



SVEUČILIŠTE U ZAGREBU
FAKULTET ELEKTROTEHNIKE I RAČUNARSTVA

Ivan Fratrić

**BIOMETRIJSKA VERIFIKACIJA OSOBA
TEMELJENA NA ZNAČAJKAMA DLANA
I LICA DOBIVENIM IZ VIDEO SEKVENCI**

DOKTORSKI RAD

Zagreb, 2011



SVEUČILIŠTE U ZAGREBU
FAKULTET ELEKTROTEHNIKE I RAČUNARSTVA

Ivan Fratrić

**BIOMETRIJSKA VERIFIKACIJA OSOBA
TEMELJENA NA ZNAČAJKAMA DLANA
I LICA DOBIVENIM IZ VIDEO SEKVENCI**

DOKTORSKI RAD

Zagreb, 2011



UNIVERSITY OF ZAGREB
FACULTY OF ELECTRICAL ENGINEERING AND COMPUTING

Ivan Fratrić

**BIOMETRIC PERSONAL VERIFICATION
BASED ON PALM AND FACE FEATURES
OBTAINED FROM VIDEO SEQUENCES**

DOCTORAL THESIS

Zagreb, 2011

Doktorski rad je izrađen na Sveučilištu u Zagrebu, Fakultetu elektrotehnike i računarstva, Zavodu za elektroniku, mikroelektroniku, računalne i inteligentne sustave.

Mentor: prof.dr.sc. Slobodan Ribarić

Doktorski rad ima 132 stranice

Doktorski rad br.:

Povjerenstvo za ocjenu i obranu doktorske disertacije:

1. Dr.sc. Zoran Kalafatić, izvanredni profesor
Sveučilište u Zagrebu, Fakultet elektrotehnike i računarstva
2. Dr.sc. Nikola Pavešić, redoviti profesor
Univerza v Ljubljani, Fakulteta za elektrotehniko
3. Dr.sc. Ivo Ipšić, redoviti profesor
Sveučilište u Rijeci, Tehnički fakultet

Datum obrane disertacije: 25. ožujka 2011. godine

Sažetak: Dlan i lice su dvije biometrijske karakteristike s velikim brojem značajki pogodnih za korištenje u biometrijskoj verifikaciji. Za razliku od većine dosadašnjih biometrijskih sustava temeljenih na značajkama dlana gdje se zahtjeva fizički kontakt korisnika i senzora, u ovoj se disertaciji razmatra korištenje značajki dlana, kao i značajki lica, dobivenih iz video sekvenci. Opisana je nova metoda lokalizacije interesnog područja dlana robusna na složene uvjete snimanja u smislu pozadine i osvjetljenja. Predložena je nova metoda izlučivanja lokalnih binarnih diskriminatornih značajki, te metoda za ocjenu kvalitete interesnog područja dlana u video sekvencama. U provedenim eksperimentima pokazana je pogodnost predloženih metoda za biometrijsku verifikaciju. Također je pokazano da se tehnikama biometrijske fuzije može ostvariti bolja točnost verifikacije u odnosu na verifikaciju koja bi se temeljila samo na značajkama dlana ili samo na značajkama lica.

Ključne riječi: Biometrija, video, dlan, lice, lokalizacija, verifikacija

Abstract: Palm and face are two biometric characteristics that contain a variety of features suitable for biometric verification. Unlike the majority of biometric systems based on palm features, which require the physical contact of a user and a biometric sensor, in this dissertation the use of palm, as well as face features, obtained from video sequences is considered. A new method of palm region of interest localization, which is robust to complex background and lighting conditions, is described. A new method for the extraction of local binary discriminatory features and a new method for quality assessment of palm region of interest in video sequences are proposed as well. The performed experiments show the feasibility of the proposed methods for the biometric verification. It has also been demonstrated that by use of biometric fusion techniques, better verification accuracy can be achieved compared to verification based only on palm or only on face features.

Key words: Biometrics, video, palm, face, localization, verification

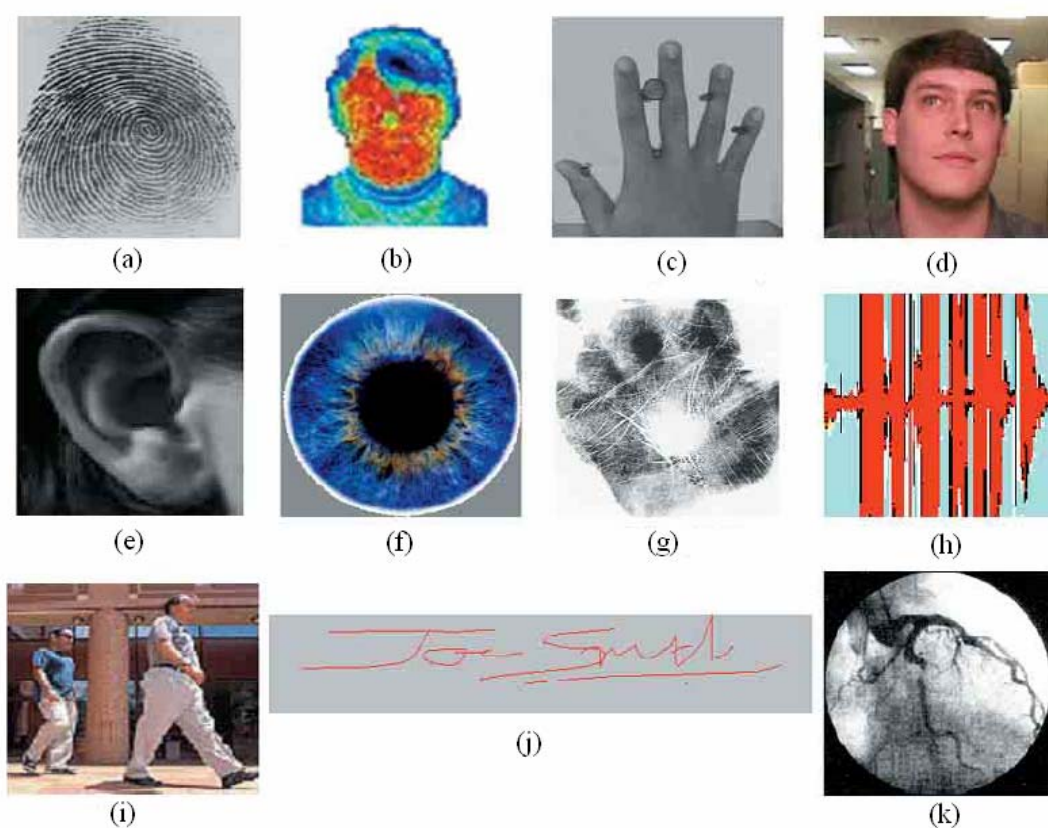
Sadržaj

1.	Uvod	3
2.	Pregled srodnih radova	9
2.1.	Sustavi za raspoznavanje dlana	9
2.2.	Sustavi za raspoznavanje lica	13
2.3.	Raspoznavanje lica u video sekvencama	17
2.4.	Beskontaktna biometrija ruke	18
2.5.	Fuzija dlana i lica	20
3.	Beskontaktna verifikacija temeljena na značajkama dlana i lica	22
4.	Lokalizacija lica i ruke	26
4.1.	Lokalizacija lica	26
4.2.	Lokalizacija ruke	32
4.2.1.	Detekcija kandidata ruke	33
4.2.2.	Lokalizacija na temelju modela	33
4.2.3.	Izlučivanje interesnog područja dlana	39
4.2.4.	Rezultati lokalizacije	40
5.	Normalizacija interesnih područja	43
6.	Izlučivanje značajki dlana i lica	46
6.1.	Analiza glavnih komponenti	46
6.1.1.	Izvod Analize glavnih komponenti	46
6.1.2.	Analiza glavnih komponenti za skup slikovnih uzoraka	49
6.2.	Linearna diskriminantna analiza	50
6.3.	Izlučivanje značajki Gaborovim filtrom	56
6.4.	Lokalna binarna LDA	59
7.	Ocjena kvalitete područja interesa dlana	63
8.	Fuzija u biometriji	67
8.1.	Fuzija na razini senzora	67
8.2.	Fuzija na razini vektora značajki	68
8.3.	Fuzija na razini mjere podudaranja	69
8.4.	Fuzija na razini odluke	71
9.	Eksperimentalna evaluacija	73
9.1.	Baza video sekvenci	73
9.2.	Baze za određivanje parametara	75
9.3.	Evaluacija metoda izlučivanja značajki	76

9.3.1.	Analiza glavnih komponenti	77
9.3.2.	Linearna diskriminantna analiza	78
9.3.3.	Gaborov filter	82
9.3.4.	Lokalna binarna LDA	85
9.3.5.	Usporedba rezultata	91
9.3.6.	Usporedba brzine izvođenja	96
9.4.	Evaluacija mjere kvalitete područja interesa dlana	98
9.5.	Evaluacije fuzije dlana i lica	102
9.5.1.	Fuzija na razini mjere podudaranja	103
9.5.2.	Fuzija na temelju klasifikatora	107
9.5.3.	Fuzija na temelju glasanja	112
10.	Zaključak	115
	Popis literature	118
	Popis oznaka	124
	Prilog A: Programska implementacija	127
A.1.	Aplikacija za lokalizaciju lica i ruke	127
A.2.	Format datoteke za zapis značajki	128
A.3.	Aplikacija za ocjenu mjere kvalitete interesnog područja dlana	129
A.4.	Biblioteka za izlučivanje značajki i evaluaciju	130

1. Uvod

Biometrija [1][2] je grana znanosti i tehnologije kojoj je cilj identificirati osobe ili verificirati identitet osoba na temelju njihovih fizioloških ili ponašajnih karakteristika. Fiziološke karakteristike korištene u biometriji uključuju otisak prsta, lice, termogram lica, dlan, geometriju ruke, šarenicu i mrežnicu oka itd. Ponašajne karakteristike uključuju govor, potpis, hod, dinamiku tipkanja itd. Slika 1 ilustrira neke fiziološke i ponašajne karakteristike korištene u biometriji.



Slika 1: Neke od biometrijskih karakteristika: (a) otisak prsta, (b) termogram lica, (c), geometrija ruke, (d) lice, (e) uho, (f) šarenica oka, (g) otisak dlana, (h), govor, (i) hod, (j) potpis, (k) krvne žile u oku. Slika preuzeta iz [2].

Prednost biometrijskih sustava nad „klasičnim“ oblicima identifikacije pomoću ključeva, lozinki, kartica i sl. je u tome što se biometrijske karakteristike ne mogu lako izgubiti, otuđiti ili zaboraviti, te biometrija zahtjeva da osoba bude prisutna na mjestu identifikacije ili verifikacije. Tako biometrija odgovara na pitanje „Tko sam ja?“ ili „Jesam li ja zaista onaj za koga se izdajem?“ umjesto na pitanja „Što znam“ ili „Što posjedujem“.

Najraširenija primjena biometrije je u sigurnosti, za kontrolu fizičkog pristupa, no biometrija se također primjenjuje u drugim područjima, kao što su forenzika ili inteligentna sučelja čovjeka i računala.

Biometrijski sustavi mogu se podijeliti na sustave za identifikaciju i verifikaciju. Zadaća sustava za identifikaciju je na temelju biometrijskih karakteristika osobe odrediti o kojoj se osobi radi, tj. takvi sustavi odgovaraju na pitanje „Tko sam ja?“. U sustavima za biometrijsku verifikaciju, korisnik sustava će, osim biometrijske karakteristike, dati sustavu i naznaku svoga identiteta (u vidu korisničkog imena, lozinke, PIN-a i sl.), te je zadaća sustava na temelju biometrijske karakteristike provjeriti radi li se zaista o danom identitetu. Drugim riječima, sustavi za verifikaciju odgovaraju na pitanje „Jesam li ja zaista onaj za koga se izdajem?“.

Biometrijska identifikacija nadalje se može podijeliti na identifikaciju u zatvorenom skupu i identifikaciju u otvorenom skupu. Kod identifikacije u zatvorenom skupu unaprijed je poznato da je osoba koja pristupa sustavu tome sustavu već unaprijed poznata i ima svoje biometrijske značajke pohranjene u bazi sustava. Zadaća identifikacije je tada samo odrediti o kojoj se osobi radi. Kod identifikacije u otvorenom skupu ovo ne mora biti slučaj te osoba koja pristupa sustavu ne mora biti poznata od prije. Biometrijski sustav za identifikaciju u otvorenom skupu bi trebao ispravno detektirati i prijaviti ovakav pokušaj identifikacije.

Biometrijski sustav obično radi u dvije faze: fazi prijave i fazi autorizacije. U fazi prijave korisnik sustava se svojim biometrijskim značajkama po prvi puta prijavljuje na sustav. Izlučene biometrijske značajke tada se spremaju u bazu podataka biometrijskog sustava. U fazi autorizacije korisnik zahtijeva od sustava identifikaciju ili verifikaciju. Pri tome se izlučene značajke uspoređuju s onima pohranjenima u bazi biometrijskog sustava u fazi prijave te biometrijski sustav na temelju ove usporedbe donosi odluku o rezultatu identifikacije ili verifikacije.

U biometrijskim sustavima moguće su dvije vrste pogrešaka. Pogreška lažnog odbijanja (engl. False Rejection, FR) nastupa kada biometrijski sustav odbije osobu koju je trebao propustiti. Druga vrsta pogreške, pogreška lažnog prihvatanja (engl. False Acceptance, FA) nastupa kada sustav pogrešno propusti osobu koju je trebao odbiti. Sukladno ovome, moguće su i dvije vrste ispravnih odluka. Ispravno prihvatanje (engl. Genuine Acceptance, GA) nastupa kada sustav ispravno prihvati osobu koju je i trebao prihvatiti. Ispravno odbijanje (engl. Genuine Rejection, GR) nastupa kada sustav ispravno odbije osobu koja nije korisnik sustava (u slučaju identifikacije) ili se sustavu pokušala prijaviti pod tuđim identitetom (u slučaju verifikacije).

1. UVOD

Na temelju ovih slučajeva moguće je izraziti osnovne mjere točnosti biometrijskih sustava. Udio lažnog prihvaćanja (engl. False Acceptance Rate, FAR) i udio lažnog odbijanja (engl. False Rejection Rate, FRR) mogu se izračunati na temelju broja ovih slučajeva na sljedeći način

$$FRR = \frac{FR}{GA + FR} \quad (1)$$

$$FAR = \frac{FA}{GR + FA} \quad (2)$$

Ove pogreške najčešće se izražavaju u postocima. Točan iznos FAR i FRR najčešće će ovisiti o nekom pragu koji sustav koristi za identifikaciju ili verifikaciju. Mijenjanjem ovog praga može se FRR smanjiti na uštrb FAR ili obrnuto. Zbog toga se točnost biometrijskog sustava često umjesto iznosom FAR i FRR za neki konkretan prag prikazuje kao ROC (engl. Receiver Operating Characteristic) krivulja. Ova krivulja na x osi ima FAR, a na y osi GAR (engl. Genuine Acceptance Rate, $GAR = 1 - FRR$). Na temelju ROC krivulje korisnik sustava može odabrati u kojoj točki želi da sustav radi. Za sigurnosne aplikacije, na primjer, poželjniji će biti manji FAR nego FRR, dok će za forenzičke aplikacije biti poželjnije imati manji FRR.

Kao jedinstvene mjere pogreške sustava najčešće se koriste udio jednakih pogrešaka (engl. Equal Error Rate, EER) i minimalni ukupni udio pogrešaka (engl. minimum Total Error Rate, mTER), odnosno polovina minimalnog ukupnog udjela pogreške ($HmTER = mTER/2$). EER se dobiva kao udio pogrešaka u točki ROC krivulje gdje su FAR i FRR jednaki. mTER je minimalna dobivena vrijednost za sumu FAR i FRR.

Sama točnost često nije dovoljna da se okarakterizira biometrijski sustav. Prilikom izgradnje ili nabavke biometrijskog sustava treba uzeti u obzir niz drugih faktora biometrijskog sustava, odnosno karakteristike na kojoj se sustav zasniva, kao što su:

- Univerzalnost – posjeduje li svaka osoba ovu karakteristiku
- Jedinstvenost – je li kod bilo koje dvije osobe ova karakteristika različita
- Stalnost – mijenja li se ova karakteristika kroz vrijeme
- Mogućnost prikupljanja – može li se ova karakteristika na jednostavan način prikupiti i prenijeti u računalu
- Učinkovitost – kolika se točnost raspoznavanja postiže uporabom ove karakteristike i koliko je brza obrada ove karakteristike

- Prihvatljivost – koliko su ljudi spremni koristiti ovu značajku za biometrijsku autorizaciju
- Mogućnost prijevare – koliko je teško prevariti sustav koji koristi ovu karakteristiku

Za razliku od jednonačinskih (engl. unimodal) biometrijskih sustava koji koriste samo jednu biometrijsku karakteristiku za identifikaciju ili verifikaciju, često je poželjno kombinirati više karakteristika u tzv. višenačinskim (engl. multimodal) biometrijskim sustavima. Time se može postići bolja točnost sustava, ali i otpornost sustava na pokušaje prijevare. U višenačinskim sustavima postoji potreba za postupcima kombiniranja više biometrijskih karakteristika što se može načiniti na više načina.

U ovome radu za biometrijsku verifikaciju razmatra se korištenje značajki dlana i lica dobivenih iz video sekvenci.

Za dlan kao biometrijsku karakteristiku karakterističan je veliki broj značajki. Shu i Zhang [3] navode više značajki otisaka dlana koje se mogu koristiti za raspoznavanje:

- geometrijske značajke, kao što su širina, visina, te površina dlana
- značajke glavnih linija dlana
- značajke bora, koje su manje i nepravilnije od glavnih linija dlana
- delta-točke, definirane kao središta područja dlana sličnih znaku delta
- minucije, koje odgovaraju značajkama koje se koriste za raspoznavanje otisaka prsta, a zahtijevaju slike visoke rezolucije za izlučivanje

Nedostatak dlana kao biometrijske karakteristike je potreba, u većini do sada razvijenih sustava, fizičkog kontakta ruke i senzora. Ovo umanjuje prihvatljivost korištenja dlana u biometriji, naročito u scenarijima s velikim brojem korisnika, zbog higijenskih ili drugih razloga. Zbog toga se u ovom radu razmatra korištenje dlana dobivenoga iz video sekvenci, bez potrebe fizičkog kontakta ruke i senzora.

Lice, kao biometrijsku karakteristiku, upravo izdvaja mogućnost jednostavnog beskontaktnog prikupljanja, čak i bez znanja osobe koja se snima. Stoga je lice „najprikupljivija“ biometrijska karakteristika, te 19% tržišnog udjela biometrijskih sustava otpada na sustave temeljene na licu [4]. Međutim, u snimkama lica dolazi do velikih varijacija u izgledu iste osobe. Ove varijacije nastaju kao posljedica osvjetljenja, različitog položaja lica, izraza lica, starenja itd. Takve degradacije na slici negativno utječu na postupke lokalizacije i raspoznavanja [5].

Za očekivati je, ukoliko dlan također snimamo kamerom na isti način kao i lice, da će i u snimkama dlana doći do velikih varijacija kao posljedica osvjetljenja, položaja ruke i sl. što će smanjiti točnost ovako snimljenih dlanova za biometrijsku verifikaciju. Također je za

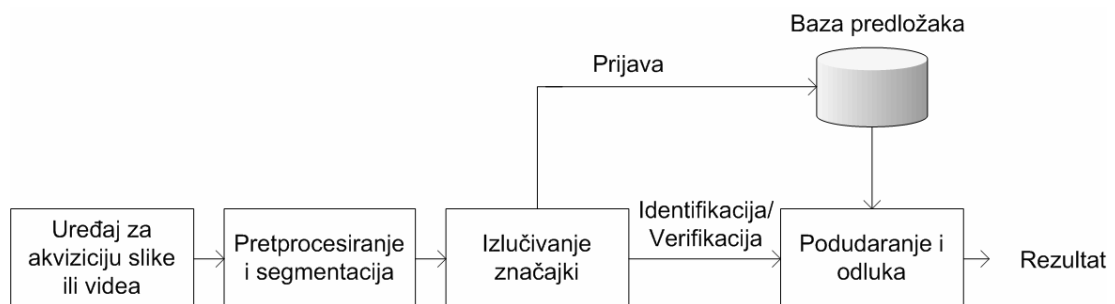
1. UVOD

očekivati, da ukoliko se akvizicija uzoraka obavlja u nekontroliranim uvjetima u smislu strukture pozadine, pokreta u pozadini i osvjetljenja, određivanje regija interesa postati će znatno veći problem nego kod snimaka dlana nastalih skenerom ili CCD kamerom u kontroliranim uvjetima.

Stoga postoji potreba za razvojem metoda lokalizacije ruke, metoda ocjene kvalitete područja interesa i metoda izlučivanja značajki koje bi radile u ovakvim uvjetima i omogućile beskontaktnu biometriju dlana na video sekvencama.

Snimanje dlana u ovakvim uvjetima također otvara mogućnost snimanja lica u istoj video sekvenci. Zbog toga je potrebno razmotriti i ispitati postupke biometrijske fuzije na ovako dobivenim uzorcima.

Biometrijski sustavi temeljeni na dlanu i licu, bez obzira o kakvoj se akviziciji radi i radi li se o video sekvencama ili statičnim slikama imaju uobičajenu zajedničku strukturu koja se može prikazati kao na slici 2.



Slika 2: Uobičajena struktura biometrijskih sustava temeljenih na dlanu i licu

Prvo je potrebno dobiti sliku ili video sekvencu ruke ili lica. Ovo se obično obavlja kamerom za lice, te kamerom ili skenerom za dlan. Nakon akvizicije, na slikama ili okvirima video sekvence provode se postupci preprocesiranja i segmentacije. Cilj ovoga postupka je određivanje gdje se na slici nalaze interesna područja (engl. Region Of Interest, ROI). Nakon lokalizacije, nad ROI se provode postupci izlučivanja značajki. Ukoliko se radi o fazi prijave, izlučene značajke se spremaju u bazu biometrijskog sustava zajedno s identitetom korisnika. Ukoliko se radi o identifikaciji ili verifikaciji, izlučene značajke se uspoređuju s onima prethodno pohranjenima u bazi sustava. Na temelju rezultata usporedbe sustav donosi odluku o prihvatanju ili odbijanju korisnika.

1. UVOD

Ostatak disertacije organiziran je kako slijedi. U poglavlju 2 dan je pregled srodnih radova. Poglavlje 3 opisuje organizaciju sustava za verifikaciju temeljenu na značajkama dlana i lica dobivenim iz video sekvenci. Poglavlje 4 opisuje metode korištene za lokalizaciju lica i ruke. U poglavlju 5 opisani su postupci normalizacije koji su primjenjivani na interesna područja dlana i lica. Poglavlje 6 opisuje metode izlučivanja značajki iz interesnih područja dlana i lica. U poglavlju 7 dan je opis predložene metode za ocjenu kvalitete interesnog područja dlana u video sekvencama. U poglavlju 8 opisani su postupci fuzije u biometriji. U poglavlju 9 opisani su provedeni eksperimenti i dani rezultati eksperimenata. Zaključak je dan u poglavlju 10.

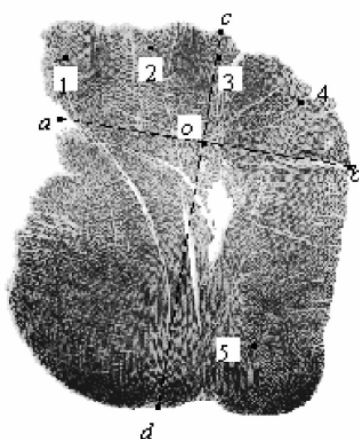
2. Pregled srodnih radova

2.1. Sustavi za raspoznavanje dlana

Sustavi za raspoznavanje dlana mogu se podijeliti na sustave s kasnijom obradom i sustave u stvarnom vremenu. Kod sustava s kasnijom obradom uzorak se uzima kao otisak dlana umočenoga u tintu, koji se zatim skenira i obrađuje. Većina ranijih sustava za raspoznavanje dlana pripadaju ovoj kategoriji [3][6][7][8]. Kod sustava u stvarnom vremenu, slika dlana uzima se uglavnom uz pomoć skenera [9] ili CCD kamere [10]. Segmentacija slika dlana dobivenih na ovaj način je vrlo jednostavna, budući da postoji veliki kontrast između intenziteta ruke i pozadine na slici.

Opisano je i nekoliko sustava u kojima se uzorci uzimaju beskontaktno. Ovakvi sustavi zahtijevaju složenije postupke segmentacije i biti će opisani u posebnom poglavlju.

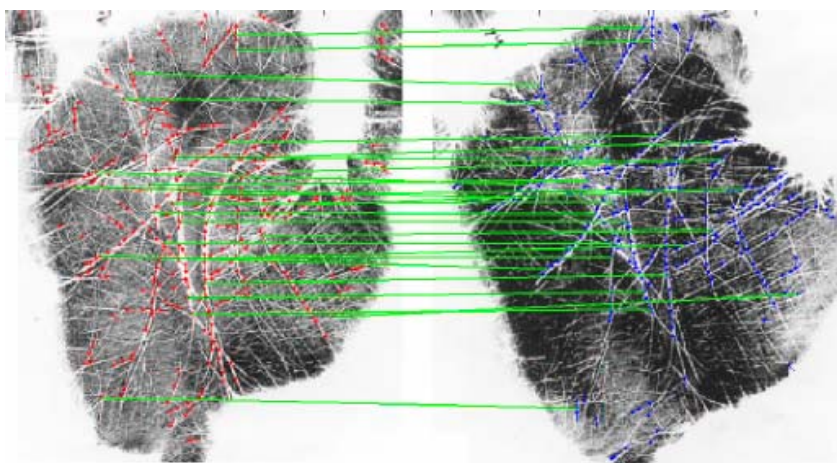
Shu i Zhang [3] su razvili sustav za raspoznavanje dlana s kasnijom obradom. U svome radu kao značajke dlana navode geometrijske značajke, glavne linije, bore, delta točke i minucije. Slika 3 prikazuje dlan korišten u ovome sustavu s označenim geometrijskim značajkama i delta točkama. Kao značajke za raspoznavanje, autori su koristili širinu i visinu dlana, te položaje nekoliko točaka duž glavnih linija dlana. Na maloj bazi na kojoj je izvršeno 48 eksperimenata poklapanja dlanova iste osobe i 844 eksperimenata poklapanja dlanova različitih osoba dobiven je rezultat 0.0% FAR i 0.0% FRR.



Slika 3: Dlan iz [3] s označenim točkama na temelju kojih se računaju geometrijske značajke (a, b, c, d) i delta točkama (1, 2, 3, 4, 5)

Zhang i Shu [6] koristili su tzv. datum točke (točke na krajevima glavnih linija dlana) kako bi poklopili koordinatne sustave različitih dlanova. Za raspoznavanje su koristili značajke linija dlana dobivene na temelju posebno izgrađenoga filtra. Sustav su testirali na maloj bazi 20 parova slika od 20 osoba. Na ovoj bazi dobiven je rezultat 0.0% FAR i 0.0% FRR.

Duta i koautori [7] su koristili točke dobivene kao rezultat binarizacije i morfoloških operacija kao značajke dlana (na slikama dlana dobivenima skeniranjem otisaka u tinti, glavne linije su mnogo svjetlije od ostatka dlana). Za svaki dlan izlučen je skup točaka i njihovih orijentacija. Razvijen je algoritam za poklapanje ovako dobivenih skupova točaka za različite dlanove. Primjer podudaranja skupova točaka za dva dlana iste osobe prikazan je na slici 4. Sustav je ispitan na skupu 30 slika od 3 osobe. Dobiven je mTER = 5.7%.



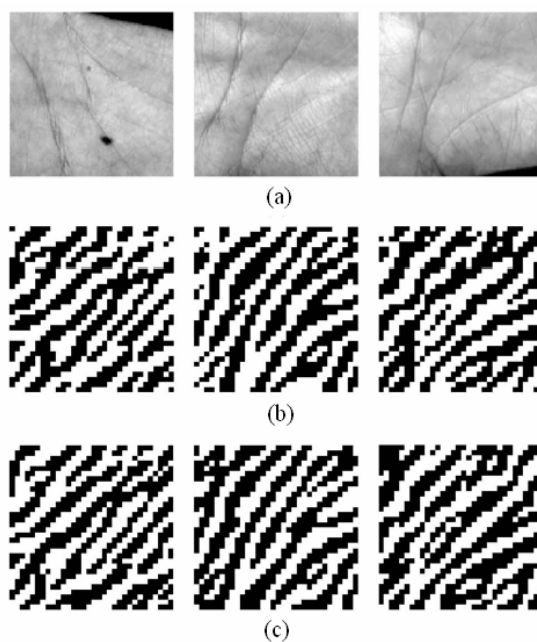
Slika 4: Podudaranje dlanova iz [7]

You i koautori [8] predlažu hijerarhijski sustav za identifikaciju osoba na temelju dlana. U prvoj fazi rada sustava, na temelju globalne energije teksture dlanova utvrđuje se mogu li oni pripadati istoj osobi. Ukoliko je odgovor pozitivan, sustav nastavlja s preciznijim podudaranjem na temelju interesnih točaka. Točnost raspoznavanja od 95% dobivena je na bazi 200 slika od 100 osoba.

Han i koautori [9] razvili su sustav za verifikaciju identiteta osoba na temelju dlana koji slike uzima pomoću stolnog skenera. Značajke dlana dobivaju se primjenom Sobelovog rubnog operatora i morfoloških operacija na području interesa čiji se položaj određuje na temelju konture ruke i položaja kažiprsta. Ispitana su dva pristupa podudaranju dlanova: usporedba s predloškom na temelju udaljenosti i neuronske mreže s propagacijom pogreške unatrag. Provedeni eksperimenti su uključivali 1000 podudaranja dlanova iste osobe i 49000

podudaranja dlanova različitih osoba. Podudaranjem na temelju predložaka ostvareni su FAR i FRR ispod 10%, a podudaranjem neuronskom mrežom ostvareni su FAR i FRR ispod 2%.

Zhang i koautori [10] razvili su sustav za identifikaciju osoba na temelju dlana koji radi u stvarnom vremenu. Slike se dobivaju pomoću CCD kamere. Uređaj za snimanje sadrži prstenasti izvor svjetla koji osigurava uniformno osvjetljenje dlana. Područje interesa dlana određuje se na temelju stabilnih točaka u udolinama između prstiju. Za izlučivanje značajki dlana koristi se Gaborov filtar. Autori koriste jedan Gaborov filtar čiji su parametri eksperimentalno određeni. Nakon filtriranja slike interesnog područja, dobiva se binarni kod (engl. palm code) na temelju binarizacije realnog i imaginarnog dijela odziva Gaborovog filtra. Na slici 5 prikazano je ovakvo izlučivanje značajki nekoliko dlanova. Binarni kodovi koji predstavljaju različite dlanove uspoređuju se na temelju Hammingove udaljenosti. Eksperimenti su provedeni na bazi koja sadrži slike dlanova 193 osobe. Baza sadrži 20 slika desnog dlana i 20 slika lijevog dlana za svaku osobu. Dobiveni su rezultati verifikacije $FRR = 2\%$ i $FAR = 0.04\%$ i rezultati identifikacije $FRR = 3\%$ i $FAR = 0.1\%$.



Slika 5: Izlučivanje značajki dlana u [10] korištenjem Gaborovog filtra. (a) Slike interesnog područja dlana, (b) binarni kod koji odgovara odzivu realnog dijela Gaborovog filtra, (c) binarni kod koji odgovara odzivu imaginarnog dijela Gaborovog filtra

U kasnijem radu [11] ovaj sustav je proširen na način da je korišteno više Gaborovih filtara s različitim parametrima. Rezultati odziva različitih filtara kombiniraju se na temelju magnitude u jedinstveni binarni kod (engl. fusion code). Za usporedbu različitih dlanova ponovno se koristi Hammingova udaljenost. Ovakav sustav ispitan je na bazi koja sadrži 20

slika lijevih i 20 slika desnih dlanova za 284 osobe. Autori navode da ovakva fuzija popravljala FRR za 15% kada je FAR = $10^{-6}\%$.

Posebnu skupinu metoda koje se koriste za raspoznavanje dlana čine metode temeljene na izgledu. Ove metode promatraju cijelu sliku interesnog područja kao vektor značajki, pri čemu su komponente vektora intenziteti pojedinih slikovnih elemenata. Ovakav izvorni vektor značajki obično se podvrgava različitim transformacijama kako bi se izvukao manji skup značajki pogodan za raspoznavanje. Na primjer, Li i koautori [12] koriste Fourierovu transformaciju za raspoznavanje osoba na temelju dlana.

Umjesto korištenja unaprijed poznatih transformacija za izlučivanje značajki, kao što je Fourierova transformacija, često je bolje koristiti transformaciju koja je naučena na uzorcima za učenje, kao što su PCA, LDA i ICA.

Tako analiza glavnih komponenti (engl. Principal Component Analysis, PCA) nalazi optimalni potprostor izvornoga prostora značajki u smislu reprezentacije uzoraka. Analiza glavnih komponenti biti će detaljno objašnjena u poglavlju 6. PCA je korištena za raspoznavanje dlanova u radu Lua i koautora [13].

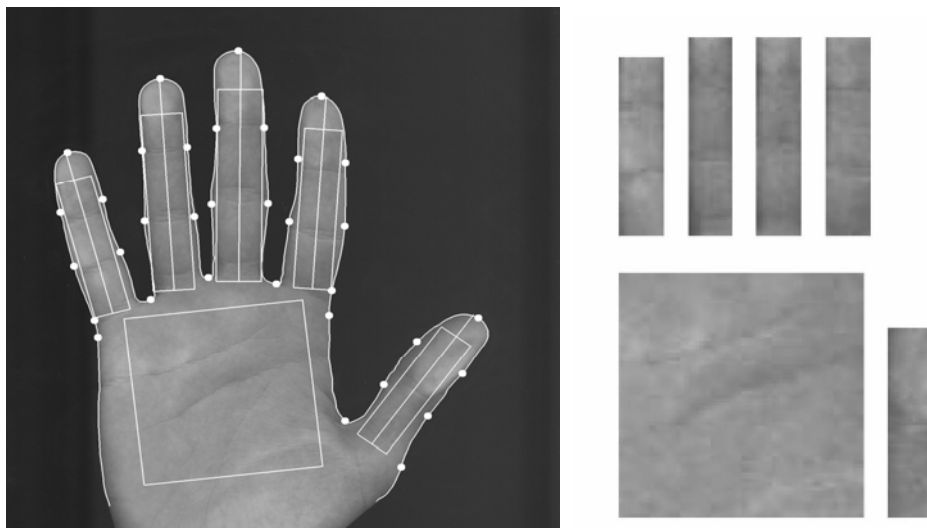
Analiza glavnih komponenti nalazi značajke pogodne za reprezentaciju uzoraka, no ne nužno i za raspoznavanje. Linearna diskriminantna analiza (engl. Linear Discriminant Analysis, LDA) nastoji naći značajke pogodne za raspoznavanje tako da nalazi potprostor izvornog prostora značajki u kome se maksimizira raspršenje uzoraka između razreda, a minimizira raspršenje uzoraka unutar pojedinih razreda. Linearna diskriminantna analiza korištena je za raspoznavanje dlanova u [14]. Linearna diskriminantna analiza biti će detaljno objašnjena u poglavlju 6.

Analiza nezavisnih komponenti (engl. Independent Component Analysis, ICA) pretpostavlja da su uzorci koje vidimo nastali kao linearna kombinacija nekih nepoznatih, statistički nezavisnih komponenti. ICA nalazi potprostor koji čine ove nezavisne komponente. ICA je korištena za raspoznavanje osoba na temelju dlana u [15].

Pored korištenja samo jednoga tipa značajki za identifikaciju ili verifikaciju identiteta osoba na temelju dlana, moguće je kombinirati više različitih značajki dlana postupcima biometrijske fuzije.

Tako je moguće kombinirati značajke dlana sa značajkama geometrije ruke kako bi se postigli bolji rezultati identifikacije ili verifikacije [16][17]. Kumar i Zhang [18] kombinirali su značajke linija dlana sa značajkama izlučenima na temelju Gaborovog filtra i značajkama dobivenima pomoću PCA. Ribarić i Fratrić [19] kombinirali su značajke izlučene pomoću

PCA ne samo iz područja dlana, nego također iz područja svakog od pet prstiju. Područja interesa ruke koja su korištena u [19] mogu se vidjeti na slici 6. Također je načinjena fuzija značajki prstiju dobivenih na temelju PCA s geometrijom prstiju [20], te fuzija koja uključuje dlan, područja prstiju kao na slici 6 te područja blizu vrha prstiju koja se koriste za dobivanje otiska prsta u slikama visoke rezolucije [21].



Slika 6: Interesna područja ruke korištena u [19]

Boja kože također je jedna od značajki dlana korištenih u biometriji. Fratrić i Ribarić [22] istražili su pogodnost različitih komponenti različitih modela boja za raspoznavanje dlanova. Eksperimentalno je utvrđeno da su normalizirana crvena i normalizirana plava komponenta u normaliziranome *rgb* modelu boja najpogodnije u smislu točnosti raspoznavanja

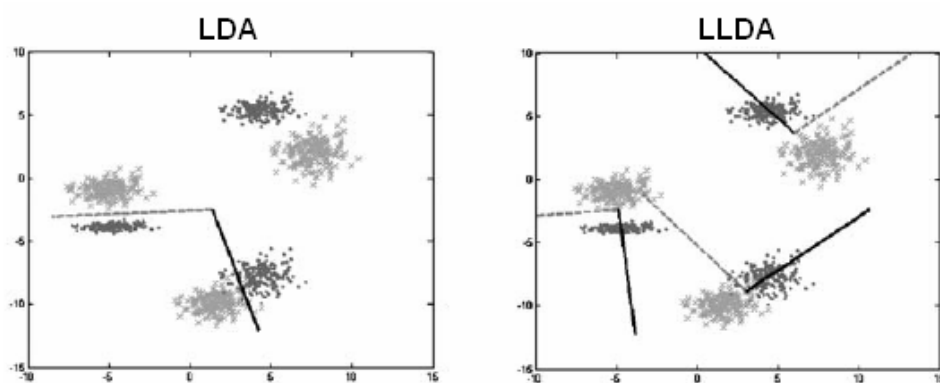
2.2. Sustavi za raspoznavanje lica

Najpopularniju skupinu metoda raspoznavanja lica čine metode temeljene na izgledu. Ove metode su u biometriji bile korištene prvo za raspoznavanje lica, a tek naknadno su prihvaćene i za raspoznavanje dlana.

Turk i Pentland [23] su 1991. koristili PCA za raspoznavanje lica. Baze potprostora dobivenoga pomoću PCA nazvali su „svojstvena lica“ (engl. *eigenfaces*).

Belhumeur i koautori [24] usporedili su PCA (*eigenfaces*) i LDA (*fisherfaces*) za raspoznavanje lica. Njihovi eksperimenti sugeriraju da će obje metode dati dobre rezultate u uvjetima kada su slike u skupu za ispitivanje slične kao slike u skupu za učenje. Međutim, LDA se nosi s promjenama u osvjetljenju i izrazu lica bolje nego PCA.

LDA je vrlo često korištena metoda za raspoznavanje lica i razvijen je niz modifikacija osnovnog postupka [25][26][27][28]. Kim i Kittler [28] predlažu lokalno linearnu diskriminantnu analizu (LLDA) za raspoznavanje lica. Ovaj postupak zamišljen je da radi u slučajevima kada je razdioba uzoraka u pojedinim razredima višemodalna. Osnovna ideja postupka je da su uzorci koji su globalno nelinearni, lokalno linearni. Postupak radi tako da prvo grupira ulazne podatke te ih predstavlja kao mješavinu Gaussovih distribucija. Zatim se maksimizira raspršenje između razreda a minimizira raspršenje unutar razreda za svaku grupu zasebno. Slika 7 pokazuje baze potprostora dobivene pomoću LDA i LLDA na jednostavnom primjeru s višemodalno distribuiranim uzorcima.



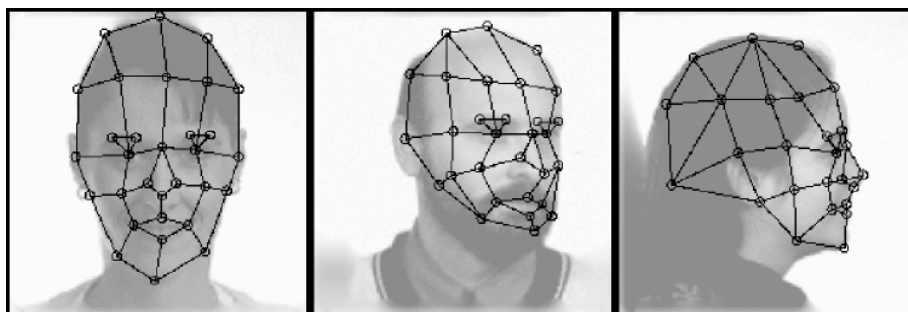
Slika 7: baze prostora dobivene pomoću LDA i LLDA [28]

Analiza nezavisnih komponenti također je bila korištena za raspoznavanje lica. Rezultati u [29] sugeriraju da ICA daje bolje rezultate nego PCA za raspoznavanje lica.

Najčešće se značajke izlučene pomoću metoda temeljenih na izgledu koriste u kombinaciji s klasifikatorima temeljenima na udaljenosti, no korišteni su i složeniji klasifikatori. Jonsson i koautori [30] koristili su klasifikator na temelju potpornih vektora nad značajkama dobivenim pomoću PCA i LDA. Meng i koautori [31] koristili su neuronske mreže kao klasifikator za značajke izlučene pomoću LDA.

Značajke izlučene na temelju Gaborovog filtra također su korištene i za raspoznavanje lica, kao i za raspoznavanje dlana. Liu [32] je izlučivao značajke lica na temelju konvolucije slike lica s više Gaborovih filtra, s različitim skalama i orijentacijama. Rezultati konvolucije s različitim Gaborovim filtrima konkatenerani su u vektor Gaborovih značajki. Nakon konvolucije, jezgrena PCA (kernel PCA) korištena je da bi se dobile najznačajnije komponente iz vektora Gaborovih značajki. U jezgrenoj PCA korištena je polinomijalna jezgra s realnim stupnjem u intervalu $<0, 1>$. Kao mjere sličnosti odnosno različitosti korišten je kut među vektorima značajki, L1 i L2 norma te Mahalanobisova udaljenost.

Također je razvijeno mnogo pristupa raspoznavanju lica koji iskorištavaju strukturu lica. Wiskott i koautori [33] predstavljaju lice kao graf. Čvorovi ovoga grafa su karakteristične točke na licu (točke smještene u zjenice oka, krajeve usta, vrh nosa itd.). Ove točke opisane su pomoću odziva različitih Gaborovih filtara sa različitim veličinama i orijentacijama. Veze čvorova grafa opisane su preko udaljenosti odgovarajućih točaka. Da bi se pronašlo lice u nepoznatoj slici, grafovi koji odgovaraju različitim licima kombinirani su u tzv. svežanj grafova (engl. bunch graph). Za različite položaje glave stvoreni su različiti grafovi, budući da sve točke nisu vidljive u svim položajima glave. Slika 8 prikazuje grafove korištene za reprezentaciju lica u različitim položajima.

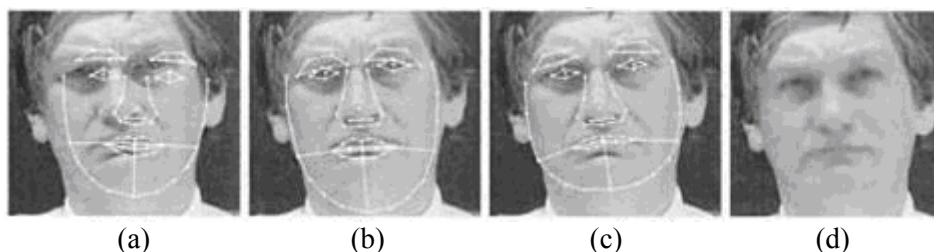


Slika 8: Grafovi korišteni za reprezentaciju lica u različitim položajima u [33]

Lica se u ovakvom sustavu uspoređuju na temelju kriterijske funkcije koja uzima u obzir udaljenost Gaborovih značajki u odgovarajućim točkama grafa, te razlike u udaljenosti točaka. Strategija podudaranja od grubljeg prema finijem korištena je da bi se našao graf koji na nepoznatoj slici najbolje odgovara grafovima pohranjenima u bazi sustava.

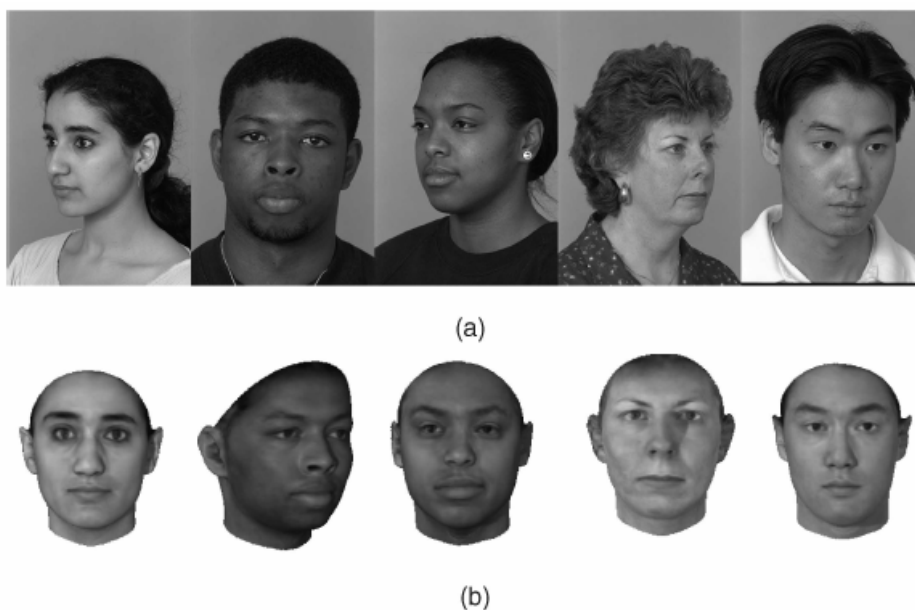
Aktivni modeli izgleda (engl. Active Appearance Models, AAM) [34] su metoda korištena za raspoznavanje lica koja radi tako da pokušava što točnije sintetizirati lice na nepoznatoj slici ugađajući parametre modela. Model lica sastoji se od oblika danog kao skup 2D točaka na licu i teksture. Da bi se odredili parametri ovakvog modela i njihova distribucija korištena je PCA na skupu za učenje. Pretpostavljeno je da oblik i tekstura nisu nezavisni, te tako isti parametar modela utječe i na oblik i na teksturu. Za pronalaženje parametara modela koji daju najmanju rekonstrukcijsku pogrešku u smislu Euklidske udaljenosti koristi se pretraživanje na slikovnoj piramidi i iterativno ugađanje parametara. Slika 9 prikazuje lice sintetizirano na temelju aktivnog modela izgleda.

