

SVEUČILIŠTE U ZAGREBU
FAKULTET ELEKTROTEHNIKE I RAČUNARSTVA

DIPLOMSKI RAD br. 1002

AUTOMATSKO RASPOZNAVANJE LICA

Ivan Renić

Zagreb, veljača 2015.

Sadržaj

UVOD	1
PREGLED PODRUČJA	2
1. METODE	6
1.1. Vektorska udaljenost između značajki.....	6
1.2. Shape unit	7
1.3. Normalizirana križna korelacija.....	8
1.4. Svojstvena lica.....	10
1.4.1. Računanje svojstvenog lica	13
1.5. Local Binary Patterns	15
2. BAZE PODATAKA	20
2.1. Face Recognition Grand Challenge	20
2.2. Multi-PIE.....	20
2.3. LFW.....	21
2.4. Plastic Surgery Face Database.....	22
2.5. YouTube Makeup Dataset.....	22
2.6. SiblingsDB.....	22
3. IMPLEMENTACIJA.....	24
3.1. Vektorska udaljenost između značajki.....	24
3.2. Shape unit	25
3.3. Normalizirana križna korelacija.....	26

3.4. Svojstvena lica.....	28
3.5. Local binary patterns	30
3.6. LBP + PCA.....	32
4. REZULTATI	34
ZAKLJUČAK.....	36
POPIS LITERATURE.....	37
DODATAK A – KORIŠTENI ALATI.....	38
SAŽETAK.....	39
ABSTRACT.....	40

UVOD

Automatsko raspoznavanje lica jedna je od najbitnijih i najpopularnijih vrsta sustava za biometrijsku identifikaciju i autorizaciju. Računalni sustav koji pruža funkciju automatskog raspoznavanja lica može biti koristan u širokom spektru ljudskog djelovanja, od informacijske sigurnosti i nadzora pa sve do industrije zabave.

Cilj ovog rada je proučiti tehnologije automatskog raspoznavanja lica i sustav za detekciju značajki lica dostupan na Zavodu za telekomunikacije te predložiti i implementirati metodu za automatsko raspoznavanje lica na bazi detektiranih značajki i teksture lica.

Prvo je dan okvirni pregled područja. Nakon toga slijedi pregled nekih od najpopularnijih metoda korištenih za automatsko raspoznavanje lica, uključujući računanje vektorske udaljenosti između značajki, normalizirane križne korelacije, metode svojstvenih lica te LBP opisnika.

Slijedi pregled nekih od popularnijih i zanimljivijih baza podataka namijenjenih za testiranje algoritama za prepoznavanje lica.

Opisan je i proces izrade testova predstavljenih metoda uz pomoć programskih alata OpenCV i VisageSDK.

U konačnici su prikazani rezultati izrađenih testova na odabranoj bazi podataka te diskusija tih rezultata.

Pregled korištenih alata dan je u Dodatku A.

PREGLED PODRUČJA

Automatsko raspoznavanje lica podrazumijeva računalnu aplikaciju koja bez navođenja korisnika verificira identitet osobe sa digitalne slike ili okvira digitalnog videa. Jedan od najčešćih načina kojim se to postiže je uspoređivanje odabranih značajki lica na dobivenoj slici i onih spremljenih u bazu podataka.

Jedan dio algoritama za raspoznavanje lica identificira značajke ili određene točke na slici lica. Primjerice, algoritam može analizirati relativnu poziciju, veličinu i oblik očiju, nosa, obraza ili čeljusti. Potom se rezultati te analize koriste za pretragu drugih slika sa odgovarajućim značajkama.

Dio algoritama funkcionira na principu normaliziranja cijele galerije slika lica te sažimanjem podataka, zadržavajući samo one podatke koji su korisni za prepoznavanje lica, nakon čega se testna slika uspoređuje sa spremljenim podacima. Jedan od najranijih uspješnih sustava za prepoznavanje lica temeljio se na metodi uspoređivanja predložaka primijenjenih samo na istaknute značajke lica, efektivno pružajući sažeti zapis lica.

Algoritmi za prepoznavanje lica mogu se podijeliti u dvije skupine. **Geometrijski** algoritmi analiziraju istaknute značajke lica, dok **fotometrijski** algoritmi pružaju statistički pristup koji interpretira sliku kao niz vrijednosti te uspoređuje vrijednosti s predlošcima kako bi se eliminirala varijanca.

Neki od popularnih algoritama za raspoznavanje lica:

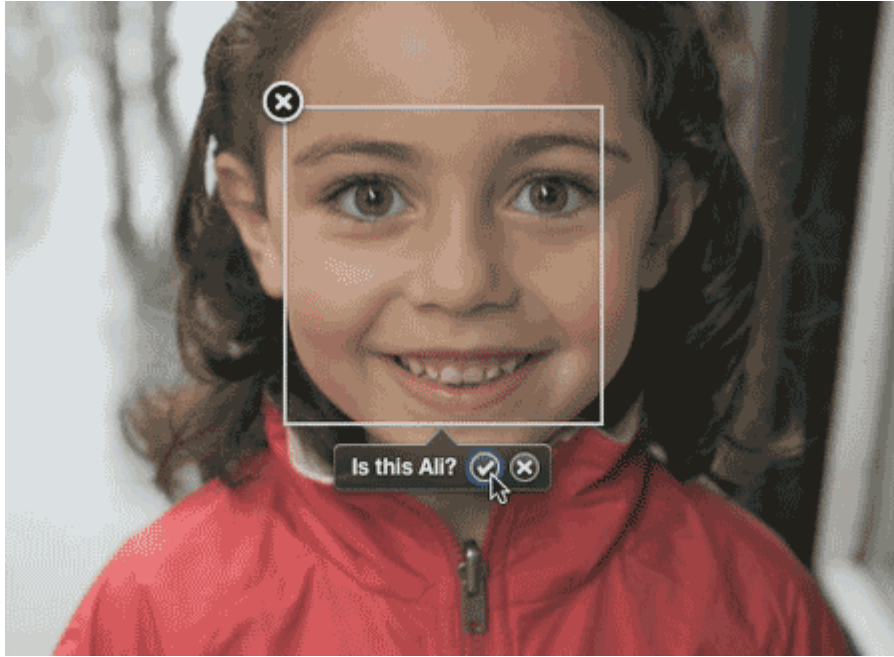
- PCA uz metodu svojstvenih lica [2]
- *Linear Discriminate Analysis* [3]
- *Elastic Bunch Graph Matching* [4]
- Skriveni Markovljev model [5]
- *Multilinear Subspace Learning* [6]

Posljednjih nekoliko godina popularnost raste metodama koje koriste 3D senzore za bilježenje informacija o obliku lica. Potom se te informacije koriste za identifikaciju istaknutih značajki lica poput oblika očnih duplji, nosa i brade, zadržavajući prostornu komponentu informacije. Prednost trodimenzionalne analize lica je robusnost s obzirom na promjene u osvjetljenju i vanjskim uvjetima te bolji rezultati u situacijama gdje se lice skenira pod kutem. Trodimenzionalna analiza moguća je zbog razvoja sofisticiranih senzora koji vrlo precizno bilježe 3D strukturu lica. Preko deset takvih senzora moguće je postaviti na jedan CMOS čip, pri čemu svaki senzor bilježi zaseban dio spektra.

Ipak, čak i savršeno bilježenje 3D strukture može biti osjetljivo na promjene u izrazima lica. Iako se javljaju neke metode koje pokušavaju zaobići taj problem, takva rješenja još uvijek nisu spremna za praktičnu uporabu.

Još jedan rastući trend je korištenje detalja na koži pri prepoznavanju lica. Analiza teksture lica pretvara jedinstvene crte, uzorke i točke na koži u varijablu koja se provjerava pri usporedbi lica s bazom podataka. Korištenje detalja na koži pri analizi u nekim slučajevima može dovesti do značajno boljih rezultata pri prepoznavanju lica, no za takve metode su potrebna kvalitetna očitavanja u visokoj rezoluciji, što u praktičnim uvjetima često nije moguće.

Appleov iPhoto, Googleova Picasa, Adobeov Photoshop Elements te Sonyjev Picture Motion Browser neki su od poznatijih softverskih alata koji pružaju mogućnosti prepoznavanja lica.



Slika 1. Apple iPhoto

Iako je postignut značajan napredak od početaka u šezdesetim godinama prošlog stoljeća, automatsko prepoznavanje lica je i dalje područje koje u praktičnim primjenama nailazi na mnoge probleme. Gotovo sve metode imaju problema sa promjenama u osvjetljenju, slikama niske rezolucije te slikama gdje je dio lica subjekta prekriven naočalama, kosom ili nekim drugim objektom. Osim toga, većina modernih metoda funkcionira vrlo dobro za frontalne slike, ali postižu znatno slabije rezultate već ako je lice pod kutem većim od 20° u odnosu na os kamere. Također, kod svih metoda postoji velika osjetljivost na promjene u izrazima lica.

S obzirom na ta ograničenja, automatsko prepoznavanje lica u ovom trenutku nije prikladno za praktičnu uporabu u većini slučajeva. Različiti sustavi za automatsko prepoznavanje lica u svrhu sigurnosti i prepoznavanja potencijalnih kriminalaca uvedeni su u Londonskoj pokrajini Newham, u Floridi te na Bostonskom aerodromu Logan, no razočaravajući rezultati doveli su do zaustavljanja rada tih sustava.

