

INDEKSIRANJE I PRETRAŽIVANJE SLIKA KORIŠTENJEM MULTIREZOLUCIONE ANALIZE

Vladimir Risojević, Zdenka Babić, Elektrotehnički fakultet u Banjaluci

Sadržaj – Porast količine slika u digitalnom formatu nije praćen odgovarajućim razvojem efikasnih metoda za automatsko indeksiranje i pretraživanje baza slika na osnovu sadržaja slika. U ovom radu prikazan je pristup opisanom problemu korištenjem multirezolucione analize. Ova reprezentacija sadrži informaciju o prostornom rasporedu obilježja na slici i ne zavisi od rezolucije polazne slike. Upoređeni su rezultati pretraživanja baza slika korištenjem Haarovog i Daubechies 4 waveleta.

1. UVOD

Danas su multimedijalni sadržaji, podaci u obliku slika, te video i audio zapisa, sastavni dio računarske svakodnevice. Neke od primjena multimedijalnih baza uključuju [1]: biomedicinske slike, digitalne biblioteke, muzeje, arhive dokumenata, arhive televizijskog i radio programa, CAD/CAM sisteme, sisteme za nadgledanje, identifikaciju kriminalaca, geografske informacione sisteme, itd. Međutim, tehnike za efikasno memorisanje, indeksiranje, klasifikaciju i pretraživanje multimedijalnog sadržaja još uvijek nisu na nivou na kojem su odgovarajuće tehnike za rad sa tradicionalnim, alfanumeričkim, tipovima podataka.

Upravo zahvaljujući zrelosti sistema za pretraživanje i obradu tekstualnih podataka prvi pristup indeksiranju i pretraživanju baza slika (ali i multimedijalnih baza uopšte) sastojao se od pridruživanja tekstualnog opisa svakoj slici [1] i njihovog pretraživanja. Pored toga što je ovaj metod zamoran i teško skalabilan, jer zahtjeva da čovjek pregleda svaku sliku i pridruži joj odgovarajući tekstualni opis, ovaj metod je i nefleksibilan. Naime, ljudsko poimanje slike je subjektivno i moguće je da dva posmatrača na različit način shvate istu sliku ili da je shvate na isti način, ali da je opišu korištenjem drugih riječi. Osim toga, postoje i neke vizuelne osobine koje su lako uočljive ali teške za opisivanje riječima.

Drugi pristup koji je dobio na značaju u poslednjih 10-15 godina je indeksiranje slika korištenjem vizuelnih obilježja koja se mogu automatski izdvojiti iz slike. Ova obilježja uključuju [1,2]: raspodjelu boja, teksturu, oblike, prostorni raspored objekata kao i međusobne odnose ivica. Ovakvi sistemi koji koriste vizuelna obilježja niskog nivoa za indeksiranje i pretraživanje slika često se nazivaju sistemi za pretraživanje slika na osnovu sadržaja. Kod ovakvog sistema se za svaku sliku u bazi izračunavaju obilježja koja odražavaju raspodjelu pomenutih vizuelnih elemenata koji se mogu direktno dobiti iz digitalne reprezentacije slike. Formulisanje upita podrazumijeva pronalaženje ili formiranje slike sa željenim vizuelnim obilježjima niskog nivoa. U ovakvoj postavci problema više nije moguće zahtijevati da kriterij za pronalaženje slika u bazi bude poklapanje (jednakost) vektora obilježja već da ti vektori obilježja opisuju semantički slične slike. Dakle, neophodno je uvesti neku metriku u prostor obilježja koja će odražavati sličnost slika predstavljenih respektivnim vektorima obilježja.

Očigledno, nedostatak ovakvog pristupa je postojanje semantičkog jaza, tj. nepostojanje jasno utvrđene korespondencije između obilježja niskog nivoa i semantike slike. Međutim, ovakav pristup je znatno fleksibilniji i skalabilniji od ručnog dodavanja opisa slikama. Jedna od metoda koja je dala dosta dobre rezultate je i indeksiranje slika pomoću wavelet transformacije [3].

