

Pretraživanje multimedijalnih sadržaja

Dragan Ivanović
dragan.ivanovic@uns.ac.rs

Katedra za informatiku, Fakultet tehničkih nauka, Novi Sad

2015.

Vrste IR po sadržajima u kolekciji

- Pretraga tekstualnih sadržaja
- Pretraga linkovanih tekstualnih sadržaja (pretraga veba)
- Pretraga multimedijalnih sadržaja
 - slika
 - zvuk
 - video

Mediji i multimediji

- Informacije mogu biti prenete u obliku teksta, slike, veb stranice, videa ili zvuka
- Svaki medij ima svoje karakteristike, prednosti i slabosti
- Odabir najprikladnijeg medija za neku svrhu je težak zadatak
- Mediji se mogu kombinovati u multimediju
- Digitalni mediji i multimediji se mogu obrađivati, a i konzumirati upotrebom odgovarajućeg hardvera i softvera
- Postoji i hipermedija - linkovani multimedijalni sadržaji (veb)

Podele medija

- Prema zavisnosti od vremena
 - Vremenski zavisni: video, animacija, zvuk
 - Statični: tekst, slika
- Prema načinu konzumiranja medija
 - linearni
 - nelinearni

Multimedija IR - motivacija

- Sve je više hardvera koji može da kreira multimedijalni sadržaj
- Sve je više softvera koji može da kreira multimedijalni sadržaj
- Sve je više multimedijalnih sadržaja
- Sve je više kolekcija multimedijalnih sadržaja i na vebu i van njega
- Teško je pronaći, odabrati, filterisati multimedijalne sadržaje
- Potrebne su nam informacije o tome šta je u multimedijalnom sadržaju

Multimedija IR - problemi

- Sirova reprezentacija podataka, bez metapodataka
- Način kreiranja upita - kako opisati šta nam treba
- Dosta redudanse - iste slike u različitoj rezoluciji i slično

Multimedija IR - cilj

- Želimo da napravimo multimedijalne sadržaje pretraživim kao što su tekstualni sadržaji
- U današnjem svetu prepunom digitalnih podataka vrednost sadržaja ne zavisi samo od kvaliteta sadržaja nego i od toga koliko ga je lako pronaći
- Potreban nam je efikasan način opisa multimedijalnih sadržaja i kao način kreiranja upita i pretraživanja ovih sadržaja

Komponente MMIR modela

- Upitni jezik
- Indeksiranje i pretraživanje

Upitni jezik

- Kako izraziti upit nad kolekcijom multimedijalnih sadržaja?
 - rečima: specifičan upitni jezik - oslanjamo se na metapodatke o medijima
 - uzorkom: tražimo medije slične datom uzorku
 - šta je kriterijum sličnosti uzorka i medija u kolekciji?
- Da li se mogu definisati neki dodatni uslovi: minimalna rezolucija tražene slike, godina objavljivanja, daj mi sve multimedije koji imaju i video i zvuk, itd.
- Da li se mogu definisati neki složeniji uslovi: daj mi sve slike koje imaju crvenu kuću
- Naravno trebala bi da postoji i mogućnost pregledanja kolekcije i navigacije

Indeksiranje i pretraživanje

- Kakav indeks koristiti za pretraživanje medija?
- Sekvencijalna pretraga nije dovoljno efikasna
 - izračunavanje sličnosti vektora u visoko-dimenzionalnom prostoru, na primer 4000 dimenzija
 - za svaku od npr. 50000 medija
 - $50000 \times 4000 =$ previše

Indeksiranje i pretraživanje

- Kakav indeks koristiti za pretraživanje medija?
- Sekvencijalna pretraga nije dovoljno efikasna
 - izračunavanje sličnosti vektora u visoko-dimenzionalnom prostoru, na primer 4000 dimenzija
 - za svaku od npr. 50000 medija
 - $50000 \times 4000 =$ previše
- Plan A: neka od struktura za višedimenzionalne vektorske prostore
 - grid file, k-d tree, quad-tree, K-D-B tree, hB-tree, R-tree
- Plan B: svođenje problema na prethodni (pretraživanje teksta)
- Plan C: specijalizovani indeksi za pojedine osobine medija

Nivoi osobina medija

- Osobine visokog nivoa - *high-level features*: sadržaj na slici, video, prepoznavanje govora, itd.

Nivoi osobina medija

- Osobine visokog nivoa - *high-level features*: sadržaj na slici, video, prepoznavanje govora, itd.
- Osobine srednjeg nivoa - *medium-level features*: detektor lica, klasifikacija regiona, žanr muzike, itd.

Nivoi osobina medija

- Osobine visokog nivoa - *high-level features*: sadržaj na slici, video, prepoznavanje govora, itd.
- Osobine srednjeg nivoa - *medium-level features*: detektor lica, klasifikacija regiona, žanr muzike, itd.
- Osobine niskog nivoa - *low-level features*: Furijeova transformacija, *texture histograms*, *colour histograms*, *shape primitives*, itd.

Multimedija IR - pristupi

- *Text based* - pretraga multimedijalnih sadržaja bazirana na tekstu kojim je opisana
- *Content based* - pretraga multimedijalnih sadržaja bazirana na sadržaju

Text based multimedija IR

- Kako izvući informacije o sadržaju medija?

