

Statistički pristup detekciji lica i lokalizacija karakterističnih tačaka lica

Cilj rada je detekcija lica i lokalizacija karakterističnih tačaka na licu (koordinatne očiju, nosa i usta). U radu su korišćene slike lica iz frontalne perspektive u boji. Detekcija lica je zasnovana na Bajesovom klasifikatoru: izvršava se klasifikacija piksela slike na dve klase (koža i pozadina), pri čemu se pretpostavlja da je najveća oblast grupisanih piksela u boji kože predstavlja lice. Funkcija gustine verovatnoće, koja se koristi za Bajesov klasifikator, je aproksimirana preko kNN algoritma na osnovu uzoraka boja skupljenih iz FERET baze podataka. Lokalizacija karakterističnih tačaka je rađena preko horizontalne i vertikalne projekcije gradijenta slike. Detekcija kože je testirana na istim uzorcima od kojih je formiran klasifikator, dok je lokalizacija karakterističnih tačaka testirana u odnosu na tačne koordinate, prema podacima iz baze. Primena ovog algoritma je raznovrsna i može se naći u algoritmima za prepoznavanje lica, autofocus algoritmima, video nadzoru itd.

Uvod

U ovom radu su posmatrana dva problema. Problem detekcije lica na osnovu statističke analize boja na slici i problem lokalizacije karakterističnih tačaka (očiju, nosa, usta). Detekcija se vrši na slikama u boji, gde je lice prikazano iz frontalne perspektive. Korišćene su slike iz FERET (The Face Recognition Technology) baze podataka. Ova baza pored velikog broja slika iz više različitih perspektiva, sadrži još podatke o koordinatama očiju, nosa, usta, što je veoma korisno prilikom obrade rezultata.

Problem 1: detekcija lica. Detekcija lica predstavlja izdvajanje regije lica na nekom snimku ili slici. Da bi se došlo do što kvalitetnijih rezultata moraju se u obzir uzeti faktori kao što su pozicija glave (od frontalne perspektive do mnogih drugih položaja lica), orijentisanost (glava naklonjena levo, desno), udaljenost od fotoaparata, kontrast i pozadina. Zbog toga su mnogi sistemi ograničeni po pitanju podataka koji se obrađuju (frontalne slike, slike u boji, one na kojima je samo jedna osoba). Do sada je razvijen veliki broj metoda detekcije: pristup preko neuronskih mreža (Rowley *et al.* 1998), teorija informisanosti (Moghaddam i Pentland 1995; Colmenarez i Huang 1997), metoda grupisanja na osnovu modela (Govindaraju 1996), metoda klasifikacije boja (Wang and Chang 1997). Robust Real-time Object Detection (Viola i Jones 2001) je dosada najkorišćeniji algoritam za detekciju lica. Detekcija lica ima široku primenu (video nadzor, praćenje lica, prepoznavanje lica, autofocus algoritmi, automatski sistemi za kontrolu pristupa, obrada slike u biometričkim sistemima, prikaz lica u 3D sistemu, prepoznavanje izraza lica itd.).

Detekcija oblasti kože se vrši na osnovu klasifikacije boja na klasu boja pozadine i klasu boja kože. Metoda klasifikacije na osnovu boja zahteva prethodnu klasifikaciju boja u dve klase. Korišćen je Bajesov klasifikator, određen preko funkcija gustine verovatnoće raspodele.

