

Podajanje slike za primer + izbira regij in značilnic za regije

Pri tem vmesniku gre za kombinacijo zgornjih dveh z dodatkom segmentacije. Uporabnik izbere sliko za primer in označi pomembne regije znotraj nje. Dobro je, če lahko uporabnik označi poljubne regije in če lahko obteži pomembnost posameznih značilnic v izbranih regijah. Tak tip vmesnika je redkejši kot iskanje z uporabo slike za primer, čeprav da boljše rezultate. Podobne rezultate lahko dobimo v sistemu, kjer je izhodišče slika za primer, nato pa izboljšamo rezultate z uporabo povratne zanke.

Kreacija lastne skice

Uporabnik sam, ponavadi s pomočjo miške, nariše skico, ki čim bolj nakazuje na iskano sliko. Sistem lahko pri procesiranju poizvedbe upošteva samo obris oz. obliko, pri večini vmesnikov pa lahko uporabnik nariše barvno skico in tako ponazori tudi zelene barve in po možnosti še njihov položaj na iskani sliki. Da so rezultati poizvedbe ustrezni skici, morajo biti barvne informacije o sliki povezane z njihovo lokacijo v skici. Takšni vmesniki ponujajo različne načine skiciranja. Uporabnik lahko izbira tanjše ali debelejšše črte, v skico lahko vstavi različne geometrijske oblike ipd. Varianta pristopa s skico so **slikovne ikone**, kjer uporabnik skice ne nariše sam, ampak jo ustvari z izbiranjem in kombiniranjem različnih preddefiniranih ikon.

4.1 Povratna zanka

Ker ne moremo realno pričakovati, da bo uporabnik v prvem iskanju našel ustrezne mere ujemanja, dober sistem omogoča interakcijo z uporabnikom, v kateri ta izrazi zadovoljstvo s predlaganimi rezultati poizvedbe. Iskanje s povratno zanko gre korak dlje od zgoraj naštetih tipov vmesnikov, saj odstrani breme ročnega uteževanja značilnic. Princip povratne zanke običajno poteka tako, da uporabnik razvršča slike, ki so rezultat poizvedbe, glede na njihovo relevantnost. Sistemu pove, katere slike so zanj relevantne in katere niso. Sistem na podlagi novih podatkov uporabniku ponudi izboljššan nabor zadetkov. Povratna zanka je lahko enkratna, ali pa se ponavlja v več zaporednih korakih. Nekateri sistemi omogočajo

uporabo povratne zanke že v samem začetku oblikovanja poizvedbe, pri nekaterih pa je potrebno najprej priklicati iz zbirke nabor zadetkov, nakar je omogočeno izboljševanje poizvedbe z uporabo povratne zanke.

4.2 Poklicne skupine kot tipični uporabniki slik

Z vsakdanjo uporabo slik se srečujejo številne poklicne skupine. Ker so te, vsaka zase, mnogo bolj homogene, kot to velja za priložnostne uporabnike slik, je možno ugotoviti, na kakšen način so slike uporabljene v nekaterih poklicih in ali to zajema tudi uporabo CBIR sistemov. Tipična področja s ključno uporabo slik so: kriminalistika (zbirke prstnih odtisov, slike prizorišč zločinov, zbirke obrazov, DNK verig, odtisov podplatov čevljev ali avtomobilskih gum in zbirke ukradenih predmetov), medicina (velike zbirke rentgenskih posnetkov), založništvo (ilustracije, spremne fotografije k člankom v časopisju) ter zgodovina in arheologija (nadomestki originala, slike izkopanin).

