

PRETRAŽIVANJE BAZA SLIKA NA OSNOVU ADAPTIVNE KLASTERIZACIJE

Slobodan Čabarkapa^{1,2}, Nenad Kojić^{1,2}, Vladan Radosavljević², Goran Zajić², Branimir Reljin²
¹Viša ICT škola, Beograd, ²Elektrotehnički fakultet, Beograd

Sadržaj – U ovom radu biće opisan sistem za pretragu baza slika gde se pomoću strategije aktivnog učenja pojačava uticaj asistencije korisnika na pretraživanje. Pored direktnog uticaja na pretraživanje, asistencija korisnika iskorišćena je i za adaptivnu klasterizaciju baza slika (ACID) na osnovu subjektivnog doživljaja. Slike koje su označene kao relevantne u poslednjem koraku pretraživanja grupisane su u klastere a zatim memorisane. Reprezentativni predstavnici klastera su, ravnopravno sa ostalim slikama iz baze, korišćeni u narednim procedurama pretraživanja. Na ovaj način deo istorije prethodnih upita uključen je u sistem obezbeđujući istovremeno brže i subjektivnije pronalaženje željenih slika. Pored navedenog, klasteri se ažuriraju nakon svake nove pretrage shodno trenutnim potrebama korisnika. Efikasnost predloženog ACID sistema je testirana na slikama iz MIT i Corel baze slika.

1. UVOD

Brz razvoj računarske tehnike uslovio je i potrebu korisnika za sve složenijim multimedijalnim sadržajima. Prenos i obrada ovakvog tipa informacija stvorio je zahteve za njihovim skladištenjem. Uzimajući u obzir količinu multimedijalnih sadržaja, nametnula se potreba za efikasnim pretraživanjem velikih baza podataka. Različiti algoritmi pretraživanja su predstavljeni u velikom broju dosadašnjih radova [1-16]. Glavna pažnja je posvećena pretraživanju slika [1-12] i multimedijalnih sadržaja uopšte [13-16]. Raniji sistemi za vizuelno pretraživanje su bili zasnovani na pretraživanju tekstualnih opisa, odnosno ključnih reči koje su manuelno dodeljivane slikama, nakon čega su se, za zadati upit, slične slike pronalazile na osnovu ključnih reči. Ovakav pristup se tradicionalno koristi u pretraživanju multimedijalnih sadržaja (klasični filmovi). Na takav način se ključnim frejmovima dodeljuju odgovarajući tekstualni opisi. Mada ove tehnike mogu biti vrlo efikasne, nekad i automatizovane, ipak imaju dva vrlo bitna nedostataka koji se najviše manifestuju pri radu sa velikim bazama podataka. Prvi je dugotrajni postupak manualnog unošenja tekstualnih opisa i pronalaženje odgovarajuće kombinacije ključnih reči u procesu pretrage. Ovaj problem se može jasno uočiti u web pretraživačima. Drugi nedostatak potiče od lingvističkih barijera i nepreciznosti pri opisivanju, kada tekstualni opis dovodi do velikih neslaganja između potreba korisnika i rezultata pretraživanja.

Da bi se prevazišli nedostaci pretraživanja baza slika na osnovu tekstualnog opisa uvedeno je pretraživanje na osnovu sadržaja slika (CBIR) [1-4]. Iz svake od slika različitim tehnikama izdvajaju se obeležja niskog nivoa kojima se opisuje boja, tekstura i oblik. Ovako izdvojena obeležja formiraju vektor obeležja koji se pridružuje odgovarajućoj slici u bazi. Procedura pretrage je zasnovana na određivanju sličnosti između upitne slike i slika u bazi, odnosno utvrđivanju rastojanja (najčešće Euklidskog) odgovarajućih vektora obeležja. Mada CBIR sistemi generalno daju zadovoljavajuće rezultate, u nekim slučajevima se

manifestuju ozbiljni nedostaci. Najveći nedostatak CBIR sistema je postojanje “semantičkog jaza” između obeležja niskog nivoa (boje, teksture, oblika) i obeležja visokog nivoa (semantičkog opisa slike).

