

Projekt iz predmeta Raspoznavanje uzoraka

Ak. god. 2013/14

# **Raspoznavanje osoba na temelju „meke“ biometrijske značajke - siluete osobe (i identifikacija)**

Autori:

Dorija Humski

Marija Mikulić

Dario Pavlović

Matija Šantl

Robert Vaser

**siječanj, 2014.**

# Sadržaj:

|  |           |
|--|-----------|
| <b>1. Projektni zadatak .....</b>  | <b>1</b>  |
| 1.1 OPIS PROJEKTOG ZADATAKA .....  | 1         |
| 1.2 PREGLED I OPIS SRODNIH RJEŠENJA .....  | 1         |
| 1.2.1 <i>Silhouette recognition using high-resolution pursuit .....</i>  | <i>2</i>  |
| 1.2.2 <i>Fourier Preprocessing for Hand Print Character Recognition .....</i>                                  | <i>2</i>  |
| 1.2.3 <i>Soft Biometrics; Human Identification using Comparative Descriptions .....</i>                        | <i>4</i>  |
| 1.2.4 <i>Silhouette-based Human Identification from Body Shape and Gait .....</i>                              | <i>5</i>  |
| 1.2.5 <i>A Gaussian Process Guided Particle Filter for Tracking 3D Human Pose in Video .....</i>               | <i>6</i>  |
| 1.2.6 <i>Automatic Monocular System for Human Fall Detection Based on Variations in Silhouette Area .....</i>  | <i>8</i>  |
| 1.2.7 <i>Human Identity and Gender Recognition From Gait Sequences With Arbitrary Walking Directions .....</i> | <i>10</i> |
| 1.2.8 <i>Probabilistic gait modelling and recognition .....</i>  | <i>12</i> |
| 1.3 KONCEPTUALNO RJEŠENJE ZADATAKA .....   | 14        |
| <b>2. Postupak rješavanja zadatka .....</b>  | <b>16</b> |
| 2.1 IZLUČIVANJE SILUETE IZ FOTOGRAFIJE .....   | 16        |
| 2.1.1 <i>Zamućivanje fotografije Gausovim filtrom .....</i>  | <i>17</i> |
| 2.1.2 <i>Uklanjanje pozadine .....</i>   | <i>17</i> |
| 2.1.3 <i>Izlučivanje siluete za klasifikaciju .....</i>  | <i>19</i> |
| 2.2 IZLUČIVANJE ZNAČAJKI IZ SILUETE .....  | 19        |
| 2.2.1 <i>Granlundovi koeficijenti .....</i>  | <i>20</i> |
| 2.2.2 <i>Huovi koeficijenti .....</i>  | <i>20</i> |
| 2.3 ODREĐIVANJE IDENTITETA OSOBE .....   | 22        |
| 2.3.1 <i>Bayesov klasifikator .....</i>  | <i>22</i> |
| 2.3.2 <i>KNN klasifikator .....</i>  | <i>23</i> |
| 2.3.3 <i>Random Forest klasifikator .....</i>  | <i>23</i> |
| <b>3. Ispitivanje rješenja .....</b>   | <b>24</b> |
| 3.1 ISPITNA BAZA .....   | 24        |
| 3.1.1 <i>Prošlogodišnja baza .....</i>   | <i>24</i> |
| 3.1.2 <i>Vlastita baza .....</i>   | <i>24</i> |
| 3.2 REZULTATI UČENJA I ISPITIVANJA .....   | 25        |
| 3.2.1 <i>Vlastita baza .....</i>   | <i>25</i> |
| 3.2.2 <i>Prošlogodišnja baza .....</i>   | <i>28</i> |
| 3.2.3 <i>Usporedba rezultata s obzirom na korištene baze .....</i>   | <i>31</i> |
| 3.3 ANALIZA REZULTATA .....  | 32        |
| <b>4. Opis programske implementacije rješenja .....</b>  | <b>33</b> |
| 4.1 STRUKTURA PROGRAMSKE IMPLEMENTACIJE .....  | 33        |
| 4.2 PARAMETRI NA ULAZU PROGRAMSKE IMPLEMENTACIJE .....   | 33        |
| 4.3 KORIŠTENJE PROGRAMSKE IMPLEMENTACIJE .....   | 35        |
| <b>5. Zaključak .....</b>  | <b>36</b> |
| <b>6. Literatura .....</b>   | <b>37</b> |

# 1. Projektni zadatak

## 1.1 Opis projektnog zadatka

Tema projektnog zadatka je raspoznavanje osobe na temelju slike. Ideja je da računalni program kao ulaz dobije fotografiju iz određenog skupa fotografija, a kao izlaz korisniku preda identifikacijsku oznaku osobe koja se nalazi na fotografiji.

Ovakav postupak potrebno je podijeliti na nekoliko dijelova. Prije svega, potrebno je iz fotografije na kojoj se nalazi osoba izlučiti siluetu te osobe, a zanemariti ostatak (pozadinu). Zbog toga se svaka fotografija mora pretvoriti u crno-bijelu fotografiju na kojoj će jednom bojom biti označena silueta osobe, a drugom pozadina. Ovo se može ostvariti tako da se fotografira sama pozadina bez osobe, a zatim pozadina s osobom u prvom planu te se takve dvije fotografije kombiniraju da bi se dobila željena crno-bijela fotografija. Silueta osobe predstavlja skup određenih „mekih“ biometrijskih značajki kao što su visina, širina ramena, širina trupa itd. Naravno, promatranjem isključivo siluete, gubi se određeni broj značajki poput boje očiju i otisaka prstiju.

Nakon što je silueta izlučena, potrebno je iz nje izvući unaprijed definirane značajke na temelju kojih će se vršiti klasifikacija. Ovakve značajke nazivamo „mekima“, a klasificiranje „meko“. Razlog tomu je što silueta daje vrlo mali broj značajki, od kojih nijedna ne ukazuje jednoznačno o kojoj se osobi radi. Također, gube se mnoge druge značajke koje bi potencijalno pridonijele identificiranju osobe, poput izgleda dlanova, lica i ostalih tjelesnih karakteristika te već spomenutih (boja očiju, otisci prstiju). Zbog svega navedenog, odabir isključivo siluete kao temelja za prepoznavanje osobe komplicira i otežava problem klasifikacije.

## 1.2 Pregled i opis srodnih rješenja

U području klasifikacije na temelju silueta postoji veliki broj radova. Kako je najveći broj tih radova temeljen na video sekvencama, a naš na statičkim slikama, odabrali smo četiri rada za koja smo smatrali da su bliska našem problemu i potencijalnom rješenju te smo ih detaljnije promotrili i saželi u nastavku.

### 1.2.1 *Silhouette recognition using high-resolution pursuit*

U ovom radu predlaže se nova metoda klasifikacije na osnovu siluete korištenjem postupka zvanog "pretraga visoke razlučivosti" (engl. *High Resolution Pursuit*).

Prvo što se u radu definira je veličina zvana CDP (engl. *Centroidal Distance Profile*) koja predstavlja jednodimenzionalni zapis siluete, to jest točaka siluete. Ideja je uzeti ekvidistantne točke na rubu objekta te za njih izračunati sljedeću veličinu:

$$f(n) = \sqrt{(x(n) - X_C)^2 + (y(n) - Y_C)^2}.$$

U ovoj formuli,  $(X_C, Y_C)$  predstavlja koordinate centroida objekta. CDP, kao jednodimenzionalan prikaz, ima taj problem što izostavlja informaciju o kutu između točke i centroida pa je iz tog razloga ta informacija naknadno dodana. Par  $(f(n), kut)$  sad nije ništa drugo doli polarna koordinata točke na rubu objekta u odnosu na centroid kao središte takvog koordinatnog sustava.

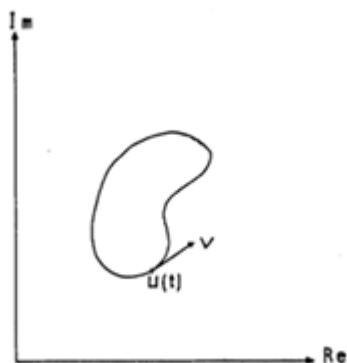
U radu se također navodi važnost otpornosti algoritma na deformacije i razne perturbacije objekta koje mogu utjecati na CDP. Kako bi doskočili tom problemu, autori su simulirali razne neželjene pojave na siluetama koristeći modele Gaussovog šuma. Podatci o CDP-u i kutu prosljeđuju se zatim algoritmu HRP koji na osnovu njih generira vektore značajki vrlo niskih dimenzija koji se zatim koriste za klasifikaciju uzoraka. Algoritam i implementacija su vrlo jednostavni, a rezultiraju jako dobrim rezultatima (boljima od mnogih kompleksnijih metoda).

### 1.2.2 *Fourier Preprocessing for Hand Print Character Recognition*

U ovom radu autor ističe kako je važan problem u raspoznavanju uzoraka naći značajke koje opisuju uzorke na optimalan način, te posebno ističe problem kod raspoznavanja znakova. Problem koji se javlja kod raspoznavanja znakova je mogućnost postojanja različitih stilova pisanja istog znaka, tj. javlja se potreba za nalaženjem značajki koje su nepromjenjive za različite stilove, tipove, širine oblika, nagib oblika i sl. Od takvih značajki se i dalje očekuje da daju dovoljno informacija za jednoznačnu klasifikaciju.

Rad se u velikoj mjeri oslanja na predstavljanju kontura pomoću Fourierovih koeficijenata, te je u nastavku dan njihov kratak pregled.

### 1.2.2.1 Opis konture Fourierovim deskriptorima



Slika 1.1 Primjer konture

Konturu zapisujemo jednodimenzionalnom kompleksnom periodičnom funkcijom  $u(t)$  koja crpi vrijednosti iz skupa kompleksnih brojeva

$$\{(X_t + jY_t): t = 0, \dots, T - 1\}$$

pri čemu je  $(X_t, Y_t)$  točka na konturi. Koeficijente diskretne Fourierove transformacije opisujemo formulom

$$a_n = \frac{1}{T} \int u(t) \cdot e^{-j2\pi \frac{nt}{T}} dt .$$

Konturu želimo opisati značajkama koje ovise samo o obliku konture, a ne o izboru početne točke obilaska, translacije, rotacije ili skaliranja.

Izbor početne točke utječe na iznos koeficijenata Fourierove transformacije na način da ih skalira faktorom  $e^{jnX_0}$ , tj.  $b_n = a_n \cdot e^{jnX_0}$ .

Ako konturu translatiramo za neki vektor  $Z$ , Fourierovi koeficijenti se promijene na sljedeći način:

$$b_n = \begin{cases} a_n + Z, n = 0 \\ a_n, \text{inače} \end{cases}$$

Rotacija izvorne konture može se uvrstiti u funkciju koja predstavlja konturu na sljedeći način:

$$d(t) = u(t) \cdot e^{j\theta}$$

tj. koeficijenti Fourierove transformacije poprimaju oblik  $b_n = a_n e^{j\theta}$ .

Ako polaznu konturu uvećamo za faktor  $R$ , koeficijenti Fourierove transformacije poprimaju oblik  $b_n = R \cdot a_n$ .

Promatrajući utjecaj pojedine transformacije na koeficijente Fourierove transformacije, njihov kumulativan utjecaj možemo prikazati izrazom

$$b_n = e^{jnX_0 + j\theta} \cdot R \cdot a_n .$$

Želimo razlikovati oblik konturu i to tako da je razlikovanje invarijantno na utjecaj izbora početne točke obilaženja konture, pomaka konture, njene rotacije i uvećanja. Rješenje tog problema dao je G.H. Granlund.

### 1.2.2.2 Opis kontura Granlundovim koeficijentima

G.H. Granlund je definirao funkcije Fourierovih koeficijenata koje eliminiraju utjecaj izbora početne točke, pomaka, rotacije i uvećanja konture:

$$d_{pq} = \frac{a_{p+1}^q \cdot a_{T+1-q}^p}{a_1^{p+q}}, \quad p = 1, 2, 3, \dots; \quad q = 2, 3, 4, \dots$$

Lako se pokaže da je takva funkcija otporna na prije spomenute transformacije. Prema tome, funkcije Fourierovih koeficijenata  $d_{pq}$  ovise samo o obliku konture.

### 1.2.3 *Soft Biometrics; Human Identification using Comparative*

#### *Descriptions*

Meke biometrijske značajke su novi oblik biometričke identifikacije koje koriste fizičke ili ponašajne karakteristike koje čovjek može prirodno objasniti. Za razliku od drugih pristupa, ovakav pristup biometrijskoj identifikaciji premošćuje semantičku razliku ljudskog opisa i biometrije.

Kako bi se omogućila identifikacija pomoću ljudskog opisa fizičkih osobina, one moraju biti precizno opisane. Često se takav opis sastoji od apsolutnih kategorijskih oznaka poput visok ili mršav. Uzevši u obzir da ljudi mogu biti netočni prilikom pogađanja mjera, oznake su uzete kao više robusna metoda izdvajanja opisa. Jedan od problema koji se javlja kod toga je da su apsolutne kategorijske oznake podložne subjektivnom dojmu. U nastavku članka je opisan novi način za izdvajanje opisa koji iskorištava proces vizualne usporedbe između subjekata.

Autori navode kako je usporedba dvaju subjekata čest prirodan proces i može biti mnogo pouzdaniji nego korištenje apsolutnih kategorijskih oznaka jer se usporedbe temelje na specifičnim oznakama te time postaju objektivnije.

Skup relativnih mjerenja koja opisuju subjekt može se precizno odrediti iz usporednih oznaka i kao takav koristiti u svrhu biometrijske identifikacije. Svaki se subjekt uspoređuje s nekim drugim čije su biometrijske značajke poznate. Relativna mjerenja prikazuju kako se svojstva subjekta razlikuju od drugih unutar populacije. Takva svojstva imaju visok stupanj korelacije sa stvarnim osobinama. Tako je npr. koeficijent korelacije između visine piksela i

stvarne visine iznosio 0,87, što pokazuje da relativna mjerenja dočaravaju i stvarne iznose fizičkih osobina.

Meke biometrijske značajke nam omogućavaju identifikaciju temeljenu na prirodnom opisu ponašajnih i fizičkih osobina. Identifikacija je moguća samo ako je opis točan i precizan. Konvencionalne forme ljudskog opisa fokusiraju se na apsolutne oznake i procjene koje mogu biti nepouzdana zbog njihove pristranosti. Usporedni opisi sadrže više razlikovnih podataka i pokazuju inovativan pristup dobavljanju robusnih opisa osoba za meke biometrijske opise.