Uporaba slik je pogosta in pomembna tudi na področju bibliotekarstva in njemu sorodnih ved. Že pri vseh zgoraj navedenih področjih se lahko z uporabo slik sooči bibliotekar v specialni knjižnici. Na področju splošnih knjižnic pa se od knjižničarja pričakuje odlično poznavanje sistemov za poizvedovanje in tudi poznavanje CBIR sistemov in njihove uporabe ne bi smelo biti izjema. Poleg tega je znana ideja, da bi se s pomočjo CBIR izgradili filtri, ki bi zaznavali prisotnost pornografskega materiala in onemogočali dostop do njega. To je pomembno področje možnega povezovanja CBIR funkcionalnosti in knjižnic, saj knjižnice svojim uporabnikom, tudi mladoletnim, nudijo prost dostop do interneta. Na področjih sorodnih ved se ne moremo izogniti asociaciji na muzeje in arhive, kjer so se tudi že izvajali projekti, povezani s CBIR funkcionalnostjo. Tako muzeji kot arhivi se soočajo z gradivom, ki se ga pogosto obdeluje izključno v digitalnem okolju, saj je obdelava originala izključena zaradi namena njegovega ohranjanja.

5 Primerjava prostodostopnih sistemov na Internetu

V nadaljevanju primerjamo večino CBIR sistemov prostodostopnih na internetu. Za širši opis priporočam ogled diplomskega dela (Kranjc, 2006). Rezultate raziskave predstavljamo v dveh velikih primerjalnih tabelah. V Tabeli 1 so združene naslednje karakteristike:

- seznam sistemov, primerjava razvojnih okolij in držav, v katerih so bili razviti,
- primerjava značilnic, ki jih uporabljajo sistemi, kateri sistemi uporabljajo katere značilnice in koliko sistemov uporablja iste oz. različne tipe značilnic,
- primerjava načinov predstavitev rezultatov, ali je podobnost številčno opredeljena.

Tabela 1: Primerjava sistemov po okolju, značilnicah in prikazu rezultatov.

Sistem	Okolje	Država	ENOSTAVNE ZNAČILNICE			Rezultati - številčno ponazorjeni
			Barva	Tekstura	Oblika	
CIRES	akademsko	ZDA	DA	DA	NE	NE
COMPASS	inštitut	Italija	DA	DA	NE	NE
CORTINA	akademsko	ZDA	DA	DA	NE	NE
FIDS	akademsko	ZDA	DA	DA	NE	NE
FIRE	akademsko	Nemčija	DA	DA	DA	NE
GIFT (VIPER)	akademsko	Švica	DA	DA	NE	DA
IMAGESCAPE	akademsko	Nizozemska	DA	DA	DA	NE
IMAGE-SEEKER	komercialno	Francija	DA	NE	DA	NE
IMEDIA	inštitut	Francija	DA	DA	DA	NE
IMGSEEK	samostojno	Brazilija	DA	NE	NE	DA
IRMA	inštitut + akademsko	Nemčija	DA	DA	DA	DA
LCPD	inštitut + akademsko	Nizozemska	NE	DA	NE	NE
PIC2SEEK	akademsko	Nizozemska	DA	NE	DA	NE
PICSOM	akademsko	Finska	DA	DA	DA	NE
PRINCETON 3D MODEL	akademsko	ZDA	NE	NE	DA	NE
QBIC (Hermitage)	komercialno	ZDA	DA	DA	DA	NE
QUICKLOOK	akademsko	Italija	DA	DA	DA	NE
RETRIEVR	komercialno	Avstrija	DA	NE	NE	NE
SHAPE SIMILARITY PROJECT	akademsko	ZDA	DA	DA	DA	NE
SIMPLICITY	akademsko	ZDA	DA	DA	DA	DA
SQUID	akademsko	VB	NE	NE	DA	DA
TILTOMO	komercialno	VB	DA	DA	NE	NE
WEBSEEK	akademsko	ZDA	DA	NE	NE	NE

5.1 Razvojna okolja

Večina sistemov (skoraj 70 %) je plod znanstvenih raziskav in izvirajo iz akademskega okolja. Skoraj vsi ti sistemi so nastali v okviru računalniških oddelkov univerz, nekateri so nastali v specializiranih centrih pod okriljem univerz, nekateri pa v povezavi z zunanjimi inštituti (2 sistema). Avtorji variirajo od študentov, ki pripravljajo diplomsko ali magistrsko nalogo in v okviru tega ustvarijo CBIR sistem, do priznanih znanstvenikov, ki okoli sebe zberejo ugledno raziskovalno skupino.

