

UNIVERZA V LJUBLJANI  
FAKULTETA ZA RAČUNALNIŠTVO IN INFORMATIKO

Metod Ribič

**Vpliv poravnave na uspešnost  
razpoznavanja uhljev**

DIPLOMSKO DELO

UNIVERZITETNI ŠTUDIJSKI PROGRAM  
PRVE STOPNJE  
RAČUNALNIŠTVO IN INFORMATIKA

MENTOR: izr. prof. dr. Peter Peer

SOMENTOR: doc. dr. Vitomir Štruc

Ljubljana, 2016

*Besedilo je oblikovano z urejevalnikom besedil L<sup>A</sup>T<sub>E</sub>X.*

Fakulteta za računalništvo in informatiko izdaja naslednjo nalogo:

Vpliv poravnave na uspešnost razpozavanja uhljev.

Tematika naloge:

Poravnava slik pred razpoznavo ima dokazan vpliv na njeno učinkovitost. Ker pa danes v biometriji težimo k uporabi slik, ki so zajete v nekontroliranem okolju, se poraja vprašanje vpliva poravnave v takšnih primerih. Razširite orodje AWE z implementacijo poravnave na podlagi soglasja naključnega vzorca ter kaskadne pozicijske regresije. Ovrednotite vpliv poravnave na bazah, ki sta zajeti v kontroliranem in nekontroliranem okolju. Argumentirajte sklepe.



*Iskreno se zahvaljujem mentorju izr. prof. dr. Petru Peeru, somentorju doc. dr. Vitomirju Štrucu in asistentu Žigu Emeršiču za vso potrpežljivost in pomoč pri izdelavi diplomske naloge. Prav tako bi se rad zahvalil celotni družini in puncu Nini, ki sta mi v času študija stali ob strani ter verjeli vame.*



# Kazalo

**Povzetek**

**Abstract**

<b>1</b>	<b>Uvod</b>	<b>1</b>
1.1	Motivacija . . . . .	1
1.2	Pregled področja . . . . .	1
1.3	Cilji . . . . .	4
<b>2</b>	<b>Orodja in knjižnice</b>	<b>5</b>
2.1	Matlab . . . . .	5
2.2	VLFeat . . . . .	6
2.3	Piotrovo Matlab orodje za računalniški vid . . . . .	6
2.4	Orodje AWE . . . . .	6
<b>3</b>	<b>Opis metod in postopkov</b>	<b>7</b>
3.1	Soglasje naključnega vzorca (RANSAC) . . . . .	7
3.2	Kaskadna pozicijska regresija (CPR) . . . . .	9
<b>4</b>	<b>Eksperimenti</b>	<b>13</b>
4.1	Soglasje naključnega vzorca (RANSAC) . . . . .	13
4.2	Kaskadna pozicijska regresija (CPR) . . . . .	16
4.3	Vrednotenje . . . . .	18

<b>5</b>	<b>Rezultati z diskusijo</b>	<b>21</b>
5.1	Soglasje naključnega vzorca (RANSAC) . . . . .	21
5.2	Kaskadna pozicijska regresija (CPR) . . . . .	25
5.3	Primerjava rezultatov z diskusijo . . . . .	30
<b>6</b>	<b>Zaključek</b>	<b>33</b>
	<b>Literatura</b>	<b>34</b>



# Seznam uporabljenih kratic

kratica	angleško	slovensko
<b>RANSAC</b>	RANdom SAMple Consensus	soglasje naključnega vzorca
<b>SIFT</b>	Scale-Invariant Feature Transform	skalno-invariantna transformacija značilk
<b>AWE</b>	Annotated Web Ears	označeni spletni uhlji
<b>IITD</b>	Indian Institute of Technology Delhi	Indijski tehnološki inštitut v Delhiju
<b>CPR</b>	Cascaded Pose Regression	kaskadna pozicijska regresija
<b>CMC</b>	Cumulative Match Curve	krivulja komulativnih zadetkov
<b>BSIF</b>	Binarized Statistical Image Features	binarizirane statične značilke slik
<b>LPQ</b>	Local Phase Quantization	lokalna fazna kvantiacija
<b>RILPQ</b>	Rotation-Invariant Local Phase Quantization	rotacijsko invariantna lokalna fazna kvantiacija
<b>POEM</b>	Patterns of Oriented Edge Magnitudes	vzorci orientiranih robnih magnitud
<b>HOG</b>	Histograms of Oriented Gradients	histogrami orientiranih gradientov
<b>DSIFT</b>	Dense Scale-Invariant Feature Transform	gosta skalno invariantna transformacija značilk



# Povzetek

**Naslov:** Vpliv poravnave na uspešnost razpoznavanja uhljev

Uhlji kot biometrična modalnost so postali pomemben vir za samodejno razpoznavo oseb, predvsem v scenarijih nadzornih aplikacij, kjer se obraz ne vidi frontalno. V preteklih letih je priljubljenost metod lokalnih deskriptorjev narasla zaradi njihove invariantnosti na osvetlitev in zakrivanje, vendar pa te metode za vhodne podatke zahtevajo poravnane in vnaprej obdelane slike, kar pa lahko predstavlja velik problem zaradi kota, pod katerim so bile slike zajete. V tem diplomskem delu smo testirali, kako poravnava uhljev z metodo soglasja naključnega vzorca (RANSAC) in metodo kaskadne pozicijske regresije (CPR) vpliva na razpoznavo oseb na podlagi uhljev. Poravnava se je izvajala na podatkovni bazi uhljev AWE. Dokazali smo, da poravnava uhljev pozitivno vpliva na razpoznavo samo v primeru, ko so obravnavane slike zajete pod majhnim nagibom in odklonom. Slike, zajete pod večjim nagibom in odklonom, pa so po naših ugotovitvah prezahtevne za poravnavo in poslabšajo rezultate razpoznavne, zaradi česar bi bilo potrebno uporabiti drugačne pristope za poravnavo.

**Ključne besede:** računalniški vid, biometrija, razpoznavanje, poravnava, uhlji.



# Abstract

**Title:** Influence of alignment on ear recognition

Ear as a biometric modality presents a viable source for automatic human recognition especially in surveillance scenarios where face is not seen frontally. In recent years local description methods have been gaining on popularity due to their invariance to illumination and occlusion. However, these methods require that images are well aligned and preprocessed as good as possible. This causes one of the greatest challenges of ear recognition: sensitivity to pose variations. In this paper we test the influence of alignment on recognition performance on images from recently presented Annotated Web Ears dataset with alignment methods Random sample consensus (RANSAC) and Cascaded Pose Regression (CPR). We prove that alignment improves recognition rate but only on images with small angle on roll and yaw axis. On other pictures RANSAC and CPR fails to align ears and recognition rate is therefore lower versus unaligned pictures. Those pictures should be addressed with more advanced alignment methods in order to improve recognition rate.

**Keywords:** computer vision, biometry, recognition, alignment, ears.



# Poglavje 1

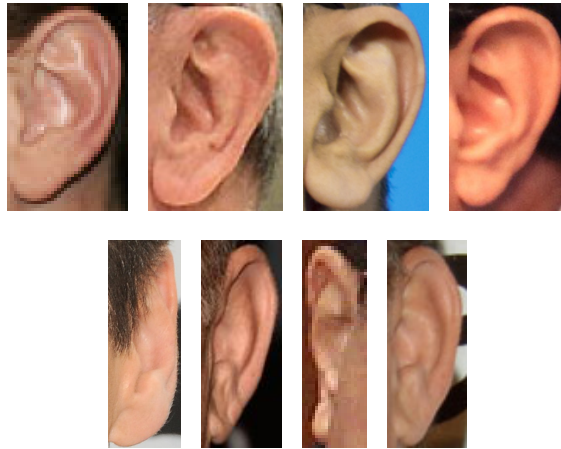
## Uvod

### 1.1 Motivacija

Potreba po razpoznavanju ljudi je v zadnjem desetletju pripeljala do odkritja novih tehnik na področju biometrije. Uhelj kot biometrični način predstavitve predstavlja obetaven primer, saj ima visoko stopnjo stalnosti, prepoznavnosti in univerzalnosti [2], vendar pa je eno izmed najtežjih, še vedno odprtih vprašanj raznolikost kota (angl. pose variation), pod katerim je bila slika uhlja zajeta – to je prikazano na sliki 1.1. Prvi korak je torej normalizacija kota in segmentacija vhodnih podatkov. V diplomski sta predstavljena koncept poravnave uhljev z metodo soglasja naključnega vzorca (angl. RANdom SAmple Consensus – RANSAC) [7] v povezavi s povprečnim uhljem in poravnava z metodo kaskadne pozicijske regresije (angl. Cascaded Pose Regression – CPR) [4].

### 1.2 Pregled področja

Na področju razpoznavanja ljudi ob pomoči uhljev je bilo v zadnjem desetletju veliko storjenega, od razvitja novih metod deskriptorjev do normalizacije in segmentacije. Nove metode so pripomogle k temu, da je skoraj vsako sliko mogoče poravnati in iz nje razpoznati objekt ter posledično tudi osebo, če je



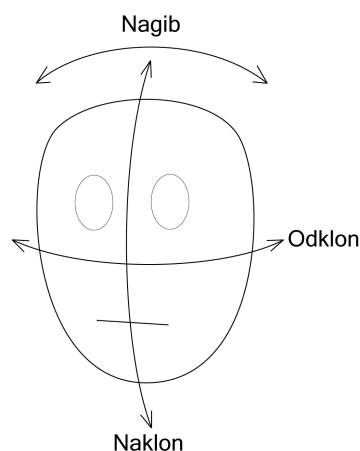
Slika 1.1: Primeri uhljev, zajetih iz različnih zornih kotov. Slike v prvi vrsti prikazujejo uhlje, ki so ugodni za poravnavo, slike v spodnji vrsti pa predstavljajo primere uhljev, ki so zelo problematični za poravnavo.

ta na sliki. Kljub temu pa je raznolikost kota, pod katerim je bila slika uhlja zajeta, še vedno odprt problem. V vseh dosedanjih raziskavah dvodimenzionalne slike poravnavaajo samo po dveh oseh in sicer po osi naklona in nagiba, medtem ko pri rotaciji po osi odklona izgubimo preveč podatkov zaradi zakrivanja. Vse tri osi so prikazane na sliki 1.2. Poleg zakrivanja pa je pri obstoječih delih [8, 11, 14, 9] možno opaziti, da je bila večina slikovnih baz, ki vsebujejo uhlje, zajetih v nadzorovanih laboratorijskih razmerah. Zaradi tega imajo slike kar nekaj skupnih karakteristik, kot so osvetlitev, parametri kamere in kot, pod katerim so bile slike zajete.