Meke biometrijske značajke također zahtijevaju manje računa u usporedbi s čvrstim biometrijskim značajkama te se time pokazuju kao dobar izbor za sustave za nadgledanje.

#### **1.2.4 *Silhouette-based Human Identification from Body Shape and Gait***

Rad opisuje izradu jednostavnog identifikatora ljudi na temelju oblika tijela i tipa hoda. Podloga ovakve klasifikacije je činjenica da ljudski hod ovisi o mnoštvu faktora koji uključuju fizikalnu građu i težinu tijela, visinu potpetice, mentalno stanje osobe itd. Općenito je poznato da su ljudi sposobni prepoznati poznanike na temelju njihovog hoda. Ugrubo, opisani postupak jednak je klasifikaciji 2D silueta ekstrahiranih iz ključnih prikaza jedne periode hoda. Biometrijske značajke koje se pritom koriste su visina, širina, proporcije dijelova tijela, duljina koraka te zamah rukom.

Metoda se sastoji od sljedećih koraka:

1. *ekstrakcija silueta*: najprije se iz fotografije oduzme pozadina uz određeno odstupanje, nakon čega slijedi filtracija fotografije kako bi se maknuli izolirani pikseli,
2. *analiza periode hoda te odabir ključnih prikaza*: analiza siluete sekvence kako bi se pronašao period i odabir prikaza koji označavaju maksimum ili minimum sinusoida,
3. *izdvajanje predložaka*: centriranje i smanjivanje ključnih prikaza na veličinu  $80 \times 128$  piksela,
4. *usporedba predložaka pomoću normalizirane unakrsne korelacije*<sup>1</sup>: usporedba predložka iz probnog skupa i predložaka iz skupa za učenje; ocjena usporedbe (C) između dva predložka ( $\vec{a}, \vec{b}$ ) dobije se izrazom (operator \* predstavlja operaciju unakrsne korelacije,  $\mu$  srednju vrijednost, a  $\sigma$  standardnu devijaciju):

---

<sup>1</sup> Unakrsna korelacija (engl. *cross-correlation*) – mjera sličnosti dva valna oblika kao funkcija vremenskog zaostatka primijenjenog na jedan od valnih oblika.

$$C(\vec{a}, \vec{b}) = \frac{\max(\alpha * \beta)}{\max(\alpha * \alpha) \max(\beta * \beta)}, \quad \text{uz } \alpha = \frac{a - \mu(a)}{\sigma(a)} \text{ i } \beta = \frac{b - \mu(b)}{\sigma(b)}$$

5. *postupak KNN („k najbližih susjeda“, engl. k nearest neighbours) koristeći kombiniranu vrijednost svih predložaka iz periode hoda: pomoću korelacijskih vrijednosti između predložaka iz probnog skupa i predložaka iz skupa za učenje računaju se četvorke  $(p_i, v_i, p_{i+1}, v_{i+1})$  koje sadrže ključne prikaze iz periode hoda (desni korak, desni međukorak, lijevi korak, lijevi međukorak); relativna vjerojatnost da predložak  $p_i$  odgovara subjektu  $s_0$  dana je izrazom*

$$R(p_i, s_0) = \frac{\max\{C(p_i, P_k^s | s = s_0)\}}{\max\{C(p_i, P_k^s)\}}$$

Klasifikacija (najbolje podudaranje)  $s^*$  dana je izrazom:

$$s^* = \operatorname{argmax}_s [R(p_i, s) + R(v_i, s) + R(p_{i+1}, s) + R(v_{i+1}, s)]$$

Testiranje otpornosti klasifikatora na šum u podacima provedeno je pomoću baza fotografija koje su snimljene iz različitih kutova, u različitim okolinama (u nekoj prostoriji i vani) te su snimljeni različiti tipovi hoda.

Rezultati su se pokazali zadovoljavajući dok su sve fotografije bile snimljene iz jednakog kuta, limit koji je posljedica klasifikacije na temelju 2D siluete. Čak se pokazalo da metoda dobro klasificira ljude čiji se hod (brzi koraci, vođenje lopte) razlikovao od tipa hoda za učenje. Klasifikator podnosi i šumovite siluete, poput onih dobivenih iz nadzornih kamera te ako je snimljen samo jedan korak (ne cijeli period).

### **1.2.5 A Gaussian Process Guided Particle Filter for Tracking 3D Human Pose in Video**

Praćenje i/ili detekcija položaja ljudskog tijela na slikama ili video zapisima je popularno područje zbog široke primjene u video nadzoru, zaštiti i interakciji čovjek-računalo. Kao primjer možemo uzeti pametan sustav video nadzora, 3D položaj tijela se može koristiti kao naznaka za zbivanja na promatranome području i tako otkriti sumnjiva ponašanja. Također, može pružati napredniji pristup interakciji čovjeka i računala u pogledu računalnih igara i proširene stvarnosti. 3D položaj osobe koji se traži na nizu slika iz videozapisa može biti koristan za biometrijske primjene prepoznavanja osobe. Takve primjene zahtijevaju jednostavne videozapise ili mirne slike kao ulaz, i kao rezultat vraćaju jednostavno rješenje, za razliku od sustava zasnovanih na markerima. Procjena položaja ljudskog tijela i sustavi za praćenje položaja tijela mogu se klasificirati prema različitim pristupima u tri grupe: diskriminativni, generativni i hibridni.



Kod generativnih metoda, procjena položaja tijela se radi pretraživanjem prostora rješenja za položajem koji najbolje opisuje gledane slike. U ovom pristupu, generativni model je konstruiran na način da mjeri koliko je hipoteza bliska promatranim značajkama. Hipoteza koja je najkonzistentnija s promatranim značajkama slike je odabrana kao izlaz, točnije, položaj tijela na slici. Filtar čestica je generativna metoda praćenja za određivanje položaja u svakoj slici videozapisa koja za procjenu trenutne slike koristi prethodnu i dosad skupljene informacije o videozapisu. S danom slikom ljudskog subjekta, postoji više 3D ljudskih položaja koji su povezani s tom slikom. Takvi nejednoznačni položaji mogu se razriješiti koristeći slike više kamera iz različitih kutova.

Kod diskriminativnih metoda, pomoću nadgledanog učenja dobivamo regresijsku funkciju koja mapira značajke slike s danim položajem tijela. Iako diskriminativne metode mogu brzo procijeniti 3D položaj, one također mogu dati netočne procjene na novim ulazima ako je učenje provedeno na malim bazama podataka. K tome, veze između značajki slike i položaja ljudskog subjekta su često višemodalne. Kao primjer možemo uzeti slučaj kad se za značajke slike koristi silueta tijela. Jedna se silueta može povezati s više od jednog položaja, što rezultira nejednoznačnošću. U takvim slučajevima, višestruki diskriminativni modeli su potrebni da naprave jedan-naprema-više veze između značajki slike i položaja.

Diskriminativne metode su značajne za specifične zadatke, poput procjene položaja ljudskog tijela, budući da se one temelje samo na mapiranju između ulaza i željenog izlaza. Generativne metode su u tom pogledu fleksibilnije jer pružaju puno više prostora za korištenje djelomičnih saznanja o rješenju i iskorištavaju model ljudskog tijela za otkrivanje prostora rješenja. Zbog različitog načina procjene rješenja, ove se dvije metode smatraju međusobno komplementarnima. Pokazalo se da hibridne generativne i diskriminativne metode imaju potencijal poboljšati performanse pronalaženja položaja tijela. Kao rezultat toga, u zadnje su vrijeme poprimile puno pažnje.

Kod hibridnih metoda, diskriminativno mapiranje funkcije se koristi za generiranje položaja prostora hipoteza. Položaj je zatim procijenjen pretraživanjem prostora hipoteza korištenjem generativne metode. Ove metode pretpostavljaju da je prostor hipoteza generiran diskriminativnim modelom uvijek točan i zbog toga ne uspijevaju savladati slučaj kada diskriminativni modeli ne daju točne rezultate.

U ovom članku autori predlažu hibridnu diskriminativnu i generativnu metodu za praćenje 3D položaja ljudskog subjekta koristeći sliku s jedne i više kamera. Za diskriminativne modele, koristimo mješavinu Gaussovog procesa regresijskih modela koji se dobivaju treniranjem Gaussovih procesa na različitim područjima prostora položaja tijela.

Takav pristup pruža efikasan način uključivanja više pouzdanih diskriminativnih predikcija u proces praćenja prilikom kojeg odbacuje one nesigurne predikcije. Prema saznanjima autora, ovo je prva hibridna metoda koja u obzir uzima prediktivnu nesigurnost diskriminativnog modela i kombinira je s generativnim modelom da bi poboljšao procjenu položaja. Vjerojatnosni izlaz Gaussovog regresijskog procesa tretira se kao jedna komponenta u prostoru hipoteza. U koraku praćenja, kombinira se taj prostor hipoteza s prostorom hipoteza koji je dobiven iz modela kretanja i s onim koji se dobije pretraživanjem optimalnog položaja korištenjem prekaljenog filtra čestica.

Autori ovog članka predlažu hibridnu metodu koja kombinira Gaussovu metodu učenja, filter čestica i prekaljivanje u svrhu praćenja 3D oblika koji predstavlja ljudsko tijelo u videozapisu. Njihov pristup, kojeg oni nazivaju prekaljeni Gaussov proces vođen filtrom čestica, sastoji se od dva koraka. U koraku učenja koriste nadzirano učenje da nauče Gaussov proces koji prima opis siluete kao ulaz i vraća više izlaza koji predstavljaju položaj koji je modeliran mješavinom Gaussovih distribucija. U koraku praćenja, distribucije izlaza položaja iz Gaussovog procesa regresije se kombiniraju s prekaljenim filtrom čestica za praćenje 3D položaja u svakoj slici videozapisa. Njihovi eksperimenti su pokazali da predložene metode ne zahtijevaju inicijalizaciju i ne gube subjekt praćenja - položaj tijela. Usporedili su svoj pristup sa standardnim prekaljenim filtrom čestica koristeći *HumaEva-I* bazu podataka i druge novije pristupe koristeći *HumanEva-II* bazu podataka. Rezultati evaluacije su pokazali da njihov pristup može uspješno pratiti 3D položaje ljudskog tijela preko duljeg vremenskog perioda i da može dati preciznije rezultate položaja tijela od prekaljenog filtra čestica.

### **1.2.6 Automatic Monocular System for Human Fall Detection Based on Variations in Silhouette Area**

Populacije starih ljudi rastu, a zbog standarda života veliki broj tih starih ljudi ostaje samo kod kuće. Jedna od najgorih nezgoda koja ih može zadesiti je pad. Iz tih razloga, autori ovog članka odlučili su ostvariti sustav za detekciju pada.

Za detekciju pada osobe postoje izgrađeni različiti uređaji; nosivi uređaji te uređaji ugrađeni u okolinu. Nosivi uređaji jednostavniji su, ali narušavaju uobičajene aktivnosti osobe. Također, postoji mogućnost da će ih osoba zaboraviti staviti. Uređaji koji se ugrađuju u okolinu neće narušavati normalne aktivnosti osobe, ali su podložni uzrokovanju čestih lažnih alarma. Drugačiji pristup detekciji pada temelji se na snimanju kamerom. Kako bi snimanje kamerom moglo narušiti privatnost osobe, autori članka iznose sljedeću ideju: u

kameri bi bio implementiran sustav za detekciju pada koji bi aktivirao alarm pa izlaz kamere ne bi bila slika, već alarm.

Autori članka testirali su sustav za detekciju pada pomoću jedne kamere, sa fiksnim područjem gledanja. Sustav za detekciju temelji se na procjeni pada pomoću silueta. Prvi korak je lociranje objekta u pokretu te određivanje siluete takvog objekta. Na temelju objekta određuju se značajke siluete i provodi se klasifikacija. Ukoliko se detektira pad, aktivira se alarm.

### 1.2.6.1 Određivanje siluete

Određivanje siluete provede na su pomoću dva algoritma: *Running Average Method* i *Modified Running Average Method*.

#### 1.2.6.1.1 *Running Average Method*

Intenzitet svakog piksela koji pripada stacionarnom dijelu kadra određen je formulom

$$b_t = \gamma I_t + (1 - \gamma)b_{t-1} \quad (3).$$

$I_t$  je trenutni intenzitet za neki kadar  $t$ ,  $b_{t-1}$  je intenzitet piksela u prethodnom kadru,  $\gamma$  je realan broj od 0 do 1. Za svaki kadar  $t$ , ukoliko je  $|I_t - b_t|$  veća od neke postavljene vrijednosti  $T$ , piksel pripada objektu u pokretu.

Izlaz je binarna slika, gdje crni dijelovi određuju stacionaran dio slike, dok su bijeli dijelovi objekt u pokretu.

#### 1.2.6.1.2 *Modified Running Average Method*

Utvrđeno je da prvi algoritam ima određenih mana. Kao dio objekta u pokretu bit će i put kojim je objekt prošao. Razlog tome leži u činjenici da algoritam obavlja provjeru za svaki piksel. Kada objekt stoji na određenoj poziciji, on postaje dio stacionarne okoline i prestaje biti dio siluete. Također, promjenom položaja mijenja se intenzitet na toj poziciji, pa će algoritam zabilježiti promjenu i taj dio postat će dio objekta u pokretu.

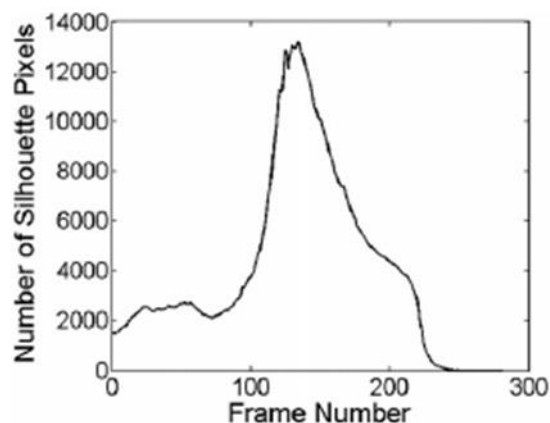
Modifikacija prethodnog algoritma prikazana je formulom

$$b_t = Mb_{t-1} + (1 - M)[\gamma I_t + (1 - \gamma)b_{t-1}] \quad (4).$$

Dodan je parametar  $M$ , gdje za objekt u pokretu  $M = 1$ , a stacionarni dio  $M = 0$ . Prema tome, modifikaciju vršimo samo ako je piksel pripadao stacionarnom dijelu kadra.