Porast popularnosti sustava za automatsko prepoznavanje lica otvorio je i raspravu o privatnosti. Razne civilne organizacije javno su izrazile vrlo negativan stav o

takvim sustavima, navodeći kako prepoznavanje lica može biti korišteno ne samo za identifikaciju pojedinca, već i za pronalazak ostalih osobnih podataka o osobi poput drugih fotografija, blog zapisa, profila na društvenim mrežama, navikama putovanja i slično. Također, navodi se kako osoba nema previše izbora ni načina za zaobilazanje prepoznavanja lica osim skrivanja samog lica, što fundamentalno mijenja dinamiku privatnosti u javnim prostorima, omogućavajući vlastima, oglašivačima ili hakerima da zlorabe dobivene informacije.



Slika 2. Microsoft Kinect također prepoznaje korisnike

1. METODE

1.1. Vektorska udaljenost između značajki

Woody Bledsoe, Helen Chan Wolf i Charles Bisson znanstvenici su koji su tijekom 1964. i 1965. godine provodili prve eksperimente vezane uz računalom potpomognuto raspoznavanje lica. Uspostavljeni problem je bio za danu sliku lica odabrati skup zapisa u bazi podataka takav da osoba iz jednog od tih zapisa odgovara osobi na danoj slici. Proces nije bio potpuno automatiziran – znanstvenici su ručno pronalazili koordinate predodređenih značajki lica na fotografijama, poput sredina zjenica, unutarnjih i vanjskih rubova očiju i sličnih, koje bi računalo potom koristilo pri izračunu. Iz tih koordinata stvara se lista dvadeset udaljenosti na temelju kojih se računa identifikator lica. Njihov tadašnji sustav je koristeći takav postupak obrađivao otprilike 40 slika po satu.

Pri izgradnji baze, ime osobe na slici povezano je sa listom izračunatih udaljenosti. U fazi testa, skup udaljenosti izračunat na novoj slici uspoređuje se sa spremljenim podacima te je rezultat udaljenost između nove slike i slike u bazi. Kao rezultat vraća se ime koje je povezano sa slikom u bazi koja u odnosu na novu sliku ima najmanju udaljenost.

Osim povijesnog značaja, metoda odabrana je i zbog činjenice da je ova metoda jedno od najintuitivnijih rješenja za problem prepoznavanja lica, iako istraživanja u području psihologije percepcije lica ne pokazuju strogu poveznicu između udaljenosti značajki na licu i procesa prepoznavanja lica.

1.2. Shape unit

Jedna od funkcija koju nudi programski alat VisageSDK pri praćenju značajki lica je izrada jednostavnog trodimenzionalnog modela promatranog lica, tzv. *candide* model. Iako nije jedna od iskušanih i ispitanih metoda automatskog raspoznavanja lica, prvenstveno zbog nepouzdanosti procjene oblika lica iz dvodimenzionalne slike, ukoliko se promatra kao produžetak prethodne metode, razumno je očekivati da će dodatak treće dimenzije pružati bolje rezultate pri promatranju geometrije lica.

Pri izradi trodimenzionalnog modela promatranog lica kreće se od početnog modela lica, nad kojim se za svaki vrh na sljedeći način računa njegova nova pozicija:

$$v'_i = v_i + f \cdot d_i \quad (1.1)$$

Oznakom v_i predstavljena je originalna poziciju vrha na početnom modelu lica, a oznakom d_i predstavljena je najveća udaljenost koju vrh može imati od svoje početne pozicije. Niz d određen je vrijednostima zapisanim u konfiguracijsku datoteku. Oznaka f predstavlja niz vrijednosti koji nazivamo *shape units* te označavaju mjeru u kojoj će se vrh pomaknuti, što znači da vrijednost te varijable određuje gdje će se u konačnici nalaziti promatrani vrh.

Budući da se *shape uniti* razlikuju od osobe do osobe, moguće ih je koristiti za direktnu usporedbu dva lica. Ipak, budući da se metoda zasniva na istoj ideji kao i prethodna, zadržava sve probleme te metode. Sve metode koje se temelje isključivo na uspoređivanju geometrije lica izrazito su osjetljive na promjene u izrazu lica, a ovdje dodatni problem predstavlja i nepoznati faktor greške pri određivanju vrijednosti *shape unita*.

1.3. Normalizirana križna korelacija

Osnovni problem koji se javlja prilikom obrade slika je određivanje pozicije danog uzorka u slici. Ovaj problem usko je vezan za problem određivanja primljenog signala u obradi signala. Problem se može podijeliti na dva osnovna slučaja:

- pozicija uzorka je nepoznata
- postoji procjena pozicije uzorka

Najčešće je potrebno riješiti oba slučaja kako bi se odredio položaj uzorka u slici. Ipak, u drugom slučaju se procjena pozicije može koristiti za značajno smanjenje kompleksnosti izračuna te se koristi u praćenju značajki u sekvencama slika.

Postoji nekoliko dobro razvijenih algoritama za praćenje značajki i početnu procjenu pozicije danog uzorka. Jedan od osnovnijih pristupa, koji se može koristiti u oba gore navedena slučaja, je sparivanje predložaka. Pozicija danog uzorka određuje se usporedbom slike i danog predloška na razini piksela. Predložak se postepeno pomiče za u piksela po x osi te v piksela po y osi slike te se usporedba računa po površini predloška na svakoj poziciji (u, v) . Normalizirana križna korelacija (u eng. literaturi *normalized cross correlation* [10]) je dobar izbor za taj izračun u većini slučajeva.

Označimo sa $f(x, y)$ vrijednost intenziteta slike f veličine $M_x \times M_y$ u točki (x, y) , $x \in \{0, \dots, M_x - 1\}, y \in \{0, \dots, M_y - 1\}$. Uzorak je predstavljen danim predloškom t veličine $N_x \times N_y$. Uobičajen način za računanje pozicije (u_{pos}, v_{pos}) uzorka u slici f je izračunati vrijednost normalizirane križne korelacije γ u svakoj točki (u, v) za sliku f i predložak t , koji je pomaknut za u piksela po x osi te v piksela po y osi. Vrijednost γ u određenoj točki računa se na sljedeći način:

$$\gamma = \frac{\sum_{x,y} (f(x, y) - \bar{f}_{u,v})(t(x - u, y - v) - \bar{t})}{\sqrt{\sum_{x,y} (f(x, y) - \bar{f}_{u,v})^2 \sum_{x,y} (t(x - u, y - v) - \bar{t})^2}} \quad (1.2)$$

Oznaku $\bar{f}_{u,v}$ koristimo za vrijednost $f(x, y)$ unutar prostora predloška t pomaknutog za (u, v) te se računa na sljedeći način:

$$\bar{f}_{u,v} = \frac{1}{N_x N_y} \sum_{x=u}^{u+N_x-1} \sum_{y=v}^{v+N_y-1} f(x, y) \quad (1.3)$$

Slične oznake, \bar{t} je srednja vrijednost predloška t . Nazivnik u jednadžbi (1.2) je varijanca funkcije slike $f(x, y) - \bar{f}_{u,v}$ koja ima srednju vrijednost nula i pomaknute funkcije predloška $t(x - u, y - v) - \bar{t}$ kojoj je srednja vrijednost također nula. Upravo zbog te normalizacije, vrijednost $\gamma(u, v)$ ne ovisi o promjenama u svjetlini ili kontrastu slike, koji su povezani sa srednjom vrijednošću i standardnom devijacijom. Željena pozicija uzorka (u_{pos}, v_{pos}) je jednaka poziciji (u_{max}, v_{max}) vrijednosti γ_{max} , maksimalnoj vrijednosti $\gamma(u, v)$. Jednadžba za računanje vrijednosti γ robusnija je od većine sličnih metoda poput jednostavne kovarijance ili zbroja apsolutnih razlika zbog normalizacije. Ipak, taj izračun je računski zahtjevan. Za nazivnik, koji normalizira cross correlation koeficijent, potrebno je za svaku točku (u, v) , $u \in \{0, \dots, M_x - N_x\}$, $v \in \{0, \dots, M_y - N_y\}$ na slici, gdje se određuje $\gamma(u, v)$, računati središnju vrijednost unutar prostora funkcije predloška $\bar{f}_{u,v}$ te energiju slike:

$$e_f(u, v) = \sum_{x=u}^{u+N_x-1} \sum_{y=v}^{v+N_y-1} (f(x, y) - \bar{f}_{u,v})^2 \quad (1.4)$$

Ako se te kalkulacije izvode na izravan način kao što je navedeno u jednadžbi, broj izračuna proporcionalan je $N_x N_y (M_x - N_x) (M_y - N_y)$, iako je energiju funkcije predloška

$$e_i(u, v) = \sum_{x=1}^{N_x} \sum_{y=1}^{N_y} (t(x, y) - \bar{t})^2 \quad (1.5)$$

i srednju vrijednost funkcije predložka \bar{t} potrebno računati samo jednom unaprijed. Za većinu praktičnih implementacija, poput aplikacija koje se izvode u stvarnom vremenu, ovakva računaska zahtjevnost nije prihvatljiva. Dio izračuna može se računati u frekvencijskoj domeni pomoću Fourierovih transformacija, ali broj kalkulacija bi i dalje ostao visok. Algoritam koji za računanje križne korelacije koristi manje kalkulacija bi uvelike povećao uporabljivost metode u sustavima koji rade u realnom vremenu.

1.4. Svojstvena lica

Metoda svojstvenih lica (u eng. literaturi *eigenfaces* [2]) je jedan od najpopularnijih i najistraživanijih pristupa računalnoj reprezentaciji ljudskog lica. Metoda se zasniva na ideji da se svako ljudsko lice može izraziti kao kombinacija određenog skupa lica koje nazivamo svojstvena lica. Taj skup lica najčešće se generira korištenjem PCA postupka nad velikim skupom slika različitih lica. Budući da nije potreban velik broj svojstvenih lica kako bi se dobila uvjerljiva aproksimacija većine lica te da se lice u bazu sprema kao niz brojevnih vrijednosti umjesto slike, reprezentacija jednog lica zauzima manje mjesta nego većina ostalih metoda.

