

Identifikacija osoba RGB-dubinskim kamerama

Ramakić, Adnan

Doctoral thesis / Disertacija

2021

Degree Grantor / Ustanova koja je dodijelila akademski / stručni stupanj: **University of Rijeka, Faculty of Engineering / Sveučilište u Rijeci, Tehnički fakultet**

Permanent link / Trajna poveznica: <https://um.nsk.hr/um:nbn:hr:190:890559>

Rights / Prava: [In copyright](#) / [Zaštićeno autorskim pravom.](#)

Download date / Datum preuzimanja: **2024-04-19**



Repository / Repozitorij:

[Repository of the University of Rijeka, Faculty of Engineering](#)



SVEUČILIŠTE U RIJECI
TEHNIČKI FAKULTET

Adnan Ramakić

**IDENTIFIKACIJA OSOBA
RGB-DUBINSKIM KAMERAMA**

DOKTORSKA DISERTACIJA

Rijeka, 2021.

SVEUČILIŠTE U RIJECI
TEHNIČKI FAKULTET

Adnan Ramakić

**IDENTIFIKACIJA OSOBA
RGB-DUBINSKIM KAMERAMA**

DOKTORSKA DISERTACIJA

Mentor: izv. prof. dr. Kristijan Lenac

Rijeka, 2021.

UNIVERSITY OF RIJEKA
FACULTY OF ENGINEERING

Adnan Ramakić

**PEOPLE IDENTIFICATION WITH
RGB-DEPTH CAMERAS**

DOCTORAL DISSERTATION

Rijeka, 2021.

Mentor disertacije: izv. prof. dr. Kristijan Lenac

Doktorska disertacija obranjena je dana _____ na Tehničkom fakultetu Sveučilišta u Rijeci pred povjerenstvom u sastavu:

1. prof. dr. sc. Ivo Ipšić - predsjednik (Sveučilište u Rijeci, Tehnički fakultet)
2. doc. dr. sc. Sandi Ljubić - član (Sveučilište u Rijeci, Tehnički fakultet)
3. prof. dr. sc. Renato Filjar - član (Veleučilište Hrvatsko zagorje Krapina)

ZAHVALA

Zahvaljujem svojem mentoru prof. dr. Kristijanu Lencu na strpljenju, uputama i motivaciji tijekom cijelog doktorskog studija, a na kraju i tijekom izrade ove doktorske disertacije. Bez njegove pomoći većina ovog ne bi bila moguća.

Posebno zahvaljujem svojim roditeljima na strpljenju i neprekidnoj podršci koja je dovela do ove točke u obrazovanju. Sve ispite, seminare, završne radove i ostalo prošli smo zajedno, kao i sve dobre i loše stvari u životu. Stoga i ne postoje riječi kojima bih im mogao zahvaliti. Također zahvaljujem i svojoj sestri na uvijek šaljivim i dobrim komentarima tijekom ovog studija, koji su uvijek popravljali raspoloženje i gurali me prema cilju.

Na kraju zahvaljujem svojoj supruzi Neiri i kćerci Adni koje su cijelo vrijeme bile moja najveća moralna podrška i najveći poticaj za završetak ovog studija.

SAŽETAK

U ovoj doktorskoj disertaciji razmotren je problem identifikacije osoba upotrebom RGB-dubinske (engl. Red, Green, Blue - Depth, RGB-D) kamere. RGB (engl. Red, Green, Blue) kamera je uređaj koji omogućava stjecanje slika u boji (RGB slike), a RGB-D kamera je uređaj koji pored slika u boji pruža i dubinske slike. Dubinska slika (još se naziva i dubinska mapa) je slika čiji elementi, pikseli, imaju kodiranu vrijednost koja predstavlja udaljenost objekta od kamere.

Osnovna motivacija istraživanja je bila realizirati pristup za identifikaciju osoba koji se može primijeniti u uslužnoj robotici, ali i u drugim područjima kao što su, primjerice, sigurnosni sustavi. Uslužni robot predstavlja vrstu robota koji djeluje djelomično ili potpuno autonomno s ciljem obavljanja postavljenih zadataka za ljude ili opremu. Napretkom tehnologije otvorene su mogućnosti zajedničkog djelovanja uslužnih robota i ljudi u istom okruženju. Uslužni roboti će biti u stanju pomagati osobi u njenom domu, bolnici, uredu ili nekom drugom mjestu. Uslužni robot često mora biti u stanju, ne samo detektirati prisustvo osobe, nego i utvrditi identitet osobe s kojom je u interakciji da bi na učinkovit način izvršio zadatak koji se pred njega postavlja. Za tu svrhu uslužni roboti mogu upotrebljavati razne vrste osjetila i kamera kojima su opremljeni. Većina uslužnih robota novije generacije opremljena je RGB-D kamerama.

Pod pretpostavkom da je analizom opažanja svoje kamere (ili više kamera) uslužni robot detektirao prisustvo osobe, *identifikacija osoba* podrazumijeva proces u kojemu robot upotrebom odgovarajućeg algoritma obavlja podudaranje značajki izlučenih iz slika jedne ili više kamera, koje se odnose na detektiranu osobu, sa značajkama predložaka (engl. Template) spremjenih u bazi osoba (engl. Gallery), a s ciljem određivanja specifične osobe. Jedan predložak u bazi osoba sadržava skup značajki (primjerice, značajke lica ili značajke hoda) koji se odnosi na određenu osobu. Pored skupa značajki svakoj osobi se mogu pridružiti i dodatne informacije poput imena i prezimena, uloge, titule i slično. Navedene dodatne informacije mogu biti relevantne u nekim primjenama jer se robot osobi može obratiti, primjerice, imenom ili prezimenom. Predlošci se spremaju u bazu osoba u postupku upisa osoba u bazu (engl. Enrollment) i predstavljaju identitet

osobe u kontekstu navedene primjene.

Tradicionalni pristup za identifikaciju osoba bio je zasnovan na primjeni RGB kamera, no danas su RGB-D kamere široko dostupne, pružaju RGB slike kao i RGB kamere, a dodatno pružaju i dubinske slike pa se danas sve više upotrebljavaju u ovom području umjesto RGB kamera. Prednost upotrebe dubinskih slika jest mogućnost izlučivanja određenih značajki koje se mogu dobiti isključivo analizom dubinskih slika. Takva značajka je, primjerice, visina osobe. U disertaciji su istražene postojeće metode za identifikaciju osoba RGB kamerama, kao i metode temeljene na RGB-D kamerama.

Znanstveni doprinos disertacije predstavlja prijedlog novih metoda za identifikaciju osoba RGB-D kamerom na temelju njihova hoda. Općenito, prednost metoda identifikacije osoba na temelju hoda jest u tome što takav način identifikacije ne zahtijeva interakciju s osobom kao što je to slučaj s većinom drugih metoda identifikacije poput identifikacije otiskom prsta, identifikacije na temelju elemenata oka (šarenica ili mrežnica oka) i drugih. Predložene metode za identifikaciju osoba na temelju hoda kombiniraju značajke koje se mogu dobiti iz RGB ili dubinskih slika (siluete osobe) i značajku dobivenu iz dubinskih slika (visina osobe).

U eksperimentalnom istraživanju vrednovane su predložene metode te su prezentirani rezultati. Tijekom istraživanja upotrebljavana su tri skupa podataka koji sadržavaju slike ili videozapise s osobama u hodu, od kojih dva javno dostupna skupa podataka koje često upotrebljavaju i drugi autori te novi, vlastiti skup podataka. Vlastiti skup podataka sadržava videozapise visoke razlučivosti za 14 osoba u hodu, a specifičan je po tome što je kreiran s RGB-D kamerom, koja ima doomet do 20 m, u vanjskom okruženju.

Iz dobivenih rezultata vidljivo je da predložene metode postižu veću točnost identifikacije osoba u odnosu na metode identifikacije osoba na temelju hoda koje su upotrebljavane za usporedbu kao što su slika energije hoda (engl. Gait Energy Image, GEI) i popunjena slika energije hoda (engl. Backfilled Gait Energy Image, BGEI).

ABSTRACT

In this dissertation, the problem of person identification with an RGB-Depth (Red, Green, Blue - Depth, RGB-D) camera was considered. An RGB camera is a device that can capture color images (RGB images). An RGB-D camera, on the other hand, is a device that provides depth images in addition to color images. A depth image (also called a depth map) is an image in which a value is encoded in each pixel that represents the distance of an object from the camera.

The main motivation was to implement a person identification approach that can be used in service robotics, but also in other areas such as security systems. A service robot is a type of robot that operates semi- or fully autonomously to perform specific tasks. Technological advancements offer many opportunities for service robots to work alongside humans in an environment. In the future, service robots will be able to help people in their homes, hospitals, offices, or other locations. A service robot often needs to not only detect the presence of a person, but also be able to determine the identity of the person it is interacting with in order to efficiently perform a specific task assigned to it. For this purpose, service robots can be equipped with and use different types of sensors and cameras. New generation service robots are usually equipped with RGB-D cameras.

In this dissertation, the term *person identification* refers to the process of matching the features extracted from the robot's camera input data related to the detected person with the features stored in the database (Templates). The database is referred to as the *gallery*. A template in the gallery contains a set of features (e.g., facial features or gait features) related to a particular person. In addition to a set of features, each person may be accompanied by additional information such as first name, last name, function, title, etc. The above additional information may be important in some applications, as the robot may address the person by first or last name, for example. The templates are stored in the gallery in a process called *enrollment*, and in this context represent the identity of a person.

In the past, RGB cameras were typically used in this area, but today RGB-D cameras are widely used and also provide depth images, so they are increasingly

used in this area instead of RGB cameras. The advantage of using depth images is the ability to obtain certain features that can only be obtained by using depth images. For example, one such feature is the height of a person. In this dissertation, existing methods for person identification using RGB cameras as well as newer methods based on RGB-D cameras have been investigated.

The scientific contribution of this dissertation is a proposal for two new methods using RGB-D camera for person identification based on gait. The advantage of the methods for gait identification is that they do not require any interaction with the person to be identified. This distinguishes them from many other methods, e.g. fingerprint, identification based on eye features (iris or retina), etc. The newly proposed methods for person identification combine features that can be obtained from RGB or depth images (silhouettes of a person) and a feature that can only be obtained from depth images (height of a person).

In the experimental study, the proposed methods were validated and the results are presented. For the study, three datasets of images and videos of people were used for gait recognition. These are two publicly available datasets that are commonly used by other authors and are well-known in the research community, and a new user-defined dataset. The new dataset contains high-resolution videos of 14 people walking and was created using an RGB-D camera that has a range of up to 20 m and is suitable for outdoor environments.

The results show that the proposed methods achieve better results than classical appearance-based gait recognition methods such as Gait Energy Image (GEI) and Backfilled Gait Energy Image (BGEI).

KLJUČNE RIJEČI / KEYWORDS

- Identifikacija osoba
- Identifikacija osoba na temelju hoda
- RGB kamera
- RGB-dubinska kamera (RGB-D kamera)
- ZED stereokamera
- Dubinska slika
- Slika energije hoda (GEI)

- People Identification
- Gait Recognition
- RGB Camera
- RGB-Depth Camera (RGB-D Camera)
- ZED Stereo Camera
- Depth Image
- Gait Energy Image (GEI)

Sadržaj

1	UVOD	1
1.1	Cilj i hipoteze istraživanja	4
1.2	Pregled dosadašnjih istraživanja	5
1.3	Znanstveni doprinos	14
1.4	Struktura doktorske disertacije	16
2	IDENTIFIKACIJA OSOBA	18
2.1	Biometrijske metode	19
2.1.1	Identifikacija na temelju otiska prsta	24
2.1.2	Identifikacija na temelju lica	26
2.1.3	Identifikacija na temelju šarenice i mrežnice oka	27
2.1.4	Identifikacija na temelju hoda	28
2.1.5	Ostale važnije biometrijske metode	31
2.1.6	Usporedba različitih biometrijskih metoda	34
3	RGB I RGB-D KAMERE	37
3.1	RGB kamere	37
3.2	RGB-D kamere	39
3.2.1	Stereovid	41
3.2.2	Time-of-flight kamere	45
3.2.3	Kamere sa strukturiranom svjetlošću	47
4	SKUPOVI PODATAKA	52
4.1	Skup podataka Casia Dataset B	52
4.2	Skup podataka TUM-GAID	54
4.3	Vlastiti skup podataka	54
5	PREDLOŽENI PRISTUPI ZA IDENTIFIKACIJU OSOBA	61
5.1	Izlučivanje značajki	63
5.1.1	Izlučivanje silueta	63

5.1.2	Normalizacija silueta	65
5.1.3	Određivanje ciklusa hoda osobe	66
5.1.4	Kreiranje GEI slika	68
5.1.5	Određivanje visine osobe iz siluete	68
5.1.6	Smanjenje dimenzionalnosti podataka	69
5.2	Opis predloženih metoda	73
5.2.1	Visina – GEI integracija (HGEI-i)	74
5.2.2	Visina – GEI fuzija (HGEI-f)	75
6	EKSPERIMENTALNI REZULTATI	78
6.1	Analiza silueta izlučenih iz RGB i dubinskih slika	78
6.2	Identifikacija osoba	81
6.2.1	Eksperimentalne postavke	82
6.2.1.1	Opis kreiranih GEI i BGEI slika	82
6.2.1.2	Razdiobe visina osoba	82
6.2.1.3	Klasifikacija	84
6.2.1.4	Testirane metode	86
6.2.2	Mjere uspješnosti	86
6.2.3	Rezultati	89
7	ZAKLJUČAK	103
	Literatura	106
	Popis slika	119
	Popis tablica	121
A	Prilog	122
A.1	Izlučivanje silueta	122
A.2	PCA i LDA analiza	126
A.3	Alfa-integracija	127
A.4	Klasifikacija	130

Životopis	133
Popis radova	135

1 UVOD

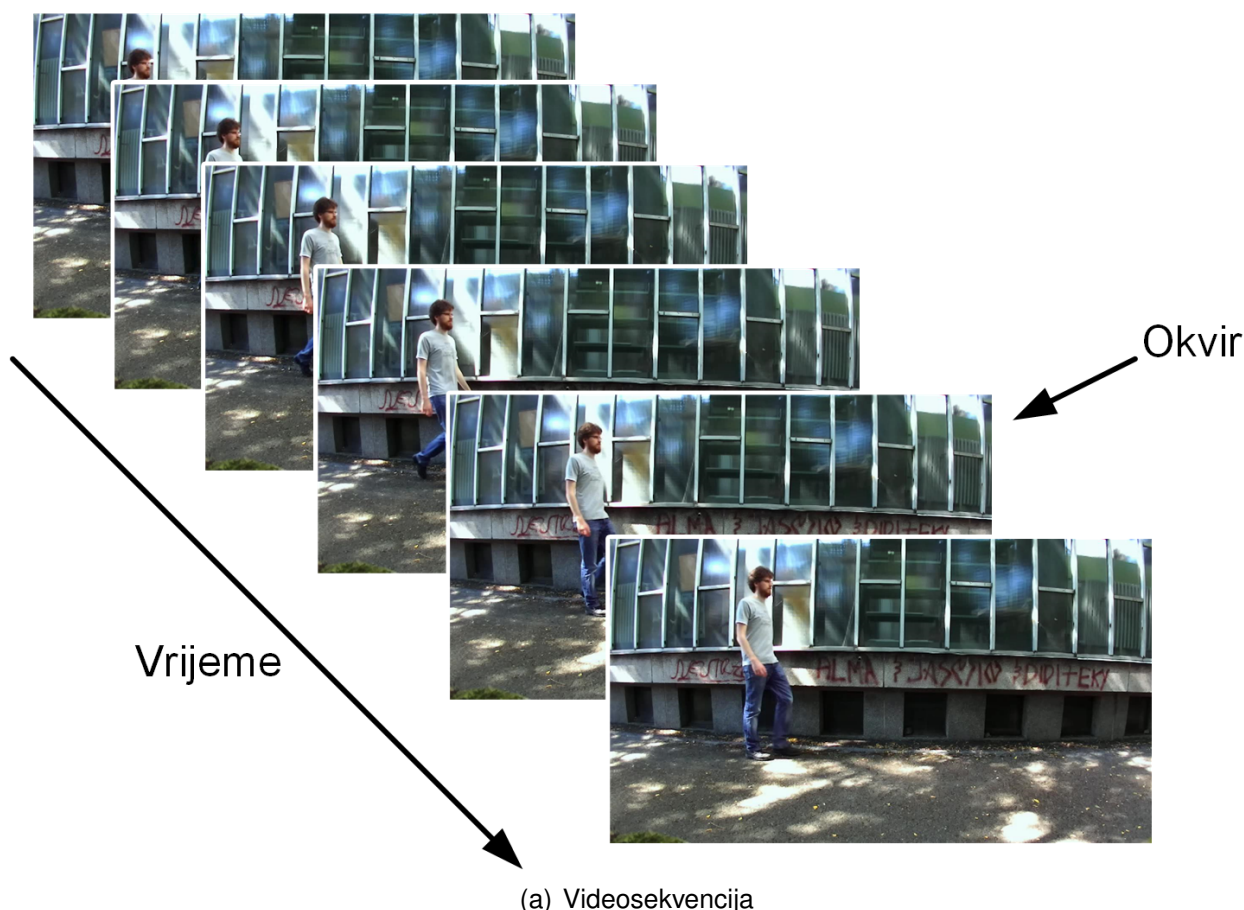
U ovoj disertaciji istražen je problem identifikacije osoba upotrebom RGB-dubinske kamere¹ (engl. **R**ed, **G**reen, **B**lue - **D**epth, skraćeno RGB-D). Danas se za identifikaciju osoba često upotrebljavaju RGB (engl. **R**ed, **G**reen, **B**lue) i RGB-D kamere. Naziv *RGB* dobiven je skraćivanjem prvih slova engleskih naziva za crvenu, zelenu i plavu boju, dok je u slučaju *RGB-D* dodano i prvo slovo engleske riječi za dubinu.

RGB kamera je kamera čiji je rezultat mjerenja digitalna slika u boji pohranjena u RGB formatu koji u svakom pikselu slike (piksel predstavlja najmanji element slike) definira vrijednost intenziteta koji odgovara crvenoj, zelenoj i plavoj boji. Kontinuirani slijed slika tj. okvira (engl. Frame) koji prikazuju određenu radnju naziva se videosekvencija. RGB-D kamera uz slike u RGB formatu pruža i dubinske slike. Dubinska slika (još se naziva i dubinska mapa) predstavlja sliku gdje je u svakom pikselu slike kodirana vrijednost koja predstavlja udaljenost objekta u sceni od kamere. Slika 1.1 prikazuje primjer videosekvencije (slika 1.1 (a)), RGB slike (slika 1.1 (b)) te dubinske slike (slika 1.1 (c)). Imajući u vidu opisano, predmet istraživanja u ovoj disertaciji bila je identifikacija osoba analizom RGB i dubinskih slika iz videosekvencija RGB-D kamera.

Osnovna motivacija istraživanja bila je realizirati pristup za identifikaciju osoba koji se može primijeniti u uslužnoj robotici, ali i u drugim područjima kao što su, primjerice, sigurnosni sustavi. Uslužna robotika posljednjih godina sve više postaje dijelom ljudskih života. Napretkom tehnologije otvorene su mogućnosti da uslužni roboti djeluju u okruženju zajedno s ljudima. Uslužni robot predstavlja vrstu robota koji djeluje djelomično ili potpuno autonomno s ciljem obavljanja postavljenih zadataka za ljude ili opremu te će uslužni roboti u budućnosti biti u stanju pomagati osobi u njenom domu, bolnici, uredu ili nekom drugom mjestu. U nekim praktičnim primjenama, uslužni robot mora biti u stanju, ne samo detektirati prisustvo osobe nego i, utvrditi identitet osobe s kojom je u interakciji da bi na efikasan način izvršio zadatak koji se pred njega postavlja. Primjerice, tijekom rada u bolnici robot mora donijeti lijek točno određenoj osobi ili prenijeti neku povjerljivu informaciju samo određenoj osobi. Uslužni roboti novije generacije danas su često opremljeni RGB-D kamerama.

Pod pretpostavkom da je analizom opažanja svoje kamere (ili više kamera) uslužni

¹Ovisno o kontekstu, u literaturi se za RGB-D kamere također upotrebljava i općeniti naziv RGB-D osjetila (engl. RGB-D sensors).



Slika 1.1: Primjer videosekvencije, RGB i dubinske slike

robot detektirao prisustvo osobe, *identifikacija osoba* podrazumijeva proces u kojemu robot upotrebom odgovarajućeg algoritma obavlja podudaranje značajki izlučenih iz ulaznih slika dobivenih od strane jedne ili više kamera, koje se odnose na detektiranu osobu, sa značajkama predložaka (engl. Template) spremljenih u bazi osoba (engl. Gallery), a s ciljem određivanja specifične osobe. Jedan predložak u bazi osoba sadržava skup značajki (primjerice, značajke hoda) koji se odnosi na jednu određenu osobu. Pored skupa značajki svakoj osobi se mogu pridružiti i dodatne informacije po-

put imena i prezimena, uloge, titule i slično. Navedene dodatne informacije mogu biti relevantne u nekim primjenama jer se robot osobi može obratiti, primjerice, imenom ili prezimenom. Predlošci se spremaju u bazu osoba u postupku upisa osoba u bazu (engl. Enrollment) i predstavljaju identitet osobe u kontekstu navedene primjene.

Postoje različite metode za identifikaciju osoba koje su danas u upotrebi, u ovisnosti o specifičnosti ciljanih primjena, a realizirane su upotrebom različitih uređaja i vrsta ulaznih podataka (primjerice, slika, zvuk itd.). Primjeri često upotrebljivanih metoda jesu identifikacija na temelju otisaka prstiju, identifikacija na temelju lica osobe, identifikacija na temelju elemenata oka kao što su šarenica i mrežnica oka, identifikacija na temelju glasa itd. Većina navedenih metoda zahtijeva interakciju s osobom na način da je potrebno doći do uređaja koji, primjerice, uzima sliku šarenice oka u svrhu podudaranja značajki iz navedene slike s onima spremljenim u bazu kako bi se identificirala osoba. To je često manjkavost jer većina navedenih metoda zahtijeva potpunu suradnju osobe koju je potrebno identificirati. Postoji skupina metoda koja ne zahtijeva izravnu interakciju s osobom, primjerice, metode identifikacije osoba na temelju lica ili hoda.

Nakon proučavanja postojećih metoda identifikacije osoba, znanstveni doprinos istraživanja ostvaren je u području identifikacije osoba na temelju hoda. Izrađen je vlastiti skup podataka koji čine videozapisi osoba u hodu visoke razlučivosti. Osobe su snimljene na većoj udaljenosti i u drugačijim uvjetima (okruženje, osvjetljenje itd.) negoli je to slučaj u dostupnim RGB-D skupovima koji se često koriste od strane drugih autora. Navedeni vlastiti skup podataka upotrebljavan je tijekom preliminarnih istraživanja te za kvalitativnu analizu silueta dobivenih iz RGB i dubinskih slika, odnosno usporedbu RGB i dubinskih slika osoba u hodu.

Predstavljena su i dva nova pristupa za identifikaciju osoba na temelju hoda u kojima se kombiniraju značajke osobe vezane uz pojavu osobe (siluete) s dodatnom značajkom (visina osobe), a koje je moguće izlučiti analizom dubinskih slika kamere tijekom ciklusa hoda osobe. Za testiranje metoda temeljenih na predloženim pristupima upotrebljavan je skup podataka koji je javno dostupan i često upotrebljavan od strane drugih autora. Javno dostupni skup podataka je upotrebljavan zbog toga što sadržava velik broj osoba u hodu s različitim varijacijama hoda kao što su normalan hod, hod s nošenjem ruksaka te varijacije u cipelama. Upotrebom većeg skupa podataka može biti

analiziran utjecaj, primjerice, veličine populacije ljudi na rezultate identifikacije osoba.

Osim u uslužnoj robotici, identifikacija osoba iz videosekvencija RGB-D kamera koristi se i u drugim područjima primjene kao što su, primjerice, sigurnosni sustavi ili autoindustrija. Rezultati ovog istraživanja mogu biti primijenjeni i u navedenim drugim područjima.

1.1 Cilj i hipoteze istraživanja

Identifikacija osobe iz videosekvencija kamere zahtjevan je problem jer se s vremenom mijenjaju informacije povezane s osobom kao što su izgled, položaj osobe u sceni, odjeća, položaj tijela, izraz lica itd., a istovremeno se mijenjaju i sama scena u kojoj se osoba nalazi kao i vanjski uvjeti koji utječu na snimku, primjerice, osvjetljenje.

Cilj istraživanja u ovoj doktorskoj disertaciji bio je razmotriti i istražiti problem identifikacije osoba iz videosekvencija RGB-D kamere i predložiti novi pristup (metodu) za identifikaciju osoba koji unaprjeđuje dosadašnje metode utemeljene na RGB, odnosno RGB-D kamerama i/ili proširuje mogućnosti njihove primjene.

Novi pristup bio bi utemeljen na postojećim pristupima koji upotrebljavaju RGB kamere, a koji bi bili izmijenjeni na način da se pored značajki izlučenih iz RGB slika upotrebljavaju i dodatne značajke izlučene iz dubinskih slika RGB-D kamere. Značajke bi bile integrirane, tj. spojene, u jednu skupinu značajki koje bi se upotrebljavale u procesu identifikacije.

U vezi s prethodno navedenim definirane su sljedeće hipoteze istraživanja:

1. Integracijom značajki izlučenih iz dubinskih slika videosekvencija RGB-D kamera sa značajkama izlučenim iz RGB slika videosekvencija RGB-D kamera može se poboljšati točnost identifikacije osoba u odnosu na metode temeljene samo na RGB kamerama.
2. Integracijom značajki izlučenih iz dubinskih slika RGB-D kamera sa značajkama izlučenim iz RGB slika istih kamera identifikacija se može primijeniti i na videosekvencijama gdje značajke izlučene iz RGB slika zasebno nisu dostatne za identifikaciju.

Većina dosad predstavljenih metoda temeljena je samo na značajkama izlučenim iz RGB slika dobivenim upotrebom RGB kamere. Značajka koja se često upotrebljava kada se koriste samo RGB slike je silueta osobe, više silueta osobe ili upotreba nekih

dijelova siluete kao što je, primjerice, donji dio siluete koji čine noge osobe. Upotrebom RGB-D kamera moguće je upotrebljavati i dubinske slike uz RGB slike. S pomoću dubinskih slika mogu se dobiti dodatne značajke kao što su visina osobe, dužina koraka, dužina ruke ili noge i slično te ih se može kombinirati sa značajkama dobivenima iz RGB slike u svrhu identifikacije osoba.

Također, dostupnost dubinskih slika pored RGB slika omogućava lakšu segmentaciju osobe u sceni, odnosno grupiranje piksela na slici koji pripadaju osobi i njihovo izdvajanje od pozadine. Upotrebom samo RGB slika segmentacija osobe može biti problematična kada se, primjerice, preklapaju boja pozadine i boja odjeće na osobi.

1.2 Pregled dosadašnjih istraživanja

Znanstvenici i istraživači predstavili su veći broj radova u dosadašnjim istraživanjima u području identifikacije osoba upotrebom RGB-D kamera. RGB i RGB-D kamere se, osim za identifikaciju osoba, često upotrebljavaju i u kontekstu detektiranja, praćenja i reidentifikacije osoba.

Detektiranje osobe odnosi se na pronalaženje osobe na slici ili videosekvenciji dok se praćenje osobe odnosi na postupak vremenskog povezivanja detekcija osobe unutar videosekvencije u svrhu, primjerice, generiranja putanje osobe.

Reidentifikacija osobe podrazumijeva ponovnu identifikaciju osobe koja je već ranije identificirana ili je pak identificirana na drugome mjestu u prostoru ili s drugim kamerama. Postoje dvije vrste reidentifikacije: kratkotrajna (engl. Short-Term Re-Identification) i dugotrajna (engl. Long-Term Re-Identification). Kratkotrajna reidentifikacija podrazumijeva reidentifikaciju unutar kratkog vremenskog razdoblja pri čemu osoba, primjerice, nosi istu odjeću kao pri zadnjoj identifikaciji. Kod dugotrajne reidentifikacije potrebno je upotrebljavati dugotrajnije značajke (primjerice, značajke lica). U radovima [6], [15], [70], [63], [29], [86], [56], [11], [67], [66], [62], [68], [75], [5], [21] i [69] istražene su detekcija, praćenje i reidentifikacija osoba.

U slučaju identifikacije osoba na temelju hoda, upotrebom RGB-D kamere, u literaturi dominiraju dva pristupa: pristup temeljen na modelu (engl. Model-based Approach) i pristup temeljen na pojavi (engl. Appearance-based Approach, Model Free). Pristup temeljen na modelu uključuje neke informacije o tijelu, primjerice, informacije o kosturu osobe, dok pristup temeljen na pojavi obično uključuje izlučivanje i upotrebu siluete

osobe. U nastavku ovog potpoglavlja opisani su važniji radovi u kojima su predloženi različiti pristupi za identifikaciju osoba na temelju hoda, a koji se upotrebljavaju ili se mogu upotrebljavati s RGB-D kamerama.

Jedan od radova koji je među najčešće spominjanima kada se radi o identifikaciji osoba na temelju hoda jest rad autora Hana i Bhanua [28]. U radu je predstavljena tada nova prostorno-vremenska reprezentacija hoda koja je nazvana slika energije hoda (engl. Gait Energy Image, GEI). GEI predstavlja slijed ljudskog kretanja u jednoj slici. U radu [28] je predstavljen i pristup za učenje i identifikaciju individualnih osobina hoda upotrebom GEI predložaka. U metodi koja je predstavljena izvorno u radu upotrebljavane su RGB slike.

Proceduru izlučivanja siluete osobe (ili silueta osoba) autori su radili kao radnju predobrade, a što kasnije uključuje i normaliziranje veličine svake siluete, smještanje visine siluete u fiksnu visinu slike te slijedno horizontalno poravnanje. Na slici 1.2 prikazani su primjeri normaliziranih i poravnatih okvira silueta u različitim sljedovima hoda. Krajnja desna slika (slika 1.2) predstavlja prosječnu sliku silueta osobe, GEI, dok ostale slike predstavljaju siluetu osobe tijekom ciklusa hoda (na primjeru triju osoba).



Slika 1.2: Primjeri normaliziranih i poravnatih okvira silueta u različitim sljedovima hoda s prosječnom slikom siluete (GEI) [28]

U radu [28] autori su razvili dva pristupa za identifikaciju osoba od predložaka, a to su direktno GEI podudaranje i statističko podudaranje GEI značajki. U direktnom GEI

podudaranju za identifikaciju osoba mjeri se sličnost između predložaka za trening i predložaka za test, dok se kod statističkog podudaranja prvo generiraju novi predlošci od predložaka za trening. Nakon toga značajke se utvrđuju iz proširenih predložaka za trening metodom analize glavnih komponenti (engl. Principal Component Analysis, PCA) za smanjenje dimenzionalnosti predložaka i metodom analiza višestruke diskriminacije (engl. Multiple Discriminant Analysis, MDA) za postizanje boljih odvojivosti klasa. Osoba je identificirana na temelju naučenih značajki. Uspješnost predloženih pristupa provjerili su na skupu podataka Sveučilišta Južna Florida (engl. University of South Florida, USF Human ID) [99].

Sivapalan i sur. [91] proširili su koncept slike energije hoda (GEI) i predstavili volumen energije hoda (engl. Gait Energy Volume, GEV). Predstavili su tehniku temeljenu na pojavi, odnosno na GEI-ju, pri čemu su predložili 3D analogiju GEI-ja te su umjesto slijeda ljudskog kretanja u jednoj slici zasnovanog na siluetama osoba radili na rekonstruiranim vokselima (pojam *voksel* generalno se odnosi na najmanji prepoznatljivi element u 3D prostoru, odnosno na jedinicu informacije koja opisuje točku u 3D prostoru; još se naziva i 3D piksel). U eksperimentima su upotrebljavali dva skupa podataka: skup podataka kretanja tijela Carnegie Mellon Sveučilišta (engl. Carnegie Mellon University Motion of Body, CMU MoBo) [26] [95] i vlastiti skup podataka koji su kreirali Kinectom. U zamišljenim eksperimentima izvršena je usporedba s GEI metodom te je utvrđeno da GEV metoda postiže veće rezultate točnosti identifikacije na temelju hoda u odnosu na GEI, a u obzir su u eksperimentima uzete sljedeće metode: GEV, GEV dobiven upotrebom prednje strane osobe, GEI dobiven upotrebom prednje i bočne strane osobe te GEI iz više pogleda.

Hofmann i sur. [32] u svojem su pristupu također proširili GEI na način da su upotrebljavali dubinske slike za izlučivanje silueta osobe (dubinski GEI). Također su formilirali novu značajku koju su definirali kao sliku dubinskog gradijentnog histograma energije (engl. Depth Gradient Histogram Energy Image, DGHEI) te su usporedili dubinski GEI, DGHEI i normalni GEI. Također su upotrebljavali novi skup podataka koji su kreirali na Tehničkom sveučilištu u Münchenu koristeći se Kinectom (engl. TUM Gait from Audio, Image and Depth (GAID), TUM-GAID) [33]. U istraživanju su provedeni eksperimenti na različitim varijacijama hoda kao što su normalno hodanje, nošenje ruksaka i cipele. Uspoređena je točnost identifikacije predložene metode (DGHEI) i

GEI-ja, dubinskog GEI-ja i GEV-a. Prema dobivenim rezultatima navedenima u radu [32] predloženi DGHEI imao je bolje rezultate u odnosu na metode koje su upotrebljavane za usporedbu. Prosječna točnost identifikacije kada se u obzir uzmu sve varijacije na TUM-GAID skupu podataka iznosila je 82 %.

Borràs i sur. [10] predstavili su DGait, novu bazu hoda dobivenu RGB-D kamerom. DGait je kreiran upotrebom Kinecta u unutarnjem okruženju i sadržava videozapise za 53 subjekta (osobe) koji hodaju u različitim smjerovima. Svaki videozapis označen je prema subjektu, spolu i godinama. Također su proveli eksperiment određivanja spola na temelju hoda osobe primjenjujući spomenuti skup podataka kako bi prikazali korisnost dubinskih slika u tom kontekstu. U eksperimentima su izlučivali 2D i 3D značajke hoda koje su zasnovane na deskriptorima oblika tijela, a upotrebljavali su algoritam klasifikacije potpornih vektora (engl. Support Vector Machine, SVM) s navedenim značajkama.

Lu i sur. [57] istraživali su problem određivanja spola i identifikacije osoba iz izlučenih silueta osoba tijekom sekvencija hoda i sa slobodnim smjerovima hodanja. Predstavili su metodu metričkog učenja zasnovanu na rijetkoj obnovi (engl. Sparse Reconstruction Based Metric Learning, SRML) za izlučivanje značajki hoda. Kreirali su i novi skup podataka u kojem su upotrebljavali Kinect u unutarnjem okruženju, a koji čine osobe u hodu sa slobodnim smjerom kretanja (engl. Advanced Digital Sciences Center – Arbitrary Walking Directions, ADSC – AWD) [93]. U istraživanju su definirali nekoliko eksperimenata te se usporedili s metodama koje su zasnovane na sličnom principu rada kao i predložena metoda, a to su analiza komponenti susjedstva (engl. Neighborhood Component Analysis, NCA) [25], velika margina najbližeg susjeda (engl. Large Margin Nearest Neighbor, LMNN) [121] i informacijsko-teorijsko metričko učenje (engl. Information-theoretic Metric Learning, ITML) [17]. Realizirali su pristup za identifikaciju osoba na temelju hoda koji je temeljen na predloženoj SRML metodi te su navedeni pristup testirali na ADSC – AWD skupu podataka.

Chattopadhyay i sur. [13] upotrebljavali su slike dobivene Kinectom za identifikaciju osoba na temelju hoda te su predstavili novu značajku koju su nazvali dubinski volumen poza (engl. Pose Depth Volume, PDV). Svoj pristup testirali su na skupu podataka koji je također kreiran upotrebom Kinecta u unutarnjem okruženju. Kombinirali su dubinske i RGB slike dobivene od Kinecta za identifikaciju osoba na temelju hoda, i to s prednje

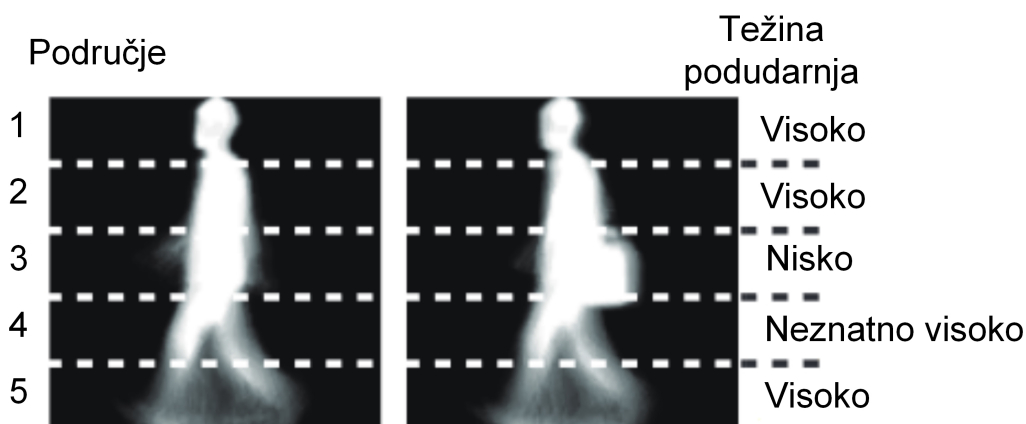
strane osobe. Predloženi pristup usporedili su s GEV i GEI metodama, pri čemu su upotrebljavali različite načine dobivanja siluete za implementaciju navedenih metoda. U slučaju metode GEI upotrebljavali su siluete dobivene iz dubinskih slika, kao i siluete dobivene iz RGB slika. U slučaju GEV-a i PDV-a upotrebljavali su dubinske siluete te dubinske registrirane siluete (odnosi se na kombiniranje RGB i dubinskih slika, gdje se silueta iz dubinskih slika projicira u RGB-u te se tako formira silueta u transformiranom prostoru). Najbolji rezultati točnosti identifikacije dobiveni su u slučaju PDV pristupa u kojem su upotrebljavane dubinske registrirane siluete. Točnost identifikacije bila je oko 80 % na vlastitom skupu podataka kreiranom Kinectom.

Arora i Srivastava [3] predstavili su prostorno-vremenski zasnovanu metodu za identifikaciju osoba u hodu te je nazvali Gaussova slika hoda (engl. Gait Gaussian Image, GGI). GGI predstavlja metodu koja se upotrebljavala za izlučivanje značajki slike hoda tijekom ciklusa hoda. Značajke koje su dobivene GGI-jem klasificirane su primjenom algoritma k-najbližih susjeda (engl. k-Nearest Neighbors, kNN) te primjenom skupa podataka Instituta za automatizaciju Kineske akademije znanosti (engl. The Institute of Automation, Chinese Academy of Sciences, Casia) - Casia Dataset B [130] [125] [102] i skupa podataka Soton [89] [98]. Predstavljenu metodu usporedili su s GEI metodom kao i s metodom koja se naziva slika entropije hoda (engl. Gait Entropy Image, GEnI) [8] gdje se u suštini na jednoj slici kodiraju nasumične vrijednosti piksela u slikama silueta kroz ciklus hoda. Na skupu podataka Soton, GGI je postizao točnost identifikacije sličnu metodama koje su upotrebljavane za usporedbu, dok je u slučaju skupa podataka Casia Dataset B GGI postigao bolje rezultate (oko 98 % točnost).

Nambiar i sur. [73] integrirali su informacije o udaljenosti objekta od kamere u pristup identifikacije osoba na temelju hoda zasnovane na silueti da bi proizveli hibridni 2D - 3D pristup identifikacije osoba na temelju hoda, i to s prednje strane, a informacije su dobivene upotrebom Kinecta, pri čemu je dobiven 3D oblak točaka osobe. Oblak točaka (engl. Point Cloud) predstavlja skup podataka točaka u prostoru gdje je svaka točka definirana x , y i z koordinatama. Značajke su izlučene razmatrajući potpuni 3D oblak točaka i njegove tri 2D projekcije (odozgo, s prednje i s bočne strane osobe). Budući da su autori u istraživanju upotrebljavali 2D i 3D podatke (odnosno značajke dobivene upotrebom navedenih podataka), razmatrali su pojedinačno navedene značajke

kao i fuziju informacija između 2D i 3D pristupa.

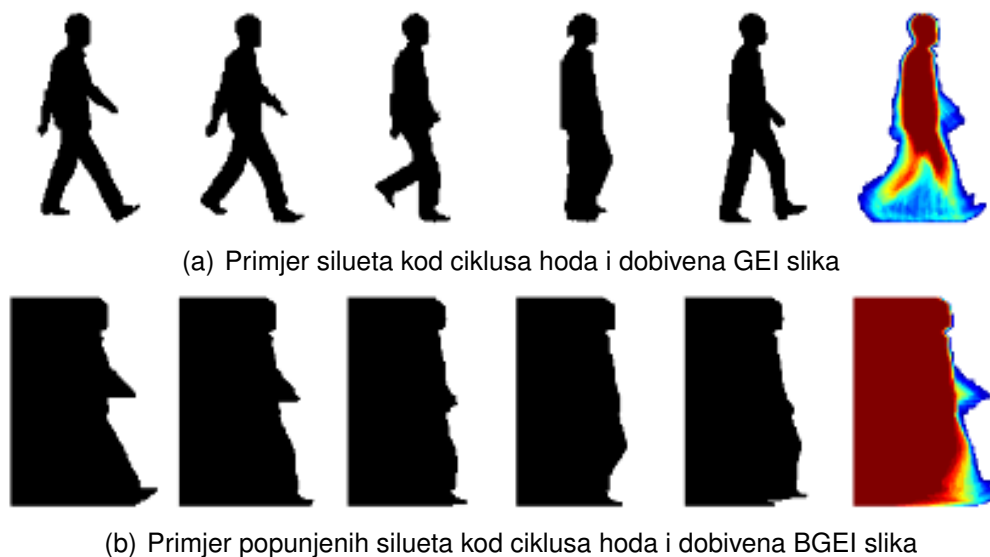
Iwashita i sur. [38] predložili su metodu prema kojoj je slika ljudskog tijela podijeljena u višestruka područja te se značajke izlučuju za svako dobiveno područje. Nakon toga radi se procjena težine podudaranja (engl. Matching Weights) za svako područje kako predlažu autori, a što je utemeljeno na sličnosti između izlučenih značajki i onih koje su spremljene u bazu podataka. Na kraju je osoba identificirana uzevši u obzir ukupnu vrijednost sličnosti podudaranja za sva područja. Primjer prosječne slike osobe u bazi podataka te prosječne slike osobe koja nosi torbu na ramenu prikazan je na slici 1.3, a ljudsko tijelo, odnosno slika podijeljena je na pet područja. Predloženu metodu eksperimentalno su vrednovali na skupovima podataka Casia Dataset B i Casia Dataset C. Predložena metoda imala je najveću točnost identifikacije u slučaju skupa podataka Casia Dataset B (97,7 %), dok je u slučaju skupa podataka Casia Dataset C točnost bila nešto manja (94 %). Također su za provjeru robusnosti predložene metode proveli eksperiment na varijacijama iz skupa podataka Casia Dataset B u smislu da su upotrebljavali varijacije s torbama (kada osoba nosi određenu torbu) (točnost identifikacije 91,9 %) i odjećom (točnost identifikacije 78 %).



Slika 1.3: Primjer prosječne slike u bazi podataka te prosječne slike osobe koja nosi torbu na ramenu – Iwashita, Uchino i Kurazume [38]

Sivapalan i sur. [92] predstavili su popunjenu sliku energije hoda (engl. Backfilled GEI, BGEI), značajku koja može biti konstruirana od silueta s pogleda sa strane (koje mogu biti dobivene iz RGB ili dubinskih slika) i silueta s prednje strane koje se dobivaju iz dubinskih slika. BGEI su uspoređivali s GEI-jem i GEV-om upotrebljavajući skup podataka Casia i vlastiti skup podataka kreiran upotrebom Kinecta. Radi na sličnim pretpostavkama kao i GEV, pri čemu prednja površina modela osobe sadržava sve

bitne informacije o hodu. Upotrebljava se samo prednja kontura siluete osobe kako bi se dobile informacije koje se upotrebljavaju u svrhu identifikacije osoba na temelju hoda. U slučaju silueta promatranih sa strane, BGEI je konstruiran na način da se popune binarne siluete (bijeke siluete na crnoj pozadini ili obrnuto) tako da se pronađe prednji najizbočeniji piksel i od njega se popunjava ostatak slike. Na slici 1.4 (a), (b) ilustriran je odnos GEI-ja i BGEI-ja, pri čemu je pod (a) prikazan ciklus hoda osobe te dobivena GEI slika, a pod (b) ekvivalentna interpretacija BGEI slike.



Slika 1.4: Primjer kreiranih GEI i ekvivalentnih BGEI slika [92]

Preis i sur. [79] predstavili su pristup temeljen na Kinectu u kojem su procjenjivali značajke tijela zajedno s dužinom koraka i brzinom hoda osobe. Upotrebljavali su informacije o kosturu osobe dobivene Kinectom te nisu direktno upotrebljavali dubinske slike i slike u boji. Budući da Kinect detektira i prati 20 različitih točaka kostura osobe, u radu su definirali 13 biometrijskih značajki za identifikaciju osobe (visina, dužina nogu, torzo, obje potkoljenice, oba bedra, obje nadlaktice, obje podlaktice, dužina koraka i brzina). Budući da je 11 značajki statično, zadnje dvije ovise o situaciji. U provedenim eksperimentima na devet osoba uspjeli su dobiti rezultate točnosti identifikacije preko 90 %.

Kumar i Babu [47] predložili su pristup koji primjenjuje 3D informacije o kosturu osobe i kovarijancu putanje točaka zglobova. Navedeni 3D kostur osobe generira se iz dubinskih slika dobivenih upotrebom Kinecta. Za potrebe provođenja eksperimentalnih istraživanja kreirali su skup podataka od 20 osoba. Proveli su nekoliko različitih eksperimenata te su u određenim postavkama dobili rezultate točnosti identifikacije veće od

90 %.

Sinha i sur. [90] predstavili su pristup za identifikaciju osoba od podataka kostura osoba, a upotrebljavali su Kinect. Njihov pristup temeljen je na kosturu osobe koji se može dobiti upotrebom Kinecta i njegova softverskog razvojnog kompleta (engl. Software Development Kit, SDK). Kako navode, njihov je pristup implementiran u tri koraka: snimanje podataka od obrazaca kretanja gledanjem kostura osobe sa strane, generiranje značajki tijekom polovice ciklusa hoda i nadziranog učenja te odabir značajki za identifikaciju osoba. Odabir značajki i klasifikacija realizirani su upotrebom adaptivnih neuronskih mreža (engl. Adaptive Neural Network, ANN).