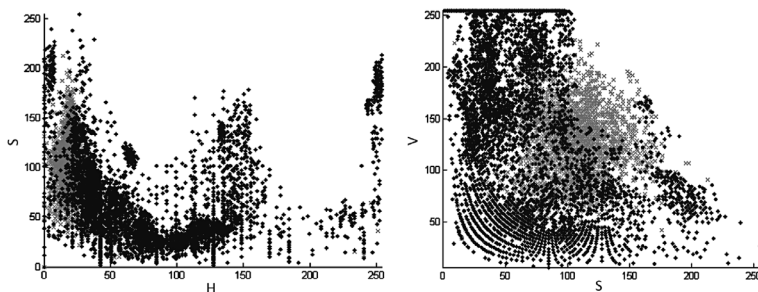
Problem 2: detekcija karakterističnih tačaka. Detekcija karakterističnih tačaka lica predstavlja izdvajanje regija lica (očiju, nosa, usta). Lokalizacijom navedene tri regije mogu se lako utvrditi i ostale regije lica (čelo, jagodice, obrazi, brada). Algoritam za lokalizaciju karakterističnih tačaka zasnovan je na ideji koja je iskorišćena za lokalizaciju očiju (Peng *et al.* 2005). Ideja je proširena na određivanje ostalih tačaka. Analizom ivica na slici su utvrđene te karakteristične tačke.

Natalija Todorčević (1994), Kragujevac, Ivana Cankara 11, učenica 2. razreda Prve kragujevačke gimnazije

MENTORI:

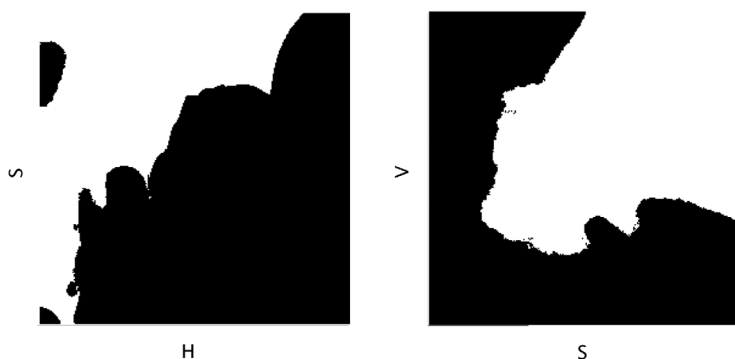
Marko Bežulj, ISP

Ivan Razumenić, student ETF-a u Beogradu



Slika 1. Projekcija uzoraka kože (sive tačke) i pozadine (crne tačke) iz HSV prostora na HS prostor (levo) i SV prostor (desno)

Figure 1. Projection of skin samples (gray dots) and background samples (black dots) from HSV space to HS space (left) and SV space (right)



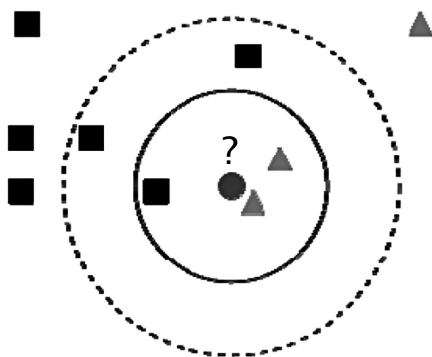
Slika 2. Prikaz klasifikatora u HS i SV sistemu. Bela boja označava kožu, crna pozadinu

Figure 2. Display of classifiers in HS and SV systems. White color represents skin, black – background

Detekcija lica

Razvijeni algoritam za detekciju lica zasnovan je na pretpostavci da će lice biti najveći povezani objekat na slici u boji kože. Stoga je prvi korak formiranje klasifikatora boja na osnovu koga će se svaki piksel sa slike okarakterisati kao koža ili pozadina, tj. šum. Obrada slike i skupljanje podataka je vršeno u HSV (Hue, Saturation, Value) sistemu boja. Prema Zarit *et al.* (1999) u HSV sistemu se dobija najveća stopa diskriminacije između klase boja kože i klase boja pozadine. Radi formiranja klasifikatora uzeti su uzorci sa slike iz baze. Označavana su polja na slikama i podaci svakog piksela (vrednosti za H, S i V) su smešteni u grupu uzoraka boja kože ili šuma u zavisnosti koju oblast sa slike dato polje obuhvata (slika 1, crna boja predstavlja uzorke boja šuma, siva uzorke boja kože). Na slici 1 se vidi da su posmatrane klase nelinearno separabilne u HS i SV prostorima.