Izraženo matematički, želimo pronaći principalne komponente distribucije lica, odnosno svojstvene vektore kovarijantne matrice skupa slika, pri čemu se svaka slika tretira kao vektor u visokodimenzionalnom prostoru. Te informacije se potom koriste za kodiranje i uspoređivanje lica.

Analiza glavnih osi, odnosno PCA (*Principal Component Analysis*), je postupak koji pomoću ortogonalnih transformacija pretvara skup linearno zavisnih varijabli u skup linearno nezavisnih varijabli, pri čemu je rezultirajući broj varijabli jednak ili manji

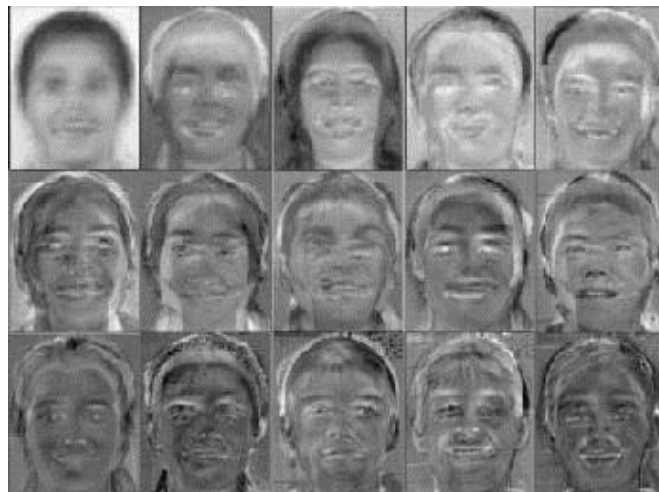
početnom. Zbog činjenice da su rezultirajuće singularne vrijednosti poredane silazno po varijanci, PCA se može koristiti za smanjenje dimenzionalnosti kod visokodimenzionalnih skupova podataka, pritom zadržavajući najveću moguću varijancu podataka. Primjerice, ukoliko se uzmu prve dvije vrijednosti, početni visokodimenzionalni skup podataka može se prikazati kao dvodimenzionalni skup podataka koji najbolje reprezentira početni skup.

Za izradu skupa svojstvenih lica potreban je početni skup slika različitih lica. Slike moraju biti jednakih veličina, uslikane u istim ili sličnim uvjetima osvjetljenja te biti normalizirane tako da značajke lica na svim slikama budu na istim mjestima. Svaka slika se potom izražava kao vektor povezivanjem svih redova piksela početne slike u jedan red. Početni skup slika izražen je kao matrica \mathbf{T} , pri čemu je svaki stupac jedna od slika iz skupa. Na temelju te matrice računa se središnja slika, odnosno središnja vrijednost početnog skupa te se oduzima od svake početne slike, odnosno stupca u matrici \mathbf{T} . Potom je potrebno izračunati vlastite vrijednosti i vlastite vektore kovarijantne matrice \mathbf{S} . Vlastiti vektori kovarijantne matrice su svojstvena lica, a budući da imaju istu dimenzionalnost kao i početne slike, oni se također mogu tumačiti kao slike.

Dobivena svojstvena lica mogu se koristiti za reprezentaciju postojećih, ali i novih lica. Moguće je projicirati novu sliku u prostor svojstvenih lica te zapisati kako se ta slika razlikuje od središnje slike početnog skupa. Vlastite vrijednosti povezane uz pojedinačno svojstveno lice predstavljaju koliko se slika razlikuje od središnjeg lica u tom smjeru. Projiciranjem slike na podskup svojstvenih vektora gube se informacije, ali taj gubitak se minimizira korištenjem PCA kako bi se zadržale vrijednosti s najvećom varijancom.

Primjerice, ukoliko se radi o slici dimenzija 100x100, rezultat projekcije će biti 10000 svojstvenih vektora. No kako je većinu lica moguće aproksimirati korištenjem već između 100 i 150 svojstvenih vektora, većinu njih koje dobijemo projiciranjem slike možemo odbaciti [8].

Budući da je metodom svojstvenih lica moguće reprezentirati velik broj subjekata koristeći relativno malu količinu podataka, metoda je efikasnija i brža od velikog broja postojećih metoda, čemu pridonosi i relativna invarijabilnost metode s obzirom na smanjenje veličina slika. Ipak, efikasnost metode drastično pada s povećanjem razlike između trening i test slika. Za dobre rezultate potrebno je dobiti slike iz kontroliranih uvjeta – slike se ne smiju naročito razlikovati u osvjetljenju, translaciji i rotaciji glave te izrazu lica.



Slika 1.1. Primjeri svojstvenih lica

Kako bi se ova metoda koristila za prepoznavanje lica, slike lica spremljene su kao nizovi težinskih vrijednosti koje opisuju doprinos svakog pojedinog svojstvenog lica toj slici. Svako novo lice projicira se u prostor svojstvenih lica kako bi se saznale težinske vrijednosti koje opisuju to lice. Te vrijednosti se potom uspoređuju s vrijednostima svake zasebne slike trening seta te se računa euklidska udaljenost kako bi se pronašla slika koja se najmanje razlikuje od dane slike.

1.4.1. Računanje svojstvenog lica

Neka je slika lica $I(x, y)$ dvodimenzionalni niz 8-bitnih vrijednosti veličine $N \times N$. Sliku također možemo smatrati vektorom dimenzije N^2 , tako da slika dimenzija 256×256 postaje vektor dimenzije 65536, odnosno točka u 65536-dimenzionalnom prostoru. To znači da bi skup slika odgovarao nizu točaka u tom prostoru. Kako su ljudska lica slična po svojoj konfiguraciji, slike neće biti nasumično raspoređene u tom prostoru, što znači da se skup slika lica može opisati podprostorom relativno male dimenzionalnosti. Ideja PCA je pronaći vektore koji najbolje predstavljaju distribuciju slika lica u cijelom prostoru slika. Svaki od tih vektora je duljine N^2 , opisuje sliku veličine $N \times N$ te je linearna kombinacija originalnih slika lica. Budući da su ovi vektori svojstveni vektori kovarijantne matrice koja odgovara originalnim slikama lica te budući da po svom izgledu slične licima, nazivamo ih svojstvena lica.

Definirajmo trening skup slika lica kao $\Gamma_1, \Gamma_2, \Gamma_3, \dots, \Gamma_M$. Prosječno lice u skupu je definirano kao $\psi = \frac{1}{M} \sum_{n=1}^M \Gamma_n$. Svako pojedino lice razlikuje se od prosječnog za vektor $\phi_i = \Gamma_i - \psi$. Nad ovim skupom vrlo velikih vektora se potom vrši PCA postupak, koji traži skup M ortonormiranih vektora \mathbf{u}_n koji najbolje opisuju distribuciju podataka. k -ti vektor \mathbf{u}_k je odabran tako da je $\lambda_k = \frac{1}{M} \sum_{n=1}^M (\mathbf{u}_k^T \phi_n)^2$ maksimum te vrijedi sljedeće:

$$\mathbf{u}_l^T \mathbf{u}_k = \delta_{lk} = \begin{cases} 1, & \text{ako je } l = k \\ 0, & \text{inače} \end{cases} \quad (1.6)$$

Vektori \mathbf{u}_k i skalari λ_k su svojstveni vektori i svojstvene vrijednosti kovarijantne matrice $C = \frac{1}{M} \sum_{n=1}^M \phi_n \phi_n^T = AA^T$, gdje je matrica A jednaka $A = [\phi_1 \phi_2 \dots \phi_M]$. Kako je matrica C dimenzija $N^2 \times N^2$, određivanje njenih svojstvenih vektora i svojstvenih vrijednosti je trenutno neostvariv zadatak za prosječne veličine slika te je potrebna računalno izvediva metoda za pronalazak svojstvenih vektora. Ako je broj točaka u

prostoru manji od dimenzije prostora, odnosno $M < N^2$, postojat će $M-1$ umjesto N^2 značajnih svojstvenih vektora, dok će ostali imati pridruženu svojstvenu vrijednost nula. U ovom slučaju moguće je pronaći N^2 -dimenzionalne svojstvene vektore pronalaženjem svojstvenih vektora $M \times M$ matrice (primjerice, pronalaskom svojstvenih vektora matrice dimenzija 16×16 umjesto matrice dimenzija 16384×16384) te uzimanjem pripadajućih linearnih kombinacija slika lica Φ_i .