### 1.2.1 Baze uhljev

Na voljo je kar nekaj baz uhljev [6], na katerih je mogoče testirati biometrične tehnike, ki temeljijo na uhljih. Kljub temu pa je večina slik v teh bazah zajetih v nadzorovanih razmerah [1], kar pomeni, da imajo vse slike podobne karakteristike, kar olajša postopek razpoznave. V nenadzorovanih razmerah je skoraj nemogoče pridobiti sliko primerne ločljivosti, ki bi bila zajeta pod idealnim kotom, saj je uhelj na človeku velikokrat zakrit z lasmi,





Slika 1.2: Prikaz prostorskih stopenj/osi nagib-naklon-odklon (angl. roll-pitch-yaw) na modelu človeške glave.

vsebuje dodatke ali pa je preveč popačen za nadaljnjo obdelavo. Zaradi teh spoznanj smo se odločili, da za potrditev hipoteze, v kateri normalizacija in segmentacija izboljšata postopek razpoznavne, uporabimo bazo označenih spletnih uhljev AWE (angl. Annotated Web Ears database). Ta baza vsebuje slike iz "divjine" (angl. "in the wild"), kar pomeni, da nimajo enakih karakteristik, zato so bližje realnemu svetu in hkrati veliko bolj zahtevne za procesiranje. Za potrditev delovanja naše implementacije metode CPR, kot je to opisano v delu A. Pflug [7], smo za testno bazo uporabili bazo indijskega tehnološkega inštituta v Delhiju – IITD I ([http://www4.comp.polyu.edu.hk/~csajaykr/IITD/Database\\_Ear.htm](http://www4.comp.polyu.edu.hk/~csajaykr/IITD/Database_Ear.htm)), ki je lažja za poravnavo, saj vsebuje le stranske slike glave, torej ima večina uhljev majhen odklon. Baza IITD I vsebuje 493 slik desnih uhljev, ki pripadajo 125-im različnim osebam in so bile zajete v laboratorijskem okolju. Povprečna velikost slike v bazi IITD I je  $272 \times 204$  slikovnih pik. Baza uhljev AWE vsebuje 1000 slik; med njimi 520 slik z levim uhljem in 480 slik z desnim uhljem, pripadajo pa 100-tim osebam. Povprečna velikost slike v bazi AWE je  $83 \times 160$  slikovnih točk [6]. Vsaki osebi pripada 10 slik, ki imajo zabeležene naslednje lastnosti:

1. koordinati tragus točke,
2. dimenziji slike,
3. podatek, ali uhelj vsebuje nakit,
4. podatek, ali je uhelj delno zakrit,
5. naklon,
6. odklon,
7. nagib.

### **1.3 Cilji**

Cilj diplomske naloge je z različnimi tehnikami dokazati, da poravnava uhljev pozitivno vpliva na razpoznavo oseb, in izpostaviti tiste karakteristike slik in uhljev, ki negativno vplivajo na poravnavo ter razpoznavo. Na koncu želimo ponuditi v uporabo pridobljeno poravnano bazo tistih uhljev iz slikovne baze AWE, ki jih je mogoče uspešno poravnati, skupaj s pripadajočimi rezultati v primerjavi z neporavnano bazo uhljev. Težavnost baze AWE bo pokazala tudi na potrebo po novih postopkih poravnave.

# Poglavje 2

## Orodja in knjižnice

Vsa orodja in knjižnice, uporabljeni v diplomski, so odprtokodni, z izjemo orodja Matlab.

### 2.1 Matlab

Diplomska naloga je v celoti izdelana v programskem jeziku Matlab (<https://www.mathworks.com/products/matlab/>). To je orodje, ki se uporablja za numerično analizo, analizo signalov, matematično modeliranje in analizo slik; z njim je mogoče izdelati grafične vmesnike ter poganjati programe, spisane v programskih jezikih C, C++, Fortranu, Javi in Pythonu. Matlab je spisan je v programskih jezikih C in Java in ga je mogoče poganjati v vseh večjih distribucijah operacijskih sistemih. Glavna prednost orodja je v optimizaciji matričnih operacij. Poleg osnovnega paketa Matlab smo uporabili tudi Matlab orodja za procesiranje slik (angl. Image Processing toolbox) in za računalniški vid (angl. Computer Vision System toolbox). Obe služita za obdelovanje slik in videa.

## 2.2 VLFeat

VLFeat [13] je odprtokodna knjižnica za orodje Matlab. Vsebuje implementacije znanih metod računalniškega vida, ki so specializirane in optimizirane za razumevanje slik ter pridobivanje in povezovanje lokalnih značilk. Za potrebe diplome smo uporabili implementacijo za skalno invariantno transformacijo značilk (angl. Scale Invariant Feature Transform – SIFT) [10], ki je dostopna na spletnem naslovu <http://www.vlfeat.org/overview/sift.html>.

## 2.3 Piotrovo Matlab orodje za računalniški vid

Piotrovo Matlab orodje za računalniški vid (angl. Piotr’s Computer Vision Matlab toolbox) [3] avtorja Piotra Dollárja je optimizirano za manipulacijo slik in videa z orodjem Matlab. Kaskadna pozicijska regresija (CPR), ki je del orodja, oceni usmeritev in pozicijo objekta na sliki na podlagi označene učne množice. Piotrovo orodje je skupaj s knjižnico CPR na voljo na spletni strani <https://pdollar.github.io/>.

## 2.4 Orodje AWE

Matlab orodje označenih spletnih uhljev AWE (angl. Annotated Web Ears toolbox) [5] se uporablja za razpoznavo oseb na podlagi uhljev in je delo Laboratorija za računalniški vid na Fakulteti za računalništvo in informatiko v Ljubljani. Orodje AWE je prosto za uporabo in je skupaj z bazo uhljev AWE na voljo na spletni strani <http://awe.fri.uni-lj.si/>.

## Poglavje 3

# Opis metod in postopkov

### 3.1 Soglasje naključnega vzorca (RANSAC)

#### 3.1.1 Predhodna obdelava

Zaradi poznejših omejitev v orodju AWE in delovanja soglasja naključnega vzorca smo morali pred obdelavo vse dimenzije slik pomnožiti s faktorjem  $n = 100/v$ , pri čemer je  $v$  višina slike. S tem smo dosegli, da so vse slike visoke 100 slikovnih pik in da se razmerje višine in širine ohrani.

#### 3.1.2 Izračun referenčnega uhlja

Vsi uhlji v postopku poravnave z metodo soglasja naključnega vzorca se poravnavaajo na en referenčni uhelj, ki pa ne sme biti kar naključno izbran iz baze slik, saj se uhlji med seboj zelo razlikujejo. Če bi za referenčni uhelj izbrali naključni uhelj iz vhodne baze, bi bile transformacije slik v veliki meri napačne, kar bi pozneje zmanjšalo uspešnost razpoznave. Za referenčni uhelj prav tako ne smemo vzeti najbolj poravnane uhlja ene osebe in vseh ostalih slik te osebe poravnati na ta uhelj, saj bi tako v postopek vnesli informacijo, kar je metodološko sporno. Ker je baza uhljev AWE označena, smo lahko iz slik, ki so označene kot popolnoma poravnane, izdelali referenčni uhelj. Vsem slikam, ki so v bazi AWE označene kot popolnoma poravnane, torej imajo



Slika 3.1: Na prvi sliki je prikazan rezultat seštevanja istoležnih slikovnih pik vseh uhljev, ki so označeni kot poravnani, sledi obrezan uhlj, ki služi kot referenčni uhlj v postopku poravnave. Tretja slika prikazuje binarno masko referenčnega uhlja, ki je bila ročno orisana, zadnja pa obrezan referenčni uhlj z binarno masko.

zanemarljiv kot na oseh nagib-naklon-odklon (angl. roll-pitch-yaw), smo sešteli istoležne slikovne pike in jih na koncu delili s številom uporabljenih slik. Ker slike niso enakih velikosti, smo uporabili enak postopek, kot je opisan v podpoglavju 3.1.1, pri čemer smo vse slike po širini dopolnili s črnimi slikovnimi pikami. S tem smo dosegli, da so bile vse slike široke enako kot najširša slika po predhodni obdelavi. Ker je večina slik ožjih od najširše slike, sta po seštevanju istoležnih slikovnih pik na končni sliki ob straneh prisotni črnini. Poleg črnine so na sliki vidne tudi vertikalne črte, ki so prav tako posledica različnih dimenzij slik. Ker so črnini in vertikalne črte motile postopek poravnave, je bil referenčni uhlj obrezan na najbolj notranjo vertikalno črto. Rezultati seštevanja, obrezan referenčni uhlj in uporabljena binarna maska so vidni na sliki 3.1.

### 3.1.3 Skalno-invariantna transformacija značilk (SIFT)

Preden metoda RANSAC oceni transformacijo, je potrebno iz vhodne slike in referenčnega uhlja pridobiti lokalne deskriptorje in njim pripadajoče točke. Za to poskrbi metoda skalno-invariantne transformacije značilk [10].