Text based multimedija IR

- Kako izvući informacije o sadržaju medija?
- Angažovati čoveka

Text based multimedija IR

- Kako izvući informacije o sadržaju medija?
- Angažovati čoveka
 - opisivanje sadržaja = formiranje metapodataka

Text based multimedija IR

- Kako izvući informacije o sadržaju medija?
- Angažovati čoveka
 - opisivanje sadržaja = formiranje metapodataka
 - opis predstavlja karakteristike visokog nivoa (high-level features)

Text based multimedija IR

- Kako izvući informacije o sadržaju medija?
- Angažovati čoveka
 - opisivanje sadržaja = formiranje metapodataka
 - opis predstavlja karakteristike visokog nivoa (**high-level features**)
 - opis koristi alfanumeričke tipove podataka: ključne reči, datume, imena, koordinate

Text based multimedija IR

- Kako izvući informacije o sadržaju medija?
- Angažovati čoveka
 - opisivanje sadržaja = formiranje metapodataka
 - opis predstavlja karakteristike visokog nivoa (**high-level features**)
 - opis koristi alfanumeričke tipove podataka: ključne reči, datume, imena, koordinate
 - za pretraživanje se koristi klasična baza podataka ili tehnike za pretraživanje teksta

Text based multimedija IR

- Kako izvući informacije o sadržaju medija?
- Angažovati čoveka
 - opisivanje sadržaja = formiranje metapodataka
 - opis predstavlja karakteristike visokog nivoa (**high-level features**)
 - opis koristi alfanumeričke tipove podataka: ključne reči, datume, imena, koordinate
 - za pretraživanje se koristi klasična baza podataka ili tehnike za pretraživanje teksta
 - skupo i dugotrajno formiranje metapodataka

Text based multimedija IR

- Kako izvući informacije o sadržaju medija?
- Angažovati čoveka
 - opisivanje sadržaja = formiranje metapodataka
 - opis predstavlja karakteristike visokog nivoa (**high-level features**)
 - opis koristi alfanumeričke tipove podataka: ključne reči, datume, imena, koordinate
 - za pretraživanje se koristi klasična baza podataka ili tehnike za pretraživanje teksta
 - skupo i dugotrajno formiranje metapodataka
 - opis zavisi od konteksta onoga ko opisuje; to može da bude različito od konteksta onoga ko pretražuje

Text based multimedija IR

- Kako izvući informacije o sadržaju medija?
- Angažovati čoveka
 - opisivanje sadržaja = formiranje metapodataka
 - opis predstavlja karakteristike visokog nivoa (**high-level features**)
 - opis koristi alfanumeričke tipove podataka: ključne reči, datume, imena, koordinate
 - za pretraživanje se koristi klasična baza podataka ili tehnike za pretraživanje teksta
 - skupo i dugotrajno formiranje metapodataka
 - opis zavisi od konteksta onoga ko opisuje; to može da bude različito od konteksta onoga ko pretražuje
- Problem se svodi na prethodni – pretraživanje teksta

Content based multimedija IR

- Kako izvući informacije o sadržaju medija?

Content based multimedija IR

- Kako izvući informacije o sadržaju medija?
- Angažovati računar

Content based multimedija IR

- Kako izvući informacije o sadržaju medija?
- Angažovati računar
 - opisivanje sadržaja = ekstrakcija osobina medija (slike, zvuka, videa)

Content based multimedija IR

- Kako izvući informacije o sadržaju medija?
- Angažovati računar
 - opisivanje sadržaja = ekstrakcija osobina medija (slike, zvuka, videa)
 - opis predstavlja osobine niskog nivoa (**low-level features**)

Content based multimedija IR

- Kako izvući informacije o sadržaju medija?
- Angažovati računar
 - opisivanje sadržaja = ekstrakcija osobina medija (slike, zvuka, videa)
 - opis predstavlja osobine niskog nivoa (**low-level features**)
 - ove osobine se **izračunavaju** na osnovu prostog sadržaja medija - na primer: piksela slike

Content based multimedija IR

- Kako izvući informacije o sadržaju medija?
- Angažovati računar
 - opisivanje sadržaja = ekstrakcija osobina medija (slike, zvuka, videa)
 - opis predstavlja osobine niskog nivoa (**low-level features**)
 - ove osobine se **izračunavaju** na osnovu prostog sadržaja medija - na primer: piksela slike
 - za pretraživanje se koristi specijalizovani indeks

Content based multimedija IR

- Kako izvući informacije o sadržaju medija?
- Angažovati računar
 - opisivanje sadržaja = ekstrakcija osobina medija (slike, zvuka, videa)
 - opis predstavlja osobine niskog nivoa (**low-level features**)
 - ove osobine se **izračunavaju** na osnovu prostog sadržaja medija - na primer: piksela slike
 - za pretraživanje se koristi specijalizovani indeks
 - brzo i jeftino kreiranje metapodataka

Content based multimedija IR

- Kako izvući informacije o sadržaju medija?
- Angažovati računar
 - opisivanje sadržaja = ekstrakcija osobina medija (slike, zvuka, videa)
 - opis predstavlja osobine niskog nivoa (**low-level features**)
 - ove osobine se **izračunavaju** na osnovu prostog sadržaja medija - na primer: piksela slike
 - za pretraživanje se koristi specijalizovani indeks
 - brzo i jeftino kreiranje metapodataka
 - manji kvalitet pretrage

Osnove

- Slike se prikazuju kao nizovi piksela, a reprezentovani su korišćenjem internog modela: kao bit-mape (bitmaps, rasterska grafika) ili vektorskom grafikom
- Generisanje piksela od modela zove se renderovanje
- Bit-mapa je niz logičkih piksela (skladištenih vrednosti boja) koji mogu biti direktno mapirani na fizičke piksele na ekranu: velike mogućnosti, dosta prostora zauzimaju, problemi u prepoznavanju oblika i transformaciji slike
- U vektorskoj grafici slike su skladištene kao matematički opis kolekcije linija, krivi i oblika (često u formi nekog XML-a): renderovanje složenije, prepoznavanje oblika i transformacije slike mnogo jednostavnije

Pretraživanje slika

- Novi Sad Priča - android aplikacija
- Google Images
- Raspolažemo velikom kolekcijom rasterskih slika
- Kako da je pretražujemo?