Dva lica mogu se usporediti na temelju usporedbe parametara modela koji najbolje sintetiziraju ta dva lica. Nedostatak ovoga pristupa je što model mora biti inicijaliziran blizu lica na slici da bi se uspješno odredili parametri modela.



Slika 9: Lice sintetizirano pomoću aktivnog modela izgleda, (a) inicijalan položaj modela, (b) položaj modela nakon 2 iteracije, (c) položaj modela nakon 19 iteracija, (d) rekonstrukcija

Drugi pristup temeljen na modelima, opisan u [35] koristi 3D model lica. Ovaj model dobiven je na temelju 3D baze lica koja sadrži informaciju o 3D obliku i teksturi. Slično kao u [34] ideja pristupa je ugoditi parametre modela kako bi se najvjernije rekonstruiralo lice na nepoznatoj slici te također parametre modela koristiti za usporedbu različitih lica. Parametri modela također su dobiveni pomoću PCA, no u ovome sustavu to je posebno učinjeno za oblik, a posebno za teksturu. Razvijeni sustav automatski procjenjuje parametre osvjetljenja, kao što su intenzitet i smjer svjetla te sjene, te koristi ovu informaciju prilikom rekonstrukcije lica na slici. Slika 10 prikazuje nekoliko parova slika lica i odgovarajućih sintetiziranih 3D modela.



Slika 10: Nekoliko parova slika lica i odgovarajućih sintetiziranih 3D modela iz [35]

2.3. Raspoznavanje lica u video sekvencama

Većina sustava za raspoznavanje lica temelji se na statičnim slikama, koje se koriste i za izgradnju baze sustava i prilikom identifikacije ili verifikacije. Međutim, u novije vrijeme, kao ulaz u biometrijski sustav, sve se češće koriste video sekvence. Chellapa i Zhou [36] razlikuju tri slučaja korištenja video sekvenci u biometrijskim sustavima:

- „still-to-video“ raspoznavanje, gdje se statične slike koriste u bazi sustava, a video sekvence za raspoznavanje
- „video-to-still“ raspoznavanje, gdje se video sekvence koriste u bazi sustava, a statične slike za raspoznavanje
- „video-to-video“ raspoznavanje, gdje se video sekvence koriste i za izgradnju baze sustava i za raspoznavanje

Prednost korištenja video sekvenci je dodatna informacija, u smislu više okvira i u smislu vremenske informacije. Također je demonstrirano da ljudi bolje raspoznaju lica iz video sekvenci nego iz statičnih slika. Međutim, tipično se u scenarijima kada se video sekvence koriste u biometrijskim sustavima nailazi na niz problema

- Kvaliteta videa često je lošija od statičnih slika koje se koriste u biometriji, bilo kao posljedica ograničenja uređaja koji se koristi za snimanje videa, bilo kao posljedica nekontroliranih uvjeta u kojima se video snima. U video sekvencama su također moguća djelomična zaklanjanja lica u pojedinim okvirima.
- Područja interesa u video sekvencama mogu biti vrlo mala, ponovo bilo kao posljedica ograničenja uređaja koji se koristi za snimanje, bilo kao posljedica nekontroliranih uvjeta snimanja.
- Varijacije unutar razreda u video sekvencama mogu biti vrlo velike. Tako se u video sekvencama mogu očekivati veće promjene u položaju i izrazu lica nego u većini scenarija u kojima se za raspoznavanje koriste statične slike.

Općenito, postoje dva pristupa raspoznavanju lica u videu. Češći pristup [37] je video sekvencu promatrati kao niz statičnih slika i svaki okvir video sekvence obrađivati nezavisno. Nakon toga, potrebno je kombinirati rezultate dobivene za pojedine okvire video sekvence, bilo na temelju glasanja (svaki okvir glasa za neku odluku), bilo na temelju kombiniranja mjere podudaranja za pojedine okvire u jedinstvenu mjeru podudaranja za cijelu video sekvencu.

Umjesto korištenja svih okvira video sekvence, može se koristiti i manji podskup okvira koji zadovoljavaju neki unaprijed definirani kriterij kvalitete [37]. Mogu se koristiti različite mjere kvalitete koje se mogu podijeliti na mjere kvalitete slike i mjere kvalitete značajki.

Na primjer, u natjecanju sustava za raspoznavanje lica u videu [38] za svaki okvir video sekvence dane su sljedeće mjere kvalitete: pouzdanost detektora, svjetlina, kontrast, fokus, broj bita po slikovnom elementu, udaljenost između očiju, vrsta osvjetljenja, uniformnost pozadine, svjetlina pozadine, refleksije, prisutnost naočala, rotacija u ravnini kamere, rotacija izvan ravnine kamere i frontalnost.

Drugi pristup raspoznavanju lica u video sekvencama iskorištava vremenski kontinuitet video sekvence. Chellapa i Zhou [36] koriste vjerojatnosni pristup koji poslove praćenja, detekcije i raspoznavanja obavlja zajedno. Stanje u vjerojatnosnom modelu koji koristi ovaj sustav obuhvaća informaciju o pokretu (vektor brzine) i informaciju o identitetu (varijablu identiteta). U svakom vremenskom odsječku procjenjuje se združena posteriorna distribucija vektora brzine i varijable identiteta, te se prosljeđuje u sljedeći vremenski odsječak.

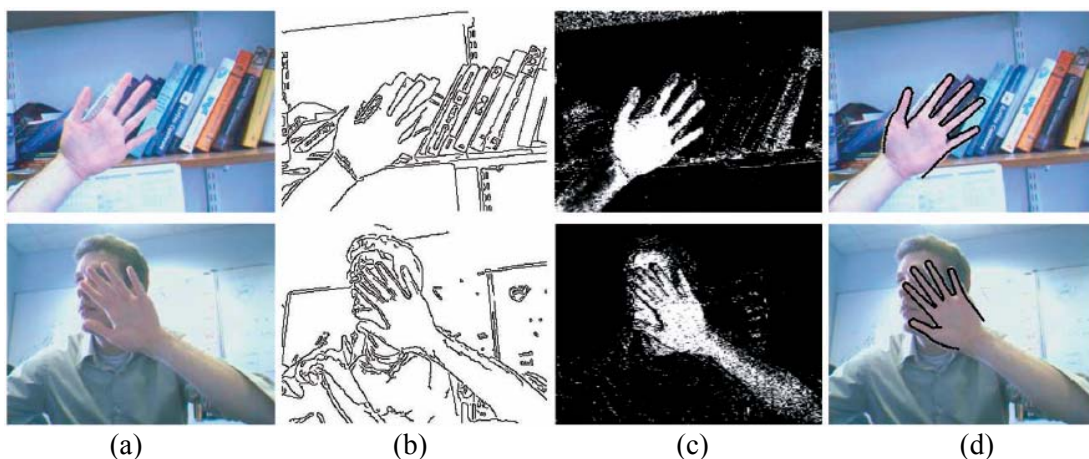
2.4. Beskontaktna biometrija ruke

Dok su biometrijski sustavi temeljeni na licu inherentno beskontaktni, beskontaktni biometrijski sustavi temeljeni na ruci počeli su se pojavljivati tek nedavno. Većina do sada razvijenih sustava temeljenih na ruci, kao što su sustavi opisani u poglavlju 2.1. zahtijevaju da korisnik stavi ruku na površinu senzora. Nadalje, u mnogim sustavima položaj ruke je kontroliran pomoću klinova. Ovo može predstavljati problem u scenarijima s velikim brojem korisnika jer neki ljudi odbijaju dirati biometrijske senzore iz higijenskih i drugih razloga. Beskontaktni biometrijski sustavi također su jednostavniji za korištenje.

S druge strane, u većini sustava slika ruke uzima se u strogo kontroliranim uvjetima (u smislu osvjetljenja i položaja ruke), tako da je segmentacija slike jednostavna. Slike snimljene beskontaktno imati će veće varijacije u položaju ruke i osvjetljenju što čini postupke segmentacije i raspoznavanja ruke u ovakvim sustavima složenijima.

Većina radova vezanih za detekciju ruke u slikama i video sekvencama vezana je za područje interakcije čovjeka i računala. Pregled tehnika korištenih u ovom području dan je u [39]. Mnoge tehnike oslanjaju se na informaciju o pokretima i heurističke pristupe za detekciju ruke. Međutim, u nekontroliranim uvjetima u kojima bi bili mogući pokreti u pozadini detekcija ruke temeljena na pokretu ne bi dala dovoljno dobre rezultate za primjenu u biometriji.

Na detekciji ruke na temelju izgleda i oblika nije rađeno toliko mnogo. Kölsch i Turk [40] za detekciju ruke koristili su detektor objekata kojega su predložili Viola i Jones [41] za detekciju lica. Ovaj detektor objekata je vrlo brz i može se primijeniti na video sekvencama u realnom vremenu. Nedostatak ovog pristupa za detekciju ruke je što se dobiva približan položaj ruke na slici, no ne i kontura ruke. Nadalje, moguće su lažne detekcije.



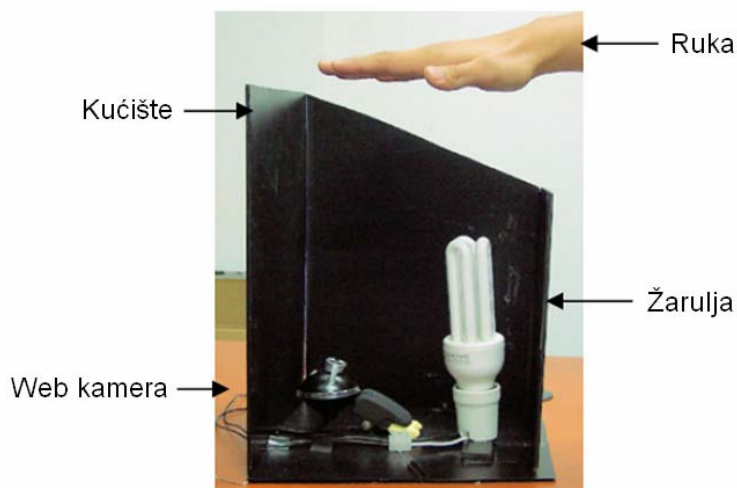
Slika 11: Primjer detekcije ruke iz [42]: (a) izvorna slika, (b) rubovi, (c) boja kože, (d) rezultat detekcije

Stegner i koautori [42] koristili su hijerarhijski pristup temeljen na modelima za nalaženje i praćenje ruke. Autori su koristili skup modela dobivenih rotirajući osnovni model ruke. Ovako generirani modeli su organizirani u stablo na temelju kuta rotacije. Organizacijom modela u stablo izbjegava se usporedba neke lokacije na slici sa svim modelima. Za usporedbu lokacije na slici i modela koriste se dva kriterija: udaljenost rubova i vjerodostojnost boje kože. Ovakav pristup nalazi približnu konturu i rotaciju ruke, ali je previše spor za primjenu u realnom vremenu. Primjer ovakve detekcije ruke može se vidjeti na slici 11.

Premda ne postoje objavljeni radovi vezani za raspoznavanje osoba na temelju ruke u potpuno nekontroliranim uvjetima, postoje dva beskontaktna sustava temeljena na ruci u polu-kontroliranim uvjetima.

Ong i koautori [43] opisali su beskontaktni biometrijski sustav temeljen na dlanu. Sustav zahtijeva da osoba koja pristupa sustavu stavi ruku iznad uređaja prikazanoga na slici 12. Kako bi se pronašlo područje interesa dlana, korištena je segmentacija na temelju boje kože i usporedbe s pragom. Značajke dlana dobivene su pomoću lokalnih binarnih uzoraka (engl. Local Binary Patterns, LBP) primijenjenih na sliku rubova dobivenu Sobelovim rubnim operatorom. Nedostaci sustava su korištenje boje kože za segmentaciju, što neće raditi

ukoliko objekti u pozadini sadrže boju kože te ne iskorištavanje video informacije. Autori su prijavili $EER = 1.52\%$ na bazi 320 osoba s 20 slika po osobi.



Slika 12: Uređaj korišten u [43] za beskontaktnu akviziciju slika ruke

Doublet i koautori [44] također su opisali beskontaktni biometrijski sustav temeljen na ruci. Autori za segmentaciju koriste aktivni model oblika, inicijaliziran na temelju segmentacije na temelju boje kože. Kao značajke ruke koristi se kombinacija oblika ruke i značajki dobivenih primjenom Gaborovog filtra na interesno područje dlana. Cijeli postupak raspoznavanja traje jednu sekundu na procesoru s taktom od 1.6 GHz, što je nedovoljno brzo za primjenu na video sekvencama. Na bazi 160 slika 16 osoba postignut je $EER = 1.85\%$. Baza je prikupljena pomoću web kamere, no autori nisu detaljno opisali uvjete u kojima je baza snimljena (u smislu osvjetljenja i složenosti pozadine).

2.5. Fuzija dlana i lica

Objavljeni radovi koji opisuju fuziju dlana i lica temelje se na statičnim slikama, te su u njima dlan i lice snimljeni različitim senzorima.

Kumar i Zhang [45] opisali su sustav koji kombinira pristup raspoznavanju dlanova opisan u [17] i pristup raspoznavanju lica temeljen na svojstvenim licima [23]. Za fuziju se koristi neuronska mreža koja na temelju individualnih mjera podudaranja donosi odluku o prihvaćanju ili odbijanju korisnika. Sustav je ispitan na bazi 40 korisnika. Dobiveni su rezultati verifikacije $mTER = 13.04\%$ za lice, $mTER = 6.90\%$ za dlan i $mTER = 1.53\%$ nakon fuzije.

Ribarić, Fratrić i Kiš [46][47] opisali su sustav koji kombinira svojstvena lica [23] s pristupom raspoznavanja dlana na temelju linija dlana. Sustav je ispitan na bazi 1488 parova

slika 241 osobe. Fuzijom na razini mjere podudaranja ostvareno je poboljšanje EER s 3.89% za dlan na 2.26% nakon fuzije.

Oba sustava koriste tzv. „himeričke“ baze, što znači da su baze dobivene uparivanjem slika dlana i lica iz različitih baza kako bi se stvorile virtualne osobe. Ovakvo stvaranje baze imati će za posljedicu neovisnost značajki dlana i lica, što ne mora vrijediti za stvarne osobe (npr. boja kože utjecati će i na dlan i lice), niti za baze u kojima su lice i dlan snimljeni istovremeno (npr. loši uvjeti osvjetljenja prilikom snimanja negativno će utjecati i na značajke dlana i na značajke lica).

3. Beskontaktna verifikacija temeljena na značajkama dlana i lica

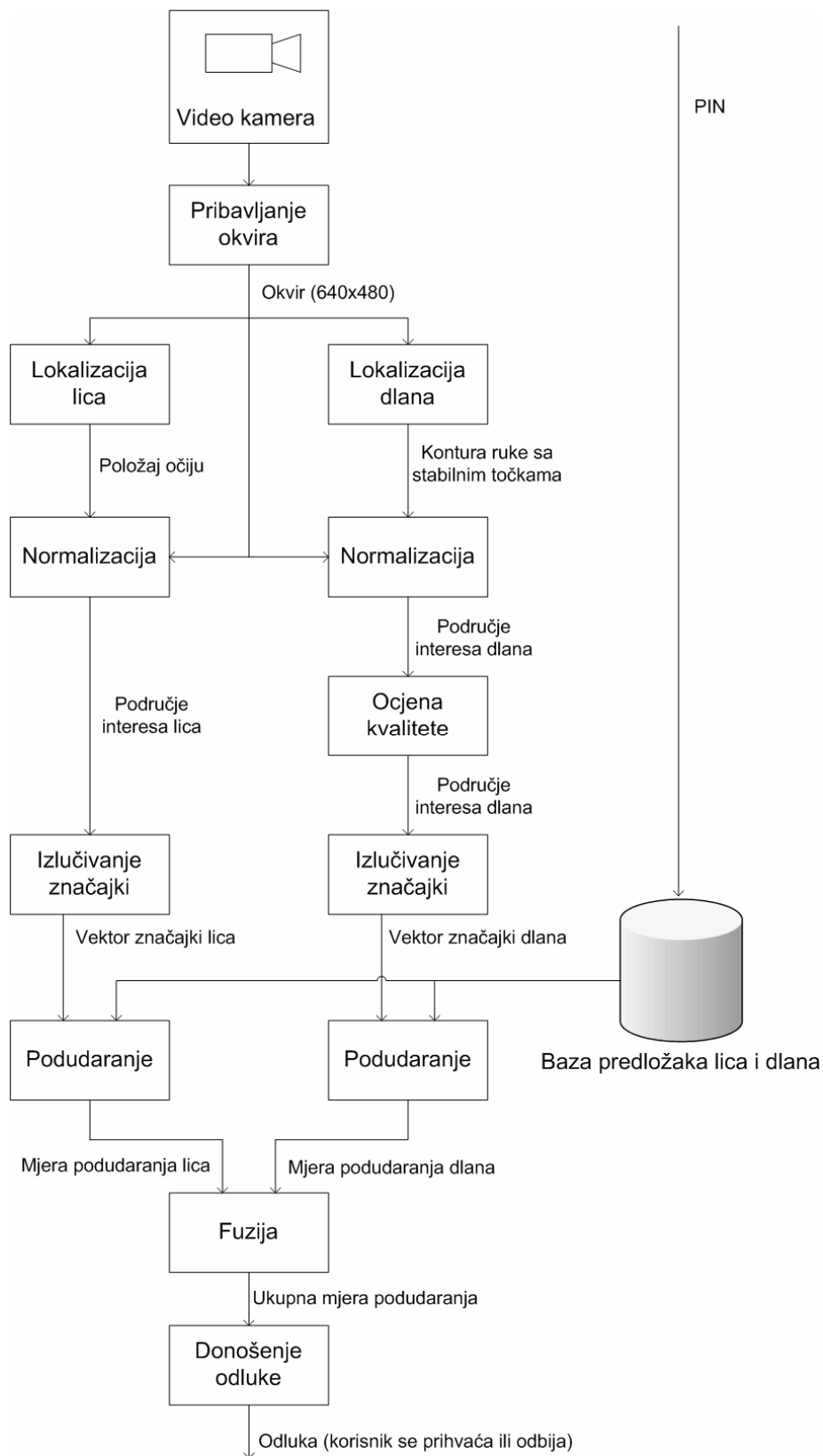
Za razliku od biometrijskih sustava temeljenih na licu, za koje je uobičajeno da se značajke dobivaju iz snimaka načinjenih video kamerom, dlan se, kao biometrijska karakteristika, do sada uglavnom koristio u sustavima koji zahtijevaju kontakt korisnika i senzora. Ovo umanjuje prihvatljivost dlana kao biometrijske karakteristike od strane korisnika, naročito u scenarijima s velikim brojem korisnika, zbog odbojnosti koju stvara ideja doticanja biometrijskog senzora, kojega je prethodno dotakao veliki broj ljudi. Tek u novije vrijeme pojavili su se sustavi u kojima se dlan snima kamerom bez doticanja dlana i senzora, ali u kontroliranim ili polu-kontroliranim uvjetima.

U nastavku rada razmatra se biometrijska verifikacija na temelju dlana i lica snimljenih video kamerom u nekontroliranim uvjetima u smislu osvjetljenja i pozadine. Dok je segmentacija slika ruke snimljenih u kontroliranim uvjetima trivijalna, korištenje snimaka dlana snimljenih u nekontroliranim uvjetima stvara potrebu za složenijim postupcima lokalizacije interesnog područja dlana, koji će biti otporni na veliki broj objekata u pozadini, uključujući i objekte boje kože, pokretne objekte u pozadini, različite uvjete osvjetljenja itd. Također, uz korištenje video sekvenci kao izvora značajki, mogu se očekivati velike varijacije u izgledu dlana i lica iste osobe, što može negativno utjecati na točnost verifikacije. Ove varijacije unutar razreda mogu nastati kao posljedica promjena položaja ruke i lica, promjena izraza lica, promjena osvjetljenja itd.

Cilj istraživanja je razvoj i evaluacija robusnih metoda lokalizacije, normalizacije, ocjene kvalitete i izlučivanja značajki, pomoću kojih bi se pokazalo može li se, i s kolikom pogreškom, ovako snimljen dlan koristiti za biometrijsku verifikaciju. Također se želi pokazati može li se na temelju fuzije dlana i lica snimljenih u istoj video sekvenci ostvariti manja pogreška verifikacije u odnosu na sustav koji bi koristio samo značajke dlana ili samo značajke lica.

Slika 13 prikazuje predloženu organizaciju sustava za biometrijsku verifikaciju temeljenu na značajkama dlana i lica dobivenim iz video sekvenci. Video sekvence dobivaju se video kamerom koja može biti postavljena u okruženju u kojemu će uvjeti osvjetljenja varirati, te se može očekivati mnoštvo drugih, moguće pokretnih, objekata u pozadini. Sa video kamere pribavljaju se okviri u rezoluciji 640x480 slikovnih elemenata, koja se pokazala dovoljno velika da dlan i lice osobe ispred kamere budu dovoljno veliki za izlučivanje značajki, a dovoljno mala da je moguće očekivati da će sustav raditi u stvarnom vremenu.

3. BESKONTAKTNA VERIFIKACIJA TEMELJENA NA ZNAČAJKAMA DLANAI LICA



Slika 13: Blok dijagram sustava za verifikaciju osoba na temelju značajki dlana i lica dobivenih iz video sekvenci

3. BESKONTAKTNA VERIFIKACIJA TEMELJENA NA ZNAČAJKAMA DLANAI LICA

Okvir video sekvence ulaz je u module za lokalizaciju lica i lokalizaciju dlana. Za lokalizaciju lica, u realizaciji sustava, preuzet je pristup Viole i Jonesa [41] koji se pokazao dovoljno robustan za primjenu u složenim uvjetima snimanja i dovoljno brz za primjenu u stvarnom vremenu. Nakon detekcije samoga lica, u području lica detektiraju se oči, čiji položaj je, zajedno s okvirom video sekvence, ulaz u modul za normalizaciju.

Za lokalizaciju dlana razvijena je nova metoda koja kombinira detekciju kandidata ruke pristupom Viole i Jonesa s verifikacijom kandidata i lokalizacijom konture ruke pristupom temeljenim na modelima. Detekcija kandidata pristupom Viole i Jonesa jamči mogućnost izvođenja u stvarnom vremenu, dok pristup temeljen na modelima određuje približnu konturu ruke i može eliminirati lažne detekcije. Izlaz modula za lokalizaciju dlana je kontura ruke sa stabilnim točkama.

Nakon lokalizacije potrebno je primijeniti postupke normalizacije kako bi dobivena područja interesa dlana i lica bila što je moguće manje osjetljiva na razlike u položaju, orijentaciji i uvjetima osvjetljenja. U modulu za normalizaciju obavlja se geometrijska normalizacija, čime se interesna područja svode na kvadratna područja predefiniranih dimenzija, te svjetlosna normalizacija. Postupak normalizacije rezultira normaliziranim slikama područja interesa veličine 64x64 slikovnih elemenata.

U slučaju dlana, područja interesa prolaze postupak ocjene kvalitete. Ocjena kvalitete potrebna je kako bi se odredilo koja područja interesa (iz kojih okvira video sekvence) su pogodna za izlučivanje značajki. Kvaliteta interesnog područja dlana mijenjati će se iz okvira u okvir zbog nepreciznosti u postupku lokalizacije, različitog položaja ruke prema kameri te različitog položaja dlana u odnosu na izvore osvjetljenja. Odbacivanje područja interesa koja ne zadovoljavaju kriterij kvalitete omogućuje poboljšanje rezultata verifikacije te ubrzanje rada ostatka sustava, budući da se u postupak izlučivanja značajki i podudaranja ne prosljeđuju sva interesna područja.

Normalizirana interesna područja ulaz su u module za izlučivanje značajki. U nastavku rada opisane su i evaluirane različite metode izlučivanja značajki lica i dlana. Kao predstavnici metoda izlučivanja globalnih značajki evaluirane su analiza glavnih komponenti i linearna diskriminantna analiza. Kao predstavnik metoda izlučivanja lokalnih značajki, evaluirano je izlučivanje binarnih značajki na temelju Gaborovog filtra. Također je opisana i evaluirana nova metoda izlučivanja značajki, nazvana lokalna binarna LDA (LBLDA), koja je zamišljena tako da kombinira dobre strane linearne diskriminantne analize i metoda temeljenih na lokalnim značajkama.

Vektori značajki dobiveni u postupku izlučivanja značajki zatim se uspoređuju s odgovarajućim vektorima značajki koji su pohranjeni u bazi biometrijskog sustava. Budući

3. BESKONTAKTNA VERIFIKACIJA TEMELJENA NA ZNAČAJKAMA DLANAI LICA

da se radi o sustavu za verifikaciju, korisnik sustava, pored biometrijskih značajki, predočava i naznaku svoga identiteta te se izlučene značajke uspoređuju samo s onima koje su u bazi sustava pohranjene upravo pod tim identitetom. Vektori značajki pohranjeni u bazi biometrijskog sustava također su dobiveni na temelju video sekvenci identičnim postupkom. Rezultati usporedbe vektora značajki su mjera podudaranja lica i mjera podudaranja dlana. Ove mjere podudaranja kombiniraju se postupcima biometrijske fuzije. Na slici 13 prikazan je sustav s fuzijom na razini mjere podudaranja, međutim načinjeni su eksperimenti i s fuzijom na razini odluke. U fuziji na razini mjere podudaranja individualne mjere podudaranja kombiniraju se u jedinstvenu mjeru podudaranja.

Na temelju ove jedinstvene mjere podudaranja (najčešće na temelju usporedbe s pragom) donosi se konačna odluka o prihvaćanju, odnosno odbijanju korisnika.

4. Lokalizacija lica i ruke

Lokalizacija lica i ruke prvi je korak, nakon akvizicije video sekvence, u sustavu za biometrijsku verifikaciju na temelju ruke i lica. Zadaća ovoga dijela sustava je odrediti postoje li ruka i lice na slici te, ukoliko postoje, gdje se nalaze interesna područja lica i dlana iz kojih će se izlučivati značajke.

Lokalizacija ruke i lica mora biti dovoljno precizna za biometrijsku verifikaciju i mora moći raditi u realnom vremenu da bi mogla biti primijenjena na video sekvencama. Pred toga, lokalizacija ruke mora kao rezultat dati konturu ruke kako bi se na temelju nje mogao odrediti položaj interesnog područja dlana.

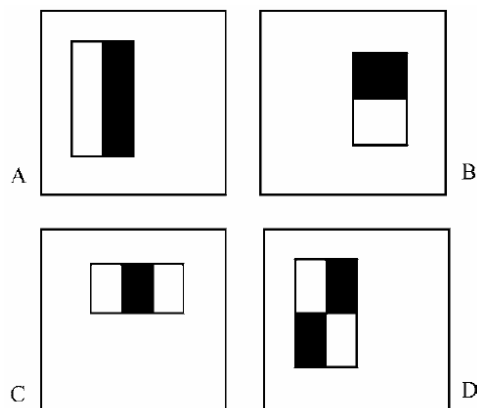
4.1. Lokalizacija lica

Za lokalizaciju lica u ovome radu preuzet je pristup koji su opisali Viola i Jones [41]. Značajka ovoga pristupa, pored dobre točnosti detekcije, je vrlo velika brzina, što ga čini pogodnim za primjenu na video sekvencama.

Ovaj detektor lica radi na slikama u nijansama sive boje i koristi pristup klizajućeg prozora: Kvadratni prozor određene veličine (obično 20x20 ili 24x24 slikovna elementa) “šeta” se po slici i za svaki položaj prozora klasifikator određuje predstavlja li dio slike unutar prozora lice ili ne. Nakon toga se prozor povećava za određeni faktor skaliranja, te se postupak ponavlja. Postupak završava kada veličina prozora postane veličina slike.

Sam klasifikator u Viola-Jones detektoru je zapravo kaskada klasifikatora, a pojedini elementi kaskade dobiveni su AdaBoost algoritmom nad velikim skupom Haarovih značajki. Izgradnja ovog klasifikatora će biti detaljno objašnjena u nastavku.

Osnova Viola-Jones detektora su Haarove značajke. Ove značajke računaju se na temelju razlike suma vrijednosti slikovnih elemenata u pravokutnim područjima. Tipovi značajki koji se koriste u Viola-Jones detektoru lica prikazani su na slici 14. Vrijednost značajke se dobiva kao suma vrijednosti slikovnih elemenata unutar bijelih pravokutnika kojoj se oduzme suma vrijednosti slikovnih elemenata unutar crnih pravokutnika.

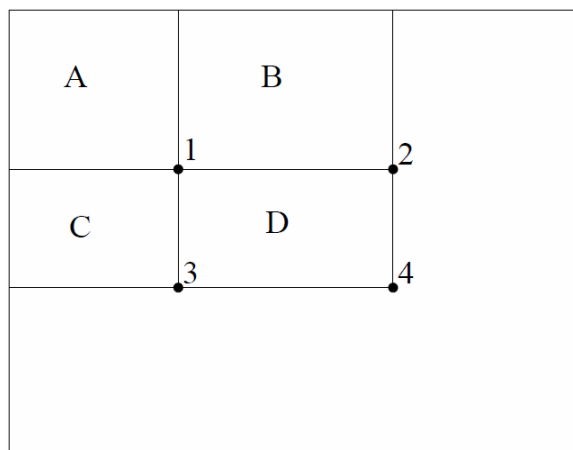


Slika 14: Tipovi Haarovih značajki korišteni u Viola-Jones detektoru lica

Karakteristika ovakvih značajki i jedan od razloga za veliku brzinu Viola-Jones detektora, je da se vrijednosti značajki mogu izračunati u $O(1)$ pomoću integralnih slika. Integralna slika je slika u kojoj se vrijednost nekog slikovnog elementa dobiva kao suma svih slikovnih elemenata u izvornoj slici koji su gore lijevo od njega, uključujući i sam taj slikovni element:

$$H(x, y) = \sum_{x' \leq x, y' \leq y} I(x', y') \quad (3)$$

Integralna slika može se izračunati u jednom prolazu preko slike. Pomoću integralne slike suma vrijednosti slikovnih elemenata u bilo kojem pravokutnom području može se dobiti na temelju vrijednosti integralne slike u vrhovima pravokutnika. Ovo ilustrira slika 15.



Slika 15: Računanje sume vrijednosti slikovnih elemenata u pravokutnom području na temelju integralne slike

Na slici 15 vrijednost integralne slike u točki 1 je suma vrijednosti slikovnih elemenata u pravokutniku A. Slično, vrijednost integralne slike u točki 2 je suma vrijednosti slikovnih

elemenata u pravokutnicima A i B. Vrijednost integralne slike u točki 3 je suma vrijednosti slikovnih elemenata u pravokutnicima A i C. Vrijednost integralne slike u točki 4 je suma vrijednosti slikovnih elemenata u pravokutnicima A, B, C i D. Iz ovoga slijedi da je suma vrijednosti slikovnih elemenata u pravokutniku D = $I(4) + I(1) - [I(2) + I(3)]$. Na temelju ove jednakosti lako se može izračunati vrijednost svih značajki sa slike 14.

Klasifikator koji koristi samo jednu Haarovu značajku smatra se slabim klasifikatorom. Slabi klasifikator h_j sastoji se od značajke f_j , praga τ_j i parametra koji kontrolira smjer nejednakosti prilikom usporedbe vrijednosti značajke s pragom, ξ_j . Slabi klasifikator h_j na prozoru x je

$$h_j(x) = \begin{cases} 1 & \text{ako } \xi_j f_j(x) < \xi_j \tau_j \\ 0 & \text{u protivnom} \end{cases} \quad (4)$$

Viola-Jones detektor koristi varijantu AdaBoost algoritma kako bi se na temelju kombinacije slabih klasifikatora izgradio jači klasifikator koji će se koristiti u jednom stupnju kaskade. Zadaća AdaBoost algoritma u Viola-Jones detektoru je naći mali podskup svih mogućih značajki koji će se koristiti i ujedno naći optimalne parametre klasifikatora. AdaBoost algoritam korišten u Viola-Jones detektoru prikazan je u nastavku.

4. LOKALIZACIJA LICA I RUKE

Zadan je skup slika za učenje i odgovarajućih klasifikacija $(x_1, o_1), \dots, (x_N, o_N)$ gdje za svaku sliku $o_i = 0, 1$ određuje radi li se o negativnom ili pozitivnom primjeru.

Inicijaliziraju se težine primjera $w_{1,i} = 1/2N^-, 1/2N^+$ za $o_i = 0, 1$ respektivno, gdje je N^- broj negativnih, a N^+ broj pozitivnih primjera.

za $k = 1, \dots, K$, gdje je K broj značajki konačnog klasifikatora

Normaliziraju se težine primjera:

$$w_{k,i} = \frac{w_{k,i}}{\sum_{j=1}^N w_{k,j}}$$

Za svaku značajku j nauči se klasifikator h_j koji koristi samo tu značajku. Ocjeni se greška klasifikatora uzimajući u obzir težine primjera:

$$e_j = \sum_i w_{k,i} |h_j(x_i) - o_i|$$

Odabere se klasifikator h_k s najmanjom pogreškom ϵ_k .

Promijene se težine

$$w_{k+1,i} = w_{k,i} \beta_k^{1-e_i}$$

gdje je $e_i = 1$ ako je primjer x_i klasificiran ispravno, 0 inače. $\beta_k = e_k / (1 - e_k)$

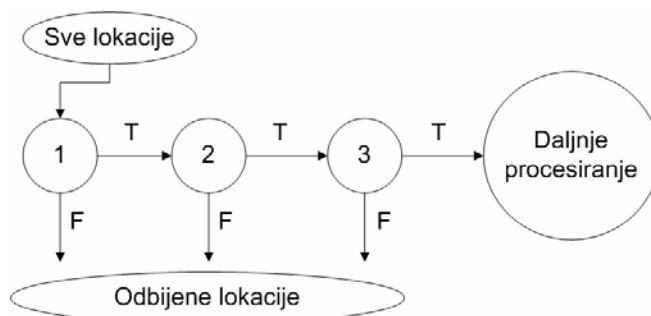
Konačan jaki klasifikator je

$$h(x) = \begin{cases} 1 & \sum_{k=1}^K \alpha_k h_k(x) \geq \frac{1}{2} \sum_{k=1}^K \alpha_k \\ 0 & \text{u protivnom} \end{cases}$$

gdje je $\alpha_k = \log(1/\beta_k)$

Algoritam 1: AdaBoost algoritam korišten u Viola-Jones detektoru [41]

Umjesto korištenja samo jednog jakog klasifikatora, Viola-Jones detektor koristi više klasifikatora povezanih u kaskadu. Zadaća svakog elementa kaskade je zadržati skoro sva (tipično više od 99,5%) prava lica, a odbaciti veliki broj (tipično više od 50%) lokacija na kojima se lice ne nalazi. Ova organizacija klasifikatora u kaskadu prikazana je na slici 16.



Slika 16: Organizacija klasifikatora u kaskadu u Viola-Jones detektoru

Ovakva kaskada klasifikatora omogućuje da se većina lokacija koje ne sadrže lice odbije u prvim elementima kaskade, a one lokacije za koje klasifikatori nisu sigurni sadrže li lice ili ne prosljeđuju se daljnjim stupnjevima kaskade na naknadne provjere. Ovakva organizacija klasifikatora, zajedno sa brzim izračunavanjem značajki, omogućuje izvođenje detekcije u stvarnom vremenu.

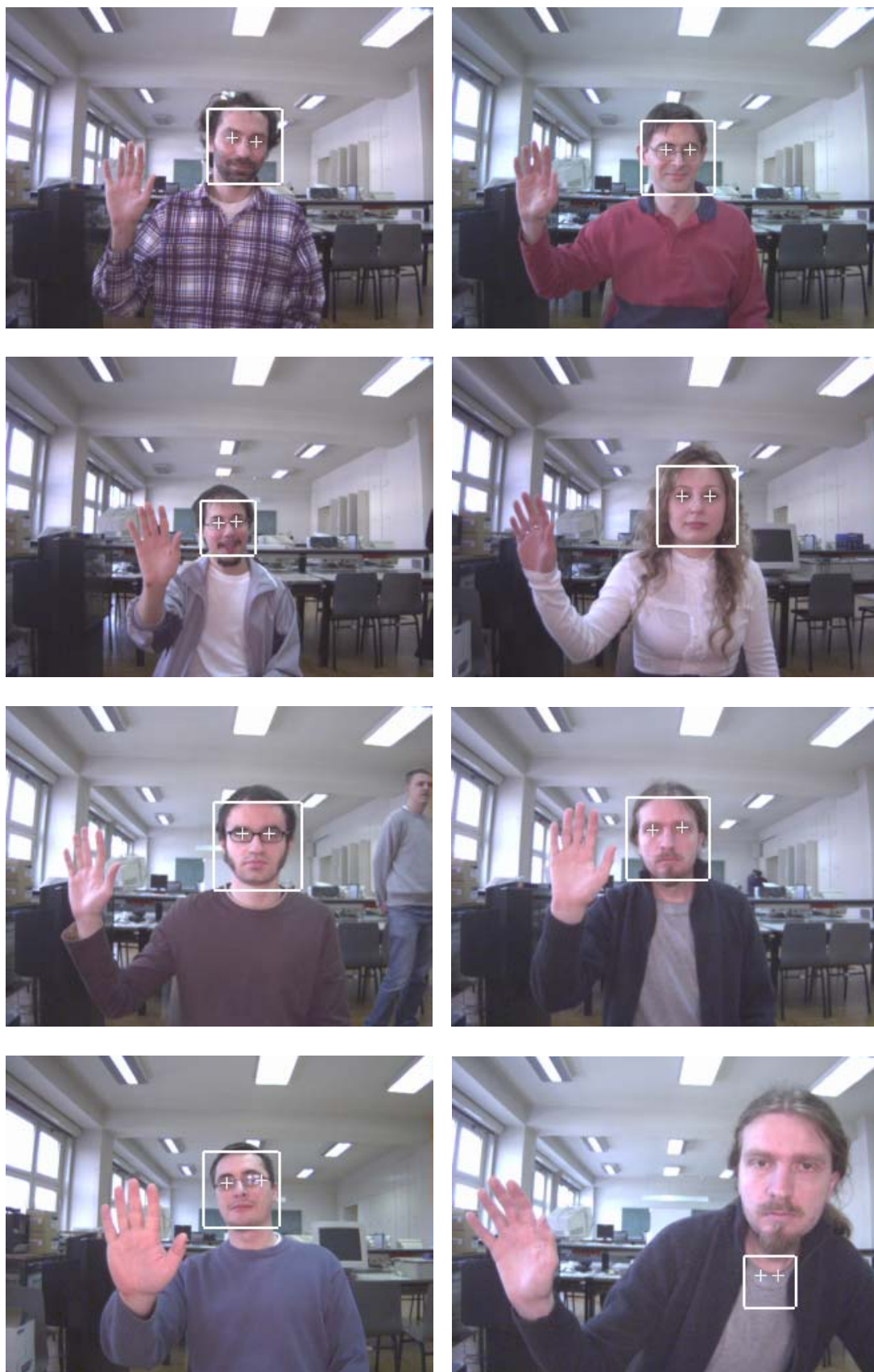
Viola-Jones detektor lica tipično neće dati samo jednu detekciju na mjestu gdje se nalazi lice, nego će dati više detekcija za susjedne, međusobno prekrivajuće lokacije. Zbog toga se bliske detekcije grupiraju u jednu. Postavljanjem praga na broj bliskih detekcija također se može smanjiti broj lažnih detekcija. U eksperimentima je minimalan broj susjednih detekcija koji mora postojati da bi se objekt smatrao licem postavljen na tri.

Viola-Jones detektor lica može naći više lica na slici ili okviru video sekvence, ukoliko ona postoje. Međutim, za primjenu u biometrijskoj verifikaciji ova funkcionalnost nije potrebna. Zbog toga je detektor modificiran na način da, ukoliko je na slici detektirano više od jednoga lica, vraća samo najveće (pretpostavlja se da će najveće lice biti ono koje je najbliže kameri). Pored Viola-Jones detektora za cijelo lice, u radu je također korišten i Viola-Jones detektor za oči. Budući da su koordinate očiju potrebne za normalizaciju slika lica, nakon što se lokalizira lice, na području slike u kojemu je nađeno lice primjenjuje se Viola-Jones detektor očiju.

U radu je korištena implementacija Viola-Jones detektora iz OpenCV biblioteke [48]. Također su korištene prethodno naučene kaskade klasifikatora za detekciju lica i očiju koje dolaze s OpenCV bibliotekom.

Slika 17 prikazuje nekoliko primjera lokalizacije lica i očiju.

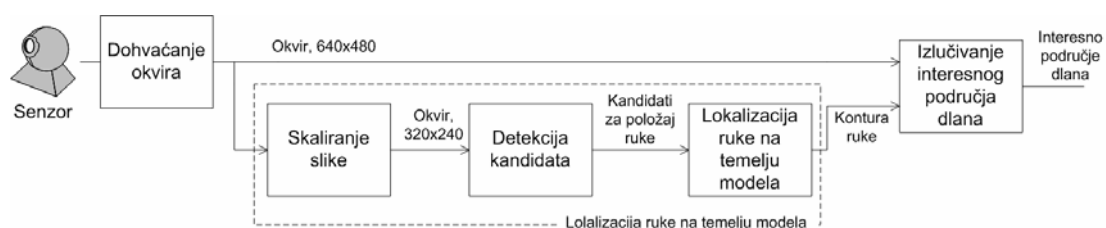
4. LOKALIZACIJA LICA I RUKE



Slika 17: Nekoliko primjera (uspješnih i neuspješnih) lokalizacije lica i očiju

4.2. Lokalizacija ruke

Za lokalizaciju ruke razvijen je podsustav [49] koji kombinira Viola-Jones detektor objekata [41] i detekciju na temelju modela [42]. Kombinacija Viola-Jones detektora i detektora na temelju modela omogućuje dovoljno brzu detekciju za primjenu u stvarnom vremenu, a također omogućuje lokalizaciju konture ruke na slici. Slika 18 prikazuje blok dijagram razvijenog podsustava za lokalizaciju ruke i izlučivanje interesnog područja dlana.



Slika 18: Blok dijagram podsustava za lokalizaciju ruke i izlučivanje interesnog područja dlana

Kao uređaj za snimanje ruke koristi se videokamera rezolucije 640x480. Ova rezolucija odabrana je zbog dobrog odnosa između kvalitete slike i vremena koje je potrebno da se obradi svaki okvir video sekvence. Svaki od okvira obrađuje se kako slijedi.

Prvo, okvir video sekvence skalira se na rezoluciju od 320x240 slikovnih elemenata kako bi se reduciralo vrijeme detekcije. Reducirani okvir (predstavljen u obliku 256 nijansi sive boje) se koristi za detekciju kandidata ruke. Za detekciju kandidata ruke koristi se Viola-Jones detektor zbog dobre točnosti i brzine obrade. Kada su pronađeni kandidati za položaj ruke, svaki kandidat se verificira poklapanjem s velikim, prethodno definiranim, skupom modela ruke. Za poklapanje kandidata s modelom korištena su dva kriterija: udaljenost rubova i vjerodostojnost boje kože. Ove dvije mjere kombiniraju se u jedinstvenu mjeru na temelju koje se određuje koji od kandidata predstavlja pravi položaj ruke, te je li ruka uopće prisutna.

Rezultat lokalizacije je kontura ruke s označenim stabilnim točkama. Ova kontura koristi se kako bi se na izvornom okviru (prije skaliranja) odredili položaj, veličina i rotacija interesnog područja dlana.

4.2.1. *Detekcija kandidata ruke*

Cilj detekcije kandidata je iz cijele slike dobiti relativno mali skup lokacija za koje je moguće da se na njima nalazi ruka. Postupak detekcije kandidata mora biti vrlo brz kako bi bilo moguće obraditi sliku u realnom vremenu.

Za detekciju kandidata korišten je Viola-Jones detektor objekata koji se pored dobre točnosti ističe velikom brzinom. Viola-Jones detektor objekata detaljno je opisan u poglavlju 4.1.

U Viola-Jones detektoru se obično, nakon detekcije objekata, obavlja grupiranje međusobno prekrivajućih detekcija. Također se eliminiraju grupe s malim brojem članova. U lokalizaciji ruke ovo grupiranje nije rađeno. Umjesto toga, svi kandidati, pa i oni međusobno prekrivajući, prosljeđuju se na podudaranje s modelima. Ovo radimo jer želimo najbolje moguće poklapanje kandidata i modela.

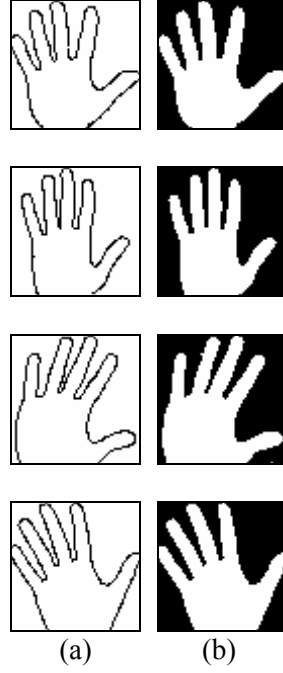
U radu je korištena implementacija Viola-Jones detektora iz OpenCV biblioteke [48].

4.2.2. *Lokalizacija na temelju modela*

U slijedećoj fazi lokalizacije ruke, svaki od kandidata dobivenih u prethodnoj fazi se provjerava tako što se poklapa s velikim brojem modela ruke. Cilj ovog procesa je daljnja verifikacija kandidata ruke, odbacivanje lažnih detekcija i odabir samo onog kandidata ruke čija je mjera poklapanja s modelom najveća.

Postupak poklapanja lokacije na slici s modelom je preuzet iz [42]. Svaki model je predstavljen na dva načina: kao kontura (skup točaka) i kao binarna maska, gdje vrijednost 1 označava ruku, a vrijednost 0 pozadinu.