Projektovana su dva Bajesova klasifikatora u dve projekcije HSV koordinatnog sistema, HS i SV (slika 2). Prilikom detekcije kažemo da piksel pripada licu ukoliko se boja tog piksela klasifikuje kao boja kože i prema jednom i prema drugom klasifikatoru.



Slika 3. Primer za k-NN klasifikaciju

Figure 3. Example of k-NN classification

Da bi se odredio Bajesov klasifikator potrebno je poznavati gustinu raspodele verovatnoće da određena boja pripada boji kože (Fukunaga 1990). Ne postoji tačan obrazac za određivanje vrednosti funkcije gustine verovatnoće, ali jedan od načina aproksimacije jeste preko k-NN algoritma (od engl. k-Nearest Neighbor). K-NN algoritam određuje najmanji poluprečnik opisan oko zadate tačke, takav da opisana



Slika 4. Morfološke operacije, s leva na desno: početna binarna slika, dilatacija, otvaranje slike, popunjavanje rupa unutar objekta, izdvojen najveći objekat

Figure 4. Morphological operations, from left to right: the initial binary images, dilation, opening images, filled holes in the objects, the largest object in the picture

kružnica sa centrom u toj tački obuhvata tačno k tačaka u datoj ravni (slika 3).

Za svaki piksel uzoraka koji predstavlja određenu boju u HS i SV (slika 1) izvršava se k -NN algoritam i dobijaju se po dva poluprečnika, jedan opisuje kružnicu koja obuhvata k tačaka boja pozadine, a drugi k tačaka boja kože. Nije vršeno istraživanje koji je optimalni parametar k , već je heurističkom metodom za vrednost parametara uzeto 30. Na osnovu toga određene su dve funkcije gustine verovatnoće čije su vrednosti obrnuto srazmerne tim poluprečnicima. Upoređivanjem vrednosti tih funkcija boja je okarakterisana kao boja pozadine ili kože:

$$\begin{aligned} -\ln F_1(x) + \ln F_2(x) < 0 &\Rightarrow x \in \omega_1 \\ -\ln F_1(x) + \ln F_2(x) > 0 &\Rightarrow x \in \omega_2. \end{aligned}$$

Grafički prikaz rešenja ovih nejednačina za predstavljene problem dat je na slici 2.

Klasifikacijom svih piksela sa slike dobija se binarna slika (slika 4). Na slici se javlja više objekata

tj. međusobno nepovezanih oblasti. Radi otklanjanja malih i spajanja bliskih objekata na slici izvršavaju se morfološke operacije nad slikom, dilatacija i erozija (slika 4). Strukturni element pri eroziji i dilataciji koji se koristi u radu je oblika diska. Nakon morfoloških operacija se određuju površine svih objekata i ostavlja se onaj najveće površine za koji se smatra da predstavlja lice (slika 4).

Detekcija karakterističnih tačaka

Nakon detekcije lica izdvaja se regija slike kojoj pripada površina detektovana kao lice. Da bi se fokusirali na ivice koje se javljaju na licu, a ne na pozadini, pozadina je maskirana (slika 5). Time su ublažene ivice koje se javljaju na pozadini, a radi eliminacije nebitnih ivica na licu slika se filtrira Gausovim filtrom (niskofrekventni filter, slika 6). Nakon toga pri određivanju gradijenta dobijaju se samo ivice oko



Slika 5. Crno bela slika, maska, maskirana slika

Figure 5.
Black and white image,
mask, masked image



Slika 6. Zamućena slika

Figure 6. Blurred image

lica, očiju, usta, nosa i neke oštrije ivice lica. Korisćena je crno-bela slika (slika 5). Filtriranje se izvršava konvolucijom slike sa matricom B :

$$B = \frac{1}{159} \cdot \begin{bmatrix} 2 & 4 & 5 & 4 & 2 \\ 4 & 9 & 12 & 9 & 4 \\ 5 & 12 & 15 & 12 & 5 \\ 4 & 9 & 12 & 9 & 4 \\ 2 & 4 & 5 & 4 & 2 \end{bmatrix} \cdot I$$

Konvolucija predstavlja local neighborhood funkciju. Local neighborhood funkcije (sliding window ili spatial filters) se intenzivno koriste u procesiranju slike u sistemima kompjuterskog vida. Primjenjuju se

za piksel na određenoj lokaciji, dok izlaz zavisi od njemu susjednih piksela. Funkcija se primenjuje za svaki piksel slike. Na taj način se dobija mutna slika (slika 6).