Pretraživanje slika

- Novi Sad Priča - android aplikacija
- Google Images
- Raspolažemo velikom kolekcijom rasterskih slika
- Kako da je pretražujemo?
- „Treba mi slika na kojoj je letnji pejzaž“

Pretraživanje slika

- Novi Sad Priča - android aplikacija
- Google Images
- Raspolažemo velikom kolekcijom rasterskih slika
- Kako da je pretražujemo?
- „Treba mi slika na kojoj je letnji pejzaž“



Semantički jaz

- Zašto je pretraživanje slika komplikovano?
- Šta je tema ove slike?
- Kojim ključnim rečima bismo je opisali?
- Koje reči bismo koristili u pretrazi?



Semantički jaz

- Problem 1: jedna slika vredi 1000 reči
- Problem 2: značenje (interpretacija sadržaja) slike je vrlo individualno i subjektivno

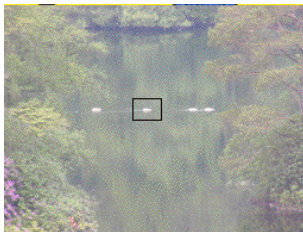
Semantički jaz

- Problem 1: jedna slika vredi 1000 reči
- Problem 2: značenje (interpretacija sadržaja) slike je vrlo individualno i subjektivno
- Po čemu su slične ove dve slike?



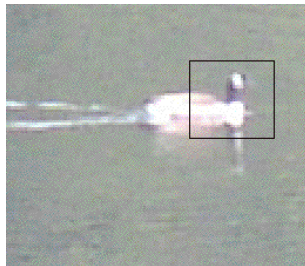
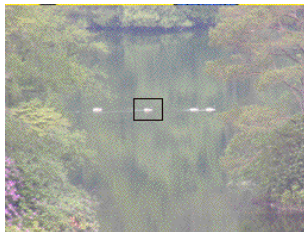
Konačna rezolucija slike

- Mi ovde vidimo patke, a računar?



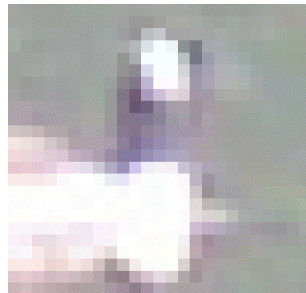
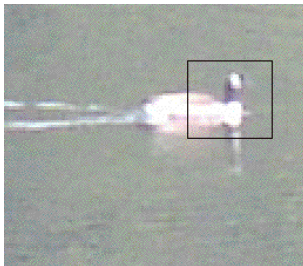
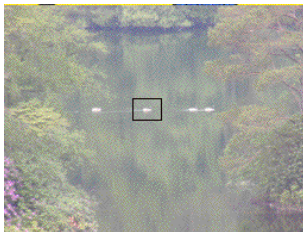
Konačna rezolucija slike

- Mi ovde vidimo patke, a računar?



Konačna rezolucija slike

- Mi ovde vidimo patke, a računar?



Kako izraziti upit nad kolekcijom slika?

- rečima: oslanjamo se na metapodatke o slikama
- slikom: tražimo slike slične datom uzorku
 - šta je kriterijum sličnosti uzorka i slike u kolekciji?

Kako izraziti upit nad kolekcijom slika?

- rečima: oslanjamo se na metapodatke o slikama
- slikom: tražimo slike slične datom uzorku
 - šta je kriterijum sličnosti uzorka i slike u kolekciji?

Upit:



Rezultat pretrage:



[15830](#) 0.00 5



[51111](#) 17.07 6



[22664](#) 17.72 6



[11222](#) 18.53 9



[44028](#) 18.58 5



[18659](#) 18.86 6



[44880](#) 18.97 6

Ekstrakcija sadržaja slike - pristup 1

- Kako izvući informacije o sadržaju slike?

Ekstrakcija sadržaja slike - pristup 1

- Kako izvući informacije o sadržaju slike?
- Angažovati čoveka

Ekstrakcija sadržaja slike - pristup 1

- Kako izvući informacije o sadržaju slike?
- Angažovati čoveka
 - opisivanje sadržaja = formiranje metapodataka

Ekstrakcija sadržaja slike - pristup 1

- Kako izvući informacije o sadržaju slike?
- Angažovati čoveka
 - opisivanje sadržaja = formiranje metapodataka
 - opis predstavlja karakteristike visokog nivoa (**high-level features**)

Ekstrakcija sadržaja slike - pristup 1

- Kako izvući informacije o sadržaju slike?
- Angažovati čoveka
 - opisivanje sadržaja = formiranje metapodataka
 - opis predstavlja karakteristike visokog nivoa (**high-level features**)
 - opis koristi alfanumeričke tipove podataka: ključne reči, datume, imena, koordinate