### 1.2.6.2 Određivanje značajki

Prilikom pada, brzina osobe se naglo povećava, osoba udara u pod i ostaje neko vrijeme u takvom položaju. Takav scenariji može se otkriti koristeći siluete prema formuli (3). Kako se brzina povećava, broj piksela siluete raste, ukoliko osoba ostaje ležati na podu, postati će dio stacionarne slike i broj piksela siluete pada na 0, prikaz na slici Slika 1.2.



Slika 1.2 Broj piksela prilikom pada

Određena su dva vektora značajki. Prvi vektor značajki određen je siluetom prema formuli (3), dok je drugi vektor značajki određen kao razlika između silueta generiranih formulom (3) i formulom (4).

### 1.2.6.3 Klasifikacija

Klasifikacija je provedena kombiniranim SVM. Kako postoje dva vektora značajki, svaki vektor značajki koristi svoj SVM, a izlaz svakog SVM-a prosljeđuje se zajedničkom SVM-u koji provodi klasifikaciju.

### 1.2.6.4 Rezultati

Predstavljeni sustav je jednostavan, zahtjeva samo jednu kameru te neće narušiti normalne aktivnosti korisnika, niti njegov privatni život. Testiranje je provedeno na javno dostupnoj bazi „*Multiple cameras fall dataset*“<sup>2</sup> i dobivena pogreška iznosi oko 5%.

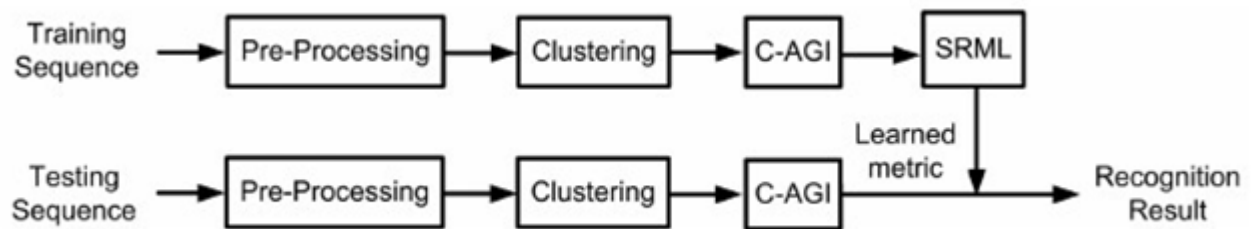
### 1.2.7 Human Identity and Gender Recognition From Gait Sequences With Arbitrary Walking Directions

Autori ovoga članka su Jiwen Lu, Gang Wang i Pierre Moulin. Ovaj članak uzeli smo u obzir kako bismo ilustrirali na koje je sve načine moguće koristiti siluete, što će se pokazati

<sup>2</sup> <http://www.iro.umontreal.ca/~labimage/Dataset/>

kasnije u tekstu. Problem koji autori obrađuju je problem prepoznavanja identiteta i spola osoba preko načina na koji se ona kreće (engl. *gait*). Autori napominju kako je problem koji oni rješavaju analiziran i prije, ali se načini rješavanja razlikuju. Naime, većina radova ovog tipa bavi se analizom scena u kojima se osobe kreću u predefinisanim smjerovima, što u principu nije realistično. Zbog toga oni, uz pomoć *Microsoft Kinecta*, generiraju novi skup podataka, tj. kratkih video sekvenci u kojima se osobe kreću u slobodnim, ne predodređenim, smjerovima. Također, autori u članku predlažu novu metodu učenja metrike koja, kako se kasnije pokazuje, bolje klasificira primjere iz tog skupa.

Njihov je pristup shematski prikazan na slici Slika 1.3. Prvo se iz svake sekvence, to jest videa, izluče siluete oduzimanjem pozadine te se one grupiraju u grozdove ili skupine (engl. *cluster*). Tad se za svaku skupinu izračuna skupno bazirana prosječna vrijednost karakteristike *gaita* (engl. *Cluster-based averaged gait image*, C-AGI) kao karakteristika te skupine. Zatim se provodi postupak učenja takav da minimizira varijancu unutar razreda, a maksimizira varijancu između razreda. Detaljniji postupak opisan je u nastavku.



**Slika 1.3 Dijagram toka predloženog pristupa**

Prvi korak je pretprocesiranje u kojem se iz svakog videa izlučuju siluete te se one potom skaliraju na rezoluciju  $64 \times 44$  kako bi bile otporne na udaljenost kamere od osobe. Ovo je logičan korak jer osoba hoda bliže ili dalje kameri, ovisno o trenutku. Zatim se ovako izlučene siluete grupiraju u grupe. Problem u ovom koraku je što broj grupa nije unaprijed poznat te se popularni algoritam „*k* srednjih vrijednosti“ ne može koristiti. Autori stoga koriste metodu zvanu *affinity propagation* koja omogućuje da se primjeri rasporede u grupe bez da se broj grupa zada unaprijed. Eksperimentalno ispada da je broj grupa uglavnom između 8 i 15.

Nakon dobivanja grupa, slijedi korak računanja gore spomenutog C-AGI prema sljedećoj formuli:

$$G_k(x, y) = \frac{1}{N_k} \cdot \sum I_{pk}(x, y)$$

gdje je  $N_k$  broj primjera u grupi  $k$ , a  $I_{pk}$   $p$ -ta silueta u  $k$ -toj grupi. Autori navode kako je ovakvo izlučivanje značajki dobro po performansama, a vrlo robusno na potencijalne šumove

i probleme kakvi prirodno nastaju u ovakvim zadacima. Ovaj se postupak ponavlja kako bi se za svaku sekvencu dobilo više C-AGI značajki. Skup za učenje radi se na način da se za svaku osobu stvori jedan vektor značajki. Taj vektor značajki sadrži kao svaku komponentu jedan izračunati  $G_k$ . Nakon dobivanja svih vektora značajki, ide se u proces učenja koji ima za cilj što više smanjiti varijacije unutar razreda, a maksimizirati varijacije među razredima. Proces se svodi na učenje pogodne metrike, krenuvši s euklidskom, i to algoritmom sličnim poznatom EM-algoritmu.

Nakon što je proces učenja proveden, provodi se postupak testiranja. Autori kažu kako su testirali tako da su slučajno podijelili bazu na dva jednaka dijela, odabrali od toga skup za učenje, a ostalo postavili kao skup za testiranje. Ovaj su postupak ponovili deset puta, a dimenziju vektora značajki pri tome postavili na 50. S obzirom da nisu prvi koji su se bavili ovom problematikom, rezultate uglavnom prezentiraju kroz usporedbu s drugima. Ispada da se s ovakvom metodom može postići točnost i do 94%. Naravno, pri tome prezentiraju i potencijalni problem koji imaju, a to je da, ako ista osoba hoda potpuno različito u dvije različite sekvence, klasifikacija postaje pogrešna. Ovo je vrlo relevantno problemu koji mi rješavamo iz razloga što, također, kako će se kasnije vidjeti, imamo problema s prepoznavanjem iste osobe u vrlo distinktnim pozama na fotografiji.

### 1.2.8 Probabilistic gait modelling and recognition

Autori ovog rada predstavljaju drugačiju metodu za raspoznavanje hoda koja se naziva probabilistički model hoda (engl. *probabilistic gait model*, PGM). Temeljne ideje ovakvog načina raspoznavanja su:

1. tretirati oblik siluete kao multivarijantnu slučajnu varijablu
2. modelirati njenu statističku distribuciju
3. iskoristiti statistički model za raspoznavanje

Silueta, koja se izluči iz sekvence videa, je binarna slika gdje svaki piksel ima vrijednost 1 s vjerojatnošću  $p$  ili 0 s vjerojatnošću  $1 - p$ . Neka je:

- $X = \{x^1, \dots, x^T\}$  skup silueta iz jedne periode hoda za određeni objekt ( $T$  - duljina periode)
- $x^t = (x_1^t, \dots, x_D^t) \in \{0,1\}^D$ , gdje je  $D$  veličina siluete (*visina*  $\times$  *širina*)

Svaki element  $x^d$  je binarna slučajna varijabla s Bernoullijevom distribucijom

$$p(x_d|\mu_d) = \mu_d^{x_d} \cdot (1 - \mu_d)^{1-x_d}$$

Cjelokupna se binarna silueta stoga opisuje izrazom

$$p(x | \mu) = \prod_{d=1}^D \mu_d^{x_d} \cdot (1 - \mu_d)^{1-x_d}$$

Bernoullijeva mješavina, koja služi za različito modeliranje silueta u periodi hoda, dana je izrazom

$$p(x|\mu, \pi) = \sum_{m=1}^M \pi_m \cdot p(x|\mu_m)$$

gdje vrijedi:

- $M$  - broj komponenata u mješavini
- $\mu = (\mu_1, \dots, \mu_M) \in [0,1]^{D \times M}$
- apriorna vjerojatnost za komponente mješavina  $\pi = (\pi_1, \dots, \pi_M)^T \in [0,1]^M$  za koju vrijedi  $\sum_{m=1}^M \pi_m = 1$

Probabilističko modeliranje hoda u 4 koraka:

1. inicijalizacija Bernoullijeve matrice  $\mu = (\mu_1, \dots, \mu_M) \in [0,1]^{D \times M}$  te koeficijenata mješavine  $\pi = (\pi_1, \dots, \pi_M)^T \in [0,1]^M$
2. evaluacija posteriornih vjerojatnosti komponenata mješavine  $\tau_m^t$ :
  - $\tau_{m,0}^t = \pi_m$
  - $\tau_{m,d+1}^t = \frac{\mu_{m,d}^{x_{d+1}^t} \cdot (1-\mu_{m,d})^{1-x_{d+1}^t} \cdot \tau_{m,d}^t}{\sum_{m=1}^M \mu_{m,d}^{x_{d+1}^t} \cdot (1-\mu_{m,d})^{1-x_{d+1}^t} \cdot \tau_{m,d}^t}, d = 0, \dots, D-1$
  - $\tau_m^t = \tau_{m,D}^t$
3. računanje novih parametara modela PGM
  - $\mu_{m,d_{novi}} = \frac{\sum_{t=1}^T \tau_m^t \cdot x_d^t}{\sum_{t=1}^T \tau_m^t}$
  - $\pi_{m_{novi}} = \frac{T_m}{T}$
4. provjera konvergencije vjerojatnosti ili parametara  $\rightarrow$  ako je prisutna divergencija, ažuriraj parametre i vrati se na korak 2.
  - $\mu_{m,d} = \mu_{m,d_{novi}}$
  - $\pi_m = \pi_{m_{novi}}$

Dobiveni parametri dalje se koriste u klasifikaciji hoda pomoću Bayesovog klasifikatora. U nastavku je dan algoritma za klasifikaciju jedne siluete te klasifikaciju cjelokupne sekvence silueta respektivno:

Klasifikacija jedne siluete:

1. inicijalizacija združene vjerojatnosti subjekta i komponente mješavine

- $\gamma_{m,0}^\omega = p(\omega) \cdot \pi_m^\omega$
- 2. rekurzivna evaluacija združene vjerojatnosti
  - $\gamma_{m,d+1}^\omega = p(\omega, m | x_{1:d+1}), \quad d = 0, \dots, D - 1$
  - $\gamma_{m,D}^\omega = p(\omega, m | x)$
- 3. marginalizacija komponente mješavine
  - $p(\omega | x) = \sum_{m=1}^M p(\omega, m | x)$
- 4. određivanje identiteta subjekta
  - $\operatorname{argmax}_{\omega=1}^C p(\omega | x)$

Klasifikacija sekvence silueta:

1. inicijalizacija posteriorne vjerojatnosti subjekta
  - $\xi_0^\omega = p(\omega)$
2. rekurzivna evaluacija posteriorne vjerojatnosti
  - $\xi_{t+1}^\omega = \frac{\frac{p(\omega | x_{t-1})}{p(\omega)} \xi_t^\omega}{\sum_{\sigma=1}^C \frac{p(\sigma | x_{t-1})}{p(\sigma)} \xi_t^\sigma}, \quad t = 0, \dots, T - 1$
  - $\xi_t^\omega = p(\omega | X)$
3. određivanje identiteta subjekta
  - $\operatorname{argmax}_{\omega=1}^C p(\omega | X)$

Najveća odlika ovako opisanog modela je velika točnost raspoznavanja čak i kod kamera male rezolucije, dok je mana velika računalna složenost cijelog procesa.

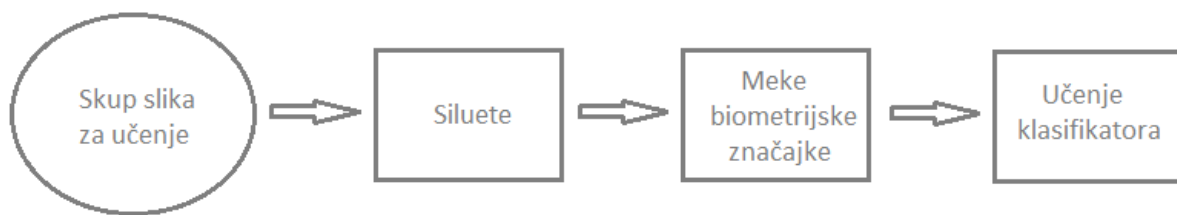
### 1.3 Konceptualno rješenje zadatka

Problem raspoznavanja osoba na temelju siluete je problem klasifikacije koji rješavamo postupcima nadziranog učenja. Postupci nadziranog učenja svode se na učenje klasifikatora i samu klasifikaciju zadanog objekta.

Učenje klasifikatora može se razložiti na 3 potproblema:

1. iz zadanog skupa za učenje grade se siluete,
2. na temelju silueta određuju se biometrijske značajke te
3. na temelju biometrijskih značajki uči se klasifikator.





**Slika 1.4 Učenje klasifikatora**

Skup za učenje sadrži slike osobe i slike pozadine. Za svaku sliku osobe, potrebno je imati sliku same pozadine. To omogućava izdvajanje osobe iz slike. Ukupan broj slika u bazi izražen je formulom

$$brojSlika = 2 \cdot [3,5] \cdot brojOsoba.$$

Iz svakog para slika (slika osobe, slika pozadine) izdvaja se osoba i gradi silueta. Silueta je binarna slika, gdje 0 označava pozadinu, a 1 osobu. Za svaku dobivenu siluetu određuje se vektor značajki, koji se izračunava korištenjem Granlundovih, kao i Huovih koeficijenata.

