

ADAPTIVNA NEURALNA MREŽA ZA PRETRAGU BAZE SLIKA

Nenad Kojić^{1,2}, Slobodan Čabarkapa^{1,2}, Goran Zajić², Branimir Reljin²

¹Viša ICT škola, Beograd, ²Elektrotehnički fakultet, Beograd

Sadržaj – U ovom radu predstaviće se postupak pretrage baza slika zasnovan na sadržaju istih uz interaktivno učešće korisnika, primenom neuralne mreže. Kao parametar sličnosti koriste se rastojanje vektora obeležja upitne slike i slika iz baze. Postupak se izvodi u dva koraka: objektivno - primenom euklidskog rastojanja i subjektivnog - na osnovu asistencije korisnika. Slike označene od strane korisnika dalje se koriste u modifikaciji vektora obeležja upitne slike i obučavanju neuralne mreže radikalne osnove. Postupak se završava kada je korisnik zadovoljan skupom ponuđenih slika. Potrebno je nekoliko iteracija da se dođe do željenog rezultata, što je potvrđeno kroz intezivne simulacije primenjene na MIT Media Laboratories bazi slika.

1. UVOD

Nagli rast multimedijalnih sadržaja koji je dodatno ubrzan razvojem interneta uslovljava potrebu za istraživanjem u oblasti obrade, prenosa, arhiviranja i pretraživanja velikih baza podataka. Pretraživanje na osnovu sadržaja postaje sve aktuelnije imajući u vidu primenu kako u profesionalne i naučne tako i u zabavne svrhe [1]-[11]. Takođe poseban akcenat stavlja se na pretraživanje multimedijalnih [12], video [13] i audio [14, 15] sekvenci.

Prvi problem u procesu pretrage predstavlja opisivanje velikog broja slika iz baze adekvatnim vektorom obeležja koji odgovara subjektivnom doživljaju krajnjeg korisnika. Nakon toga potrebno je pronaći meru sličnosti među vektorima obeležja koja odgovara meri sličnost vizuelnog sadržaja. Kako je ova dva uslova teško objediniti koristi se neuralna mreža zbog mogućnosti upotrebe pethodno stečenih znanja u procesu pretrage baze. Kao dodatni uticaj se uzima i asistencija korisnika, koji će na osnovu svoje vizuelne percepcije odabirom relevantnih slika, direktno modifikovati vektor obeležja upitne slike i izvršiti promenu parametara neuralne mreže.

U ovom radu predstaviće se postupak pretrage baze slika zasnovan na sadržaju uz asistenciju korisnika primenom neuralne mreže radikalne osnove. Iako je izbor i minimizacija vektora obeležja od presudnog značaja, glavni akcenat u našem istraživanju je postavljen na projektovanju neuralne mreže i uvođenje uticaja korisnika u cilju efikasnije pretrage baze slika.

Ovaj rad je organizovan kroz sledeća poglavlja: U poglavlju II je dat opis sistema pretrage i mehanizma odlučivanja. Rezultati simulacije su predstavljeni u poglavlju III. Poglavlje IV sadrži zaključke i dalje smernice u istraživanju.

2. OPIS SISTEMA

2.1 Pregled sistema

Pretraživanje baze slika sa uticajem korisnika je klasične strukture. Ovo podrazumeva da se vektor obeležja generiše za svaku od slika u bazi. Prvi korak je definisanje upitne slike iz postojeće baze. Dalja pretraga zasniva se na dva odvojena procesa. U prvom, koji je objektivni, vrši se poređenje