Ekstrakcija sadržaja slike - pristup 1

- Kako izvući informacije o sadržaju slike?
- Angažovati čoveka
 - opisivanje sadržaja = formiranje metapodataka
 - opis predstavlja karakteristike visokog nivoa (**high-level features**)
 - opis koristi alfanumeričke tipove podataka: ključne reči, datume, imena, koordinate
 - za pretraživanje se koristi klasična baza podataka ili tehnike za pretraživanje teksta

Ekstrakcija sadržaja slike - pristup 1

- Kako izvući informacije o sadržaju slike?
- Angažovati čoveka
 - opisivanje sadržaja = formiranje metapodataka
 - opis predstavlja karakteristike visokog nivoa (**high-level features**)
 - opis koristi alfanumeričke tipove podataka: ključne reči, datume, imena, koordinate
 - za pretraživanje se koristi klasična baza podataka ili tehnike za pretraživanje teksta
 - skupo i dugotrajno formiranje metapodataka

Ekstrakcija sadržaja slike - pristup 1

- Kako izvući informacije o sadržaju slike?
- Angažovati čoveka
 - opisivanje sadržaja = formiranje metapodataka
 - opis predstavlja karakteristike visokog nivoa (**high-level features**)
 - opis koristi alfanumeričke tipove podataka: ključne reči, datume, imena, koordinate
 - za pretraživanje se koristi klasična baza podataka ili tehnike za pretraživanje teksta
 - skupo i dugotrajno formiranje metapodataka
 - opis zavisi od konteksta onoga ko opisuje; to može da bude različito od konteksta onoga ko pretražuje

Ekstrakcija sadržaja slike - pristup 1

- Kako izvući informacije o sadržaju slike?
- Angažovati čoveka
 - opisivanje sadržaja = formiranje metapodataka
 - opis predstavlja karakteristike visokog nivoa (**high-level features**)
 - opis koristi alfanumeričke tipove podataka: ključne reči, datume, imena, koordinate
 - za pretraživanje se koristi klasična baza podataka ili tehnike za pretraživanje teksta
 - skupo i dugotrajno formiranje metapodataka
 - opis zavisi od konteksta onoga ko opisuje; to može da bude različito od konteksta onoga ko pretražuje
- Problem se svodi na prethodni – pretraživanje teksta

Ekstrakcija sadržaja slike - pristup 2

- Kako izvući informacije o sadržaju slike?

Ekstrakcija sadržaja slike - pristup 2

- Kako izvući informacije o sadržaju slike?
- Angažovati računar

Ekstrakcija sadržaja slike - pristup 2

- Kako izvući informacije o sadržaju slike?
- Angažovati računar
 - opisivanje sadržaja = ekstrakcija osobina slike

Ekstrakcija sadržaja slike - pristup 2

- Kako izvući informacije o sadržaju slike?
- Angažovati računar
 - opisivanje sadržaja = ekstrakcija osobina slike
 - opis predstavlja osobine niskog nivoa (**low-level features**)

Ekstrakcija sadržaja slike - pristup 2

- Kako izvući informacije o sadržaju slike?
- Angažovati računar
 - opisivanje sadržaja = ekstrakcija osobina slike
 - opis predstavlja osobine niskog nivoa ([low-level features](#))
 - ove osobine se **izračunavaju** na osnovu prostog sadržaja slike - piksela

Ekstrakcija sadržaja slike - pristup 2

- Kako izvući informacije o sadržaju slike?
- Angažovati računar
 - opisivanje sadržaja = ekstrakcija osobina slike
 - opis predstavlja osobine niskog nivoa (**low-level features**)
 - ove osobine se **izračunavaju** na osnovu prostog sadržaja slike - piksela
 - za pretraživanje se koristi specijalizovani indeks

Ekstrakcija sadržaja slike - pristup 2

- Kako izvući informacije o sadržaju slike?
- Angažovati računar
 - opisivanje sadržaja = ekstrakcija osobina slike
 - opis predstavlja osobine niskog nivoa (**low-level features**)
 - ove osobine se **izračunavaju** na osnovu prostog sadržaja slike - piksela
 - za pretraživanje se koristi specijalizovani indeks
 - brzo i jeftino kreiranje metapodataka

Ekstrakcija sadržaja slike - pristup 2

- Kako izvući informacije o sadržaju slike?
- Angažovati računar
 - opisivanje sadržaja = ekstrakcija osobina slike
 - opis predstavlja osobine niskog nivoa (**low-level features**)
 - ove osobine se **izračunavaju** na osnovu prostog sadržaja slike - piksela
 - za pretraživanje se koristi specijalizovani indeks
 - brzo i jeftino kreiranje metapodataka
 - manji kvalitet pretrage

Osobine niskog nivoa

- Najčešće osobine niskog nivoa
 - boja
 - oblik
 - tekstura

Segmentacija slike

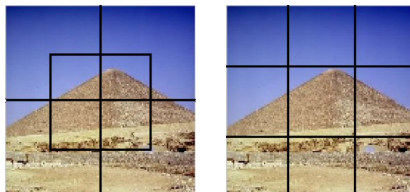
- Da li osobine izračunavamo za celu sliku ili pojedinačno po delovima?
 - **globalne osobine**: za celu sliku
 - **lokalne osobine**: po delovima