Za razliku od [42] gdje su modeli ruke dobiveni rotirajući nekoliko osnovnih modela koji predstavljaju osnovne geste ruke, ovdje su modeli dobiveni na temelju baze 1872 ruke koja je snimljena stolnim skenerom u rezoluciji 180 dpi. Ova baza uključuje snimke 243 osobe. Snimke sadrže značajne varijacije u rotaciji ruke te položaju prstiju. Da bi se formirali modeli, ruke su izrezane iz snimaka i skalirane na rezoluciju 64x64 slikovna elementa. Slika 19 prikazuje neke od ovako dobivenih modela ruke.



Slika 19: Neki od modela korišteni za poklapanje s kandidatima ruke, (a) konture, (b) odgovarajuće binarne maske

Za poklapanje kandidata i modela korištene su dvije mjere: udaljenost rubova i vjerodostojnost boje kože. Udaljenost rubova računa se kao prosječna vrijednost kvadrata minimalnih udaljenosti dva skupa točaka, A i B:

$$d(A, B) = \frac{1}{N_a} \sum_{a \in A} \min_{b \in B} \|a - b\|^2 \quad (5)$$

gdje je N_a broj točaka u skupu A. U našem slučaju, A je skup točaka koji čine konturu modela, a B je skup rubnih elemenata na izvornoj slici. Rubovi na izvornoj slici dobiveni su na temelju Cannyjevog detektora rubova [50]. Računanje udaljenosti iz gornje jednadžbe može se ubrzati ako se prethodno izračuna polje istih dimenzija kao i izvorna slika, gdje će vrijednost polja za neki element odgovarati udaljenosti toga elementa do najbližeg rubnog elementa.

Vjerodostojnost boje kože, uz dati model x , računa se kao

$$p(ruka \mid Model\ x) = \prod_{k \in S(Model\ x)} p^s(I(k)) \prod_{k \in \bar{S}(Model\ x)} p^{bg}(I(k)) \quad (6)$$

gdje $S(Model\ x)$ označava skup točaka unutar ruke u modelu x , $\bar{S}(Model\ x)$ označava skup točaka izvan ruke u modelu x , $I(k)$ je boja slikovnog elementa k , a p^s i p^{bg} su distribucija boje kože i distribucija boje pozadine.

Distribucije boje kože i pozadine preuzete su iz [51]. Ove distribucije dobivene su u obliku histograma u RGB prostoru izračunatog na temelju velikog broja ručno segmentiranih slika. Rezolucija R, G i B komponenti ovoga histograma je 32x32x32.

Vjerodostojnost iz jednadžbe 6 lakše je izračunati kao logaritam

$$\log(p(ruka | Model\ x)) = \sum_{k \in S(Model\ x)} \log(p^s(I(k))) + \sum_{k \in \bar{S}(Model\ x)} (\log p^{bg}(I(k))) \quad (7)$$

Jedinstvena mjera podudaranja između lokacije na slici i modela, s , dobiva se tako da se udaljenost rubova i vjerodostojnost boje kože prvo normaliziraju, a zatim se normalizirane vrijednosti zbroje. Za normalizaciju individualnih mjera podudaranja korištena je min-max normalizacija.

Kada brzina lokalizacije ne bi bila važna, u postupku lokalizacije mogli bismo iterativno izračunati mjeru podudaranja svakog kandidata sa svim modelima i uzeti model s najvećom mjerom podudaranja kao rezultat lokalizacije. Međutim, ovaj proces bio bi vrlo spor. Umjesto toga, svaki kandidat se poklapa samo s podskupom modela koji su organizirani u tro-razinsko stablo. Korijski čvor stabla sadrži sve modele, a ostali čvorovi stabla sadrže grupu modela i njihov prototip. Čvorovi na nekoj razini stabla dobiveni su grupirajući modele sadržane u čvoru roditelju na temelju modificiranog jednostavnog heurističkog algoritma grupiranja. Čvorovi na trećoj razini predstavljaju individualne modele.

Postupak izgradnje stabla može se opisati kao rekurzivna procedura dana u slijedećem pseudokodu.

4. LOKALIZACIJA LICA I RUKE

```
izgradi_stablo(čvor, razina, karakteristična_udaljenost)
    ako je razina <= 2
        Grupiraj modele iz trenutnog čvora na temelju
        karakteristične udaljenosti;
        Stvori djecu trenutnog čvora. Svaki čvor dijete
        neka sadrži modele iz jedne grupe i njihov
        prototip;
        Za svako dijete
            izgradi_stablo(dijete, razina+1,
            karakteristična_udaljenost/2);
    inače
        stvori djecu trenutnog čvora. Svaki čvor dijete
        neka sadrži samo jedan model iz trenutnog čvora;
```

Algoritam 2: Izgradnja stabla modela

Modificirani jednostavan heuristički algoritam grupiranja na temelju karakteristične udaljenosti dan je u nastavku

```
Postavi prvi model kao prototip prve grupe;
Za svaki model m
    Ako je udaljenost modela m od svih prototipa veća od
    karakteristične udaljenosti
        Stvori novu grupu i postavi m kao njezin prototip;
Za svaki model m
    Za svaku grupu G
        Ako je udaljenost prototipa od G i modela m manja
        od karakteristične udaljenosti
            Stavi m u G;
```

Algoritam 3: Modificirani heuristički algoritam grupiranja na temelju karakteristične udaljenosti

Potrebno je primijetiti da jedan model, prema gornjem algoritmu, može biti smješten u više različitih grupa. Mjera udaljenosti dvaju modela računa se prema jednadžbi 5. Karakteristična udaljenost na prvoj razini stabla je, na temelju eksperimenata, postavljena na vrijednost 8. Kada je stablo izgrađeno, podudaranje kandidata i stabla modela može se opisati slijedećim algoritmom.

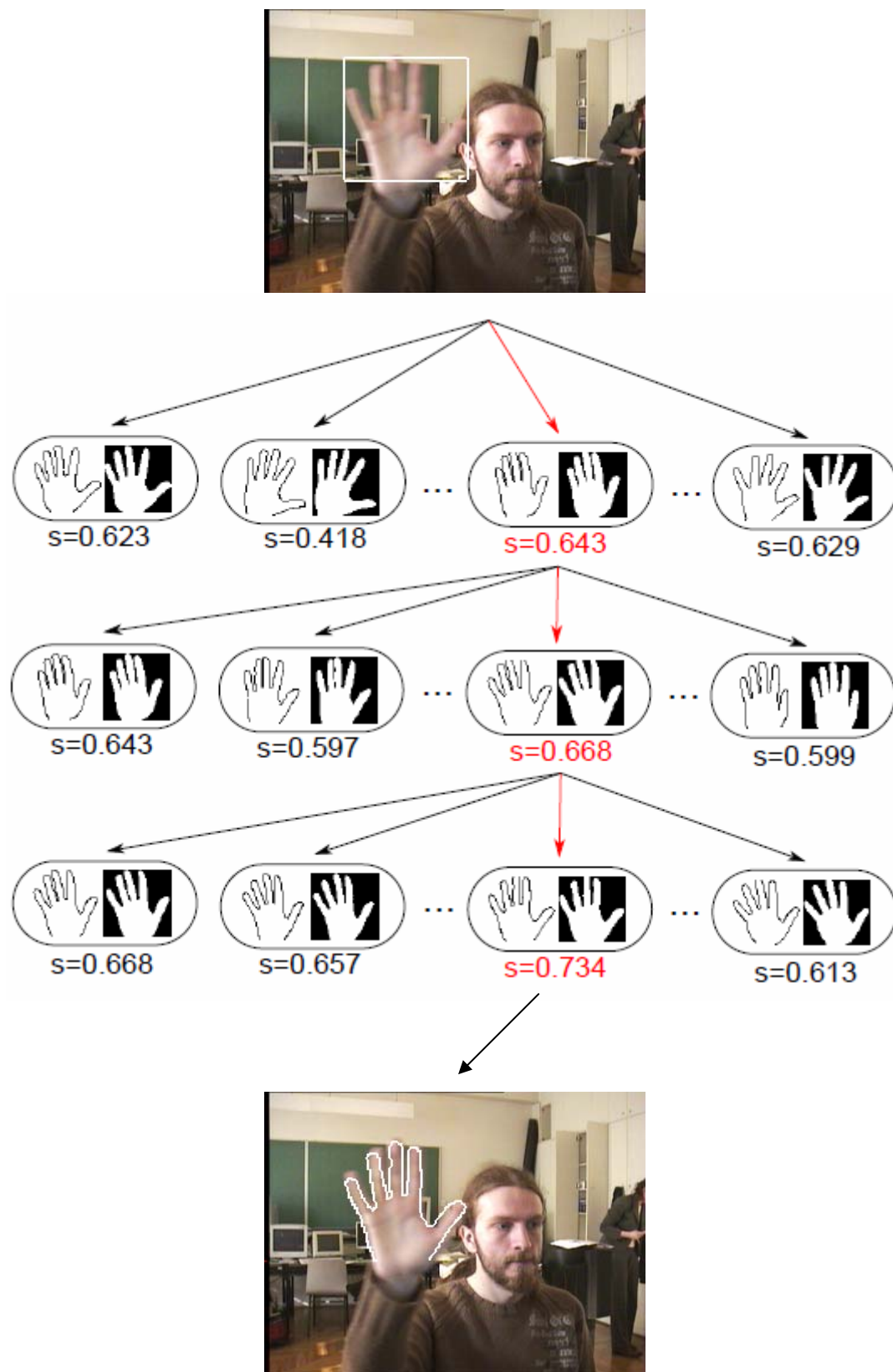
1. Postavi trenutni čvor na korijen.
2. Izračunaj mjeru podudaranja kandidata i prototipa za svako dijete trenutnog čvora.
3. Postavi trenutni čvor na dijete s najvećom mjerom podudaranja.
4. Ako trenutni čvor nije list idi na korak 2.
5. Model s najvećom mjerom podudaranja je onaj sadržan u trenutnom čvoru.

Algoritam 4: Podudaranje kandidata i stabla modela

Primjer podudaranja kandidata sa stablom može se vidjeti na slici 20. Najprije se model uspoređuje sa svim prototipima u čvorovima na 1. razini stabla. Najbolja ostvarena mjera podudaranja na 1. razini stabla je $s = 0.643$. U slijedećem koraku (na razini 2) promatramo sve čvorove djecu čvora koji je u prethodnom koraku ostvario najveću mjeru podudaranja. Također računamo mjeru podudaranja kandidata i svih prototipa u ovim čvorovima. U ovom koraku najbolja ostvarena mjera podudaranja je $s = 0.668$. Sada promatramo sve čvorove djecu čvora koji je ostvario tu mjeru podudaranja. Budući da smo došli do 3. razine stabla, sada će svaki čvor sadržavati samo jedan model, a ne grupu modela i njihov prototip kao u prethodnim koracima. Uspoređujemo kandidata i modele u čvorovima u razini 3 koji su djeca najboljeg čvora iz prethodnog koraka. Najbolja ostvarena mjera podudaranja je $s = 0.734$. Model s kojime je ostvarena ova mjera podudaranja je rezultat, što znači da se kontura toga modela smatra konturom ruke na slici.

Ovo poklapanje potrebno je učiniti za svakog kandidata ruke. Na kraju se odabire kandidat s najvećom mjerom podudaranja. Ukoliko je ova mjera podudaranja manja od praga T (postavljen na temelju eksperimenata na vrijednost 1, polovina očekivane maksimalne vrijednosti mjere podudaranja), smatrati ćemo da trenutni okvir video sekvence ne sadrži ruku. Inače, smatra se da ruka postoji i da je kontura ruke u slici dana konturom modela s najvećom mjerom podudaranja. Svrha usporedbe s pragom je eliminirati lažne detekcije, te slučajeve u kojima ruka nije u položaju pogodnom za izlučivanje značajki dlana (npr. prevelika rotacija u bilo kojem smjeru, poluzatvorena ili zatvorena šaka itd.).

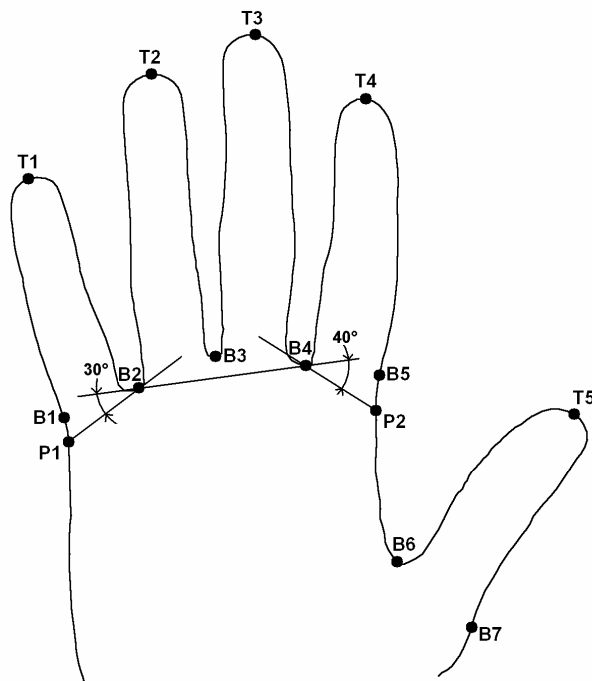
Također su načinjeni eksperimenti sa zmijama [52] koje su bile inicijalizirane na konturu modela s najvećom mjerom podudaranja. Međutim, primijećeno je da zmije nisu dovoljno robusne u uvjetima složene pozadine, naročito ako se u pozadini mogu očekivati snažni rubovi. Stoga se u praksi bolji rezultati dobivaju ako se jednostavno uzme kontura modela s najvećom mjerom podudaranja, bez naknadnog ugađanja.



Slika 20: Primjer podudaranja kandidata ruke i stabla modela

4.2.3. Izlučivanje interesnog područja dlana

Izlučivanje interesnog područja dlana obavlja se na izvornom, neskalinom okviru video sekvence (rezolucije 640x480). Ulaz u modul za izlučivanje interesnog područja dlana, pored okvira video sekvence, je kontura modela s najvećom mjerom podudaranja. Na konturi svakog modela prethodno su označene stabilne točke.



Slika 21: Kontura ruke s označenim stabilnim točkama

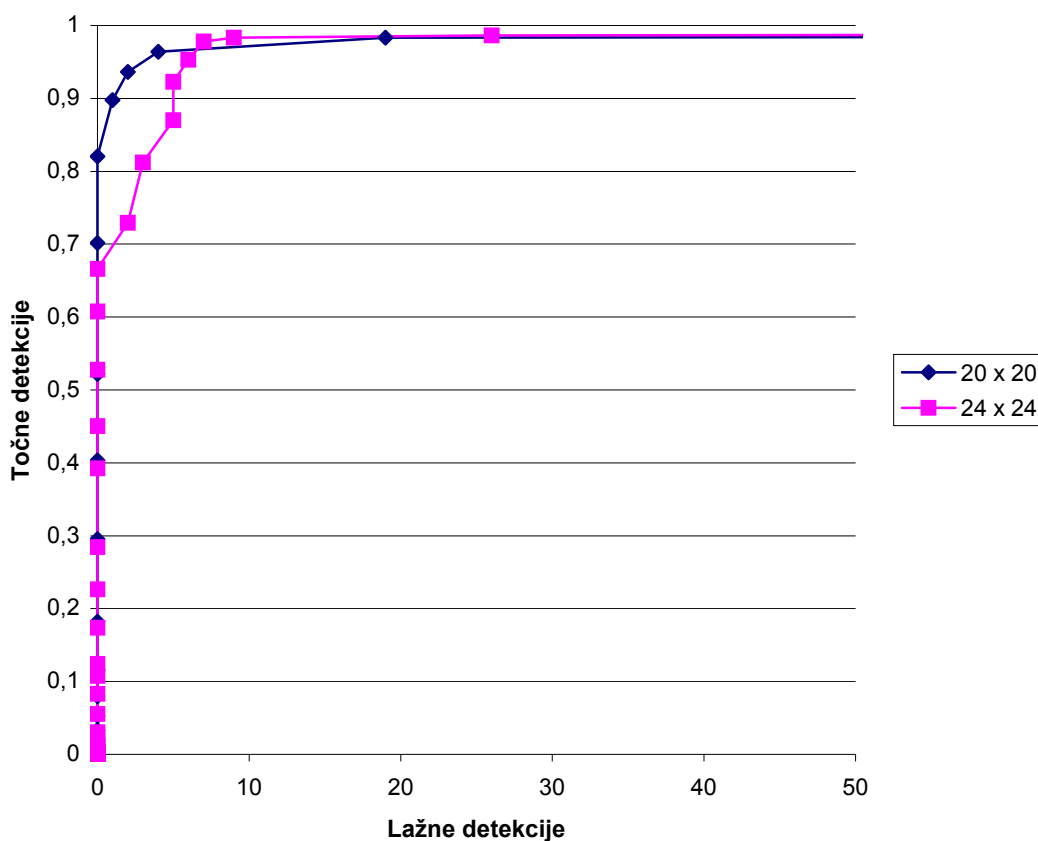
Za nalaženje stabilnih točaka na konturama modela preuzet je postupak iz [19]. Slika 21 prikazuje konturu ruke s označenim stabilnim točkama. Točke T1, T2, T3, T4, T5, te B2, B3, B4 i B6 (vrhovi prstiju i udoline između prstiju) dobivene su kao lokalni maksimumi i minimumi na konturi. Točke B1, B5 i B7 dobivene su zahtijevajući da udaljenost ovih točaka i vrha odgovarajućeg prsta bude jednaka kao udaljenost vrha prsta od udoline s druge strane tog istog prsta. Točke P1 i P2 dobivene su kao na slici 21.

Interesno područje dlana definirano je kao kvadratno područje čija dva vrha padaju u polovišta dužina P1-B2 i B4-P2.

4.2.4. Rezultati lokalizacije

Točnost lokalizacije evaluirana je na bazi video sekvenci ruke. Baza je podijeljena u skup za učenje i skup za ispitivanje. Skup za učenje sadrži 32 video sekvence koje sadrže ruke 10 osoba. Na video sekvencama za učenje ručno je označeno 2138 primjera ruke za učenje Viola-Jones detektora. Skup za ispitivanje sadrži 10 video sekvenci 5 osoba. Osobe u bazi za ispitivanje nisu iste osobe kao osobe u bazi za učenje. Za ispitivanje Viola-Jones detektora, u bazi za ispitivanje ručno su označene ruke u 362 okvira. Baza je dobivena na temelju dvije kamere: CCD kamere visoke kvalitete i CMOS web kamere niske kvalitete slike. Sekvence su prikupljene u različitim svjetlosnim uvjetima i sadrže složenu pozadinu.

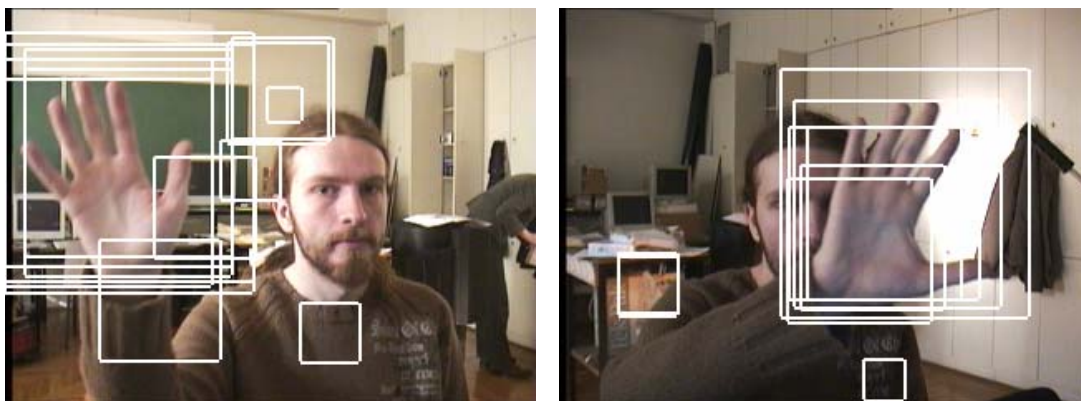
Prvo je ispitana točnost detekcije samo na temelju Viola-Jones detektora, a zatim točnost cijelog postupka lokalizacije temeljenog na modelima. Viola-Jones detektor ispitivan je za dvije osnovne veličine prozora detekcije: 20x20 i 24x24 slikovnih elemenata. Rezultati su prikazani u obliku ROC krivulje na slici 22.



Slika 22: ROC krivulja za detekciju ruke Viola-Jones detektorom za dvije veličine prozora

Kao što se može vidjeti iz slike, Viola-Jones detektor ostvaruje visoku točnost detekcije za obje veličine prozora. U daljnjem radu odabran je prozor manje veličine (20x20) zbog veće

brzine detekcije s manjim prozorom. Detektor uspješno detektira 98.3% ruku uz 19 lažnih detekcija, te uspješno detektira 99.7% ruku uz 713 lažnih detekcija. Razvijeni sustav radi u ovoj drugoj točki na ROC krivulji, s najvećim brojem lažnih detekcija, ali također s najvećim brojem ispravnih detekcija. Lažne detekcije mogu se eliminirati u postupku podudaranja s modelima. Slika 23 prikazuje nekoliko primjera detekcije kandidata.



Slika 23: Primjeri detekcije kandidata ruke Viola-Jones detektorom

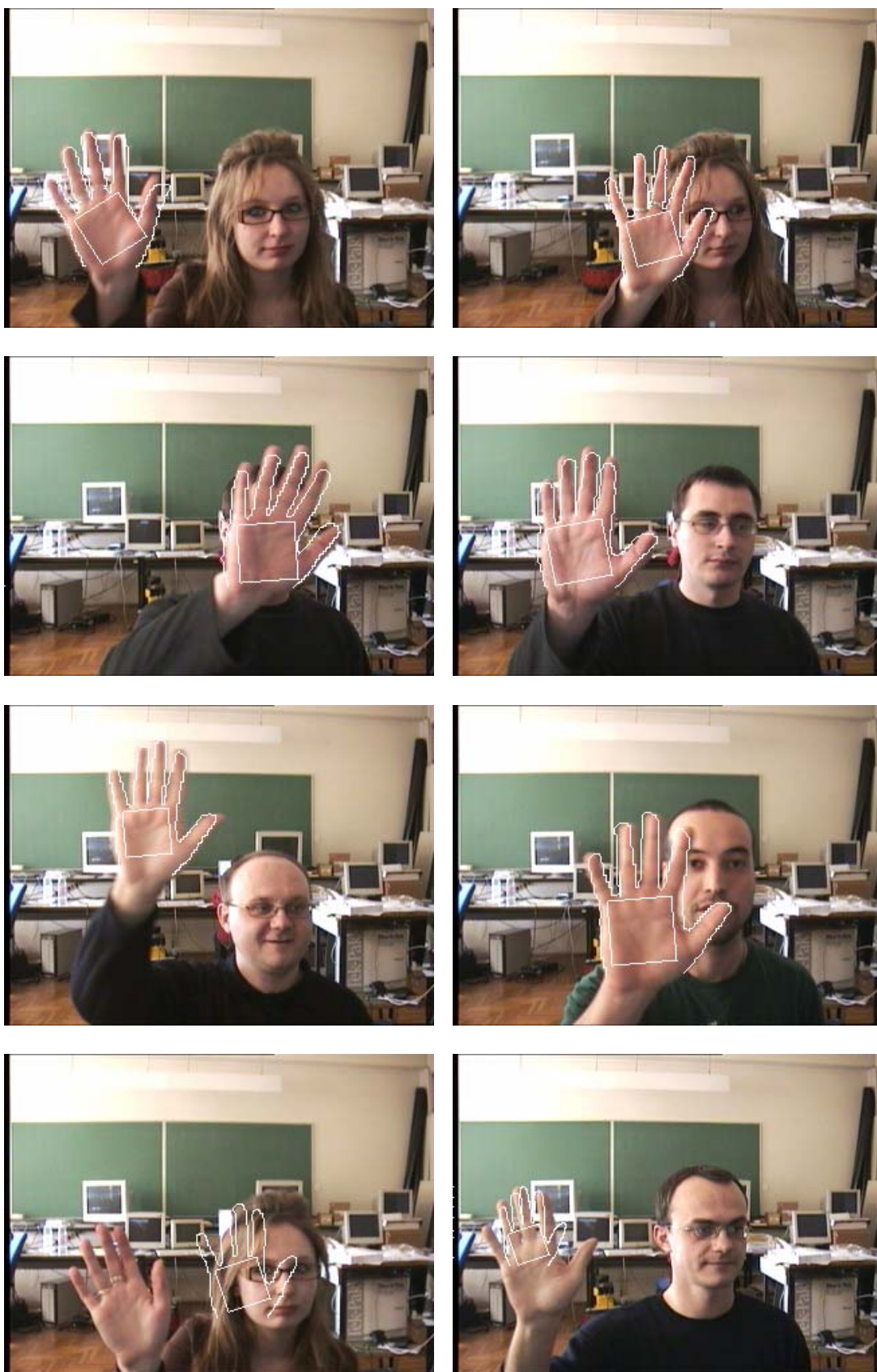
U drugom eksperimentu provedeno je ispitivanje cijelog postupka lokalizacije temeljene na modelima. Slika 24 prikazuje nekoliko uspješnih i neuspješnih rezultata lokalizacije.

Kao što se može vidjeti iz slike 24, lokalizacija uspješno radi u uvjetima složene pozadine, čak i ako je ruka postavljena preko drugog područja boje kože, kao što je lice. Greške obično nastaju ili kao posljedica neprihvatljivog položaja ruke ili kao posljedica postojanja područja na slici koja su boje kože i imaju rubove sličnoga oblika kao rubovi ruke.

Da bi se mogla izmjeriti točnost lokalizacije, definirati ćemo točnu lokalizaciju kao lokalizaciju u kojoj detektirano interesno područje dlana pada unutar dlana na izvornom okviru video sekvence. Uz ovako definiranu točnu lokalizaciju, na skupu uzoraka za ispitivanje, sustav ispravno lokalizira ruku u 96.7% slučajeva.

Na računalu s četverojezgrenim procesorom na 2.4 GHz, u jednodretvenoj implementaciji (koja koristi samo jednu jezgru procesora), obrada jednoga okvira video sekvence traje prosječno 0.13s.

4. LOKALIZACIJA LICA I RUKE



Slika 24: Nekoliko primjera uspješne i neuspješne lokalizacije ruke

5. Normalizacija interesnih područja

Nakon lokalizacije dlana i lica, a prije izlučivanja značajki, potrebno je provesti postupke normalizacije na interesnim područjima. Cilj postupaka normalizacije je priprema interesnih područja za izlučivanje značajki te smanjivanje varijacija unutar razreda u interesnim područjima.

Normalizacija u ovome radu uključuje geometrijsku normalizaciju, svjetlosnu normalizaciju te, u slučaju lica, uklanjanje pozadine.

Geometrijska normalizacija provodi se jer se interesna područja na slici pojavljuju u različitim veličinama i različitim rotacijama. U sklopu geometrijske normalizacije obavlja se transformacija interesnog područja koja uključuje translaciju, rotaciju te skaliranje primjenom kojih se dobiva slika interesnog područja veličine 64x64 slikovnih elemenata.

U slučaju dlana, geometrijska normalizacija će jednostavno rotirati i skalirati kvadratno područje dlana čiji je položaj određen kao u poglavlju 4.2.3. Slika 25 (a) prikazuje nekoliko interesnih područja dlana nakon geometrijske normalizacije.

U slučaju lica, geometrijska normalizacija provodi se tako da se izračuna i primjeni afina transformacija koja će, na normaliziranoj slici, uvijek oči transformirati u iste točke. Kod normalizacije lica, u ovome koraku se također provodi i uklanjanje pozadine. Pretpostavlja se da je lice eliptičnog oblika, pa se uklanjanje pozadine obavlja tako da se, na normaliziranoj slici, svi slikovni elementi izvan elipse visine 64 i širine 45 slikovnih elemenata postave na crnu boju. Slika 26 (a) prikazuje nekoliko interesnih područja lica nakon geometrijske normalizacije.

Nakon geometrijske normalizacije i uklanjanja pozadine, slijedeći korak je normalizacija svjetline čiji je cilj smanjiti varijacije u slikama interesnih područja nastale kao posljedica promjena uvjeta osvjetljenja. Normalizacija svjetline provodi se postupkom poklapanja histograma (engl. histogram matching) [19][53].

U ovom postupku odabire se ciljni histogram $G(l)$ za normalizirano područje interesa dlana i lica. U razvijenom sustavu kao ciljni histogrami dlana i lica uzeti su histogrami prve slike dlana i prve slike lica iz baza dlana i lica koje su dobivene u kontroliranim uvjetima osvjetljenja. Svako interesno područje potrebno je transformirati tako da njegov histogram (opisan histogramskom funkcijom $H(l)$, $l = 0, 1, 2, \dots, 255$; gdje je l intenzitet sivila) postane jednak ciljnom histogramu (opisanom histogramskom funkcijom $G(l)$).

Za ovaj postupak prvo je potrebno pronaći funkcije $f_{H \rightarrow U}(l)$ i $f_{G \rightarrow U}(l)$ koje preslikavaju $H(l)$ i $G(l)$ u histogram uniformne distribucije $U(l)$.

$$f_{H \rightarrow U}(l) = \frac{\sum_{j=0}^l H(j)}{\sum_{j=0}^{L-1} H(j)} \quad (8)$$

$$f_{G \rightarrow U}(l) = \frac{\sum_{j=0}^l G(j)}{\sum_{j=0}^{L-1} G(j)} \quad (9)$$

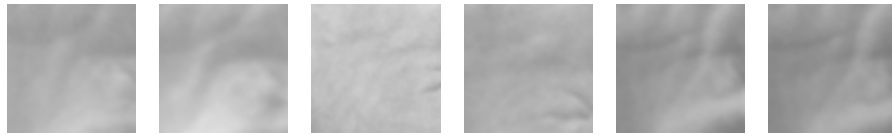
gdje L predstavlja broj razina intenziteta u slikama (256 za slike korištene u implementiranome sustavu).

Da bismo pronašli traženu funkciju preslikavanja $f_{H \rightarrow G}(l)$ koja preslikava histogram slike u ciljni histogram moramo pronaći inverznu funkciju od $f_{G \rightarrow U}(l)$ koju ćemo označiti kao $f_{U \rightarrow G}(l)$. Ova funkcija može se jednostavno implementirati pretraživanjem svih vrijednosti funkcije $f_{G \rightarrow U}(l)$. Kada imamo ovu funkciju, traženu funkciju $f_{H \rightarrow G}(l)$ možemo jednostavno izraziti kao

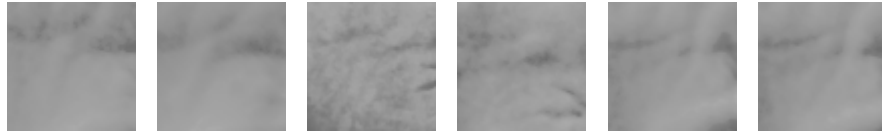
$$f_{H \rightarrow G}(l) = f_{U \rightarrow G}(f_{H \rightarrow U}(l)), \quad l = 0, 1, \dots, L-1 \quad (10)$$

Slika 25 prikazuje nekoliko interesnih područja dlana prije i nakon normalizacije svjetline.

Slika 26 prikazuje nekoliko interesnih područja lica prije i nakon normalizacije svjetline.



(a)



(b)

Slika 25: Nekoliko interesnih područja dlana nakon (a) geometrijske normalizacije i (b) normalizacije svjetline



(a)



(b)

Slika 26: Nekoliko interesnih područja lica nakon (a) geometrijske normalizacije i (b) normalizacije svjetline

6. Izlučivanje značajki dlana i lica

U ovome poglavlju biti će opisane metode izlučivanja značajki dlana i lica koje će kasnije biti evaluirane na dlanu i licu dobivenima iz video sekvenci. Kao predstavnici metoda temeljenih na izgledu biti će opisane analiza glavnih komponenti i linearna diskriminantna analiza. Kao predstavnik metoda temeljenih na lokalnim značajkama biti će opisano izlučivanje značajki pomoću Gaborovog filtra. Također će biti opisana nova metoda izlučivanja značajki, lokalna binarna LDA (LBLDA), koja je zamišljena tako da kombinira dobre strane linearne diskriminantne analize i metoda temeljenih na lokalnim značajkama.

6.1. Analiza glavnih komponenti

Analiza glavnih komponenti (engl. Principal Component Analysis, PCA) [54], također poznata i pod nazivom Karhunen-Loève transformacija te Hotellingova transformacija, je metoda kojoj je osnovna svrha naći potprostor izvornog prostora optimalan za reprezentaciju danog skupa uzoraka u smislu da je udaljenost tako kodiranih uzoraka kada ih se vrati u izvorni prostor i izvornih uzoraka minimalna.

PCA se često koristi kao tehnika za izlučivanje značajki u raspoznavanju uzoraka, pa tako i u biometriji. Do sada je uspješno primijenjena na raspoznavanje lica [23] te dlanova [13]. Iako je poznato da ova metoda ne nalazi nužno one komponente koje su optimalne za raspoznavanje (nego za reprezentaciju) rezultati dobiveni uz pomoć ove metode izlučivanja značajki pokazali su se dobrim u mnogim primjenama.

6.1.1. Izvod Analize glavnih komponenti

Pretpostavimo da nam je na raspolaganju skup n -dimenzionalnih uzoraka organiziran u matricu \mathbf{X} tako da je svaki stupac matrice \mathbf{X} jedan od uzoraka u skupu,

$$\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N] \quad (11)$$

gdje su \mathbf{x}_i , $i = 1, 2, \dots, N$ uzorci, a N broj uzoraka u skupu. Pretpostavimo da su uzorci prethodno normalizirani tako da imaju središte u ishodištu koordinatnog sustava. Razmotrimo transformaciju danu kao vektor stupac \mathbf{w} koji svaki uzorak projicira u jednodimenzionalan prostor,

$$\mathbf{y} = \mathbf{w}^T \mathbf{X} \quad (12)$$

gdje je \mathbf{y} vektor redak koji sadrži projekcije pojedinih uzoraka iz \mathbf{X} . Cilj PCA je naći takvu transformaciju \mathbf{w} nakon koje će varijanca projekcija biti maksimalna, uz uvjet da vektor \mathbf{w} ima jediničnu normu (inače bi norma vektora \mathbf{w} utjecala na varijancu projiciranih uzoraka). To možemo zapisati kao

$$J^{PCA}(\mathbf{w}) = \frac{1}{N} \mathbf{y} \mathbf{y}^T = \frac{1}{N} \mathbf{w}^T \mathbf{X} \mathbf{X}^T \mathbf{w} = \mathbf{w}^T \mathbf{C} \mathbf{w} \quad (13)$$

uz uvjet $\mathbf{w}^T \mathbf{w} = 1$

gdje je \mathbf{C} kovarijacijska matrica uzoraka. Ovo je problem optimizacije uz uvjet koji se rješava upotrebom Lagrangeovih multiplikatora. Tvorimo Lagrangeovu funkciju kao

$$L^{PCA}(\mathbf{w}) = \mathbf{w}^T \mathbf{C} \mathbf{w} + \lambda \mathbf{w}^T \mathbf{w} \quad (14)$$

te tražimo optimalno rješenje u točkama gdje je gradijent ove funkcije jednak nuli,

$$\frac{\partial L^{PCA}(\mathbf{w})}{\partial \mathbf{w}} = (\mathbf{C} + \mathbf{C}^T) \mathbf{w} + 2\lambda \mathbf{w} = 2\mathbf{C} \mathbf{w} + 2\lambda \mathbf{w} \quad (15)$$

Može se uočiti da su rješenje ovog problema zapravo svojstveni vektori matrice \mathbf{C} ,

$$\mathbf{C} \mathbf{w} = \lambda \mathbf{w} \quad (16)$$

pri čemu je Lagrangeov koeficijent λ promijenio predznak. Slijedi da je rješenje koje maksimizira kriterijsku funkciju i daje optimalnu projekciju uzoraka u smislu maksimizacije varijance svojstveni vektor kovarijacijske matrice uzoraka s najvećom odgovarajućom svojstvenom vrijednosti.

Također, nismo ograničeni na korištenje samo jedne projekcije. Umjesto korištenja samo jednog vektora \mathbf{w} koji odgovara najvećoj svojstvenoj vrijednosti možemo uzeti N_{PCA} najvećih svojstvenih vektora matrice \mathbf{C} i poslagati ih u transformacijsku matricu \mathbf{W} ,

$$\mathbf{W} = [\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_{N_{PCA}}] \quad (17)$$

Pri tome transformacija izgleda kao

$$\mathbf{Y} = \mathbf{W}^T \mathbf{X} \quad (18)$$

pri čemu stupci matrice \mathbf{Y} odgovaraju transformiranim stupcima matrice \mathbf{X} . Transformirani uzorci sada imaju N_{PCA} komponenti. Kada bismo takve transformirane uzorke vratili u izvorni prostor inverznom transformacijom

$$\hat{\mathbf{X}} = \mathbf{W} \mathbf{Y} \quad (19)$$

Euklidska udaljenosti između tih i originalnih uzoraka bila bi minimalna u smislu da ne postoji niti jedna druga transformacija koja bi uz istu redukciju dimenzionalnosti prostora dala manju udaljenost.

Projiciramo li skup uzoraka \mathbf{X} koristeći samo jednu glavnu komponentu danu kao vektor \mathbf{w} s odgovarajućom svojstvenom vrijednosti λ tada se varijanca projiciranih uzoraka može izraziti kao

$$E\{y^2\} = E\{\mathbf{w}^T \mathbf{x} \mathbf{x}^T \mathbf{w}\} = \mathbf{w}^T \mathbf{C} \mathbf{w} = \lambda \quad (20)$$

Ovo slijedi iz činjenice da je \mathbf{w} svojstveni vektor od \mathbf{C} kojoj odgovara svojstvena vrijednost λ . Uz pomoć ove jednadžbe srednju kvadratnu pogrešku pri rekonstrukciji skupa uzoraka \mathbf{X} koristeći N_{PCA} najvećih svojstvenih vektora možemo zapisati kao

$$E\{\mathcal{E}^2\} = E\left\{\left(\sum_{i=N_{\text{PCA}}+1}^n y_i \mathbf{w}_i\right)^T \left(\sum_{j=N_{\text{PCA}}+1}^n y_j \mathbf{w}_j\right)\right\} = \sum_{i=N_{\text{PCA}}+1}^n E\{y_i^2\} = \sum_{i=N_{\text{PCA}}+1}^n \lambda_i \quad (21)$$

Oдавде se vidi direktna povezanost između dobivenih svojstvenih vrijednosti i srednje kvadratne pogreške u rekonstrukciji.

6.1.2. Analiza glavnih komponenti za skup slikovnih uzoraka

Obzirom da će PCA u eksperimentalnom dijelu ovoga rada biti korištena za izlučivanje značajki iz dva različita područja interesa, u ovom dijelu dan je općeniti PCA algoritam za skup slikovnih uzoraka.

Pretpostavimo da nam je na raspolaganju N slika, a svaka slika sadrži n elemenata. Prvo je potrebno svaku od N slika prikazati kao vektor stupac. Označiti ćemo te vektore kao $\mathbf{s}_1, \dots, \mathbf{s}_N$.

Potrebno je naći srednju vrijednost ovih vektora kao

$$\mathbf{m} = \frac{1}{N} \sum_{i=1}^N \mathbf{s}_i \quad (22)$$

a zatim oduzeti tu srednju vrijednost od vektora koji predstavljaju slike da bi dobili nove vrijednosti s kojima ćemo računati u daljnjem postupku.

$$\mathbf{a}_i = \mathbf{s}_i - \mathbf{m} \quad i = 1, 2, \dots, N \quad (23)$$

Te vektore organizirati ćemo u matricu \mathbf{A}

$$\mathbf{A} = [\mathbf{a}_1 \mathbf{a}_2 \dots \mathbf{a}_N] \quad (24)$$

Matrica \mathbf{A} dimenzija je $n \times N$. Da bismo izračunali optimalan potprostor u smislu PCA moramo naći svojstvene vektore \mathbf{v}_k kovarijacijske matrice uzoraka koju dobivamo kao

$$\mathbf{C} = \mathbf{A}\mathbf{A}^T \quad (25)$$

Međutim, matrica \mathbf{C} dimenzija je $n \times n$, što je u praksi često veliko. Na primjer, računamo li PCA za skup slika veličine 64×64 slikovnih elemenata matrica \mathbf{C} imati će dimenzije 4096×4096 . Računanje svojstvenih vektora ovakve matrice biti će vrlo dugotrajno – do nekoliko sati na modernom računalu.

Budući da u većini primjena vrijedi $N < n$ (broj slika na raspolaganju manji je od broja slikovnih elemenata svake slike) možemo, umjesto računanja svojstvenih vektora matrice \mathbf{C} izračunati svojstvene vektore matrice \mathbf{C}' [23]:

$$\mathbf{C}' = \mathbf{A}^T \mathbf{A} \quad (26)$$

Ova matrica dimenzija je $N \times N$. Kada izračunamo svojstvene vektore \mathbf{v}'_k i njima odgovarajuće svojstvene vrijednosti λ'_k matrice \mathbf{C}' , možemo izračunati svojstvene vektore \mathbf{v}_k i svojstvene vrijednosti λ_k matrice \mathbf{C} kao

$$\lambda_k = \lambda'_k \quad k = 1, 2, \dots, N \quad (27)$$

$$\mathbf{v}_k = \frac{1}{\sqrt{\lambda_k}} \mathbf{A} \mathbf{v}'_k \quad k = 1, 2, \dots, N \quad (28)$$

Traženi potprostor u kojemu ćemo predstaviti uzorke biti će razapet vektorima \mathbf{v}_k koji odgovaraju najvećim svojstvenim vrijednostima λ_k .

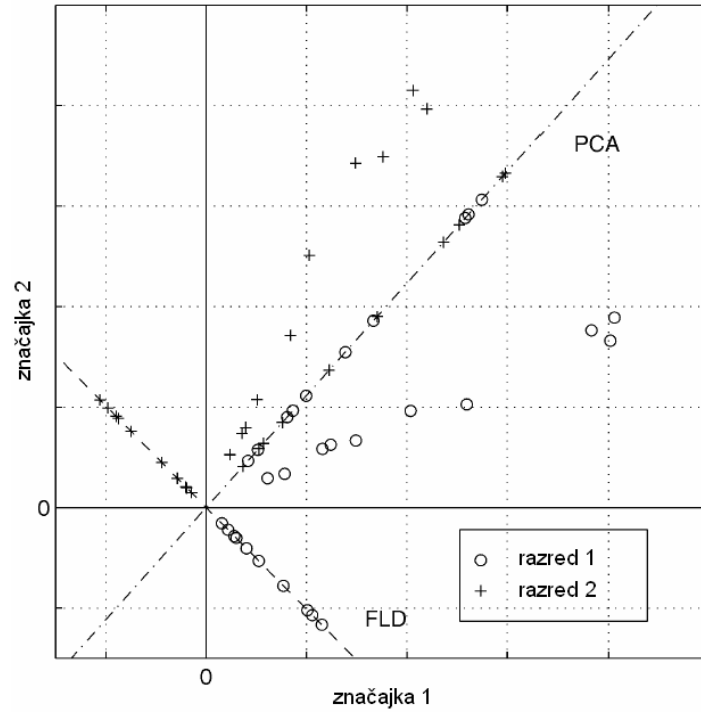
Svaku sliku možemo projicirati u ovaj potprostor kao

$$\eta_k = \mathbf{v}_k^T \cdot \mathbf{a}_s \quad k = 1, 2, \dots, N_{PCA} \quad (29)$$

Gdje je N_{PCA} odabrana dimenzionalnost prostora značajki, a $\mathbf{a}_s = \mathbf{s}_s - \mathbf{m}$, gdje je \mathbf{s}_s vektor koji predstavlja sliku koju želimo prikazati u ovom prostoru. Komponente η_k ; $k = 1, \dots, N_{PCA}$ čine vektor koji ćemo koristiti za reprezentaciju slike.

6.2. Linearna diskriminantna analiza

Linearna diskriminantna analiza [55], (engl. Linear Discriminant Analysis, LDA), poznata još i kao Fisherova linearna diskriminanta (engl. Fisher Linear Discriminant, FLD) je metoda redukcije prostora značajki koja želi postići maksimalnu odvojivost razreda u novom prostoru. Za razliku od PCA koja ne traži eksplicitno potprostor pogodan za raspoznavanje, već za reprezentaciju, LDA traži prostor pogodan za raspoznavanje uzoraka u smislu da se maksimizira kriterij izražen kao omjer raspršenja između razreda i raspršenja unutar razreda. Slika 27, preuzeta iz [24], tako demonstrira jedan slučaj kada redukcijom dvodimenzionalnog prostora u jednodimenzionalni PCA ne uspijeva pronaći potprostor u kome su razredi razdvojivi, dok LDA uspijeva.