Uzmimo svojstvene vektore v_i matrice $A^T A$ takve da vrijedi sljedeće:

$$A^T A v_i = \mu_i v_i. \quad (1.7)$$

Ukoliko pomnožimo obje strane matricom A , dobivamo sljedeće:

$$A A^T A v_i = \mu_i A v_i \quad (1.8)$$

Iz gornje jednadžbe vidljivo je da su $A v_i$ svojstveni vektori matrice $C = A A^T$. Potom stvaramo $M \times M$ matricu $L = A^T A$, pri čemu vrijedi $L_{mn} = \phi_m^T \phi_n$ te nalazimo M svojstvenih vektora matrice L koje označavamo sa v_l . Ti vektori određuju linearnu kombinaciju M slika iz trening skupa slika koje formiraju svojstvena lica u_l .

$$u_l = \sum_{k=1}^M v_{lk} \phi_k, \quad l = 1, \dots, M \quad (1.9)$$

Ovim postupkom količina računanja značajno je smanjena, od reda broja piksela na slici (N^2) do reda broja slika u trening skupu (M). U praktičnim slučajevima broj slika u trening skupu će biti relativno manji te će izračun svojstvenih vektora postati izvediv. Pridružene svojstvene vrijednosti omogućavaju rangiranje svojstvenih vektora prema njihovoj korisnosti u karakteriziranju varijacije između slika. Nakon odabira M' značajnih svojstvenih vektora s najvećim pridruženim svojstvenim vrijednostima, vektori razapinju M' -dimenzionalni prostor originalnog N^2 prostora. Nova slika lica Γ projicira se u novi prostor M' na sljedeći način:

$$\omega_k = u_k^T (\Gamma - \Psi), \quad k = 1, \dots, M' \quad (1.10)$$

Iz toga formiramo vektor $\Omega^T = [\omega_1, \omega_2, \dots, \omega_{M'}]$ koji opisuje doprinos svakog svojstvenog lica u prikazu ulazne slike lica, pri čemu se skup svojstvenih lica tretira kao bazni skup za slike lica. Taj vektor se potom može koristiti u algoritmima raspoznavanja uzoraka kako bi se pronašlo koja od predefiniranih klasa lica najbolje opisuje lice. Najjednostavnija metoda je pronaći klasu lica k koja minimizira Euklidsku udaljenost:

$$\epsilon_k^2 = \|(\Omega - \Omega_k)\|^2 \quad (1.11)$$

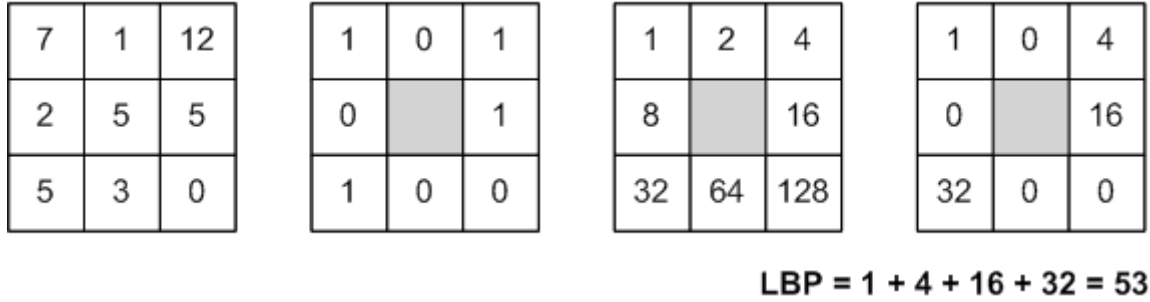
1.5. Local Binary Patterns

Local Binary Pattern (LBP) je jednostavan i efikasan operator nad teksturama koji pridaje oznake pikselima slike u obliku binarnih brojeva koji predstavljaju odnos piksela u odnosu na susjedne piksele. Zbog jednostavnosti izračuna te sposobnosti razlikovanja slika, LBP je postao jedan od popularnih pristupa problemu raspoznavanja lica. Možda najbitnije svojstvo LBP operatora u stvarnoj primjeni je njegova robusnost s obzirom na promjene u crno-bijeloj slici uzrokovane uvjetima osvjetljenja. Drugo bitno svojstvo je jednostavnost izračuna, zbog čega je moguće koristiti metodu u stvarnom vremenu.

Osnovna ideja iza LBP operatora je da se dvodimenzionalne površinske teksture mogu opisati pomoću dvije komplementarne mjere – lokalni prostorni uzorci te kontrast sivih tonova. Originalni operator određuje oznake za piksele slike pomoću binarnih brojeva koje stvara uspoređivanjem susjednih piksela s centralnim pikselom. Histogram ovih 256 različitih oznaka potom se može koristiti kao deskriptor teksture. Označimo sa P broj točaka iz susjedstva koje uzimamo u obzir,

a sa R radijus promatranog susjedstva te označimo susjedstvo sa (P, R). Tada se vrijednost LBP koda piksela računa na sljedeći način:

$$LBP_{P,R} = \sum_{p=0}^{P-1} s(g_p - g_c)2^p, \quad s(x) = \begin{cases} 1, & \text{ako je } x \geq 0 \\ 0, & \text{inače} \end{cases} \quad (1.12)$$



Slika 1.2. Izračun LBP koda piksela

Osnovni LBP operator proširen je tako da koristi susjedstva različitih veličina. Umjesto osam direktnih susjeda piksela, može se koristiti proizvoljni broj susjeda. Korištenje kružnog susjedstva te bilinearna interpolacija vrijednosti na necjelobrojnim koordinatama omogućavaju uzimanje bilo kojeg radijusa te broja piksela u susjedstvu. Varijanca sivih tonova lokalnog susjedstva može se koristiti kao komplementarna mjera. Uzmimo monokromatsku sliku $I(x, y)$ te sa g_c označimo razinu sive proizvoljnog piksela (x, y) , odnosno $g_c = I(x, y)$. Također, neka g_p označava vrijednost sive boje uzorkovane točke na jednoliko raspoređenom kružnom susjedstvu sa P točaka i radijusa R oko točke (x, y) . Tada vrijedi:

$$g_p = I(x_p, y_p), \quad p = 0, \dots, P - 1 \quad (1.13)$$

$$x_p = x + R \cos\left(\frac{2\pi p}{P}\right) \quad (1.14)$$

$$y_p = y - R \sin\left(\frac{2\pi p}{P}\right) \quad (1.15)$$

Pretpostavimo da je lokalna tekstura slike $I(x,y)$ karakterizirana distribucijom vrijednosti sive $P+1$ piksela:

$$T = t(g_c, g_0, g_1, \dots, g_{P-1}) \quad (1.16)$$

Vrijednost središnjeg piksela moguće je oduzeti od susjednih vrijednosti bez gubitka informacija:

$$T = t(g_c, g_0 - g_c, g_1 - g_c, \dots, g_{P-1} - g_c) \quad (1.17)$$

Distribuciju je moguće aproksimirati ako pretpostavimo da je središnji piksel statistički nezavisan od razlika, zbog čega možemo faktorizirati distribuciju:

$$T \approx t(g_c)t(g_0 - g_c, g_1 - g_c, \dots, g_{P-1} - g_c) \quad (1.18)$$

U kontekstu analiziranja lokalnih uzoraka, prvi faktor ne sadrži nikakvu korisnu informaciju, pa za modeliranje lokalne teksture možemo koristiti samo ostatak:

$$t(g_0 - g_c, g_1 - g_c, \dots, g_{P-1} - g_c) \quad (1.19)$$

Još jedno proširenje originalnog operatora je definiranje takozvanih *ujednačenih uzoraka*, koji se mogu koristiti za smanjenje dužine vektora značajki te za implementaciju deskriptora koji je invarijantan u odnosu na rotaciju slike. Proširenje je nastalo iz činjenice da se neki binarni uzorci pojavljuju češće od drugih. Binarni uzorak je ujednačen ukoliko sadrži najviše dvije tranzicije iz 0 i 1 ili obratno kad ga se promatra cirkularno. Primjerice, uzorci 00000000 (0 tranzicija), 00011000 (2 tranzicije) i 11000001 (2 tranzicije) su ujednačeni, dok uzorci 10100000 (4 tranzicije), 01101010 (6 tranzicija) i 10101010 (8 tranzicija) nisu. Pri računanju LBP vrijednosti, ujednačeni uzorci se koriste tako da postoji zasebna oznaka za svaki ujednačeni uzorak, a svi neujednačeni uzorci koriste istu oznaku. U praksi se pokazalo da je

malo manje od 90% uzoraka ujednačeno ukoliko se koristi susjedstvo od 8 točaka i radijusa 1, odnosno oko 70% u susjedstvu od 16 točaka radijusa 2 [11].

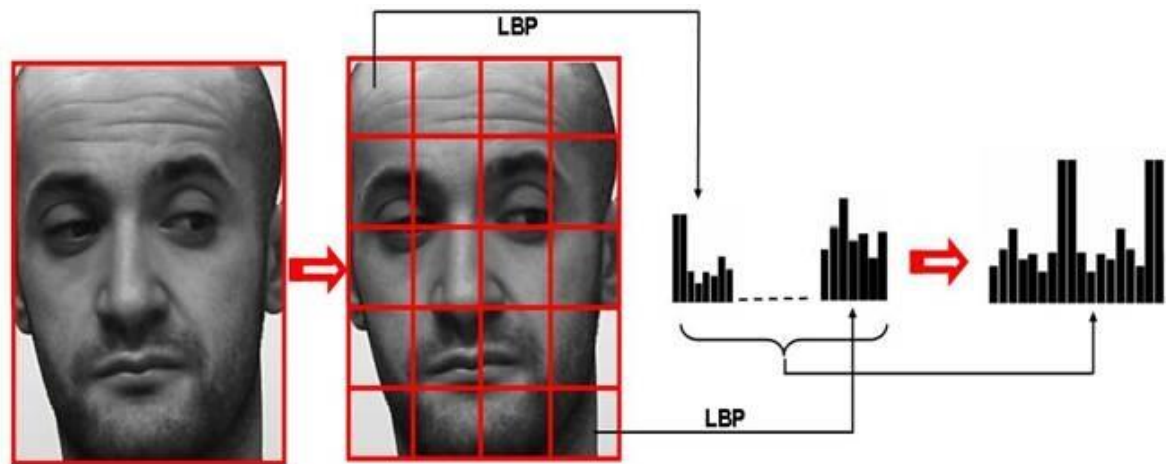
Nakon što se stvori slika $f_l(x, y)$ koja se sastoji od LBP oznaka, histogram se definira ovako:

$$H_i = \sum_{x,y} I\{f_l(x, y) = i\}, \quad i = 0, \dots, n - 1 \quad (1.20)$$

Pri tome je n broj različitih oznaka koje su nastale kao rezultat LBP-a te je $I\{A\}$ jednak 1 ako je A istinit, odnosno 0 ako nije. Ukoliko se uspoređuju slike (ili dijelovi slika) različitih veličina, histogrami moraju biti normalizirani:

$$N_i = \frac{H_i}{\sum_{j=0}^{n-1} H_j} \quad (1.21)$$

Pri korištenju LBP pristupa za klasifikaciju tekstura, pojave LBP oznaka u slici se skupljaju u histogram. Klasifikacija je potom izvršava pronalaskom najbližeg postojećeg histograma. Ovakav pristup nije prikladan pri rješavanju problema prepoznavanja lica jer se na taj način gube prostorne informacije. Jedan od načina na koji se prostorne informacije mogu zadržati prilikom LBP pristupa je izgradnja nekoliko lokalnih deskriptora lica koji se potom kombiniraju u jedan globalni deskriptor. Takve metode su istovremeno i robusnije s obzirom na varijacije u položaju, odnosno rotaciji lica te osvjetljenju nego metode koje u obzir uzimaju samo cjelovitu sliku. U osnovi, ideja takvog pristupa je podijeliti sliku lica u lokalne regije te odrediti LBP deskriptore za svaku regiju zasebno. Dobiveni deskriptori se potom ulančavaju u histogram koji je globalni deskriptor lica.



Slika 1.3. Računanje lokalnih histograma

Ovakav histogram opisuje lice na tri različite razine lokalnosti – LBP oznake za histogram sadrže informacije o uzorcima na razini piksela, oznake su sumirane nad malim dijelovima slike kako bi stvorile informaciju na regionalnoj razini, a ulančavanje regionalnih histograma u jedan daje nam opis na razini cijelog lica. Lokalne regije na koje se slike dijele ne moraju biti kvadratnog oblika, ne moraju sve biti istog oblika te ne moraju nužno pokrivati cijelu površinu slike. Također je moguće da se neke regije preklapaju.

2. BAZE PODATAKA

2.1. Face Recognition Grand Challenge

Face Recognition Grand Challenge je evaluacija za sustave za automatsko prepoznavanje lica provedena od strane američkog Nacionalnog instituta za standarde i tehnologiju između 2004. i 2006. godine. Primarni cilj evaluacije bio je promicanje i napredak tehnologije za prepoznavanje lica. Evaluacija se sastojala od šest progresivno težih zadataka te je istovremeno pokrivala sve pristupe problemu. Popratna baza podataka koja se koristila pri evaluaciji sastoji se od 50 000 zapisa podijeljenih u trening i validacijske setove. Trening setovi namijenjeni su za treniranje razvijenih algoritama, dok su validacijski setovi namijenjeni za procjenu performansa algoritma u laboratorijskim uvjetima. Baza se sastoji fotografija 4003 osobe, pri čemu za svaku osobu postoje četiri slike u kontroliranim uvjetima, dvije slike u nekontroliranim uvjetima te jedne trodimenzionalne slike. Slike izrađene u kontroliranim uvjetima podrazumijevaju frontalne slike pri dva različita studijska postava osvjetljenja te dva izraza lica – neutralan izraz te osmijeh. Slike izrađene u nekontroliranim uvjetima podrazumijevaju frontalne slike pri različitim, nekontroliranim uvjetima osvjetljenja te dva izraza lica – neutralan izraz i osmijeh. Trodimenzionalna slika izrađena je u studijskim uvjetima osvjetljenja.

2.2. Multi-PIE

Multi-PIE baza slika lica stvorena je na američkom sveučilištu Carnegie Mellon. Sastoji se od preko 750 000 slika koje pokrivaju 337 subjekata uslikanih iz petnaest pozicija uz devetnaest različitih uvjeta osvjetljenja te nekoliko različitih izraza lica.

Uz tako velik broj visokorezolucijskih slika lica, ovo je jedna od najopsežnijih baza u primjeni.



Slika 2.1. Različite pozicije kamere u Multi-PIE bazi podataka

2.3. LFW

Labeled Faces in the Wild je baza slika lica dizajnirana sa idejom proučavanja problema automatskog prepoznavanja lica u nekontroliranim uvjetima. Baza sadrži više od 13 000 slika koje pokrivaju preko 5000 subjekata, od kojih je 1680 pokriveno s više od jedne slike. Baza je sastavljena od fotografija javnih osoba sakupljenih s interneta, zbog čega se slike često drastično razlikuju u poziciji i rotaciji lica, izrazima lica, uvjetima osvjetljenja te ostalim vanjskim faktorima. S obzirom na to, testiranje uz LFW bazu podataka smatra se jednim od najtežih testova vezanih uz automatsko prepoznavanje lica, a rezultati testiranja često su bliži rezultatima u stvarnim uvjetima nego korištenjem drugih baza podataka.

2.4. Plastic Surgery Face Database

Baza podataka koja sadrži 1800 slika koje pokrivaju 900 subjekata. Ono što razlikuje ovu bazu od ostalih je činjenica da se sastoji isključivo od slika lica prije i nakon plastičnih operacija. Cilj baze podataka je omogućiti istraživanje utjecaja operacija na performanse sustava za automatsko prepoznavanje lica. S obzirom na to, baza se sastoji od slika koje pokrivaju različite operacije poput rinoplastije i blefaroplastike. Slike su izrađene u studijskim uvjetima osvjetljenja uz neutralne izraze lica. Baza sadrži 519 parova slika koji se odnose na lokalne operacije te 381 par slika koji se odnose na općenitije plastične operacije poput face liftinga.

2.5. YouTube Makeup Dataset

YMU baza podataka namijenjena za proučavanje utjecaja kozmetičkih proizvoda na performanse sustava za automatsko prepoznavanje lica, a sastoji se od slika preuzetih iz raznih YouTube uputa za šminkanje. Baza pokriva 150 subjekata te za svakog od njih sadrži dvije slike prije te dvije slike nakon nanošenja šminke. Količina šminke, izrazi lica, pozicija i kut glave te uvjeti osvjetljenja razlikuju se između subjekata, a ponekad i između različitih slika istog subjekta.

2.6. SiblingsDB

SiblingsDB sadrži dva različita skupa slika braće i sestara. Baza nije namijenjena za testiranje sposobnosti sustava da prepozna istu osobu na dvije različite slike, već je cilj testirati može li sustav prepoznati subjekte koji su u srodstvu. Prvi skup slika

zove se *HQFaces* te se sastoji od slika 184 subjekata. Slike su izrađene u visokoj rezoluciji uz jednaku pozadinu te kontrolirane uvjete osvjetljenja. Najveća razlika u godinama unutar jednog para subjekata je 30, a prosječna 4.6 godina.

Drugi skup slika zove se *LQFaces* te se sastoji od 196 subjekata. Ove slike preuzete su s interneta, a velik broj subjekata su poznate osobe. Ove slike razlikuju se u rezoluciji, pozama, izrazima lica te uvjetima osvjetljenja.



Slika 2.2. Primjeri slika u SiblingsDB bazi podataka

3. IMPLEMENTACIJA

Praktični dio rada bavi se implementacijom i testiranjem nekih od popularnijih metoda za automatsko prepoznavanje lica. Metode navedene u poglavlju 2 ovog rada implementirane su u C++ programskom jeziku uz korištenje OpenCV i VisageSDK programskih alata. OpenCV korišten je za manipulaciju i prikaz fotografija, dok je VisageSDK korišten za pronalaženje značajki lica na učitanim slikama. Za testiranje odabranih metoda korištena je Labeled Faces in the Wild baza slika lica. Svaki algoritam prolazi kroz trening fazu, gdje kroz algoritam prolaze slike iz skupa galerije za koje su identiteti osoba poznati. Potom se algoritam testira tako da se za slike iz testnog skupa pokušava odrediti odgovarajuća slika iste osobe iz skupa galerije. Uspješnost algoritma promatra se kroz postotak ispravno uparenih slika.

3.1. Vektorska udaljenost između značajki

Za svaku sliku lica iz zadanog skupa galerije vrši se detekcija lica pomoću prikladnih VisageSDK metoda. Budući da su slike u bazi takve da je navedena osoba uvijek u središtu slike, ukoliko se detekcijom prepozna više lica na slici, odbacuju se sva osim onog najbližeg središtu. U zasebnu varijablu sprema se udaljenost između središta zjenica oba oka. Potom se računaju udaljenosti između odabranih značajki lica te ih se normalizira dijeljenjem njihove vrijednosti s vrijednošću udaljenosti između središta zjenica oba oka. Izračunati skup vrijednosti sprema se u pričuvnu memoriju kao svojevrsni identifikator lica.

Nakon što se navedeni proces izvrši za svaku sliku lica iz galerije, sličan postupak vrši se za slike iz testnog skupa. Nakon što se za sliku iz testnog skupa izračunaju i

normaliziraju sve predodređene udaljenosti, dobiveni skup vrijednosti uspoređuje se sa spremljenim identifikatorima lica. Uspoređuju se vrijednosti odgovarajućih udaljenosti testne slike i slike iz galerije, a akumulirana razlika sprema se u zasebnu varijablu. Slika iz galerije za koju je akumulirana razlika najmanja odabire se kao par testne slike.