vektora obeležja upitne slike sa vektorima slika iz baze na osnovu euklidskog rastojanja. Nakon toga program nudi grupu slika koje su poređane u opadajućem redosledu u odnosu na upitnu sliku. Drugi korak podrazumeva uvođenje uticaj korisnika na poboljšanje strategije pretraživanja. Korisnik određuje broj slika, na primer C , koje očekuje da mu program prikaže kao najbliže upitnoj. Ove slike se prikazuju na ekranu nakon čega korisnik po svom subjektivnom mišljenju obeležava grupu slika najslabijih upitnoj slici. Na osnovu ovih vektora obeležja ovako označenih slika vrši se promena koeficijenata neuralne mreže radikalne osnove modifikujući poziciju centara i širine Gausovih krivi. Ovako promenjena RBF se koristi u sledećoj iteraciji subjektivne pretrage. Proces se zatim ponavlja sve dok korisnik ne bude zadovoljan ponuđenim rezultatima. U praksi se pokazalo da je nekoliko iteracija sasvim dovoljno sto je pokazano velikim brojem simulacija.

2.2 Detaljniji prikaz sistema

1) Faza predprocesiranja: U ovoj fazi potrebno je izvršiti indeksiranje svih slika u bazi brojevima $i=1,2,\dots,I$. Svakoj slici iz baze pridružen je vektor obeležja $F_i=[F_{i1}, F_{i2}, \dots, F_{iJ}]$ koji je opisuje boju i teksturu. Uticaj oblika se neće razmatrati u ovoj pretrazi. U prvom delu vektora obeležja su komponente dobijene na osnovu osam dominantnih boja i HSV prostora boja. Tekstura je opisana u drugom delu vektora obeležja pomoću koeficijenata dobijenih Gaborovom transformacijom [9], i obeležja co-occurrence matrice. Svi vektori obeležja su smeštene u odgovarajuću matricu obeležja $F=\{F_i\}=\{F(i,j)\}$, dimenzije $I \times J$.

Generalno posmatrano koeficijenti matrice obeležja $F(i,j)$ $i=1,2,\dots,I$ i $j=1,2,\dots,J$ mogu imati potpuno različite vrednosti u zavisnosti da li pripadaju elementima koji opisuju boju to jest teksturu. Komponente sa većom varijansom će biti dominantne u postupku računanja Euklidskog rastojanja i mogu izazvati maskirajući efekat. Da bi se sprečio uticaj obeležja sa velikom varijansom i istovremeno omogućio podjednak uticaj svih komponenti svi koeficijenti $F(i,j)$ u matrici obeležja se u svakoj koloni skaliraju težinskim faktorom $W1_j$

$$W1_j = \frac{1}{\text{mean}(F_j)} \log_2 \left(\text{std} \left(\frac{F_j}{\text{mean}(F_j)} \right) + 2 \right). \quad (1)$$

Termin $\text{mean}(F_j)$ je srednja vrednost svih elemenata j -te kolone matrice F , $\text{std}(\cdot)$ je standardna devijacija. Ovako normalizovana matrica, na vrednosti od nula do jedan, se snima i dalje koristi za izračunavanje Euklidskog rastojanja. Ovo rastojanje predstavlja meru sličnosti između upitne slike i slika iz baze. Ovakav vid skaliranja predložen je [5] u cilju pronalaženja najslabije slike. Razlike između objektivne i subjektivne sličnosti dve slike u ovom radu pokušaćemo da se reše primenom težinskog Euklidskog rastojanja, u prvoj iteraciji, i uticajem asistencije korisnika na rad neuralne mreže radikalne osnove. Primenom težinskih koeficijenata koji su u korelaciji sa odgovarajućim komponentama vektora