Slika 27: Primjer nalaženja potprostora pomoću PCA i LDA za skup dvodimenzionalnih uzoraka dvaju razreda. U ovom slučaju LDA nalazi potprostor u kome su razredi odvojivi dok PCA to ne uspijeva. Slika preuzeta iz [24]

Pretpostavimo da želimo naći samo jednu projekciju (tj. reducirati prostor u jednodimenzionalni) za skup uzoraka \mathbf{x}_j , $j = 1, 2, \dots, N$ iz dva razreda, ω_1 i ω_2 , koja maksimizira zadani kriterij. Za svaki razred i možemo izračunati središte \mathbf{m}_i . Projekcija svakog uzorka računa se kao

$$y_i = \mathbf{w}^T \mathbf{x}_i \quad (30)$$

gdje je \mathbf{w} vektor stupac koji sadrži smjer projekcije. Srednja vrijednost projiciranih uzoraka za svaki razred dobiva se kao

$$\tilde{m}_i = \frac{1}{N_i} \sum_{y \in \omega_i} y \quad (31)$$

gdje je N_i broj uzoraka iz razreda ω_i . Varijancu projiciranih uzoraka za svaki razred dobivamo kao

$$\tilde{\sigma}_i^2 = \sum_{y \in \omega_i} (y - \tilde{m}_i)^2 \quad (32)$$

Želimo maksimizirati kriterijsku funkciju

$$J^{LDA}(\mathbf{w}) = \frac{(\tilde{m}_1 - \tilde{m}_2)^2}{\tilde{\sigma}_1^2 + \tilde{\sigma}_2^2} \quad (33)$$

tj. želimo postići da su istovremeno uzorci iz različitih razreda međusobno maksimalno udaljeni dok su uzorci iz istog razreda međusobno maksimalno kompaktni. Brojnik možemo raspisati kao

$$(\tilde{m}_1 - \tilde{m}_2)^2 = (\mathbf{w}^T m_1 + \mathbf{w}^T m_2)^2 = \mathbf{w}^T (\mathbf{m}_1 - \mathbf{m}_2)(\mathbf{m}_1 - \mathbf{m}_2)^T \mathbf{w} = \mathbf{w}^T \mathbf{S}_B \mathbf{w} \quad (34)$$

Gdje smo sa \mathbf{S}_B označili matricu raspršenja između razreda.

$$\mathbf{S}_B = (\mathbf{m}_1 - \mathbf{m}_2)(\mathbf{m}_1 - \mathbf{m}_2)^T \quad (35)$$

Slično, varijancu projiciranih uzoraka za svaki razred možemo raspisati kao

$$\tilde{\sigma}_i^2 = \sum_{\mathbf{x} \in \omega_i} (\mathbf{w}^T \mathbf{x} - \mathbf{w}^T \mathbf{m}_i)^2 = \sum_{\mathbf{x} \in \omega_i} \mathbf{w}^T (\mathbf{x} - \mathbf{m}_i)(\mathbf{x} - \mathbf{m}_i)^T \mathbf{w} = \mathbf{w}^T \mathbf{S}_i \mathbf{w} \quad (36)$$

gdje smo sa \mathbf{S}_i označili matricu raspršenja razreda ω_i .

$$\mathbf{S}_i = \sum_{\mathbf{x} \in \omega_i} (\mathbf{x} - \mathbf{m}_i)(\mathbf{x} - \mathbf{m}_i)^T \quad (37)$$

Kao \mathbf{S}_W označiti ćemo ukupnu matricu raspršenja unutar razreda

$$\mathbf{S}_W = \mathbf{S}_1 + \mathbf{S}_2 \quad (38)$$

Sada možemo kriterijsku funkciju zapisati u obliku

$$J^{LDA}(\mathbf{w}) = \frac{\mathbf{w}^T \mathbf{S}_B \mathbf{w}}{\mathbf{w}^T \mathbf{S}_W \mathbf{w}} \quad (39)$$

Ovaj problem poznat je kao generalizirani Rayleighov kvocijent. Rješenje problema može se dobiti izjednačavanjem gradijenta kriterijske funkcije s nulom:

$$\frac{\partial J^{LDA}(\mathbf{w})}{\partial \mathbf{w}} = \frac{(\mathbf{w}^T \mathbf{S}_W \mathbf{w})(2\mathbf{S}_B \mathbf{w}) - (\mathbf{w}^T \mathbf{S}_B \mathbf{w})(2\mathbf{S}_W \mathbf{w})}{(\mathbf{w}^T \mathbf{S}_W \mathbf{w})^2} = \vec{0} \quad (40)$$

Vidimo da, ukoliko želimo da gornji razlomak bude jednak nul-vektoru, tada brojnik treba biti nul-vektor

$$(\mathbf{w}^T \mathbf{S}_W \mathbf{w})(2\mathbf{S}_B \mathbf{w}) - (\mathbf{w}^T \mathbf{S}_B \mathbf{w})(2\mathbf{S}_W \mathbf{w}) = \vec{0} \quad (41)$$

odnosno, prebacivanjem drugog člana na desnu stranu jednadžbe

$$(\mathbf{w}^T \mathbf{S}_W \mathbf{w})(2\mathbf{S}_B \mathbf{w}) = (\mathbf{w}^T \mathbf{S}_B \mathbf{w})(2\mathbf{S}_W \mathbf{w}) \quad (42)$$

Dijeljenjem s dva i množenjem s $(\mathbf{w}^T \mathbf{S}_W \mathbf{w})^{-1}$ dobiva se

$$\mathbf{S}_W \mathbf{w} (\mathbf{w}^T \mathbf{S}_B \mathbf{w}) (\mathbf{w}^T \mathbf{S}_W \mathbf{w})^{-1} = \mathbf{S}_B \mathbf{w} \quad (43)$$

Može se primijetiti da je $(\mathbf{w}^T \mathbf{S}_B \mathbf{w})(\mathbf{w}^T \mathbf{S}_W \mathbf{w})^{-1}$ skalar. Uvodimo supstituciju

$$\lambda = (\mathbf{w}^T \mathbf{S}_B \mathbf{w})(\mathbf{w}^T \mathbf{S}_W \mathbf{w})^{-1} \quad (44)$$

i kao rješenje optimizacijskog problema dobivamo

$$\mathbf{S}_B \mathbf{w} = \lambda \mathbf{S}_W \mathbf{w} \quad (45)$$

Ovo je generalizirani problem svojstvenih vektora (engl. generalized eigenvector problem), što znači da su rješenje optimizacijskog problema generalizirani svojstveni vektori sustava matrica \mathbf{S}_B , \mathbf{S}_W .

Ako je matrica \mathbf{S}_W nesingularna možemo ovo pretvoriti u običan problem traženja svojstvenih vektora kao

$$\mathbf{S}_W^{-1}\mathbf{S}_B\mathbf{w} = \lambda\mathbf{w} \quad (46)$$

Za slučaj kada tražimo samo jedan smjer projekcije za dva razreda postoji i pojednostavljeno rješenje. Budući da nas zanima samo smjer i budući da je $\mathbf{S}_B\mathbf{w}$ uvijek u smjeru vektora ($\mathbf{m}_1 - \mathbf{m}_2$) dobivamo za ovaj slučaj rješenje za smjer projekcije

$$\mathbf{w} = \mathbf{S}_W^{-1}(\mathbf{m}_1 - \mathbf{m}_2) \quad (47)$$

Za više razreda i višedimenzionalne prostore moguće je dobiti, slično kao i kod PCA, više projekcija koji čine potprostor optimalan za raspoznavanje u smislu maksimizacije raspršenja između razreda i minimizacije raspršenja unutar razreda. Ovo poopćenje LDA često se naziva i MDA (engl. Multiple Discriminant Analysis).

Postupak je vrlo sličan onome za dva razreda, a jedina razlika je što se sada matrice raspršenja između razreda, \mathbf{S}_B , i unutar razreda, \mathbf{S}_W , računaju drugačije:

$$\begin{aligned} \mathbf{S}_B &= \sum_{i=1}^{N_C} N_i (\mathbf{m}_i - \mathbf{m})(\mathbf{m}_i - \mathbf{m})^T \\ \mathbf{S}_W &= \sum_{i=1}^{N_C} \sum_{\mathbf{x} \in \omega_i} (\mathbf{x} - \mathbf{m}_i)(\mathbf{x} - \mathbf{m}_i)^T \end{aligned} \quad (48)$$

gdje je \mathbf{m} središte svih uzoraka, N_C broj razreda, a N_i broj uzoraka razreda ω_i . Kriterijska funkcija identična je kao za slučaj sa dva razreda (jednadžba 39), te su rješenja ponovno generalizirani svojstveni vektori sustava matrica \mathbf{S}_B , \mathbf{S}_W . U slučaju više razreda ovo nije moguće pojednostaviti kao u slučaju s dva razreda te je potrebno zaista izračunati generalizirane svojstvene vektore da bi se problem riješio.

Kada riješimo ovaj problem možemo uzeti N_{LDA} svojstvenih vektora s najvećim svojstvenim vrijednostima te od njih stvoriti potprostor u koji ćemo projicirati sve uzorke i u postupku raspoznavanja koristiti takvu reprezentaciju uzoraka.

Čest problem u primjenama LDA je singularnost matrice \mathbf{S}_W . Ovaj problem poznat je kao problem maloga broja uzoraka (engl. small sample size problem). Ukoliko je matrica \mathbf{S}_W singularna, što se događa u slučajevima kada je $n > N - N_C$ gdje je n dimenzionalnost uzoraka, nije moguće izravno izračunati generalizirane svojstvene vektore. Taj problem najčešće se rješava redukcijom dimenzionalnosti prostora pomoću PCA prije uporabe LDA. Taj pristup korišten je i u eksperimentalnome dijelu ovoga rada.

Međutim, u literaturi su predloženi i drugi načini dobivanja optimalne transformacije, bez potrebe prethodnog smanjenja dimenzionalnosti uzoraka, kao što su direktna LDA [26], diskriminatorski zajednički vektori [27] i regularizirana LDA [25].

Direktna LDA (DLDA) [26] radi tako da prvo eliminira nul-prostor matrice \mathbf{S}_B (Nul prostor matrice \mathbf{S}_B su svi vektori \mathbf{w} za koje vrijedi $\mathbf{S}_B \mathbf{w} = \mathbf{0}$). Na taj se način zapravo eliminira potprostor izvornog prostora značajki u kojemu ne postoji razlika između središta razreda. Nakon toga, u preostalom prostoru se nalazi optimalna transformacija u smislu LDA kriterijske funkcije.

Ideja diskriminatorskih zajedničkih vektora [27] je komplementarna ideji direktne LDA. Umjesto eliminiranja nul-prostora matrice \mathbf{S}_B diskriminatorski zajednički vektori koriste samo nul-prostor matrice \mathbf{S}_W . Ovo se radi na način da se odbacuju sve značajke u smjeru svojstvenih vektora matrice \mathbf{S}_W za koje su svojstvene vrijednosti različite od nule. Nakon toga, provođenjem PCA na preostalim značajkama maksimizira se modificirana LDA kriterijska funkcija

$$J^{LDA}(\mathbf{w}) = \frac{\mathbf{w}^T \mathbf{S}_B \mathbf{w}}{\mathbf{w}^T \mathbf{S}_T \mathbf{w}} \quad (49)$$

gdje je \mathbf{S}_T ukupna matrica raspršenja.

Regularizirana LDA (RLDA) [25] također koristi ovu modificiranu kriterijsku funkciju, a problem singularnosti matrice \mathbf{S}_T rješava na način da joj se doda jedinična matrica pomnožena s nekim malim skalarom ϵ , koji se naziva regularizacijski parametar. Ovom modifikacijom matrica \mathbf{S}_T postaje nesingularna.

6.3. Izlučivanje značajki Gaborovim filtrom

Gaborov filtar [10][56][57] je često korišten u obradi slike i zvuka, segmentaciji slike i za izlučivanje značajki za raspoznavanje. Impulsni odziv ovog filtra je harmonička funkcija pomnožena s Gaussovom funkcijom. Gaborov filtar pogodan je za izlučivanje značajki zbog svoga svojstva da je istovremeno osjetljiv u prostornoj i frekvencijskoj domeni.

2D Gaborov filtar, koji se koristi za izlučivanje značajki slike, definiran je kao

$$G(x, y, \theta, u, \sigma) = \frac{1}{2\pi\sigma^2} \exp\left\{-\frac{x^2 + y^2}{2\sigma^2}\right\} \exp\{2\pi i(ux \cos \theta + uy \sin \theta)\} \quad (50)$$

Parametri filtra su

- u - frekvencija sinusoide
- σ - standardna devijacija Gaussove funkcije
- θ - parametar koji kontrolira orijentaciju filtra.

Slika 28 prikazuje nekoliko primjera Gaborovih filtara s različitim parametrom θ . Budući da je Gaborov filtar kompleksan, svaki filtar se sastoji od realnog i imaginarnog dijela.

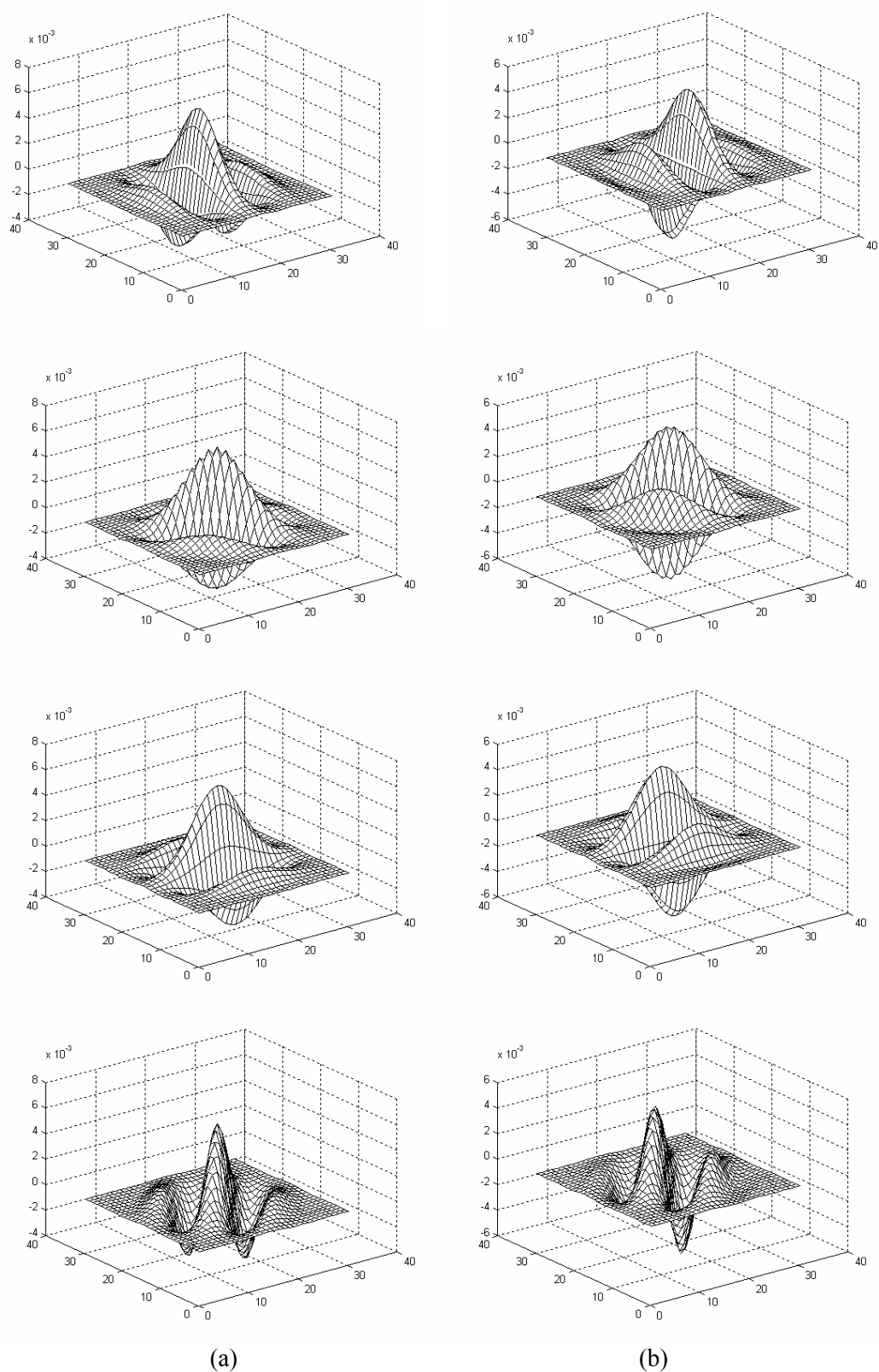
Za primjenu u raspoznavanju, kako filtar ne bi bio osjetljiv na razlike u intenzitetu svjetline slika, najčešće se filtar svodi na srednju vrijednost 0 (eliminira se DC komponenta filtra)

$$\tilde{G}(x, y, \theta, u, \sigma) = G(x, y, \theta, u, \sigma) - \frac{\sum_{i=-n}^n \sum_{j=-n}^n G(i, j, \theta, u, \sigma)}{(2n+1)^2} \quad (51)$$

gdje je $(2n+1) \times (2n+1)$ veličina diskretnog filtra.

Postoje različiti načini na koje se Gaborov filtar može primijeniti za izlučivanje značajki [10][32][56][57]. U ovome radu Gaborov filtar korišten je da bi se izlučile binarne značajke dlana i lica, na sličan način kako je to učinjeno za izlučivanje značajki za raspoznavanje osoba na temelju šarenice oka [56] i dlana [10].

U ovim primjenama binarni kod se dobiva kodiranjem faze odziva Gaborovog filtra. Slika 29 prikazuje ovo kodiranje. Binarni kod dobiva se tako da se za svaku točku u kojoj se promatra odziv filtra uzme samo predznak realnog i predznak imaginarnog dijela odziva.

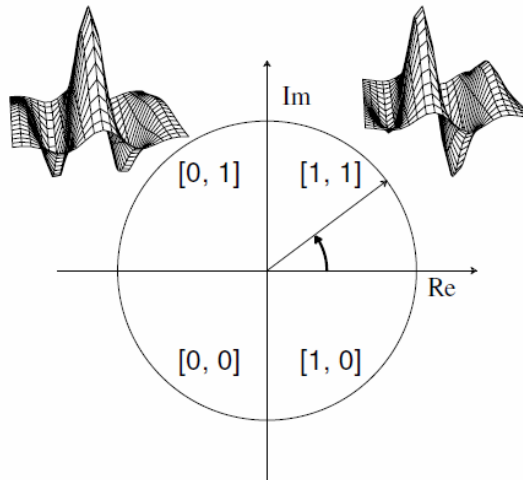


Slika 28: Primjeri Gaborovih filtara s različitim parametrom orijentacije, (a) realni dio, (b) imaginarni dio

Ovako dobiveni binarni kod za dvije slike može se usporediti koristeći normaliziranu Hammingovu udaljenost. Za dva binarna vektora značajki, \mathbf{b}_1 i \mathbf{b}_2 duljine n , normalizirana Hammingova udaljenost je

$$d_{Hamming}(\mathbf{b}_1, \mathbf{b}_2) = \frac{\sum_{i=1}^N \mathbf{b}_1[i] \otimes \mathbf{b}_2[i]}{n} \quad (52)$$

gdje simbol \otimes označava operaciju isključivo ili na bitovima. Da bi se ovakvom usporedbom ostvarilo najbolje moguće poklapanje bilo bi potrebno da su slike interesnih područja iz kojih se izlučuju značajke savršeno poravnate. Kako to obično nije slučaj, poklapanje se obavlja tako da se izračuna vektor značajki koji odgovara translaciji jednoga od područja interesa i zatim ponovno izračuna Hammingova udaljenost. Kao rezultat poklapanja uzima se najmanja Hammingova udaljenost za bilo koju translaciju (unutar nekog intervala, obično ± 2 po x i ± 2 po y osi).



Slika 29: Kodiranje faze odziva Gaborovog filtra

6.4. Lokalna binarna LDA

U ovome poglavlju opisana je nova metoda izlučivanja značajki iz biometrijskih slika, nazvana lokalna binarna LDA (LBLDA), koja je osmišljena tako da kombinira dobre strane linearne diskriminantne analize i metoda koje se temelje na lokalnim značajkama. LBLDA se temelji na dvije osnovne ideje: nalaženju lokalnih značajki koje nose najveću diskriminatornu vrijednost i korištenju binarnih značajki umjesto točnih vrijednosti.

Prednost lokalnih značajki, u odnosu na globalne, je manja osjetljivost na promjene u osvjetljenju. Lokalne značajke u LBLDA dobivaju se tako da se slika podijeli na skup međusobno prekrivajućih područja, a zatim se izvede LDA za svako područje zasebno. Ovo omogućuje nalaženje optimalnih lokalnih značajki u smislu LDA kriterijske funkcije. Na temelju ove kriterijske funkcije također je moguće koristiti više značajki iz onih područja koja sadrže više diskriminatorne informacije.

Binarne značajke mogu se dobiti uzimanjem predznaka realnih značajki. Sun i Tan [58] opisali su neke prednosti korištenja binarnih značajki naspram točnih vrijednosti:

- (i) Mjerenja izražena kao točne vrijednosti su osjetljiva na promjene osvjetljenja, zamućenje, šum, deformacije i druge degradacije na slici. Precizni modeli temeljeni na točnim mjerenjima korisni su za reprezentaciju slike, no ne nužno za raspoznavanje.
- (ii) Binarne značajke su jednostavnije za zapisati i obraditi na računalu
- (iii) Postoji biološka motivacija za korištenje binarnih značajki. Na primjer, istraživanja [59] sugeriraju da neuroni vrlo brzo ulaze u zasićenje s promjenom intenziteta vizualnih podražaja. Ovo sugerira da je kod vizualne percepcije važnija polarnost kontrasta od intenziteta.

Prvi korak LBLDA je podjela slike na skup međusobno prekrivajućih područja. Ovo je najlakše izvesti metodom klizajućeg prozora: Prozor veličine $p \times p$ slikovnih elemenata postavi se u gornji lijevi kut slike. Prvo područje sastoji se od svih slikovnih elemenata koji padaju unutar ovog prozora. Prozor se zatim translatira za $q \leq p$ slikovnih elemenata udesno i, kada prozor izađe iz okvira slike, pomiče se q slikovnih elemenata dolje i poravnava uz lijevi rub slike. Postupak završava kada donji desni kut prozora dosegne donji desni kut slike. Svaki položaj prozora definira jedno od N_R područja R_r , $r = 1, 2, \dots, N_R$, gdje se svako područje sastoji od $p \times p$ slikovnih elemenata.

6. IZLUČIVANJE ZNAČAJKI DLANA I LICA

Za svako područje R_r i svaku sliku u skupu za učenje tvori se vektor \mathbf{x}_i^r , $i = 1, 2, \dots, N$; $r = 1, 2, \dots, N_R$ na način da se intenziteti svih slikovnih elemenata slike i koji upadaju u područje R_r poredaju u vektor. Veličina vektora \mathbf{x}_i^r je $p \times p$. Za svako područje R_r :

- (i) izvede se lokalna PCA na vektorima \mathbf{x}_i^r , $i = 1, 2, \dots, N$. Ovo rezultira potprostorom \mathbf{W}_{PCA}^r . Projiciramo svaki od vektora \mathbf{x}_i^r u ovaj potprostor i tako dobivamo vektore \mathbf{z}_i^r . Veličina vektora \mathbf{z}_i^r je N_{PCA} , $N_{PCA} \leq \min(N-1, p \times p)$.
- (ii) izvede se LDA na vektorima \mathbf{z}_i^r , $i = 1, 2, \dots, N$. Ovim postupkom dobivamo potprostor \mathbf{W}_{LDA}^r .
- (iii) Množeći transformacijske matrice \mathbf{W}_{PCA}^r i \mathbf{W}_{LDA}^r dobivamo konačni potprostor za područje R_r , $\mathbf{W}_{PCA+LDA}^r$. Ovaj potprostor razapet je LDA baznim vektorima \mathbf{w}_j^r , $j = 1, 2, \dots, N_{LDA}$. N_{LDA} je dimenzionalnost potprostora. $N_{LDA} = \min(N_C - 1, p \times p, N_{PCA})$. Uz svaki od vektora \mathbf{w}_j^r također ćemo zabilježiti vrijednost odgovarajuće svojstvene vrijednosti λ_j^r .

Sada imamo skup $N_{LDA} \times N_R$ vektora \mathbf{w}_j^r i odgovarajuće svojstvene vrijednosti λ_j^r . $N_{LDA} \times N_R$ može biti prilično veliki broj. Na primjer, za slike veličine 64×64 slikovnih elemenata uz $p = 16$ i $q = 8$, može se dobiti do 12544 vektora \mathbf{w}_j^r . Stoga je potreban način da se iz $N_{LDA} \times N_R$ vektora \mathbf{w}_j^r odabere manji podskup koji će omogućiti izlučivanje značajki pogodnih za verifikaciju.

Svojstvene vrijednosti λ_j^r daju informaciju o dobroti vektora \mathbf{w}_j^r u smislu LDA kriterijske funkcije (jednadžba 39). Preciznije, može se pokazati da su svojstvene vrijednosti upravo iznosi kriterijske funkcije za odgovarajuće svojstvene vektore.

Da bismo ovo pokazali, pretpostavimo da su \mathbf{w}_1 i λ_1 jedno rješenje optimizacije LDA kriterijske funkcije i odgovarajuća svojstvena vrijednost. To znači da je \mathbf{w}_1 generalizirani svojstveni vektor i vrijedi

$$\mathbf{S}_B \mathbf{w}_1 = \lambda_1 \mathbf{S}_W \mathbf{w}_1 \quad (53)$$

Iz ovoga slijedi da je vrijednost kriterijske funkcije za vektor \mathbf{w}_1 jednaka

$$J^{LDA}(\mathbf{w}_1) = \frac{\mathbf{w}_1^T \mathbf{S}_B \mathbf{w}_1}{\mathbf{w}_1^T \mathbf{S}_W \mathbf{w}_1} = \frac{\mathbf{w}_1^T (\lambda_1 \mathbf{S}_W \mathbf{w}_1)}{\mathbf{w}_1^T \mathbf{S}_W \mathbf{w}_1} = \frac{\lambda_1 \mathbf{w}_1^T \mathbf{S}_B \mathbf{w}_1}{\mathbf{w}_1^T \mathbf{S}_W \mathbf{w}_1} = \lambda_1 \quad (54)$$

U skladu s ovim rješenjem, da bismo dobili najviše diskriminatorne značajke, vektore \mathbf{w}_j^r sortiramo prema padajućem redoslijedu odgovarajućih svojstvenih vrijednosti. λ_j^r . Lokalne

baze potprostora koje ćemo koristiti za izlučivanje značajki dobivamo tako da uzmemo prvih $N_{\text{LBLDA}} \leq N_R \times N_{\text{LDA}}$ vektora \mathbf{w}_j^r iz ovog padajućeg slijeda.

Ovakvim postupkom odabira lokalnih baza možemo uzeti više značajki iz područja na slici koja su više diskriminatorna i manje značajki iz područja koja su manje diskriminatorna. Također je moguće da se iz nekih područja na slici uopće neće izlučivati značajke.

Naposljetku, lokalne baze organiziramo u strukturu podataka koju nazivamo lokalni potprostor. Lokalni potprostor se sastoji od N_{LBLDA} zapisa, gdje svaki zapis sadrži indeks područja r i lokalnu bazu \mathbf{w}_k , gdje je \mathbf{w}_k k -ti vektor u padajućem slijedu vektora \mathbf{w}_j^r .

U fazi verifikacije, koristeći lokalni potprostor, iz nepoznate slike I možemo izlučiti N_{LBLDA} -dimenzionalni vektor značajki \mathbf{y} kako slijedi. Prvo ćemo sliku I podijeliti na N_R područja na identičan način kako je to učinjeno sa slikama u skupu za učenje. k -tu komponentu vektora značajki, y_k , dobiti ćemo tako da izračunamo skalarni produkt \mathbf{w}_k i $(p \times p)$ -dimenzionalnog vektora dobivenog kao konkatencijom intenziteta slikovnih elemenata slike u području R_r , gdje \mathbf{w}_k i r čine k -ti zapis lokalnog potprostora.

Da bismo dobili binarni vektor značajki, \mathbf{b} , uzeti ćemo predznake komponenti vektora \mathbf{y} ($b_k = 1$ za $y_k > 0$; inače $b_k = 0$). Korištenjem binarnog vektora značajki naspram izvornog vektora značajki s realnim vrijednostima, u eksperimentima je ostvareno značajno povećanje točnosti raspoznavanja. Uzimanjem predznaka komponenti vektora \mathbf{y} zapravo koristimo samo informaciju o tome je li korelacija između slikovnih elemenata u području R_r i lokalne baze \mathbf{w}_k pozitivna ili negativna, a zanemarujemo točan iznos korelacije.

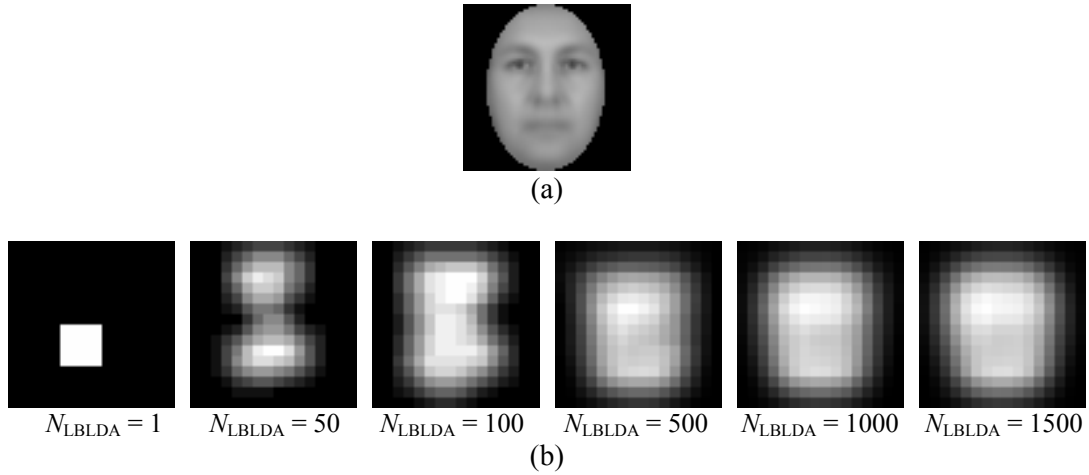
Alternativan način interpretacije ovakvih lokalnih značajki je da ih se promatra kao odzive filtera. Umjesto da koristimo predefinirane filtre, kao što je Gaborov filter, filtre učimo na skupu slika za učenje, i to za svaku lokaciju na slici posebno, tako da dobiveni filteri naglašavaju razlike između razreda dok potiskuju varijacije unutar istoga razreda. Da bismo izlučili značajke, na svako područje u slici primijenimo odgovarajuće filtre i uzimamo samo predznak odziva na sličan način kako je to rađeno s Gaborovim filtrom za izlučivanje značajki šarenice oka [56] i dlana [10].

Dva binarna vektora značajki koji odgovaraju dvjema slikama dlana ili lica mogu se usporediti na temelju Hammingove udaljenosti.

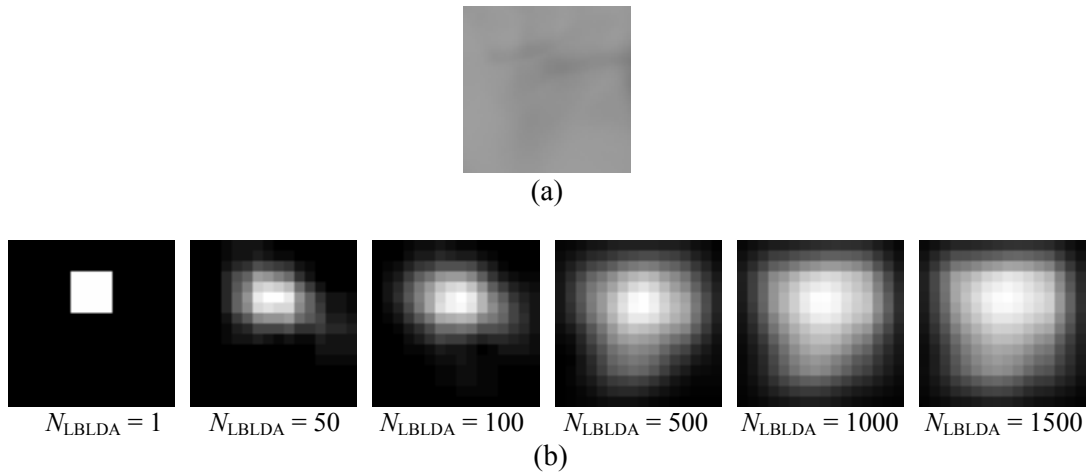
LBLDA će uzeti više značajki iz područja na slici koja sadrže više diskriminatorne informacije. U nastavku su prikazana ovakva područja za dvije baze statičnih slika lica i dlana. Kao baza lica korištena je XM2VTS baza [60] koja sadrži 2360 slika 295 osoba (8 slika po osobi). Baza dlana dobivena je stolnim skenerom, a sadrži 1330 slika 133 osobe.

6. IZLUČIVANJE ZNAČAJKI DLANA I LICA

Slika 30 prikazuje vizualizaciju broja značajki koje LBLDA koristi iz pojedinih područja slike, ako je naučena na bazi lica. Dano je nekoliko slika koje odgovaraju različitom ukupnom broju značajki (N_{LBLDA}). Svjetliji dijelovi slike odgovaraju područjima s kojih je uzeto više značajki, dok crni dijelovi slike odgovaraju područjima s kojih nije uzeta niti jedna značajka. Ista vizualizacija dana je za bazu dlanova na slici 31.



Slika 30: (a) Srednja slika lica iz baze, (b) Vizualizacija broja značajki iz pojedinih područja. Svjetliji dijelovi slike odgovaraju područjima s kojih je uzeto više značajki, dok crni dijelovi slike odgovaraju područjima s kojih nije uzeta niti jedna značajka.



Slika 31: (a) Srednja slika dlana iz baze, (b) Vizualizacija broja značajki iz pojedinih područja. Svjetliji dijelovi slike odgovaraju područjima s kojih je uzeto više značajki, dok crni dijelovi slike odgovaraju područjima s kojih nije uzeta niti jedna značajka.

Iz slike 30 vidi se da se najveći broj značajki uzima iz područja očiju, obrva, nosa i usta što odgovara ljudskoj percepciji značajki lica. Iz slike 31 vidi se da se najveći dio značajki uzima iz gornjeg dijela dlana gdje se najčešće nalaze glavne linije dlana.

7. Ocjena kvalitete područja interesa dlana

Za razliku od biometrijske verifikacije uporabom statičnih slika, gdje se iz jedne snimke može dobiti samo jedan biometrijski predložak, u verifikaciji na temelju video sekvenci iz jedne video sekvence moguće je dobiti skup predložaka (maksimalno onoliko predložaka koliko ima okvira u video sekvenci). Verifikaciju je moguće provesti koristeći informaciju iz svih predložaka (svih okvira video sekvence), no ponekad može biti bolje koristiti samo pojedine okvire video sekvenci za izlučivanje značajki. Stoga je potrebna mjera kvalitete na temelju koje bi se odabrali okviri video sekvence iz kojih bi se mogle izlučiti značajke pogodne za biometrijsku verifikaciju. Uporabom mjere kvalitete moguće je odbaciti okvire video sekvenci koji, bilo zbog degradacija na slici (na primjer, zamućenje nastalo zbog gibanja), bilo zbog nepreciznosti postupaka lokalizacije, ne predstavljaju dobar izvor značajki. Druga prednost korištenja metoda ocjene kvalitete je povećanje brzine verifikacije jer se smanjuje broj usporedbi koje je potrebno načiniti prilikom verifikacije.

U sklopu ovoga rada razvijena je metoda za ocjenu kvalitete interesnog područja dlana u video sekvencama. Umjesto korištenja općenitih metoda za ocjenu kvalitete slike (npr. oštrina, kontrast), koje ne bi mogle razlučiti kvalitetnu sliku dlana od kvalitetne slike koja, pored dlana, sadrži i dio pozadine, metoda za ocjenu kvalitete prilagođena je upravo objektu čija se kvaliteta želi ocijeniti. Ideja metode ocjene kvalitete je načiniti parametarski model koji opisuje kvalitetan dlan. Za nepoznatu sliku potrebno je naći parametre modela kojima se, što je moguće točnije, rekonstruira nepoznata slika. Ukoliko je pogreška pri rekonstrukciji mala, rekonstrukcija je dobra i slika odgovara modelu kvalitetnog dlana, pa se stoga može pretpostaviti da ona i sadrži kvalitetan dlan. Ukoliko je rekonstrukcijska pogreška velika, slika se ne može dobro opisati modelom i postoji velika vjerojatnost da slika ne predstavlja kvalitetan dlan.

U razvijenoj metodi, model za kvalitetan dlan dobiven je uporabom analize glavnih komponenti na velikoj bazi (110 osoba, 5 slika po osobi) dlanova dobivenih pomoću stolnog skenera. Stolni skener daje vrlo kvalitetnu sliku dlana u kontroliranim svjetlosnim uvjetima te omogućuje vrlo precizno određivanje područja interesa. Stoga možemo reći da baza korištena za učenje modela sadrži samo kvalitetne dlanove.

Postupkom analize glavnih komponenti nad bazom kvalitetnih dlanova dobiva se srednja vrijednost, \mathbf{m} , i matrica \mathbf{W} čiji su stupci baze prostora koji daje PCA. Na temelju eksperimenata, dimenzionalnost PCA prostora postavljena je na 200, što znači da matrica \mathbf{W} korištena u radu ima 200 stupaca i 4096 redaka (slike su dimenzije 64 x 64 slikovnih

elemenata). Ovo također znači da model kvalitetnog područja interesa dlana ima 200 parametara.

Parametri modela koji najbolje rekonstruiraju nepoznatu sliku, prikazanu kao vektor stupac \mathbf{x} , mogu se dobiti projekcijom u PCA potprostor kao

$$\mathbf{y} = \mathbf{W}^T (\mathbf{x} - \mathbf{m}) \quad (55)$$

gdje je \mathbf{y} vektor stupac koji sadrži parametre modela. Uz dane parametre modela, rekonstrukciju je moguće izračunati kao

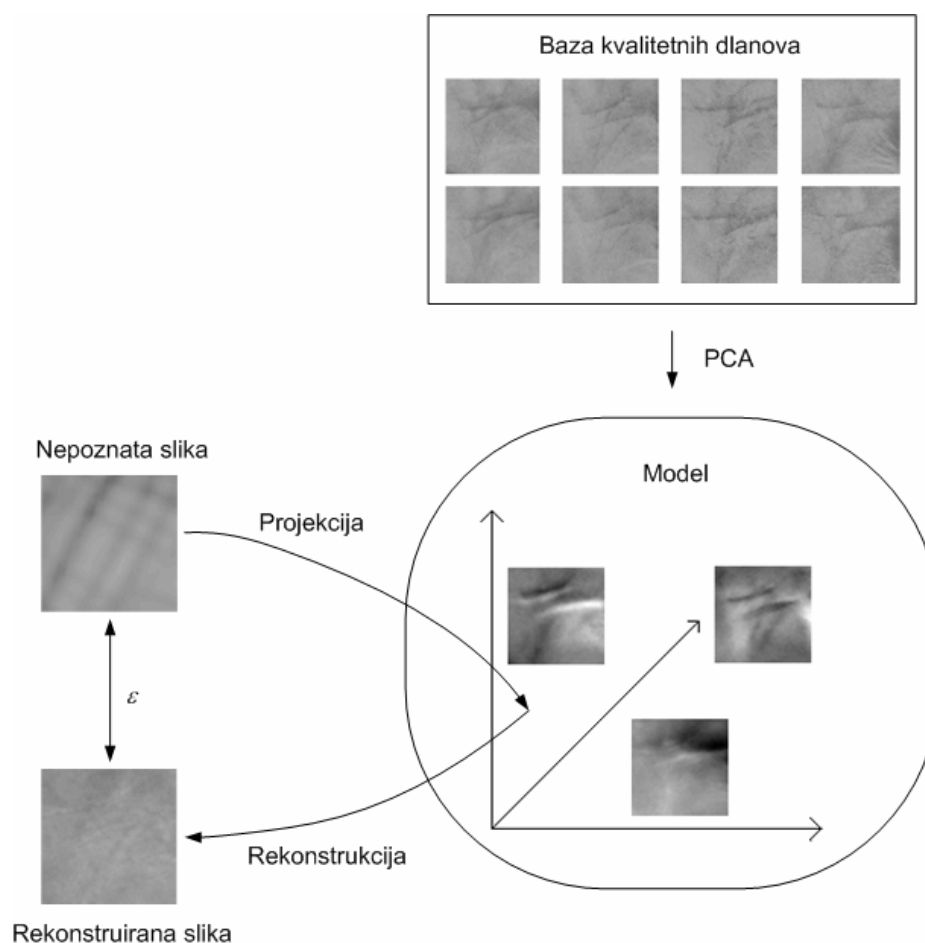
$$\mathbf{x}' = \mathbf{W}\mathbf{y} + \mathbf{m} \quad (56)$$

pri čemu je rekonstrukcijska pogreška ε jednaka.

$$\varepsilon = \|\mathbf{x} - \mathbf{x}'\| = \|\mathbf{x} - \mathbf{W}\mathbf{y} - \mathbf{m}\| = \|\mathbf{x} - \mathbf{m} - \mathbf{W}\mathbf{W}^T (\mathbf{x} - \mathbf{m})\| \quad (57)$$

Manja rekonstrukcijska pogreška znači bolje poklapanje s modelom i veću sličnost s kvalitetnim dlanovima, dok će veća rekonstrukcijska pogreška značiti veću različitost od kvalitetnih dlanova. Stoga se ε može koristiti kao mjera nekvalitete dlana, odnosno $-\varepsilon$ kao mjera kvalitete.






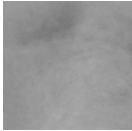

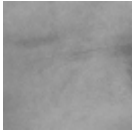
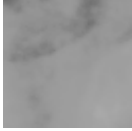

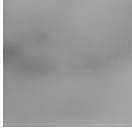

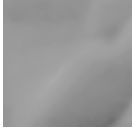

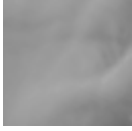

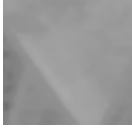

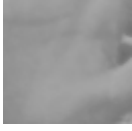
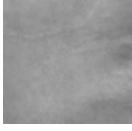
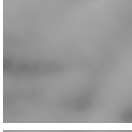

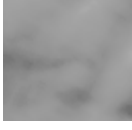

Slika 32 prikazuje postupak izgradnje modela kvalitetnih dlanova i ocjene kvalitete nepoznate slike. Prethodno je primjenom PCA na bazi kvalitetnih dlanova izgrađen model kvalitetnih dlanova. Postupak ocjene kvalitete za nepoznatu sliku je slijedeći: Nepoznata slika se projicira u PCA prostor (nalaze se optimalni parametri modela u smislu rekonstrukcije). Nakon toga se slika rekonstruira iz PCA prostora natrag u izvorni prostor slike (na temelju parametara modela). Ocjena nekvalitete dlana je pogreška u rekonstrukciji tj. Euklidska udaljenost izvorne i rekonstruirane slike (slike su predstavljene kao vektori).



Slika 32: Postupak izgradnje modela kvalitetnih dlanova i ocjene kvalitete nepoznate slike

Slika 33 prikazuje više slika područja interesa dlana dobivenih postupkom lokalizacije na bazi video sekvenci. Za svako područje interesa prikazana je slika koja se dobiva rekonstrukcijom gore opisanim postupkom, te je dana pogreška prilikom rekonstrukcije (mjera nekvalitete). Interesna područja sortirana su prema mjeri nekvalitete. Kao što se može vidjeti iz slike, interesna područja dlana koja sadrže samo dlan i na kojima su dobro vidljive linije dlana imaju manju rekonstrukcijsku pogrešku (tipično 230-260), dok interesna područja na kojima je dlan lokaliziran neprecizno ili uopće ne sadrže dlan imaju veću rekonstrukcijsku pogrešku (tipično 280-350). Većina rekonstrukcijskih pogrešaka nalazi se u relativno malom rasponu (230-350). Razlog tome je taj što su prilikom izgradnje PCA prostora eliminirane komponente s manjim svojstvenim vrijednostima, pa će čak i najkvalitetniji dlanovi imati neku rekonstrukcijsku pogrešku. Međutim, ukoliko se ne gledaju apsolutne vrijednosti, nego odnosi mjera kvalitete, može se primijetiti da će kvalitetniji dlanovi najčešće imati niže rekonstrukcijske pogreške od manje kvalitetnih, čak i ako uspoređujemo mjere kvalitete za dlanove različitih osoba.

7. OCJENA KVALITETE PODRUČJA INTERESA DLANA

		$\varepsilon = 236.88$
		$\varepsilon = 240.59$
		$\varepsilon = 241.42$
		$\varepsilon = 244.102$
		$\varepsilon = 249.03$
		$\varepsilon = 261.01$
		$\varepsilon = 273.58$
		$\varepsilon = 276.47$
		$\varepsilon = 305.93$
		$\varepsilon = 306.32$
		$\varepsilon = 331.42$
		$\varepsilon = 335.01$
a)	b)	c)

Slika 33: Nekoliko primjera ocjene kvalitete, a) izvorna slika interesnog područja, b) rekonstruirana slika interesnog područja, c) rekonstrukcijska pogreška

8. Fuzija u biometriji

Biometrijski sustavi koji koriste samo jednu biometrijsku značajku nazivaju se jednonačinski (engl. unimodal) sustavi. Za razliku od korištenja samo jedne biometrijske značajke, u višenačinskim (engl. multimodal) sustavima [2][61] moguće je kombinirati više različitih značajki (na primjer značajke lica i značajke dlana). Ovime se često može postići manja pogreška sustava, ali i bolja otpornost sustava na pokušaje prijevare jer bi napadač morao lažirati više različitih biometrijskih značajki. Pored toga, u sustavima koji koriste lice kao biometrijsku karakteristiku, moguće je pohraniti snimku korištenu za verifikaciju, što otvara mogućnost kasnijeg uvida i identificiranja napadača.

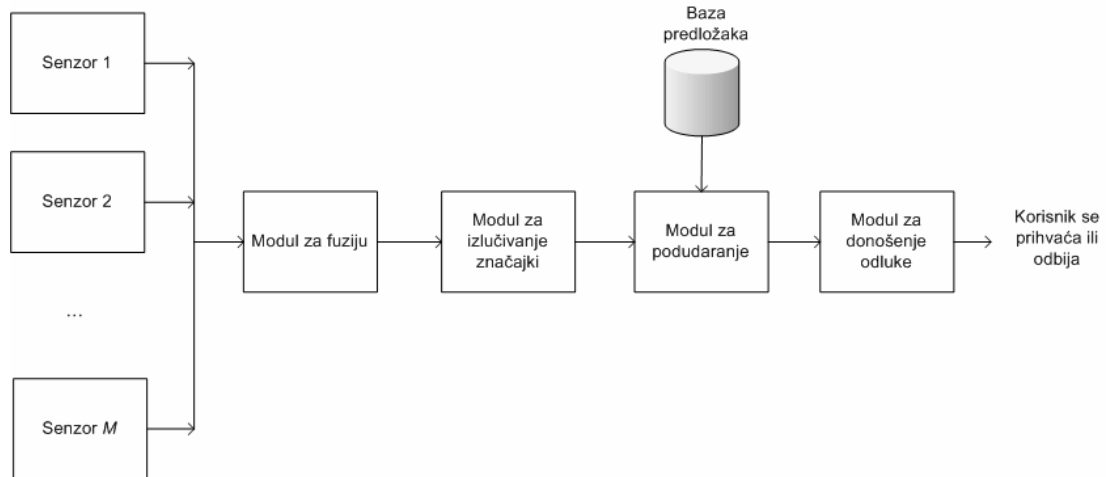
Kombiniranje više biometrijskih značajki, tj. fuziju u biometriji u višenačinskim biometrijskim sustavima moguće je načiniti na više različitih razina [61]:

- na razini senzora
- na razini vektora značajki
- na razini mjere podudaranja
- na razini odluke

8.1. Fuzija na razini senzora

U fuziji na razini senzora izlazi više različitih senzora kombiniraju se kako bi se dobili novi podaci koji će biti ulaz u modul za izlučivanje značajki. Na primjer, kombiniranjem izlaza koji daje video kamera i izlaza senzora koji daje 3D dubinsku informaciju moguće je dobiti 3D teksturiranu sliku objekta koji se snima.

