



Univerzitet u Nišu
Mašinski Fakultet



Ivan T. Ćirić

**Inteligentno upravljanje mobilnim robotima na
osnovu neuro-fazi-genetskog prepoznavanja
objekata i praćenja ljudi u robotskoj viziji**

— doktorska disertacija —

Niš, 2015.



University of Niš
Faculty of Mechanical Engineering



Ivan T. Ćirić

**Intelligent Control of Mobile Robots based on
Neuro-Fuzzy-Gentic Object Recognition and Human
Tracking in Robot Vision System**

— doctoral dissertation —

Niš, 2015.

Ani i Milici

Podaci o doktorskoj disertaciji

1. Autor	
Ime i prezime	Ivan Ćirić
Datum i mesto rođenja	06.06.1980. godine u Prokuplju
2. Doktorska disertacija	
Naslov	Inteligentno upravljanje mobilnim robotima na osnovu neuro-fazi-genetskog prepoznavanja objekata i praćenja ljudi u robotskoj viziji
Broj strana	169
Broj slika	96
Broj tabela	8
Broj literaturnih citata	120
Naučno polje	Tehničko-tehnološke nauke
Naučna oblast	Mašinsko inženjerstvo
Uža naučna oblast	Automatsko upravljanje i robotika
UDK broj	621-52:007.52(043.3) 004.85:004.8.032.26]:591.87+510.644(043.3) 004.932:621.397(043.03)
CERIF	T125, T111, P176
Mentor	dr Žarko Ćojbašić, red. prof. Mašinskog fakulteta Univerziteta u Nišu
Članovi komisije za pregled, ocenu i odbranu doktorske disertacije	1) dr Dragan Antić, red. prof., Elektronski fakultet, Univerzitet u Nišu 2) dr Žarko Ćojbašić, red. prof., Mašinski fakultet, Univerzitet u Nišu 3) dr Vlastimir Nikolić, red. prof., Mašinski fakultet, Univerzitet u Nišu 4) dr Danijela Ristić, doc., Mašinski fakultet, Univerzitet u Nišu 5) dr Stevan Stankovski, red. prof., Fakultet tehničkih nauka, Univerzitet u Novom Sadu
Datum odbrane	15.12.2015. godine

Dissertation data

1. Author	
Name and Surname	Ivan Ćirić
Date and place of birth	06/06/1980, Prokuplje
2. Doctoral dissertation	
Title	Intelligent Control of Mobile Robots based on Neuro-Fuzzy-Gentic Object Recognition and Human Tracking in Robot Vision System
N° pages	169
N° figures	96
N° tables	8
N° references	120
Scientific field	Engineering sciences and technology
Scientific field	Mechanical engineering
Narrow scientific field	Automatic control and robotics
UDK	621-52:007.52(043.3) 004.85:004.8.032.26]:591.87+510.644(043.3) 004.932:621.397(043.03)
CERIF	T125, T111, P176
Mentor	dr Žarko Ćojbašić, full professor, Faculty of Mechanical Engineering, University of Niš
Members of Commission for thesis examination, evaluation and defense	1) dr Dragan Antić, full professor., Faculty of Electronic Engineering, University of Niš 2) dr Žarko Ćojbašić, full professor, Faculty of Mechanical Engineering, University of Niš 3) dr Vlastimir Nikolić, full professor, Faculty of Mechanical Engineering, University of Niš 4) dr Danijela Ristić, doc., Faculty of Mechanical Engineering, University of Niš 5) dr Stevan Stankovski, full professor, Faculty of Technical Sciences, University of Novi Sad
Date of public presentation	15/12/2015

Zahvalnica

Ova teza je nastala kao rezultat mog višegodišnjeg istraživanja koje sam obavio na Mašinskom fakultetu u Nišu i na Institutu za automatiku Univerziteta u Bremenu. Istraživanja su sprovedena u okviru bilateralnih projekata između Mašinskog fakulteta Univerziteta u Nišu i Instituta za automatiku Univerziteta u Bremenu koji su realizovani u okviru programa „PPP Serbien“. Rukovodilac ovih projekata sa srpske strane je prof. dr Vlastimir Nikolić na čijoj sam ogromnoj pomoći i podršci izuzetno zahvalan. Zahvalnost dugujem i svom mentoru, prof. dr Žarku Čojbašiću koji me je za sve vreme istraživanja i izrade ove teze vodio i usmeravao. Zahvalio bih se i kolegici docentu dr Danijeli Ristić Durrant na velikoj pomoći tokom istraživanja koje sam sproveo u Bremenu, kao i Adrianu Leuu i Siniši Slavniću sa Instituta za Automatiku Univerziteta u Bremenu. Bez njihove pomoći, znanja i sugestija ne bih mogao završiti ovu tezu. Zahvalio bih i kolegama sa Mašinskog fakulteta, Emini Petrović, Milošu Simonoviću, Miši Tomiću i dr Milanu Baniću, koji su mi tokom istraživanja i izrade ove disertacije nesebično pomagali. Posebnu zahvalnost dugujem svom dragom kolegi i kumu, dr Ivanu Pavloviću, na svojoj pomoći i podršci koju mi je pružio kako u Bremenu, tako i u Nišu.

Konačno, zahvaljujem ćerki Ani i supruzi Milici na ljubavi, podršci i neizmernom strpljenju koje su mi pružili tokom istraživanja i samog pisanja ove teze.

Rezime

Sistem robotske vizije se u poslednje vreme sve češće javlja kao važan sastavni deo robotskih sistema i predstavlja osnovni izvor informacija na osnovu koga robotski sistem vrši percepciju okoline u cilju njegove odgovarajuće interakcije sa objektima i ljudima. I pored velikog broja prednosti vizuelni sistem ima i nedostatke koji se ogledaju u maloj robusnosti sistema vizije u odnosu na različite spoljašnje uticaje, jako komplikovanoj analizi prenatrpanih scena ili potrebi prepoznavanja objekta prilikom varijacije njegove forme. S obzirom da aktuelno stanje istraživanja u svetu ukazuju na veliki potencijal razvoja i implementacije algoritama zasnovanih na veštačkoj inteligenciji i mašinskom učenju kod sistema robotske vizije, a u cilju prepoznavanja objekata i praćenje ljudi, ovaj naučnoistraživački pravac predstavlja jednu od polaznih oblasti istraživanja u okviru rada na predmetnoj doktorskoj disertaciji.

Tokom istraživanja bilo je potrebno razviti sistem vizije kojim bi robot dobio neophodne informacije za pouzdano i autonomno funkcionisanje u dinamičnom okruženju. Sistem vizije mora da omogući ekstrakciju i klasifikaciju osobina različitih objekata koji su prisutni u sceni bez potrebe za velikim bazama podataka, koje zahteva većina postojećih robotskih sistema. U tom cilju razvijen je veći broj novih algoritama inteligentnog mobilnog robota zasnovanih na neuronskim mrežama, fazi logici i genetskim algoritmima, kao i drugim alatima iz domena veštačke inteligencije kojima sistem robotske vizije detektuje objekte i ljude, prati ih i na osnovu toga izvršava unapred definisane zadatke.

Istraživanja predstavljena u ovoj tezi su usmerena na razvoj mobilne robotske platforme i njenog upravljačkog sistema na bazi inteligentnih algoritama robotske vizije u cilju rešavanja problema prepoznavanja tj. klasifikacije objekata i ljudi u laboratorijskom okruženju. Inicijalna istraživanja vršena su sa robotskim sistemom kod koga sistem vizije daje osnovne podatke za upravljanje manipulatorom, a razvijeni algoritmi su bili polazna osnova za razvoj algoritama za detekciju i praćenje ljudi mobilnim robotom.

Inicijalno istraživanje podrazumevalo je prepoznavanje objekata. U tu svrhu korišćena je monokularna kamera kod mobilne robotske platforme DaNI i stereo kamera kod robotske platforme sa manipulatorom FRIEND koja se koristi za pomoć osobama sa invaliditetom i razvijeni su algoritmi segmentacije i inteligentni sistemi klasifikacije kako bi ove robotske platforme mogle da interaguju s željenim objektima.

Kasnija istraživanja podrazumevala su prepoznavanje i praćenje ljudi, te su razvijene eksperimentalne robotske platforme sa stereo kamerom, sa termovizijskom kamerom i sa 3D

senzorom sa struktuiranom svetlošću. U ovoj disertaciji predstavljeno je unapređenje modula robotske vizije klasifikatorom zasnovanim na neuronskoj mreži i predstavljen je koncept primene genetskih algoritama za optimalan izbor ulaznih karakteristika u klasifikator. Primena termovizijske kamere u robotskim sistemima za prepoznavanje i praćenje ljudi nije rešenje koje se često sreće u naučnim istraživanjima, pa je bilo potrebno razviti novi algoritam fazi segmentaciju termovizijske slike, kao i genetski algoritama za optimalno podešavanje parametara segmentacije slike dobijene termovizijskom kamerom. Razvijen je i softver zasnovan na metodi potpunih vektora za klasifikaciju objekata sa termovizijske slike kojim se zapravo vrši prepoznavanje ljudi u sceni.

U cilju unapređenja praćenja razvijena je nelinearna auto-regresivne neuronska mreža za ocenu i predviđanje položaja čoveka snimljenog stereo kamerom. Konačno predložena su i rešenja inteligentnog upravljanja mobilnom robotskom platformom za praćenje ljudi koji informacije dobija od 3D senzora sa struktuiranom svetlošću, a ovakva mobilna robotska platforma može da se koristi i za prepoznavanje gestova.

Rezultati predstavljeni u ovoj doktorskoj disertaciji potvrđuju da je primenom savremenih metoda i algoritama iz domena veštačke inteligencije moguće razviti sistem upravljanja zasnovanog na robotskoj viziji koji omogućava mobilnom robotu da detektuje, prepozna i prostorno locira objekte i ljude, što dalje omogućava inteligentnu interakciju ovog robotskog sistema sa okolinom. Implementacijom više tehnika veštačke inteligencije, kao što su neuronske mreže, fazi logika i genetski algoritmi, u proces razvoja inteligentnog upravljačkog sistema moguće je obezbediti autonomno i pouzdano prepoznavanje objekata i ljudi i njihovo praćenje koje je superiorno u odnosu na postojeća rešenja.

Ključne reči: robotska vizija, upravljanje mobilnim robotom, veštačka inteligencija, mašinsko učenje, neuronske mreže, genetski algoritmi, fazi logika, SVM, NARX, digitalna obrada slike, segmentacija slike, klasifikacija, stereo vizija, termovizija.

Abstract

Robot Vision system is lately an important part of the robot system and it becomes main source of the data used for robot environment perception in order to adequate interaction between robot and surrounding objects and humans. There are numerous advantages of the vision system, but there are also disadvantages like unrobustness of the vision system considering external influences, very complex analysis of cluttered scenes or object recognition when object shape changes. Since current state of the art shows growing potential of development, improvement and implementation of robot vision system based on artificial intelligence and machine learning, for object recognition and human tracking, this field of scientific research is used as a basis for research presented in this doctoral dissertation.

During the research it was mandatory to develop vision system that collects data and enables robot to reliably and autonomously function in dynamical environment. Vision system needs to enable feature extraction and classification of different objects from the scene without large databases commonly used in modern robot systems. Therefore, several new algorithms based on neural networks, fuzzy logic and genetic algorithms (as well as some other artificial intelligence algorithms) were developed and used in intelligent mobile robot platform for object and human detection, tracking and interaction.

Presented research is guided towards development of mobile robot platform and its control system based on intelligent robot vision algorithms for object and human recognition (classification) and tracking in lab environment. Initial research were done with robot platform where vision system is used as main sensor for manipulator control, and developed algorithms were basis for development of algorithms for human detection and tracking by mobile robot.

Initial research was done in order to enable object recognition for DaNI mobile robot platform with monocular camera and FRIEND robot system with manipulator equipped with stereo camera. For adequate interaction of robot with surrounding objects segmentation algorithms and intelligent classification algorithms were developed.

Further research was done for human tracking, so experimental mobile robot platforms with stereo camera, thermal camera and RGBD sensor were developed. In this dissertation improvement of robot vision module by neural network classificatory is presented as well as concept of GA-optimal neural network input selection. Implementation of thermal camera for human tracking into mobile robot system is not very common, so development of new fuzzy-segmentation algorithm and new genetic algorithm for optimal segmentation parameter tuning

of thermal image was done. Also, software based on Support Vector Machine (SVM) was developed for classification of objects in thermal image (human detection in thermal image).

As an improvement of human tracking algorithm, in this dissertation nonlinear autoregressive neural network is suggested for prediction and estimation of human position filmed by stereo camera. Finally, some solutions for intelligent control of mobile robot platform equipped with RGBD sensor for human tracking were suggested and with adequate adaptation this platform can be used for gesture recognition.

Results presented in this doctoral dissertation confirm that implementation of modern methods and algorithms from the palette of artificial intelligence and machine learning tools allows development of vision based control system that enables mobile robot to detect, recognize and locate objects and humans, that further enables intelligent interaction of robot system with surrounding environment. Implementation of more than one artificial intelligence technique (like artificial neural networks, fuzzy logic and genetic algorithms) into intelligent control system development enables autonomous and reliable object and human recognition and tracking that is superior compared to current solutions.

Key words: Robot Vision, Mobile Robot Control, Artificial Intelligence, Machine Learning, Artificial Neural Network, Genetic Algorithms, Fuzzy Logic, Support Vector Machine, NARX, Digital Image Processing, Image segmentation, Classification, Stereo Vision, Thermal Imaging.

Sadržaj

1. UVOD	1
1.1 Motivacija	2
1.2. Problem	2
1.3. Predložena rešenja	3
1.4. Publikacije	4
1.5. Pregled	9
2. DIGITALNA AKVIZICIJA I OBRADA SLIKE U ROBOTICI	10
2.1. Digitalne slike	11
2.2. Akvizicija digitalne slike	12
2.3. Boja u kompjuterskoj viziji	13
2.4. Analiza i obrada slike u prostornom i frekventnom domenu	18
2.5. Senzori za prepoznavanje okoline i 3D senzori	20
2.6. 3D i Stereo vizija	23
2.7. Segmentacija slike	24
3. DETEKCIJA I PREPOZNAVANJE OBJEKATA	26
3.1. Pregled istraživanja	26
3.2. Stereovizija i kalibracija stereo kamere	29
3.3. Određivanje mape dispariteta i dubinske slike	32
3.4. Sistemi za robusnu segmentaciju slika u robotskoj viziji	37
3.4.1. Segmentacija zasnovana na određivanju praga	37
3.4.2. Segmentacija u zatvorenom kolu	41
3.4.3. Segmentacija mape dispariteta.	45
3.5. Određivanje karakteristika segmentirane slike	51
3.6. Klasifikacija i prepoznavanje objekata na osnovu karakteristika	53

4. INTELIGENTNI SISTEMI ZA SEGMENTACIJU SLIKE I PREPOZNAVANJE OBJEKATA	55
4.1. Inteligentno prepoznavanje i praćenje objekata eksperimentalnom robotskom platformom sa monokularnom kamerom	57
4.1.1. Eksperimentalna robotska platforma DaNI sa monokularnom kamerom	57
4.1.2. Segmentacija slike dobijene monokularnom kamerom robotske platforme	59
4.1.3. Klasifikacija i prepoznavanje objekata	60
4.1.4. Praćenje objekata u sceni	64
4.2. Eksperimentalna robotska platforma FRIEND sa stereo kamerom	66
4.2.1. Segmentacija slike u povratnoj sprezi i izdvajanje karakteristika za klasifikaciju	71
4.2.2. ANFIS klasifikator za prepoznavanje objekata u scenariju FRIEND robotske platforme	74
4.2.3. Klasifikator zasnovan na neuronskoj mreži za prepoznavanje objekata u scenariju FRIEND robotske platforme	78
4.2.4. Naivni Bajesov klasifikator za prepoznavanje objekata mobilnom robotskom platformom FRIEND	82
4.3. Prepoznavanje objekata robotskom platformom FRIEND na osnovu karakteristika segmentirane mape dispariteta	86
4.3.1. Segmentacija mape dispariteta i određivanje karakteristika segmentiranih objekata	87
4.3.2. Neuronska mreža za klasifikaciju zahtevnog seta podataka	90
4.4. Inteligentna adaptivna segmentacija slike	93
5. INTELIGENTNI SISTEMI ZA DETEKCIJU I PREPOZNAVANJE LJUDI U ROBOTSKOJ VIZIJI	96
5.1. Pregled istraživanja	96
5.2. Mobilna robotska platforma sa stereo kamerom za prepoznavanje i praćenje ljudi	97
5.2.1. Segmentacija slike i izdvajanje karakteristika u cilju prepoznavanja ljudi	99
5.2.2. Klasifikacija u cilju prepoznavanja ljudi	103
5.2.3. Verifikacija rezultata i integracija senzora	107
5.3. Optimalan izbor ulaznih veličina klasifikatora	109

5.3.1. Kriterijum za odabir ulaznih veličina klasifikatora	109
5.3.2. Optimalan izbor ulaznog vektora karakteristika neuronske mreže primenom genetskih algoritama	111
6. INTELIGENTNI ROBOTSKI SISTEM ZA DETEKCIJU I PRAĆENJE LJUDI NA OSNOVU SLIKE DOBIJENE TERMOVIZIJSKOM KAMEROM	113
6.1. Pregled istraživanja	114
6.2. Mobilna robotska platforma DaNI sa Termovizijskom kamerom	115
6.3. Segmentacija termovizijske slike i određivanje karakteristika segmentiranog objekta	118
6.3.1. Fazi segmentacija	119
6.3.2. Optimizacija parametara segmentacije slike primenom genetskih algoritama	123
6.3.3. Izdvajanje karakteristika segmentiranih binarnih objekata	125
6.4. Inteligentni sistemi za klasifikaciju objekata sa termovizijske slike	126
6.4.1. Metod potpornih vektora za detekciju ljudi	126
6.5. Evaluacija rezultata	127
7. INTELIGENTNI SISTEMI ZA PRAĆENJE LJUDI, PREPOZNAVANJE POZA I GESTOVA	130
7.1. Inteligentni sistem za ocenu i predviđanje položaja čoveka	131
7.1.1. Predviđanje vremenskih serija	132
7.1.2. Nelinearna autoregresivna neuronska mreža za ocenu i predviđanje položaja čoveka	133
7.1.3. Obuka mreže i verifikacija rezultata	136
7.2 Razvoj mobilne robotske platforme sa 3D senzorom za praćenje ljudi i prepoznavanje poza	139
7.2.1. Mobilna robotska platforma DaNI sa 3D senzorom	140
7.3. Upravljanje mobilnom robotskom platformom sa 3D senzorom	142
7.3.1. Prepoznavanje i praćenje ljudi 3D senzorom	143

7.3.3. Verifikacija rezultata praćenja čoveka mobilnom robotskom platformom sa 3D senzorom	147
7.4. Prepoznavanje gestova 3D senzorom	149
8. ZAKLJUČCI I PRAVCI DALJIH ISTRAŽIVANJA	154
9. LITERATURA	157

Spisak tabela

Tabela 1 Pregled korišćenih rešenja u objavljenim radovima	8
Tabela 2. Rezultati segmentacije objekata i ekstrakcije karakteristika	74
Tabela 3 Rezultati klasifikacije	85
Tabela 4 Karakteristike segmentiranih objekata	90
Tabela 5 Ekstrahovane karakteristike ljudi u sceni	102
Tabela 6 Ekstrahovane karakteristike objekata koji nisu ljudi u sceni	103
Tabela 7 Test slučajevi različitih vektora karakteristika korišćenih za obučavanje klasifikatora	110
Tabela 8. Specifikacije Asus Xtion live 3D senzora	141

Spisak slika

Slika 2.1 Blok dijagram tipične aplikacije kompjuterske vizije [22]	10
Slika 2.2 Primer digitalne slike i diskretnih vrednosti piksela i koordinatnog sistema slike [25]	11
Slika 2.3 Razlaganje svetlosti na optičkoj prizmi [23]	14
Slika 2.4. Vidljivi spektar svetlosti [23]	14
Slika 2.5 Fotosenzitivne ćelije u mrežnjači ljudskog oka [27]	15
Slika 2.6 Spektralne osetljivosti ćelija ljudskog oka [23]	15
Slika 2.7 Primer Bajerovog obrasca na senzoru kamere [26]	16
Slika 2.8 RGB prostor boja [23]	17
Slika 2.9 HSV prostor boja a) i YCbCr prostor boja b) [28]	17
Slika 2.10 Histogram crno-bele slike i ujednačavanje histograma [24]	19
Slika 2.11 Princip rada kamera sa struktuiranom svetlošću [29]	21
Slika 2.12 Microsoft Kinect [29]	21
a) b)	22
Slika 2.13 a) Infracrvena slika snimljena Kinect senzorom, b) uzorak emitovane svetlosti na ravnoj površini [29]	22
Slika 2.14 Obojeni oblak tačaka snimljen Kinect-om [29]	23
Slika 2.15 Originalne i segmentirane slike [25]	25
Slika 3.1 Prikaz Stereovizijske kamere sa konvergentnim optičkim osama [30]	30
Slika 3.2 Stereovizijska kamera sa paralelnim optičkim osama [30]	30
Slika 3.3 Primer stereo ispravljanja slike a) Originalne leva i desna slika i b) Stereo ispravljene leva i desna slika [25]	32
Slika 3.4 Primer formiranja mape dispariteta, a) ispravljena leva slika, b) ispravljena desna slika, c) crno bela slika koja predstavlja mapu dispariteta, d) slika u boji koja predstavlja mapu dispariteta [25]	34
Slika 3.5 Uticaj izbora veličine bloka na mapu dispariteta, a) i b) originalne ispravljene leva i desna slika, c) blok veličine 9x9 piksela, d) blok veličine 15x15 piksela i e) blok veličine 21x21 piksela [25]	36
Slika 3.6 a) Sekvencijalna upravljačka struktura obrade slike i b) kaskadna upravljačka struktura obrade slike [22]	41

Slika 3.7 a) Idealna segmentacija, b) segmentacija sa šumom i c) segmentacija sa prekidima [22]	42
Slika 3.8 Kriva zavisnosti mere nesigurnosti I od inkrimenta segmentacije i kod crno-bele slike	44
Slika 3.9 Šema algoritma segmentacije s povratnom spregom [46, 48]	44
Slika 3.10 Detekcija regiona a) binarna slika i b) razdvojeni regioni [25]	45
Slika 3.11 Primer algoritma detekcije regiona na binarnoj slici a) prvi korak b) drugi korak [25]	46
Slika 3.12 Segmentacija mape dispariteta [25]	47
Slika 3.13 Segmentacija test scene [25]	49
Slika 3.14 Segmentacija mape dispariteta, konveksno popunjavanje okvira i ručno segmentirani objekti[47]	50
Slika 3.15 Blok dijagram klasifikacije objekta na osnovu karakteristika	53
Slika 4.1 DaNI robot.	58
Slika 4.2 Mobilna robotska platforma sa senzorom robotske vizije.	58
Slika 4.3. Program za segmentaciju slike dobijene senzorom robotske vizije	59
Slika 4.4 NXT roboti robogator i šuterbot.	60
Slika 4.5 Arhitektura ANFIS klasifikatora [71]	61
Slika 4.6 Algoritam za upravljanje mobilnom robotskom platformom a) i potprogram za praćenje koordinata detektovanog objekta b)	65
Slika 4.7 Mobilna robotska platforma u akciji	65
Slika 4.8 Promena x koordinate objekta koje mobilna robotska platforma prati	66
Slika 4.9 FRIEND robotska platforma [46]	68
Slika 4.9 Bumblebee® 2 stereo kamera i upravljivo postolje sa 2 stepena slobode kretanja	69
Slika 4.10 Robotska platforma i karakterističan kućni scenario	70
Slika 4.11 MASSiVE upravljačka arhitektura sistema FRIEND [44, 47]	71
Slika 4.12 Različito osvetljene scene u kojima je potrebno prepoznati objekte zarad dalje manipulacije	72
Slika 4.13 Merenje osvetljenja scene	73
Slika 4.14 ANFIS struktura	75
Slika 4.15 Fazi površ dobijena nakon obučavanja ANFIS strukture	76
Slika 4.16 Rezultati klasifikacije primenom ANFIS klasifikatora	77
Slika 4.17 Struktura klasifikatora zasnovanog na neuronskoj mreži	79

Slika 4.18. Rezultati obučavanja klasifikatora zasnovanog na neuronskoj mreži	81
Slika 4.19 Rezultati klasifikacije neuronskom mrežom	82
Slika 4.20 Trening set za klasifikaciju	83
Slika 4.21 Arhitektura naivnog Bajesovog klasifikatora	83
Slika 4.22 Graf Bajesovog klasifikatora - svakoj labeli odgovara određeni set karakteristika - atributa	84
Slika 4.23 FRIEND robot i radno okruženje	86
Slika 4.24 Scena sa većim brojem različito obojenih objekata ispred robotske platforme FRIEND	87
Slika 4.25 Određivanje mape dispariteta scene na osnovu slika dobijenih stereo kamerom	88
Slika 4.26 Originalna leva slika a), rezultat segmentacije mape dispariteta b), rezultat konveksnog popunjavanja c) i originalni objekat u poligonalnom okviru	88
Slika 4.27 Arhitektura višeslojnog perceptrona	90
Slika 4.28 Konfuzione matrice klasifikatora	91
Slika 4.29 Operativne karakteristike prijemnika	92
Slika 4.27 Histogrami HSV slike	94
Slika.5.1 Šema mobilne robotske platforme za praćenje ljudi [25]	98
Slika 5.2 Mobilna robotska platforma sa stereo kamerom za praćenje ljudi	98
Slika 5.3 Testiranje upravljačkih algoritama mobilne robotske platforme sa stereo kamerom za praćenje ljudi	99
Slika 5.4 Formiranje mape dispariteta na osnovu stereo slike korišćene za prepoznavanje ljudi	101
Slika 5.5 Segmentacija mape dispariteta za prepoznavanje ljudi	101
Slika 5.6 Arhitektura klasifikatora zasnovanog na neuronskoj mreži za detekciju ljudi	104
Slika 5.7 Konfuzione matrice klasifikatora	105
Slika 5.8 Minimizacija greške kroz iteracije obučavanja neuronske mreže	106
Slika 5.9 Operativne karakteristike prijemnika	107
Slika 5.10 Detekcija čoveka u sceni primenom neuronske mreže	108
Slika 5.11 Mobilna robotska platforma prati čoveka a) i prikaz modula robotske vizije koji se u tom trenutku izvršava na udaljenom računaru b)	108
Slika 5.12 Šema genetskog algoritma za optimalno određivanje vektora ulaza klasifikatora	111
Slika 6.1 Šema robotske platforme sa termovizijskom kamerom za praćenje ljudi	116
Slika 6.2 FLIR E50 termovizijska kamera (a) i termovizijska slika (b)	117

Slika 6.3 Termovizijska slika i vrednosti piksela za izabrano područje	118
Slika 6.4 Algoritam fazi segmentacije termovizijske slike	122
Slika 6.5 Rezultati segmentacije termovizijske slike a) nakon primene ručno podešenog praga b), nakon primene fazi segmentacije c) i promene parametara fazi segmentacije d)	122
Slika 6.6 Struktura genetskog algoritma	124
Slika 6.7 Termovizijska slika a), ručno segmentirana slika b) i segmentirana slika sa optimalno određenim graničnim pragom c)	125
Slika 6.8 SVM klasifikator	127
Slika 6.9 Praćenje osobe mobilnim robotom	128
Slika 7.1 Predviđanje vremenskih serija black-box modelom	133
Slika 7.2 Neuronska mreža sa prostiranjem signala unapred a) i rekurentna neuronska mreža b) [116]	134
Slika 7.3 Rešavanje problema black-box modeliranja sistema primenom rekurentne neuronske mreže [116]	135
Slika 7.4 Topologija NAR neuronske mreže za predviđanje i ocenu jedne koordinate položaja čoveka u prostoru	136
Slika 7.5 Praćenje čoveka koji hoda po šablonu nacrtanom na podu [25]	137
Slika 7.6 Promena X koordinate čoveka tokom vremena a), promena Y koordinate čoveka b) i promena Z koordinate čoveka u prostoru c)	139
Slika 7.7 Asus Xtion PRO live 3D senzor	140
Slika 7.8 DaNI robot opremljen laptopom i Asus Xtion PRO LIVE kamerom i njihovi koordinatni sistemi	141
Slika 7.9 Konekcija OpenNI-a i LabVIEW-a	142
Slika 7.10 Simulink Model za prikupljanje podataka i detekciju	143
Slika 7.11 Tačke od interesa detektovane Skeleton blokom	144
Slika 7.12 Šema praćenja	145
Slika 7.13 Šema genetske optimizacije parametara PID kontrolera	146
Slika 7.14. LabView program za upravljanje DANI robotom u cilju praćenja čoveka	147
Slika 7.15 Brzina levog točka DaNI mobilnog robota	148
Slika 7.16 Brzina desnog točka DaNI mobilnog robota	148
Slika 7.17 Distanca dobijena sistemom robotske vizije i distanca dobijena iz modula za praćenje	149
Slika 7.18 Slika skelete sa zglobovima [120]	150

Slika 7.19 Proces formiranja zglobova

151

Slika 7.20 Skelet iscrtan na osnovu podataka dobijenih od Kinect uređaja

152

1. UVOD

U današnje vreme roboti se pomeraju iz fabričkih proizvodnih linija i postepeno se uvode u sve oblasti svakodnevnog ljudskog života. Ovo vodi do novih izazova u istraživanju robota sa adekvatnom veštačkom inteligencijom u cilju postizanja prirodne i intuitivne interakcije između ljudskog bića i robotskog sistema. Roboti koji rade zajedno sa ljudima ne treba samo da identifikuju ljude već i da prate njihovo kretanje kako bi se izbegao sudar sa njima ili kako bi bilo omogućeno praćenje njihove trajektorije od strane robota. Kada je interakcija sa ljudima neminovna, sigurnost i pouzdanost su kritični uslovi za široku primenu robota. Ljudska detekcija, identifikacija i praćenje je veoma aktivna oblast istraživanja ne samo u robotskoj viziji već generalno u računarskoj viziji zbog neophodnosti aplikacija kao što su na primer video nadzor i napredni sistemi pomoći vozačima.

Međutim, iako je mnogo toga urađeno u nekoliko prethodnih godina, ima prostora za unapređenje ovih sistema. Naime, detekcija objekata i ljudi na slikama mora da bude robusna u odnosu na veliki broj spoljašnjih uticaja kao što su različiti uslovi osvetljenja kako u otvorenom tako i u zatvorenom prostoru i razlika u izgledu objekata i ljudskih bića. Takođe, robot koji deluje u ljudskom okruženju treba da prepozna ljude između drugih objekata u okruženju. Često je robotsko okruženje dinamično i robot mora da napravi razliku između čoveka koji je upravo ušao u njegovo vidno polje i ljudi koji se već nalaze u vidnom polju robota. Problem lokalizacija ljudi u trodimenzionalnom (3D) prostoru je od ključnog značaja, ali i pored savremenih tehnoloških dostignuća koja omogućavaju korišćenje različitih naprednih senzora za ekstrakciju 3D informacija, još uvek su brojne karakteristike senzora vizije u dobijanju detaljne informacije nepouzdan.

1.1 Motivacija

Sistem robotske vizije se u poslednje vreme sve češće javlja kao važan sastavni deo robotskih sistema i predstavlja osnovni izvor informacija na osnovu koga robotski sistem vrši percepciju okoline u cilju njegove odgovarajuće interakcije sa objektima i ljudima. U velikom broju slučajeva, kamera i sistem digitalne obrade slike su osnovni senzor za dobijanje informacije o objektima nad kojima je potrebno izvršiti određeni zadatak u robotskim sistemima. Međutim, i pored velikog broja prednosti vizuelni sistem ima i određenih nedostataka. Ti nedostaci se ogledaju u maloj robusnosti sistema vizije u odnosu na različite spoljašnje uticaje, kao što je promena osvetljenja pri akviziciji slike objekta koji se analizira. Pored promenljivog osvetljenja, sistem vizije u robotici, čiji su osnovni funkcionalni elementi prepoznavanje objekta i razumevanje slike, treba da bude robusan u odnosu na prisustvo mnogo različitih objekata u sceni. Takođe, sistem treba da da pouzdan rezultat i prilikom varijacije forme analiziranog objekta kao i u slučaju posmatranja istog objekta pod različitim okolnostima (npr. isti objekat može drugačije izgledati na slici prilikom promene osvetljenja, ugla snimanja ili orijentacije posmatranog objekta).

S obzirom da aktuelno stanje istraživanja u svetu ukazuju na veliki potencijal razvoja i implementacije algoritama zasnovanih na veštačkoj inteligenciji i mašinskom učenju kod sistema robotske vizije, a u cilju prepoznavanja objekata i praćenje ljudi, ovaj naučnoistraživački pravac predstavlja jednu od polaznih oblasti istraživanja u okviru rada na predmetnoj doktorskoj disertaciji.

1.2. Problem

I pored velikog broja publikovanih metoda za analizu slike, prepoznavanje objekata, prepoznavanje i praćenje ljudi, do sada nisu razvijene dovoljno pouzdane metode za primenu u scenarijima gde je neophodna pouzdana interakcija robota sa okolinom i ljudima. Potrebno je razviti inovativan sistem vizije kako bi robot dobio neophodne informacije za pouzdano i autonomno funkcionisanje u dinamičnom okruženju. Sistem vizije mora da omogući ekstrakciju i klasifikaciju osobina različitih objekata koji su prisutni u sceni bez potrebe za velikim bazama podataka, koje zahteva većina postojećih robotskih sistema.

Predmet istraživanja odnosi se na razvoj i primenu novih algoritama inteligentnog mobilnog robota koji primenom neuro-fazi genetskih alata veštačke inteligencije u robotskoj viziji detektuje objekte i ljude, prati ih i na osnovu toga izvršava unapred definisane zadatke.

U užem smislu, istraživanja predstavljena u ovoj tezi su usmerena na razvoj mobilne robotske platforme i njenog upravljačkog sistema na bazi neuro-fazi-genetskih sistema u raznim sistemima robotske vizije (vizija monokularnom kamerom, stereo kamerom, 3D senzorom sa struktuiranom svetlošću) u cilju rešavanja problema prepoznavanja tj. klasifikacije objekata i ljudi u laboratorijskom okruženju. Kako su početna istraživanja vršena sa robotskim sistemom kod koga sistem vizije daje osnovne podatke za upravljanje manipulatorom, a razvijeni algoritmi su bili polazna osnova za razvoj algoritama za detekciju i praćenje ljudi mobilnim robotom, predstavljani su i naučni rezultati dobijeni tokom istraživanja obavljenog na ovom robotskom sistemu.

Implementacijom više tehnika veštačke inteligencije, kao što su neuronske mreže, fazi logika i genetski algoritmi, u inteligentni upravljački sistem potrebno je razviti odgovarajuće module vizije i lokalnog upravljanja mobilnom robotskom platformom koja ima sposobnost detekcije, prepoznavanja i prostornog lociranja objekata i ljudi kao i odgovarajuću inteligentnu interakciju ovog robotskog sistema sa njima

1.3. Predložena rešenja

U ovoj doktorskoj disertaciji biće predstavljen veći broj novih algoritama čiji je glavni zadatak uspešno detektovanje i praćenje objekata i ljudi u robotskom okruženju. Osnovu ovih algoritama čine alati iz domena veštačke inteligencije, odnosno neuronske mreže, fazi logika, genetski algoritmi, metod potpunih vektora i naivni Bajesov klasifikator. Tokom istraživanja korišćena su 4 različite vrste senzora vizije koje su služile kao osnovni izvor informacija o okolini robota, pa će za svaki od sistema ponaosob u tezi biti predstavljeni odgovarajući algoritmi.

Inicijalno istraživanje podrazumevalo je prepoznavanje objekata. U tu svrhu korišćena je monokularna kamera i predložen je algoritam segmentacije i inteligentni sistem klasifikacije kako bi mobilna robotska platforma mogla da interaguje s željenim objektima.

Nakon toga su razmatrani savremeni algoritmi segmentacije slike stereo kamere, pa su predloženi inteligentni algoritmi koji bi ovu segmentaciju dodatno unapredili, kao i odgovarajući inteligentni algoritmi klasifikacije koji su integrisani u sistem vizije. U ovu svrhu

korišćen je robotski sistem FRENDO za pomoć osobama sa invaliditetom, razvijen na Univerzitetu u Bremenu. I pored toga što robotski sistem FRIEND nije mobilni robot, algoritmi razvijani za njegov sistem vizije, čija je uloga percepcija objekata u okruženju zarad upravljanja manipulatorom, korišćeni su kao polazna osnova za sva dalja istraživanja, te će u ovoj doktorskoj disertaciji biti predstavljeni i rezultati vezani za razvoj i unapređenje njegovog sistema vizije. Razvijeni algoritmi su potom uz odgovarajuće modifikacije implementirani u sisteme upravljanja mobilne robotske platforme za prepoznavanje ljudi.

Kasnija istraživanja podrazumevala su prepoznavanje i praćenje ljudi. Razvijene su eksperimentalne robotske platforme sa stereo kamerom, sa termovizijskom kamerom i sa 3D senzorom sa struktuiranom svetlošću. Kod sistema sa stereo kamerom će biti predstavljeno unapređenje modula robotske vizije klasifikatorom zasnovanim na neuronskoj mreži i biće predstavljen koncept primene genetskih algoritama za optimalan izbor ulaznih karakteristika u klasifikator.

Primena termovizijske kamere u robotskim sistemima za prepoznavanje i praćenje ljudi nije često prisutna u naučnim istraživanjima, pa će u ovoj tezi biti predstavljeni novi algoritmi za fazi segmentaciju slike dobijene termovizijskom kamerom, kao i razvoj genetskog algoritma za optimalno podešavanje parametara segmentacije slike dobijene termovizijskom kamerom. Razvijen je i softver zasnovan na metodi potpornih vektora kojim se vrši prepoznavanje ljudi u sceni.

Biće predstavljen razvoj nelinearne auto-regresivne neuronske mreže koja služi za ocenu i predviđanje položaja čoveka snimljenog stereo kamerom. Konačno predložena su i rešenja inteligentnog upravljanja mobilnom robotskom platformom za praćenje ljudi koji informacije dobija od 3D senzora sa struktuiranom svetlošću, a ovakva mobilna robotska platforma može da se koristi i za prepoznavanje gestova.

1.4. Publikacije

Neki od naučnih rezultati predstavljenih u ovoj tezi publikovani su u okviru 17 naučnih radova čiji je autor ujedno i autor ove teze, a štampani su u međunarodnim i domaćim časopisima i predstavljeni na međunarodnim i domaćim skupovima, te štampani u zbornicima radova:

- [1] Ćirić, Ivan T; Ćojbašić, Žarko M; Nikolić, Vlastimir D; Igić, Tomislav S; Turšnek, Branko AJ; (2014) “Intelligent optimal control of thermal vision-based Person-Following Robot Platform”, *Thermal Science*,18,3,957-966,2014, (M23);
- [2] Ćirić, Ivan; Ćojbašić, Žarko; Nikolić, Vlastimir; Živković, Predrag; Petković, Dušan; Tomić, Mladen; Tomić, Miša; (2014) “Thermal Vision Integration In Mobile Robot Vision System., *Annals of the Faculty of Engineering Hunedoara-International Journal of Engineering*”,12,2,2014, (M51);
- [3] Ćirić Ivan, Ćojbašić Žarko, Miša Tomić, Milan Pavlović, Vukašin Pavlović, Ivan Pavlović, Vlastimir Nikolić (2012) “Intelligent Control of DaNi Robot Based on Robot Vision and Object Recognition”, *Facta Universitatis, Series: Automatic Control And Robotics*, Vol.11, No 2, 2012 pp. 129 – 140 (M52);
- [4] Petrović, E., Ćojbašić, Ž., Ristić-Durrant, D., Nikolić, V., Ćirić, I., & Matić, S. (2013). “Kalman Filter and Narx Neural Network for Robot Vision Based Human Tracking” *Facta Universitatis, Series: Automatic Control and Robotics*, Vol. 12, No 1, pp. 43-51. (M52);
- [5] Ćirić Ivan , Žarko Ćojbašić, Vlastimir Nikolic, Milica Ćirić, Predrag Živković, Ivan Pavlović, Emina Petrovic, (2015) “Intelligent Control of Mobile Robot for Object Recognition And Grasping”, 12th International Conference on Accomplishments in Electrical and Mechanical Engineering and Information Technology DEMI 2015, Banjaluka, BiH (M33);
- [6] Ćirić Ivan, Žarko Ćojbašić, Vlastimir Nikolić, Milica Ćirić, Mladen Tomić, Emina Petrović, Miloš Simonović, (2015) “Neural Network Prediction of Person Position for Human Following Mobile Robot Platform”, 12th International Conference on Accomplishments in Electrical and Mechanical Engineering and Information Technology DEMI 2015, Banjaluka, BiH (M33);
- [7] Ćirić Ivan, Žarko Ćojbašić, Miloš Simonović, Milica Ćirić, Ivan Pavlović, Emina Petrović, Vlastimir Nikolić, (2014) “Intelligent System for Object Recognition in Stereo Vision Based Robotic Applications”, ISSN 1313-4264, Sozopol, Bulgaria, pp.143-149. (M33);
- [8] Petrović Emina, Ivan Ćirić, Vlastimir Nikolić,(2014) “Human Tracking by a DaNI Mobile Robot based on Extended Kalman Filter”, ISSN 1313-4264, Sozopol, Bulgaria, pp.136-142. (M33);

- [9] Ćirić Ivan, Žarko Ćojbašić , Vlastimir Nikolić , Dragan Antić, (2013) “Computationally Intelligent System for Thermal Vision People Detection and Tracking in Robotic Applications”, Proceedings of the 11th International Conference on Telecommunications in Modern Satellite, Cable and Broadcasting Services - TELSIS 2013, Niš, Serbia, 2013. (M33);
- [10] Ćirić, I., Ćojbašić, Ž., Nikolić, V., Igić, T., (2013) “Intelligent Control System for Thermal Vision-Based Person-Following Robot Platform”, Proceedings, 16th Symposium on Thermal Science and Engineering SIMTERM, Sokobanja, Serbia, 2013, pp. 640-648 (M33);
- [11] Petrović E., Tomić M., Nikolić V., Ćojbašić Ž., Pavlović V., Ćirić I., (2013) “Human Tracking With a Person Following Robot Based on Extended Kalman Filter”, 2nd International Conference (Mechanical Engineering in the 21st century), Niš, Serbia. (M33);
- [12] Ćirić, Ž. Ćojbašić, V. Nikolić, P. Živković, D. Petković, M. Tomić, M. Tomić, (2013) “Thermal Vision Integration in Mobile Robot Vision System”, 11 International Conference on Accomplishments in Electrical and Mechanical Engineering and Information Technology DEMI 2013, Banjaluka, BiH (M33);
- [13] Ćirić, Ž. Ćojbašić, M. Tomić, M. Pavlović, V. Pavlović, (2012) “Computationally Intelligent Object Recognition for DaNI Robot Vision”, Proceedings of XI International SAUM Conference, Niš, Serbia, 2012, pp.132 – 135, (M33);
- [14] Ćojbašić Ž., V. Nikolić, D. Ristić-Durant, I. Ćirić, E. Petrović, S. Matić, (2012) “Neural Networks Based Human Tracking for Robot Vision”, Proceedings of XI International SAUM Conference, Niš, Serbia, 2012, pp.164 – 167, (M33);
- [15] Nikolić V., Ž. Ćojbašić, D. Ristić-Durant, E. Petrović, S. Matić, I. Ćirić, (2012) “Kalman Filter For Robot Vision-Based Human Tracking”, Proceedings of 29th DANUBIA-ADRIA Symposium on Advances in Experimental Mechanics, Belgrade, Serbia, 2012., pp. 178-181 (M33);
- [16] Ćojbašić Žarko, Danijela Ristić-Durrant, Ivan Ćirić, Sorin M. Grigorescu, Axel Gräser, Vlastimir Nikolić (2013) “Reliable Computationally Intelligent Object Recognition and Human Tracking in Robotic Vision”, Zbornik 57. konferencije ETRAN, Zlatibor, 3-6. juna 2013, str. RO3.6.1-5 (M63);
- [17] Ćirić Milica, Ćirić Ivan, Milan Gocić, (2014) “Application of Naive Bayes Classifier in Stereo-Vision Based Object Recognition”, The Second National Conference on

Information Theory and Complex Systems, TINKOS 2014, Niš, Serbia, pp 39-40 (M63);

Osim ovih radova, neki od rezultata predstavljenih u ovoj tezi našli su se i u 4 tehnička rešenja, verifikovana od strane recenzenata čije su recenzije usvojene na Nastavno-naučnom veću Mašinskog fakulteta i potom verifikovane od strane Matičnog odbora za mašinstvo Ministarstva prosvete, nauke i tehnološkog razvoja Republike Srbije:

- [18] Žarko Čojbašić, Vlastimir Nikolić, Ivan Ćirić, Ivan Pavlović, “Neuro-fazi klasifikator objekata kod robotske vizije složenog rehabilitacionog robotskog sistema FRIEND” (M85);
- [19] Ivan Ćirić, Žarko Čojbašić, Vlastimir Nikolić, Ivan Pavlović, Emina Petrović, Milica Ćirić, Miloš Simonović, “Mobilna robotska platforma sa termovizijskom kamerom za detekciju i praćenje ljudi” (M81);
- [20] Vlastimir Nikolić, Žarko Čojbašić, Ivan Ćirić, Miša Tomić, Emina Petrović, Andrija Milojević, “Inteligentno upravljanje mobilnom robotskom platformom zasnovano na robotskoj viziji i prepoznavanju objekata” (M85);
- [21] Vlastimir Nikolić, Žarko Čojbašić, Emina Petrović, Ivan Ćirić, Miša Tomić, Miloš Simonović, Vukašin Pavlović, “Sistem za detekciju i praćenje ljudi mobilnim robotom zasnovan na robotskoj viziji i proširenom kalmanovom filteru” (M85).

Predstavljeni radovi i tehnička rešenja pristupaju rešavanju problema projektovanja upravljačkog sistema mobilnog robota zasnovanog na robotskoj viziji na savremen način, a akcenat je stavljen na tehnike iz domena računarske inteligencije. U istraživanjima je korišćeno više robotskih platformi sa sensorima vizije koji su podrazumevali monokularnu kameru, stereo kameru, termovizijsku kameru i 3D senzor sa struktiranom svetlošću, pa su za svaki od njih razvijeni odgovarajući inteligentni algoritmi vizije i upravljanja.

Radi preglednosti sprovedenih istraživanja u tabeli 1 predstavljeni su korišćeni algoritmi i metodologija u navedenim radovima.

Tabela 1 Pregled korišćenih rešenja u objavljenim radovima

rad	Senzor vizije				Algoritam segmentacije				Klasifikator				Filter	
	web kam.	Stereo kam.	Termo kam.	3D senzor	s povr. spreg.	Mapa disp.	Fazi	GA	ANFIS	NN	SVM	NBC	NAR NN	Kalman
[1]			x					x			x			
[2]			x				x			x				
[3]	x								x	x				
[4]		x				x							x	x
[5]		x			x	x				x				
[6]		x				x				x			x	
[7]		x			x	x			x	x				
[8]				x										x
[9]			x				x		x					
[10]			x					x			x			
[11]				x										x
[12]			x				x			x				
[13]	x								x	x				
[14]		x				x							x	x
[15]		x				x								x
[16]		x			x				x	x				
[17]		x			x	x						x		
[18]		x			x				x					
[19]			x					x			x			
[20]	x								x	x				
[21]				x										x

Kako se vidi iz tabele, najviše je radova je vezano za korišćenje slike dobijene stereo kamerom kao osnovne informacije za raspoznavanje okoline od strane robota. Razlog je taj što su ovakve kamere korišćene na dve robotske platforme, od kojih jedna služi za pomoć osobama sa invaliditetom i cilj sistema vizije je raspoznavanje objekata kojima manipulator treba da manipuliše, dok je druga mobilna robotska platforma namenjena pomoći ljudi u hazardnim okruženjima, te je osnovni zadatak modula robotske vizije u upravljačkom sistemu detekcija i praćenje ljudi.