obeležja opisuje se uticaj pojedinih komponenata na meru sličnosti. Kao takav ovaj metod je primenjen u sistemima Photobook [1], QBIC [2], Virage [3], NETRA [4]. Bitan uticaj povratne sprege, ostvaruje se asistencijom korisnika koji definiše grupu sličnih slika zadatom upitu. Dve osnovne karakteristike mreža radijalne osnove u CBIR sistemima su primenjene i u ovom radu: pomeranje upitnog vektora [5] i promena širine Gausove krive [6,7]. Pomeranje upita vrši se sekvencijalnom modifikacijom koja se zasniva na približavanju regionima subjektivno relevantnih slika a istovremeno udaljavaju od regiona nerelevantnih. Glavni cilj promene širine je da se modifikuje upitni vektor obeležja koji će najbolje opisati sličnost između relevantnih slika i pojačati njihov uticaj u odnosu na objektivnu meru sličnosti (Euklidsko rastojanje). U [7] *Probabilistic Feature Relevance Learning* (PFRL) predstavljena je metoda koja određuje relevantnost određenih obeležja na osnovu asistencije korisnika. U našem istraživanju predlažemo novu metodu koja konbinuje nelinearnu promenu parametara funkcije radijalne osnove (RBF) [10], PFRL kao u [7] i pomeranje upitnog vektora unutar interaktivnog CBIR sistema.

2) Inicijalna pretraga: Na početku procedure pretrage korisnik bira određenu sliku a sistem automatski dohvata iz baze ranije izračunati njen vektor obeležja koji nazivamo *query feature vector*, \mathbf{F}_q . Nakon toga računa se objektivna mera sličnosti, primenom Euklidskog rastojanja, između zahtevane slike i slika iz baze:

$$d_i = d(\mathbf{F}_q, \mathbf{F}_i) = \sqrt{\sum_{j=1}^J (F_{qj} - F_{ij})^2}, \quad i = 1, 2, \dots, I \quad (2)$$

gde \mathbf{F}_q predstavlja vektor obeležja upitne slike, a \mathbf{F}_i vektor obeležja pridružen i-toj slici u bazi. Od korisnika se zahteva da definiše broj slika C , na osnovu čega će sistem slektovati i prikazati toliko najbližih po Eulidskom rastojanju.

3) Asistencija korisnika: Iz prvih C prikazanih slika korisnik prepoznaje i označava R slika kao relevantne koje su subjektivno najbližije upitnoj. Ostalih $N=C-R$ su automatski označene kao nerelevantne. Vektori obeležja ovih slika su iskorišćeni za modifikaciju upitnog vektora obeležja sa subjektivne tačke gledišta. Ova modifikacija, bazirana na asistenciji korisnika, je izvedena primenom standardne *Rocchio* jednačina [5]

$$\hat{\mathbf{F}}_q = \mathbf{F}_q + \alpha_R (\bar{\mathbf{F}}_R - \mathbf{F}_q) - \alpha_N (\bar{\mathbf{F}}_N - \mathbf{F}_q) \quad (3)$$

gde je \mathbf{F}_q prethodni vektor obeležja upitne slike a $\hat{\mathbf{F}}_q$ je modifikovni vektor, dok su $\bar{\mathbf{F}}_R$ i $\bar{\mathbf{F}}_N$ srednje vrednosti vektora obeležja relevantnih i nerelavantnih slika respektivno. Pozitivne konstante α_R i α_N određuju uticaj relevantnih i nerelevantnih slika na modifikaciju vektora obeležja upitne slike. Dobijeni vektor obeležja koristi se za unapređivanje procedure pretrage sa subjektivne tačke gledišta, kao što je ranije opisano. S obzirom da su relevantne slike najčešće u nekoj meri klasterizovane za razliku od nerelevantnih koje su heterogene, uticaj relevantnih slika je značajniji pri korekciji upitnog vektora. Iz tog razloga koristimo relaciju $\alpha_R \gg \alpha_N$.

U našem radu koristiće se 1D Gausova RBF gde je svaka i-ta komponenta vektora obeležja predstavljena sa

$$S_i(\mathbf{F}_i, \hat{\mathbf{F}}_q) = \sum_{j=1}^J \exp\left(-\frac{(F_{ij} - \hat{F}_{qj})^2}{2\sigma_j^2}\right), \quad i = 1, 2, \dots, I \quad (4)$$

Amplituda S_i predstavlja sličnost između i-tih komponenata vektora obeležja \mathbf{F}_i i modifikovanog upitnog vektora $\hat{\mathbf{F}}_q$, posle asistencije korisnika. Naravno, najveća sličnost se postiže pri $\mathbf{F}_i = \hat{\mathbf{F}}_q$.