Kastaniotis i sur. [43] istraživali su probleme identifikacije osoba na temelju hoda te su s tim u vezi predložili pristup koji je originalno razvijen za prepoznavanje aktivnosti osoba i primijenili su ga na hod. Kako navode, novi pristup omogućuje postizanje složenih prikaza sekvencija hoda, što dovodi do efikasnog izražavanja dinamičkih svojstava sekvencija ljudskog hoda. Upotrebljavali su skup podataka koji čini 30 osoba u hodu, a koji je kreiran Kinectom. Tijekom eksperimenta na temelju upotrebljavanog skupa podataka postigli su rezultate točnosti identifikacije veće i od 90 %.

Rida i sur. [83] predložili su metodu izlučivanja značajki koja može odabrati važne diskriminativne značajke za identifikaciju, a u cilju smanjenja utjecaja kovarijanci te poboljšanja uspješnosti identifikacije. Eksperimentalna istraživanja provodili su upotrebljavajući skup podataka Casia Dataset B, pri čemu su u obzir uzeli varijacije normalno, nošenje torbe te nošenje kaputa i dobili prosječne rezultate točno identificiranih osoba za sve navedene varijacije iznad 81 %.

Chattopadhyay i sur. [14] upotrebljavali su dubinske slike dobivene Kinectom za rješavanje problema zaklonjenosti (do koje, primjerice, dolazi kada se prati neki objekt koji u nekom trenutku skriva, odnosno zaklanja drugi objekt; često se upotrebljava i pojam okluzija, engl. Occlusion) u identifikaciji osoba na temelju hoda s prednje strane. Razmatrali su situacije u kojima su kamere postavljene na vrhu ulaza i izlaza iz područja pod nadzorom (tako da se dobije pogled na osobu s prednje i sa stražnje strane). U provedenim eksperimentima dobili su rezultate točnosti identifikacije osoba oko 76 %.

Arora i sur. [2] predstavili su značajke slike hoda temeljene na teoriji skupa informacija (skup informacija proizlazi iz nejasnog skupa u cilju predstavljanja nesigurnosti

u vrijednostima izvora informacija s pomoću funkcije entropije) te su formulirali naziv slika informacije hoda (engl. Gait Information Image, GII). GII je izveden primjenom koncepta informacijskog skupa na okvire u jednom ciklusu hoda te su izlučene dvije značajke, nazvane GII sa značajkom energije (engl. Gait information Image with Energy Feature, GII-EF) i GII sa značajkom sigmoida (engl. Gait Information Image with Sigmoid Feature, GII-SF). Za identifikaciju na temelju hoda primijenjen je algoritam klasifikacije najbližih susjeda (engl. Nearest Neighbour Algorithm). Predložene značajke testirane su na skupu podataka Casia Dataset B, Soton te skupu podataka Instituta za znanstvena i industrijska istraživanja Sveučilišta Osaka (engl. Institute of Scientific and Industrial Research, Osaka University, OU – ISIR Treadmill B) [59] [97] s velikim varijacijama u uvjetima odjeće (skup podataka gdje osobe hodaju na traci). Također, eksperimenti su provedeni i na skupu podataka OU – ISIR Treadmill A s blagim varijacijama u brzini hoda radi demonstracije robusnosti predloženih značajki. Rezultati identifikacije dobiveni s navedenim skupovima podataka razlikuju se te ovise o primijenjenim eksperimentalnim postavkama (primjerice, za skup podataka Casia Dataset B rezultati su od 45 % do 98 % u slučaju GII-SF za jedan tip postavki).

López – Fernández i sur. [54] u istraživanju su predložili novi pristup za identifikaciju osoba na temelju hoda s više pogleda (engl. Multi-View) koji je fokusiran na identifikiranje osoba koje hodaju u slobodnim putanjama. Navedeni pristup, kako navode, funkcionira s osobama koje hodaju slobodno, a da to ne utječe na proces identifikacije čak i kad se radi o zakrivljenim putanjama hoda. S tim u vezi, predložili su novi deskriptor hoda koji se može nositi s promjenama rotacije, a da pritom pruža dovoljno informacija od hoda. Eksperimentalna istraživanja proveli su na dvama skupovima (AVA Multi-View Dataset [55] [94] i Sveučilište Kyushu 4D Gait Database [37] [96]) podataka te su pokazali da navedeni pristup u nekim postavkama postiže rezultate točnosti identifikacije iznad 90 %.

Castro i sur. [12] istraživali su upotrebu konvolucijskih neuronskih mreža (engl. Convolutional Neural Networks, CNN) za učenje deskriptora značajki kretanja u svrhu identifikacije osoba na temelju hoda. Upotrebljavali su skup podataka razvijen na Tehničkom sveučilištu u Münchenu (TUM-GAID) za eksperimentalnu evaluaciju. Autori su u istraživanju proveli tri tipa eksperimenta. Glavni eksperiment je uključivao identifikaciju osoba na temelju hoda u uvjetima s promjenom odjeće i nošenjem različitih

dodataka (primjerice, torbe). Drugi eksperiment u obzir je uzimao promjene u vremenu (pojava osobe u različnim vremenskim razdobljima), a treći eksperiment analizirao je određivanje spola osoba na temelju hoda. U glavnom eksperimentu upotrebljavali su različite algoritme klasifikacije (kao što su SVM i NN) s različitim mjerama uspješnosti te su u nekim slučajevima dobili rezultate točnosti više i od 95 %.

Alotaibi i Mahmood [1] razvili su CNN arhitekturu za identifikaciju osoba na temelju hoda i provodili eksperiment na skupu podataka Casia Dataset B. Kako navode, predložena arhitektura manje je osjetljiva na nekoliko slučajeva uobičajenih varijacija i zaklonjenosti koje utječu na uspješnost identifikacije na temelju hoda, a može podnijeti i relativno male skupove podataka. Tijekom istraživanja provedena su tri eksperimenta na različitim varijacijama (normalan hod, nošenje ruksaka i varijacije u odjeći) skupa podataka Casia Dataset B, a pritom je točnost identifikacije u slučajevima kada je kut između kamere i osoba od 90 stupnjeva i kada osobe normalno hodaju prelazila 90 %.

Lishani i sur. [52] predstavili su pristup izlučivanja značajki na temelju Haralick značajki (Haralick značajke predstavljaju deskriptore teksture u analizi slika te opisuju povezanost intenziteta piksela koji se nalaze jedan do drugog) koje su izlučene od GEI-ja. Navedene značajke izlučene su lokalno te vertikalni ili horizontalni GEI lokalno dijele na dva ili tri jednaka područja od interesa. Za evaluaciju pristupa upotrebljavali su skup podataka Casia Dataset B, a pritom su primjenjivali različite varijacije iz skupa (normalan hod, nošenje torbe i odjeća). Dobili su prosječnu točnost identifikacije za različite varijacije do 80 %.

U istraživanju autora Portillo-Portillo i sur. [78] predstavljen je pristup za identifikaciju osoba na temelju hoda koji upotrebljava GEI i direktnu linearnu diskriminantnu analizu (engl. Direct Linear Discriminant Analysis, DLDA), a koji može utvrditi identitet osobe koja se analizira neovisno o ulaznom kutu.

1.3 Znanstveni doprinos

Znanstveni doprinos predloženog istraživanja utemeljen je na:

- izradi skupa podataka, koji sadržava videozapise visoke razlučivosti za istraživanje metoda za identifikaciju osoba na temelju hoda, upotrebom RGB-D kamere u vanjskom okruženju
- prijedlogu novih pristupa i metoda za identifikaciju osoba na temelju hoda koje

poboljšavaju točnost identifikacije osoba u usporedbi s poznatim metodama identifikacije osoba na temelju hoda upotrebom RGB-D kamere.

Osnovanost stavke koja se odnosi na izradu vlastitog skupa podataka temeljena je na činjenici da većina dosad predstavljenih i javno dostupnih skupova podataka sadržava dubinske slike osoba u hodu dobivene upotrebom Microsoftova Kinect uređaja. Radi tehničkih ograničenja Microsoft Kinect uređaj nije pogodan za snimanje u vanjskom okruženju i ima ograničen domet (obično do 5 m).

Snimanje osoba na većoj udaljenosti od kamere omogućava snimanje više ciklusa hoda u vidnom polju. Veća udaljenost također više odgovara scenariju nadzora šireg područja, što je jedna od primarnih primjena identifikacije osoba na temelju hoda. Snimanje osoba u hodu u vanjskom okruženju s nejednolikom pozadinom i u svjetlosnim uvjetima koji se razlikuju od onih u postojećim skupovima podataka omogućava dodatnu analizu učinkovitosti metoda za segmentaciju osoba i izlučivanja cjelovitih silueta osobe te analizu i usporedbu silueta koje se izlučuju iz RGB i dubinskih slika u navedenim uvjetima. Iz navedenih razloga, cilj je bio kreirati vlastiti novi skup podataka koji bi se razlikovao u odnosu na postojeće po tome da su snimane osobe koje hodaju na većoj udaljenosti od kamere, u vanjskom okruženju i da sadržavaju snimke visoke razlučivosti snimljene u standardnom formatu. Izrada novog skupa podataka temeljila bi se na metodologiji koju su koristili i drugi javno dostupni skupovi podataka osoba u hodu koji sadržavaju dubinske i RGB slike, specifično TUM-GAID [33] skup podataka.

Imajući u vidu navedeno, izrađen je novi skup podataka s pomoću RGB-D kamere većeg dometa koji ispunjava zadane ciljeve. Skup podataka sadržava RGB i dubinske slike i videozapise osoba u hodu visoke razlučivosti, snimljene u vanjskom okruženju. Skup je dostupan na zahtjev svim istraživačima, a u budućnosti se planira i njegova nadogradnja.

Kada je riječ o stavci koja se odnosi na realizaciju pristupa koji će poboljšati točnost identifikacije osoba u usporedbi s poznatim metodama identifikacije osoba na temelju hoda, u disertaciji su predložena dva nova pristupa za identifikaciju osoba na temelju hoda koji kombiniraju značajke izlučene iz RGB slika sa značajkama izlučenim iz dubinskih slika. Na osnovu navedenih predloženih pristupa, za potrebe eksperimentalnog istraživanja i usporedbe s drugim metodama, realizirane su metode koje proširuju GEI i BGEI metodu.

Predložene metode su pri eksperimentalnom vrednovanju uspoređene s poznatim metodama identifikacije osoba na temelju hoda koje se danas često primjenjuju. Tijekom eksperimentalnog vrednovanja razmatrane su različite veličine populacija ljudi, s različitim razdiobama visina i varijacijama hoda. Na upotrijebljenom velikom, javno dostupnom te često korištenom od strane drugih autora skupu podataka dvije nove predložene metode redovito su postizale veću prosječnu točnost identifikacije u odnosu na metode koje su upotrebljavane za usporedbu.

1.4 Struktura doktorske disertacije

Disertacija se sastoji od sedam poglavlja s pripadajućim potpoglavljima. Prvo poglavlje jest uvodno poglavlje u kojem je ukratko opisana problematika kojom se bavi ova disertacija, kao i cilj i hipoteze istraživanja. Detaljno je prikazan pregled dosadašnjih istraživanja, pri čemu su razmotreni i opisani najvažniji radovi iz područja kojemu pripadaju disertacija i istraživanje. Na koncu uvodnog dijela naveden je znanstveni doprinos ove disertacije.

U drugom poglavlju opisane su metode za identifikaciju osoba. Osobe se, općenito gledajući, identificiraju na temelju nekih biometrijskih osobina ili primjenom nekih nebiometrijskih načina. Prvo su ukratko opisani pojmovi identiteta i identifikacije osoba, a potom su detaljnije obrađene biometrijske metode jer je metoda koja je predmet istraživanja u ovoj disertaciji, metoda identifikacije osoba na temelju hoda, biometrijska metoda. Na kraju poglavlja prikazana je usporedba biometrijskih metoda kako bi se stekao uvid u prednosti i nedostatke pojedine metode i usporedila metoda identifikacije temeljena na hodu s drugim metodama identifikacije.

U sljedećem, trećem, poglavlju opisane su RGB i RGB-D kamere. U ovom poglavlju opisani su najpoznatiji modeli boja te osjetila slike koja se upotrebljavaju u RGB kamerama. Vezano za RGB-D kamere opisane su aktivne i pasivne tehnike za dobivanje dubinskih slika. Pored toga, opisani su stereovid, *time-of-flight* (TOF) kamere te kamere sa strukturiranom svjetlošću.

U četvrtom poglavlju detaljnije su opisani skupovi podataka koji su upotrebljavani tijekom istraživanja i koji su relevantni za područje istraživanja. Navedene skupove podataka između ostalog čine slike ili videozapisi osoba u hodu i ti se skupovi često upotrebljavaju tijekom eksperimentalnih istraživanja. Opisana su tri skupa podataka,

od kojih su dva javno dostupna i često upotrebljavana od strane drugih autora, a jedan je vlastito kreirani skup podataka. Detaljnije je opisana i stereokamera koja je upotrebljavana za izradu vlastitog skupa podataka.

U petom poglavlju opisani su predloženi pristupi, tj. nove predložene metode za identifikaciju osoba na temelju hoda. Dodatno su opisani i postupci izlučivanja značajki koje su upotrebljavane s predloženim pristupima.

U šestom poglavlju prikazani su rezultati eksperimentalnog istraživanja. U ovom poglavlju opisana je kvalitativna analiza silueta dobivenih iz RGB i dubinskih slika. Također su opisane eksperimentalne postavke te su prikazani dobiveni rezultati vrednovanja predloženih metoda.

Sedmo poglavlje sadržava zaključna razmatranja.

2 IDENTIFIKACIJA OSOBA

Osobe se mogu identificirati na različite načine, odnosno primjenom različitih metoda utemeljenih na upotrebi značajki koje su svojstvene određenoj osobi (primjerice, značajke otiska prsta, oka, lica itd.). Primjena i tip metode ovise o potrebama korisnika. Mnogi ljudi svakodnevno se susreću s različitim metodama identifikacije, odnosno sustavima koji su na njima utemeljeni, primjerice, tijekom prelaska državnih granica, u zračnim lukama diljem svijeta, pri ulasku u različite ustanove, tijekom interakcije s robotom itd.

Identitet predstavlja ukupnost nepromjenjivih osobina koja čine određenu osobu ili predmet, a prema kojima se ona/ono može razlikovati od svih drugih. Navedeni skup osobina predstavlja individualnost. Identifikacija osoba je utvrđivanje istovjetnosti nepoznatog s otprije poznatim, na temelju određenih identifikacijskih osobina, odnosno predstavlja postupak usporedbe određenog broja identifikacijskih osobina, pri čemu se ustanovljava podudarnost ili različitost između objekata koji se uspoređuju [80]. Identifikacijske osobine u kontekstu računalnog vida i istraživanja u ovoj disertaciji odnose se na značajke, izlučene iz RGB i dubinskih slika, određene osobe.

Identifikacija osobe podrazumijeva utvrđivanje identiteta osobe s obzirom da se osoba ne predstavlja nikakvim identitetom. S druge strane, postoji i verifikacija osobe gdje se obavlja provjera identiteta s kojim se osoba predstavlja. Okosnicu metoda koje se upotrebljavaju za identifikaciju ili verifikaciju osoba obično čine faktori koji se nazivaju faktori ovjere ili faktori autentičnosti. Navedeni faktori ovjere opisani su sljedećim izrazima:

- „nešto što osoba zna” (engl. *Something You Know*)
- „nešto što osoba posjeduje” (engl. *Something You Have*)
- „nešto što osoba jest” (engl. *Something You Are*).

„Nešto što osoba zna” odnosi se na neki podatak koji samo osoba zna i koji se najčešće upotrebljava u procesu verifikacije. Obično uključuje korištenje PIN-a (engl. Personal Identification Number), lozinke, neke ključne riječi itd. Jednostavnost metoda koje se temelje na upotrebi takvih elemenata leži u činjenici da ne zahtijevaju nikakav poseban uređaj te su stoga vrlo jednostavne za realizaciju. Sigurnosni problem se javlja jer velik broj osoba upotrebljava jednostavne lozinke i PIN-ove te je zlonamjernim

korisnicima vrlo lako doći do tih podataka.

„*Nešto što osoba posjeduje*” uključuje posjedovanje nekog određenog predmeta koji se upotrebljava u svrhu identifikacije ili verifikacije. Upotrebljavaju se magnetne kartice, pametne kartice, uređaji s univerzalnom serijskom sabirnicom (engl. Universal Serial Bus, USB), mobilni telefoni, tokeni itd. Danas se sve više upotrebljavaju i pametni satovi te pametne narukvice.

„*Nešto što osoba jest*” jednoznačno određuje svaku osobu. Tu spadaju otisak prsta, elementi oka (šarenica, mrežnica), dlan, geometrija ruke itd. Prema tome, metode koje su na tome utemeljene su biometrijske metode, a one su detaljnije opisane u drugom dijelu ovog poglavlja budući da su predloženi pristupi, odnosno metode, temeljeni na hodu osobe što spada u biometrijske osobine.

2.1 Biometrijske metode

Biometrija je definirana kao svaka mjerljiva tjelesna ili fiziološka osobina ili osobina ponašanja koja se može upotrebljavati za identifikaciju osobe [36]. Biometrijske metode općenito se mogu podijeliti u dvije skupine: fiziološke metode i metode biometrije temeljene na ponašanju. Prema tome, primjeri fizioloških biometrija jesu otisak prsta, geometrija ruke, lice osobe, oko (šarenica, mrežnica), obrasci vena ruke, dok su primjeri biometrije ponašanja glas, potpis, dinamika pritiska, tj. način tipkanja na tipkovnici te hod, tj. način hoda [36].

Nisu sve fiziološke osobine i osobine ponašanja pogodne za upotrebu u biometriji. Pogodnost se ocjenjuje na temelju više sljedećih kriterija (tablica 2.1) [36]:

- univerzalnost
- posebnost
- trajnost
- naplativost
- uspješnost
- prihvatljivost
- otpornost na zaoblazhenje.

Opisani kriteriji se često još nazivaju i sedam stupova biometrije (engl. Seven Pillars of Biometrics) [36]. Iako nema biometrijske metode koja jednako dobro ispunjava svih sedam kriterija, neke metode ispunjavaju uvjete bolje od drugih te su stoga po-

Tablica 2.1: Sedam stupova biometrije [36]

NAZIV STUPA BIOMETRIJE	OPIS
Univerzalnost	Sve osobe trebaju imati odgovarajuće osobine.
Posebnost	Osobine svake osobe trebale bi biti dovoljno različite kako bi se osobe međusobno razlikovale.
Trajnost	Osobine bi trebale ostati uglavnom nepromijenjene tijekom života osobe.
Naplativost	Naplativost se odnosi na osobinu koja bi trebala biti prezentirana i mjerena kvantitativno na relativno jednostavan način.
Uspješnost	Uspješnost se odnosi na razinu točnosti i brzinu identifikacije sustava s obzirom na okolišne i operativne faktore.
Prihvatljivost	Prihvatljivost predstavlja želju svake osobe da prihvati osobinu za upotrebu u svrhu identifikacije.
Otpornost na zaobilaženje	Otpornost na zaobilaženje se odnosi na stupanj težine zaobilaženja sustava.

godnije za primjenu u različitim situacijama. Primjerice, otisak prsta i šarenica oka dat će točnije rezultate tijekom identifikacije nego, recimo, potpis i dinamika tipkanja na tipkovnici te su stoga pogodniji za primjenu kada je potrebna veća i pouzdanija zaštita [36].

Sustav biometrijske identifikacije obično ima dvije faze: fazu registracije i fazu identifikacije. Faza registracije je faza u kojoj biometrijski sustav upotrebom određenog uređaja (primjerice, RGB ili RGB-D kamera) izrađuje digitalnu reprezentaciju biometrijskih osobina neke osobe. Sustav izlučuje skup značajki koji predstavlja istaknute biometrijske osobine osobe te pohranjuje taj skup značajki u bazu podataka zajedno s identifikatorom koji povezuje skup značajki s osobom. Navedeni skup značajki naziva se predložak. S druge strane, u fazi identifikacije sustav ponovno prikuplja biometrijske osobine osobe, izlučuje skup značajki te ga uspoređuje s predlošcima u bazi podataka [40]. Faza identifikacije podrazumijeva identificiranje osobe, pri čemu se uzima novi uzorak biometrijskih osobina, izlučuju se značajke te se uspoređuju s onima u bazi podataka, tj. s predlošcima. Tako se utvrđuje identitet osobe koja prolazi proces identifikacije [40].

Prema [36] biometrijski sustav ima četiri koraka:

- registraciju
- pohranu
- prikupljanje
- podudaranje.

Biometrijski sustav može biti, tj. može raditi u dvama načinima rada. Ta dva načina jesu [36] [40]:

- verifikacija ili ovjera
- identifikacija.

Verifikacija identiteta osobe uspoređuje upitni biometrijski uzorak (novi predložak) samo s predloškom koji odgovara identitetu kojim se osoba predstavlja, što se naziva podudaranjem *jedan-prema-jedan* (engl. One-to-One Match). Identifikacija osobe je postupak u kojemu se osoba ne predstavlja određenim identitetom, nego se upitni biometrijski uzorak uspoređuje sa svim predlošcima u bazi podataka, a naziva se podudaranje *jedan-prema-više* (engl. One-to-Many Match) [40]. Na slici 2.1 prikazane su neke biometrijske osobine koje se upotrebljavaju u procesu identifikacije.

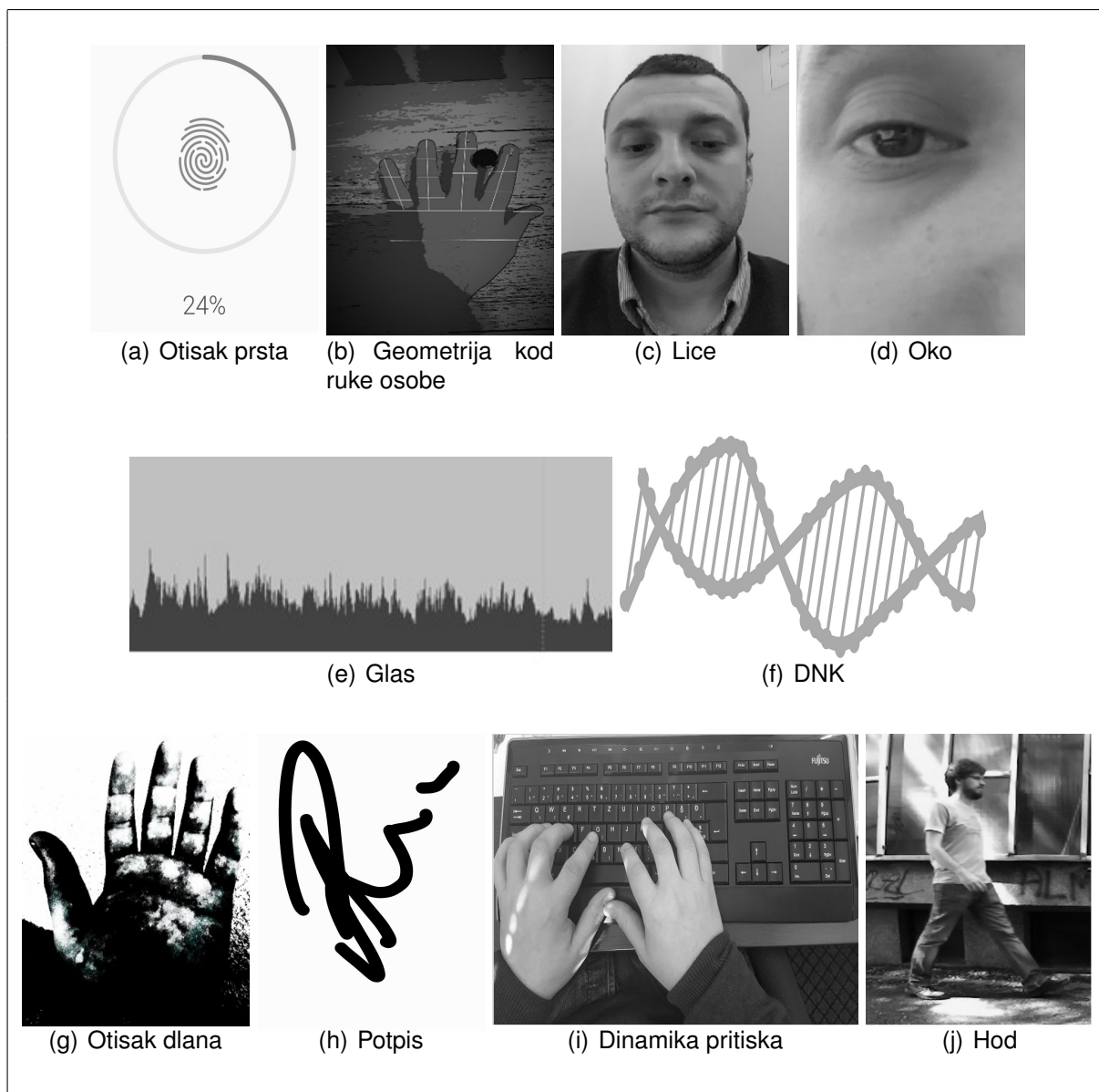
U [87] biometrijske metode klasificirane su s korisničkog i biometrijskog gledišta kao:

- statičke metode
- dinamičke metode.

Za neke biometrijske metode osoba mora prezentirati fiksnu, nepromjenjivu osobinu koja je uvijek prisutna, neovisna o bilo kojoj akciji osobe, tj. s korisničke strane nije potrebna posebna akcija koju je potrebno obaviti za njezino prezentiranje. Takve osobine nazivaju se statičke biometrijske osobine, a metoda biometrijske identifikacije statička metoda [87]. Neki od primjera navedenih metoda jesu metode koje upotrebljavaju [87]:

- otisak prsta
- lice
- mjere geometrije ruke
- oko (šarenica ili mrežnica oka)
- vene.

S druge strane, kako je navedeno u [87], biometrijske osobine koje su povezane s određenom akcijom osobe jesu dinamičke osobine, a biometrijska metoda naziva se



Slika 2.1: Biometrijske osobine koje se upotrebljavaju pri identifikaciji

dinamička metoda. Često se nazivaju i osobinama ponašanja. Primjeri navedenih metoda jesu metode koje upotrebljavaju [87]:

- govor
- dinamiku potpisa
- dinamiku pritiska na tipku.

Navedene metode upotrebljavaju se u različitim sustavima identifikacije i u različitim područjima ljudske djelatnosti, a odabir neke određene metode ovisi o potrebama korisnika. Najčešće metode koje se upotrebljavaju jesu metode temeljene na otisku prsta, licu te elementima oka.

Biometrijska identifikacija predstavlja statistički proces obuhvaćen varijacijama unutar klasa, a između registracije i kasnijeg prikupljanja. Primjerice, lozinka pri unosu može biti točna ili netočna, dok kod biometrijskih sustava ne postoji 100 %-tna mjerna točnost te su svi oni podložni različitim pogreškama (primjerice, neuspjeh upisivanja, neuspjeh prikupljanja, lažna pogreška prihvaćanja ili odbijanja itd.) [36]. Kada se govori o prethodno spomenutim neuspjesima i pogreškama, potrebno je navesti i sljedeće pojmove:

- stopa neuspjeha upisivanja (engl. Failure To Enroll Rate, FTE)
- stopa neuspjeha prikupljanja (engl. Failure To Acquire Rate, FTA)
- stopa lažnog prihvaćanja (engl. False Accept Rate, FAR)
- stopa lažnog odbijanja (engl. False Reject Rate, FRR).

FTE se odnosi na poteškoće koje može imati osoba pri upisivanju u sustav. Razlog tih poteškoća često je kontrola kvalitete sustava. FTA podrazumijeva poteškoće prikupljanja biometrijskih osobina osobe pri kasnijoj upotrebi sustava [36].

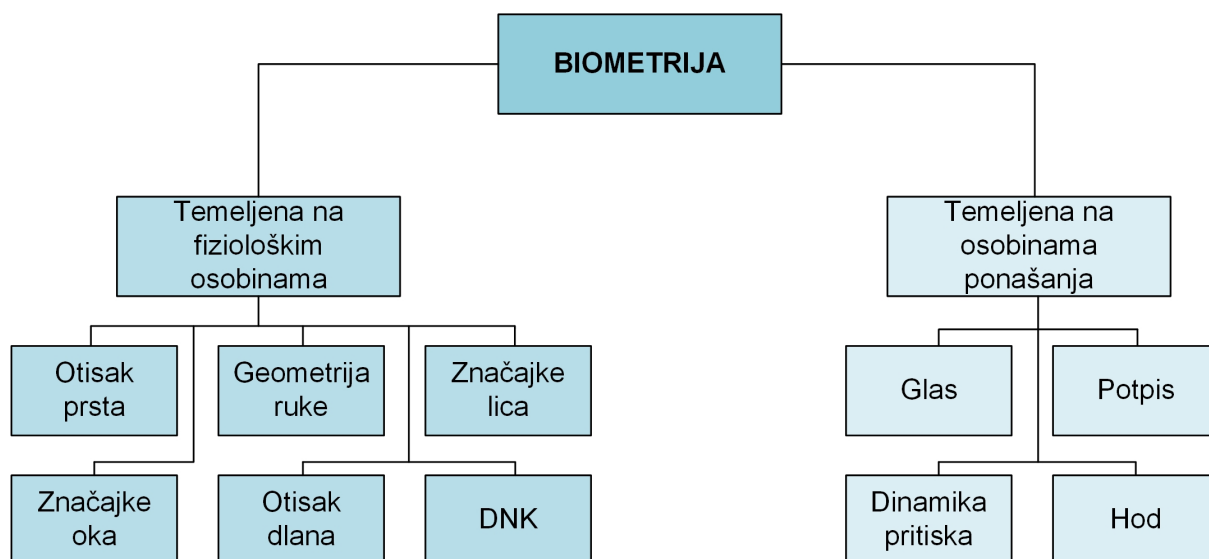
Prije nego što se definiraju FAR i FRR, potrebno je opisati što je to pogreška lažnog prihvaćanja, odnosno odbijanja. Pogreška lažnog prihvaćanja odnosi se na pogrešku kada se stečeni predložak osobe koji nije u sustavu pogreškom podudara s upisanim predloškom druge osobe i te pogreške mogu kompromitirati sigurnost i integritet samog sustava. S druge strane, do pogreške lažnog odbijanja dolazi kada se stečeni predložak osobe ne podudara s upisanim predloškom za tu osobu. U vezi s tim pogreškama postoje stope pogreške, FAR i FRR. Te dvije stope obrnuto su proporcionalne, što znači da smanjenjem jedne dolazi do povećanja druge i obrnuto. Na FAR i FRR direktno utječe odabrani prag sustava (engl. Threshold) (odnosi se na definirani prag sustava, primjerice, iznad kojeg je osoba prihvaćena), koji je i sam funkcija posebne aplikacije te treba postojati kompromis između tih stopa za svaku aplikaciju [36].

Kao što je već spomenuto, biometrijske metode općenito se mogu podijeliti u dvije skupine:

- biometrijske metode temeljene na fiziološkim osobinama
- biometrijske metode temeljene na osobinama ponašanja.

U vezi s navedenom podjelom biometrijskih metoda, na slici 2.2 prikazane su najčešće metode koje se primjenjuju u okviru te dvije skupine, a u nastavku su i opisane najčešće

metode koje se primjenjuju pri identifikaciji osoba iz navedenih skupina. U nastavku ovog potpoglavlja detaljnije su opisane metode koje se danas najčešće primjenjuju u svrhu identifikacije osoba, kao što su identifikacija temeljena na otisku prsta, identifikacija na temelju lica, identifikacija na temelju šarenice ili mrežnice oka, uz predmetnu metodu identifikacije na temelju hoda. Kraće su opisane i ostale važnije biometrijske metode.



Slika 2.2: Podjela najčešće primjenjivanih biometrijskih metoda

2.1.1 Identifikacija na temelju otiska prsta

Identifikacija temeljena na otisku prsta predstavlja jednu od najčešće upotrebljivanih biometrijskih metoda koja ima različita područja primjene. Otisci prsta predstavljaju jedinstvene obrasce koje čine grebeni trenja (engl. Friction Ridges) i brazde (engl. Furrows) koje se nalaze na prstima i palčevima [72]. Obrasci grebena trenja grupirani su u tri različita tipa [72]:

- petlje (engl. Loops)
- spirale (engl. Whorls)
- lukove (engl. Arches).

Petlje predstavljaju otiske koji su savijeni prema natrag te formiraju oblik petlje. Spirale formiraju kružne oblike poput sitnih virova. Postoje četiri skupine spirala: jednostavne koje su u osnovi koncentrični krugovi, središnja jamičasta petlja (petlja s virom na kraju), dvostruka petlja (dvije petlje koje kreiraju obrazac u obliku slova S) i slučajna

petlja koja je nepravilnog oblika. Lukovi u osnovi predstavljaju obrasce u obliku valova te uključuju obične i jelovite lukove [72].

Identifikacija temeljena na otisku prsta jedna je od najstarijih metoda za identifikaciju osoba te se može upotrebljavati za verifikaciju i identifikaciju osoba [87]. Sustavi otiska prsta uključuju AFIS (engl. Automated Fingerprint Identification System) i ne-AFIS sustave [36]. AFIS sustave zanimljivo je definirao Komarinski [46] u svojoj knjizi u kojoj je razložio definiciju po svakoj riječi koja čini naziv ovog sustava. Komarinski je to opisao na sljedeći način. Proces automatizacije (A) eliminira potrebu za ispisanim klasifikatorom za lociranje kartica otiska prsta iz datoteke i usporedbu dvije fizičke kartice. Baza podataka koja se može pretraživati sadržava slike otisaka prstiju (F) koji su uzeti od osobe s pomoću kartica otisaka prstiju ili elektroničkim putem, nekim od uređaja predviđenih za tu radnju (skener otiska prsta). Proces identifikacije (I) odvija se kada je od osobe uzet otisak prsta te se taj otisak prsta traži u bazi podataka. Definira se sustavom (S) jer upotrebljava računala i softvere te može ostvarivati interakciju s nekim podsustavima ili pak drugim sustavima.

AFIS sustavi obično se upotrebljavaju u svrhu provedbe državnih zakona i obično uključuju usporedbu otisaka prsta mogućeg osumnjičenika s onima u bazi podataka te se obično svih 10 otisaka prstiju osobe sakupljaju i uspoređuju s onima u bazi podataka. S druge strane, ne-AFIS sustavi potpuno su automatizirani i provode se u stvarnom vremenu te nije nužno upotrebljavati svih 10 prstiju [36].

Kako navode autori u [39], sustav identifikacije temeljen na otisku prsta sastoji se od triju osnovnih faza:

- prikupljanje
- izlučivanja značajki
- donošenja odluke.

Postoje dva pristupa koja se obično upotrebljavaju u slučaju podudaranja otiska prsta [87]:

- detaljno podudaranje (engl. Minutiae Matching)
- podudaranje uzoraka (engl. Pattern Matching).

Detaljno podudaranje analizira završetke i račvanja grebena, dok podudaranje uzoraka predstavlja više makroskopski pristup u kojemu se promatraju tokovi grebena u smislu petlji, spirala i lukova [87].

Sustavi identifikacije koji su temeljeni na otisku prsta svoju primjenu imaju u različitim područjima. Jedan od takvih sustava upotrebljava se i na sveprisutnim pametnim telefonima. Upotrebljava se za prijavu u uređaj, ali i za pristup različitim aplikacijama (primjerice, aplikacija povezana s bankarskim uslugama).

2.1.2 Identifikacija na temelju lica

Identifikacija na temelju lica je biometrijska metoda koja je danas često u upotrebi. Iako je načelno metoda u kojoj je potrebna interakcija s osobom, sustav temeljen na ovoj metodi može se realizirati i tako da osoba nije upoznata s time da prolazi proces identifikacije, tj. bez interakcije s osobom.

Kada se želi identificirati određena osoba, neophodno je uzeti sliku osobe na kojoj se jasno vidi lice, izlučiti značajke lica sa slike te ih usporediti sa značajkama koje su spremljene u bazi podataka. Drugim riječima potrebno je kreirati novi predložak koji sadržava značajke lica osobe i usporediti ga s predlošcima u bazi podataka.

Efikasnost sustava temeljenog na licu osobe uvelike ovisi o kvaliteti slike. Slika lica obično se izlučuje s veće slike koja na sebi ima brojne druge strukture i obrasce, moguće čak i više različitih osoba. Sustav će u najvećoj mogućoj mjeri normalizirati sliku kako bi značajke slike što bolje odgovarale značajkama slika u bazi podataka. Nakon što se slika normalizira, izlučuju se značajke i stvara se predložak koji se uspoređuje s onima u bazi podataka [34].

Ovisno o tome radi li se o procesu verifikacije ili identifikacije osobe, sustav nudi dva moguća izlaza, osoba je identificirana ili osoba nije identificirana. Ako osoba nije identificirana, za to može postojati nekoliko razloga: osoba se možda pokušava lažno predstaviti ili sustav griješi. Ako sustav griješi, radi se o lažnom odbijanju (engl. False Reject, FR). Sustav također može i lažno prihvatiti osobu, tj. pozitivno je identificirati, što se naziva lažno prihvaćanje (engl. False Accept, FA) [34].

Kada se uzme slika osobe, proces identifikacije na temelju lica obično ima četiri međusobno povezane faze [34]:

- detekciju lica
- normalizaciju
- izlučivanje značajki
- podudaranje (identifikacija na temelju lica).

U procesu identifikacije temeljenom na licu osobe obično se snimaju i uspoređuju prostorno-geometrijske značajke lica, što uključuje lokaciju i oblik atributa lica kao što su oči, obrve, nos, usne, brada te njihove prostorne veze [36].

Sustav koji upotrebljava značajke lica za identifikaciju i verifikaciju osoba počeo se prije nekoliko godina upotrebljavati i u području pametnih telefona. Sustav se najčešće upotrebljava za prijavu u uređaj, ali se upotrebljava i za neke druge namjene, kao što je verifikacija u svrhu plaćanja itd.

2.1.3 Identifikacija na temelju šarenice i mrežnice oka

Pri upotrebi oka za procese biometrijske identifikacije osoba, upotrebljavaju se dva elementa oka: šarenica i mrežnica. S tim u vezi postoje dvije metode:

- identifikacija temeljena na šarenici oka (engl. Iris Recognition)
- identifikacija temeljena na mrežnici oka (engl. Retina Scanning).

Šarenica (engl. Iris) predstavlja obojani dio oka koji je jedinstven po boji, strukturi i uzorku. Šarenica oka jedinstvena je za svaku osobu te je zbog toga pogodna za upotrebu u procesima identifikacije osoba. Mrežnica (engl. Retina) je višeslojno osjetilno tkivo stražnje očne jabučice na koje pada svjetlost [35], a radi se o dijelu oka koji prima svjetlosne podražaje te ih kemijskim procesom mijenja u električne impulse koji se vode prema mozgu.

U svojstvenom sustavu identifikacije koji upotrebljava šarenicu, slika oka prvo prolazi kroz proces inicijalne obrade da bi se dobila segmentirana i normalizirana slika šarenice. Zatim se njezine teksture analiziraju i kodiraju da bi se dobio predložak koji sadržava značajke šarenice [22]. Kako navodi Wildes [122], važna pitanja tijekom razvoja i implementacije sustava za identifikaciju temeljenu na šarenici oka mogu se podijeliti u tri dijela:

1. prikupljanje slika
2. lokalizaciju šarenice (izlučivanje značajki i kreiranje predloška)
3. podudaranje izlučenih značajki s onima u bazi podataka.

Kod sustava koji upotrebljava mrežnicu oka proces identifikacije ili verifikacije sličan je procesu kod šarenice oka i uključuje [131]:

1. prikupljanje i obradu slika
2. izlučivanje značajki i kreiranje predloška

3. podudaranje izlučenih značajki s onima u bazi podataka.

Prikupljanje slika u metodi koja upotrebljava mrežnicu oka radi se s pomoću uređaja za skeniranje mrežnice. Mrežnica predstavlja sloj kompleksnih krvnih žila i živčanih stanica na stražnjem dijelu oka te se stoga ova identifikacija temelji na analizi slika kompleksnih uzoraka krvnih žila mrežnice. Proces se odvija na način da se prvo dobije slika mrežnice s pomoću uređaja za skeniranje koji primjenjuje infracrvenu svjetlost niskog intenziteta kako bi osvijetlio mrežnicu te se daljnjom obradom analizira slika krvnih žila. Potom se izlučuju značajke i pohranjuju u predloške koji se poslije uspoređuju u procesu podudaranja [131].

Mrežnica i šarenica su jedinstvene kod svake osobe (čak i kod identičnih blizanaca) te se u normalnim zdravstvenim uvjetima ne mijenjaju do smrti. Obje metode mogu se upotrebljavati na gotovo svim ljudima, ali određene bolesti vezane za šarenicu i mrežnicu mogu u određenoj mjeri utjecati na pouzdanost metode. Identifikacija temeljena na šarenici upotrebljava kamere slične običnim kamerama, dok se za mrežnicu upotrebljavaju specijalizirane kamere koje se općenito upotrebljavaju u oftalmologiji i koje su dosta skuplje. Danas se metoda temeljena na šarenici oka općenito više primjenjuje u odnosu na metodu identifikacije u kojoj se upotrebljava mrežnica oka.

U slučaju metode koja upotrebljava mrežnicu oka teže je provesti lažno predstavljanje i zloupotrebljavati ovu metodu s obzirom na to da se radi o unutarnjem dijelu oka gdje se za skeniranje upotrebljavaju specijalizirani uređaji. Kad je riječ o metodi koja upotrebljava šarenicu oka, zabilježeni su neki slučajevi u kojima se upotrebljava slika oka visoke kvalitete, kontaktne leće gdje je ispisan uzorak šarenice itd. [131].

2.1.4 Identifikacija na temelju hoda

Mnoge biometrijske metode identifikacije zahtijevaju određenu interakciju s osobom. Da bi se osoba identificirala otiskom prsta potrebno je na uređaj postaviti prst kako bi se izlučile značajke. Slično vrijedi i u drugim metodama koje upotrebljavaju značajke s ruke osobe. Da bi se osoba identificirala na temelju šarenice ili mrežnice oka, također je potrebna interakcija s osobom jer osoba mora stati i postaviti oko pred uređaj kako bi se snimila slika u svrhu izlučivanja značajki. Slično vrijedi i u metodama koje upotrebljavaju osobine ponašanja kao što su glas, dinamika pritiska na tipkovnici i metoda u kojoj se upotrebljava potpis. Metoda koja zasigurno ne zahtijeva interakciju s oso-

bom jest metoda koja upotrebljava osobine hoda. Koristeći se hodom za identifikaciju osoba specifična osoba može prolaziti proces identifikacije, a da uopće nije svjesna toga. Merriam-Webster [103] hod (engl. Gait) definira kao način hodanja ili kretanja na nogama. Štoviše, Lee i Grimson [49] u svojem radu navode da je ljudski hod osobina osobe koju, među ostalim, određuju težina, dužina ekstremiteta, obuća i držanje u kombinaciji sa svojstvenim pokretima.

U dosadašnjim istraživanjima dominiraju dva pristupa kada se radi o identifikaciji na temelju hoda:

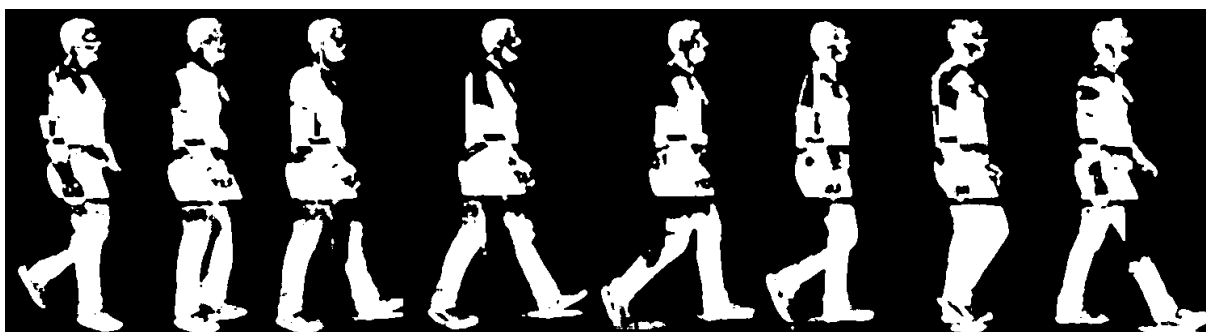
- pristup temeljen na modelu
- pristup temeljen na pojavi, pristup koji ne upotrebljava model.

U pristupu temeljenom na modelu obično se upotrebljavaju neke osobine tijela i dinamika kretanja osobe, tj. modeliranjem tih osobina dobivaju se značajke koje se upotrebljavaju u procesu identifikacije. Te značajke često uključuju dužinu ruku, dužinu nogu, dužinu koraka itd. Kad je riječ o ovom pristupu, često se u različitim radovima upotrebljavaju informacije o kosturu. Nedostatak ovog pristupa jest njegova kompleksnost, a s druge strane prednost mu je u tome što je manje podložan različitim promjenama u pojavi (primjerice, odjeća, okruženje i slično). S druge strane, u pristupu temeljenom na pojavi direktno se upotrebljavaju značajke koje se dobiju izlučivanjem siluete osobe, a prate se značajke hoda promatranjem obrasca hodanja na toj silueti osobe. Implementacija ovog pristupa relativno je jednostavnija, ali je podložnija različitim promjenama u pojavi.

Značajke vezane za hod mogu se izlučiti na različite načine. Danas se za tu namjenu najčešće upotrebljavaju RGB i RGB-D kamere. Kada se upotrebljavaju RGB kamere, to obično uključuje izlučivanje silueta osoba iz skupova podataka koji su kreirani tim kamerama. Kada se upotrebljavaju RGB-D kamere, siluete se izlučuju iz RGB ili dubinskih slika, a često se kombiniraju i značajke dobivene iz RGB i dubinskih slika u procesu identifikacije osoba na temelju hoda. Uz siluete osoba, korištenjem RGB-D kamera kreiraju se i različiti modeli koji se analiziraju i upotrebljavaju tijekom identifikacije osoba na temelju hoda. Posao je znatno lakši kada se upotrebljava uređaj kao što je Kinect pa se može vrlo jednostavno ocrtati kostur osobe te analizirati odvojeno značajke kao što su dužina ruke, dužina noge, dužina torza itd., koje se mogu upotrebljavati u pristupima temeljenima na modelu. Oba se pristupa danas aktivno upotrebljavaju i

razvijaju, implementirajući u nove pristupe nove značajke i načine upotrebe značajki.