Izračunati vektori značajki prosljeđuju se programu za učenje klasifikatora. Prilikom izrade projekta odlučili smo paralelno naučiti i zatim ispitati tri klasifikatora: Bayesov, KNN i *Random Forest*.

## 2. Postupak rješavanja zadatka

### 2.1 Izlučivanje siluete iz fotografije

Kako bismo izlučili siluetu osobe iz fotografije, potrebno je znati gdje se na njoj nalazi osoba u odnosu na pozadinu. Stoga, u ovom se koraku kao ulaz predaju dva parametra:

1. putanja do fotografije na kojoj se nalazi osoba u nekom okruženju
2. putanja do fotografije na kojoj se nalazi prazna pozadina - samo okruženje, bez osobe.

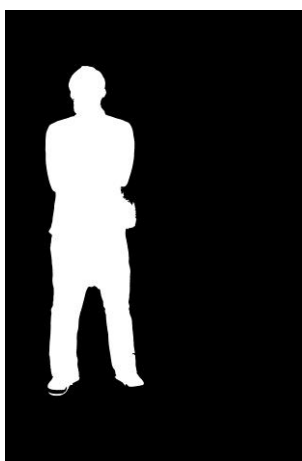


Slika 2.1 Osoba u okruženju



Slika 2.2 Samo okruženje

Osnovna ideja ovog koraka jest pronaći piksele u kojima se dane fotografije razlikuju i na temelju toga izlučiti siluetu osobe, koja se predaje kao izlaz.



Slika 2.3 Primjer izdvojene siluete

Iako je ideja prilično jednostavna, njenu izvedbu kompliciraju čimbenici poput potencijalnih razlika u osvjetljenju prilikom fotografiranja same pozadine i osobe, način na koji prisustvo osobe u nekom okruženju utječe na to okruženje (npr. sjene) itd. Stoga je prvo

potrebno obraditi dane fotografije kako bismo maksimalno poništili utjecaj tih čimbenika na rezultate klasifikacije.

Budući da će se pri objašnjavanju rješenja često spominjati fotografija na kojoj je osoba i fotografija pozadine, ubuduće ćemo ih zvati fotografija O („fotografija osobe“) i fotografija P („fotografija pozadine“).

### 2.1.1 Zamućivanje fotografije Gaussovim filtrom

Svrha zamućivanja fotografije jest umanjiti utjecaj već spomenutih čimbenika koji mogu smanjiti točnost klasifikacije. Ideja algoritma jest primijeniti Gaussovu funkciju na svaki piksel fotografije, a rezultat je fotografija s manje detalja.

Matematički, izračunavanje novih vrijednosti piksela izvodi se formulom

$$G(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}}.$$

U ovom razmatranju, smatra se da gornji lijevi kut fotografije ima oznaku (0,0), a parametri  $x$  i  $y$  predstavljaju odmak pojedinog piksela udesno odnosno prema dolje u odnosu na gornji lijevi kut. Nadalje,  $\sigma$  predstavlja standardnu devijaciju Gaussove distribucije.

Rezultat ove operacije jest površina čije su konture koncentrične kružnice distribuirane Gaussovom distribucijom iz središnje točke. Vrijednosti distribucije se zatim koriste kako bi se izgradila konvolucijska matrica koja se potom primijeni na originalnu sliku, što za posljedicu ima da je nova vrijednost svakog piksela težinska sredina njegovih susjeda.

Ovaj pomoćni algoritam se koristi za uklanjanje smetnji koje bi utjecale na izlučivanje siluete pri postupku uklanjanja pozadine iz fotografije O.

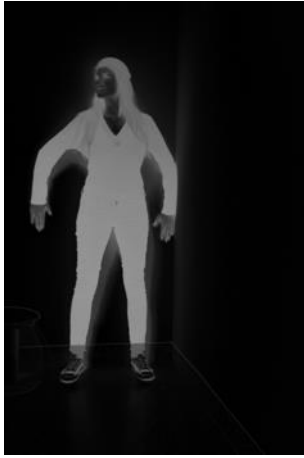
### 2.1.2 Uklanjanje pozadine

Uklanjanje pozadine se vrši na jedan od dva načina, koji su opisani u nastavku. Isprobano je uklanjanje pozadine korištenjem rastava po bojama te korištenjem ljestvice sivih tonova. U oba slučaja se kao ulaz predaju fotografije O i P, a kao izlaz se dobije binarna fotografija u kojoj je crnim tonom označeno podudaranje između O i P, a bijelim mjesta gdje se one razlikuju.

Pri objašnjavanju postupaka koristit će se matematičke formule u kojima ćemo koristiti sljedeće oznake:

- gornji lijevi kut fotografije ima oznaku (0,0), a parametri  $x$  i  $y$  predstavljaju odmak pojedinog piksela udesno odnosno prema dolje u odnosu na gornji lijevi kut,
- indeksi  $r$ ,  $g$  i  $b$  predstavljaju crvenu, zelenu i plavu komponentu fotografija,

- indeks *gray* označava fotografiju u sivim tonovima, te
- *diff* označava fotografiju koja se dobije oduzimanjem fotografije P od O.



Slika 2.4 Uklanjanje pozadine korištenjem ljestvice sivih tonova



Slika 2.5 Uklanjanje pozadine korištenjem rastava po bojama

### 2.1.2.1 Uklanjanje pozadine korištenjem ljestvice sivih tonova

Ovaj postupak započinje pretvaranjem boja na fotografijama P i O u odgovarajuće sive tonove, što se izvodi po formuli

$$X_{gray}(x, y) = X_r(x, y) \cdot 0,299 + X_g(x, y) \cdot 0,587 + X_b(x, y) \cdot 0,114, \quad X = O, P.$$

Ovom formulom ograničava se mogući raspon tonova na interval  $[0, 255]$  te se postiže razlikovanje piksela čiji bi zbrojevi  $r$ ,  $g$  i  $b$  komponenti bili jednaki.

Nakon obavljene pretvorbe, obje fotografije se zamute Gausovim filtrom i potom se napravi nova fotografija koja predstavlja njihovu razliku, a dobije se primjenom formule

$$diff(x, y) = \max(\min(O_{gray}(x, y) - P_{gray}(x, y), 255), 0).$$

Dobivenu fotografiju još jednom zamutimo Gausovim filtrom te predamo kao izlaz.

### 2.1.2.2 Uklanjanje pozadine korištenjem rastava po bojama

Kod ovog pristupa uklanjanje pozadine započinje rastavom fotografija O i P na crvenu, plavu i zelenu komponentu (*engl.* RGB). Ovakav rastav pridonosi smanjenju utjecaja vanjskih čimbenika na odvajanje pozadine od osobe. Za svaku komponentu se provede oduzimanje fotografije P od fotografije O po formuli

$$diff_i(x, y) = |O_i(x, y) - P_i(x, y)|, \quad i = r, g, b.$$

Nakon oduzimanja se nastale komponente zamute Gausovim filtrom i ujedine natrag u jednu sliku sivih tonova zbrajanjem po formuli

$$diff(x, y) = \max(\min(diff_r(x, y) + diff_g(x, y) + diff_b(x, y), 255), 0).$$

Razlog ograničenja zbrajanja jest raspon RGB ljestvice (0-255). Ona se potom još jednom zamuti Gaussovim filtrom i zatim preda kao izlaz.

### 2.1.3 Izlučivanje siluete za klasifikaciju

Nakon što smo iz fotografije izuzeli pozadinu, dobili smo fotografiju sivih tonova koja predstavlja područja u kojima se razlikuju P i O. Kako bismo izlučili siluetu koju želimo klasificirati, potrebno je načiniti binarnu fotografiju te odrediti koja od prisutnih silueta jest ona koju želimo klasificirati. Naime, može se dogoditi da na binarnoj fotografiji imamo više silueta koje se mogu protumačiti kao predmet klasifikacije, uslijed pogreške pri oduzimanju pozadine zbog raznih smetnji, prisustva više od jedne osobe na fotografiji, dodatnog predmeta koji se pojavio u pozadini ili zbog nekog drugog razloga. Kako očekujemo da će takve pogreške generirati relativno male siluete u odnosu na siluetu osobe koju želimo klasificirati, pretpostavljamo da je silueta koju tražimo ona silueta s fotografije čija kontura ima najveći opseg.

#### 2.1.3.1 Pretvaranje fotografije sivih tonova u binarnu fotografiju

Kako bismo fotografiju sivih tonova pretvorili u binarnu (crno-bijelu) fotografiju, uz samu fotografiju *diff* kao ulaz dajemo i graničnu vrijednost (*threshold*) za određivanje boje pojedinog piksela.

Dalje je ideja vrlo jednostavna, za svaki piksel se primjeni funkcija

$$bin(x, y) = \begin{cases} 0, & diff(x, y) \leq threshold \\ 1, & diff(x, y) > threshold \end{cases}, \text{ gdje } bin \text{ označava binarnu fotografiju.}$$

#### 2.1.3.2 Određivanje konture siluete za klasifikaciju

Za izlučivanje kontura koristi se algoritam opisan u (1). Nakon što smo pronašli sve konture koje se pojavljuju na fotografiji, pronađemo skup piksela koji čine konturu najveće duljine. Pronađene piksele zatim iskoristimo kako bismo iscrtali binarnu fotografiju koja sadrži samo siluetu koju oni opisuju.

## 2.2 Izlučivanje značajki iz siluete

Izlučivanje značajki iz siluete jest postupak kojim se stvori vektor značajki spreman za korištenje od strane klasifikatora. Kao ulaz ovog koraka predaje se generirana binarna fotografija, a izlaz jest vektor značajki siluete. Nezavisno su implementirana dva načina za stvaranje vektora značajki: Granlundovi koeficijenti i Huovi koeficijenti.

### 2.2.1 Granlundovi koeficijenti

Fourierovi koeficijenti su jedan od najpopularnijih načina za raspoznavanje oblika na temelju njihovih kontura. Ideja postupka jest gledati na konturu oblika kao na krivulju koja se može prikazati periodičnom funkcijom  $u(t)$  koja poprima vrijednosti iz skupa kompleksnih brojeva  $\{(x_t + jy_t): t = 0, 1, \dots, T-1\}$ . U navedenoj formuli  $T$  predstavlja broj točaka koje će se koristiti za opis krivulje, dok  $x_t$  i  $y_t$  predstavljaju koordinate pojedine točke.

Funkcija  $u(t)$  predstavlja kretanje po krivulji od neke početne točke i može se izraziti diskretnom Fourierovom transformacijom kao

$$u(t) = \sum_{n=0}^{T-1} a_n \cdot e^{j2\pi \frac{nt}{T}}$$

gdje oznaka  $a_n$  predstavlja koeficijente Fourierove transformacije:

$$a_n = \frac{1}{T} \sum_{t=0}^{T-1} d(t) \cdot e^{-j2\pi \frac{nt}{T}}$$

Ovaj postupak ovisan je o obliku konture, izboru početne točke, translaciji, rotaciji i povećanju konture, što njegovu uporabu čini kompliciranom. Stoga je Granlund u (2) definirao funkcije koje eliminiraju te utjecaje:

$$d_{mn} = \frac{a_{m+1}^n \cdot a_{T+1-n}^m}{a_1^{m+n}}$$

Ovako dobiveni Granlundovi koeficijenti predstavljaju vektor značajki koji se koristi za klasifikaciju siluete.

U svojoj implementaciji rješenja koristimo Granlundove koeficijente  $d_{12}$ ,  $d_{13}$ ,  $d_{14}$ ,  $d_{22}$ ,  $d_{23}$ ,  $d_{24}$ ,  $d_{32}$ ,  $d_{33}$  i  $d_{34}$ .

### 2.2.2 Huovi koeficijenti

Drugi način za odrediti vektor značajki koji će se koristiti za klasifikaciju osobe jest pomoću Huovih koeficijenata, odnosno momenata invarijanti.

Momenti invarijanti se često upotrebljavaju kod raspoznavanja uzoraka na slikama zbog svojstva otpornosti značajki na rotaciju, translaciju i skaliranje. Momenti i povezane invarijante su područje istraživanja karakterizacije uzoraka na slikama u širokom spektru primjena. Dobro poznati momenti uključuju geometrijske momente, zernike momente, rotacijske momente i kompleksne momente.

Momente invarijanti je prvi uveo Hu, koji je izveo šest apsolutnih ortogonalnih invarijanata i jednu iskrivljenu ortogonalnu invarijantu koja se temelji na algebarskim

invarijantama, koje ne samo da ne ovise o položaju, veličini i orijentaciji već su i nezavisni o paralelnim projekcijama. Momenti invarijanti su se pokazali kao adekvatna mjera za praćenje uzoraka na slikama uzevši u obzir translaciju, skaliranje i rotaciju istih na slikama (pod pretpostavkom da su slike kontinuirane funkcije i da nema utjecaja šuma). Najčešća primjena momenata invarijanti su raspoznavanje uzoraka na slikama, registriranje slika i rekonstrukcija slika. Slike u praktičnim primjenama često nisu kontinuirane već diskretne funkcije s prisutnošću šuma. Ustanovljeno je da su momenti višeg reda podložniji pogrešci uzrokovanoj šumom. Numeričke pogreške kod računanja isto nisu rijetkost. Promjena dimenzija slika uzrokuje interpolaciju ili brisanje pojedinih piksela. Rotacija slike također narušava promjenu funkcije koja opisuje sliku jer se može dogoditi zaokruživanje vrijednosti koordinata ili piksela.

Dvodimenzijski moment reda  $(p + q)$  definiran je na sljedeći način:

$$m_{pq} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x^p y^q f(x, y) dx dy, \text{ uz } p, q = 0, 1, 2, \dots$$

Ako je funkcija  $f(x, y)$ , koja opisuje sliku, kontinuirana ograničena funkcija uređenih parova, postoje momenti svih redova i niz momenata  $m_{pq}$  jedinstveno definiran sa  $f(x, y)$ , vrijedi i obrat. Svojstvo invarijantnosti se postiže korištenjem središnjih momenata, koji su definirani na sljedeći način:

$$\mu_{pq} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - x')^p (y - y')^q f(x, y) dx dy, \text{ uz } p, q = 0, 1, 2, \dots$$

pri čemu vrijedi  $x' = \frac{m_{10}}{m_{00}}$  i  $y' = \frac{m_{01}}{m_{00}}$ . Piksela na točki  $(x', y')$  centroid je slike opisane funkcijom  $f(x, y)$ .