Rezultati testiranja ove metode tek su zanemarivo bolji od nasumičnog pogađanja identiteta osobe na slici. S obzirom na to da su razlike u udaljenostima između značajki promatranih na dva različita lica često manje nego razlike u udaljenostima između dva različita izraza istog lica, ova metoda je vrlo osjetljiva na promjene u izrazu lica te će u većini slučajeva upariti dva najbližija izraza lica. S druge strane, budući da se metoda oslanja isključivo na udaljenosti između značajki, izrazito je otporna na promjene u osvjetljenju i sličnim vanjskim uvjetima. Druga prednost algoritma je jednostavnost izračuna, a samim time i brzina.

Broj udaljenosti između značajki koje se uzimaju u obzir ima zanemariv utjecaj na rezultate – čak i u najboljem slučaju, rezultati su znatno lošiji od svih ostalih isprobanih metoda.

3.2. Shape unit

Proces pripreme skupa galerije svodi se na detekciju lica na slici te spremanje *shape unit* težinskih vrijednosti za lice najbliže središtu slike. Potom se za svaku sliku iz testnog skupa provodi isti postupak te se težinske vrijednosti slike iz testnog skupa uspoređuju sa spremljenim težinskim vrijednostima skupa galerije. Uparuju se slike kod kojih je razlika između težinskih vrijednosti slike iz skupa galerije i slike testnog skupa najmanja.

Iako su korištenjem ove metode rezultati i dalje značajno lošiji od nekih naprednijih metoda, zanimljivo je uočiti da su višestruko bolji od rezultata metode vektorskih udaljenosti između značajki.

3.3. Normalizirana križna korelacija

U ovom slučaju koristi se mogućnost VisageSDK alata da pronađeno lice na slici normalizira na način da se značajke lica uvijek smjeste na isto mjesto na slici. Iako se tim postupkom naizgled gube vrijedne informacije topologije lica, korištenje normaliziranog lica daje bolje rezultate u svim metodama koje se ne baziraju isključivo na mjerenju udaljenosti između značajki lica. Normalizacija lica umanjuje negativne efekte različitih pozicija i rotacija lica te, u nešto manjoj mjeri, izraza lica.



Slika 3.1. Normalizirana slika lica glumice Jennifer Garner

Postupak pripreme zadanog skupa galerije za ovu metodu svodi se na detekciju lica te normaliziranja lica najbližeg središtu slike. Nakon što se isto ponovi za sliku iz testnog skupa, nad dvije slike vrši se uparivanje predložaka. Predložak se postepeno pomiče za u piksela po x osi te v piksela po y osi slike na kojoj se traži predložak te se za svaku poziciju (u, v) predloška računa mjera γ u kojoj predložak odgovara slici. Izračun mjere γ prikazan je u (1.2). Rezultirajuća matrica koja sadrži

vrijednosti mjere γ u svakoj točki slike može se interpretirati kao crno-bijela slika. Točka u kojoj je dobivena slika najsvjetlija (odnosno, ćelija matrice u kojoj je vrijednost mjere γ najveća) je točka u kojoj se predložak najviše podudara sa slikom.



Slika 3.2. NCC uparivanje predložaka

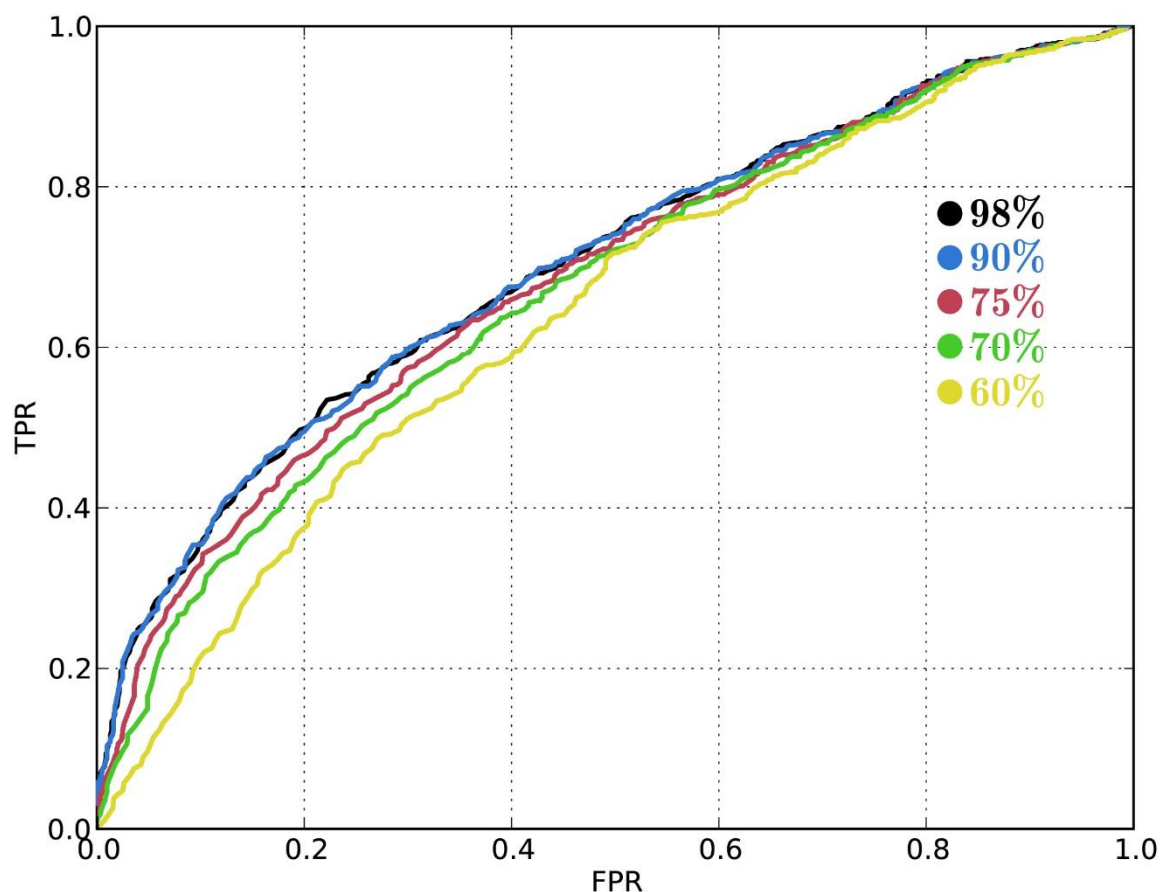
Uparivanje predložaka često je dugotrajan i računski zahtjevan proces, no budući da je u ovom slučaju veličina slike iz skupa galerije jednaka veličini slike iz testnog skupa, neće biti nikakvog pomicanja te je potrebno mjeru γ računati samo jednom da bi se dobila mjera sličnosti dviju slika. S obzirom na to, dodatne optimizacije nisu potrebne.

Rezultati ove metode osjetno su bolji od prethodno isprobanih metoda, no nedovoljno dobri za praktičnu primjenu. Budući da se pri uparivanju predložaka u obzir uzimaju pojedinačni pikseli, a geometrija značajki se zanemaruje, metoda je osjetljiva na promjene u osvjetljenju. Ipak, u ovom slučaju utjecaj tih promjena djelomično je umanjen zbog činjenice da je normalizirana slika monokromatska.

Jedine optimizacije koje je moguće primijeniti na ovu metodu vezane su uz brzinu izvođenja algoritma pa je daljnje istraživanje metode suvišno.

3.4. Svojstvena lica

Pri implementaciji ove metode korištene su metode računanja svojstvenih lica te projiciranja slike u prostor niže dimenzionalnosti dostupne u OpenCV programskoj knjižnici. Priprema slika iz zadanog skupa galerije potrebna za izvođenje tih metoda uključuje detekciju lica te normaliziranja lica najbližeg središtu slike, nakon čega se sve normalizirane slike spremaju u jednu vektorsku matricu. Potom se nad dobivenom matricom računa PCA postupak. Pri tome je smanjenje dimenzionalnosti moguće odrediti na dva načina – odabirom željenog konačnog broja dimenzija ili odabirom postotka informacija koji se želi sačuvati u ciljnom prostoru. Dimenzionalnost je značajno smanjena čak i kad je postotak zadržanih informacija vrlo visok. Utjecaj odabranog postotka koji se želi zadržati prikazan je na slici 3.3.



Slika 3.3. Utjecaj postotka zadržane informacije na rezultate

Rezultati PCA postupka su prostor niže dimenzionalnosti u koji je potom moguće projicirati sliku lica, središnje lice skupa te svojstvene vrijednosti i vektori tog prostora. Nakon što je postupak završen, potrebno je svaku sliku iz zadanog skupa galerije projicirati u dobiveni PCA prostor.

Pripremanje slika iz testnog skupa vrši se na isti način – nakon detekcije lica odabire se lice najbliže središtu te se slika normalizira. Tako normalizirana slika potom se projicira u PCA prostor. Budući da projicirane slike imaju oblik vektorske matrice, uspoređivanje slika iz testnog skupa i skupa galerije svodi se na računanje akumulirane razlike između dva vektora. Uparuju se slike s najmanjom akumuliranom razlikom.

Korištenje ove metode daje puno bolje rezultate od prethodno isprobanih. U slučaju kontroliranih uvjeta te malih razlika u rotaciji glave, pomaku glave i izrazima lica, ova metoda može se uz zadovoljavajuće rezultate koristiti čak i u aplikacijama s realnim primjenama. Ipak, budući da je PCA postupak alat koji je moguće primijeniti na bilo koji set podataka, ova metoda najčešće se koristi u kombinaciji s nekom drugom metodom.