1.5. Pregled

Ova teza podeljena je u osam poglavlja. Prvo poglavlje čini uvod i u njemu su predstavljene osnovni ciljevi istraživanja. U drugom poglavlju su predstavljene osnove kompjuterske, odnosno robotske vizije koje podrazumevaju akviziciju i obradu slike. U trećem poglavlju predstavljene su napredni algoritmi obrade slike i stereo slike u cilju prepoznavanja objekata. Ovi algoritmi su podrazumevali formiranje mape dispariteta i segmentacione algoritme koji su bili osnova za dalje unapređenje alatima iz domena računarske inteligencije.

U četvrtom poglavlju predstavljene su dve robotske platforme kod kojih je osnovni zadatak bio razviti algoritme za prepoznavanje objekata. U tu svrhu razvijeni su ANFIS klasifikator i klasifikator zasnovan na neuronskim mrežama. Osim toga, u ovom poglavlju su predložena unapređenja postojećeg algoritma klasifikacije s povratnom spregom koja podrazumevaju inteligentnu adaptivnu segmentaciju.

Peto poglavlje bavi se razvojem inteligentnog sistema za detekciju ljudi zasnovanog na neuronskoj mreži za klasifikaciju. Ovde je predstavljen i novi koncept primene genetskog algoritma za odabir optimalnog vektora karakteristika kao ulaza u klasifikator.

U šestom poglavlju se kao alternativa najčešće korišćenim sensorima vizije predlaže primena termovizijske kamere. Razvijena su dva algoritma segmentacije od kojih se jedan bazira na fazi logici, dok se kod drugog parametri podešavaju genetskim algoritmima. Predstavljen je i metod potpunih vektora za klasifikaciju segmentiranih objekata na ljude i one koji to nisu.

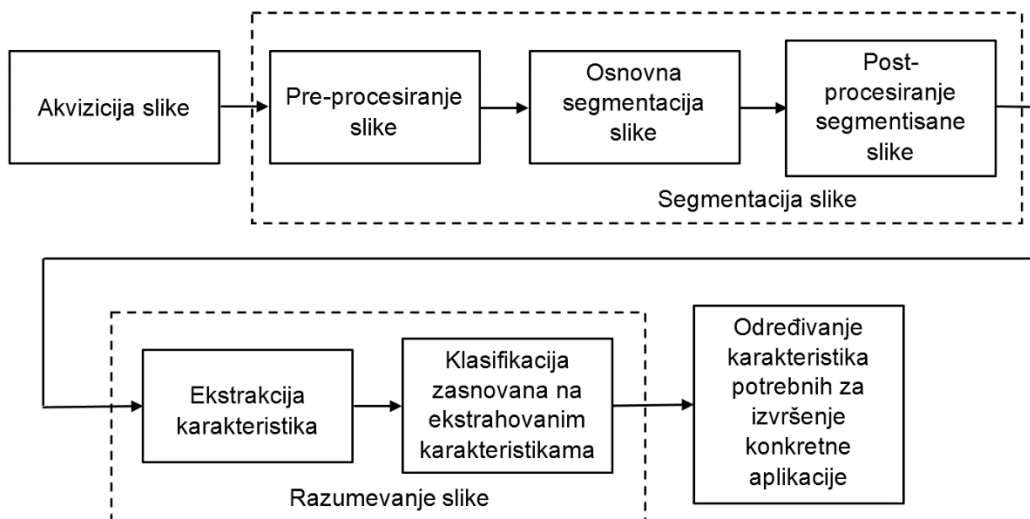
Sedmo poglavlje se bavi problemom praćenja ljudi i prepoznavanja njihovih gestova. Najpre je predstavljena nelinearna auto-regresivna neuronska mreža za predikciju i estimaciju položaja čoveka koja se može primeniti na eksperimentalnoj robotskoj platformi predstavljenoj u petoj glavi. Potom je predstavljena robotska platforma sa 3D sensorom, kod koje je prepoznavanje ljudi rešeno hardverski, pa je bilo samo potrebno razviti sistem praćenja, upravljanja i prepoznavanja gestova čoveka kako bi se omogućila uspešna interakcija robota i čoveka.

U osmom poglavlju su predstavljeni zaključci i pravci daljih istraživanja. Na kraju rada data je lista korišćene literature.

2. DIGITALNA AKVIZICIJA I OBRADA SLIKE U ROBOTICI

Prepoznavanje objekata zasnovano na viziji se primenjuje u mnogim robotskim sistemima i aplikacijama vezanim za kompjutersku viziju poput 3D rekonstrukcije objekata u cilju manipulacije u servisnoj robotici.

I pored toga što različite aplikacije imaju različite ciljeve, njihovi algoritmi se sastoje od svih ili većine koraka prikazanih na dijagramu 2.1.



Slika. 2.1 Blok dijagram tipične aplikacije kompjuterske vizije [22]

Akvizicija slike je prvi korak kod aplikacije za prepoznavanje objekata zasnovano na viziji gde se digitalna kamera koristi kao osnovni izvor informacija koje sadrže scenu u kome se nalazi objekat interesovanja. Nakon toga pristupa se segmentaciji slike, gde se slika "pojednostavljuje" i priprema za prepoznavanje objekata, odnosno uopšteno gledano razumevanje digitalne slike [23,24]. Segmentacija slike se obično vrši u tri koraka, pre-procesiranje slike, sama segmentacija i post-procesiranje slike. Pre-procesiranje se vrši u cilju

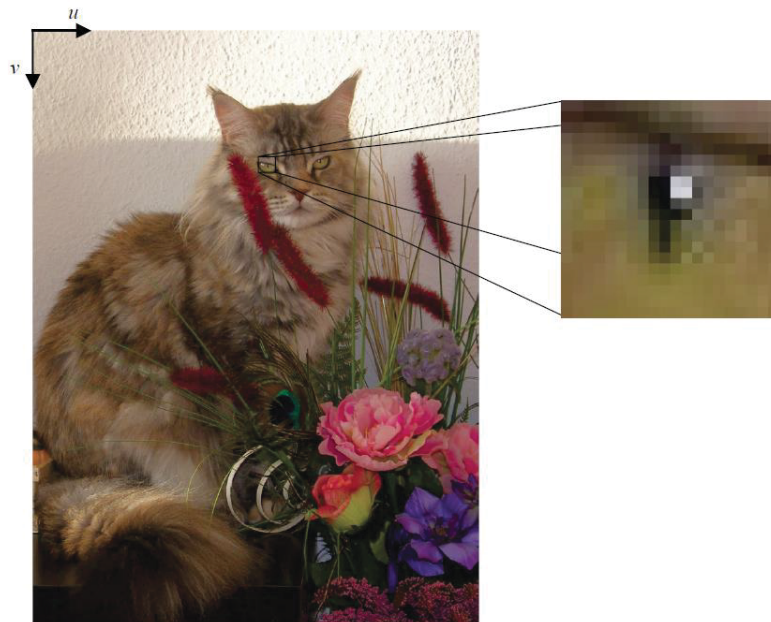
poboljšanja kvaliteta slike (najčešće primenom raznih filtera) ili kako bi se odredila oblast interesovanja (engl. Region of Interest - ROI) tj. smanjila oblast segmentacije. Nakon primene adekvatnog algoritma segmentacije, moguće je vršiti post-procesiranje čime bi se npr. uklonili elementi koji nisu od interesa za klasifikaciju ili se mogu primeniti filteri koji obezbeđuju bolju rekonstrukciju objekata.

Deo algoritma koji omogućava razumevanje slike sastoji se iz dva bloka, i to bloka za određivanje karakteristika i klasifikatora zasnovanog na karakteristikama. U zavisnosti od specifične primene, moguće je odrediti različite karakteristike koje formiraju ulazni vektor za klasifikator. Nakon klasifikacije prepoznat je objekat od interesa i određuje se karakteristika (poput 3D koordinata centra mase objekta, 3D koordinata rekonstruisanog objekta i sl.) koja predstavlja izlaz sistema robotske vizije.

U koliko se ovaj algoritam primeni na frejmove (uzastopne slike) neke video sekvence, moguće je pratiti objekat, odnosno njegovu izlaznu karakteristiku u vremenu i prostoru

2.1. Digitalne slike

Osnovni izvor informacija kod prepoznavanja objekata čine digitalne slike. Digitalna slika, poput one predstavljene na slici 2.2, je digitalna reprezentacija dvodimenzionalne slike koja se sastoji od konačnog broja skupova digitalnih vrednosti koji se nazivaju pikseli (engl. pixel) i koji su organizovani matrično, sa konačnim brojem vrsta i kolona.



Slika 2.2 Primer digitalne slike i diskretnih vrednosti piksela i koordinatnog sistema slike [25]

Pikseli su najmanji individualni elementi slike koji sadrže digitalne podatke o intenzitetu osnovnih boja u datoj tački. Svaki piksel je definisan dvema koordinatama koje jednoznačno određuju njegovu lokaciju u odnosu na koordinatni sistem slike. Koordinatni početak se najčešće nalazi u gornjem levom uglu digitalne slike, gde je u osa u pravcu vrsta i usmerena na desno, dok je v osa u pravcu kolona slike i usmerena na dole, kao što je prikazano na slici 2.2. Ovakav izbor koordinatnog sistema slike je korišćen i u ostalim poglavljima ove teze. Na uvećanom segmentu prikazanom na slici 2.2 jasno se mogu uočiti vrste i kolone čime se jasno vidi diskretna priroda digitalne slike.

2.2. Akvizicija digitalne slike

Algoritmi za prepoznavanje koji su zasnovani na viziji podrazumevaju procesiranje slika koje su snimljene digitalnom kamerom. Ime je izvedeno od latinskog *camera obscura* (mračna komora), što je bio rani mehanizam za projektovanje slika u kome je cela prostorija funkcionisala kao unutrašnjost modernih fotografskih kamera, osim što tada nije postojao način da se slika zabeleži već se ručno crtala [23, 24, 25].

Savremene kamere koriste isti princip tokom akvizicije slike tako što projektuju svetlo reflektovano o objekat bilo na fotografski film ili digitalni senzor. Međutim, sada svetlost ne prolazi kroz običan mali otvor, već kroz različita sočiva, što omogućava bolje osvetljenje i adekvatno fokusiranje slike. Kamera može da snima vidljivi spektar ili bilo koji drugi deo elektromagnetnog spektra.

U proteklih nekoliko decenija zavladao je ogromno interesovanje za digitalnu fotografiju kako zbog razumnih cena uređaja, tako i zbog visokog kvaliteta slika i gotovo besplatnog snimanja i skladištenja, te je stoga i ogromno interesovanje za digitalnom obradom slike. Ako se snimi serija slika, ovaj niz se naziva video i sastoji se od više uzastopnih frejmova odnosno slika.

Tokom akviziciju slike svetlost ulazi kroz objektiv kamere i fokusira se na senzor slike. Posle izvesnog vremena (vreme ekspozicije), ili se mehaničkim zatvaračem poklapa senzor tako da svetlost više ne da pada na njega ili se podaci slike očitaju vrlo brzo tako da je mehanički zatvarač nepotreban. Vreme ekspozicije utiče na osvetljenost slike, jer duže vreme ekspozicije

omogućava da više svetla padne na senzor, što čini sliku svetlijom. Ostali važni parametri akvizicije slike su veličina blende i osetljivost senzora na svetlost.

Pojam veličina blende se odnosi na prečnik otvora kroz koji svetlost putuje do senzora dok je zatvarač otvoren. Veći prečnik je poželjan u većini situacija jer omogućava kameri da dobro radi i u uslovima slabog osvetljenja. Međutim, manji prečnik omogućava manji prolaz svetlosti do senzora, što može biti korisno u vrlo svetlim uslovima i uz minimalnu brzinu zatvarača. Takođe, kroz manji otvaranja upadni ugao svetlosnih zraka je skoro isti i za objekte koji su blizu kamere i za one dalje, što čini veći deo scene oštrim.

Osetljivost senzora na svetlost predstavlja minimalnu količinu svetlosti koja je potrebna da bi senzor uočio razliku od potpunog mraka. Osetljiv senzor može dobro funkcionisati u uslovima slabog osvetljenja, ali to košta i zahteva dobar objektiv. Standardni senzori daju tamnije slike u uslovima slabog osvetljenja, koje se elektronski pojačava, što pojačava i šum. Ovaj pojačani šum je jedan od razloga zašto se dobijaju različiti rezultati kada se isti algoritam obrade slike primeni na različite frejmove istog video snimka, iako oni ljudskom oku deluju gotovo identično.

Digitalne slike u boji su uglavnom dobijene postavljanjem tzv. Bajerovog filtera (engl. Bayer) [26] preko senzora slike, koji inače detektuje samo intenzitet svetlosti. Bayer filter je mozaik filter koji za svaki pojedinačni piksel omogućava prolaz samo crvenog, zelenog ili plavog dela spektra svetlosti do površine senzora. Nakon toga se boja piksela može odrediti tzv. de-Bajerovim algoritmom. Međutim, prilikom primene u uslovima slabog osvetljenja ili kada je potreban veliki broj frejmova u sekundi (engl. frame rate), ponekad se prave crno-beli snimci (engl. grayscale) i Bayerov filter se izostavlja kako bi se povećala količinu svetlosti koja stiže do senzora.

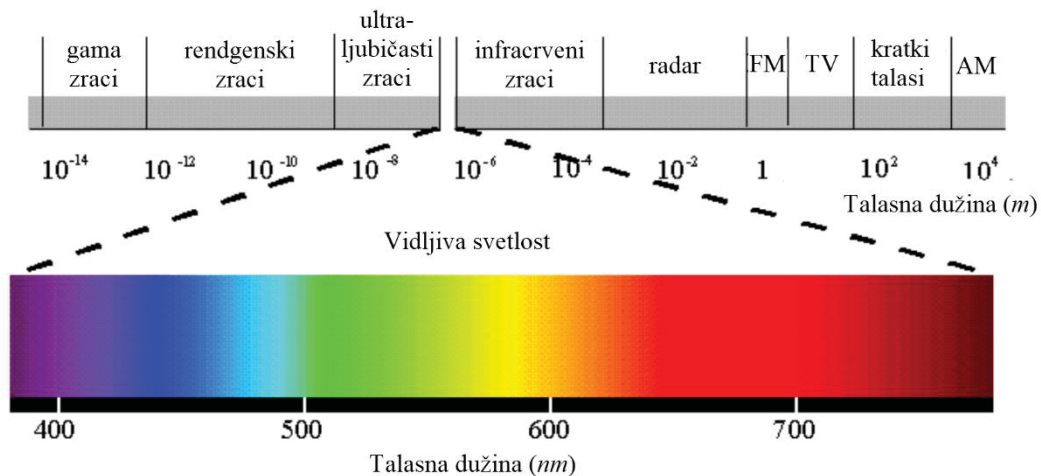
2.3. Boja u kompjuterskoj viziji

Proces percipiranja i interpretiranja boje u ljudskom mozgu je psiho-fiziološki fenomen koji nije još sasvim shvaćen, ali se priroda boje može opisati na formalnoj bazi koja je poduprta eksperimentalnim i teoretskim rezultatima. Prvo ozbiljno istraživanje boje zabeleženo je u 17-om veku kada je Isak Njutn otkrio da kada sunčeva svetlost obasja staklenu prizmu, na izlazu iz nje ne dobijamo isto obojenu svetlost već celi spektar boja od ljubičaste do crvene, kao što se vidi na slici 2.3.



Slika 2.3 Razlaganje svetlosti na optičkoj prizmi [23]

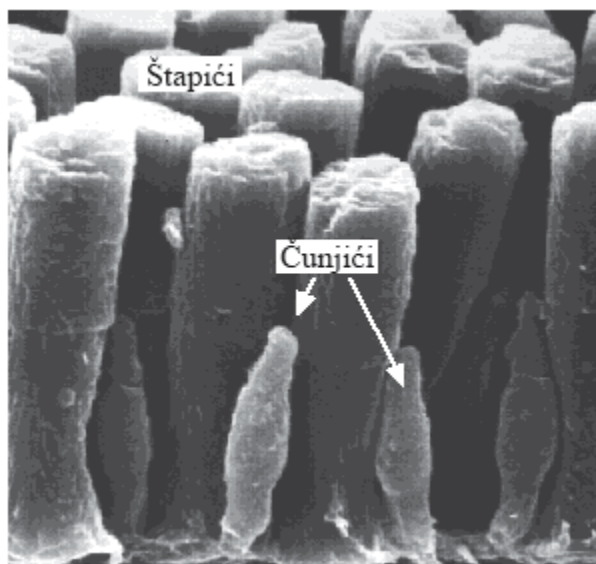
Otkriveno je da vidljivu svetlost koju ljudi mogu percipirati čini relativno uzak spektar frekvencija u elektromagnetnom spektru koji se može videti na slici 2.4.



Slika 2.4. Vidljivi spektar svetlosti [23]

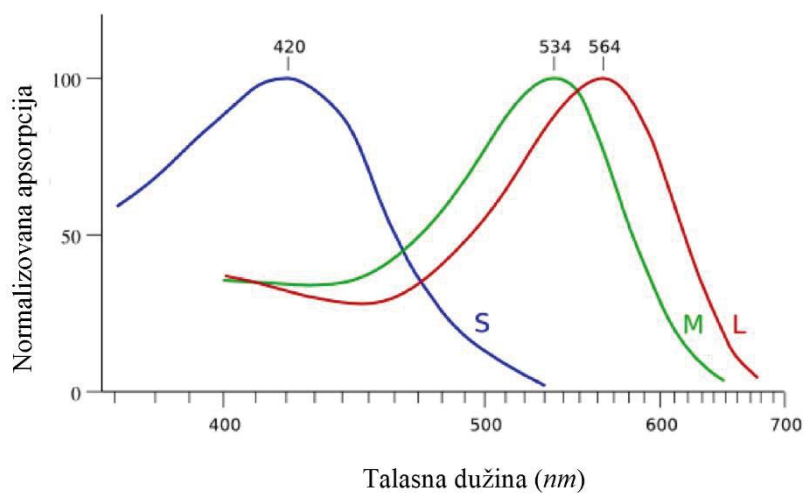
Ono što ljudi i neke životinje percipiraju kao boju je zapravo reflektovana svetlost s površine nekog objekta. Zbog toga je izučavanje svetlosti izuzetno bitno kako za nauku o bojama, tako i za korišćenje boje u kompjuterskoj viziji. Ako svetlost nije obojena, tada je ona ahromatska i promatramo je isključivo kao intenzitet.

Foto osetljivi deo ljudskog oka se naziva mrežnjača i sadrži fotosenzitivne ćelije nazvane prema njihovom obliku, čunjići i štapići, koje se mogu videti na slici 2.5.



Slika 2.5. Fotosenzitivne ćelije u mrežnjači ljudskog oka [27]

Štapići su odgovorni za osetljivost na intenzitet svetla kao i za percepciju granica i rubova u vidljivoj sceni, dok su čunjići odgovorni za percepciju boja. Postoje tri tipa čunjića S, M i L tip. Eksperimentalnom metodom je utvrđeno da je njihova osetljivost najveća na talasnim dužinama od 420, 534 i 564nm, koje približno odgovaraju plavoj, zelenoj i crvenoj boji kao što se vidi na slici 2.6.

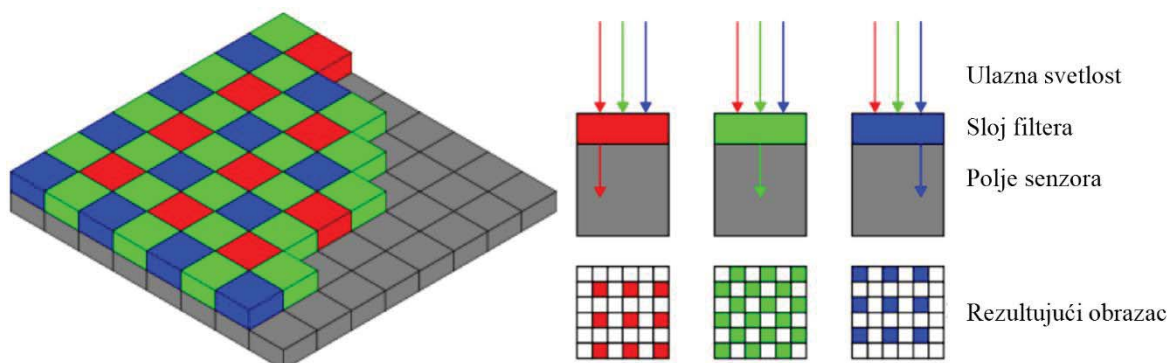


Slika 2.6. Spektralne osetljivosti ćelija ljudskog oka [23]

Otkriveno je stoga da se ljudska percepcija boja temelji na tripletu boja RGB prostora boja, te je stoga taj prostor boja postao osnovni sistem boja i za digitalnu obradu slike.

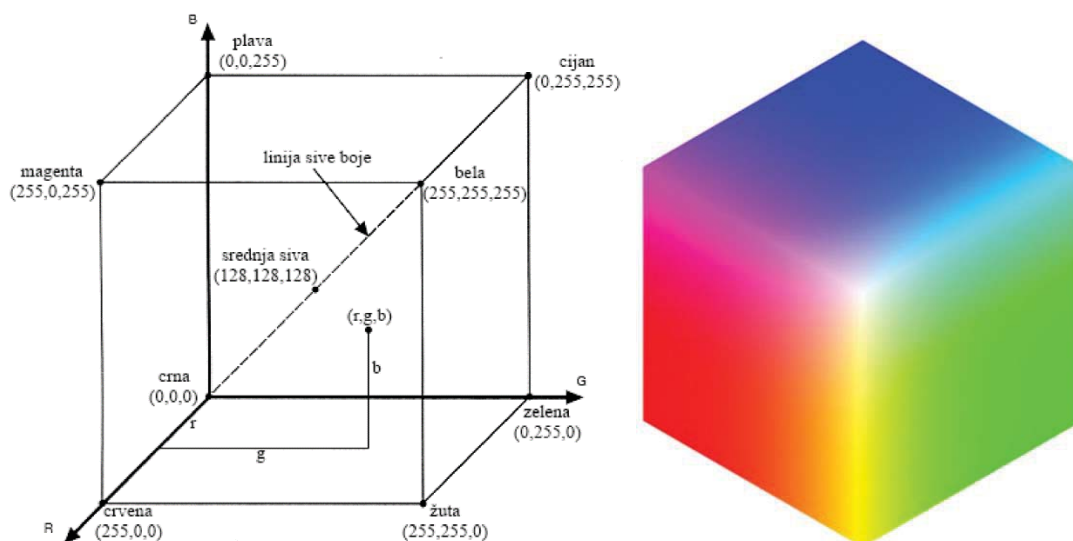
Detaljnijom analizom mrežnjače je utvrđeno da prosečno ljudsko oko ima oko 6 do 7 miliona čunjića i da je otprilike 65% njih je osetljivo na crvenu svetlost, 33% na zelenu svetlost, a samo 2% na plavu (iako su ti čunjići najosetljiviji).

Sistemi kompjuterske vizije za akviziciju podataka iz okoline uglavnom koriste digitalne kamere bazirane na CCD ili CMOS tehnologijama koje pretvaraju svetlost koja pada na senzor kamere prvo u analogni signal, a nakon toga u digitalni [23]. Senzori u digitalnim kamerama se sastoje od 3 tipa fotoosetljivih filtera (R, G i B) raspoređenih u određenom obrascu da bi se što bolje aproksimovao rad ljudskog oka. Primer obrasca filtera je npr. Bajerov filter koji možemo videti na sledećoj slici, a čiji obrazac sadrži 50% zelenih, 25% crvenih i 25% plavih ćelija (Slika 2.7).



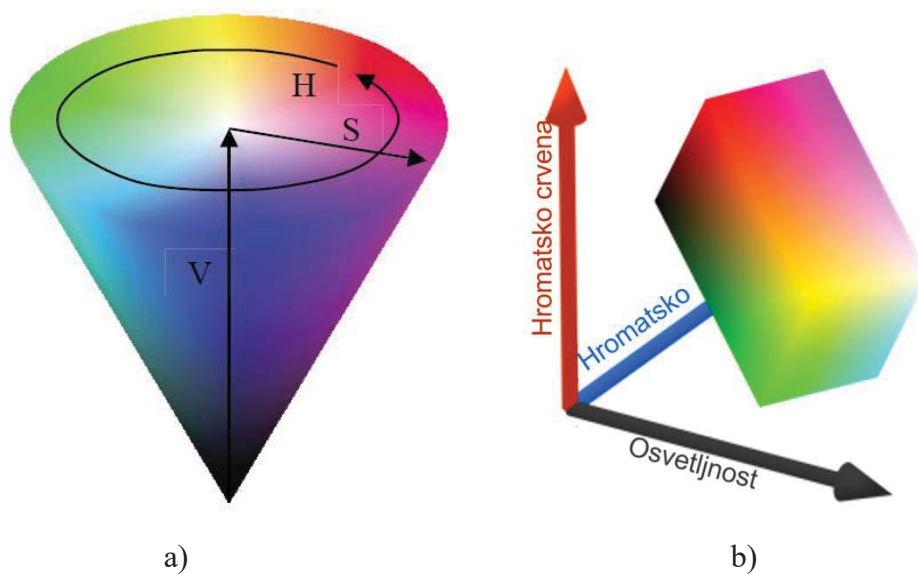
Slika 2.7 Primer Bajerovog obrasca na senzoru kamere [26]

U RGB prostoru boja svaka boja je sastavljena od primarnih spektralnih komponenti crvene, zelene i plave boje. Ovaj se model bazira na Dekartovom koordinatnom sistemu kao što je prikazano na slici 2.8.



Slika 2.8 RGB prostor boja [23]

Kao što se može videti na slici 2.8, primarne boje RGB prostora se nalaze na osima koordinatnog sistema dok su na uglovima sekundarne boje (cijan, magenta i žuta). Koordinatni početak predstavlja crnu boja, dok je njemu najudaljeniji ugao bela boja. Zbog svoje jednostavnosti RGB prostor boja je najrašireniji u upotrebi, ali radi se o prostoru koji nije perceptualno uniforman što znači da euklidska udaljenost između dve boje u tom prostoru ne odgovaraju perceptualnoj udaljenosti tih dveju boja kod ljudske percepcije boja, niti je ovakav prostor boja intuitivan s obzirom na ljudski način razmišljanja o bojama.



Slika 2.9 HSV prostor boja a) i YCbCr prostor boja b) [28]

Još jedan popularan široko zastupljen prostor boja je tzv. HSV (engl. Hue, Saturation, Value) prostor koji ima oblik kupe i može se koristiti za segmentaciju boja. Njegove tri komponente su nijansa (H), zasićenje (S) i vrednost (V) i ovaj prostor boja prikazan je na slici 2.9 a). Treći prostor boja koji je u širokoj upotrebi je YCbCr, prikazan na slici 2.9 b), koji je zapravo digitalni ekvivalent YUV prostoru boja koji se koristi(o) u prenosu TV signala.

Osim njih, u obradi slike se često koriste crno-bele (engl. grayscale) slike ukoliko je bitno samo osvetljenje. Konverzija iz RGB slike u crno-belu sliku se vrši primenom sledeće jednačine, koja računa percepciju svetline različitih boja u ljudskom oku:

$$Y = 0.299R + 0.587G + 0.114B \quad 2.1.$$

gde Y predstavlja osvetljenje crno-belog piksela, a R , G i B predstavljaju intenzitete crvene, zelene i plave boje odgovarajućeg piksela u boji.

2.4. Analiza i obrada slike u prostornom i frekventnom domenu

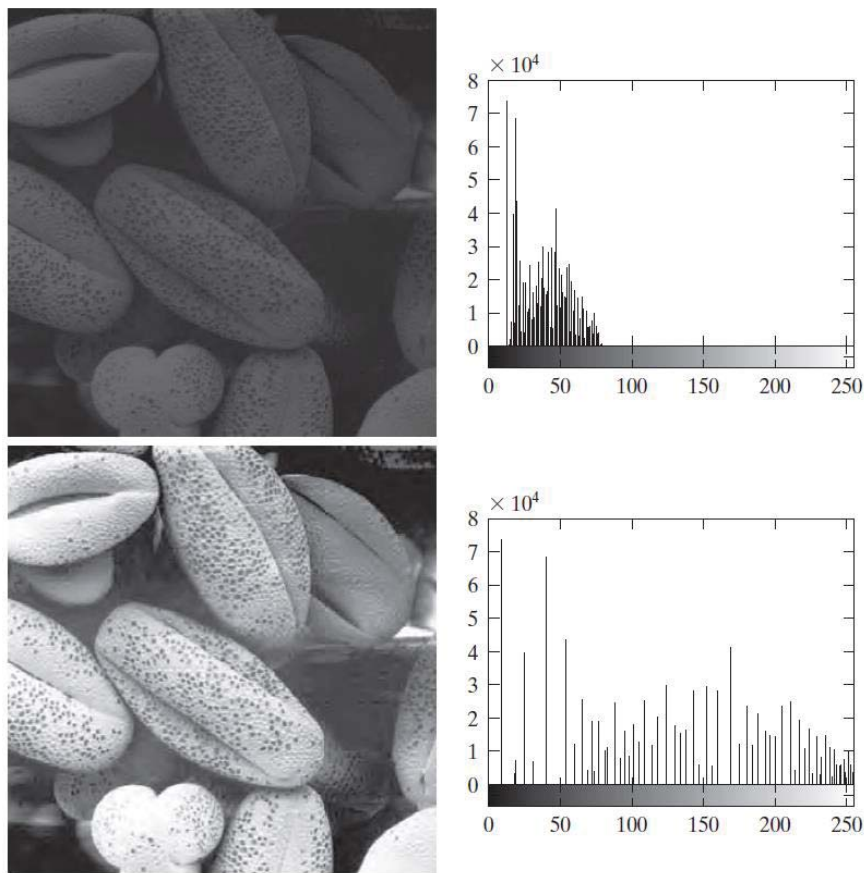
Analiza i obrada slika fokusira se na dva glavna zadatka, i to na poboljšanje informacija sa slike zarad ljudskog razumevanja i procesiranje slike zarad čuvanja, prenosa i korišćenja u mašinskoj, kompjuterskoj ili robotskoj viziji. Teško je odrediti kada prestaje procesiranje slike, a počinje kompjuterska vizija.

Obrada i analiza slika može se vršiti bilo u prostornom domenu, bilo u frekventnom domenu. U prostornom domenu se sve operacije vrše nad pikselima slike koji su direktno dobijene akvizicijom, vrše se operacije nad matricama koje predstavljaju sliku ili se dodaju filteri u vidu maski matrica.

Tehnike obrade slike u frekventnom domenu se baziraju na operacijama nad furijeovom transformacijom slike (najčešće crno-bele slike), gde frekvence predstavljaju količinu promene nijanse sive (intenziteta boje) u zavisnosti od rastojanja. Visoke frekvence karakterišu velike promene u nijansama sive na malim rastojanjima i najčešće odgovaraju ivicama i šumu, dok niske frekvence karakterišu male promene nijansi sive i odgovaraju pozadini ili teksturi površina. Pri procesiranju slike u frekventnom domenu jako često se primenjuju filteri visoke frekvence (engl. High pass filter) koji uklanja niskofrekventne komponente i filteri niske

frekvence (engl. Low pass filter) koji propuštaju komponente niske frekvence i eliminišu komponente visoke frekvence.

Za procesiranje slike jedna od izuzetno bitnih karakteristika je i histogram. Kod crno-belih slika se ovaj histogram sastoji od odgovarajućih nivoa sive, odnosno grafik predstavlja koliko puta se svaka nijansa sive pojavila na slici (slika 2.10). Kod tamnih slika su nivoi sive grupisani bliže koordinatnom početku, kod svetlih su oni grupisani dosta daleko od koordinatnog početka grafika, dok se kod slika sa dobrim kontrastom raspoređeni duž celog opsega histograma.



Slika 2.10 Histogram crno-bele slike i ujednačavanje histograma [24]

Kako bi izoštrili sliku, potrebno je njen histogram "razvući" duž celog opsega, što se može učiniti pomoću dva metoda, istezanjem histograma (engl. Histogram Stretching) ili ujednačavanjem histograma (engl. Histogram Equalization, slika 2.10) koji je potpuno automatizovana procedura [24].

2.5. Senzori za prepoznavanje okoline i 3D senzori

Za potrebe lokalizacije robota i inteligentne interakcije robota sa okolinom i ljudima mogu se koristiti po prirodi različiti senzori kao što su ultrazvučni i laserski, ili senzori robotske vizije poput optičkih kamera, time-of-flight kamera i kamera sa strukturiranom svetlošću. Kao senzori robotske vizije u servisnoj i rehabilitacionoj robotici posebno su prikladni 3D senzori [29].

3D laserski senzor (engl. Laser Range Finder, LRF) u svakom trenutku skenira samo jednu tačku u prostoru. Bazira se na merenju vremena koje je potrebno laserskom zraku da pređe put od senzora do tačke u prostoru koju merimo i natrag do senzora. Uz pretpostavku da je brzina svetlosti konstantna i poznata, udaljenost se računa pomoću izraza:

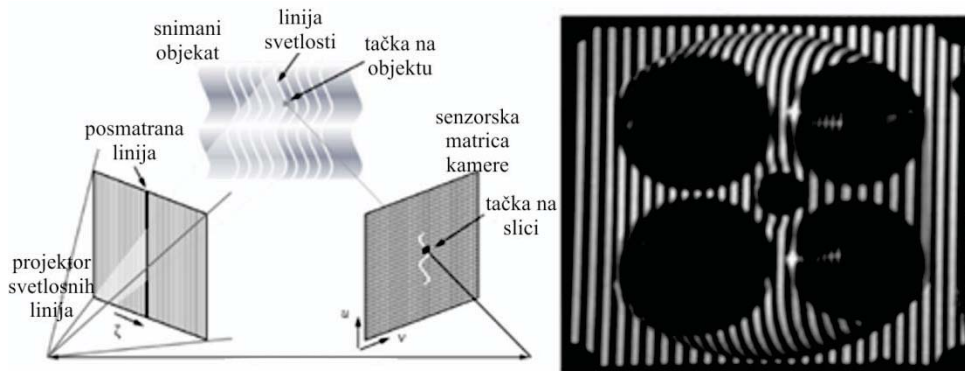
$$s = \frac{c \cdot t}{2} \quad 2.2.$$

gde je s pređeni put, c je brzina svetlosti, a t proteklo vreme. Korišćenjem rotirajućih ogledala ovaj senzor može relativno brzo skenirati celu scenu. Nedostaci se ogledaju u visokoj ceni i u osetljivosti na reflektivne površine. Ultrazvučni senzor funkcioniše po istom principu.

3D senzorskih tehnologija u praktičnoj primeni ima puno buduću da se uglavnom ne radi o novim tehnologijama, a primena zavisi prvenstveno od zahteva koji se pred njih stave. Trenutno su opšte prihvaćene u medicini, geodeziji, industriji zabave i računarskoj viziji. Ni jedna od tehnologija za 3D senzorku nije razvijena s isključivom primenom u robotici, već je uglavnom robotika preuzela te tehnologije iz drugih grana tehnike.

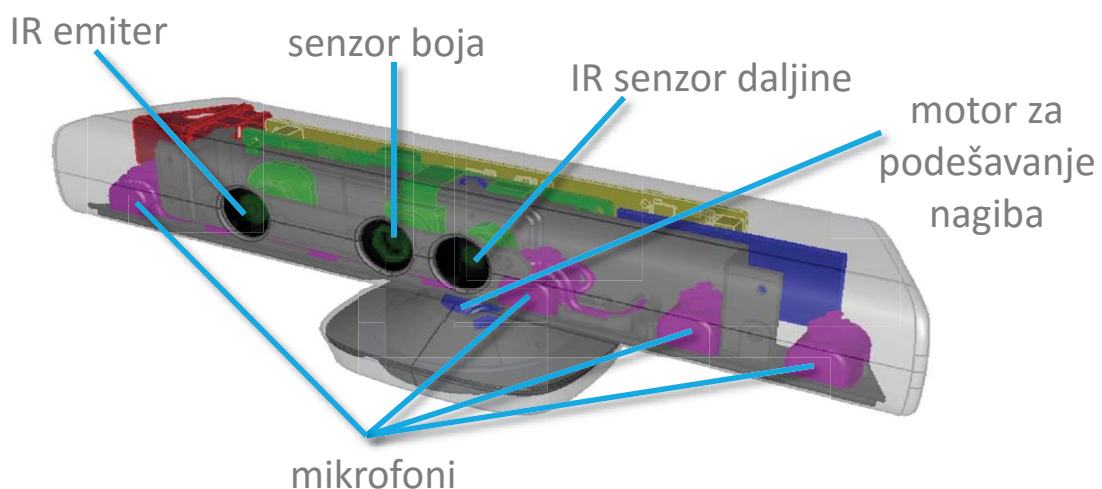
Time-of-flight kamere [29] predstavljaju napredniji oblik laserskih senzora za merenje udaljenosti laserom koje su bazirane se na istom principu kao i LRF senzor tako da meri vreme putovanja svetlosnih zraka, s tom razlikom što je kod ovih kamera u uređaj ugrađeno 2D polje senzora, pa se istovremeno mere podaci o dubini cele scene. Noviji senzori obično daju dubinsku sliku veličine 320x240 tačaka, dok brzina osvežavanja slike može biti i do 100Hz. Različite tehnologije ovih kamera omogućavaju široke opsege snimanja koji se kreću od nekoliko cm do nekoliko kilometara uz rezoluciju od oko 1 cm. Relativno su skupe i osetljive na intenzitet prirodne svetlosti. U slučaju korišćenja više ovakvih senzora u isto vreme može doći do pojave interference, a moguća je i značajne greška merenja prilikom skeniranja okoline s više reflektujućih površina.

3D senzori sa strukturiranom svetlošću projektuju određeni uzorak svetlosti na prostor koji se želi snimiti i zatim snimaju reflektovanu svetlosti pomoću kamere. Uočavanjem pomaka u očekivanoj slici uzorka može se odrediti udaljenost od kamere kao što se vidi na slici 2.11, gde se koristi uzorak svetlosnih linija za detekciju pomaka.



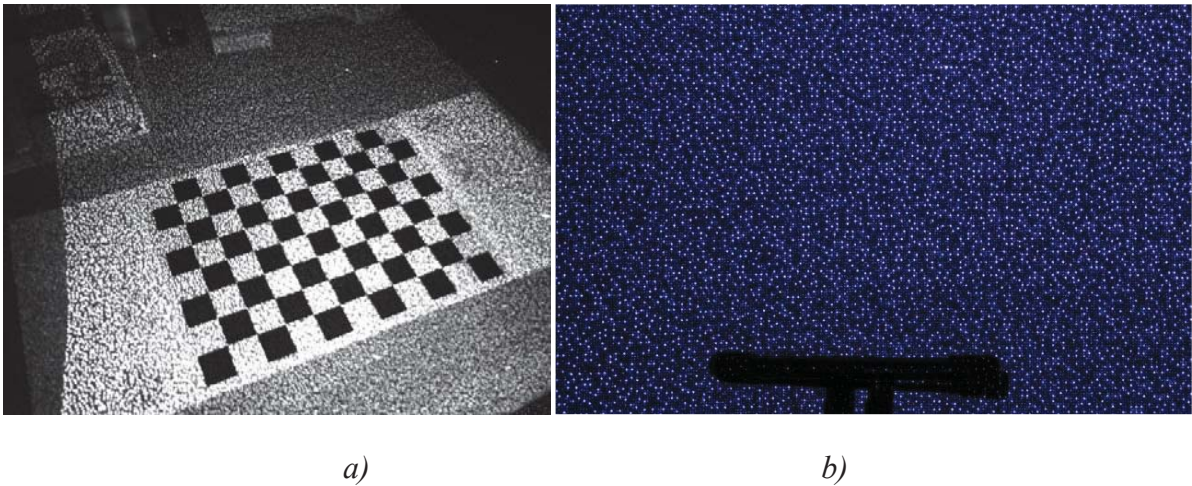
Slika 2.11 Princip rada kamera sa struktuiranom svetlošću [29]

Relativno im je mali domet i imaju velikih problema u slučaju skeniranja reflektujućih površina. Primer najnovijeg komercijalno dostupnog 3D senzora ovog tipa je Microsoft Kinect prikazan na slici 2.12. Originalno zamišljen kao dodatak kućnog sistema zabave, pokazao se izuzetno koristan i u druge svrhe. Identično rešenje sa neznatno izmenjenim karakteristikama poseduje ASUS Xtion Live 3D senzor i koristi iste softverske biblioteke kao i Microsoft Kinect.



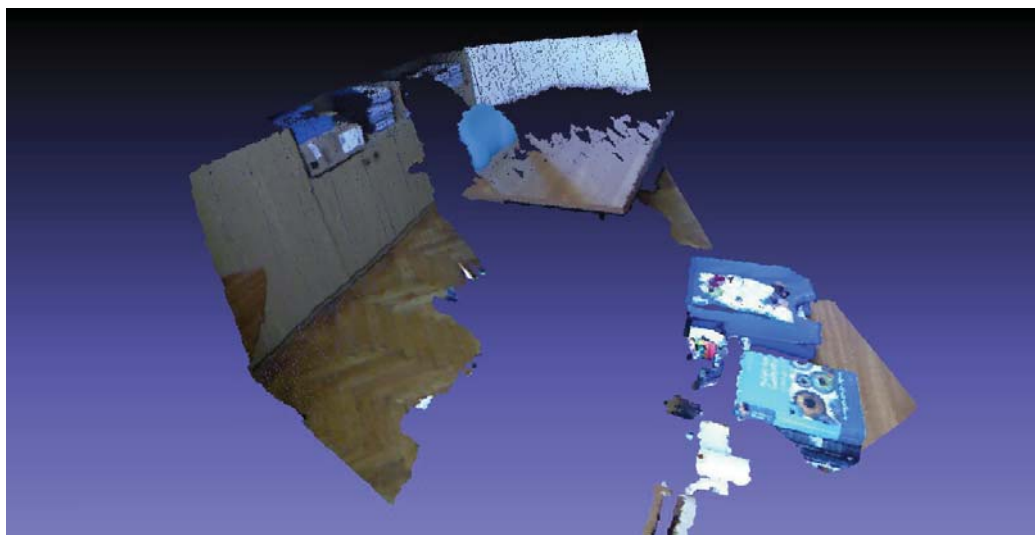
Slika 2.12 Microsoft Kinect [29]

Kinect senzor funkcioniše tako što projektuje uzorak tačaka iz infracrvenog spektra i potom snima refleksiju svetlosti na sceni pomoću infracrvene kamere. Uzorak svetlosti na ravnoj površini i na tipičnoj sceni je vidljiv na slici 2.13. Upoređujući položaj snimljene tačke s određenim referentnim položajem odnosno modelom vrši se triangulacija i izračunava se dubina te tačke, odnosno rastojanje između snzora i same tačke u prostoru.



Slika 2.13. a) Infracrvena slika snimljena Kinect senzorom, b) uzorak emitovane svetlosti na ravnoj površini [29]

Kinect daje izuzetno gustu dubinsku sliku rezolucije 640x480 tačaka pri osvežavanju slike od 30Hz. Dubinska slika je 11-bitna što rezultira s 2048 merljivih nivoa dubine. Kinect senzor ima širinu horizontalnog pogleda od 58° i vertikalnog od 45°. Prema specifikacijama proizvođača preporučeno radno područje je od 0.7m do 5m. Uz Kinect senzor je moguće koristiti službeni SDK (engl. Software Development Kit). Korišćenjem dobijene dubinske slike i RGB slike snimljene kamerom koja je takođe ugrađena u kućište Kinect-a, omogućeno je stvaranje oblaka obojenih tačaka u 3D prostoru kao što se vidi na slici 2.14.



Slika 2.14 Obojeni oblak tačaka snimljen Kinect-om [29]

2.6. 3D i Stereo vizija

U robotici se najčešće koristi optičke video kamere kako bi robot došao do informacija o svojoj okolini, bilo da se radi o monokularnom sistemu, tj. sistemu zasnovanom na jednoj kameri ili stereo sistemu s dve kamere, a postje sistemi i sa više kamera. Pomoću monokularnih sistema podatke o dubini možemo dobiti jedino korišćenjem određenih algoritama kao što su npr. algoritmi za određivanje „strukture iz pokreta“ (engl. structure from motion). Korišćenjem stereo vizije s dve ili više kamera dobijamo dva ili više pogleda na istu scenu i mogućnost dobijanja informacija o dubini scene na relativno jednostavan način. Koristeći pravila određenih geometrijskih ograničenja uslovljenih „pinhole“ modelom kamere i triangulacijom možemo doći do položaja neke tačke u 3D prostoru [30].

Najveći problem kod stereo vizije je određivanje korespodencija između pogleda jedne i druge kamere. Posmatranje iste scene pomoću minimalno dve kamere i pronalaženje odnosa među slikama omogućava određivanje mape dispariteta čije vrednosti predstavljaju razlike koordinata projekcije tačaka u prostoru na levoj i desnoj slici. Iz takve mape dispariteta moguće je izračunati udaljenost tačaka vidljivih na slikama u odnosu na stereo sistem kamera. Dubinske informacije dobijene na ovakav način nam omogućuju stvaranje dubinskih mapa.

2.7. Segmentacija slike

Svrha segmentacije slike je da pojednostavi sliku, podeli je u pod-regione ili setove piksela kako bi sadržaj bio razumljiviji i jednostavniji za računarsku analizu. Drugačije rečeno, slika kompleksne scene se transformiše procesom segmentacije kako bi se izdvojili elementi bitni za konkretnu aplikaciju i kako bi se uklonili nebitni detalji sa slike. Cilj je dakle izdvojiti objekte od interesa od pozadine tako što se svakom pikselu slike dodeli labela tako da pikseli sa istom labelom dele istu specifičnu vizuelnu karakteristiku [31, 32].

U zavisnosti od primene, postoji veliki broj algoritama za segmentaciju koji se baziraju ili na oblastima ili na granicama između oblasti, a mogu se vršiti u otvorenom ili zatvorenom kolu. Kako bi se slika pripremila za segmentaciju, često je neophodno primeniti neku vrstu pre-procesiranja.