Procedura pretrage se znatno poboljšava uvođenjem asistencije korisnika. Korisnik iz grupe ponuđenih slika, koje su objektivno najbliže upitnoj, određuje subjektivno najbliže slike i označava ih na odgovarajući način. Na osnovu označenih slika menjaju se težinski koeficijenti izdvojenih obeležja prema subjektivnoj percepciji vizuelnog sadržaja. Ovakva procedura poznata pod nazivom *asistencija korisnika* (RF), kombinovana sa neuralnim mrežama i/ili fazi sistemima, na efikasan način premošćava semantički jaz između obeležja niskog nivoa i obeležja visokog nivoa odnosno semantičkog opisa slike [5-16].

U ovom radu biće opisan CBIR sistem uz asistenciju korisnika. Strategija aktivnog učenja koristi i pozitivne i negativne primere da bi povećala uticaj korisnikove asistencije [5-12]. U predloženom sistemu uloga korisnika je iskorišćena ne samo za podešavanje vektora obeležja već i za adaptivnu klasterizaciju baza slika (ACID) koja je u skladu sa potrebama korisnika. Klasteri se memorišu i njihovi reprezentativni predstavnici se koriste u daljoj proceduri pretraživanja. Na ovaj način deo istorije prethodnih pretraga postaje deo procesa pretraživanja obezbeđujući brže i subjektivnije pretraživanje. Efikasnog predloženog ACID sistema je testirana na slikama iz MIT i Corel baze slika [17,18].

Rad je organizovan na sledeći način: Sekcija 2 sadrži kratak pregled trenutnog stanja u CBIR sistemima baziranih na asistenciji korisnika. Pregled ACID sistema predstavljen je u sekciji 2 a eksperimentalni rezultati u sekciji 3. Zaključne napomene su date u sekciji 4.

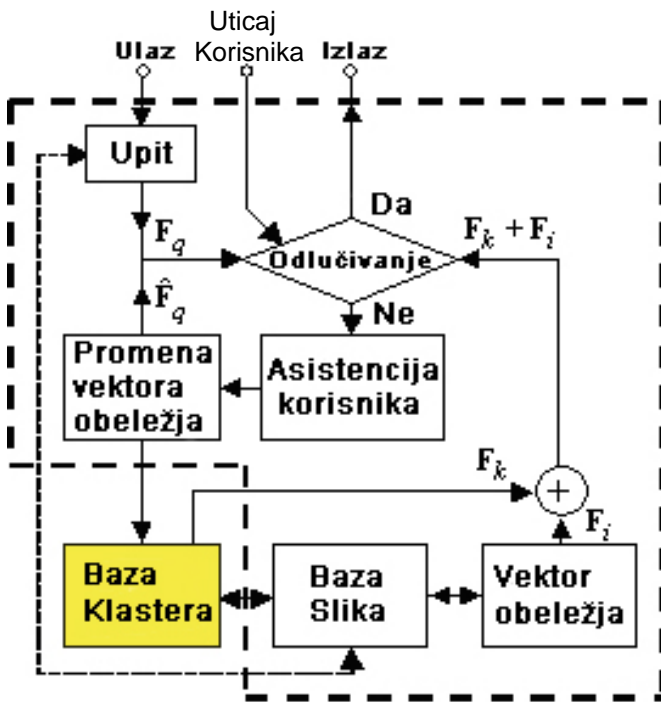
Semantički jaz u CBIR sistemima se može ublažiti uvođenjem težinskih koeficijenata, što je primenjeno u sistemima Photobook [1], QBIC [2], Virage [3], NETRA [4]. Uveliko se primenjuju dve osnovne RF strategije u CBIR sistemima: pomeranje upita [5] i uvođenje težinskih rastojanja [6]. U [7] je predstavljen *Probabilistic Feature Relevance Learning* (PFRL) metod u kome se određuje relevantnost pojedinih obeležja na osnovu povratne informacije dobijene od korisnika. Takođe predloženo je nekoliko metoda za dalje unapređenje RF strategije. U [8] je prezentovan sistem za pretraživanje baza slika sa automatskom modifikacijom upita. Adaptivan CBIR sistem prilagođen potrebama korisnika je opisan u [9]. U nedavno objavljenom radu [10] predstavljen je metod za pretraživanje slika na osnovu regiona, nazvan FRIP (*Finding Region In the Picture*) koji je baziran samo na obeležjima označenih regiona. U svim CBIR sistemima procedura pretrage je vremenski vrlo zahtevan proces, posebno ako se pretražuju velike baze slika. U našem prethodnom radu [11] opisan je

interaktivni metod za pretragu koji kombinuje i mogućnost nelinearnog modelovanja pomoću funkcija radialne osnove (RBF) [12], relevantnost pojedinih obeležja [7] kao i pomeranje upita. U ovom radu predstavljen je ACID sistem gde adaptivna klasterizacija slika iz baze doprinosi bržoj i subjektivnijoj pretrazi.