Detektovanje ivica se može izvršiti prostim diferenciranjem, međutim uzimajući u obzir činjenicu da je slika uvek u izvesnom stepenu zahvaćena šumom, to nije moguće. Ivice su određivane preko gradijenta slike uz pomoć Sobel operatora. Suština ovog operatora se sastoji u određivanju nivoa odstupanja vrednosti nekog piksela od njemu susjednih piksela. To se određuje preko dve kernel matrice (G_x i G_y), uz pomoć jedne se dolazi do vertikalnog, tj. uz pomoć druge do horizontalnog gradijenta slike.

$$G_x = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix}, G_y = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix}$$

Kombinacijom horizontalnog i vertikalnog gradijenta, po formuli $G = \sqrt{G_x^2 + G_y^2}$ dobija se gradijent slike (slika 7). Vertikalna i horizontalna projekcija predstavljaju sabiranje piksela po horizontalnoj, odnosno vertikalnoj osi. Pre određivanja koordinata vrednosti horizontalne i vertikalne projekcije su filtrirane median filterom. Zbog ivica same slike, granične vrednosti grafika imaju veoma velike vrednosti i te vrednosti se moraju zanemariti („preskočiti”) (slika 8a)

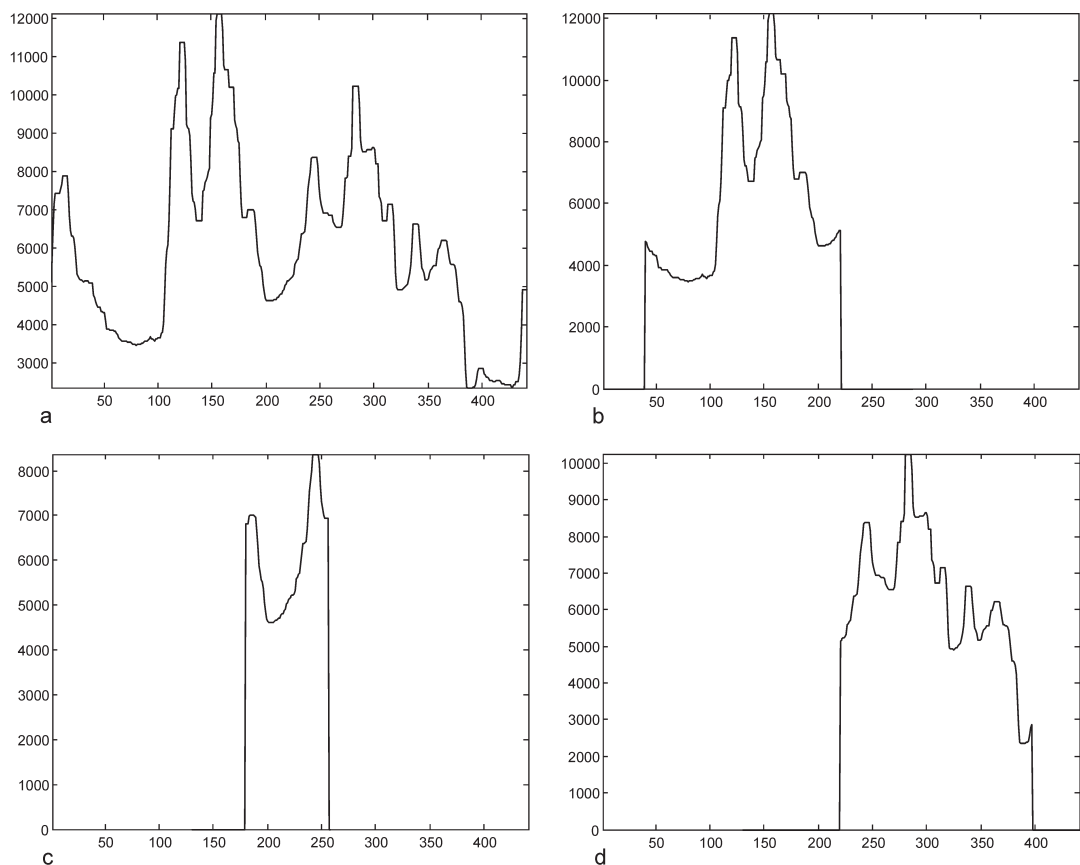
Kao što je rečeno, cilj detekcije karakterističnih tačaka lica jeste da odredi koordinate:

- očiju: $(x_{\text{lefteye}}, y_{\text{eye}})$,
desno oko $(x_{\text{righteye}}, y_{\text{eye}})$
- nosa $(x_{\text{nosemouth}}, y_{\text{nose}})$
- usta $(x_{\text{nosemouth}}, y_{\text{mouth}})$



Slika 7. Horizontalni gradijent, vertikalni gradijent i gradijent slike

Figure 7. Horizontal gradient, vertical gradient and gradient of the picture



Slika 8. a – Horizontalna projekcija; b – oblast očiju; c – oblast nosa; d – oblast usta

Figure 8. a – Horizontal projection; b – eyes area; c – area of the nose; d – area of the mouth

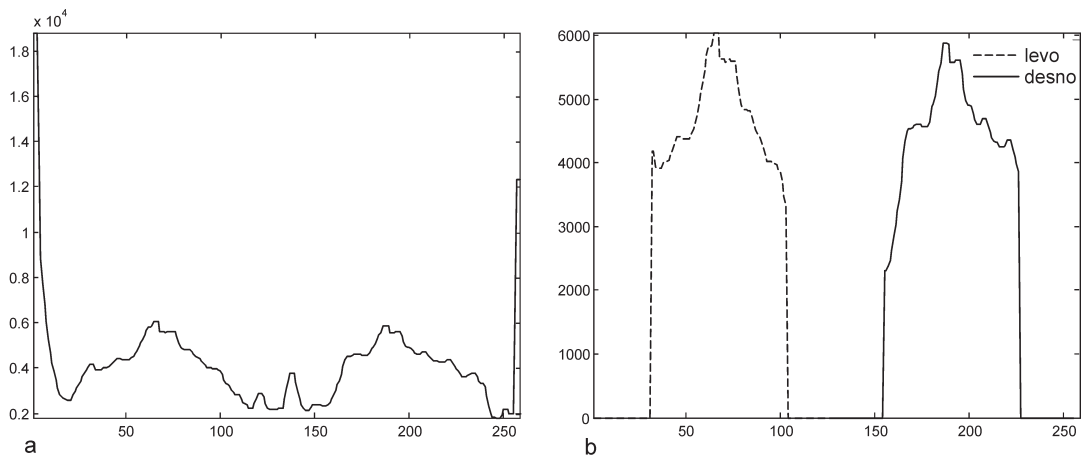
Te koordinate se određuju na sledeći način:

- y_{eye} se računa kao lokacija maksimuma na gornjoj polovini grafika horizontalne projekcije (što predstavlja horizontalnu projekciju predela lica gde se pretpostavlja da se nalaze oči) – slika 8b.
- y_{mouth} se računa kao lokacija maksimuma na donjoj polovini grafika horizontalne projekcije (horizontalna projekciju predela lica gde se pretpostavlja da su usta) – slika 8d).
- y_{nose} se računa kao lokacija maksimuma između dva lokalna minimuma između y_{eye} i y_{mouth} . Ovakav pristup je uveden, jer često slike imaju dosta „lažnih” maksimuma u predelu očiju – slika 8c

- $x_{lefteye}$ se računa kao lokacija maksimuma na levoj polovini grafika vertikalne projekcije (što predstavlja vertikalnu projekciju predela gde se pretpostavlja da se nalazi levo oko) – slika 9b
- $x_{righteye}$ se računa kao lokacija maksimuma na desnoj polovini grafika vertikalne projekcije (vertikalna projekcija predela lica gde se pretpostavlja da se nalazi desno oko) – slika 9b.