Na slici 2.15 prikazani su različiti tipovi segmentacije objekata. Slika 2.15 a) predstavlja originalnu sliku, dok slika 2.15 b) dobijena binarnom segmentacijom slike 2.15 a) kod koje je uklonjena pozadina. Slika 2.15 d) je tipičan primer segmentacije po boji slike 2.15 c), dok 2.15 f) predstavlja segmentaciju objekata na osnovu dubinske slike scene prikazane na slici 2.15 e).



a)



b)



c)



d)



e)



f)

Slika 2.15 Originalne i segmentirane slike [25]

3. DETEKCIJA I PREPOZNAVANJE OBJEKATA

U cilju detekcije i prepoznavanja objekata u robotici se uglavnom koriste sistemi koji se zasnivaju na tehnikama računarske vizije ili sistemi koji podatke dobijaju od 3D laserskog skenera, a često se koriste i oba senzora istovremeno. U suštini dva različita problema se mogu javiti prilikom manipulacije objektima u servisnoj robotici i potpuno je različit pristup rešavanju ovih problema. Naime, u koliko se u sceni nalaze unapred poznati objekti za koje imamo 3D modele traži se najbolje preklapanje snimljenog oblaka 3D tačaka sa modelom. U koliko se srećemo sa novim objektima ili nemamo unapred poznate 3D modele objekata, zadatak poprima sasvim drugi kontekst, pa se bilo vrši 3D rekonstrukcija ili se određuju tačke u prostoru koje su pogodne za hvatanje robotskom rukom. Kako je zadatak razmatranih robotskih sistema bio ili da se kreću ili da manipulišu objektima u realnom okruženju, a ne u strogo kontrolisanim uslovima, svakako da će se sretati sa novim objektima, a i u određenoj klasi objekata nalaziće se fizički različiti objekti. Ovo otežava zadatak upravljanja robotom i iziskuje pristup koji u nekim svojim segmentima mora koristiti tehnike iz domena veštačke inteligencije i mašinskog učenja.

3.1. Pregled istraživanja

U radu [33] autori se bave prepoznavanjem aktivnosti ljudi i njihovom interakcijom sa objektima u sceni na osnovu podataka dobijenih sa RGB-D (engl. Red-Green-Blue-Depth) kamere (poput Microsoft Kinect-a). Veliki je akcenat u ovom radu na prepoznavanju i praćenju objekata, kako bi se odredila njihova dostupnost robotu. Prepoznavanje i praćenje objekata kao i njihova interakcija sa ljudima se obavlja po algoritmu koji je već predstavljen u uvodnim razmatranjima drugog poglavlja ove teze. Nakon temporalne segmentacije frejmova, određuju se karakteristike vezane za objekte i ljude u sceni, nakon čega se prepoznavanje objekata i aktivnosti ljudi vrši primenom SSVM (engl. Structural Support Vector Machine - strukturni metod potpornih vektora) klasifikatora.

Osim ovog pristupa javljaju se i ideje [34] primene proprioceptivnih i audio senzora u cilju prepoznavanja objekata, ali se tek integracijom sa senzorom vizije može govoriti o autonomnom sistemu koji može manipulirati objektima.. Primenom kNN (k najbližih suseda,

engl. k Nearest neighbour) klasifikatora na podatke dobijene sa ovih senzora autori u radu [34] predstavljaju novu metodologiju prepoznavanja objekata.

Autori u svom radu [35] predstavljaju koncept upravljanja robotom u kućnim uslovima na osnovu podataka sa stereovizijske kamere. Neuronska mreža se koristi za određivanje težina koje optimalno određuju oblast od interesa (ROI), a u kome se nalazi određeni objekat, a zatim se vrši segmentacija kojom se povećava rezolucija slike, tj. broj detalja, koji karakterišu određeni ROI. Tek se tada za taj segment slike određuje dubinska slika, koja omogućava kvalitetnu 3D reprodukciju objekta i samim tim odgovarajuću manipulaciju.

Pre segmentacije i modeliranja objekata ne dolazi do interakcije robota i objekta. Eventualno se kamera nalazi na robotskoj ruci i pasivno snima objekat u cilju segmentacije, prepoznavanja i modeliranja. U koliko robot treba da interaguje sa nepoznatim objektima u radu [36] autori predlažu aktivnu segmentaciju i modeliranje objekata kroz kontrolisan fizički kontakt između robota i objekta. Ovakvim pristupom moguća je autonomna i precizna segmentacija novih objekata, čak i u "pretrpanim" scenama, njihova reprodukcija i kontinualno on-line učenje robota. Ipak ovakav metod je bio ograničen na prepoznavanje i manipulaciju simetričnim i skoro-simetričnim objektima.

U koliko je okruženje je strogo kontrolisano, moguće je u prenatrpanim scenama prepoznati objekte i striktnim prepoznavanjem 3D geometrije, ukoliko su nam 3D modeli objekata predefinisani, kako su autori pokazali u [37]. Osim prepoznavanja objekata, autori predloženom metodom kao izlaz imaju i pozu objekta, što omogućava jednostavnu manipulaciju, odnosno hvatanje objekta. Prepoznavanje objekata, odnosno striktno poklapanje 3D geometrije se u njihovom radu odvija u dve faze, off-line pre-procesiranje i on-line prepoznavanje, kojim se određuje sličnost između modela i objekata.

Kao što se može videti iz [37, 38], poza objekta je, ukoliko objekat nije simetričan, od izuzetnog značaja za adekvatnu robotsku manipulaciju. Kako bi se prevazišla nepouzdanost koju unose senzori vizije i senzori za 3D mapiranje prostora, uvođenje podataka sa taktilnih senzora može obezbediti pouzdanost pri određivanju poze i pozicije objekta, posebno u pretrpanim scenama gde se objekti međusobno zaklanjaju [38].

Pozu objekta je kod unapred poznatih objekata moguće relativno jednostavno oceniti u koliko se 3d senzor kreće i poznate su nam brzine i ubrzanja tačaka snimljenog objekta [39], ali je zbog nepouzdanosti sistema vizije i šumova ovo jako teško realizovati na realnom sistemu.

Jedan od najvažnijih segmenata servisne i industrijske robotike koja koristi senzore vizije kao primarne senzore jeste rekonstrukcija 3D scene [40]. U radu [40] autori daju osvrt na ovu problematiku iz ugla primene sistema kompjuterske vizije u proizvodnji.

U servisnoj robotici neophodna je brzo i robusno prepoznavanje objekta, kao i fleksibilna adaptacija robota na okruženje. Ovom problemu se u radu [41] pristupa primenom dva sistema 2D vizije kod robotske ćelije sa dve robotske ruke koje imaju po 6 stepeni slobode kretanja, a koja je obučavana da služi i servira piće. Jedna kamera je fiksna i montirana tako da posmatra celu scenu, dok je druga na kraju jedne robotske ruke. U cilju prepoznavanja objekata korišćeni su algoritmi detekcije okvira, detekcije ivica i poklapanja sa unapred poznatim oblicima.

Hvatanje novih objekata, za koje ne postoji prethodno napravljen 3D model je jako zahtevan problem. U radu [42] su autori predstavili pristup određivanja tačaka na nepoznatom objektu koje su pogodne za hvatanje robotskom rukom, koji se zasniva na Bajesovom modelu verovatnoće. Za obučavanje su koristili 2500 sintetičkih podataka i kasnije su predložene algoritme testirali sa dva različita robota, čiji su sistemi vizije bili Bumblebee stereovizijske kamere.

Istraživanja predstavljena u ovoj tezi se oslanjaju na zajednički naučni rad koje su izveli istraživači sa Mašinskog fakulteta Univerziteta u Nišu i istraživači sa Instituta za automatiku Univerziteta u Bremenu u okviru dva bilateralna projekta, "Robusni sistem vizije u rehabilitacionoj robotici" i "Novi pristup detekciji i praćenju ljudi u robotici". Kao članovi jedne od vodećih naučno-istraživačkih institucija na polju rehabilitacione robotike u Nemačkoj, istraživači na čelu sa šefom instituta, prof. Graeserom, su razvili mobilnu robotsku platformu za pomoć osobama sa invaliditetom, FRIENDII [43], kojoj su dodali sistem stereo vizije i segmentaciju u zatvorenoj sprezi kako bi poboljšali prepoznavanje objekata [44]. Osnovna ideja je uvođenje povratne sprege u obradu slike, gde se u skladu sa podacima dobijenim povratnom spregom menjaju parametri procesiranja slike, odnosno segmentacije [45,46]. Podaci dobijeni nakon ovakve upravljive segmentacije se mogu upotrebiti za rekonstrukciju objekata [47], ili za klasifikaciju objekata [48]. U radu [48] se klasifikacija vrši nakon segmentacije u zatvorenoj sprezi i određivanja karakteristika proporcionalnosti i konektivnosti primenom ANFIS klasifikatora.

Sistem vizije kod robotske platforme za pomoć osobama sa invaliditetom FRIEND je osnova za razvoj sistema upravljanja ovom robotskom platformom [49, 50].

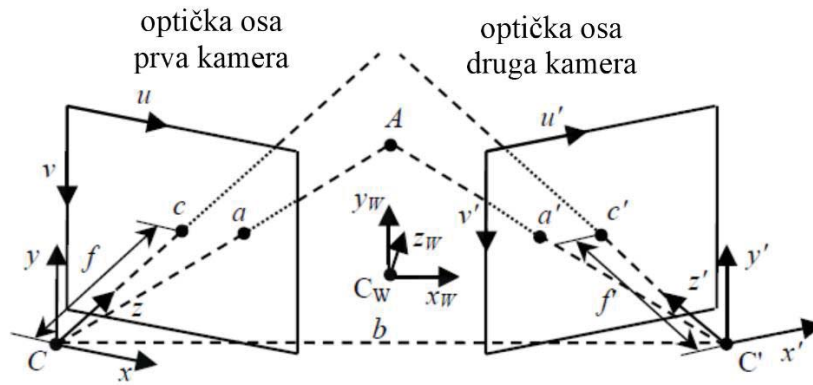
Kako su istraživanja predstavljena u ovoj tezi sprovedena u okviru pomenutih zajedničkih projekata, a vršeni su eksperimenti na FRIEND II robotskoj platformi, u nastavku ovog poglavlja biće predstavljeni osnovni algoritmi korišćeni u upravljanju sistemom vizije ovog robotskog sistema, počev od akvizicije stereo slike, formiranja mape dispariteta, preko segmentacije, odnosno segmentacija, do ekstrahovanja karakteristika i klasifikacije. Sistem vizije ove robotske platforme činio osovu za sva dalja istraživanja koja su predstavljena u ovoj tezi.

3.2. Stereovizija i kalibracija stereo kamere

Glavni nedostatak monokularnih kamera je što se prilikom projekcije 3D scene na 2D površinu slike gubi podatak o rastojanju objekata od kamere. I pored toga što postoje neki metodi za ocenu dubinske slike na osnovu senki i unapred poznatih informacija o sceni, ovi metodi su izuzetno neprecizni i nepouzdati, te se ne mogu primeniti kod robotskih aplikacija. Dubinske informacije se mogu preciznije oceniti ukoliko se koriste dve ili više kamere i poznato je njihovo međusobno rastojanje i usmerenje. Sistem dve kamere se naziva stereo kamera i funkcioniše slično ljudskom vidu za ocenu dubine. Dubina se ocenjuje na osnovu razlike između projekcija slike na dve ravni slika, tj. triangulacije [25, 30, 51].

U zavisnosti od usmerenja dve kamere, stereo kamera može biti konvergentna, u koliko kamere gledaju jedna ka drugoj i njihove optičke ose se seku, kao što je prikazano na slici 3.1, divergentna, u koliko se optičke ose ne seku i gledaju jedna od druge, ili može biti paralelna, u koliko su optičke ose kamere međusobno paralelne.

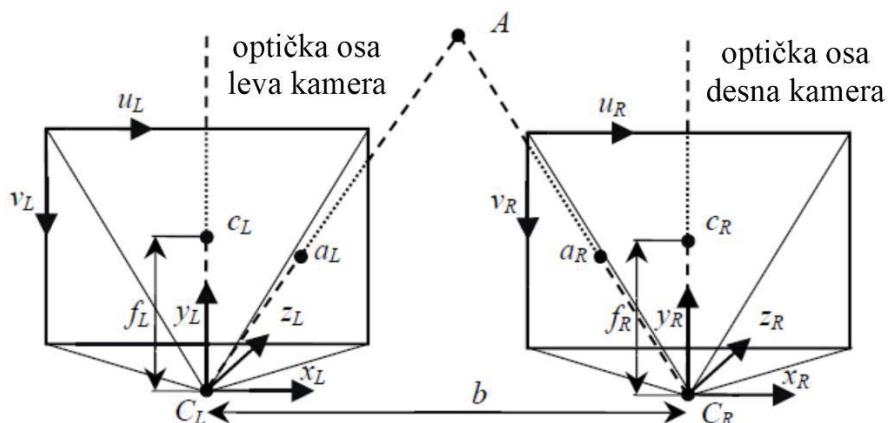
Paralelna konfiguracija stereo kamere omogućava projektovanje mape dispariteta, koja predstavlja crno-belu sliku kod koje intenzitet osvetljenja piksela odgovara rastojanju tog piksela od kamere. Ipak, projektovanje mape dispariteta iziskuje da žižne daljine sočiva obe kamere budu identične, kako bi obe ravni slika ležale na istoj geometrijskoj ravni.



Slika 3.1 Prikaz Stereovizijske kamere sa konvergentnim optičkim osama [30]

Kod stereo kamere svaka od dve kamere ima po dva koordinatna sistema, koordinatni sistem slike (u, v) i koordinatni sistem kamere (x, y, z), koji su najčešće indeksirani različito, kako bi se znalo o kojoj je kameri reč. Dodatni parametar koji je bitan kod stereovizijske kamere jeste rastojanje između centara dve kamere, osnovna linija b . Presek osnovne linije sa dve ravni slika daje dva epipola. Sa slike 3.2 se može videti da centri dve kamere sa tačkom A , koja je tačka u prostoru koja se posmatra - snima, formiraju ravan. Ova ravan se naziva epipolarna ravan i njen presek sa dve ravni slika formira dve epipolarne linije, koje se nalaze na ravnima slika. Sa slike se vidi da se dve međusobno odgovarajuće tačke a i a' , a koje predstavljaju projekciju tačke A , nalaze na odgovarajućim epipolarnim linijama

U ovoj tezi je korišćena stereo kamera sa paralelnim optičkim osama, čiji je šematski prikaz funkcionisanja predstavljen na slici 3.2, gde su sa indeksom L obeležene sve ose i karakteristične tačke leve kamere, dok su sa indeksom R označene kod desne kamere.



Slika 3.2 Stereovizijska kamera sa paralelnim optičkim osama [30]

Kod stereo kamera uobičajeno je da se koordinatni sistem celog senzora vizije veže za centar leve kamere. Stereo kamere sa paralelnim optičkim osama imaju specifične vrednosti rotacione i translacione matrice, koje opisuju rotaciju, odnosno translaciju u odnosu na koordinatni sistem celog senzora, i u idealnom slučaju su:

$$R_L = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad t_L = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

3.1.

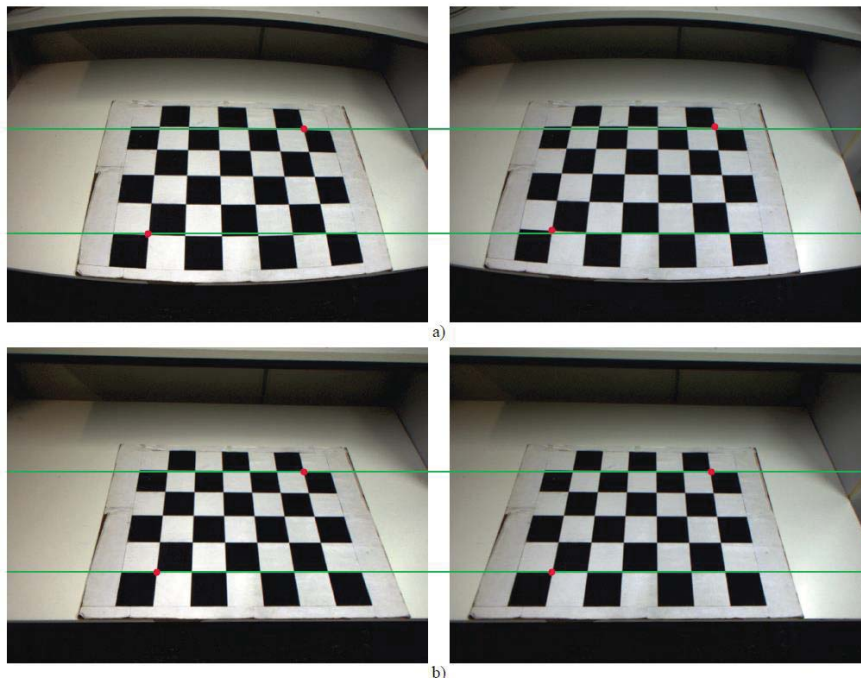
$$R_R = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad t_R = \begin{bmatrix} -b \\ 0 \\ 0 \end{bmatrix}$$

gde su R_L i t_L rotaciona matrica i translacioni vektor leve kamere, a R_R i t_R rotaciona matrica i translacioni vektor desne kamere. Osnovna linija b je upravo rastojanje za koje je desna kamera translirana duž x ose, kao što se može videti na slici 3.2.

Kako se zbog nesvršenosti sočiva kamera u realnim uslovima ne može javiti ovakav idealni slučaj, parametri rotacionih matrica i translacionih vektora se moraju oceniti kalibracijom stereo kamere. Postoje standardizovane metode kalibracije koje koriste šablon šahovske table. Ovakvo ocenjivanje parametara omogućavaju epipolarna geometrija i činjenica da međusobno odgovarajuće tačke moraju da leže na epipolarnim linijama. Tako se uglovi šablona šahovske table sa leve slike koriste za određivanje odgovarajućih epipolarnih linija desne slike. Potom se određuje rastojanje između svakog od detektovanih uglova šablona i odgovarajućih epipolarnih linija. Isti postupak se ponavlja i s desna na levo, pa se statističkim metodama koje podrazumevaju određivanje srednjih vrednosti i devijacija rastojanja konačno ocenjuju parametri i vrši kalibrisanje kamere.

Kao što je već spomenuto, za izračunavanje mape dispariteta neophodno je da optičke ose budu idealno paralelne i da žižne daljine kamera u sistemu budu identične. Kako ni ove uslove nije moguće postići u realnim uslovima, primenjuje se i metod stereo ispravljanja slike. Ovim se metodom generišu slike koje izgledaju kao da su snimljene idealnim sistemom, ali što je realni sistem bliži idealnom, to su i rezultati ispravljanja stereo slike bolji.

Osnovna karakteristika stereo ispravljenih slika jeste da su epipolarne linije paralelne u osi slika i da projekcije iste tačke moraju imati istu v koordinatu na obe slike. Na slici 3.3 prikazana je stereo slika pre i nakon ispravljanja, gde je očigledno da odgovarajuće tačke na levoj i desnoj slici imaju istu v koordinatu.



Slika 3.3 Primer stereo ispravljanja slike a) Originalne leva i desna slika i b) Stereo ispravljene leva i desna slika [25]

Stereo ispravljanje slike može se izvršiti promenom parametara matrica rotacije i vektora translacije tokom procesa kalibracije. I pored toga što je triangulacijom moguće rekonstruisati položaj tačke u prostoru bez obzira na položaj optičkih osa kamera sistema, stereo ispravljanje dovodi do toga da sve odgovarajuće tačke leže na istoj liniji, što značajno ubrzava proces pretrage tačaka i korišćenjem adekvatnih metoda moguće je brzo odrediti mapu dispariteta.

3.3. Određivanje mape dispariteta i dubinske slike

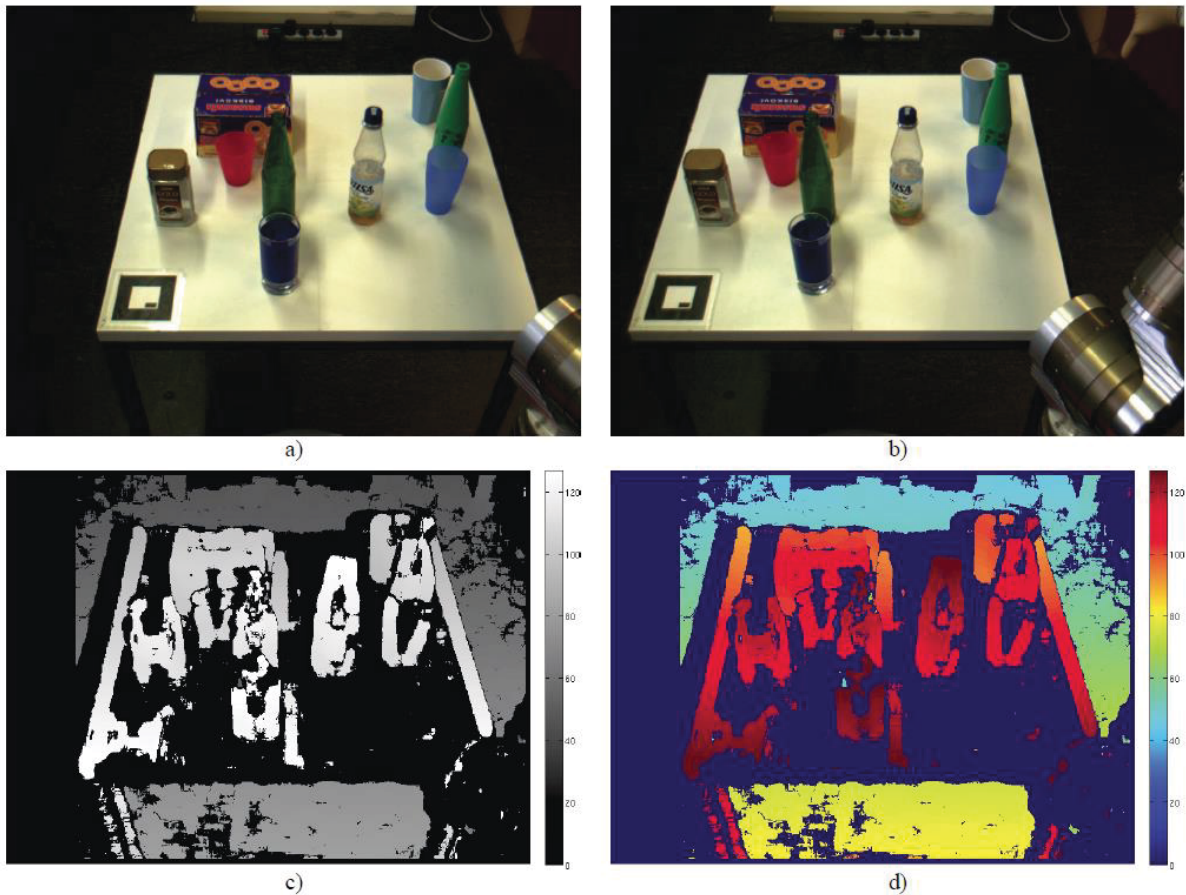
Bez obzira na orijentaciju optičkih osa kamera, 3D pozicija tačke $A(x, y, z)$ se može odrediti stereo triangulacijom, koristeći projekcije tačke A na dve slike, $a(u, v)$ i $a'(u', v')$, kao i projekcione matrice P i P' , rešavanjem sledeće jednačine:

$$\begin{bmatrix} uP_3^T - P_1^T \\ vP_3^T - P_2^T \\ u'P_3'^T - P_1'^T \\ v'P_3'^T - P_2'^T \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \quad 3.2.$$

gde se P_n^T i $P_n'^T$ odnose na n -tu vrstu transponovanih projekcionih matrica leve i desne kamere respektivno P^T i P'^T .

Kako je već naznačeno, stereo ispravljanje slike omogućava izračunavanje tzv. mape dispariteta. Mapa dispariteta predstavlja sliku kod koje vrednost svakog piksela odgovara rastojanju između projekcija tačke u prostoru na slike leve i desne kamere, odnosno njihovoj razlici, što je zapravo obrnuto proporcionalno rastojanju tačke u prostoru od stereo kamere. Svako projekciji tačke na levoj slici sa koordinatama $a(u,v)$ odgovara vrednost dispariteta $d(u,v)$ na mapi dispariteta. Za vizuelizaciju mape dispariteta najčešće se koriste crno-bele slike, kod kojih je intenzitet osvetljenja piksela obrnuto proporcionalan (tamno su oblasti dalje od kamere, svetlo oblasti bliže kameri). Osim ovog načina, moguće je predstaviti mapu dispariteta na slici u boji i koristiti određenu paletu, npr. plavo-crvenu paletu, gde su crvenom bojom označene stvari koje su najbliže kameri (toplo), plavom udaljeni objekti (hladno), a nijansama između njih preko zelene i žute objekti između. Ovakve mape dispariteta prikazane su na slici 3.4.

Najveći problem u određivanju mape dispariteta jeste pronalaženje međusobno odgovarajućih tačaka na levoj i desnoj slici, tzv. stereo uparivanje. Postoji više načina određivanja mape dispariteta koji se mogu podeliti u dve kategorije, lokalni i globalni metodi. Lokalni metodi koriste lokalne susedne piksele odgovarajućem pikselu leve slike kako bi pronašli parnjak na desnoj slici. Globalni metodi koriste ekstrahovane karakteristike sa obe cele slike i uparuju ih, pa na osnovu tih podataka formiraju mapu dispariteta. Postoje i hibridni metodi koji predstavljaju kombinaciju dva unapred pomenuta metoda.



Slika 3.4 Primer formiranja mape disariteteta, a) ispravljena leva slika, b) ispravljena desna slika, c) crno bela slika koja predstavlja mapu disariteteta, d) slika u boji koja predstavlja mapu disariteteta [25]

Lokalni metodi su vrlo brzi u odnosu na globalne metode, ali ne mogu izračunati vrednosti disariteteta za objekte koji se međusobno zaklanjaju na slici ili u slučaju da su objekti isuviše blizu kameri, jer lokalni metodi koriste samo ograničen broj piksela u okolini zadatog piksela za uparivanje. Sa druge strane, globalni metodi prevazilaze ove probleme jer vrše pretragu po celoj slici, ali su relativno spori kako zbog niza potrebnih optimizacija, tako i zbog upotrebe svih piksela prilikom određivanja vrednosti disariteteta samo jednog piksela.

Tokom istraživanja primenjivani su i lokalni i globalni metod formiranja mape disariteteta. U početnoj fazi istraživanja je korišćen globalni metod, gde se tek nakon segmentacije leve i desne slike i klasifikacije objekata, čime se određivala oblast od interesa, triangulacijom određivalo rastojanje željenog objekta od kamere, a samim tim i robota.

U kasnijoj fazi primenjena je tzv. "block matching" metoda određivanja mape dispariteta, odnosno metoda preklapanje blokova. Ovakav algoritam se realizuje brzo, što je od izuzetnog značaja za prepoznavanje i izvršavanje akcija u realnom vremenu, a pokazao se upotrebljiv u velikom broju scenarija. Kod ovog metoda se $n \times n$ blok piksela sa leve slike, gde je piksel od interesa u sredini, upoređuje sa svim mogućim blokovima iste dimenzije desne slike i nalazi se parnjak. Kako je piksel od interesa u sredini bloka, n je uvek neparan broj. Kako je stereo slika ispravljena, primenjuju se i dva ograničenja kako bi se proces ubrzao. Naime, centralni piksel bloka sa leve slike leži na istoj liniji paralelnoj u osi kao i traženi piksel na desnoj slici, tj. ovi pikseli imaju istu koordinatu v . Osim toga, u koordinata centralnog piksela sa desne slike je uvek manja od onog na levoj slici. Ako uzmemo u obzir ova ograničenja, najbolje poklapanje blokova može se odrediti na osnovu funkcije cilja koja predstavlja sumu apsolutnih razlika (SAD) vrednosti piksela blokova sa leve i desne slike:

$$SAD(u, v, d) = \sum_{i=u-\frac{n}{2}, j=v-\frac{n}{2}}^{i=u+\frac{n}{2}, j=v+\frac{n}{2}} |a_L(u+i, v+j) - a_R(u+i-d, v+j)|, \quad d = \overline{0..d_{\max}} \quad 3.3.$$

gde su u i v koordinate tačke leve slike za koju se određuje vrednost dispariteta, d je trenutna vrednost dispariteta, n je veličina bloka koji se koristi za stereo uparivanje, dok a_L i a_R predstavljaju vrednosti piksela leve i desne slike.

Vrednost SAD se računa za sve vrednosti mogućeg dispariteta od 0 do maksimalne očekivane vrednosti dispariteta d_{\max} , a najbolje poklapanje postiže se sa minimalnom vrednošću funkcije SAD :

$$d(u, v) = \text{MIN}(SAD(u, v, d)), \quad d = \overline{0..d_{\max}} \quad 3.4.$$

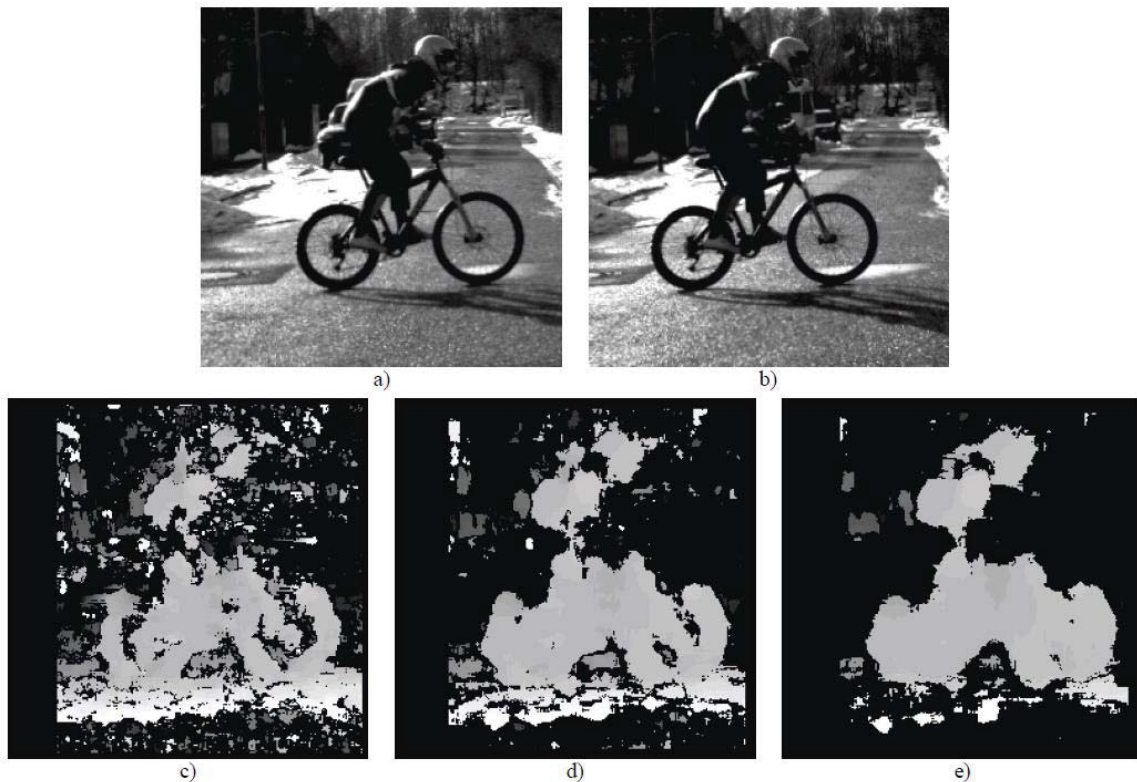
U koliko je minimalna suma veća od neke unapred zadate granice, do čega može doći usled zaklanjanja objekata ili perspektive, nije moguće upariti tačke. U koliko su dve različite minimalne sume prilikom pretrage relativno slične znači da su na desnoj slici pronađene dve različite tačke sa vrednosti dispariteta koje odgovaraju tački na levoj slici, pa nije moguće jedinstveno odrediti tačnu vrednost dispariteta zadate tačke. Do ovoga može doći ukoliko je neka oblast veća od zadanog bloka na slici iste boje, odnosno bez teksture. Moguće je primeniti neke metode post-procesiranja slike kako bi se eliminisale oblasti bez teksture ili zaklonjeni objekti.

Nakon što algoritam pronađe koordinate međusobno odgovarajućih tačaka $a_L(u_L, v_L)$ i $a_R(u_R, v_R)$, vrednost dispariteta d se određuje kao:

$$d = u_L - u_R \quad 3.5.$$

gde su u_L i u_R koordinate međusobno odgovarajućih tačaka u levoj i desnoj slici.

Značajnu ulogu u primeni ovog algoritma igra i veličina bloka, jer se premalim blokom ne može prebroditi problem određivanja mape dispariteta kod slika bez dovoljno teksture, dok se primenom prevelikog bloka ne može postići uparivanje usled različitih perspektiva. Problem sa prevelikim blokom izražen je kod objekata koje su blizu kameri. Uticaj veličine bloka na rezultate određivanja dubinske slike moguće je videti na slici 3.5.



Slika 3.5 Uticaj izbora veličine bloka na mapu dispariteta, a) i b) originalne ispravljene leva i desna slika, c) blok veličine 9x9 piksela, d) blok veličine 15x15 piksela i e) blok veličine 21x21 piksela [25]

Oblasti za koje nije bilo moguće upariti piksele predstavljeni su na slikama mape dispariteta crnom bojom. Kod mape dispariteta gde je korišćen manji blok je objekat na slici

tačnije predstavljen, ali više je oblasti sa šumom i više je neuparenih oblasti. Prilikom primene većeg bloka manje je neuparenih piksela i manje je šuma na mapi dispariteta, ali oblik objekta nije više korektan jer je došlo do takozvanog efekta udebljanja, koji je karakterističan kod primene velikih blokova.

Nakon što nam je za neku tačku $a_L(u_L, v_L)$ poznata vrednost dispariteta d , moguće je rekonstruisati 3D poziciju te tačke u prostoru, $A(x, y, z)$, u odnosu na koordinatni sistem kamere pomoću sledećih jednačina:

$$x = \frac{b \cdot (u_L - u_{cL})}{d}, \quad y = \frac{b \cdot (v_{cL} - v_L)}{d}, \quad z = \frac{b \cdot f}{d} \quad 3.6.$$

gde je f žižna daljina kamere, b je osnovna linija stereo kamere, a $c(u_{cL}, v_{cL})$ je centralna tačka leve kamere.

3.4. Sistemi za robusnu segmentaciju slika u robotskoj viziji

U cilju prepoznavanja objekata se u računarskoj i robotskoj viziji sreće veliki broj algoritama za segmentaciju slike, a među njima postoji i određeni broj inteligentnih algoritama segmentacije. Ipak, ne postoji jedinstvena metodologija izbora algoritma segmentacije, već se svaki sistem računarske (odnosno robotske) vizije projektuje zasebno, u zavisnosti od senzora vizije koji se koriste, namene i zadataka koje robot treba da obavlja.

Postoje dva različita pristupa segmentaciji slike, segmentacijom se mogu određivati oblasti odnosno regioni slike ili se mogu određivati granice tih regiona. U ovoj tezi je korišćena tzv. binarizacija, gde je izlaz segmentacije binarna slika koja ima isključivo bele i crne piksele. Beli pikseli odgovaraju segmentiranom objektu, dok crni pikseli odgovaraju pozadini. Može se govoriti i o dva principa segmentacije, u otvorenom i zatvorenom kolu.

3.4.1. Segmentacija zasnovana na određivanju praga

Prilikom segmentacije koja se vrši u cilju određivanja oblasti na slici vrši se grupisanje piksela po zajedničkim karakteristikama, poput boje, intenziteta osvetljenja, teksture i sl. Najčešće primenjivani metod jeste metod određivanja praga (engl. thresholding). Pregled 40 tehnika segmentacije slike određivanjem praga predstavljen je u radu [52]. Tehnike opisane u ovom radu mogu se uspešno primenjivati na različite sisteme, od analize slike dokumenta u

cilju prepoznavanja teksta, preko analize scene gde se traži ciljni objekat ili se vrši kontrola kvaliteta, do neinvazivne metode testiranja i ispitivanja. Algoritmi segmentacije za neinvazivna testiranja i ispitivanja su razvijeni prvenstveno za analizu slika dobijenih specijalnim uređajima, poput ultrazvučnih slika, termo-slika, endoskopskih slika, slika dobijenih laserskom mikroskopijom, rendgenskih snimaka i sl. Sezgin *et. al.* su u svom radu razvrstali metode određivanja praga na 6 kategorija, i to:

- metodi zasnovane na analizi oblika histograma, gde se analiziraju ekstremne vrednosti i prevojne tačke histograma;

- metodi zasnovani na klasterima, gde se elementi crno-bele slike razdvajaju na dva klastera, gde je jedan pozadina, a drugi klaster je sam objekat;

- metodi koji se zasnivaju na entropiji regiona pozadine i objekta;

- metodi koji se zasnivaju na atributima objekta;

- prostorno orijentisani metodi koji koriste verovatnoću raspodele piksela i korelacije između samih piksela slike

- lokalni metodi, koji u zavisnosti od dela slike za svaki piksel ponaosob određuju prag.

Jedan od najjednostavnijih metoda jeste određivanja praga histograma, gde se vrednosti piksela mogu razdeliti globalnim pragom T_G . Pikseli koji pripadaju segmentiranom objektu na originalnoj slici imaju vrednost manju od T_G , dok pikseli koji su deo pozadine imaju vrednost veću ili jednaku vrednosti praga T_G :

$$t_G = \begin{cases} 1, & \text{if } f(x, y) \geq T_G \\ 0, & \text{if } f(x, y) < T_G \end{cases} \quad 3.7.$$

gde je $t_G(x,y)$ vrednost piksela binarne slike.

Ovaj metod je primenjiv u koliko je na slici samo jedan objekat koji je potrebno segmentisati i pozadina je uniformna. Kako bi se ova ograničenja prevazišla može se definisati interval praga $T=[T_{low}, T_{high}]$, pa je vrednost segmentiranog piksela $t(x,y)$:

$$t(x, y) = \begin{cases} 1, & \text{if } f(x, y) \in T \\ 0, & \text{if } f(x, y) \notin T \end{cases} \quad 3.8.$$

gde je $f(x,y)$ vrednost piksela u koordinati (x,y) , a T_{low} i T_{high} su donja i gornja granica praga primenjena na histogram slike definisane sa $f(x,y)$.

Optimalno određivanje praga histograma ili intervala praga histograma je izazovan i zahtevan problem. U cilju rešavanja ovog problema koriste se automatizovani metodi poput adaptivnog određivanja praga, gde se "maska" kreće po slici i za svaki položaj maske računa se srednja vrednost piksela obuhvaćenih maskom. Još jedan značajan algoritam je tzv. Otsu metod, koji optimalni prag određuje zahvaljujući varijaciji vrednosti piksela objekta i pozadine.

U koliko je potrebno segmentisati sliku na kojoj se nalaze jednobojni ili relativno uniformni objekti od interesa, moguće je segmentisati sliku i koristeći informacije sa slike predstavljene HSV modelom boja, gde digitalna slika predstavljena u dve ravni, tj. dvema matricama, i prva je slika nijansi $f_H(x,y)$, a druga slika je slika zasićenja $f_S(x,y)$.

Kod slike nijansi vrednost piksela se kreće u intervalu $[0, 2\pi]$ ili $[0, 359]$, jer je komponenta nijanse u HSV prostoru boja zapravo ugao H kruga (slika 2.9 a) i svaki od 360 inkrimenata odgovara određenoj nijansi. Kako bi se odredile razlike između različito obojenih objekata, komponenta nijanse je razdeljena u klase boja čije su vrednosti iz intervala $[T_{low}, T_{high}]$:

$$C_l \in [T_{low}, T_{high}], \quad 3.9.$$

gde l predstavlja klasu boje, a T_{low} i T_{high} su minimalna i maksimalna vrednost boje piksela objekta. Primenom istog principa kao kod metode određivanja praga histograma, binarna vrednost piksela segmentirane slike $t_H(x,y)$ određuje se prema sledećem izrazu:

$$t_H(x, y) = \begin{cases} 1, & \text{if } f_H(x, y) \in C_l \\ 0, & \text{if } f_H(x, y) \notin C_l \end{cases}, \quad 3.10.$$

odnosno, u koliko vrednost piksela sa slike nijansi pripada određenoj klasi boje, taj piksel pripada objektu od interesa.

Komponenta zasićenja se u segmentaciji koristi kao maska koja služi za određivanje novih oblasti od interesa, [23], a binarna vrednost segmentiranog piksela dobijenog iz slike zasićenja je:

$$t_s(x, y) = \begin{cases} 1, & \text{if } f_s(x, y) \geq 0,1 \cdot f_s \text{ MAX} \\ 0, & \text{if } f_s(x, y) < 0,1 \cdot f_s \text{ MAX} \end{cases} \quad 3.11.$$

pa je finalni rezultat segmentacije:

$$t(x, y) = t_H(x, y) \wedge t_S(x, y), \quad 3.12.$$

gde je vrednost segmentiranog piksela $t(x,y)$ jednak jedinici, tj. pripada željenom objektu, samo u koliko su ispunjena oba uslova pripadnosti segmentiranom regionu (logičko I). I u ovom slučaju se problem svodi na valjano određivanje intervala praga.

Osim opisanih metoda kod kojih se određuje da li neki piksel pripada ili ne pripada određenoj oblasti na slici, često zastupljene metode se baziraju na određivanju granica objekta, odnosno metodi detekcije ivica. Ovim metodama se detektuju nagle promene vrednosti intenziteta piksela na slici i rezultat segmentacije je binarna slika koja u prvi plan izbacuje ivice sa slike, odnosno delove slike gde dolazi do naglih prelaza.

Osnovna ideja kod detekcije ivice jeste određivanje gradijenta pomoću parcijalnih izvoda prvog ili drugog reda. Gradijent slike $f(x,y)$ u tački (x,y) je:

$$\nabla f = \begin{bmatrix} G_x \\ G_y \end{bmatrix} = \begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{bmatrix} \quad 3.13.$$

Izračunavanje gradijenta vrši se pomeranjem maske po slici. Nakon toga se formira binarna segmentirana slika određivanjem praga koji se primenjuje na matricu gradijenata. Kritičan je ponovo adekvatan izbor praga, jer u koliko je vrednost praga isuviše mala binarna segmentirana slika sadržaće lažno pozitivne ivice, dok će se u slučaju prevelike vrednosti praga javiti lažno negativne ivice. Prilikom optimalnog određivanja praga postoje tri funkcije cilja koje je potrebno ispuniti [53]:

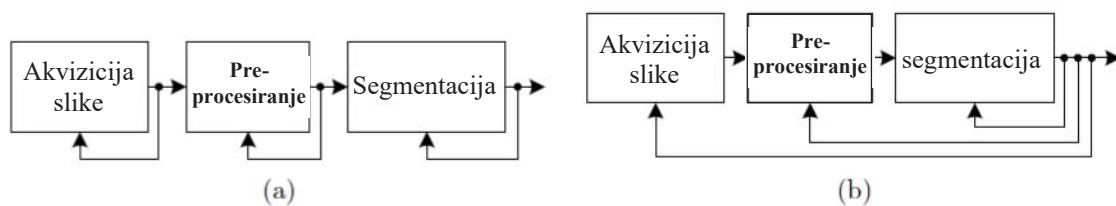
- potrebno je pronaći sve ivice;
- potrebno je da se ivice poklope sa stvarnim ivicama;
- potrebno je da ivice budu što tanje, tj. ne treba da umesto jedne stvarne ivice na binarnoj segmentiranoj slici postoje dve ili više ivica.

U slučaju da je potrebno segmentisati pretrpane scene, često se kombinuju metode detekcije ivice i segmentacije oblasti kao što je predstavljeno u radu [54] i prepoznavanje objekata se vrši istovremeno sa segmentacijom.

3.4.2. Segmentacija u zatvorenom kolu

Algoritam za prepoznavanje objekata u ovom radu zasniva se na robusnoj segmentaciji slike sa povratnom spregom, razvijenom od strane istraživačkog tima iz Bremena, koji se prvi put javlja u doktorskoj disertaciji [22]. Kako prethodno navedeni metodi daju dobre rezultate samo u slučaju strogo kontrolisanih uslova, npr. konstantnog osvetljenja, bilo je neophodno primeniti drukčiji pristup. U realnim uslovima eksploatacije osvetljenje nije konstantno, već se menja u zavisnosti od doba dana, vremenskih uslova, ili u zavisnosti od prisustva veštačkih izvora osvetljenja, pa se samim tim ne mogu ručno i jednoznačno odrediti parametri segmentacije, već se ovi parametri moraju non-stop menjati i prilagođavati trenutnoj sceni kako bi se segmentacijom izdvojio objekat od interesa.

Kako bi se postigla automatska adaptacija parametara segmentacije, u disertaciji [22] i disertaciji [46] predložene su dve upravljačke strukture sa povratnom spregom, sekvencijalna i kaskadna, koje su prikazane na slici 3.6



Slika 3.6 a) Sekvencijalna upravljačka struktura obrade slike i b) kaskadna upravljačka struktura obrade slike [22]

Kod sekvencijalne segmentacije se nakon svakog bloka procesiranja uvodi povratna sprega, te se na osnovu povratne sprege vrši korekcija parametara tog nivoa procesiranja, dok se kod kaskadne segmentacije vrši podešavanje parametara svih parametara svih nivoa procesiranja na osnovu povratne sprege tek nakon što se celokupno procesiranje izvrši.

Kao povratna informacija pri procesiranju s povratno spregom se u ovim tezama predlaže mera kvaliteta izlazne segmentirane binarne slike, odnosno tzv. dvodimenzionalna (2D) mera entropije. Ova mera 2D entropije, koja je u istraživanju korišćena za segmentaciju oblasti, kvantifikuje stepen međusobne povezanosti piksela segmentiranog objekta:

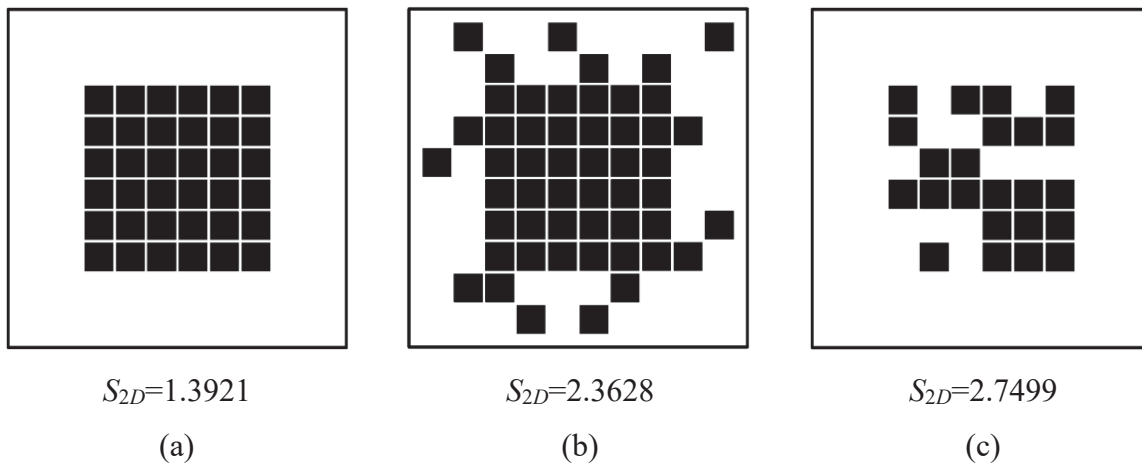
$$S = - \sum_{i=0}^8 p_{(1,i)} \log_2 p_{(1,i)} \quad 3.14.$$

gde je $p(1,i)$ relativna frekvencija, odnosno ocena verovatnoće da segmentirani piksel koji ima binarnu vrednost 1 (beli piksel) bude okružen sa i segmentiranih piksela u svom okruženju od 8 piksela:

$$P_{(1,i)} = \frac{\text{broj segmentiranih piksela okružen sa } i \text{ segmentiranih piksela}}{\text{ukupan broj segmentiranih piksela na slici}}. \quad 3.15$$

Ova mera verovatnoće se kod segmentiranih objekata koristi i kao karakteristika za kasniju klasifikaciju. Dvodimenzionalna entropija S može se posmatrati kao mera neuređenosti binarne segmentirane slike i s porastom 2D entropije rase i neuređenost (prekidi ili šum) binarne slike. Jasno je da se problem optimalno izbora parametara procesiranja slike, odnosno segmentacije u našem slučaju, svodi na minimizaciju veličine S .