Standarda devijacija, σ_j , određuje širinu Gausove krive, a u stvari reflektuje se na značaj pojedinačnih obeležja. Ako je obeležje više relevantno vrednost σ_j će biti mala i dozvoliće veliku osetljivost na bilo kakvu promenu razlike između upitne slike i slika iz baze. Nasuprot tome, velike vrednosti σ_j vode ka nerelevantnim obeležjima pošto je tada amplituda eksponencijalne funkcije (4) blizu 1 na osnovu razlike između upitne slike i slika iz baze.

U PFRL metodu [7] slike dobijene pretragom uz asistenciju korisnika, koriste se za izračunavanje značajnosti lokalnih obeležja. Mera značajnosti obeležja modifikovanog upitnog vektora $\hat{\mathbf{F}}_q$ može se tumačiti kao verovatnoća da vektori obeležja slika koje su označene kao relevantne odgovaraju modifikovanom upitnom vektoru

$$r_j(\mathbf{F}_R) = E[f | F_{Rj} = F_{qj}] \quad (5)$$

Parametar r_j izračunat iz jednačine (5) se potom koristi za izračunavanje težinskih koeficijenata $W2_j$, koji su povezani sa σ_j sledećom jednačinom

$$W2_j = \frac{e^{Kr_j}}{\frac{1}{J} \sum_{j=1}^J e^{Kr_j}} = \frac{1}{2\sigma_j^2} \quad (6)$$

Parametar K u jednačini (6) kontroliše uticaj parametra r_j na težinski koeficijent $W2_j$ i odabran je ekseprimentalno kao što je predloženo u [7].

Funkcije S_i izračunavaju se posle zamene jednačina (3) i (6) u (4) posle čega se koriste za određivanje sličnosti slika u novom subjektivnom procesu pretrage. Amplituda funkcije S_i je sortirana po opadajućem redosledu i prvih C slika će biti prikazano na ekranu. Nakon ovoga korisnik ponova označava R relevantnih i N nerelevantnih slika, čime koriguje RBF i nastavlja pretragu shodno svojim zahtevima. Proces se ponavlja sve dok korisnik ne bude zadovoljan rezultatima pretrage ili ako proces divergira što znači da se nijedna od subjektivno relevantnih slika ne nalazi unutar prvih C pozicija posle korekcije RBF (4). U drugom slučaju proces se vraća na prethodnu iteraciju.