Manj sistemov nastane v drugih raziskovalnih ustanovah, kot so inštituti (4 sistemi). Komercialnih sistemov je bistveno več, toda niso prosto dostopni, ampak so zaprti sistemi. V raziskavo so tako vključeni štirje komercialni sistemi, od katerih je en demonstracijski sistem (Image-seeker) z nepopolno CBIR funkcionalnostjo, eden je sistem, ki ni več aktualen, pa še vedno živi v zbirki muzeja Hermitage, dva pa sta eksperimentalna sistema (Tiltomo in Retrievr), ki za svoj razvoj uporabljata spletno skladišče slik Flickr, in sta v tej eksperimentalni različici prosto dostopna, saj je njuna uporaba v interesu avtorjev.

Edini sistem, ki ne prihaja iz ZDA ali Evrope, je imgSeek, ki prihaja iz Brazilije, a je razvit na osnovi dela (algoritma) ameriških znanstvenikov. Sicer je izvor sistemov po državah tak, da 8 sistemov (35 %) prihaja iz ZDA, 14 sistemov (60 %) pa iz različnih evropskih držav, pri čemer ni nobena država zastopana več kot z dvema sistemoma. Vsi sistemi, ki prihajajo iz ZDA so nastali v akademskem okolju, kar je posledica zgodnje raziskovalne dejavnosti na tem področju, ki se je odvijala prav tam. Številni sistemi, ki so nastali v devetdesetih, niso več aktualni in niso vključeni v raziskavo, prihajajo pa v veliki večini prav tako iz ZDA.

5.2 Poizvedovanje v sistemih

Enostavne značilnice, ki jih uporabljajo posamezni sistemi, so razvidne iz Tabele 1. Potrjuje se predpostavka, da je barva najpogosteje uporabljena značilnica, saj značilnic barve ne uporabljajo le trije (specializirani) sistemi. Edini sistem, ki uporablja barvo kot edino značilnico, je WeebSEEk, imgSeek in retrievr uporabljata poleg barve še *robove*, kot kompleksno značilnico. Podobno število sistemov uporablja značilnice teksture in oblike, kar je zanimivo, saj so značilnice oblike načeloma težje izračunljive in se včasih uvrščajo med kompleksnejše značilnice (Johansson, 2000). Sistem LCPD se v celoti zanaša na značilnice teksture.

Načini oblikovanja poizvedb so prikazani v Tabeli 2. Iz nje je razvidno, katere in koliko različnih možnosti zastavljanja poizvedb omogočajo uporabniški vmes-

niki posameznih sistemov. Poudarek je na preglednosti prikaza in dobri primerljivosti podatkov.

Tabela 2: Primerjava sistemov po načinih oblikovanja poizvedbe.

Sistem	Slika za primer			Ročno obteževanje značilnic	Slika za primer + izbira regije	Lastna skica	Povratna zanka
	Iz zbirke	Iz svetovnega spleta	Z osebne računalnika				
CIRES	DA	NE	NE	DA	NE	NE	DA
COMPASS	DA	NE	NE	DA	NE	NE	DA
CORTINA	DA	NE	NE	NE	NE	NE	NE
FIDS	DA	NE	NE	DA	NE	NE	NE
FIRE	DA	NE	DA	NE	NE	NE	DA
GIFT (VIPER)	NE	NE	NE	NE	NE	NE	DA
IMAGESCAPE	NE	NE	NE	NE	NE	DA	NE
IMAGE-SEEKER	DA	NE	NE	NE	NE	NE	NE
IMEDIA	DA	NE	DA	NE	DA	NE	DA
IMGSEEK	DA	DA	DA	NE	NE	DA	NE
IRMA	DA	NE	NE	NE	NE	NE	DA
LCPD	DA	NE	NE	NE	NE	NE	DA
PIC2SEEK	DA	DA	NE	NE	NE	NE	NE
PICSOM	DA	NE	NE	DA	NE	NE	DA
PRINCETON 3D MODEL	NE	NE	NE	NE	NE	DA	NE
QBIC (Hermitage)	DA	NE	NE	DA	NE	DA	NE
QUICKLOOK	NE	NE	NE	NE	DA	NE	DA
RETRIEVR	NE	DA	DA	NE	NE	DA	NE
SHAPE SIMILARITY PROJECT	DA	NE	NE	NE	NE	DA	NE
SIMPLICITY	DA	DA	NE	NE	NE	NE	NE
SQUID	DA	NE	NE	NE	NE	NE	NE
TILTOMO	DA	NE	NE	NE	NE	NE	NE
WEBSEEK	DA	NE	NE	DA	NE	NE	DA