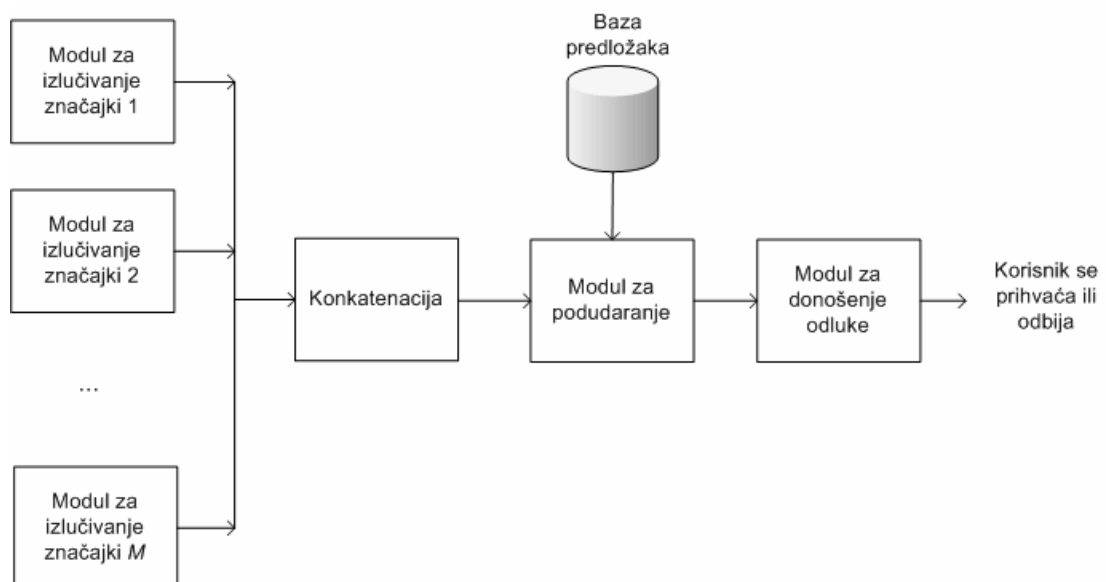
Model sustava koji koristi fuziju na razini senzora prikazan je na Slici 34.



Slika 34: Model biometrijskog sustava s fuzijom na razini senzora

8.2. Fuzija na razini vektora značajki

U fuziji na razini vektora značajki, potrebno je da svaki od modula za izlučivanje značajki daje izlaz u obliku vektora značajki. Individualni vektori značajki tada se konkatenuiraju u jedinstveni vektor značajki. Ovakvi vektori značajki tada se mogu podudarati s vektorima u bazi sustava koji su dobiveni na identičan način. Dio modela biometrijskog sustava koji koristi fuziju na razini vektora značajki prikazan je na Slici 35.



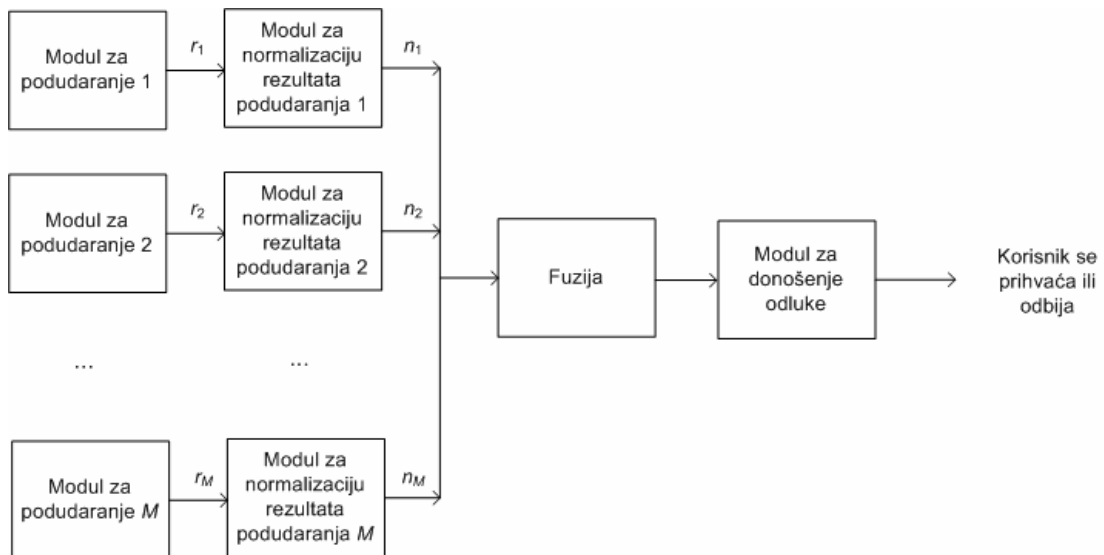
Slika 35: Model dijela biometrijskog sustava s fuzijom na razini vektora značajki

8.3. Fuzija na razini mjere podudaranja

U fuziji na razini mjere podudaranja mjere podudaranja individualnih biometrijskih značajki kombiniraju se u jedinstvenu mjeru podudaranja. Ovo je najčešće korištena razina fuzije zbog jednostavnosti i dobrih rezultata koji se pomoću nje mogu ostvariti [62].

Kittler i Alkoot [63] teorijski su pokazali da se fuzijom na razini mjere podudaranja dobivaju bolji rezultati od fuzije na razini odluke ako greške klasifikatora slijede Gaussovu razdiobu i različite značajke korištene u višenačinskom biometrijskom sustavu su međusobno statistički nezavisne. Taj rezultat su eksperimentalno pokazali na primjerima fuzije značajki lica i govora.

Slika 36 prikazuje model dijela biometrijskog sustava koji koristi fuziju na razini mjere podudaranja. U fuziji na razini mjere podudaranja svaki od modula za podudaranje producira mjeru sličnosti ili različitosti biometrijskog predložka trenutnog korisnika i biometrijskih predložaka spremljenih u bazi predložaka. Ove mjere označiti ćemo kao r_1, r_2, \dots, r_N gdje je N broj modula za podudaranje. Ove mjere je prije same fuzije potrebno normalizirati. O razlozima i tehnikama normalizacije više će biti rečeno kasnije. Normalizirane mjere sličnosti (normalizacijom se uvijek dobivaju mjere sličnosti) označiti ćemo kao n_1, n_2, \dots, n_N . Nakon normalizacije, u modulu za fuziju se, na temelju nekog od pravila fuzije, računa ukupna mjera podudaranja (engl. Total Similarity Measure, *TSM*).



Slika 36: Model dijela biometrijskog sustava s fuzijom na razini mjere podudaranja

Neka od pravila fuzije koja se koriste u fuziji na razini mjere podudaranje za dobivanje TSM su

- Suma

$$TSM = \sum_{i=1}^N n_i \quad (71)$$

- Težinska suma

$$TSM = \sum_{i=1}^N w_i n_i \quad (72)$$

Kod težinske sume uobičajeno je da je suma težinskih faktora w_i , $i = 1, 2, N$ jednaka jedinici. U ovoj vrsti fuzije problem predstavlja optimalno odrediti težinske faktore. Često se oni postavljaju tako da budu proporcionalni rezultatima raspoznavanja odgovarajućih jednonačinskih sustava ili obrnuto proporcionalni EER-u odgovarajućih jednonačinskih sustava.

- Minimum

$$TSM = \min(n_i), \quad i = 1, 2, \dots, N \quad (73)$$

- Maksimum

$$TSM = \max(n_i), \quad i = 1, 2, \dots, N \quad (74)$$

- Produkt

$$TSM = \prod_{i=1}^N n_i \quad (75)$$

Nakon fuzije, u modulu za donošenje odluke se na temelju TSM donosi odluka hoće li se trenutni korisnik prihvatiti ili odbiti.

U višenačinskom sustavu s fuzijom na razini mjere podudaranja najčešće nije moguće izravno kombinirati mjere koje daju pojedini moduli za podudaranje. Neki od razloga za ovo su [62]:

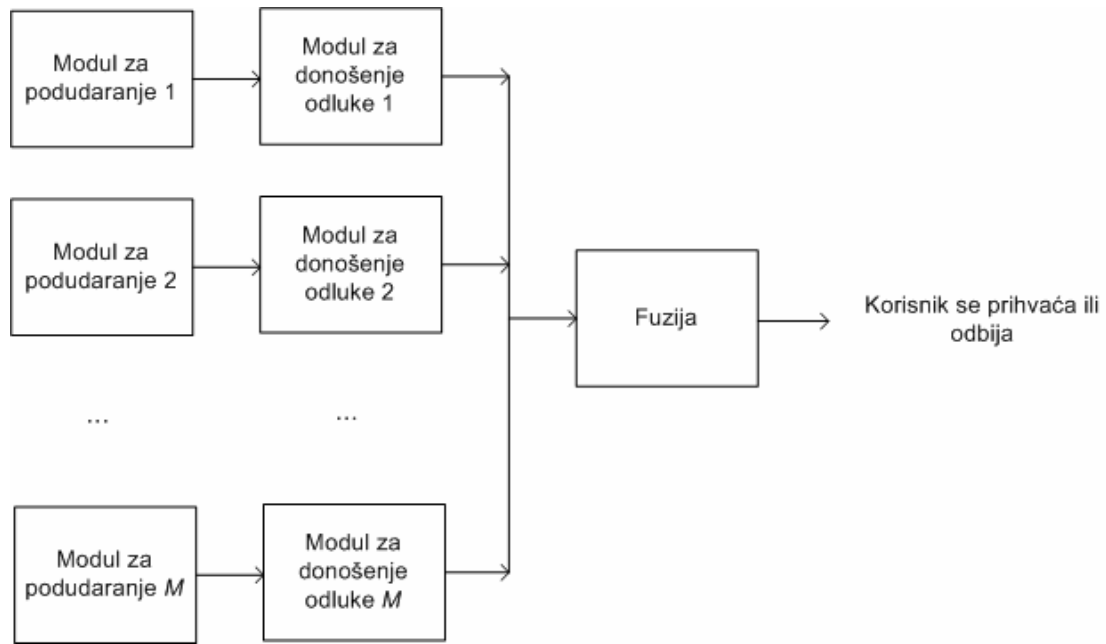
- (i) Mjere podudaranja za različite značajke mogu biti predstavljane na različit način. Neki moduli za podudaranje mogu davati mjere sličnosti dok drugi, unutar istoga sustava, mogu davati mjere različitosti. Normalizacija uvijek rezultira mjerama sličnosti.
- (ii) Izlazi modula za podudaranje mogu biti u različitim numeričkim opsezima. Većina tehnika normalizacije rezultira mjerama sličnosti u intervalu $[0, 1]$.
- (iii) Rezultati podudaranja uzoraka iste osobe i uzoraka različitih osoba za različite značajke neće uvijek pratiti iste statističke distribucije.

Ovo su razlozi zašto je prije fuzije potrebno normalizirati mjere podudaranja. Normalizaciju je moguće izvesti na temelju Bayesovog teorema te brojnim heurističkim tehnikama od kojih su neke min-max, z-score, median-MAD, dvostruka sigmoidalna funkcija, tangens hiperbolni [62], te po dijelu linearna funkcija normalizacije [64]. Parametre funkcija koje se koriste za normalizaciju u pravilu je potrebno naučiti na skupu za učenje.

U eksperimentalnom dijelu rada korištena je min-max normalizacija zbog malog broja parametara i jer u velikom broju slučajeva daje dobre rezultate [62][65]. Min-max je jednostavna linearna normalizacijska funkcija čiji se parametri odrede na način da funkcija za najbolje podudaranje na skupu za učenje daje vrijednost 1, a za najlošije podudaranje na skupu za učenje vrijednost 0.

8.4. Fuzija na razini odluke

U fuziji na razini odluke, odluke donesene na temelju individualnih biometrijskih karakteristika kombiniraju se u jedinstvenu odluku o prihvatanju ili odbijanju korisnika sustava. Dio modela biometrijskog sustava koji koristi fuziju na razini odluke prikazan je na slici 37.



Slika 37: Dio modela biometrijskog sustava s fuzijom na razini odluke

Neki od načina donošenja konačne odluke u sustavima s fuzijom na temelju odluke su

- Glasanje. Donosi se ona odluka za koju glasa najveći broj modula za donošenje odluke.
- Težinsko glasanje. Odluka se donosi slično kao kod glasanja, samo što je svakom modulu za donošenje odluke dodijeljena težina koja govori koliko njegov glas vrijedi. Težine se mogu podesiti tako da za svaki modul budu proporcionalne vjerojatnosti da taj modul daje ispravnu odluku.
- Fuzija na temelju klasifikatora. U ovom slučaju odluku donosi klasifikator koji vektor značajki koji se sastoji od individualnih mjera podudaranja klasificira u razred klijenata ili uljeza. Fuzija na temelju klasifikatora nije fuzija na razini odluke u strogom smislu, jer se u fuziji na razine odluke individualne odluke kombiniraju u jedinstvenu odluku, dok se ovdje individualne mjere podudaranja kombiniraju u jedinstvenu odluku. Ipak, u većini radova gdje je ovakva fuzija korištena autori je navode kao fuziju na razini odluke [66][67].
- Stabla odluke. Na temelju rezultata na nekom skupu za učenje moguće je izgraditi stablo odluke koje će se koristiti u radu sustava.

9. Eksperimentalna evaluacija

9.1. Baza video sekvenci

Za potrebe eksperimenata prikupljena je baza video sekvenci dlana i lica. Baza sadrži video sekvence 87 osoba. Za svaku osobu snimljene su dvije video sekvence s vremenskim razmakom između snimanja tipično od tjedan do mjesec dana. Osobe u bazi su za vrijeme snimanja bile djelatnici i studenti Fakulteta Elektrotehnike i Računarstva. Tipična dob osoba u bazi je između 20 i 40 godina, a od 87 osoba ukupno u bazi, 11 su žene a 76 muškarci. Tipična duljina video sekvenci je između 3 i 5 sekundi.

Baza je prikupljena videokamerom u boji rezolucije 640 x 480 slikovnih elemenata. Videokamera snima video sekvence brzinom od trideset okvira u sekundi. Sve video sekvence su pohranjene u nekomprimiranom obliku kako bi se izbjegle negativne posljedice kompresije na točnost verifikacije. Iza kamere za vrijeme snimanja bio je postavljen izvor osvjtljenja čije su uloge bile

- (i) Osigurati da su ruka i lice dovoljno osvijetljeni kako bi se mogle dobiti kvalitetne snimke s kratkim trajanjem ekspozicije. Veće vrijeme ekspozicije dovodi do pojave zamućenja pokretnih objekata u video sekvenci, što može negativno utjecati na kvalitetu značajki dlana i lica
- (ii) Osigurati da na sceni postoji barem jedan izvor svjetla usmjeren prema dlanu kako bi bile vidljive linije dlana

Baza je prikupljena na slijedeći način: Osobama su na snimanju dane upute da sjednu ispred kamere, tijekom snimanja gledaju u kameru i lagano mahnu rukom s površinom dlana okrenutom prema kameri. Osobama s dugim rukavima koji su im prekrivali površinu dlana rečeno je da rukav povuku kako bi bila vidljiva cijela površina dlana. Pored ovoga nisu bile davane nikakve daljnje upute. Osobe koje su nosile naočale ili prstenje zadržale bi ih i za vrijeme snimanja. Snimanje je rađeno u složenim uvjetima u pogledu pozadine i vanjskog osvjtljenja. Pozadina sadrži velik broj objekata, moguće su druge osobe u pokretu u pozadini te drugi objekti boje kože pored lica i ruke osobe koja se snima. Također su primjetni različiti uvjeti osvjtljenja u različitim snimkama iste osobe (premda uređaj za snimanje sadrži izvor svjetla, utjecaj drugih izvora svjetla kao što je sunčeva svjetlost koja dolazi kroz prozore bio je promjenjiv).

Slika 38 prikazuje primjere iz nekoliko video sekvenci ove baze. Kao što se može vidjeti iz slike, u bazi postoje varijacije u izgledu iste osobe nastale kao posljedica promjene frizure, izraza lica, prisustva i odsustva naočala itd.



Slika 38: Primjeri okvira video sekvenci iz baze. Okviri u istom retku pripadaju snimkama iste osobe. Okviri u lijevom stupcu uzeti su iz prvog, a okviri u desnom stupcu iz drugog snimanja iste osobe

9.2. Baze za određivanje parametara

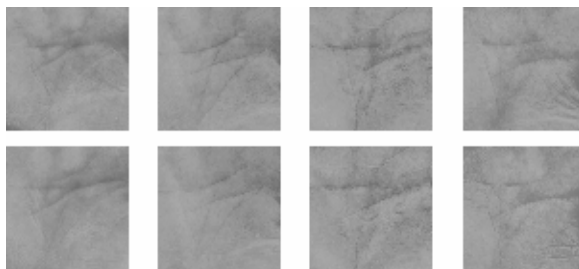
Parametri metoda koje će biti korištene u eksperimentima verifikacije dlana i lica u video sekvencama utvrđeni su na temelju preliminarnih eksperimenata raspoznavanja provedenih na bazama statičnih slika lica i dlana.

Kao baza lica korištena je XM2VTS baza statičnih slika lica [60]. Ova baza sadrži 2360 slika 295 osoba (8 slika po osobi). Slike su snimljene u četiri vremenski odvojena snimanja, a u svakom snimanju dobivene su dvije slike. Prije eksperimenata raspoznavanja, ove su slike bile normalizirane postupkom opisanom u poglavlju 5. Četiri slike svake osobe (slike iz prva dva snimanja) bile su korištene za učenje, a preostale slike su bile korištene za raspoznavanje. Slika 39 prikazuje nekoliko normaliziranih lica iz XM2VTS baze.



Slika 39: Nekoliko normaliziranih lica iz XM2VTS baze. Slike u istom stupcu pripadaju istoj osobi.

Kao baza dlanova korištena je baza 1330 slika 133 osobe (10 slika po osobi) snimljena stolnim skenerom. Prije raspoznavanja lokalizirano je i normalizirano područje interesa dlana postupkom opisanom u poglavlju 5. Baza je podijeljena tako da su tri slike svake osobe korištene za učenje, a preostalih 7 za ispitivanje. Baza je podijeljena na takav način da je relativno mali broj slika u odnosu na raspoloživi broj slika po osobi korišten za učenje da bi se naglasile razlike u rezultatima za različite parametre. Ukoliko uzmemo veći broj slika za učenje, sa svim vrijednostima parametara dobiti ćemo bolje rezultate raspoznavanja, što će u slučaju korištene baze dlanova značiti da će mnogi rezultati postati međusobno vrlo slični i biti će teško izvesti zaključke o dobroti nekih vrijednosti parametara. Slika 40 prikazuje nekoliko normaliziranih dlanova iz korištene baze.



Slika 40: Nekoliko normaliziranih dlanova iz baze statičnih slika dlanova. Slike u istom stupcu pripadaju istoj osobi.

9.3. Evaluacija metoda izlučivanja značajki

Evaluacija metoda izlučivanja značajki provedena je na slijedeći način. Prva video sekvenca za svaku osobu korištena je za učenje i generiranje predložaka u bazi (skup za učenje), a druga video sekvenca za ispitivanje. Nakon postupka učenja izlučene su značajke iz svih okvira video sekvenci iz skupa za učenje, te su odgovarajući vektori značajki spremljeni u bazu predložaka. Svaka video sekvenca iz skupa za ispitivanje uspoređena je sa svakom video sekvencom iz skupa za učenje na slijedeći način.

Za svaki okvir i video sekvence za ispitivanje na kojoj je nađena ruka ili lice dobiva se odgovarajući vektor značajki \mathbf{v}_i . \mathbf{v}_i se uspoređuje sa svim vektorima značajki dobivenih iz video sekvence iz skupa za učenje i bilježi se najmanja udaljenost d_i . Nakon što je ovo učinjeno za sve okvire u video sekvenci za ispitivanje, konačna mjera podudaranja dobiva se kombinirajući udaljenosti d_i . U eksperimentima su neovisno razmatrana tri načina kombiniranja udaljenosti pojedinih okvira:

- (i) MIN – uzima se najmanja udaljenost d_i kao konačna mjera podudaranja
- (ii) MEAN – kao konačna mjera podudaranja video sekvence uzima se srednja vrijednost mjera podudaranja pojedinih okvira
- (iii) MEDIAN – kao konačna mjera podudaranja video sekvence uzima se medijan mjera podudaranja pojedinih okvira

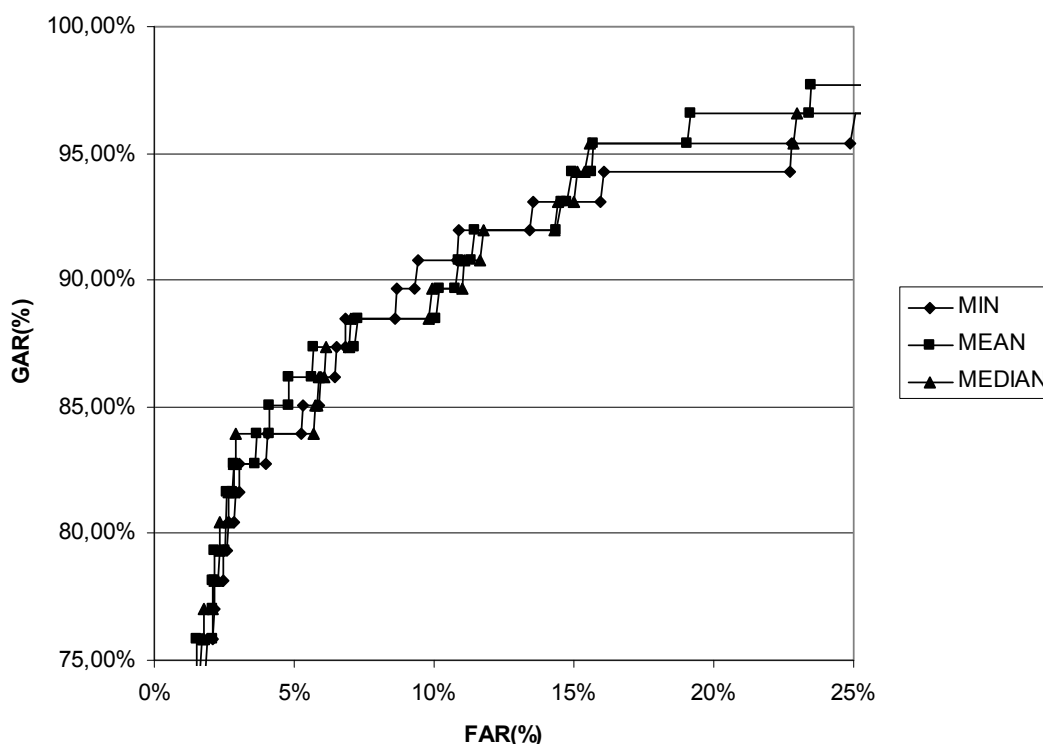
Konačna odluka o verifikaciji dobiva se usporedbom konačne mjere podudaranja s pragom t . Variranjem vrijednosti praga dobiva se ROC krivulja.

Ovakav eksperiment sastoji se od 87 poklapanja video sekvenci istih osoba i 7482 poklapanja video sekvenci različitih osoba.

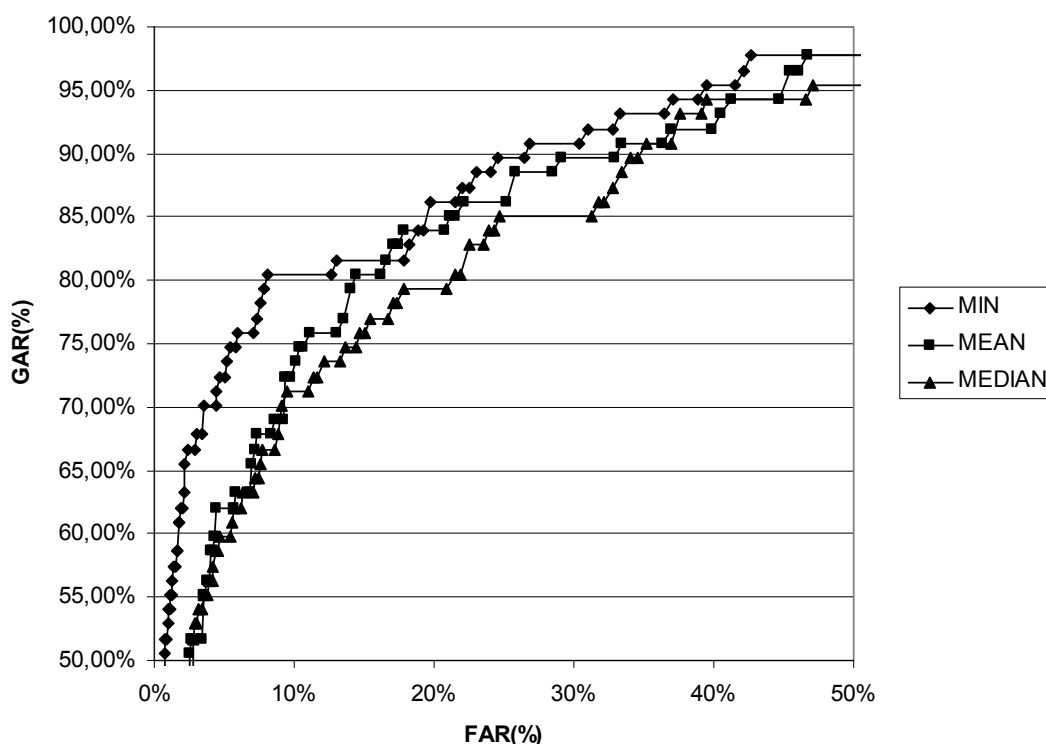
Parametri metoda izlučivanja značajki u svim metodama dobiveni su na temelju eksperimenata raspoznavanja na bazama statičnih slika lica i ruke.

9.3.1. Analiza glavnih komponenti

ROC krivulje koje odgovaraju eksperimentu verifikacije sa značajkama dobivenima pomoću analize glavnih komponenti mogu se vidjeti na slici 41 za lice i slici 42 za dlan. Za lice je korišteno 140 PCA komponenti (140-dimenzionalni vektori značajki), a za dlan je korišteno 100 PCA komponenti, u skladu s rezultatima na bazama za određivanje parametara. Kao što se može vidjeti na slikama, za ove značajke nije bilo velike razlike u rezultatima za različite načine kombiniranja mjera podudaranja pojedinih okvira. Vidljivo je da su rezultati verifikacije na temelju lica značajno bolji od rezultata verifikacije na temelju dlana. Za značajke lica ostvaren je EER = 9.41% i HmTER = 9.18%. Za značajke dlana ostvaren je EER = 17.24% i HmTER = 13.83%.



Slika 41: ROC krivulje za verifikaciju lica na temelju značajki dobivenih analizom glavnih komponenti



Slika 42: ROC krivulje za verifikaciju dlana na temelju značajki dobivenih analizom glavnih komponenti

9.3.2. Linearna diskriminantna analiza

Za značajke dobivene linearnom diskriminantnom analizom razmatrane su dvije mjere podudaranja: Euklidska udaljenost i normalizirana korelacija (kosinus kuta među vektorima značajki). Normalizirana korelacija razmatrana je kao mjera podudaranja jer je u literaturi [68] eksperimentalno pokazano da se uporabom ove mjere podudaranja s LDA značajkama dobivaju bolji rezultati verifikacije.

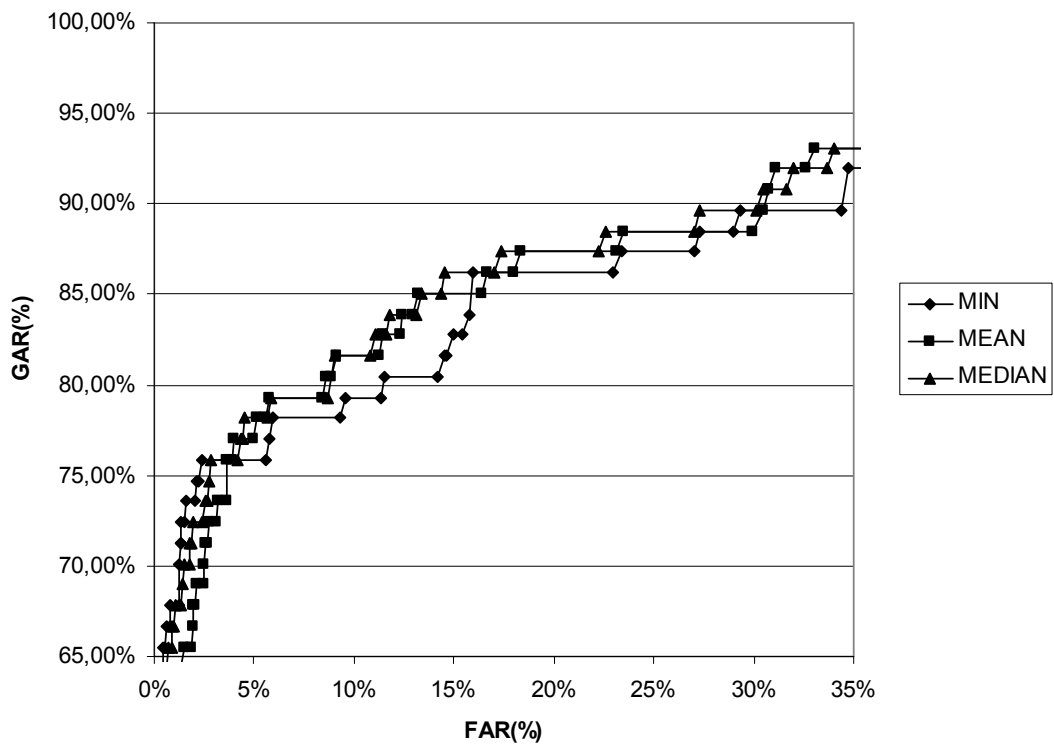
Slike 43 i 44 prikazuju ROC krivulje za verifikaciju lica na temelju Euklidske udaljenosti i normalizirane korelacije. Slike 45 i 46 prikazuju ROC krivulje za verifikaciju dlana na temelju Euklidske udaljenosti i normalizirane korelacije.

Parametri metoda (duljina vektora značajki nakon PCA, N_{PCA} , i duljina vektora značajki nakon LDA, N_{LDA}) korišteni u eksperimentu mogu se vidjeti u tablici 1. Parametri su dobiveni na temelju eksperimenata na bazama za određivanje parametara. U tablici su također navedene vrijednosti EER i HmTER dobivene u eksperimentima.

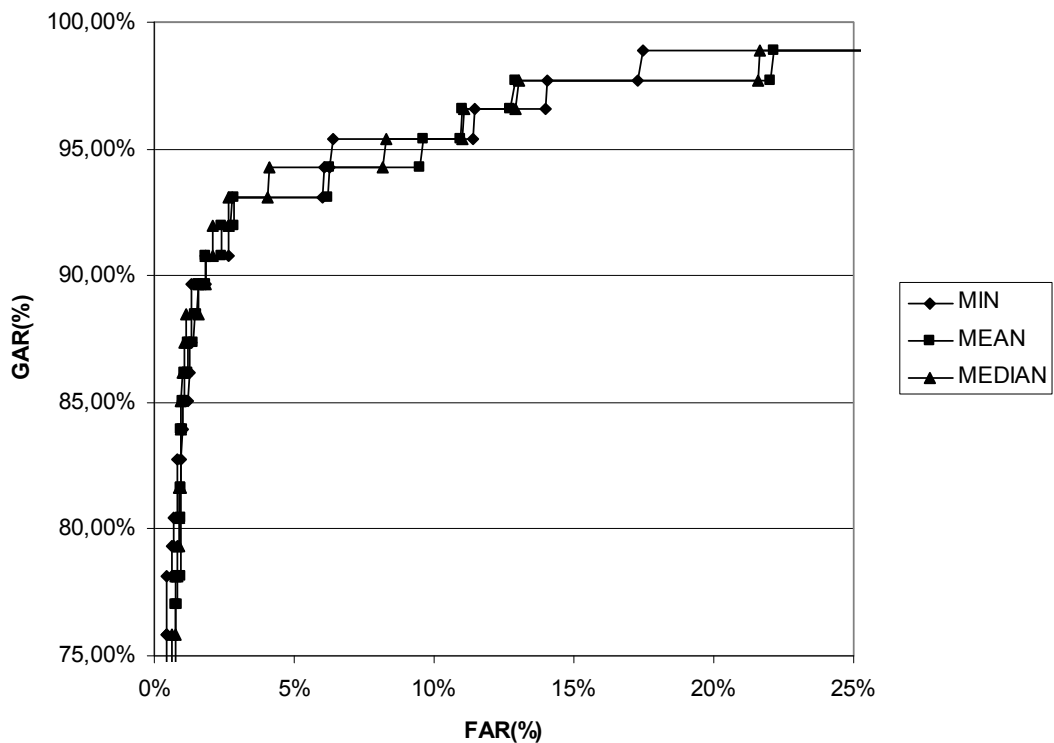
Značajka	Mjera podudaranja	N_{PCA}	N_{LDA}	EER	HmTER
Lice	Euklidska udaljenost	200	40	14.51%	13.16%
	normalizirana korelacija	300	80	5.75%	4.79%
Dlan	Euklidska udaljenost	100	60	15.09%	14.17%
	normalizirana korelacija	150	70	6.51%	5.45%

Tablica 1: Parametri i rezultati (EER i HmTER) za verifikaciju lica i dlana na temelju linearne diskriminantne analize

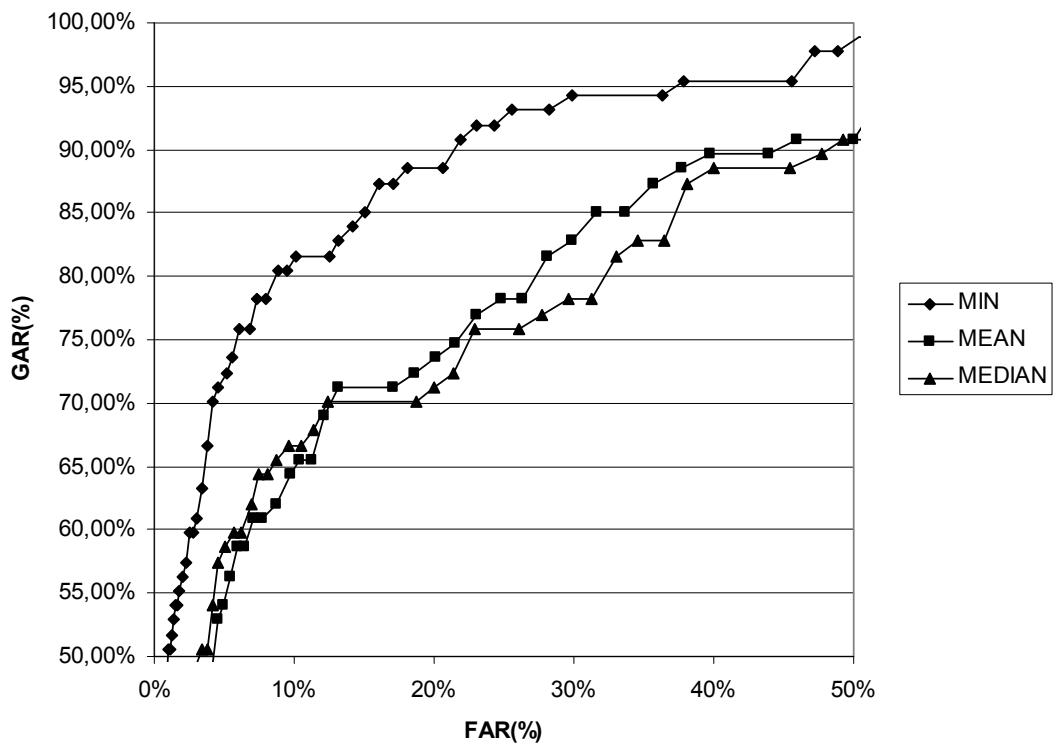
Vidljivo je da su pogreške verifikacije dobivene pomoću normalizirane korelacije znatno manje od pogrešaka uz korištenje Euklidske udaljenosti (npr. EER za dlan pada s 15.09% na 6.51%) što je u skladu s rezultatima prijavljenima u literaturi [68]. Također se može vidjeti da je uz uporabu ovih značajki razlika pogrešaka verifikacije između dlana i lica vrlo mala, što znači da se samo na temelju dlana snimljenim video kamerom može ostvariti približno ista točnost verifikacije kao i na temelju lica. Uzevši u obzir lošu kvalitetu video sekvenci i uvjete snimanje, točnost verifikacije na temelju dlana uz $EER = 6.51\%$ i $HmTER = 5.45\%$ je iznenađujuće dobar rezultat.



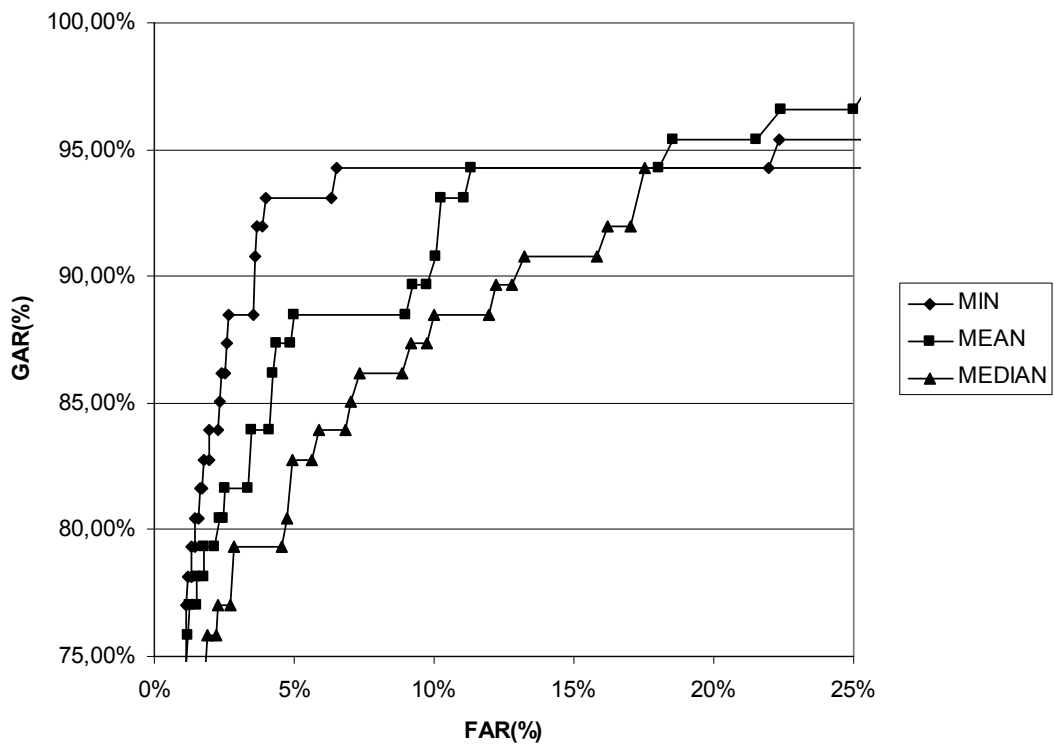
Slika 43: ROC krivulje za verifikaciju lica na temelju LDA značajki i Euklidske udaljenosti



Slika 44: ROC krivulje za verifikaciju lica na temelju LDA značajki i normalizirane korelacije



Slika 45: ROC krivulje za verifikaciju dlana na temelju LDA značajki i Euklidske udaljenosti



Slika 46: ROC krivulje za verifikaciju dlana na temelju LDA značajki i normalizirane korelacije

9.3.3. Gaborov filter

Gaborov filter korišten je kako bi se dobio binarni vektor značajki za reprezentaciju dlana i lica (odzivi realnog i imaginarnog dijela Gaborovog filtra binariziraju se na način da se kodira samo predznak odziva). Korišten je jedan Gaborov filter za dlan i jedan Gaborov filter za lice, a parametri ovih dvaju filtera eksperimentalno su određeni na temelju eksperimenata na bazama statičnih slika dlana i lica. U eksperimentima je posebno evaluirana usporedba vektora značajki bez njihove translacije prilikom podudaranja, te usporedba vektora značajki koja uključuje translaciju jednog od interesnih područja za ± 2 po x i ± 2 po y osi, pri čemu se kao rezultat podudaranja uzima najbolje podudaranje za bilo koju translaciju (vidi poglavlje 6.3.).

Slike 47 i 48 prikazuju ROC krivulje za verifikaciju lica na temelju Gaborovog filtra bez i sa translacijom ROI. Slike 49 i 50 prikazuju ROC krivulje za verifikaciju dlana na temelju Gaborovog filtra bez i sa translacijom ROI.

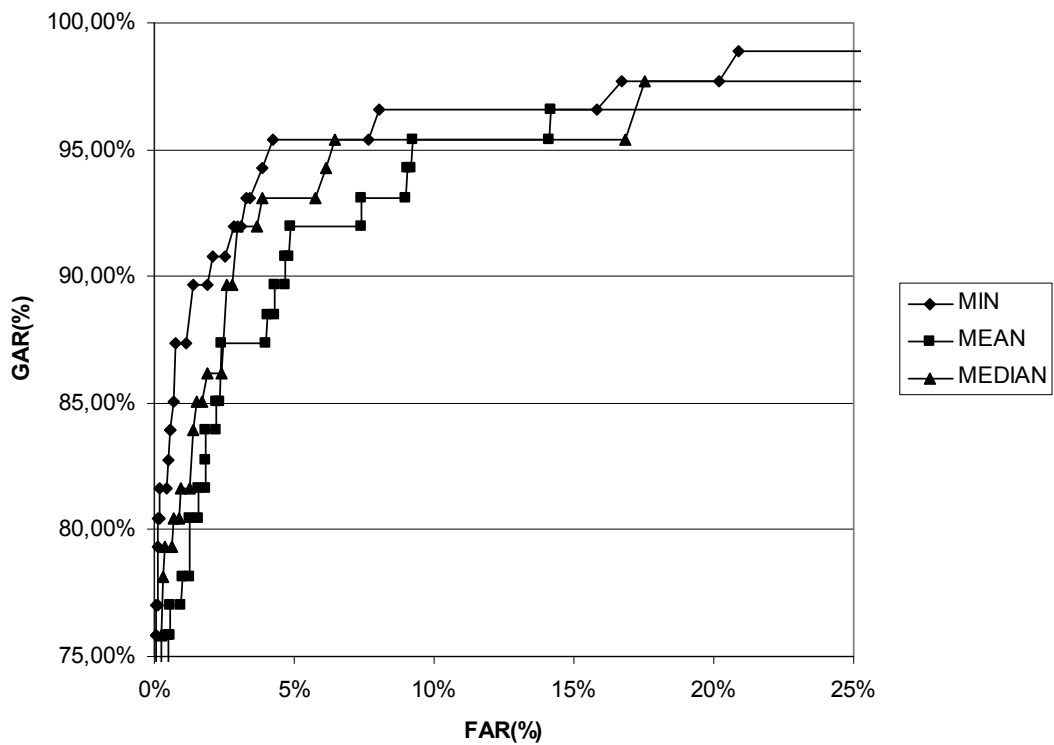
Parametri Gaborovih filtera korištenih u eksperimentu mogu se vidjeti u tablici 2. U tablici su također navedene vrijednosti EER i HmTER dobivene u eksperimentima.

Značajke	Podudaranje	σ	θ	u	EER	HmTER
Lice	bez translacije ROI	1	$3\pi/4$	0.083	4.60%	4.43%
	s translacijom ROI	1	$3\pi/4$	0.083	5.52%	4.78%
Dlan	bez translacije ROI	1	0	0.127	9.20%	8.30%
	s translacijom ROI	1	0	0.127	8.05%	6.92%

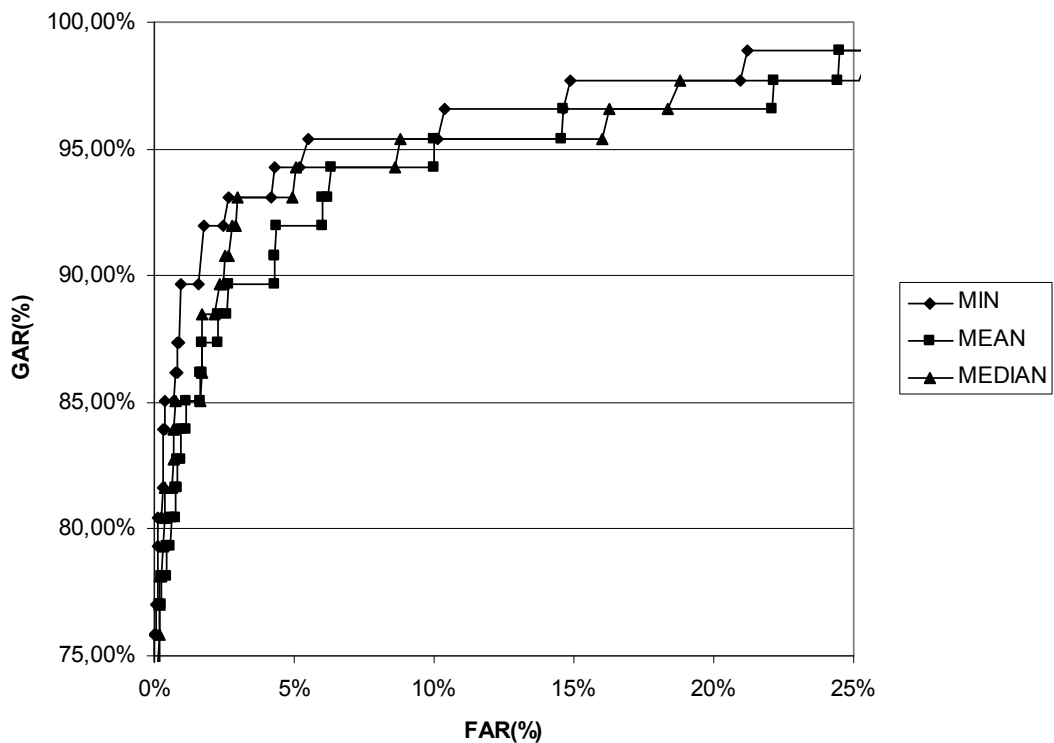
Tablica 2: Parametri i rezultati (EER i HmTER) za verifikaciju lica i dlana na temelju Gaborovog filtra

Kao što se može vidjeti iz slika i tablice, uporabom Gaborovog filtra mogu se ostvariti dobri rezultati verifikacije. Na primjer, za verifikaciju na temelju lica dobivaju se nešto bolji rezultati nego s LDA i normaliziranom korelacijom. Razlog tome je što se Gaborovim filtrom izlučuju lokalne značajke koje će biti manje osjetljive na promjene osvjetljenja i druge lokalne promjene nego što će to biti značajke dobivene pomoću LDA.