Globalne osobine

- Vrednosti predstavljaju prosek za celu sliku
- Gubi se razlikovanje prednjeg plana i pozadine
- Ne odslikava čovekovo poimanje slike
- Računski jednostavno
- Postoje (relativno) uspešne implementacije

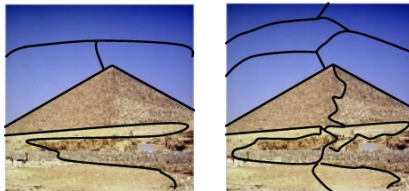
Način segmentacije

- **Tile-based:** pravilan raspored delova
- Slični problemi kao kod globalnih osobina
- Računski jednostavno
- Neke šeme su se pokazale dobro u praksi



Način segmentacije

- **Region-based**: regioni se formiraju na osnovu sadržaja
- Podela slike na vizuelno koherentne zone
- Može da identifikuje značajne objekte
- Računski složeno
- Nepouzđano

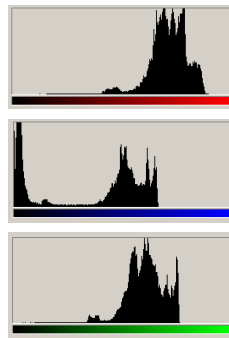


Boja kao osobina slike

- Izračuna se **signatura** regiona/slike na osnovu boje piksela
- Veći broj različitih osobina
 - histogram boja
 - vektor koherencije boja
 - ...

Histogram boja

- Prebrojimo koliko ima piksela za svaku boju u slici
- Od dobijenih vrednosti sastavimo histogram
- Za kolor slike to uradimo za svaki kanal (npr. R, G i B)



Histogram boja

- Histogram boja je proizveo 3 niza *integer*-a
- Normalizujemo njihove vrednosti (da ne zavise od veličine slike)
- Spojimo sve u jedan niz \rightarrow signatura
- Histogram boja je prilično neosetljiv na male promene
- Sve slike moraju imati isti model boja (RGB, HSV, ...)

Histogram boja

- Histogram boja je proizveo 3 niza *integer*-a
- Normalizujemo njihove vrednosti (da ne zavise od veličine slike)
- Spojimo sve u jedan niz \rightarrow signatura
- Histogram boja je prilično neosetljiv na male promene
- Sve slike moraju imati isti model boja (RGB, HSV, ...)
- Pretraživanje = poređenje signatura u vektorskom prostoru

Vektor koherencije boja

- Piksele klasifikujemo u grupe po boji
 - pikseli sa vrednošću u opsegu $[0, 10] \rightarrow$ klasa 1, $[11, 20] \rightarrow 2 \dots$

22	10	21	22	15	16
24	21	13	20	14	17
23	17	38	23	17	16
25	25	22	14	15	21
27	22	12	11	21	20
24	21	10	12	22	23

Vektor koherencije boja

- Piksele klasifikujemo u grupe po boji
 - pikseli sa vrednošću u opsegu $[0, 10] \rightarrow$ klasa 1, $[11, 20] \rightarrow 2 \dots$

$$\begin{bmatrix} 22 & 10 & 21 & 22 & 15 & 16 \\ 24 & 21 & 13 & 20 & 14 & 17 \\ 23 & 17 & 38 & 23 & 17 & 16 \\ 25 & 25 & 22 & 14 & 15 & 21 \\ 27 & 22 & 12 & 11 & 21 & 20 \\ 24 & 21 & 10 & 12 & 22 & 23 \end{bmatrix} \rightarrow \begin{bmatrix} 2 & 1 & 2 & 2 & 1 & 1 \\ 2 & 2 & 1 & 2 & 1 & 1 \\ 2 & 1 & 3 & 2 & 1 & 1 \\ 2 & 2 & 2 & 1 & 1 & 2 \\ 2 & 2 & 1 & 1 & 2 & 2 \\ 2 & 2 & 1 & 1 & 2 & 2 \end{bmatrix}$$

Vektor koherencije boja

- Piksele klasifikujemo u grupe po boji
 - pikseli sa vrednošću u opsegu $[0, 10] \rightarrow$ klasa 1, $[11, 20] \rightarrow 2 \dots$
- Uočimo regione obojene istom klasom (pikseli povezani 8-susedstvom)

$$\begin{bmatrix} 22 & 10 & 21 & 22 & 15 & 16 \\ 24 & 21 & 13 & 20 & 14 & 17 \\ 23 & 17 & 38 & 23 & 17 & 16 \\ 25 & 25 & 22 & 14 & 15 & 21 \\ 27 & 22 & 12 & 11 & 21 & 20 \\ 24 & 21 & 10 & 12 & 22 & 23 \end{bmatrix} \rightarrow \begin{bmatrix} 2 & 1 & 2 & 2 & 1 & 1 \\ 2 & 2 & 1 & 2 & 1 & 1 \\ 2 & 1 & 3 & 2 & 1 & 1 \\ 2 & 2 & 2 & 1 & 1 & 2 \\ 2 & 2 & 1 & 1 & 2 & 2 \\ 2 & 2 & 1 & 1 & 2 & 2 \end{bmatrix}$$