Rezultati potrjujejo domnevo, da je osnovni način podajanja vnosa pri CBIR poizvedovanju slika za primer. Tak način podpira kar 87 % sistemov. Pri tem je daleč najbolj pogosto podajanje slike za primer iz obstoječe zbirke, ki jo uporablja sistem (omogoča 78 % sistemov), redkejše pa podajanje slike iz osebnega računalnika (omogoča 17 % sistemov) in s podajanjem URL-ja na povezavo s svetovnega spleta (omogoča 13 % sistemov).

Manj sistemov omogoča druge načine oblikovanja poizvedb. Sedem sistemov (30 %) omogoča uporabniku vnos lastne skice, pri čemer en sistem omogoča

uporabo preddefiniranih ikon (ImageScape). Šest sistemov omogoča uporabniku, da ročno obteži posamezne značilnice, pogosto se to zgodi v kombinaciji z uporabo slike za primer. Najmanj sistemov (dva) omogoča iskanje s sliko za primer in izbiranjem regije v sliki. QuickLook omogoča tovrsten način, toda uporabnik mora predhodno v nastavitvah sistema določiti, kateri segmenti slike bodo upoštevani.

En sistem (Viper) ne omogoča nobenega od zgoraj naštetih načinov. V njem je možno poizvedovanje zgolj z uporabo principa povratne zanke. Ta princip sicer uporablja 10 sistemov, kar je manj kot polovica vseh sistemov. To je presenetljivo, saj povratna zanka po naši oceni izjemno pozitivno vpliva na skladnost uporabnikovih pričakovanj in rezultatov, ki jih na osnovi njegovega vnosa vrne sistem. Uporabnik namreč izrazi nezadovoljstvo nad rezultati oziroma zanj tipično vedenje, kakšni naj bodo rezultati, in na ta način se manjša semantični prepad.

Pri predstavitvi rezultatov ugotavljamo, da so v vseh sistemih rezultati razvrščeni po podobnosti. Iz Tabele 1 je razvidno, kateri sistemi ob tem podajo tudi številčne vrednosti za podobnost. Takih sistemov je le pet oz. 22 %. Po naši oceni gre za informacijo, ki je lahko za uporabnika zelo uporabna, saj mu omogoča boljše razumevanje rezultatov. Rangiranje, ko v zbirki ni podobnih slik, lahko uporabnika pusti v misli, da je sistem slab, številčno rangiranje pa mu omogoča ustreznejše vrednotenje rezultatov.

5.3 Uporaba sistemov

Posamezni sistemi se zelo razlikujejo po načinu uporabe. Večina sistemov obstaja samo v spletni demonstraciji, bodisi da jih avtorji razvijajo z možnostjo kasnejše uporabe (npr. Tiltomo) bodisi so že deloma zastareli (spletna demonstracija pa še vedno deluje in pogosto je še vedno možno kontaktirati avtorje). Nekateri sistemi so odprtokodnega značaja, in jih lahko v svojih aplikacijah uporablja kdor hoče (npr. imgSeek).

Nekateri sistemi so presegle demonstracijski namen, in se uporabljajo v širšem kontekstu, npr. v organizacijah, kot so inštituti in podjetja. Bodisi so to sistemi, ki so bili razviti z namenom konkretne uporabe in si gresta razvoj in uporaba z roko v roki (npr. IRMA), bodisi so to sistemi, ki so izrazito komercialne narave (npr. Image-seeker), bodisi so to sistemi, kot je SIMPLcity, kjer je bil prisoten močan interes razvijalcev za različne oblike konkretne uporabe.