### 3.1.4 Soglasje naključnega vzorca (RANSAC)

RANSAC iterativno povezuje pare točk/deskriptorje, ki jih je izločil SIFT, in jih poskuša postaviti v navidezno ravnino. Metoda v vsaki izmed  $t$  iteracij naključno izbere  $n$  parov točk ( $n$  je odvisen od vrste transformacije) in iz njih oceni transformacijo [7]. Večje kot je število iteracij, bolj natančna je transformacija, vendar hkrati porabimo več virov.

### 3.1.5 Poravnava

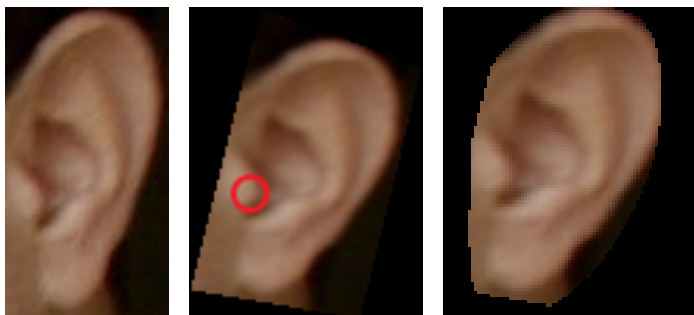
Ob pomoči transformacijske matrike, pridobljene z metodo RANSAC, smo nato transformirali vhodno sliko. Rezultat je slika, na kateri je vhodni uhelj poravnan enako, kot je poravnan povprečni uhelj. Če metoda RANSAC ni uspela oceniti transformacijske matrike, je slika po transformaciji močno popačena, zaradi česar smo jo izločili iz nadaljnjega postopka.

### 3.1.6 Maskiranje

Da smo se izognili šumom, ki nastanejo pri transformaciji, smo uhelj glede na točko tragus obrezali z binarno masko, ki je bila ročno določena na referenčnem uhlju. Poravnava, točka tragus in uporaba binarne maske so prikazani na sliki 3.2, referenčni uhelj pa na sliki 3.1.

## 3.2 Kaskadna pozicijska regresija (CPR)

Kaskadna pozicijska regresija je metoda, ki iz označene učne množice oceni orientacijo objekta na sliki in ga označi z vnaprej definiranim modelom – to je bila v našem primeru elipsa. Za izračun/ocenitev transformacijske matrike metoda CPR uporablja zgolj informacije lokalne osvetlitve med deskriptorjem in naučenim modelom [11]. Avtor metode je v članku [4] predstavil delovanje metode na modelu človeškega obraza in nekaj živalih (ribah, kuščarjih in miših). Ravno zaradi uspešne ocenitve transformacij poze ne glede na raznolikost testnih množic smo se odločili metodo preizkusiti na uhljih.



Slika 3.2: Tri stopnje v postopku poravnave. Prva slika je vhodna, druga prikazuje rezultat poravnave, tretja pa rezultat, obrezan z binarno masko. Ta je bila ročno orisana na referenčnem uhlju in uporabljena na vseh poravnanih uhljih. Na drugi sliki je z rdečim krogom označena točka tragus, glede na katero se poravnan uhelj obreže z binarno masko.

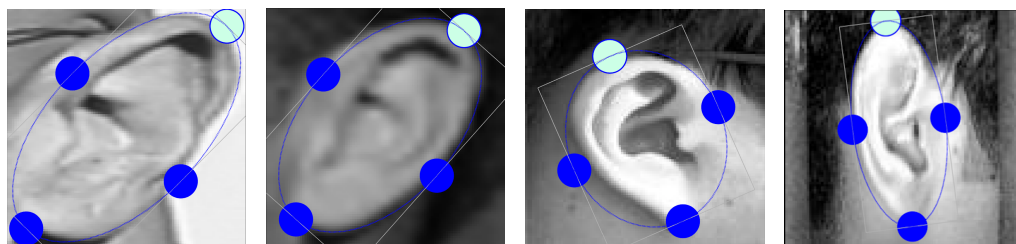
### 3.2.1 Učna množica in označevanje

Metoda CPR potrebuje za uspešno delovanje vnaprej definirano učno množico, na kateri ročno označimo uhlje z modelom – elipso. Učna množica je lahko majhna (min. 100 slik) [4], kar je idealno za bazo slik, kot je AWE, saj vsebuje malo ugodnih slik za poravnavo. Elipsa se mora čim bolj prilegati zunanjemu delu uhlja, kar je v nekaterih primerih nemogoče bodisi zaradi rotacije uhlja ali pa zaradi njegove unikatne oblike. V takšnih primerih elipse ni mogoče popolnoma prilagoditi legi uhlja, zaradi česar vnesemo napako že v učno množico. Nekaj označenih uhljev iz učne množice je prikazanih na sliki 3.3, na kateri je viden problem prilagajanja elips.

### 3.2.2 Učenje

Na začetku je model, v našem primeru elipsa, inicializiran na naključni lokaciji na sliki. Metoda nato iterativno optimizira projekcijsko matriko tako, da je razlika znotraj parov pozicijskih značilk (angl. pose features) minimalna [11]. Pozicijska značilka je razlika med črno-belimi vrednostmi slikovnih pik dveh kontrolnih lokacij na sliki, ki so vnaprej fiksno določene.





Slika 3.3: Na slikah je nekaj primerov označenih uhljev, pri katerih je moč opaziti, da elipse ni vedno mogoče popolnoma prilagoditi legi uhlja, bodisi zaradi njegove lege bodisi oblike. To pa pozneje lahko vpliva na postopek ugotavljanja lege in usmeritve uhlja na testni množici.

### 3.2.3 Vhodni podatki

Metoda CPR kot vhodni podatek sprejme tri argumente. Prvi argument je tabela neporavnanih slik, drugi je model, ki ga metoda poskuša pozicionirati glede na ugotovljeno orientacijo objekta na sliki, tretji parameter pa je naučen regresor, ki ga pridobimo v fazi učenja metode.

### 3.2.4 Rezultat pozicioniranja modela

Rezultat je središče elipse in točki, ki definirata elipso. Prav tako dobimo tudi orientacijo elipse, ki je predstavljena s kotom  $\theta$  (slika 3.4) v radianih. Pomeni naklon, za katerega elipsa odstopa od svoje navpične lege.



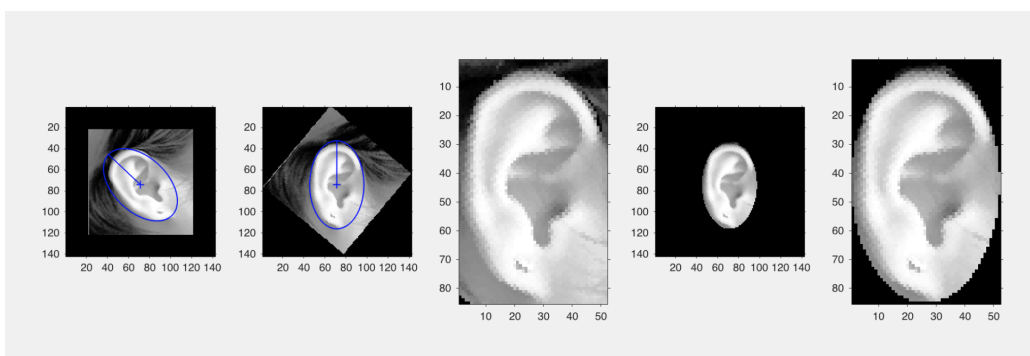
Slika 3.4: Na sliki sta vidna središče elipse in njena orientacija (kot  $\theta$ ) ter kot  $\delta$ , za katerega pozneje rotiramo sliko.

### 3.2.5 Poravnava, segmentacija in maskiranje

Ko metoda CPR oceni kot rotacije  $\theta$ , lahko sliko ustrezno zavrtimo za kot  $\delta$ , ki je definiran z enačbo:

$$\delta = \frac{\pi}{2} - \theta.$$

Po poravnavi uhlja in elipse za kot  $\delta$ , dobimo poravnan uhelj po osi naklona. Zaradi rotacije se na novi sliki v ogliščih pojavijo črni predeli, ki negativno vplivajo na postopek razpoznave. Da bi te črnine minimalizirali, je bil uhelj obrezan glede na elipsi očitani pravokotnik. Rezultat poravnave sta dve množici. Prva vsebuje obrezane slike glede na elipsi očitani pravokotnik, v drugi pa so slike še dodatno maskirane z binarno masko, ki je orisana glede na ocenjeno elipso. Postopek poravnave, segmentacije in maskiranja je prikazan na sliki 3.5.



Slika 3.5: Na sliki je prikazan postopek poravnave z metodo CPR. Ta ob pomoči naučenega regresorja najprej oceni orientacijo in središče elipse, nato pa sledi rotacija slike za kot  $\delta$ . Tu se poravnava zaključi; sledita še obrezovanje ob pomoči elipsi očitane pravokotnika in maskiranje.

Odprikodna implementacija naše metode je na voljo na naslovu <https://github.com/metodribic/ear-alignment-cpr>.

# Poglavje 4

## Eksperimenti

### 4.1 Soglasje naključnega vzorca (RANSAC)

#### 4.1.1 Podatki

Za potrebe testiranja smo uporabili vse slike levih uhljev iz baze uhljev AWE, saj je v bazi 40 (4 % celotne baze) levih uhljev več kot desnih. Ker so slike v bazi zajete pod različnimi koti, sta bili za potrebe testiranja določeni dve ločeni množici, ki sta služili za ocenitev uspešnosti metode. Prva množica vsebuje le slike, ki imajo nagib in odklon manjša od  $10^\circ$  ter poljuben naklon. V drugi množici pa so vse slike ne glede na velikost nagiba, naklona in odklona. Prva množica vsebuje 358 slik, druga pa 520. Slike, ki jih metoda RANSAC ni uspela poravnati, so bile izločene. Prav tako smo izločili slike, ki so bile po poravnavi še edine preostale slike posamezne osebe. V nasprotnem primeru bi namreč v postopku razpoznave to edino sliko primerjali samo s sabo, kar je nesmiselno, saj bi vedno dobili popolno ujemanje. Na koncu je v prvi množici ostalo 105, v drugi pa 163 slik. Vsaka poravnana slika v prvi in drugi množici ima svoj neporavnani par v eni izmed originalnih množic. Ti dve množici služita kot temelj za določanje vpliva poravnave na uspešnost razpoznave.