U ovom radu prikazano je poređenje rezultata koji se postižu indeksiranjem i pretraživanjem baze slika korištenjem Haarovog i waveleta iz Daubechies familije. Haarova baza se jednostavno i efikasno implementira, i pogodna je za otkrivanje diskontinuiteta u signalima. Sa druge strane, Daubechies baza je pogodna za analiziranje neprekidnih signala. U drugom odjeljku su ukratko prikazane osnovne ideje multirezolucione analize na primjeru jednodimenzionalnog signala, a zatim je pokazano kako se ovi koncepti proširuju na dvodimenzionalni slučaj. U trećem odjeljku opisan je način indeksiranja slika korištenjem wavelet transformacije i metrika koja je korištena za ocjenu sličnosti slika. Detalji implementacije algoritma i rezultati testiranja prikazani su u četvrtom odjeljku. Rad se završava zaključkom nakon čega slijedi spisak korištene literature.

2. MULTIREZOLUCIONA ANALIZA

Korištenjem wavelet transformacije moguće je sliku (ili signal uopšte) analizirati na različitim rezolucijama [4]. Neka je $L^2(\mathbf{R})$ vektorski prostor kvadratno-integrabilnih realnih funkcija. Ukoliko signal $f(x)$ ima konačnu energiju, onda je $f(x) \in L^2(\mathbf{R})$. Neka su elementi vektorskog prostora $V^j \subset L^2(\mathbf{R})$ aproksimacije svih funkcija iz $L^2(\mathbf{R})$ na rezoluciji 2^j . Lako se može pokazati da aproksimacija signala na rezoluciji 2^{j+1} sadrži sve informacije potrebne za određivanje aproksimacije na nižoj rezoluciji 2^j . Iz ovoga slijedi uređenje vektorskih prostora

$$\forall j \in \mathbf{Z}, \quad V^j \subset V^{j+1} \quad (1)$$

Za reprezentaciju aproksimacije signala na rezoluciji 2^j neophodno je odrediti bazu vektorskog prostora V^j . Može se pokazati [4] da postoji jedinstvena funkcija $\phi(x) \in L^2(\mathbf{R})$ takva da je skup funkcija:

$$\phi_i^j(x) = 2^{-j/2} \phi(2^j x - i), \quad i \in \mathbf{Z} \quad (2)$$

ortonormirana baza prostora V^j . Funkcija $\phi(x)$ se naziva skalirajuća funkcija.

Očigledno, pri prelazu na nižu rezoluciju gubi se određena informacija o izgledu signala. Razlika između aproksimacija

na rezolucijama 2^{j+1} i 2^j naziva se *detalji signala* na rezoluciji 2^j . Detalji signala na rezoluciji 2^j očigledno ne pripadaju prostoru \mathbf{V}^j . Detalji signala u stvari pripadaju prostoru $\mathbf{W}^j \subset \mathbf{V}^{j+1}$, koji predstavlja ortogonalni komplement prostora \mathbf{V}^j u \mathbf{V}^{j+1} .

Da bi se mogli odrediti detalji signala na rezoluciji 2^j potrebno je pronaći bazu vektorskog prostora \mathbf{W}^j . Slično kao za prostor \mathbf{V}^j , postoji jedinstvena funkcija $\psi(x) \in \mathbf{L}^2(\mathbf{R})$ takva da je skup funkcija:

$$\psi_i^j(x) = 2^{-j/2} \psi(2^j x - i), \quad i \in \mathbf{Z} \quad (3)$$

ortonormirana baza prostora \mathbf{W}^j . Funkcija $\psi(x)$ se naziva ortogonalni majka wavelet [4]. Familija funkcija $2^{-j/2} \psi(2^j x - i)$, $(i, j) \in \mathbf{Z}^2$ je ortonormirana baza prostora $\mathbf{L}^2(\mathbf{R})$. Wavelet transformacijom signala $f(x) \in \mathbf{L}^2(\mathbf{R})$ se upravo dobijaju koeficijenti reprezentacije signala u opisanoj bazi. Iz ovoga slijedi da se wavelet transformacijom signal može predstaviti pomoću linearne kombinacije grube aproksimacije i detalja na sve većim rezolucijama. Drugim riječima, uzimanjem u obzir samo aproksimacije i detalja zaključno sa rezolucijom npr. 2^j dobijamo aproksimaciju signala na rezoluciji 2^j .