2. PREGLED ACID SISTEMA

Blok šema ACID sistema prikazana je na slici 1. Deo sistema označen isprekidanom linijom je identičan sistemu prezentovanom u [11]. Obojeni blok (donji levi ugao slike) označen sa "Baza klastera" je osnovni deo ACID sistema. Slike iz baze su indeksirane na odgovarajući način, npr. brojevima $i=1,2,\dots,I$. Iz svake slike je izdvojen vektor obeležja $\mathbf{F}_i = [F_{i1}, F_{i2}, \dots, F_{iJ}]$ sa J koordinata, kojim su opisana obeležja niskog nivoa (boja, tekstura i oblik). Da bi se sprečila dominacija obeležja sa velikom varijansom i osigurao jednak uticaj svih komponenti, svaka koordinata F_{ij} u vektoru obeležja je skalirana po kolonama kao što je opisano u [11]. Ovako normalizovani vektori obeležja se čuvaju i koriste za računanje rastojanja (Euklidskog ili nekog drugog) koje predstavlja meru sličnosti između upitne slike (opisane vektorom obeležja \mathbf{F}_q) i slika iz baze. Upitna slika može biti slika iz baze ili iz nekog drugog izvora.

Strategija pretrage zavisi od klastera i pruža dve mogućnosti: pretragu bez korišćenja klastera (WCS) i pretragu sa korišćenjem klastera (SWC).



Slika 1. Prikaz sistema za pretraživanje baza slika na osnovu adaptivne klasterizacije

2.1 Pretraga bez korišćenja klastera

Na početku slike nisu klasterizovane i inicijalna pretraga se obavlja u WCS modu. Prvi korak je čisto objektivni: slike iz baze se sortiraju u rastućem redosledu na osnovu rastojanja (Euklidskog ili nekog drugog) od upitne slike. U drugom koraku uvodimo RF. Od prvih B slika ($B=20$ u našem slučaju) koje su objektivno najbliže upitnoj, korisnik

prepoznaje $R \leq B$ slika i označava ih kao relevantne (subjektivno najbliže upitnoj). Preostalih $N = (B - R)$ slika se označavaju kao nerelevantne. Vektori obeležja relevantnih i nerelevantnih slika se koriste za modifikaciju vektora upitne slike [12], [11] na osnovu relacije:

$$\hat{\mathbf{F}}_q = \mathbf{F}_q + \alpha_R (\bar{\mathbf{F}}_R - \mathbf{F}_q) - \alpha_N (\bar{\mathbf{F}}_N - \mathbf{F}_q) \quad (1)$$

$\hat{\mathbf{F}}_q$ označava modifikovani upitni vektor \mathbf{F}_q , a $\bar{\mathbf{F}}_R$ i $\bar{\mathbf{F}}_N$ predstavljaju srednje vrednosti vektora obeležja relevantnih i nerelevantnih slika, respektivno. Pozitivnim konstantama $\alpha_R > \alpha_N$ određuje se stepen uticaja označenih slika na modifikovanje upitnog vektora. U našem radu korišćene su vrednosti $\alpha_R = 0.9$ i $\alpha_N = 0.1$. Modifikovani upitni vektor $\hat{\mathbf{F}}_q$ prilagođava parametre RBF neuralne mreže menjajući odgovarajuće Gausove funkcije G_{ij}

$$G_{ij} = \exp\left(-\frac{(F_{ij} - \hat{F}_{qj})^2}{2\sigma_{ij}^2}\right), i = 1, 2, \dots, I; j = 1, 2, \dots, J \quad (2)$$

Relevantnost pojedinih obeležja se kontroliše standardnom devijacijom σ_{ij} . Što je obeležje relevantnije, vrednost odgovarajuće standardne devijacije σ_{ij} je manja da bi se povećala osetljivost na razliku između upitne i slika u bazi. Suprotno, velike vrednosti standardne devijacije odgovaraju nerelevantnijim obeležjima jer je u tom slučaju G_{ij} približno jedan bez obzira na postojanje razlike između upitne i slika u bazi. U ovom radu parametar σ_{ij} je estimiran korišćenjem PFRL metoda [7].