Glede na sam namen uporabe gre izpostaviti še eno vrsto sistemov. Gre za sisteme, katerih osnovni namen je soočanje z veliko količino slik, ki so prisotne na internetu. Tovrstni pionirski sistem je bil WebSEEk, danes pa iskanje v več kot

10.000.000 slik s spleta omogoča Cortina. Tovrstni sistemi avtomatsko indeksirajo velike količine slik s svetovnega spleta po hibridnem principu, pri čemer spretno uporabljajo strukturo spletnih dokumentov.

5.4 Omejitve

Sistemi so lahko omejeni zaradi različnih razlogov:

- neuporabe povratne zanke (npr. SIMPLcity),
- prepočasnega delovanja (npr. IRMA),
- premajhne ali premalo heterogene zbirke slik (razen če je taka s posebnim namenom),
- nezadovoljivega uporabniškega vmesnika (npr. LCPD),
- pretirane zapletenosti (npr. COMPASS),
- zastarelosti (npr. ImageScape),
- preslabih navodil za uporabo (npr. Pic2Seek),
- neskladja med uporabnikovimi pričakovanji in dejanskimi rezultati, ki pa ni nujno odraz slabega algoritma, ampak je lahko v eksperimentalnih sistemih pogosto odraz premajhne zbirke (že navedeno v tretji alineji) ali pa posledica slabih mer.

Uporabniški vmesniki pogosto delujejo v javanskem okolju, kar v veliko primerih pomeni tudi počasno delovanje. Sistemi, kot je Viper, kažejo smernice za razvoj hitrejših vmesnikov, napisanih v skriptnem jeziku PHP, in s tem bolj prilagojenih podajanju spletnih vsebin.

5.5 Trendi

Znanstveniki, ki raziskujejo na področju računalniškega vida, so se v veliki meri že preusmerili stran od sistemov, ki delujejo na principih enostavnih značilnic, in iščejo naprednejše rešitve. Nadaljuje pa se tudi razvoj tovrstnih sistemov. Ta razvoj poteka v dve smeri. Na eni strani so sistemi, ki se razvijajo za specifičen namen uporabe v homogenih zbirkah, na tipičnih področjih, kot sta medicina in kriminalistika. Na drugi strani gre razvoj v smer razvoja hibridnih sistemov, v katerih se avtomatsko združita vsebina in tekstovni opis slik.

Trend razvoja slednjih je še posebej aktualen, saj se ta razvoj odvija zdaj, zanimiv pa je tudi za našo stroko, saj se navezuje na indeksiranje spletnih vsebin.

Retriever in Tiltomo sta močno povečala zanimanje za uporabo CBIR sistemov med splošnimi uporabniki slik. Če se zgodi, da Flickr vgradi CBIR sistem v svoj spletni servis, kar je povsem verjetno, bomo doživeli veliko povečanje števila zainteresiranih uporabnikov in v tržnem okolju interneta bo to vodilo v nove raziskave in predvsem v nove načine uporabe, predvsem imamo v mislih velike spletne iskalnike, kot je Google.

Tak razvoj lahko vodi tudi v še bolj intenzivno znanstvenoraziskovalno dejavnost na področju CBIR. V vsakem primeru pa bodo vidiki konkretne uporabe sistemov vedno bolj zanimivi za našo stroko, kot teorija, ki je za algoritmi njihovega delovanja.

6 Zaključek

CBIR je princip, ki je in bo ostal soudeležen pri soočanju z velikimi zbirkami slik. Če ima danes pogosto še stransko vlogo, je iz raziskovalne vneme, ki ga obdaja, jasno razvidna naraščajoča potreba po tovrstnem pristopu. Principi CBIR so že dolgo trdno umeščeni v področja, kot je biomedicinska informatika, a glavne omejitve CBIR sistemov se navezujejo na uporabo v velikih heterogenih zbirkah slik. Ker CBIR sistemi še vedno niso zmožni uspešno zajeti logičnih konceptov in semantičnih pomenov iz slik, so z uporabniškega stališča pogosto bolj ali manj neuporabni. Znanstvena skupnost vidi rešitev predvsem v iskanju naprednejših rešitev, določen segment raziskovalcev (med njimi tudi akademski) pa se osredotoča na razvoj hibridnih sistemov, kjer se tekstovni opisi, tako kot vsebina, avtomatsko pridobijo iz strukture spletnih dokumentov in ovrednotijo skozi algoritem.