Opisani koncepti wavelet transformacije se jednostavno proširuju na slučaj dvodimenzionalnih signala u prostoru $\mathbf{L}^2(\mathbf{R}^2)$ [4]. Posebno značajan je slučaj separabilne dvodimenzionalne transformacije koja omogućava transformaciju dvodimenzionalnog signala sukcesivnim jednodimenzionalnim transformacijama po redovima i po kolonama.

Wavelet transformacija ima nekoliko osobina koje je čine pogodnom za reprezentaciju vizuelnih obilježja slike [3]:

- Najveći dio energije signala sadržan je u relativno malom broju koeficijenata sa najvećom magnitudom tako da se samo pomoću tog relativno malog broja koeficijenata može dobiti vjerna reprezentacija slike.
- Pošto je reprezentacija slike multirezoluciona, informacija sadržana u koeficijentima ne zavisi od rezolucije polazne slike.
- Korištenjem wavelet transformacije mogu se izdvojiti informacije o ivicama na slici.
- Obilježja se relativno lako izračunavaju.

Najjednostavniji waveleti pripadaju Haarovoj bazi. Skalirajuća funkcija ove baze je Hevisajdova funkcija, a majka wavelet je funkcija:

$$\psi(x) = \begin{cases} 1, & 0 \leq x < 1/2 \\ -1 & 1/2 \leq x < 1. \\ 0 & \text{inače} \end{cases} \quad (4)$$

Haarovi waveleti očigledno nisu neprekidne funkcije i samim tim nisu pogodni za analizu neprekidnih funkcija sa eventualno neprekidnim izvodima.

Najpoznatiji predstavnici neprekidnih waveleta su waveleti iz Daubechies familije. Skalirajuća funkcija ove familije $\phi_r(x) \in \mathbf{L}^2(\mathbf{R})$, $r \in \mathbf{Z}$, ima osobinu da je sekvenca $\{\phi_r(x - k) | k \in \mathbf{Z}\}$ ortonormirana. Majka waveleti ove familije imaju slične osobine. Ovi waveleti imaju kompaktan interval na kojem je njihova vrijednost različita od nule i imaju $r/5$ neprekidnih izvoda. Takođe je i

$$\int_{-\infty}^{\infty} \psi_r(x) dx = \dots = \int_{-\infty}^{\infty} x^r \psi_r(x) dx = 0. \quad (5)$$

Ovakva baza je efikasnija za analizu neprekidnih funkcija od Haarove baze.

3. INDEKSIRANJE I PRETRAŽIVANJE KOLEKCIJE SLIKA

Kao obilježja za kompaktnu reprezentaciju slika u bazi usvojili smo wavelet koeficijente dobijene dekompozicijom slike [3]. Kompletna dekompozicija slike zahtijeva onoliki broj koeficijenata koliko piksela ima polazna slika. Međutim, već je pomenuto da se veoma dobra reprezentacija slike može dobiti korištenjem relativno malog broja koeficijenata. Iz ovog razloga ne koristimo sve dobijene koeficijente kao obilježja, već za reprezentaciju slika biramo samo određeni broj koeficijenata sa najvećom magnitudom. Slike se kod pretraživanja ne rekonstruišu već se mjera njihove udaljenosti (sličnosti) izračunava direktno korištenjem reprezentacije slika u wavelet domenu.