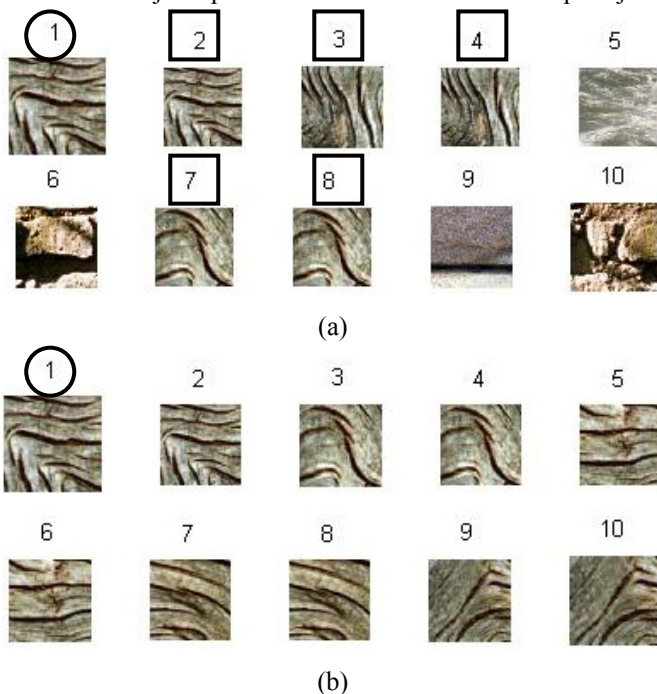
3. REZULTATI SIMULACIJE

U našem istraživanju korišćena je MIT Media Laboratories [16] baza slika. Iz potreba da se obezbede male slike radi bržeg procesiranja i velika baza slika, originalne slike dimenzija 512x512 piksela podeljene su na nepreklopajuće blokove diomenzija 128x128 piksela. Na ovaj način kreirano je I=3500 slika koje čine inicijalnu bazu slika. Kao što je pomenuto u sekciji 2.2.1 vektori obeležja koji su dodeljeni svakoj slici sadrže J=258 koordinata opisujući boju i teksturu slike. Pre pretrage, od strane korisnika, određen je broj C=10 slika koje će se prikazati kao najbliže upitnoj slici. Nakon mnogobrojnih simulacija ovde će biti prikazano samo nekoliko karakterističnih primera. U svim primerima prvi skup slika je dobijen posle prve iteracije primenom objektivne mere sličnosti dok drugi skup slika predstavlja

rezultat pretrage nakon primene asistencije korisnika. Prikazane slike su označene brojevima od 1 do 10 što odgovara njihovoj odaljenosti od upitne slike. Upitna slika je dodatno označena krugom, dok su slike koje su subjektivno izabrane kao relevantne označene pravougaonicima.

A. Primer 1.

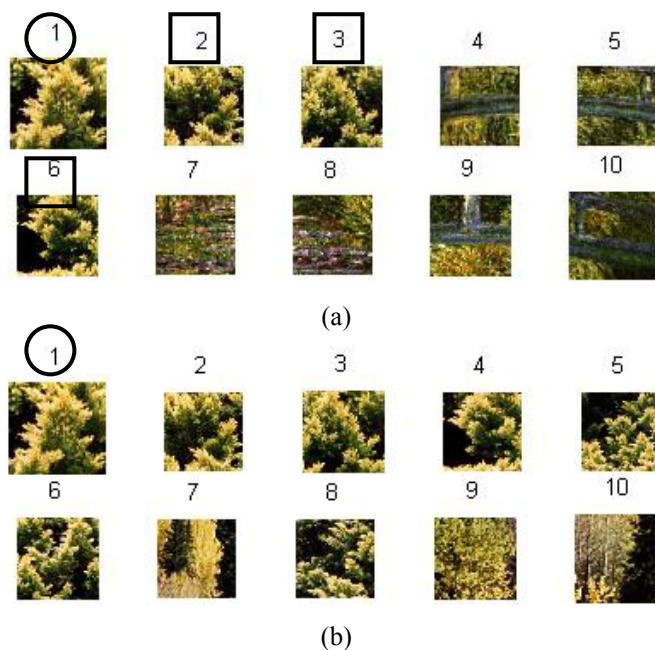
Na slici 1 prikazan je prvi primer koji ilustruje rezultate pretrage. Posle prve iteracije koja predstavlja samo objektivno merenje sličnosti bazirano na Euklidskom rastojanju na slici 1.a je prikazano prvih C=10 slika koje su najbliže upitnoj. Na prvom mestu se našla upitna slika, kao što se i očekivalo. Slike koje su direktno slične upitnoj (slike 2, 3 i 6) selektovane su od strane korisnika kao što je prikazano na slici 1.a. Na osnovu ovih slika upitni vektor je modifikovan i rezultati dobijeni nakon asistencije korisnika su prikazani na slici 1.b. Kao što se može videti upitna slika se zadržala na prvoj poziciji, slike na pozicijama 2 i 3 koje je korisnik odabrao zadržale svoje pozicije, dok je slika sa pozicije 6 prešla na poziciju 4. Kao što se može videti i ostale slike koje su prikazane na slici 1.b slične su upitnoj.