Poznavanje problematike CBIR je pomembno za bibliotekarje, saj se kot informacijski strokovnjaki umeščamo v sisteme pretoka informacij, tudi slikovnih. Poleg splošnega poznavanja sistemov vidimo še nekaj razlogov, zaradi katerih bi morali poznati CBIR problematiko. En je uporabniški vidik, povsem možno je, da bo v bližnji prihodnosti interakcija z uporabniki zahtevala od nas osnovno poznavanje CBIR sistemov. Drug je vidik informacijske pismenosti, ki zajema tudi zmožnost uporabe različnih sistemov za poizvedovanje, tretji je vidik proste dostopnosti informacij, ki zajema tudi omogočanje spletnih vsebin in v povezavi s tem tudi omogočanje uporabe prosto dostopnih CBIR sistemov. Pomemben vidik pa je tudi bodoča integracija CBIR sistema v digitalno knjižnico, ki bo sicer verjetno potekala v izvedbi računalniškega strokovnjaka, a zagotovo tudi ob sodelovanju informacijskega strokovnjaka. V vsakem primeru je pomembno, da smo bibliotekarji v koraku s časom in da imamo ustrezna znanja za delo s kakršnikoli sistemi za poizvedovanje.

Viri

1. BIRD, C. L., ELLIOTT, P. J., HAYWARD, P. M. (1999). Content-based retrieval for European image libraries. [online]. V *Challenge of image retrieval, Newcastle upon Tyne, UK, 25-26 Februar 2006*. Pridobljeno 6. 10. 2005 s spletne strani <http://ewic.bcs.org/conferences/1999/imageret/papers/paper2.pdf>
2. CAWKELL, T. (2000). Image indexing and retrieval by content. *Information Services and Use*, vol. 20, no. 1, str. 49-58.
3. CHANG, S-F., CHEN W., SUNDARAM, H. (1998). Semantic visual templates: linking visual features to semantics. [online]. V *International Conference on Image Processing (ICIP 98)*, 4-7 Oct 1998, str. 531-535. Pridobljeno 6. 5. 2005 s spletne strani <http://ame2.asu.edu/faculty/hs/pubs/icip98.pdf>
4. CHANG, S-F., SMITH, J. R., BEIGI, M., BENITEZ, A. (1997). Visual information retrieval from large distributed online repositories. *Communications of the ACM*, vol. 40, no. 12, str. 63-71.
5. *Content-Based Image Retrieval (CBIR) of Biomedical Images: a report to the Board of Scientific Counselors*. (2002). [online]. Communications Engineering Branch, Lister Hill National Center for Biomedical Communications, National Library of Medicine. Pridobljeno 16. 8. 2005 s spletne strani <http://archive.nlm.nih.gov/pubs/reports/bosc02/>
6. DEL BIMBO, A. (1999). *Visual information retrieval*. San Francisco: Morgan Kaufmann Pub.
7. EAKINS, J. P. (1996). *Automatic image content retrieval: are we getting anywhere?* [online]. Newcastle: Department of computing. University of Northumbria. Pridobljeno 24. 4. 2005 s spletne strani <http://www.unn.ac.uk/iidr/papers/elvira3.ps.Z>
8. EAKINS, J. P., GRAHAM, M. E. (1999). *Content-based image retrieval: a report to the JISC technology applications programme*. [online]. Newcastle: Institute for image data research. University of Northumbria. Pridobljeno 27. 4. 2005 s spletne strani <http://www.unn.ac.uk/iidr/CBIR/report.html>
9. FINN, R. (1996). *Querying by image content*. [online]. Pridobljeno 8. 6. 2005 s spletne strani http://domino.research.ibm.com/comm/wwwr_thinkresearch.nsf/pages/image396.html
10. FORSYTH, D. A. (1999). Computer vision tools for finding images and video sequences. *Library trends*, vol. 48, no. 2, str. 326-355.
11. GERADTS, Z. (2002). *Content-based information retrieval from forensic image databases*. [online]. Utrecht: Proefschrift Universiteit. Pridobljeno 21. 6. 2005 s spletne strani <http://forensic.to/Dissertation.pdf>
12. GEVERS, T., SMEULDERS, A. (2004). Image search engine: an overview. V *Emerging topics in computer vision*. Upper Sadle River: Prentice Hall.