Na slici 2.3 prikazan je primjer ciklusa hoda za jednu osobu na izlučenim siluetama osobe. Ciklus hoda odnosi se na slijed događaja tijekom kretanja između trenutka dok je noga na tlu pa do sljedećeg kontakta te iste noge s tlom. Upravo se ciklus hoda osobe obično analizira kod metoda identifikacije na temelju hoda. Različite metode iskorištavaju ciklus hoda na različite načine. Primjerice, jedna od najpoznatijih metoda u kontekstu identifikacije osoba na temelju hoda, GEI metoda, upotrebljava ciklus hoda tako da se on objedini u jednoj slici. Drugim riječima, GEI je slika prosječnih silueta osobe tijekom ciklusa hoda koji je prikazan na jednoj slici (primjer je prikazan na slici 1.2). Tijekom upotrebe silueta osoba siluete je obično potrebno normalizirati na jednaku veličinu te poravnati jer osoba obično ne hoda na jednakoj udaljenosti od kamere pa ni dobivene siluete nisu jednake veličine. Siluete se obično izlučuju primjenom različitih metoda, kao što je oduzimanje pozadine (engl. Background Subtraction), što se detaljno opisuje u poglavlju 5.



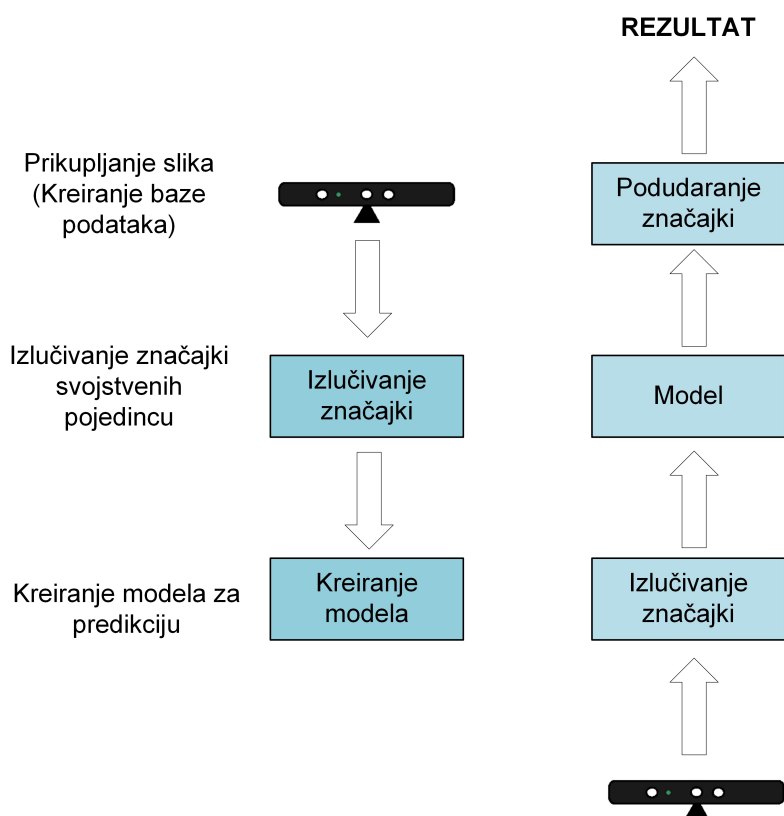
Slika 2.3: Prikaz ciklusa hoda osobe

Sustav identifikacije osoba na temelju hoda može se realizirati u nekoliko koraka (slično kao i u prethodnim metodama):

- prikupljanje slika (kreiranje skupa podataka)
- izlučivanje značajki
- podudaranje značajki.

Detaljniji nacrt sustava može se prikazati kao na slici 2.4.

Kao što se može vidjeti na temelju slike 2.4, sustav se može realizirati tako da se realiziraju dvije faze. Prva faza jest faza registracije u sustav. Drugim riječima, dolazi do prikupljanja slika, odnosno kreiranja baze podataka i izlučivanja svojstvenih značajki vezanih za svaku osobu. U drugoj fazi, fazi identifikacije, uzimaju se biome-



Slika 2.4: Sustav identifikacije na temelju hoda

triske osobine osobe koju se želi identificirati, tj. izlučuju se značajke te se uspoređuju sa značajkama osoba u bazi podataka. Ovisno o tom procesu dobiva se informacija o tome je li osoba identificirana.

2.1.5 Ostale važnije biometrijske metode

Uz prethodno opisane metode u ovom potpoglavlju su ukratko opisane i sljedeće metode koje se također primjenjuju u biometriji:

- identifikacija na temelju deoksiribonukleinske kiseline (DNK)
- identifikacija na temelju dlana
- identifikacija na temelju geometrije ruke
- identifikacija na temelju glasa
- identifikacija prema načinu tipkanja na tipkovnici
- identifikacija na temelju potpisa osobe
- identifikacija na temelju vena osobe
- identifikacija na temelju uha
- identifikacija na temelju termografije lica

- identifikacija na temelju mirisa.

Jedna od metoda koja se u posljednje vrijeme često primjenjuje pri identifikaciji osoba (primjerice, u sigurnosnim sustavima) jest metoda temeljena na deoksiribonukleinskoj kiselini, DNK (engl. Deoxyribonucleic Acid, DNA). DNK je nukleinska kiselina koja je formirana kao dvostruka spiralna zavojnica. Iako se DNK metoda uvelike primjenjuje u različitim područjima ljudske djelatnosti, svoju najveću primjenu postigla je u provedbi zakona (primjerice, gdje se osoba najčešće identificira u svrhu dokazivanja nevinosti ili krivice u vezi s počinjenjem kaznenog djela). DNK je jedinstvena značajka svake osobe, što je čini pogodnom za upotrebu u biometriji. Kako bi se prikupio DNK svake osobe, sustavu treba više vremena jer, prije svega, treba prikupiti fizičke uzorke DNK-a kao što su krv, slina, kosa itd. Prednost upotrebe ove metode jest u činjenici da ne postoje dvije osobe s identičnim DNK profilom (osim identičnih blizanaca). U nedostatke ove metode ubrajaju se činjenica da se DNK podudaranje ne obavlja u stvarnom vremenu, da se radi o metodi koja obično zahtijeva kontakt s osobom jer se uzorak mora fizički uzeti (u nekim drugim metodama upotrebljavaju se samo, primjerice, slika ili videosekvencija) te da se radi o vrlo sporoj metodi [48]. Danas se DNK metoda najčešće upotrebljava u provedbi zakonodavstva, utvrđivanju očinstva, identifikaciji nestalih ili umrlih osoba itd.

Dlanovi ruku upotrebljavaju se u različitim sustavima za identifikaciju, a imaju obrasce grebena i doline, slično kao i kod otiska prsta. Za ovu namjenu, kao i u procesu identifikacije na temelju otiska prsta, upotrebljavaju se različiti uređaji zasnovani na upotrebi optičkog, kapacitivnog, ultrazvučnog te termalnog osjetila. Ovisno o tim upotrebljavanim uređajima, tj. o razlučivosti uhvaćene slike, slike mogu sadržavati brojne elemente dlana, uključujući već navedene grebene i doline, linije, nabore, pa i geometriju ruke. Slično kao i u procesu identifikacije na temelju otiska prsta, izlučuju se detalji ili obrasci koji se upotrebljavaju za stvaranje predložaka, s time da predložak, ovisno o sustavu, može biti izrađen od cijelog dlana ili njegova jednog dijela [36]. Kako je navedeno u [87], s dlana se mogu iskoristiti različite informacije: geometrijske značajke kao što su širina, visina itd., značajke glavne linije, značajke nabora, značajke točaka delte, značajke detalja, grebena dlana.

Geometrija ruke, kako navode autori Ross i Jain [84], odnosi se na geometrijsku strukturu ruke koja uključuje dužinu i širinu prsta na različitim lokacijama, širinu i deb-

ljinu dlana itd. Metoda koja se koristi geometrijom ruke obično uključuje jednostavnu obradu rezultirajućih značajki. Ova metoda, za razliku od metode temeljene na otisku prsta, obično ne uključuje izlučivanje detaljnih značajki ruke. Sustavi koji upotrebljavaju geometriju ruke nisu novost te su dostupni već od 1970-ih godina [84].

Identifikacija na temelju glasa predstavlja još jednu biometrijsku metodu u kojoj se upotrebljava glas osobe za identifikaciju ili verifikaciju. Postoje dva tipa sustava koja se upotrebljavaju u ovom kontekstu:

- sustav ovisan o tekstu
- sustav neovisan o tekstu.

U sustavu koji je ovisan o tekstu izgovara se točno predefinirana fraza te se ta fraza obično mora ponoviti nekoliko puta pa zato ovaj proces traje nešto duže. U slučaju sustava koji je neovisan o tekstu fraza nije obavezna [36] [87].

Metoda temeljena na načinu tipkanja predstavlja zanimljiv pristup identifikaciji osoba, a radi se o metodi koja se i danas primjenjuje. Identifikacija prema načinu tipkanja na tipkovnici predstavlja proces analiziranja načina na koji osoba tipka na tipkovnici promatranjem unosa s tipkovnice tisuću puta u sekundi kako bi se osoba identificirala prema ritmu naviknutih uzoraka tipkanja [65].

Način na koji osoba piše još je jedan od načina identifikacije osoba. Svaka osoba ima svojstven način pisanja, što se može primijeniti i na proces identifikacije. U osnovi, postoje dva pristupa identifikacije potpisom [87] [117]:

- dinamički potpis
- statički potpis.

Pri identifikaciji dinamičkim potpisom osoba svoj potpis unosi na digitalizirani uređaj koji u stvarnom vremenu analizira rukopis dok osoba piše. Ovaj proces još se naziva i *online* identifikacija. Kod identifikacije statičkim potpisom, osoba se potpisuje na papir, zatim se potpis digitalizira skenerom ili kamerom i pohranjuje u obliku slike, a biometrijski sustav identificira osobu analiziranjem oblika potpisa na slici. Statički potpis još se naziva i *offline* identifikacija [117].

Obrasci vena predstavljaju mrežu krvnih žila osobe te se razlikuju od osobe do osobe, što ih čini pogodnima za primjenu u sustavima identifikacije osoba. Ideja je nastala još 1992. godine, ali se u zadnje vrijeme istraživači ponovno više fokusiraju na taj način identifikacije. Navedeni obrasci vena mogu se dobiti upotrebom infracrvenih

uređaja, i to na dva načina: kontaktom i bez kontakta. U potonjem načinu nema potrebe da osoba dodiruje uređaj, dok u kontaktnom načinu osoba mora ostvariti kontakt s uređajem. Najčešće se primjenjuju vene u ruci osobe, ali se sustav može realizirati i korištenjem obrazaca vena iz nekog drugog dijela tijela [123].

Identifikacija na temelju uha zasnovana je na analizi vanjskog oblika uha, ušne školjke i strukture kostiju. Identifikacija na temelju termografije lica mjeri količinu termalnog zračenja koju emitira lice osobe, dok je identifikacija na temelju mirisa zasnovana na značajkama mirisa koji manifestira osoba [36].

2.1.6 Usporedba različitih biometrijskih metoda

Svaka od prethodno navedenih metoda ima određene prednosti i nedostatke te se mogu primjenjivati u različitim kontekstima u ovisnosti o potrebama korisnika. Jain i sur. [41] u svojem su radu usporedili biometrijske metode, odnosno biometrijske identifikatore na kojima su zasnovane biometrijske metode. Usporedba se temelji na sedam biometrijskih kriterija, tj. na sedam stupova biometrije (tablica 2.1) te je prikazana u tablici 2.2. Također, u istraživanju [41] navode da se primjena biometrije i biometrijskih metoda može podijeliti u tri glavne skupine [41]:

- komercijalne primjene (prijava u računalnu mrežu, sigurnost podataka itd.)
- primjene u državnim službama (vozačka dozvola, prijelazi državnih granica, popravne ustanove itd.)
- primjene u forenzici (identifikacije tijela, kriminalne istrage itd.).

Ako se analizira tablica 2.2, vidljivo je da među pojedinim metodama ima dosta različitosti pogledaju li se biometrijski kriteriji. Metode u kojima se upotrebljavaju otisak prsta i dlana prema biometrijskim kriterijima imaju jednake rezultate, a opet se danas ne primjenjuju u istoj mjeri te je metoda u kojoj se upotrebljava otisak prsta dominantnije korištena biometrijska metoda. Općenito gledajući, otisak prsta vrlo je pouzdana metoda identifikacije i danas spada u najčešće upotrebljavane metode identifikacije.

U slučaju primjene elemenata oka, metoda u kojoj se upotrebljava šarenica oka ima bolje općenite rezultate po biometrijskim kriterijima u odnosu na metodu u kojoj se upotrebljava mrežnica oka. Zbog boljih rezultata te zbog veće dostupnosti uređaja koji se upotrebljava uz metodu identifikacije temeljenu na šarenici, ova je metoda općenito isplativija za implementaciju. Primjerice, metoda u kojoj se upotrebljava šarenica oka

Tablica 2.2: Usporedba različitih biometrijskih metoda [41]

Biom. identifikator	Univ.	Pos.	Traj.	Naplat.	Perf.	Prih.	Zaob.
Otisak prsta	Srednja	Visoka	Visoka	Srednja	Visoka	Srednja	Srednja
Lice	Visoka	Niska	Srednja	Visoka	Niska	Visoka	Visoka
Šarenica oka	Visoka	Visoka	Visoka	Srednja	Visoka	Niska	Niska
Mrežnica oka	Visoka	Visoka	Srednja	Niska	Visoka	Niska	Niska
Hod	Srednja	Niska	Niska	Visoka	Niska	Visoka	Srednja
Ostale metode							
DNK	Visoka	Visoka	Visoka	Niska	Visoka	Niska	Niska
Dlan	Srednja	Visoka	Visoka	Srednja	Visoka	Srednja	Srednja
Geometrija ruke	Srednja	Srednja	Srednja	Visoka	Srednja	Srednja	Srednja
Glas	Srednja	Niska	Niska	Srednja	Niska	Visoka	Visoka
Dinamika tipkanja	Niska	Niska	Niska	Srednja	Niska	Srednja	Srednja
Potpis osobe	Niska	Niska	Niska	Visoka	Niska	Visoka	Visoka
Vene ruke	Srednja	Srednja	Srednja	Srednja	Srednja	Srednja	Niska
Uho	Srednja	Srednja	Visoka	Srednja	Srednja	Visoka	Srednja
Termografija lica	Visoka	Visoka	Niska	Visoka	Srednja	Visoka	Niska
Miris	Visoka	Visoka	Visoka	Niska	Niska	Srednja	Niska

implementirana je i na nekim pametnim telefonima, dok metoda u kojoj se upotrebljava mrežnica oka zahtijeva specijalizirane uređaje koji se upotrebljavaju u oftalmologiji. Metode identifikacije temeljene na značajkama oka, neovisno o tome radi li se o šarenici ili mrežnici, predstavljaju i jedne od najsigurnijih i najpouzdanijih biometrijskih metoda.

Metoda u kojoj se upotrebljava lice osobe, tj. značajke lica, postiže dobre rezultate te se danas sve više upotrebljava u različitim područjima, a u budućnosti će se vjerojatno upotrebljavati još više jer se zavaljujući širokoj dostupnosti kamera visoke razlučivosti ova metoda može realizirati i bez potrebe za interakcijom s osobom. Drugim riječima, osoba se može identificirati i na većoj udaljenosti, čak i ako ta osoba nije ni svjesna da prolazi proces identifikacije. Važna je i činjenica da se podatci koji se upotrebljavaju u ovoj metodi (prvenstveno se misli na slike s kojih se izlučuju značajke) mogu dobiti upotrebom danas u velikoj mjeri dostupnih RGB i RGB-D kamera, što uvelike doprinosi daljnjem razvoju ove metode.

Prema biometrijskim kriterijima, DNK metoda spada u pouzdanu i sigurnu metodu, no nedostatak joj je to što zahtijeva visok stupanj interakcije s osobom te što je proces identifikacije dosta spor. Zbog toga se DNK metoda danas najčešće primjenjuje u provedbi zakonodavstva. Biometrijske metode temeljene na ponašanju, kao što su dinamika tipkanja, metoda temeljena na glasu te metoda temeljena na potpisu osobe, po biometrijskim kriterijima postižu nešto lošije rezultate u odnosu na biometrijske metode

temeljene na fiziološkim osobinama. Međutim, navedene metode danas se također primjenjuju i aktivno istražuju. U slučaju ostalih metoda, i one se danas upotrebljavaju u različitim djelatnostima, a implementacija i primjena neke od njih ovisi o potrebama krajnjih korisnika.

Metoda identifikacije koja se temelji na hodu prema navedenim kriterijima spada u središnji rang metoda (po pouzdanosti), no glavna je prednost ove metode to što ne zahtijeva nikakvu interakciju s osobom, što je često ključno. Naime, osoba može prolaziti proces identifikacije, a da toga nije ni svjesna. Značajke neophodne za ovu metodu mogu se dobiti analizom podataka RGB ili RGB-D kamera koje su danas široko dostupne na tržištu. Metoda identifikacije osoba na temelju hoda danas se sve više istražuje te su zadnjih godina razvijeni i predstavljeni mnogi različiti pristupi.

3 RGB I RGB-D KAMERE

U ovom poglavlju opisane su RGB i RGB-D kamere. Opisani su najpoznatiji modeli boja kao i najpoznatija osjetila slike koja se upotrebljavaju u RGB kamerama. Vezano za RGB-D kamere opisane su aktivne i pasivne tehnike za dobivanje dubinskih slika. Opisani su i stereovid, *time-of-flight* kamere kao i kamere sa strukturiranom svjetlošću.

3.1 RGB kamere

RGB kamere danas imaju široku primjenu u različitim područjima, pa se tako upotrebljavaju i u istraživanjima čija je tema identifikacija osoba. Velik broj istraživanja te mnogi predloženi pristupi koji se tiču identifikacije osoba na temelju hoda ili, primjerice, identifikacije osoba na temelju lica upotrebljavaju RGB kamere.

RGB kamera je uređaj koji omogućava stjecanje slika u boji (RGB slike), a upotrebljava RGB model boja. Model boja predstavlja strukturirani sustav za stvaranje cijelog niza boja od nekoliko primarnih boja [82]. RGB model boja predstavlja model koji upotrebljava crvenu, zelenu i plavu boju kako bi se kreirao niz različitih boja. Naziv *RGB* je izveden od prvih slova engleskih naziva za spomenute boje (engl. **R**ed, **G**reen, **B**lue). Model RGB jest aditivni model boja u kojem se kombiniranjem crvene, zelene i plave primarne svjetlosti dobiva bijela svjetlost, a ima različite primjene kao što su digitalni zasloni (TV, računalni zasloni itd.), digitalne kamere itd.

Najpoznatiji modeli boja su RGB i CMYK (engl. **C**yan, **M**agenta, **Y**ellow, **K**ey). Po krata CMYK također je sastavljena od prvih slova engleskih naziva za cijan, magenta, žutu te crnu boju. Model boja CMYK je dizajniran da bi se podržalo tiskanje boja na bijeli papir. CMYK model je supstraktivni model u kojem se tinte u boji oduzimaju od originalne bijele površine [82].

Ključni dio svake kamere jest osjetilo slike. Osjetilo slike predstavlja električni sklop koji sadržava milijune diskretnih fotodetekcijskih mjesta koja se nazivaju pikseli i koji svjetlost koja pada na njihovu površinu konvertiraju u niz električnih signala [19]. Da bi se omogućilo slikanje u boji niz filtara boje (engl. Color-Filter Array, CFA) obično se postavlja u odgovarajućem obrascu na niz piksela osjetila slike; primjenom takva filtra svaki piksel producira signal koji odgovara jednoj od triju boja (crvenoj, zelenoj ili plavoj boji) [19].

Jednako tako, analogni podatci piksela očitavaju se s osjetila slike te se digitaliziraju s pomoću analogno-digitalnog pretvarača. Da bi se proizvela slika u punoj boji (slika s crvenim, zelenim i plavim vrijednostima za svaki piksel), radi se prostorna interpolacija (ta se operacija još naziva *demosaicking*). Signal se digitalno obrađuje da bi se izjednačila bijela boja te da bi se korigirala boja i umanjili efekti neispravnih piksela i nesavršenosti optike. Slika se dalje komprimira i pohranjuje u memoriju [19].

U današnje vrijeme dominiraju dvije kategorije osjetila slike: nabojski vezani elementi (engl. Charge-Coupled Device, CCD) i komplementarni metal-oksidni poluvodič (engl. Complementary Metal Oxide Semiconductor, CMOS). Obje kategorije osjetila konvertiraju svjetlost u električni naboj i procesuiraju ga u električni signal. Radi se o pikseliziranim metal-oksidnim poluvodičima koji akumuliraju naboj signala u svakom pikselu proporcionalno intenzitetu lokalnog osvjetljenja. CCD prenosi naboj paketa svakog piksela uzastopno u zajedničku izlaznu strukturu koja naboj pretvara u napon. Kod CMOS osjetila slike konvertiranje naboja u napon odvija se u svakom pikselu. Ta razlika u tehnikama očitavanja ima značajne implikacije za arhitekturu osjetila, mogućnosti i ograničenja [53].

U svojoj najjednostavnijoj implementaciji CCD struktura sastoji se od serije usko razmaknutih elektroda odvojenih od temeljne podloge poluvodiča tankim izolacijskim oksidnim slojem. Kada se na elektrodu primijeni napon, formira se područje iscrpljivanja u poluvodiču neposredno ispod njega. Područje iscrpljivanja je potencijalna jama koja može pohraniti paket električnog naboja. Pulsiranjem elektroda u prikladnom slijedu potencijalna se jama, a time i njezin paket naboja, može transferirati kroz poluvodič. Registar pomaka može se formirati dodavanjem sklopova za umetanje i detektiranje paketa naboja. Iako je izvorno zamišljen kao uređaj za pohranu digitalnih podataka, ima li se u vidu da potencijalna jama može pohraniti promjenjive količine naboja, on također može prenijeti i analogne signale te se CCD koncept primjenjuje i na funkcije analogne obrade signala, uključujući jednostavne linije kašnjenja i poprečne filtre. Međutim, CCD je najveću primjenu našao u osjetilima slike [118].

CMOS osjetilo slike upotrebljava istu tehnologiju silicijskog čipa koja se upotrebljava u mikroprocesorskim sustavima. Kod CMOS tehnologije milijuni tranzistora mogu se integrirati na jednom silicijskom sklopu, što otvara mogućnost integriranja velikog niza piksela, svakog s vlastitom fotodiodom i tranzistorima za očitavanje, zajedno sa

svom pomoćnom elektronikom potrebnom za adresiranje niza, baferiranje analognog videosignala, pa čak i digitalizacija spremna za obradu, pohranjivanje ili prikaz. CMOS osjetila rade na znatno nižem naponu u odnosu na CCD (to je obično u vrijednostima između 1,8 V i 5 V). Velika funkcionalna integracija i rad na nižim naponima otvaraju mogućnost široke primjene CMOS osjetila gdje se zahtijeva kompaktnost, duži vijek trajanja baterije i mali troškovi proizvodnje. S tim u vezi CMOS osjetila slike danas se u velikoj mjeri upotrebljavaju u digitalnim kamerama i mobilnim telefonima. CMOS osjetila slike danas su najčešće CMOS osjetila s aktivnim pikselima (engl. Active Pixel Sensor, APS) [118].

Do početka 1990-ih godina osjetila s pasivnim pikselima (engl. Passive Pixel Sensor, PPS) bila su primarna tehnologija CMOS osjetila. Budući da su PPS osjetila imala dosta slabija obilježja uspješnosti u odnosu na CCD osjetila, aktivno se radilo na razvoju modernih CMOS osjetila s aktivnim pikselima. Zaključak je bio da se dodavanjem pojačala na svaki piksel značajno poboljšava osjetilo i da se eliminiraju nedostaci PPS osjetila [19]. Osjetilo s aktivnim pikselima (APS) se definira kao osjetilo s jednim aktivnim tranzistorom ili više aktivnih tranzistora lociranih u svakom pikselu [23]. APS sadržava aktivno pojačalo [24]. APS tehnologija čuva poželjne attribute CCD osjetila kao što su visoka osjetljivost, visoka vjernost signala i veliki formati nizova. Više o navedenim tehnologijama može se pročitati u radovima [23] i [24].

Kad se govori o dvama opisanim osjetilima slike, CCD osjetila se obično upotrebljavaju u kamerama čiji je cilj visoka kvaliteta slika s manje šuma. Usto, iznimno su osjetljiva na svjetlost. CMOS osjetila slike obično imaju slabiju kvalitetu slika s manjom razlučivosti i slabije su osjetljiva na svjetlost. Međutim, CMOS osjetila troše manje energije u odnosu na CCD osjetila te su jeftinija. CMOS osjetila danas već po obilježjima uspješnosti u nekim izvedbama mogu parirati CCD osjetilima.

3.2 RGB-D kamere

Posljednjih godina pojavio se velik broj tzv. 3D kamera koje, uz uobičajenu 2D kameru, u istom kućištu imaju i dubinsku kameru. Jedan od najpoznatijih takvih uređaja jest Microsoft Kinect koji se, osim primarne namjene za korištenje s igračom konzolom, upotrebljavao u velikom broju istraživanja, primjerice, u području uslužne robotike ili istraživanjima vezanim za identifikaciju osoba. Ovisno o tehnici koja se upotrebljava

za dobivanje dubinskih slika od scene, uređaji koji pružaju RGB i dubinske slike u stvarnom vremenu mogu biti podijeljeni na aktivne i pasivne. Kod aktivnih uređaja upotrebljava se aktivan izvor energije (primjerice, projektor) da bi se emitirala laserska zraka ili pak određeni uzorak svjetlosti i prijamnik koji hvata svjetlost reflektiranu od objekta iz scene.

Primjeri aktivnih uređaja su kamere zasnovane na strukturiranoj svjetlosti i *time-of-flight* (TOF) kamere [88]. U slučaju uređaja koji su zasnovani na strukturiranoj svjetlosti svjetlosni se obrazac projicira na scenu te se mjeri deformacija obrasca kako bi se rekonstruirala scena. *Time-of-flight* uređaji emitiraju svjetlosni signal i mjere vrijeme potrebno da navedeni svjetlosni signal prođe put od uređaja do objekta i natrag. Primjeri aktivnih uređaja prikazani su na slici 3.1 (a),(b),(c) i (d).

Prednost uređaja zasnovanih na strukturiranoj svjetlosti i *time-of-flight* tehnikama jest u tome što oni istovremeno hvataju cijelu scenu, što ih čini pogodnima za upotrebu u procesu identifikacije osoba na temelju hoda, a i u drugim primjenama s dinamičnom scenom. Nedostatak tih uređaja jest ograničen domet zbog snage projektor (iz sigurnosnih razloga) koji u uređajima niže cijene obično iznosi oko 3 m do 5 m. Problematična je i upotreba tih uređaja u vanjskim okruženjima za vrijeme jake sunčeve svjetlosti jer svjetlost nadvladava projicirane obrasce, što ih često čini neuporabljivima u ovim situacijama. Jednako tako, identični uređaji često upotrebljavaju istu frekvenciju i obrasce te se njihova područja djelovanja ne bi trebala preklapati kako bi se izbjegla interferencija. Kada se radi o dubinskim slikama, one često sadržavaju šum, artefakte te nedostajuće piksele.

S druge strane, u slučaju pasivnih uređaja, ne upotrebljava se nikakav aktivni projektor, nego se uređaj, da bi se dobila slika scene, oslanja na svjetlost koja je prirodno prisutna u sceni. Svjetlost može biti sunčeva ili iz nekog umjetnog izvora. U ovom slučaju izračun dubinskih slika nije temeljen na strukturi izvora svjetlosti ni nekim vremenskim mjerenjima, nego se dubinske slike obično dobivaju analizom slika dvije ili više RGB ili infracrvenih kamera upotrebom tehnike obrade zasnovane na stereovidu (potpoglavlje 3.2.1). Ovisno o razlučivosti, osnovnoj udaljenosti između kamera i svjetlosnih uvjeta u sceni, stereovid nije intrinzično ograničen u dometu za razliku od aktivnih uređaja te može dobiti dubinske slike na većim udaljenostima. Jednako tako, bez problema se mogu upotrebljavati u vanjskim okruženjima. Nedostaci pasivnih uređaja

uključuju računalnu složenost izračuna koji se koriste u tehnikama zasnovanim na stereovidu i prisutnost neželjenih artefakata u dubinskim slikama, veća osjetljivost na refleksiju, sjene i druge značajke u sceni. Koraci tijekom obrade neobrađenih dubinskih slika slični su koracima kod aktivnih uređaja te ih je potrebno prikladno podesiti pri instalaciji i posebno za određene primjene. Neki od primjera pasivnih uređaja prikazani su na slici 3.1 (e),(f).

U nastavku ovog potpoglavlja detaljnije su opisani stereovid, *time-of-flight* kamere i kamere sa strukturiranom svjetlošću. Kod kamera sa strukturiranom svjetlošću detaljnije je opisan i Kinect uređaj, verzije jedan, jer se radi o uređaju koji se danas često upotrebljava u području kojem pripadaju ova disertacija i istraživanje.

3.2.1 Stereovid

Cilj stereovida jest izračun dubinske slike scene analizom značajki dviju dvodimenzionalnih slika iste scene dobivenih iz različitih točaka u prostoru. Može se postići upotrebom jedne kamere ili više kamera. Na slici 3.1 (e),(f) prikazani su primjeri kamera zasnovanih na stereovidu, a na slici 3.2 primjer koncepta stereovida. Na slici je prikazana ciljana točka T koja se vidi na strani optičkih centara O_L i O_D dviju kamera. Točke T_L i T_D predstavljaju odgovarajuće intenzitete piksela u lijevoj i desnoj ravnini slike za istu scenu točke T . U sljedećem dijelu disertacije opisani su najvažniji pojmovi povezani sa stereovidom.

Epipolarna geometrija (engl. Epipolar Geometry) je unutarnja projekcijska geometrija između dvaju pogleda. Neovisna je o strukturi scene, a ovisna o unutarnjim parametrima kamere i relativnog položaja [31]. Osnovica (engl. Baseline) predstavlja liniju između centara projekcije. Epipolarna ravnina (engl. Epipolar Plane) je ravnina koja sadržava osnovicu, dok je epipolarna linija (engl. Epipolar Line) sjecište epipolarne ravnine s ravninom slike. Epipol (engl. Epipol) je točka presjeka linije koja spaja centre kamere, tj. osnovicu s ravninom slike. Drugim riječima, epipol je slika u jednom pogledu centra kamere drugog pogleda [31]. Slika 3.3 prikazuje epipolarnu geometriju.

Fundamentalna matrica (engl. Fundamental Matrix, F) predstavlja algebarsku reprezentaciju epipolarne geometrije [31]. Navedena fundamentalna matrica F je 3×3 matrica koja povezuje odgovarajuće točke u stereoslikama, dok se esencijalna matrica (engl. Essential Matrix, E) odnosi na matricu koja povezuje kalibrirane odgo-



(a) Kinect verzije 1



(b) Kinect verzije 2 [104]



(c) Asus Xtion [100]



(d) Basler TOF kamera [101]

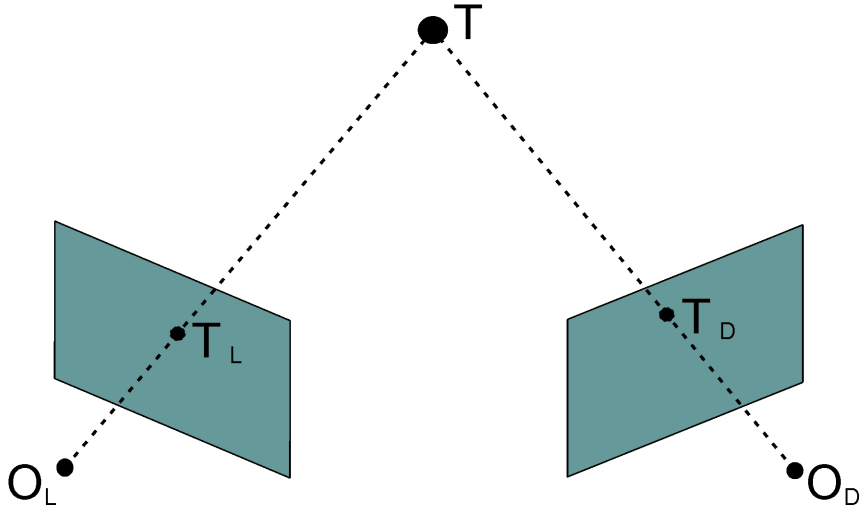


(e) Ricoh stereokamera [107]

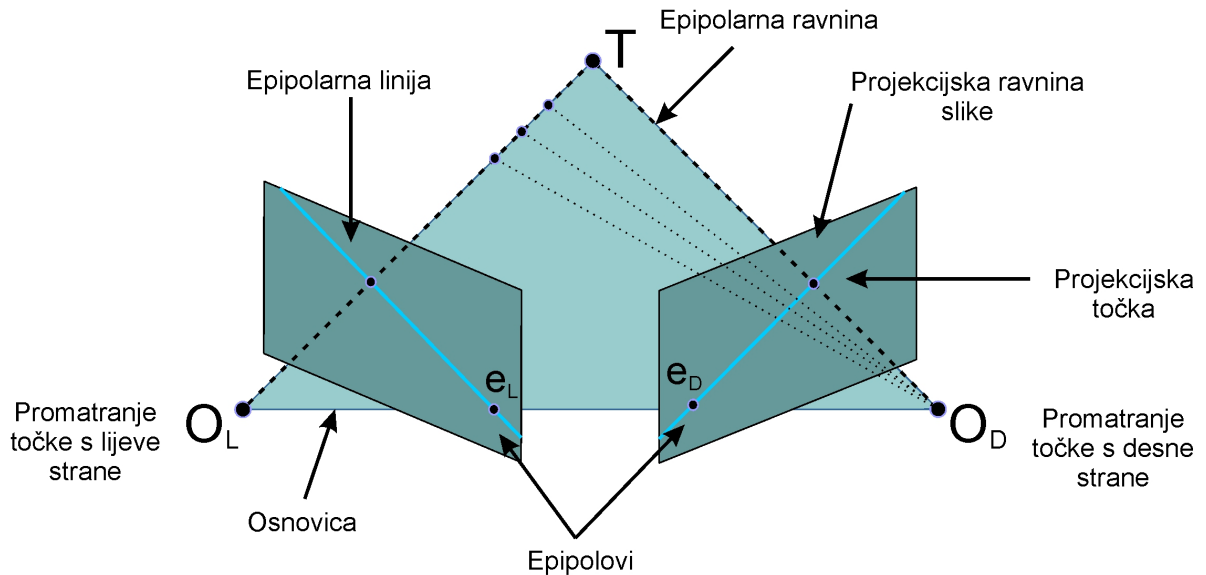


(f) ZED stereokamera [112]

Slika 3.1: Primjeri RGB-D kamera



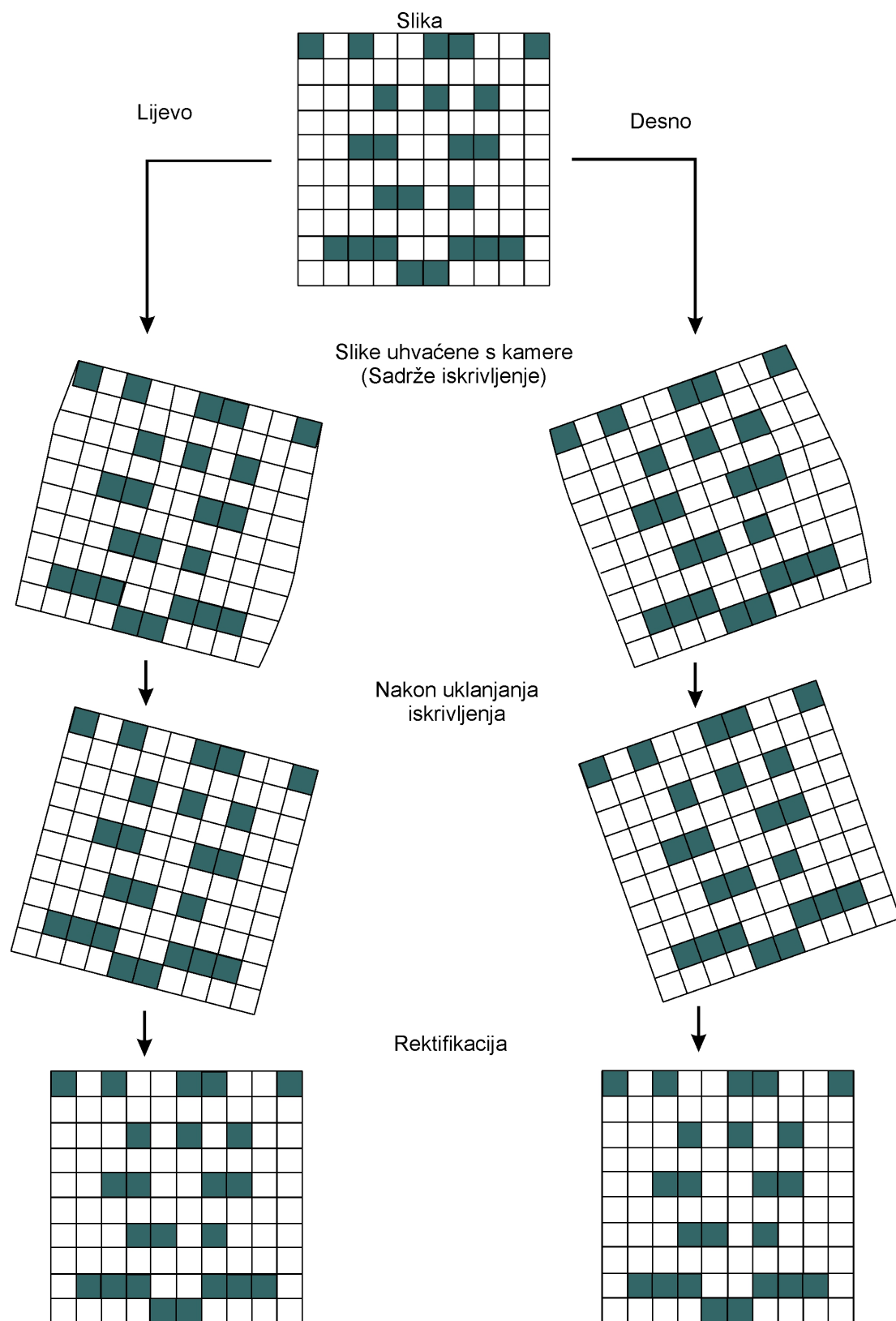
Slika 3.2: Stereovid



Slika 3.3: Epipolarna geometrija

varajuće točke slike. Pretpostavimo da su $x_1 \leftrightarrow x_2$ odgovarajuće točke u paru slika. Tada je $x_2^t F x_1 = 0$. Esencijalna matrica može biti definirana iz temeljne matrice kao $E = K_2^t F K_1$, a K je matrica kalibracije kamere definirana unutarnjim parametrima kamere [126].

Stereo rektifikacija (engl. Rectification) predstavlja transformaciju ravnine svake slike tako da konjugirane epipolarne linije postaju kolinearne i paralelne s jednom od osi slike [115], tj. kako u svojem radu navodi Nozick [74], rektifikacija stereoslike sastoji se u transformiranju dviju slika iste scene tako da za svaku korespondenciju piksela između dvaju pogleda spomenuti pikseli budu poravnati duž horizontalne osi slike. Na slici 3.4 prikazan je proces rektifikacije.



Slika 3.4: Stereo rektifikacija

Disparitet (engl. Disparity) predstavlja pomak, tj. razliku u poziciji između odgovarajućih točaka na dvjema slikama, one s lijeve i desne strane koju vidi stereokamera.

Na slici 3.5 prikazan je primjer mape dispariteta.



Slika 3.5: Mapa dispariteta

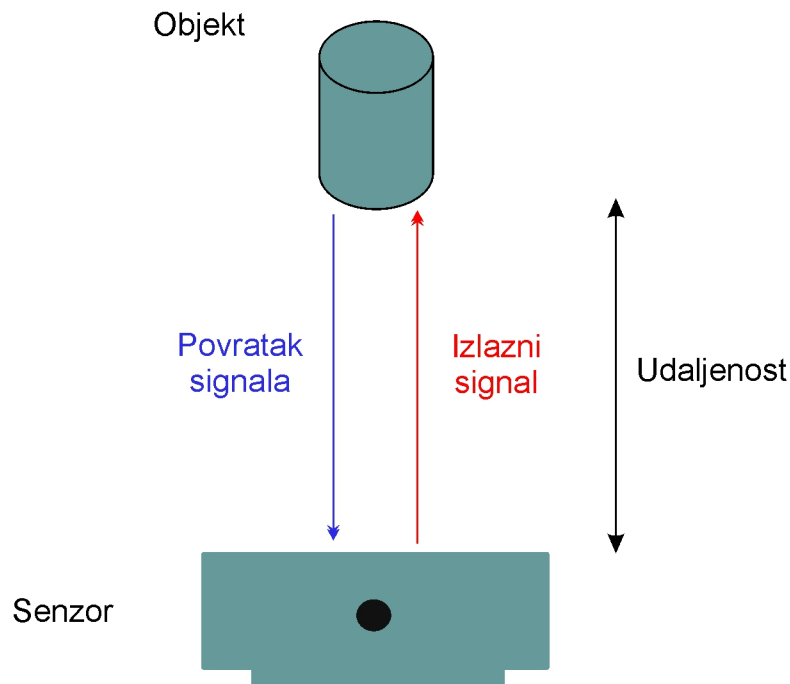
Kalibracija stereokamere podrazumijeva proces procjene parametara kamere te obično uključuje procjenu intrinzičnih ili unutarnjih (engl. Intrinsic) i ekstrinzičnih ili vanjskih (engl. Extrinsic) parametara. Intrinzični parametri su parametri kamere koji uključuju žarišnu duljinu, optički centar, iskrivljenje, dok su ekstrinzični parametri pozicija i orijentacija kamere u prostoru.

Podudaranje (engl. Matching) podrazumijeva pronalazak odgovarajućih točaka na slikama, dok je triangulacija (engl. Triangulation) utvrđivanje 3D koordinata na temelju 2D odgovarajućih točaka.

3.2.2 Time-of-flight kamere

Time-of-flight kamera emitira svjetlosni signal i mjeri vrijeme koje je potrebno da navedeni svjetlosni signal prođe put od kamere do objekta i natrag kako bi se utvrdila udaljenost između objekta i kamere. Na slici 3.6 prikazan je princip rada TOF kamere. TOF kamere se danas aktivno upotrebljavaju u robotskim sustavima i području računalnog vida.

Princip rada TOF kamere, prikazan na slici 3.6, jest sljedeći. Kamera emitira signal,



Slika 3.6: Princip rada TOF kamere

najčešće svjetlosni, koji putuje približno brzinom svjetlosti koja iznosi $3 \times 10^8 \text{ m/s}$. Obi-
 lježimo radi primjera udaljenost između objekta i kamere s d . Vrijeme koje je potrebno
 da signal prijeđe udaljenost jest t . Budući da signal prijeđe dvije udaljenosti d (jednu
 kada putuje od kamere i drugu kada se vraća), dužina prijeđenog puta onda je $2d$.
 Veza između udaljenosti i vremena može se napisati kao [71]:

$$2d = c \cdot t \quad (3.1)$$

$$d = \frac{c \cdot t}{2} \quad (3.2)$$

Primjeri TOF kamera su Panasonic D – Imager i Basler TOF kamera. Basler TOF
 kamera je prikazana na slici 3.1 (d). Danas na nekim pametnim telefonima također
 postoji implementirana TOF kamera. TOF kamere produciraju dubinske slike čiji pikseli
 imaju kodiranu udaljenost do odgovarajuće točke u sceni, a mogu se upotrebljavati za
 procjenu 3D struktura bez potrebe za tradicionalnim algoritmima koji se upotrebljavaju
 u računalnom vidu [30].

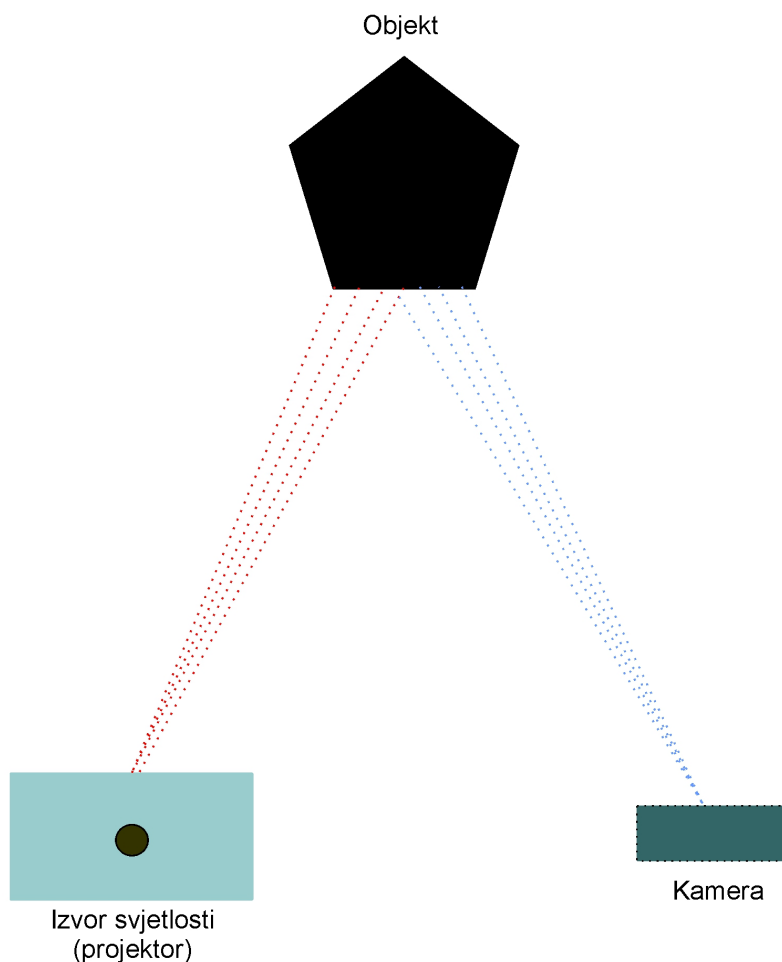
3.2.3 Kamere sa strukturiranom svjetlošću

Kamere sa strukturiranom svjetlošću postoje već neko vrijeme na tržištu, no veliku popularnost u istraživanjima doživjele su 2010. godine kad je predstavljen Microsoft Kinect, verzije jedan (postoji i Kinect verzije dva), koji je zasnovan na PrimeSense tehnologiji [129]. Uz Kinect, postoje još neki uređaji koji se uvelike upotrebljavaju, kao što je Asus Xtion. Asusov uređaj i Microsoftov Kinect (verzije jedan) prikazani su na slici 3.1 (a),(c).

Sustavi skeniranja zasnovani na strukturiranoj svjetlosti projiciraju različite svjetlosne obrasce ili strukture te analiziraju kako su ti obrasci iskrivljeni, tj. deformirani od strane scene za dobivanje 3D oblika objekta. Potencijalno brzo prikupljanje podataka, način bez kontakta, velika točnost mjerenja te dostupnost danas različitih kamera čine ovu tehnologiju privlačnom za upotrebu u različitim sferama ljudske djelatnosti. Ovaj sustav radi na sličnom principu kao i sustavi stereovida, samo što je kamera u ovom slučaju zamijenjena projektorom. Da bi se zaobišao problem oslanjanja na varijacije u teksturi objekta koji se skenira, projektor može projicirati i svoje specijalne teksture ili obrasce na objekt, a ti se obrasci nazivaju strukturiranim kodiranjima [9]. Ilustracija sustava sa strukturiranom svjetlošću prikazana je na slici 3.7.