Postupak indeksiranja se odvija na sledeći način. Pošto slike potencijalno mogu biti različitih dimenzija prvi korak je normalizacija dimenzija slike na 128×128 piksela. Na ovaj način se gubi informacija o odnosu dimenzija slike, ali se postiže određena invarijantnost na skaliranje pošto su dimenzije svih slika prije indeksiranja normalizovane na isti iznos. U sledećem koraku se vrši wavelet dekompozicija date slike. Ukoliko se radi o slici u boji, dekompozicija se odvojeno primjenjuje na svaki od tri kanala izabranog kolor modela. U slučaju grayscale slika, dekompozicija se primjenjuje direktno na vrijednosti piksela. Kao što je pomenuto, ne čuvaju se svi koeficijenti dekompozicije već se za svaki kolor kanal zadržava samo određen broj koeficijenata sa najvećom magnitudom. Za ostale koeficijente se uzima vrijednost nula. Na ovaj način, slika se može predstaviti rijetko popunjenom matricom što ovu reprezentaciju čini pogodnom i sa stanovišta memorijskog prostora potrebnog za čuvanje vektora obilježja.

Kao mjera udaljenosti slika izabrana je klasična L^1 metrika u prostoru obilježja od koje se polazi i u [3]. Neka su Q i T dvije slike. Njihova udaljenost data je izrazom

$$d(Q, T) = \sum_{i, j} |\tilde{Q}(i, j) - \tilde{T}(i, j)|, \quad (6)$$



Slika 1. Primjer pretraživanja korištenjem Haarovog waveleta. Upit je u gornjem lijevom uglu. Slike su skalirane radi preglednijeg prikazivanja.



Slika 2. Primjer pretraživanja korištenjem Daubechies 4 waveleta. Upit je u gornjem lijevom uglu. Slike su skalirane radi preglednijeg prikazivanja.

gdje su \tilde{Q} i \tilde{T} wavelet koeficijenti slika Q i T , respektivno. U slučaju da se radi o slikama u boji, sumiranje se vrši i po sva tri kolor kanala.

$$d(Q, T) = \sum_{\substack{i, j \\ \tilde{Q}(i, j) \neq 0}} |\tilde{Q}(i, j) - \tilde{T}(i, j)|. \quad (7)$$

Wavelet koeficijenti slika predstavljaju detalje u slikama na različitim rezolucijama i sa različitim orijentacijama. Pošto se u ljudskoj percepciji različita vizuelna obilježja, koja mozak "izračunava", uzimaju u obzir sa različitom težinom, u [3] su određenim grupama koeficijenata dodijeljene različite težine. Analitičko izračunavanje ovih težina je nemoguće na sadašnjem stepenu razvoja nauke tako da preostaje njihovo određivanje nadgledanim obučavanjem na osnovu mjera sličnosti između slika iz trening skupa koje su odredili ljudski subjekti. Međutim, u ljudskoj percepciji pomenute težine zavise od sadržaja slike koja se posmatra te će ovako određene težine neminovno imati otklon prema oblasti pokrivenoj trening skupom. Da bi se ovaj otklon izbjegao neophodno je da sistem za pretraživanje slika na osnovu sličnosti ima i mehanizam *relevance feedback* [1] koji omogućava kontinuirano prilagođavanje ovih težina u toku rada sistema čime se u suštini proširuje trening skup. Pošto je cilj ovog rada samo poređenje tipova waveleta iskorištenih za indeksiranje slika težine se ne dodjeljuju koeficijentima, tj. koristi se klasična L^1 norma data jednačinom (6).

Pretpostavimo da slika Q predstavlja upit postavljen sistemu. Da bi se u dobijenim rezultatima pretraživanja dobile slike koje posjeduju obilježja koja odgovaraju obilježjima koja posjeduje upit, u [4] se u sumi u jednačini (6) uzimaju samo članovi u kojima je $\tilde{Q}(i, j) \neq 0$. Sada jednačina (6) dobija oblik (7).