Vektor koherencije boja

- Piksele klasifikujemo u grupe po boji
 - pikseli sa vrednošću u opsegu $[0, 10] \rightarrow$ klasa 1, $[11, 20] \rightarrow 2 \dots$
- Uočimo regione obojene istom klasom (pikseli povezani 8-susedstvom)

$$\begin{bmatrix} 22 & 10 & 21 & 22 & 15 & 16 \\ 24 & 21 & 13 & 20 & 14 & 17 \\ 23 & 17 & 38 & 23 & 17 & 16 \\ 25 & 25 & 22 & 14 & 15 & 21 \\ 27 & 22 & 12 & 11 & 21 & 20 \\ 24 & 21 & 10 & 12 & 22 & 23 \end{bmatrix} \rightarrow \begin{bmatrix} 2 & 1 & 2 & 2 & 1 & 1 \\ 2 & 2 & 1 & 2 & 1 & 1 \\ 2 & 1 & 3 & 2 & 1 & 1 \\ 2 & 2 & 2 & 1 & 1 & 2 \\ 2 & 2 & 1 & 1 & 2 & 2 \\ 2 & 2 & 1 & 1 & 2 & 2 \end{bmatrix} \rightarrow \begin{bmatrix} BCBBA \\ BBCBA \\ BCDBA \\ BBBAE \\ BBAAE \\ BBAAE \end{bmatrix}$$

Vektor koherencije boja

- Piksele klasifikujemo u grupe po boji
 - pikseli sa vrednošću u opsegu $[0, 10] \rightarrow$ klasa 1, $[11, 20] \rightarrow 2 \dots$
- Uočimo regione obojene istom klasom (pikseli povezani 8-susedstvom)
- Označimo piksele kao koherentne ili nekoherentne
 - koherentan piksel je deo regiona piksela iste klase

$$\begin{bmatrix} 22 & 10 & 21 & 22 & 15 & 16 \\ 24 & 21 & 13 & 20 & 14 & 17 \\ 23 & 17 & 38 & 23 & 17 & 16 \\ 25 & 25 & 22 & 14 & 15 & 21 \\ 27 & 22 & 12 & 11 & 21 & 20 \\ 24 & 21 & 10 & 12 & 22 & 23 \end{bmatrix} \rightarrow \begin{bmatrix} 2 & 1 & 2 & 2 & 1 & 1 \\ 2 & 2 & 1 & 2 & 1 & 1 \\ 2 & 1 & 3 & 2 & 1 & 1 \\ 2 & 2 & 2 & 1 & 1 & 2 \\ 2 & 2 & 1 & 1 & 2 & 2 \\ 2 & 2 & 1 & 1 & 2 & 2 \end{bmatrix} \rightarrow \begin{bmatrix} BCBBA \\ BBCBA \\ BCDBA \\ BBBAE \\ BBAAE \\ BBAAE \end{bmatrix}$$

Vektor koherencije boja

- Piksele klasifikujemo u grupe po boji
 - pikseli sa vrednošću u opsegu $[0, 10] \rightarrow$ klasa 1, $[11, 20] \rightarrow 2 \dots$
- Uočimo regione obojene istom klasom (pikseli povezani 8-susedstvom)
- Označimo piksele kao koherentne ili nekoherentne
 - koherentan piksel je deo regiona piksela iste klase
- Formiramo binarne vektore po horizontali ili vertikali

$$\begin{bmatrix} 22 & 10 & 21 & 22 & 15 & 16 \\ 24 & 21 & 13 & 20 & 14 & 17 \\ 23 & 17 & 38 & 23 & 17 & 16 \\ 25 & 25 & 22 & 14 & 15 & 21 \\ 27 & 22 & 12 & 11 & 21 & 20 \\ 24 & 21 & 10 & 12 & 22 & 23 \end{bmatrix} \rightarrow \begin{bmatrix} 2 & 1 & 2 & 2 & 1 & 1 \\ 2 & 2 & 1 & 2 & 1 & 1 \\ 2 & 1 & 3 & 2 & 1 & 1 \\ 2 & 2 & 2 & 1 & 1 & 2 \\ 2 & 2 & 1 & 1 & 2 & 2 \\ 2 & 2 & 1 & 1 & 2 & 2 \end{bmatrix} \rightarrow \begin{bmatrix} BCBBA \\ BBCBA \\ BCDBA \\ BBBAE \\ BBAAE \\ BBAAE \end{bmatrix}$$

Boja nije uvek dovoljna



Aktivne konture

- M. Kass, A. Witkin, and D. Terzopoulos. *Snakes: Active contour Models*. Intl.J. Computer Vision, vol. 1, pp. 321 - 331, 1987.
- Generiše krive koje se kreću unutar slike i traže granicu objekta
- Zavise od početne pozicije
- Mogu da završe u lokalnom minimumu

Gradient Vector Flow

- C. Xu and J.L. Prince. *Snakes, Shapes, and Gradient Vector Flow*. IEEE Transactions on Image Processing, 359-369, March 1998
- Usavršene aktivne konture

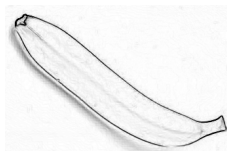
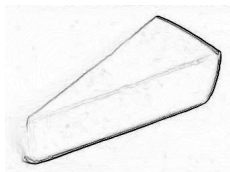
Furijeovi deskriptori

- Hannu Kauppinen et.al. *An Experimental Comparison of Autoregressive and Fourier Based Descriptors in 2D Shape Classification*. IEEE Transactions on Pattern Recognition and Machine Intelligence Vol. 17 No 2 Feb, 1995.
- Izračunava varijantu Furijeove transformacije za ivice objekta
- Otporan na geometrijske transformacije i šum