13. GROSKY, W. I., STANCHEV, P. L. (2000). An image data model. V *Proceedings of the fourth international conference on advances in visual information systems, lecture notes in computer science*. [online]. London: Springer-Verlag. Pridobljeno 8. 5. 2005 s spletne strani <http://www.kettering.edu/~pstanche/Visua2000.pdf>
14. GUPTA A., SANTINI, S, JAIN R. (1997). In search of information in visual media. *Communcations of the ACM*, vol. 40, no. 12, str. 35-42.
15. HOLT, B., WEISS, K., NIBLACK, W., FLICKNER, M., PETKOVIC, D. (1997). The QBIC project in the department of art and art history at UC Davis. [online]. V *Proceedings of the sixtieth annual meeting of American Society for Information Science, November 1-6, 1997, Washington DC*. Pridobljeno 5. 5. 2005 s spletne strani <http://www.asis.org/annual-97/holt.htm>
16. HOVE, L-J. (2004). *Extending image retrieval systems with thesaurus for shapes: master thesis*. [online]. Bergen: Institute for information and media science. University of Bergen. Pridobljeno 17. 7. 2005 s spletne strani <http://www.nik.no/2004/bidrag/Hove.pdf>
17. JACOBS, C. E., FINKELSTEIN., A., SALESIN, D. (1995). Fast multiresolution image querying. [online]. V *Proceedings of SIGGRAPH 95*. Pridobljeno 20. 6. 2005 s spletne strani <http://grail.cs.washington.edu/projects/query/mrquery.pdf>
18. JAIN, R. (1997). Visual information management. *Communcations of the ACM*, vol. 40, no. 12, str. 31-32.
19. JOHANSSON, B. (2000). *A survey on: content based image search in image databases*. [online]. Pridobljeno 2005-04-16]. S spletne strani <http://www.cvl.isy.liu.se/ScOut/TechRep/Papers/LiTHISYR2215.pdf>
20. KONAK, E. S. (2002). *A content-based image retrieval system for texture and color queries: a thesis*. [online]. Bilkent: Department of computer engineering, Bilkent university. [Pridobljeno 2005-06-10]. S spletne strani <http://www.cs.bilkent.edu.tr/tech-reports/2002/BU-CE-0212.pdf>
21. KRANJIC, T. (2006). *Poizvedovanje po slikah na podlagi vsebine: s pregledom obstoječih prosto dostopnih sistemov: diplomsko delo*. Ljubljana: [samozal T. Kranjc].
22. LEHMANN et al. (2004). Content-based image retrieval in medical applications. [online]. *Methods Inf Med*, vol. 43, str. 354-361. Pridobljeno 21. 3. 2006 s spletne strani: [http://phobos.imib.rwth-aachen.de/irma/ps-pdf/MIM_2004-43\(4\)354-361.pdf](http://phobos.imib.rwth-aachen.de/irma/ps-pdf/MIM_2004-43(4)354-361.pdf)
23. LU, G., WILLIAMS, B., YOU, C. (2001). An effective World Wide Web image search engine. *Journal of Information Science*, vol. 27, no. 1, str. 27-37.
24. MÜLLER, H., MICHOUX, N., BANDON D., GEISSBUHLER, A. (2004). A review of content-based image retrieval systems in medical applications: clinical benefits and future directions. *International Journal of Medical Informatics*, vol. 73, str. 1-23.