Radi pojednostavljenja prikazane su sintetički generisane binarne slike jednog te istog segmentiranog objekta kod koga je segmentacija idealna (a), sa šumom (b) i sa prekidima (c). Nakon što se za ove slike odredi 2D entropija na osnovu jednačine 3.14., vidi se da je entropija najniža kod idealno segmentirane slike.



Slika 3.7 a) Idealna segmentacija, b) segmentacija sa šumom i c) segmentacija sa prekidima [22]

Segmentacija se vrši na osnovu slike nijansi jer je osnovna RGB slika prvo prebačena u HSV domen. Slika nijansi sadrži čistu informaciju o boji osnovne slike. Kako je reč o

segmentaciji po boji, svakom pikselu koji pripada odgovarajućem opsegu nijansi C_l dodeljuje se logička jedinica (beli pikseli), dok se pikselima izvan tog intervala dodeljuje logička nula (na binarnoj segmentiranoj slici predstavljeni su crnim pikselima). Problem se dakle svodi na određivanje intervala:

$$C_l \in [T_{low}, T_{high}], \quad 3.16.$$

kako bi kod segmentirane slike bio ispunjen kriterijum:

$$t_H(x, y) = \begin{cases} 1, & \text{if } f_H(x, y) \in C_l \\ 0, & \text{if } f_H(x, y) \notin C_l \end{cases} \quad 3.17.$$

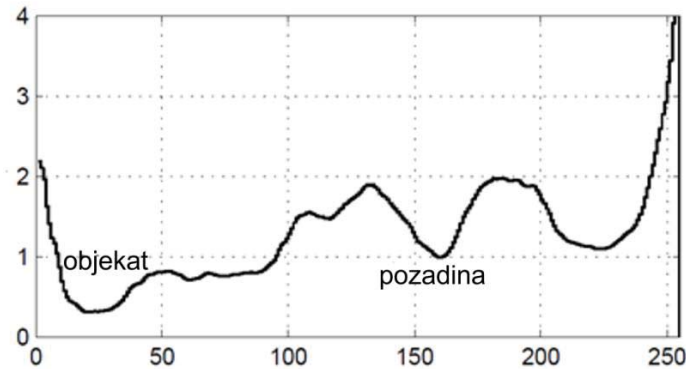
Kako bi se postigla odgovarajuća povezanost piksela uvodi se promenljiva i koja predstavlja inkrimet segmentacije, pa je interval praga segmentacije promenljiv, tj. $C_{li} \in [T_{min} + i, T_{max} + i]$. Za svaku vrednost promenljive i potrebno je izračunati meru entropije za 8 povezanih piksela, tj.:

$$I = -\log_2 p_8, I(0) = 0 \quad 3.18.$$

gde je p_8 verovatnoća javljanja segmentiranog piksela okruženog sa 8 segmentiranih piksela (u okolini od 8 piksela):

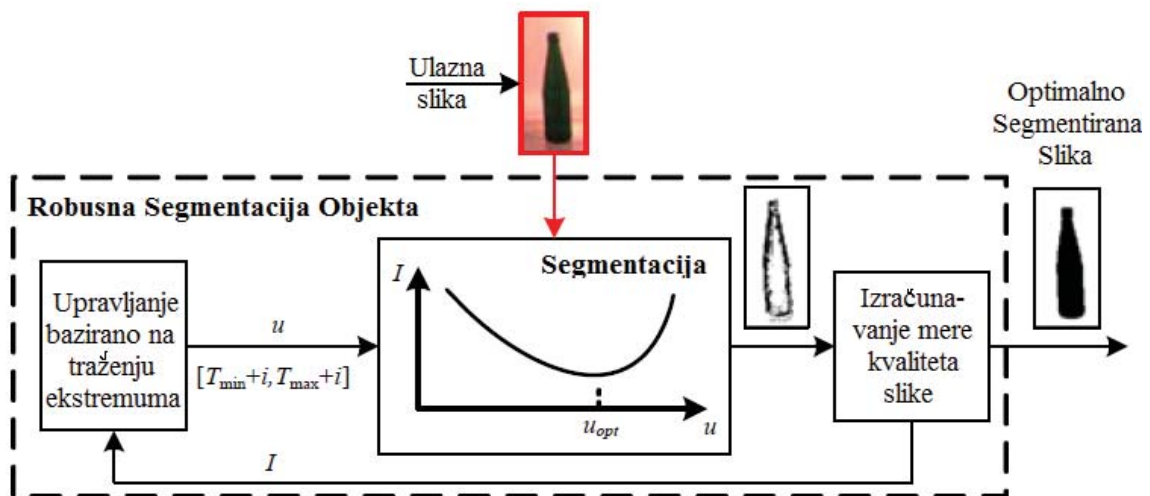
$$p_8 = \frac{\text{broj segmentiranih piksela okružen sa 8 segmentiranih piksela}}{\text{ukupan broj segmentiranih piksela}}. \quad 3.19.$$

Nakon toga se sa krive zavisnosti bira jedna ili više vrednosti inkrimeta i za koje kriva promene mere entropije I ima lokalni minimum (slika 3.8). U radovima [47, 48] je kao početni interval segmentacije uzimano da je $[T_{min}, T_{max}] = [0, 20]$, dok je interval promene inkrimeta bio $[0, 179]$ kod HSV slike, odnosno $[0, 255]$ kod crno-bele slike, ili svake boje pojedinačno kod RGB slike.



Slika 3.8 Kriva zavisnosti mere nesigurnosti I od inkrimenta segmentacije i kod crno-bele slike

Kako je ovaj algoritam baziran na ideji uvođenja povratne sprege koja je osnova teorije upravljanja, otuda i naziv segmentacije u zatvorenom kolu, tj. segmentacije s povratnom spregom. Šema algoritma segmentacije prikazana je na slici 3.9.



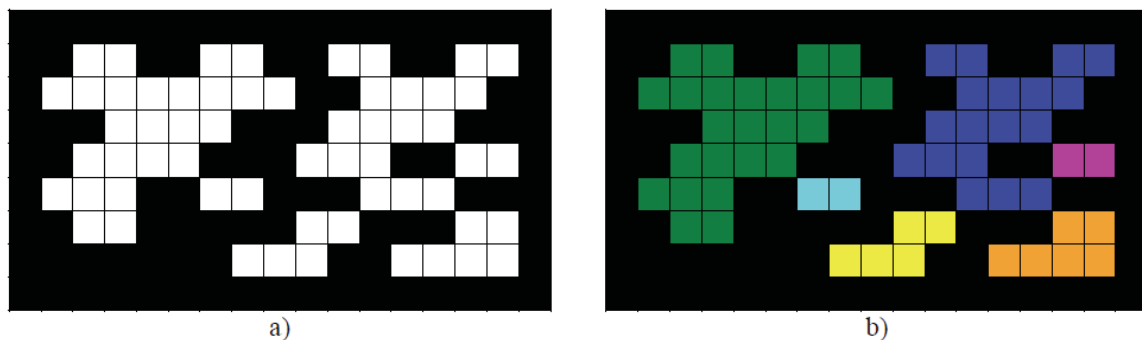
Slika 3.9 Šema algoritma segmentacije s povratnom spregom [46, 48]

Osim slike nijansi može se kao osnovni ili dodatni ulaz u algoritam segmentacije iskoristiti i slika intenziteta ili slika zasićenja. Ovaj ritup je prvi put primenjen u cilju prepoznavanja teksta [22].

3.4.3. Segmentacija mape dispariteta.

Dubinska slika može se koristiti kao dodatni izvor informacija pri segmentaciji objekata [55], pa se primenom CRF (engl. Conditional Random Fields) metode nakon uvođenja opsega praga dubine može vršiti segmentacija objekata u vizuelno kompleksnim okruženjima. Dubinska slika i mapa dispariteta su relativno slične slike, jer mapa dispariteta sadrži dubinsku informaciju za svaki piksel, pa se uvođenje praga ili intervala praga na dubinsku informaciju može smatrati osnovnom idejom segmentacije mape dispariteta. Moguće je i kombinovati segmentirane oblasti i segmentirane ivice dobijene segmentacijom dubinske slike [56] kako bi se pronašli objekti u sceni.

U odeljku 3.3 je objašnjeno kako se formira mapa dispariteta na osnovu stereo slike, ali kako bi se odredile karakteristike objekata i izvršila klasifikacija neophodno je najpre segmentirati mapu dispariteta, odnosno međusobno razdvojiti objekte i razdvojiti ih od pozadine. U ovoj tezi korišćena je nova metoda razvijena od strane Leu-a [25] koja se javlja i u nekim ranijim radovima grupe autora iz Bremena [47, 55]. Ideja je potekla od poznatog algoritma za obeležavanje odvojenih objekata na binarnoj slici, odnosno algoritma za detekciju regiona (engl. blob detection) binarne slike i može se primeniti kod povezanosti sa 4 susedna piksela ili povezanosti sa 8 susednih piksela. Kod povezanosti sa 4 susedna piksela posmatraju se pikseli samo iznad, ispod, levo i desno od piksela koji trenutno posmatramo. Kod povezanosti 8 susednih piksela posmatra se svih 8 piksela koji se nalaze u okolini trenutnog piksela. Na slici 3.10 prikazani su segmentirani regioni u slučaju da se posmatra povezanost 4 piksela. Da se razmatrala povezanost 8 susednih piksela, na slici bi bila odvojena samo dva objekta.



Slika 3.10 Detekcija regiona a) binarna slika i b) razdvojeni regioni [25]

Označavanje piksela se vrši u dva koraka. U prvom koraku se označavaju pikseli na sledeći način:

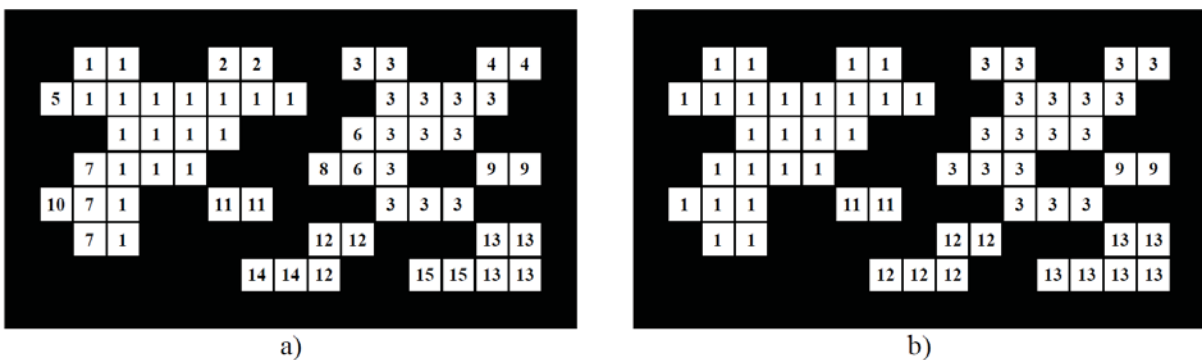
$$l_{u,v} = \begin{cases} l_{u-1,v}, & l_{u-1,v} \neq 0 \wedge l_{u,v-1} = 0 \\ l_{u,v-1}, & l_{u,v-1} \neq 0 \wedge l_{u-1,v} = 0 \\ \min(l_{u-1,v}, l_{u,v-1}), & l_{u-1,v} \neq 0 \wedge l_{u,v-1} \neq 0 \\ u \text{ suprotnom sledeća vrednost labele} \end{cases} \quad 3.20.$$

gde je $l_{u,v}$ oznaka trenutnog piksela, a $l_{u-1,v}$ i $l_{u,v-1}$ su oznake piksela levo i piksela iznad trenutno posmatranog piksela. Osim toga, tokom prvog koraka se formira lista ekvivalenata eq iz koje se može videti da dva susedna elementa imaju različite vrednosti oznake tako što se formiraju nove vrednosti liste za trenutne piksele kod kojih gornji susedni i levi susedni piksel imaju različite vrednosti oznake.

U drugom koraku se konačno formiraju oznake piksela na osnovu liste ekvivalenata:

$$l_{u,v} = \min(eq[l_{u,v}]), \quad u = \overline{0..imWidth - 1}, \quad v = \overline{0..imHeight - 1} \quad 3.21.$$

gde je $l_{u,v}$ oznaka trenutnog piksela sa koordinatama (u,v) , a $imWidth$ i $imHeight$ su širina i visina slike (u pikselima). Primer dva prolaza algoritma detekcije regiona prikazan je na slici 3.11.



Slika 3.11 Primer algoritma detekcije regiona na binarnoj slici a) prvi korak b) drugi korak

[25]

Kako se mape dispariteta mogu predstaviti kao crno-bele a ne binarne slike, bilo je potrebno modifikovati prethodni algoritam. Pretpostavka je da susedni pikseli jednog te istog objekta na mapi dispariteta imaju sličnu vrednost dispariteta, dok se na ivicama objekta

disparitet naglo menja, pa se uvodi prag dispariteta i modifikovana jednačina određivanja oznake piksela sada je:

$$l_{u,v} = \begin{cases} l_{u-1,v}, & p_{u,v} - p_{u-1,v} < Th \quad \wedge \quad l_{u-1,v} \neq 0 \wedge l_{u,v-1} = 0 \\ l_{u,v-1}, & p_{u,v} - p_{u,v-1} < Th \quad \wedge \quad l_{u-1,v} = 0 \wedge l_{u,v-1} \neq 0 \\ \min(l_{u-1,v}, l_{u,v-1}), & (p_{u-1,v} \wedge p_{u,v-1} \text{ su OK}) \wedge l_{u-1,v} \neq 0 \wedge l_{u,v-1} \neq 0 \\ & \text{u suprotnom sledeća vrednost labele} \end{cases} \quad 3.22.$$

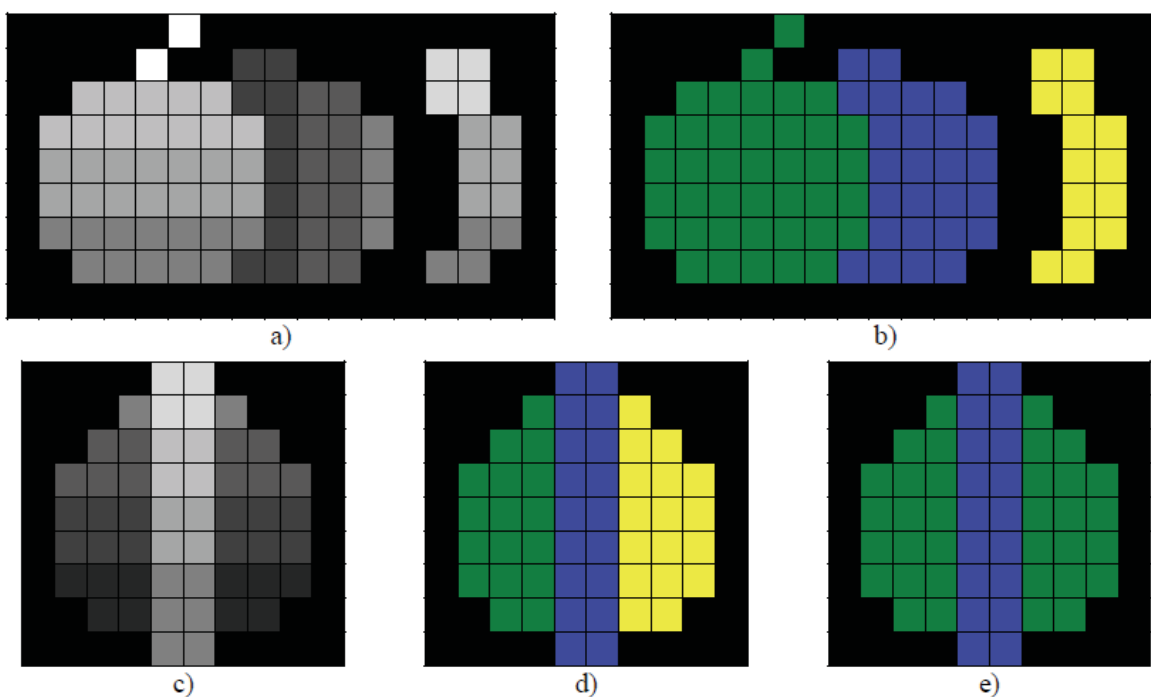
gde su $p_{u,v}$ vrednosti intenziteta piksela crno-bele slike sa koordinatama u, v , a izraz:

$$(p_{u-1,v} \wedge p_{u,v-1} \text{ su OK}) \wedge l_{u-1,v} \neq 0 \wedge l_{u,v-1} \neq 0 \quad 3.23.$$

zapravo je kraća forma izraza:

$$p_{u,v} - p_{u-1,v} < Th \quad \wedge \quad p_{u,v} - p_{u,v-1} < Th \quad \wedge \quad l_{u-1,v} \neq 0 \wedge l_{u,v-1} \neq 0 \quad 3.24.$$

Prag dispariteta Th se menja od piksela do piksela, zavisi od dispariteta trenutnog piksela i iznosi 10% od vrednosti dispariteta. Na slici 3.12 prikazana je segmentacija mape dispariteta sintetičke slike.



Slika 3.12 Segmentacija mape dispariteta [25]

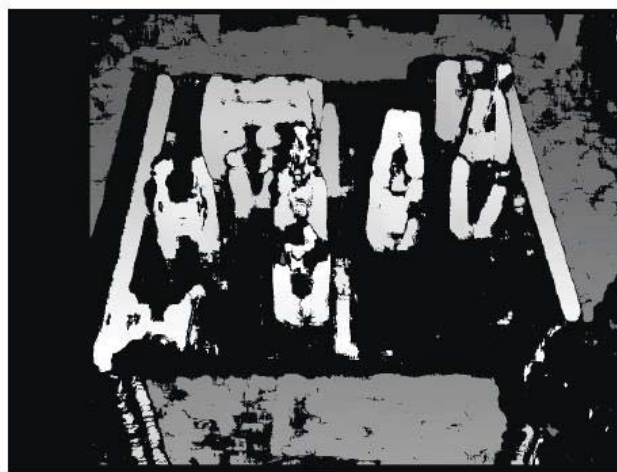
Osim osnovnog algoritma ubačen je i jedan dodatni korak post-procesiranja, da se prostorno bliski objekti sa istom ili približno istom vrednošću dispariteta grupišu u jedan region (Slika 3.12 e), čime se postiže uspešna segmentacija objekata na "pretrpanoj" sceni, gde se objekti međusobno zaklanjaju.

Na slici 3.13 prikazana je segmentacija mape dispariteta slike scene koja se može sresti u servisnoj robotici.

Kako bi se dodatno poboljšala dubinska segmentacija, odnosno dodali segmenti koji nedostaju, dodat je još jedan element post-procesiranja segmentirane slike, konveksno popunjavanje okvira (engl. convex hull fitting), prikazano na slici 3.14. Pri formiranju mape dispariteta često nedostaju elementi jednobojnih objekata ili objekata koji nemaju izraženu teksturu, pa otuda praznine u segmentiranim objektima sa mape dispariteta. Eksperimenti pokazuju da se i pored toga na segmentiranoj slici ivice objekta jasno vide, pa je popunjavanje objekta kako bi se formirao konveksni poligon (objekat) deluje kao logičan korak.



b)



d)



f)

Slika 3.13 Segmentacija test scene [25]



Slika 3.14 Segmentacija mape dispariteta, konveksno popunjavanje okvira i ručno segmentirani objekti[47]

Algoritam pronalazi tačke na segmentiranom objektu koje kada se povežu pravim linijama formiraju konveksan poligon. Na ovaj način su segmentirani objekti mnogo približniji ručno segmentiranim objektima. Konveksni okvir koji treba popuniti može se definisati kao skup tačaka međusobno povezanih pravim linijama koje ispunjavaju sledeći kriterijum:

$$\{\sum_{i=1}^{|S|} \alpha_i x_i \mid (\forall i: \alpha_i \geq 0) \wedge \sum_{i=1}^{|S|} \alpha_i = 1\}, \quad 3.25.$$

gde je S broj piksela ivice segmentiranog objekta, x_i je koordinata i -tog piksela ivice segmentiranog objekta, a α_i pozitivan broj.

Ova metoda segmentacije dubinske slike može se koristiti za definisanje regiona od interesa koji je mnogo bolje definisan nego li pravougaonim graničnim okvirom. Može se videti da nakon segmentacije mape dispariteta i konveksnog popunjavanja dolazi do efekta „zadebljanja“ objekata, kao posledica „block matching“ metode formiranja mape dispariteta. Naime, segmentirana oblast sadrži osim samog objekta i deo pozadine, pa se u radu [47] u ovako definisanom region od interesa dodatno vrši segmentacija po boji u zatvorenoj sprezi kako bi se omogućila adekvatna 3D rekonstrukcija objekta. Osim toga, prepoznavanje objekata se može vršiti direktno iz segmentiranih slika mape dispariteta.

3.5. Određivanje karakteristika segmentirane slike

Nakon uspešne segmentacije, u cilju prepoznavanja objekata potrebno je razviti klasifikator koji bi određivao pripadnost nekog segmentiranog regiona odgovarajućoj klasi. Najzastupljeniji klasifikatori koji služe za prepoznavanje objekata ne klasifikuju objekte direktno iz segmentirane slike, već koriste određene karakteristike segmentiranog objekta. Određivanje, odnosno ekstrakcija karakteristika slike je među-operacija, tj. spona između procesiranja slika niskog nivoa i procesiranja slika visokog nivoa [56,57,58]. Ove karakteristike treba da opišu međusobnu zavisnost povezanih segmentiranih piksela. Najčešće korišćeni su zapravo karakteristike koje opisuju oblik objekta, poput proporcionalnosti regiona, povezanosti piksela, Hu-ovih invarijantnih momenata oblasti i sl [1,2,3].

Proporcionalnost regiona međusobno povezanih piksela predstavlja odnos visine h i širine regiona w koji obuhvata segmentirane piksele:

$$P = h \cdot w^{-1} \quad 3.26.$$

Još jedna karakteristika, o kojoj je već bilo reči u odeljku 3.4.2 jeste konektivnost, odnosno prostorna povezanost piksela. Kako je ova mera bila u toku istraživanja lako dostupna jer se koristila i u segmentaciji slike, povezanost sa 8 susednih piksela uzeta je kao jedna od karakteristika koja je ulaz u klasifikator za prepoznavanje objekata. Prostorna povezanost piksela segmentiranog objekta i njegovih 8 susednih piksela može da se izrazi na sledeći način:

$$I = -\log_2 p_8, I(0) = 0 \quad 3.27.$$

gde je p_8 verovatnoća segmentiranog piksela, okruženog sa 8 segmentiranih piksela u okolini od 8 piksela:

$$p_8 = \frac{\text{broj segmentiranih piksela okružen sa 8 segmentiranih piksela}}{\text{ukupan broj segmentiranih piksela}} \quad 3.28.$$

Iz jednačine 3.27. se vidi da za malu verovatnoću p_8 , koja odražava razbacane, nepovezane segmentirane piksele, dobija se velika vrednost za I . Sa druge strane, ako je p_8 blizu 1, dobija se mala vrednost za I što znači da je na segmentiranoj slici region povezanih piksela.

Hu-ovi momenti su invarijantni koeficijenti određuju se pomoću momenata segmentiranog regiona. Kod digitalne slike funkcija intenziteta $f(x, y)$, momenta reda $(p + q)$ je:

$$m_{pq} = \sum_x \sum_y x^p y^q f(x, y) \quad 3.29.$$

gde su x i y koordinate piksela razmatranog regiona slike. Centralni momenti μ_{pq} su:

$$\mu_{pq} = \sum_x \sum_y (x - \bar{x})^p (y - \bar{y})^q f(x, y) \quad 3.30.$$

gde je $p, q = 1, 2, 3, \dots$, $\bar{x} = m_{10} / m_{00}$, $\bar{y} = m_{01} / m_{00}$. Kod binarne slike dobijene nakon segmentacije je:

$$f(x, y) = \begin{cases} 1, & \text{if } f(x, y) \in C \\ 0, & \text{if } f(x, y) \notin C \end{cases} \quad 3.31.$$

gde je C subinterval gornjeg i donjeg graničnog praga $C \in [T_{low}, T_{high}]$.

Hu-ovi momenti su:

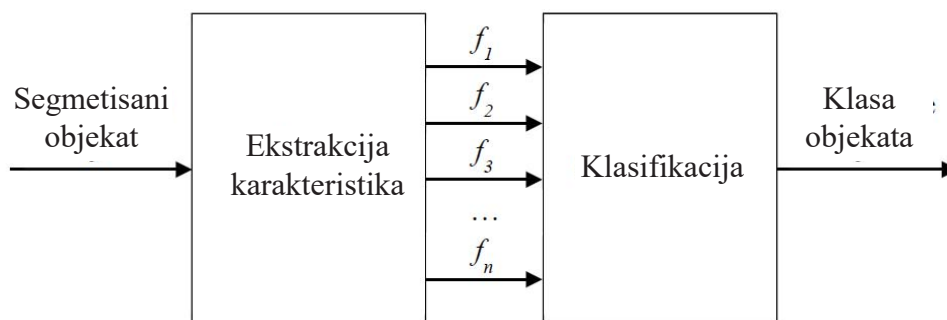
$$\begin{aligned} I_1 &= \eta_{20} + \eta_{02} \\ I_2 &= (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2 \\ I_3 &= (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} - \eta_{03})^2 \\ I_4 &= (\eta_{30} - \eta_{12})^2 + (\eta_{21} - \eta_{03})^2 \\ I_5 &= (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} - \eta_{03})^2] \\ &\quad + (3\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \\ I_6 &= (\eta_{20} - \eta_{02})[(\eta_{30} + \eta_{12})^2 - (\eta_{21} - \eta_{03})^2] + 4\eta_{11}(\eta_{30} - \eta_{12})(\eta_{21} - \eta_{03}) \\ I_7 &= (3\eta_{21} - \eta_{03})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} - \eta_{03})^2] \\ &\quad - (\eta_{30} - 3\eta_{12})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \\ I_8 &= \eta_{11}[(\eta_{30} + \eta_{12})^2 - (\eta_{03} - \eta_{21})^2] - (\eta_{20} - \eta_{02})(\eta_{30} + \eta_{12})(\eta_{03} + \eta_{21}) \end{aligned} \quad 3.32.$$

gde je η_{pq} normalizovani centralni moment:

$$\eta_{pq} = \mu_{pq} \cdot \mu_{00}^{-\frac{p+q}{2}} \quad 3.33.$$

3.6. Klasifikacija i prepoznavanje objekata na osnovu karakteristika

Konačni ishod obrade slike jeste klasifikacija ekstrahovanih karakteristika, odnosno prepoznavanje objekata. Vektoru ekstrahovanih karakteristika jednog objekta sa slike se primenom jednog ili više alata iz domena mašinskog učenja dodeljuje oznaka neke određene klase, a algoritam koji vrši ovo dodeljivanje oznake naziva se klasifikator. Na slici 3.15 je prikazan princip klasifikacije objekata na osnovu ekstrahovanih karakteristika.



Slika 3.15 Blok dijagram klasifikacije objekta na osnovu karakteristika

Klasifikacija se može, kao i mašinsko učenje uopšte, vršiti nadgledano i nenadgledano. Kod nadgledanog učenja se obučavanje vrši na osnovu seta vektora kojima su unapred dodeljene oznake klase, pa se tek tako obučeni klasifikator koristi za klasifikaciju novih neoznačenih vektora karakteristika. Kod nenadgledanog učenja klasifikator sam formira klase na osnovu sličnosti vektora karakteristika.

Kako je već pomenuto, veliki je broj tehnika mašinskog učenja koje se mogu primeniti u klasifikaciji segmentiranih objekata, od relativno jednostavnih poput klasifikatora najmanjeg rastojanja ili naivnog Bajesovog klasifikatora [17], preko nešto kompleksnijih poput k-najbližeg suseda, algoritma adaptivnog podsticaja (engl. adaboost) [59] i metode potpunih vektora (SVM) [1,10, 19] do sofisticiranih neuronskih mreža [2,5,6,20] i neuro-fazi sistema [3,7,9,18]. Javlja se i ideja kaskadno postavljenih klasifikatora [59], gde prvo jednostavni

klasifikatori odbacuju elemente koji pripadaju pozadini, a kasnije komplikovaniji klasifikatori određuju kojoj klasi određeni segmentirani objekat pripada. Osim toga, u koliko želimo koristiti npr. SVM klasifikator za klasifikaciju gde postoji više od dve klase objekata, neophodno je kaskadno postaviti SVM klasifikatore [25].

U ovom radu su u narednom poglavlju primenjeni različiti klasifikatori i upoređene su njihove performanse na zadatom eksperimentalnom skupu koji predstavlja slike iz svakodnevnog okruženja robota za pomoć osobama sa invaliditetom.

Klasifikacija objekata u ovom radu ima trostruku ulogu. Kod eksperimentalne robotske platforme sa robotskom rukom koja je detaljno opisana u poglavlju 4 ove teze klasifikacija objekata omogućava izvršenje zadataka koji se pred robota postave. S druge strane, informacija koju dobijemo od klasifikatora može se iskoristiti kao još jedna korisna informacija pri 3D rekonstrukciji objekta u 3D virtuelnom prostoru. Konačno, rekonstrukcija objekta u 3D virtuelnom prostoru i informacija o vrsti objekta omogućava adekvatno hvatanje od strane robotske ruke.

4. Inteligentni sistemi za segmentaciju slike i prepoznavanje objekata

Sistem kompjuterske vizije, odnosno robotske vizije kao osnovne elemente ima akviziciju slike, segmentaciju i razumevanje slike, tj. u ovom slučaju prepoznavanje objekata. Osnovni algoritmi za segmentaciju i klasifikaciju su već su predstavljani u prethodnom poglavlju, ali je primena alata iz domena veštačke inteligencije i mašinskog učenja samo spomenuta.

I pored toga što su mnogi alati veštačke inteligencije i mašinskog učenja isti, mašinsko učenje i veštačka inteligencija nisu sinonimi. Mašinsko učenje je podoblast veštačke inteligencije i proisteklo je iz proučavanja prepoznavanja obrazaca i teorije računskog učenja, te korišćenja statistike u cilju učenja na osnovu prethodno dostupnih podataka. Izuzetno bitna distinkcija između mašinskog učenja i veštačke inteligencije nalazi se u cilju njihovog operisanja: dok veštačka inteligencija ima za cilj ne samo da imitira ljudsko razmišljanje kroz učenje, već i da to bude prožeto apstraktnim razmišljanjem, predstavljanjem znanja i rasuđivanjem, mašinsko učenje je samo usmereno ka stvaranju sistema koji može da uči iz prošlih iskustava. Mašinsko učenje je znatno bliže data mining-u i statističkoj analizi.

Osim pojmova veštačke inteligencije i mašinskog učenja u literaturi se često mogu sresti i pojmovi računarska inteligencija i meki račun (engl. soft computing). Računarska inteligencija podrazumeva alate koji omogućavaju računaru inteligentno učenje i ponašanje, dok meki račun podrazumeva alate koji omogućavaju analizu i obradu nepotpunih, nepouzdatih, nepreciznih ili aproksimiranih podataka, a koji su ponovo inspirisani ljudskim razmišljanjem. Ponovo su u pitanju isti alati kao i kod veštačke inteligencije i mašinskog učenja, poput neuronskih mreža, fazi logike, metaheurističkih metoda optimizacije, metode potpornih vektora, Bajesove mreže i sl.

Alati koji se najčešće koriste pri segmentaciji jesu razni optimizacioni algoritmi. Kod svih dosad spomenutih metoda segmentacije problem se svodio na optimalan izbor praga ili intervala praga. Metaheuristički metodi, poput metoda evolutivnog računarstva se mogu

koristiti kao alati optimizacije praga. Evolutivno računarstvo podrazumeva dve grupe optimizacionih algoritama:

- algoritmi zasnovani na inteligenciji roja (engl. swarm intelligence algorithms), poput optimizacionog algoritma inspirisanog kolonijom mrava (engl. ant colony optimization - ACO), rojem pčela (engl. bee colony optimization - BCO), traganjem bakterija za hranom (engl. bacterial foraging optimization - BFO) ili gomilom čestica (engl. particle swarm optimization - PSO);
- algoritmi zasnovani na evolutivnoj genetici, poput genetskih algoritama i genetskog programiranja.

Tako Abbasgholipour et al koriste genetski algoritam za određivanje intervala pragova H, S i V komponente slike kako bi uspešno izvršili segmentaciju slike suvog grožđa u mašini za razvrstavanje ovog prehrambenog proizvoda [60]. Moguće je koristiti algoritme zasnovane na inteligenciji roja i za pretragu, odnosno nalaženje optimalne putanje, pa se tako Mullen et al. u svom istraživanju [61] koriste optimizacijom inspirisanom kolonijom mrava u cilju nalaženja ivica digitalne slike, gde se agenti, odnosno "mravi", kreću po digitalnoj slici u potrazi za ivicama odnosno promenom gradijenta i za sobom ostavljaju feromone. Mapa feromona u ovom slučaju može se iskoristiti kao segmentirana slika.

U cilju uspešne segmentacije digitalne slike kod koje je teško i zahtevno ručno definisati regione od interesa i segmentisati sliku, bilo je i pokušaja polu-nadgledanog određivanja klastera [62], gde su samo neke informacije o konačno segmentiranoj slici bile ručno unete. Ovakav pristup omogućuje segmentaciju i definisanje oblasti ljudskog mozga na osnovu magnetno-rezonantne slike mozga.

Kako je kod obrade slike jako teško čak i ručno jasno definisati granice objekta, fazi logika se nameće kao jedan od logičnih izbora pri segmentaciji. Izlaz iz modula za segmentaciju ipak mora biti binarna slika sa jasno definisanim ivicama, ali odlučivanje o tome da li neki piksel, ivica ili region pripada objektu ili ne može vršiti fazi logika. Tako Zhao et al. u svom istraživanju [63] predlažu fazi klasterovanje pri segmentaciji, čime se znatno popravljaju rezultati pri segmentaciji slika koje imaju puno šuma. Ovo je bilo moguće zahvaljujući adaptivnom algoritmu podešavanja parametara fazi klasterovanja koji je razvijen od strane ovih istraživača. Kombinacija inteligentne optimizacije i fazi segmentacije istraživali su Tao et al. u radu [64] gde su vršili segmentaciju fazifikacijom entropije slike, dok su kao optimizacioni algoritam koristili ACO.

Prilikom klasifikacije objekata na osnovu ekstrahovanih karakteristika, a prvenstveno zbog nelinearnosti i nepouzdanosti pri određivanju karakteristika segmentiranog objekta, neophodno je koristiti neku od metoda klasifikacije iz domena računarske inteligencije. Tako su Viola i Jones [59] koristili kaskadno postavljene AdaBoost klasifikatore, koji su detektovali lica na slici klasifikujući jednostavne karakteristika slike, poput karakteristike pravougaonog okvira oko očiju na crno-belaj slici lica gde su oči tamnije od obraza i karakteristike da su oči tamnije od nosa. Ovakav algoritam koji omogućava detekciju lica na slici danas je u širokoj upotrebi kako kod softvera za digitalne foto aparate, tako i kod vizuelnih sistema za nadzor, bezbednost i praćenje.

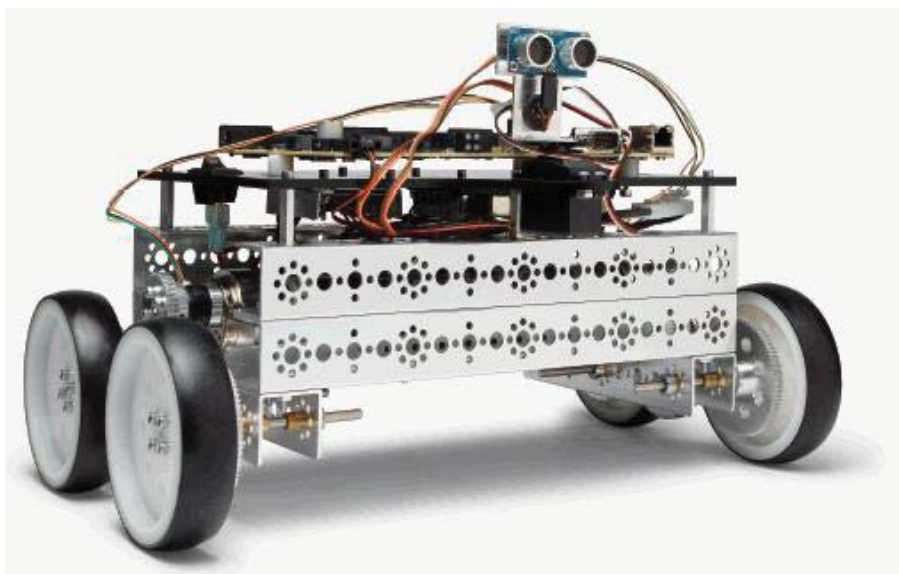
4.1. Inteligentno prepoznavanje i praćenje objekata eksperimentalnom robotskom platformom sa monokularnom kamerom

Za razvoj pouzdanog, računski inteligentnog prepoznavanja objekta u robotskoj viziji prilikom istraživanja [3,13,20] korišćena je eksperimentalna robotska platforma bazirana na robotu National Instruments Robotics Starter Kit 1.0 i veb kameri - "DaNI-C sistem".

Sistem vizije je koristi relativno jednostavnu segmentaciju po boji uvođenjem praga, koji se pokazao dovoljno robustan za primenu u laboratorijskim uslovima. Zbog nelinearnosti problema razvijani su klasifikatori zasnovani na metodama iz domena računarske inteligencije, neuro klasifikator i neuro-fazi klasifikator. U ovoj tezi je predstavljen samo neuro-fazi klasifikator jer je razvijena fazi struktura implementirana u upravljački algoritam robota i bilo je moguće testirati sistem u realnom okruženju i realnom vremenu.

4.1.1. Eksperimentalna robotska platforma DaNI sa monokularnom kamerom

Pokretna robotska platforma NI Robotics Starter Kit 1.0, poznat kao DaNI robot, (Slika 4.1) poseduje senzore, motore, i NI Single-Board RIO ploču za kontrolu. Korišćeni mobilni robot poseduje četiri točka koji se pogone sa dva motora, i opremljen je ultrazvučnim sensorom za merenje daljine.



Slika 4.1 DaNI robot.

Tokom istraživanja razvijen je sistem sa robotskom vizijom za prepoznavanje objekata, gde monokularna kamera omogućuje vizualne informacije o okolini robota, pa je zbog toga njen izbor od izuzetnog značaja za ceo sistem. Monokularna veb kamera je povezana sa laptopom, koji preko Lan kabla komunicira sa robotom (Slika 4.2).

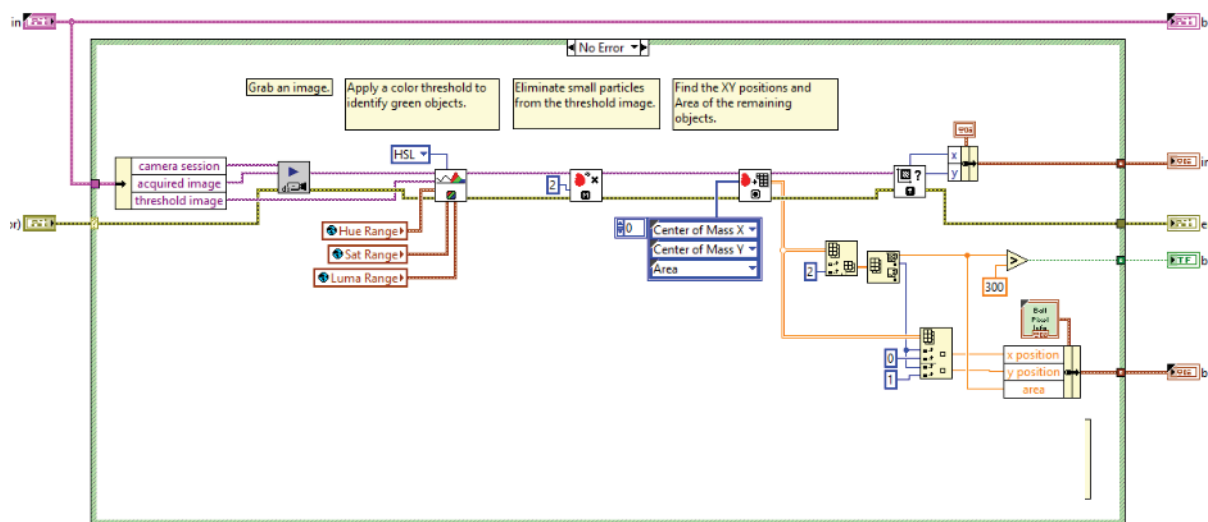


Slika 4.2 Mobilna robotska platforma sa senzorom robotske vizije.

Sva softverska rešenja za obradu slike, klasifikaciju i upravljanje mobilnom robotskom platformom sa robotskom vizijom realizovana su u softverskom paketu LabVIEW 2011 kompanije National Instruments. Osim toga, tokom istraživanja korišćena je i ostala dostupna oprema iz Laboratorije za upravljanje sistemima (LUPS) Mašinskog fakulteta Univerziteta u Nišu.

4.1.2. Segmentacija slike dobijene monokularnom kamerom robotske platforme

Kako bi se dobili valjani podaci za dalju obradu, klasifikaciju i konačno upravljanje mobilnom robotskom platformom, bilo je neophodno prvo razviti softver za obradu i segmentaciju slika dobijenih senzorom robotske vizije. Segmentacija objekata iz pozadine vrši se segmentacijom po boji oblasti u takozvanim Hjuovim slikama, koje sadrže informacije o originalnim RGB bojama (crvena, zelena, plava). Najpre je bilo potrebno prebaciti digitalnu sliku iz RGB domena u HSV domen boja, odnosno HSL (Hue, Saturation, Luma) sliku. Nakon toga se primenom opsega praga sa slike izdvajaju objekti obojeni jednom određenom bojom. U inicijalnoj fazi bilo je potrebno da softver detektuje jednobojni objekat, eliminiše male segmentirane elemente i odredi centar mase segmentiranog objekta kako bi robot mogao da prati objekat. Softver za segmentaciju slike koji je razvijen u programskom paketu LabView prikazan je na slici 4.3.



Slika 4.3. Program za segmentaciju slike dobijene senzorom robotske vizije

U cilju postizanja dobre segmentacije objekata neophodno je bilo podesiti intervale segmentacije objekta kako bi sistem vizije bio robustan, jer se menjaju npr. uslovi osvetljenja. Pronalaženje optimalnog intervala segmentacije koji obezbeđuje visok kvalitet slike segmentiranog objekta koji će biti odgovarajuća za određivanje karakteristika objekta i njegovo prepoznavanje vršen je ručno, jer se istraživanje odvijalo u kontrolisanim uslovima.

Kasnije modifikacije vršene su u cilju prepoznavanja određenog objekta koji je bilo potrebno pratiti. Rezultati segmentacije predstavljaju onoliki broj binarnih slika koliko ima objekata u sceni. Kako bi međusobno razdvojili segmentirane objekte sa binarne slike primenjen je algoritam detekcije regiona binarne slike koji je već opisan u odeljku 3.4.3. Nakon toga se pristupa ekstrakciji karakteristika i klasifikaciji objekata.

4.1.3. Klasifikacija i prepoznavanje objekata

Nakon segmentacije moguće je odrediti različite karakteristike objekata, odnosno regiona sa binarne slike, kao što je pokazano u odeljku 3.5. prethodnog poglavlja. Izbor karakteristika segmentiranog objekta koje su ulaz u klasifikator zavisi od specifičnosti scenarija i same aplikacije. Kako je predstavljeni sistem razvijan za prepoznavanje malog broja klasa objekata, bilo je dovoljno kao karakteristike objekata koristiti boju objekta, proporcionalnost i konektivnost piksela, koji su već predstavljeni u odeljku 3.5 prethodnog poglavlja. Boja je kao karakteristika već iskorišćena kao kriterijum segmentacije, pa se za klasifikaciju kao ulazni vektor koriste samo dve promenljive - proporcionalnost i konektivnost.

U ovom istraživanju je testirana efektivnost mobilnog robotskog sistema za radni scenario kod koga mobilni robotski sistem da prepozna je jednog od dva Lego NXT robota, koji su sklopljeni u različite oblike, i prati ga. Jedan NXT robot je sklopljen kao robogator, a drugi kao šuterbot (Slika. 4.4).



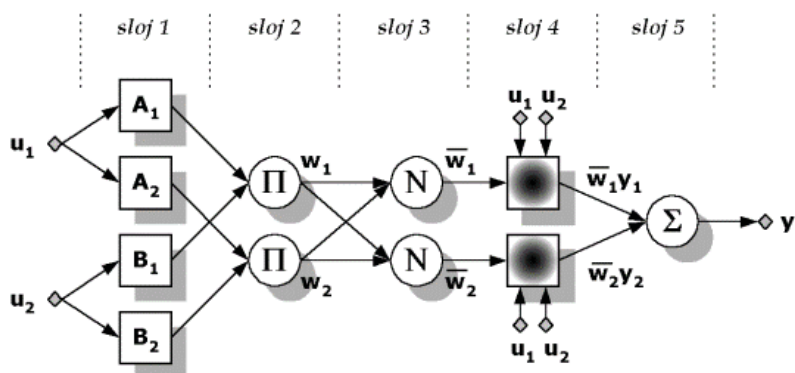
Slika 4.4 NXT roboti robogator i šuterbot.

Da bi omogućili mobilnom robotu da ispuni ove zadatke, sistem robotske vizije mora da bude sposoban da prepozna robogatora i šuterbota među drugim predmetima, i da odredi njihove lokacije sa dovoljnom preciznošću. Oba Lego robota su pretežno belo - svetlo sivi, pa se ove boje koriste pri određivanju granica praga za segmentaciju.

U koliko kamera frontalno snima nepokretne objekte koji su različitih oblika i dimenzija dovoljno je koristiti samo proporcionalnost kao kriterijum klasifikacije, jer je i pored šuma koji se javlja na slici i greški koju sa sobom povlači segmentacija (zbog senki i promene osvetljenja) jasno uočljiva razlika između objekata i dovoljno je samo uvesti odgovarajući prag ili opseg praga. Kako je zadatak bio praćenje pokretnih objekata, ti objekti tokom kretanja menjaju orijentaciju u odnosu na kameru, pa samim tim njihova projekcija na digitalnoj slici menja oblik. Zbog toga je bilo neophodno uvesti još jednu ulaznu promenljivu - povezanost i projektovati nelinearni klasifikator - neuro-fazi (engl. Adaptive Neuro-Fuzzy Inference System, ANFIS) klasifikator [65-70].

Neuronske mreže su često korišćene kao alternativa statističkim metodama klasifikacije obrazaca jer zahtevaju manje vremena proračuna i ne zahtevaju postojanje velikih baza podataka. Neuro-fazi struktura može da se posmatra kao neuronska mreža sa radijalnim bazičnim funkcijama, ali predstavlja hibridni algoritam i fazi logike i veštačkih neuronskih mreža čiji hibridni algoritam učenja omogućava poboljšane rezultate obuke u odnosu na druge slične metode. Ovu strukturu karakteriše neosetljivost na nepouzdanosti u sistemu obuke i može se izboriti sa velikim brojem ulaznih parametara [71].