Subjektivna sličnost između slika iz baze i modifikovanog upita određuje se na osnovu sume:

$$S_i(\mathbf{F}_i, \hat{\mathbf{F}}_q) = \sum_{j=1}^J G_{ij}, i = 1, 2, \dots, I \quad (3)$$

Vrednost S_i označava stepen sličnosti i -te slike iz baze sa upitnom (veća vrednost – veća subjektivna sličnost). Izračunate S_i se sortiraju u opadajući poredak i koriste u narednom koraku pretrage. Na osnovu sortiranih vrednosti S_i određuje se skup B najbližih slika i prikazuje korisniku. Korisnik označava relevantne slike čime utiče na promenu parametara RBF mreže. Proces se ponavlja dok korisnik ne bude zadovoljan rezultatima pretrage, a u slučaju divergencije proces se prekida.

2.2 Pretraga sa klasterima

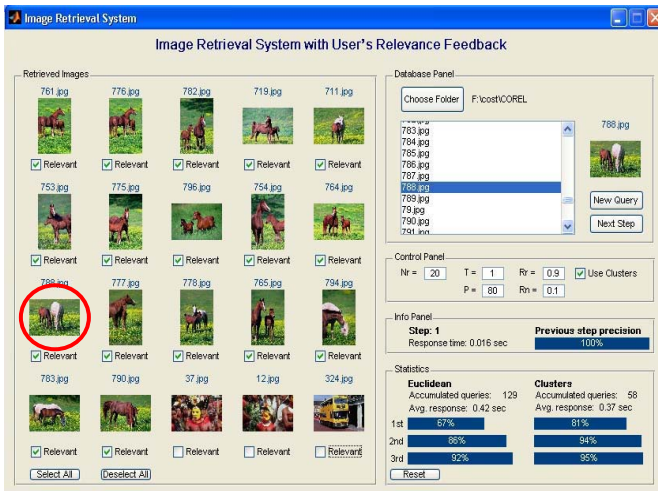
Osnovna ideja ACID sistema je uvođenje uticaja istorije prethodnih upita na proces pretrage. Skup relevantnih slika dobijenih u jednoj RF sesiji se pamti kao k -ti klaster sa $L_k=R(k)$ elemenata, $R(k)$ predstavlja broj relevantnih slika u k -toj iteraciji. Svaki klaster $k=1, 2, \dots, K$, je reprezentovan sa: vektorom obeležja poslednjeg modifikovanog upita, označen sa \mathbf{F}_k , odgovarajućim vektorom standardnih devijacija $\Sigma_k = \{\sigma_{kj}\}$, i brojem $S_k(\mathbf{F}_{Lk}, \hat{\mathbf{F}}_q)$ izračunatim za poslednju relevantnu sliku u sesiji prema formulama (2) i (3).

SWC mod rada podrazumeva da se u prvom prolazu upitna slika poredi sa reprezentima postojećih klastera,

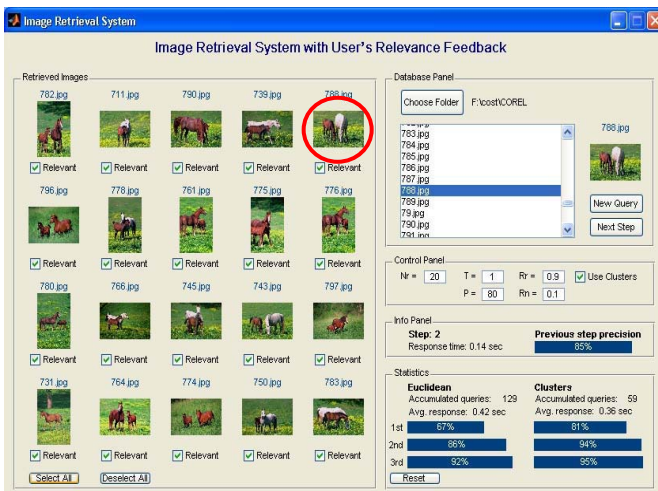
odnosno izračunava mera sličnosti $S_q(\mathbf{F}_k, \mathbf{F}_q)$. Ako je dobijena vrednost S_q veća od vrednosti S_k tada upit pripada k -tom klasteru. Ovakav zaključak može biti izveden za više klastera, pa uvodimo dodatno odlučivanje na bazi relativne sličnosti $RS_k = (S_q - S_k) / S_k, k = 1, 2, \dots, K$.