Kako se u dosadašnjim istraživanjima dosad najčešće upotrebljavao Kinect, verzije jedan, koji spada u osjetilne uređaje manjeg dometa zasnovane na strukturiranoj svjetlosti, u nastavku ovog potpoglavlja detaljnije je opisan Kinect uređaj. Microsoft je 2010. godine predstavio Kinect za potrebe korištenja s XBOX 360 konzolom, a sve u cilju interakcije korisnika s konzolom, pokretima, bez potrebe za korištenjem kontrolera. Microsoft je naknadno izdao i verziju za korištenje s računalom i odgovarajući softverski razvojni komplet. Ranije su također upotrebljavani različiti uređaji ovog tipa u različitim djelatnostima, ali je pojava Kinecta verzije jedan (slika 3.1 (a)) znatno popularizirala navedene uređaje. Razlog tome bila je pristupačnost Kinecta i cijenom i tehnologijom koju je upotrebljavao, pa je Kinect postao popularan i u akademskoj i istraživačkoj zajednici gdje se upotrebljava u mnogim istraživačkim projektima.

Budući da uređaji manjeg dometa nisu pogodni za upotrebu u svim uvjetima, a u većina dosadašnjih istraživanja upotrebljavani su upravo ti uređaji, u ovoj disertaciji glavni je fokus na istraživanju i primjeni uređaja, odnosno kamere, većeg dometa. Primjerice, Kinect može detektirati osobu koja stoji u dometu od 0,8 m do 4 m, dok je



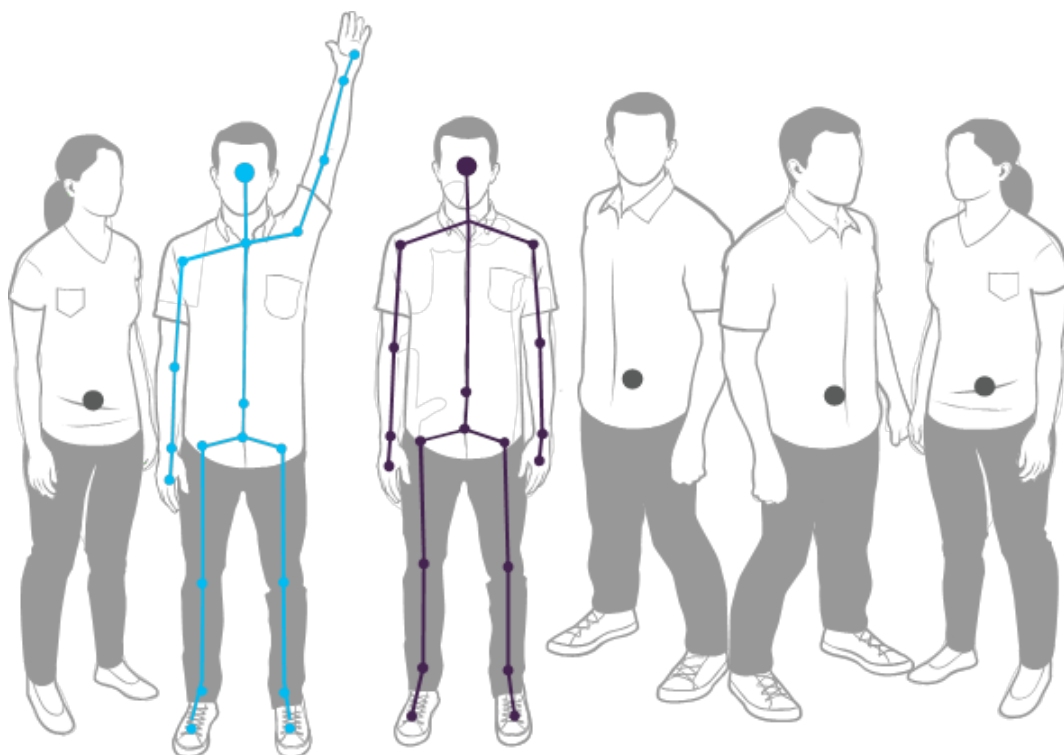
Slika 3.7: Ilustracija sustava sa strukturiranom svjetlošću

praktični domet, koji se pokazao u praksi, negdje između 1,2 m i 3,5 m u zadanom načinu rada (engl. Default Mode Range). U načinu rada za blizinu (engl. Near Mode) Kinect uređaj je osjetljiviji na bliže objekte te domet iznosi između 0,4 m i 3 m (0,8 m i 2,5 m za praktičnu uporabu).

Kinect je detaljno opisao autor Zhang [129] koji navodi da Kinect objedinjuje nekoliko hardverskih komponenti kao što su dubinsko osjetilo, kamera u boji i niz od četiriju mikrofona. Dubinsko osjetilo sastoji se od infracrvenog (engl. Infrared, IR) projektora u kombinaciji s IR kamerom, a dubinska tehnologija licencirala je izraelska kompanija PrimeSense [129] koja je danas dio kompanije Apple. IR projektor je IR laser koji prolazi kroz difrakcijsku rešetku i pretvara se u skup IR točaka [129]. Kako je navedeno na službenoj stranici Microsofta, Kinect servisi podržavaju značajke kao što su dubinska slika, RGB slika, nagib uređaja te omogućavaju praćenje kostura osobe. Moguće je specificirati razlučivost dubinske i RGB kamere neovisno, datotekom za konfiguriranje,

kao i način rada dubinske kamere.

Kod dubinskog osjetila navodi se domet između 0,8 m i 4 m u zadanom načinu rada. Usto se može primjenjivati i tzv. način za blizinu koji omogućava domet do 3 m [105]. Jedna od zanimljivijih stvari u vezi s upotrebom Kinecta jest opcija praćenja kostura osobe. Ti kosturni podatci sastoje se od skupa zglobova i Kinect ih detektira 20 na osobi. Tu su uključeni desna ruka osobe, desni ručni zglob, desni lakat, desno rame, desni kuk, desno koljeno, desni gležanj, desna noga, glava, centar između ramena, kičma, centar između kukova, lijeva ruka, lijevi ručni zglob, lijevi lakat, lijevo rame, lijevi kuk, lijevo koljeno, lijevi gležanj i lijeva noga. Također, Kinect koristeći se IR kamerom, može detektirati do šest osoba u vidnom polju uređaja (slika 3.8), a dvije osobe može detaljno pratiti [105]. Na slikama 3.9, odnosno 3.10. prikazana je i dubinska slika dobivena Kinectom kao i opcija praćenja kostura na osobi.



Slika 3.8: Ilustracija Kinectove mogućnosti praćenja šest osoba [105]

Iako je krajem 2017. Microsoft najavio da prekida proizvodnju Kinecta on se i dalje upotrebljava u mnogim istraživanjima. Osim toga, neovisno o razvoju klasičnog Kinecta, Microsoft danas nudi Azure Kinect DK, razvojni komplet s naprednim osjetilima za računalni vid i govorne modele [106].

Iz pregleda dosadašnjih istraživanja (potpoglavlje 1.2) vidi se da je Kinect uređaj



Slika 3.9: Dubinska slika dobivena Kinectom



Slika 3.10: Ilustracija praćenja kostura osobe Kinectom

upotrebljavan u najvećem broju dosad predstavljenih pristupa identifikacije osoba na temelju hoda. Za kreiranje vlastitog skupa podataka tijekom istraživanja prezentira-

nih u ovoj disertaciji upotrebljavana je stereokamera kompanije StereoLabs pod nazivom ZED [112]. ZED stereokamera i navedeni vlastiti skup podataka opisani su u sljedećem poglavlju ove disertacije.

4 SKUPOVI PODATAKA

Tijekom provedbe istraživanja upotrebljavana su tri skupa podataka: Casia Dataset B [130] [125] [102], TUM Gait from Audio, Image and Depth (GAID), TUM-GAID [33] te vlastiti skup podataka. Casia Dataset B i vlastiti skup podataka upotrebljavani su u preliminarnim istraživanjima i istraživanjima povezanima s analizom dobivenih silueta iz RGB i dubinskih slika.

Skup podataka TUM-GAID upotrebljavan je u eksperimentu s predloženim pristupima jer sadržava velik broj osoba u hodu, s različitim varijacijama kao što su normalno hodanje, nošenje ruksaka te varijacije u cipelama. TUM-GAID skup je bio najpogodniji za provođenje istraživanja jer su mogli biti razmotreni različiti aspekti koji mogu utjecati na točnost identifikacije kao što su podjela na različite veličine populacija ljudi, različite razdiobe visina te različite varijacije u hodu.

4.1 Skup podataka Casia Dataset B

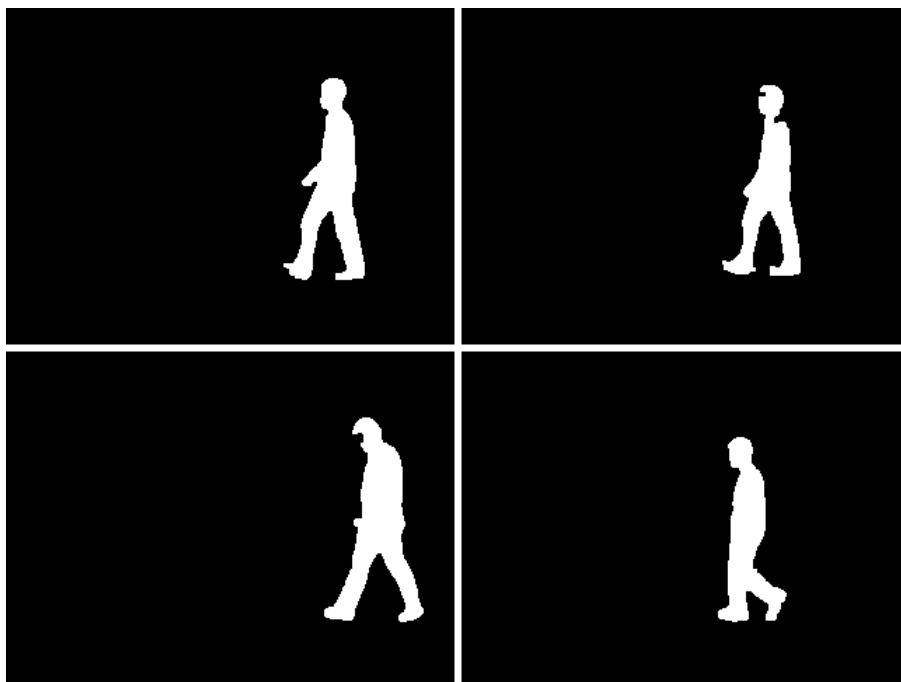
Skup podataka Casia Dataset B pripada velikom skupu podataka Casia. Skup podataka Casia podijeljen je na skupove A, B i C. Skup A (Casia Dataset A) kreiran je 2001. godine i sadržava 20 osoba, a svaka osoba ima 12 sekvencija slika, tj. četiri sekvencije za svaki od triju smjerova (paralelno, 45 i 90 stupnjeva). Kreiran je u vanjskom okruženju. Veličina ovog skupa jest oko 2,2 GB, a više detalja o njemu može se vidjeti na poveznici na koju se upućuje pod [102].

U slučaju skupa B (Casia Dataset B), koji je upotrebljavan za istraživanja predstavljena u ovoj disertaciji, radi se o većem skupu podataka osoba u hodu. Čine ga 124 subjekta čiji su podatci snimani iz 11 kutova u unutarnjem okruženju i razvrstani su u tri varijacije: promjene u odjeći, nošenje torbe i normalno hodanje [102]. Cijeli skup podataka može se dobiti ako se slijedi uputa na stranici na koju se upućuje pod [102]. Na slici 4.1 prikazan je primjer osobe u hodu iz ovog skupa s različitim varijacijama (normalno hodanje, varijacija u odjeći i nošenje torbe).

U svrhu istraživanja predstavljenog u ovoj disertaciji upotrebljavane su samo javno dostupne siluete osoba koje se također mogu pronaći na poveznici na koju se upućuje pod [102]. Primjer silueta osoba iz skupa podataka Casia Dataset B koje su upotrebljavane u istraživanju prikazan je na slici 4.2. Ovaj skup kreiran je 2005. godine.



Slika 4.1: Primjer osobe u hodu iz skupa podataka Casia Dataset B



Slika 4.2: Primjer silueta za četiri osobe iz skupa podataka Casia Dataset B

Skup C (Casia Dataset C) je kreiran s pomoću infracrvene (toplinske) kamere te sadržava 153 osobe s četiri načina hodanja: normalno hodanje, sporo hodanje, brzo hodanje i normalno hodanje uz nošenje torbe. Videozapisi su kreirani 2005. godine noću i u vanjskom okruženju. Više informacija i o ovom skupu također se može pronaći na poveznici na koju se upućuje pod [102].

Pored silueta koje se mogu slobodno preuzeti s navedene stranice, mogu se dobiti i GEI slike ako se odgovornoj osobi pošalje zahtjev. Cijeli skup Casia, tj. svi videozapisi i sve siluete, ima veličinu od oko 10 GB.

4.2 Skup podataka TUM-GAID

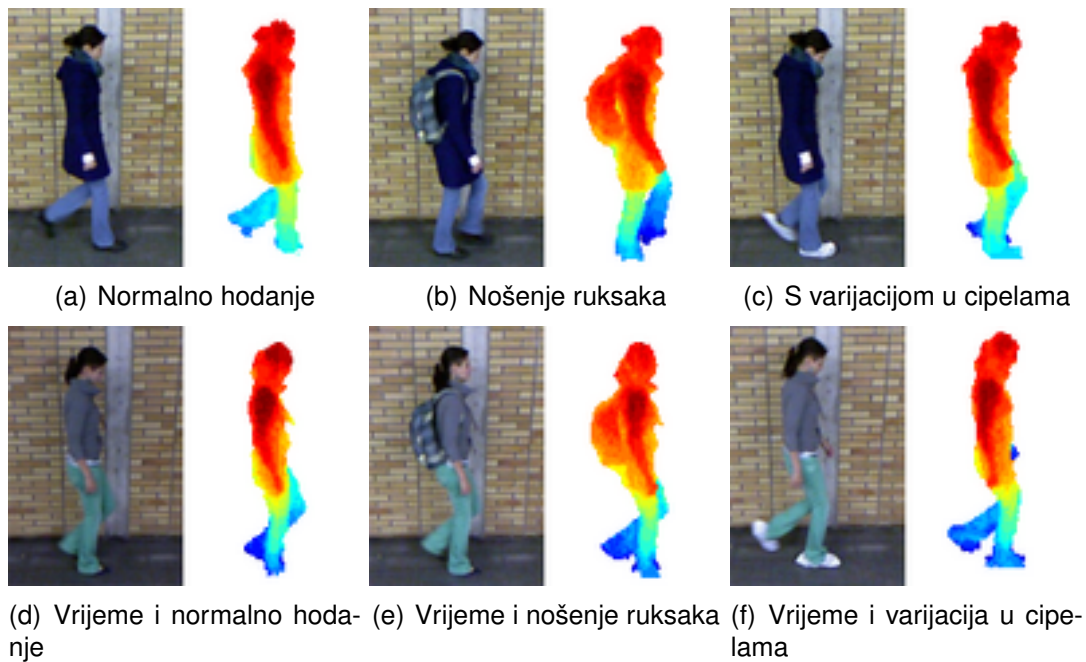
TUM-GAID [33] je velik skup podataka dostupan na zahtjev u svrhu istraživanja. Sadržava 305 osoba, od kojih su 186 muških, a 119 ženskih osoba. Prema dobi obuhvaća osobe između 18 i 55 godina, s prosjekom od 24,8 godina. Prema visini obuhvaća osobe između 152 cm i 195 cm. Osobe su snimane u dvama navratima, u siječnju i travnju 2012. godine.

Za kreiranje ovog skupa podataka upotrijebljen je Kinect, a skup sadržava osobe u trima varijacijama: normalno hodanje, varijacija s ruksakom te varijacija s cipelama. Varijacije s ruksacima pogodne su za proučavanje utjecaja degradacije silueta na uspješnost identifikacije osoba na temelju hoda, dok su varijacije s cipelama pogodne za proučavanje uspješnosti metoda za identifikacije koje su zasnovane na modalitetima audiotokova. Za svaku od navedenih 305 osoba postoji šest sekvencija u normalnom hodu, dvije s ruksakom i dvije s varijacijama u cipelama, tj. ukupno 10 snimaka za svaku osobu. Dodatno su 32 osobe snimane i drugi put za dodatnih 10 snimaka (s obzirom na vrijeme i odjeću). Prema tome su varijacije obilježene kao *normalno* - *N*, *ruksak* - *R*, *cipele* - *C*, *vrijeme normalno* - *VN*, *vrijeme ruksak* - *VR* i *vrijeme cipele* - *VC*. Za svaku su osobu spremljeni podatci kao što su spol, godine, visina i tip cipela. Primjeri iz skupa podataka TUM-GAID prikazani su na slici 4.3.

Kako se dalje navodi u radu [33], da bi se uređaj iskoristio na najbolji mogući način i da bi se izbjegla njegova ograničenja, skup podataka kreiran je u kampusu Tehničkog sveučilišta u Münchenu u hodniku širokom 3,5 m. Kinect je bio postavljen na visinu od 1,9 m i usmjeren prema dolje, ugrubo pod kutom od 13 stupnjeva. Osobe su hodale normalno u odnosu na liniju vida, na udaljenosti od oko 3 m, blizu suprotnog zida, što znači da su osobe prelazile udaljenost od približno 4 m kroz vidno polje. Osobe urade između 1,5 i 2,5 ciklusa hoda u svakoj snimljenoj sekvenciji. Svaka sekvencija traje približno 2-3 s.

4.3 Vlastiti skup podataka

Tijekom istraživanja izrađen je i upotrebljavan vlastiti skup podataka. Osnovna motivacija bila je kreirati skup podataka koji sadržava osobe u hodu koje se kreću na većoj udaljenosti od kamere u vanjskom okruženju.



Slika 4.3: Primjeri okvira iz skupa podataka TUM-GAID na jednoj osobi u šest različitih varijacija

Snimanje osoba na većoj udaljenosti od kamere omogućava snimanje više ciklusa hoda u vidnom polju. Veća udaljenost također više odgovara scenariju nadzora šireg područja, što je jedna od primarnih primjena identifikacije osoba na temelju hoda. U vanjskom okruženju uvjeti osvjetljenja mogu značajno utjecati na snimku. Snimanje osoba u hodu u vanjskom okruženju s nejednolikom pozadinom i u uvjetima osvjetljenja, koji se razlikuju u odnosu na kontrolirano osvjetljenje postojećih javno dostupnih skupova podataka, omogućava dodatnu analizu učinkovitosti metoda za segmentaciju osoba, analizu i usporedbu silueta koje se izlučuju iz RGB i dubinskih slika i u konačnici analizu robusnosti metoda identifikacije osoba na temelju hoda u navedenim uvjetima. Cilj je bio izraditi skup podataka koji sadržava snimke RGB-D kamerom navedenih obilježja za testiranje metoda identifikacije osoba na temelju hoda.

Preostali zahtjevi su slični zahtjevima koje su imali i ostali skupovi podataka osoba u hodu. Skup mora sadržavati snimke različitih osoba u hodu, nema preklapanja osoba u slici, svaka snimka mora sadržavati samo jednu osobu i mora biti snimljen cijeli ciklus hoda. Snimke moraju omogućiti izlučivanje RGB i dubinskih slika, a poželjno je da su u skupu podataka dostupne snimke istih osoba snimane iz različitih kutova u odnosu na kameru. Metodologija izrade je pratila onu od TUM-GAID skupa podataka [33]. Za kreiranje vlastitog skupa podataka upotrijebljena je ZED stereokamera [112].

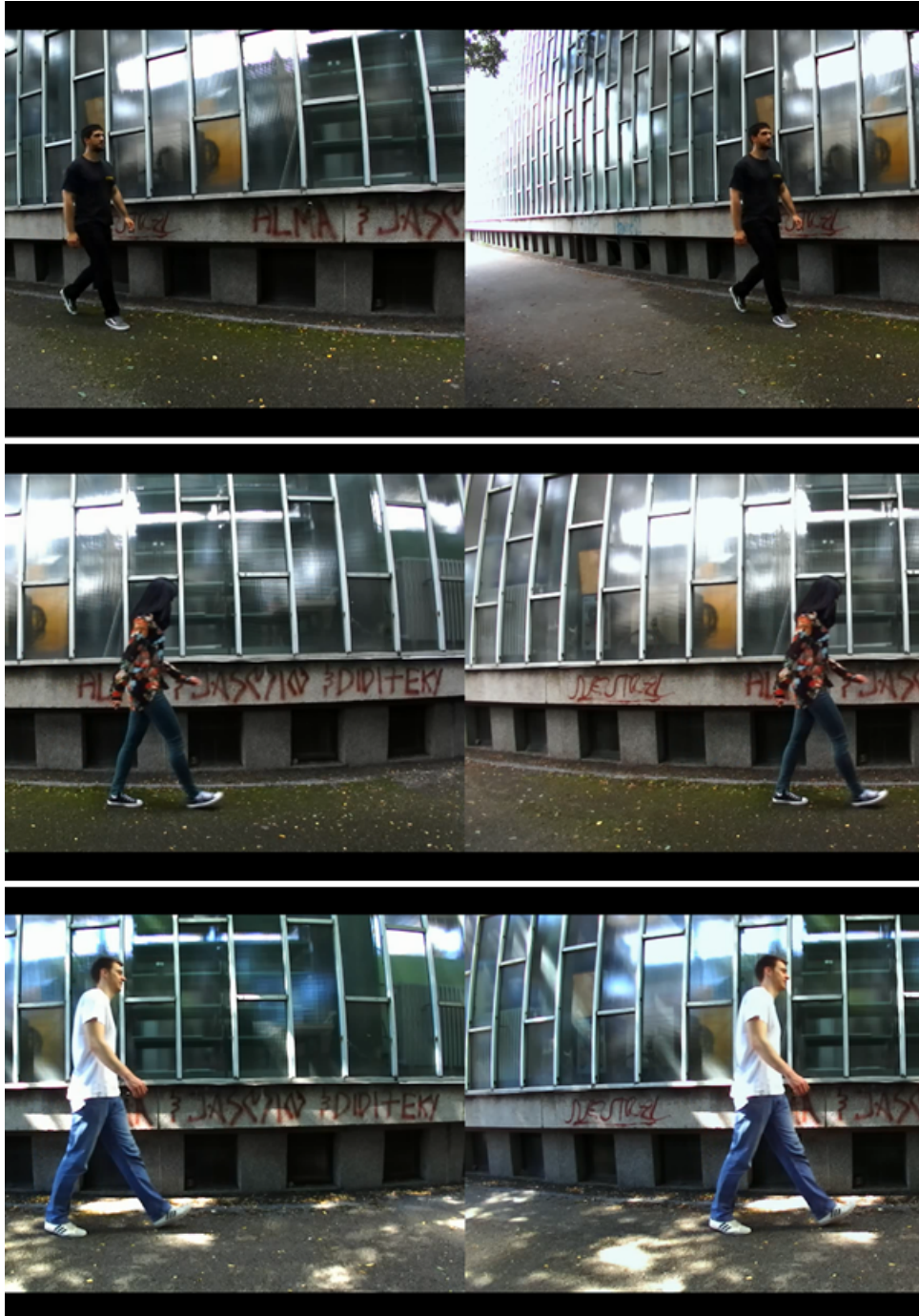
Novokreirani skup podataka čine snimke 14 osoba u normalnom hodu bez dodatka koje osobe nose kao što su, primjerice, torba ili ruksak. Skup je sniman po vedrom vremenu ispred Tehničkog fakulteta Sveučilišta u Rijeci. Skup je podijeljen tako da ga čini sedam muških i sedam ženskih osoba. Sve su osobe u ranim dvadesetim godinama.

Sadržaj skupa podataka čine videozapisi visoke razlučivosti (2560x720) snimani iz triju različitih kutova s po trima ponavljanjima za svaki. Kutovi snimanja su 60, 90 i 120 stupnjeva s udaljenostima od 2 m do 10 m. Sve su snimke pohranjene u ZED *.svo* formatu, ali su kreirane i pohranjene snimke u standardnom *.avi* formatu. ZED *.svo* format je korišten s obzirom na upotrebljavanu kameru, dok su snimke u *.avi* formatu kreirane i pohranjene s obzirom na druge skupove podataka koji su obično videozapise kreirali i pohranjivali u *.avi* formatu.

Skup podataka je kreiran tako da je za svaku od navedenih 14 osoba dostupno po devet datoteka u *.svo* formatu, što znači da ukupno ima 126 takvih datoteka. Jednako tako, za svaku od navedenih 14 osoba napravljeno je devet videozapisa u *.avi* formatu koji sadržavaju snimke s lijevog i desnog toka u jednom videozapisu (kao na slici 4.4). Na slici 4.4 prikazan je primjer pokrenutog videozapisa za tri različite osobe gdje je videozapis podijeljen na lijevi i desni tok stereokamere. Uz te videozapise, za svaku od 14 osoba kreiran je i videozapis u *.avi* formatu u kojem je odvojen lijevi i desni tok, tj. snimka s lijevog i desnog toka stereokamere, tako da su za svaku snimku dostupna po dva *.avi* videozapisa. U tom slučaju sada za 14 osoba postoji 252 videozapisa, tj. za svaku osobu po 18 videozapisa (za lijevi i desni tok kamere). Ilustracija je na slici 4.5 na kojoj se može vidjeti primjer za tri osobe za koje je pokrenut samo lijevi ili desni tok stereokamere. Uz navedene videozapise, za svaku od 14 osoba dostupne su RGB slike i slike koje sadržavaju samo siluete osoba. Veličina cjelokupnog vlastitog skupa podataka jest 66 GB.

Primjer videozapisa *.svo* formata iz skupa podataka u razvojnom okruženju prikazan je na slici 4.6. Iz prethodno navedenih videozapisa iz vlastitog skupa podataka dobivene su i siluete potrebne za provedbu eksperimentalnih istraživanja. Primjer silueta iz vlastitog skupa podataka prikazan je na slici 4.7.

ZED stereokamera je izabrana jer ima veći domet u odnosu na kamere korištene u javno dostupnim skupovima podataka osoba u hodu koji sadržavaju dubinske slike. U



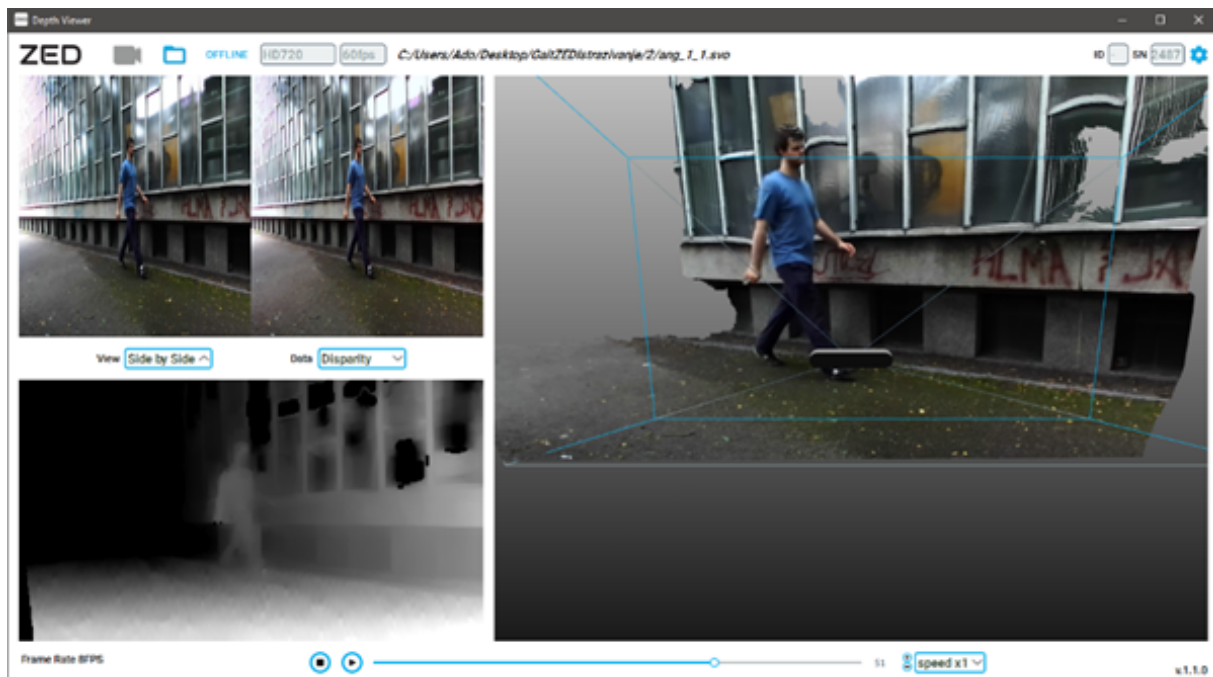
Slika 4.4: Stereo okviri za nekoliko osoba iz vlastitog skupa podataka kreirani ZED stereo kamerom

velikom broju predstavljenih radova autori su često upotrebljavali neki javno dostupan skup (primjerice, vrlo često Casia Dataset B) uz koji su upotrebljavali i neki manji vlastiti skup podataka kreiran Kinectom koji im je služio za provjeru robusnosti predložene metode.

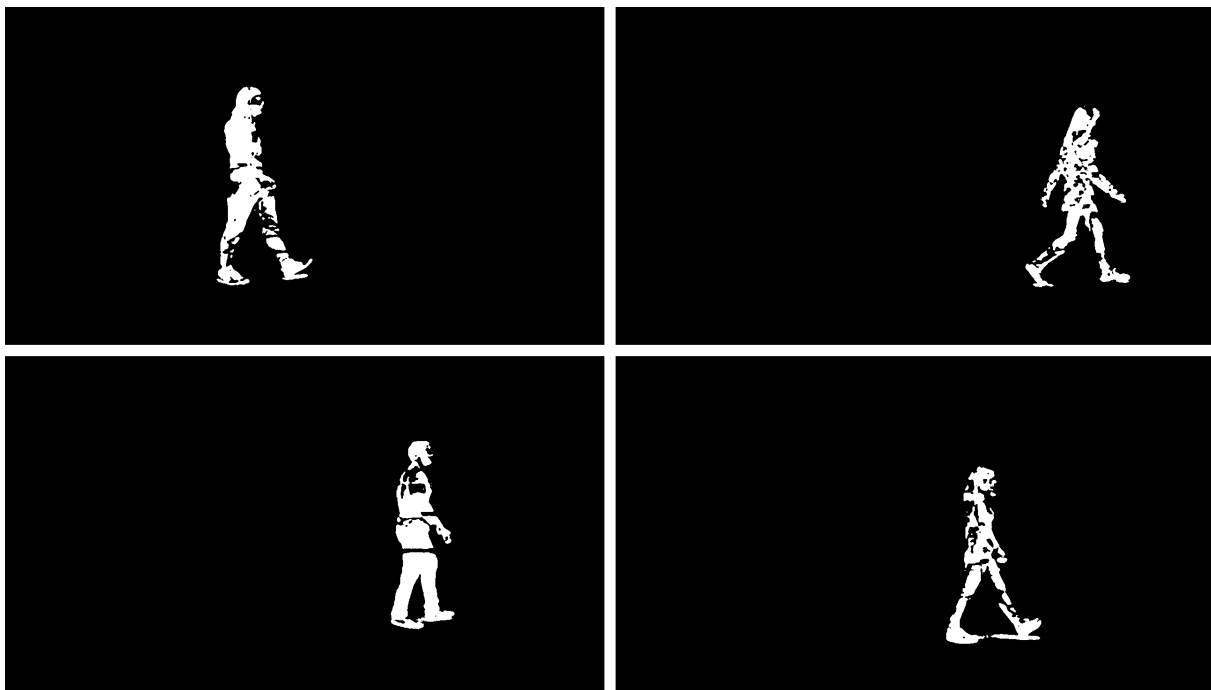
Budući da Kinect zbog tehnologije koju upotrebljava nije pogodan za vanjska okruženja, bio je potreban uređaj koji se mogao efikasno upotrebljavati u vanjskom okruženju



Slika 4.5: Primjer videozapisa s jednog toka stereo kamere za nekoliko osoba iz vlastitog skupa podataka kreirani ZED stereo kamerom



Slika 4.6: Prikaz RGB-D slike u razvojnom okruženju (ZED Depth Viewer)



Slika 4.7: Prikaz silueta izlučenih iz vlastitog skupa podataka

i to na većim udaljenostima, a da pritom pruža i dubinske slike. Na početku istraživanja, tijekom analize dostupnih kamera na tržištu, ZED stereokamera spadala je u kategoriju novih kamera na tržištu, a nudila je sve ono što je bilo potrebno za kreiranje skupa podataka i provedbu preliminarnih istraživanja. ZED stereokamera je prikazana a slici 3.1 (f).

ZED stereokamera proizvodi *side-by-side* videozapis u boji visoke razlučivosti, a na računalo se spaja USB sučeljem verzije 3.0. Spomenuti videozapis sadržava dva sinkronizirana videotoka, lijevi i desni. Pritom se taj videozapis u boji u sprezi sa ZED softverskim razvojnim kompletom (ZED SDK) na uređaju domaćinu upotrebljava za kreiranje dubinskih slika [112]. Za upotrebu ZED SDK potrebna je platforma NVIDIA za paralelno računanje i programski model (engl. Compute Unified Device Architecture, CUDA) [77] u svrhu izračunavanja dubinskih slika u stvarnom vremenu na grafičkom procesoru (engl. Graphics Processing Unit, GPU) računala domaćina. Zbog toga ZED SDK zahtijeva NVIDIA GPU [112]. ZED SDK je dostupan na službenoj stranici za operacijske sustave Windows i Linux, a posebna verzija nudi se i za ugradbeni sustav Jetson [116].

Minimalni hardverski zahtjevi uključuju dvojezgreni (engl. Dual-Core) 2,3 GHz procesor ili brži, 4 GB RAM memorije, grafičku karticu NVIDIA i USB 2.0 priključak. Preporučena konfiguracija jest četverojezgreni (engl. Quad-Core) 2,7 GHz procesor ili brži, 8 GB RAM memorije, grafička kartica NVIDIA GTX1060 ili brža, i USB 3.0 priključak.

5 PREDLOŽENI PRISTUPI ZA IDENTIFIKACIJU OSOBA

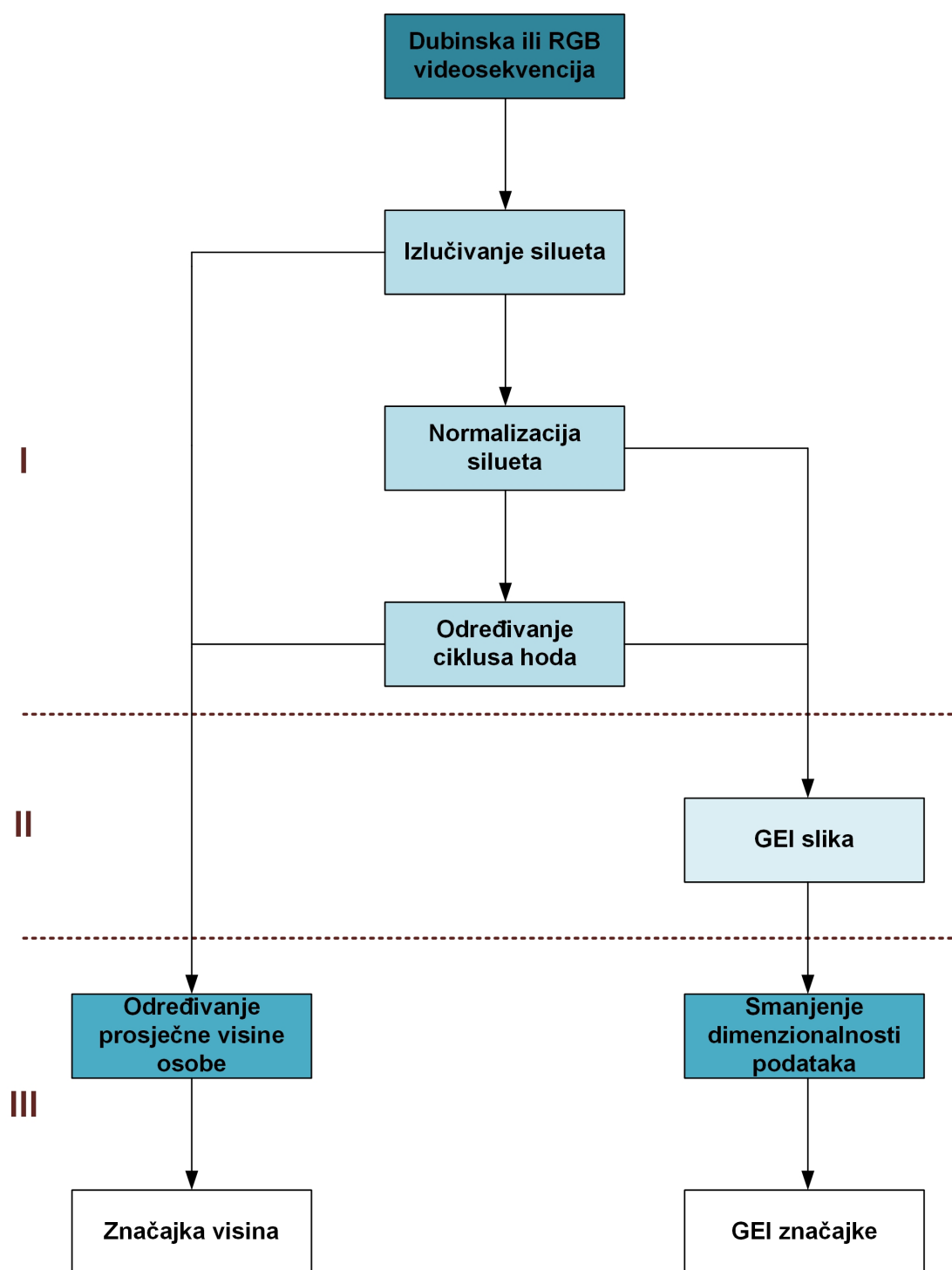
U ovom dijelu disertacije opisana su dva nova predložena pristupa za identifikaciju osoba na temelju hoda. Radi se o pristupima koji proširuju metode identifikacije osoba na temelju hoda, a koje su temeljene na pojavi, od kojih je najpoznatija slika energije hoda ili skraćeno GEI [28].

Pristupi su zamišljeni na način da se uz značajke vezane uz pojavu osobe, primjerice GEI značajke, u procesu identifikacije upotrebljava i visina osobe kao dodatna značajka. Značajku visina osobe moguće je izlučiti analizom dubinskih slika RGB-D kamere tijekom ciklusa hoda osobe.

Predloženi pristupi primijenjeni na GEI metodi nazvani su HGEI-i (Height-GEI Integration) i HGEI-f (Height-GEI Fusion) metoda, a nazivi predloženih metoda nastali su kombiniranjem prvog slova engleskog naziva za visinu i skraćenice za GEI metodu, uz dodatno slovo koje se odnosi na način kombiniranja značajki. Predloženi pristupi mogu se primjenjivati i za proširivanje drugih metoda koje, poput GEI metode, upotrebljavaju značajke temeljene na pojavi osobe. U disertaciji je tako proširena i BGEI metoda [92] gdje se također upotrebljavaju dubinske slike kako bi se izlučile siluete osoba neovisno o smjeru hoda osoba u odnosu na kameru (mogu biti dobivene s prednje i s bočne strane osobe). U tom se slučaju proširene metode nazivaju HBGEI-i i HBGEI-f metoda.

Predloženi pristupi se mogu primijeniti na postojećim metodama temeljenim na pojavi neovisno od postupaka koje te metode koriste za izlučivanje silueta osobe i određivanje ciklusa hoda. Može se tako izabrati metodu koja je najprikladnija za specifičnu primjenu i/ili samo prilagoditi postupke obrade slika i izlučivanja značajki prema primjeni i uvjetima u sceni. Na slici 5.1 prikazani su osnovni postupci izlučivanja značajki u slučaju kada se predloženi pristupi primjene na GEI metodi.

Gornji dio slike (I) prikazuje preliminarne korake obrade slika i izlučivanja značajki. Za te korake mogu se upotrebljavati metode koje su najpogodnije za specifičnu primjenu, uvjete u sceni, svojstva kamere i slično. U središnjem dijelu slike (II), prikazan je korak svojstven metodi koja se proširuje, odnosno stvaranje GEI slika u ovom slučaju (primjerice, kada se proširuje BGEI metoda u tom koraku se stvaraju BGEI slike). U donjem dijelu slike (III) prikazani su koraci koji su dio pristupa predloženih u ovoj diser-



Slika 5.1: Koraci pri izlučivanju značajki

taciji.

U nastavku je, u potpoglavlju 5.1, prikazan općeniti opis postupaka za obradu slika i izlučivanje silueta osoba koji se tipično upotrebljavaju u metodama identifikacije osoba na temelju hoda koje su temeljene na pojavi osobe. Također je opisan i postupak stvaranja GEI slika. U istom potpoglavlju su zatim detaljnije opisani postupci određivanja

prosječne visine osobe i smanjenja dimenzionalnosti podataka koji su svojstveni za predložene pristupe.

Postupci vezani uz izlučivanje značajki mogu biti identični za oba predložena pristupa pošto se isti razlikuju samo u načinu na koji se dobivene značajke kombiniraju. Nakon opisa postupaka vezanih uz izlučivanje značajki, u potpoglavlju 5.2 predstavljena su oba predložena pristupa pojedinačno i opisane njihove razlike kroz prikaz HGEI-i i HGEI-f metoda. Implementacijski detalji i konfiguracije parametara korištene tijekom istraživanja navedeni su u prilogu A.

5.1 Izlučivanje značajki

U ovom potpoglavlju opisani su postupci za izlučivanje značajki iz dubinskih ili RGB videosekvencija koji se upotrebljavaju u metodama identifikacije osoba na temelju hoda temeljenim na pojavi. Postupci su opisani slijedom prikazanim na slici 5.1.

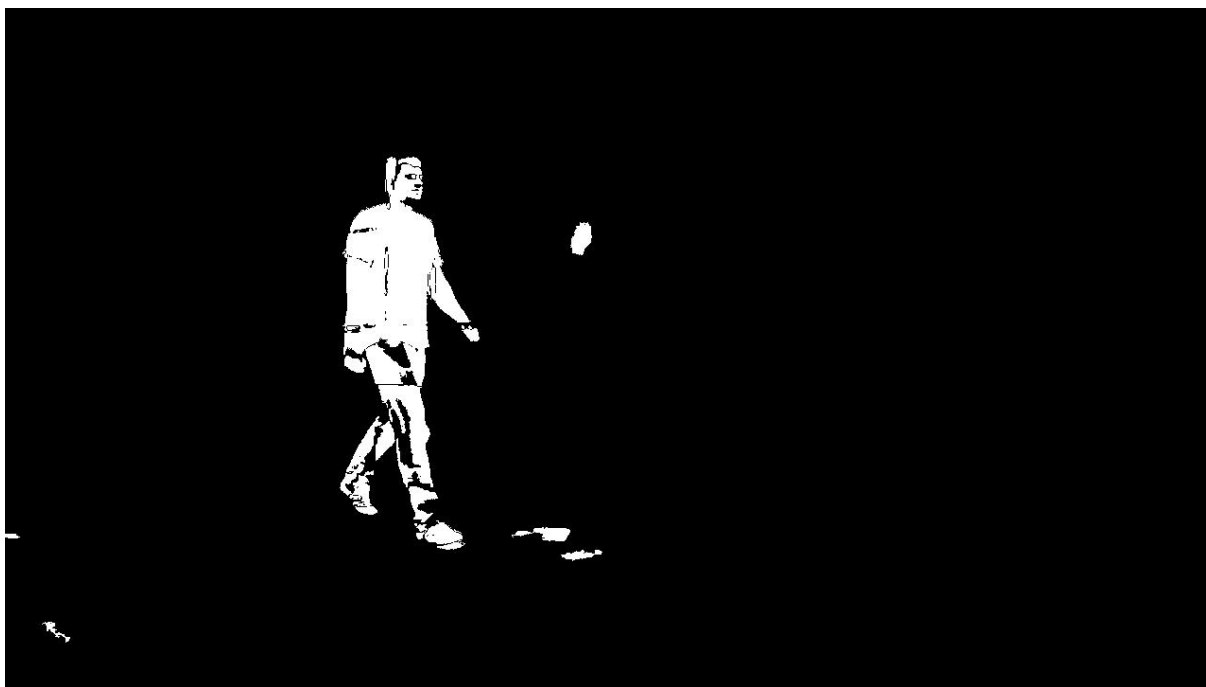
5.1.1 Izlučivanje silueta

Da bi se izlučile siluete osoba, primjenjuju se različiti postupci. To obično uključuje metodu poput oduzimanja pozadine (engl. Background Subtraction) i metodu binarnog ograničavanja (engl. Thresholding). Oduzimanje pozadine predstavlja tehniku za segmentiranje objekata od interesa u sceni [61], pri čemu se obavlja oduzimanje između slike koja sadržava objekt od interesa u sceni i modela pozadine. Model pozadine koji se oduzima može biti zadan, odnosno statički postavljen, ili se može dinamički određivati.

Za oduzimanje pozadine obično se primjenjuju poznate metode kao što su modeli Gaussove mješavine (engl. Gaussian Mixture Models, GMM) [114] [42], modeliranje pozadine i oduzimanje po konstrukcijama kodne knjige (engl. Background Modeling and Subtraction by Codebook Construction) [45] te procjena gustoće jezgra (engl. Kernel Density Estimation, KDE) [20].

Nakon oduzimanja modela pozadine od slike primjenjuje se binarno ograničavanje za postavljanje vrijednosti intenziteta piksela na slici. Radi se o metodi koja je zasnovana na postavljanju vrijednosti intenziteta piksela na slici (u odnosu na zadanu vrijednost) na način da se pikseli ispod određene zadane vrijednosti konvertiraju u crno, dok se oni čija je vrijednost iznad određene zadane vrijednosti konvertiraju u bijelo.

Rezultat oduzimanja pozadine i primjene metode binarnog ograničavanja jest binarna slika, odnosno maska prednje strane (engl. Foreground Mask). Binarna slika jest slika koja sadržava piksele koji imaju samo dvije moguće vrijednosti, 0 ili 1 (ili 255). Pritom 0 predstavlja crno, a 1 (ili 255) bijelo. Primjer dobivene binarne slike, odnosno maske prednje strane na kojoj se vidi silueta osobe, i to na vlastitom skupu podataka, prikazan je na slici 5.2. Na binarnim slikama koje sadržavaju siluete osoba obično se primjenjuju standardne metode za uklanjanje šuma sa slika primjenom, primjerice, linearnih filtera ili medijan filtra.