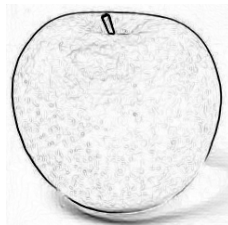
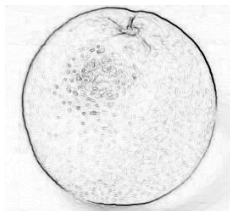
Oblik nije uvek dovoljan

Ovo se lako klasifikuje po obliku:



Oblik nije uvek dovoljan

Ovo se teško klasifikuje po obliku:



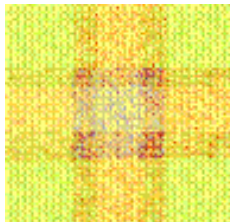
Boja i oblik zajedno ne moraju biti dovoljni

Ovo se teško klasifikuje i po obliku i po boji:



Pojam teksture

Ove dve slike su različite po boji, ali slične po teksturi:



Pojam teksture

- Tekstura je matematički opis ponavljajućeg šablona u slici
 - glatko
 - peskovito
 - zrnasto
 - trakasto

Matrica pojavljivanja

- P_{ij} : prebrojimo koliko puta se piksel boje i pojavljuje u smeru p u odnosu na piksel boje j
- Na primer: ako slika ima tri boje: 0, 1 i 2 i smer je „dole desno“

$$slika = \begin{bmatrix} 0 & 0 & 0 & 1 & 2 \\ 1 & 1 & 0 & 1 & 1 \\ 2 & 2 & 1 & 0 & 0 \\ 1 & 1 & 0 & 2 & 0 \\ 0 & 0 & 1 & 0 & 1 \end{bmatrix} \Rightarrow P_{ij} = \begin{bmatrix} 4 & 2 & 1 \\ 2 & 3 & 2 \\ 0 & 2 & 0 \end{bmatrix}$$

ISO/IEC 13249-5:2001

- „SQL/MM Still Image standard“
- Definiše standardne objektno-relacione tipove podataka za pretraživanje slika po sadržaju
 - SI_StillImage
 - SI_AverageColor
 - SI_Color
 - SI_ColorHistogram
 - SI_FeatureList
 - SI_PositionalColor
 - SI_Texture

Oracle interMedia

- Proširenje Oracle relacione baze podataka za rad sa slikama, zvukom i videom
- Slike se smeštaju u BLOB kolone baze podataka
- Podrška za više formata
- Ekstrakcija metapodataka zavisno od formata
- Konverzija formata
- Pretraživanje po sadržaju

Oracle interMedia

- Indeksiranje slike obuhvata:
 - segmentaciju slike
 - analizu boje
 - analizu oblika
 - analizu teksture
- Rezultat ovih koraka je vektor sa 3000-4000 elemenata
→ signatura
- Upit se izražava slikom-uzorkom
- Izračunavanje upita = poređenje signatura u vektorskom prostoru

Kakav indeks koristiti za pretraživanje slika?

- Sekvencijalna pretraga nije dovoljno efikasna
 - izračunavanje sličnosti vektora u 4000-dimenzionalnom prostoru
 - za svaku od npr. 50000 slika
 - $50000 \times 4000 =$ previše

Kakav indeks koristiti za pretraživanje slika?

- Sekvencijalna pretraga nije dovoljno efikasna
 - izračunavanje sličnosti vektora u 4000-dimenzionalnom prostoru
 - za svaku od npr. 50000 slika
 - $50000 \times 4000 =$ previše
- Plan A: neka od struktura za višedimenzionalne vektorske prostore
 - grid file, k-d tree, quad-tree, K-D-B tree, hB-tree, R-tree
- Plan B: svođenje problema na prethodni (pretraživanje teksta)
- Plan C: specijalizovani indeksi za pojedine osobine slike

Plan B

- Koristi mehanizme za pretraživanje teksta: invertovani indeks i tf-idf težine
- Prilagodi reprezentaciju osobina slike tako da se mogu smestiti u rečnik Lucene indeksa
- LIRE: Lucene Image REtrieval
 - <http://www.semanticmetadata.net/lire/>
 - koristi tri osobine iz MPEG-7 standarda: ScalableColor, ColorLayout i EdgeHistogram
 - dodaje još jednu svoju: Auto Color Correlogram

Osnove

- Zvuk se proizvodi konverzijom energije u talase koji se prostiru kroz neki etar, talasi dolaze do uha gde se konvertuju u nervni impuls što mozak detektuje kao zvuk
- Ljudsko uho može detektovati zvuk frekvencije između 20Hz i 20kHz (zavisi i od godina)
- Složen i subjektivan fenomen, teško je napraviti precizan model zvuka
- Waveform (predstava talasima) zvuka prikazuje promenu amplitude zvuka kroz vreme
- MIDI - format za skladištenje instrukcija kako će se zvuk proizvesti (nešto kao vektorska grafika kod slike)
- Percepcija zvuka ima i psihološku dimenziju: isti zvuk prvo tiše pa glasnije, Vaše ime na kraju glasne sale, Stereo zvuk, utisak da je nešto desno, odnosno levo, itd.