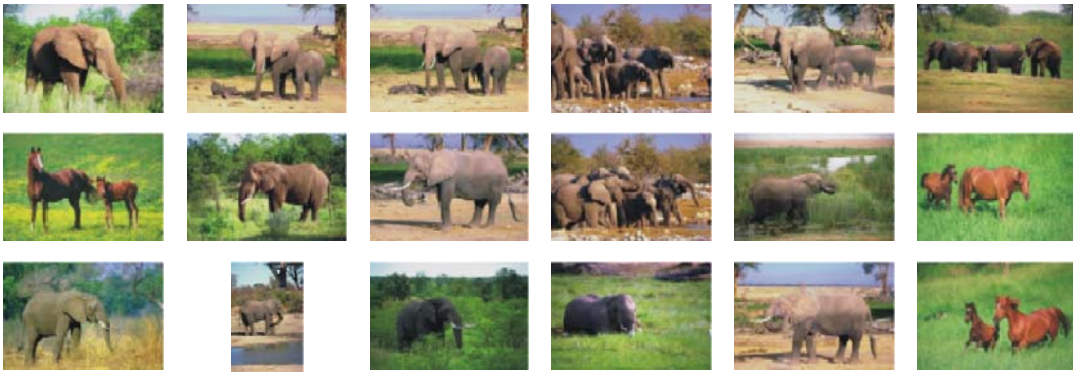
Kao rezultati pretraživanja dobijaju se slike iz baze sortirane prema rastućoj vrijednosti udaljenosti od slike-upita.

4. REZULTATI

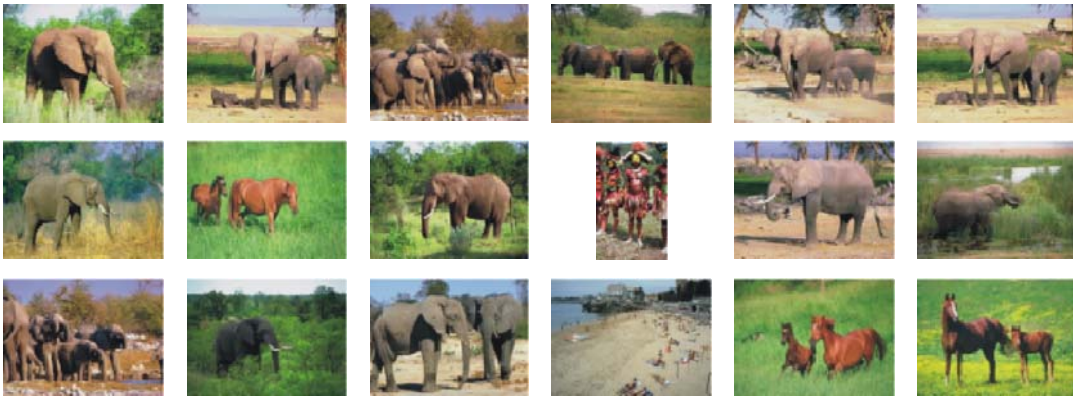
Opisani algoritmi za indeksiranje i pretraživanje kolekcije slika implementirani su korištenjem MATLAB-a, Image Processing Toolboxa i Wavelet Toolboxa i testirani su na kolekciji slika prikupljenoj od strane istraživača na Pennsylvania State univerzitetu [5]. Kolekcija se sastoji od 1000 slika dimenzija 384*256, odnosno, 256*384 piksela u JPEG formatu. Indeksiranje kolekcije traje 5-6 minuta na PC računaru sa jednim 1,6GHz Pentium IV procesorom i 512MB RAM-a. Obrada jednog upita i pretraživanje indeksirane kolekcije traje oko 10 sekundi.

Pošto je za pretraživanje slika po sličnosti bitno da se uzme u obzir i perceptualna sličnost između boja objekata na slici, RGB model nije dobar izbor za jedan ovakav sistem. Umjesto toga u ovom radu je korišten HSV model kod kojeg su perceptualno sličnije boje bliže u trodimenzionalnom prostoru kojim je model predstavljen. Ovo je provjereno i eksperimentalno te je HSV model zadržan prilikom poređenja wavelet baza.

Za reprezentaciju slika u obje wavelet baze zadržano je po 60 koeficijenta sa najvećom magnitudom iz svakog kolor kanala. Eksperimenti su pokazali da se značajnijim smanjenjem ili povećanjem broja koeficijenata performanse umanjuju. Kod smanjenja broja koeficijenata razlog za to je očigledno odbacivanje važnih vizuelnih atributa slike, dok se povećanjem broja koeficijenata ne dobija poboljšanje performansi zato što se tada uzima u obzir više detalja i odvraća pažnja od semantike slike.