U konkretnom slučaju korišćena je ANFIS struktura sa pet slojeva (slika 4.5), gde su elementi istog sloja u ANFIS mreži su iz istih funkcionalnih familija. Oznaka O_i^j obeležava izlaz i -tog elementa u sloju j .



Slika 4.5 Arhitektura ANFIS klasifikatora [71]

Oznaka O_i^j obeležava izlaz i -tog elementa u sloju j .

Sloj 1: Svaki element u ovom sloju generiše stepen pripadnosti za fazi skup koji je pridružen lingvističkoj labeli. Na primer, funkcija i -tog elementa može biti:

$$O_i^1 = \mu_{A_i}(u_1) = \max \left[\min \left(\frac{u_1 - a}{b - a}, 1, \frac{d - u_1}{d - c} \right), 0 \right], \quad 4.1.$$

gde je u_1 ulaz u elementat i , A_i je lingvistička labela (veliko, srednje, malo) pridružena tom elementu (odnosno odgovarajućem fazi skupu) a $\{a, b, c, d\}$ je parametar set koji opisuje funkciju pripadnosti, koja je u ovom slučaju definisana kao trapezna. Parametri u ovom sloju nazivaju se uzročni parametri, jer se nalaze u uzročnim delovima odgovarajućih pravila.

Sloj 2: Svaki elementat u ovom sloju računa nivo aktivacije odnosno jačinu okidanja pravila putem množenja:

$$O_i^2 = w_i = \mu_{A_i}(u_1) \times \mu_{B_i}(u_2), \quad i = 1, 2 \quad 4.2.$$

Sloj 3: Elementat i u ovom sloju računa odnos nivoa aktivacije i -tog pravila (jačine okidanja) prema sumi nivoa aktivacija svih pravila:

$$O_i^3 = \bar{w}_i = \frac{w_i}{w_1 + w_2}, \quad i = 1, 2 \quad 4.3.$$

Sloj 4: Elementat i u ovom sloju ima sledeću funkciju:

$$O_i^4 = \bar{w}_i y_i = \bar{w}_i (p_i u_1 + q_i u_2 + r_i), \quad 4.4.$$

gde je w_i izlaz iz sloja 3, a $\{p_i, q_i, r_i\}$ je parametar set. Parametri u ovom sloju označavaju se imenom posledični parametri, jer oni figurišu u posledičnim delovima odgovarajućih fazi pravila.

Sloj 5: Jedini elementat u ovom sloju računa ukupni izlaz ANFIS strukture kao zbir svih dolaznih signala:

$$O_i^3 = \text{ukupni izlaz} = \sum_i \bar{w}_i y_i = \frac{\sum w_i y_i}{\sum w_i}. \quad 4.5.$$

Hibridni algoritam obučavanja ANFIS sistema sastoji se od dva sukcesivna dela, metoda povratne propagacije greške i metoda najmanjih kvadrata. Metod povratne propagacije, odnosno najstrmijeg spusta (BP/GD) proračunava signale greške (definisana kao izvod kvadratne greške u odnosu na izlaz svakog elementa) rekurzivno od izlaznog sloja unazad ka ulaznim elementima. Rekurzivnim metodom najmanjih kvadrata (RLSE) se određuje optimalni set posledičnih parametara. Sa fiksiranim vrednostima uzročnih parametara izlaz se može izraziti kao linearna kombinacija posledičnih parametara:

$$y = \bar{w}_1 y_1 + \bar{w}_2 y_2 = (\bar{w}_1 u_1) p_1 + (\bar{w}_1 u_2) q_1 + (\bar{w}_2 u_1) p_2 + (\bar{w}_2 u_2) q_2 + (\bar{w}_1) r_1 + (\bar{w}_2) r_2. \quad 4.6.$$

Početna struktura klasifikatora sa 14 pravila, i 12 do 14 primarnih fazi skupova u prostoru ulaza i izlaza, izabrana je kao optimalna. Ovo je urađeno sa 464 parova za treniranje. Isti skup parova se koristi za treniranje klasifikatora, tj. za podešavanje parametara. Obučavanje klasifikatora je vršeno za manuelno označene segmentirane regione različitih objekata koji pripadaju dvema klasama, "Robogator" i "Shooterbot", a konačni izlaz treniranja je formirana fazi struktura. Testiranje razvijenog klasifikatora izvršeno je korišćenjem deskriptora drugih 232 parova objekata. Obučavanje i testiranje klasifikatora izvršeno je u "off line" režimu u softveru koji je razvijen u laboratoriji LUPS, a kasnije je fazi struktura prebačena u LabView. Različiti uslovi osvetljenja, kao i pogledi kamere izazivaju razlike u vrednostima deskriptora jednog istog objekta.

Testiranje je takođe urađeno i u eksperimentu sa mobilnom robotskom platformom u njenom realnom okruženju. Eksperiment je rađen za različite uslove osvetljenja, u vrednostima odgovarajućih osvetljenju u tamnoj sobi (15lx) do vrednosti osvetljenja u kancelariji prema Evropskom UNI EN 12464 (500lx). Dobijeni rezultati klasifikacije objekata su vrlo dobri, s' obzirom da je tačnost oko 95%. Greške u klasifikovanju objekta su se javile samo onda kada je objekat bio značajno zaklonjen, što je izazvalo netačne vrednosti za karakteristiku povezanosti segmentiranog objekta.

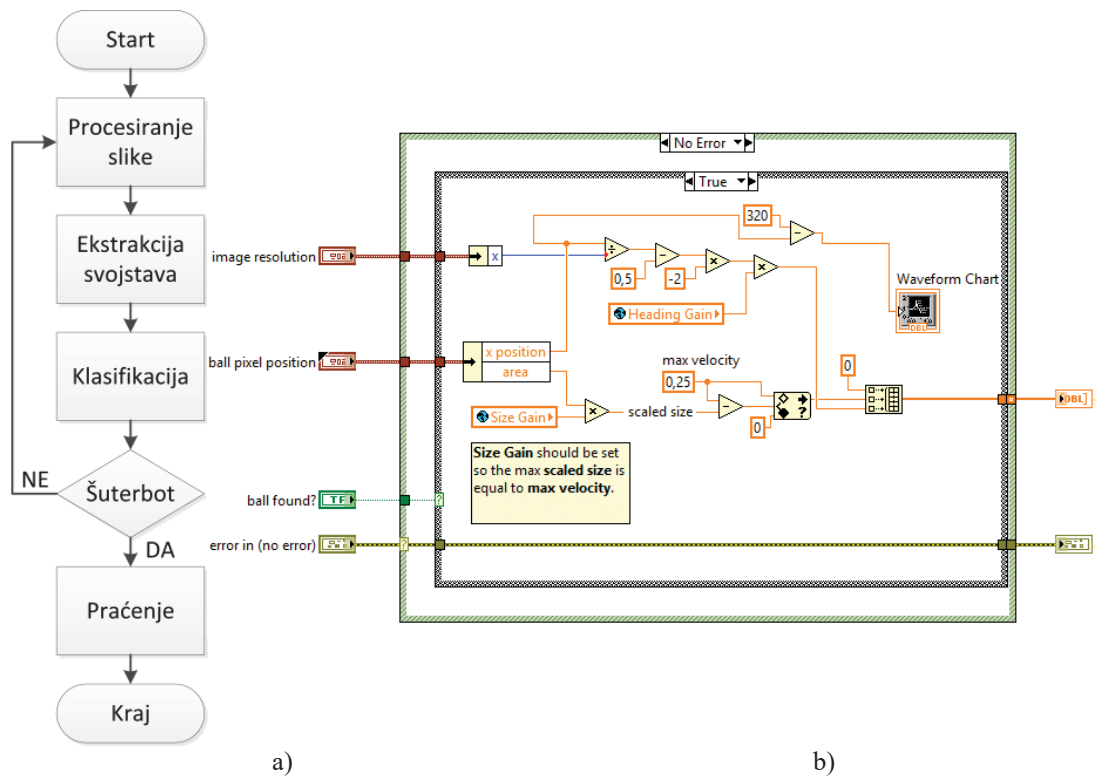
4.1.4. Praćenje objekata u sceni

Cilj praćenja objekta u robotskoj viziji jeste detekcija objekta u svakoj slici (frejmu) video snimka koji pravi senzor vizije. Ovo nam zapravo omogućava potprogram za prepoznavanje objekata koji se izvršava pri svakom novom frejmu, pa se time iz frejma u frejm prati položaj objekta i donose odgovarajuće upravljačke odluke.

Mobilna robotska platforma prepoznaje objekte i prati objekat iz prve klase na osnovu algoritma prikazanog na slici 4.6 a). Nakon pokretanja programa, prvo se vrši obrada slike. Robot dobija sliku iz svog okruženja i onda se vrši obrada i segmentacija slike po prethodno opisanom algoritmu. Iz segmentirane slike program izdvaja karakteristike objekata. Korišćenjem tih karakteristika klasifikator prepoznaje objekte na slici. Ako na slici ne postoji objekat prve klase, zahteva se nova slika, i pomenuti koraci se ponavljaju. Ako klasifikator prepozna objekat prve klase, onda mobilni robot počinje da ga prati, tj. upravljački sistem donosi zaključke o pokretanju motora koji pogone točkove DaNI robota. Postupak se ponavlja iz frejma u frejm.

Na slici 4.6 b) predstavljen je potprogram razvijen u LabView-u za praćenje koordinata detektovanog objekta. Na osnovu predstavljenih algoritama u programskom paketu LabView razvijen je softver koji omogućava inteligentno upravljanje mobilnom robotskom platformom.

Mobilna robotska platforma u akciji predstavljena je na slici 4.7. Razvijeni softver objedinjuje module za obradu slike, prepoznavanje i praćenje objekata i upravljanje kretanjem robota.



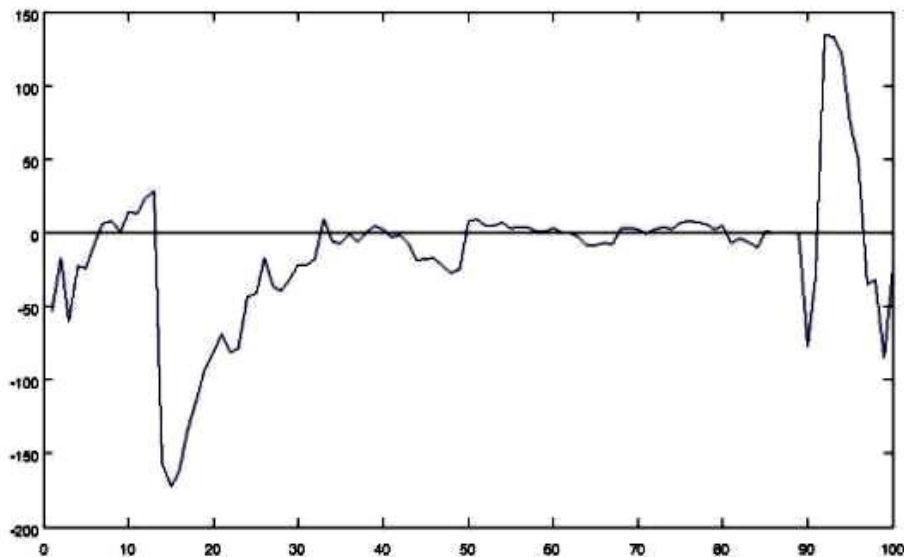
Slika 4.6. Algoritam za upravljanje mobilnom robotskom platformom a) i potprogram za praćenje koordinata detektovanog objekta b)



Slika 4.7. Mobilna robotska platforma u akciji

Da bi analizirali performanse sistema za prepoznavanja i praćenja šuterbota, iskoristili smo podatke dobijene iz LabVIEW-a. Dijagram predstavljen na slici 4.8 prikazuje promenu x koordinate šuterbota (merene u pikselima) u toku vremena, dobijena na osnovu pogleda kamere.

Na ovom dijagramu X osa predstavlja vreme, a Y osa x koordinatu objekta koji mobilni robot treba da prati. Na osnovu tog podatka mobilni robot prati objekat tako što odlučuje na koju stranu da krene u nameri da minimizuje apsolutnu vrednost greške, odnosno x koordinate. Tokom eksperimenta dva puta je šuterbot naglo promenio smer kretanja, nakon 13 sekundi od početnog trenutka eksperimenta i u 89. sekundi.



Slika 4.8 Promena x koordinate objekta koje mobilna robotska platforma prati

Imajući u vidu činjenicu da implementirani neuro-fazi klasifikator može da prepoznaje različite tipove objekata i donosi odluke koje su pouzdane i realne, predloženi algoritam se može implementirati u različitim zadacima u robotskoj viziji.

Predstavljeni metod praćenja objekta se može dodatno popraviti dodavanjem odgovarajućeg filtera kretanja, poput Kalmanovog filtera. Ovakav filter vrši ocenu trenutnog položaja objekta interpolacijom snimljenih podataka ili na osnovu prethodno snimljenih podataka vrši predviđanje narednog položaja objekta. Ograničenje koje se javlja kod Kalmanovog filtera jeste njegova promenljivost isključivo na praćenje objekata koji se kreću konstantnom brzinom ili imaju konstantno ubrzanje.

4.2. Eksperimentalna robotska platforma FRIEND sa stereo kamerom

Tokom poslednjih 15 godina je na Institutu za automatiku Univerziteta u Bremenu razvijana robotska platforma za pomoć osobama sa invaliditetom kroz veliki broj različitih

projekata. Robotska platforma za pomoć osobama sa invaliditetom FRIEND (Functional Robot arm with FIENDly interface for Disabled people) se tokom godina menjala, ali je uvek na električna invalidska kolica montirana robotska ruka, sistem vizije, računarsko-upravljački sistem i različiti vidovi korisničkog interfejsa [43-50].

Ova robotska platforma, tj. njen sistem vizije bio je osnov razvoja svih ostalih sistema vizije predstavljenih u ovoj tezi, te je stoga posebna pažnja posvećena rezultatima istraživanja dobijenih implementacijom inteligentnih algoritama u sistem vizije ovog robotskog sistema.

Različite hardverske komponente mogu se razvrstati u 4 klase:

- senzori, potrebni za razumevanje okoline;
- aktuatori, koji služe za izvođenje akcija koje od robota zatraži korisnik;
- ulazno-izlazni uređaji, za interakciju robota sa korisnikom;
- računarski sistem, za akviziciju i obradu podataka, kao i planiranje i upravljanje.

Za razvoj ovakvog sistema bila je potrebna saradnja istraživača iz različitih oblasti, od onih koji se bave računarskom vizijom, preko robotičara i automatičara do istraživača koji se bave neuro-rehabilitacijom. U okviru bilateralnog projekta između Mašinskog fakulteta Univerziteta u Nišu i Instituta za automatiku Univerziteta u Bremenu pod nazivom "Robusni sistem vizije u rehabilitacionoj robotici" korišćena je treća generacija robotske platforme FRIEND, koja je prikazana na slici 4.9.

Osnovna senzorska komponenta robotske platforme jeste modul računarske vizije sa Bumblebee® 2 stereo kamerom (Slika 4.9). Stereo kamera poseduje dva 1/3" Sony® progressive scan CCD ICX204 senzora koji daju dve uparene RGB slike rezolucija 1024x768 piksela sa maksimalno 20 frejmova u sekundi i kvadratnim pikselima veličine $4,65\mu m$. Žižna daljina ovih senzora je $6mm$, horizontalno polje gledanja (engl. Horizontal Field of View - HFOV) je 43° , a rastojanje između dva sočiva je $120mm$.



Slika 4.9 FRIEND robotska platforma [46]



Slika 4.9 Bumblebee® 2 stereo kamera i upravljivo postolje sa 2 stepena slobode kretanja

Bumblebee® stereo kamera je prekalibrisana, tako da se na slikama ne javlja distorzija i izobličenje objekata. Konverziju iz analognog signala u digitalne slike vrši 12-to bitni analogno-digitalni konvertor, dok je komunikacija između kamere i računara serijska i ostvaruje se preko 6-to pinskog IEEE-1394a FireWire interfejsa. Različiti parametri poput ekspozicije, osvetljenja, otvora blende i sl. se mogu podešavati bilo manuelno, bilo automatski.

Stereo kamera montirana je na upravljivo postolje sa dva stepena slobode kretanja koje može zaokretati kameru gore-dole i levo-desno (engl. Pan-Tilt Head - PTH) koje se nalazi iznad i iza korisnika, tako da kamera može da posmatra scenu ispred robotske platforme uključujući i robotsku ruku i poslužavnik koji se nalazi ispred korisnika. Ugao gledanja kamere može se menjati promenom orijentacije upravljivog postolja.

Kao upravljivo postolje u sistemu FRIEND korišćena je Schunk® Power Cube servo-električna PTH jedinica sa dva stepena slobode kretanja, i postolje se može zaokretati za 180° gore-dole i za 180° levo-desno. Za precizno pozicioniranje ova jedinica koristi dva inkrementalna enkodera sa rezolucijom od 2000 inkremenata po rotaciji, dok je za komunikaciju sa računarom zadužen CAN bus interfejs.

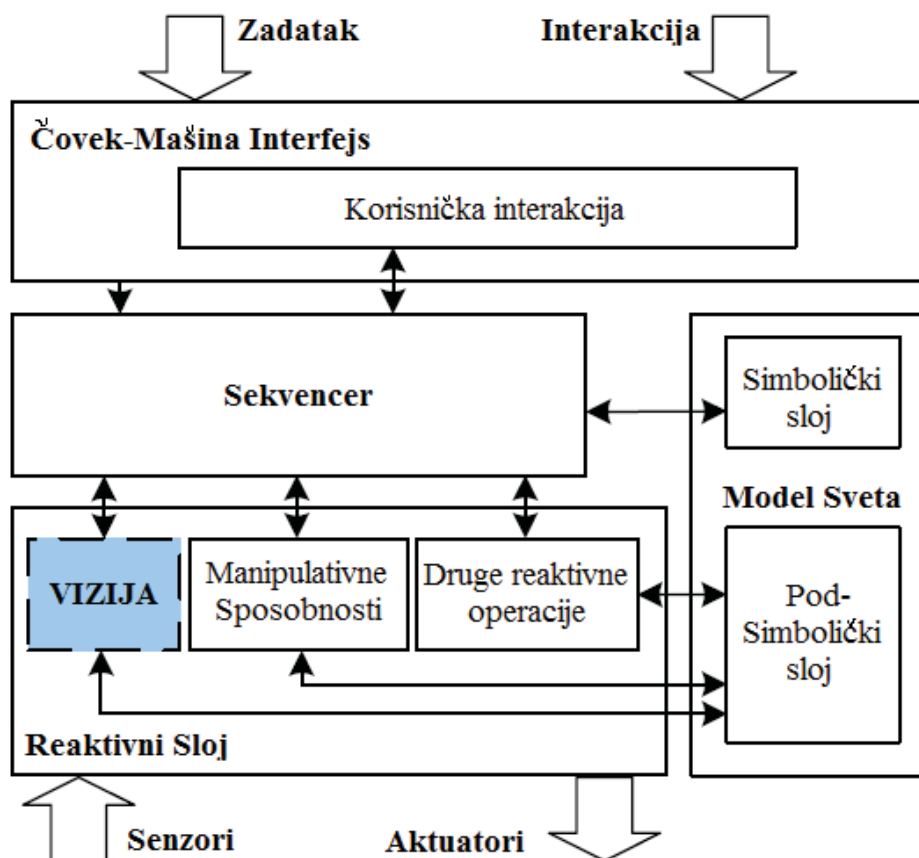
Osnovni zadatak ove robotske platforme bio je pomoći osobi sa hendikepom odnosno invaliditetom u obavljanju svakodnevnih aktivnosti, prvenstveno onih koje osoba obavlja u kućnom okruženju i odnose se na pripremu i posluživanje hrane i pića (slika 4.10). Aktivnosti su podrazumevale rukovanje tipičnim predmetima u domaćinstvu, poput otvaranja i zatvaranja frižidera, otvaranja i zatvaranja mikrotalasne pećnice, hvatanje i manipulacija čašama, flašama, poslužavnikom sa hranom i sl.



Slika 4.10 Robotska platforma i karakterističan kućni scenario

Kako bi realizacija ovih zadataka bila moguća, razvijena je upravljačka struktura nazvana MASSiVE (engl. MultiLayer Architecture for SemiAutonomous Service Robots with Verified Task Execution, slika 4.11). Ova upravljačka arhitektura sastoji se iz 4 segmenta:

- Segment za interakciju robota i čoveka (engl. Human-Machine Interface - HMI);
- Sekvencer, koji planira redosled operacija;
- Reaktivni sloj, algoritam koji prevodi informacije dobijene kao komande od strane korisnika i informacije dobijene sa senzora u akcije robota ;
- Model okruženja.



Slika 4.11 MASSiVE upravljačka arhitektura sistema FRIEND [44, 47]

Sistem vizije je integrisan u reaktivni sloj i omogućava prepoznavanje okruženja. S jedne strane šalje informacije sekvenceru, u cilju planiranja akcija, dok s druge strane komunicira s modelom okruženja i koristi podatke poput karakteristika objekata kako bi izveo adekvatnu klasifikaciju i prepoznavanje objekata, ali i šalje podatke u cilju 3D rekonstrukcije okruženja. Ovi podaci se kasnije koriste za planiranje putanje robotske ruke i za hvatanje objekata.

4.2.1. Segmentacija slike u povratnoj sprezi i izdvajanje karakteristika za klasifikaciju

Algoritam za prepoznavanje objekata koristi karakteristike binarnih slika objekata dobijenih robusnom segmentacijom slike sa povratnom spregom, razvijenom od strane istraživačkog tima iz Bremena [22, 44, 45, 48, 47] koja u realnim uslovima eksploatacije gde npr. osvetljenje nije konstantno daje prihvatljive rezultate. Pronalaženjem lokalnih minimuma

na dijagramu promene konektivnosti određuju se intervali pragova kod kojih svaki definiše zaseban uniformno obojeni objekat.

Scenario koji je korišćen podrazumevao je otvoreni frižider koji se nalazi ispred robotske platforme i u njemu različite čaše, flaše i poslužavnik sa drškom (slika 4.12 a) u jednom eksperimentu, odnosno samo čaše i flaše u drugom (slika 4.12 b).



a)

b)

Slika 4.12 Različito osvetljene scene u kojima je potrebno prepoznati objekte zarad dalje manipulacije

Tokom eksperimenata su objekti raspoređivani na različite načine i vršena je promena osvetljenja. Robusnost algoritma za segmentaciju slike testirana je pri prirodnom osvetljenju, pri zamačivanju prostorije tamnim zavesama i kombinacijom sa 4 vrste različitih veštačkih izvora osvetljenja, gde se osvetljenje u sceni menjalo od 15lx do 500lx, što je mereno ATP® DT-1309meračem osvetljenja (slika 4.13).



Slika 4.13 Merenje osvetljenja scene







Kako se u sceni javljaju samo tri (odnosno u drugom eksperimentu samo dve) klase objekata, za uspešnu klasifikaciju nije bio potreban veliki broj karakteristika, već je osnovna karakteristika korišćena za klasifikaciju bila proporcionalnost, tj. odnos između visine i širine binarnog segmentiranog objekta (već predstavljena u odeljku 3.5 i definisana jednačinom 3.26.). Kao druga karakteristika prirodno se nameće konektivnost osam susednih piksela (već predstavljena u odeljku 3.5 i definisana jednačinom 3.27.), jer je ova karakteristika već računata pri segmentaciji slike, a opisuje i "razuđenost" oblika binarnog segmentiranog objekta.

Na osnovu snimljenih 200 slika uspešno je segmentirano 430 objekata koje je trebalo prepoznati, tj. obučiti klasifikator kako bi prepoznao kojoj od 3 klase objekat pripada. U drugom eksperimentu je slikano preko 600 slika, odakle je segmentacijom izdvojeno 810 objekata, odnosno određeno je 810 vektora karakteristika koje je trebalo razvrstati u dve klase. Ovde je bilo potrebno odrediti da li je segmentirani objekat čaša ili flaša, a objekti su slikani pri osvetljenju od 23lx, 105lx, 240lx i 340lx.

Potrebno je istaći da u oba eksperimenta i pored toga što je akvizicija vršena stereo kamerom, samo su leve slike segmentirane i kasnije je vršeno određivanje karakteristika segmentiranih binarnih objekata.

U tabeli 2 prikazani su rezultati segmentacije jedne te iste scene snimljene pod različitim uglovima i pri različitom osvetljenju i određene su vrednosti konektivnosti i proporcionalnosti.

Tabela 2. Rezultati segmentacije objekata i ekstrakcije karakteristika.

Flaša	Čaša	Ručica poslužavnika
 $I=0.2053$ $p_r=3.5778$	 $I=0.0779$ $p_r=1.7051$	 $I= 0.2835$ $p_r = 0.2153$
 $I=0.2277$ $p_r =3.2727$	 $I = 0.2973$ $p_r = 1.8431$	 $I= 0.4318$ $p_r = 0.4019$

Može se uočiti da nisu sve slike uspešno segmentirane, i pored robusnog segmentacionog algoritma, prvenstveno zbog činjenice da su se u pojedinim scenama neki objekti međusobno zaklanjali druge, bacali senke jedni preko drugih ili je jednostavno osvetljenje bilo loše, pa segmentacija po boji nije mogla biti odrađena. Kako je primarni cilj ovih eksperimenata bio zapravo testiranje algoritma segmentacije i priprema trening skupa klasifikatora, odbacivanje nekih loše segmentiranih slika je zapravo bilo korisno za pripremu aplikacije koja će se koristiti u realnom okruženju i u realnom vremenu.

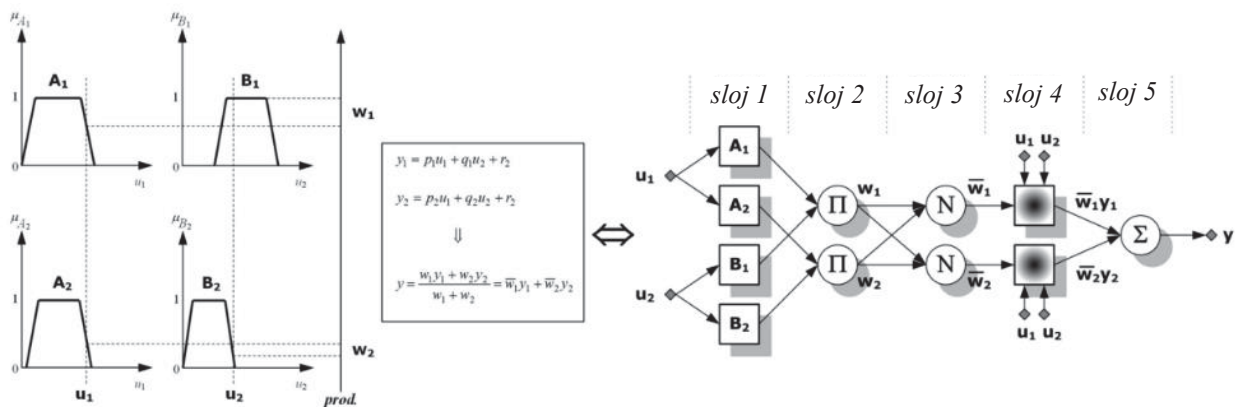
4.2.2. ANFIS klasifikator za prepoznavanje objekata u scenariju FRIEND robotske platforme

U radovima [16,48] i tehničkom rešenju [18] su predstavljeni klasifikatori koji omogućavaju prepoznavanje objekata a zasnovani su na tehnikama računarske inteligencije. Kasnije se ova ideja pojavljuje i dodatno je razrađivana u [5,7]. Razvijeni klasifikatori su zasnovani na ANFIS strukturi i neuronskoj mreži, a obučeni su i testirani vektorima karakteristika dobijenih segmentacijom preko 800 objekata (kombinovanjem objekata dobijenih segmentacijom pri prvom i drugom eksperimentu).

ANFIS (Adaptive-Network-Based Fuzzy Inference System - fazi sistem za zaključivanje zasnovan na adaptivnoj mreži), je primer TSK fazi sistema kome adaptacioni mehanizam podešava parametre. ANFIS se naziva i neuro-fazi sistemom, jer je mehanizam za adaptaciju njegovih parametara sličan mehanizmu učenja kod neuronskih mreža. Načelno se ANFIS koristi za klasifikaciju u problemima vezanim za procesiranje signala (i slika je jedan

vid signala koji se dovodi do računara), a posebno je česta primena u medicinske svrhe npr. za analizu EEG signala [65] ili za analizu ECG signala [66], a u cilju rane detekcije bolesti.

Klasifikacija slika, tj. objekata sa slike primenom ANFIS klasifikatora (slika 4.14) nije novina i često se sreće u medicinskim aplikacijama [67-70], gde se na slikama dobijenim magnetnom rezonancom nalaze razne vrste kancerogenih ćelija. Analizom ovih radova jasno se uočava superiornost ANFIS klasifikatora u odnosu na relativno jednostavne algoritme klasifikacije zasnovane na mašinskom učenju, poput kNN klasifikatora ili fazi klasteringa [67], ali su evidentni i nedostaci u odnosu na SVM klasifikator [68]. Naravno, ovo poređenje klasifikatora ne može se generalizovati, jer detekcija i prepoznavanje objekata na slici zavisi od vrste slike (klasifikacija je u ovim radovima vršena na osnovu MR slike), načina procesiranja i segmentacije, kao i same primene.



Slika 4.14 ANFIS struktura

Softver za prepoznavanje objekata baziran je na ANFIS strukturi sa 5 slojeva, korišćen je hibridni algoritam učenja, (back-propagation metod za nelinearne parametre i metod najmanjih kvadrata za linearne parametre) a Modified Mountain Clustering (MMC) tehnike su korišćene za određivanje početne strukture neuro-fazi klasifikatora. Svrha klasterovanja jeste prirodno odrediti slične skupove podataka, koji kasnije služe kao jezgra fazi skupova

U koliko se posmatra skup n normalizovanih tačaka podataka $\{x_1, x_2, \dots, x_n\}$ u M -dimenzionom prostoru, mera potencijala tačke podataka može se definisati kao:

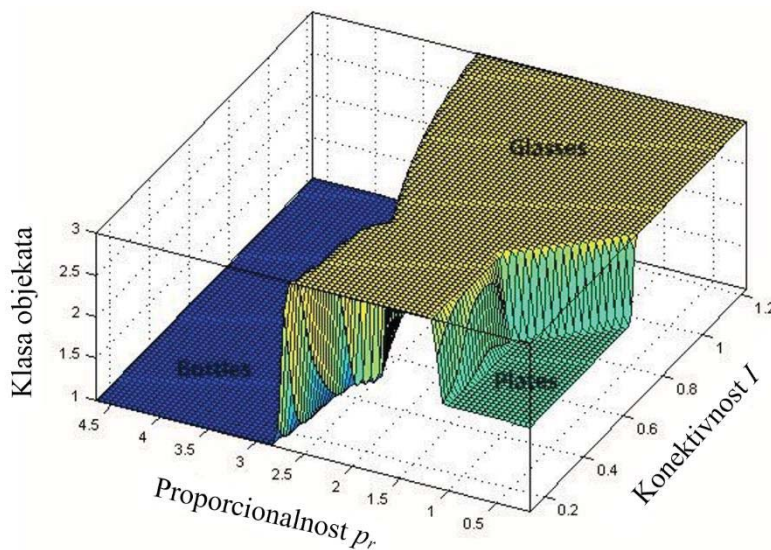
$$P_i = \sum_{j=1}^n \exp\left(-\alpha \|x_i - x_j\|^2\right), \quad \alpha = 4/r_a^2 \quad 4.7.$$

Konstanta r_a je radijus koji definiše susedne tačke. Nakon što se izračuna potencijal svake tačke, tačka sa najvećim potencijalom se usvaja kao prvi centar klastera. Ako je x_1^* prvi centar klastera sa potencijalom P_1^* , onda se menja potencijal ostalih tačaka i računaju se na sledeći način:

$$P_i \leftarrow P_i - P_i^* \exp\left(-\beta \|x_i - x_j\|^2\right), \quad \beta = 4/r_b^2, \quad 4.8.$$

gde je r_b pozitivna konstanta veća od r_a , kako bi se izbeglo nagomilavanje centara klastera (najčešće je $r_b = 1.5r_a$).

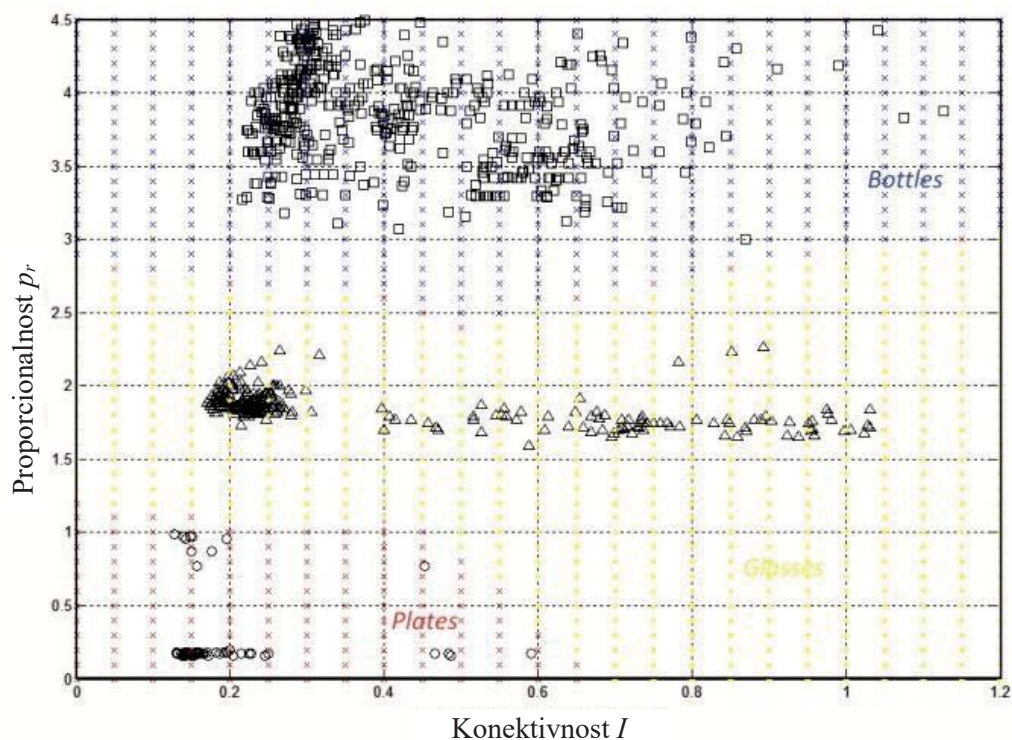
Početna struktura klasifikatora sa 14 pravila, i 12 do 14 primarnih fazi skupova u prostoru ulaza i izlaza, izabrana je kao optimalna. Ovo je urađeno sa 430 parova za treniranje. Isti skup parova se koristi za treniranje klasifikatora, tj. za podešavanje parametara. Obučavanje klasifikatora je vršeno za manuelno označene segmentirane regione različitih objekata koji pripadaju trima klasama, "Glasses", "Bottles" i "Plates", a konačni izlaz treniranja je formirana fazi struktura. Testiranje razvijenog klasifikatora izvršeno je korišćenjem deskriptora drugih 215 vektora karakteristika, dok je treći skup od 215 karakteristika korišćen za validaciju rezultata. Na slici 4.15 predstavljena je fazi površ dobijena nakon obučavanja klasifikatora.



Slika 4.15 Fazi površ dobijena nakon obučavanja ANFIS strukture

Na slici 4.16 prikazani su rezultati obuke ANFIS klasifikatora. Kako je ulazni vektor karakteristika imao samo dva elementa, konektivnost i proporcionalnost, moguće je predstaviti rezultate klasifikacije na 2D grafiku. Rezultati klasifikacije su jako dobri i čak i uz promenu osvetljenja i povremeno međusobno zaklanjanje objekata je pri verifikaciji tačnost bila veća od 95%.

U koliko se razmatraju performanse klasifikatora moraju se u obzir uzeti dva međusobno suprotstavljena cilja. S jedne strane klasifikator mora da poseduje zadovoljavajuću tačnost kako bi pravilno prepoznao objekte, dok je s druge strane vreme potrebno za klasifikaciju kritičan resurs, jer robot treba da prepoznaje objekte u realnom vremenu. ANFIS klasifikator uspešno ispunjava oba kriterijuma klasifikacije, jer nakon off-line obuke daju zadovoljavajuću tačnost uz minimalno vreme potrebno za klasifikaciju.



Slika 4.16 Rezultati klasifikacije primenom ANFIS klasifikatora

4.2.3. Klasifikator zasnovan na neuronskoj mreži za prepoznavanje objekata u scenariju FRIEND robotske platforme

Veštačke neuronske mreže (ANN) ili samo neuronske mreže (NN) su računarski modeli koji su nastali kao pokušaj matematičke formalizacije strukture ljudskog mozga. To su modeli koji omogućavaju sprovođenje zadataka kao što su učenje, klasifikacija i optimizacija. Takođe, neuro mreže se karakterišu masivnim paralelizmom operacija što ih čini veoma efikasnim. Modeli veštačkih neuronskih mreža mogu se shvatiti kao drugačiji način reprezentacije znanja, koji se razlikuje od mapiranja visokog nivoa (pravila, semantičke mreže, okviri, itd.). Neuro mreže predstavljaju jedinstvenu metodologiju kojom se znanje prikuplja iz skupova podataka za obučavanje i smešta u distribuiranom obliku u strukturu mreže.

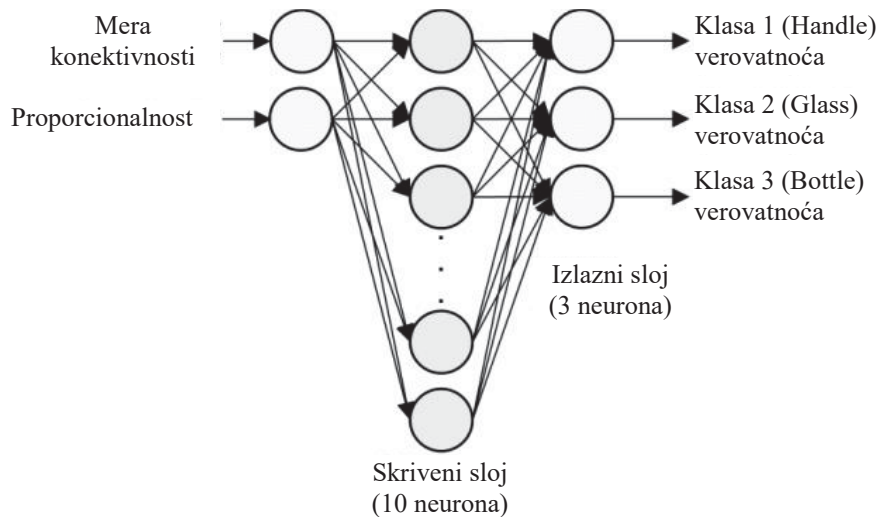
Osnovni elementi koji čine model veštačkih neuronskih mreža su:

- Skup procesnih elemenata, zvanih neuroni, ćelijski procesni elementi, itd.
- Svaki elemenat i prima povorku ulaza (iz okoline ili od drugih procesnih elemenata), koji mogu imati vrednosti u raznovrsnim domenima. Ako je skup vrednosti $\{0, 1\}$, neuron se naziva binarnim. Ako je skup vrednosti $\{-1, 1\}$ neuron je bipolarni. Takođe često su korišćeni slučajevi kod kojih je skup mogućih vrednosti kontinualni interval kao $[0, 1]$ ili $[-1, 1]$.
- Neuroni su međusobno povezani vezama, koje se takođe nazivaju konekcije ili sinapse, koje su u pravilu jednosmerne iako postoje i modeli sa simetričnim konekcijama. Svaka konekcija sprovodi komunikaciju između dva elementa i ima pridruženu težinu ili težinski faktor w_{ij} , koji određuje efikasnost sa kojom signali i -tog neurona i utiču na j -ti neuron.
- Pravilo agregacije ulaza, koje određuje nivo ekscitacije elementa i .
- Aktivaciona funkcija f koja indicira stanje ili nivo aktivacije a_i , prema nivou ekscitacije, koji se može propagirati kao ulaz drugim neuronima.
- Spoljni ulaz za svaki element, ili trend θ_i , se takođe primenjuje u mnogim slučajevima. On predstavlja prag koji ukupni signal koji stiže do neurona mora da premaši da bi ga aktivirao.

U ovom istraživanju je kao klasifikator korišćena standardna neuronska mreža sa prostiranjem signala u napred i sa tri sloja, ulazni sloj, skriveni sloj sa 10 skrivenih neurona i izlazni sloj (slika 4.17). Ulaz u neuronsku mrežu čini vektor karakteristika čija su dva elementa konektivnost i proporcionalnost, dok je izlaz iz neuronske mreže verovatnoća da segmentirani binarni objekat pripada svakoj od tri klase: čašama, flašama ili ručicama poslužavnika.

Za obučavanje je korišćen algoritam povratne propagacije greške (EBP), koji je zapravo generalizacija delta pravila. Osnovna ideja je da se ponovo kvantifikuje postojeća greška kroz razliku između izlaza mreže i ispravnog izlaza. Teškoća leži u činjenici da postoji veliki broj

konekcija koje uvećavaju grešku, koji se ne odnose samo na neurone izlaznog sloja. Posebno, pošto nema informacije o ulozi koju igraju skriveni neuroni, njihov doprinos ukupnoj grešci se ne može direktno izračunati.



Slika 4.17 Struktura klasifikatora zasnovanog na neuronskoj mreži

Problem se može razrešiti ako se pretpostavi da greška koja se uočava u izlaznom sloju nastaje primarno usled akcije skrivenih elemenata lociranih neposredno u prethodnom sloju, i tako nadalje. Proces adaptacije definisan je sa dve etape koje se ponavljaju dok se obučavanje ne smatra završenim. U prvoj etapi nivo aktivacije svih neurona u mreži se procenjuje zadržavanjem fiksnih vrednosti konekcija, što omogućava da se odredi nivo postojeće greške. U drugoj fazi ova greška se propagira povratno od sloja do sloja, uspešno modifikujući težine koje u će u narednoj etapi služiti za izračunavanje nove greške. Ovaj proces je ono što daje ime čitavoj metodi.

Kod višeslojne mreže sa n ulaznih neurona, h skrivenih neurona i m izlaznih neurona aktivaciona funkcija je diferencijabilna funkcija stvarnog ulaza u neuron:

$$a_j^p = F_j(e_j^p), \quad 4.9.$$

gde je:

$$e_j^p = \sum_i w_{ij} a_i^p + \theta_j, \quad 4.10.$$

tj. stvarni ulaz u j -ti element za p -ti primer.

Nakon prve epohe u kojoj je procenjena greška mreže, svaki težinski faktor se podešava na osnovu izraza:

$$\Delta_p w_{ij} = \gamma \delta_j^p a_i^p. \quad 4.11.$$

Parametar δ_j^p dobija različite vrednosti za slučaj izlaznih i internih elemenata. Počinje se postavljanjem težina koje povezuju elemente skrivenih slojeva i izlaznog sloja. U tom slučaju je:

$$\delta_j^p = (d_j^p - a_j^p) F_j'(e_j^p), \quad 4.12.$$

za svaki izlazni element j . Ako se razmatra skriveni sloj, vrednosti δ_j^p se računaju prema izrazu:

$$\delta_j^p = F_j'(e_j^p) \cdot \sum_{h=1}^{N_a} \delta_h^p \cdot w_{hj}. \quad 4.13.$$

Jednačine (3.32) i (3.33) daju rekurzivnu proceduru za računanje delta za sve elemente u mreži. Ova procedura čini generalizovano delta pravilo za algoritam povratne propagacije greške (EBP).

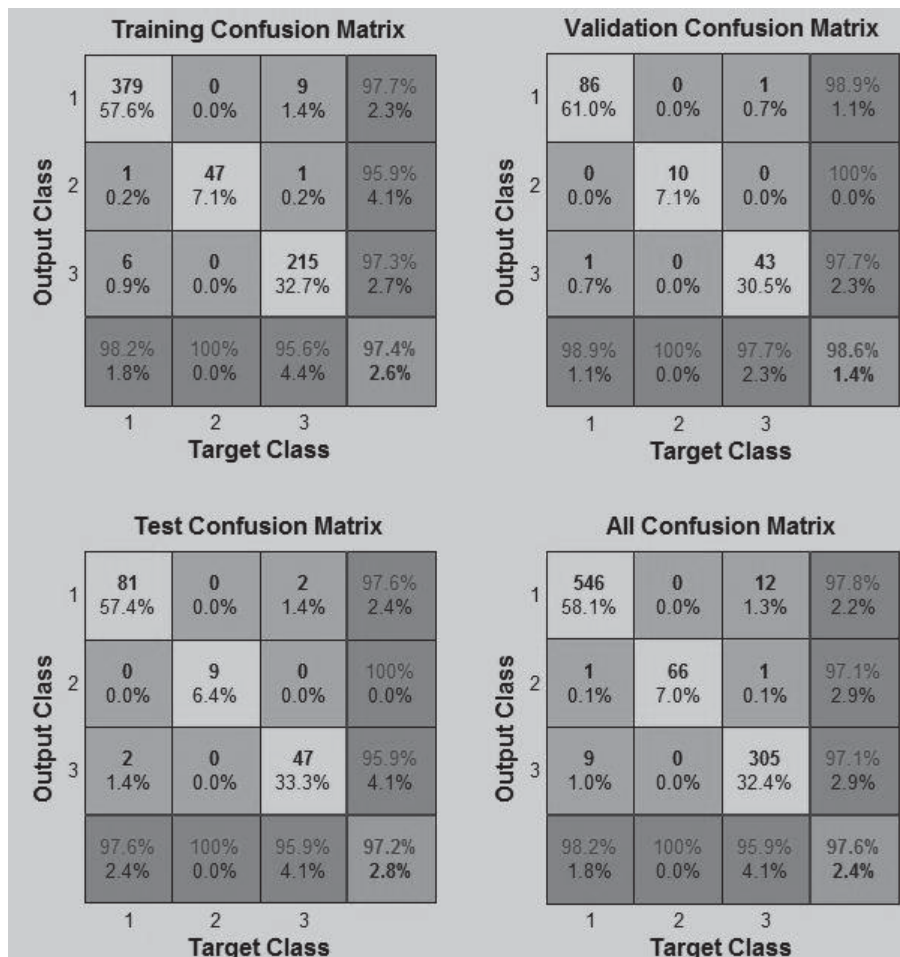
Metod najstrmijeg spusta zahteva da se podešavanje vrši kroz minimalne korake, ali zbog praktičnih efekata zanimljivo je razmotriti veliku brzinu učenja koja omogućava brzo obučavanje. Problem moguće pojave oscilacija za velike vrednosti izbegava se dodavanjem momentnog člana:

$$\Delta w_{ij}(t+1) = \gamma \delta_i^p a_j^p + \alpha \Delta w_{ij}(t) \quad 4.14.$$

gde je α konstanta momenta, koja meri važnost prethodnog podešenja u odnosu na trenutno.