Slika 5.2: Primjer maske prednje strane dobivene iz vlastitog skupa podataka

Za oduzimanje pozadine tijekom istraživanja s vlastitim skupom podataka u ovoj su disertaciji primijenjene metode zasnovane na GMM metodi. GMM se općenito upotrebljava za grupiranje podataka odnosno za kreiranje klastera podataka (engl. Data Clustering). Klaster (engl. Cluster) predstavlja skupinu srodnih podataka. Kod GMM-a klasteri su modelirani kao Gaussove razdiobe. S obzirom da je slika predstavljena kao niz piksela gdje je svaki piksel predstavljen skalarom (ili vektorom) koji opisuje intenzitet (ili boju) GMM metoda je pogodna za podjelu piksela u slične segmente odnosno grupiranje piksela sa sličnim vrijednostima intenziteta.

Neka je X_t vrijednost boje ili intenziteta za određeni piksel slike promatran u trenutku t , povijest piksela $X_1, \dots, X_t$ se modelira kao mješavina K Gaussovih razdioba.

Vjerojatnost promatranja određene boje ili vrijednosti intenziteta u trenutku t izražava se kao [7]:

$$P(X_t) = \sum_{i=1}^K \omega_{i,t} \eta(X_t, \mu_{i,t}, \Sigma_{i,t}) \quad (5.1)$$

pri čemu su: K broj Gaussovih klastera, $\omega_{i,t}$ je težinska vrijednost vezana za i -ti klaster u trenutku t , $\mu_{i,t}$ i $\Sigma_{i,t}$ su srednja vrijednost i matrica kovarijance (engl. Covariance Matrix), η je Gaussova funkcija gustoće. Zbog učinkovitosti izračuna matrica kovarijance $\Sigma_{i,t}$ se pretpostavlja da je izotropijska (matrica kovarijance je izotropijska ako je proporcionalna jediničnoj matrici) i dijagonalna te se opisuje formulom [7]:

$$\Sigma_{i,t} = \sigma_{i,t}^2 I \quad (5.2)$$

gdje je σ^2 varijanca. Detaljan opis izlučivanja silueta upotrebom različitih programskih paketa opisan je u prilogu A.1. U skupu podataka TUM-GAID dostupne su siluete osoba te ih nije potrebno izlučivati.

5.1.2 Normalizacija silueta

Vrlo rijetko osoba hoda na istoj udaljenosti od kamere što znači da, ako se mijenja udaljenost od kamere, veličina siluete osobe u slici također varira. Zbog toga se sve siluete obično normaliziraju, odnosno svode na istu visinu. Primjerice, ako su siluete osoba izlučene iz videosekvencije potrebno je proći kroz svaki okvir te izvršiti normalizaciju što često uključuje definiranje područja od interesa (engl. Region Of Interest, ROI) tako da se sa slika koje sadržavaju siluete osoba uzme samo dio slike neposredno oko siluete osobe. Obično se oko siluete ocrta pravokutnik koji se naziva granični pravokutnik (engl. Bounding Box). Granični pravokutnik se može definirati tako da se na slici odredi objekt od interesa, što je u ovom slučaju silueta osobe, na kojem se pronađu najizbočeniji pikseli i od tih piksela se definira udaljenost kojom prolaze linije pravokutnika.

Neka je u okviru O_t u vremenu t detektirana osoba s graničnim pravokutnikom B_t i neka je silueta osobe označena sa S_t . Da bi se registrirala silueta osobe, središte siluete P_t (referentna točka odnosno piksel s koordinatama x_0 i y_0) u vremenu t se računa kao [18]:

$$P_t(x_0, y_0) = \left(\frac{M_{10}}{M_{00}}, \frac{M_{01}}{M_{00}} \right) \quad (5.3)$$

pri čemu su M_{ij} momenti binarne slike siluete. Navedeni momenti binarne slike siluete su definirani kao:

$$M_{ij} = \sum_x \sum_y x^i y^j S_t(x, y) \quad (5.4)$$

Slika siluete S postavi se na fiksnu visinu h , da bi se dobila jednaka visina za sve siluete, te se obreže s lijeve i desne strane kako bi sve slike bile jednake širine. Primjer originalne slike (lijeva slika) i normalizirane slike (desna slika, manja slika) prikazan je na slici 5.3.



Slika 5.3: Prikaz normalizacije silueta

Definiranje područja od interesa često je potrebno i zato što, primjerice, dobivena maska prednje strane koja sadržava siluetu osobe sadržava još određenih manjih objekata (vidi se i na primjeru na slici 5.2) koje je potrebno ukloniti te se izlučivanjem samo dijela slike koji sadržava siluetu osobe navedeni nepotrebni dijelovi odbacuju.

5.1.3 Određivanje ciklusa hoda osobe

Jedna videosekvencija može uključivati nekoliko ciklusa hoda osobe. Ciklus hoda, kao što je navedeno u [76], može biti mjereno od bilo kojeg sljedećeg događaja istom nogom, tj. od jednog poteza nogom pa do sljedećeg poteza istom nogom. Dvije faze su svojstvene za ciklus hoda, faza držanja (engl. Stance Phase) i faza zamaha (engl.

Swing Phase). Faza držanja obuhvaća vrijeme dok je noga na zemlji i približno 60 % ciklusa hoda odvija se u ovoj fazi, dok je u fazi zamaha noga u zraku [76].

Neka je dana sekvencija silueta osobe, ciklus hoda može biti detektiran tako da se izračuna normalizirana autokorelacija (engl. Normalized Auto Correlation, NAC) za sekvenciju hoda duž vremenske osi [128]:

$$NAC(N) = \frac{\sum_{x,y} \sum_{n=0}^{N_{uk}-N-1} S(x,y,n)S(x,y,n+N)}{\sqrt{\sum_{x,y} \sum_{n=0}^{N_{uk}-N-1} S(x,y,n)^2} \sqrt{\sum_{x,y} \sum_{n=0}^{N_{uk}-N-1} S(x,y,n+N)^2}} \quad (5.5)$$

pri čemu su: $NAC(N)$ autokorelacija za N pomak okvira, N_{uk} je ukupan broj okvira u svakoj sekvenciji hoda, $S(x,y,n)$ je vrijednost piksela na poziciji (x,y) n -tog okvira.

Ciklus hoda se može odrediti kao:

$$T_c = \arg \max_N NAC(N) \quad (5.6)$$

Ciklus hoda također može biti procijenjen na temelju analize vremenskog slijeda širine i visine siluete. Širina siluete osobe predstavljena je kao funkcija vremena, odnosno signal, pri čemu os X predstavlja vrijeme, a os Y amplitudu signala. Ova funkcija vremena je periodična funkcija izraženih vrhova i dolina: ovisno o kretanju osobe, vrhovi i doline šire se i skupljaju kako se noge osobe šire i ponovno približavaju jedna drugoj tijekom ciklusa hoda. Vrhovi i doline signala odgovaraju označenim siluetama u određenim pozama (primjerice, gdje osoba ima obje noge na tlu, raširene, i gdje je prva noga desna noga opisano je određenom amplitudom).

Kada se osoba analizira s prednje strane, širina siluete ne daje toliko informacija, ali visina osobe kao funkcija vremena vrlo je korisna jer vrhovi i doline mogu pokazivati kada je osoba s dvije noge na tlu (primjerice, kada su noge raširene i kada je desna noga naprijed), a kada je, recimo, jedna noga u zraku (primjerice, lijeva noga je u zraku i prolazi neposredno pored desne noge). I u tom slučaju os X predstavlja vrijeme, a os Y predstavlja amplitudu signala te vrhovi i doline signala odgovaraju označenim siluetama u određenim pozama, kao i u slučaju kada se upotrebljava širina siluete [16] [119].

5.1.4 Kreiranje GEI slika

GEI slika u osnovi predstavlja slijed ljudskog kretanja u jednoj slici, odnosno višestruke siluete koje su normalizirane, poravnate te vremenski usrednjene i prikazane u jednoj slici. GEI metodu predstavili su autori Ju Han i Bir Bhanu [28].

Da bi se kreirale GEI slike, idealno bi bilo koristiti se uvijek jednim cijelim ciklusom hoda osobe, ali pritom nije toliko važno gdje je točno definiran početak ciklusa hoda sve dok se ustraje na konzistentnosti da se koristi cijeli ciklus hoda. GEI slika izlučena tijekom jednog ciklusa hoda definirana je formulom:

$$G(x, y) = \frac{1}{N} \sum_{t=1}^N S(x, y, t) \quad (5.7)$$

pri čemu je N broj okvira sa siluetama u ciklusu hoda, t predstavlja broj okvira u ciklusu hoda u trenutku vremena, $S(x, y)$ je originalna slika siluete s (x, y) vrijednostima u 2D koordinatama slike. Primjeri GEI slika za nekoliko skupova podataka prikazani su na slici 5.4.

5.1.5 Određivanje visine osobe iz siluete

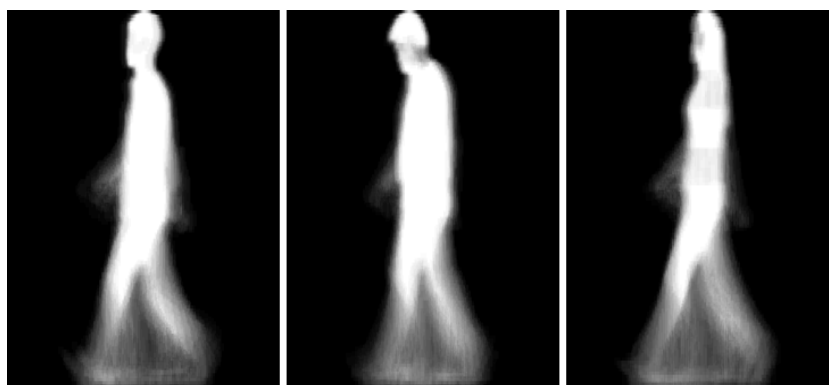
Nakon što su dostupne siluete osoba, može se procijeniti visina osobe za svaku osobu upotrebljavajući informacije o udaljenosti osobe od kamere. Procjena visine osobe se može odrediti jer je udaljenost svake točke siluete od kamere kodirana u dubinskoj slici. Na slici 5.5 (a) prikazana je dubinska slika sa siluetom osobe, dok je na slici 5.5 (b) prikazan odnos stvarne udaljenosti objekta i procijenjene vrijednosti dubinske kamere.

U ovom slučaju značajka visina osobe određena je kao prosječna vrijednost procijenjenih visina silueta osobe u svakom okviru kroz cijeli ciklus hoda. Primjerice, visina osobe u jednom okviru može se odrediti kao visina područja od interesa. Opisuje se i sljedećom formulom:

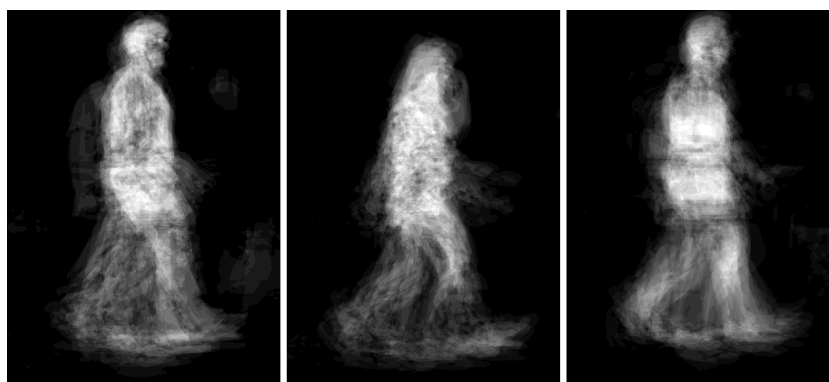
$$H(t) = \frac{1}{N} \sum_{t=1}^N h(t) \quad (5.8)$$

pri čemu je $h(t)$ procijenjena visina osobe u jednom okviru.

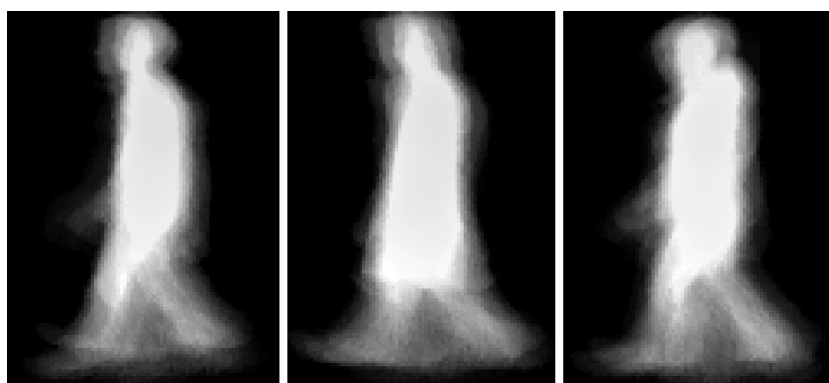
Točnost procijenjenih visina ovisi o različitim faktorima kao što su tip i tehnologija dubinskog osjetila, uvjeti okruženja, trenutačno stanje kretanja osobe itd. S obzirom



(a) Casia Dataset B



(b) Vlastiti skup podataka



(c) TUM-GAID

Slika 5.4: GEI slike za tri osobe na primjeru triju skupova podataka

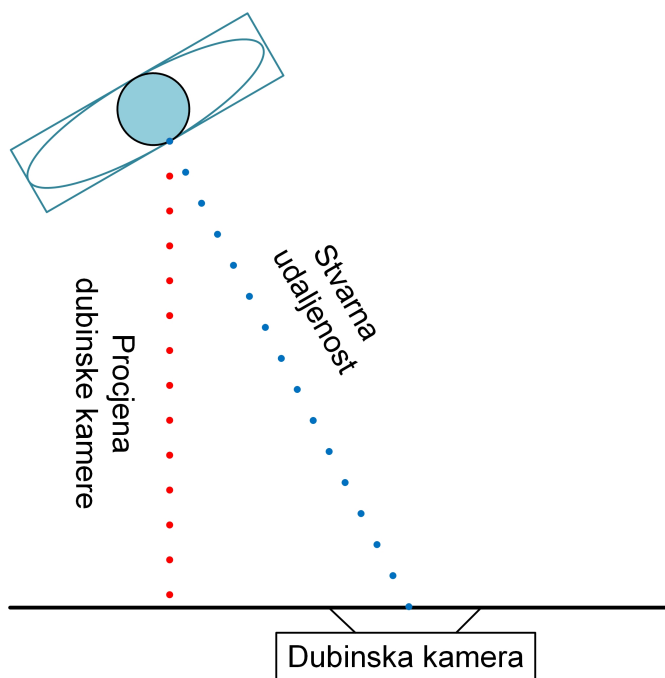
na to da se procjena visine izvodi nad pojedinačnim siluetama ciklusa hoda koje se upotrebljavaju i za izračun GEI slike, izračun se može obavljati istovremeno. U odnosu na ostale korake za obradu slika i izlučivanje značajki, zanemariva je vremenska složenost procjene visine siluete.

5.1.6 Smanjenje dimenzionalnosti podataka

Primjerice, kod predloženog pristupa gdje se radi spajanje značajki prije klasifikacije, odnosno HGEI-i metode, informacije o visini osobe dodaju se kao jedna od značajki



(a) Udaljenost od kamere kodirana u dubinskoj slici

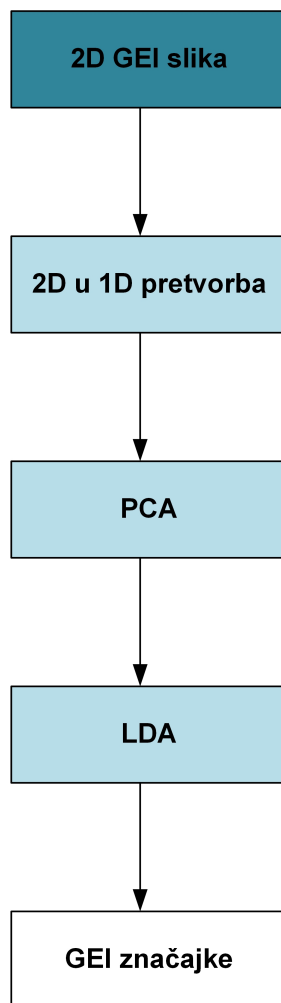


(b) Odnos sa stvarnom udaljenosti

Slika 5.5: Prikaz dubinske slike gdje je kodirana informacija o udaljenosti, odnos između stvarne udaljenosti i procjene dubinske kamere

uz dobivene značajke GEI slike. Prije nego se to učini, smanjuje se dimenzionalnost podataka primjenom analize glavnih komponenti (engl. Principal Component Analysis, PCA) i linearne diskriminantne analize (engl. Linear Discriminant Analysis, LDA). Proces koji ovo ilustrira prikazan je na slici 5.6.

Kada je dostupna GEI slika (ili GEI slike), da bi se dobile značajke koje se upotrebljavaju tijekom procesa identifikacije, GEI slika se transformira, odnosno upotrebljava



Slika 5.6: Koraci pri smanjenju dimenzionalnosti podataka

kao matrica podataka. Ta matrica podataka predstavlja GEI sliku opisanu s pomoću brojeva u redovima i stupcima, pri čemu svaki broj predstavlja vrijednost intenziteta pojedinog piksela slike.

PCA predstavlja metodu za smanjenje dimenzionalnosti velikih skupova podataka, što funkcionira tako da se varijable velikog skupa podataka pretvore u manji skup podataka koji sadržava većinu informacija većeg skupa. Drugim riječima, PCA se upotrebljava za projekciju podataka iz višedimenzionalnog prostora u prostor nižih dimenzija. LDA također spada u metode za smanjenje dimenzionalnosti, s tim da je cilj postizanje veće odvojivosti između klasa značajki. Često se primjenjuje kao radnja predobrade u različitim primjenama strojnog učenja. Te se metode često kombiniraju pa se nakon PCA-a provodi LDA. Primjenom PCA-a i LDA-a dobivaju se GEI značajke koje se upotrebljavaju tijekom procesa identifikacije. Na temelju preliminarnih eksperimentalnih istraživanja u vezi s vremenom izračuna i točnosti te ciljane primjene, broj komponenti

koji se izlučuje postavljen je na 50 za PCA i 10 za LDA.

PCA se općenito može realizirati kroz nekoliko koraka [113]. Ti koraci uključuju:

1. definiranje skupa podataka koji će se koristiti
2. računanje srednje vrijednosti za svaku dimenziju skupa podataka te njeno oduzimanje
3. računanje matrice kovarijance
4. računanje svojstvenih vektora i svojstvenih vrijednosti (engl. Eigenvectors, Eigenvalues)
5. sortiranje dobivenih svojstvenih vektora u padajućem redoslijedu i odabir svojstvenih vektora s najvećim svojstvenim vrijednostima te kreiranje vektora značajki odnosno matrice koju čine svojstveni vektori
6. transformiranje u novi skup podataka.

Neka postoji višedimenzionalni skup podataka X gdje je $X = [X_1, X_2, \dots, X_n]$, $i = 1, \dots, n$. Za dati skup podataka X potrebno je za svaku dimenziju izračunati srednju vrijednost te je oduzeti od pojedine vrijednosti za svaku dimenziju. Nakon toga potrebno je izračunati matricu kovarijance C (često se za matricu kovarijance upotrebljava i oznaka Σ) cijelog skupa podataka X .

Kada se pronađu svojstveni vektori iz matrice kovarijance oni se sortiraju po svojoj vrijednosti od najveće do najmanje vrijednosti po značaju. U vezi s tim potrebno je kreirati vektor značajki odnosno matricu vektora. Radi se na način da se uzmu svojstveni vektori koji se žele zadržati (po važnosti) i tako formira matrica.

Vektor značajki W se opisuje na sljedeći način:

$$W = (e_1 e_2 \dots e_n), i = 1, \dots, n \quad (5.9)$$

gdje e_i predstavlja svojstveni vektor.

U zadnjem koraku se vektor značajki koristi da bi se uzorci transformirali u novi potprostor (cilj PCA je naći potprostor izvornog prostora koji je optimalan za reprezentaciju skupa uzoraka) koristeći formulu:

$$y = W^T x \quad (5.10)$$

pri čemu je x $d \times 1$ dimenzionalni vektor koji predstavlja jedan uzorak, a y je tran-

sformirani uzorak u novom potprostoru.

LDA se može realizirati kroz tri koraka:

1. Računanje razdvojenosti između klasa (poznato je još pod nazivom varijanca između klasa ili matrica između klasa) (engl. Between-Class Variance, Between-Class Matrix) što je definirano kao udaljenost između srednje vrijednosti različitih klasa.

Definira se na sljedeći način:

$$S_b = \sum_{i=1}^n N_i (\bar{x}_i - \bar{x})(\bar{x}_i - \bar{x})^T \quad (5.11)$$

pri čemu su: S_b matrica raspršenja, N_i je broj uzoraka klase i , $\bar{x}$ je ukupna srednja vrijednost, a $\bar{x}_i$ srednja vrijednost klase i .

2. Računanje varijance unutar klase (engl. Within-Class Variance, Within-Class Matrix) odnosno udaljenosti između srednje vrijednosti i uzoraka svake klase.

Definirano je sljedećom formulom:

$$S_w = \sum_{i=1}^n (N_i - 1) S_i = \sum_{i=1}^n \sum_{j=1}^{N_i} (x_{i,j} - \bar{x}_i)(x_{i,j} - \bar{x}_i)^T \quad (5.12)$$

pri čemu su S_w matrica raspršenja unutar klase, a S_i matrica raspršenja za klasu i .

3. Kreiranje prostora niže dimenzije koji maksimizira varijancu između klasa, a minimizira varijancu unutar klase. Osnovni cilj LDA je naći matricu projekcije P_{lda} koja se definira kao:

$$P_{lda} = \arg \max_P \frac{|P^T S_b P|}{|P^T S_w P|} \quad (5.13)$$

pri čemu je P projekcija prostora niže dimenzije.

5.2 Opis predloženih metoda

U ovoj doktorskoj disertaciji predstavljena su dva pristupa koja proširuju metodu identifikacije osoba na temelju hoda koja je temeljena na pojavi. Riječ je o metodama HGEI-i i HGEI-f u kojima se GEI metoda proširuje upotrebom dodatne značajke visina osobe koja se izlučuje analizom silueta osobe iz dubinskih slika.

Identifikacija osoba s navedenim predstavljenim metodama je realizirana upotrebljavajući proces klasifikacije (engl. Classification). Klasifikacija je proces predviđanja kojoj klasi pripadaju određeni ulazni podatci odnosno, u ovom konkretnom slučaju, kojem predlošku osoba spremljenih u bazu u postupku upisa osoba u bazu, pripadaju ulazne značajke. Klase se još ponekad nazivaju ciljevi, oznake ili kategorije.

Razlika između navedenih dviju novih metoda jest u načinu kombiniranja značajki tijekom klasifikacije. Metoda HGEI-i vrši ranu fuziju na razini značajke, dok je metoda HGEI-f zasnovana na kasnoj fuziji na razini odluke [64]. Rana fuzija odnosi se na fuziju značajki, pri čemu se svojstvene značajke dobivene iz različitih izvora, primjerice, vizualne značajke, slušne značajke, tekstualne značajke itd. (ili, u slučaju predloženih metoda, GEI značajke i značajka visine), kombiniraju te se potom koriste za klasifikaciju primjenom nekog od algoritama klasifikacije (primjerice, kNN). U slučaju kasne fuzije klasifikacija se provodi zasebno za svaku od navedenih svojstvenih značajki te se dobiveni rezultati potom integriraju.

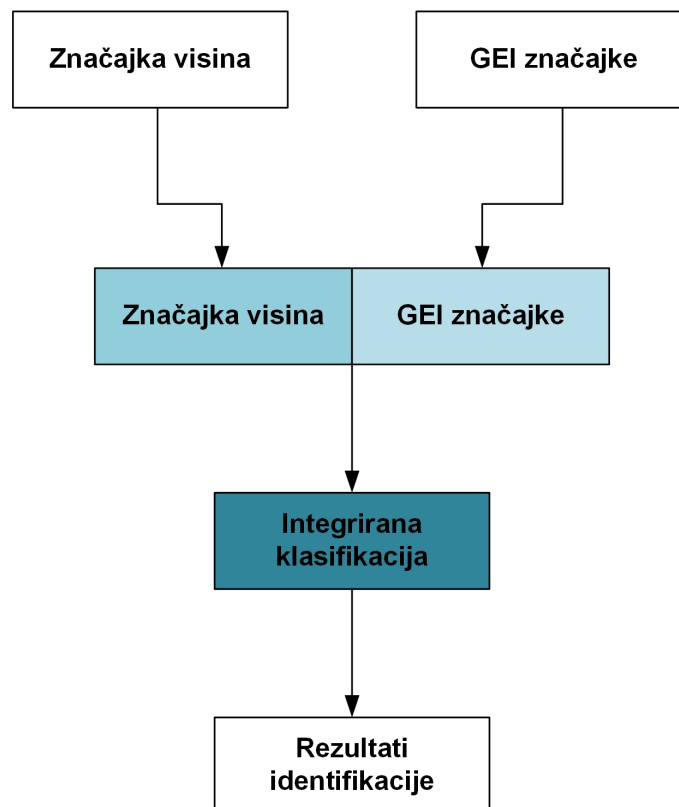
Inicijalni koraci obrade koji uključuju izlučivanje silueta osobe, normalizaciju silueta, određivanje ciklusa hoda, kreiranje GEI slika, određivanje visine osobe te smanjenje dimenzionalnosti podataka, a koji su opisani u potpoglavlju 5.1, su identični za obje predložene metode.

Na isti način su realizirane i metode HBGEI-i i HBGEI-f s tom razlikom da su u odnosu na dvije predložene metode umjesto GEI slika upotrebljavane BGEI slike.

5.2.1 Visina – GEI integracija (HGEI-i)

U predloženoj metodi HGEI-i značajka visine se dodaje kao jedna od značajki uz značajke GEI slika te se vrši integrirana klasifikacija. Prikaz koraka kod HGEI-i metode je ilustriran na slici 5.7.

Tijekom izvršavanja HGEI-i metode prvo se izlučuju značajka visine te GEI značajke. Tijek izlučivanja značajki prikazan je na slici 5.1, a opis postupaka je u potpoglavlju 5.1. GEI značajke dobivene nakon postupka smanjenja dimenzionalnosti spajaju se sa značajkom visina osobe u vektor značajki koji predstavlja ulazne podatke za kNN algoritam klasifikacije (može se primijeniti i neki drugi algoritam klasifikacije, primjerice, SVM).



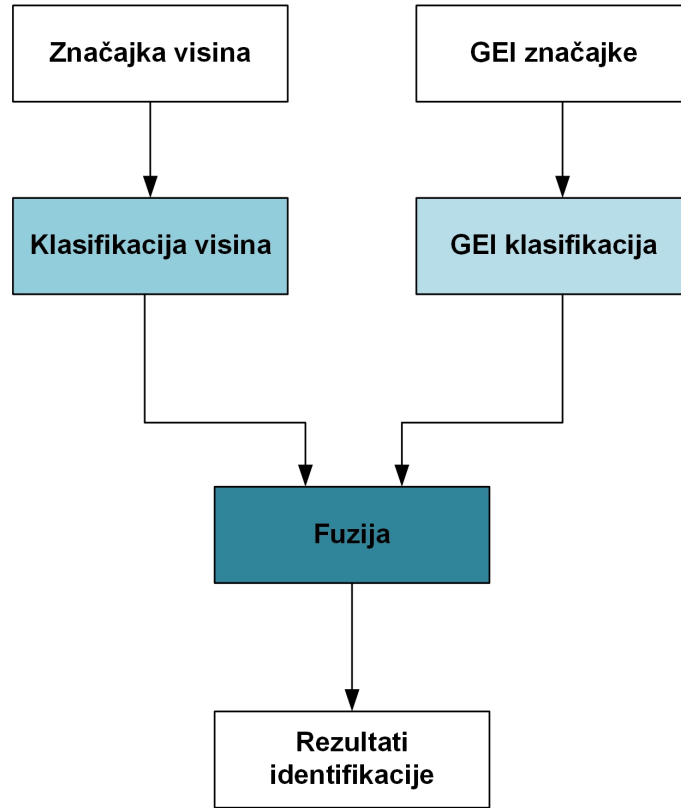
Slika 5.7: Koraci obrade pri integraciji GEI značajki i visine osobe

5.2.2 Visina – GEI fuzija (HGEI-f)

Za razliku od metode HGEI-i gdje se značajka visine pridodaje GEI značajkama te se zajedno koriste kao ulaz algoritmu za klasifikaciju, kod metode HGEI-f značajke se razmatraju odvojeno. Značajka visina osobe i GEI značajke zasebno se koriste za klasifikaciju te se za svaki uzorak radi jedna predikcija. Prikaz koraka kod metode HGEI-f je na slici 5.8.

Tijekom realizacije metode HGEI-f potrebno je, kao i u slučaju metode HGEI-i, prvo izlučiti značajke koje se upotrebljavaju. I u ovom slučaju potrebno je izlučiti značajku visine za svaku osobu te GEI značajke. Tijek izlučivanja značajki isti je kao i kod metode HGEI-i te je prikazan na slici 5.1 (opis postupaka u potpoglavlju 5.1).

Kada su dostupne značajka visina osobe i GEI značajke one se zasebno koriste kao ulazni podatci kNN algoritmu za klasifikaciju. Klasifikacija se može obavljati i nekim drugim algoritmom, primjerice, SVM. Nakon klasifikacije svakog tipa značajki provodi se fuzija dobivenih rezultata. Kako bi se odredili optimalni parametri za fuziju, metoda koja je zasnovana na *alfa-integraciji* (engl. Alpha-Integration) je upotrebljavana, što je predstavljeno i detaljno opisano u radu [85].



Slika 5.8: Koraci obrade pri fuziji GEI značajki i visine osobe

Ukoliko postoje dvije klase podataka (engl. The Two-Class Problem) radi se o problemu koji predstavlja test između hipoteza H_1 i H_0 . Za dati skup detektora D koji rade na tim hipotezama dobije se skup rezultata s_i , $i = 1, \dots, D$. Rezultati su u rasponu od 0 do 1 gdje veća vrijednost s_i odgovara većoj izglednosti hipoteze H_1 . Cilj alfa-integracije je optimalno integrirati dobivene rezultate u jedinstveni rezultat s_α što je opisano sljedećom formulom [85]:

$$s_\alpha(s = [s_1 \dots s_D]^T) = \begin{cases} \left(\sum_{i=1}^D \omega_i \cdot s_i^{\frac{1-\alpha}{2}} \right)^{\frac{2}{1-\alpha}}, & \alpha \neq 1 \\ \exp \left(\sum_{i=1}^D \omega_i \cdot \log(s_i) \right), & \alpha = 1 \end{cases} \quad (5.14)$$

gdje su α i ω_i parametri koji se optimiziraju.

Ukoliko se radi o problemu s više klasa može se izvršiti modificiranje alfa-integracije s dvije klase. Modificiranje navedene alfa-integracije s dvije klase u generalni slučaj ($K \geq 2$) može se dobiti radeći alfa-integraciju odvojeno na rezultatima koji odgovaraju svakoj klasi. Ukoliko postoji K klasa gdje je $k = 1, \dots, K$ i D klasifikatora, i -ti klasifikator će rezultirati vektorom rezultata $s_i = [s_{1i}, \dots, s_{Ki}]^T$, $i = 1, \dots, D$. Ukoliko se pretpostavi

da su rezultati normalizirani na jediničnu sumu $\sum_{k=1}^K s_{ki} = 1$ vektor identifikator istinske klase je definiran kao $y = [y_1, \dots, y_K]^T$ gdje je y_k jednako 1 ako je istinska klasa k i 0 ako nije. Parametri za integriranje rezultata koji odgovaraju klasi k su α_k i ω_{ki} , $i = 1, \dots, D$. Za dati skup rezultata $S = [s_1, \dots, s_D]$ integracijska funkcija (formula 5.14) se može primijeniti direktno na svaku klasu [85]:

$$s_{\alpha_k}(r_k) = \begin{cases} \left(\sum_{i=1}^D \omega_{ki} \cdot s_{ki}^{\frac{1-\alpha_k}{2}} \right)^{\frac{2}{1-\alpha_k}}, & \alpha_k \neq 1 \\ \exp \left(\sum_{i=1}^D \omega_{ki} \cdot \log(s_{ki}) \right), & \alpha_k = 1 \end{cases} \quad k = 1, \dots, K \quad (5.15)$$

gdje je r_k^T k -ti red matrice S . Na ovaj način problem više klasa s K klasa je konvertiran u K odvojenih problema s dvije klase.

Implementacijski detalji alfa-integracije primijenjene u HGEI-f metodi opisani su u prilogu A.3.

6 EKSPERIMENTALNI REZULTATI

Cilj eksperimentalnog istraživanja bio je provjeriti i analizirati učinkovitost predloženih pristupa, odnosno predloženih metoda, upotrebom većeg skupa podataka koji sadržava snimke osoba u hodu. S obzirom da predložene metode upotrebljavaju značajke izlučene iz silueta osoba, a iste se s RGB-D kamerama mogu izlučiti iz RGB ili dubinskih slika, u prvom dijelu ovog poglavlja opisana je analiza silueta izlučenih iz RGB i dubinskih slika u kojoj su prikazani primjeri dobivenih silueta, uz naglasak na poteškoće koje se mogu javiti kada se siluete izlučuju iz RGB ili dubinske slike.

U drugom dijelu ovog poglavlja prikazani su rezultati testiranja predloženih metoda na skupu podataka TUM-GAID. Za razliku od vlastitog skupa podataka, TUM-GAID skup podataka pogodan je za testiranje predloženih metoda jer sadržava velik broj osoba u hodu s različitim varijacijama hoda kao što su normalan hod, hod s nošenjem ruksaka te varijacije u cipelama. Upotrebom skupa podataka TUM-GAID mogao se analizirati utjecaj, primjerice, veličine populacije ljudi ili različite razdiobe visina na rezultate identifikacije osoba. Jednako tako, ovaj skup podataka često je upotrebljavan od strane drugih autora pa se buduća istraživanja mogu jednostavno usporediti s predloženim metodama u ovoj disertaciji. S tim u vezi, drugi dio ovog poglavlja prikazuje eksperimentalne postavke koje su upotrebljavane tijekom istraživanja, mjere uspješnosti koje su upotrebljavane te, na kraju, dobivene rezultate.

6.1 Analiza silueta izlučenih iz RGB i dubinskih slika

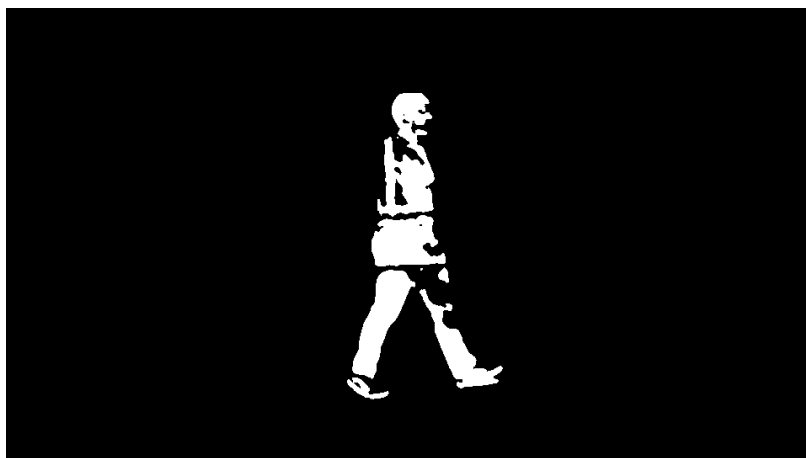
Tijekom istraživanja izvedena je kvalitativna analiza silueta dobivenih iz RGB i dubinskih slika. U prošlosti su RGB slike dominantno upotrebljavane za izlučivanje silueta osoba jer su RGB kamere bile zastupljenije u odnosu na RGB-D kamere. Međutim, u zadnje se vrijeme sve više upotrebljavaju i RGB-D kamere pa se siluete osoba mogu izlučiti iz dubinskih ili RGB slika. Dubinska slika može biti korisna i za dobivanje dodatnih značajki (primjerice, visine osobe ili dužine koraka). Za svih 14 osoba iz vlastitog skupa podataka izlučene su siluete osoba koristeći RGB i dubinske slike s ciljem njihove kvalitativne analize. Također su analizirane i siluete iz skupa podataka Casia Dataset B kao i siluete iz skupa podataka TUM-GAID.

Izlučivanje silueta iz RGB slika ponekad može biti problematično. Često se događa

da, primjerice, na RGB slikama postoji sličnost između objekta i pozadine što otežava postupak oduzimanja pozadine i dovodi do toga da se dobiva silueta koja nije cjelovita. Primjer siluete osobe gdje je izražen ovaj problem prikazan je na slici 6.1 (b). Silueta je izlučena upotrebljavajući metodu GMM za oduzimanje pozadine i metodu binarnog ograničavanja, a postupak je detaljnije opisan u prilogu A.1. Na slici 6.1 prikazan je primjer RGB slike i dobivene siluete osobe na vlastitom skupu podataka.



(a) RGB slika



(b) Dobivena silueta

Slika 6.1: Primjer dobivene siluete iz RGB slike na vlastitom skupu podataka

Ako se pogleda primjer na slici 6.1 (b), vidljivo je da na silueti nedostaju određena područja (u predjelu nogu i gornjeg dijela tijela), odnosno da ona nije cjelovita. Razlog tome, najvećim dijelom, u ovom je slučaju sličnost između boje pozadine i boje odjeće koju nosi osoba. Problem se naziva kamuflaža boje (engl. Color Camouflage). Kamuflaža boje nije jedini problem koji se može javiti tijekom upotrebe RGB slika u svrhu dobivanja silueta osobe. Također se mogu javiti i, primjerice, problemi zbog promjene

osvjetljenja ili objekata koji se kreću u pozadini, čime se mijenja pozadina [58].

Problem kamuflaže boje nije toliko izražen kada se upotrebljavaju dubinske slike, a siluete koje se dobiju upotrebom dubinskih slika uglavnom su cjelovite. Ako se pogleda primjer prikazan na slici 6.2 (b), vidljivo je da je silueta osobe dobivena iz dubinske slike u ovom slučaju cjelovita, tj. nema nedostajućih područja.



(a) Dubinska slika



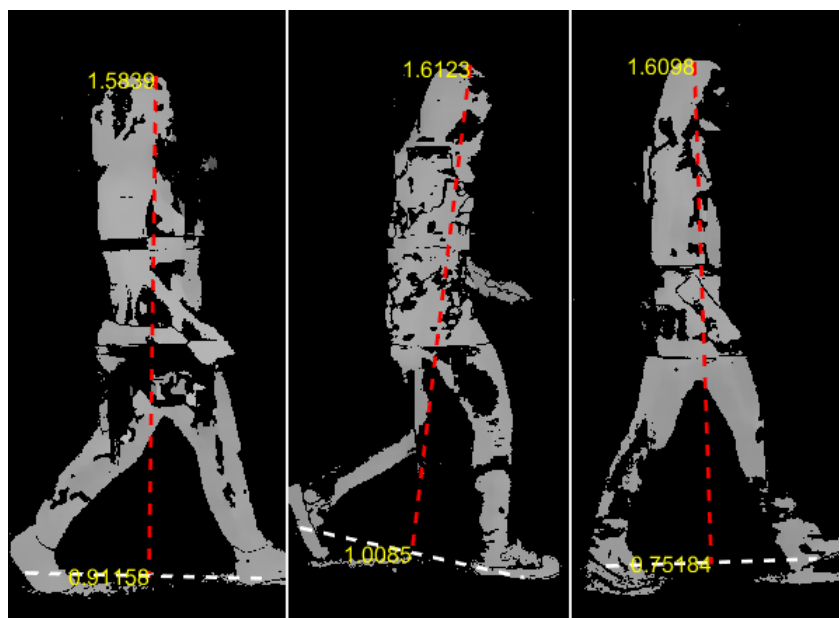
(b) Dobivena silueta

Slika 6.2: Primjer dobivene siluete iz dubinske slike [120]

I kad je riječ o dubinskim slikama, mogu se također javiti određeni problemi koji utječu na izlučivanje silueta (ili, općenito, nekog drugog objekta) kao što su, primjerice, kamuflaža dubine (engl. Depth Camouflage) kada su objekti prednje strane vrlo blizu pozadini u dubini pa osjetilo daje iste vrijednosti dubine za prednju stranu i pozadinu,

što otežava segmentaciju koja je zasnovana samo na dubini. Probleme stvaraju i dubinske sjene, zrcalni objekti u sceni koji utječu na refleksiju IR zraka osjetila, kada su objekti pozadine ili prednje strane preblizu, odnosno predaleko od osjetila pa osjetilo ne može izmjeriti dubinu [58].

Još jedan primjer gdje je dubinska slika korisna prikazan je na slici 6.3. Radi se o primjeru iz vlastitog skupa podataka gdje su visina osobe i dužina koraka izračunati kao udaljenost između ručno zabilježenih točaka na silueti osobe upotrebljavajući naredbu *norm* u Matlabu [81]. Na slici se vidi da se, iako se radi o siluetama koje imaju nedostataka, mogu izlučiti značajke kao što su visina osobe i dužina koraka. Izlučivanje dodatnih značajki je korisno jer iste se mogu upotrebljavati kao značajke za identifikaciju osoba, samostalno ili u kombinaciji s drugim dostupnim značajkama.



Slika 6.3: Primjer određivanja visine osobe i dužine koraka

6.2 Identifikacija osoba

U ovom potpoglavlju opisane su upotrebljavane eksperimentalne postavke, mjere uspješnosti kao i dobiveni rezultati za predložene metode i metode koje su upotrebljavane za usporedbu.

6.2.1 Eksperimentalne postavke

Rezultati klasifikacije predloženim metodama, a prvenstveno točnost klasifikacije, ovise o veličini populacije ljudi. Jednako tako, uz broj osoba, točnost klasifikacije ovisi i o razdiobi visina osoba. Stoga su za potrebe testiranja predloženih metoda upotrebljavane različite veličine populacija te različite razdiobe visina. Budući da su u skupu podataka TUM-GAID dostupne i različite varijacije hoda kao što je normalan hod, hod uz nošenje dodataka kao što je ruksak te varijacije u cipelama, dodatno se testirao i utjecaj tih varijacija. U ovom dijelu disertacije opisane su kreirane GEI i BGEI slike koje su upotrebljavane u nekim od testiranih metoda, upotrijebljene razdiobe visina te postavke koje su upotrebljavane tijekom eksperimenta.

6.2.1.1 Opis kreiranih GEI i BGEI slika

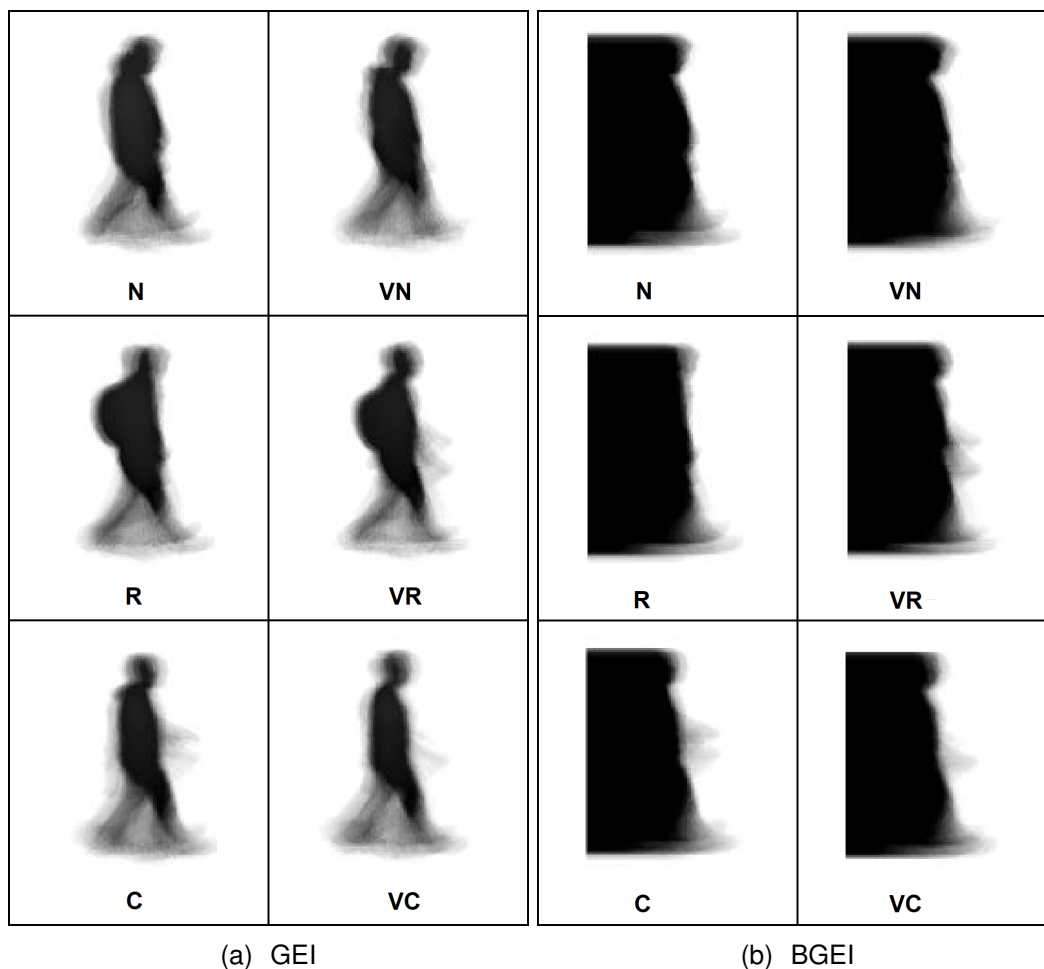
Za kreiranje GEI i BGEI slika upotrebljavane su siluete koje su izlučene iz dubinskih slika. Postupak izlučivanja silueta kao i ostali koraci koje je potrebno realizirati da bi se dobile GEI slike opisani su detaljnije u potpoglavlju 5.1. U slučaju kreiranja BGEI slika postupak je sličan kao kod GEI slika s tom razlikom da se izlučene siluete koje se upotrebljavaju tijekom kreiranja BGEI slika moraju popuniti na način da se pronađe prednji najizbočeniji piksel i od njega se popunjava ostatak slike [92]. Na slici 6.4 prikazane su kreirane GEI i BGEI slike. Kreirano je ukupno 3370 GEI i BGEI slika, što je detaljnije prikazano u tablici 6.1. Razlučivost navedenih slika upotrijebljenih u eksperimentu jest 88 x 128.