### 4.1.2 Metoda SIFT

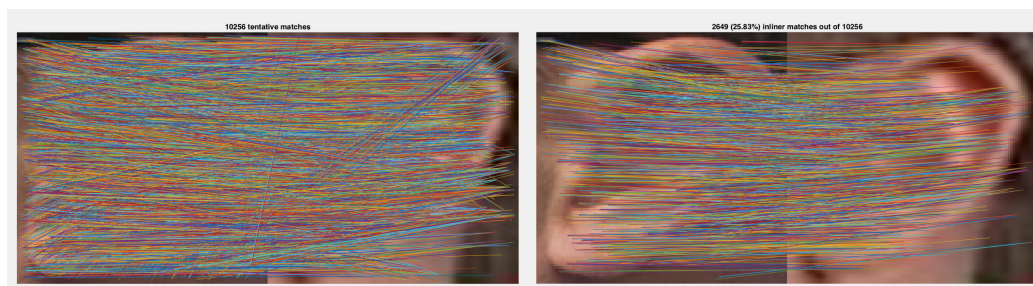
Ker so slike v slikovni bazi AWE zelo majhnih dimenzij, je bilo pri metodi SIFT potrebno zelo natančno nastaviti vhodne parametre. V nasprotnem primeru metoda ni pridobila dovolj parov točk in njim ustreznih deskriptorjev. Parametri metode SIFT so sledeči [12]:

- Meja izbora vrhov (angl. peak selection threshold): 0.
- Meja ne-robnege izbora (angl. non-edge selection threshold): 1000.
- Minimum drugostopenjske normalizacije deskriptorjev pred normalizacijo (angl. minimum l2-norm of the descriptors before normalization): 2 (deskriptorji pod mejo so nastavljeni na 0).
- Število stopenj na oktavo v skalarnem prostoru razlik Gaussove piramide (angl. number of levels per octave of the DoG scale space): 200.
- Faktor povečave deskriptorjev (angl. descriptor magnification factor): 5.

Vsi parametri so bili nastavljeni eksperimentalno.

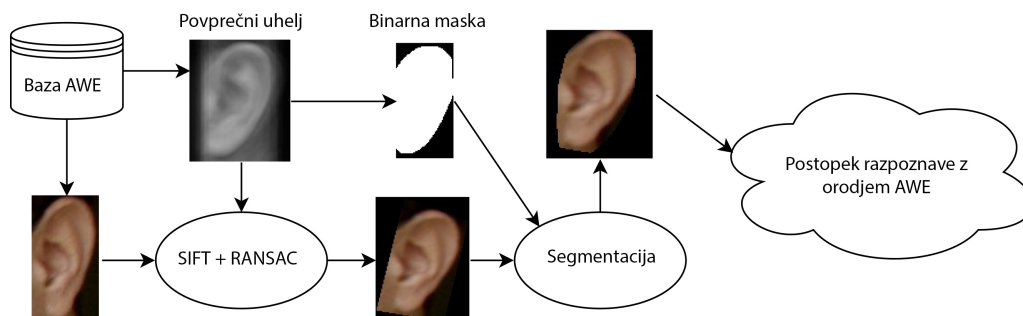
### 4.1.3 Metoda RANSAC in poravnava

Kot je opisno v poglavju 3.1.4, RANSAC v vsaki izmed  $t$  iteracij izbere  $n$  točk in glede na te točke oceni geometrijsko transformacijo. Ker so slike v bazi uhljev AWE zajete pod različnimi koti in ker predstavljajo tridimenzionalne objekte, je potrebno  $n$  nastaviti na 4, saj je to minimum za ocenitev projekcijske transformacije. Število iteracij  $t$ , ki jih RANSAC opravi za ocenitev transformacije, je v našem primeru 5000, po potrebi pa je  $t$  povečan na 10000 in nazadnje na 25000. Če RANSAC po 25000 iteracijah ni uspel poravnati slike z vsaj 40-odstotno uspešnostjo, jo zavržemo. Uspešnost je definirana kot  $n_{ugodni}/n_{vsi}$ , kjer je  $n_{ugodni}$  število parov točk, ki jih pridobimo z metodo SIFT in jih je RANSAC uspešno postavil na ravnino,  $n_{vsi}$  pa je število vseh parov



Slika 4.1: Na levi sliki so med sabo povezani vsi pari točk, ki smo jih pridobili z metodo SIFT. Na desni sliki pa so prikazane povezave parov točk, ki jih je metoda RANSAC uspela postaviti na ravnino. Te točke v veliki meri predstavljajo istoležeče točke na obeh slikah.

točk. Vsi parametri, razen parametra  $n$ , so bili nastavljeni eksperimentalno. Primer vseh zaznanih parov točk in pari točk, ki ostanejo po procesiranju z metodo RANSAC, so vidni na sliki 4.1. Če metoda RANSAC uspe pridobiti ustrezno transformacijsko matriko, se ta uporabi za transformacijo originalne slike. Celoten proces od poravnave do razpoznave je razviden na sliki 4.2.



Slika 4.2: Proces poravnave uhljev z metodo RANSAC.

## 4.2 Kaskadna pozicijska regresija (CPR)

### 4.2.1 Podatki

Ker so izsledki analize pri metodi RANSAC pokazali, da za poravnavo uhljev iz baze AWE niso primerne vse slike, smo se odločili, da za poravnavo pri metodi CPR uporabimo zgolj tiste, ki so za poravnavo primerne. Slika velja za primerno, če ima nagib in odklon manjša od  $10^\circ$ . Takšnih slik je 358, kar je 68 % originalne podmnožice levih uhljev. Ker je A. Pflug v svojem delu [11] za poravnavo baze UND-J2 uporabila metodo CPR, smo najprej preverili delovanje metode na bazi IITD I. Ta je zelo podobna bazi UND-J2 in je hkrati lažja za poravnavo kot baza AWE. Ker metoda potrebuje vsaj 100 slik za učenje, smo se odločili, da iz vsake baze, tako AWE kot IITD I, naključno izberemo  $\frac{1}{3}$  slik. S tem zagotovimo zadostno število slik za učenje, prav tako pa ostane dovolj slik za testiranje.

### 4.2.2 Metoda CPR

Metoda CPR poleg vhodnih podatkov, ki so opisani v poglavju 3.2.3, zahteva definicijo nekaterih parametrov, ključnih za uspešno delovanje metode. Vsi parametri, naštetih v nadaljevanju, so bili nastavljeni v skladu s priporočili avtorja metode [4]:

- Število kaskad: 512.
- Število množic značilk (angl. ferns): 64.
- Število pozicijsko indeksiranih značilk (angl. pose-index features): 512.
- Regresor je inicializiran 50-krat.

### 4.2.3 Poravnava

Za postopek poravnave je potrebno najprej pridobiti naučen regresor, ki ga pozneje uporabimo za samodejno ugotavljanje pozicije in orientacije uhlja. Učenje se izvaja na označeni učni množici, ki je opisana v podpoglavju 4.2.1. Naučeni regresor uporabimo na testni množici. Tako pridobimo množico elips s pripadajočimi podatki. Sledi iteracija čez vse slike v testni množici, za vsako sliko pa se izvedejo operacije v enakem vrstnem redu, kot so navedene v nadaljevanju:

1. Pretvorba kota  $\theta$  iz radianov v stopinje:  $1^\circ = \frac{180}{\pi}$ .
2. Izračun kota  $\delta$ :  $\delta = 90^\circ - \theta$ .
3. Izris elipse na sliko vključno s središčem in premico, ki prikazuje usmerjenost elipse (primer je prikazan na sliki 3.5).
4. Izračun binarne maske na podlagi izrisane elipse.
5. Kopija slike, na kateri je uporabljena binarna maska.
6. Izračun koordinat in velikosti očrtanega pravokotnika.
  - (a) Dvakratnik najmanjše razdalje od središča elipse do njenega loka je enak krajši stranici očrtanega pravokotnika.
  - (b) Dvakratnik dolžine premice, ki prikazuje usmerjenost elipse, je enak daljši stranici očrtanega pravokotnika.
  - (c) Koordinate pravokotnika se izračunajo iz središča elipse in dolžin stranic očrtanega pravokotnika.
7. Obrezovanje slike glede na očrtan pravokotnik.
8. Obrezovanje slike, ki smo jo pridobili v koraku 5, v katerem je bila uporabljena binarna maska.
9. Shranjevanje slike tako, da se struktura slik po poravnavi ohrani.

## 4.3 Vrednotenje

Za ocenitev razpoznavne na poravnanih uhljih smo uporabili orodje AWE. Pridobljene množice slik smo testirali z vsemi deskriptorji, ki so v orodju na voljo. Za ocenitev vpliva poravnave na uspešnost razpoznavanja smo uporabili prvi rang stopnje razpoznave (angl. Rank 1 recognition rate – Rank-1). Poleg vsakega rezultata je predstavljena tudi enakostna stopnja napake (angl. Equal Error Rate – EER). ERR je vrednost, pri kateri je delež napačno sprejetih vzorcev (angl. False Acceptance Rate – FAR ) enak deležu napačno zavrnjenih vzorcev (angl. False Rejection Rate – FRR ). FAR in FRR sta verjetnosti, da sistem napačno sprejme (FAR) ali zavrne (FRR) osebi dostop na podlagi biometrične modalnosti. Višji kot je Rank-1 in nižji kot je EER, bolj natančna je razpoznavna.