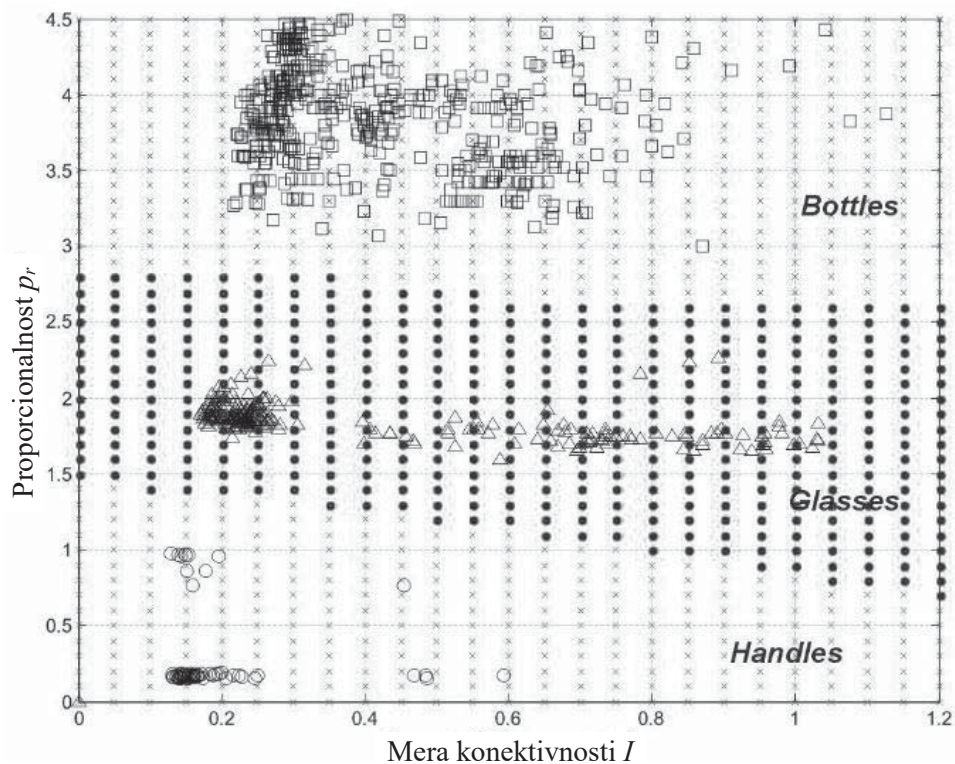
Zahtevni set podataka od 940 tripleta (dva elementa ulaznog vektora i izlaz) je razdeljen slučajnim izborom na 658 semplova za obučavanje mreže, 141 za validaciju mreže i 141 za

testiranje mreže. Na slici 4.18 prikazani su rezultati obučavanja neuronske mreže, odnosno konfuzione matrice trening seta, validacionog seta, test seta podataka i zbirna konfuziona matrica.



Slika 4.18. Rezultati obučavanja klasifikatora zasnovanog na neuronskoj mreži

Na osnovu konfuzionih matrica može se zaključiti da je prepoznavanje čak i za ovako zahtevan set podataka klasifikator zasnovan na neuronskoj mreži izuzetno uspešno, jer je tačnost klasifikatora veća od 97%. Kako bi rezultati klasifikacije neuronskom mrežom bili uporedivi sa klasifikacijom primenom ANFIS klasifikatora, na slici 4.19 je prikazan dijagram sa rezultatima klasifikacije neuronskom mrežom.

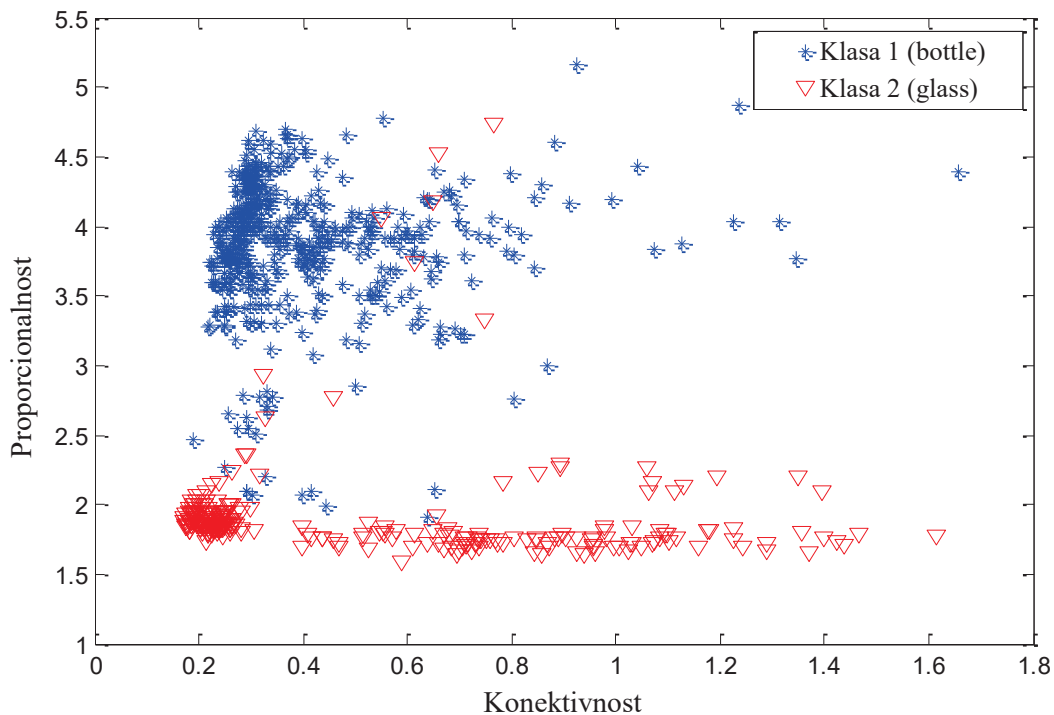


Slika 4.19 Rezultati klasifikacije neuronskom mrežom

I ANFIS klasifikator i klasifikator zasnovan na neuronskoj mreži tokom testiranja pokazali su visoku robusnost, ali se njihov pun potencijal može iskoristiti tek uvođenjem većeg broja klasa objekata ili uvođenjem većeg broja karakteristika kao elemenata ulaznog vektora. Potrebno je naglasiti da se nakon obučavanja klasifikacija izvršava izuzetno brzo, te nije potrebno unapređenje hardvera kako bi se omogućila real-time interakcija robota sa okolinom.

4.2.4. Naivni Bajesov klasifikator za prepoznavanje objekata mobilnom robotskom platformom FRIEND

Prilikom drugog eksperimenta rađenog mobilnom robotskom platformom FRIEND je usled promene osvetljenja i međusobnog zaklanjanja objekata algoritam za segmentaciju dao loše rezultate, što se može videti i sa slike 4.20. Usled okulzije se delovi objekata na slici ne vide, pa otuda i nelogične vrednosti proporcionalnosti segmentiranih binarnih objekata. Kako bi se prevazišao ovaj problem primenjen je naivni Bajesov klasifikator, koji bi trebalo da bude manje osetljiv na ovakve podatke [17].



Slika 4.20 Trening set za klasifikaciju

Prilikom projektovanja ANFIS i ANN klasifikatora su ovi podaci ručno odbačeni iz trening seta. Pri klasifikaciji je korišćena je arhitektura klasifikatora kod koga vektor ulaza koji ima dva elementa - konektivnost i proporcionalnost, dok su izlazi verovatnoće da objekat pripada jednoj od dve klase - čašama ili flašama (Slika 4.21).



Slika 4.21 Arhitektura naivnog Bajesovog klasifikatora

Bajesova metoda predstavlja verovatnosno okruženje za rešavanje problema klasifikacije. Metoda se zasniva na Bajesovskoj teoremi koja kombinuje prethodno znanje o klasama sa novim znanjima dobijenim iz podataka za učenje. Bajesova teorema glasi:

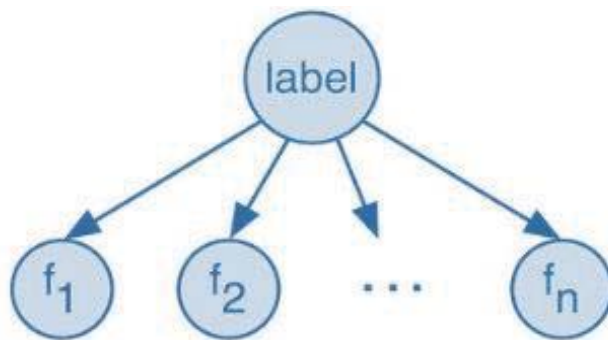
$$P(h | D) = \frac{P(D | h)P(h)}{P(D)} \quad 4.15.$$

pri čemu su $P(h | D)$ i $P(D | h)$ uslovne verovatnoće.

Naivni Bajesov klasifikator (Slika 4.22) klasifikuje nepoznate primere na osnovu Bajesove teoreme i podataka za učenje (na osnovu dodeljenih labela) ali uz jednu značajnu pretpostavku a to je nezavisnost atributa (f_1, f_2, \dots, f_n).

$$P(c | i) \propto P(c) \prod_{1 \leq k \leq n_i} P(f_k | c) \quad 4.16.$$

Za diskretne attribute verovatnoća $P(f_k | c)$, gde je $k=1, \dots, n$, se generalno može lako izračunati iz podataka za učenje.



Slika 4.22 Graf Bajesovog klasifikatora - svakoj labeli odgovara određeni set karakteristika - atributa

Dostupno je bilo 810 objekata, odnosno određeno je 810 vektora karakteristika koje je trebalo razvrstati u dve klase, a objekti su slikani pri osvetljenju od 23lx, 105lx, 240lx i 340lx. Svi objekti su ručno označeni, tj. dodeljene su im labele. Deset puta su slučajno izabrani vektori ulaza i njima odgovarajući vektori izlaza i na osnovu tih podataka je vršeno obučavanje klasifikatora, dok su preostali vektori ulaza i njima odgovarajući vektori izlaza korišćeni za testiranje. Prilikom testiranja određivana je preciznost, opoziv, F1 skor i tačnost. Rezultati su prikazani u tabeli 3.

Tabela 3 Rezultati klasifikacije

Iteracija	Klasa	Preciznost	Opoziv	F1 skor	Tačnost
1	“Flaša”	68.60%	83.84%	75.45%	66.87%
	“Čaša”	61.90%	40.63%	49.06%	
2	“Flaša”	69.29%	88.89%	77.88%	69.33%
	“Čaša”	69.44%	39.06%	50.00%	
3	“Flaša”	65.91%	87.88%	75.32%	65.03%
	“Čaša”	61.29%	29.69%	40.00%	
4	“Flaša”	73.98%	91.92%	81.98%	75.46%
	“Čaša”	80.00%	50.00%	61.54%	
5	“Flaša”	69.77%	90.91%	78.95%	70.55%
	“Čaša”	73.53%	39.06%	51.02%	
6	“Flaša”	68.85%	84.85%	76.02%	67.48%
	“Čaša”	63.41%	40.63%	49.52%	
7	“Flaša”	70.87%	90.91%	79.65%	71.78%
	“Čaša”	75.00%	42.19%	54.00%	
8	“Flaša”	68.75%	88.89%	77.53%	68.71%
	“Čaša”	68.57%	37.50%	48.48%	
9	“Flaša”	69.17%	92.93%	79.31%	70.55%
	“Čaša”	76.67%	35.94%	48.94%	
10	“Flaša”	68.50%	87.88%	76.99%	68.10%
	“Čaša”	66.67%	37.50%	48.00%	

Na osnovu prikazanih podataka iz tabele vidi se da je izbor trening skupa od izuzetne važnosti pri projektovanju klasifikatora, a razlike u vrednostima tačnosti mogu biti gotovo 10%, što je značajna razlika. Do ovakvih rezultata došlo je zbog loše segmentacije, nikakav drukčiji klasifikator ne bi mogao da da značajnije unapređenje i jedini način za prevazilaženje ovog problema jeste uvođenje dodatnih karakteristika u ulazni vektor klasifikatora, odnosno dodatnih informacija dobijenih iz slike.

Klasifikator je tolerantan na izolovani šum jer ga on jednostavno "izravna" srednjim vrednostima pri određivanju uslovnih verovatnoća. Takođe, prevazilazi problem nedostajućih vrednosti tako što ignoriše takve primere u procesu izračunavanja procene verovatnoće. Klasifikator dosta dobro toleriše irelevantne attribute, ali s druge strane visoko korelisani atributi mogu značajno da ugroze kvalitet ovog klasifikatora jer za njih pretpostavka o nezavisnosti nikako ne važi.

4.3. Prepoznavanje objekata robotskom platformom FRIEND na osnovu karakteristika segmentirane mape dispariteta

Kako bi se pokazao pun potencijal klasifikatora zasnovanih na alatima iz domena veštačke inteligencije, izvršen je eksperiment gde je korišćeno više objekata i analizirana je prenatrpana scena (Slika 4.23).



Slika 4.23 FRIEND robot i radno okruženje

Bilo je potrebno klasifikovati objekte u 4 klase, odnosno odrediti da li je jedan od konkretnih segmentiranih objekata čaša, flaša, tegla ili tetrapak. U sceni su se nalazile još i loptica, kutija sa žitaricama, a postojale su dve različite čaše, plastična žuta i staklena čaša.

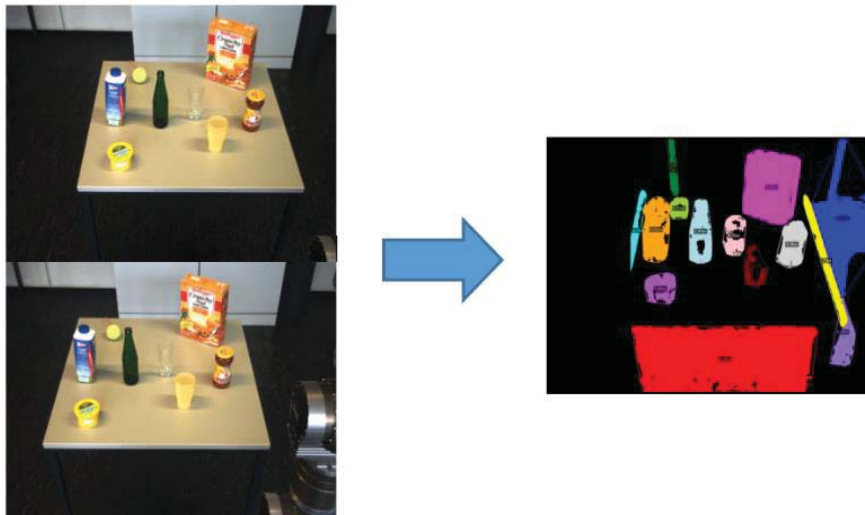
4.3.1. Segmentacija mape dispariteta i određivanje karakteristika segmentiranih objekata

Kako je u sceni koja se nalazi ispred robotske platforme FRIEND (Slika 4.24) bilo više objekata, a objekti uglavnom nisu bili uniformno obojeni, segmentacija po boji ne bi sama po sebi mogla da da zadovoljavajuće rezultate, jer bi bilo potrebno grupisati različito obojene regione istog objekta kako bi uspešno mogli rekonstruisati njegov položaj u prostoru i oblik. Ovde je razrađena i modifikovana ideja predstavljena u radu grupe autora iz Bremena [47], da se vrši segmentacija mape dispariteta, i da se objekat traži unutar poligona okvira koji se dobije ovom segmentacijom.



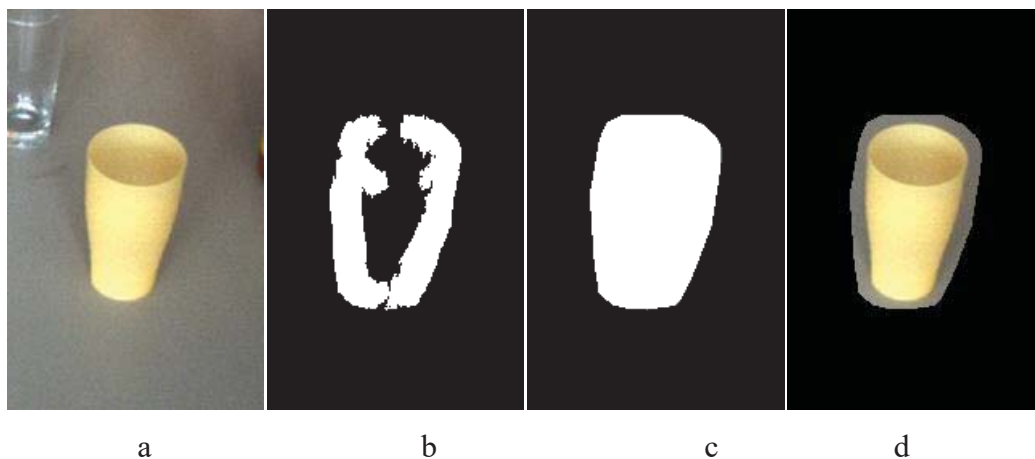
Slika 4.24 Scena sa većim brojem različito obojenih objekata ispred robotske platforme FRIEND

Jasno je da je usled neuniformne obojenosti objekata bilo moguće metodom preklapanja blokova opisanoj u odeljku 3.3 odrediti dosta preciznu mapu dispariteta (slika 4.25). Sa slike se vidi da je formiranje mape dispariteta samo delimično uspešno za jednoboje predmete na slici (zeleni flaša i žuta čaša). U cilju prevazilaženja ovog problema primenom konveksnog popunjavanja formiraju se objekti koji se mogu koristiti kao poligonalni okviri unutar koga se zaista nalazi objekat od interesa.



Slika 4.25 Određivanje mape dispariteta scene na osnovu slika dobijenih stereo kamerom

Segmentirana binarna slika objekta dobijena segmentacionom metodom opisanom u odeljku 3.4.3 ove teze u ovom slučaju ne daje realističnu reprezentaciju objekta, prvenstveno zbog efekta "zadebljanja" objekta (Slika 4.26) do koga dolazi zbog metode preklapanja blokova koja je korišćena za formiranje mape dispariteta.



Slika 4.26 Originalna leva slika a), rezultat segmentacije mape dispariteta b), rezultat konveksnog popunjavanja c) i originalni objekat u poligonalnom okviru

Softver korišćen za dubinsku segmentaciju "HiSpeed" razvijen je od strane Leu-a [25,55], korišćen je za detekciju i prepoznavanje prepreka na putu i povećanje nivoa bezbednosti u saobraćaju. Osnov softvera su algoritmi iz OpenCV biblioteke (Open Source Computer Vision Library). Ova biblioteka sadrži preko 2500 optimizovanih algoritama među






kojima su i najsavremeniji algoritmi računarske vizije i mašinskog učenja, a poseban akcent u izradi ovih algoritama stavljen je na efikasnost izračunavanja i izvršavanje aplikacija u realnom vremenu.

U radu [47] se nakon određivanja poligonalnog okvira vrši segmentacija po boji u zatvorenoj sprezi po algoritmu opisanom u odeljku 3.4.2 ove teze, pa se zatim grupišu segmentirani regioni kako bi se omogućila osnova za valjanu 3D rekonstrukciju objekta i konačno formiranje 3D modela okruženja robota. Ovaj model kasnije omogućava planiranje putanje i hvatanje objekata, a u cilju izvršavanja zadatka.

Ideja koja je predstavljena u ovoj tezi, a javlja se i u radovima [5,7] jeste prepoznavanje objekata koji se nalaze u poligonalnom okviru samo nakon segmentacije mape dispariteta i konveksnog popunjavanja. Ovom metodom moguće je brže pretražiti prostor slike i vršiti 3D rekonstrukciju samo onih objekata koji su od interesa za robotsku akciju (npr. samo rekonstruisati čašu i flašu u koliko robot treba da posluži piće). Za planiranje putanje potrebni su i ostali objekti u sceni, ali nije potrebna njihova verna 3D rekonstrukcija, već se mogu koristiti i segmentirani regioni iz mape dispariteta za definisanje regiona sa kojima robotska ruka ne sme da se sudari.

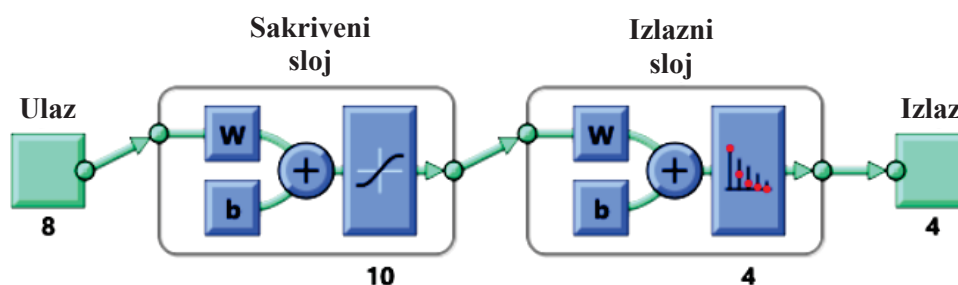
Kako su objekti u sceni sada relativno sličnih dimenzija, a nakon konveksnog popunjavanja im je svima konektivnost približno ista, bilo je neophodno uvesti nove karakteristike koje opisuju oblik segmentiranog objekta. U ovom istraživanju, osim proporcionalnosti korišćeno je i 7 Hu-ovih invarijantnih momenata, pa su rezultati ekstrahovanja karakteristika iz segmentirane slike za 4 objekta iz jedne karakteristične scene predstavljeni u tabeli 4.

Tabela 4 Karakteristike segmentiranih objekata

	Flaša (Class 1)	Tetrapak (Class 2)	Čaša (Class 3)	Tegla (Class 4)
				
p_t	2.7338	2.3343	1.1534	1.4623
I_1	0.7468	0.7424	0.7558	0.7479
I_2	2.6109	2.5318	2.6456	2.6100
I_3	6.6204	5.1435	7.2942	6.6940
I_4	5.6006	5.3548	5.9980	5.7246
I_5	11.561	10.7033	12.8466	12.3733
I_6	6.9190	7.0355	7.3234	7.0721
I_7	11.770	10.5647	13.6693	11.7935

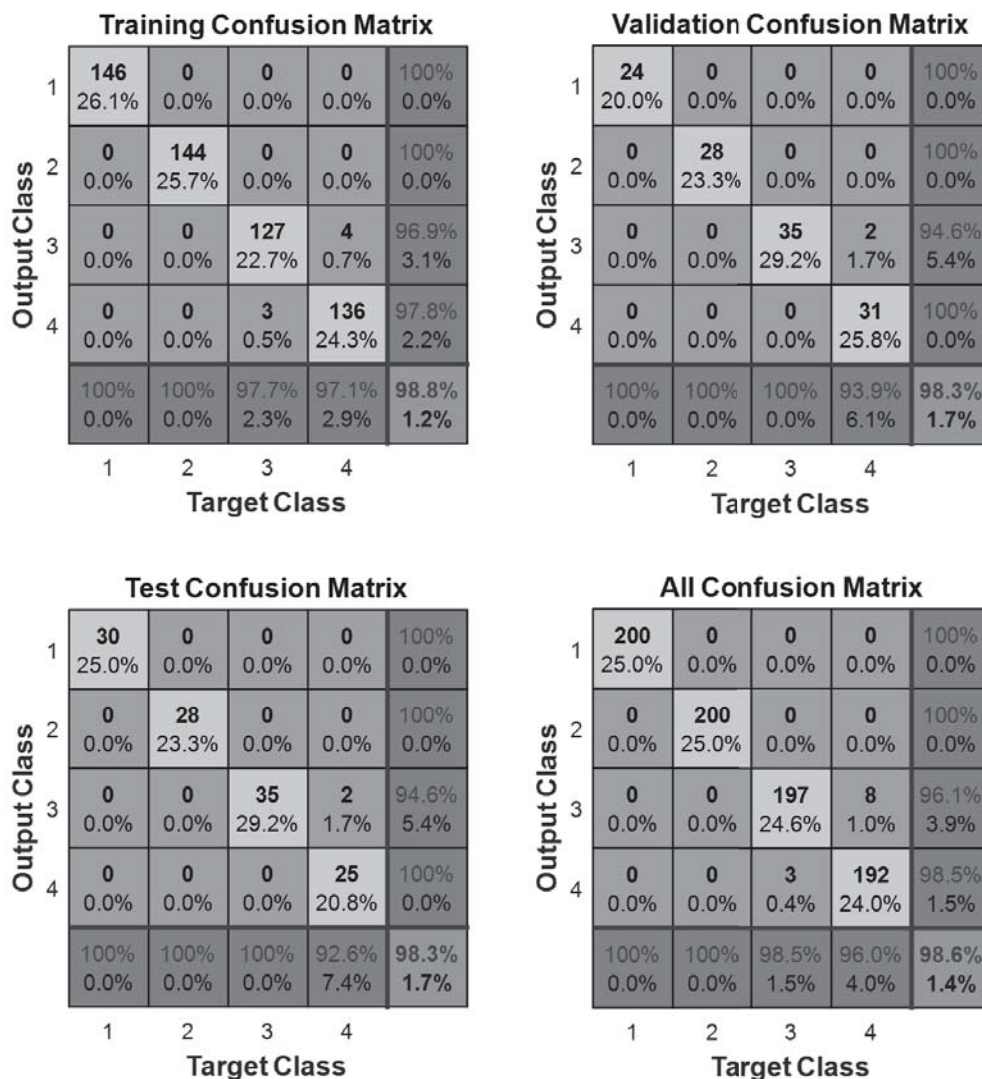
4.3.2. Neuronska mreža za klasifikaciju zahtevnog seta podataka

U klasifikacije objekata kompleksnog seta podataka predložen je klasifikator zasnovan na neuronskoj mreži za prepoznavanje obrazaca sa deset neurona u skrivenom sloju, čija je arhitektura prikazana na slici 4.27. U pitanju je dvoslojna mreža sa prostiranjem signala unapred, gde su neuroni u skrivenom sloju sa sigmoidnom aktivacionom funkcijom, dok su u izlaznom sloju sa tzv. "softmax" funkcijom. Ulaz u klasifikator čini vektor koji sadrži 8 elemenata, tj. karakteristika, i to proporcionalnost i 7 Hu-ovih momenata, dok se vektor izlaza sastoji od 4 elementa koji predstavljaju verovatnoću da objekat sa određenim karakteristikama pripada određenoj klasi.



Slika 4.27 Arhitektura višeslojnog perceptrona

Obučavanje, validacija i testiranje vršeno je sa 800 objekata kojima su ručno određene labele. Slučajno je izabrano 560 setova podataka za obučavanje, a po 120 slučajno izabranih setova je korišćeno za validaciju i testiranje. Za obučavanje je korišćena povratna propagacija greške i metod najstrmijeg spusta. Rezultati obučavanja, validacije i testiranja mreže mogu se videti na konfuzionim matricama prikazanim na slici 4.28, odnosno na slici 4.29 koja predstavlja operativne karakteristike prijemnika.

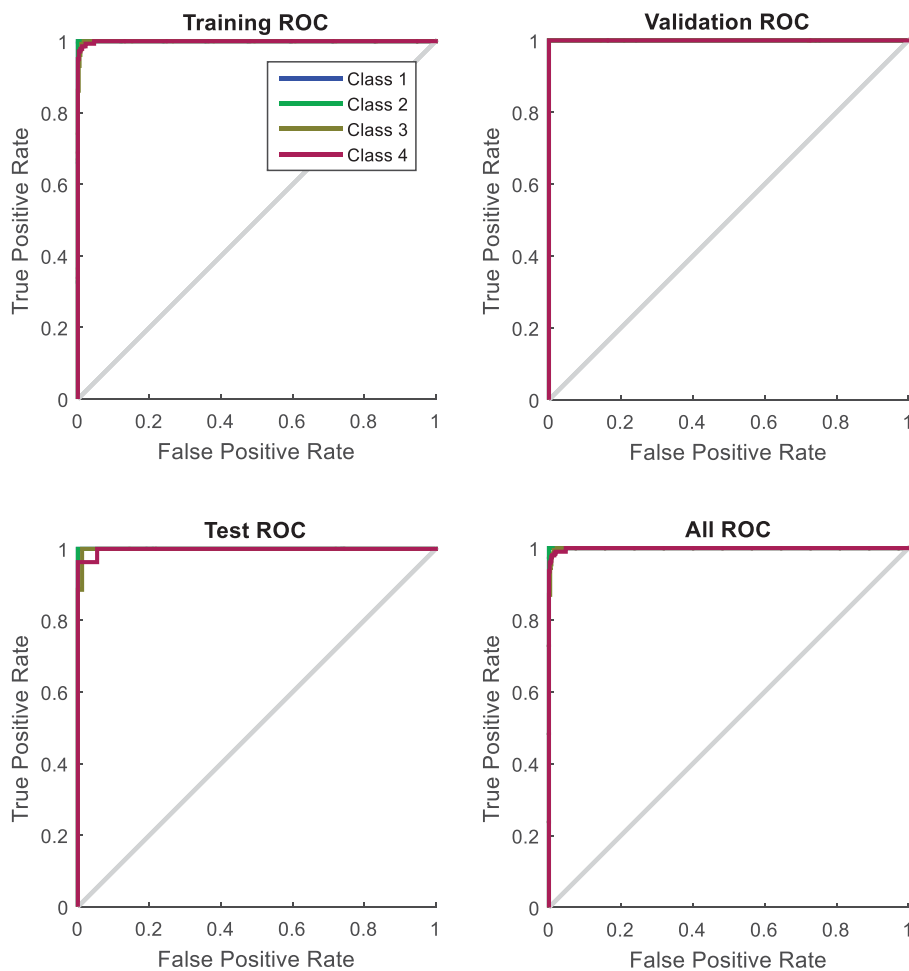


Slika 4.28. Konfuzione matrice klasifikatora

Obučavanje predloženog klasifikatora izvršeno je u 30 iteracija i rezultati klasifikacije su izuzetno dobri. Tačnost obučenog klasifikatora je preko 98%, a može se dodatno povećati

ubacivanjem dopunskih neurona u skriveni sloj. Izlaz iz klasifikatora je verovatnoća da objekat koji ispitujemo i čije smo karakteristike ekstrahovali pripada svakoj od 4 klase. Iz tabele 3 se vidi da su međusobne razlike između karakteristika segmentiranih objekata relativno male, i tek njihova kombinacija daje dobar osnov za klasifikaciju.

Primenom ovog klasifikatora se čak i u slučaju da ne pogodi jedan poligonalni okvir u kome je objekat od interesa smanjuje oblast pretrage slike (slika 4.26 d). Konačni broj poligonalnih okvira će imati verovatnoću veću od granične da se u njima nalazi traženi objekat, pa se onda u njima može vršiti segmentacija po boji i zatim neka od metoda klasifikacije predstavljene u odeljku 4.2. ili 3D rekonstrukcija i prepoznavanje na osnovu pretrage i poklapanja sa unapred poznatim modelom.



Slika 4.29 Operativne karakteristike prijemnika

Ovakva 3D rekonstrukcija je od izuzetnog značaja pri planiranju putanje robotske ruke i pri rešavanju hvatanja objekata robotskom šakom, a značaj uspešnog prepoznavanja objekata

je višestruk. S jedne strane se prepoznavanjem objekata smanjuje oblast pretrage slike i omogućava brža 3D rekonstrukcija, dok se s druge strane informacije o tipu objekta mogu iskoristiti pri hvatanju objekta za estimaciju sile kojom robotska šaka treba da uhvati željeni predmet.

4.4. Inteligentna adaptivna segmentacija slike

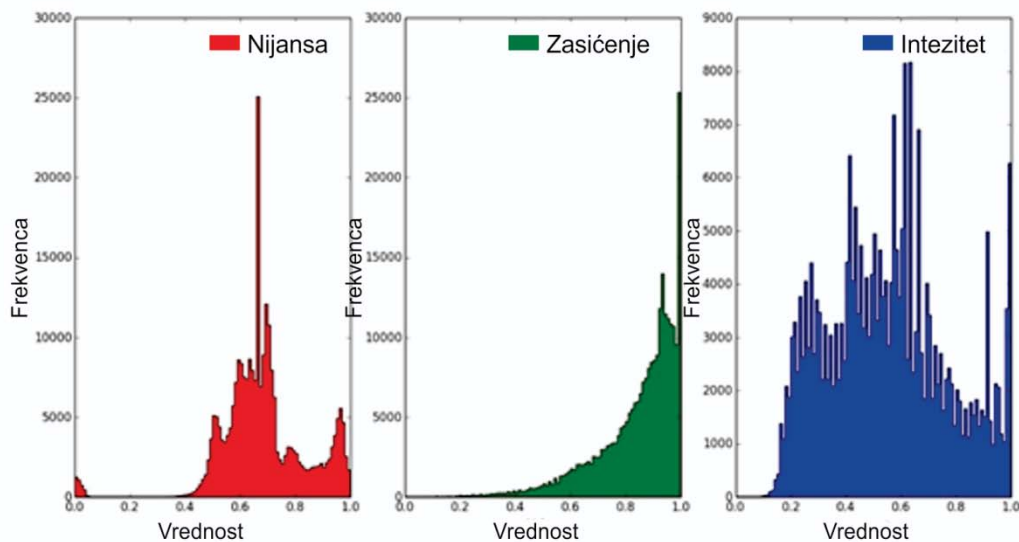
Na osnovu svega izloženog u ovom poglavlju vidi se da klasifikacija nije kritična, već je segmentacija ta koja u pojedinim frejmovima ne daje dobre rezultate, pa se nikakvim klasifikatorom ne mogu nadomestiti greške koje u sistem uvodi algoritam segmentacije. Posebno su kritični uslovi sa slabim osvetljenjem, ispod 50lx, dok se problem okulzija između objekata u slici može rešiti konveksnim popunjavanjem objekata dobijenih segmentacijom mape dispariteta.

Jedino na šta se može uticati pri segmentaciji sa povratnom spregom HSV slike jeste širina intervala praga koja je pri segmentaciji scene iz kuhinje iznosila $[T_{\min}, T_{\max}] = [0, 20]$ [44,47,48]. Ideja koja se krije iza inteligentne adaptivne segmentacije koja je predložena u ovom radu jeste promena širine praga, tj. parametra T_{\max} u zavisnosti od osvetljenja. Za svako pojedinačno osvetljenje scene potrebno je odrediti optimalnu veličinu intervala praga, odnosno T_{\max} . Kako se ovo svodi na optimizacioni problem koji se ne mora rešavati u realnom vremenu, već se izračunavanje ovih vrednosti može vršiti u off line režimu, potrebno je pronaći minimum funkcije cilja koja se može definisati kao razlika između manuelno segmentirane binarne slike i binarne slike dobijene segmentacijom po boji sa povratnom spregom ako T_{\max} variramo po intervalu [10,30]. Manuelna segmentacija bi se pri određenoj vrednosti osvetljenja mogla izvršiti na n slika, pa bi onda funkcija cilja mogla da se definiše kao:

$$\min_{T_{low}, T_{high}} \sum_{i=1}^n \sum_{u=1}^{1024} \sum_{v=1}^{768} \left\| f_i(u, v, T_{\max}) - \hat{f}_i(u, v) \right\|, \quad 4.17.$$

gde je $\hat{f}_i(u, v)$ vrednost piksela manuelno segmentirane binarne slike sa koordinatama (u, v) , a $f_i(u, v, T_{\max})$ predstavlja vrednost piksela binarne slike dobijene segmentacijom po boji sa povratnom spregom sa koordinatama (u, v) i za širinu opsega praga T_{\max} . Kako je rezolucija slike 1024x768 piksela, u i v koordinate se kreću od 1 do 1024, odnosno od 1 do 768.

Sada se pomoću seta optimalno određenih širina intervala pragova može obučiti neuronska mreža koja bi vršila interpolaciju i preporučivala optimalnu vrednost širine praga za određeno osvetljenje scene. Ovakvo rešenje podrazumevalo bi dodatni hardver kojim je potrebno proširiti robotski sistem. Kako je osvetljenje mereno u samoj sceni ovakvo rešenje nije posebno praktično jer podrazumeva da robotska ruka najpre u scenu postavi senzor, izmeri osvetljenje, a tek potom pokrene modul robotske vizije. Kako bi se ovaj problem prevazišao može se potpuno izbeći dodavanje novog senzora i koristiti direktno podaci sa stereo kamere. Naime, podatke o osvetljenju scene vidimo sa histograma HSV slike, odnosno histograma nijanse, histograma zasićenja i histograma intenziteta (slika 4.27).



Slika 4.27 Histogrami HSV slike

Maksimalne vrednosti histograma, srednje vrednosti histograma i raspodele odnosno oblici histograma nam govore dovoljno o osvetljenju u sceni, pa se javlja ideja korišćenja ovih podataka u cilju optimalnog određivanja širine intervala praga.

Još jedna ideja koja se javlja kod segmentacije u povratnoj sprezi jeste optimalan izbor većeg broja inkrimenata i za koje funkcija promene konektivnosti ima lokalne minimume i spajanje tako segmentiranih binarnih objekata u jedan. Na ovaj način se uspešno mogu segmentisati neuniformno obojeni objekti. Predloženi algoritam podrazumeva prepoznavanje objekta na osnovu segmentirane mape dispariteta, nakon čega bi sistem za preporuku odabrao odgovarajuće lokalne minimume, tj. inkrimente i , čime bi se izdvojili objekti čijim spajanjem

se formira realna reprezentacija objekta koja omogućava tačnu 3D rekonstrukciju objekta u prostoru. Za formiranje sistema odlučivanja može se koristiti neuronska mreža, ali se podaci za obuku neuronske mreže ne mogu formirati direktno, već je potrebno koristeći neki metaheuristički algoritam, poput genetskog algoritma optimalno odrediti lokalne minimume na osnovu razlike između binarne slike formirane ručnom segmentacijom i binarne slike formirane predloženim inteligentnim adaptivnim algoritmom segmentacije.

Rezultati segmentacije bi se pri segmentaciji mape dispariteta mogli popraviti optimalnim odabirom veličine bloka za formiranje mape dispariteta. U daljem istraživanju bi jedna od interesantnih tema mogla da bude optimizacija veličine bloka za formiranje mape dispariteta, odnosno projektovanje inteligentnog adaptivnog sistema za određivanje veličine bloka.

5. Inteligentni sistemi za detekciju i prepoznavanje ljudi u robotskoj viziji

Kako je sve veći broj aplikacija robota u svakodnevnom životu, javlja se potreba za kohabitacijom robota i ljudi, te je neophodno da savremeni servisni roboti detektuju ljude i mogu da ih prate. Roboti izlaze iz proizvodnih hala, što zahteva razvoj sistema intuitivne interakcije između robota i čoveka.

5.1. Pregled istraživanja

Problem razvoja mobilne robotske platforme za praćenje ljudi se usled naglog napretka računarske tehnike i robotike javlja sve češće, a primena ovakve mobilne robotske platforme može da bude jako raznovrsna, od robotskih sistema za nadzor [72], preko robota koji se koriste u rehabilitacione svrhe, servisnih robota za pomoć starim osobama ili osobama sa invaliditetom, do robota koji treba da rade zajedno sa čovekom u raščišćavanju hazardnih i kontaminiranih sredina [73]. Problem razvoja ovakvog sistema može se podeliti na dva dela, tj. detekciju ljudi u sceni i praćenje ljudi [72]. Ovo poglavlje se bavi prvenstveno detekcijom ljudi u sceni.

Najčešće se javlja ideja prepoznavanja lica [74] primenom „Gaborovih talasića“ i skrivenog Markovljevog modela, dok je dodato i prepoznavanje boje odeće gornjeg dela tela kako bi se međusobno razlikovao veći broj ljudi u sceni. Lokalizaciju čoveka može omogućiti i jedna monokularna kamera [75], ali je najpre stereo kamerom i SLAM metodom (simultana lokalizacija i mapiranje prostora) formirana 3D mapa prostora, pa se tek onda monokularnom kamerom detektuju promene u okruženju i samim tim ljudi.

Javlja se i ideja kombinacije formiranja 3D mape okruženja stereo kamerom, u kojoj se naknadno traže lica, i ponovo se boja odeće koristi kako bi robot razlikovao veći broj ljudi u sceni [76]. Jasno je da za uspešnu lokalizaciju nije dovoljna monokularna kamera, već se koriste stereo sistemi kamera. Stereo sistemi ne mora da podrazumevaju dve identične monokularne

kamere, već je moguća i kombinacija infra-crvene kamere za daljinsko osmatranje (FIR) i kamere za dnevno snimanje [77]. Korišćenjem kombinacije ove dve kamere povećava se robusnost sistema vizije.

Sistemi kompjuterske vizije za prepoznavanje ljudi su doživeli ekspanziju pojavom sistema vizuelnih bezbednosnih sistema i automatske detekcije pešaka kod automobila. Ovakvi sistemi podrazumevaju i stereo sistem infra-crvenih kamera [78].

U sistemima robotske vizije se javljaju i ideje formiranja mape pokreta [79] iz koje se određuju karakteristike, koje u kombinaciji sa „mapama istaknutosti“ formiraju prostor pretrae klasifikatora. Primenom algoritama za redukovanje dimenzija karakteristika i klasifikatora zasnovanog na metodi potpornih vektora se detektuju ljudi bez obzira na količinu kretanja njihovog kretanja. Ovakvi sistemi su primenjivi u aplikacijama poput video nadzora, detekcije pešaka i robotskoj viziji.

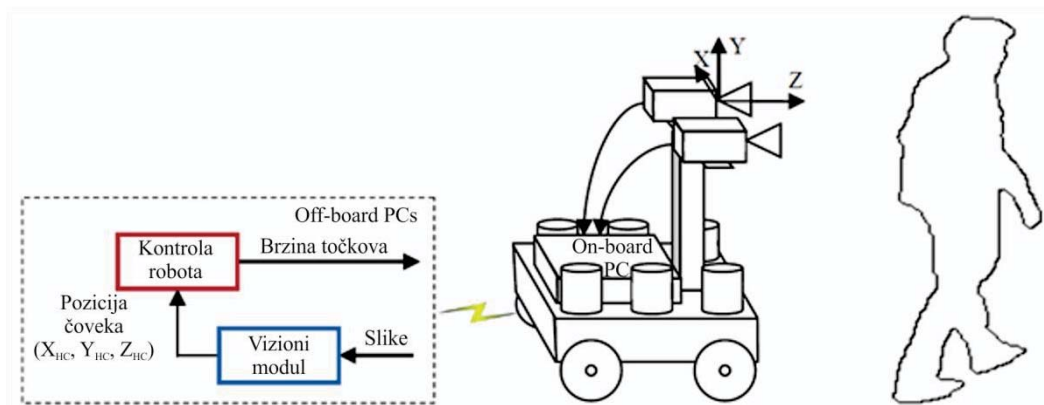
Stereo kamera koja snima osobu ispred interaktivne table [80] može služiti kao osnovni izvor informacija interaktivnog edukacionog sistema, pa se nakon formiranja mape dispariteta mogu izdvojiti karakteristike potrebne za detekciju i lokalizaciju čoveka. Ovakva aplikacija je pogodna i zbog same interaktivne table a koju može da se projektuje matrica koja služi za kalibraciju i rektifikaciju stereo kamere.

Osim ideje detekcije ljudske glave i boje garderobe koju čovek nosi [74,76,81], korišćenjem eliptičkog modela glave i metoda maksimalne približnosti (engl. Maximum likelihood) se uspešno mogu detektovati ljudi u sceni, a sistem je neosetljiv na promenu osvetljenja i rotaciju glava ljudi.

Kako lice čoveka nije uvek vidljivo u sceni jer robot često prati čoveka s leđa, ili u hazardnim sredinama čovek nosi zaštitno odelo ili masku, pristup detekcije ljudi na osnovu detekcije lica ne može da da zadovoljavajuće rezultate. Zbog toga je bilo potrebno razviti novi pristup rešavanju ovog problema. Deo rezultata istraživanja koji su predstavljeni u ovom poglavlju publikovani su u radovima na kojima je autor ove disertacije koautor [4,6,14,15].

5.2. Mobilna robotska platforma sa stereo kamerom za prepoznavanje i praćenje ljudi

U cilju adekvatne interakcije između robota i čoveka prvi zadatak koji mobilna robotska platforma treba da ispuni jeste da u sceni koju snima sistem vizije robot prepozna čoveka i da odredi distancu između čoveka i sebe. Šema arhitekture mobilne robotske platforme sa stereo kamerom kojoj je cilj praćenje čoveka prikazana je na slici 5.1.



Slika.5.1 Šema mobilne robotske platforme za praćenje ljudi [25]

Ova mobilna robotska platforma (slika 5.2) koja je testirana i korišćena u eksperimentalne svrhe razvijena je na Institutu za automatiku Univerziteta u Bremenu, a algoritmi za detekciju i praćenje ljudi razvijani su u okviru bilateralnog projekta na kome su učestvovali istraživači sa Mašinskog fakulteta Univerziteta u Nišu, među kojima je i autor ove teze.



Slika 5.2 Mobilna robotska platforma sa stereo kamerom za praćenje ljudi

Ova mobilna robotska platforma poseduje robotsku ruku sa 7 stepeni slobode kretanja, stereo kameru i laserski skener. Novina u sistemu jeste industrijski laserski skener SICK S3000 koji služi u bezbednosne svrhe i ima podesiv opseg zaštite koji može da varira od 4 do 7 metara. Za ovakav laserski skener postoje gotova softverska rešenja koja služe za prepoznavanje ruke, noge ili tela čoveka (slika 5.3).

Robotska platforma je opremljena Point Grey Bumblebee XB3 stereo kamerom, koja može da snima slike u rezoluciji od 1280x960 piksela, a zbog hardverskih ograničenja snimano je 12 stereo frejmova u sekundi.