Audio žanrovi

- Bazirani na govoru: radio program, telefonska komunikacija, snimljeni razgovori
- Bazirana na muzici: instrumentalna, vokalna (pevanje)
- Ostalo: alarm, zvuci iz prirode

Zvuci bazirani na govoru

- Pretraga radio vesti, snimljenih razgovora, predavanja
- Iste reči izgovorene od iste osobe pod istim uslovima mogu proizvesti prilično različite talase
- Kolike su tek onda razlike kada su u pitanju različiti ljudi
- Koliko glasno je nešto izgovoreno, akcenat, harmonici, itd.
- Zvuk je vrlo složen fenomen, ne čujemo svi isto, subjektivnost u slušanju (poznate termine bolje čujemo od drugih, naše ime na kraju bučne sale)

Detekcija osobina u govoru

- Sadržaj: fonemi, *one-best word recognition*, *n-best*
- Identitet: identifikacija govornika, podela zvuka po različitim govornicima
- Jezičke osobine: jezik, dijalekt, akcenat
- Ostale osobine: okruženje, kanal (stereo, surround), itd.

Kako *speech recognition* radi

- Tri koraka:
 - Prebacivanje *waveform*-i u foneme (delove reči, slogove)
 - Segmentacija reči, kako foneme grupisati u reči, gde je kraj jedne reči i početak druge
 - Detekcija koja je reč izgovorena
- Sva tri koraka se obično obučavaju nadgledanim mašinskim učenjem

Pretraživanje zvučnih zapisa bazirano na uzorku

- Pretraživanje zvučnih zapisa baziranih na muzici
- Iz zvučnih zapis izvući osobine i vektore osobina indeksirati nekom strukturom (npr. R-stablom)
- Definirati meru sličnosti dva zapisa
- Klasterovati zapise u MBR (minimal bounding rectangle)
- Iz uzorka izvući osobine i formirati vektor osobina
- Za vektor osobina upita utvrđuje se kom klasteru pripada, odnosno utvrđuje se MBR koji se vraća ili najbliži MBR

Osobine u zvučnim zapisima

- Preuzeto iz magistarskog rada: Aleksandar Kovačević, Adaptivni sistem za pretraživanje zvučnih zapisa, Fakultet tehničkih nauka u Novom Sadu (2006)
- *Mean square* - procena apmlitude zvučnog signala
- *ZeroCrossings* - broj promena znaka zvučnog signala
- *Spectrum[32]* - prvih 32 koeficijenta dobijena Diskretnom Furijeovom Transformacijom (DFT) signala
- *SpectralSum* - sumu razlika susednih članova niza Spectrum
- *Beats (per minute, BPM)* - procena ritma zvučnog zapisa
- *Avg FFT Delta[32]* - Članovi niza *Avg FFT Delta* su razlike susednih elemenata niza Spectrum

Osobine u zvučnim zapisima

- *Haar[64]* - Niz 64 koeficijenta dobijena Harovom wavelet transformacijom
- *Song seconds* - Vreme trajanja zvučnog zapisa u sekundama
- *Energy difference (ED)* - signal se deli na delove (frejm), za svaki deo računa se energija, *Energy difference* je suma razlika energija za sve frejmove zvučnog zapisa
- *Energy Zero Crossing* - događa se kad vrednosti energije dva uzastopna frejma (*Energy difference*) imaju različit znak, *Energy Zero Crossings* je broj ovih događaja za zvučni zapis

Definisanje mere sličnosti dva zvučna zapisa

- Formirati obučavajući skup, odnosno mali deo kolekcije zvučnih zapisa podeliti na klase
- Iskoristiti ovaj obučavajući skup za algoritam (npr. genetski algoritam) za tjuniranje prostora osobina
- Rezultat tjuniranja - koeficijenti koji se stavljaju ispred osobina kako bi se optimizovala metrika sličnosti

Osnove

- Niz frejmova, odnosno rasterskih slika
- Medij koji je doživeo revoluciju razvojem ICT, digitalni video je danas normalna pojava, digitalne kamere su u telefonima, laptop računarima, itd.
- Video sa zvukom i tekstom je multimedija, ali se pod pojmom video IR najčešće podrazumeva pretraga kolekcije videa koji sadrži i zvučne i tekstualne komponente (u formi *subtitle-a*)

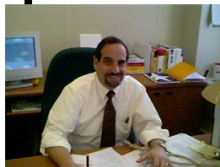
Primena video IR

- Postoje velike baza video zapisa
- Potreba za pretragom ovih kolekcija je sve veća razvojem ICT
- Primer: BBC arhiva u toku jedne godine primi više od 500,000 upita čiji bi odgovor mogao da bude video ili deo videa, YouTube primi još više upita
- Ogroman prostor za napredak i u *text-based* video IR, a posebno u *content-based* video IR

Kompleksnost problema

- Veliki broj osobina koji se mogu ekstrahovati iz videa
- Kretanje objekata u kadrovima, kretanje kamere, promena zoom-a, itd.
- Koristi se: OCR, speech recognition, face recognition, scene recognition, itd.
- Odgovor ne treba da bude ceo video, nego njegov deo koji odgovara informacionoj potrebi korisnika (obično 20-40 sekundi) - zbog ovoga je segmentacija scena jako bitno

Upotreba OCR i speech recognition tehnike



Query:

Find pictures of Harry Hertz, Director of the National Quality Program, NIST



Speech:

We're looking for people that have a broad range of expertise that have business knowledge that have knowledge on quality management on quality improvement and in particular ...

OCR:

H,arry Hertz a Director aro 7 wa-
 ,i,,ty Program
 .Harrry Hertz. a Director

..

Informaciona potreba