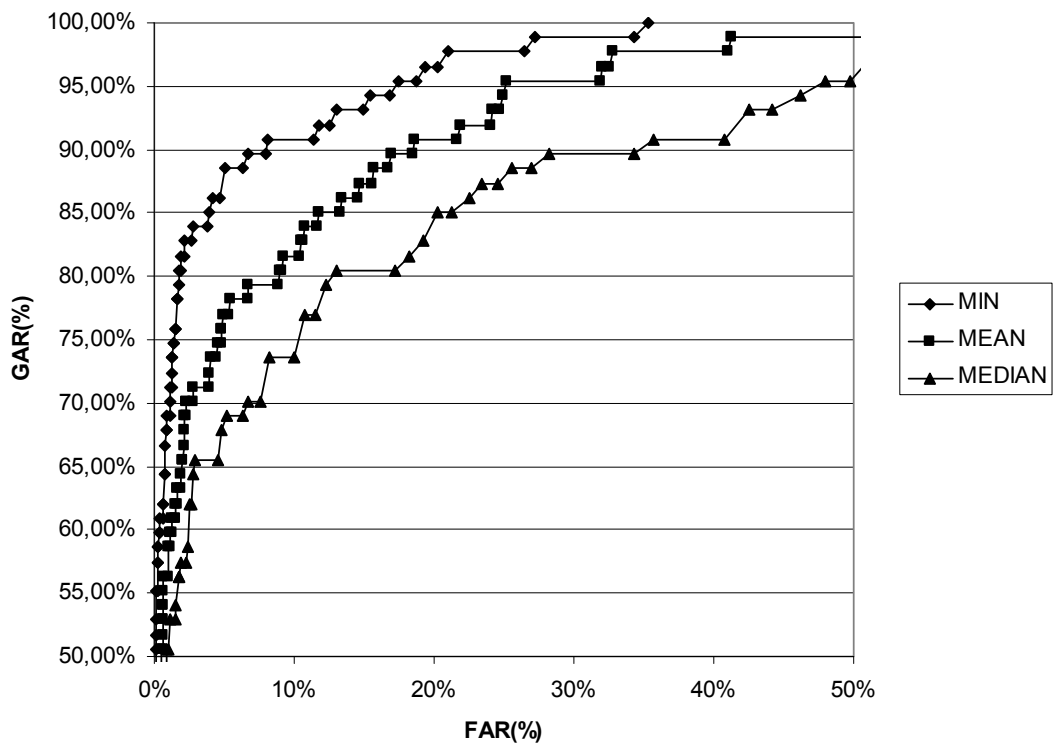
Može se vidjeti da se uz podudaranje s translacijom ROI pogreška verifikacije za dlan poboljšava, dok za lice to nije slučaj. Razlog za ovo se može tražiti u preciznijoj lokalizaciji lica (na temelju očiju), dok će kod dlana položaj detektiranoga područja interesa više varirati. Stoga će mehanizmi koji bi učinili podudaranje invarijantnim na translaciju biti korisni u verifikaciji dlana, ali ne nužno i u verifikaciji lica.



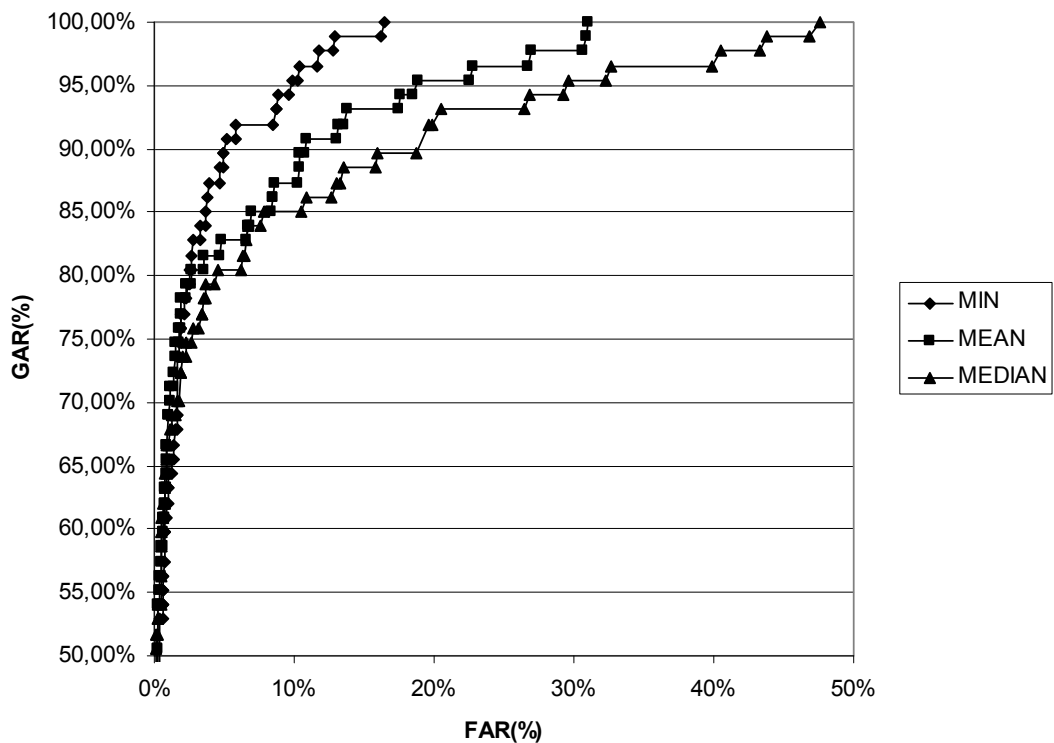
Slika 47: ROC krivulje za verifikaciju lica na temelju Gaborovog filtra bez translacije ROI



Slika 48: ROC krivulje za verifikaciju lica na temelju Gaborovog filtra s translacijom ROI



Slika 49: ROC krivulje za verifikaciju dlana na temelju Gaborovog filtra bez translacije ROI



Slika 50: ROC krivulje za verifikaciju dlana na temelju Gaborovog filtra s translacijom ROI

9.3.4. Lokalna binarna LDA

Parametri LBLDA korišteni u eksperimentima verifikacije dobiveni su na bazama statičnih slika dlana i lica. Lokalna binarna LDA ima tri parametra:

- (i) p – određuje širinu i visinu prozora u slikovnim elementima
- (ii) q – određuje za koliko se slikovnih elemenata prozor translacija u svakom koraku kako bi se formiralo slijedeće područje
- (iii) N_{PCA} – određuje dimenzionalnost na koju se reduciraju lokalni uzorci prije primjene LDA. Ako je $N_{PCA} = p \times p$, ne dolazi do redukcije dimenzionalnosti

Na bazama za utvrđivanje parametara načinjena je serija eksperimenata raspoznavanja (identifikacije u zatvorenome skupu) za različite kombinacije ovih parametara. Za svaku kombinaciju parametara također je tražena optimalna duljina vektora značajki (N_{LBLDA}) i zabilježen je najbolji ostvareni rezultat raspoznavanja zajedno s odgovarajućom duljinom vektora značajki. Eksperimenti su provedeni koristeći 1-NN klasifikator s Hammingovom udaljenosti.

Rezultati eksperimenata prikazani su u tablici 3 za lice i tablici 4 za dlan. Na temelju tablica može se izvući nekoliko zaključaka. Prvo, rezultati raspoznavanja su bolji ako se područja međusobno prekrivaju nego ako se područja ne prekrivaju. Ako se q smanjuje na $p/2$ ili $p/4$ točnost raspoznavanja se povećava, budući da se dobiva više diskriminatorskih značajki. Međutim, također se povećava i duljina vektora značajki. U nekim slučajevima smanjenjem q na $p/8$ mogu se ostvariti još bolji rezultati (na primjer za lice kada je $p=16$), ali to vodi drastičnom povećanju duljine vektora značajki (u primjeru s 1500 na 7300).

Drugo, u većini slučajeva najbolji rezultati ostvareni su uz $N_{PCA} = 100$ ili $N_{PCA} = 150$. Daljnje povećanje parametra N_{PCA} obično vodi smanjenju točnosti raspoznavanja. Ovo možemo interpretirati na slijedeći način. U slučaju kada joj se da mnogo stupnjeva slobode, LDA, isto kao i druge metode nadgledanoga učenja, daje bolje rezultate na skupu za učenje nego na skupu za ispitivanje. Često je bolje ograničiti dimenzionalnost podataka koji su ulaz u LDA kako bi se ostvarila bolja generalizacija.

Treće, optimalna veličina prozora je $p = 16$ za bazu lica i $p = 8$ ili $p = 16$ za bazu dlanova. Zanimljivo je primijetiti da se s vrijednostima parametara $p = 16$ i $q = 4$ dobivaju najbolji rezultati za obje baze (bez dramatičnog povećanja vektora značajki). Iako se ne može općenito tvrditi da će ovo biti ujedno najbolje vrijednosti parametara za ostale baze, ove vrijednosti daju dobru procjenu optimalnih parametara za korištenje u eksperimentima na bazi video sekvenci.

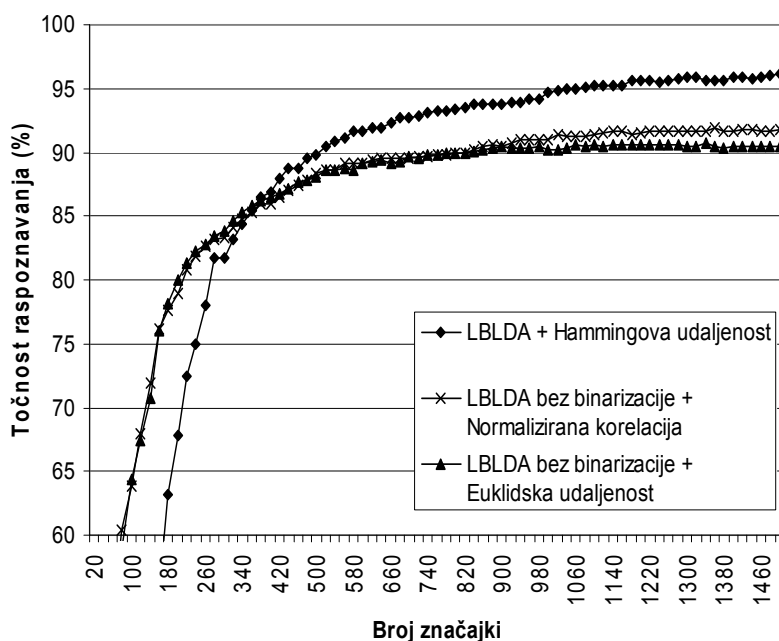
Veličina prozora p	Korak translacije prozora q	Broj PCA komponenti N_{PCA}	N_{LBLDA} za najbolju točnost raspoznavanja	Najveća točnost raspoznavanja
8	8	64	400	91.44%
8	4	64	1500	94.32%
8	2	64	4000	95.17%
16	16	50	300	91.10%
16	16	100	300	91.53%
16	16	150	400	90.25%
16	16	200	500	86.95%
16	16	256	200	86.19%
16	8	50	700	94.92%
16	8	100	1000	95.25%
16	8	150	700	94.83%
16	8	200	900	94.49%
16	8	256	1200	94.32%
16	4	50	1800	95.51%
16	4	100	1500	96.02%
16	4	150	1500	96.19%
16	4	200	1800	95.51%
16	4	256	1700	95.51%
16	2	50	5100	95.76%
16	2	100	7300	96.44%
16	2	150	7600	96.44%
16	2	200	9200	96.27%
16	2	256	9100	96.19%
32	32	100	200	88.56%
32	32	200	200	87.12%
32	32	400	300	80.93%
32	32	800	400	69.15%
32	16	100	400	93.64%
32	16	200	400	93.98%
32	16	400	600	92.03%
32	16	800	900	82.12%
32	8	100	800	94.58%
32	8	200	800	95.34%
32	8	400	700	93.73%
32	8	800	1600	86.10%

Tablica 3: Rezultati raspoznavanja lica na XM2VTS bazi za različite kombinacije LBLDA parametara

Veličina prozora p	Korak translacije prozora q	Broj PCA komponenti N_{PCA}	N_{LBLDA} za najveću točnost raspoznavanja	Najveća točnost raspoznavanja
8	8	64	500	98.81%
8	4	64	1900	99.68%
8	2	64	5800	99.78%
16	16	50	700	99.24%
16	16	100	400	99.03%
16	16	150	700	97.84%
16	16	200	500	96.43%
16	16	256	700	87.68%
16	8	50	2100	99.68%
16	8	100	400	99.57%
16	8	150	2000	99.14%
16	8	200	700	98.49%
16	8	256	3000	95.46%
16	4	50	1900	99.68%
16	4	100	1300	99.78%
16	4	150	1600	99.57%
16	4	200	3700	99.14%
16	4	256	2900	97.30%
16	2	50	5400	99.57%
16	2	100	8700	99.68%
16	2	150	5500	99.57%
16	2	200	7500	99.03%
16	2	256	7700	97.62%
32	16	100	400	99.57%
32	16	150	500	99.35%
32	16	200	400	99.14%
32	16	250	1500	98.59%
32	32	100	300	99.03%
32	32	150	300	98.81%
32	32	200	700	98.27%
32	32	250	400	97.19%
32	8	100	1300	99.68%
32	8	150	1500	99.24%
32	8	200	1000	99.35%
32	8	250	800	98.92%

Tablica 4: Rezultati raspoznavanja dlana na bazi statičnih slika dlana za različite kombinacije LBLDA parametara

Pored eksperimenta u kojem su određeni parametri LBLDA, na XM2VTS bazi također je načinjen eksperiment koji pokazuje učinak binarizacije i korištenja različitih mjera podudaranja na rezultate koji se postižu pomoću LBLDA značajki. Slika 51 prikazuje točnost raspoznavanja za LBLDA (uz Hammingovu udaljenost kao mjeru podudaranja), te LBLDA bez binarizacije uz Euklidsku udaljenosti i normaliziranu korelaciju kao mjere podudaranja. Iz slike se može vidjeti da binarizacija značajki poboljšava rezultate raspoznavanja u odnosu na značajke u obliku točnih vrijednosti, kada se kao mjera podudaranja koriste Euklidska udaljenost ili normalizirana korelacija. Uz normaliziranu korelaciju dobiva se nešto bolja točnost raspoznavanja nego uz Euklidsku udaljenost, slično kao kod „klasične“ LDA.



Slika 51: Usporedba točnosti raspoznavanja koja se dobiva pomoću LBLDA (s Hammingovom udaljenosti), te pomoću LBLDA bez binarizacije s Euklidskom udaljenosti i normaliziranom korelacijom.

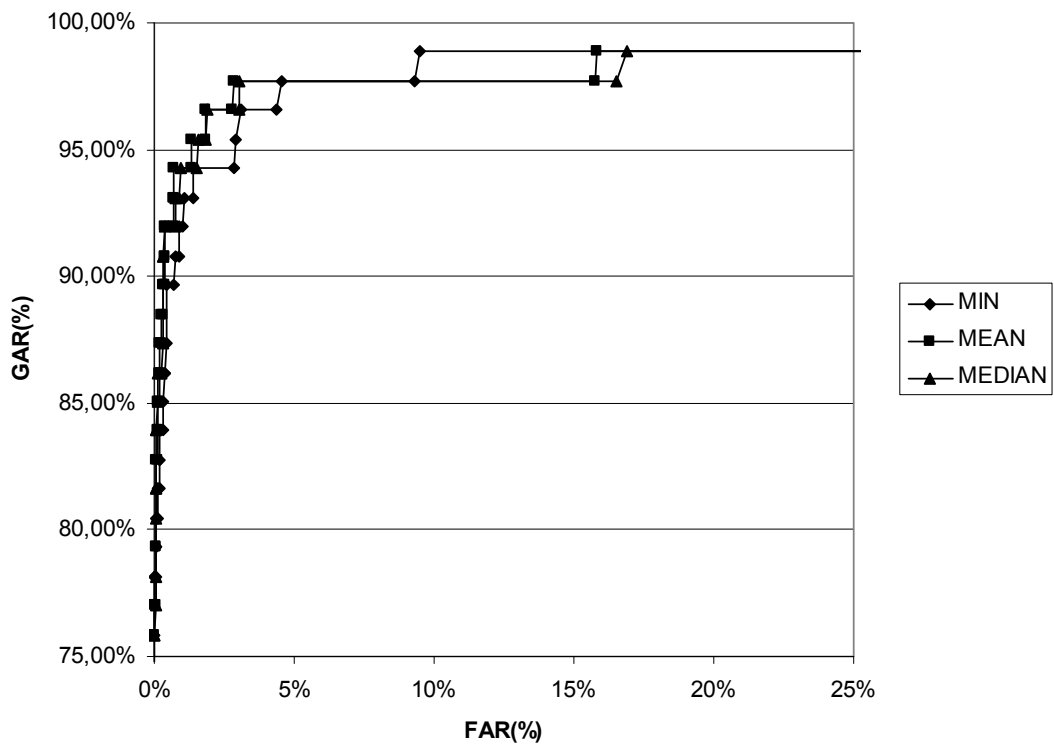
Tablica 5 prikazuje parametre korištene u eksperimentima verifikacije na bazi video sekvenci, zajedno s odgovarajućim EER i HmTER vrijednostima. Slike 52 i 53 prikazuju ROC krivulje za verifikaciju lica i verifikaciju dlana.

Značajke	p	t	N_{PCA}	N_{LBLDA}	EER	HmTER
Lice	16	4	150	1500	2.83%	2.57%
Dlan	16	4	100	1300	10.73%	10.54%

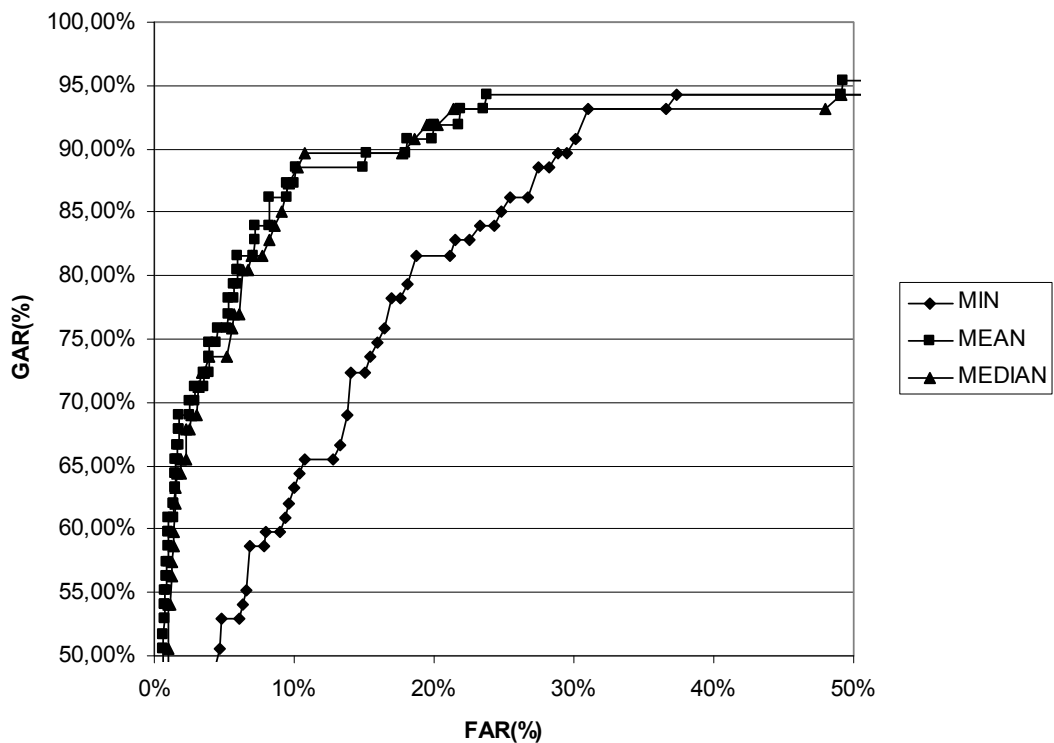
Tablica 5: Parametri i rezultati (EER i HmTER) za verifikaciju lica i dlana na temelju lokalne binarne LDA

Kao što se može vidjeti iz tablice i slika, korištenjem LBLDA značajki ostvareni su najbolji rezultati verifikacije za lice na bazi video sekvenci s EER = 2.83% i HmTER = 2.57%, te je time smanjena pogreška u odnosu na verifikaciju koja koristi značajke dobivene na temelju LDA (EER = 5.75%) ili značajke dobivene na temelju Gaborovog filtra (EER = 4.60%). Ovime je pokazana pogodnost lokalnih binarnih značajki za biometrijsku verifikaciju. Lokalne binarne značajke biti će manje osjetljive na lokalne promjene koje na licu mogu nastati kao posljedica promjene osvjetljenja, izraza lica, frizure itd.

Za verifikaciju na temelju dlana, korištenjem lokalnih binarnih značajki dobivenih na temelju LBLDA nije smanjena pogreška verifikacije u odnosu na verifikaciju na temelju značajki dobivenih „klasičnom“ LDA ili na temelju značajki dobivenih Gaborovim filtrom. S jedne strane, budući da dlan nije toliko podložan lokalnim promjenama izgleda koliko lice, nije za očekivati da će korištenje lokalnih značajki poboljšati rezultate verifikacije u tolikoj mjeri, koliko je to bilo za lice. S druge strane, rezultati postignuti Gaborovim filtrom također su, u slučaju dlana, bolji od onih postignutih pomoću LBLDA. Usporedimo li Gaborov filter i LBLDA, vidimo da i jedna i druga metoda daju lokalne binarne značajke. LBLDA značajke također možemo smatrati lokalnim filtrima, gdje je oblik svakog filtra dobiven na temelju učenja, posebno za svaku lokaciju na slici. To znači da je moguće u postupku učenja LBLDA dobiti lokalne filtre koji bi bili identični Gaborovom filtru, ukoliko bi se pokazalo da upravo takve značajke optimiziraju kriterijsku funkciju korištenu u LBLDA. Budući da su se u slučaju dlana značajke koje daje LBLDA pokazale lošijima od značajki dobivenih na temelju Gaborovog filtra, može se zaključiti da u postupku učenja nije bilo moguće naći značajke koje bi nosile više diskriminatorne informacije (u smislu Fisherovog kriterija) od onih dobivenih Gaborovim filtrom. Postupak učenja za LBLDA ovisiti će o korištenome skupu za učenje. U ovom slučaju, za učenje LBLDA koristila se po jedna video sekvenca za učenje po osobi. Moguće je da zbog niza faktora, kao što su nepreciznosti u lokalizaciji, vremenska bliskost uzoraka iste osobe i sl. skup korišten za učenje nije bio odgovarajući u smislu pronalaženja lokalnih baza koje bi dale dobre rezultate i na neviđenim video sekvencama.



Slika 52: ROC krivulje za verifikaciju lica na temelju LBLDA značajki



Slika 53: ROC krivulje za verifikaciju dlana na temelju LBLDA značajki

9.3.5. *Usporedba rezultata*

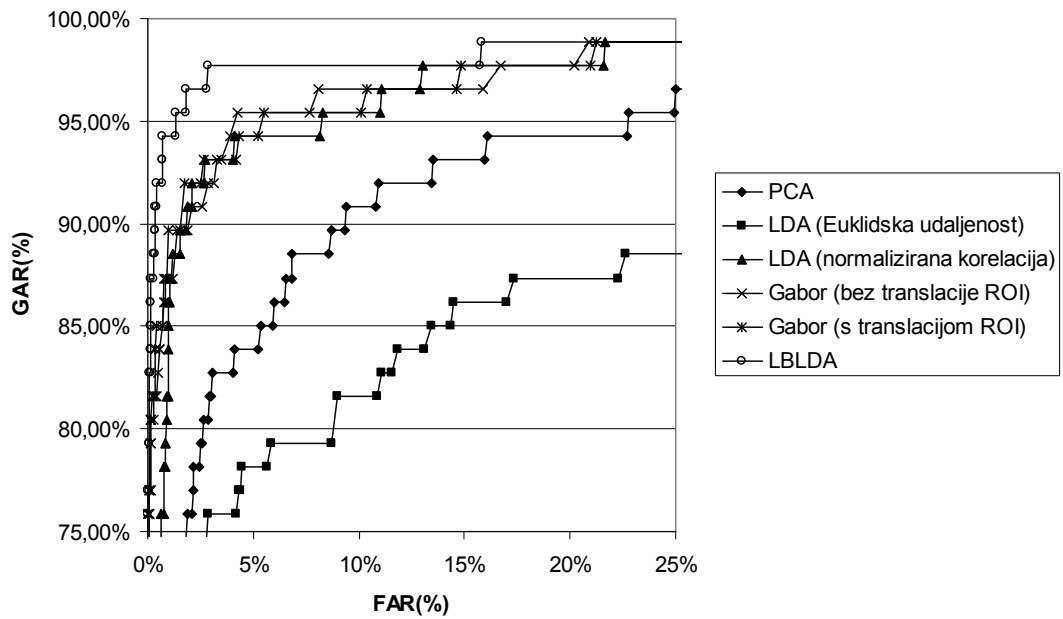
Na slici 54 prikazana je usporedba ROC krivulja za različite metode izlučivanja značajki lica. Slika 55 prikazuje istu usporedbu za dlan. Usporedba EER i HmTER vrijednosti dana je na slici 56 za lice i slici 57 za dlan, te u tablici 6.

Kao što se može vidjeti iz slika, za verifikaciju na temelju lica, značajke dobivene pomoću LBLDA pokazale su se najbolje u smislu najmanjih pogrešaka verifikacije. Kao što se može vidjeti na slici 54 uz isti GAR uporabom LBLDA značajki dobiva se osjetno manji FAR u odnosu na ostale metode. Ovo pokazuje pogodnost LBLDA značajki za verifikaciju na temelju lica.

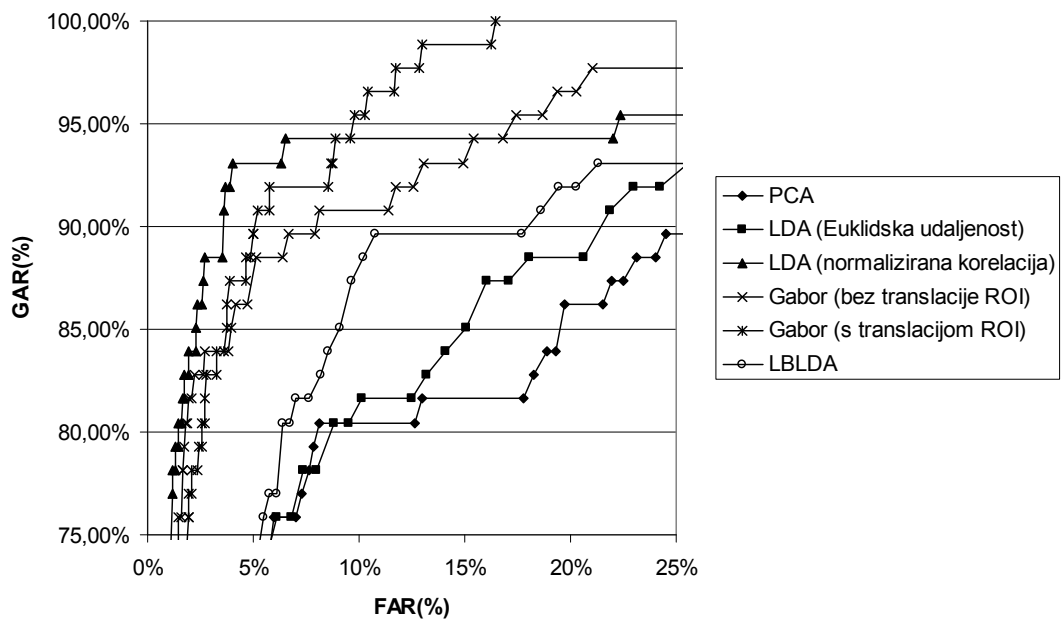
U eksperimentima verifikacije na temelju dlana najbolji rezultati ostvareni su pomoću LDA uz korištenje normalizirane korelacije za usporedbu vektora značajki. Ovdje je zanimljivo za primijetiti da je LDA, kao metoda za izlučivanje globalnih značajki, dala bolje rezultate verifikacije u odnosu na metode izlučivanja lokalnih značajki. Razlog za ovo može se naći u samoj strukturi dlana koji se, u usporedbi s licem, ponaša više kao čvrsto tijelo i nije toliko podložan lokalnim promjenama. Varijacije osvjetljenja, koje su jedan od argumenata za korištenje lokalnih značajki, također ne predstavljaju toliko problem kod dlana, jer će i u različitim uvjetima osvjetljenja cijela površina dlana u jednom okviru video sekvence biti osvijetljena približno ravnomjerno (možemo reći da je dlan na video sekvencama približno planaran).

Kao druga metoda za verifikaciju na temelju dlana po pogreškama verifikacije pokazao se Gaborov filter, koji je u slučaju dlana dao bolje rezultate od LBLDA, premda se obje metode temelje na lokalnim binarnim značajkama. Razlozi za ovakav poredak mogu se tražiti u neadekvatnom skupu za učenje LBLDA i u obliku Gaborovog filtra koji je pogodan za naglašavanje diskriminatornih značajki dlana, kao što su glavne linije dlana.

Pored eksperimenata verifikacije za sve metode načinjen je eksperiment klasifikacije (identifikacija u zatvorenom skupu). Rezultati eksperimenta mogu se vidjeti na slici 58 za lice i slici 59 za dlan. Poredak metoda u smislu točnosti raspoznavanja je u većini slučajeva ostao isti kao poredak u smislu točnosti verifikacije, osim u slučaju raspoznavanja dlana gdje se pokazalo da se značajkama dobivenim Gaborovim filtrom može ostvariti bolja točnost raspoznavanja nego sa LDA značajkama koje su se pokazale bolje za verifikaciju.



Slika 54: ROC krivulje za različite metode izlučivanja značajki za verifikaciju lica



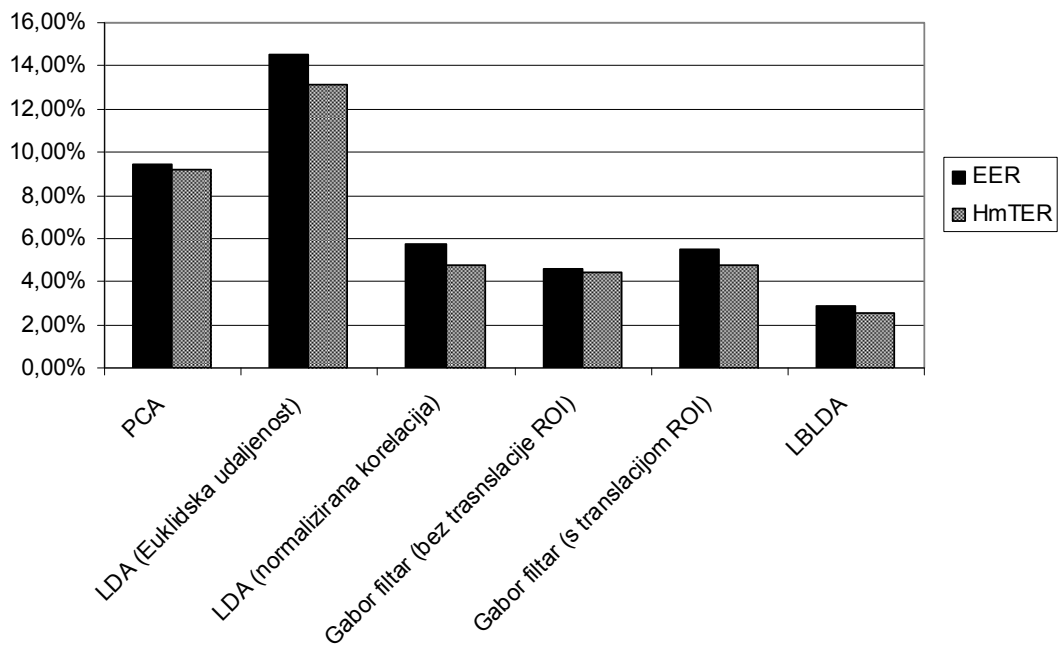
Slika 55: ROC krivulje za različite metode izlučivanja značajki za verifikaciju dlana

Značajka	Metoda izlučivanja značajki	EER	HmTER
Lice	PCA	9.41%	9.18%
	LDA (Euklidska udaljenost)	14.51%	13.16%
	LDA (Normalizirana korelacija)	5.75%	4.79%
	Gaborov filter (bez translacije ROI)	4.60%	4.43%
	Gaborov filter (s translacijom ROI)	5,52%	4.78%
	LBLDA	2,83%	2.57%
Dlan	PCA	17.24%	13.83%
	LDA (Euklidska udaljenost)	15.09%	14.17%
	LDA (Normalizirana korelacija)	6.51%	5.45%
	Gaborov filter (bez translacije ROI)	9.20%	8.30%
	Gaborov filter (s translacijom ROI)	8.05%	6.92%
	LBLDA	10.73%	10.54%

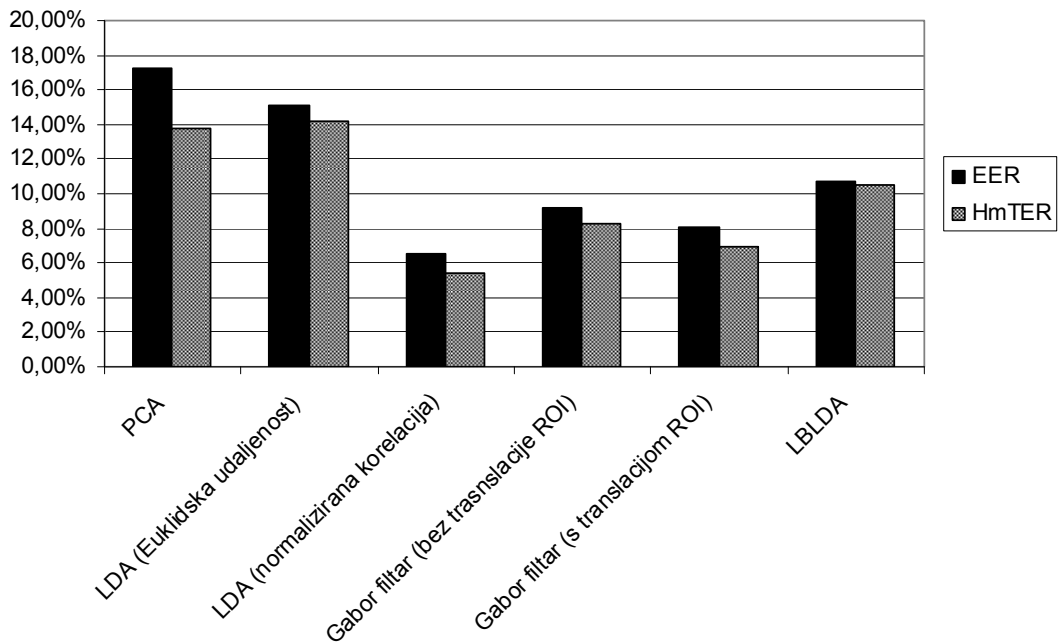
Tablica 6: Usporedba EER i HmTER vrijednosti za različite metode izlučivanja značajki

Značajka	Metoda izlučivanja značajki	Točnost klasifikacije
Lice	PCA	80.46%
	LDA (Euklidska udaljenost)	73.56%
	LDA (Normalizirana korelacija)	85.05%
	Gaborov filter (bez translacije ROI)	89.65%
	Gaborov filter (s translacijom ROI)	89.65%
	LBLDA	93.10%
Dlan	PCA	70.11%
	LDA (Euklidska udaljenost)	67.81%
	LDA (Normalizirana korelacija)	80.45%
	Gaborov filter (bez translacije ROI)	82.75%
	Gaborov filter (s translacijom ROI)	90.80%
	LBLDA	73.56%

Tablica 7: Usporedba točnosti raspoznavanja za različite metode izlučivanja značajki

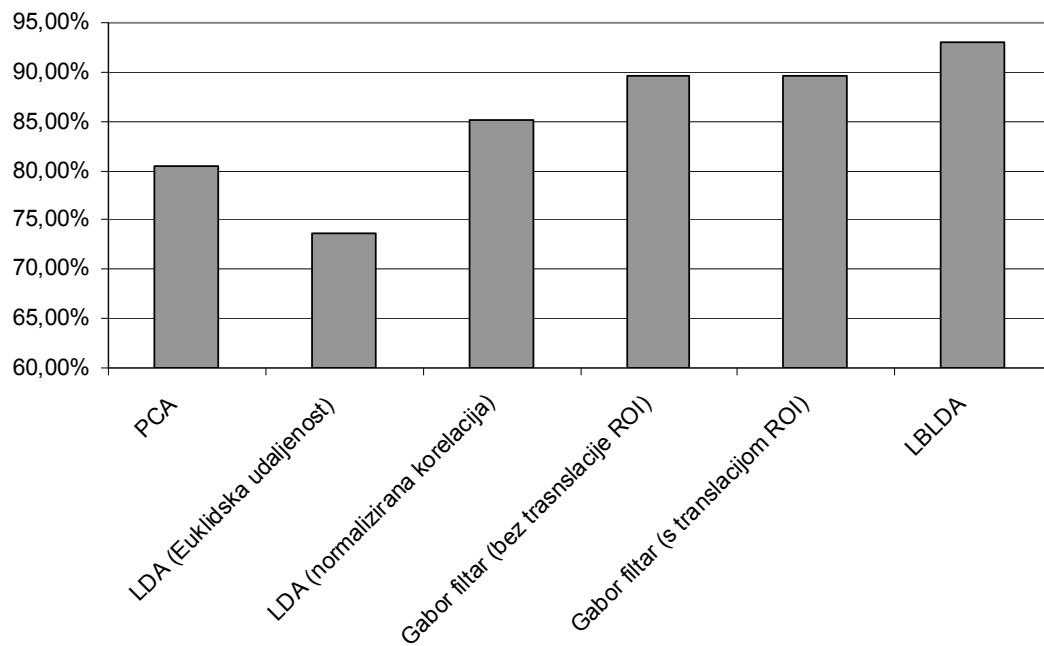


Slika 56: Usporedba EER i HmTER vrijednosti za različite metode izlučivanja značajki za verifikaciju na temelju lica

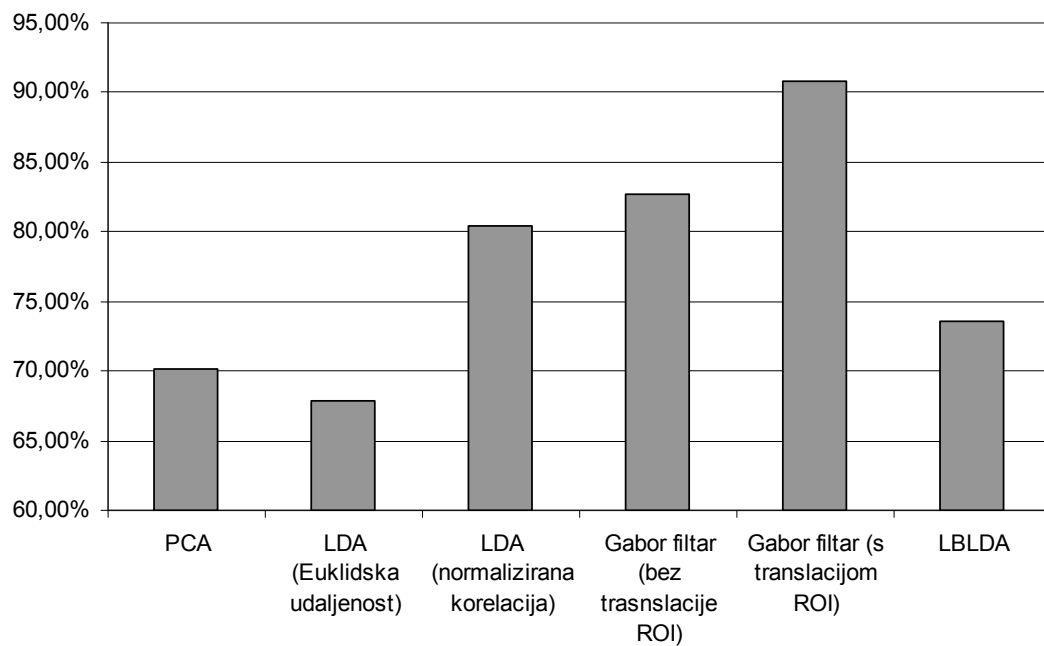


Slika 57: Usporedba EER i HmTER vrijednosti za različite metode izlučivanja značajki za verifikaciju na temelju dlana

9. EKSPERIMENTALNA EVALUACIJA



Slika 58: Točnost raspoznavanja lica za različite metode izlučivanja značajki



Slika 59: Točnost raspoznavanja dlana za različite metode izlučivanja značajki

9.3.6. *Usporedba brzine izvođenja*

Izvođenje za svaku od metoda izlučivanja značajki može se podijeliti u tri faze. U ovom poglavlju brzina za svaku od faza će biti evaluirana zasebno, te će za svaku od faza biti dane usporedbe za različite metode izlučivanja značajki.

Prva faza je faza učenja koju je potrebno izvesti samo jednom, na skupu za učenje. Ovo je najčešće vremenski najzahtjevnija faza. U drugoj fazi se značajke izlučuju iz slike. Ovo je potrebno izvesti jednom za svaku sliku, odnosno jednom za svaki okvir video sekvence u kome je pronađeno lice ili dlan. Konačno, potrebno je usporediti različite vektore značajki. Broj usporedbi ovisiti će o broju vektora značajki koje su pohranjene u bazi za svakog korisnika.

Brzina faze učenja mjerena je na bazi statičnih slika lica koja sadrži 2360 slika 295 osoba. Brzina faze izlučivanja značajki i usporedbe vektora značajki neće ovisiti o veličini baze.

Tablica 8 prikazuje vrijeme izvođenja za svaku od faza različitih metoda izlučivanja značajki. Sve metode implementirane su u programskom jeziku C++, a vrijeme izvođenja mjereno je na Intel Core 2 Quad procesoru s radnim taktom na 2.4GHz. Sve implementacije su jednodretvene, što znači da iskorištavaju samo jednu jezgru procesora.

Metoda izlučivanja značajki i parametri / mjera podudaranja	Vrijeme učenja	Vrijeme izlučivanja značajki	Vrijeme usporedbe vektora značajki
PCA, 100 značajki / Euklidska udaljenost	233s	0.73ms	0.20ns
LDA, 300 PCA značajki, 100 LDA značajki / normalizirana korelacija	241s	0.73ms	0.46ns
LBLDA, $p = 16$, $t = 8$, NPCA = 100, NLBLDA = 1000 / Hammingova udaljenost	34s	0.59ms	0.16ns
LBLDA, $p = 16$, $t = 4$, NPCA = 150, NLBLDA = 1500 / Hammingova udaljenost	139s	1.40ms	0.22ns
Gaborov filter, veličina filtra 7x7 / Hammingova udaljenost bez translacija ROI	0.0s	3.11ms	0.32ns
Gaborov filter, veličina filtra 7x7 / Hammingova udaljenost s translacijama ROI u rasponu ± 2 po x i ± 2 po y osi slike	0.0s	3.11ms	8.05ns

Tablica 8: Vrijeme izvođenja za pojedine faze za različite metode izlučivanja značajki

Kao što se može vidjeti iz tablice, ukoliko izuzmemo Gaborov filter koji ne zahtijeva učenje, LBLDA ima najmanje vrijeme učenja. Razlog tome je što LBLDA prilikom postupka učenja ne obavlja operacije na velikim matricama koje uključuju podatke o cijeloj slici, kao što to čine PCA i LDA. Također, bez obzira što su binarni vektori značajki koji se koriste u LBLDA i izlučivanju značajki Gaborovim filtrom višestruko veći od realnih vektora značajki koji se koriste u PCA i LDA, vremena usporedbe binarnih vektora značajki su manja ili usporediva s vremenima usporedbe realnih vektora značajki. Razlog za ovo je mogućnost vrlo brzog računanja Hammingove udaljenosti pomoću binarnih operacija i tablica pretraživanja. Iznimka od ovoga je usporedba binarnih vektora značajki dobivenih Gaborovim filtrom ukoliko koristimo različite translacije ROI prilikom usporedbe, jer svaka translacija znači dodatno poklapanje i usporedbu vektora značajki. Gaborov filter također ima najveće vrijeme izlučivanja značajki zbog spore operacije konvolucije slike sa filtrom.

9.4. Evaluacija mjere kvalitete područja interesa dlana

Da bi se evaluirala mjera kvalitete načinjena je serija eksperimenata na slijedeći način. Za svaku video sekvencu prvo su izlučena područja interesa dlana. Nakon toga izračunata je mjera kvalitete za svako područje interesa. Konačno, iz svake video sekvence odbačen je određeni postotak područja interesa tako da su zadržana područja interesa s najvećom mjerom kvalitete.