3.5. Local binary patterns

Priprema slika iz skupa galerije ovdje se viši na isti način kao i dosad – detektiraju se lica na slici, odabire se ono najbliže središtu slike te se odabrano lice normalizira. Budući da se u sklopu LBP postupka izrađuje histogram slike, normalizacija vrijednosti monokromatske verzije slike bitan je korak.

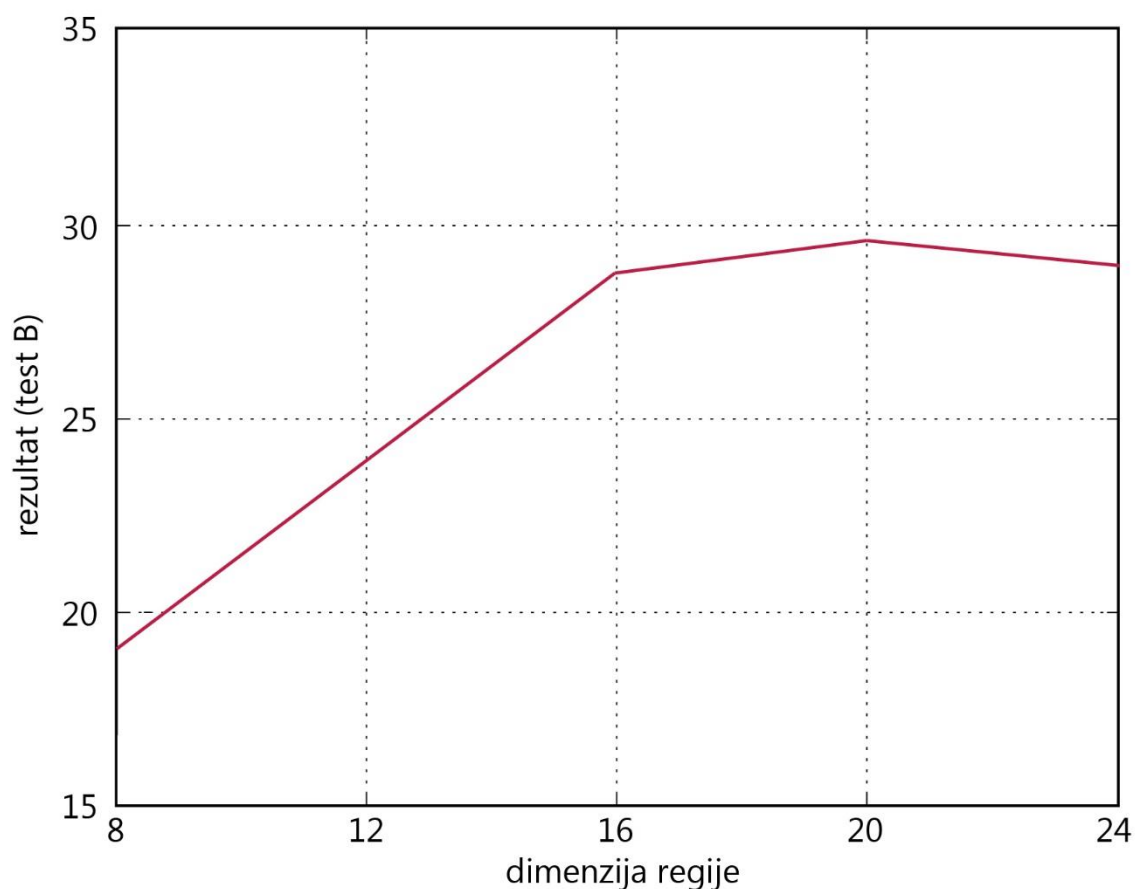
Za svaku sliku u skupu galerije stvaramo matricu koju potom punimo odgovarajućim LBP kodovima te slike koristeći (1.12). Nakon toga se računaju histogrami dobivenih matrica te se isti spremaju u memoriju.

Postupak pripreme i obrade slika iz testnog skupa je isti. Parovi se određuju računanjem distance između histograma pojedinih slika iz seta galerije i testnog seta. Uparuju se one slike između kojih je distanca najmanja.

Ukoliko se nad cijelom površinom slike računa jedan histogram, rezultati dobiveni korištenjem ove metode gotovo su potpuno beskorisni. Tako dobiveni rezultati tek su malo bolji od rezultata dobivenih metodom vektorske udaljenosti između značajki te znatno lošiji od rezultata dobivenih metodom svojstvenih lica.

Bolje rješenje je računati lokalizirane histograme te ih naposljetku povezati u jedan globalni histogram. Na taj način zadržavaju se lokalne, ali i globalne prostorne informacije, a veličina zapisa jedne slike i dalje je mala.

Pri tome je potrebno odrediti dimenzije pojedinačne regije nad kojom se računa histogram, kao i broj piksela preklapanja susjednih regija. U praktičnim testovima pokazalo se da preklapanje regija povećava vrijeme izračuna, a ne donosi primjetnu razliku u rezultatima. S druge strane, dimenzije pojedinačne regije pokazale su se kao bitan faktor u primjeni metode. Utjecaj dimenzije pojedinačne regije na rezultate prikazan je na slici 3.4. Zanimljivo je uočiti blagi pad rezultata kad dimenzija dovoljno poraste, što se slaže s rezultatima pri računanju jednog histograma nad cijelom slikom – taj slučaj možemo promatrati kao lokalizirani histogram kojemu je dimenzija regije jednaka dimenziji slike.



Slika 3.4. Utjecaj dimenzije regije na rezultate

Zbog činjenice da je robusna u odnosu na razlike u uvjetima osvjetljenja, LBP metoda je pogodna za korištenje pri nekontroliranim uvjetima, što se vidi iz rezultata dobivenih testiranjem na LFW bazi slika.

3.6. LBP + PCA

Implementacija kombinacije LBP deskriptora lica i metode svojstvenih lica jednostavna je zbog činjenice da je moguće primijeniti PCA postupak na regionalne histograme dobivene kao rezultat LBP postupka.

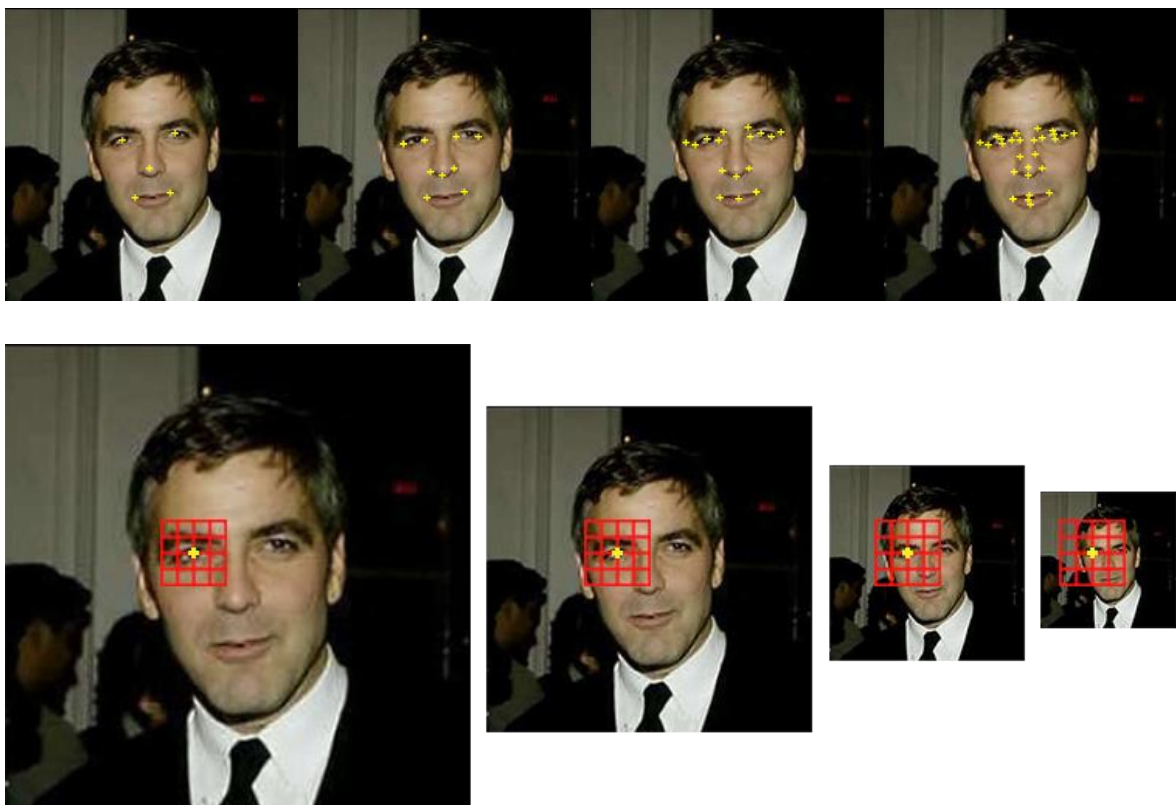
Za svaku sliku u zadanom skupu galerije računa se LBP deskriptor normaliziranog lica, nakon čega se računaju lokalizirani histogrami koji se na kraju povezuju u vektorsku matricu koja sadržava podatke o licu. Matrice pojedinih lica se u konačnici povezuju u jednu vektorsku matricu koja sadržava podatke o svim licima u bazi podataka. Nad tim podacima potom se izvodi PCA postupak kako bi se odredio nižedimenzionalan prostor u koji se projiciraju slike. Potom se sve slike iz skupa galerije projiciraju u rezultirajući prostor, a dobiveni rezultati spremaju se u memoriju u obliku vektorskih matrica.