Središnji momenti  $\mu_{pq}$ , koji su izračunati koristeći centroid slike  $f(x, y)$ , ekvivalentni su momentima  $m_{pq}$ , čiji je centar pomaknut u centroid slike. Upravo su zbog toga središnji momenti invarijante slike na translaciju.

Invarijante na skaliranje možemo dobiti normalizacijom. Normalizirani oblik središnjih momenata opisan je izrazom  $\eta = \frac{\mu_{pq}}{\mu_{00}^{\wp}}$ , uz  $\wp = \frac{p+q+2}{2}$  i  $p + q = 2, 3, \dots$

Usredotočivši se na normalizirane središnje momente, Hu je uveo sedam momenata invarijanti:

$$\begin{aligned} \phi_1 &= \eta_{20} + \eta_{02} \\ \phi_2 &= (\eta_{20} - \eta_{02})^2 + 4 \cdot \eta_{11}^2 \\ \phi_3 &= (\eta_{30} - 3 \cdot \eta_{12})^2 + (3 \cdot \eta_{21} - \mu_{03})^2 \\ \phi_4 &= (\eta_{30} + \eta_{12})^2 + (\eta_{21} + \mu_{03})^2 \end{aligned}$$

$$\begin{aligned}
\phi_5 &= (\eta_{30} - 3 \cdot \eta_{12}) \cdot (\eta_{30} + \eta_{12}) \cdot [(\eta_{30} + \eta_{12})^2 - 3 \cdot (\eta_{21} + \eta_{03})^2] + (3 \cdot \eta_{21} - \eta_{03}) \\
&\quad \cdot (\eta_{21} + \eta_{03}) \cdot [3 \cdot (\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \\
\phi_6 &= (\eta_{20} - \eta_{02}) \cdot [(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] + 4 \cdot \eta_{11} \cdot (\eta_{30} + \eta_{12}) \cdot (\eta_{21} + \eta_{03}) \\
\phi_7 &= (3 \cdot \eta_{21} - \eta_{03}) \cdot (\eta_{30} + \eta_{12}) \cdot [(\eta_{30} + \eta_{12})^2 - 3 \cdot (\eta_{21} + \eta_{03})^2] - (\eta_{30} - 3 \cdot \eta_{12}) \\
&\quad \cdot (\eta_{21} + \eta_{03}) \cdot [3 \cdot (\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2]
\end{aligned}$$

Tih je sedam momenata invarijanti veoma korisnih zbog svojstva da ostaju nepromijenjeni usred rotacije, translacije ili skaliranja.

## 2.3 Određivanje identiteta osobe

Određivanje identiteta osobe vrši se klasificiranjem dobivenih značajki. Kao ulaz se dobije vektor značajki koji predstavlja siluetu osobe, a pretpostavlja se da postoji baza u kojoj su zapisane tipične vrijednosti vektora značajki za određeni skup osoba. Dobiveni vektor se uspoređi s bazom te se na temelju te usporedbe klasificira kao određena osoba iz baze ili kao osoba koja ne postoji u bazi (u slučaju da sličnost niti s jednom osobom u bazi nije dovoljno velika da bi se moglo pretpostaviti da se radi o toj osobi). Kao izlaz se preda identifikacijska oznaka osobe.

### 2.3.1 Bayesov klasifikator

Bayesov klasifikator najpoznatiji je klasifikacijski model. Njegovu osnovu čini Bayesovo pravilo, pomoću kojega za svaku klasu možemo procijeniti vjerojatnost da neki primjer pripada toj klasi.

Bayesovo pravilo glasi:

$$P(C_j|x) = \frac{p(x|C_j) \cdot p(C_j)}{\sum_{k=1}^K p(x|C_k) \cdot p(C_k)}$$

Dana formula predstavlja aposteriornu vjerojatnost da primjer  $x$  pripada klasi  $C_j$ .  $p(C_j)$  predstavlja apriornu vjerojatnost klase,  $p(x|C_j)$  uvjetnu gustoću, a  $K$  broj klasa. Stoga, određivanje kojoj od  $K$  klasa pripada primjer  $x$  svodi se na  $\operatorname{argmax}_{C_k} (p(x|C_k) \cdot p(C_k))$ .

Bayesov klasifikator je parametarski model, dakle pretpostavlja da su podaci raspodijeljeni po nekoj teorijskoj vjerojatnosnoj razdiobi. Primijenjena implementacija pretpostavlja raspodjelu po normalnoj (Gaussovoj) razdiobi.

Prilikom učenja, algoritam procjenjuje srednje vrijednosti vektora i kovarijacijsku matricu za svaku klasu, te koristi te podatke za klasificiranje ulaznih podataka.



### 2.3.2 KNN klasifikator

Često opisivan kao klasifikator koji „uči na primjerima“, KNN (engl. *K Nearest Neighbours*) pamti primjere za učenje te na temelju njih vrši klasificiranje novih primjera. Kao dodatni parametar klasifikatoru se preda  $k$ , broj susjeda koji će se uzimati u obzir pri klasifikaciji novog primjera.

Klasifikacija se vrši tako da se pronade  $k$  zapamćenih vektora koji su najbliži novom vektoru, a potom se vrši „glasanje“. Proglasit će se da novi ulazni vektor pripada klasi  $C_j$  ako je među  $k$  najbližijih vektora najviše onih koji pripadaju klasi  $C_j$ .

### 2.3.3 *Random Forest* klasifikator

*Random Forest* ili *Random Trees* je klasifikator koji prilikom učenja stvara više stabala odluke (opcionalni parametar pri stvaranju klasifikatora ili se koristi pretpostavljena vrijednost od maksimalno 10 stabala) te kasnije primjere klasificira na temelju rezultata klasifikacije svakog pojedinog stabla.

Ako imamo  $N$  primjera za učenje i  $M$  komponenata u vektoru značajki, stvaranje svakog pojedinog stabla vrši se na sljedeći način:

1. zada se  $m$ , broj komponenata koje će se koristiti pri donošenju odluke u pojedinom čvoru stabla ( $m$  bi trebao biti puno manji od  $M$ )
2. dio primjera za učenje se koristi kako bi se naučilo stablo, a ostali za procjenu pogreške klasifikacije
3. u svakom čvoru stabla se na temelju  $m$  nasumičnih komponenti odredi najbolja podjela.

Kad na provjeru dođe novi primjer, on se propusti kroz stablo i klasificira se prema klasi završnog čvora u koji dospije. Postupak se provede nad svim stablima te se konačna klasa primjera odredi „glasanjem“ svih stabala.

### **3. Ispitivanje rješenja**

#### **3.1 Ispitna baza**

Pri ispitivanju rješenja koristili smo dvije baze fotografija . S obzirom da postoji baza od prošlogodišnjeg sličnog projekta, od njenih autora zatražili smo i dobili dozvolu da je koristimo, a osim toga, napravili smo i vlastitu bazu. Obje baze sastoje se od fotografija 11 osoba u zatvorenom prostoru.

##### **3.1.1 Prošlogodišnja baza**

Baza sadrži 248 slika od ukupno 24 osobe, no za potrebe testiranja, iskoristili smo njen podskup (kao što je već rečeno, iskoristili smo fotografije 11 osoba). Slike su slikane digitalnim fotoaparatom razlučivosti  $3648 \times 2736$  piksela postavljenim na visini od 156 cm i udaljenim 319 cm od osobe. Slike su naknadno reducirane na 30% originalne veličine i takve spremljene u bazu, što rezultira slikama razlučivosti od  $821 \times 1094$  piksela. Takva razlučivost pokazala se dovoljno dobrom za izlučivanje kvalitetne siluete, također i značajno smanjuje ukupnu veličinu baze. Baza sadrži fotografije koje su uglavnom slične jedne drugima (ista osoba u pozama koje se minimalno razlikuju).

##### **3.1.2 Vlastita baza**

Ova baza dobivena je pomoću digitalnog fotoaparata Nikon D3100, a fotografije su originalne rezolucije  $3072 \times 4608$  piksela. S obzirom na to da je to vrlo visoka rezolucija, odlučili smo smanjiti istu, primarno iz razloga što su nam tako velike fotografije nepotrebno usporavale izvođenje programa, a smanjenjem rezolucije ne gubimo informaciju koju trebamo. Zbog toga smo slike smanjili na 25% originalne veličine i radili sa slikama rezolucije  $768 \times 1152$  piksela.

Za razliku od prošlogodišnje baze, odlučili smo da bi bilo dobro iskušati i kako bitno različite poze utječu na klasifikaciju. Stoga, naša je baza takva da su fotografije iste osobe većinom različite u smislu da na jednoj fotografiji osoba stoji potpuno uspravno s rukama uz tijelo, a već na idućoj ima potpuno raširene ruke i noge. Tako možemo usporediti koliki je pad postotka točnosti prepoznavanja, što će biti pokazano kasnije, u poglavlju s rezultatima.

Također, naše su fotografije snimane na nekoliko lokacija s različitim pozadinama, iz dva razloga. Prvi je teškoća pronalaženja optimalnog mjesta na kojem bi sjene bile

minimalne, a rasvjeta dovoljno dobra. Drugi je kako bismo vidjeli hoće li različitost pozadina imati kakvog utjecaja na finalne rezultate.

## 3.2 Rezultati učenja i ispitivanja

Uvodna napomena: za metode koje koriste rastav ljestvicom sivih tonova za oduzimanje pozadine granična vrijednost 120 bila je prevelika da bi se mogla naći kontura na fotografiji. Kao prag koristili smo medijan. U tablicama koje slijede prikazani su postoci točnosti prepoznavanja za pojedine parametre i klasifikatore.

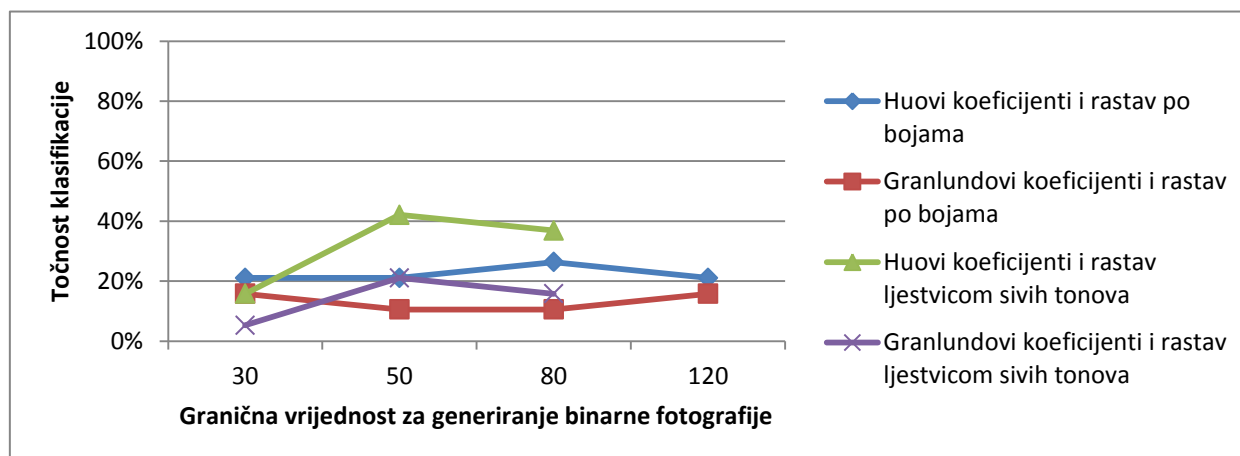
### 3.2.1 Vlastita baza

U ovom potpoglavlju prikazani su rezultati ispitivanja baze koju smo napravili ove godine.

#### 3.2.1.1 Pregled po klasifikatorima

|                                |                          | Granične vrijednosti za generiranje binarne fotografije |         |         |               |
|--------------------------------|--------------------------|---|---------|---------|---------------|
|                                |                          | 30  | 50      | 80      | 120           |
| Rastav po bojama               | Huovi koeficijenti       | 0,21053   | 0,21053 | 0,26316 | 0,21053       |
|                                | Granlundovi koeficijenti | 0,15789   | 0,10526 | 0,10526 | 0,15789       |
| Rastav ljestvicom sivih tonova | Huovi koeficijenti       | 0,15789   | 0,42105 | 0,36842 | Prevelik prag |
|                                | Granlundovi koeficijenti | 0,05263   | 0,21053 | 0,15789 | Prevelik prag |

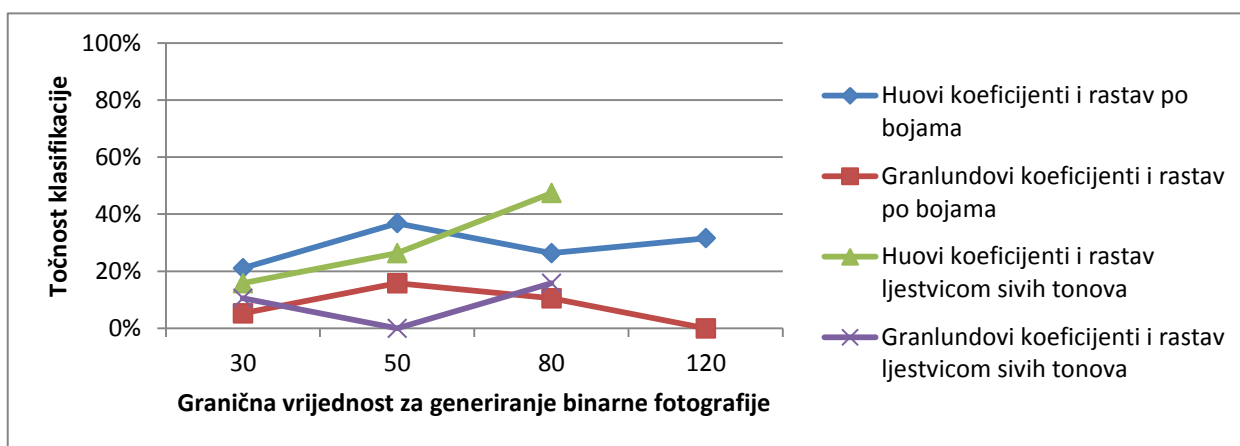
Tablica 3.1 Točnost Bayesovog klasifikatora