Tablica 6.1: Broj kreiranih GEI i BGEI slika za skup podataka TUM-GAID

Tip	GEI	BGEI	Tip	GEI	BGEI
Normalno	1830	1830	Vrijeme normalno	192	192
Ruksak	610	610	Vrijeme ruksak	64	64
Cipele	610	610	Vrijeme cipele	64	64

6.2.1.2 Razdiobe visina osoba

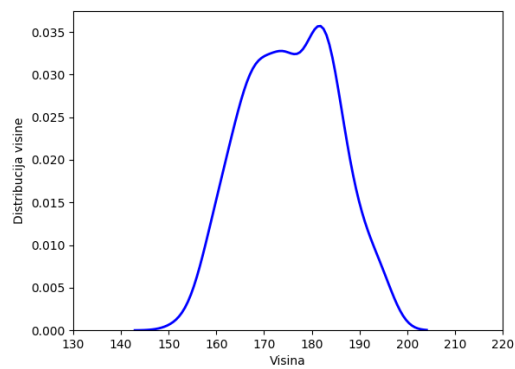
U klasifikacijskim metodama koje koriste značajku visina osobe točnost identifikacije mijenja se ovisno o veličini populacije i razdiobi visina te populacije. Kako bi se analizirao učinak značajke visina osobe na rezultate klasifikacije, simulirano je ukupno



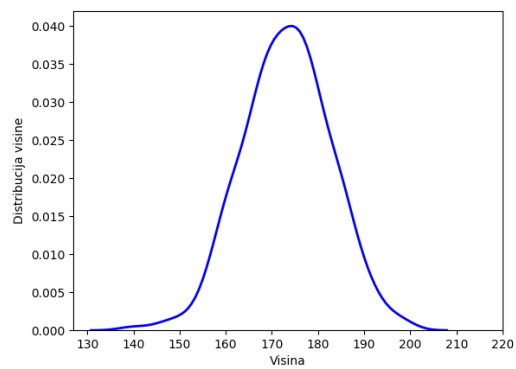
Slika 6.4: Prikaz dobivenih GEI i BGEI slika

pet razdioba visina populacije prema pretpostavljenom interesu za neke određene primjene. Te razdiobe nazvane su: *Stvarne*, *Gauss (normalna razdioba)*, *Odrasli*, *Godine* i *Slične*. Prikaz je na slici 6.5.

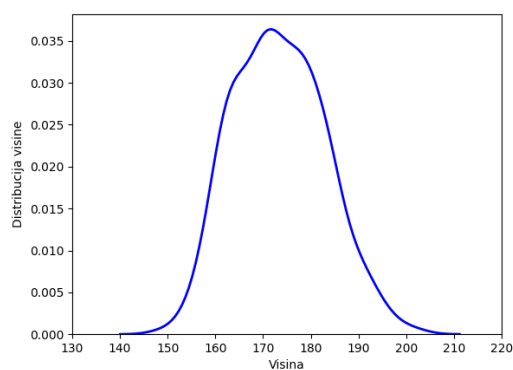
U slučaju razdiobe *Stvarne*, razdioba visina odgovara razdiobi stvarnih visina osoba u skupu podataka TUM-GAID. Kod *Gauss* razdiobe visina radi se o čistoj Gaussovoj razdiobi, a upotrijebljena je radi jednostavnog testiranja i usporedivosti. U provedenim eksperimentima odabrana prosječna visina osobe i odgovarajuća standardna devijacija iznosile su 173,2 cm i 9,5 cm. Kod razdiobe *Odrasli* radi se o razdiobi visina koja predstavlja populaciju odraslih osoba, dok se u slučaju *Godine* radi o razdiobi visina osoba svih godina. *Slične* predstavlja razdiobu visina osoba sličnih visina. Pogreška procjene visina simulirana je u skladu s rezultatima prikazanim u [44]. Na procijenjenu značajku visine svakog uzorka dodana je slučajna pogreška od $\pm 0,5 \%$ njegove vrijednosti.



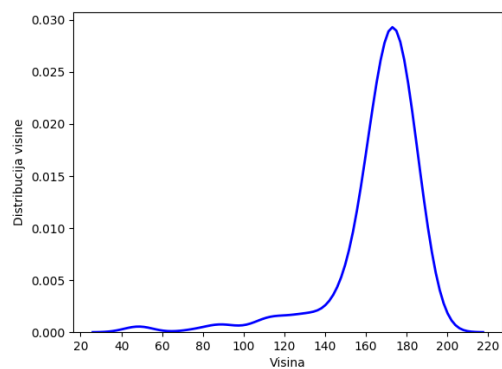
(a) Stvarne



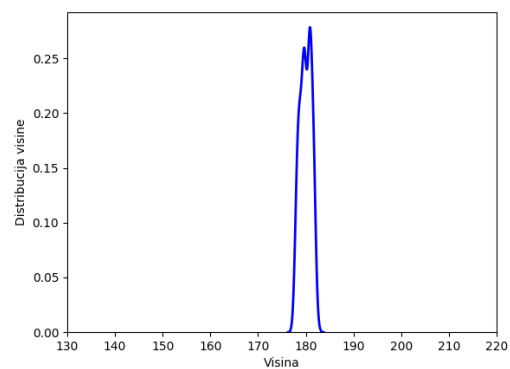
(b) Gauss



(c) Odrasli



(d) Godine



(e) Slične

Slika 6.5: Razdiobe visina upotrebljavane u eksperimentima

6.2.1.3 Klasifikacija

Tijekom istraživanja korišten je kNN algoritam klasifikacije. Odabir ovog algoritma rezultirao je s obzirom na usporedbu s algoritmom klasifikacije potpornim vektorima (engl. Support Vector Classification, SVC) [27], algoritmom klasifikacije slučajnih šuma (engl. Random Forest Classification, RFC) [51] i algoritmom Gaussov naivni

Bayes (engl. Gaussian Naive Bayes, GNB) [60].

Usporedba navedenih algoritama je realizirana upotrebom značajke visina osobe, GEI značajki nakon smanjenja dimenzionalnosti kao i spojenih značajki metode HGEI-i na populaciji ljudi koju čini 305 osoba iz skupa podataka TUM-GAID. Upotrebljavane su visine iz razdiobe visina *Stvarne*. Algoritam kNN je imao bolje rezultate u odnosu na ostale uspoređene algoritme. Rezultati usporedbe algoritama s GEI značajkama nakon smanjenja dimenzionalnosti prikazani su u tablici 6.2.

Tablica 6.2: Usporedba klasifikacijskih algoritama koristeći GEI značajke

Algoritam	Točnost	Preciznost	Odziv	F1 mjera	Kapa
kNN	0.97	0.97	0.97	0.96	0.97
SVC	0.85	0.90	0.85	0.87	0.85
RFC	0.95	0.96	0.95	0.94	0.95
GNB	0.81	0.84	0.81	0.80	0.81

Upotrebljavane su mjere uspješnosti *Točnost* (engl. *Accuracy*), *Preciznost* (engl. *Precision*), *Odziv* (engl. *Recall*), *F1 mjera* (engl. *F1 Score*, *F Score*, *F Measure*) i *Kapa* (engl. *Kappa*) koje su detaljnije opisane u potpoglavlju 6.2.2. Tijekom korištenja navedenih algoritama za usporedbu skup podataka podijeljen je na podskupove za trening i test, slično kao kod [32]. Podskup za trening sastojao se od prvih četiriju GEI silueta iz varijacije *normalno*, tj. gdje se osoba kreće normalno, dok je testni podskup sadržavao sve ostale GEI siluete: dvije iz varijacije *normalno*, dvije iz varijacije *ruksak* i dvije iz varijacije *cipele* za svih 305 osoba u skupu podataka. Podskup za test također se sastojao od šest varijacija *vrijeme normalno*, dvije za *vrijeme ruksak* te dvije varijacije *vrijeme cipele* za svaku od 32 osobe u podskupu snimanom drugi put. Skupovi za trening i test balansirani su, tj. u oba slučaja upotrebljavan je isti broj primjeraka od svake klase.

Također su analizirani i pristupi temeljeni na neuronskim mrežama i dubokom učenju [1], [50], [124], [127]. S upotrjebljenim mjerama uspješnosti pristupi temeljeni na konvolucijskim neuronskim mrežama [1], [50] i generativnim kontradiktornim mrežama (engl. Generative Adversarial Networks, GAN) [124] primijenjeni na Casia Dataset B skupu podataka imali su slične rezultate onima koji su na istom skupu podataka dobiveni upotrebom kNN algoritma.

Nakon što je odabran kNN algoritam klasifikacije, provedeni su eksperimenti sa svim testiranim metodama opisanim u narednom potpoglavlju. Eksperimenti su prove-

deni na tri različite skupine populacija ljudi koje se razlikuju u veličini populacije te su definirane kao mala, srednja i velika populacija. Malu populaciju čine 32 osobe, srednju 155 osoba, dok veliku čini 305 osoba. Velika populacija uzeta je kao maksimalan broj dostupnih osoba u korištenom skupu podataka. Druge dvije veličine populacija odabrane su u skladu s načinom odabira koji su primjenjivali autori u [32] [33]. U velikoj populaciji izvedena je identifikacija na 305 osoba; dakle, postoji 305 klasa klasifikacije. U ostala dva slučaja s malom i srednjom populacijom ima 155 klasa, odnosno 32 klase.

6.2.1.4 Testirane metode

Tijekom istraživanja testirane su predložene metode kao i metode za usporedbu. Metode koje su korištene za usporedbu su metoda koja koristi isključivo značajku visina osobe za klasifikaciju, GEI [28] metoda koja koristi isključivo značajke dobivene iz GEI slika te BGEI [92] metoda koja koristi BGEI značajke dobivene iz BGEI slika.

Pored predloženih metoda, HGEI-i i HGEI-f, koje kombiniraju značajke visina osobe i GEI značajke, kako je to opisano u poglavlju 5, dodatno je za usporedbu upotrijebljena i implementacija predloženih pristupa s BGEI metodom (HBGEI-i i HBGEI-f metoda). Metode HBGEI-i i HBGEI-f realiziraju se kao i opisane dvije predložene metode (HGEI-i i HGEI-f), s tom razlikom da se umjesto GEI slika upotrebljavaju BGEI slike.

U konačnici metode koje su ispitane jesu:

- Visina osobe
- GEI metoda
- BGEI metoda
- visina-GEI integracija (HGEI-i)
- visina-GEI fuzija (HGEI-f)
- visina-BGEI integracija (HBGEI-i)
- visina-BGEI fuzija (HBGEI-f).

6.2.2 Mjere uspješnosti

Tijekom eksperimentalnih istraživanja upotrebljavane su sljedeće mjere uspješnosti testiranih metoda: *Točnost*, *F1 mjera* i *Kapa*.

Točnost predstavlja mjeru uspješnosti koja se odnosi na omjer točno predviđenih

zapažanja u cjelokupnom broju zapažanja. Ova mjera uspješnosti najčešće se upotrebljava za provjeru uspješnosti predloženih pristupa odnosno metoda u ovom području identifikacije osoba na temelju hoda. Kada postoje samo dvije klase klasifikacije (naziva se binarna klasifikacija) opisuje se sljedećom formulom:

$$Točnost = \frac{TP + TN}{TP + TN + FP + FN} \quad (6.1)$$

pri čemu su: TP – *istinito pozitivan* (odnosi se na točno predviđene pozitivne vrijednosti, što, drugim riječima, znači da u ovom slučaju algoritam klasifikacije točno predvidi pozitivnu klasu; primjerice, neka postoje dvije klase, *osoba A* i *osoba B* gdje je u slučaju TP osoba A točno identificirana kao osoba A), TN – *istinito negativan* (u ovom slučaju algoritam klasifikacije točno predvidi negativnu klasu; primjerice, osoba B je točno identificirana kao osoba B, a ne kao osoba A), FP – *lažno pozitivan* (u ovom slučaju algoritam klasifikacije netočno predvidi da se radi o pozitivnoj klasi – ustvari se radi o negativnoj; primjerice, osoba B je pogrešno identificirana kao osoba A), FN – *lažno negativan* (u ovom slučaju algoritam klasifikacije netočno predvidi da se radi o negativnoj klasi – ustvari se radi o pozitivnoj klasi; primjerice, osoba A je pogrešno identificirana kao osoba B).

U slučaju kada postoji više klasa (više od dvije) klasifikacije računa se točnost koja se opisuje formulom:

$$Točnost = \frac{\sum_{i=1}^n \frac{TP_i + TN_i}{TP_i + TN_i + FP_i + FN_i}}{n} \quad (6.2)$$

pri čemu su: TP_i , TN_i , FP_i i FN_i vezani za specifičnu klasu (primjerice, za klasu *Osoba A*), a n je ukupan broj klasa.

$F1$ mjera odnosi se na mjeru uspješnosti koja u obzir uzima *Preciznost* (engl. *Precision*) i *Odziv* (engl. *Recall*) te se opisuje sljedećom formulom:

$$F1 \text{ mjera} = 2 \cdot \frac{Preciznost \cdot Odziv}{Preciznost + Odziv} \quad (6.3)$$

Preciznost se odnosi na omjer točno predviđenih pozitivnih zapažanja i ukupnog broja pozitivnih zapažanja. Opisuje se sljedećom formulom:

$$Preciznost = \frac{TP}{TP + FP} \quad (6.4)$$

U slučaju kada postoji više klasa opisuje se formulama:

$$Preciznost_M = \frac{\sum_{i=1}^n \frac{TP_i}{TP_i + FP_i}}{n} \quad (6.5)$$

$$Preciznost_\mu = \frac{\sum_{i=1}^n TP_i}{\sum_{i=1}^n (TP_i + FP_i)} \quad (6.6)$$

Može se računati na dva načina, na *makro* (M) (formula 6.5) i na *mikro* (μ) (formula 6.6) razini (isto važi i u slučaju kada se računa *Odziv*). U načinu *makro* (oznaka $Preciznost_M$) se računa *Preciznost* za svaku klasu te se dobivena suma dijeli s ukupnim brojem klasa, dok se u načinu *mikro* (oznaka $Preciznost_\mu$) uzimaju kumulativne vrijednosti za TP , TN , FP i FN tijekom računanja. Drugim riječima, kod načina *makro* će se izračunati vrijednosti neovisno za svaku klasu te nakon toga prosječna vrijednost dok se u načinu *mikro* dobije vrijednost od TP , TN , FP i FN svake pojedine klase. *Makro* način tretira sve klase jednako i pogodan je za sveukupne provjere pristupa koristeći skupove podataka, dok je *mikro* način pogodan za analizu u slučaju nebalansiranih klasa (primjerice, kada postoji više primjeraka jedne klase u odnosu na druge klase).

Primjerice, $Preciznost_M$ i $Preciznost_\mu$ mogu se napisati i sljedećim formulama:

$$Preciznost_M = \frac{Preciznost_1 + Preciznost_2 + \dots + Preciznost_n}{n} \quad (6.7)$$

$$Preciznost_\mu = \frac{TP_1 + TP_2 + \dots + TP_n}{TP_1 + TP_2 + \dots + TP_n + FP_1 + FP_2 + \dots + FP_n} \quad (6.8)$$

Odziv se odnosi na omjer točno predviđenih pozitivnih zapažanja i svih pozitivnih zapažanja. Opisuje se formulom:

$$Odziv = \frac{TP}{TP + FN} \quad (6.9)$$

U slučaju više klasa opisuje se formulama:

$$Odziv_M = \frac{\sum_{i=1}^n \frac{TP_i}{TP_i + FN_i}}{n} \quad (6.10)$$

$$Odziv_\mu = \frac{\sum_{i=1}^n TP_i}{\sum_{i=1}^n (TP_i + FN_i)} \quad (6.11)$$

Kapa ili kako se još naziva Cohenova kapa (engl. Cohen's Kappa) odnosi se na mjerenje pouzdanosti stupnja slaganja između procjenitelja (ili se još mogu nazvati, primjerice, suci, promatrači itd.), gdje svaki procjenitelj klasificira N stavki odnosno određeni podatak u C međusobno isključivih kategorija. Do navedene pouzdanosti dolazi kada procjenitelji daju isti rezultat procjene za istu stavku podatka. Primjerice, ako postoje dva procjenitelja odgovorna za mjerenje neke varijable u okviru neke ljestvice kategorija, potrebno je odrediti slažu li se ti procjenitelji. Definira se sljedećom formulom:

$$\kappa = \frac{p_o - p_e}{1 - p_e} = 1 - \frac{1 - p_o}{1 - p_e} \quad (6.12)$$

pri čemu su: p_o relativno slaganje promatranja između procjenitelja i p_e hipotetička vjerojatnoća šanse slaganja. Navedene p_o i p_e definirane su sljedećim formulama:

$$p_o = \frac{\text{Broj slaganja promatranja između procjenitelja}}{\text{Ukupan broj promatranja}} \quad (6.13)$$

$$p_e = \frac{1}{N^2} \sum_C n_{C1} n_{C2} \quad (6.14)$$

pri čemu su: C kategorije, N broj promatranja odnosno stavki za kategoriziranje, a n_C broj puta odnosno koliko puta je određeni procjenitelj predvidio kategoriju C .

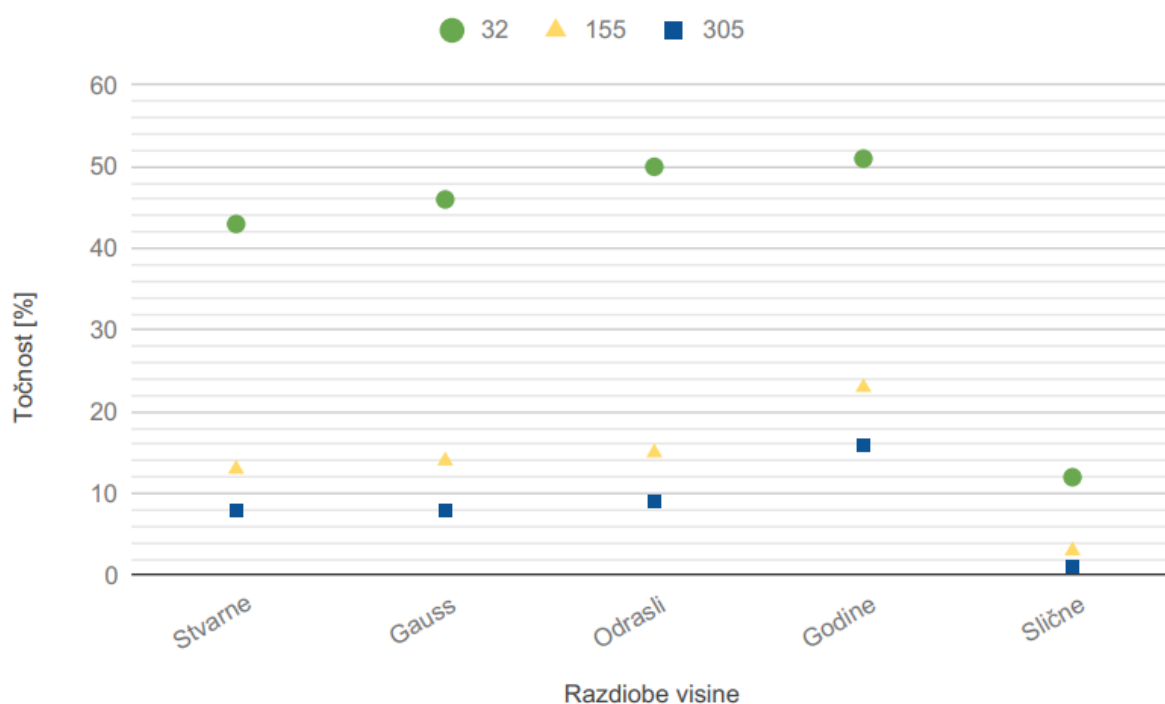
6.2.3 Rezultati

U tablicama 6.3, 6.4, 6.5, 6.6 i 6.7 prikazani su rezultati klasifikacije testiranih metoda na skupu podataka TUM-GAID sa *Stvarne*, *Slične*, *Gauss*, *Odrasli* i *Godine* razdiobama visina. Upotrebljavane su varijacije *normalno*, *ruksak*, *cipele*, *vrijeme normalno*, *vrijeme ruksak* i *vrijeme cipele* za malu, srednju i veliku populaciju ljudi. Jednako tako,

prikazane su *Točnost*, *F1 mjera* i *Kapa* mjere uspješnosti s algoritmom klasifikacije kNN.

Prikazane tablice (6.3, 6.4, 6.5, 6.6 i 6.7) s dobivenim rezultatima organizirane su tako da je za svaku varijaciju iz skupa (*normalno*, *ruksak*, *cipele*, *vrijeme normalno*, *vrijeme ruksak* i *vrijeme cipele*) odvojeno prikazan rezultat. Dobiveni rezultati prikazani su za samostalno korištenje značajke visina osobe i primijenjene metode (GEI, BGEI, HGEI-i, HGEI-f, HBGEI-i, HBGEI-f). Rezultati su odvojeni za tri upotrijebljene veličine populacije i za upotrijebljene mjere uspješnosti. Za svaku metodu i samostalnu upotrebu značajke visina rezultati su složeni horizontalno, pri čemu se prve tri vrijednosti odnose na malu veličinu populacije (za tri navedene mjere uspješnosti, *Točnost*, *F1 mjera* i *Kapa*), potom se sljedeće tri vrijednosti odnose na srednju populaciju te zadnje tri vrijednosti na veliku populaciju, a za iste mjere uspješnosti.

Efekti koje razdioba visina i veličina populacije imaju na točnost identifikacije kada se samostalno upotrebljava visina osobe kao značajka prikazani su na slici 6.6. Iz navedene se slike može zaključiti da su rezultati očekivano bolji što je manja populacija ljudi. Vidi se da je za malu populaciju točnost identifikacije najveća, a opada kako se populacija ljudi povećava.



Slika 6.6: Rezultati klasifikacije značajke visina za pet različitih razdioba visina i tri različite veličine populacije

Tablica 6.3: Rezultati klasifikacije testiranih metoda na razdiobi visina *Stvarne*

Veličina populacije		32			155			305		
Mjera uspješnosti	Točnost	F1 mjera	Kapa	Točnost	F1 mjera	Kapa	Točnost	F1 mjera	Kapa	
Normalno (N) varijacija										
Visina	0.43	0.4	0.41	0.13	0.11	0.13	0.08	0.07	0.08	
GEI	0.97	0.96	0.97	0.99	0.99	0.99	0.97	0.96	0.97	
HGEI-i	1	1	1	1	1	1	0.99	0.99	0.99	
HGEI-f	0.97	0.97	0.97	0.89	0.88	0.89	0.77	0.74	0.77	
BGEI	0.98	0.98	0.98	0.94	0.93	0.94	0.88	0.86	0.88	
HBGEI-i	1	1	1	0.98	0.98	0.98	0.97	0.96	0.97	
HBGEI-f	0.97	0.96	0.97	0.87	0.86	0.87	0.75	0.72	0.75	
Ruksak (R) varijacija										
Visina	0.43	0.4	0.42	0.13	0.12	0.13	0.08	0.07	0.08	
GEI	0.97	0.97	0.97	0.89	0.86	0.89	0.8	0.78	0.8	
HGEI-i	1	1	1	0.96	0.95	0.96	0.93	0.91	0.93	
HGEI-f	0.96	0.95	0.96	0.83	0.81	0.83	0.67	0.63	0.67	
BGEI	0.84	0.82	0.84	0.74	0.71	0.74	0.63	0.59	0.63	
HBGEI-i	0.96	0.96	0.96	0.9	0.88	0.9	0.86	0.84	0.86	
HBGEI-f	0.93	0.92	0.93	0.8	0.78	0.8	0.65	0.61	0.65	
Cipele (C) varijacija										
Visina	0.42	0.39	0.4	0.13	0.12	0.13	0.08	0.07	0.08	
GEI	0.95	0.94	0.95	0.89	0.88	0.89	0.83	0.81	0.83	
HGEI-i	1	1	1	0.96	0.96	0.96	0.93	0.93	0.93	
HGEI-f	0.96	0.95	0.96	0.84	0.82	0.84	0.69	0.65	0.69	
BGEI	0.91	0.9	0.9	0.78	0.75	0.78	0.7	0.66	0.7	
HBGEI-i	0.98	0.98	0.98	0.92	0.91	0.92	0.87	0.85	0.87	
HBGEI-f	0.94	0.93	0.94	0.81	0.79	0.81	0.66	0.62	0.66	
Vrijeme normalno (VN) varijacija										
Visina	0.43	0.42	0.42	0.13	0.06	0.12	0.07	0.02	0.06	
GEI	0.44	0.38	0.42	0.27	0.16	0.26	0.21	0.09	0.2	
HGEI-i	0.64	0.58	0.62	0.42	0.26	0.41	0.35	0.17	0.34	
HGEI-f	0.68	0.65	0.67	0.4	0.24	0.39	0.25	0.1	0.24	
BGEI	0.35	0.3	0.33	0.23	0.1	0.22	0.2	0.07	0.19	
HBGEI-i	0.67	0.61	0.66	0.42	0.24	0.41	0.33	0.15	0.32	
HBGEI-f	0.65	0.61	0.64	0.41	0.23	0.4	0.3	0.13	0.29	
Vrijeme ruksak (VR) varijacija										
Visina	0.43	0.4	0.41	0.14	0.08	0.13	0.07	0.04	0.06	
GEI	0.34	0.27	0.32	0.17	0.09	0.16	0.11	0.05	0.1	
HGEI-i	0.65	0.58	0.63	0.45	0.32	0.44	0.33	0.2	0.32	
HGEI-f	0.65	0.59	0.64	0.43	0.29	0.42	0.24	0.13	0.23	
BGEI	0.22	0.17	0.19	0.12	0.06	0.12	0.09	0.04	0.09	
HBGEI-i	0.62	0.57	0.61	0.43	0.27	0.42	0.3	0.15	0.29	
HBGEI-f	0.61	0.55	0.60	0.42	0.28	0.41	0.31	0.18	0.30	
Vrijeme cipele (VC) varijacija										
Visina	0.44	0.41	0.42	0.13	0.07	0.12	0.07	0.04	0.07	
GEI	0.33	0.26	0.31	0.19	0.11	0.18	0.17	0.08	0.16	
HGEI-i	0.63	0.57	0.62	0.32	0.2	0.31	0.24	0.13	0.23	
HGEI-f	0.7	0.69	0.69	0.38	0.25	0.37	0.25	0.14	0.24	
BGEI	0.33	0.25	0.31	0.17	0.1	0.16	0.09	0.05	0.09	
HBGEI-i	0.6	0.53	0.59	0.37	0.24	0.36	0.28	0.15	0.27	
HBGEI-f	0.58	0.52	0.56	0.39	0.25	0.38	0.28	0.15	0.27	

Tablica 6.4: Rezultati klasifikacije testiranih metoda na razdiobi visina *Slične*

Veličina populacije		32			155			305		
Mjera uspješnosti	Točnost	F1 mjera	Kapa	Točnost	F1 mjera	Kapa	Točnost	F1 mjera	Kapa	
Normalno (N) varijacija										
Visina	0.12	0.1	0.09	0.03	0.02	0.02	0.01	0.01	0.01	
GEI	0.97	0.96	0.97	0.99	0.99	0.99	0.97	0.96	0.97	
HGEI-i	0.98	0.98	0.98	0.99	0.99	0.99	0.97	0.97	0.97	
HGEI-f	0.82	0.79	0.82	0.57	0.52	0.57	0.42	0.36	0.41	
BGEI	0.98	0.98	0.98	0.94	0.93	0.94	0.88	0.86	0.88	
HBGEI-i	0.98	0.98	0.98	0.95	0.95	0.95	0.9	0.89	0.9	
HBGEI-f	0.82	0.8	0.82	0.54	0.49	0.54	0.39	0.34	0.39	
Ruksak (R) varijacija										
Visina	0.12	0.11	0.1	0.03	0.02	0.02	0.01	0.01	0.01	
GEI	0.97	0.97	0.97	0.89	0.86	0.89	0.8	0.78	0.8	
HGEI-i	0.97	0.96	0.97	0.91	0.89	0.91	0.84	0.81	0.84	
HGEI-f	0.79	0.76	0.78	0.49	0.44	0.49	0.33	0.28	0.33	
BGEI	0.84	0.82	0.84	0.74	0.71	0.74	0.63	0.59	0.63	
HBGEI-i	0.88	0.87	0.88	0.8	0.77	0.8	0.7	0.66	0.7	
HBGEI-f	0.77	0.75	0.76	0.45	0.4	0.44	0.29	0.25	0.29	
Cipele (C) varijacija										
Visina	0.12	0.1	0.09	0.03	0.02	0.02	0.01	0.01	0.01	
GEI	0.95	0.94	0.95	0.89	0.88	0.89	0.83	0.81	0.83	
HGEI-i	0.97	0.96	0.97	0.92	0.91	0.92	0.86	0.85	0.86	
HGEI-f	0.8	0.78	0.8	0.5	0.46	0.5	0.35	0.29	0.35	
BGEI	0.91	0.9	0.9	0.78	0.75	0.78	0.7	0.66	0.7	
HBGEI-i	0.96	0.96	0.96	0.83	0.8	0.83	0.76	0.73	0.75	
HBGEI-f	0.79	0.76	0.78	0.48	0.43	0.47	0.32	0.27	0.32	
Vrijeme normalno (VN) varijacija										
Visina	0.12	0.11	0.09	0.03	0.01	0.03	0.02	0.01	0.02	
GEI	0.44	0.38	0.42	0.27	0.16	0.26	0.21	0.09	0.2	
HGEI-i	0.52	0.45	0.51	0.32	0.18	0.31	0.25	0.11	0.24	
HGEI-f	0.48	0.42	0.46	0.24	0.11	0.23	0.12	0.05	0.12	
BGEI	0.35	0.3	0.33	0.23	0.1	0.22	0.2	0.07	0.19	
HBGEI-i	0.42	0.36	0.4	0.28	0.13	0.27	0.24	0.09	0.23	
HBGEI-f	0.47	0.44	0.45	0.19	0.09	0.18	0.11	0.04	0.1	
Vrijeme ruksak (VR) varijacija										
Visina	0.12	0.11	0.09	0.03	0.02	0.02	0.02	0.01	0.02	
GEI	0.34	0.27	0.32	0.17	0.09	0.16	0.11	0.05	0.1	
HGEI-i	0.44	0.36	0.43	0.26	0.14	0.25	0.15	0.07	0.14	
HGEI-f	0.44	0.39	0.43	0.23	0.14	0.22	0.12	0.06	0.12	
BGEI	0.22	0.17	0.19	0.12	0.06	0.12	0.09	0.04	0.09	
HBGEI-i	0.33	0.28	0.31	0.18	0.1	0.17	0.13	0.06	0.13	
HBGEI-f	0.47	0.41	0.45	0.18	0.1	0.18	0.09	0.05	0.09	
Vrijeme cipele (VC) varijacija										
Visina	0.12	0.1	0.09	0.03	0.02	0.03	0.02	0.01	0.01	
GEI	0.33	0.26	0.31	0.19	0.11	0.18	0.17	0.08	0.16	
HGEI-i	0.4	0.32	0.38	0.23	0.15	0.22	0.2	0.11	0.19	
HGEI-f	0.46	0.41	0.45	0.23	0.14	0.22	0.12	0.07	0.12	
BGEI	0.33	0.25	0.31	0.17	0.1	0.16	0.09	0.05	0.09	
HBGEI-i	0.41	0.33	0.39	0.22	0.11	0.21	0.15	0.08	0.14	
HBGEI-f	0.41	0.36	0.39	0.17	0.1	0.16	0.1	0.05	0.1	

Tablica 6.5: Rezultati klasifikacije testiranih metoda na razdiobi visina *Gauss*

Veličina populacije		32			155			305		
Mjera uspješnosti	Točnost	F1 mjera	Kapa	Točnost	F1 mjera	Kapa	Točnost	F1 mjera	Kapa	
Normalno (N) varijacija										
Visina	0.46	0.42	0.44	0.14	0.12	0.13	0.08	0.07	0.08	
GEI	0.97	0.96	0.97	0.99	0.99	0.99	0.97	0.96	0.97	
HGEI-i	1	1	1	1	1	1	0.99	0.99	0.99	
HGEI-f	0.99	0.99	0.99	0.91	0.89	0.91	0.80	0.78	0.80	
BGEI	0.98	0.98	0.98	0.94	0.93	0.94	0.88	0.86	0.88	
HBGEI-i	1	1	1	0.98	0.97	0.98	0.97	0.96	0.97	
HBGEI-f	0.99	0.99	0.99	0.90	0.89	0.90	0.78	0.75	0.78	
Ruksak (R) varijacija										
Visina	0.44	0.41	0.43	0.14	0.12	0.13	0.08	0.07	0.07	
GEI	0.97	0.97	0.97	0.89	0.86	0.89	0.80	0.78	0.80	
HGEI-i	1	1	1	0.98	0.97	0.98	0.95	0.95	0.95	
HGEI-f	0.98	0.98	0.98	0.86	0.84	0.86	0.72	0.68	0.71	
BGEI	0.84	0.82	0.84	0.74	0.71	0.74	0.63	0.59	0.63	
HBGEI-i	1	0.99	1	0.92	0.91	0.92	0.84	0.82	0.84	
HBGEI-f	0.98	0.97	0.98	0.85	0.83	0.85	0.69	0.65	0.69	
Cipele (C) varijacija										
Visina	0.45	0.42	0.43	0.14	0.12	0.13	0.08	0.07	0.07	
GEI	0.95	0.94	0.95	0.89	0.88	0.89	0.83	0.81	0.83	
HGEI-i	0.99	0.99	0.99	0.98	0.98	0.98	0.95	0.94	0.95	
HGEI-f	0.98	0.98	0.98	0.86	0.84	0.86	0.73	0.70	0.73	
BGEI	0.91	0.90	0.90	0.78	0.75	0.78	0.70	0.66	0.70	
HBGEI-i	0.99	0.99	0.99	0.93	0.92	0.93	0.89	0.87	0.89	
HBGEI-f	0.98	0.98	0.98	0.86	0.84	0.86	0.71	0.67	0.71	
Vrijeme normalno (VN) varijacija										
Visina	0.45	0.44	0.43	0.14	0.06	0.13	0.08	0.03	0.08	
GEI	0.44	0.38	0.42	0.27	0.16	0.26	0.21	0.09	0.20	
HGEI-i	0.63	0.56	0.62	0.48	0.29	0.47	0.39	0.19	0.39	
HGEI-f	0.70	0.65	0.69	0.53	0.32	0.52	0.36	0.17	0.36	
BGEI	0.35	0.30	0.33	0.23	0.10	0.22	0.20	0.07	0.19	
HBGEI-i	0.66	0.61	0.65	0.46	0.24	0.45	0.35	0.16	0.34	
HBGEI-f	0.68	0.62	0.67	0.50	0.27	0.49	0.38	0.16	0.37	
Vrijeme ruksak (VR) varijacija										
Visina	0.45	0.42	0.43	0.14	0.08	0.13	0.08	0.04	0.08	
GEI	0.34	0.27	0.32	0.17	0.09	0.16	0.11	0.05	0.10	
HGEI-i	0.59	0.52	0.58	0.45	0.29	0.44	0.35	0.20	0.34	
HGEI-f	0.67	0.60	0.66	0.51	0.35	0.50	0.34	0.19	0.33	
BGEI	0.22	0.17	0.19	0.12	0.06	0.12	0.09	0.04	0.09	
HBGEI-i	0.68	0.63	0.67	0.40	0.24	0.39	0.29	0.16	0.28	
HBGEI-f	0.67	0.61	0.66	0.44	0.27	0.43	0.32	0.18	0.31	
Vrijeme cipele (VC) varijacija										
Visina	0.46	0.43	0.44	0.14	0.08	0.13	0.08	0.04	0.07	
GEI	0.33	0.26	0.31	0.19	0.11	0.18	0.17	0.08	0.16	
HGEI-i	0.64	0.56	0.62	0.41	0.26	0.40	0.34	0.20	0.33	
HGEI-f	0.72	0.68	0.72	0.52	0.36	0.51	0.35	0.21	0.34	
BGEI	0.33	0.25	0.31	0.17	0.10	0.16	0.09	0.05	0.09	
HBGEI-i	0.62	0.54	0.61	0.45	0.28	0.44	0.35	0.19	0.34	
HBGEI-f	0.64	0.57	0.63	0.45	0.29	0.44	0.31	0.16	0.30	

Tablica 6.6: Rezultati klasifikacije testiranih metoda na razdiobi visina *Odrasli*

Veličina populacije		32			155			305		
Mjera uspješnosti	Točnost	F1 mjera	Kapa	Točnost	F1 mjera	Kapa	Točnost	F1 mjera	Kapa	
Normalno (N) varijacija										
Visina	0.50	0.47	0.48	0.15	0.14	0.15	0.09	0.08	0.08	
GEI	0.97	0.96	0.97	0.99	0.99	0.99	0.97	0.96	0.97	
HGEI-i	0.98	0.98	0.98	1	1	1	0.99	0.99	0.99	
HGEI-f	0.97	0.96	0.97	0.92	0.91	0.92	0.82	0.79	0.82	
BGEI	0.98	0.98	0.98	0.94	0.93	0.94	0.88	0.86	0.88	
HBGEI-i	1	1	1	0.99	0.99	0.99	0.96	0.95	0.96	
HBGEI-f	0.98	0.98	0.98	0.90	0.89	0.90	0.80	0.77	0.80	
Ruksak (R) varijacija										
Visina	0.50	0.47	0.48	0.15	0.14	0.15	0.09	0.07	0.08	
GEI	0.97	0.97	0.97	0.89	0.86	0.89	0.80	0.78	0.80	
HGEI-i	0.97	0.97	0.97	0.98	0.98	0.98	0.95	0.94	0.95	
HGEI-f	0.96	0.95	0.95	0.88	0.86	0.88	0.75	0.71	0.74	
BGEI	0.84	0.82	0.84	0.74	0.71	0.74	0.63	0.59	0.63	
HBGEI-i	0.95	0.94	0.95	0.91	0.89	0.91	0.86	0.84	0.86	
HBGEI-f	0.96	0.96	0.96	0.83	0.80	0.83	0.70	0.66	0.70	
Cipele (C) varijacija										
Visina	0.50	0.47	0.48	0.16	0.14	0.15	0.09	0.07	0.08	
GEI	0.95	0.94	0.95	0.89	0.88	0.89	0.83	0.81	0.83	
HGEI-i	0.98	0.98	0.98	0.97	0.96	0.97	0.94	0.93	0.94	
HGEI-f	0.96	0.96	0.96	0.87	0.86	0.87	0.75	0.71	0.75	
BGEI	0.91	0.90	0.90	0.78	0.75	0.78	0.70	0.66	0.70	
HBGEI-i	0.98	0.98	0.98	0.93	0.92	0.93	0.88	0.87	0.88	
HBGEI-f	0.97	0.96	0.96	0.85	0.83	0.85	0.72	0.69	0.72	
Vrijeme normalno (VN) varijacija										
Visina	0.50	0.49	0.49	0.12	0.05	0.12	0.08	0.03	0.07	
GEI	0.44	0.38	0.42	0.27	0.16	0.26	0.21	0.09	0.20	
HGEI-i	0.63	0.57	0.62	0.51	0.31	0.50	0.43	0.21	0.42	
HGEI-f	0.65	0.59	0.64	0.54	0.32	0.53	0.40	0.18	0.39	
BGEI	0.35	0.30	0.33	0.23	0.10	0.22	0.20	0.07	0.19	
HBGEI-i	0.66	0.60	0.65	0.43	0.24	0.42	0.40	0.19	0.39	
HBGEI-f	0.66	0.62	0.65	0.48	0.25	0.47	0.38	0.16	0.37	
Vrijeme ruksak (VR) varijacija										
Visina	0.50	0.47	0.48	0.14	0.08	0.13	0.08	0.04	0.07	
GEI	0.34	0.27	0.32	0.17	0.09	0.16	0.11	0.05	0.10	
HGEI-i	0.62	0.55	0.60	0.43	0.27	0.42	0.37	0.21	0.36	
HGEI-f	0.61	0.54	0.59	0.55	0.38	0.54	0.38	0.22	0.37	
BGEI	0.22	0.17	0.19	0.12	0.06	0.12	0.09	0.04	0.09	
HBGEI-i	0.62	0.57	0.60	0.31	0.18	0.30	0.23	0.12	0.22	
HBGEI-f	0.64	0.61	0.62	0.46	0.30	0.46	0.34	0.19	0.33	
Vrijeme cipele (VC) varijacija										
Visina	0.50	0.47	0.48	0.13	0.07	0.12	0.08	0.04	0.07	
GEI	0.33	0.26	0.31	0.19	0.11	0.18	0.17	0.08	0.16	
HGEI-i	0.64	0.56	0.63	0.52	0.35	0.51	0.42	0.24	0.41	
HGEI-f	0.68	0.62	0.67	0.55	0.39	0.54	0.38	0.23	0.37	
BGEI	0.33	0.25	0.31	0.17	0.10	0.16	0.09	0.05	0.09	
HBGEI-i	0.55	0.49	0.54	0.37	0.21	0.36	0.33	0.17	0.32	
HBGEI-f	0.57	0.51	0.56	0.43	0.27	0.42	0.33	0.18	0.32	

Tablica 6.7: Rezultati klasifikacije testiranih metoda na razdiobi visina *Godine*

Veličina populacije		32			155			305		
Mjera uspješnosti	Točnost	F1 mjera	Kapa	Točnost	F1 mjera	Kapa	Točnost	F1 mjera	Kapa	
Normalno (N) varijacija										
Visina	0.51	0.47	0.49	0.23	0.21	0.22	0.16	0.15	0.16	
GEI	0.97	0.96	0.97	0.99	0.99	0.99	0.97	0.96	0.97	
HGEI-i	1	1	1	1	1	1	0.99	0.99	0.99	
HGEI-f	0.98	0.98	0.98	0.93	0.92	0.93	0.85	0.83	0.85	
BGEI	0.98	0.98	0.98	0.94	0.93	0.94	0.88	0.86	0.88	
HBGEI-i	1	1	1	0.99	0.99	0.99	0.97	0.97	0.97	
HBGEI-f	0.96	0.96	0.96	0.92	0.91	0.92	0.83	0.80	0.83	
Ruksak (R) varijacija										
Visina	0.50	0.47	0.48	0.23	0.21	0.22	0.16	0.15	0.16	
GEI	0.97	0.97	0.97	0.89	0.86	0.89	0.80	0.78	0.80	
HGEI-i	1	1	1	0.98	0.98	0.98	0.96	0.95	0.96	
HGEI-f	0.96	0.96	0.96	0.89	0.87	0.89	0.79	0.76	0.79	
BGEI	0.84	0.82	0.84	0.74	0.71	0.74	0.63	0.59	0.63	
HBGEI-i	0.97	0.96	0.96	0.91	0.90	0.91	0.88	0.87	0.88	
HBGEI-f	0.94	0.94	0.94	0.85	0.83	0.85	0.75	0.72	0.75	
Cipele (C) varijacija										
Visina	0.51	0.48	0.49	0.22	0.20	0.22	0.16	0.15	0.16	
GEI	0.95	0.94	0.95	0.89	0.88	0.89	0.83	0.81	0.83	
HGEI-i	1	1	1	0.98	0.97	0.98	0.95	0.94	0.95	
HGEI-f	0.97	0.97	0.97	0.90	0.89	0.90	0.80	0.78	0.80	
BGEI	0.91	0.90	0.90	0.78	0.75	0.78	0.70	0.66	0.70	
HBGEI-i	0.98	0.98	0.98	0.95	0.94	0.95	0.90	0.88	0.90	
HBGEI-f	0.94	0.93	0.93	0.87	0.85	0.87	0.76	0.73	0.76	
Vrijeme normalno (VN) varijacija										
Visina	0.51	0.50	0.50	0.20	0.08	0.20	0.12	0.04	0.12	
GEI	0.44	0.38	0.42	0.27	0.16	0.26	0.21	0.09	0.20	
HGEI-i	0.69	0.63	0.68	0.52	0.33	0.51	0.43	0.22	0.42	
HGEI-f	0.64	0.57	0.63	0.52	0.31	0.51	0.40	0.19	0.39	
BGEI	0.35	0.30	0.33	0.23	0.10	0.22	0.20	0.07	0.19	
HBGEI-i	0.64	0.59	0.63	0.47	0.25	0.46	0.41	0.19	0.41	
HBGEI-f	0.64	0.59	0.63	0.44	0.25	0.43	0.37	0.16	0.36	
Vrijeme ruksak (VR) varijacija										
Visina	0.50	0.47	0.49	0.20	0.11	0.19	0.13	0.06	0.12	
GEI	0.34	0.27	0.32	0.17	0.09	0.16	0.11	0.05	0.10	
HGEI-i	0.63	0.55	0.62	0.48	0.33	0.47	0.38	0.23	0.37	
HGEI-f	0.64	0.57	0.63	0.53	0.38	0.52	0.40	0.25	0.39	
BGEI	0.22	0.17	0.19	0.12	0.06	0.12	0.09	0.04	0.09	
HBGEI-i	0.65	0.60	0.63	0.38	0.23	0.37	0.30	0.17	0.29	
HBGEI-f	0.66	0.61	0.65	0.42	0.27	0.41	0.35	0.20	0.34	
Vrijeme cipele (VC) varijacija										
Visina	0.51	0.48	0.50	0.19	0.11	0.19	0.13	0.06	0.12	
GEI	0.33	0.26	0.31	0.19	0.11	0.18	0.17	0.08	0.16	
HGEI-i	0.67	0.61	0.66	0.43	0.29	0.42	0.38	0.22	0.37	
HGEI-f	0.65	0.59	0.64	0.50	0.34	0.49	0.39	0.24	0.39	
BGEI	0.33	0.25	0.31	0.17	0.10	0.16	0.09	0.05	0.09	
HBGEI-i	0.59	0.51	0.57	0.41	0.23	0.40	0.37	0.20	0.36	
HBGEI-f	0.58	0.52	0.57	0.41	0.25	0.40	0.32	0.18	0.31	

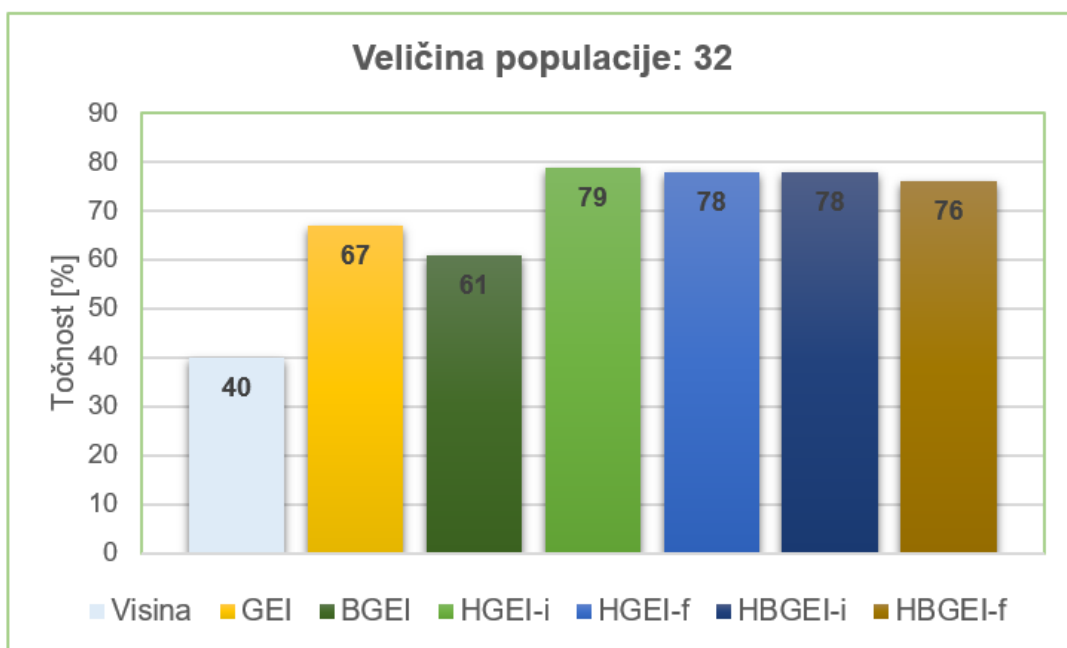
Kada se uzmu u obzir sve razdiobe visina i sve varijacije hoda, metoda zasnovana na značajki visina osobe ostvarila je prosječnu točnost od 40 % za malu, 13 % za srednju i 8 % za veliku populaciju (usporedni rezultati za sve metode su prikazani na slikama 6.7, 6.8 i 6.9). Najbolji rezultati točnosti identifikacije su u slučaju razdiobe visina *Godine*, dok su najlošiji rezultati u slučaju razdiobe visina *Slične* što vrijedi za sve korištene populacije ljudi.