Slika 1. Slike dobijene nakon primene objektivne mere sličnosti Euklidovim rastojanjem (a), prikaz slika dobijenih primenom neuralne mreže uz asistenciju korisnika (b).

B. Primer 2.

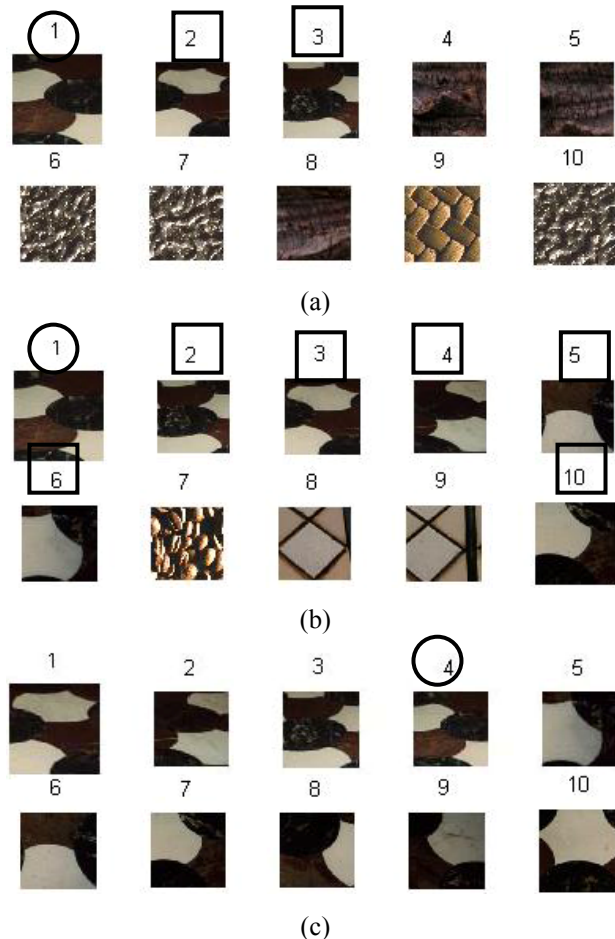
Drugi primer prikazan na slici 2.a. prikazuje da objektivno merenje sličnosti bazirano na Euklidskom rastojanju nije dovoljno za rangiranje i pretragu slika. Posle prve iteracije koja je prikazana na slici 2.a upitna slika se nalazi na prvom mestu. Korisnik odabire slike koje su nasličnije upitnoj (pozicije 2, 3, 4, 7 i 8), dok ostale slike koje se prikazuju kao najbliže subjektivno su vrlo različite od upitne. Nakon asistencije korisnika svih deset prikazanih slika su subjektivno slične (slika 2.b) ali su slike sa pozicija 3 i 4 na slici 2.a izgubile svoje mesto pri novom rangiranju najbližih slika zbog pomeranja centra pri modifikaciji upitnog vektora. Možemo videti da su sve slike na slici 2.b slične upitnoj.



Slika 2. Slike vraćene u prvom koraku primenom Euklidskog rastojanja (a), slike vraćene u drugoj iteraciji nakon asistencije korisnika (b).

C. Primer 3.

Treći primer prikazan na slici 3. naglašava značaj asistencije korisnika. U ovom primeru su neophodne dve asistencije korisnika (slike 2.b i 2.c) da bi svih deset slika postale vrlo slične upitnoj. Vidimo da je upitna slika izgubila prvu poziciju i pomerila se na četvrtu poziciju što je posledica modifikaciji upitnog vektora.



Slika 3. Slike pronađene na osnovu primene Euklidskog rastojanja (a), grupa slika koja je vraćena nakon asistencije korisnika u prvoj (b) i drugoj iteraciji (c).