Slika 5.3 Testiranje upravljačkih algoritama mobilne robotske platforme sa stereo kamerom za praćenje ljudi

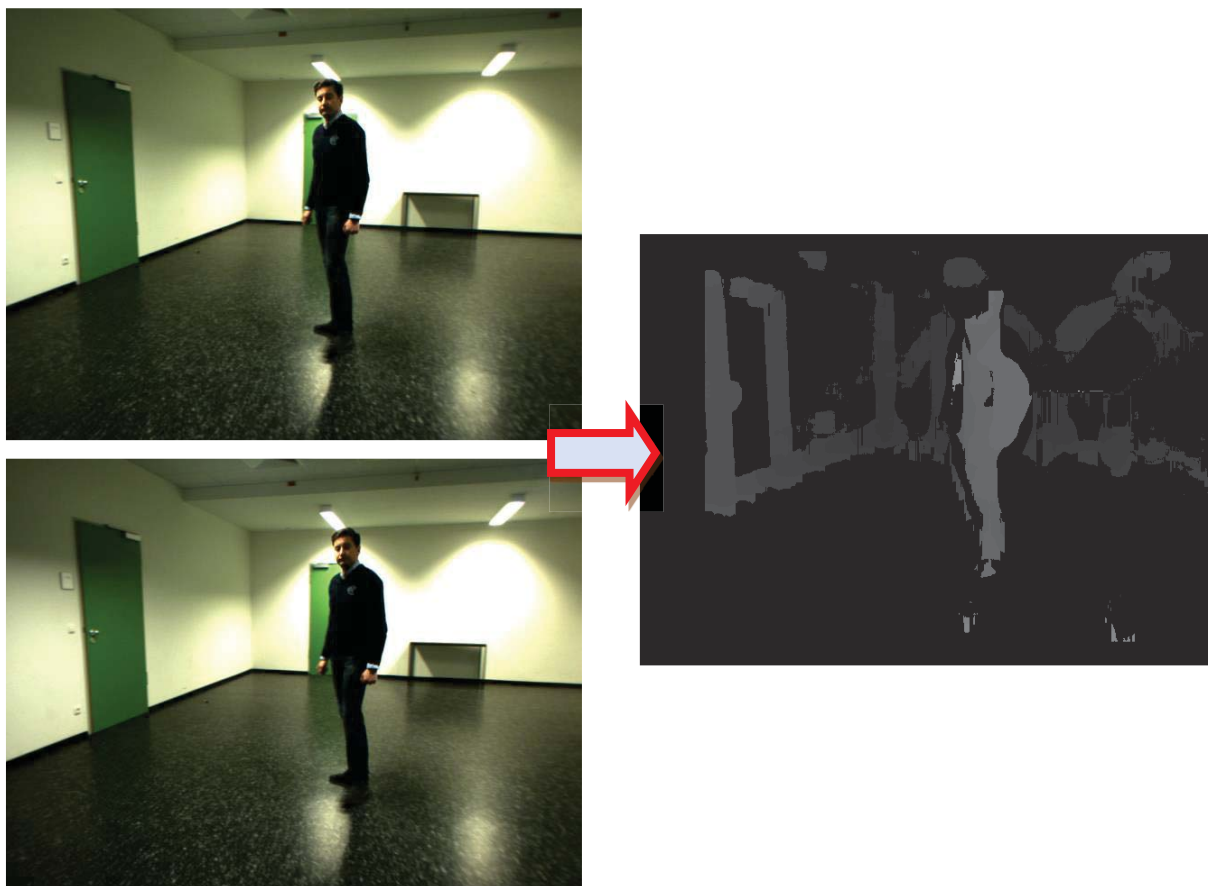
5.2.1. Segmentacija slike i izdvajanje karakteristika u cilju prepoznavanja ljudi

Kao algoritam segmentacije slike u sistemu vizije za prepoznavanje i praćenje čoveka se segmentacija po boji ne može razmatrati kao dovoljno robustan algoritam, jer namena ove mobilne robotske platforme nije da uvek posmatra čoveka spreda, već da ga prati, u kom slučaju je čovek okrenut leđima robotu. Osim toga, namena ovog robota jeste da pomaže čoveku u hazardnim sredinama, gde čovek nosi zaštitno odelo, pa traženje piksela u boji kože na slici

ne može da da zadovoljavajuće rezultate. Zbog toga se u ovom slučaju primenjuje segmentacija mape dispariteta, po algoritmu koji je predstavljen u odeljku 3.4.3. Kako se čovek kreće u sceni, tako se i morfologija segmentirane binarne slike menja, pa problem prepoznavanja ljudi u sceni kompleksniji u odnosu na probleme prepoznavanja objekata. Zbog promene morfologije nije potrebno u ovom slučaju vršiti konveksno popunjavanje, jer se time gube informacije o obliku koji je karakterističan za ljude. Osim toga, algoritam formiranja mape dispariteta je proširen modulom za uklanjanje pozadine, odnosno podloge na kojoj stoji čovek. Ovo je bilo neophodno učiniti kako kasnije tokom segmentacije ne bi došlo do spajanja podloge, čoveka i objekata u neposrednoj blizini čoveka u jedan objekat, čime bi detekcija čoveka na osnovu segmentirane binarne slike bila praktično nemoguća. Ovaj algoritam predstavljen je u radu [25]. Rezultat formiranja mape dispariteta uz uklanjanje podloge predstavljeno je na slici 5.4.

Sada se vrši segmentacija mape dispariteta po već poznatom algoritmu i formiraju se oblasti od interesa, za koje je potrebno odrediti karakteristike na osnovu kojih se može izvršiti klasifikacija (slika 5.5).

Za sve oblasti od interesa binarne segmentirane slike mape dispariteta moguće je odrediti sve karakteristike predstavljene u odeljku 3.4.5. ove teze. Ipak, konektivnost piksela nije referentna vrednost pri klasifikaciji čoveka, kako zbog promene morfologije čoveka tokom kretanja, tako i zbog nesavršenosti algoritama formiranja mape dispariteta i segmentacije mape dispariteta. U tabeli 4 predstavljene su ekstrahovane karakteristike binarnih segmentiranih slika ljudi u različitim frejmovima jednog video klipa snimljenog stereo kamerom, dok su u tabeli 5 predstavljene karakteristike segmentiranih objekata koji nisu ljudi, a javili su se u sceni. Karakteristike koje su određivane su stvarna visina i širina objekta (u tabelama 5 i 6 predstavljena je samo proporcionalnost) i prvih 7 Hu-ovih invarijantnih momenata.



Slika 5.4 Formiranje mape dispariteta na osnovu stereo slike korišćene za prepoznavanje ljudi



Slika 5.5 Segmentacija mape dispariteta za prepoznavanje ljudi

Ove karakteristike su korišćene za detekciju ljudi u sceni, odnosno klasifikaciju.

Tabela 5 Ekstrahovane karakteristike ljudi u sceni

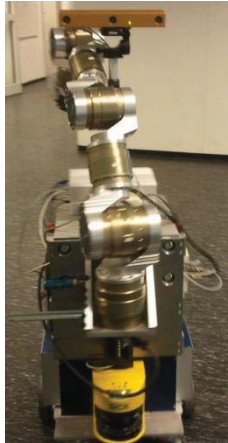





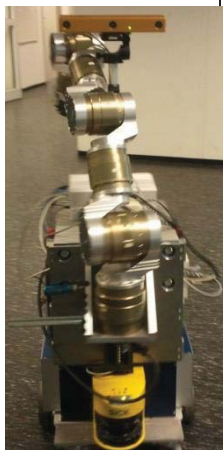





					
	Čovek 1	Čovek 2	Čovek 3	Čovek 4	Čovek 5
p	2.82	2.70	2.57	2.51	2.39
I_1	0.33473	0.3476	0.322443	0.321882	0.31998
I_2	0.0557013	0.0669781	0.0573314	0.0571661	0.0563461
I_3	0.00126679	0.00038232	0.00077274	0.000960961	0.392684
I_4	8.40634	0.000108095	5.76663	5.48374	0.00021946
I_5	-0.00073858	-5.93832	0.00090811	0.000445	0.00016667
I_6	-1.23022	2.31904	1.12871	0.882666	5.20611
I_7	0.00078271	0.00185407	0.00100215	0.00035926	0.00064111

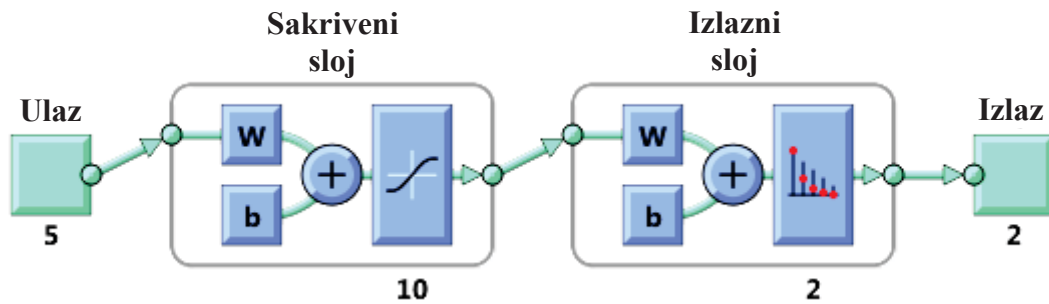
Tabela 6 Ekstrahirane karakteristike objekata koji nisu ljudi u sceni

					
	Objekat 1	Objekat 2	Objekat 3	Objekat 4	Objekat 5
p	2.11	0.92	0.23	2.01	2.17
I_1	0.475794	0.694843	0.778063	0.397016	1.02095
I_2	0.157128	0.107268	0.556262	0.0585865	0.715118
I_3	0.0381782	0.161428	0.00258805	0.0237106	0.521932
I_4	0.0318163	0.0140506	0.00133926	0.0107419	0.403269
I_5	-0.0676176	-0.0639997	0.00291517	-0.049058	0.860055
I_6	0.0126046	-0.0033753	0.000945317	0.0024049	0.32578
I_7	-0.0000191	0.00008128	0.00000035023	-0.000063	0.0271802

5.2.2. Klasifikacija u cilju prepoznavanja ljudi

U cilju klasifikacije objekata u cilju detekcije ljudi predložen je klasifikator zasnovan na neuronskoj mreži za prepoznavanje obrazaca sa deset neurona u skrivenom sloju, čija je arhitektura prikazana na slici 5.6. U pitanju je dvoslojna mreža sa prostiranjem signala unapred, gde su neuroni u skrivenom sloju sa sigmoidnom aktivacionom funkcijom, dok su u izlaznom sloju sa tzv. "softmax" funkcijom. Na osnovu ekspertskog znanja kolega iz Bremena, kao ulaz klasifikatora predloženo je pet karakteristika binarnog segmentiranog objekta, a to su visina h i širina objekta w i prva 3 Hu-ova invarijantna momenta:

$$\begin{aligned}
 I_1 &= \eta_{02} + \eta_{20} \\
 I_2 &= (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2 \\
 I_3 &= (\eta_{30} - \eta_{12})^2 + (3\eta_{21} - \eta_{03})^2.
 \end{aligned}
 \tag{5.1}$$



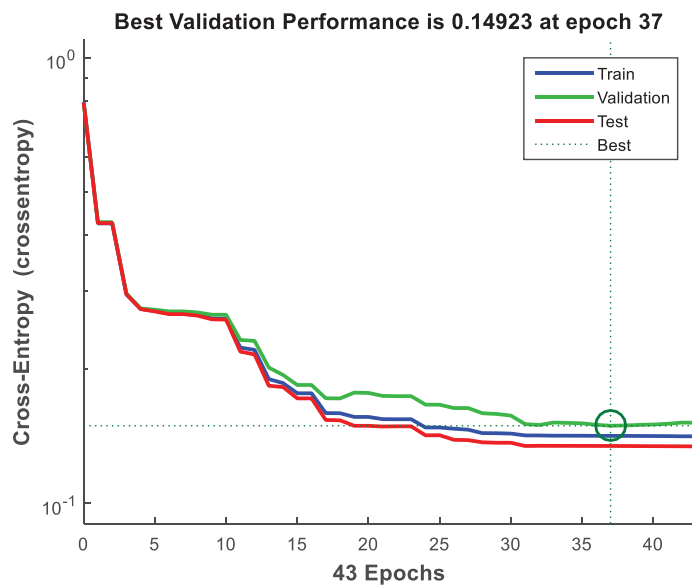
Slika 5.6 Arhitektura klasifikatora zasnovanog na neuronskoj mreži za detekciju ljudi

Obučavanje, validacija i testiranje vršeno je sa 3474 objekata kojima su ručno određene labele. Slučajno je izabrano 70% setova podataka za obučavanje, a po 15% slučajno izabranih setova je korišćeno za validaciju i testiranje. Za obučavanje je korišćena povratna propagacija greške i metod najstrmijeg spusta. Rezultati obučavanja, validacije i testiranja mreže mogu se videti na konfuzionim matricama prikazanim na slici 5.7, odnosno na slici 5.8 gde se može videti minimizacija greške kroz iteracije obučavanja neuronske mreže.



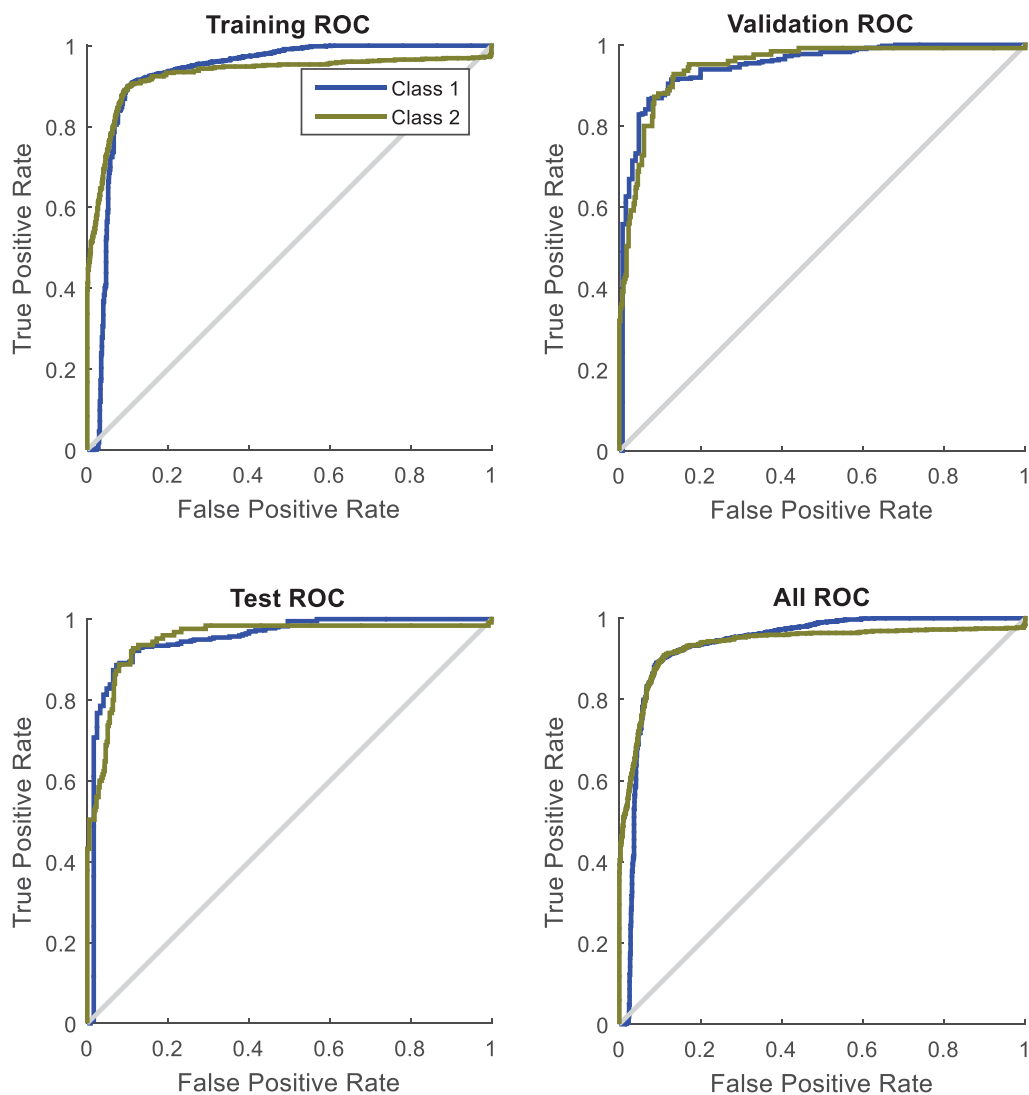
Slika 5.7 Konfuzione matrice klasifikatora

Obučavanje predloženog klasifikatora izvršeno je kroz 90 iteracija, a rezultati klasifikacije su zadovoljavajući. Tačnost obučenog klasifikatora je preko 89%, a može se dodatno povećati ubacivanjem dopunskih neurona u skriveni sloj. Izlaz iz klasifikatora je verovatnoća da objekat koji ispitujemo i čije smo karakteristike ekstrahovali pripada klasi "čovjek" i "objekat koji nije čovek".



Slika 5.8 Minimizacija greške kroz iteracije obučavanja neuronske mreže

Do greške pri klasifikaciji dolazi zbog toga što je moguće da se u sceni jave objekti koji su nakon segmentacije dosta slični binarnim slikama čoveka u hodu ili je omanuo algoritam za segmentaciju mape dispariteta. Druga nepovoljnost jeste da objekti koji nisu ljudi mogu biti najrazličitijeg oblika, pa bi povećanje broja objekata koji pripadaju ovoj klasi pri treniranju dodatno mogao popraviti rezultate klasifikacije. Osnovni zadatak ove mobilne robotske platforme jeste praćenje ljudi, pa je prihvatljivo da mobilna robotska platforma u pojedinim frejmovima ne detektuje ljude, jer se ovaj problem može prevazići uvođenjem odgovarajućeg filtera koji će vršiti estimaciju i predikciju položaja čoveka u sceni. Ovakvo rešenje predloženo je u poglavlju 7 ove teze.



Slika 5.9 Operativne karakteristike prijemnika

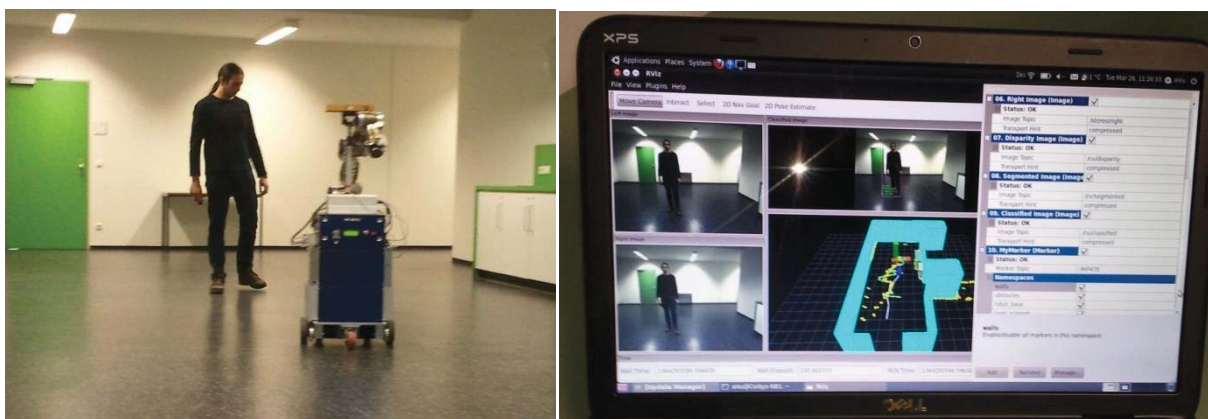
5.2.3. Verifikacija rezultata i integracija senzora

Predloženi klasifikator zasnovan na neuronskoj mreži daje dobre rezultate, pa je implementacijom ovakvog klasifikatora u sistem vizije moguće detektovati čoveka i pratiti ga (slika 5.10). Performanse mobilne robotske platforme dodatno se mogu popraviti uvođenjem algoritma za ocenu i predviđanje položaja čoveka u sceni.



Slika 5.10 Detekcija čoveka u sceni primenom neuronske mreže

Kako je već napomenuto, mobilna robotska platforma korišćena u ovom istraživanju opremljena je laserskim skenerom. Uloga ovog senzora je prvenstveno bezbednosna, jer sprečava da dođe do kolizije robota i čoveka i robota i prepreka uopšte. Međutim, podaci dobijeni sa ovog senzora mogu se koristiti i za mapiranje prostora (slika 5.11), kao i za verifikaciju rezultata detekcije ljudi i praćenja.



a)

b)

Slika 5.11 Mobilna robotska platforma prati čoveka a) i prikaz modula robotske vizije koji se u tom trenutku izvršava na udaljenom računaru b)

U nekim istraživanjima [83,84,85] se ovakvi senzori koriste za prepoznavanje objekata u situacijama kada sistem robotske vizije zakaže, a javlja se i ideja prepoznavanja ljudi, jer laserski skener može da detektuje noge ljudi koji se nalaze ispred njega. Integracija podataka dobijenih ovim senzorom i senzora robotske vizije [85] može se formirati upravljački sistem koji čini mobilnog robota izuzetno autonomnim. Podaci dobijeni laserskim skenerom se, prema tome, mogu koristiti kao dodatni izvor informacija u sistemu za prepoznavanje i praćenje ljudi. Tako se u radu [86] podaci dobijeni sa laserskog senzora koriste za detekciju nogu, dok se podaci dobijeni sa kamere koriste za detekciju lica, čime se povećava robusnost sistema.

5.3. Optimalan izbor ulaznih veličina klasifikatora

Tokom istraživanja se kao kritičan pokazao izbor odgovarajućeg vektora karakteristika koji je ulazni vektor klasifikatora. Moguće je koristiti karakteristike koje nedovoljno dobro karakterišu klasu ili proces uopšte, ili s druge strane koristiti kao elemente vektora ulaza karakteristike koje ni na koji način ne utiču na tačnost klasifikatora, ili uopšteno sistema mašinskog učenja.

Ovaj problem je danas jako aktuelan, jer i pored svih svojih dobrih osobina, *black-box* modeli nas mogu dovesti u zabludu po pitanju uticaja određenih karakteristika na sam sistem. Jedan od prvih radovima koji se bave ovim problemom jeste [87]. Metode za optimalan izbor vektora ulaznih karakteristika se uglavnom zasnivaju na određivanju uticajnih parametara i korelacija među samim karakteristikama [88], ali je u poslednje vreme sve veći broj radova koji predlažu primenu metaheurističkih metoda optimizacije za izbor optimalnog vektora ulaznih karakteristika klasifikatora, odnosno sistema mašinskog učenja [89,90,91]. U ovom odeljku biće predstavljena optimizaciona metoda zasnovana na genetskom algoritmima, kojom se formira optimalni vektor karakteristika koji je ulaz u klasifikator zasnovan na neuronskoj mreži.

5.3.1. Kriterijum za odabir ulaznih veličina klasifikatora

Ideja koja je ovde predstavljena rezultat je analize istraživanja ekspertskeg tima iz Bremena. Naime, oni su tokom istraživanja formirali sedam test slučajeva kako bi odredili optimalni set ulaznog vektora karakteristika. Test slučajevi izbora karakteristika predstavljeni su u tabeli 7.

Tabela 7 Test slučajevi različitih vektora karakteristika korišćenih za obučavanje klasifikatora

	Vektor karakteristika
Test 1	$I_1, I_2, I_3, I_4, I_5, I_6, I_7$
Test 2	I_1, I_2, I_3
Test 3	I_1, I_3, I_5
Test 4	Širina w , visina h , I_1, I_2, I_3
Test 5	Širina, visina, I_1, I_2, I_3 (samo glava)
Test 6	Vrednost 640 piksela (slika je skalirana na 16x40 piksela)
Test 7	Širina w , visina h , Vrednost 640 piksela

Nakon obučavanja i upoređenja tačnosti klasifikatora ustanovljeno je da najbolje rezultate daju klasifikatori čiji su ulazi definisani testom 4, 6 i 7. Zbog svoje relativne jednostavnosti se u radovima ove grupe pri projektovanju klasifikatora koriste kao ulazni vektori visina, širina i prva 3 Hu-ova momenta.

Međutim, ovi testovi ne pokrivaju sve moguće kombinacije formiranja ulaznog vektora, pa se ne može sa sigurnošću tvrditi da je predloženi vektor karakteristika optimalan. Parametri koji su dostupni u istraživanju su širina, visina i 7 Hu-ovih invarijantnih momenata, i vektor svih karakteristika je $x = [w, h, I_1, I_2, I_3, I_4, I_5, I_6, I_7]$, pa ima 512 mogućih kombinacija potencijalnog vektora karakteristika.

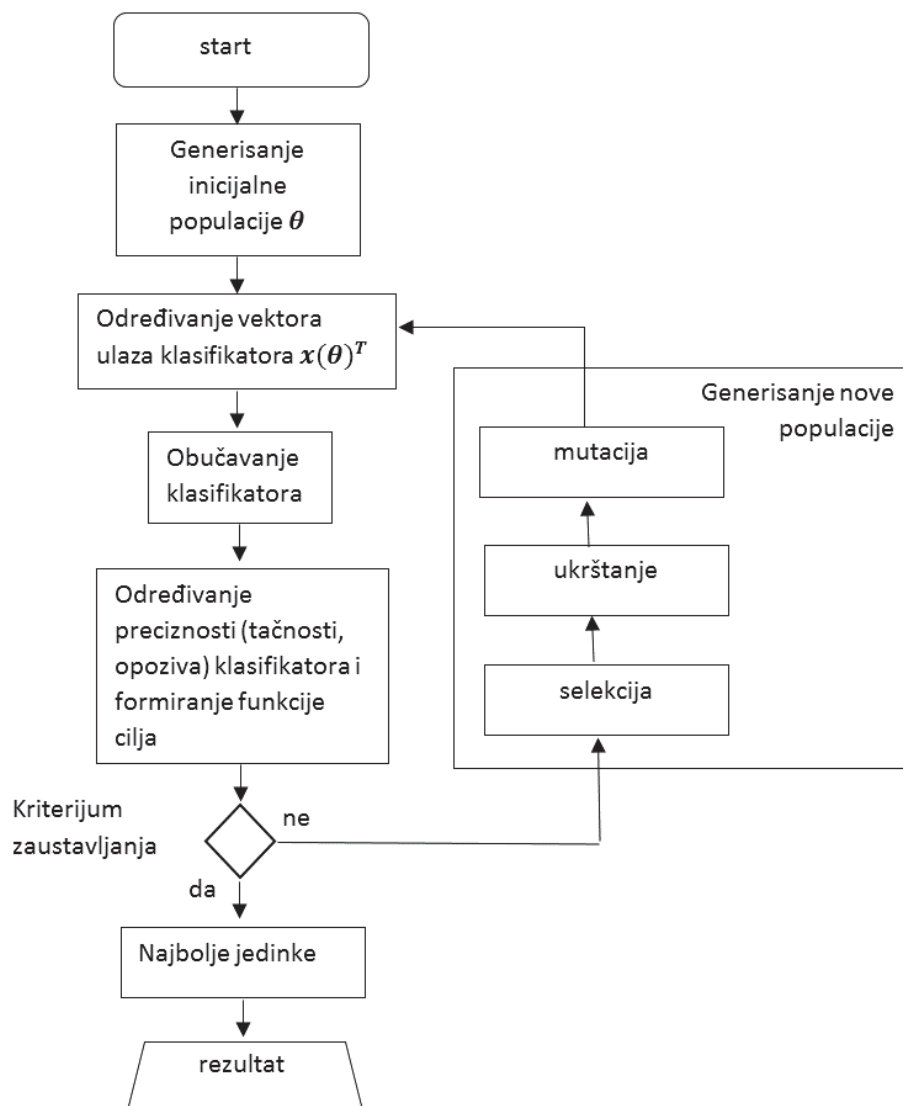
Kako bi se odredilo optimalno rešenje nije dovoljno samo koristiti tačnost klasifikatora, već u obzir treba uzeti i broj karakteristika, jer što je veći broj ulaznih karakteristika ceo algoritam vizije se sporije izvršava. Dakle problem se svodi na višekriterijumsku optimizaciju, a cilj je naći vektor karakteristika sa najmanjim brojem karakteristika koji će dati najveću tačnost, odnosno dva međusobno suprotstavljena uslova.

5.3.2. Optimalan izbor ulaznog vektora karakteristika neuronske mreže primenom genetskih algoritama

Ovaj problem može se rešiti primenom genetskih algoritama. Potrebno je pronaći binarni niz $\theta = [\theta_1, \theta_2, \theta_3, \theta_4, \theta_5, \theta_6, \theta_7, \theta_8, \theta_9]$ takav da se vektor karakteristika može formirati kao:

$$x(\theta)^T = \theta \cdot I(9) \cdot x^T, \quad 5.2.$$

gde je $I(9)$ jedinična matrica dimenzija 9x9. Šema funkcionisanja genetskog algoritma predstavljena je na slici 5.12.



Slika 5.12 Šema genetskog algoritma za optimalno određivanje vektora ulaza klasifikatora

Najveći je problem zapravo definisati funkciju cilja. U tu svrhu mogu nam poslužiti merljive performanse klasifikatora poput tačnosti, preciznosti ili opoziva. Osim toga, broj elemenata vektora ulaza klasifikatora igra bitnu ulogu u brzini izvršavanja algoritma za segmentaciju. U predloženom rešenju se problem definisanja funkcije cilja svodi na linearnu kombinaciju dva cilja, a to su tačnost koju ima neuronska mreža i broj elemenata koji čine vektor karakteristika:

$$f(\boldsymbol{\theta}) = a_1 \cdot (1 - NNtAcc(\boldsymbol{\theta})) + a_2 \sum_{i=1}^9 \theta_i, \quad 5.3.$$

gde su a_1 i a_2 težinski faktori, a $NNtAcc(\boldsymbol{\theta})$ je normalizovana tačnost koja se postiže tokom obuke neuronske mreže.

Potrebno je iskustveno odrediti težinske koeficijente i potom genetskim algoritmom odrediti minimum funkcije cilja $f(\boldsymbol{\theta})$. Osim tačnosti klasifikatora u zavisnosti od namene mogu se koristiti i veličine poput preciznosti, opoziva ili F1 skora, a mogu se uvesti i ograničenja (npr. da je minimalni broj elemenata vektora karakteristika 3 i sl.).

Osim toga, svaka od karakteristika zahteva različito vreme izračunavanja od strane računara, pa je moguće svakom elementu vektora ulaza ponaosob dodeliti težinske koeficijente, čime bi se prednost dala karakteristikama za čije izračunavanje je potrebno manje vremena.

Pravci budućih istraživanja podrazumevaju dalji razvoj i uopštenje metodologije određivanja optimalnog ulaznog vektora karakteristika kod sistema mašinskog učenja.

6. Inteligentni robotski sistem za detekciju i praćenje ljudi na osnovu slike dobijene termovizijskom kamerom

I pored velikog broja prednosti, konvencionalni sistemi vizije imaju i određene nedostatke koji se ogledaju u maloj robusnosti sistema vizije u odnosu na različite spoljašnje uticaje, kao što su promena osvetljenja pri akviziciji slike objekta koji se analizira, kolizija objekata u sceni, orijentacija objekta koji se posmatra, refleksija i sl. S obzirom da aktuelna istraživanja u svetu ukazuju na veliki broj mogućnosti daljeg istraživanja razvoja i primene inteligentnih sistema robotske vizije za prepoznavanje objekata i odgovarajuću interakciju robota sa njima, u ovom poglavlju predstavljena je mobilna robotska platforma sa termovizijskom kamerom i implementiranim softverskim i upravljačkim rešenja kojima se mogu prevazići neki od predočenih problema.

Postoje neki opšteprihvaćeni algoritmi koji se fokusiraju na praćenje ljudi na osnovu podataka dobijenih sa nekih termičkih senzora, ali takvi algoritmi još uvek nisu zaživeli u robotskim sistemima. Inspiraciju za ovakve algoritme možemo naći u prirodi gde mnoge životinje i insekti koriste termoreceptore da bi se hranili, lovili i opstajali. U sistemu kompjuterske vizije se polazi od pretpostavke da je objekat od interesa topliji (ili hladniji) od okruženja, što omogućava tehničkom sistemu da vidi u potpunom mraku ili vidi boje izvan vidljivog spektra. Termalne infracrvene kamere detektuju relativne razlike u količini termičke energije koja je emitovana, odnosno reflektovana od objekata u sceni. Dok god su termičke karakteristike objekta neznatno različite u odnosu na zračenje pozadine, uočava se kontrast između objekta i okruženja.

Kako bi odredili položaj čoveka na osnovu informacija prikupljenih sa termovizijske kamere i omogućili supervizijsko upravljanje robotskom platformom, neophodno je pre-procesirati i segmentisati termičku sliku. Ideja koja stoji iza svih algoritama vezanih za segmentaciju termičke slike jeste određivanje graničnog praga kako bi se izdvojila oblast od interesa (engl. Region of interest - ROI). Ipak, kako bi se postigla odgovarajuća tačnost i

robusnost sistema, optimalno određivanje graničnog praga segmentacije nije jednostavan zadatak i često se koristi određivanje graničnog praga u više nivoa ili adaptivno odrađivanje graničnog praga koji su zasnovani na neuronskim mrežama ili fazi logici.

Sistem koji je razvijen koristi različite metode računarske inteligencije kako bi se prevazišli problemi koje donosi uvođenje sistema termovizije u mobilnu robotsku platformu. Robot na osnovu informacija dobijenih sa termovizijske kamere detektuje čoveka u scenariju i određuje njegove koordinate, koje se potom šalju do kontrolera kretanja koji upravlja radom motora koji pokreću točkove robotske platforme. Sistem vizije detektuje prisustvo čoveka (ljudi) u okruženju robota na osnovu termovizijske slike koristeći algoritme za detekciju i praćenje ljudi koji su razvijeni i predstavljeni u radovima autora ove teze [1,2,9,10,12,19], a mogu naći primenu u velikom broju aplikacija.

U okviru ovog istraživanja neophodno je bilo razviti veći broj modula koji po sadržaju obuhvataju više aktuelnih naučnoistraživačkih pravaca i odnose se na:

- Razvoj robotske mobilne platforme sa termovizijskom kamerom;
- Razvoj softvera za robusnu segmentaciju i obradu slike dobijene termovizijskom kamerom zasnovanog na genetskim algoritmima;
- Razvoj softvera za detekciju i praćenje ljudi;
- Razvoj novog integrisanog inteligentnog sistema upravljanja mobilnim robotom.

Istraživanje je vršeno u Laboratoriji za upravljanje sistemima Mašinskog fakulteta Univerziteta u Nišu u okviru bilateralnih projekata Ministarstva prosvete, nauke i tehnološkog razvoja i DAAD-a koji se realizuju u okviru programa „PPP Serbien“ pod nazivom “A novel approach to human detection and tracking in robotics” i „Robust action recognition for human-robot synergy” (2015-2016). Određene rezultate moguće je primeniti kod sličnih mobilnih robotskih platformi, ili u sistemima kompjuterske vizije, poput sistema za video nadzor i sistema za bezbednost u automobilima.

6.1. Pregled istraživanja

Većina sistema za prepoznavanje ljudi na osnovu vizije razmatra nemobilne aplikacije, tj. sisteme za praćenje ili verifikaciju identiteta, gde se prepoznavanje osoba može lako rešiti metodama uklanjanja pozadine koje se ne mogu primeniti na mobilne sisteme. Postojeći sistemi za prepoznavanje ljudi na mobilnim robotima uobičajeno koriste algoritme za prepoznavanje boje kože i lica. Međutim, ovi pristupi pretpostavljaju frontalnu poziciju osobe i zahtevaju da

osoba nije previše daleko. Termovizija pomaže da se neki od problema koji se tiču senzora vizije u boji savladaju, jer ljudi imaju različit termalni profil u poređenju s neživim objektima.

Jedna od ideja za prepoznavanje kretanja ljudi u sceni na osnovu termovizijske kamere jeste praćenje optičkog protoka u cilju uklanjanja pozadine i segmentacije ljudi u sceni [92]. Ovakvim algoritmom moguće je da autonomna mobilna robotska platforma detektuje i prati provalnika. I pored toga što je ova metoda zasnovana na sličnim principima kao i uklanjanje pozadine, daje dobre rezultate i kada mobilna platforma stoji i kada se kreće.

Pri praćenju ljudi mobilnom robotskom platformom sa termovizijskom kamerom se zbog specifičnosti aplikacije u cilju detekcije ljudi, tj. klasifikacije, mogu koristiti i karakteristike koje se ne sreću kod detekcije ljudi monokularnom ili stereo kamerom. Tako se u radu [93] koristi eliptički model kontura i pravougaone karakteristike intenziteta kako bi se AdaBoost klasifikatorom detektovali ljudi u sceni.

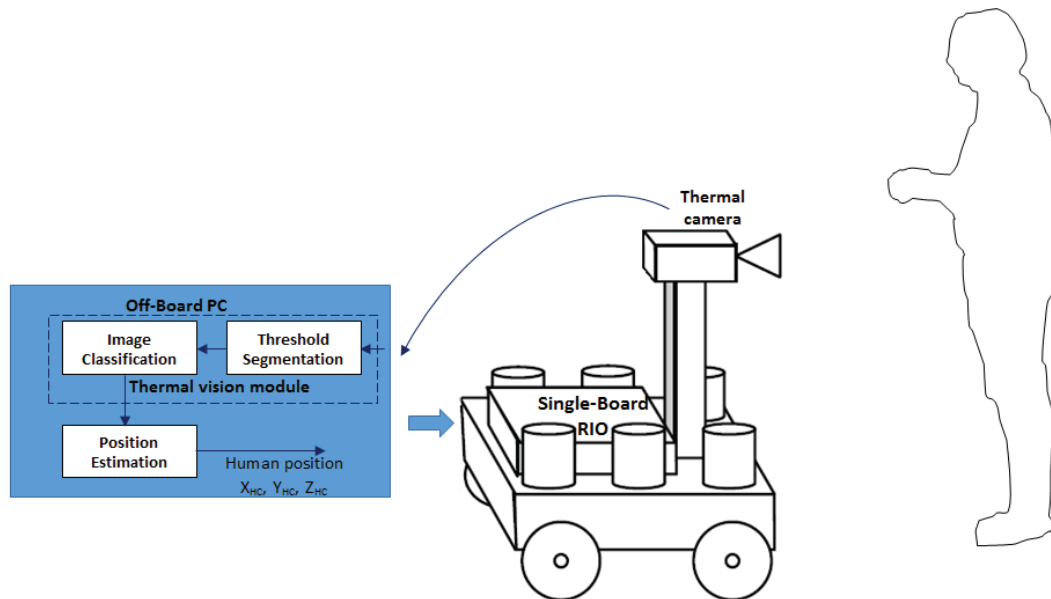
Kombinacija termovizijske kamere i senzora vizije omogućava mobilnim robotskim platformama mapiranje prostora i prepoznavanje objekata [94]. Posebno je značajna robusnost ovakvog sistema sa kombinacijom ovih senzora u raznim scenama koje se mogu javiti u hazardnim situacijama, poput noćne scene, zadimljene scene ili požara [95]. Javlja se i ideja uvođenja fazi sistema [96] za fuziju informacija sa senzora vizije i termovizijske kamere, a u cilju pronalaženja i praćenja meta bespilotnim letelicama.

Pri praćenju ljudi mobilnom robotskom platformom sa termovizijskom kamerom se primenom odgovarajućih filtara kojima bi se vršila estimacija i predikcija položaja ljudi u sceni dodatno mogao unaprediti sam sistem [93,96].

6.2. Mobilna robotska platforma DaNI sa Termovizijskom kamerom

Za razvoj pouzdanog, računski inteligentnog prepoznavanja objekta u robotskoj viziji u predloženom tehničkom rešenju je korišćen National Instruments Robotics Starter Kit 1.0 – DaNI robot opremljen termovizijskom kamerom FLIR E50. Termovizijska kamera nije mogla direktno da se poveže sa upravljačkom pločom mobilnog robota (Single Board RIO), niti je sofisticirani softver za segmentaciju i klasifikaciju direktno mogao biti prebačen na ploču, pa je dodatni personalni računar korišćen za obradu slike dobijene termovizijskom kamerom i viši nivo upravljanja. Funkcionalna šema robotske platforme sa termovizijskom kamerom za praćenje ljudi prikazana je na slici 6.1.

Slika dobijena termovizijskom kamerom je nakon pre-procesiranja segmentirana, a potom modul za klasifikaciju određuje da li je segmentirani objekat čovek ili ne. Algoritmi za segmentaciju i klasifikaciju su dva glavna dela termovizijskog modula koji je deo sofisticiranog upravljanja najvišeg nivoa. Nakon što je čovek detektovan, jednostavan algoritam koji uzima u prosečnu visinu čoveka kao referentnu vrednost ocenjuje relativni položaj čoveka u odnosu na robota. Koordinate položaja čoveka se potom šalju na niži upravljački nivo, gde PID kontroleri upravljaju elektromotorima koji pogone točkove robotske platforme.



Slika 6.1. Šema robotske platforme sa termovizijskom kamerom za praćenje ljudi

NI Robotics Starter Kit 1.0, poznat kao DaNI robot (slika 4.1) je mobilna robotska platforma koja obuhvata senzore, motore, i NI Single-Board RIO hardver za integrisani sistem upravljanja. DaNI je robot sa 4 točka koji je pogonjen sa dva motora i opremljen ultrazvučnim senzorom za merenje daljine.

Integrisani sistem upravljanja i akvizicije NI sbRIO-9632 integriše procesor koji radi u realnom vremenu i ulazno-izlazne uređaje na jednostavnoj štampanoj ploči. Karakteriše se industrijskim procesorom od 400 MHz, 2M Xilinx Spartan FPGA, 110 3.3 V (5 V tolerancije/TTL kompatibilan) digitalnom U/I linijom, 32/16 diferencijalnim 16-bitnim analognim ulaznim kanalima na 250 kS/s, i sa četiri 16-bitna analogna izlazna kanala na 100 kS/s.

DaNI ima tri konektora za proširenje U/I portova korišćenjem NI C rednih U/I modula. SbRIO-9632 obezbeđuje opseg radne temperature od -20 do 55°C, i uključuje od 19 do 30 VDC

ulazni opseg dovoda energije, 128 MB DRAM za integrisani rad, i 256 MB memorije za skladištenje programa i podataka za logovanje.

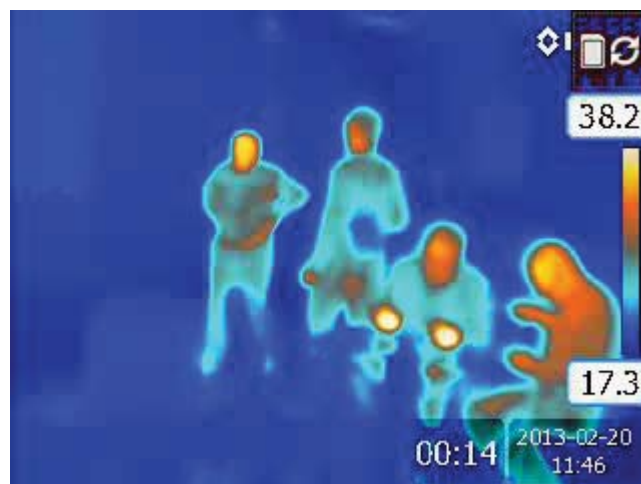
Mobilna robotska platforma ima ugrađen 10/100 Mbit/s Ethernet port koji može da se koristi za programsku komunikaciju preko mreže i ugrađene Web (HTTP) i File (FTP) servere. Programiranje samog uređaja sbRIO-9632 je realizovano preko LabView grafičkog razvojnog okruženja. Procesor u realnom vremenu upravlja LabView Real-Time modulom na Wind River VxWorks operativnom sistemu u realnom vremenu (RTOS) za ekstremnu pouzdanost i određenost. LabVIEW sadrži ugrađene drajvere i API-je za upravljanjem podacima za prenos između FPGA procesora u realnom vremenu.

Termovizijske kamere detektuju zračenje u infracrvenom opsegu elektromagnetnog spektra (oko 9–14 μm) i proizvode slike takvog zračenja koje nazivamo termogramima. Kako se infracrveno zračenje emituje od strane svih objekata iznad apsolutne nule a u skladu sa zakonom o zračenju crnog tela, termovizija omogućava da se vidi neko okruženje sa ili bez vidljive iluminacije. Količina zračenja emitovanog od strane objekata se povećava sa temperaturom. Zato, termovizija dozvoljava da se vide varijacije u temperaturi. Kada se gleda kroz termovizijsku kameru, topli objekti su značajno naznačeni u odnosu na hladniju pozadinu [97]. Ljudi i toplokrvne životinje postaju lako vidljivi u odnosu na okruženje, i preko noći i preko dana.

Eksperimentalna mobilna robotska platforma je opremljena nizom senzora uključujući i termovizijsku kameru FLIR E50 prikazanu na slici 6.2 a). Kamera može detektovati infracrveno zračenje i konvertovati ovu informaciju u sliku gde svaki piksel odgovara vrednosti temperature (slika 6.2 b).



a)

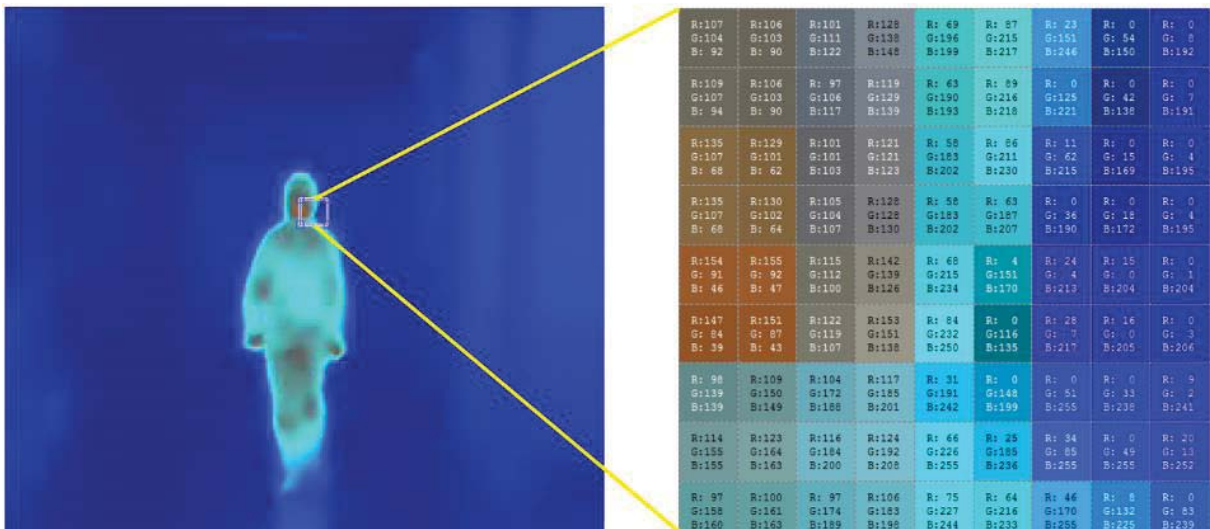


b)

Slika 6.2 FLIR E50 termovizijska kamera (a) i termovizijska slika (b)

Rezolucija termovizijske kamere je 240×180 piksela, tako da je ukupan broj piksela 43,200. Termovizijska osetljivost kamere je manja od 0.05°C a tačnost čitanja je $\pm 2^{\circ}\text{C}$ ili $\pm 2\%$ unutar temperaturnog opsega od -20°C do 650°C . Učestalost frejmova je 60 Hz a operativno vreme 4 sata. Masa kamere je 0.825 kg što nije zanemarljivo ali ne utiče značajno na dinamiku mobilne robotske platforme

Robot radi u zatvorenom prostoru bez posebnih ograničenja. Ljudi koji su učestvovali u eksperimentu su zamoljeni da hodaju ispred robota dok ne prepozna i realizuje dva različita autonomna ponašanja: prepoznavanje, praćenje i obilazak prepreka i praćenje čoveka. U našim eksperimentima vidljivi opseg slike je ekvivalentan opsegu temperature od 17.3 do 38.2°C pri čemu je izabrana jedna o šest predefinisanih paleta boja (slika 6.2, 6.3).



Slika 6.3. Termovizijska slika i vrednosti piksela za izabrano područje

6.3. Segmentacija termovizijske slike i određivanje karakteristika segmentiranog objekta

Da bi prepoznali i pratili osobu na termovizijskoj slici tj. nizu frejmova, potrebno je segmentisati termovizijsku sliku i za to koristimo napredne algoritme iz domena veštačke inteligencije u cilju odgovarajućeg određivanja graničnog praga segmentacije [2,12]. U istraživanju su kao inteligentni alati koji pomažu segmentaciju korišćeni genetski algoritmi [1,10,19] i fazi logika [9].

Termovizijska kamera može da detektuje infracrveno zračenje i konvertuje informaciju u sliku gde svaki piksel odgovara vrednosti temperature, pri čemu na crno-beljoj slici tamnije nijanse odgovaraju nižoj temperaturi dok oblasti koje čine svetliji pikseli odgovaraju regionima sa višom temperaturom.

6.3.1. Fazi segmentacija

U realnom scenariju na uspešno rešavanje problema određivanja položaja čoveka na osnovu termovizijske slike utiču razni faktori, jer se ispred osobe može naći prepreka koja na slici zaklanja neki deo čoveka, u slici se dve osobe mogu preklopiti ako prođu jedna pored druge, a vrlo često dolazi i do refleksije. Osim toga potrebno je u obzir uzeti i šum koji se javlja u termovizijskoj slici. Zbog svega navedenog se javlja ideja uvođenja alata iz domena računarske inteligencije koji bi omogućili uspešnu klasifikaciju i prepoznavanje ljudi. Kod termovizijske slike koja koristi paletu koja se kreće od plave ka žutoj (slika 6.2) i sa opsegom temperature od 15 do 40 stepeni, vrlo je lako da uočiti ljude na slici. S druge strane, računaru ovaj zadatak nije trivijalan. Kako je teško definisati jasnu granicu, tj. prag, kojim bi se izdvojio čovek sa slike, ideja primene fazi logike se nameće kao prirodno rešenje ovog problema. Klasična teorija skupova (takozvanih *crisp* skupova) polazi od stava da neki element x iz univerzalnog skupa X pripada ili ne pripada skupu A koji je definisan na skupu X . Teorija fazi skupova uvodi fundamentalno novi pojam, kontinualnu funkciju pripadnosti $\mu_A(x)$. Ova funkcija pokazuje koliko ispunjava uslov pripadnosti skupu A . U teoriji *crisp* skupova ona može da ima jednu od dve vrednosti, 0 i 1, tj. element pripada ili ne pripada skupu. U teoriji fazi skupova funkcija pripadnosti može da ima bilo koju vrednost između 0 i 1.