25. QUACK, T., MÖNOCH, U., THIELE, L., MANJUNATH, B.S. (2004). Cortina: a system for large-scale, content-based web image retrieval. [online]. V *Proceedings of the 12th annual ACM international conference on Multimedia*. Pridobljeno 17. 3. 2006 s spletne strani: <http://www.vision.ee.ethz.ch/~tquack/quack04.pdf>
26. RUI, Y, HUANG, T. S., CHANG, S-F. (1999). Image retrieval: current techniques, promising directions and open issues. [online]. *Journal of Visual Communication and Image Representation*, vol. 10, str. 39-62. Pridobljeno 18. 5. 2005 s spletne strani: http://www.cs.princeton.edu/courses/archive/spr05/cos598E/bib/rui99_cbir_survey.pdf
27. STANCHEV, P. L. (2001). Content-based image retrieval systems. [online]. *CompSysTech, Bulgarian computer science conference, 21-22.06.2001, Sofia*. Pridobljeno 21. 6. 2005 s spletne strani <http://www.kettering.edu/~pstanche/comosys2001.pdf>
28. STANCHEV, P. L., GREEN, D. (2002). Current state and research trend in the image database systems. [online]. *Mathematics and education in mathematics*. Pridobljeno 6. 8. 2005 s spletne strani <http://www.kettering.edu/~pstanche/borovez.pdf>
29. STICKER, M., DIMAI, A. (1996). Color indexing with weak spatial constraints. V *Symposium on Electronic Imaging: Science and Technology - Storage & Retrieval for Image and Video Databases IV*, str. 29-41.
30. SWAIN, M. J., BALLARD, D. H. (1991). Color indexing [online]. *International Journal of Computer Vision*, vol. 7, no. 1, str. 11-32. Pridobljeno 4. 10. 2005 s spletne strani: <http://www.stanford.edu/~simonb/papers/Swain%20Ballard%20Color%20Indexing%20n231141541p1211g.pdf>
31. VASCONCELOS, N., KUNT, M. (2001). Content-based retrieval from image databases: current solutions and future directions. [online]. V *IEEE International Conference on Image Processing, Thessaloniki, Greece, October 2001*. Pridobljeno 11. 8. 2005 s spletne strani http://www.svcl.ucsd.edu/~nuno/ICIP01/Vasconcelos_Kunt.pdf
32. VELTKAMP, R. C., TANASE, M. (2002). *Content-based image retrieval systems: a survey*. [online]. Pridobljeno 27. 4. 2005 s spletne strani <http://give-lab.cs.uu.nl/cbirsurvey/cbir-survey.pdf>
33. WANG, J. Z., LI, J., WIEDERHOLD, G. (2001). SIMPLicity: Semantics-Sensitive Integrated Machine for Picture Libraries [online]. V *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 23, no. 9, str. 947-963. Pridobljeno 24. 6. 2005 s spletne strani <http://www-db.stanford.edu/~wangz/project/imsearch/SIMPLicity/TPAMI/wang2.pdf>
34. WEES, M. (2002). *How to find images of art on the internet?* Zaandijk: [samozal. M. van Wees].

35. YANAI, K., SHINDO M., NOSHITA, K. (2004). A fast image-gathering system from World-Wide-Web using a PC cluster [online]. *Image and vision computing*, vol. 22, str. 59-71. Pridobljeno 22. 3. 2006 s spletne strani <http://skynet.liacs.nl/downloads/MIR.RL1/Yanai.Shindo.Noshita.IVC04.pdf>
-

Tomaž Kranjc je diplomant Oddelka za bibliotekarstvo, informacijsko znanost in knjigarstvo na Filozofski fakulteti Univerze v Ljubljani.

Naslov: Kuzeletova 8, 1000 Ljubljana

Naslov elektronske pošte: tkranjc@gmail.com

Izr. prof. dr. Jasna Maver je predavateljica na Oddelku za bibliotekarstvo, informacijsko znanost in knjigarstvo na Filozofski fakulteti Univerze v Ljubljani.

Naslov: Pleteršnikova 26, 1000 Ljubljana

Naslov elektronske pošte: jasna.maver@ff.uni-lj.si