Slika 3. Primjer pretraživanja korištenjem Haarovog waveleta. Upit je u gornjem lijevom uglu. Slike su skalirane radi preglednijeg prikazivanja.



Slika 4. Primjer pretraživanja korištenjem Daubechies 4 waveleta. Upit je u gornjem lijevom uglu. Slike su skalirane radi preglednijeg prikazivanja.

Eksperimentalno su upoređeni indeksiranje i pretraživanje kolekcije korištenjem Haarovog i Daubechies 4 (DB4) waveleta i karakteristični primjeri rezultata su prikazani na slikama 1-4. Može se uočiti da su rezultati dobijeni korištenjem oba waveleta dosta dobri. Međutim, korištenjem Haarovog waveleta dobijamo za nijansu bolje rezultate nego korištenjem DB4 waveleta. Razlog za ovo je vjerovatno postojanje jasno definisanih ivica na objektima, odnosno, ukoliko se slika posmatra kao funkcija, postojanje tačaka prekida.

5. ZAKLJUČAK

U ovom radu opisan je pristup indeksiranju i pretraživanju slika na osnovu obilježja dobijenih wavelet transformacijom. Upoređeno je indeksiranje i pretraživanje korištenjem 60 koeficijenata sa najvećom magnitudom dobijenih dekompozicijom slike Haarovim i DB4 waveletima. Eksperimentalni rezultati su pokazali da oba pristupa daju dosta dobre rezultate uz malu prednost pristupa korištenjem Haarovog waveleta.

Dalji razvoj uključuje dodjelu različitih težina koeficijentima koji se dobijaju dekompozicijom komponenti kolor-modela kao i određenim nivoima dekompozicije. Inicijalne vrijednosti ovih težina mogu se odrediti nadgledanim obučavanjem ali bi trebalo i omogućiti njihovu adaptaciju u toku interakcije sistema sa korisnikom.

Diskretna wavelet transformacija nije invarijantna na affine transformacije, niti postoji neki ekvivalent teoreme pomaka iz Furijeove transformacije. Zanimljivo bi bilo ispitati robusnost ovog metoda u odnosu na affine transformacije, translaciju i rotaciju, kao i mogućnosti uvođenja invarijantnosti na ove transformacije u vektor obilježja.

LITERATURA

- [1] B. Furht, S. W. Smoliar and H-J. Zhang, *Video and Image Processing in Multimedia Systems*, Kluwer Academic Publishers, 1995.
- [2] R. Schettini, G. Ciocca and S. Zuffi, "A survey on methods for colour image indexing and retrieval in image databases", In *Color Imaging Science: Exploiting Digital Media*, R. Luo and L. MacDonald, Eds. New York: J. Wiley, 2001.
- [3] C. E. Jacobs, A. Finkelstein, and D. H. Salesin, "Fast muliresolution image querying", in *SIGGRAPH 95 Conference Proceedings*, pp. 227-286, 1995.
- [4] S. G. Mallat, "A Theory for Multiresolution Signal Decomposition: The Wavelet Representation", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11(7):674-693, July 1989.
- [5] James Z. Wang, Jia Li, Gio Wiederhold, "SIMPLiCity: Semantics-sensitive Integrated Matching for Picture Libraries," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol 23, no.9, pp. 947-963, 2001.

Abstract – The increase in the amount of images available in digital formats has not been accompanied by the matching development of methods for efficient and effective content-based automatic indexing and retrieval in image databases. In this paper an approach to this problem using multiresolution analysis is presented. This representation contains spatial layout of features and does not depends on the resolution of the original image. The results of searching using Haar and Daubechies' 4 wavelet bases are presented..

INDEXING AND RETRIEVAL OF IMAGES USING MULTIREOLUTION ANALYSIS
Vladimir Risojević, Zdenka Babić