Preostala područja interesa korištena su za izlučivanje značajki i verifikaciju. Ovo je načinjeno za različite postotke zadržanih područja interesa (10%, 20%, ..., 90%). Ukoliko je postotak zadržanih područja interesa 100%, tada se koriste sva područja interesa, te su rezultati identični kao oni predstavljeni u prethodnom poglavlju. Ovakav eksperiment načinjen je za četiri metode izlučivanja značajki:

- Analizu glavnih komponenti (PCA)
- Linearnu diskriminantnu analizu (LDA) uz korištenje normalizirane korelacije kao mjere podudaranja
- Gaborov filtar (uz korištenje više translacija područja interesa)
- Lokalnu binarnu LDA (LBLDA)

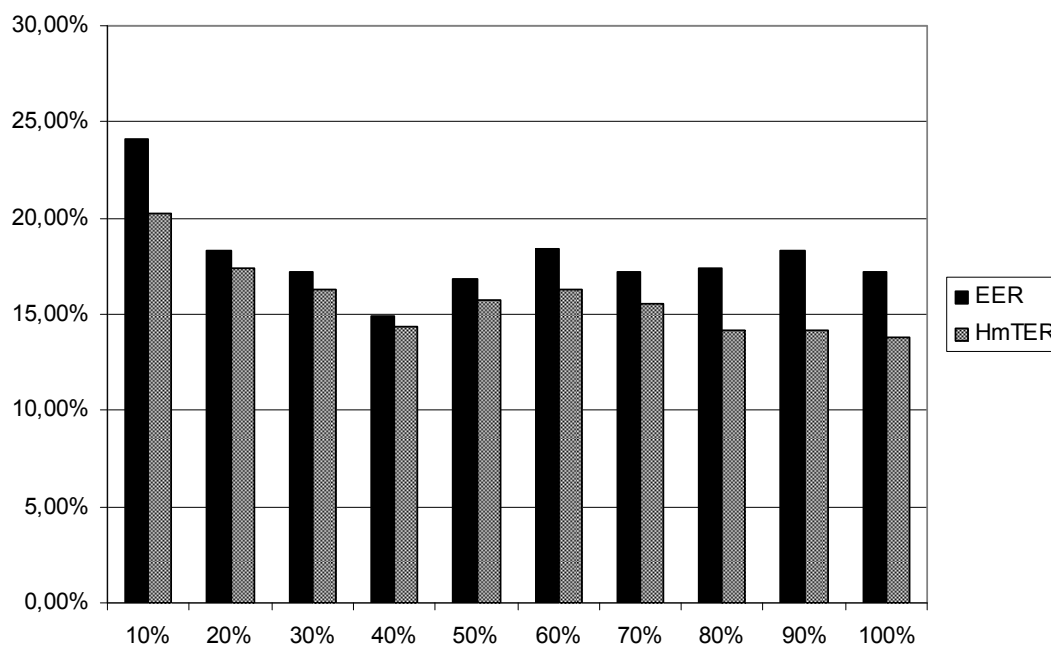
Rezultati eksperimenata prikazani su kao EER i HmTER za različite postotke zadržanih područja interesa.

Rezultati eksperimenta prikazani su na slici 60 za PCA, slici 61 za LDA, slici 62 za Gaborov filtar i slici 63 za LBLDA. Kao što se može vidjeti iz slika, različite metode izlučivanja značajki različito reagiraju na uklanjanje dijela interesnih područja u skladu s mjerom kvalitete. Zajedničko za sve metoda je da, ukoliko se postupno eliminira do 60% interesnih područja, pogreška verifikacije lagano raste, ili čak opada. Za razliku od toga, ukoliko eliminiramo 70% ili više interesnih područja, za sve metode pogreška verifikacije počinje rasti vrlo strmo, budući da ovime eliminiramo preveliki dio interesnih područja pogodnih za verifikaciju.

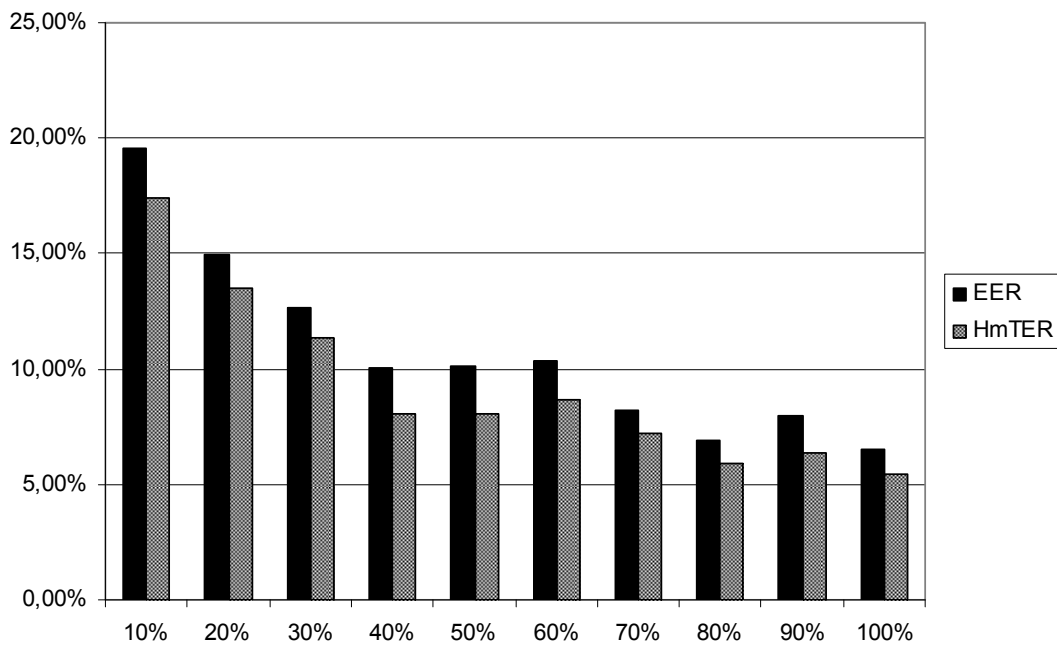
PCA i Gaborov filtar, u kombinaciji s mjerom kvalitete, pokazali su se vrlo uspješni, jer se, uz uklanjanje dijela interesnih područja, ovim metodama može ostvariti bolji rezultat nego ukoliko se koriste sva interesna područja, uz povećanje brzine verifikacije. Posebno dobrom pokazalo se korištenje mjere kvalitete u kombinaciji s izlučivanjem značajki Gaborovim filtrom. Uz zadržavanje 60% interesnih područja u skladu s njihovom mjerom kvalitete moguće je ostvariti $EER = 6.08\%$, što je najbolji rezultat na bazi video sekvenci korištenoj u eksperimentima, za bilo koju metodu izlučivanja značajki. Za Gaborov filtar moguće je ostvariti približno iste pogreške verifikacije kao za sva interesna područja ukoliko se zadrži

samo 40% interesnih područja. Za PCA se sličan rezultat dobiva uz zadržavanje samo 30% interesnih područja.

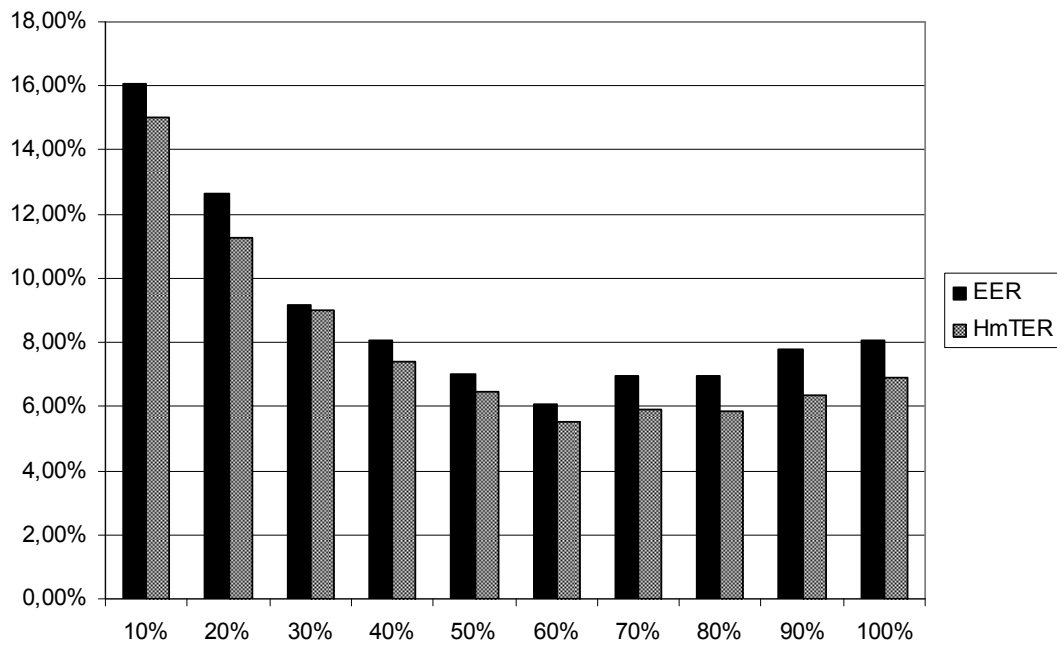
LDA i LBLDA u kombinaciji s uklanjanjem dijela interesnih područja na temelju mjere kvalitete dale su lošije rezultate, što bi mogla biti posljedica korištenja značajki koje nije lako modelirati na temelju izgleda kvalitetnih dlanova. Metoda za ocjenu mjere kvalitete zato bi se trebala koristiti u kombinaciji s izlučivanjem značajki na temelju analize glavnih komponenti ili Gaborovog filtra (na temelju kojega se dobivaju najbolji rezultati verifikacije), ne u kombinaciji s ostalim metodama na kojima su provedeni eksperimenti. Za neku drugu metodu izlučivanja značajki, prvo bi eksperimentalno trebalo utvrditi pogodnost korištenja ocjene kvalitete za odabir područja interesa.



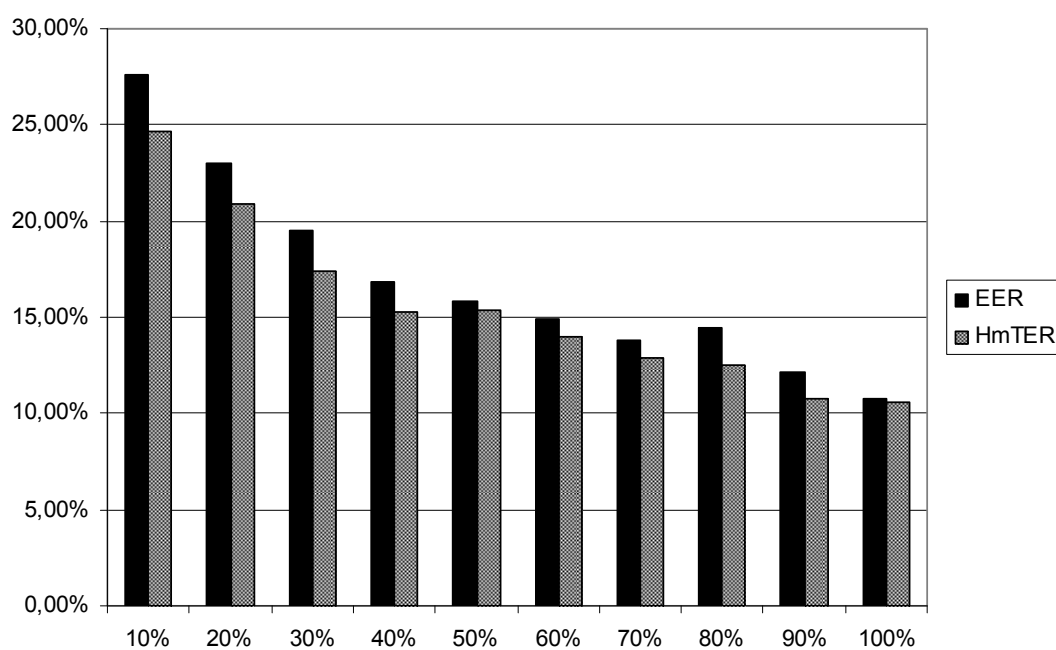
Slika 60: Pogreške verifikacije za različite postotke zadržanog ukupnog broja interesnih područja u skladu s mjerom kvalitete, uz korištenje PCA za izlučivanje značajki



Slika 61: Pogreške verifikacije za različite postotke zadržanog ukupnog broja interesnih područja u skladu s mjerom kvalitete, uz korištenje LDA za izlučivanje značajki



Slika 62: Pogreške verifikacije za različite postotke zadržanog ukupnog broja interesnih područja u skladu s mjerom kvalitete, uz korištenje Gaborovog filtra za izlučivanje značajki



Slika 63: Pogreške verifikacije za različite postotke zadržanog ukupnog broja interesnih područja u skladu s mjerom kvalitete, uz korištenje LBLDA za izlučivanje značajki

9.5. Evaluacije fuzije dlana i lica

Kako bi se pokazalo može li se fuzijom dlana i lica ostvariti bolji rezultat verifikacije u odnosu na verifikaciju temeljenu samo na licu, odnosno samo na dlanu, načinjen je niz eksperimenata fuzije. Fuzija je primijenjena na više razina: na razini podudaranja i na razini odluke, gdje je primijenjena fuzija na temelju klasifikatora i fuzija na temelju glasanja.

Fuzija na razini senzora nije pogodna za razvijeni sustav, budući da sustav koristi samo jedan senzor. Fuzija na razini vektora značajki također nije pogodna jer se u video sekvencama dlan i lice obrađuju neovisno. To znači da je moguće da u nekom okviru video sekvence bude detektirano lice, ali ne bude detektirana ruka i obrnuto. Ukoliko bismo željeli ipak primijeniti fuziju na razini vektora značajki, za najbolje rezultate bilo bi potrebno upariti vektore značajki lica iz svih okvira s vektorima značajki dlana iz svih okvira, što bi dalo preveliki broj vektora značajki i preveliko vrijeme usporedbe za primjenu u stvarnom vremenu. Alternativno, mogli bismo odabrati za svaku video sekvencu samo jedan vektor značajki za lice i jedan vektor značajki za dlan, međutim ovo bi vodilo pogoršanju pogrešaka verifikacije.

U eksperimentima su razmatrane slijedeće značajke

- Značajke lica dobivene na temelju LDA (uz korištenje normalizirane korelacije kao mjere podudaranja)
- Značajke lica dobivene na temelju Gaborovog filtra (uz korištenje normalizirane Hammingove udaljenosti kao mjere podudaranja)
- Značajke lica dobivene na temelju lokalne binarne LDA (uz korištenje normalizirane Hammingove udaljenosti kao mjere podudaranja)
- Značajke dlana dobivene na temelju LDA (uz korištenje normalizirane korelacije kao mjere podudaranja)
- Značajke dlana dobivene na temelju Gaborovog filtra (uz korištenje normalizirane Hammingove udaljenosti kao mjere podudaranja)
- Značajke dlana dobivene na temelju lokalne binarne LDA (uz korištenje normalizirane Hammingove udaljenosti kao mjere podudaranja)

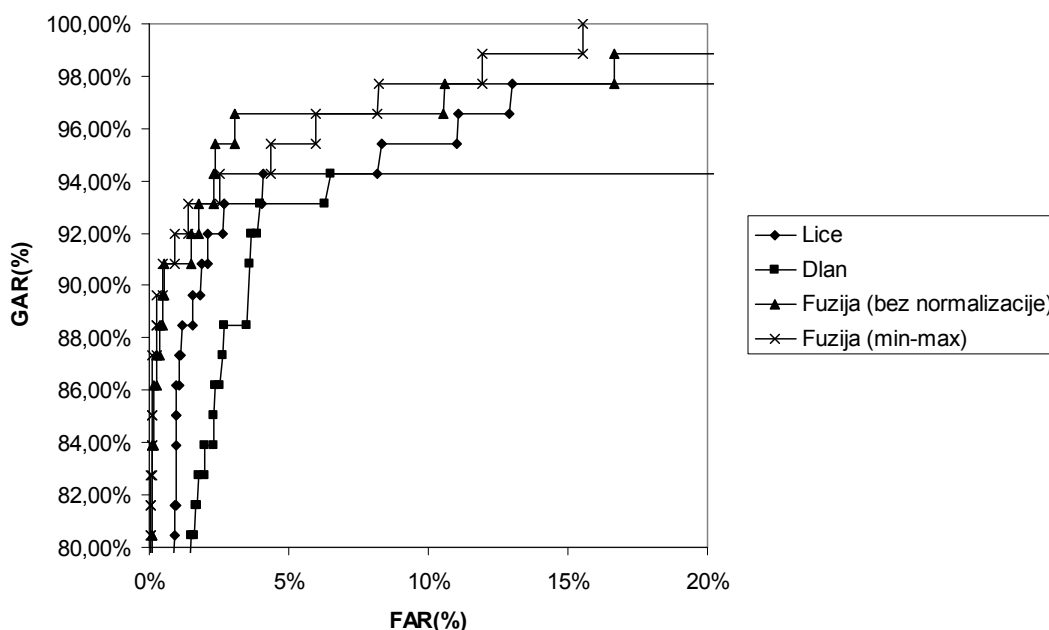
Ostale značajke koje su korištene u eksperimentima samo s licem i samo s dlanom (npr. PCA značajke) nisu razmatrane u eksperimentima fuzije jer daju značajno lošije rezultate od navedenih metoda.

9.5.1. Fuzija na razini mjere podudaranja

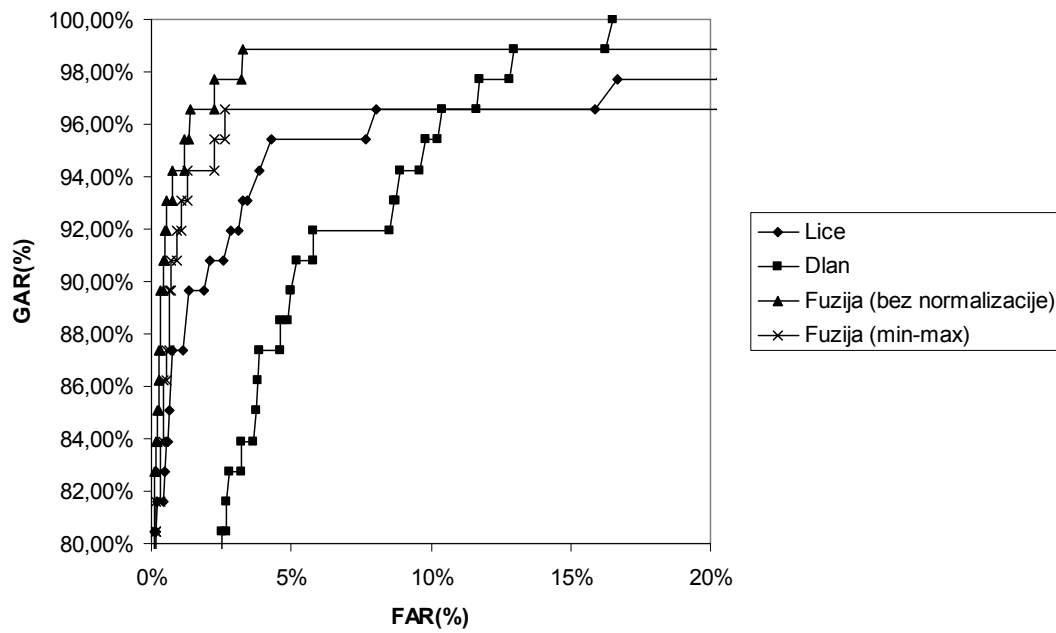
U eksperimentima fuzije na razini mjere podudaranja korištena je jednostavna suma kao pravilo fuzije. Odluka o prihvatanju ili odbijanju korisnika tada se donosi na temelju usporedbe sume s pragom. Za svaku kombinaciju značajki načinjeni su eksperimenti

- bez normalizacije. Budući da su sve razmatrane mjere podudaranja već normalizirane u intervalu $[-1, 1]$ (normalizirana korelacija) ili $[0, 1]$ (normalizirana Hammingova udaljenost) nije nužno potrebno raditi dodatni korak normalizacije
- Uz korištenje min-max normalizacije. Min-max normalizacija odabrana je zbog svoje jednostavnosti i jer u velikom broju slučajeva daje dobre rezultate [62][65].

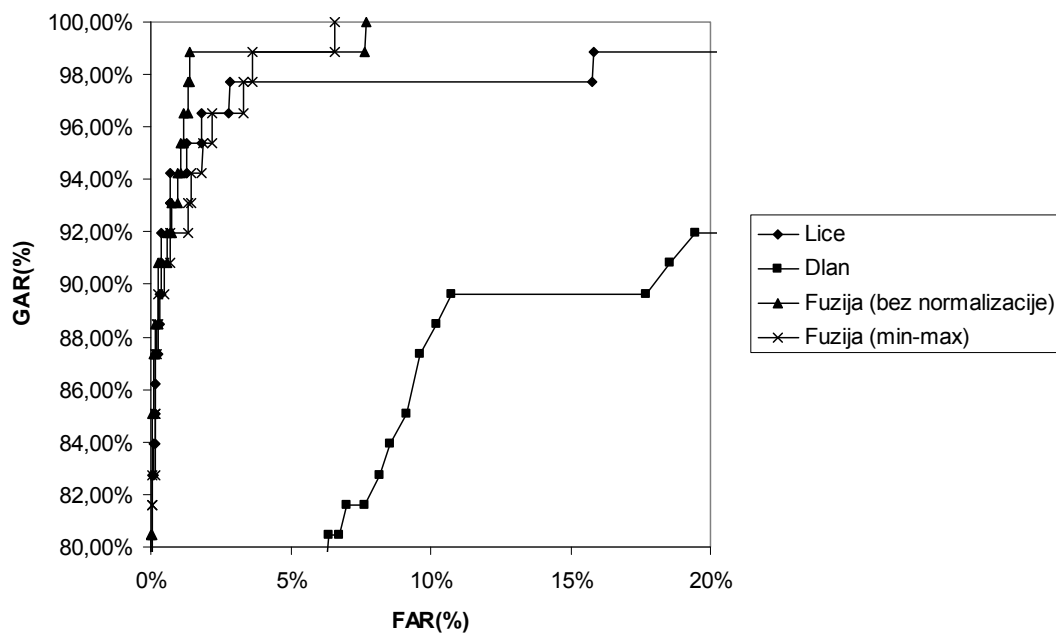
U prvoj seriji eksperimenata načinjena je fuzija značajki dlana i lica dobivenih istom metodom izlučivanja značajki. Tako slika 64 prikazuje ROC krivulje za fuziju značajki dlana i lica dobivenih na temelju LDA, slika 65 prikazuje ROC krivulje za fuziju značajki dlana i lica dobivenih na temelju Gaborovog filtra, a slika 66 prikazuje ROC krivulje za fuziju značajki dlana i lica dobivenih na temelju lokalne binarne LDA.



Slika 64: ROC krivulje za fuziju značajki dlana i lica dobivenih na temelju LDA



Slika 65: ROC krivulje za fuziju značajki dlana i lica dobivenih na temelju Gaborovog filtra



Slika 66: ROC krivulje za fuziju značajki dlana i lica dobivenih na temelju lokalne binarne LDA

Kao što se može vidjeti iz slika, fuzija u sva tri slučaja poboljšava rezultate verifikacije. Također se može vidjeti da se za fuziju istovjetnih metoda izlučivanja značajki bolji rezultati dobivaju bez normalizacije nego uz min-max normalizaciju.

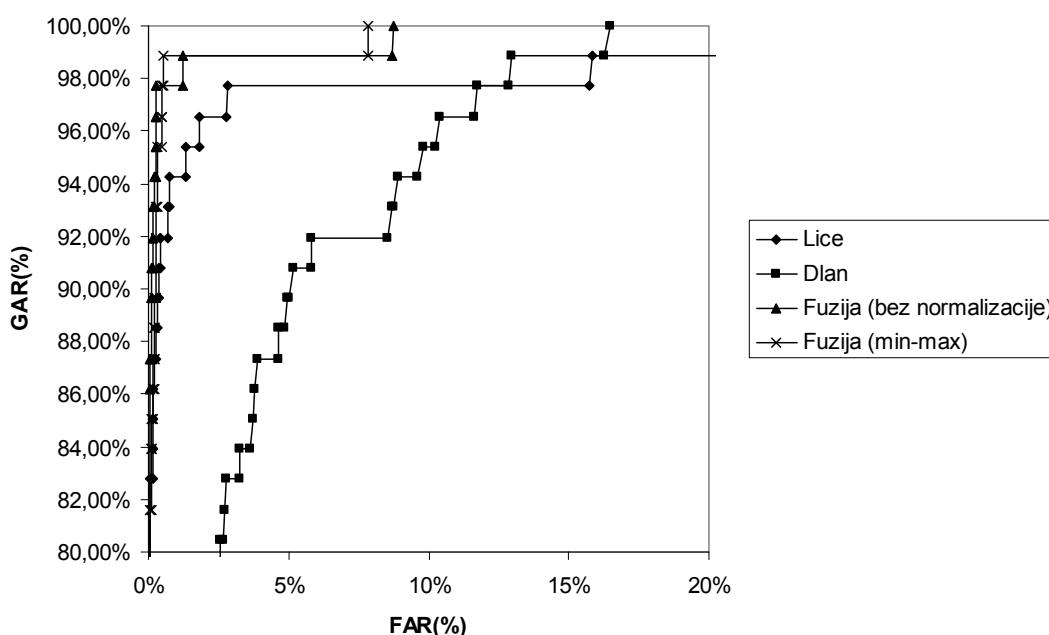
9. EKSPERIMENTALNA EVALUACIJA

Slijedeće je načinjena fuzija na temelju svih mogućih kombinacija metoda izlučivanja značajki dlana i lica. Za svaku kombinaciju značajki zabilježeni su EER i HmTER. Rezultati eksperimenta prikazani su u tablici 9.

Kao što se može vidjeti iz tablice, najmanje ostvarene pogreške verifikacije odnose se na kombinaciju značajki lica dobivenih na temelju LBLDA i značajki dlana dobivenih na temelju Gaborovog filtra. ROC krivulja za ovu kombinaciju prikazana je na slici 67. Ova kombinacija značajki daje najmanje pogreške verifikacije i kada se ne koristi normalizacija prije fuzije i kada se koristi min-max normalizacija. Korištenjem min-max normalizacije dobiva se nešto bolji rezultat s $EER = 1.14\%$ i $HmTER = 0.84\%$. Ove pogreške su manje u odnosu na pogreške ostvarene koristeći samo značajke lica, odnosno samo značajke dlana, što znači da se fuzijom na razini mjere podudaranja može ostvariti poboljšanje točnosti verifikacije u odnosu na verifikaciju koja koristi samo jedan tip značajki.

Značajke lica korištene u fuziji	Značajke dlana korištene u fuziji	Pogreška verifikacije nakon fuzije (bez normalizacije)	Pogreška verifikacije nakon fuzije (min-max normalizacija)
LDA	LDA	EER = 3.45% HmTER = 3.25%	EER = 4.60% HmTER = 4.14%
LDA	Gaborov filter	EER = 4.60% HmTER = 4.05%	EER = 2.29% HmTER = 1.88%
LDA	LBLDA	EER = 5.75% HmTER = 4.33%	EER = 5.75% HmTER = 4.96%
Gaborov filter	LDA	EER = 5.75% HmTER = 3.21%	EER = 4.60% HmTER = 3.19%
Gaborov filter	Gaborov filter	EER = 2.29% HmTER = 2.20%	EER = 2.29% HmTER = 1.90%
Gaborov filter	LBLDA	EER = 4.60% HmTER = 3.13%	EER = 4.60% HmTER = 3.68%
LBLDA	LDA	EER = 4.60% HmTER = 3.03%	EER = 3.45% HmTER = 2.73%
LBLDA	Gaborov filter	EER = 1.21% HmTER = 1.18%	EER = 1.14% HmTER = 0.84%
LBLDA	LBLDA	EER = 1.39% HmTER = 1.26%	EER = 3.30% HmTER = 2.39%

Tablica 9: Pogreške verifikacije za fuziju različitih značajki lica s različitim značajkama dlana na razini mjere podudaranja



Slika 67: ROC krivulje za fuziju značajki lica dobivenih pomoću LBLDA i značajki dlana dobivenih pomoću Gaborovog filtra

U slijedećem nizu eksperimenata pokušano je kombinirati više različitih značajki dlana i više različitih značajki lica, ne bi li se korištenjem više značajki iste modalnosti ostvarila bolja točnost verifikacije. Rezultati eksperimenata prikazani su u tablici 10.

U tablici su podebljani najbolji rezultat koji se odnosi samo na dlan, najbolji rezultat koji se odnosi samo na lice, te najbolji rezultat koji se odnosi na kombinaciju značajki koja uključuje i značajke dlana i značajke lica. Kao što se može vidjeti iz tablice, fuzijom značajki dlana dobivenih na temelju LDA i Gaborovog filtra moguće je postići smanjenje pogreške verifikacije u odnosu na korištenje samo LDA značajki dlana. Slično, fuzijom LBLDA značajki i značajki dobivenih na temelju Gaborovog filtra, moguće je smanjiti pogreške verifikacije za lice. Međutim, u eksperimentima koji uključuju i značajke dlana i značajke lica nije dobiveno poboljšanje u odnosu na fuziju koja uključuje samo značajke lica dobivene pomoću LBLDA i značajke dlana dobivene pomoću Gaborovog filtra.

Značajke korištene u fuziji	Pogreška verifikacije nakon fuzije (bez normalizacije)	Pogreška verifikacije nakon fuzije (min-max normalizacija)
LDA (dlan) + Gaborov filter (dlan)	EER = 5.74% HmTER = 4.70%	EER = 5.02% HmTER = 4.48%
LDA (dlan) + LBLDA (dlan)	EER = 6.89% HmTER = 5.40%	EER = 6.89% HmTER = 6.11%
Gaborov filter (dlan) + LBLDA (dlan)	EER = 9.01% HmTER = 8.02%	EER = 9.19% HmTER = 8.12%
LDA (dlan) + Gaborov filter (dlan) + LBLDA (dlan)	EER = 6.72% HmTER = 4.93%	EER = 6.89% HmTER = 5.18%
LBLDA (lice) + Gaborov filter (lice)	EER = 2.29% HmTER = 1.90%	EER = 2.29% HmTER = 2.03%
LDA (lice) + Gaborov filter (lice) + LBLDA (lice)	EER = 4.59% HmTER = 3.67%	EER = 3.23% HmTER = 2.74%
LBLDA (lice) + Gaborov filter (lice) + LDA (dlan) + Gaborov filter (dlan)	EER = 3.95% HmTER = 3.00%	EER = 3.06% HmTER = 2.42%
LBLDA (lice) + Gaborov filter (lice) + LBLDA (dlan) + Gaborov filter (dlan)	EER = 1.16% HmTER = 1.16%	EER = 3.20% HmTER = 2.35%
LDA (lice) + Gaborov filter (lice) + LBLDA (lice) + LDA (dlan) + Gaborov filter (dlan) + LBLDA (dlan)	EER = 4.59% HmTER = 3.07%	EER = 4.59% HmTER = 3.15%

Tablica 10: Pogreške verifikacije za različite kombinacije značajki korištene u fuziji

9.5.2. Fuzija na temelju klasifikatora

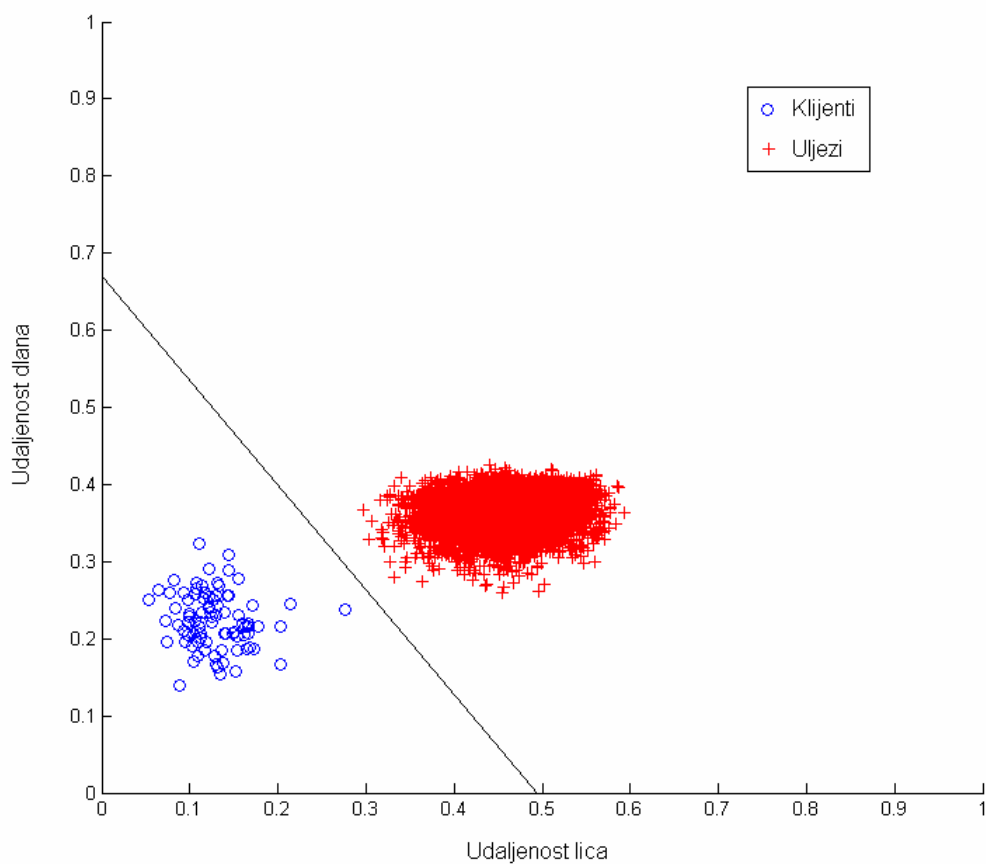
Kao klasifikator u fuziji je korišten stroj s potpornim vektorima (engl. Support Vector Machine, SVM) [69]. Korištena je gotova implementacija stroja s potpornim vektorima iz biblioteke LibSVM [70].

Za fuziju na temelju klasifikatora potreban je skup za učenje na kojemu je moguće dobiti mjere podudaranja kljenata i mjere podudaranja uljeza na kojima će se učiti klasifikator. Obzirom da se baza korištena u eksperimentima sastoji od samo dvije video sekvence, sekvence za učenje i sekvence za ispitivanje, mjere podudaranja za učenje dobivene su iz sekvenci za učenje na slijedeći način. Video sekvenca za učenje svake osobe podijeljena je na dva dijela, od kojih svaki sadrži upola manje okvira u odnosu na izvornu video sekvencu.

Mjere podudaranja korištene za učenje dobivene su usporedbom značajki iz druge polovice sekvence sa značajkama iz prve polovice sekvence za sve sekvence u bazi.

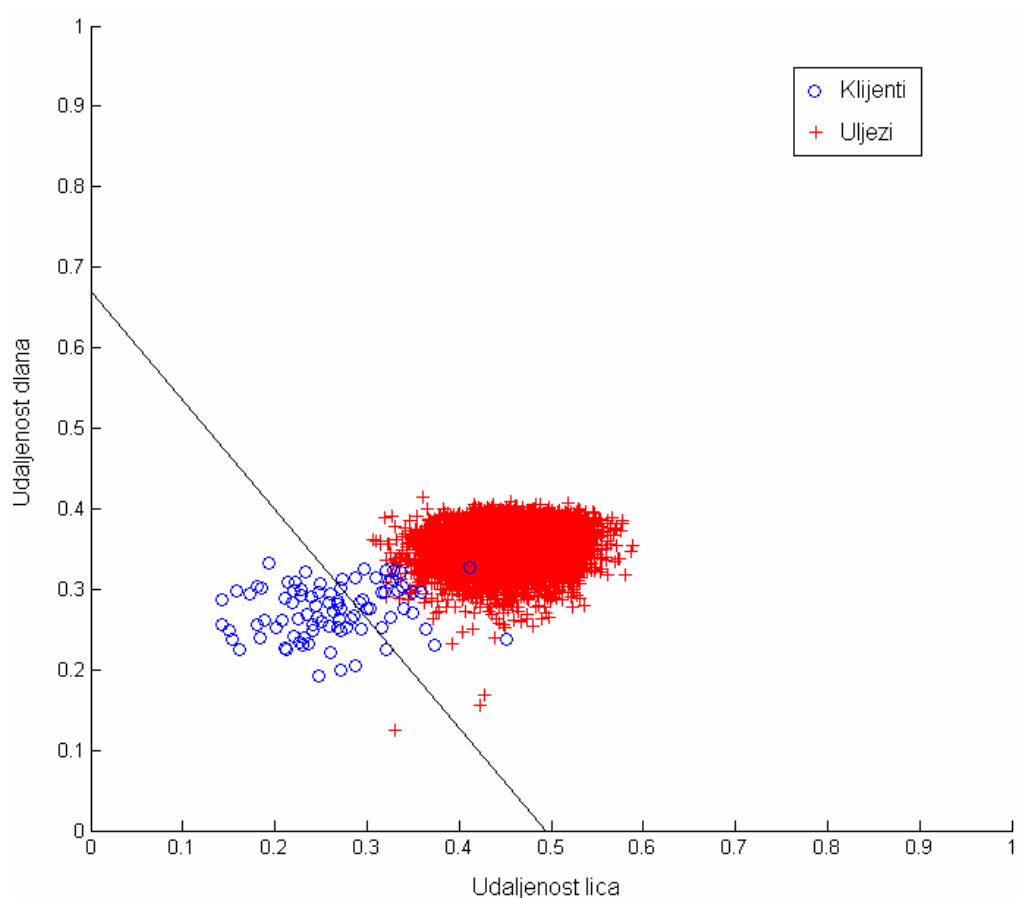
Mjere podudaranja dobivene na ovakav način u eksperimentima su se pokazale neprikladne za učenje klasifikatora, jer ne odražavaju dobro distribuciju mjera podudaranja klijenata koja se dobiva usporedbom različitih video sekvenci (snimljenih s vremenskim odmakom).

Slika 68 prikazuje distribuciju mjera udaljenosti lica i dlana za učenje, dobivenih na temelju LBLDA značajki za lice i Gaborovog filtra za dlan. Na slici je također prikazana granica između razreda klijenata i uljeza koju daje SVM.



Slika 68: Distribucija mjera udaljenosti lica i dlana za učenje s granicom razreda koju daje SVM za ovakve uzorke. Mjere udaljenosti dobivene su na temelju LBLDA značajki za lice i na temelju Gaborovog filtra za dlan.

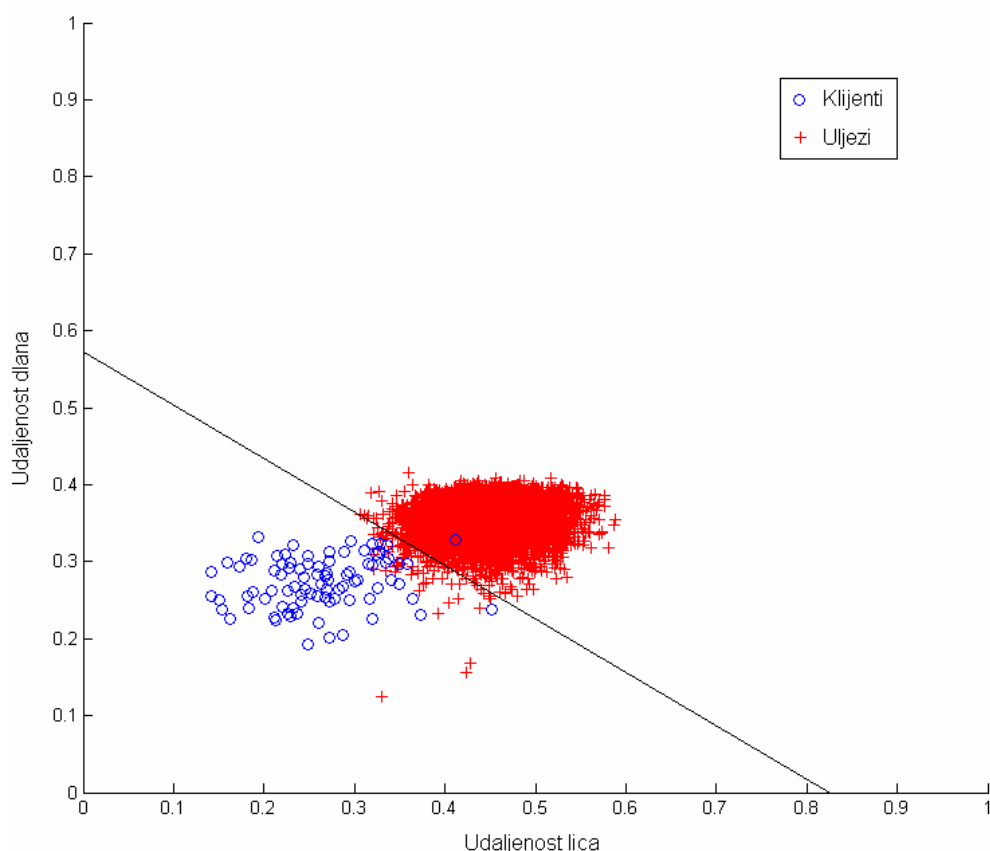
Slika 69 prikazuje distribuciju mjera udaljenosti lica i dlana za ispitivanje sa granicom razreda koja je dobivena na temelju mjera udaljenosti za učenje. Na slici se vidi puno veće preklapanje razreda klijenata i uljeza nego što je to bio slučaj na slici 68. Naročito je velika razlika u distribuciji mjera udaljenosti klijenata, koja je sada više pomaknuta prema mjerama udaljenosti koje odgovaraju uljezima. Ukoliko se ova granica razreda primijeni za verifikaciju, dobiva se rezultat $FRR = 31.03\%$, $FAR = 0.01\%$. Kao što je i bilo očekivano, dobiva se pretjerano veliki FRR budući da distribucija mjera podudaranja klijenata korištena za učenje ne predstavlja dobro distribuciju mjera podudaranja klijenata na skupu za ispitivanje.



Slika 69: Distribucija mjera udaljenosti lica i dlana za ispitivanje s granicom razreda koju daje SVM za mjere udaljenosti dobivene na temelju skupa za učenje. Mjere udaljenosti dobivene su na temelju LBLDA značajki za lice i na temelju Gaborovog filtra za dlan.

Zbog neprikladnosti mjera podudaranja dobivenih samo na temelju video sekvence za učenje, u daljnjim eksperimentima za učenje SVM klasifikatora korištene su mjere podudaranja dobivene na temelju skupa za ispitivanje. Slika 70 prikazuje distribuciju mjera udaljenosti lica i dlana dobivenih na temelju skupa za ispitivanje, također za LBLDA značajke lica i značajke dlana dobivene na temelju Gaborovog filtra. Na slici je prikazana granica između razreda klijenata i uljeza koju daje SVM za ovakve uzorke.

Pogreške verifikacije nakon fuzije za ovako izvedene eksperimente prikazane su u tablici 11.



Slika 70: Distribucija mjera udaljenosti lica i dlana za ispitivanje s granicom razreda koju daje SVM za ovakve uzorke. Mjere udaljenosti dobivene su na temelju LBLDA značajki za lica i na temelju Gaborovog filtra za dlan.

Značajke lica korištene u fuziji	Značajke dlana korištene u fuziji	Pogreška verifikacije nakon fuzije
LDA	LDA	FRR = 6.90% FAR = 2.17%
LDA	Gaborov filter	FRR = 1.14% FAR = 2.35%
LDA	LBLDA	FRR = 6.90% FAR = 4.24%
Gaborov filter	LDA	FRR = 5.75% FAR = 1.55%
Gaborov filter	Gaborov filter	FRR = 2.30% FAR = 2.35%
Gaborov filter	LBLDA	FRR = 4.60% FAR = 3.34%
LBLDA	LDA	FRR = 3.44% FAR = 1.31%
LBLDA	Gaborov filter	FRR = 1.14% FAR = 1.04%
LBLDA	LBLDA	FRR = 1.14% FAR = 2.59%

Tablica 11: Pogreške verifikacije za fuziju različitih značajki lica s različitim značajkama dlana pomoću SVM klasifikatora

U tablici je podebljano označen rezultat s najmanjom ukupnom pogreškom, koji se odnosi na fuziju LBLDA značajki lica i značajki dlana dobivenih na temelju Gaborovog filtra. Međutim, ovaj rezultat, čak i uz učenje stroja s potpornim vektorima na video sekvencama za ispitivanje, nije bolji od rezultata ostvarenih fuzijom na razini mjere podudaranja (tablica 9). Ipak, i ovaj je rezultat bolji od onih dobivenih ako se razmatra jednonačinska verifikacija (verifikacija samo na temelju lica ili samo na temelju dlana).

Pored korištenja linearnog stroja s potpornim vektorima, za najbolji slučaj fuzije također je korišten stroj s potpornim vektorima s jezgrenom funkcijom koja odgovara polinomu drugoga stupnja. U ovom eksperimentu ostvaren je gotovo identičan rezultat kao onaj ostvaren s linearnim SVM-om (FRR = 1.14%; FAR = 1.06%).

9.5.3. Fuzija na temelju glasanja

U fuziji na temelju glasanja, na temelju svake od značajki pojedinačno se donosi odluka radi li se o klijentu ili uljezu. Ova odluka donosi se na temelju usporedbe mjere podudaranja s pragom. Glasovi koji odgovaraju pojedinim značajkama, a glasaju za klijente, se zbrajaju. Suma glasova za klijente uspoređuje se s pragom t . Ukoliko je suma glasova veća ili jednaka t , korisnik se proglašava klijentom, a u protivnom se odbija.

U prvoj seriji eksperimenata s fuzijom na temelju glasanja, razmatrana je fuzija dvaju značajki, od kojih jedna odgovara licu, a druga dlanu. Ovakav eksperiment je načinjen za sve moguće kombinacije metoda izlučivanja značajki. U eksperimentima je evaluirano korištenje praga $t = 1$ (što znači da je dovoljno da jedna značajka glasa za klijenta da bi se korisnik proglasio klijentom) i praga $t = 2$ (što znači da obje značajke moraju glasati za klijenta da bi se korisnik proglasio klijentom).