### 4.3.1 Deskriptorji

Orodje AWE ima za izločanje značilke na voljo 8 različnih metod deskriptorjev (LBP, BSIF, LPQ, RILPQ, POEM, HOG, DSIFT in Gabor), ki iz slike pridobijo značilke za vhodno sliko. Značilka je vektor oziroma opisnik za neko območje na sliki in opisuje njegove lastnosti. Pri ugotavljanju, ali sta si dve sliki med sabo podobni, si pomagamo z razdaljami/razlikami med deskriptorji. Po priporočilu avtorjev orodja AWE [5] pri vseh deskriptorjih uporabljamo  $\chi^2$  razdaljo, ki je definirana kot:

$$d(i, 1) = \sum_{i=1}^m \sum_{j=1}^p d(i, 1) + \frac{(XI(1, j) - (\frac{XI(1, j) + XJ(i, j)}{2}))^2}{\frac{XI(1, j) + XJ(i, j)}{2}},$$

kjer je  $m$  enak številu komponent v vektorju  $XJ$ ,  $p$  pa je dimenzija vzorcev, ki so prisotni v vektorjih  $XI$  in  $XJ$ .  $XI$  in  $XJ$  vsebujeta deskriptorje, izločene iz slik  $A_i$  in  $A_j$ . Prav tako mora biti zadoščen pogoj:

$$\frac{XI(1, j) + XJ(i, j)}{2} > 0.$$



### 4.3.2 Prvi rang stopnje razpoznavne (Rank-1) in krivulja kumulativnih zadetkov (CMC)

Vse pripadajoče razdalje med slikami postavimo v matriko, kot je prikazano na sliki 4.3.  $A_i$  predstavlja  $i$ -to sliko v bazi,  $d_{i,j}$  pa izračunano razdaljo med deskriptorjema slik  $A_i$  in  $A_j$ . Ko imamo matriko podobnosti sestavljeno, jo uredimo po velikosti in izločimo primerjavo slike same s sabo, saj je brezpredmetna. Ker vemo, da vsaka slika  $A_i$  pripada nekemu razredu  $C_l$ , primerjamo razred vzorca/slike v urejeni matriki z razredom slike  $A_i$ . Če se ujemata, pomeni, da sliki predstavljata isti razred, kar v našem primeru pomeni, da pripadata isti osebi. Koliko slik v vrstici primerjamo, je odvisno od tega, kateri rang želimo. Če je to Rank-1, opazujemo samo en vzorec, če gre za Rank- $n$ , pa gledamo prvih  $n$  vzorcev. Razmerje med pravilnimi razredi in vsemi vzorci predstavlja krivulja kumulativnih zadetkov (angl. Cumulative Match Curve – CMC).

$$\begin{array}{c}
 A_1 \quad A_2 \quad \dots \quad A_{n-1} \quad A_n \\
 \begin{array}{c}
 A_1 \\
 A_2 \\
 \vdots \\
 A_{n-1} \\
 A_n
 \end{array}
 \begin{pmatrix}
 1 & d_{1,2} & \dots & d_{1,n-1} & d_{1,n} \\
 d_{2,1} & 1 & \dots & d_{2,n-1} & d_{2,n} \\
 \vdots & \vdots & \ddots & \vdots & \vdots \\
 d_{n-1,1} & d_{n-1,2} & \dots & 1 & d_{n-1,n} \\
 d_{n,1} & d_{n,2} & \dots & d_{n,n-1} & 1
 \end{pmatrix}
 \end{array}$$

Slika 4.3: Matrika podobnosti, ki je osnova za Rank-1.

### 4.3.3 RANSAC

Na obeh množicah sta bila izvedena dva ločena testa. Prvi je bil izveden na originalni množici, drugi pa na enaki množici, le da so bili uhlji poravnani. Rezultati neporavnanih uhljev so služili kot temelj za ocenitev uspešnosti vpliva poravnave na razpoznavo.

#### 4.3.4 CPR

Pri metodi CPR smo za poravnavo uporabili dve bazi: IITD I in AWE. Ker v literaturi nismo zaznali uporabe metode RANSAC v povezavi s poravnavo uhljev, smo pri metodi RANSAC uporabili zgolj bazo AWE, medtem ko je A. Pflug v svojem delu [11] uporabila poravnavo ob pomoči metode CPR. Ker smo imeli izsledke njene raziskave, v kateri je uporabila bazo, podobno bazi IITD I, smo vedeli, kakšne rezultate pričakovati. Poravnava baze IITD I je pokazala, da naša implementacija metode CPR deluje pravilno za poravnavo uhljev, hkrati pa so rezultati poravnane baze IITD I služili za primerjavo poravnave lahke in težke baze. Tako na bazi IITD I kot na bazi AWE smo izvedli tri neodvisne teste:

1. Test originalne množice slik.
2. Test poravnane množice slik, obrezane glede na elipsi očrtan pravokotnik.
3. Test poravnane množice slik z uporabo binarne maske, obrezane glede na elipsi očrtan pravokotnik (slika 3.5).

Za izločitev značilk v postopku razpoznavne smo uporabili enake deskriptorje kot pri metodi RANSAC.

# Poglavje 5

## Rezultati z diskusijo

### 5.1 Soglasje naključnega vzorca (RANSAC)

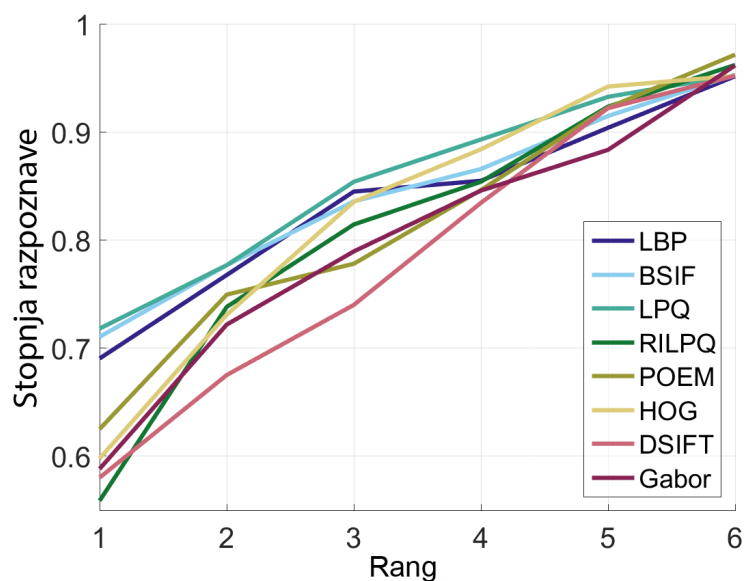
Rezultati prve množice (slike z nagibom in odklonom, manjšima od  $10^\circ$ ), opisani v poglavju 4.1.1, so pokazali, da poravnani uhlji z metodo RANSAC v splošnem ne izboljšajo razpoznavne, saj so rezultati ti presegli rezultate neporavnane množice zgolj pri deskriptorju LPQ. Rezultat, pridobljen z LPQ deskriptorjem, se pri Rank-1 zviša s 66.3 % na 71.8 %. Pri vseh ostalih deskriptorjih je občutno slabši. Pri drugi množici (vse slike) niti LPQ, ki je pokazal dobre rezultate pri prvi množici, ne doseže rezultatov neporavnane množice. To potrjuje, da slik, ki imajo odklon večji od  $10^\circ$ , ni mogoče uspešno poravnati. Zgodi se tudi, da uhlj z metodo RANSAC kljub velikem kotu po vseh definiranih kriterijih uspešno poravnamo, v resnici pa izvedemo transformacijo, ki mu popolnoma spremeni karakteristike. To je razlog, da so rezultati poravnanih uhljev v drugi množici, ki vsebuje vseh 520 slik, občutno slabši. Rezultat takšne poravnave je prikazan na sliki 5.1. Rezultati prve množice poravnanih in neporavnanih uhljev so predstavljeni v tabeli 5.1 ter na slikah 5.2 in 5.3. Rezultati druge množice, ki vsebuje vse slike neodvisno od kota, pa so prikazani v tabeli 5.2 ter na slikah 5.4 in 5.5.

Tabela 5.1: Rezultati razpoznavne slik prve množice, ki vsebuje slike z majhnima in srednjima nagibom ter odklonom. Nepravilna množica slik je služila kot temelj za ugotavljanje vpliva poravnave na uspešnost razpoznavanja uhljev.

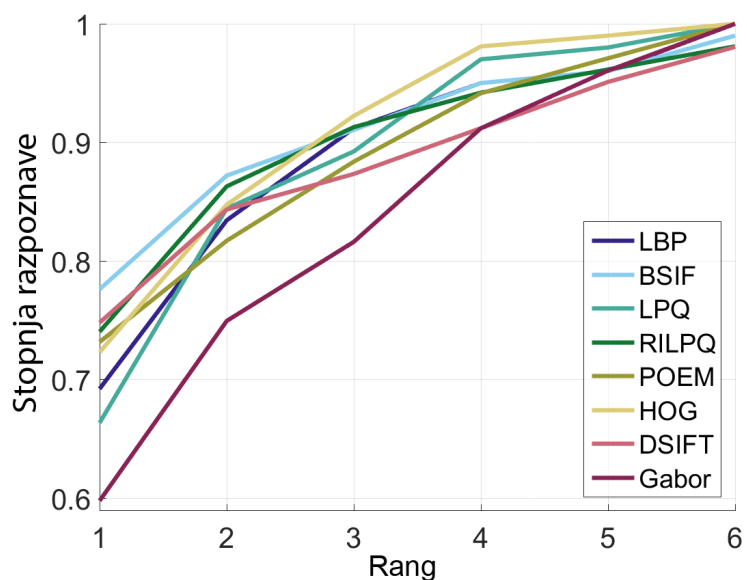
Deskriptorji	Poravnane slike		Neporavnane slike	
	Rank-1 (%)	EER (%)	Rank-1 (%)	EER (%)
LBP	69.0 ± 14.4	25.3 ± 11.9	69.2 ± 8.7	26.1 ± 13.2
BSIF	71.0 ± 13.7	23.2 ± 10.2	77.6 ± 13.5	22.1 ± 9.5
LPQ	71.8 ± 17.0	26.3 ± 13.2	66.3 ± 11.0	27.7 ± 13.1
RILPQ	55.9 ± 14.0	29.8 ± 10.2	74.0 ± 9.4	24.2 ± 10.1
POEM	62.5 ± 17.8	26.9 ± 11.2	73.2 ± 7.6	21.0 ± 7.5
HOG	59.8 ± 19.5	30.5 ± 14.4	72.3 ± 10.7	22.9 ± 11.0
DSIFT	58.0 ± 20.1	32.0 ± 12.9	74.8 ± 11.2	24.0 ± 6.9
Gabor	58.8 ± 12.1	31.5 ± 9.4	59.8 ± 11.2	27.9 ± 5.6



Slika 5.1: Na sliki je prikazana transformacija leve slike v desno, pri čemer je poravnava v željeno dvodimenzionalno ravnino nemogoča, saj nimamo dovolj podatkov, da bi uhlje lahko rekonstruirali. Zaradi poravnave uhlja kljub pomanjkanju informacij nastane popolnoma nov uhelj, ki je komaj še podoben originalnemu.