Graf 3.1 Točnost Bayesovog klasifikatora

|                                |                          | Granične vrijednosti za generiranje binarne fotografije |         |         |               |
|--------------------------------|--------------------------|---|---------|---------|---------------|
|                                |                          | 30  | 50      | 80      | 120           |
| Rastav po bojama               | Huovi koeficijenti       | 0,21053   | 0,36842 | 0,26316 | 0,31579       |
|                                | Granlundovi koeficijenti | 0,05263   | 0,15789 | 0,10526 | 0             |
| Rastav ljestvicom sivih tonova | Huovi koeficijenti       | 0,15789   | 0,26316 | 0,47368 | Prevelik prag |
|                                | Granlundovi koeficijenti | 0,10526   | 0       | 0,15789 | Prevelik prag |

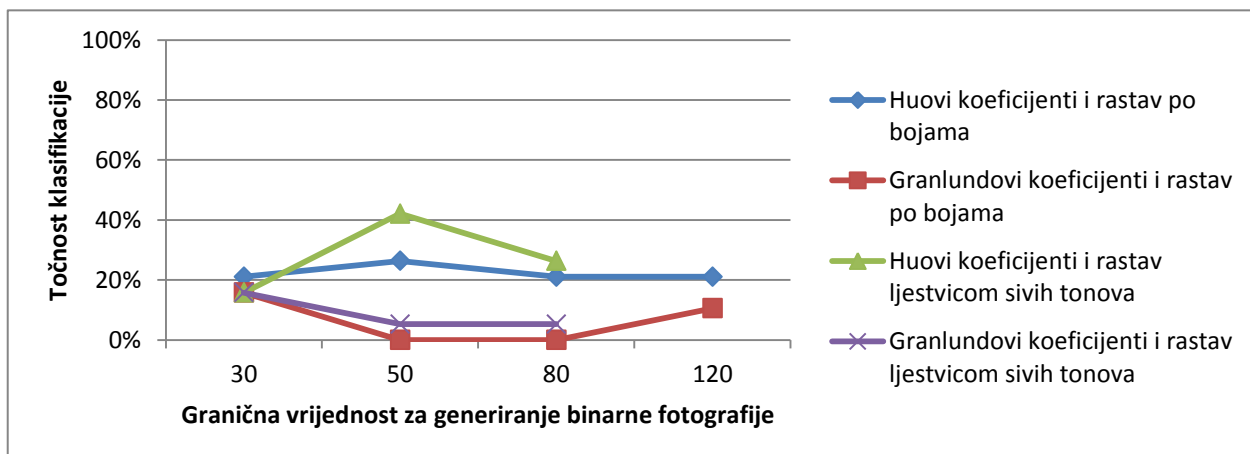
Tablica 3.2 Točnost klasifikatora KNN



Graf 3.2 Točnost klasifikatora KNN

|                                |                          | Granične vrijednosti za generiranje binarne fotografije |         |         |               |
|--------------------------------|--------------------------|---|---------|---------|---------------|
|                                |                          | 30  | 50      | 80      | 120           |
| Rastav po bojama               | Huovi koeficijenti       | 0,21053   | 0,26316 | 0,21053 | 0,21053       |
|                                | Granlundovi koeficijenti | 0,15789   | 0       | 0       | 0,10526       |
| Rastav ljestvicom sivih tonova | Huovi koeficijenti       | 0,15789   | 0,42105 | 0,26316 | Prevelik prag |
|                                | Granlundovi koeficijenti | 0,15789   | 0,05263 | 0,05263 | Prevelik prag |

Tablica 3.3 Točnost klasifikatora *Random Forest*



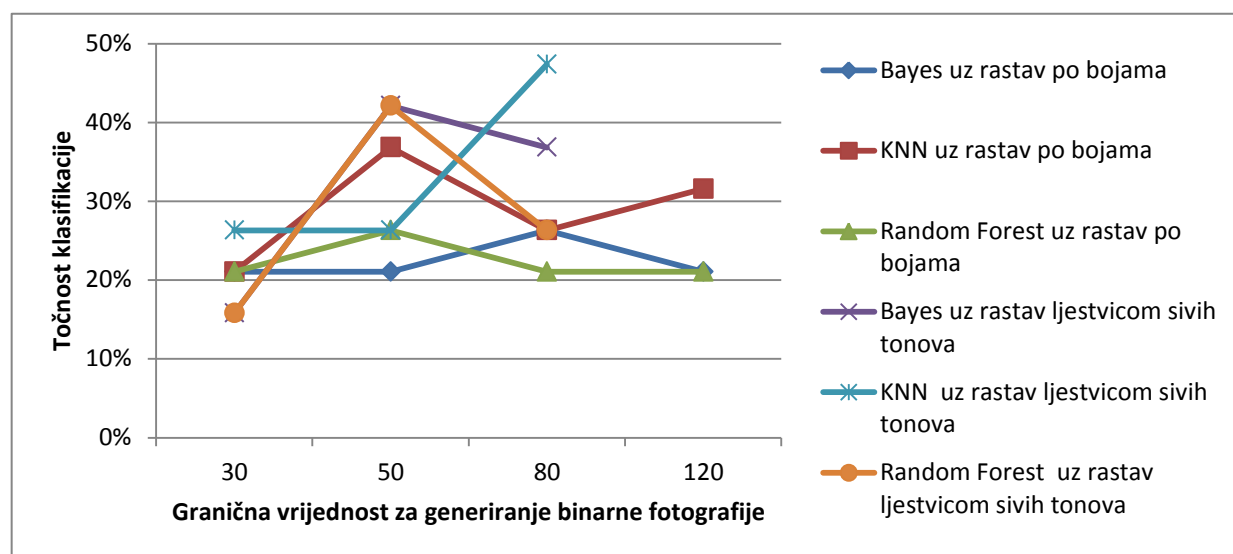
Graf 3.3 Točnost klasifikatora *Random Forest*

### 3.2.1.2 Usporedba rezultata za Huove koeficijente

Budući da su se Huovi koeficijenti pokazali kao bolji odabir za izlučivanje značajki iz siluete, napravljena je usporedba s obzirom na pristup generiranju fotografije osobe.

|                                |                      | Granična vrijednost za generiranje binarne fotografije |         |         |               |
|--------------------------------|----------------------|--|---------|---------|---------------|
|                                |                      | 30   | 50      | 80      | 120           |
| Rastav po bojama               | Bayes                | 0,21053  | 0,21053 | 0,26316 | 0,21053       |
|                                | KNN                  | 0,21053  | 0,36842 | 0,26316 | 0,31579       |
|                                | <i>Random Forest</i> | 0,21053  | 0,26316 | 0,21053 | 0,21053       |
| Rastav ljestvicom sivih tonova | Bayes                | 0,15789  | 0,42105 | 0,36842 | Prevelik prag |
|                                | KNN                  | 0,26316  | 0,26316 | 0,47368 | Prevelik prag |
|                                | <i>Random Forest</i> | 0,15789  | 0,42105 | 0,26316 | Prevelik prag |

Tablica 3.4 Usporedba točnosti klasifikatora za Huove koeficijente



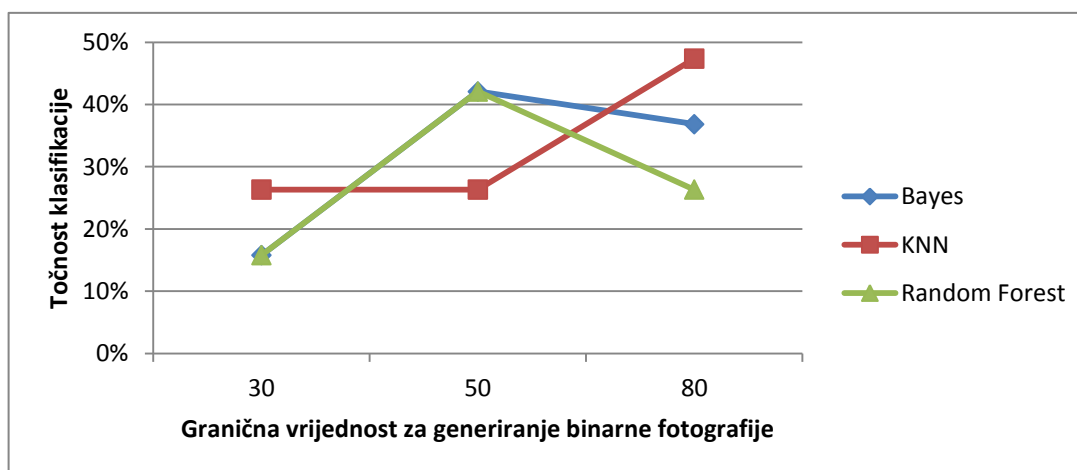
Graf 3.4 Usporedba točnosti klasifikatora za Huove koeficijente

### 3.2.1.3 Usporedba klasifikatora za Huove koeficijente i oduzimanje pozadine korištenjem rastava po ljestvici sivih tonova

Kako se ova kombinacija parametara pokazala najuspješnijom, napravljena je usporedba točnosti klasifikatora za njih.

|                      | Granična vrijednost generiranje binarne fotografije |         |         |
|----------------------|---|---------|---------|
|                      | 30  | 50      | 80      |
| Bayes                | 0,15789   | 0,42105 | 0,36842 |
| KNN                  | 0,26316   | 0,26316 | 0,47368 |
| <i>Random Forest</i> | 0,15789   | 0,42105 | 0,26316 |

Tablica 3.5 Usporedba klasifikatora uz Huove koeficijente i rastav ljestvicom sivih tonova



Graf 3.5 Usporedba klasifikatora za Huove koeficijente i rastav ljestvicom sivih tonova

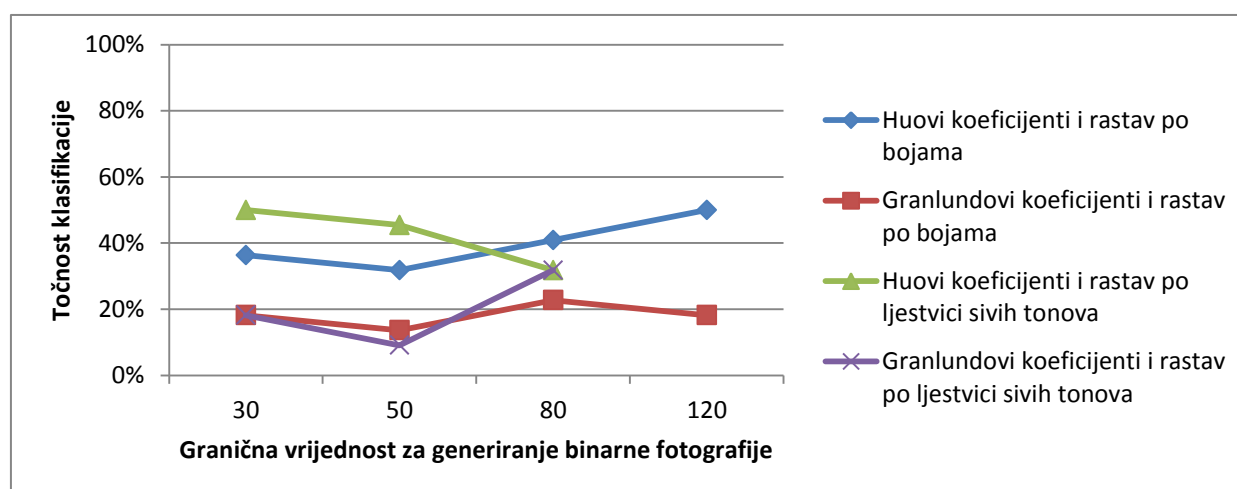
### 3.2.2 Prošlogodišnja baza

U ovom potpoglavlju prikazani su rezultati ispitivanja baze koju smo posudili od kolega koji su sličan projekt napravili prošle godine.

#### 3.2.2.1 Pregled po klasifikatorima

|                                |                          | Granične vrijednosti za generiranje binarne fotografije |         |         |               |
|--------------------------------|--------------------------|---|---------|---------|---------------|
|                                |                          | 30  | 50      | 80      | 120           |
| Rastav po bojama               | Huovi koeficijenti       | 0,36364   | 0,31818 | 0,40909 | 0,5           |
|                                | Granlundovi koeficijenti | 0,18182   | 0,13636 | 0,22727 | 0,18182       |
| Rastav ljestvicom sivih tonova | Huovi koeficijenti       | 0,5   | 0,45455 | 0,31818 | Prevelik prag |
|                                | Granlundovi koeficijenti | 0,18182   | 0,09091 | 0,31818 | Prevelik prag |

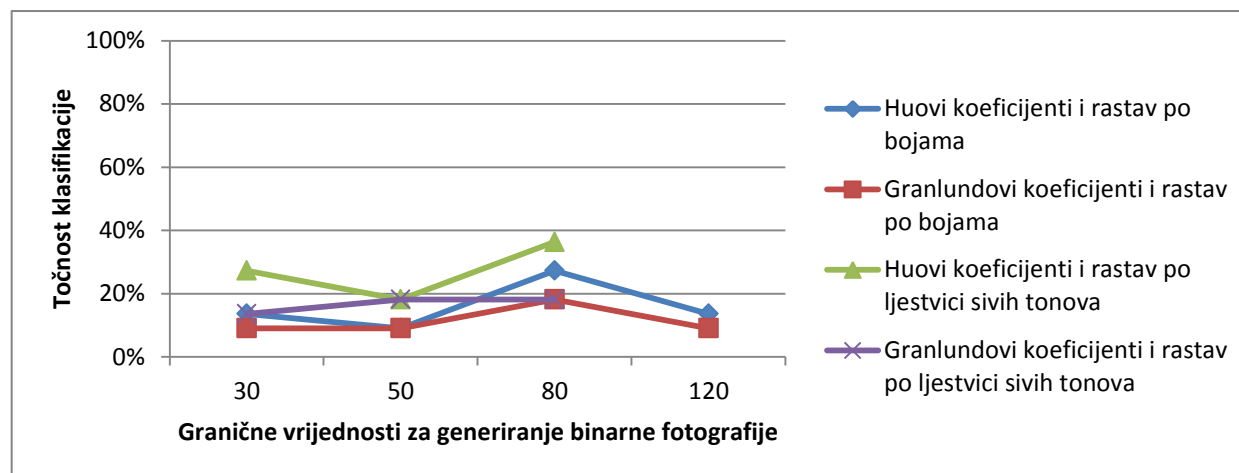
Tablica 3.6 Točnost Bayesovog klasifikatora