Rezultati eksperimenata potvrđuju očekivane rezultate za GEI i BGEI metodu. GEI metoda ima prosječnu točnost identifikacije za sve razdiobe visina i varijacije hoda od 67 % za malu, 57 % za srednju te 52 % za veliku populaciju. U slučaju BGEI metode, ti rezultati iznose 61 % za malu, 50 % za srednju i 43 % za veliku populaciju. U svim veličinama populacija točnost identifikacije GEI metode veća je u odnosu na BGEI metodu. Uzmu li se u obzir sve varijacije i sve veličine populacija, rezultati točnosti identifikacije su veći za oko 7 % u korist GEI metode.

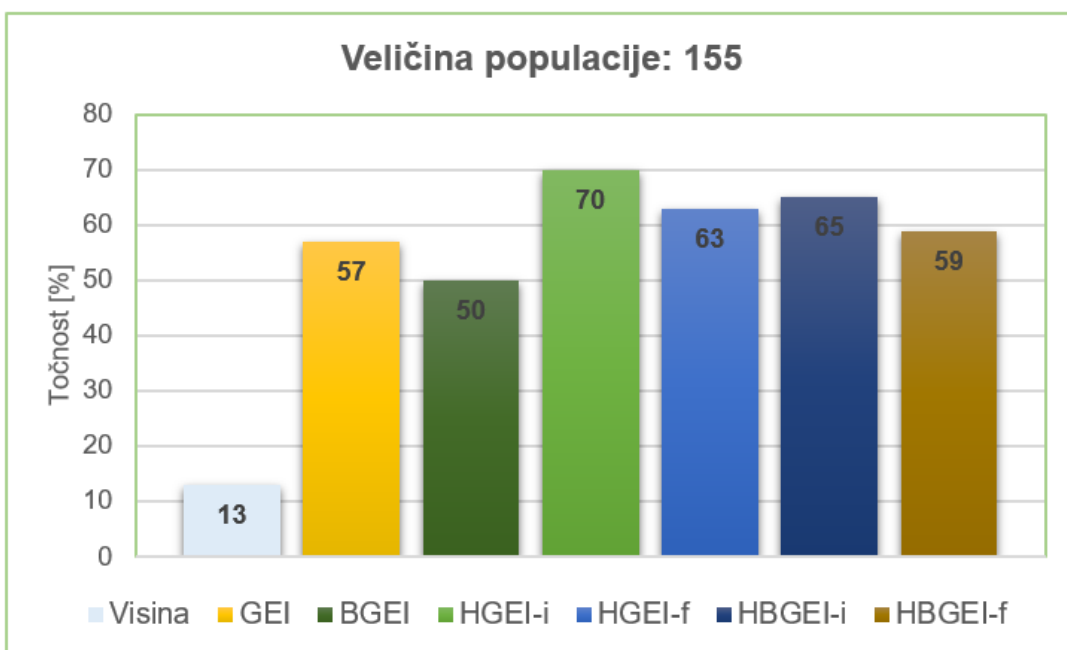
Predstavljani pristupi koji kombiniraju značajku visina osobe sa značajkama dobivenima metodom temeljenom na pojavi (HGEI-i, HGEI-f, HBGEI-i, HBGEI-f metode) znatno poboljšavaju točnost identifikacije u odnosu na samostalnu primjenu metoda. Kod dvije predložene metode, HGEI-i ima sljedeću točnost identifikacije : 79 % za malu populaciju, 70 % za srednju i 64 % za veliku populaciju kada se u obzir uzmu sve razdiobe visina i sve varijacije hoda. U slučaju metode HGEI-f, ti su rezultati 78 % za malu, 63 % za srednju te 49 % za veliku populaciju.

Prosječna točnost identifikacije u slučaju metoda HBGEI-i i HBGEI-f iznosi 78 % za malu, 65 % za srednju te 59 % za veliku populaciju kada je u pitanju HBGEI-i. Kod metode HBGEI-f, rezultati su 76 % za malu, 59 % za srednju i 47 % za veliku populaciju. Prethodno navedeni rezultati za sve primijenjene metode ilustrirani su na slikama 6.7, 6.8 i 6.9.

Klasifikacija koja je zasnovana samo na visini osobe ima najlošije rezultate, što je izrazito očekivano za veliku populaciju jer velik broj osoba ima slične visine koje je teško precizno odrediti. GEI metoda ima bolje rezultate u odnosu na BGEI metodu. Na temelju prethodno navedenoga i rezultata prikazanih u tablicama može se reći da primjena predloženih metoda (HGEI-i i HGEI-f), odnosno realizacije rane i kasne fuzije primjenom GEI-ja i visine osobe, poboljšava točnost identifikacije u odnosu na samostalnu primjenu metode kao što je GEI i značajke kao što je visina osobe. Jasno se



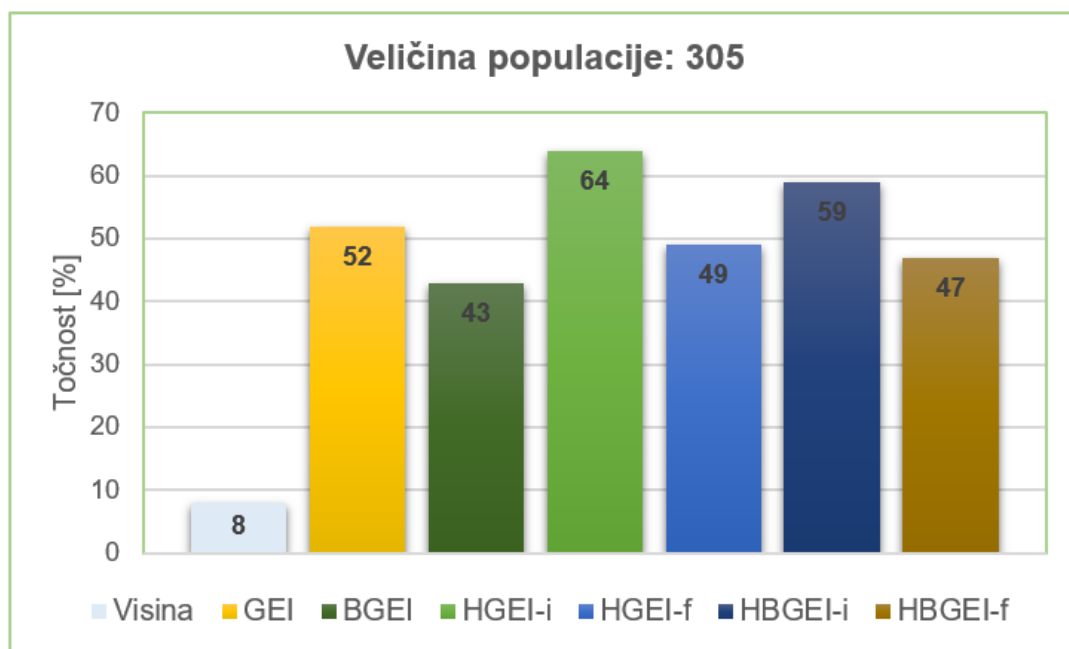
Slika 6.7: Prosječna točnost identifikacije uspoređenih metoda za malu veličinu populacije (32 osobe)



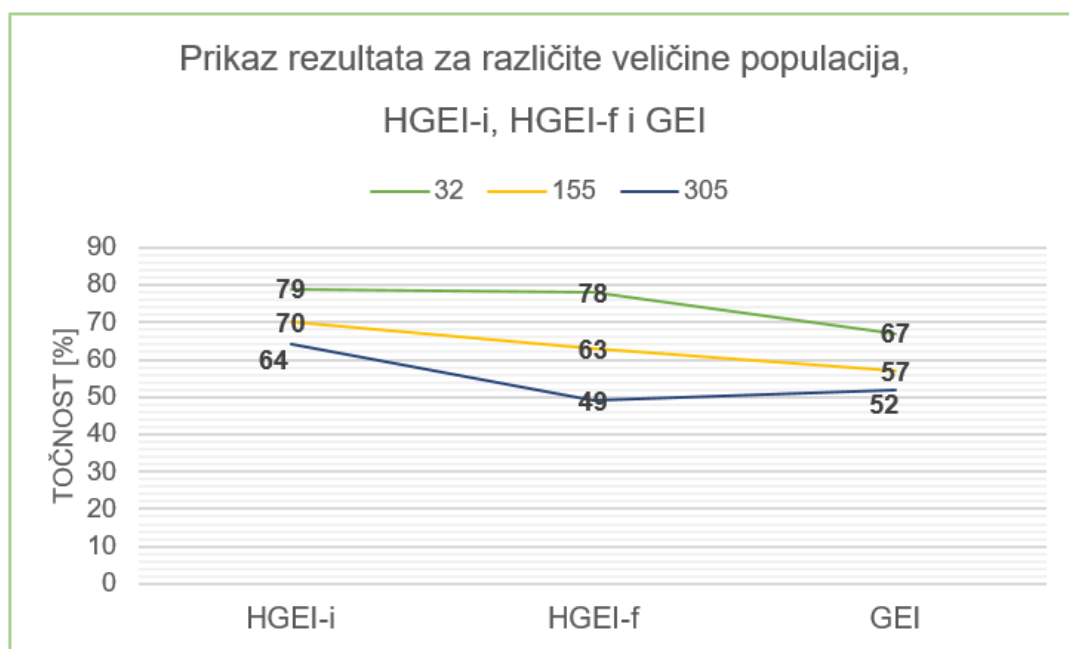
Slika 6.8: Prosječna točnost identifikacije uspoređenih metoda za srednju veličinu populacije (155 osoba)

može vidjeti na slici 6.10 gdje su rezultati prosječne točnosti identifikacije prikazani za različite veličine populacija ljudi za metode GEI, HGEI-i i HGEI-f.

Kada se razmotre dvije predložene metode, općenito gledajući, HGEI-i nadmašuje HGEI-f i može biti preferirani prvi izbor za upotrebu u većini slučajeva. Razlika u



Slika 6.9: Prosječna točnost identifikacije uspoređenih metoda za veliku veličinu populacije (305 osoba)

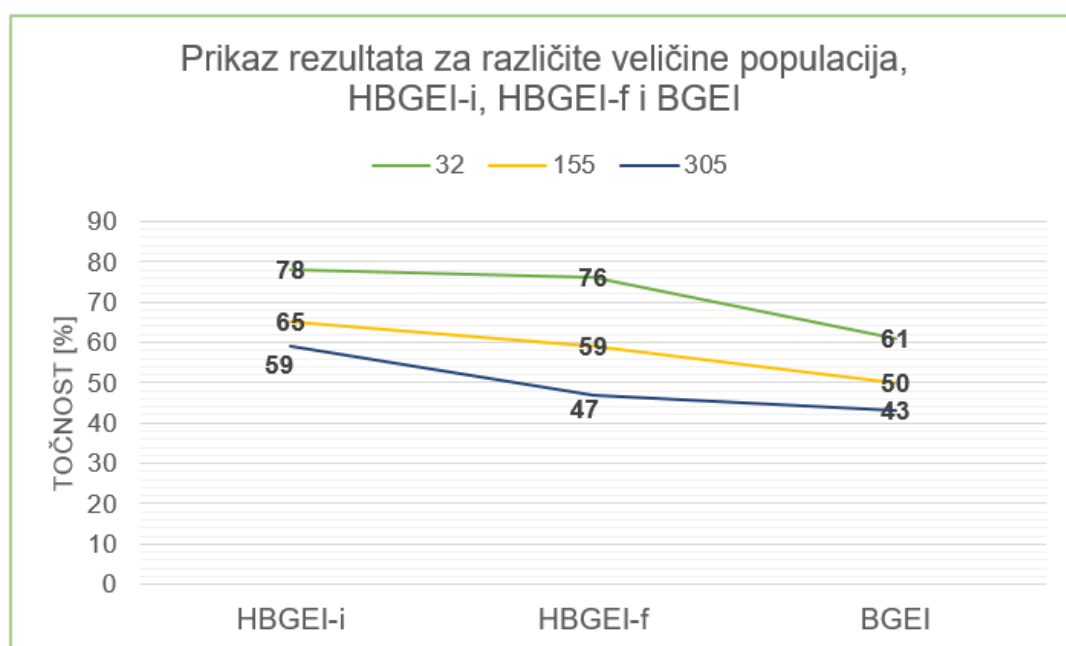


Slika 6.10: Usporedba između dobivenih rezultata na maloj, srednjoj i velikoj populaciji za predložene metode (HGEI-i i HGEI-f) i GEI metodu

točnosti između dvaju predloženih metoda veća je ovisno o porastu populacije, u korist HGEI-i metode. U nekim slučajevima, primjerice, kada je manja populacija ljudi, točnost identifikacije zasnovana na visini osobe može biti velika te u primjeni kasnije fuzije može rezultirati povećanjem točnosti tako da HGEI-f metoda može nadmašiti

HGEI-i metodu.

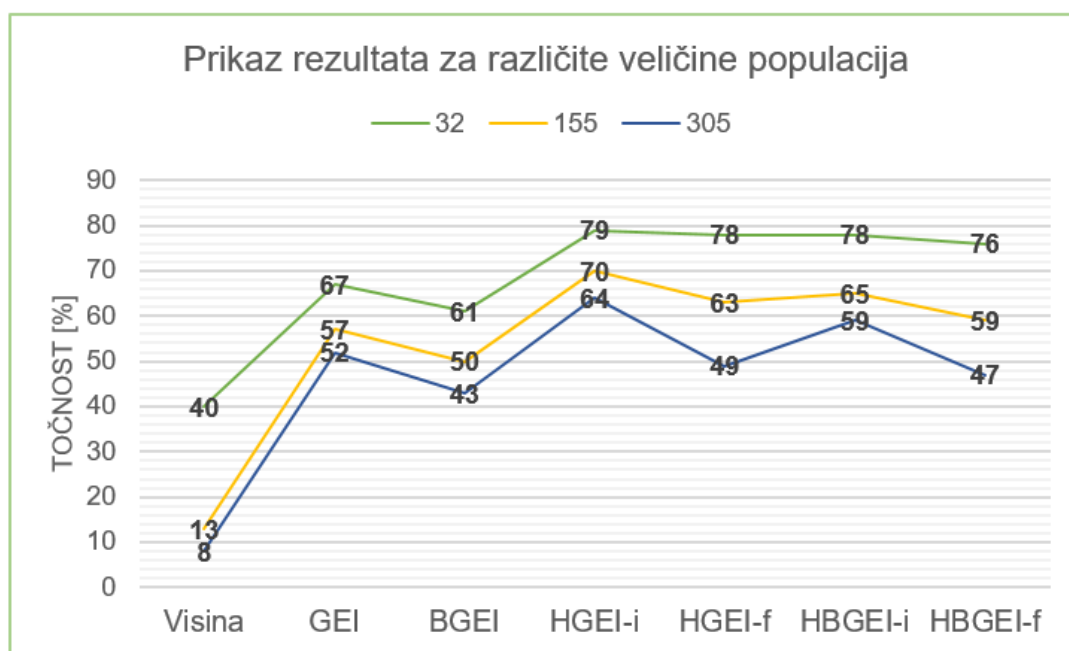
Jednako tako, samostalna primjena BGEI metode koja je, u suštini, nastala na GEI metodi i predstavlja jednu varijaciju GEI metode također ima lošije rezultate u odnosu na primjenu predloženih pristupa zasnovanih na realizaciji fuzije informacija. Realizacija fuzije informacija između BGEI metode i visine osobe te metode koje su nastale na temelju tog (HBGEI-i i HBGEI-f) također pokazuju poboljšanje rezultata u odnosu na samostalnu primjenu BGEI metode. U oba pristupa odnosno metode u kojima se upotrebljava fuzija informacija rezultati su bolji u usporedbi s rezultatima u odnosu na samostalnu primjenu BGEI metode što se vidi i na slici 6.11.



Slika 6.11: Usporedba dobivenih rezultata na maloj, srednjoj i velikoj populaciji za HBGEI-i, HBGEI-f i BGEI metodu

Kada se pogledaju dobiveni rezultati za HBGEI-i i HBGEI-f, vidljivo je da su oni lošiji u odnosu na dvije predložene metode, HGEI-i i HGEI-f. Očekivano je da u svim slučajevima rezultati opadaju ovisno o porastu populacije pa su tako u svim metodama najbolji rezultati u slučaju male populacije, dok porastom populacije ti rezultati opadaju. Radi usporedbe i preglednosti, svi rezultati za primjenjivane metode prikazani su na slici 6.12.

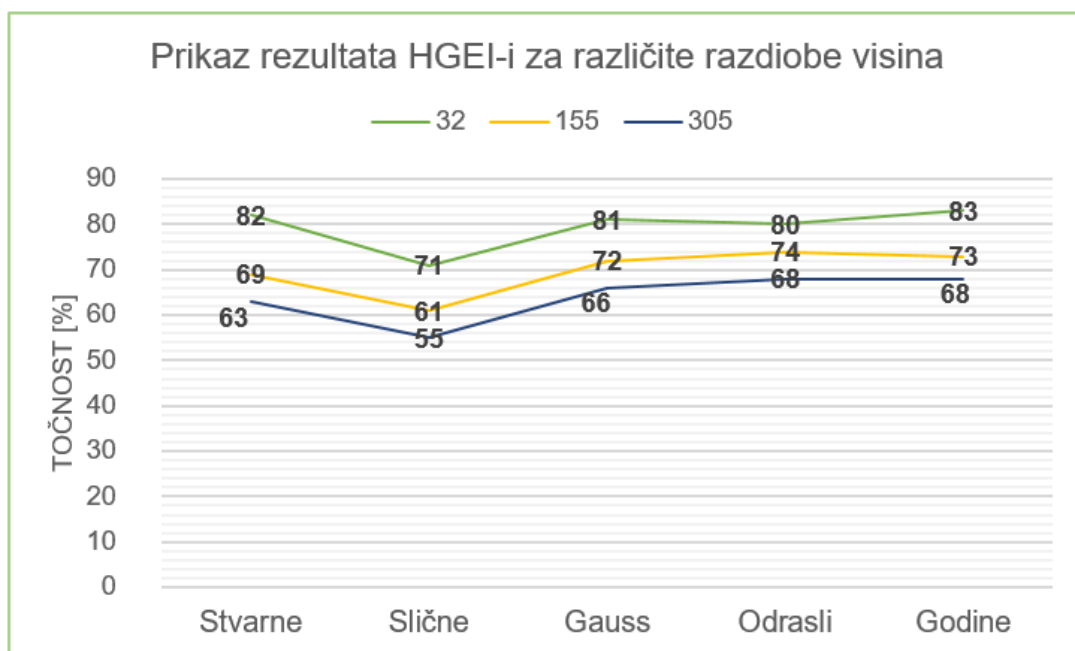
Dobiveni rezultati po razdiobama visina za dvije predložene metode su kako slijedi. Za razdiobu visina *Stvarne* (tablica 6.3) u slučaju HGEI-i metode točnost identifikacije iznosi 82 % za malu, 69 % za srednju te 63 % za veliku populaciju. Kod HGEI-f metode



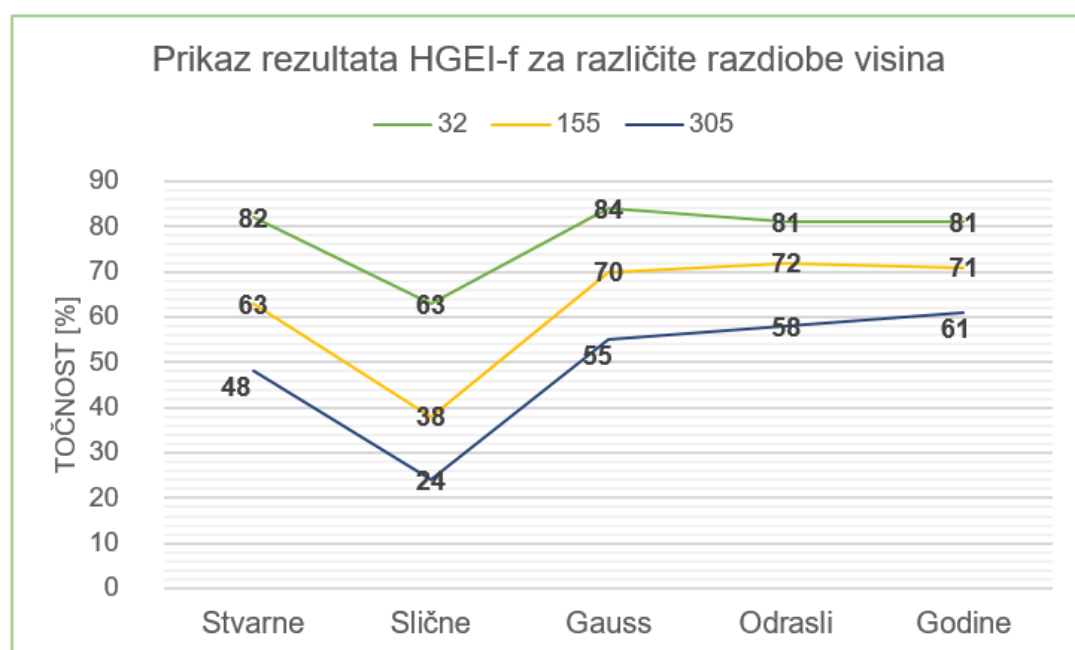
Slika 6.12: Usporedba između dobivenih rezultata na maloj, srednjoj i velikoj populaciji za sve primijenjene metode

točnost identifikacije je 82 % za malu, 63 % za srednju i 48 % za veliku populaciju. Kod razdiobe visina *Slične* (tablica 6.4), rezultati su 71 %, 61 % i 55 % za malu, srednju i veliku populaciju u slučaju HGEI-i metode. U slučaju HGEI-f metode ti rezultati iznose 63 %, 38 % i 24 % za malu, srednju i veliku populaciju. Razdioba visina *Gauss* (tablica 6.5) ima 81 %, 72 % i 66 % za tri navedene populacije i HGEI-i metodu. Za razdiobu visina *Gauss* i HGEI-f metodu rezultati su u slučaju male populacije 84 %, a za srednju i veliku populaciju 70 %, odnosno 55 %. Pri primjeni razdiobe visina *Odrasli* (tablica 6.6) za malu, srednju i veliku populaciju rezultati za HGEI-i iznose 80 %, 74 % i 68 %, dok su za HGEI-f 81 %, 72 % i 58 %. Na kraju, za razdiobu visina *Godine* (tablica 6.7) rezultat za malu populaciju iznosi 83 %, dok su za srednju i veliku populaciju rezultati 73 % i 68 % u slučaju HGEI-i, dok za HGEI-f iznose 81 %, 71 % i 61 %. Navedeni rezultati su prikazani na slikama 6.13 i 6.14.

Imajući u vidu sve rezultate opisane u tekstu i prikazane u tablicama 6.3, 6.4, 6.5, 6.6 i 6.7, može se zaključiti da su navedene metode najučinkovitije i najtočnije kada se radi o manjem skupu ljudi, odnosno o manjoj populaciji ljudi. Vidljivo je i u navedenim tablicama ako se pogledaju prosječni rezultati kroz sve razdiobe visina i sve veličine populacija ljudi: u maloj populaciji ljudi od 32 osobe postignuti su bolji prosječni rezultati u odnosu na srednju (155 osoba) i veliku (305 osoba) populaciju ljudi. Rezultati su



Slika 6.13: Prikaz rezultata HGEI-i po razdiobama visina za malu, srednju i veliku populaciju



Slika 6.14: Prikaz rezultata HGEI-f po razdiobama visina za malu, srednju i veliku populaciju

očekivani s obzirom na to da se povećanjem populacije povećava i broj sličnih osoba po svojstvima hoda i po visini osobe. Tome u prilog idu i rezultati s razdiobom visina *Slične* (tablica 6.4) koji su najlošiji jer navedena razdioba sadržava najveći broj osoba slične visine, za razliku od, primjerice, razdiobe visina *Godine* (tablica 6.7) koja ima

znatno bolje rezultate jer obuhvaća veći broj osoba različitih visina. Tijekom identifikacije najbolje bi bilo da su sve osobe različite visine, što gotovo nikad nije slučaj jer uvijek postoje osobe slične visine, što posebno dolazi do izražaja u većoj populaciji ljudi.

Idealan slučaj tijekom identifikacije bi bio i da osoba hoda normalno bez određenih dodataka što u praksi obično nije tako. Osobe često nose određene dodatke kao što su ruksak ili neki drugi tip torbe, nose različite tipove kapa, različite oblike i debljine jakni itd., što utječe na postupke određivanja i izlučivanja siluete osobe, određivanja visine osobe i u konačnici na točnost identifikacije osoba. Zbog toga primjena jedne značajke u svrhu identifikacije osoba često zna biti manjkava, zbog čega je poželjno kombinirati različite značajke te realizirati različite načine fuzije informacija s dobivenim značajkama iz više izvora. U prethodno opisanom tekstu kao i u tablicama 6.3, 6.4, 6.5, 6.6 i 6.7 vidi se da kombiniranje različitih značajki (kao u predloženim metodama HGEI-i i HGEI-f, pa i u drugim realiziranim izvedbama kao što su HBGEI-i i HBGEI-f) doprinosi povećanju točnosti identifikacije osoba u odnosu na samostalnu primjenu jednog tipa značajki (primjerice, samo GEI ili BGEI, samo visina osobe itd.).

7 ZAKLJUČAK

U ovoj disertaciji istražen je problem identifikacije osoba upotrebom RGB-D kamere. U odnosu na klasične RGB kamere koje pružaju RGB slike, RGB-D kamere dodatno pružaju i dubinske slike iz kojih se može procijeniti udaljenost objekata od kamere. Analizom dubinskih slika moguće je izlučiti dodatne značajke koje se mogu upotrebljavati tijekom identifikacije osoba. Tijekom istraživanja u ovoj disertaciji za potrebe identifikacije osoba, pored značajki koje se izlučuju iz RGB slika, upotrebljavana je i značajka visina osobe koja se izlučuje iz dubinskih slika.

Znanstveni doprinos provedenog istraživanja predstavlja:

- prijedlog dva nova pristupa i dvije nove metode za identifikaciju osoba na temelju hoda koje poboljšavaju točnost identifikacije osoba u usporedbi s poznatim metodama identifikacije osoba na temelju hoda
- izrađeni skup podataka, koji sadržava videozapise visoke razlučivosti za istraživanje metoda za identifikaciju osoba na temelju hoda, upotrebom RGB-D kamere u vanjskom okruženju.

Svaka osoba tijekom ciklusa hoda kreira različite obrasce koji se mogu iskoristiti u procesu identifikacije. Prednost metoda identifikacije osoba na temelju hoda je to što osoba može biti identificirana, a da uopće nije ni svjesna da prolazi proces identifikacije. Tradicionalno metode za identifikaciju osoba na temelju hoda kao jedan od prvih koraka obrade izlučuju siluete osoba iz slika odnosno pojedinačnih okvira videosekvencija. S RGB-D kamerama, koje istovremeno daju i RGB i dubinske slike, siluete osoba mogu se također izlučiti i iz dubinskih slika pored uobičajenog načina izlučivanja iz RGB slika. U tom se slučaju, budući da se može procijeniti udaljenost osobe od kamere, može procijeniti i visina osobe i upotrebljavati je kao značajku za identifikaciju osoba. Radi se o značajki koja je konceptualno jednostavna i koja ne zahtijeva složenu računalnu obradu te se može procijeniti u stvarnom vremenu.

Na temelju tog uvida, predložene su dvije nove metode koje kombiniraju značajke temeljene na pojavi kao što je GEI sa značajkom visina osobe procijenjene upotrebom dubinskih slika. Predložene metode su nazvane HGEI-i i HGEI-f. Metoda HGEI-i upotrebljava ranu fuziju informacija, pri čemu se značajke spajaju prije koraka klasifikacije, dok se u metodi HGEI-f primjenjuje kasna fuzija u kojoj se rezultati klasifikacije

na pojedinačnim značajkama optimalno integriraju s alfa-integracija metodom.

Eksperimentom su vrednovane metode i predstavljeni su dobiveni rezultati, a pritom su tijekom eksperimenata u obzir uzete populacije ljudi različitih veličina i s različitim razdiobama visina. Postignuti rezultati s predloženim metodama direktno potvrđuju hipotezu istraživanja u kojoj je navedeno da *integracijom značajki izlučenih iz dubinskih slika videosekvencija RGB-D kamera sa značajkama izlučenim iz RGB slika videosekvencija RGB-D kamera može poboljšati točnost identifikacije osoba u odnosu na metode temeljene samo na RGB kamerama*. Prema dobivenim rezultatima, dvije predložene metode u kojima se integriraju značajke koje se dobivaju iz RGB i dubinskih slika, su ostvarile veću prosječnu točnost identifikacije u odnosu na metode koje su upotrebljavane za usporedbu. Metode koje su upotrebljavane za usporedbu, GEI metoda i izvedba BGEI metode gdje se koriste siluete osobe promatrane sa strane, su dosad obično realizirane upotrebljavajući RGB slike.

Da bi se dodatno provjerila učinkovitost predloženih pristupa tj. upotreba rane i kasne fuzije informacija kreirane su i metode HBGEI-i i HBGEI-f. Navedene metode, HBGEI-i i HBGEI-f, realizirane su na način kao i nove predložene metode HGEI-i i HGEI-f s tom razlikom da upotrebljavaju BGEI slike iz RGB-D kamera umjesto GEI slika. Prosječni rezultati točnosti identifikacije su također bolji u korist metoda HBGEI-i i HBGEI-f u odnosu na samostalnu upotrebu BGEI metode.

Za potrebe istraživanja upotrijebljeni su dostupni RGB i RGB-D skupovi podataka za identifikaciju osoba (Casia Dataset B i TUM-GAID), a koji se često koriste i od strane drugih autora. Uz njih je realiziran i vlastiti skup podataka koji se, u odnosu na trenutačno dostupne skupove podataka RGB-D, razlikuje u tome što je sniman dalekosežnom RGB-D kamerom u vanjskom okruženju. Radi se o manjem skupu podataka koji čine videozapisi za 14 osoba u hodu. Glavna prednost vlastitog skupa podataka jesu visoka razlučivost, sadržane dubinske slike te snimanje u vanjskom okruženju gdje su osobe na većoj udaljenosti od kamere, a što ga čini pogodnim za dodatne provjere robusnosti različitih metoda.

Analizirane su siluete osoba dobivene iz RGB slika vlastitog skupa podataka. Siluete osobe dobivene iz RGB slika nisu bile cjelovite u situacijama kada je zbog sličnosti boje odjeće koju nosi osoba i boje pozadine segmentacija osobe u sceni bila otežana. Međutim, s pomoću dubinskih slika mogu se izlučiti dodatne značajke koje se mogu

upotrebljavati zajedno sa značajkama iz RGB slika i tako omogućiti točniju identifikaciju osobe. Analizom eksperimentalnih rezultata pružen je uvid u kojoj mjeri upotreba dodatnih značajki izlučenih iz dubinskih slika poput visine osobe utječe na točnost identifikacije u odnosu na identifikaciju upotrebom samo silueta osoba dobivenih iz RGB slika. Time je potvrđena hipoteza istraživanja u kojoj je navedeno da *integracijom značajki izlučenih iz dubinskih slika RGB-D kamera sa značajkama izlučenim iz RGB slika istih kamera identifikacija se može primijeniti i na videosekvencijama gdje značajke izlučene iz RGB slika zasebno nisu dostatne za identifikaciju.*

Identifikacija osoba iz videosekvencija kamera, a posebno iz danas široko dostupnih RGB-D kamera, kao i rezultati istraživanja predstavljeni u ovoj disertaciji, mogu biti primijenjeni u različitim područjima kao što su uslužna robotika, autoindustrija ili sigurnosni sustavi.

Literatura

- [1] M. Alotaibi and A. Mahmood, "Improved Gait Recognition Based on Specialized Deep Convolutional Neural Network," *Computer Vision and Image Understanding*, vol. 164, pp. 103–110, 2017.
- [2] P. Arora, M. Hanmandlu, and S. Srivastava, "Gait Based Authentication Using Gait Information Image Features," *Pattern Recognition Letters*, vol. 68, pp. 336–342, 2015.
- [3] P. Arora and S. Srivastava, "Gait Recognition Using Gait Gaussian Image," in *2nd International Conference Signal Processing and Integrated Networks (SPIN)*. IEEE, 2015, pp. 791–794.
- [4] C. Audet and J. E. Dennis Jr, "Mesh Adaptive Direct Search Algorithms for Constrained Optimization," *SIAM Journal on Optimization*, vol. 17, no. 1, pp. 188–217, 2006.
- [5] D. Baltieri, R. Vezzani, and R. Cucchiara, "Mapping Appearance Descriptors on 3D Body Models for People Re-identification," *International Journal of Computer Vision*, vol. 111, no. 3, pp. 345–364, 2015.
- [6] I. B. Barbosa, M. Cristani, A. Del Bue, L. Bazzani, and V. Murino, "Re-identification with RGB-D Sensors," in *12th European Conference on Computer Vision (ECCV)*. Springer, 2012, pp. 433–442.
- [7] O. Barnich, "Motion Detection and Human Recognition in Video Sequences," Ph.D. dissertation, Ph. D. thesis, University of Liège, Belgium, 2010.
- [8] K. Bashir, T. Xiang, and S. Gong, "Gait Recognition Using Gait Entropy Image," 2009.
- [9] T. Bell, B. Li, and S. Zhang, "Structured Light Techniques and Applications," *Wiley Encyclopedia of Electrical and Electronics Engineering*, 2016.
- [10] R. Borràs, À. Lapedriza, and L. Igual, "Depth Information in Human Gait Analysis: An Experimental Study on Gender Recognition," *Image Analysis and Recognition*, pp. 98–105, 2012.

- [11] D. Brščić, T. Kanda, T. Ikeda, and T. Miyashita, "Person Tracking in Large Public Spaces Using 3-D Range Sensors," *IEEE Transactions on Human-Machine Systems*, vol. 43, no. 6, pp. 522–534, 2013.
- [12] F. M. Castro, M. J. Marín-Jiménez, N. Guil, and N. P. de la Blanca, "Automatic Learning of Gait Signatures for People Identification," in *International Work-Conference on Artificial Neural Networks*. Springer, 2017, pp. 257–270.
- [13] P. Chattopadhyay, A. Roy, S. Sural, and J. Mukhopadhyay, "Pose Depth Volume Extraction from RGB-D Streams for Frontal Gait Recognition," *Journal of Visual Communication and Image Representation*, vol. 25, no. 1, pp. 53–63, 2014.
- [14] P. Chattopadhyay, S. Sural, and J. Mukherjee, "Frontal Gait Recognition from Occluded Scenes," *Pattern Recognition Letters*, vol. 63, pp. 9–15, 2015.
- [15] W. Choi, C. Pantofaru, and S. Savarese, "Detecting and Tracking People Using an RGB-D Camera via Multiple Detector Fusion," in *International Conference on Computer Vision Workshops (ICCV Workshops)*. IEEE, 2011, pp. 1076–1083.
- [16] R. T. Collins, R. Gross, and J. Shi, "Silhouette-based Human Identification from Body Shape and Gait," in *Proceedings of 5th IEEE International Conference on Automatic Face Gesture Recognition*. IEEE, 2002, pp. 366–371.
- [17] J. V. Davis, B. Kulis, P. Jain, S. Sra, and I. S. Dhillon, "Information-theoretic Metric Learning," in *Proceedings of the 24th International Conference on Machine Learning*, 2007, pp. 209–216.
- [18] T. D. Do, V. H. Nguyen, and H. Kim, "Real-time and Robust Multiple-view Gender Classification Using Gait Features in Video Surveillance," *Pattern Analysis and Applications*, vol. 23, no. 1, pp. 399–413, 2020.
- [19] A. El Gamal and H. Eltoukhy, "CMOS Image Sensors," *IEEE Circuits and Devices Magazine*, vol. 21, no. 3, pp. 6–20, 2005.
- [20] A. Elgammal, R. Duraiswami, D. Harwood, and L. S. Davis, "Background and Foreground Modeling Using Nonparametric Kernel Density Estimation for Visual Surveillance," *Proceedings of the IEEE*, vol. 90, no. 7, pp. 1151–1163, 2002.

- [21] A. Farooq, A. Jalal, and S. Kamal, "Dense RGB-D Map-based Human Tracking and Activity Recognition Using Skin Joints Features and Self-organizing Map," *KSII Transactions on Internet & Information Systems*, vol. 9, no. 5, 2015.
- [22] N. Feddaoui, H. Mahersia, and K. Hamrouni, "Improving Iris Recognition Performance Using Quality Measures," in *Advanced Biometric Technologies*. InTech, 2011.
- [23] E. R. Fossum, "Active Pixel Sensors: Are CCDs Dinosaurs?" in *Charge-Coupled Devices and Solid State Optical Sensors III*, vol. 1900. International Society for Optics and Photonics, 1993, pp. 2–14.
- [24] E. R. Fossum, "CMOS Image Sensors: Electronic Camera-on-a-Chip," *IEEE Transactions on Electron Devices*, vol. 44, no. 10, pp. 1689–1698, 1997.
- [25] J. Goldberger, G. E. Hinton, S. T. Roweis, and R. R. Salakhutdinov, "Neighbourhood Components Analysis," in *Advances in Neural Information Processing Systems*, 2005, pp. 513–520.
- [26] R. Gross and J. Shi, "The CMU Motion of Body (MoBo) Database," Carnegie Mellon University, Pittsburgh, PA, Tech. Rep. CMU-RI-TR-01-18, June 2001.
- [27] S. R. Gunn, "Support Vector Machines for Classification and Regression," *ISIS Technical Report*, vol. 14, no. 1, pp. 5–16, 1998.
- [28] J. Han and B. Bhanu, "Individual Recognition Using Gait Energy Image," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 2, pp. 316–322, 2006.
- [29] J. Han, L. Shao, D. Xu, and J. Shotton, "Enhanced Computer Vision with Microsoft Kinect Sensor: A Review," *IEEE Transactions on Cybernetics*, vol. 43, no. 5, pp. 1318–1334, 2013.
- [30] M. Hansard, S. Lee, O. Choi, and R. P. Horaud, *Time-of-Flight Cameras: Principles, Methods and Applications*. Springer Science & Business Media, 2012.
- [31] R. I. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*, 2nd ed. Cambridge University Press, ISBN: 0521540518, 2004.

- [32] M. Hofmann, S. Bachmann, and G. Rigoll, "2.5D Gait Biometrics Using the Depth Gradient Histogram Energy Image," in *5th International Conference Biometrics: Theory, Applications and Systems (BTAS)*. IEEE, 2012, pp. 399–403.
- [33] M. Hofmann, J. Geiger, S. Bachmann, B. Schuller, and G. Rigoll, "The TUM Gait from Audio, Image and Depth (GAID) Database: Multimodal Recognition of Subjects and Traits," *Journal of Visual Communication and Image Representation*, vol. 25, no. 1, pp. 195–206, 2014.
- [34] L. Introna and H. Nissenbaum, "Facial Recognition Technology-A Survey of Policy and Implementation Issues," 2010.
- [35] K. Irsch and D. L. Guyton, "Anatomy of Eyes," 2009.
- [36] Irsko vijeće za bioetiku, "Biometrics: Enhancing Security or Invading Privacy? Opinion," Dublin, 2009.
- [37] Y. Iwashita, R. Baba, K. Ogawara, and R. Kurazume, "Person Identification from Spatio-temporal 3D Gait," in *2010 International Conference on Emerging Security Technologies*. IEEE, 2010, pp. 30–35.
- [38] Y. Iwashita, K. Uchino, and R. Kurazume, "Gait-based Person Identification Robust to Changes in Appearance," *Sensors*, vol. 13, no. 6, pp. 7884–7901, 2013.
- [39] A. Jain and S. Pankanti, "Automated Fingerprint Identification and Imaging Systems," *Advances in Fingerprint Technology*, p. 275, 2001.
- [40] A. K. Jain, K. Nandakumar, and A. Ross, "50 Years of Biometric Research: Accomplishments, Challenges, and Opportunities," *Pattern Recognition Letters*, vol. 79, pp. 80–105, 2016.
- [41] A. K. Jain, A. Ross, and S. Prabhakar, "An Introduction to Biometric Recognition," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 14, no. 1, pp. 4–20, 2004.
- [42] P. KaewTraKulPong and R. Bowden, "An Improved Adaptive Background Mixture Model for Real-time Tracking with Shadow Detection," in *Video-based Surveillance Systems*. Springer, 2002, pp. 135–144.

- [43] D. Kastaniotis, I. Theodorakopoulos, C. Theoharatos, G. Economou, and S. Fotopoulos, "A Framework for Gait-based Recognition Using Kinect," *Pattern Recognition Letters*, vol. 68, pp. 327–335, 2015.
- [44] H.-J. Kim, Y.-H. Park, and S.-K. Kwon, "A Method of Estimating the Human Height Using Depth Images," *Journal of the Korea Industrial Information Systems Research*, vol. 22, no. 2, pp. 53–60, 2017.
- [45] K. Kim, T. H. Chalidabhongse, D. Harwood, and L. Davis, "Background Modeling and Subtraction by Codebook Construction," in *2004 International Conference on Image Processing, 2004. ICIP'04.*, vol. 5. IEEE, 2004, pp. 3061–3064.
- [46] P. Komarinski, *Automated Fingerprint Identification Systems (AFIS)*. Academic Press, 2005.
- [47] M. Kumar and R. V. Babu, "Human Gait Recognition Using Depth Camera: A Covariance Based Approach," in *Proceedings of the 8th Indian Conference on Computer Vision, Graphics and Image Processing*. ACM, 2012, p. 20.
- [48] C. Le and R. Jain, "A Survey of Biometrics Security Systems," *EEUU. Washington University in St. Louis*, 2009.
- [49] L. Lee and W. E. L. Grimson, "Gait Analysis for Recognition and Classification," in *Automatic Face and Gesture Recognition, Proceedings of the 5th IEEE International Conference*. IEEE, 2002, pp. 155–162.
- [50] R. Liao, C. Cao, E. B. Garcia, S. Yu, and Y. Huang, "Pose-based Temporal-spatial Network (PTSN) for Gait Recognition with Carrying and Clothing Variations," in *Chinese Conference on Biometric Recognition*. Springer, 2017, pp. 474–483.
- [51] A. Liaw and M. Wiener, "Classification and Regression by randomForest," *R News*, vol. 2, no. 3, pp. 18–22, 2002.
- [52] A. O. Lishani, L. Boubchir, E. Khalifa, and A. Bouridane, "Human Gait Recognition Based on Haralick Features," *Signal, Image and Video Processing*, vol. 11, no. 6, pp. 1123–1130, 2017.

- [53] D. Litwiller, "CCD vs. CMOS," *Photonics Spectra*, vol. 35, no. 1, pp. 154–158, 2001.
- [54] D. López-Fernández, F. J. Madrid-Cuevas, A. Carmona-Poyato, R. Muñoz-Salinas, and R. Medina-Carnicer, "A New Approach for Multi-view Gait Recognition on Unconstrained Paths," *Journal of Visual Communication and Image Representation*, vol. 38, pp. 396–406, 2016.
- [55] D. López-Fernández, F. J. Madrid-Cuevas, Á. Carmona-Poyato, M. J. Marín-Jiménez, and R. Muñoz-Salinas, "The AVA Multi-view Dataset for Gait Recognition," in *International Workshop on Activity Monitoring by Multiple Distributed Sensing*. Springer, 2014, pp. 26–39.
- [56] J. Lorenzo-Navarro, M. Castrillón-Santana, and D. Hernández-Sosa, "On the Use of Simple Geometric Descriptors Provided by RGB-D Sensors for Re-identification," *Sensors*, vol. 13, no. 7, pp. 8222–8238, 2013.
- [57] J. Lu, G. Wang, and P. Moulin, "Human Identity and Gender Recognition from Gait Sequences with Arbitrary Walking Directions," *IEEE Transactions on Information Forensics and Security*, vol. 9, no. 1, pp. 51–61, 2014.
- [58] L. Maddalena and A. Petrosino, "Background Subtraction for Moving Object Detection in RGBD Data: A Survey," *Journal of Imaging*, vol. 4, no. 5, p. 71, 2018.
- [59] Y. Makihara, H. Mannami, A. Tsuji, M. A. Hossain, K. Sugiura, A. Mori, and Y. Yagi, "The OU-ISIR Gait Database Comprising the Treadmill Dataset," *IPSJ Transactions on Computer Vision and Applications*, vol. 4, pp. 53–62, 2012.
- [60] C. D. Manning, H. Schütze, and P. Raghavan, *Introduction to Information Retrieval*. Cambridge University Press, 2008.
- [61] A. M. McIvor, "Background Subtraction Techniques," *Proceedings of Image and Vision Computing*, vol. 4, pp. 3099–3104, 2000.
- [62] A. Mogelmose, T. B. Moeslund, and K. Nasrollahi, "Multimodal Person Re-identification Using RGB-D Sensors and a Transient Identification Database," in *International Workshop on Biometrics and Forensics (IWBF)*. IEEE, 2013, pp. 1–4.