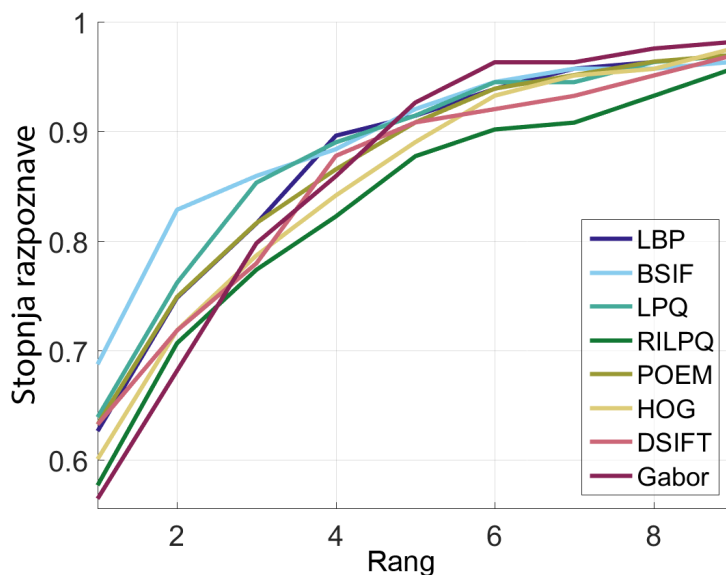
Slika 5.2: Na grafu so prikazani rezultati glede na rang razpoznave poravnanih slik iz prve množice, ki vsebuje le slike z majhnima ter srednjima nagibom in odklonom.



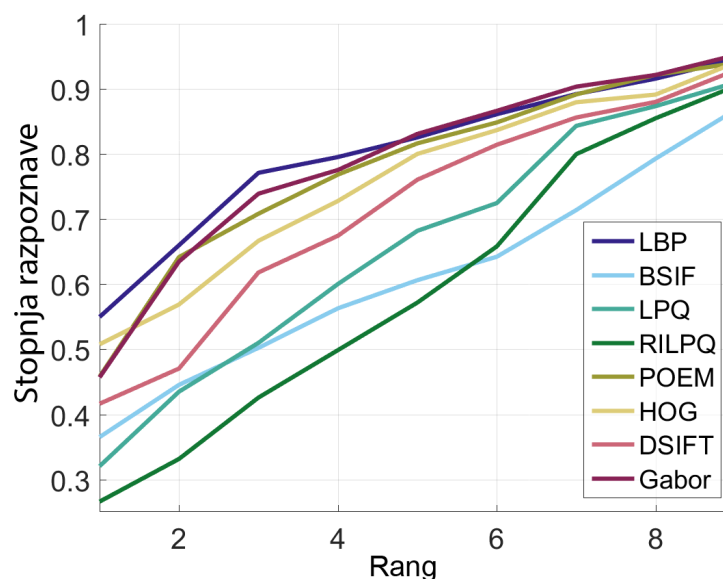
Slika 5.3: Na grafu so prikazani rezultati razpoznavne nepravilnih slik iz prve množice, ki vsebuje le slike z majhnima do srednjima nagibom in odklonom.

Tabela 5.2: Rezultati razpoznavne slik druge množice, ki vsebuje vse slike brez omejitve kota. Nепорavnana množica slik je služila kot temelj za ugotavljanje vpliva poravnave na uspešnost razpoznavanja uhljev.

Deskriptorji	Poravnane slike		Neporavnane slike	
	Rank-1 (%)	EER (%)	Rank-1 (%)	EER (%)
LBP	$54.97 \pm 11.5$	$34.0 \pm 7.69$	$62.6 \pm 10.1$	$26.37 \pm 10.3$
BSIF	$36.5 \pm 12.1$	$47.0 \pm 12.9$	$68.7 \pm 10.5$	$25.2 \pm 10.3$
LPQ	$32.0 \pm 8.3$	$47.9 \pm 9.9$	$63.9 \pm 10.5$	$24.7 \pm 11.8$
RILPQ	$26.5 \pm 11.0$	$46.8 \pm 6.1$	$57.7 \pm 11.4$	$25.0 \pm 12.9$
POEM	$45.8 \pm 12.3$	$29.1 \pm 5.1$	$63.2 \pm 13.9$	$26.4 \pm 11.2$
HOG	$50.8 \pm 13.9$	$33.0 \pm 9.9$	$60.1 \pm 5.8$	$25.9 \pm 9.79$
DSIFT	$41.6 \pm 15.7$	$37.0 \pm 10.9$	$63.2 \pm 8.9$	$25.4 \pm 6.2$
Gabor	$45.7 \pm 12.2$	$29.9 \pm 10.1$	$56.46 \pm 12.6$	$26.4 \pm 7.7$



Slika 5.4: Graf prikazuje rezultate glede na rang razpoznave za neporavnane slike iz druge množice, ki vsebuje vse slike iz podmnožice levih uhljev baze AWE.



Slika 5.5: Graf prikazuje rezultate glede na rang razpoznave za poravnane slike iz druge množice, ki vsebuje vse slike iz podmnožice levih uhljev baze AWE.

## 5.2 Kaskadna pozicijska regresija (CPR)

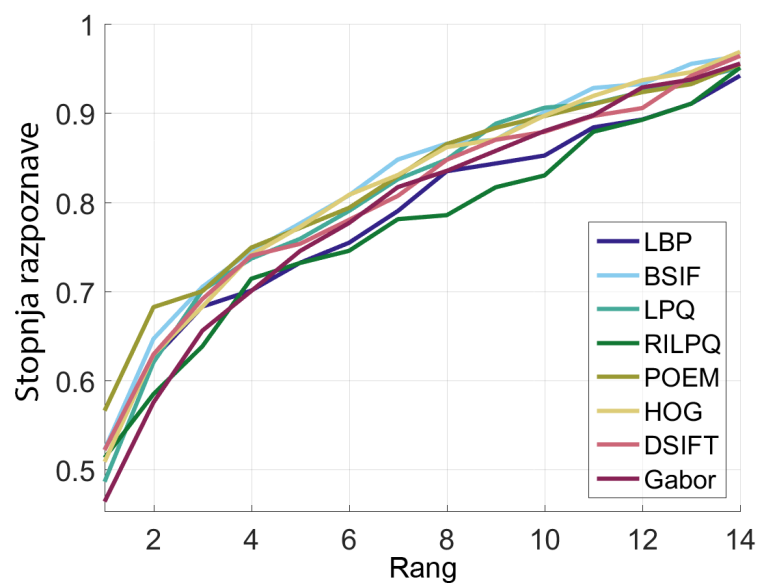
Kot je opisano v poglavju 4.2.1, smo evalvacijo poravnanih uhljev z metodo CPR izvedli na 238-ih slikah levih uhljev iz baze AWE (vsi levi uhlji, ki imajo nagib in odklon manjša od  $10^\circ$ ) ter na celotni bazi IITD I. Rezultati so potrdili izsledke, ugotovljene pri metodi RANSAC, da je baza AWE zahtevna za poravnavo, saj niti izločitev slik, katerih odklon je večji od  $10^\circ$ , ne pripomore k izboljšanju rezultata razpoznave pri poravnavi z metodo CPR. Rezultati poravnave in maskiranja so vidni v tabeli 5.3 ter slikah 5.6, 5.7 in 5.8. Iz rezultatov se dobro vidi, da niti maskiranje in normalizacija ne pripomoreta k boljši učinkovitosti razpoznave. Ravno nasprotno pa so pokazali rezultati baze IITD I, kjer poravnava skupaj z maskiranjem in normalizacijo (tako ločeno kot tudi skupaj) občutno izboljša postopek razpoznave. Rezultati so se pri nekaterih deskriptorjih izboljšali za več kot 10 %. Vsi rezultati baze

IITD I so prikazani v tabeli 5.4 ter slikah 5.9, 5.10 in 5.11. Zanimiva je ugotovitev, da maskiranje pri bazi IITD I ne pripomore k izboljšanju rezultatov poravnave. Razlog za to je dejstvo, da že v učno množico vnesemo minimalno napako, ker elips ne moremo popolnoma prilagoditi uhlju, iz česar sledi, da tudi naučen regresor ne more vseh elips popolnoma prilagoditi uhljem. Zaradi tega slike po uporabi binarne maske vsebujejo dele slik, ki pripadajo okolici in ne uhlju, kar pa negativno vpliva na razpoznavo.