Značajke lica korištene u fuziji	Značajke dlana korištene u fuziji	Pogreška verifikacije nakon fuzije ($t = 1$)	Pogreška verifikacije nakon fuzije ($t = 2$)
LDA	LDA	EER = 3.45% HmTER = 3.11%	EER = 4.71% HmTER = 3.94%
LDA	Gaborov filter	EER = 3.41% HmTER = 2.77%	EER = 2.35% HmTER = 2.33%
LDA	LBLDA	EER = 4.60% HmTER = 4.22%	EER = 5.75% HmTER = 4.48%
Gaborov filter	LDA	EER = 3.45% HmTER = 2.37%	EER = 4.60% HmTER = 3.70%
Gaborov filter	Gaborov filter	EER = 3.45% HmTER = 3.37%	EER = 3.45% HmTER = 2.89%
Gaborov filter	LBLDA	EER = 3.67% HmTER = 3.48%	EER = 4.60% HmTER = 4.09%
LBLDA	LDA	EER = 2.09% HmTER = 1.62%	EER = 3.34% HmTER = 2.32%
LBLDA	Gaborov filter	EER = 1.39% HmTER = 1.27%	EER = 2.29% HmTER = 1.48%
LBLDA	LBLDA	EER = 2.29% HmTER = 1.83%	EER = 2.29% HmTER = 2.29%

Tablica 12: Pogreške verifikacije za različite kombinacije značajki u fuziji na temelju glasanja

Mijenjanjem individualnih pragova na mjere sličnosti može se dobiti veliki raspon rezultata u terminima FAR i FRR. Za svaku kombinaciju metoda izlučivanja značajki na ovaj je način nađen najmanji EER i HmTER. Rezultati eksperimenata prikazani su u tablici 12.

Usporedimo li ove rezultate s rezultatima ostvarenim fuzijom na razini mjere podudaranja, vidi se da u nekim slučajevima fuzija na razini odluke daje bolje, a u nekima lošije rezultate. Najbolja kombinacija značajki u smislu pogreške verifikacije je i dalje LBLDA značajke lica i značajke dlana dobivene na temelju Gaborovog filtra. Uz ovu kombinaciju značajki dobivaju se nešto lošiji rezultati nego fuzijom na razini mjere podudaranja, no još uvijek se radi o manjim pogreškama od onih ostvarenih jednonačinskim sustavom.

U drugom eksperimentu načinjena je fuzija svih šest značajki razmatranih u fuziju, pri čemu se u istom eksperimentu koristi više različitih značajki lica i više različitih značajki dlana. U ovome eksperimentu pragovi na mjere podudaranja za individualne značajke postavljeni su, u dvije varijante eksperimenta, na slijedeći način

varijanta a: pragovi su postavljeni na vrijednosti kojima se dobiva EER, tj. vrijednosti su postavljene tako da su pogreške FAR i FRR individualnih modalnosti izjednačene.

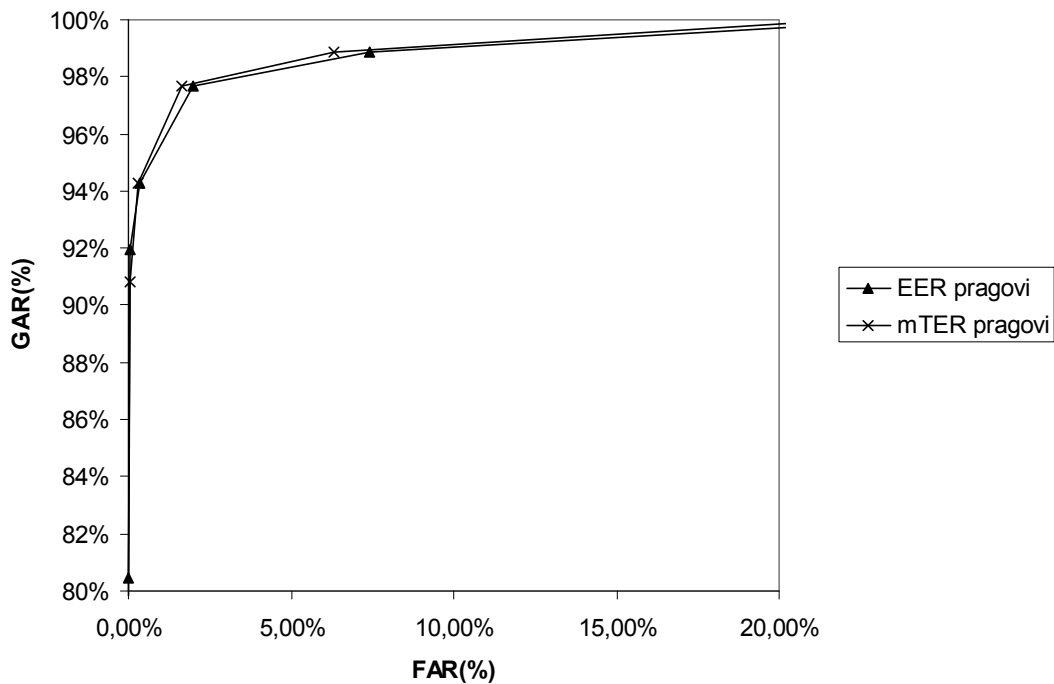
varijanta b: pragovi su postavljeni na vrijednosti kojima se za individualne modalnosti dobiva najmanje ukupna pogreška (mTER)

Načinjeni su eksperimenti za vrijednosti praga odluke $t = 1, 2, \dots, 6$. Rezultati eksperimenta prikazani su u tablici 13, te u obliku ROC krivulje na slici 71.

prag odluke t	Pogreška verifikacije (EER pragovi)	Pogreška verifikacije (mTER pragovi)
1	FRR = 0.00% FAR = 24.20%	FRR = 0% FAR = 22.03%
2	FRR = 1.14% FAR = 7.43%	FRR = 1.14% FAR = 6.29%
3	FRR = 2.29% FAR = 1.99%	FRR = 2.29% FAR = 1.64%
4	FRR = 5.74% FAR = 0.34%	FRR = 5.74% FAR = 0.28%
5	FRR = 8.04% FAR = 0.05%	FRR = 9.19% FAR = 0.02%
6	FRR = 19.54% FAR = 0%	FRR = 20.68% FAR = 0%

Tablica 13: Pogreške verifikacije u eksperimentu fuzije svih značajki lica i dlana na temelju glasanja

Kao što se može vidjeti iz tablice i slike, oba načina određivanja individualnih pragova daju vrlo slične rezultate nakon fuzije. I u ovom eksperimentu fuzijom se dobivaju manje pogreške u odnosu na jednonačinske sustave. Međutim, uključivanjem više od jedne metode izlučivanja značajki dlana i više od jedne metode izlučivanja značajki lica u fuziju nije ostvareno poboljšanje rezultata, što znači da je za optimalni rezultat fuzije dovoljno koristiti samo jednu metodu izlučivanja značajki za dlan i samo jednu za lice.



Slika 71: ROC krivulja za eksperiment fuzije svih značajki lica i dlana na temelju glasanja

10. Zaključak

Dlan i lice su dvije biometrijske karakteristike s velikim brojem značajki pogodnih za korištenje u biometriji. Za razliku od lica kao biometrijske karakteristike, za koje je karakteristično beskontaktno uzimanje uzoraka, većina dosadašnjih sustava temeljenih na dlanu zahtijeva kontakt ruke i senzora što umanjuje prihvatljivost takvih sustava od strane korisnika, naročito u scenarijima u kojima biometrijski sustav koristi veliki broj različitih osoba.

U sklopu disertacije istražena je mogućnost korištenja dlana i lica snimljenih video kamerom u nekontroliranim uvjetima u smislu pozadine i osvjetljenja za biometrijsku verifikaciju. Kako bi se omogućila ovakva beskontaktna verifikacija, razvijene su i evaluirane metode lokalizacije dlana i lica, normalizacije, ocjene kvalitete i izlučivanja značajki. Također su evaluirane metode biometrijske fuzije koje omogućuju smanjenje pogreške verifikacije kombinirajući značajke dlana sa značajkama lica snimljenim u istoj video sekvenci.

Metode lokalizacije koje bi se koristile u ovakvom sustavu moraju raditi u uvjetima nekontrolirane pozadine i promjenjivog osvjetljenja gdje se mogu očekivati drugi objekti boje kože u pozadini, pokretni objekti u pozadini itd. Metode korištene u verifikaciji također moraju biti robusne na promjene u izgledu dlana i lica iste osobe koje će postojati u snimkama načinjenima video kamerom, kao što su razlike u položaju dlana i lica, promjene izraza lica, promjene osvjetljenja itd. Sve metode moraju raditi dovoljno brzo za primjenu u stvarnom vremenu kako bi bilo moguće iskoristiti što više informacija iz video sekvence.

Doprinosi disertacije su

- Metoda za brzu lokalizaciju ruke temeljena na modelima
- Metoda izlučivanja lokalnih binarnih diskriminatornih značajki iz biometrijskih slika
- Metoda za ocjenu kvalitete područja interesa dlana u video sekvencama
- Evaluacija i usporedba metoda izlučivanja značajki dlana i lica dobivenih iz video sekvenci
- Evaluacija metoda biometrijske fuzije na značajkama dlana i lica dobivenim iz video sekvenci

Razvijena metoda lokalizacije dlana kombinira detekciju kandidata ruke postupkom Viola i Jonesa [41] i verifikaciju kandidata postupkom temeljenim na modelima. Korištenje Viola-Jones detektora omogućuje brzu detekciju, a pristup temeljen na modelima eliminaciju lažnih detekcija i približno određivanje konture ruke koja se zatim koristi za lokalizaciju interesnog područja dlana. Eksperimentalno je utvrđena visoka točnost ovakve lokalizacije

dlana od 96.7%. Brzina izvođenja od 0.13 sekundi po okviru u jednodretvenoj implementaciji otvara mogućnost korištenja ovakve lokalizacije na video sekvencama u stvarnom vremenu te predstavlja značajno ubrzanje u odnosu na lokalizaciju koja koristi samo pristup temeljen na modelima [42].

Razvijena je nova metoda izlučivanja značajki, nazvana lokalna binarna LDA (LBLDA), koja nalazi lokalne binarne značajke pogodne za raspoznavanje u smislu da maksimiziraju raspršenje između razreda, a minimiziraju raspršenje unutar istoga razreda. Korištenjem lokalnih binarnih značajki želi se smanjiti osjetljivost verifikacije na promjene u osvjetljenju i druge lokalne promjene. Metoda je evaluirana na bazi video sekvenci dlana i lica koja sadrži video sekvence 87 osoba. Također su, na istoj bazi, evaluirane metode izlučivanja značajki koje se temelje na analizi glavnih komponenti, linearnoj diskriminantnoj analizi te Gaborovom filtru. Eksperimenti su pokazali da lokalna binarna LDA daje bolje rezultate od ostalih evaluiranih metoda za verifikaciju i raspoznavanje lica. Uz pomoć značajki lica dobivenih lokalnom binarnom LDA ostvaren je, u eksperimentu verifikacije, EER = 2.83%, što je poboljšanje u odnosu na rezultate ostvarene pomoću Gaborovog filtra (EER = 4.60%) i „klasične“ LDA (EER = 5.75%).

U eksperimentu verifikacije dlana bolji rezultati verifikacije dobiveni su pomoću „klasične“ LDA (EER = 6.51%) i Gaborovog filtra (EER = 8.05%). Razlozi za bolju uspješnost LDA, kao globalne metode, na dlanu nago za licu mogu se tražiti u manjoj podložnosti dlana lokalnim promjenama. S druge strane, oblik Gaborovog filtra pogodan je za naglašavanje struktura na dlanu, kao što su linije dlana. Premda je u teoriji moguće da LBLDA nakon postupka učenja rezultira lokalnim značajkama koje bi bile identične Gaborovom filtru ukoliko bi se pokazalo da upravo takve lokalne značajke maksimiziraju Fisherov kriterij na skupu za učenje, u ovom slučaju su se značajke koje daje LBLDA pokazale lošije. Stoga se nameće zaključak da je skup za učenje, bilo zbog ekstremnih vrijednosti, bilo zbog nepreciznosti lokalizacije, bilo zbog samo jedne video sekvence za učenje po osobi, u slučaju dlana bio neadekvatan za učenje lokalnih značajki te je zato bilo bolje koristiti značajke predefiniranog oblika.

Razvijena je metoda ocjene kvalitete interesnog područja dlana u video sekvencama kako bi se, umjesto korištenja svih okvira video sekvence za izlučivanje značajki, mogao odabrati podskup okvira iz kojih će se izlučivati značajke pogodne za raspoznavanje. Eksperimenti su pokazali da se na temelju ove metode iz video sekvence mogu odabrati interesna područja dlana pogodna za verifikaciju i time poboljšati točnost verifikacije, ukoliko se kao metode izlučivanja značajki koriste analiza glavnih komponenti ili Gaborov filter. Uz korištenje 60% interesnih područja za izlučivanje značajki koja su odabrana na temelju mjere kvalitete,

pogreška verifikacije za značajke izlučene Gaborovim filtrom smanjuje se s $EER = 8.05\%$ na $EER = 6.08\%$. Pored smanjenja pogreške verifikacije, odbacivanjem dijela interesnih područja na temelju mjere kvalitete također se postiže veća brzina verifikacije jer se smanjuje potreban broj usporedbi vektora značajki iz trenutne video sekvence s vektorima značajki iz baze sustava. Korištenje metode ocjene kvalitete nije umanjilo pogreške verifikacije dobivene na temelju svih metoda izlučivanja značajki, uključujući LBLDA. Općenito, pogodnost korištenja ocjene kvalitete u kombinaciji s nekom novom metodom izlučivanja značajki potrebno je eksperimentalno utvrditi.

U sklopu eksperimenata evaluirane su metode biometrijske fuzije na razini mjere podudaranja i na razini odluke primijenjene na dlanu i licu dobivenim iz iste video sekvence. Eksperimenti su pokazali da se fuzijom može smanjiti pogreška verifikacije u odnosu na verifikaciju koja se temelji samo na licu, odnosno samo na dlanu. Fuzijom na razini mjere podudaranja ostvareni su slični rezultati kao s fuzijom na razini odluke. Najbolji rezultati fuzije ostvareni su kombinirajući značajke lica dobivene pomoću LBLDA i značajke dlana dobivene pomoću Gaborovog filtra. Fuzijom na razini mjere podudaranja ostvaren je rezultat verifikacije s $EER = 1.14\%$ i $HmTER = 0.84\%$ čime je ostvareno smanjenje pogreške verifikacije u odnosu na verifikaciju temeljenu na dlanu ($EER = 6.08\%$) i verifikaciju temeljenu na licu ($EER = 2.83\%$) i pokazana pogodnost značajki dlana i lica iz video sekvenci za primjenu u sustavima za biometrijsku verifikaciju.

Ukoliko se ovi rezultati usporede s rezultatima ostvarenima u postojećim sustavima u kojima se dlan snima bez kontakta sa senzorom [43][44], može se primijetiti da su pojedinačne pogreške lica i dlana lošije u odnosu na pogreške verifikacije dlana u ovim sustavima, što je i očekivano budući da se u njima dlan snima u kontroliranim ili polu-kontroliranim uvjetima te je moguće dobiti kvalitetniju i preciznije lokaliziranu sliku interesnog područja. Međutim, snimanje dlana u nekontroliranim uvjetima otvara mogućnost korištenja lica iz iste video sekvence, a pogreške verifikacije dobivene u eksperimentima fuzije manje su u odnosu na pogreške prijavljene u postojećim sustavima.

Korištenje lokalnih binarnih značajki dalo je obećavajuće rezultate u verifikaciji na temelju značajki lica. Stoga će ovakve značajke biti predmetom i budućeg istraživanja. Konkretno, potrebno je provesti istraživanja vezana za binarizaciju značajki i utvrditi je li moguće koristiti neki drugi prag binarizacije ili neki drugi način kodiranja značajki kojime bi se mogli ostvariti bolji rezultati. Jedan od smjerova također bi mogao biti razmotriti alternative postupku učenja ovakvih značajki kojima bi se možda mogla postići bolja svojstva generalizacije i veća robusnost na pogreške i nepreciznosti u skupu za učenje.

Popis literature

- [1] A. K. Jain, A. Ross, S. Prabhakar, *An Introduction to Biometric Recognition*, IEEE Transactions on Circuits and Systems for Video Technology, Special Issue on Image- and Video-Based Biometrics 14(1), 2004., str. 4-20.
- [2] A. K. Jain, A. Ross, *Multibiometric Systems*, Communications of the ACM, Special Issue on Multimodal Interfaces 47(1), 2004., str. 34-40.
- [3] W. Shu, D. Zhang, *Automated Personal Identification by Palmprint*, Optical Engineering 37(8), 1998., str. 2359-2362.
- [4] *Biometric Market and Industry Report 2006-2010*, International Biometric Group, <http://www.biometricgroup.com/>
- [5] N. Pavešić, I. Fratrić, S. Ribarić, *Degradation of the XM2VTS Database Face Images*, Proc. 2nd COST 275 Workshop, Biometrics on the Internet: Fundamentals, Advances and Applications, str. 17-21, Vigo, Španjolska, ožujak 2004.
- [6] D. Zhang, W. Shu, *Two Novel Characteristics in Palmprint Verification: Datum Point Invariance and Line Feature Matching*, Pattern Recognition 32(4), 1999., str. 691-702.
- [7] N. Duta, A. K. Jain, K. V. Mardia, *Matching of Palmprints*, Pattern Recognition Letters 23(4), 2001., str. 477-485.
- [8] J. You, W. Li, D. Zhang, *Hierarchical Palmprint Identification via Multiple Feature Extraction*, Pattern Recognition 35(4), 2002., str. 847-859.
- [9] C. C. Han, H. L. Cheng, K. C. Fan, C. L. Lin, *Personal Authentication Using Palmprint Features*, Pattern Recognition 36(2), 2003., str. 371-381.
- [10] D. Zhang, W.K. Kong, J. You, M. Wong, *Online Palm Print Identification*, IEEE Transactions on Pattern Analysis and Machine Intelligence 25(2), 2003., str. 1041-1050.
- [11] A. Kong, D. Zhang, M. Kamel, *Palmprint identification using feature-level fusion*, Pattern Recognition, 39(3), 2006., str. 478-487.
- [12] W. Li, D. Zhang, Z. Xu, *Palmprint Identification by Fourier Transform*, International Journal of Pattern Recognition and Artificial Intelligence 16(4), 2002., str. 417-432.

- [13] G. Lu, D. Zhang, K. Wang, *Palmprint Recognition Using Eigenpalms Features*, Pattern Recognition Letters 24(9-10), 2003., str. 1463-1467.
- [14] X. Wu, D. Zhang, K. Wang, *Fisherpalms Based Palmprint Recognition*, Pattern Recognition Letters 24(15), 2003., str. 2829-2838.
- [15] T. Connie, A. Teoh, M. Goh, D. Ngo, *Palmprint recognition with PCA and ICA*, Proc. Image and Vision Computing, str. 227-232, Novi Zeland, studeni 2003.
- [16] S. Ribarić, D. Ribarić, N. Pavešić, *Multimodal Biometric User-identification System for Network-based Applications*, IEE Proceedings on Vision, Image and Signal Processing 150(6), 2003., str. 409-416.
- [17] A. Kumar, D. C. M. Wong, H. Shen, A. K. Jain, *Personal verification using palmprint and hand geometry biometric*, Proc. International Conference on Audio- and Video-based Person Authentication, str. 668-675, Guildford, UK, lipanj 2003.
- [18] A. Kumar, D. Zhang, *Personal authentication using multiple palmprint representation*, Pattern Recognition 38(10), 2005., str.1695-1704.
- [19] S. Ribarić, I. Fratrić, *A Biometric Identification System Based on Eigenpalm and Eigenfinger Features*, IEEE Transactions on Pattern Analysis and Machine Intelligence 27(11), 2005., str.1698-1709.
- [20] S. Ribarić, I. Fratrić, *An Online Biometric Authentication System Based on Eigenfingers and Finger-Geometry*, Proc. 13th European Signal Processing Conference, Antalya, Turska, rujan 2005.
- [21] N. Pavešić, T. Savič, S. Ribarić, I. Fratrić, *A Prototype of a Multimodal Hand-based Verification System with an Aliveness-Detection Module*, Annals of Telecommunications 62(1-2), Special Issue on Multimodal Biometrics, 2007., str. 1655-1680.
- [22] I. Fratrić, S. Ribarić, *Colour-Based Palmprint Verification – an Experiment*, Proc. 14th IEEE Mediterranean Electrotechnical Conference, str. 890-895, Ajaccio, Francuska, svibanj 2008.
- [23] M. Turk, A. Pentland, *Eigenfaces for Recognition*, Journal of Cognitive Neurosience 3(1), 1991., str. 71-86.
- [24] P.N. Belhumeur, J.P. Hespanha, D.J. Kriegman, *Eigenfaces vs. Fisherfaces: Recognition Using Class Specific Linear Projection*, IEEE Transactions on Pattern Analysis and Machine Intelligence 19(7), 1997., str. 711-720.

- [25] D.Q. Dai, P.C. Yuen, *Regularized discriminant analysis and its application to face recognition*, Pattern Recognition 36(3), 2003., str. 845–847.
- [26] H. Yu, J. Yang, *A Direct LDA Algorithm for High-Dimensional Data with Application to Face Recognition*, Pattern Recognition 34(10), 2001., str. 2067–2070.
- [27] H. Cevikalp, M. Neamtu, M. Wilkes, A. Barkana, *Discriminative common vectors for face recognition*, IEEE Transactions on Pattern Analysis and Machine Intelligence 27(1), 2005., str. 4–13.
- [28] T.K. Kim, J. Kittler, *Locally linear discriminant analysis for multimodally distributed classes for face recognition with a single model image*, IEEE Transactions on Pattern Analysis and Machine Intelligence 27(3), 2005., str. 318–327.
- [29] M.S. Bartlett, J.R. Movellan, T.J. Sejnowski, *Face Recognition by Independent Component Analysis*, IEEE Transactions on Neural Networks 13(6), 2002., str. 1450–1464.
- [30] K. Jonsson, J. Kittler, Y.P. Li, J. Matas, *Support vector machines for face authentication*, Image and Vision Computing 20(5-6), 2002., str. 369–375.
- [31] J.E. Meng, W. Shiqian, L. Juwei, L.T. Hock, *Face recognition with radial basis function (RBF) neural networks*, IEEE Transactions on Neural Networks 13(3), 2002., str. 697–710.
- [32] C. Liu, *Gabor-Based Kernel PCA with Fractional Power Polynomial Models for Face Recognition*, IEEE Transactions on Pattern Analysis and Machine Intelligence 26(5), 2004., str. 572–581.
- [33] L. Wiskott, J.M. Fellous, N. Krueger, C. von der Malsburg, *Face Recognition by Elastic Bunch Graph Matching*, poglavlje u knjizi *Intelligent Biometric Techniques in Fingerprint and Face Recognition*, L.C. Jain et al., urednici, str. 355–396, CRC Press, 1999.
- [34] T. Cootes, C. Taylor, H. Kang, V. Petrović, *Modeling Facial Shape and Appearance*, poglavlje u knjizi *Handbook of Face Recognition*, S.Z. Li, A.K. Jain, urednici, str. 39–63, Springer 2004.
- [35] V. Blanz, T. Vetter, *Face Recognition Based on Fitting a 3D Morphable Model*, IEEE Transactions on Pattern Analysis and Machine Intelligence 25(9), 2003., str. 1063–1074.

- [36] R. Chellapa, S. K. Zhou, *Face Tracking and Recognition from Video*, poglavlje u knjizi *Handbook of Face Recognition*, S.Z. Li, A.K. Jain, urednici, str. 169-192, Springer 2004.
- [37] W. Zhao, R. Chellappa, A. Rosenfeld and P.J. Phillips, *Face Recognition: A Literature Survey*, ACM Computing Surveys 25(4), 2003., str. 399-458.
- [38] N. Poh, C.H. Chan, J. Kittler, S. Marcel, C. Mc Cool, E. Argones Rua, J.L. Alba Castro, M. Villegas, R. Paredes, V. Struc, N. Pavesic, A.A. Salah, H. Fang, N. Costen, *Face Video Competition*, Proc. 3rd IAPR/IEEE International Conference on Biometrics, str. 715-724, Alghero, Italija, lipanj 2009.
- [39] A. Erol, G. Bebis, M. Nicolescu, R. Boyle, X. Twombly, *A Review on Vision-Based Full DOF Hand Motion Estimation*, Proc. IEEE Workshop on Vision for Human-Computer Interaction, str. 75-82, San Diego, SAD, lipanj 2005.
- [40] M. Kölsch, M., Turk, *Robust Hand Detection*, Proc. IEEE International Conference on Automatic Face and Gesture Recognition, str. 614-619, Seul, Južna Koreja, svibanj 2004.
- [41] P. Viola, M.J. Jones, *Robust Real-Time Face Detection*, International Journal of Computer Vision 57(2), 2004., str. 137-154.
- [42] B. Stenger, A. Thayananthan, P.H.S. Torr, R. Cipolla, *Model-based hand tracking using a hierarchical Bayesian filter*, IEEE Transactions on Pattern Analysis and Machine Intelligence 28(9), 2006., str. 1372–1384.
- [43] M.G.K. Ong, T. Connie, A.B.J. Teoh, *Touch-less palm print biometrics: Novel design and implementation*, Image Vision Computing 26(12), 2008., str. 1551-1560.
- [44] J. Doublet, O. Lepetit, M. Revenu, *Contact-less Hand Recognition using shape and texture features*, Proc. 8th International Conference on Signal Processing, Guilin, Kina, studeni 2006.
- [45] A. Kumar, D. Zhang, *Integrating Palmprint with Face for User Authentication*, Proc. Multi Modal User Authentication Workshop, str. 107-112, Santa Barbara, SAD, prosinac 2003.
- [46] S. Ribarić, I. Fratrić, K. Kiš, *A Biometric Verification System Based on the Fusion of Palmprint and Face Features*, Proc. 4th International Symposium on Image and Signal Processing and Analysis, str. 12-17, Zagreb, Hrvatska, rujan 2005.

- [47] S. Ribarić, I. Fratrić, K. Kiš, *A Novel Biometric Personal Verification System Based on the Combination of Palmprints and Faces*, Informatica 19(1), 2008., str. 81-100.
- [48] G. Bradski, A. Kaehler, *Learning OpenCV: Computer Vision with the OpenCV Library*, O'Reilly Press, 2008.
- [49] I. Fratrić, S. Ribarić, *Real-Time Model-Based Hand Localization for Unsupervised Palmar Image Acquisition*, Lecture Notes in Computer Science 5558: Advances in Biometrics, 2009., str.1280-1289.
- [50] J. Canny, *A Computational Approach To Edge Detection*, IEEE Transactions on Pattern Analysis and Machine Intelligence 8(6), 1986., str. 679-714.
- [51] M.J. Jones, J.M. Rehg, *Statistical color models with application to skin detection*, International Journal of Computer Vision 46(1), 2002., str. 81-96.
- [52] M. Kass, A. Witkin, D. Terzopoulos, *Snakes: active contour models*, International Journal of Computer Vision 1(4), 1988, str. 259–268.
- [53] R. C. Gonzales, R. E. Woods, *Digital Image Processing*, Addison Wesley Publishing Company, Inc., 1993.
- [54] K. Fukunaga, *Introduction to Statistical Pattern Recognition*, 2. izdanje., Academic Press, New York, 1990.
- [55] R.O. Duda, P.E. Hart, D. G. Stork, *Pattern Classification*, 2. izdanje, John Wiley & Sons, 2001.
- [56] J. Daugman, *High confidence visual recognition of persons by a test of statistical independence*, IEEE Transactions on Pattern Analysis and Machine Intelligence 15(11), 1993., str. 1148-1161.
- [57] Á. Serrano, I.M. de Diego, C. Conde, E. Cabello, *Recent advances in face biometrics with Gabor wavelets: A review*, Pattern Recognition Letters 31(5), 2010., str. 372-381.
- [58] Z. Sun, T., Tan, *Ordinal Measures for Iris Recognition*, IEEE Transactions on Pattern Analysis and Machine Intelligence 31(12), 2009., str. 2211-2226.
- [59] G.C DeAngelis, I. Ohzawa, R.D. Freeman, *Spatiotemporal Organization of Simple-Cell Receptive Fields in the Cat's Striate Cortex, I. General Characteristics and Postnatal Development*, Journal of Neurophysiology 69(4), 1993., str. 1091-1117.

- [60] K. Messer, J. Matas, J. Kittler, J. Luetttin, G. Maitre, *XM2VTSDB: The extended M2VTS database*, Proc. 2nd International Conference on Audio and Video-based Biometric Person Authentication, str. 72-77, Washington, SAD, ožujak 1999.
- [61] A. Ross, A. K. Jain, *Information Fusion in Biometrics*, Pattern Recognition Letters 24(13), 2003., str. 2115-2125.
- [62] A. K. Jain, K. Nandakumar, A. Ross, *Score Normalization in Multimodal Biometric Systems*, Pattern Recognition 38(12), 2005., str. 2270-2285.
- [63] J. Kittler, F.M. Alkoot, *Sum Versus Vote Fusion in Multiple Classifier Systems*, IEEE Transactions on Pattern Analysis and Machine Intelligence 25(1), 2003., str. 110-115.
- [64] S. Ribarić, I. Fratrić, *A Matching-Score Normalization Technique for Multimodal Biometric Systems*, Proc. 3rd COST 275 Workshop: Biometrics on the Internet, str. 55-58, Hatfield, UK, listopad 2005.
- [65] S. Ribarić, I. Fratrić, *Experimental Evaluation of Matching-Score Normalization Techniques on Different Multimodal Biometric Systems*, Proc. 13th IEEE Mediterranean Electrotechnical Conference, str. 498-501, Malaga, Španjolska, svibanj 2006.
- [66] J. Czyz, M. Sadeghi, J. Kittler, L. Vandendorpe, *Decision Fusion for Face Authentication*, Proc. 1st International Conference on Biometric Authentication, str. 686-693, Hong Kong, lipanj 2004.
- [67] S. Prabhakar, A. K. Jain, *Decision-level Fusion in Fingerprint Verification* Pattern Recognition 35(4), 2002., str. 861-874.
- [68] J. Kittler, Y. Li, J. Matas, *On Matching Scores for LDA-based Face Verification*, Proc. 11th British Machine Vision Conference, str. 42-51, University of Bristol, rujan 2000.
- [69] C.J.C. Burges, *A Tutorial on Support Vector Machines for Pattern Recognition*, Data Mining and Knowledge Discovery 2(2), 1998., str. 121–167.
- [70] C.C. Chang, C.J. Lin, *LIBSVM: a library for support vector machines*, 2001., <http://www.csie.ntu.edu.tw/~cjlin/libsvm>

Popis oznaka

GA	Broj ispravno prihvaćenih klijenata (engl. Genuine Acceptance)
GR	Broj ispravno odbijenih uljeza (engl. Genuine Rejection)
FR	Broj pogrešno odbijenih klijenata (engl. False Rejection)
FA	Broj pogrešno prihvaćenih uljeza (engl. False Acceptance)
FAR	Udio pogrešno prihvaćenih uljeza (engl. False Acceptance Rate)
FRR	Udio pogrešno odbijenih klijenata (engl. False Rejection Rate)
EER	Udio jednakih pogrešaka, pogreška kada je $FAR = FRR$
mTER	Minimalni ukupni udio pogrešaka, $FAR + FRR$
HmTER	Polovina minimalnog ukupnog udjela pogrešaka

I	slika
x	x koordinata slike
y	y koordinata slike
II	integralna slika
f_j	j -ta Haarova značajka
h_j	j -ti slabi klasifikator temeljen na jednoj Haarovoj značajki
τ_j	prag j -tog slabog klasifikatora
e_k	pogreška klasifikatora u k -tom koraku
o_i	željeni izlaz klasifikatora za uzorak i
K	broj značajki klasifikatora
N^+	broj uzoraka koji predstavljaju traženi objekt
N^-	broj uzoraka koji ne predstavljaju traženi objekt
$d(A,B)$	udaljenost skupova točaka A i B

$H(l)$	histogramska funkcija $l = 0, 1, 2, \dots, 255$
$f_{H \rightarrow G}(l)$	funkcija koja histogram H transformira u histogram G

\mathbf{x}	vektor značajki
\mathbf{X}	matrica čiji su stupci vektori značajki
\mathbf{m}	srednja vrijednost uzoraka
\mathbf{m}_i	srednja vrijednost uzoraka iz razreda i

POPIS OZNAKA

\mathbf{S}_B	matrica raspršenja između razreda
\mathbf{S}_i	matrica raspršenja razreda i
\mathbf{S}_W	matrica raspršenja unutar razreda
\mathbf{C}	kovarijacijska matrica uzoraka
\mathbf{y}	transformirani vektor značajki
\mathbf{Y}	matrica čiji su stupci transformirani vektori značajki
\mathbf{b}	binarni vektor značajki
\mathbf{w}	bazni vektor linearne transformacije
\mathbf{W}	transformacijska matrica
\mathbf{W}_{PCA}	transformacijska matrica dobivena analizom glavnih komponenti
\mathbf{W}_{LDA}	transformacijska matrica dobivena linearnom diskriminantnom analizom
n	izvorna dimenzionalnost uzoraka
N	broj uzoraka
N_C	broj razreda
N_{PCA}	broj PCA značajki
N_{LDA}	broj LDA značajki
\mathbf{v}	svojstveni vektor
λ	svojstvena vrijednost
ε	rekonstrukcijska pogreška
\mathbf{s}	slika prikazana u obliku vektora čiji su elementi intenziteti slikovnih elemenata slike
\mathbf{a}	slika prikazana u obliku vektora kojemu je oduzeta srednja vrijednost svih uzoraka
\mathbf{A}	matrica čiji su stupci vektori \mathbf{a} pojedinih uzoraka
R_r	područje na slici, $r = 1, 2, \dots, N_R$
N_R	broj područja na koje je slika podijeljena
p	veličina prozora u slikovnim elementima
q	korak translacije prozora
\mathbf{x}_i'	i -ti vektor značajki iz područja R_r
\mathbf{z}_i'	i -ti vektor značajki iz područja R_r nakon smanjenja dimenzionalnosti analizom glavnih komponenti

POPIS OZNAKA

$\mathbf{W}_{\text{PCA}}^r$	transformacijska matrica dobivena analizom glavnih komponenti na području R_r
$\mathbf{W}_{\text{LDA}}^r$	transformacijska matrica dobivena linearnom diskriminantnom analizom na području R_r
\mathbf{w}_j^r	j -ti svojstveni vektor za područje R_r
λ_j^r	j -ta svojstvena vrijednost za područje R_r
$G(x, y, \theta, u, \sigma)$	Gaborov filter s parametrima u, σ, θ
u	frekvencija sinusoide
σ	standardna devijacija
θ	parametar orijentacije
d_{Hamming}	Hammingova udaljenost
M	broj karakteristika koje se kombiniraju u višenačinskom sustavu
r_i	sirova mjera podudaranja za karakteristiku i
n_i	normalizirana mjera podudaranja za karakteristiku i
t	prag verifikacije
t_i	prag verifikacije za karakteristiku i
TSM	ukupna mjera podudaranja (engl. Total Similarity Measure)

Prilog A: Programska implementacija

U sklopu rada na disertaciji razvijene su aplikacije za lokalizaciju ruke i lica u video sekvencama, ocjenu kvalitete područja interesa dlana te biblioteka za izlučivanje značajki i evaluaciju. U nastavku ovoga priloga pojedinačno su opisane aplikacije i biblioteka te korišteni formati zapisa podataka.

Aplikacije i biblioteka implementirani su u programskom jeziku C++. Na priloženom DVD-u nalazi se izvorni kod te izvršne verzije aplikacija kompajlirane za Microsoft Windows operacijski sustav.

A.1. Aplikacija za lokalizaciju lica i ruke

Aplikacija za lokalizaciju lica i ruke obrađuje video sekvencu ili niz slika i kao izlaz vraća skup normaliziranih područja interesa lica i dlana. Aplikacija se poziva kao

```
localize.exe [v|i] <ulazna datoteka> <izlazna datoteka>
```

Ukoliko je prvi parametar 'v', tada ulazna datoteka sadrži video sekvencu, a ukoliko je prvi parametar 'i', tada je ulazna datoteka tekstualna datoteka u kojoj je svaki redak put do jedne statične slike. Primjer takve tekstualne datoteke koja sadrži niz od pet slika je

```
F:\baza\slike\slika001.bmp  
F:\baza\slike\slika002.bmp  
F:\baza\slike\slika003.bmp  
F:\baza\slike\slika004.bmp  
F:\baza\slike\slika005.bmp
```

Slike mogu biti u slijedećim nekomprimiranim formatima zapisa:

- .bmp (bitmap, podržan je samo nekomprimirani oblik datoteke)
- .pgm (portable gray map)
- .ppm (portable pixel map)
- .raw (format zapisa koji ne sadrži zaglavlja već samo sirove podatke o intenzitetu pojedinih slikovnih elemenata)

Ukoliko je ulaz video sekvenca, ona može biti u bilo kojem formatu za koji je na računalu instaliran DirectShow-kompatibilan dekodir.

Izlazna datoteka je u obliku formata datoteke za zapis značajki. Ovaj format datoteke opisan je u slijedećem poglavlju.

Primjer poziva aplikacije za lokalizaciju lica i ruke je

```
localize.exe v video.avi roi.f
```

U ovome primjeru aplikacija će lokalizirati lice i dlan u video sekvenci u datoteci 'video.avi', a dobivena normalizirana područja interesa biti će spremljena u datoteku 'roi.f'.

A.2. Format datoteke za zapis značajki

Zbog lakšeg zapisa velikog broja vektora značajki koji se mogu dobiti iz jedne video sekvence stvoren je poseban format zapisa značajki pomoću kojega se u jednoj datoteci mogu zapisati sve značajke izlučene iz jedne video sekvence, uz informacije o tipu značajki i o samom uzorku.

Na početku datoteke nalazi se zaglavlje slijedećeg formata

Ime	Tip	Veličina	Opis
Potpis	Niz znakova	4 okteta	4 okteta koja uvijek trebaju imati ascii vrijednost 'FEAT'
Oznaka razreda	Niz znakova	20 okteta	Niz znakova koji jednoznačno označava osobu kojoj uzorak pripada
Oznaka uzorka	Niz znakova	20 okteta	Niz znakova koji jednoznačno označava uzorak unutar određenog razreda (npr. redni broj uzorka)
Broj predložaka	32-bitni broj	4 okteta	Broj predložaka. Svaki predložak može predstavljati jedan vektor značajki

Tablica A1: Format zapisa zaglavlja datoteka za zapis biometrijskih značajki

Odmah nakon zaglavlja slijedi polje koje sadrži informacije o svakom predlošku u datoteci. Broj zapisa ovog polja određen je zaglavljem, a svaki zapis je slijedećeg formata.

Ime	Tip	Veličina	Opis
Tip	Niz znakova	10 okteta	Niz znakova koji jednoznačno označava tip biometrijske značajke. Npr. vrijednost 'PALM-RAW' označava da zapis sadrži sirove vrijednosti intenziteta iz normalizirane slike dlana. Vrijednost 'FACE-RAW' označava da zapis sadrži sirove vrijednosti intenziteta iz normalizirane slike lica.
Veličina	32-bitni broj	4 okteta	Veličina predloška za ovaj tip značajki u oktetima.
Pomak	32-bitni broj	4 okteta	Pomak u oktetima od početka datoteke do položaja gdje je zapisan biometrijski predložak koji odgovara navedenom tipu značajke

Tablica A2: Format zapisa opisnog polja datoteka za čuvanje biometrijskih značajki

Nakon ovoga polja slijede zapisi pravih biometrijskih predložaka, koji mogu biti vektori značajki ili neki drugi zapisi. Veličina zapisa i njihov položaj u datoteci određeni su poljem iz tablice A2.

A.3. Aplikacija za ocjenu mjere kvalitete interesnog područja dlana

Aplikacija za ocjenu mjere kvalitete interesnog područja dlana analizira skup interesnih područja dlana te vraća određeni postotak izvornih interesnih područja koja imaju najveću mjeru kvalitete. Aplikacija se poziva kao

```
quality.exe <ulazna datoteka> <izlazna datoteka> <postotak>
```

I ulazna i izlazna datoteka su u formatu datoteke za zapis značajki (format datoteke opisan u prethodnom poglavlju). Svako od interesnih područja dlana je jedan od zapisa u ovoj datoteci s tipom postavljenim na vrijednost 'PALM-RAW'. Argument `postotak` određuje koliki će postotak ukupnog broja interesnih područja dlana biti zadržan u izlaznoj datoteci.

Primjer poziva aplikacije je

```
quality.exe 'svi_roi.f' 'odabrani_roi.f' 40
```


U ovom primjeru aplikacija će analizirati interesna područja dlana pohranjena u datoteci 'svi_roi.f', odabrati će 40% interesnih područja s najvećom mjerom kvalitete i njih će spremiti u datoteku 'odabrani_roi.f'.

A.4. Biblioteka za izlučivanje značajki i evaluaciju

Biblioteka za izlučivanje značajki i evaluaciju sadrži klase za ulazno/izlazne operacije na podacima, učenje i primjenu metoda izlučivanja značajki te provođenje eksperimenata. Najvažnije klase u ovoj biblioteci navedene su u nastavku.

Image – ova klasa sadrži podatke o slici te osnovne metode za manipulaciju slikom. Najvažnije funkcije su `GetPixelGray` i `GetPixelColor` koje vraćaju vrijednosti pojedinog slikovnog elementa, te funkcije `SetPixelGray` i `SetPixelColor` koje postavljaju vrijednosti pojedinog slikovnog elementa.

ImageIO – ova klasa sadrži metode za čitanje i pisanje slika u datoteke u nekoliko nekomprimiranih formata slikovnog zapisa: .bmp, .ppm, .pgm i .raw. Najvažnije metode su `LoadImage` za učitavanje slike, te `SaveImage` za spremanje slike u datoteku.

FeatureFile – klasa koja služi za učitavanje i spremanje datoteka u formatu datoteka za zapis značajki. Ovaj format datoteka opisan je u poglavlju A.2.

Sample – ova klasa sadrži vektor značajki za jedan uzorak. Za svaki uzorak, osim vektora značajki, pamti se ime datoteke za čitanje i spremanje te ime razreda kojemu uzorak pripada. Klasa sadrži funkcije za učitavanje vektora značajki iz datoteka i spremanje u datoteku.

SampleSet – ova klasa sadrži kolekciju više vektora značajki koji mogu pripadati različitim osobama. Također su implementirane funkcije za čitanje iz datoteke. Ova klasa podržava i neke naprednije operacije nad skupom uzoraka kao što su pretvaranje u matricu, računanje srednje vrijednosti uzoraka, računanje matrica raspršenja unutar razreda i između razreda itd.

Matrix – implementira matricu i osnovne operacije nad matricama. Također sadrži funkcije za pisanje matrice u datoteku i čitanje iz datoteke.

SubspaceGenerator – apstraktna klasa koja služi za generiranje (globalnog) potprostora na temelju skupa uzoraka. Jedina funkcija u klasi je `GenerateSubspace` čijim se pozivom pokreće učenje potprostora za različite transformacije.

PCASubspaceGenerator – nasljeđuje `SubspaceGenerator`, služi za generiranje potprostora analizom glavnih komponenti.

LDASubspaceGenerator – nasljeđuje `SubspaceGenerator`, služi za generiranje potprostora linearnom diskriminantnom analizom.

Subspace – sadrži podatke o (globalnom) potprostoru koji generira `SubspaceGenerator`. Također sadrži metode za zapis potprostora u datoteku i čitanje potprostora iz datoteke.

SubspaceProjector – ova klasa sadrži metode za projiciranje uzorka ili skupa uzoraka u neki potprostor. Najvažnije metode su `ProjectSample` koja projicira jedan vektor značajki u zadani potprostor i `ProjectSampleSet` koja projicira skup vektora značajki u zadani potprostor.

LocalSubspaceGenerator – klasa koja služi za generiranje lokalnog potprostora na temelju lokalne binarne LDA. Najvažnija funkcija je `GenerateSubspace` koja računa potprostor na temelju skupa uzoraka za učenje.

LocalSubspace – sadrži podatke o lokalnom potprostoru koji je generirao `LocalSubspaceGenerator`. Također sadrži metode za zapis ovog potprostora u datoteku i čitanje iz datoteke.

LocalSubspaceProjector – ova klasa sadrži metode za dobivanje vektora (lokalnih) značajki na temelju lokalnog potprostora koji je izračunat na temelju lokalne binarne LDA.

GaborFilter – implementira izlučivanje značajki na temelju Gaborovog filtra. Moguće je generirati filter na temelju parametara (funkcija `ConstructFilter`) i primijeniti filter na sliku (funkcija `FilterImage`).

OneNNClassifier – implementacija klasifikatora na temelju najmanje udaljenosti.

Podržane su različite mjere podudaranja:

- Euklidska udaljenost
- Normalizirana korelacija
- Hammingova udaljenost

Klasa sadrži funkcije za provođenje eksperimenata raspoznavanja (funkcija `classificationTest`) i verifikacije (funkcija `verificationTest`).

Životopis

Ivan Fratrić rođen je 28. svibnja 1980. godine u Zagrebu. Osnovnu i srednju školu završio je u Slavonskom Brodu. 1998. godine upisao je Fakultet elektrotehnike i računarstva Sveučilišta u Zagrebu. Diplomirao je 2003. godine u okviru završetka studija s naglaskom na znanstveno-istraživačkom radu. Magistrirao je 2007. godine na istom fakultetu. Radio je kao suradnik na Zavodu za elektroniku, mikroelektroniku, računalne i inteligentne sustave na Fakultetu elektrotehnike i računarstva Sveučilišta u Zagrebu te kao softver inženjer u Ericsson Nikola Tesla, d.d. Od lipnja 2004. godine zaposlen je kao znanstveni novak na Zavodu za elektroniku, mikroelektroniku, računalne i inteligentne sustave na Fakultetu elektrotehnike i računarstva Sveučilišta u Zagrebu. Njegovi istraživački interesi obuhvaćaju raspoznavanje uzorka, računalni vid i umjetnu inteligenciju. U koautorstvu je objavio sedam radova na međunarodnim znanstvenim skupovima i tri rada u časopisima indeksiranim u bazi CC.

Biography

Ivan Fratrić was born on May, 28th, 1980 in Zagreb. He finished elementary school and high school in Slavonski Brod. He entered Faculty of Electrical Engineering and Computing at the University of Zagreb in 1998. He graduated in 2003 with the emphasis on scientific and research work. He got his masters degree in 2007 from the same faculty. He worked as research associate at the Department of Electronics, Microelectronics, Computer and Intelligent Systems at the Faculty of Electrical Engineering and Computing at the University of Zagreb and as a software engineer in Ericsson Nikola Tesla, d.d. From June 2004 he works as a researcher at the Department of Electronics, Microelectronics, Computer and Intelligent Systems at the Faculty of Electrical Engineering and Computing at the University of Zagreb. His research interests include pattern recognition, computer vision and machine intelligence. As a co-author he published seven papers at the international scientific gatherings and three papers in the journals indexed in the CC database.