Smatra se da je horizontalna koordinata nosa i usta tačno na sredini između očiju, odnosno:

$$x_{nosemouth} = \frac{x_{lefteye} + x_{righteye}}{2}.$$



Slika 9. a – Vertikalna projekcija; b – vertikalna projekcija oblasti levog i desnog oka

Figure 9. a – Vertical projection; b – vertical projection area of the left and right eye

Rezultati

Formiranje klasifikatora i detekcija

Klasifikator za segmentaciju kože formiran je na osnovu 10% od ukupnog broja prikupljenih uzoraka (broj uzoraka iz klase boje kože oko $13 \cdot 10^6$, broj uzoraka boje pozadine oko $7 \cdot 10^6$), a testiran na svim uzorcima. Smatrano je da je regija kože na slikama u proseku bila 40×40 piksela, dok je regija pozadine bila oko 100×100 piksela. Tačnost projektovanog klasifikatora predstavlja procenat uzoraka boja koje su tačno klasifikovao pri klasifikaciji boja pozadine, odnosno pri klasifikaciji klase boja kože (tabela 1). Na osnovu procenata uzoraka boja koje su iskorišćene za formiranje klasifikatora, rezultati se mogu smatrati zadovoljavajućim.

Tabela 1. Preciznost klasifikatora prilikom klasifikacije

	Boja pozadine	Boja kože
Preciznost	79%	95%

Detekcija karakterističnih tačaka

Za svaku sliku u bazi postoje podaci sa tačnim koordinatama očiju, nosa i usta, koje su korišćene za testiranje algoritma za određivanje tačaka. Normirana

greška je definisana kao razdaljina između tačnih i merenih koordinata normirana sa tačnim rastojanjem između očiju (tabela 2).

Tabela 2. Srednje vrednosti normiranih grešaka

	levo oko	desno oko	nos	usta
y-koordinata 10%	10%	28%	22%	
x-koordinata 9%	11%	8%	8%	

Normiranje se vrši radi mogućnosti upoređivanja rezultata sa dve različite slike kod kojih lice nije iste dimenzije, pa samim tim i iste vrednosti razdaljine koordinata na različitim slikama ne predstavljaju istu grešku. U tabeli 3 prikazano je u koliko merenja je normirana greška procentualno ispod praga od 10, 20, odnosno 30 procenata, dok tabela 4 prikazuje medijan, maksimalnu i minimalnu vrednost normirane greške. Iz tabela se može uočiti da je odstupanje y-koordinate nosa od tačne vrednosti najveće, jer y-koordinata zavisi od prethodno utvrđenih koordinata očiju i usta.

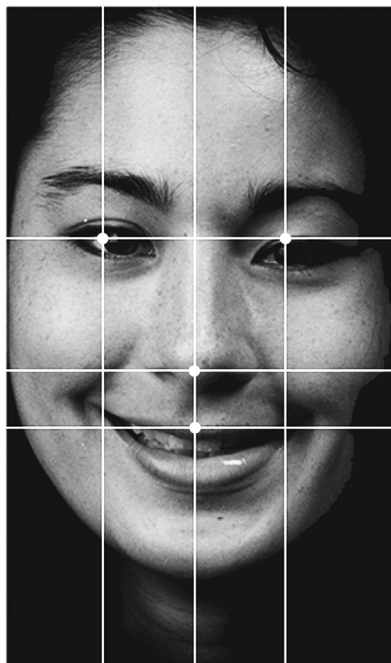
Na slici 10 može se videti fotografija na kojoj je detektovano lice i karakteristične tačke. Bele linije predstavljaju mesta gde su određeni maksimumi horizontalne i vertikalne projekcije, a tačke određene položaje očiju, nosa i usta.