D. Performanse sistema

Performanse sistema su određene srednjom efikasnošću pretrage

$$RR = \frac{R}{C} \times 100\% \quad (7)$$

U izrazu (7) R predstavlja broj relevantnih slika izabranih od strane korisnika, C ($C=10$) predstavlja broj slika na koje se primenjuje asistencija korisnika. Vrednost RR predstavlja srednju vrednost na uzorku od 60 različitih upitnih slika I iznosi za prvu iteraciju (samo objektivno merenje) 70.5%, za drugu iteraciju 87.2% i posle treće iteracije 89.8%. Naši rezultati su upoređljivi sa rezultatima pri pretraživanju tekstura objavljenim u radovima [5], [7] i [10]. Napominjemo da ova poređenja imaju samo kvalitativno značenje pošto procedure pretrage nisu izvođene pod istim uslovima.

4. ZAKLJUČAK

Problem pretraživanja audio, video i multimedijalnog sadržaja predstavlja pronalaženje odgovarajućeg algoritma koji sa jedne strane definiše kombinaciju različitih obeležja koja opisuju sadržaj materijala i vrše njegovo indeksiranje a sa druge strane obezbeđuju efikasan sistem prepoznavanja sličnosti u bazama podataka. Procedura pretrage može se unaprediti korišćenjem neuralnih mreža i ljudskom asistencijom. U našem istraživanju glavna pažnja je posvećena promeni stanja i parametara neuralne mreže pod uticajem korisnika, kao jednog od efikasnih alata za pretraživanje. Naši inicijalni rezultati su poredljivi sa rezultatima koji su već objavljeni u literaturi.

Dalja istraživanja biće usmerena u pravcu unapređivanja procedure pretrage u smislu redukcije broja komponenata vektora obeležja i na korišćenje semantičkih deskriptora.

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, 28(9), 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, no. 5, 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, no. 1/2, pp 150-164, 1999.
- [8] R. Brunelli, and O. Mich, "Image retrieval by examples", IEEE Trans. Multimedia, vol. 2, no. 3, pp. 164-171, Sept. 2000
- [9] D. Zhang, A. Wong, M. Indrawan, and G. Lu, "Content-based image retrieval using Gabor texture features", Monash University Churchill, Victoria, Australia
- [10] P. Muneesawang, and L. Guan, "An interactive approach for CBIR using a network of radial basis functions", IEEE Trans. Multimedia, vol. 6, no. 5, pp. 703-716, Oct. 2004
- [11] K.-M. Lee, and W. N. Street, "Cluster-driven refinement for contentbased digital image retrieval", IEEE Trans. Multimedia, vol. 6, no.6, pp. 817-827, Dec. 2004
- [12] G. Liu, "Techniques and data structures for efficient multimedia retrieval based on similarity", IEEE Trans. Multimedia, vol. 4, no. 3, pp. 372-384, Sept. 2002
- [13] 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.
- [14] M. Clausen, and F. Kurth, "A unified approach to content-based and fault-tolerant music recognition", IEEE Trans. Multimedia, vol. 6, no. 5, pp. 717-731, Oct. 2000
- [15] V. Đorđević, N. Reljin, and I. Reljin, "Identifying and retrieving of audio sequences by using wavelet descriptors and neural network with user's assistance", submitted to Conf. EUROCON-2005, Belgrade, 21-24 Nov. 2005
- [16] <http://vismod.media.mit.edu/pub/VisTex/>

Abstract – Retrieval of images, based on similarities between feature vectors of querying image and those from database, is considered. The searching procedure was performed through the two basic steps: an objective one, based on the Euclidean distances and a subjective one based on the user's relevance feedback. Images recognized from user as the best matched to a query are labeled and used for updating the query feature vector through a RBF (radial basis function) neural network. The searching process is repeated from such subjectively refined feature vectors. In practice, several iterative steps are sufficient, as confirmed by intensive simulations.

ADAPTIVE NEURAL NETWORK FOR IMAGE RETRIEVAL

Nenad Kojić, Slobodan Čabarkapa,
Goran Zajić, Branimir Reljin