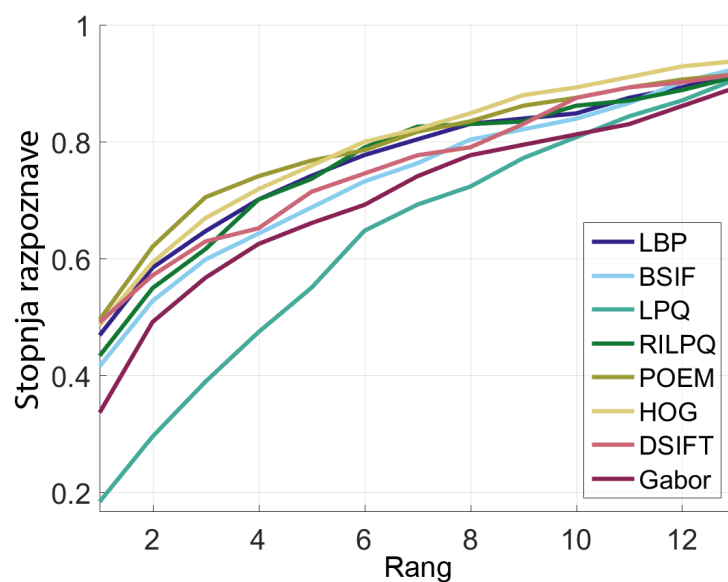
Tabela 5.3: Rezultati razpoznave za bazo slik AWE, ki je bila poravnana z metodo CPR. V prvem stolpcu so prikazani rezultati razpoznave za neporavnano množico slik. V drugem stolpcu so rezultati razpoznave za poravnane in obrezane slike glede na elipsi očitran pravokotnik. V zadnjem stolpcu pa so rezultati razpoznave poravnanih, obrezanih in maskiranih slik.

Deskriptorji	Neporavnane slike		Poravnane in obrezane slike		Poravnane, obrezane in maskirane slike	
	Rank-1 (%)	EER (%)	Rank-1 (%)	EER (%)	Rank-1 (%)	EER (%)
LBP	52.2 ± 10.8	30.4 ± 7.7	46.9 ± 11.0	35.7 ± 6.5	49.9 ± 8.5	32.7 ± 6.6
BSIF	52.2 ± 13.8	30.2 ± 6.7	41.53 ± 8.2	37.9 ± 9.01	34.8 ± 6.1	40.2 ± 8.4
LPQ	48.6 ± 12.6	30.9 ± 6.1	18.4 ± 7.8	39.8 ± 6.2	23.2 ± 6.7	52.0 ± 6.7
RILPQ	51.3 ± 14.9	33.5 ± 9.3	43.3 ± 12.0	32.2 ± 4.2	35.7 ± 4.9	37.7 ± 5.3
POEM	56.6 ± 12.0	28.0 ± 5.9	49.5 ± 11.5	32.2 ± 7.6	48.6 ± 9.2	32.0 ± 4.9
HOG	50.9 ± 6.4	30.0 ± 9.7	48.6 ± 15.7	30.6 ± 8.1	43.3 ± 4.2	32.5 ± 7.1
DSIFT	52.2 ± 10.2	29.9 ± 7.2	49.0 ± 13.3	37.1 ± 11.6	45.9 ± 9.2	32.6 ± 7.7
Gabor	46.4 ± 6.4	30.3 ± 5.1	33.6 ± 7.7	36.5 ± 8.8	37.0 ± 9.4	36.8 ± 9.8

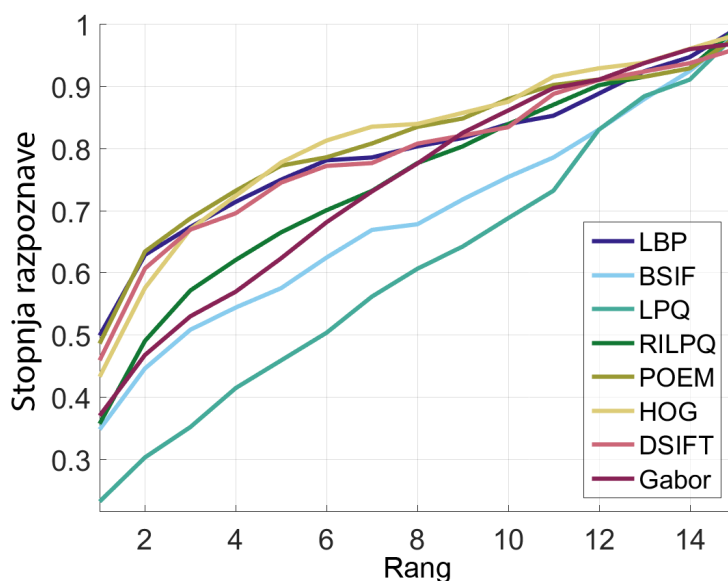




Slika 5.6: Graf uspešnosti razpoznave glede na rang razpoznave pri neporavnani bazi AWE z metodo CPR.



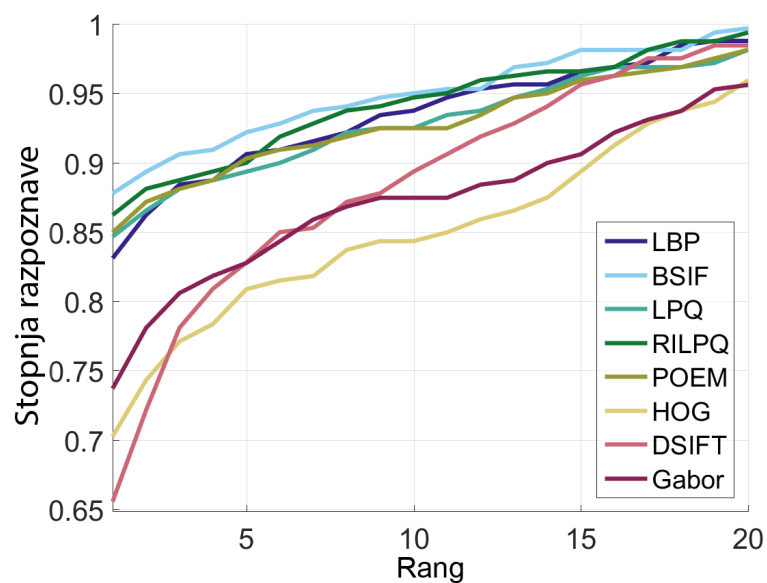
Slika 5.7: Graf uspešnosti razpoznave glede na rang razpoznave pri poravnani in obrezani testni množici baze AWE z metodo CPR.



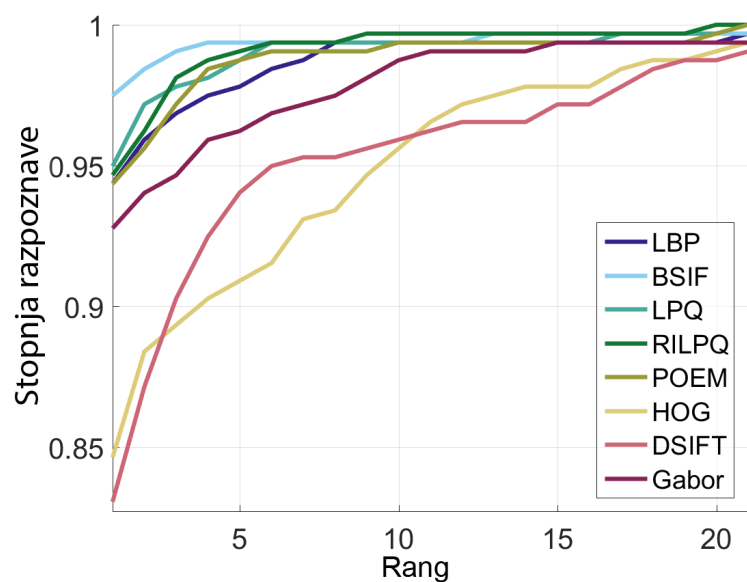
Slika 5.8: Graf uspešnosti razpoznave glede na rang razpoznave pri poravnani, obrezani in maskirani testni množici baze AWE z metodo CPR.

Tabela 5.4: Rezultati razpoznave za bazo slik IITD I, ki je bila poravnana z metodo CPR. V prvem stolpcu so rezultati razpoznave za neporavnane slike. V drugem stolpcu so prikazani rezultati razpoznave za poravnane in obrezane slike glede na elipsi očrtan pravokotnik. V zadnjem stolpcu pa so rezultati razpoznave poravnanih, obrezanih in maskiranih slik.

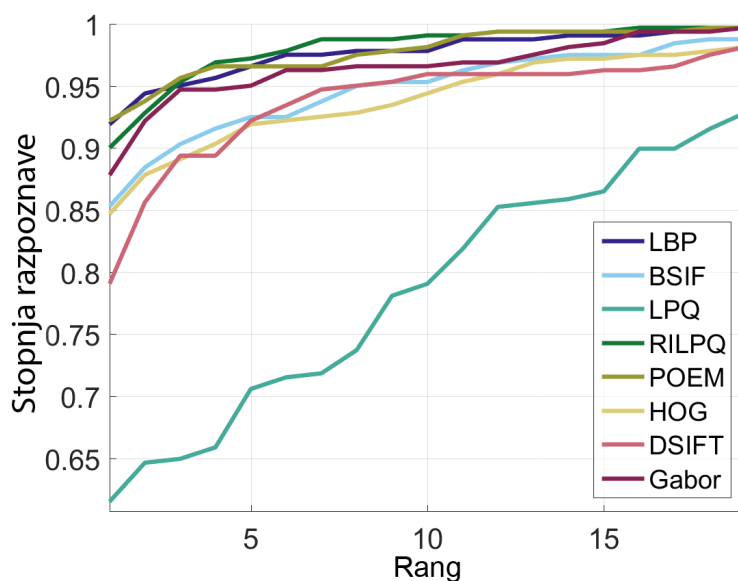
Deskriptorji	Neporavnane slike		Poravnane in obrezane slike		Poravnane, obrezane in maskirane slike	
	Rank-1 (%)	EER (%)	Rank-1 (%)	EER (%)	Rank-1 (%)	EER (%)
LBP	83.1 ± 7.9	12.9 ± 4.3	94.4 ± 2.1	8.2 ± 2.8	91.9 ± 4.1	9.4 ± 4.3
BSIF	87.8 ± 5.4	11.9 ± 4.9	97.5 ± 1.6	6.0 ± 2.4	85.3 ± 5.4	14.1 ± 5.1
LPQ	84.7 ± 6.9	14.7 ± 4.6	95.0 ± 1.9	7.5 ± 2.8	61.5 ± 6.8	34.8 ± 7.2
RILPQ	86.2 ± 4.7	13.2 ± 4.1	94.7 ± 2.2	6.3 ± 1.9	90.0 ± 3.7	10.2 ± 5.0
POEM	85.0 ± 4.6	13.1 ± 4.1	94.4 ± 2.4	7.1 ± 2.4	92.2 ± 2.6	8.8 ± 3.9
HOG	70.3 ± 7.2	20.7 ± 3.8	84.6 ± 3.3	12.6 ± 5.0	84.7 ± 2.8	14.0 ± 5.6
DSIFT	65.5 ± 9.5	21.4 ± 6.0	83.1 ± 4.4	14.8 ± 5.9	79.1 ± 6.3	16.3 ± 4.8
Gabor	73.7 ± 8.3	19.1 ± 5.9	92.8 ± 3.8	10.7 ± 3.1	87.8 ± 2.9	11.6 ± 3.3



Slika 5.9: Graf uspešnosti razpoznavne glede na rang razpoznave pri neporavnani testni množici baze IITD I z metodo CPR.