(a)



(b)



(c)

Slika 2. Primeri pretraga baze slika pomoću ACID sistema: (a) Inicijalna pretraga u WCS modu; (b) Prvi korak u SWC modu; (c) Drugi korak u SWC modu.

Klaster koji odgovara najvećoj vrednosti RS_k je "pobednik" i prvih $B/2$ slika pridruženih tom klasteru se prikazuje korisniku. Dodatnih $B/2$ slika se određuje na osnovu sortiranog Euklidskog rastojanja upitne i slika u bazi. Dalja pretraga se obavlja na način opisan u glavi 2.1, u WCS modu sa asistencijom korisnika kroz RBF neuralnu mrežu.

Klasteri se formiraju automatski, posle svake iteracije pretrage. Da bi se sprečila redundantost, novoformirani klasteri se poreda sa postojećim. Na osnovu rezultata eksperimenata odlučeno je da ako novi klaster sadrži više od 70% slika nekog postojećeg klastera, tada se novi klaster snima a postojeći briše iz memorije. Na ovaj način omogućeno je ubacivanje novih slika u klasteru u slučaju proširenja baze.

3. EKSPERIMENTALNI REZULTATI

Algoritam je testiran na referentnim MIT i Corel bazama slika [17,18]. Na slici 2 prikazana je jedna RF sesija. Može se primetiti da upitna slika (dodatno označena) može izgubiti prvu poziciju zbog pomeranja centara funkcija radialne osnove posle modifikacije upita.

Kao meru performanse sistema uvodimo odnos $RR=R/B$ broja subjektivno relevantnih slika (R) u poslednjem koraku pretrage i broja slika (B) koje se prikazuju korisniku. Bez primene ACID sistema odnos RR za 57 različitih upita iznosi 70.1% posle prvog koraka (objektivna mera sličnosti), posle drugog koraka 86.3% i posle trećeg 89.5% [11]. Primenom ACID sistema performanse se generalno povećavaju: procedura pretrage je brža za oko 20%, a preciznost (izražena odnosom RR) veća za oko 15%, u odnosu na rad bez ACID sistema. Dobijeni rezultati su uporedivi (ili bolji) sa do sada objavljenim rezultatima pretrage referentnih baza [5], [7] i [12]. Treba primetiti da poređenja rezultata imaju samo kvalitativno značenje pošto se procedure pretrage ne izvode u jednakim uslovima.

4. ZAKLJUČAK

Prezentovan je ACID sistem za pretraživanje baza slika na osnovu sadržaja sa asistencijom korisnika i adaptivnom klasterizacijom. Klasteri se formiraju automatski, posle svake procedure pretrage. Takođe, reprezentivni klastera su subjektivno orijentisani tako da je procedura pretrage brza, a semantički jaz manji. Šta više, interaktivno ažuriranje baze klastera pruža mogućnost neograničenog proširenja baze slika.

Zahvalnost: Istraživanja u ovom radu su delimično podržana od strane evropskog projekta COST Action 292 "Semantic Multimodal Analysis of Digital Media".

LITERATURA

- [1] A. Pentland, R. Picard, and S. Sclaroff, "Photobook: Tools for content-based manipulation of image databases", in *Proc SPIE Conf. Storage and Retrieval of Still Image and Video Databases II*, pp 34-47, San Jose, CA, USA, 1994.
- [2] M. Flickner, H. Sawhney, W. Niblack, J. Ashley, Q. Huang, and B. Dom, "Query by image and video content: The QBIC system", *IEEE Computer*, vol. 28, 1995.