- [63] A. Mogelmose, C. Bahnsen, T. Moeslund, A. Clapés, and S. Escalera, “Tri-modal Person Re-identification with RGB, Depth and Thermal Features,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPR)*, 2013, pp. 301–307.
- [64] M. Mohandes, M. Deriche, and S. O. Aliyu, “Classifiers Combination Techniques: A Comprehensive Review,” *IEEE Access*, vol. 6, pp. 19 626–19 639, 2018.
- [65] F. Monroe and A. D. Rubin, “Keystroke Dynamics as a Biometric for Authentication,” *Future Generation Computer Systems*, vol. 16, no. 4, pp. 351–359, 2000.
- [66] M. Munaro, A. Basso, A. Fossati, L. Van Gool, and E. Menegatti, “3D Reconstruction of Freely Moving Persons for Re-identification with a Depth Sensor,” in *International Conference on Robotics and Automation (ICRA)*. IEEE, 2014, pp. 4512–4519.
- [67] M. Munaro, A. Fossati, A. Basso, E. Menegatti, and L. Van Gool, “One-shot Person Re-identification with a Consumer Depth Camera,” in *Advances in Computer Vision and Pattern Recognition, Person Re – Identification*. Springer, 2014, pp. 161–181.
- [68] M. Munaro, S. Ghidoni, D. T. Dizmen, and E. Menegatti, “A Feature-based Approach to People Re-identification Using Skeleton Keypoints,” in *International Conference on Robotics and Automation (ICRA)*. IEEE, 2014, pp. 5644–5651.
- [69] M. Munaro, C. Lewis, D. Chambers, P. Hvass, and E. Menegatti, “RGB-D Human Detection and Tracking for Industrial Environments,” in *Intelligent Autonomous Systems 13*. Springer, 2016, pp. 1655–1668.
- [70] M. Munaro and E. Menegatti, “Fast RGB-D People Tracking for Service Robots,” *Autonomous Robots*, vol. 37, no. 3, pp. 227–242, 2014.
- [71] C. D. Mutto, P. Zanuttigh, and G. M. Cortelazzo, *Time-of-Flight Cameras and Microsoft Kinect (TM)*. Springer Publishing Company, Incorporated, 2012.
- [72] “Nacionalni forenzički znanstveno - tehnološki centar, Internacionalno Sveučilište Florida, Web portal: A Simplified Guide to Forensic Science,”

Poveznica: <http://www.forensicsciencesimplified.org/prints/Fingerprints.pdf>,
(Posjeta: 18.12.2019. godine).

- [73] A. M. Nambiar, P. L. Correia, and L. D. Soares, "Frontal Gait Recognition Combining 2D and 3D Data," in *Proceedings of the on Multimedia and Security*. ACM, 2012, pp. 145–150.
- [74] V. Nozick, "Multiple View Image Rectification," in *1st IEEE International Symposium on Access Spaces (ISAS)*. IEEE, 2011, pp. 277–282.
- [75] F. Pala, R. Satta, G. Fumera, and F. Roli, "Multimodal Person Reidentification Using RGB-D Cameras," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 26, no. 4, pp. 788–799, 2016.
- [76] J. Perry and J. R. Davids, "Gait Analysis: Normal and Pathological Function," *Journal of Pediatric Orthopaedics*, vol. 12, no. 6, p. 815, 1992.
- [77] "Platforma NVIDIA CUDA," Poveznica: <https://docs.nvidia.com/cuda/index.html>,
(Posjeta: 03.09.2021. godine).
- [78] J. Portillo-Portillo, R. Leyva, V. Sanchez, G. Sanchez-Perez, H. Perez-Meana, J. Olivares-Mercado, K. Toscano-Medina, and M. Nakano-Miyatake, "A View-invariant Gait Recognition Algorithm Based on a Joint-direct Linear Discriminant Analysis," *Applied Intelligence*, vol. 48, no. 5, pp. 1200–1217, 2018.
- [79] J. Preis, M. Kessel, M. Werner, and C. Linnhoff-Popien, "Gait Recognition with Kinect," in *1st International Workshop on Kinect in Pervasive Computing*. New Castle, UK, 2012, pp. 1–4.
- [80] Ž. Radmilović, "Biometrijska identifikacija," *Policija i sigurnost*, vol. 17, no. 3-4, pp. 159–180, 2008.
- [81] A. Ramakić, D. Sušan, K. Lenac, and Z. Bundalo, "Depth-based Real-time Gait Recognition," *Journal of Circuits, Systems and Computers*, vol. 29, no. 16, 2020.
- [82] T.-M. Rhyne, "Applying Color Theory to Digital Media and Visualization," 2016.

- [83] I. Rida, S. Almaadeed, and A. Bouridane, "Gait Recognition Based on Modified Phase-only Correlation," *Signal, Image and Video Processing*, vol. 10, no. 3, pp. 463–470, 2016.
- [84] A. Ross, A. Jain, and S. Pankati, "A Prototype Hand Geometry-based Verification System," in *Proceedings of 2nd Conference on Audio and Video-based Biometric Person Authentication*, 1999, pp. 166–171.
- [85] G. Safont, A. Salazar, and L. Vergara, "Multiclass Alpha Integration of Scores from Multiple Classifiers," *Neural Computation*, vol. 31, no. 4, pp. 806–825, 2019.
- [86] R. Satta, F. Pala, G. Fumera, and F. Roli, "Real-time Appearance-based Person Re-identification Over Multiple Kinect Cameras," in *8th International Conference on Computer Vision Theory and Applications (VISAPP)*, 2013, pp. 407–410.
- [87] D. Scheuermann, S. Schwiderski–Grosche, and B. Struif, "Usability of Biometrics in Relation to Electronic Signatures," GMD – German National Research Center for Information Technology Institute for Secure Telecooperation (SIT).
- [88] J. Schöning and G. Heidemann, "Taxonomy of 3D Sensors-A Survey of State-of-the-Art Consumer 3D-Reconstruction Sensors and their Field of Applications," in *VISIGRAPP (3: VISAPP)*, 2016, pp. 194–199.
- [89] J. D. Shutler, M. G. Grant, M. S. Nixon, and J. N. Carter, "On a Large Sequence-based Human Gait Database," in *Applications and Science in Soft Computing*. Springer, 2004, pp. 339–346.
- [90] A. Sinha, K. Chakravarty, and B. Bhowmick, "Person Identification Using Skeleton Information from Kinect," in *Proceedings of the 6th International Conference on Advances in Computer – Human Interactions*, 2013, pp. 101–108.
- [91] S. Sivapalan, D. Chen, S. Denman, S. Sridharan, and C. Fookes, "Gait Energy Volumes and Frontal Gait Recognition Using Depth Images," in *International Joint Conference on Biometrics (IJCB)*. IEEE, 2011, pp. 1–6.
- [92] S. Sivapalan, D. Chen, S. Denman, S. Sridharan, and C. Fookes, "The Backfilled GEI-A Cross-Capture Modality Gait Feature for Frontal and Side-view Gait Re-

- cognition,” in *International Conference Digital Image Computing Techniques and Applications (DICTA)*. IEEE, 2012, pp. 1–8.
- [93] “Skup podataka ADSC – AWD,” Poveznica: <https://sites.google.com/site/adscawdgait/>, (Posjeta: 31.08.2021. godine).
- [94] “Skup podataka AVA Multi-view Dataset,” Poveznica: <https://www.uco.es/investiga/grupos/ava/node/41>, (Posjeta: 31.08.2021. godine).
- [95] “Skup podataka CMU MoBo,” Poveznica: <https://www.ri.cmu.edu/publications/the-cmu-motion-of-body-mobo-database/>, (Posjeta: 31.08.2021. godine).
- [96] “Skup podataka Kyushu University 4D Gait Database,” Poveznica: <http://robotics.ait.kyushu-u.ac.jp/yumi/db.html>, (Posjeta: 31.08.2021. godine).
- [97] “Skup podataka OU – ISIR Treadmill,” Poveznica: <http://www.am.sanken.osaka-u.ac.jp/BiometricDB/GaitTM.html>, (Posjeta: 31.08.2021. godine).
- [98] “Skup podataka Soton,” Poveznica: <http://www.gait.ecs.soton.ac.uk>, (Posjeta: 31.08.2021. godine).
- [99] “Skup podataka USF Human ID,” Poveznica: http://www.eng.usf.edu/cvprg/Gait_Data.html, (Posjeta: 31.08.2021. godine).
- [100] “Službena web stranica Asus,” Poveznica: https://www.asus.com/3D-Sensor/Xtion_PRO, (Posjeta: 23.12.2019. godine).
- [101] “Službena web stranica Basler,” Poveznica: <https://www.baslerweb.com>, (Posjeta: 23.12.2019. godine).
- [102] “Službena web stranica Centra za istraživanje biometrije i sigurnosti, Institut automatizacije, Kineska akademija znanosti,” Poveznica: <http://www.cbsr.ia.ac.cn/english/Gait%20Databases.asp>, (Posjeta: 23.12.2019. godine).
- [103] “Službena web stranica Merriam–Webster,” Poveznica: <https://www.merriam-webster.com/dictionary/gait>, (Posjeta: 23.12.2019. godine).
- [104] “Službena web stranica Microsoft,” Poveznica: <https://docs.microsoft.com/en-us/archive/blogs/kinectforwindows/pre-order-your-kinect-for-windows-v2-sensor-starting-today>, (Posjeta: 23.12.2019. godine).

- [105] "Službena web stranica Microsoft," Poveznice: <https://msdn.microsoft.com/enus/library/hh438998.aspx>, <https://msdn.microsoft.com/enus/library/hh973074.aspx>, (Posjeta: 23.12.2019. godine).
- [106] "Službena web stranica Microsoft Azure," Poveznica: <https://azure.microsoft.com/en-in/services/kinect-dk/>, (Posjeta: 27.12.2019. godine).
- [107] "Službena web stranica Ricoh," Poveznica: https://www.ricoh.com/technology/tech/045_stereocamera, (Posjeta: 23.12.2019. godine).
- [108] "Službena web stranica scikit-learn.org, GNB," Poveznica: https://scikit-learn.org/stable/modules/naive_bayes.html#gaussian-naive-bayes, (Posjeta: 08.04.2021. godine).
- [109] "Službena web stranica scikit-learn.org, kNN," Poveznica: <https://scikit-learn.org/stable/modules/generated/sklearn.neighbors.KNeighborsClassifier.html>, (Posjeta: 08.04.2021. godine).
- [110] "Službena web stranica scikit-learn.org, RFC," Poveznica: <https://scikit-learn.org/stable/modules/generated/sklearn.ensemble.RandomForestClassifier.html>, (Posjeta: 08.04.2021. godine).
- [111] "Službena web stranica scikit-learn.org, SVC," Poveznica: <https://scikit-learn.org/stable/modules/generated/sklearn.svm.SVC.html#sklearn.svm.SVC>, (Posjeta: 08.04.2021. godine).
- [112] "Službena web stranica StereoLabs," Poveznica: <https://www.stereolabs.com/>, (Posjeta: 22.12.2019. godine).
- [113] L. I. Smith, "A Tutorial on Principal Components Analysis," Tech. Rep., 2002.
- [114] C. Stauffer and W. E. L. Grimson, "Adaptive Background Mixture Models for Real-time Tracking," in *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition (Cat. No PR00149)*, vol. 2. IEEE, 1999, pp. 246–252.
- [115] "Sveučilište Edinburgh, Škola informatike," Poveznica: <http://homepages.inf.ed.ac.uk/rbf/CVonline>, (Posjeta: 23.12.2019. godine).

- [116] "Ugradbeni sustav Jetson," Poveznica: <https://www.nvidia.com/en-us/autonomous-machines/embedded-systems/>, (Posjeta: 03.09.2021. godine).
- [117] D. Verma and S. Dubey, "Static Signature Recognition System for User Authentication Based Two Level Cog, Hough Transform and Neural Network," *International Journal of Engineering Sciences & Emerging Technologies*, vol. 6, no. 3, pp. 335–343, 2013.
- [118] N. Waltham, "CCD and CMOS Sensors," in *Observing Photons in Space*. Springer, 2013, pp. 423–442.
- [119] L. Wang, T. Tan, H. Ning, and W. Hu, "Silhouette Analysis-based Gait Recognition for Human Identification," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 25, no. 12, pp. 1505–1518, 2003.
- [120] "Web stranica grauonline.de," Poveznica: http://grauonline.de/wordpress/?page_id=3065, (Posjeta: 31.10.2020. godine).
- [121] K. Q. Weinberger, J. Blitzer, and L. K. Saul, "Distance Metric Learning for Large Margin Nearest Neighbor Classification," in *Advances in Neural Information Processing Systems*, 2006, pp. 1473–1480.
- [122] R. P. Wildes, "Iris Recognition: An Emerging Biometric Technology," *Proceedings of the IEEE*, vol. 85, no. 9, pp. 1348–1363, 1997.
- [123] L. Xueyan and G. Shuxu, "The Fourth Biometric-Vein Recognition," in *Pattern Recognition Techniques, Technology and Applications*. InTech, 2008.
- [124] S. Yu, H. Chen, E. B. Garcia Reyes, and N. Poh, "Gaitgan: Invariant Gait Feature Extraction Using Generative Adversarial Networks," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, 2017, pp. 30–37.
- [125] S. Yu, D. Tan, and T. Tan, "A Framework for Evaluating the Effect of View Angle, Clothing and Carrying Condition on Gait Recognition," in *18th International Conference on Pattern Recognition (ICPR)*, vol. 4. IEEE, 2006, pp. 441–444.

- [126] A. Zeng, "Princeton Sveučilište, Odjel za računalne nauke," Poveznica: <https://www.cs.princeton.edu/~andyz/algorithms.pdf>, (Posjeta: 23.12.2019. godine).
- [127] C. Zhang, W. Liu, H. Ma, and H. Fu, "Siamese Neural Network Based Gait Recognition for Human Identification," in *2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2016, pp. 2832–2836.
- [128] X. Zhang, S. Sun, C. Li, X. Zhao, and Y. Hu, "Deepgait: A Learning Deep Convolutional Representation for Gait Recognition," in *Chinese Conference on Biometric Recognition*. Springer, 2017, pp. 447–456.
- [129] Z. Zhang, "Microsoft Kinect Sensor and its Effect," *IEEE Multimedia*, vol. 19, no. 2, pp. 4–10, 2012.
- [130] S. Zheng, J. Zhang, K. Huang, R. He, and T. Tan, "Robust View Transformation Model for Gait Recognition," in *International Conference on Image Processing (ICIP)*. IEEE, 2011, pp. 2073–2076.
- [131] M. F. Zibran, "Eye Based Authentication: Iris and Retina Recognition," Technical Report# 2011-04, University of Saskatchewan, Tech. Rep., 2009.

Popis slika

1.1	Primjer videosekvencije, RGB i dubinske slike	2
1.2	Primjeri normaliziranih i poravnatih okvira silueta u različitim sljedovima hoda s prosječnom slikom siluete (GEI) [28]	6
1.3	Primjer prosječne slike u bazi podataka te prosječne slike osobe koja nosi torbu na ramenu – Iwashita, Uchino i Kurazume [38]	10
1.4	Primjer kreiranih GEI i ekvivalentnih BGEI slika [92]	11
2.1	Biometrijske osobine koje se upotrebljavaju pri identifikaciji	22
2.2	Podjela najčešće primjenjivanih biometrijskih metoda	24
2.3	Prikaz ciklusa hoda osobe	30
2.4	Sustav identifikacije na temelju hoda	31
3.1	Primjeri RGB-D kamera	42
3.2	Stereovid	43
3.3	Epipolarna geometrija	43
3.4	Stereo rektifikacija	44
3.5	Mapa dispariteta	45
3.6	Princip rada TOF kamere	46
3.7	Ilustracija sustava sa strukturiranom svjetlošću	48
3.8	Ilustracija Kinectove mogućnosti praćenja šest osoba [105]	49
3.9	Dubinska slika dobivena Kinectom	50
3.10	Ilustracija praćenja kostura osobe Kinectom	50
4.1	Primjer osobe u hodu iz skupa podataka Casia Dataset B	53
4.2	Primjer silueta za četiri osobe iz skupa podataka Casia Dataset B	53
4.3	Primjeri okvira iz skupa podataka TUM-GAID na jednoj osobi u šest različitih varijacija	55
4.4	Stereo okviri za nekoliko osoba iz vlastitog skupa podataka kreirani ZED stereo kamerom	57
4.5	Primjer videozapisa s jednog toka stereo kamere za nekoliko osoba iz vlastitog skupa podataka kreirani ZED stereo kamerom	58
4.6	Prikaz RGB-D slike u razvojnom okruženju (ZED Depth Viewer)	59
4.7	Prikaz silueta izlučenih iz vlastitog skupa podataka	59

5.1	Koraci pri izlučivanju značajki	62
5.2	Primjer maske prednje strane dobivene iz vlastitog skupa podataka . . .	64
5.3	Prikaz normalizacije silueta	66
5.4	GEI slike za tri osobe na primjeru triju skupova podataka	69
5.5	Prikaz dubinske slike gdje je kodirana informacija o udaljenosti, odnos između stvarne udaljenosti i procjene dubinske kamere	70
5.6	Koraci pri smanjenju dimenzionalnosti podataka	71
5.7	Koraci obrade pri integraciji GEI značajki i visine osobe	75
5.8	Koraci obrade pri fuziji GEI značajki i visine osobe	76
6.1	Primjer dobivene siluete iz RGB slike na vlastitom skupu podataka . . .	79
6.2	Primjer dobivene siluete iz dubinske slike [120]	80
6.3	Primjer određivanja visine osobe i dužine koraka	81
6.4	Prikaz dobivenih GEI i BGEI slika	83
6.5	Razdiobe visina upotrebljavane u eksperimentima	84
6.6	Rezultati klasifikacije značajke visina za pet različitih razdioba visina i tri različite veličine populacije	90
6.7	Prosječna točnost identifikacije uspoređenih metoda za malu veličinu populacije (32 osobe)	97
6.8	Prosječna točnost identifikacije uspoređenih metoda za srednju veličinu populacije (155 osoba)	97
6.9	Prosječna točnost identifikacije uspoređenih metoda za veliku veličinu populacije (305 osoba)	98
6.10	Usporedba između dobivenih rezultata na maloj, srednjoj i velikoj popu- laciji za predložene metode (HGEI-i i HGEI-f) i GEI metodu	98
6.11	Usporedba dobivenih rezultata na maloj, srednjoj i velikoj populaciji za HBGEI-i, HBGEI-f i BGEI metodu	99
6.12	Usporedba između dobivenih rezultata na maloj, srednjoj i velikoj popu- laciji za sve primijenjene metode	100
6.13	Prikaz rezultata HGEI-i po razdiobama visina za malu, srednju i veliku populaciju	101
6.14	Prikaz rezultata HGEI-f po razdiobama visina za malu, srednju i veliku populaciju	101

Popis tablica

2.1	Sedam stupova biometrije [36]	20
2.2	Usporedba različitih biometrijskih metoda [41]	35
6.1	Broj kreiranih GEI i BGEI slika za skup podataka TUM-GAID	82
6.2	Usporedba klasifikacijskih algoritama koristeći GEI značajke	85
6.3	Rezultati klasifikacije testiranih metoda na razdiobi visina <i>Stvarne</i>	91
6.4	Rezultati klasifikacije testiranih metoda na razdiobi visina <i>Slične</i>	92
6.5	Rezultati klasifikacije testiranih metoda na razdiobi visina <i>Gauss</i>	93
6.6	Rezultati klasifikacije testiranih metoda na razdiobi visina <i>Odrasli</i>	94
6.7	Rezultati klasifikacije testiranih metoda na razdiobi visina <i>Godine</i>	95

A Prilog

IMPLEMENTACIJSKI DETALJI POSTUPAKA U PREDLOŽENIM METODAMA

U nastavku su opisani implementacijski detalji vezani uz postupke upotrebljavane tijekom istraživanja i za realizaciju predloženih metoda. Za odabrane postupke navode se prema potrebi upotrebljavani alati, okruženje za razvoj, upotrebljavane knjižnice funkcija, izvodi programskog koda za bitne dijelove postupaka te odabrani konfiguracijski parametri.

A.1 Izlučivanje silueta

U ovoj doktorskoj disertaciji za potrebe izlučivanja silueta osoba upotrebljavani su programski paket Matlab i programski jezik Python. Za izlučivanje silueta iz RGB slika realiziran je program u Matlab programskom paketu. Pored toga, isti program je implementiran i s Python programskim jezikom i dostupnim knjižnicama funkcija.

U slučaju programskog paketa Matlab, postoji skupina alata za računalni vid (engl. Matlab Computer Vision System Toolbox) koja integrira različite funkcionalnosti za tu namjenu i koja se upotrebljavala u navedenom programu. Program Matlab realiziran je tako da je u prvom koraku potrebno učitati, odnosno odabrati videozapis (u *.avi* ili nekom sličnom formatu) ili videosekvenciju slika, dok je ostali dio rada programa automatiziran. Učitavanje videozapisa se, primjerice, radi na sljedeći način:

```
filenameVideoChooser = uigetfile ( ' *.avi ', ' Odaberite ! ' );
```

Umjesto videozapisa ili videosekvencije može se učitati i jedna slika, a može se raditi i u stvarnom vremenu. Jedna slika se učitava na sljedeći način (upotrebljava se funkcija *imread*):

```
I = imread (N);
```

pri čemu je *N* naziv slike (primjerice, *'Slika1.png'*). U slučaju kada se upotrebljava videozapis program također ide kroz svaki okvir koji u suštini predstavlja jednu sliku, a isto je i u slučaju videosekvencije. U navedenom Matlab programu upotrebljavaju se ugrađeni alati *vision.ForegroundDetector* i *vision.BlobAnalysis* iz navedene skupine alata Matlab za računalni vid. Dovoljno ih je definirati na sljedeći način:

```

fgDetector = vision.ForegroundDetector(...
'NumTrainingFrames', 10, 'InitialVariance', 0.05);
blobAnalyzer = vision.BlobAnalysis('AreaOutputPort',
true, ...
'MinimumBlobArea', 70, 'CentroidOutputPort', true);

```

pri čemu su: *NumTrainingFrames* broj početnih okvira za treniranje modela pozadine, *InitialVariance* početna varijanca modela mješavine, *AreaOutputPort* odnosi se na blob područje (engl. Binary Large Object) - skupina povezanih piksela na binarnoj slici, *MinimumBlobArea* odnosi se na minimalno blob područje u pikselima i *CentroidOutputPort* odnosi se na koordinate blob centra. Navedena dva alata (preciznije, radi se o funkcijama) upotrebljavaju se u svrhu detektiranja prednje strane i određivanja jesu li određeni pikseli dio pozadine ili prednje strane te određivanja svojstava povezanih područja. Navedeni upotrebljavani alat *vision.ForegroundDetector* razvijen je na GMM metodi te radi tako što se okvir u boji ili u nijansama sive boje uspoređuje s modelom pozadine da bi se utvrdilo jesu li određeni pikseli dio pozadine ili prednje strane. Pri identifikaciji osoba na temelju hoda cilj je tim postupkom dobiti što bolju siluetu osobe na binarnoj slici. Da bi ostala samo silueta osobe, potrebno je ukloniti i sve manje objekte s binarne slike koji također često ostanu na slici. Za to je korištena ugrađena funkcija u Matlabu (naziv funkcije je *bwareaopen*) koja uklanja sve objekte koji imaju manje od zadanog broja piksela. Navedena funkcija se upotrebljava na sljedeći način:

```
I1 = bwareaopen(I, P);
```

gdje su: *I1* nova slika na kojoj su uklonjeni neželjeni objekti, *I* je slika s koje se uklanjaju objekti, dok je *P* parametar (broj) koji definira da će se sa slike ukloniti svi objekti koji imaju manje od tog broja piksela (primjerice, *P* je za vlastiti skup podataka obično bio 50).

I nakon primjene navedene funkcije na binarnim slikama znaju ostati određeni objekti koji nisu potrebni pa je, kako bi se izlučila samo silueta osobe, upotrebljavana i funkcija koja izlučuje objekte s binarne slike prema zadanoj veličini (naziv funkcije je *bwareafilt*). Navedena funkcija se upotrebljava na sljedeći način:

```
I1 = bwareafilt(I, B);
```

pri čemu su: *I1* nova slika koja sadržava samo željene objekte, *I* slika s koje se

izlučuju željeni objekti i B je broj objekata koji se izlučuju (primjerice, $B = 2$ gdje se izlučuju dva najveća objekta na slici).

Budući da silueta osobe ima određenih nedostajućih područja, ta se područja što je moguće više popunjavaju s pomoću ugrađene funkcije *imfill* u Matlabu. Navedena funkcija se upotrebljava na sljedeći način:

```
I1 = imfill(I, L);
```

pri čemu su: $I1$ nova slika gdje su popunjene praznine na slici, I slika na kojoj se popunjavaju praznine, L su linearni indeksi koji identificiraju lokacije piksela.

Ako se upotrebljava videozapis, siluete se izlučuju kroz svaki okvir videozapisa i tako se dobiju siluete kroz cijeli ciklus hoda. Ako se učitava samo slika, onda se izlučuje silueta osobe samo za tu sliku. Isti je princip i za videosekvenciju slika. Spremanje svake slike vrši se funkcijom *imwrite*.

```
imwrite(I, fullfile(L, N));
```

pri čemu su: I slika koja se sprema, L je lokacija na računalu gdje se slika sprema i N je naziv pod kojim se slika sprema.

Prethodno opisani postupak izlučivanja siluete (ili silueta) odnosi se na jedan videozapis ili jednu videosekvenciju, odnosno na jednu osobu. Ako se želi izlučiti silueta za narednu osobu, potrebno je ponovo učitati videozapis ili videosekvenciju koji sadržava drugu osobu u hodu; ponavlja se cijeli postupak. Također se može učitati i videozapis ili videosekvencija koji sadržava više osoba u hodu te se u tom slučaju izlučuju siluete za svaku detektiranu osobu.

Drugi program za izlučivanje silueta osoba realiziran je s pomoću programskog jezika Python, zajedno s knjižnicom OpenCV (engl. Open Source Computer Vision Library). OpenCV je softverska knjižnica otvorenog koda (engl. Open Source) za računalni vid i strojno učenje. Python program ima istu funkcionalnost kao prethodno navedeni program realiziran u programskom paketu Matlab (rezultat izvršavanja u oba programa jest maska prednje strane, odnosno, u konačnici, siluete osobe). Da bi se definirao, primjerice, videozapis koji će se upotrebljavati koristi se naredna linija koda:

```
VZ = cv.VideoCapture(N)
```

pri čemu N predstavlja naziv videozapisa (primjerice, 'Osoba1.avi').

U ovom programu upotrebljavana je klasa *cv::BackgroundSubtractorMOG2* te funkcija *createBackgroundSubtractorMOG2()*. Navedena funkcija se upotrebljava na sljedeći način:

```
FG = cv.createBackgroundSubtractorMOG2 ( )
```

Upotrebljavala se metoda za oduzimanje pozadine zasnovana na modelu Gaussove mješavine (engl. Mixture of Gaussian 2, MOG2). Umjesto metode MOG2 u programu je istražena upotreba i metode MOG (engl. Mixture of Gaussian) i metode kNN (upotrebljavaju se funkcije *createBackgroundSubtractorMOG()* ili *createBackgroundSubtractorKNN()*), no MOG2 je davao najbolje rezultate u smislu cjelovitosti dobivene siluete osobe s manje šuma na slikama. Navedeni Python program razvijen je u cilju usporedbe dobivenih rezultata u odnosu na program razvijen u programskom paketu Matlab, a rezultati su u suštini identični.

U slučaju stereovida (općenito o stereovidu i pojmovima vezanim za stereovid može se vidjeti u potpoglavlju 3.2.1) to je uključivalo razvijanje programa u programskom paketu Matlab. Da bi se efikasno realizirao postupak izlučivanja silueta osoba, potrebno je imati stereoparametre, odnosno vrijednosti koje se odnose na kalibraciju kamere. Realizacija programa koji izlučuje siluete osoba, a čiji su ulazni podatci slike s lijeve i desne strane stereokamere, radi na način da se učitava videozapis (umjesto videozapisa mogu se učitati videosekvencije ili samo dvije slike, s lijeve i s desne strane stereokamere) koji sadržava lijevi i desni tok, odnosno sliku s lijeve i desne strane te je proces dalje automatiziran. Videozapisi se učitavaju na način kako je to već prikazano za Matlab na početku ovog priloga. Isti princip je i kada se upotrebljavaju videosekvencije. Ako se učitava videozapis, onda se proces odvija kroz svaki okvir videozapisa. Za potrebnu rektifikaciju slika koje se dobiju s lijeve i desne strane upotrebljava se funkcija *rectifyStereoImages* koja kao parametre uzima lijevu i desnu sliku te stereoparametre. Definira se na sljedeći način:

```
[S1,S2] = rectifyStereoImages (I1 , I2 ,SP);
```

pri čemu su: *S1* i *S2* slike *I1* i *I2* nad kojima je izvršena rektifikacija dok su *SP* stereoparametri.

Prethodni postupak se radi kako bi se eliminiralo iskrivljenje. Nakon toga dobiju se dvije slike na kojim je izvršena rektifikacija te se može izračunati mapa dispariteta. Za to se upotrebljavala funkcija *disparity* (u starijim verzijama Matlaba upotrebljava se

navedena funkcija, dok se danas preporučuje upotreba funkcija *disparityBM* i *disparitySGM*). Upotrebljavanoj funkciji se kao parametri prosleđuju slike s lijeve i desne strane na kojima je izvršena rektifikacija, ali u nijansama sive boje. Radi se na sljedeći način:

```
disparityMap = disparitySGM( rgb2gray(S1), rgb2gray(S2) );
```

pri čemu su: *S1* i *S2* slike nad kojima je izvršena rektifikacija, a koje su u nijansama sive boje (za pretvorbu RGB slike u sliku koja je u nijansama sive boje upotrebljava se funkcija *rgb2gray*).

Dobivena slika mape dispariteta izgleda kao na slici 3.5. Na toj slici mape dispariteta potrebno je podesiti kontrast definiranjem graničnih vrijednosti.

A.2 PCA i LDA analiza

Za PCA (engl. Principal Component Analysis) i LDA (engl. Linear Discriminant Analysis) analizu upotrebljavana je *scikit-learn* knjižnica funkcija u Python programskom jeziku. Radi se o knjižnici otvorenog koda koja se upotrebljava za strojno učenje u Python programskom jeziku. Navedena *scikit-learn* knjižnica podržava nadzirano i nenadzirano učenje, a pruža različite alate za podešavanje, odabir i evaluaciju modela, početnu obradu podataka itd.

Pored navedene *scikit-learn* knjižnice funkcija upotrebljavani su i dodatni pomoćni paketi za Python, poput *NumPy* i *pandas*. *NumPy* predstavlja Python knjižnicu za rad s nizovima, a posjeduje i funkcije za rad u području linearne algebre, Furijerove transformacije i matrica. *Pandas* se koristi za manipuliranje i analizu podataka.

```
import numpy as np
import pandas as pd
from sklearn.decomposition import PCA
from sklearn.discriminant_analysis import
LinearDiscriminantAnalysis
```

Nakon što se postave parametri

```
PCA_components = [50]
LDA_components = [10]
```

GEI (ili BGEI, u ovisnosti od metode) slike se vektoriziraju.

```
image_vectors[ggt] = np.vstack((image_vectors[ggt],
image.reshape(1, -1)))
```

Nakon toga, pozivom funkcije iz paketa *sklearn.decomposition* se računaju PCA komponente vektoriziranih slika za cijeli skup podataka.

```
pca = PCA(n_components=p_components)
```

Zatim se pozivom i *sklearn.discriminant_analysis* računaju LDA komponente između dobivenih PCA komponenti.

```
lda = LinearDiscriminantAnalysis(n_components=
l_components)
```

A.3 Alfa-integracija

Alfa-integracija implementirana je također u Python programskom jeziku. Upotrebljavala se također knjižnica *scikit-learn*. U postupku alfa integracije težinski se faktori za fuziju optimiziraju za svaku klasu zasebno. Ulaz optimizacijskom algoritmu su vektori koji su dobiveni kao rezultat zasebne klasifikacije s GEI značajkama i sa značajkom visina osobe s vjerojatnostima podudaranja za svaku klasu.

Za optimizaciju se upotrebljavao MADS (engl. Mesh Adaptive Direct Search Algorithm, MADS) [4] optimizacijski algoritam.

```
mads = MADS(params['evaluations'],
params['objective'].penalty,
params['lower_bounds'],
params['upper_bounds'])
mads.run()
```

Maksimalni broj evaluacija

```
params['evaluations']
```

je postavljen na 1500, a upotrebljavaju se gore navedene funkcije za određivanje pogreške i ograničenja.

```
params['objective'].penalty
params['lower_bounds']
params['upper_bounds']
```

Vrijednost alfa parametara je između -1 i 3 kao što je definirano u radu [85]. Optimiziraju se tri parametra po svakoj klasi (alfa, w1 i w2). Parametri upotrebljavani u implementaciji su:

```
lower_bounds = [0.0, 0.0, -1.0]
upper_bounds = [1.0, 1.0, 3.0]
```

Za izračun *sscore* i pogreške koja se minimizira u postupku optimizacije upotrebljavala se funkcija:

```
calculate_s_score(self, alpha, w1, w2, h_pred, g_pred)
```

Za određivanje pogreške upotrebljavala se funkcija:

```
penalty(self, opt_vars)
```

Cijeli izvod koda koji prikazuje upotrebljavanu implementaciju prikazan je u nastavku.

```
class Objective():
    def __init__(self, true_labels, height_probabilities,
                  gait_probabilities):
        self.__true_labels = true_labels
        self.__height_probabilities = height_probabilities
        self.__gait_probabilities = gait_probabilities
        self.__solutions = []

    def __calculate_s_score(self, alpha, w1, w2, h_pred, g_pred):
        if alpha == 1:
            if h_pred == 0 or g_pred == 0:
                s_score = 0
            else:
                inner = (w1 * math.log(h_pred)) +
                        (w2 * math.log(g_pred))
                s_score = math.exp(inner)
        else:
            inner_exp = (1 - alpha) / 2
            outer_exp = 2 / (1 - alpha)
```

```

        s_score = (((w1 * (h_pred ** inner_exp)) +
                     (w2 * (g_pred ** inner_exp)))) ** outer_exp

    return s_score

def __get_error(self, alpha, w1, w2, h_pred, g_pred,
                label, error_type="LMSE"):
    s_score = self.__calculate_s_score(alpha, w1, w2, h_pred,
                                       g_pred)

    if error_type == "LMSE":
        error = (label - s_score) ** 2
    elif error_type == "MPE":
        try:
            error = math.e ** ((label * math.log(s_score)) +
                               ((1. - label) * math.log(1. - s_score)))
        except:
            error = 0
    else:
        error = 0

    return error

def penalty(self, opt_vars):
    w1 = opt_vars[0]
    w2 = opt_vars[1]
    alpha = opt_vars[2]

    errors = []
    for sample in range(len(self.__true_labels)):
        label_sample = self.__true_labels[sample]
        h_sample = self.__height_probabilities[sample]

```

```

g_sample = self.__gait_probabilities[sample]

errors.append(self.__get_error(alpha, w1, w2, h_sample,
g_sample, label_sample, error_type="LMSE"))

error = sum(errors) / len(errors)

# save parameters and error value
self.__solutions.append({'w1': opt_vars[0],
'w2': opt_vars[1],
'alpha': opt_vars[2], 'error': error})

return error

@property
def solutions(self):
    return self.__solutions

```

A.4 Klasifikacija

Za proces klasifikacije ispitano je više algoritama koji su dostupni u knjižnici *scikit-learn*, a tu su uključeni: algoritam klasifikacije potpornim vektorima (engl. Support Vector Classification, SVC) [27] [111], algoritam klasifikacije slučajnih šuma (engl. Random Forest Classification, RFC) [51] [110], algoritam Gaussov naivni Bayes (engl. Gaussian Naive Bayes, GNB) [60] [108] i k-najbližih susjeda (engl. k-Nearest Neighbors, kNN) [109].

SVC implementacija je temeljena na *libsvm*, gdje se vrijeme prilagođavanja skalira barem kvadratno s brojem uzoraka te može biti nepraktičan nakon više desetaka tisuća uzoraka. SVC je klasa koja omogućava binarnu klasifikaciju i klasifikaciju s više klasa [111].

Slučajna šuma je metaprocjenitelj koji odgovara brojnim klasifikatorima stabla odluke na različitim poduzorcima skupa podataka i koristi prosječnu vrijednost u cilju poboljšanja točnosti predviđanja i kontrolu prekomjernog uklapanja. Kod slučajnih šuma

svako stablo u ansamblu gradi se od uzorka izvučenog zamjenom iz skupa za trening [110].

GNB implementira Gaussov naivni Bayes algoritam za klasifikaciju [108].

Klasifikacija koja je zasnovana na susjedima je vrsta učenja koja je zasnovana na primjercima, a koja ne pokušava konstruirati općeniti unutarnji model nego jednostavno pohranjuje primjerke trening podataka. Klasifikacija se izračunava prostom većinom glasova najbližih susjeda svake točke. Drugim riječima, točki koja je upit dodjeljuje se klasa koja ima najviše predstavnika unutar najbližih susjeda točke [109].

```
from sklearn.neighbors import KNeighborsClassifier
from sklearn.ensemble import RandomForestClassifier
from sklearn.naive_bayes import GaussianNB
from sklearn.svm import SVC
```

Na osnovu preliminarne analize [108] [109] [110] [111] i eksperimentalne provjere navedenih algoritama odabran je kNN za klasifikaciju. U implementaciji je upotrebljavana *KNeighborsClassifier* iz knjižnice funkcija *scikit-learn*.

Parametar težina postavljen je na *'distance'* koji predstavlja udaljenost točaka od središta klastera. Dio koda koji prikazuje implementaciju prikazan je u nastavku.

```
classifier_data = {
    'knn': {
        'classifier': KNeighborsClassifier,
        'params': {
            'weights': 'distance',
            'n_neighbors': 0
        }
    },
    'rfc': {
        'classifier': RandomForestClassifier,
        'params': {
            'n_estimators': 0
        }
    },
    'svc': {
```

```
    'classifier': SVC,  
    'params': {  
        'probability': True,  
        'gamma': 'auto'  
    }  
},  
'nb': {  
    'classifier': GaussianNB,  
    'params': {  
    }  
},  
}
```

Životopis

Adnan Ramakić rođen je 1. lipnja 1986. godine u Bihaću, Bosna i Hercegovina. Osnovnu školu završio je u Bihaću 2000. godine s odličnim uspjehom. Nakon toga upisuje Srednju elektrotehničku i drvoprerađivačku školu, također u Bihaću, koju završava 2004. godine s odličnim uspjehom.

Od 2004. do 2009. godine studirao je na Tehničkom fakultetu Univerziteta u Bihaću, na elektrotehničkom odsjeku, smjer: informatika. U travnju 2009. godine diplomirao je s najvišom ocjenom (10) i stekao zvanje diplomirani inženjer elektrotehnike, smjer: informatika.

Nakon završetka studija zapošljava se u Centru za promociju lokalnog razvoja (PLOD Centar) u Bihaću, gdje radi na poslovima administratora informacijskog sustava. Također je bio uključen i u različite projekte iz područja turizma. Istovremeno radi i kao vanjski suradnik (asistent) na Tehničkom fakultetu Univerziteta u Bihaću, na predmetu Baze podataka. Potom se 2010. godine zapošljava u Rektoratu Univerziteta u Bihaću na radno mjesto administratora servisa. Usto je radio i kao vanjski suradnik na Pedagoškom fakultetu Univerziteta u Bihaću, na predmetima Računarske mreže i Informatika.

Tijekom 2011. godine upisuje drugi ciklus studija na Elektrotehničkom fakultetu Univerziteta u Banja Luci koji završava u roku i s najvišom ocjenom (10) brani završni magistarski rad pod nazivom „*Mogućnosti i implementacija zaštite podataka u mikrorračunarskim sistemima i njihovim mrežama*”. Godine 2014. upisuje poslijediplomski doktorski studij na Tehničkom fakultetu Sveučilišta u Rijeci.

Trenutačno je stalno zaposlen u Rektoratu Univerziteta u Bihaću na radnome mjestu stručnog savjetnika za IT poslove. Angažiran je i kao viši asistent (područje Informacione tehnologije i informacioni sistemi) na Tehničkom fakultetu Univerziteta u Bihaću. Predmeti na kojima je dosad bio angažiran su: Osnove programskog jezika Java, Algoritmi i strukture podataka, Objektno-orijentirano programiranje, Sigurnost računarskih sistema, Objektno-orijentirane baze podataka, Elektrotehnika II, Osnove telekomunikacija, Osnove informacionih sistema i Projektiranje mikroprocesorskih sistema.

Ima Oracle certifikate *Database Programming with PL/SQL* i *Database Design and Programming with SQL*.

Aktivno se služi engleskim jezikom.

Područja iz računarstva i informatike s kojim se aktivno bavi su objektno-orijentirano programiranje (C++, Java), sigurnost sustava, razvoj desktop i web aplikacija, računalne mreže, arhitektura računala i računalni hardver, a njegovo su trenutačno područje interesa umjetna inteligencija, strojno učenje, algoritmi strojnog učenja, obrada slika i prepoznavanje uzoraka.

Autor je 18 znanstveno-stručnih radova indeksiranih u relevantnim bazama podataka.

Oženjen je i otac je jednog djeteta.

Popis radova

1. Adnan Ramakić, Zlatko Bundalo (2013): DATA PROTECTION IN MICROCOMPUTER SYSTEMS AND NETWORKS, 11th Internacional Conference on Accomplishments in Electrical and Mechanical Engineering and Information Technology DEMI 2013, Banja Luka, Bosnia and Herzegovina
2. Adnan Ramakić, Zlatko Bundalo (2013): SOFTVERSKO-HARDVERSKA ZAŠTITA PODATAKA U RAČUNARSKIM SISTEMIMA, 9th International Conference Research and Modernization of Production RIM 2013, Budva, Montenegro
3. Adnan Ramakić (2013): KORIŠTENJE NAPREDNIH JAVA TEHNOLOGIJA PRI IZRADI WEB APLIKACIJA, 9th International Conference Research and Modernization of Production RIM 2013, Budva, Montenegro
4. Adnan Ramakić, Zlatko Bundalo (2014): DATA PROTECTION IN MICROCOMPUTER SYSTEMS AND NETWORKS, ACTA TECHNICA CORVINIENSIS – BULLETIN of ENGINEERING, Faculty of Engineering, University of Timisoara, Issue April-June 2014.
5. Adnan Ramakić, Zlatko Bundalo (2014): COMBINED HARDWARE AND CRYPTOGRAPHIC USER DATA PROTECTION IN PERSONAL COMPUTER SYSTEMS AND NETWORKS, 23rd International Electrotechnical and Computer Science Conference ERK, IEEE Slovenia Section together with Faculty of Electrical Engineering University of Ljubljana and other Slovenian professional societies, Portorož, Slovenia, September 22-24, 2014.
6. Adnan Ramakić, Zlatko Bundalo, Dušanka Bundalo (2015): USING MARKOV CHAINS IN COMPUTER SYSTEMS SECURITY APPLICATIONS, 24th International Electrotechnical and Computer Science Conference ERK, IEEE Slovenia Section together with Faculty of Electrical Engineering University of Ljubljana and other Slovenian professional societies, Portorož, Slovenia, September 21-23, 2015.
7. Adnan Ramakić, Zlatko Bundalo (2015): SIGURNOST I IMPLEMENTACIJA ZAŠTITE KOD JAVA BAZIRANIH WEB APLIKACIJA, 10th International Conference Research and Modernization of Production RIM 2015, Dubrovnik, Croatia
8. Amel Toroman, Adnan Ramakić (2015): MATEMATIČKO MODELIRANJE I SI-

MULACIJA ISTOSMJERNOG MOTORA U PROGRAMSKOM PAKETU MATLAB, 10th International Conference Research and Modernization of Production RIM 2015, Dubrovnik, Croatia

9. Adnan Ramakić, Amel Toroman (2016): SISTEMI KUĆNE AUTOMATIZACIJE TEMELJENI NA UPRAVLJANJU POMOĆU GOVORA, 3rd International Conference "New Technologies NT-2016", Mostar, Bosnia and Herzegovina, May 13-14, 2016.
10. Adnan Ramakić, Kristijan Lenac (2016): SHORT-TERM PEOPLE RE- IDENTIFICATION USING IMAGE CLASSIFICATION, International Conference on Smart Systems and Technologies (SST) 2016, Faculty of Electrical Engineering, Computer Science and Information Technology in Osijek, Josip Juraj Strossmayer University of Osijek, Osijek, Croatia (IEEE Region 8, IEEE Croatia Section), October 12-14, 2016. (IEEE Xplore Digital Library)
11. Zlatko Bundalo, Adnan Ramakić, Dušanka Bundalo (2016): INCREASING DESKTOP APPLICATION AND USER DATA PROTECTION USING SMARTPHONE, 24th Telecommunications Forum (TELFOR) 2016, Telecommunications Society - TS, Belgrade, ETF - School of Electrical Engineering, University of Belgrade, IEEE Serbia and Montenegro COM CHAPTER (IEEE Region 8), Belgrade, Serbia, November 22-23, 2016. (IEEE Xplore Digital Library)
12. Adnan Ramakić, Amel Toroman (2017): PEOPLE IDENTIFICATION USING KINECT SENSOR, International Conference on Biological and Medical Engineering (CMBEBIH) 2017, Medical and Biological Society in Bosnia and Herzegovina (DMBIUBIH), Sarajevo, Bosnia and Herzegovina, March 16-18, 2017. (IFMBE Proceedings by Springer Nature)
13. Adnan Ramakić, Diego Sušan, Kristijan Lenac (2017): GEOLOCATION USING GOOGLE VISION API, 10th Annual RIN Baška GNSS Conference, The Royal Institute of Navigation, London, UK, Faculty of Maritime Studies, University of Rijeka and Faculty of Transport and Traffic Sciences, University of Zagreb, Baška, Krk Island, Croatia, May 8-10, 2016. (Proceedings is published in 2017 on 11th Annual RIN Baška GNSS Conference)
14. Adnan Ramakić, Zlatko Bundalo, Dušanka Bundalo, Bojan Rakita (2019): PEOPLE IDENTIFICATION BASED ON PERSON IMAGE AND ADDITIONAL PHY-

- SICAL PARAMETERS COMPARISON, TEM Journal, 8(1), 87-93, February 2019.
15. Zlatko Bundalo, Momčilo Veljko, Dušanka Bundalo, Goran Kuzmić, Mirko Sajić and Adnan Ramakić (2019): ENERGY EFFICIENT EMBEDDED SYSTEMS FOR LED LIGHTING CONTROL IN TRAFFIC, The 8th Mediterranean Conference on Embedded Computing – MECO 2019, June 10-14, 2019.
 16. Adnan Ramakić, Zlatko Bundalo, Dušanka Bundalo (2019): A METHOD FOR HUMAN GAIT RECOGNITION FROM VIDEO STREAMS USING SILHOUETTE, HEIGHT AND STEP LENGTH, Journal of Circuits, Systems and Computers, Vol 29, No. 7, 15 June 2020.
 17. Kristijan Lenac, Diego Sušan, Adnan Ramakić, Domagoj Pinčić (2019): EXTENDING APPEARANCE BASED GAIT RECOGNITION WITH DEPTH DATA, Appl. Sci. 9, no. 24: 5529.
 18. Adnan Ramakić, Diego Sušan, Kristijan Lenac, Zlatko Bundalo (2020): DEPTH-BASED REAL-TIME GAIT RECOGNITION, Journal of Circuits, Systems and Computers, Vol 29, No. 16, 31 December 2020.