Slika 5.10: Graf uspešnosti razpoznavne glede na rang razpoznave pri poravnani in obrezani testni množici baze IITD I z metodo CPR.



Slika 5.11: Graf uspešnosti razpoznavne glede na rang razpoznavne pri poravnani, obrezani in maskirani testni množici baze IITD I z metodo CPR.

## 5.3 Primerjava rezultatov z diskusijo

### 5.3.1 Ugotovitve

Dokazali smo, da slike, poravnane z metodo RANSAC, v splošnem ne izboljšajo rezultatov razpoznavne. Do tega spoznanja smo prišli ob ugotovitvi, da dobimo boljše rezultate na poravnani kot na neporavnani množici zgolj pri deskriptorju LPQ. Poleg njega je moral veljati pogoj, da imajo slike nagib in odklon manjša od  $10^\circ$ . V nasprotnem primeru izvedemo transformacijo, katere rezultat je popolnoma nov uhelj, ki v najslabšem primeru ne vsebuje niti podobnosti, vidnih s prostim očesom. Kot nadgradjo metode RANSAC smo po zgledu članka A. Pflug [11] preizkusili metodo CPR nad bazo AWE in IITD I. Rezultati obeh metod so potrdili, da poravnava baze, kot je slikovna baza AWE, negativno vpliva na postopek razpoznavne, saj so slike v tej bazi prezahtevne za poravnavo. Na drugi strani pa je test poravnave z metodo

CPR na bazi IITD I pokazal, da se odstotek razpoznavne signifikantno izboljša, če je baza enostavna za poravnavo. Iz teh ugotovitev je moč sklepati, da so metode poravnave uporabne zgolj na slikah, ki so bile zajete pod zelo majhnima nagibom in odklonom, saj v nasprotnem primeru nimamo dovolj informacij za uspešno poravnavo.

### 5.3.2 Mogoče izboljšave

Če bi želeli poravnavati poljubne slike neodvisno od kotov njihovega zajema, bi morali pri transformaciji ohraniti le vidne dele uhlja, saj sicer izvedemo razteg slikovnih pik, kar pa v splošnem pomeni napačno transformacijo. To bi bilo mogoče, če bi znali po končani transformaciji zaznati, kateri del uhlja je bil viden že na originalni sliki in kateri deli so bili zgolj interpolirani/raztegnjeni. Če bi bili sposobni zaznati ta območja, bi jih lahko zapolnili s črnino ali barvo, ki se najmanjkrat pojavi na slikah. Ko bi imeli poravnane uhlje, katerih novi deli bi bili posebno označeni, bi morali izdelati samo še deskriptor, ki bi ta območja ignoriral, kar bi pomenilo, da bi se v postopku poravnave upoštevali zgolj deli uhljev, ki so bili vidni na prvotni sliki. Poravnava poljubnih uhljev brez omejitev kota bi bila mogoča tudi v primeru, da bi imeli tridimenzionalno sliko uhlja vsake osebe, saj bi tako vedeli, kako izgleda uhelj v katerikoli poziciji. V tem primeru bi morali oceniti kote na vseh treh oseh na vhodni sliki, obrniti tridimenzionalni model skladno s temi koti in primerjati vhodni uhelj s sliko, ki je vidna opazovalcu (kameri) tridimenzionalnega modela. Če bi bilo ujemanje zadostno, bi lahko z veliko verjetnostjo trdili, da vhodni uhelj pripada določeni osebi in bi lahko zapolnili morebitne manjkajoče dele uhlja. Potrebno bi bilo postaviti visoko mejo, nad katero bi lahko z gotovostjo trdili, da določen uhelj pripada določeni osebi, saj bi v nasprotnem primeru v sliko vnesli napačne informacije, ki pa bi lahko zelo negativno vplivale na postopek razpoznavne. Ta postopek bi bil verjetno precej dolgotrajen, saj bi morali vsako vhodno sliko primerjati z vsemi tridimenzionalnimi modeli, ki jih imamo, in uporabiti tistega, ki ima najvišje ujemanje in je hkrati nad dovoljeno mejo minimalnega ujemanja.



# Poglavje 6

## Zaključek

V diplomski nalogi smo dokazali, da poravnava pozitivno vpliva na postopek razpoznavne pod pogojem, da so slike ugodne za poravnavo, kar pomeni, da imajo nagib in odklon manjša od  $10^\circ$ . Prav tako je pomembno, kateri deskriptor izberemo za evalvacijo slik, pomembna faktorja pri poravnavi sta tudi velikost in ločljivost slik. Manjše slike je težje uspešno transformirati, saj z vsako transformacijo popačimo sliko oziroma izgubimo pomembne informacije. Normalizacija slik je v vseh primerih izboljšala rezultat razpoznavne, saj smo omejili območje zanimanja (angl. region of interest – ROI), medtem ko je maskiranje rezultate izboljšalo le pri bazi AWE, v kateri so slike zelo majhnih ločljivosti. Pri bazi IITD I je maskiranje sicer izboljšalo rezultate razpoznavne, a so bili rezultati brez maskiranja občutno boljši. To dokazuje, da že najmanjša napaka pri maskiranju poravnanih uhljev negativno vpliva na postopek razpoznavne.

Kot je bilo omenjeno že v prejšnjem poglavju, je prostora za izboljšave še kar nekaj. Najprej bi bilo potrebno izdelati zanesljivo metodo za uspešno poravnavo zgolj tistih delov uhlja, ki so vidni na vhodni sliki, oziroma bi nevidne dele zapolnila z relevantnimi informacijami. Če bi to dosegli, bi lahko poravnavali vse uhlje ne glede na omejitve, ki smo jih dokazali v tej diplomski. Vsekakor so uhlji dobra alternativa obrazom, ki so največkrat upo-

rabljeni za identifikacijo oseb v videonadzornih sistemih. Za uspešno poravnavo in pozneje razpoznavo bi bilo potrebno izdelati zanesljivo metodo, ki bi kljub majhni ločljivosti in poljubnemu kotu zajema slike zagotovila ustrezno poravnavo in razpoznavo. Prav tako bi z ustreznim sistemom, ki bi zanesljivo zaznal in izluščil uhlje iz slike ali videa, lahko zagotovili popolnoma samodejno zaznavanje, poravnavanje in razpoznavanje uhljev, kar bi lahko s pridom uporabljali za forenzične, nadzorne, pa tudi komercialne namene.



# Literatura

- [1] Ayman Abaza, Arun Ross, Christina Hebert, Mary Ann F Harrison, and Mark S Nixon. A survey on ear biometrics. *ACM Computing Surveys*, 45(2):22, 2013.
- [2] Kyong Chang, Kevin W Bowyer, Sudeep Sarkar, and Barnabas Victor. Comparison and combination of ear and face images in appearance-based biometrics. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 25(9):1160–1165, 2003.
- [3] Piotr Dollár. Piotr’s Computer Vision Matlab Toolbox (PMT). <https://github.com/pdollar/toolbox>. [Dostopano: 1. 7. 2016].
- [4] Piotr Dollár, Peter Welinder, and Pietro Perona. Cascaded pose regression. In *Computer Vision and Pattern Recognition (CVPR), IEEE Conference on*, pages 1078–1085, 2010.
- [5] Žiga Emeršič and Peter Peer. Toolbox for ear biometric recognition evaluation. In *EUROCON International Conference on Computer as a Tool (EUROCON), IEEE*, pages 1–6. IEEE, 2015.
- [6] Žiga Emeršič, Vitomir Štruc, and Peter Peer. Ear recognition: More than a survey. *Accepted in Neurocomputing*, 2016.
- [7] Martin A Fischler and Robert C Bolles. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM*, 24(6):381–395, 1981.

- 
- [8] Esther Gonzalez, Luis Alvarez, and Luis Mazonara. Normalization and feature extraction on ear images. In *IEEE International Carnahan Conference on Security Technology (ICCST)*, pages 97–104, 2012.
- [9] Ajay Kumar and Chenye Wu. Automated human identification using ear imaging. *Pattern Recognition*, 45(3):956–968, 2012.
- [10] David G. Lowe. Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*, 60(2):91–110, 2004.
- [11] Anika Pflug and Christoph Busch. Segmentation and normalization of human ears using cascaded pose regression. In *Nordic Conference on Secure IT Systems, (NordSec)*, pages 261–272, 2014.
- [12] VLFeat. VLFeat - Scale Invariant Feature Transform (SIFT) API Documentation. <http://www.vlfeat.org/api/sift.html>. [Dostopano: 3. 5. 2016].
- [13] VLFeat. VLFeat open source library. <http://www.vlfeat.org/>. [Dostopano: 20. 4. 2016].
- [14] Li Yuan and Zhi-chun Mu. Ear recognition based on 2d images. In *IEEE International Conference on Biometrics: Theory, Applications, and Systems (BTAS)*, pages 1–5, 2007.