Graf 3.6 Točnost Bayesovog klasifikatora

|                                |                          | Granične vrijednosti za generiranje binarne fotografije |         |         |               |
|--------------------------------|--------------------------|---|---------|---------|---------------|
|                                |                          | 30  | 50      | 80      | 120           |
| Rastav po bojama               | Huovi koeficijenti       | 0,13636   | 0,09091 | 0,27273 | 0,13636       |
|                                | Granlundovi koeficijenti | 0,09091   | 0,09091 | 0,18182 | 0,09091       |
| Rastav ljestvicom sivih tonova | Huovi koeficijenti       | 0,27273   | 0,18182 | 0,36364 | Prevelik prag |
|                                | Granlundovi koeficijenti | 0,13636   | 0,18182 | 0,18182 | Prevelik prag |

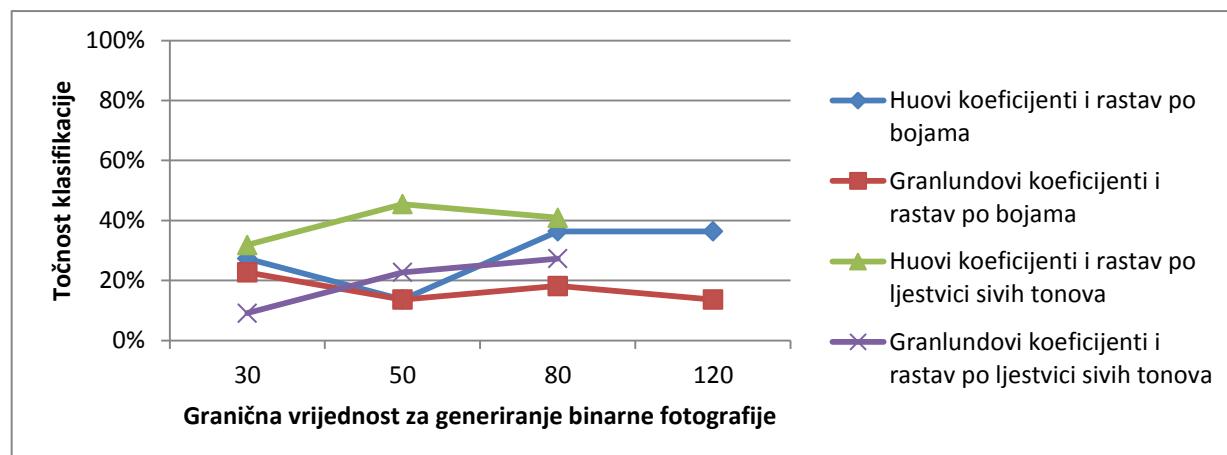
Tablica 3.7 Točnost klasifikatora KNN



Graf 3.7 Točnost klasifikatora KNN

|                                |                          | Granične vrijednosti za generiranje binarne fotografije |         |         |               |
|--------------------------------|--------------------------|---|---------|---------|---------------|
|                                |                          | 30  | 50      | 80      | 120           |
| Rastav po bojama               | Huovi koeficijenti       | 0,27273   | 0,13636 | 0,36364 | 0,36364       |
|                                | Granlundovi koeficijenti | 0,22727   | 0,13636 | 0,18182 | 0,13636       |
| Rastav ljestvicom sivih tonova | Huovi koeficijenti       | 0,31818   | 0,45455 | 0,40909 | Prevelik prag |
|                                | Granlundovi koeficijenti | 0,09091   | 0,22727 | 0,27273 | Prevelik prag |

Tablica 3.8 Točnost klasifikatora *Random Forest*



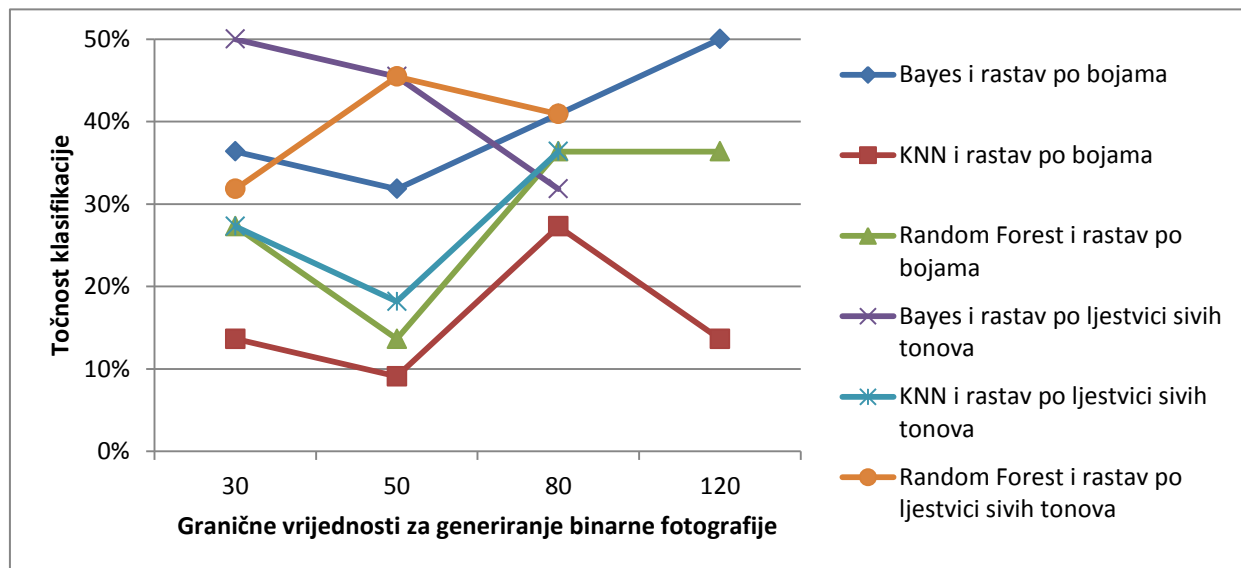
Graf 3.8 Točnost klasifikatora *Random Forest*

### 3.2.2.2 Usporedba rezultata za Huove koeficijente

Kao i kod prošle baze, pokazalo se da Huovi koeficijenti daju bolje rezultate nego Granlundovi, pa je u nastavku dana usporedba točnosti klasifikatora za Huove koeficijente s obzirom na pristup oduzimanju pozadine za generiranje siluete.

|                                |                      | Granična vrijednost za generiranje binarne fotografije |         |         |               |
|--------------------------------|----------------------|--|---------|---------|---------------|
|                                |                      | 30   | 50      | 80      | 120           |
| Rastav po bojama               | Bayes                | 0,36364  | 0,31818 | 0,40909 | 0,5           |
|                                | KNN                  | 0,13636  | 0,09091 | 0,27273 | 0,13636       |
|                                | <i>Random Forest</i> | 0,27273  | 0,13636 | 0,36364 | 0,36364       |
| Rastav ljestvicom sivih tonova | Bayes                | 0,5  | 0,45455 | 0,31818 | Prevelik prag |
|                                | KNN                  | 0,27273  | 0,18182 | 0,36364 | Prevelik prag |
|                                | <i>Random Forest</i> | 0,31818  | 0,45455 | 0,40909 | Prevelik prag |

Tablica 3.9 Usporedba točnosti klasifikatora za Huove koeficijente



Graf 3.9 Usporedba točnosti klasifikatora za Huove koeficijente

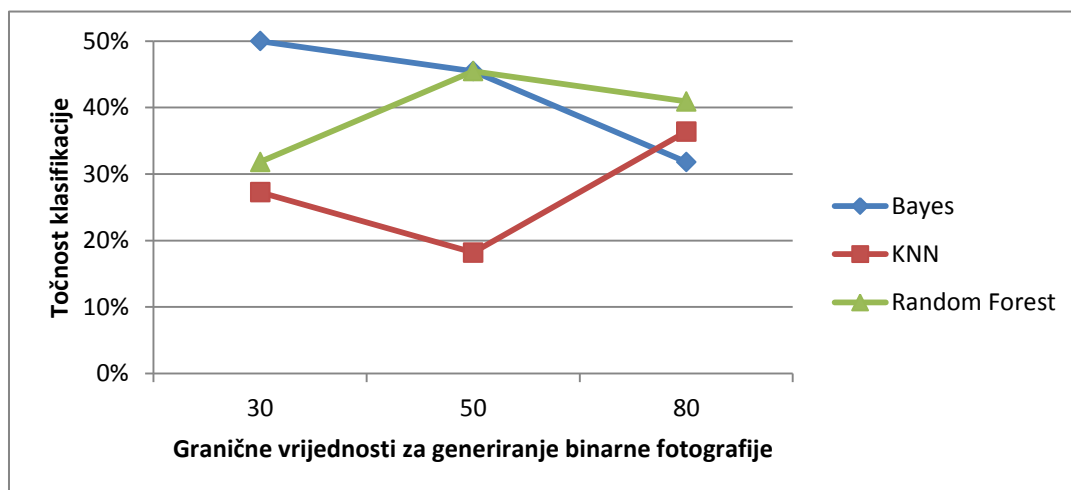
### 3.2.2.3 Usporedba klasifikatora za Huove koeficijente i oduzimanje pozadine korištenjem rastava po ljestvici sivih tonova

Ispitivanja su i na ovoj bazi pokazala kako je kombinacija parametara u podnaslovu ona koja daje najtočnije rezultate, pa navodimo usporedbu klasifikatora s obzirom na njih.

|                      | Granična vrijednost za generiranje binarne fotografije |         |         |
|----------------------|--|---------|---------|
|                      | 30   | 50      | 80      |
| Bayes                | 0,5  | 0,45455 | 0,31818 |
| KNN                  | 0,27273  | 0,18182 | 0,36364 |
| <i>Random Forest</i> | 0,31818  | 0,45455 | 0,40909 |

Tablica 3.10 Usporedba klasifikatora uz Huove koeficijente i rastav ljestvicom sivih tonova





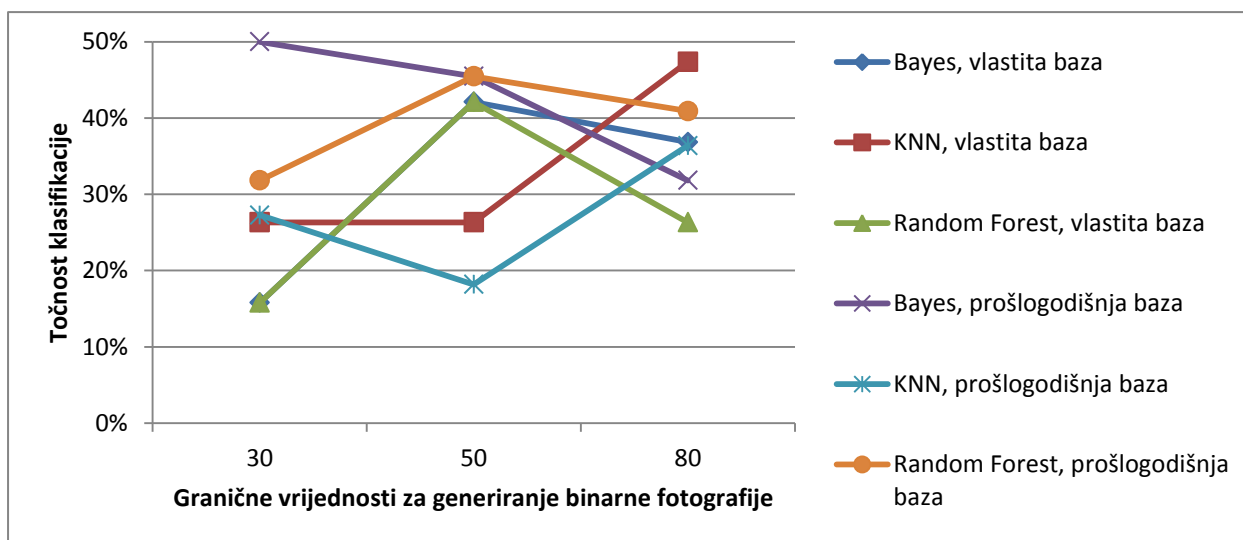
Graf 3.10 Usporedba klasifikatora za Huove koeficijente i rastav ljestvicom sivih tonova

### 3.2.3 Usporedba rezultata s obzirom na korištene baze

Budući da su se za obje baze kao najuspješnija kombinacija parametara pokazali Huovi koeficijenti i rastav ljestvicom sivih tonova, u nastavku je napravljena usporedba točnosti klasifikatora s obzirom na korištenu bazu.

|                     |               | Granična vrijednost za generiranje binarne fotografije |         |         |
|---------------------|---------------|--|---------|---------|
|                     |               | 30   | 50      | 80      |
| Vlastita baza       | Bayes         | 0,15789  | 0,42105 | 0,36842 |
|                     | KNN           | 0,26316  | 0,26316 | 0,47368 |
|                     | Random Forest | 0,15789  | 0,42105 | 0,26316 |
| Prošlogodišnja baza | Bayes         | 0,5  | 0,45455 | 0,31818 |
|                     | KNN           | 0,27273  | 0,18182 | 0,36364 |
|                     | Random Forest | 0,31818  | 0,45455 | 0,40909 |

Tablica 3.11 Usporedba klasifikatora s obzirom na bazu



Graf 3.11 Usporedba klasifikatora s obzirom na bazu

### 3.3 Analiza rezultata

Iz rezultata se vidi kako problem nije trivijalan. Općenito u opisanom rješenju točnost prepoznavanja nikad nije veća od 50%. Smatramo da ti rezultati ukazuju na složenost ovog problema i na potrebu pristupa problemu iz mnogo različitih stajališta. U rješavanju problema postoji mnogo parametara koji se moraju ručno podešavati i kombinirati, a o njima najviše ovisi točnost prepoznavanja.

Prag za generiranje binarne fotografije značajan je parametar, što se može vidjeti na sljedećem primjeru:



Slika 3.1 Originalna fotografija



Slika 3.2 Silueta uz prag 20



Slika 3.3 Silueta uz prag 40



Slika 3.4 Silueta uz prag 100

Vidimo da, ovisno o tom parametru, dobijemo vrlo različite konture, što značajno utječe na generiranje vektora značajki za klasifikator.