Tabela 3. Udeo broja vrednosti za koje su normirane greške za x i y koordinate manje od 10, 20 i 30 procenata

	< 10%		< 20%		< 30%	
	x	y	x	y	x	
Levo oko	78%	80%	88%	87%	90%	90%
Desno oko	68%	78%	86%	87%	87%	94%
Nos	80%	5%	83%	44%	96%	66%
Usta	73%	50%	89%	56%	96%	61%

Tabela 4. Medijan, minimalna i maksimalna vrednost normirane greške

	Levo oko		Desno oko		Nos		Usta	
	x	y	x	y	x	y	x	
Maksimum	77%	152%	59%	153%	57%	163%	53%	115%
Minimum	0%	0%	0%	0%	0%	1%	0%	0%
Medijan	6%	4%	6%	5%	4%	22%	7%	10%



Slika 10. Detektovano lice i karakteristične tačke

Figure 10. Detected face and the characteristic points

Zaključak

Metod statističke detekcije lica na osnovu boja se pokazao kao dobar način za detekciju lica. Međutim ukoliko je boja pozadine, kose ili odeće slična boji kože, ako se pored kože lica vide na primer i ramena, tada se i oblast ramena detektuje kao lice, što kasnije, prilikom lokalizacije karakterističnih tačaka izaziva veliku grešku. Greška prilikom lokalizacije karakterističnih tačaka se uglavnom javlja kao posledica pogrešno detektovanog lica.

U daljem radu bi se mogao primeniti neki drugi algoritam za detekciju. Takođe se može unaprediti statistički pristup detekciji tako što bi se projektovao trodimenzionalni HSV klasifikator, ili bi se vršila napredna redukcija dimenzija iz HSV u neki novi dvo-dimenzioni ili jednodimenzioni prostor.

Literatura

Colmenarez A., Huang T.1997 Face detection with information-based maximum discrimination. U *Proc. CVPR, Puerto Rico*, str. 782

Govindaraju V. 1996. Locating human faces in photographs. *International Journal of Computer Vision*, **19** (2): 129.

Fukunaga K. 1990. *Introduction Statistical Pattern Recognition*. Academic Press

Moghaddam B., Pentland A. 1995. Maximum likelihood detection of faces and hands. In *Proc. Workshop on Automatic Face and Gesture Recognition, Zurich*, str. 122-128

Oliver N., Pentland A., Berard F. 1997. Lafter: Lips and face real time tracker. U *Proc. CVPR*, str. 123.

Peng K., Chen L., Ruan S., Kukharev G. 2005. A Robust and Efficient Algorithm for Eye Detection on Gray Intensity Face. *Proc. of 3rd International Conference on Advances in Pattern Recognition, LNCS*, 3687: 302.

Rowley H. A., Baluja S., Kanade T. 1998. Neural network-based face detection. *Trans. on PAMI*, **20** (1): 23.

Viola P., Jones M. 2004. Robust real-time object detection. *2nd Int. Workshop Stat. Comput. Theories Vis., Vancouver, Canada*

Wang H., Chang S-F. 1997. A highly efficient system for automatic face region detection in mpeg video. *Trans. Circuits and systems for video technology*, **7** (4): 615.

<http://suraj.lums.edu.pk/~ccs504m04/csvt99.pdf>

Zarit B. D., Super B. J., Quek F. K. H. 1999. Comparison of five color models in skin pixel classification. U *ICCV'99 Int'l Workshop on recognition, analysis and tracking of faces and gestures in Real-Time systems*, str. 58.

Natalija Todorčević

Statistical Approach for Face Detection Algorithm and Localization of Characteristic Face Points Using Image Gradient Projections

This paper shows implementation of statistical approach for single face detection algorithm and algorithm for finding characteristic points (nose, eyes and mouth) on frontal face image using horizontal and vertical image projection. Face detection is based on Bayes classifier, it classify image pixels into two classes (skin and noise) and then biggest suggest that face is biggest connected cloud of skin classified pixels. Density probability functions, used for Bayes classifier, are approximated with KNN algorithm on data collected from FERET image database. Skin detection is tested versus same collected data, while algorithm for finding characteristic points is tested versus their known coordinates, available in used database. Purpose of these algorithms is at face recognition systems. 