Priprema testnog skupa izgleda slično – za svaku sliku u skupu računa se LBP deskriptor normaliziranog lica te lokalizirani histogrami koji se povezuju u jednu vektorsku matricu. Nakon toga se dobivena matrica projicira u prethodno određeni PCA prostor te dobivamo vektorsku matricu koja opisuje lice te je jednakih dimenzija kao matrice koje opisuju lica iz skupa galerije.

Računaju se udaljenosti između projiciranih lokaliziranih histograma te se uparuju oni između kojih je ta udaljenost najmanja.

Rezultati testiranja kombinacije ovih dviju metoda osjetno su bolji od svih prethodno testiranih metoda, uključujući obje te metode pojedinačno. Metoda svojstvenih lica osjetljiva je na promjene u vanjskim uvjetima poput osvjetljenja, pa je intuitivno da korištenje LBP deskriptora lica, koji je izrazito robustan u odnosu na varijabilne uvjete osvjetljenja, daje nezanemarivo bolje rezultate pri radu s bazom podataka poput Labeled Faces in the Wild.

Jedna od predloženih modifikacija pri korištenju kombinacije ovih metoda je primjena metode samo na određenim ključnim točkama oko istaknutih značajki lica umjesto na području cijelog lica te pritom za svaku točku zapisati nekoliko histograma za okvire različitih veličina i rezolucija. Pri tome najveći okvir može pokrivati i do polovice cijelog lica kako bi se zadržale informacije o licu na globalnoj razini.



Slika 3.5. Uzorkovanje okvira različitih veličina oko značajki

4. REZULTATI

Za potrebe testiranja metoda na Labeled Faces in the Wild bazi podataka izrađena su četiri testna slučaja. Testovi A, B i C pokrivaju 20, 200, odnosno 1000 nasumično odabranih osoba iz baze podataka, pri čemu su slike odabranih osoba za skup galerije i testni skup također određene nasumičnim odabirom. Test D pokriva 200 osoba, ali osobe su, kao i pripadajuće slike za skup galerije i testni skup, odabrane ručno. Naime, kako se LFW baza sastoji od slika javnih osoba prikupljenih s interneta, nezanemariv broj slika je suboptimalan za primjenu predstavljenih metoda – velike su razlike u uvjetima osvjetljenja, pozama, rotaciji glave, izrazima lica, količini šminke, dok je kod nekih slika lice djelomično prekriveno naočalama, rukom ili nekim drugim predmetom. Nekoliko primjera takvih slika nalazi se na slici 4.1.



Slika 4.1. Primjeri problematičnih slika

U tablici 4.1. prikazani su rezultati testiranja u obliku postotaka ispravnih uparivanja.

Metoda	A (20)	B (200)	C (1000)	D (200*)
Vektorska udaljenost	20,00	2,22	0,77	7,83
Shape units	30,00	7,78	4,60	20,87
NCC	50,00	12,70	11,61	36,09
PCA	55,00	25,82	20,90	66,09
LBP	55,00	29,83	25,14	68,04
LBP + PCA	65,00	41,34	31,83	71,00

Tablica 4.1.

Iz rezultata je vidljivo koliki je napredak ostvaren modernim metodama u odnosu na prvu isprobanu metodu – uspoređivanje vektorskih udaljenosti između značajki na licu.

Bitno je uočiti i značajnu razliku između rezultata testova B i D. Iako se sastoje od jednakog broja različitih osoba, rezultati testa D konzistentno su bitno bolji od rezultata testa B. Ta razlika ukazuje na velik utjecaj koji uvjeti poput osvjetljenja, pozicije, izraza lica i slično imaju na rezultate. Test D je ručno odabran, no i dalje se radi o slikama izrađenim pri nekontroliranim uvjetima, što znači da su razlike u istim tim uvjetima nezanemarive čak i u takvom testu. Samim time, test D je najbolja aproksimacija rezultata u većini stvarnih aplikacija.

Metoda koja je u svim testovima pokazala najbolje rezultate zapravo je kombinacija dviju najboljih prethodno testiranih metoda. Prijedlog za daljnje poboljšanje rezultata je proširenje metode primjenom na manja područja oko bitnih značajki lica umjesto pridavanja jednakog značaja cijelom licu.

ZAKLJUČAK

Lice je zbog količine informacija koje nosi o identitetu i emocijama osobe primarna točka pažnje u socijalnoj interakciji. Ljudska sposobnost prepoznavanja lica je izvanredna – uspijevamo prepoznati tisuće lica na temelju kratkog pogleda čak i ukoliko drugu osobu nismo vidjeli godinama.

Računalni modeli prepoznavanja lica interesantni su ne samo zbog teorijskog razumijevanja procesa identifikacije, već i zbog praktične primjenjivosti u širokom spektru ljudske djelatnosti. Područje automatskog raspoznavanja lica doživjelo je značajan rast i razvoj u posljednjih nekoliko desetljeća, a u posljednjih deset godina u komercijalnoj upotrebi pojavile su se brojne tehnologije koje u svom radu koriste prepoznavanje lica.

Ovaj rad pokazuje rezultate testiranja nekih od često korištenih metoda u procesu računalne identifikacije lica. Iz rezultata je očita bitnost kontroliranih uvjeta za dobivanje dobrih rezultata.

Iako su prikazani rezultati obećavajući, trenutne metode računalne identifikacije lica primjenjive su samo uz kontrolirane uvjete osvjetljenja te što manje odstupanja od frontalne pozicije i neutralnog izraza lica.

POPIS LITERATURE

- [1] Chen, D., Cao, X., Wen, F., Sun, J. Blessing of Dimensionality: High-dimensional Feature and Its Efficient Compression for Face Verification. Kina: University of Science and Technology of China, 2013.
- [2] Turk, M. A., Pentland, A. P. Face Recognition Using Eigenfaces. SAD: Massachusetts Institute of Technology, 1991.
- [3] Welling, M. Fisher Linear Discriminant Analysis. Kanada: University of Toronto, 2005.
- [4] Wiskott, L., Fellous, J. M., Krüger, N., Malsburg, C. Face Recognition by Elastic Bunch Graph Matching. Njemačka: Institut für Neuroinformatik, Ruhr-Universität Bochum, 1996.
- [5] Salah, A. A., Bicego, M., Akarun, L., Grosso, E., Tistarelli, M. Hidden Markov Model-based face recognition using selective attention. Turska: Boğaziçi University Istanbul, 2007.
- [6] Lu, H. Multilinear Subspace Learning for Face and Gait Recognition. Kanada: University of Toronto, 2008.
- [7] Sayeed, F., Hanmandlu, M., Ansari, A. Q. Face Recognition using Segmental Euclidean Distance. Defence Science Journal. Svezak 61, br. 5 (2011), str. 431-442
- [8] Turk, M., Pentland, A. Eigenfaces for Recognition. SAD: Massachusetts Institute of Technology, 1991.
- [9] Ahonen, T., Hadid, A., Pietikäinen, M. Face Description with Local Binary Patterns: Application to Face Recognition. Finska: University of Oulu, 2006.
- [10] Brunelli, R., Poggio, T. Face Recognition: Features versus Templates. IEEE Transactions on Pattern Analysis & Machine Intelligence. Svezak 15, br. 10 (1993), str. 1042-1052
- [11] Ojala, T., Pietikäinen, M., Mäenpää. Multiresolution Gray Scale and Rotation Invariant Texture Classification with Local Binary Patterns. Finska: University of Oulu, 2002.
- [12] OpenCV, <http://opencv.org/>, siječanj 2015.
- [13] VisageSDK, <http://www.visagetechologies.com/products/visagesdk/>, siječanj 2015.
- [14] Visual Studio, <http://www.visualstudio.com/>, siječanj 2015.

DODATAK A – KORIŠTENI ALATI

U tablici su prikazani alati korišteni pri izradi završnog rada.

Alat	Opis	Dostupnost	Referenca
C++	Programski jezik	-	-
OpenCV	Programska knjižnica	Otvoreni kod	[12]
VisageSDK	Programska knjižnica	Komercijalna licenca	[13]
Visual Studio	Razvojna okolina	Komercijalna licenca	[14]

SAŽETAK

Automatsko raspoznavanje lica

Ovaj rad daje pregled trenutnog stanja područja automatskog raspoznavanja lica uz fokus na nekoliko najpopularnijih metoda koje omogućavaju računalnu identifikaciju lica – vektorska udaljenost između značajki, shape units, normalized cross correlation, PCA te LBP metoda. Metode su predstavljene na teorijskoj razini, nakon čega su prikazane i neke od dostupnih baza slika lica. Implementacija je izvedena u C++ programskom jeziku uz korištenje OpenCV i VisageSDK programskih alata te Labeled Faces in the Wild bazu podataka, a koraci implementacije prikazani su u sklopu ovog rada. Rezultati testiranja opisanih metoda također su prikazani i komentirani.

Ključne riječi: automatsko raspoznavanje lica, vektorska udaljenost značajki lica, značajke lica, normalized cross correlation, PCA, LBP, svojstveno lice, OpenCV, VisageSDK, LFW

ABSTRACT

Automatic face recognition

This paper provides an overview of the state in the field of facial recognition, with focus on some of the more popular methods used in the field – vector distance between facial features, shape units, normalized cross correlation, principal component analysis and local binary pattern methods. The methods are described on a theoretical level, followed by an overview of some of the available databases related to facial recognition. The described methods have been tested against the Labeled Faces in the Wild database by developing a test environment in C++ with the use of OpenCV and VisageSDK tools. Development steps have been presented, as well as the results.

Keywords: automatic face recognition, vector distance between facial features, facial features, normalized cross correlation, principal component analysis, local binary patterns, eigenface, OpenCV, VisageSDK, labeled faces in the wild