Iz prikazanih rezultata također vidimo kako su se Huovi koeficijenti pokazali boljim odabirom pri izlučivanju vektora značajki iz siluete od Granlundovih. To je suprotno našim očekivanjima i pretpostavljamo da je problem u tome što nismo uspjeli pronaći adekvatnu kombinaciju značajki za Granlundove koeficijente.

Nadalje, pri oduzimanju pozadine kao bolji izbor pokazala se metoda rastava ljestvicom sivih tonova. Smatramo da je tome tako jer se tim postupkom bolje neutraliziraju sjene, koje su same po sebi sive, a rastavom na RGB komponente dobivaju više na značaju no što bi trebale.

Konačno, kao što smo i pretpostavljali, točnost prepoznavanja je veća za prošlogodišnju bazu, gdje su osobe na fotografijama u relativno sličnim pozama. Smatramo da je različitost poza u kojima su fotografirane osobe, kao i različite pozadine te raznovrsni artefakti u pozadini, u bazi koju smo stvorili pri izradi projekta pridonijela smanjenoj točnosti klasifikacije u odnosu na korištenje prošlogodišnje baze.

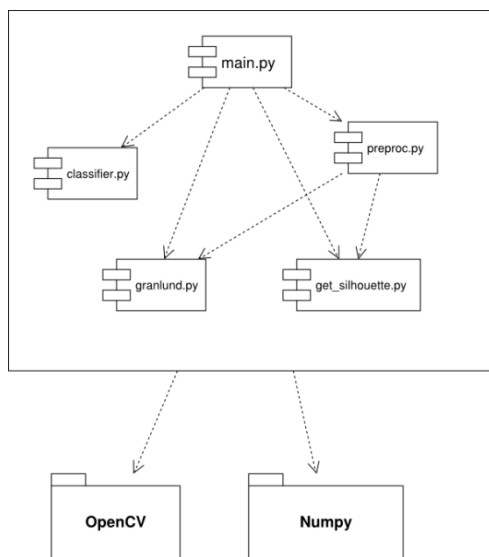
## 4. Opis programske implementacije rješenja

### 4.1 Struktura programske implementacije

Projekt je implementiran u programskom jeziku Python (verzija 2.7) uz korištenje biblioteke *OpenCV* (Python modul *cv2*, verzija 2.4.7) te pomoćnog modula *numpy*.

Implementacija je podijeljena između 5 modula:

1. *main.py* - glavni modul iz kojeg se pokreće cijeli program i ima dvije važne funkcije:
  - *train* - obavlja treniranje klasifikatora nad dobivenim skupom za učenje
  - *predict* - kasnije čeka na ulazne fotografije i pomoću naučenih klasifikatora vrši njihovu klasifikaciju i računa statistiku točnosti
2. *preproc.py* - modul koji se poziva iz *main.py*, a služi tome da pripremi podatke za učenje za klasifikatore. Naime, on procesira sve ulazne fotografije i njihove pozadine te zove module *get\_silhouette* i *granlund* preko kojih dohvaća siluetu i značajke te vektore značajki sprema u listu koja se kasnije šalje klasifikatorima na učenje
3. *get\_silhouette.py* - modul koji dohvaća i obrađuje fotografije te izlučuje siluete koje vraća modulu *preproc* iz kojeg se i poziva
4. *granlund.py* - modul koji služi tome da se iz silueta generiraju vektori značajki, Hu momenti i Granlundovi koeficijenti; poziva se iz modula *preproc* te kao parametar prima siluetu, a vraća vektor značajki za tu siluetu
5. *classifier.py* - modul koji služi kao omotač za razne klasifikatore iz *OpenCV*-a; konkretno, modul pruža sučelja prema Bayesovom klasifikatoru, KNN-u i *Random Forest* klasifikatoru.



Slika 4.1 Strukturni dijagram programa

## 4.2 Parametri na ulazu programske implementacije

Parametri s kojima se pokreće program opisani su u nastavku. Program ima tri obavezna parametra bez kojih ga neće biti moguće pokrenuti, te X neobaveznih koji služe za podešavanje i odabir načina izvođenja programa.

Obavezni parametri:

- *path* - putanja do direktorija sa skupom za učenje. Pretpostavlja se da je struktura tog direktorija sljedeća: za svaku osobu postoji poddirektorij u kojem se nalaze dva direktorija, *pics/* i *back/*. U direktoriju *pics/* nalaze se fotografije osobe. Svaka fotografija osobe ima u direktoriju *back/* odgovarajuću istoimenu fotografiju koja odgovara pozadini fotografije na kojoj se nalazi osoba. Program će kao ime osobe, pa prema tome i razred u koji se ona klasificira, uzeti ime direktorija s fotografijama te osobe, dakle direktorija u kojem se nalaze *pics/* i *path/*
- *method* - metoda koja se upotrebljava za izlučivanje značajki sa silueta; može biti *hu* ili *granlund*
- *threshold* - parametar koji regulira prag za određivanje crne i bijele boje pri generiranju binarne fotografije. Naime, kod izlučivanja siluete, ideja je da sve točke koje imaju intenzitet manji od mediana ili srednje vrijednosti intenziteta zbrojenog s *thresholdom* budu crne, a ostale bijele. Ovaj parametar nije obavezan samo u jednom slučaju koji je opisan u nastavku.

Parametri navedeni u nastavku su dodatni parametri koji utječu na načina izvođenja programa:

- *optparam* - ovaj parametar smo dodali zbog jedne ideje. Naime, naš program podržava varijabilno zadavanje praga kroz datoteku. Što to znači? To znači da korisnik može sam pokušati procijeniti kolika je optimalna vrijednost parametra *threshold* za svaku fotografiju koju ima, zapisati to u datoteku *parameters.txt* i staviti je u direktoriji svake osobe. Tad će program prilikom učenja abecednim redom učitavati fotografije i pripadajuće *threshold*-e za svaku. Pretpostavka je da će za svaku fotografiju *threshold* biti zapisan u novom redu datoteke *parameters.txt* i da će njihov redoslijed pratiti abecedni redoslijed fotografija. Primjer takve datoteke dan je u bazi na DVD-u. Ako je ovaj parametar postavljen na 0, to je jednako kao da i nije postavljan, a ako je postavljen na 1, program će tražiti *threshold* za svaku fotografiju i ignorirati ulazni parametar *threshold* ako je postavljen (on u tom slučaju može biti izostavljen).

- *approach* - parametar regulira koristimo li rastav po bojama ili po ljestvici sivih tonova za izlučivanje siluete. Dozvoljene vrijednosti su *c* (engl. *colour*) i *g* (engl. *grayscale*). Postavljen je na *c* ako se drukčije ne postavi.
  - *threshtype* - za parametar *threshold* regulira koristi li se medijan ili srednja vrijednost intenziteta kao mjera. Postavljen je početno na medijan.
  - *nmax* - regulira maksimalan broj susjeda koji KNN može gledati prilikom klasifikacije.
  - *nclass* - regulira koliko susjeda se zaista gleda kod KNN-a prilikom klasifikacije. Mora biti manji ili jednak *nmax*.
  - *maxdepth* - maksimalna dubina stabla kod *Random Forest* klasifikatora.
  - *criteria* - kriterij zaustavljanja kod učenja *Random Forest* klasifikatora. Može biti *cv.TERM\_CRITERIA\_MAX\_ITER* ili *cv.TERM\_CRITERIA\_EPS* ili oboje.
  - *maxtrees* - maksimalan broj stabala u *Random Forest* šumi.
  - *maxerror* - maksimalna pogreška prije stajanja s učenjem *Random Forest* klasifikatora.
- Zadnja dva parametra uzimaju se u obzir ovisno o parametru *criteria*.

### 4.3 Korištenje programske implementacije

Program je predviđen za izvođenje na Linux ili Mac OS operacijskim sustavima. Program se pokreće iz terminala sa sljedećim obaveznim parametrima (za ostale, pokrenuti *main.py* s opcijom *-h*):

```
--path=<put_do_skupa_za_učenje>
--method=hu|granlund - odabir metode izlučivanja značajki: Hu momenti ili Granlundovi koeficijenti
--threshold=<broj> - prag koji prilikom izlučivanja siluete određuje koje točke postaju bijele (silueta) ili crne (pozadina).
```

Svi ostali parametri imaju pretpostavljene vrijednosti koje se mogu vidjeti pokretanjem *main.py* uz opciju *-h*.

Primjer osnovnog pokretanja programa:

```
python src/main.py --path=baza_2013 --method=hu --threshold=50
```

Nakon što program nauči podatke koje dobije, moguće ga je testirati. Program će nakon faze učenja obavijestiti korisnika da je učenje gotovo i čekati da mu se na standardni ulaz preda putanja do fotografije, ime osobe te putanja do fotografije pozadine. Na temelju ovoga program će dati klasifikaciju i procijeniti je li ona točna uspoređujući ime osobe koje je producirao svaki klasifikator s imenom koje mu se zadalo kao ispravno. Kad se želi završiti s programom, dovoljno je utipkati *-1*.

## 5. Zaključak

Cilj ovog projekta bio je implementirati prepoznavanje osobe na temelju njene siluete. Za početak smo našli i analizirali nekoliko srodnih radova kako bismo odlučili u kojem smjeru krenuti sa svojom implementacijom.

Prilikom implementacije, odlučili smo napraviti modularan program. To će reći, odlučili smo napraviti nekoliko međusobno nezavisnih implementacija pojedinih koraka rješavanja problema i testirati kako oni međusobno djeluju. Posljedica toga jest da smo implementirali dvije metode za oduzimanje pozadine iz fotografije (rastav fotografija na RGB komponente odnosno korištenjem ljestvice sivih tonova), dvije metode izvlačenja vektora značajki iz dobivene siluete (Granlundovi i Huovi koeficijenti), te tri klasifikatora (Bayesov, KNN i *Random Forest*).

Kako bismo napravili što raznovrsnije testiranje implementiranog rješenja, koristili smo bazu fotografija koju su napravili kolege pri izradi sličnog projekta prošle godine, ali smo i napravili svoju bazu koja u obzir uzima više čimbenika (različite poze osoba, artefakti u pozadini, različiti načini osvjetljenja).

Nakon što smo proveli testiranje i analizirali rezultate, uvidjeli smo da su oni uglavnom u skladu s našim očekivanjima. Problem prepoznavanja osobe samo na temelju njene siluete nije nimalo jednostavan, što se vidi po tome da ni u kojem slučaju nemamo točnost klasifikacije veću od 50%. Jedino što nije bilo u skladu s našim očekivanjima je veća točnost pri uporabi Huovih koeficijenata u odnosu na Granlundove. Slutimo da je problem u tome što nismo uspjeli pronaći odgovarajuće Granlundove koeficijente i/ili kombinacije s ostalim parametrima programa.

Budući da je prepoznavanje konkretne osobe samo na temelju njene siluete veoma neistraženo područje, točnost klasifikacije od 50% ocjenjujemo kao uspješnu, uz napomenu kako prostora za poboljšanje, naravno, ima.

Kao poboljšanje naše implementacije mogli bi se istražiti drugi koeficijenti i parametri za izlučivanje vektora značajki pri korištenju Granlundovih koeficijenata. Usto, mogle bi se testirati metode prepoznavanja poput prepoznavanja na temelju proporcija tijela, kao i neki drugi klasifikatori (primjerice *SVM*) ili metode njihovog učenja (*backpropagation* algoritam za učenje neuronske mreže, neuronske mreže učene genetskim algoritmima itd.).

## 6. Literatura

1. *Topological Structural Analysis of Digitized Binary Images by Border Following*. **Suzuki, S. i Abe, K.** 1985, CVGIP, str. 32-46. URL:  
<http://www.sciencedirect.com/science/article/pii/0734189X85900167>.
2. *Fourier Preprocessing for Hand Print Character Recognition*. **Granlund, G. H.** 1972. IEEE Transactions on Computers. str. 195-201. URL:  
<http://ieeexplore.ieee.org/xpl/articleDetails.jsp?tp=&arnumber=5008926>.
3. *Silhouette recognition using high-resolution pursuit*. **Jaggi, Seema, i dr., i dr.** 1999, Pattern Recognition, str. 753-771. URL:  
[http://ssg.mit.edu/~willsky/publ\\_pdfs/141\\_pub\\_PR.pdf](http://ssg.mit.edu/~willsky/publ_pdfs/141_pub_PR.pdf).
4. *Soft Biometrics; Human Identification using Comparative Descriptions*. **Reid, D., Nixon, M. i Stevenage, S.** 2013, IEEE Transactions on Pattern Analysis and Machine Intelligence, str. 1. URL: <http://ieeexplore.ieee.org/xpl/articleDetails.jsp?reload=true&arnumber=6654143>.
5. *Automatic Monocular System for Human Fall Detection Based on Variations in Silhouette Area*. **Mirmahboub, B., i dr., i dr.** 2, 2013, IEEE Transactions on Biomedical Engineering, Svez. 60, str. 427-436. URL:  
<http://ieeexplore.ieee.org/xpl/articleDetails.jsp?tp=&arnumber=6357229>.
6. *Silhouette-Based Human Identification from Body Shape and Gait*. **Collins, R.T., Gross, R. i Shi, J.** Washington, DC, USA : an., 2002. Fifth IEEE International Conference on Automatic Face and Gesture Recognition. str. 366-371. URL:  
<http://www.computer.org/csdl/proceedings/fg/2002/1602/00/16020366-abs.html>.
7. *1.2.3 Human Identity and Gender Recognition From Gait Sequences With Arbitrary Walking Directions*. **Lu, J., Wang, G. i Moulin, P.** 1, 2014, IEEE Transactions on Information Forensics and Security, Svez. 9, str. 51-61. URL:  
<http://ieeexplore.ieee.org/xpl/articleDetails.jsp?tp=&arnumber=6671367>.
8. *A Gaussian Process Guided Particle Filter for Tracking 3D Human Pose in Video*. **Sedai, S., Bennamoun, M. i Huynh, D.Q.** 11, 2013, IEEE Transactions on Image Processing, Svez. 22, str. 4286-4300. URL:  
<http://ieeexplore.ieee.org/xpl/articleDetails.jsp?tp=&arnumber=6553289>.
9. *Probabilistic gait modelling and recognition*. **Hong, S., Lee, H. i Kim, E.** 1, 2013, IET Computer Vision, Svez. 7, str. 56-70. URL:  
<http://ieeexplore.ieee.org/xpl/articleDetails.jsp?tp=&arnumber=6518026>.