- [3] J. R. Back, C. Fuller, A. Gupta, A. Hampapur, B. Horowitz, R. Jain, and C. F. Shu, "The virage image search engine: An open framework for image management", in *Proc. SPIE Conf. Storage and Retrieval for Still Image and Video Databases IV*, pp. 76-87, San Jose, CA, USA, February 1996.
- [4] W. Y. Ma and B. S. Manjunath, "NETRA: A toolbox for navigating large image databases", in *Proc. IEEE Int. Conf. on Image Processing*, 1997.
- [5] Y. Rui, T. Huang, and S. Mehrotra. Content-based image retrieval with relevance feedback in MARS", in *Proc. IEEE Int. Conf. on Image Processing*, pp 815–818, 1997.
- [6] Y. Rui and T. Huang, "Relevance feedback: A power tool for interactive content-based image retrieval", *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 8, pp 644–655, 1998.
- [7] J. Peng, B. Bhanu, and S. Qing, "Probabilistic feature relevance learning for content-based image retrieval", *Computer Vision and Image Understanding*, vol. 75, pp. 150–164, 1999.
- [8] G. Aggarwal, T. Ashwin, S. Ghosal, "An image retrieval system with automatic query modification", *IEEE Trans. Multimedia*, vol. 4, pp. 201-214, June 2002.
- [9] K.-M. Lee, and W. N. Street, "Cluster-driven refinement for content-based digital image retrieval", *IEEE Trans. Multimedia*, vol. 6, pp. 817-827, December 2004.
- [10] B. Ko, H. Byun, "FRIP: A region-based image retrieval tool using automatic image segmentation and stepwise Boolean AND matching", *IEEE Trans. Multimedia*, vol. 7, pp. 105-113, February 2005.
- [11] S. Čabarkapa, N. Kojić, V. Radosavljević, G. Zajić, Branimir Reljin, "Adaptive content-based image retrieval with relevance feedback", in *Proc. Conf. EUROCON-2005*, Vol. 1, pp. 147-150, Belgrade, 21-24 November 2005.
- [12] P. Muneesawang, and L. Guan, "An interactive approach for CBIR using a network of radial basis functions", *IEEE Trans. Multimedia*, vol. 6, pp. 703-716, October 2004.
- [13] G. Liu, "Techniques and data structures for efficient multimedia retrieval based on similarity", *IEEE Trans. Multimedia*, vol. 4, pp. 372-384, September 2002
- [14] J. Calic, N. Campbell, A. Calway, M. Mirmehdi, T. Burghardt, S. Hannuna, C. Kong, S. Porter, N. Canagarajah, and D. Bull, "Towards intelligent content based retrieval of wildlife videos", in *Proc. 6th Int. Workshop WIAMIS-05*, EFPL, Montreux, Switzerland, April 2005.
- [15] M. Clausen, and F. Kurth, "A unified approach to content-based and fault-tolerant music recognition", *IEEE Trans. Multimedia*, vol. 6, pp. 717-731, October 2000.
- [16] V. Đorđević, N. Reljin, and I. Reljin, "Identifying and retrieving of audio sequences by using wavelet descriptors and neural network with user's assistance", in *Proc. IEEE Conf. EUROCON-2005*, Vol. 1, pp. 167-170, Belgrade, 21-24 November 2005.
- [17] <http://vismod.media.mit.edu/pub/VisTex/>.
- [18] Corel Gallery Magic 65000 (1999), www.corel.com.

Abstract - A content-based image retrieval system where an active learning strategy is used to gain relevance feedback (RF) is described. User's feedback is used also for adaptive clustering of image database (ACID) according to particular subjective needs. Images labeled as relevant in last step of retrieving session are gathered in clusters, then memorized, and their representative members are used in further searching procedure instead of all images from database. In this way some history of previous retrieving is embedded into a searching process enabling faster and more subjective retrieval. Moreover, clusters are adaptively updated after each retrieving session, following actual user's needs. The efficiency of proposed ACID system is tested over images from MIT and Corel databases.

AN IMPROVED IMAGE RETRIEVAL SYSTEM BASED ON ADAPTIVE CLUSTERIZATION

Slobodan Čabarkapa, Nenad Kojić, Vladan Radosavljević ,
Goran Zajić, Branimir Reljin