Formalno, fazi skup A se definiše kao skup uređenih parova: $x \in X, A = \{(x, \mu_A(x)) | x \in X\}$, gde je x član skupa A , $\mu_A(x)$ predstavlja funkciju pripadnosti elementa x skupu A . Na ovaj način se uvodi neodređenost, koja potiče od osobine ljudi da stvari ne posmatraju na diskretan, egzaktna način. Fazi skup može se uprostiti sledećom predstavom:

$$A = \bigcup_{i=1}^n \mu_i / x, \quad 6.1.$$

Slične operacije koje su definisane nad *crisp* skupovima mogu se definisati i nad fazi skupovima. Dok kod klasičnih relacija imamo mogućnosti da dva ili više elemenata iz domena

budu ili ne budu u relaciji, kod fazi relacije imamo slučaj da elementi mogu imati veći ili manji stepen međusobne povezanosti. Time se dozvoljavaju različiti stepeni interakcije između elemenata. Može se reći da fazi relacija zapravo predstavlja višedimenzionalni fazi skup, kod koga funkcija pripadnosti predstavlja višedimenzionalnu funkciju $f = \mu_R(x_1, \dots, x_n)$ tako da fazi relaciju R u multidimenzionom prostoru $X_1 \times X_2 \times \dots \times X_n$ definišemo kao:

$$R = \{\mu_R(x_1, \dots, x_n)/(x_1, \dots, x_n) | x_1 \in X_1, \dots, x_n \in X_n\}. \quad 6.2.$$

Na isti način kao što teorija *crisp* skupova služi kao osnova za klasičnu logiku, teorija fazi skupova služi kao osnova za fazi logiku. Osnovni pojam fazi logike, kao i logike *crisp* skupova, je tvrđenje (iskaz) koje ima oblik x is A . Za razliku od logike *crisp* skupova koja polazi od stava da je tvrđenje istinito ili lažno fazi logika uključuje i mogućnost neodređenosti uvodeći stepen istinitosti. Kao i kod klasične logike, i u fazi logici operacije preseka, unije i komplementa imaju svoje korespodente u veznicima ‘**and**’, ‘**or**’ i ‘**not**’, tako da se pomoću njih mogu kombinovati pojedina tvrđenja. Kao primer posmatrajmo sledeća tvrđenja p : x_1 is A_1 and x_2 is A_2 , gde su A_1 i A_2 fazi skupovi koji imaju funkcije pripadnosti $\mu_{A_1}(x)$ i $\mu_{A_2}(x)$. Prethodno tvrđenje možemo predstaviti kao relaciju P sa funkcijom pripadnosti:

$$\mu_P(x_1, x_2) = T(\mu_{A_1}(x), \mu_{A_2}(x)), \quad 6.3.$$

gde je T opšta T -norma. Kombinovanjem tvrđenja dobijaju se pretpostavke (premise) u fazi pravilima. Tvrđenja se mogu odnositi na jednu ili na više lingvističkih promenljivih. U slučaju da se i jedna i druga tvrdnja odnose na istu lingvističku promenljivu imamo slučaj da se kombinovanjem fazi skupova dobija kao rezultat - fazi skup. U slučaju kada se tvrdnje odnose na više različitih lingvističkih promenljivih, konjunkcija u premisi se može prikazati u obliku fazi relacije.

Fazi **pravilo** predstavlja ‘**if-then**’ iskaz gde su premisa (pretpostavka) i konsekvencija (posledica) sastavljeni od fazi tvrđenja. Premisa se, kao što smo naveli, obično sastoji iz više tvrđenja povezanih logičkim veznicima ‘**and**’ i ‘**or**’. Navodimo primer jednog fazi pravila **R**: **if x_1 is A_1 and x_2 is A_2 then y is B** , gde su ‘ x_1 is A_1 ’ i ‘ x_2 is A_2 ’ tvrđenja u premisi, a ‘ y is B ’ u konsekvenci.

Funkcionalan fazi sistem mora da sadrži više od jednog fazi lingvističkog pravila r_k , gde je $k=1, \dots, N_r$. Kombinovanjem ovih pravila (poznato kao **agregacija**), dobija se kompaktna

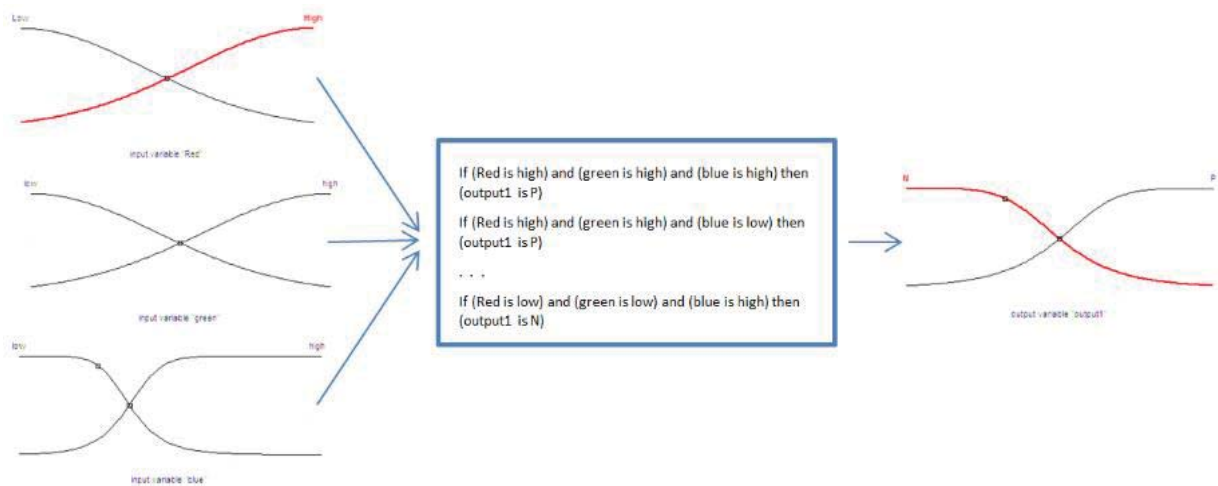
matematička predstavu celokupne baze znanja. U zavisnosti od tipa implikacije koja je korišćena, agregacija se svodi na neku od osnovnih logičkih operacija (konjunkciju ili disjunkciju, odnosno T- ili S-normu), tj.:

$$R = \bigcap_k R_k, \quad 6.4.$$

ili:

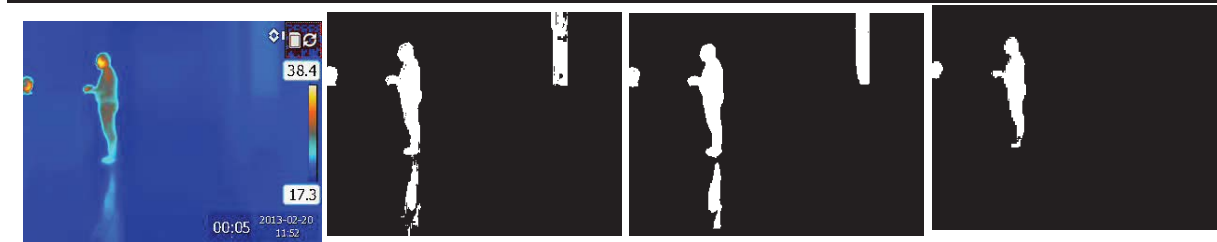
$$R = \bigcup_k R_k. \quad 6.5.$$

Teorija fazi skupova i fazi logike poslužila je kao osnova za projektovanje fazi segmentacionog algoritma prikazanog na slici 6.4. Termovizijsku sliku sa ovom paletom boja računar zapravo vidi kao i svaku drugu RGB sliku, tj. 3 matrice od kojih prva predstavlja intenzitet crvene boje svakog piksela na slici, druga intenzitet zelene i treća intenzitet plave boje svakog piksela na slici. Kako bi simulirali ljudsko razmišljanje, razvijen je niz fazi pravila kojim se izdvaja objekat čija je temperatura odgovara temperaturi čoveka. U zavisnosti od odeće koju nosi, strujanja vazduha, osvetljenja i mnogih drugih parametara temperatura čoveka izmerena termovizijskom kamerom varira i teško je jasno odrediti njen opseg. Izvršena je fazifikacija osnovnih boja i intenzitet svake poje podeljen je na dve oblasti - niska i visoka pripadnost određenoj boji, sa Gausovom funkcijom pripadnosti fazi skupu. Zatim se na osnovu fazi pravila donosi se zaključak da li neki piksel treba ili ne treba izdvojiti, tj. da li pripada pozadini ili objektu od interesa.



Slika 6.4 Algoritam fazi segmentacije termovizijske slike

Konačni izlaz iz algoritma za segmentaciju jeste binarna slika na kojoj su izdvojeni objekti koji bi mogli da predstavljaju ljude. Rezultati segmentacije manuelnim određivanjem praga i fazi segmentacije za dva različita podešavanja funkcija pripadnosti ulaza i izlaza fazi sistema prikazani su na slici 6.5.



Slika 6.5 Rezultati segmentacije termovizijske slike a) nakon primene ručno podešenog praga b), nakon primene fazi segmentacije c) i promene parametara fazi segmentacije d)

U ovom slučaju su parametri fazi sistema vršeni ručno, ali je moguće koristiti i neki optimizacioni algoritam za određivanje parametara fazi sistema. Kako je problem nelinearan, moguće je uspešno koristiti neki od metaheurističkih alata optimizacije, poput PSO algoritma [98] ili genetske algoritme [Nik12] za određivanje parametara fazi sistema za segmentaciju.

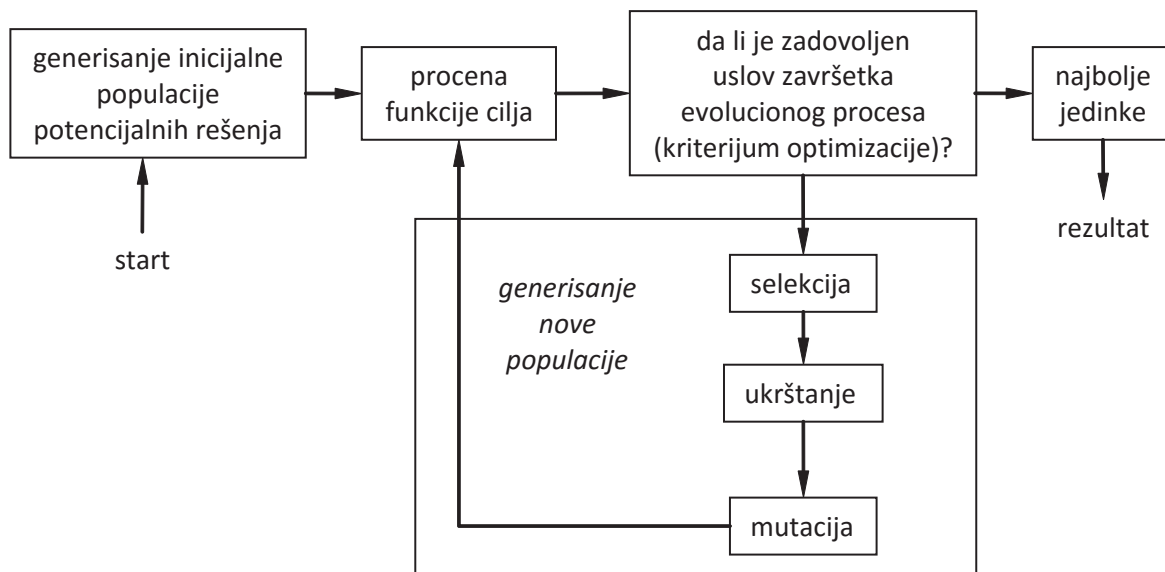
6.3.2. Optimizacija parametara segmentacije slike primenom genetskih algoritama

Snimljena su četiri video snimka u zatvorenom prostoru za različita stvarna scenarija, pri čemu su svi napravljeni pod različitim osvetljenjem gde postoje prepreke ispred osobe, dve osobe se preklapaju kada se približavaju jedna drugoj i moguća je pojava refleksije. Prva ideja je bila da se napravi klasifikator zasnovan na karakteristikama ručno segmentiranih slika. Međutim, u stvarnom svetu algoritam za segmentaciju neće biti tačan i dobar pa nam je potrebno da implementiramo rešenje koje će biti bliže stvarnom problemu. Zato smo razmatrali ideju određivanja jednog skupa parametara graničnog praga koji je primenljiv na svaki scenario. Optimizacija pomoću genetičkog algoritma, zasnovana na 50 slučajno izabranih i ručno segmentiranih slika za svaki od 4 eksperimentalna termovizijska snimaka, se koristi za određivanje optimalne vrednosti parametara graničnog praga. Problem se svodi na višeciljnu optimizaciju sa dva parametra, T_{high} i T_{low} gornjom i donjom granicom praga, uključujući 200 ciljnih funkcija za istovremenu optimizaciju. Ciljne funkcije su sledeće:

$$g_i(T_{low}, T_{high}) = \|\mathbf{y}(T_{low}, T_{high}) - \hat{\mathbf{y}}_i\| \quad 6.6.$$

gde i ima vrednosti od 1 do 200, vektor \mathbf{y}_i i $\hat{\mathbf{y}}_i$ su veličine 4320 i čine ih vrste matrice \mathbf{A} i $\hat{\mathbf{A}}$ respektivno.

Matrica \mathbf{A} je 240x180 matrica koja predstavlja segmentnu sliku, vrednost svakog elementa matrice može biti 0 ili 1 i zavisi od gornje T_{high} i donje T_{low} granice praga. Matrica $\hat{\mathbf{A}}$ je 240x180 matrica binarne slike koja predstavlja ručno segmentiranu sliku. Termovizijska slika, ručno segmentirana slika i slika segmentiranog graničnog praga su pokazani na Slici 5. Kako problem segmentacije termovizijske slike iziskuje uključivanje nekih inteligentnih i heurističkih algoritama optimizacije autori su odlučili da testiraju genetski algoritam (GA) za ove potrebe. Implementacija GA zahteva određivanje 6 osnovnih stvari: predstavljanje hromozoma, izbor funkcije, genetičke operatore, funkcije inicijalizacije, prekida i evaluacije. Struktura korišćenog genetskog algoritma prikazana je na slici 6.6.



Slika 6.6 Struktura genetskog algoritma

Na početku genetske optimizacije se inicijalizuje populacija tj. određen broj jedinki. Inicijalna populacija hromozoma se generiše slučajnim odabirom rešenja iz domena traženja. Za svako rešenje se određuje vrednost funkcije cilja koji predstavlja meru kvaliteta (fitness). Postupkom selekcije se iz stare populacije se sada formira nova populacija izdvajanjem boljih jedinki. Tokom selekcije loše jedinke odumiru, a bolje opstaju. Neki članovi ove nove populacije podvrgnuti su uticajima genetskih operatora koji iz njih formiraju nove jedinke.

Ukrštanjem se prenose svojstva roditelja na potomke. Mutacija stvara novu jedinku menjajući manji deo genetskog materijala. Kada se zadovolji uslov zaustavljanja, posle određenog broja generacija, najbolji član trenutne populacije predstavlja rešenje koje bi trebalo da bude blizu optimalnog.

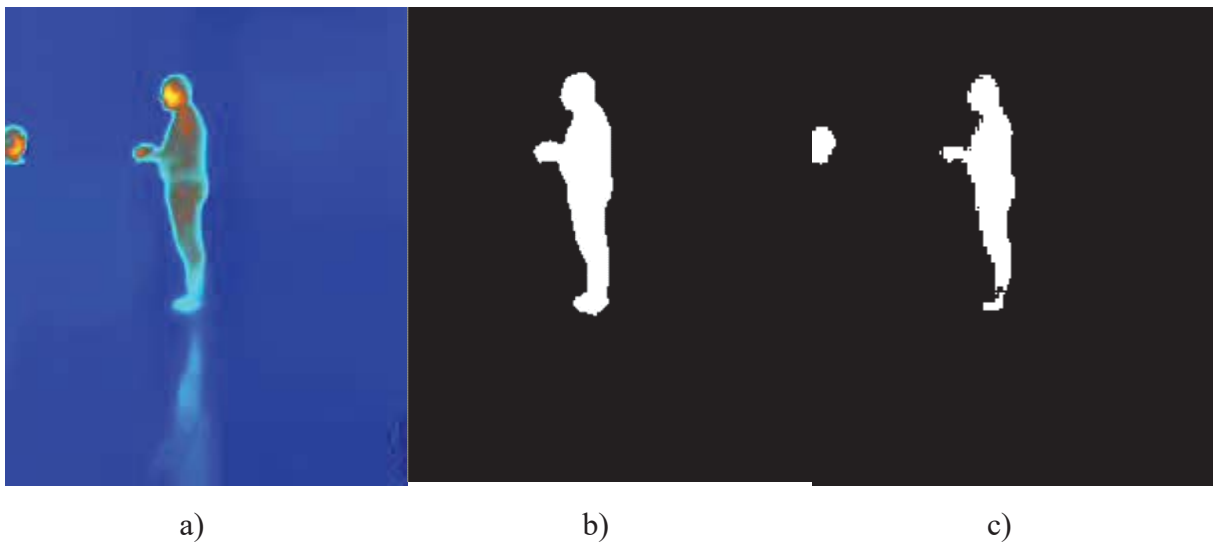
Funkcija cilja se projektuje skalarizacijom višeciljnog problema optimizacije veoma poznatom metodom linearne skalarizacije:

$$\min_{T_{low}, T_{high}} \sum_{i=01}^{200} w_i g_i(T_{low}, T_{high}) \quad 6.7.$$

gde su težinski koeficijenti ciljeva $w_i > 0$ parametri skalarizacije, a u našem slučaju $w_i = 1$, sve dok je svaki frejm svakog videa jednakog značaja.

U implementiranom algoritmu se koristi populacija od 30 jedinki sa tri elitne jedinke a inicijalna populacija je slučajno generisana pri čemu je inicijalni interval $[0,255]$. Realizovana je funkcija ukrštanja sa rasipanjem, koja pravi slučajni binarni vektor a onda bira gene gde je vektor 1 od prvog roditelja a vektor 0 od drugog roditelja i kombinuje gene za formiranje deteta. Svi članovi su podložni mutaciji osim elite. Operator mutacije je adaptivan. Kao rezultat *offline* podešavanja genetskih parametara, određene su optimalne, gornja T_{high} i donja T_{low} , granice graničnog praga.

Rezultati segmentacije prikazani su na slici 6.7.



Slika 6.7 Termovizijska slika a), ručno segmentirana slika b) i segmentirana slika sa optimalno određenim graničnim pragom c)

6.3.3. Izdvajanje karakteristika segmentiranih binarnih objekata

Nakon uspešne segmentacije, u nekim frejmovima se javljaju osim ljudi i objekti slične temperature ili se javljaju neke smetnje poput refleksije koje dovode do prikazivanja segmentiranih objekata koji nisu ljudi. Kako bi robotska platforma pratila samo ljude i zanemarivala poremećaje, bilo je neophodno razviti klasifikator koji određuje da li je segmentirani objekat čovek ili ne. Za razvoj klasifikatora neophodno je odrediti ulaze, a najčešće korišćeni su zapravo karakteristike koje opisuju oblik objekta, poput proporcionalnosti, konektivnosti, Hu-ovih invarijantnih momenata i sl. Izabrani ulazi u

razvijeni klasifikator su, nakon analize osjetljivosti, proporcionalnost P i prva tri Hu-ova momenta I_1 , I_2 i I_3 .

6.4. Inteligentni sistemi za klasifikaciju objekata sa termovizijske slike

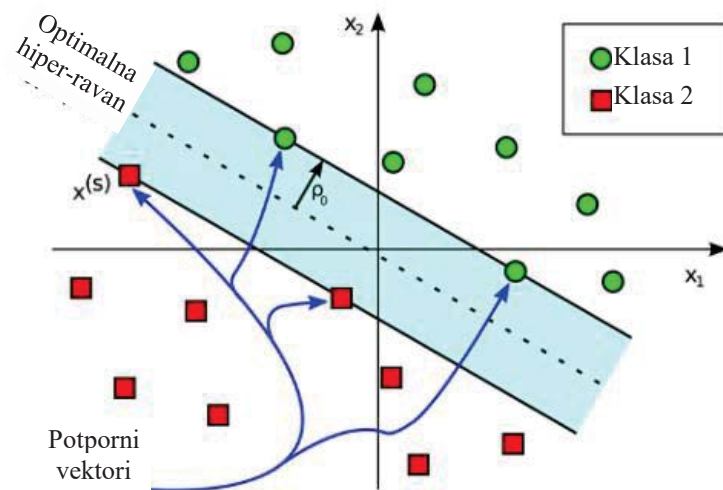
U istraživanju su za klasifikaciju korišćeni različiti mehanizmi klasifikacije koji se zasnivaju na veštačkoj inteligenciji, poput metode potpornih vektora [1,10,19], klasifikatora zasnovanog na neuronskoj mreži [2,12] i ANFIS klasifikator [9]. Kako su ANFIS klasifikator i klasifikator zasnovan na neuronskoj mreži vrlo slični klasifikatorima predstavljenim u poglavlju 4 ove teze, u ovom poglavlju će akcenat biti stavljen na metod potpornih vektora, koji upravo i razdvaja set podataka na dve klase, tj. segmentirane objekte koji su ljudi i one koji to nisu.

Za klasifikaciju segmentiranih regiona razvijen je SVM klasifikator čiji su ulazi proporcionalnost P i prva tri Hu-ova momenta I_1 , I_2 i I_3 . Kako je u pitanju sistem koji treba da radi u realnom okruženju i realnom vremenu, značajno je da klasifikacija bude pouzdana i efikasna. Ovo podrazumeva da klasifikator prepoznaje ljude s visokom tačnošću, koristi malo kompjuterskih resursa i vremenski nije zahtevna.

6.4.1. Metod potpornih vektora za detekciju ljudi

Metod potpornih vektora (engl. *Support Vector Machine* - SVM) predstavlja model mašinskog učenja s nadzorom koji analizira podatke i prepoznaje šablone i koristi se za klasifikaciju i analizu regresije. Odlikuje ga izuzetna robusnost što ga čini pogodnim za veliki broj aplikacija [100].

Kada se koriste za klasifikaciju, razdvajaju binarno numerisane podatke pomoću hiper ravni koja je na maksimalnoj udaljenosti od različitih klasa (slika 6.8). Kada nije moguće linearno razdvajanje, koristi se metod s jezgrima koji vrši nelinearno mapiranje prostora značajnih ulaza.



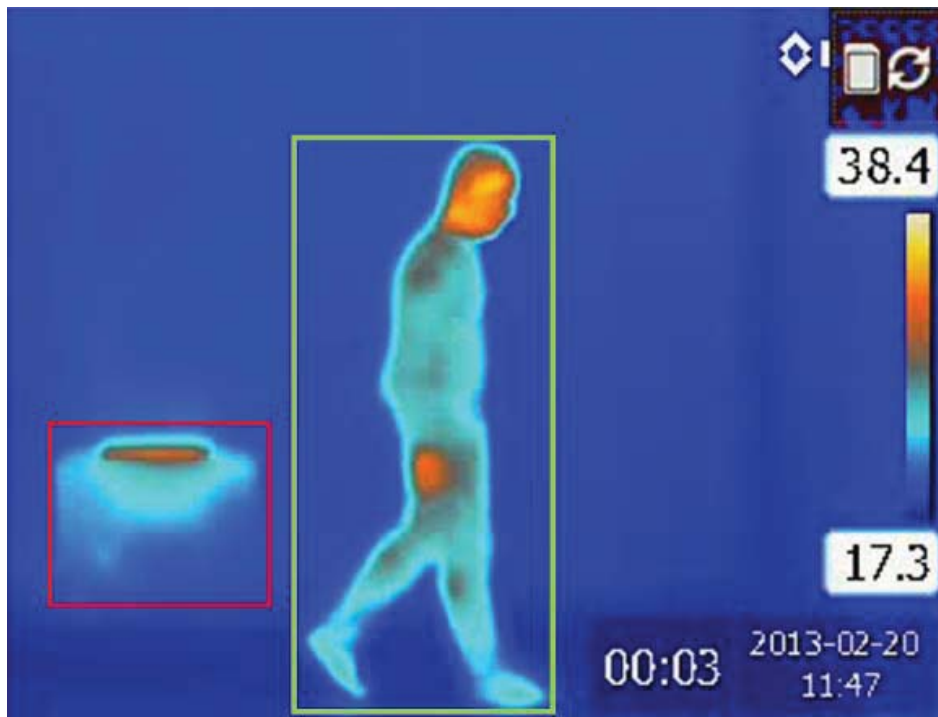
Slika 6.8 SVM klasifikator

Za razvoj klasifikatora korišćeni su podaci dobijeni iz dvesta slučajno izabranih frejmova iz video sekvenci dobijenih nakon 4 eksperimenata u kojima je korišćeno različito osvetljenje, različiti ljudi i različiti zatvoreni prostori. Nakon segmentacije i ručnog obeležavanja objekata formirano je 512 objekata koja karakterišu po 4 značajne karakteristike i koji su korišćeni za trening, validaciju i testiranje SVM klasifikatora.

6.5. Evaluacija rezultata

Problem evaluacije sistema za praćenje postao je aktuelan poslednjih godina u kompjuterskoj viziji. Opšta saglasnost je da ne postoji jedinstvena mera koja bi mogla da prikaže dovoljno dobro kvalitet celog sistema. Za adekvatnu evaluaciju bitno je koristiti različitu metriku kvantifikovanja performansi za različite aspekte sistema. Valjano određene karakteristike ponašanja omogućavaju: optimizaciju parametara algoritma, proveru performansi algoritma za praćenje za različite tipove podataka, kvantitativno upoređivanje različitih algoritama, podršku razvoja algoritma i odlučuje o konsenzusu između različitih aspekata sistema. Evaluaciona procedura zahteva da su dostupni podaci o sistemu, što je u slučaju video podataka veoma težak, monoton i zahtevan proces. Bilo je pokušaja da se poboljša i automatizuje proces evaluacije pomoću nekog drugog algoritma koji bi približno odredio regione od interesa, koji bi se kasnije ručno poboljšavali, i čime bi se omogućilo obučavanje sistema ili pak na osnovu sistema koji bi potpuno automatski vršili procenu na

osnovu boje i pokreta. Na slici 6.9 prikazano je uspešno prepoznavanje i praćenje čoveka robotom sa termovizijskom kamerom.



Slika 6.9. Praćenje osobe mobilnim robotom

Primenom savremenih metoda i algoritama iz domena veštačke inteligencije, razvijen je upravljački sistem inteligentnog robotskog sistema sa termovizijskom kamerom koji ima sposobnosti detekcije, prepoznavanja, prostornog lociranja i praćenja ljudi u zatvorenom prostoru kao i odgovarajuću interakciju ovog robotskog sistema sa ljudima. Implementacijom više tehnika veštačke inteligencije, kao što su genetski algoritmi i SVM klasifikacija u proces razvoja inteligentnog upravljačkog sistema postignuto je autonomno i pouzdano prepoznavanje ljudi i njihovo praćenje koje pokazuje u odnosu na postojeća rešenja.

S obzirom na izrazito multidisciplinarni karakter naučnoistraživačke oblasti predloženog rešenja koja podrazumeva značajan uticaj matematičkih, mehaničkih, softverskih i senzorsko-elektronskih aspekata, neophodno je bilo primeniti analitički pristup za rešavanje problema obrade i analize slike dobijene termovizijskom kamerom i ekstrakcije i klasifikacije osobina različitih objekata u scenariju. Primenjen je i eksperimentalni metod razvoja hibridnih sistema za prepoznavanje i praćenje ljudi, gde je SVM klasifikator obučen eksperimentalnim podacima kako bi vršio prepoznavanje, estimaciju položaja i praćenje ljudi u zatvorenom

prostoru. Integracijom razvijenih sistema sa sistemom za upravljanje kretanjem robota koji je ranije razvijen u Laboratoriji za upravljanje sistemima dobijen je inteligentni sistem upravljanja koji omogućava interakciju mobilne robotske platforme sa ljudima.

7. Inteligentni sistemi za praćenje ljudi, prepoznavanje poza i gestova

Kako je pokazano u poglavlju 5 ove teze, zbog promene morfologije čoveka tokom kretanja bilo je neophodno razviti sistem za estimaciju i predikciju položaja čoveka. U ovom poglavlju biće predstavljen koncept razvoja inteligentnog sistema zasnovanog na nelinearnoj autoregresivnoj neuronskoj mreži za ocenu položaja čoveka i kratkoročno predviđanje budućih položaja čoveka u prostoru.

Osim toga, biće predstavljena i mobilna robotska platforma sa 3D senzorom koji uz odgovarajuće predložene algoritme omogućava pouzdano praćenje ljudi i prepoznavanje poza i gestova.

Kako je već napomenuo u radovima koji se bave praćenjem ljudi u robotici, sam problem praćenja se realizuje detekcijom ljudi u svakom frejmu, ali se preporučuje dodatno korišćenje filtera koji omogućuju ocenu i predviđanje položaja ljudi. U svrhu predviđanja trajektorija ljudi javlja se ideja korišćenja Kalmanovog filtera [73], neuronske mreže [101] ili korišćenja algoritma učenja i skrivenog Markovljevog modela u kombinaciji sa tipičnim socijalnim obrascima kretanja ljudi [102].

Prepoznavanje akcija ljudi je takođe zanimljiv problem i može predstavljati deo modula za praćenje ljudi u sistemu vizije, i zahteva razvoj klasifikatora koji pomažu u prepoznavanju konkretnih akcija [103,104]. Prepoznavanje i predviđanje akcija ljudi se u tehničkim sistemima svode na prepoznavanje poza, gestova i kretanja ljudi [105], a u cilju prepoznavanja obrazaca ponašanja ljudi. Za prepoznavanje poza potrebno je formirati 3D model (ili 3D model skeleta) čoveka [106] što kod sistema stereo vizije predstavlja svojevrsan problem zbog nepreciznosti modela.

Na ovim osnovama su zasnovani i specijalizovani sistemi za automatsko prepoznavanje glasovnog jezika kojim se koriste gluvo-neme osobe ili osobe sa oštećenim sluhom [107,108]. Za prepoznavanje gestova [109] neophodan je algoritam učenja koji se najčešće svodi na regresiju najmanjih kvadrata i formiranje klasifikatora zasnovanog na verovatnoći. Konačno,

gestovi ljudi mogu se koristiti kako bi se upravljalo robotom [110], a ovakav način interakcije robota i čoveka može se dodatno unaprediti modulom za prepoznavanje govora.

7.1. Inteligentni sistem za ocenu i predviđanje položaja čoveka

Tokom eksperimenata se pokazalo da i pored robusnosti algoritama za segmentaciju i klasifikaciju sistem u pojedinim trenucima ne prepozna čoveka u sceni i pored njegove prisutnosti. U koliko je čovek suviše blizu nekom objektu, segmentacioni algoritam koji nalazi oblasti od interesa iz mape dispariteta će segmentisati čoveka i taj objekat kao jedinstveni objekat čije će ekstrahovane karakteristike biti bitno različite od karakteristika čoveka, pa će i sam klasifikator pogrešiti i neće prepoznati čoveka.

Osim toga deo upravljačkog softvera mobilnog robota sa stereo kamerom podrazumeva izvršavanje algoritma za formiranja mape dispariteta, algoritma segmentacije mape dispariteta, ekstrakciju karakteristika i konačno klasifikaciju. Svaki od ovih segmenata upravljačkog bloka vizije zahteva odgovarajuće računarske resurse i vreme izračunavanja nije zanemarljivo. Kako je bitno da se praćenje vrši u realnom vremenu potrebni su značajni računarski resursi kako bi se ovaj proces sa svim svojim segmentima izvršio između dva frejma. Kako stereo kamera snima 30 frejmova u sekundi, ceo proces obrade slike mora da se izvrši za manje od 33,3 *ms*.

Ova dva problema dovela su do potrebe da se uvede algoritam za praćenje čoveka, odnosno filter koji vrši estimaciju i predikciju položaja čoveka u sceni, čija uloga može biti dvojaka. S jedne strane ovakav algoritam može oceniti položaj čoveka u trenutku kada sistem vizije iz nekog razloga izgubi čoveka iz vida. Rešenje problema estimacije primenom Kalmanovog filtera može se naći u radu [73], dok je autor ove teze ujedno koautor radova [8, 11, 15, 21] koji obrađuju ovaj problem. Kao alternativa Kalmanovom filteru u radu [4,14] predložene su rekurentne neuronske mreže i upoređen je njihov učinak sa Kalmanovim filterom.

Drugi problem, tj. vreme obrade slike može se rešiti smanjenjem oblasti pretrage. Primenom neuronske mreže [6] može se predvideti položaj čoveka u narednom frejmu, čime se može značajno smanjiti oblast pretrage u prostoru.

Kako se promene položaja čoveka u prostoru menjaju u zavisnosti od vremena, a ocena trenutnog položaja, odnosno predviđanje narednog položaja, vrši se isključivo analizom prethodnih položaja čoveka u prostoru, jasno je da se ovaj problem svodi na poznati problem predviđanje vremenskih serija.

7.1.1. Predviđanje vremenskih serija

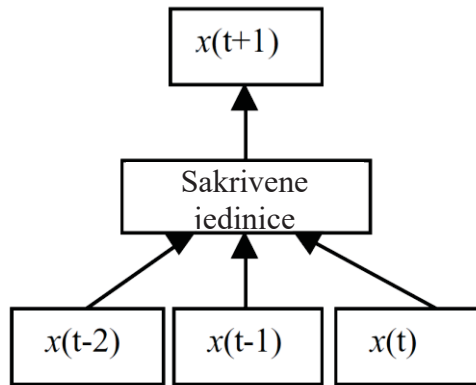
U današnje vreme predviđanje vremenskih serija je veoma aktuelan problem sa primenom u različitim oblastima. Vremenska serija predstavlja sekvencu vrednosti, dobijenih merenjima koja se obično izvode u uzastopnim vremenskim trenucima koji su međusobno ekvidistantni. U opštem slučaju, vremenska serija može ispoljavati nelinearnost, nestacionarnost, periodičnost, prisustvo haotične komponente i prisustvo šuma.

Vremenska serija je linearna ukoliko se njene buduće vrednosti mogu predstaviti preko linearne funkcije njenih prethodnih vrednosti. Ukoliko se statistička svojstva (srednja vrednost i varijansa) vremenske serije ne menjaju tokom vremena ona se smatra stacionarnom. Vremenska serija koja sadrži komponente koje se periodično ponavljaju, smatra se periodičnom. Specijalni slučaj ovoga su takozvane sezonske vremenske serije koje imaju ponovljivo i predvidljivo pomeranje oko trenda u intervalu od jedne godine ili manjem. Pored toga vremenska serija može sadržati haotičnu komponentu i šum, koji može biti prisutan samo u njenim pojedinim delovima ili u celoj vremenskoj seriji.

Zbog mogućnosti da vremenske serije u opštem slučaju poseduju kombinaciju svih prethodno navedenih osobina, oblast njihovog predviđanja je veoma zahtevna sa stanovišta razvoja novih i unapređenja postojećih predikcionih metoda.

Konvencionalne metode za predikciju vremenskih serija, aktuelne tokom 70-tih i 80-tih godina dvadesetog veka, među kojima su najpoznatije linearna regresija, Kalmanovi filteri, *Box-Jenkins* ARIMA i *exponential smoothing* ne mogu uvek da obezbede dovoljno precizna predviđanja kada je proces koji generiše vremensku seriju nelinearan, nestacionaran i kada njegova svojstva nisu unapred poznata. Pored toga, veoma često sam izbor predikcione metode i određivanje njenih parametara zavisi od detaljnog poznavanja svojstava procesa koji generiše vremensku seriju.

U cilju prevazilaženja ovih ograničenja i pojednostavljenja izbora modela, u poslednje dve decenije, modeli zasnovani na nadgledanom mašinskom učenju su se pozicionirali kao konkurentni “standardnim” modelima. Ovi modeli, u literaturi poznati još i kao *black-box*, odnosno *data-driven* modeli, koriste podatke o vremenskoj seriji iz prošlosti kako bi “naučili” stohastičke zavisnosti između ulaza i izlaza modela, i na osnovu njih predviđali buduće vrednosti (slika 7.1).



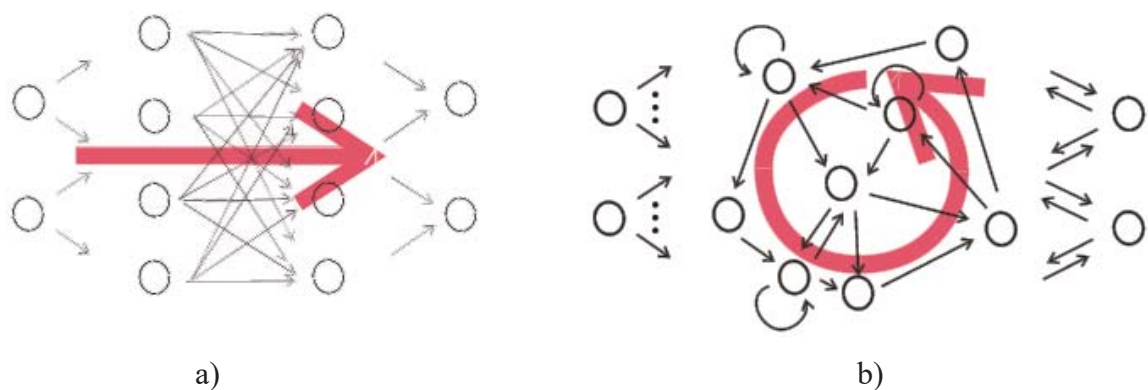
Slika 7.1 Predviđanje vremenskih serija black-box modelom

U osnovi, kod svih metoda nadgledanog mašinskog učenja, problem predikcije vremenskih serija se može posmatrati sa stanovišta obučavanja modela koji uspostavlja preslikavanje između ulaznih i izlaznih promenljivih. Nakon uspostavljanja takvog preslikavanja, formirani model se može koristiti za predviđanje budućih vrednosti na osnovu prethodnih i tekućih vrednosti.

Često korišćeni alati iz domena mašinskog učenja za predikciju vremenskih serija su neuronske mreže. Njihova primena u kratkoročnom i dugoročnom predviđanju nelinearnih vremenskih serija sa prisustvom nepouzdanosti i šumom može se sresti u radovima koji se bave najrazličitijim problemima, od predviđanja zagađenja vazduha [111], preko predviđanja haotičnih vremenskih serija kretanja na berzi [112] i promene kursa određenih valuta [113], do prediktivnog upravljanja kompleksnim procesima u hemijskom postrojenju [114] i lokalizacije mobilnog robota [115].

7.1.2. Nelinearna autoregresivna neuronska mreža za ocenu i predviđanje položaja čoveka

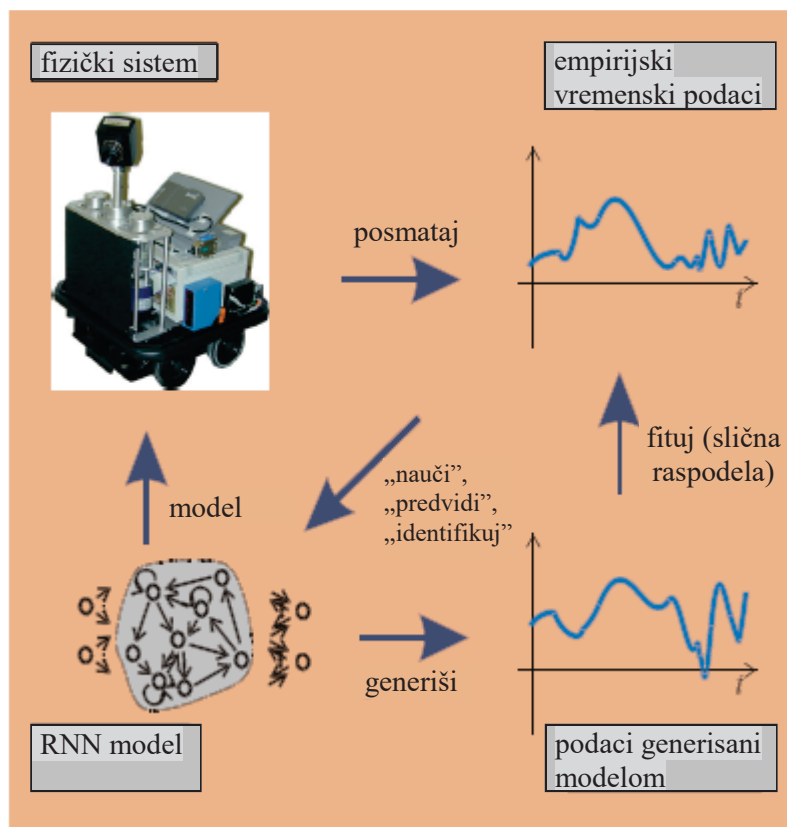
U prethodno navedenim radovima koji se bave predikcijom i ocenom vremenskih serija primenom neuronskih mreža se zbog prirode samog problema primenjuju rekurentne neuronske mreže. Kod neuronskih mreža sa prostiranjem signala unapred je aktivacija neurona “provedena” kroz mrežu od ulaznih do izlaznih jedinica (s leva na desno, slika 7.2 a), dok se kod rekurentne neuronske mreže javlja bar jedna zatvorena putanja sinaptičkih veza (slika 7.2 b).



*Slika 7.2 Neuronska mreža sa prostiranjem signala unapred a)
i rekurentna neuronska mreža b) [116]*

Sve biološke neuronske mreže su povratne, pa se primena ovakve mreže nameće kao rešenje problema kod kojih neuronske mreže sa prostiranjem signala unapred ne daju zadovoljavajuće rezultate. Ovim alatom se jako uspešno mogu rešiti problemi formiranja black-box modeli i koristiti se u procesiranju signala [116]. Slika 7.3 predstavlja šemu rešavanja problema identifikacije i formiranja modela fizičkog sistema, ocene stanja sistema i obučavanja mreže.

Uopšteno, rekurentna neuronska mreža može se koristiti za identifikaciju sistema, filtriranje podataka, kao obrazac klasifikacije, za nasumično sekvencijalno modeliranje, kompresiju podataka i sl.



Slika 7.3 Rešavanje problema black-box modeliranja sistema primenom rekurentne neuronske mreže [116]

Najjednostavnija neuronska mreža koja može poslužiti za predviđanje vremenskih serija jeste nelinearna autoregresivna mreža (NAR). Ovim pristupom se koristeći znanja iz prošlosti predviđa budućnost. Ovakva mreža poseduje sposobnost da uči na osnovu podataka iz svog okruženja iterativnim procesom prilagođavanja sinaptičkih težina i pragova. Nelinearna autoregresivna metoda je nastala od linearne autoregresivne metode, a u cilju rešavanja realnih problema kod kojih se javljaju različiti tipovi nelinearnosti.

Odnos ulaza i izlaza linearnog autoregresivnog modela može se definisati sledećom jednačinom:

$$r(k) = \sum_{i=1}^p a(i)r(k-i) + e(k) \quad 7.1.$$

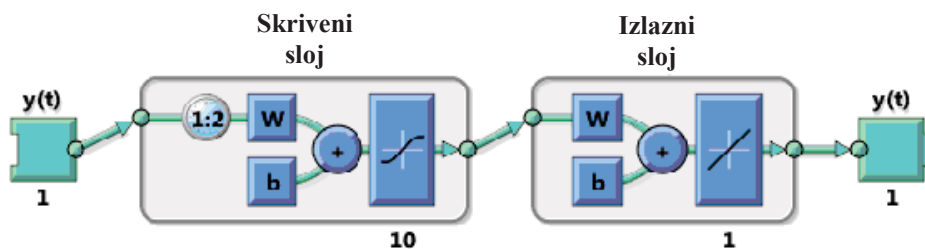
gde je trenutni ulazni signal $r(k)$ jednako sumi prethodnih stanja $r(k-i)$ pomnoženih odgovarajućim težinskim faktorima $a(i)$. Komponenta $e(k)$ predstavlja signal greške, dok je p red sistema. Težinski faktori $a(i)$ se određuju primenom Yule-Walkerove jednačine [117,118], minimizacijom nezavisnog ravnomerno raspoređenog signala greške. Ovi težinski faktori su

konstantni za sve semplove signala $r(k)$, ali je ovo primenljivo samo kod stacionarnih signala. Kod nestacionarnih signala se ovi težinski faktori menjaju tokom vremena usled pojave različitih tipova nelinearnosti. Kako bi se prevazišao ovaj problem, linearni odnos ulaza i izlaza može biti proširen na nelinearan odnos, no u tom slučaju javlja se problem ocene parametara modela. Nelinearni autoregresivni model reda p predstavljen je sledećom jednačinom:

$$r(k) = g(r(k-1), \dots, r(k-p)) + e(n), \quad 7.2.$$

gde je $g(x)$ nepoznata nelinearna funkcija, koja predstavlja nelinearne odnose trenutnog signala i p prethodnih signala.

Za rešavanje konkretnog problema praćenja ljudi razvijene su 3 nelinearne autoregresivne neuronske mreže od kojih svaka prati po jednu koordinatu (X, Y i Z) centra mase čoveka u prostoru. Svaka mreža ima jedan skriveni sloj sa 10 neurona, a njihova topologija predstavljena je na slici 7.4.



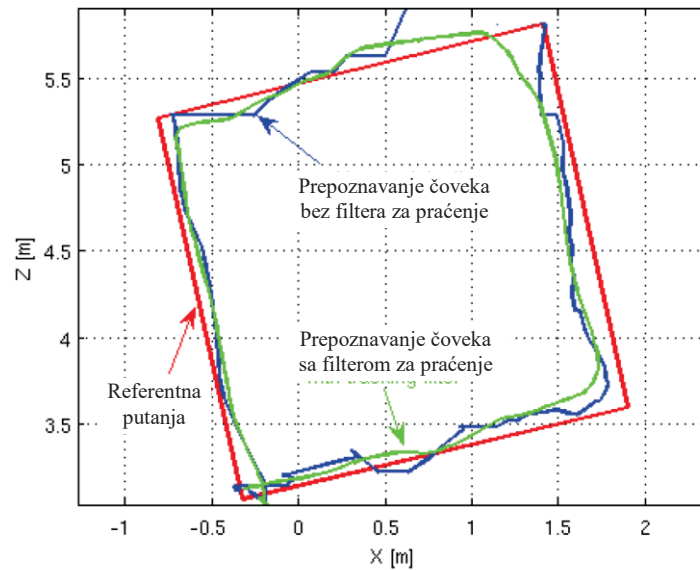
Slika 7.4 Topologija NAR neuronske mreže za predviđanje i ocenu jedne koordinate položaja čoveka u prostoru

Kako se vidi sa slike topologije mreže, za predviđanje narednog stanja koriste se trenutno i dva prethodna stanja.

7.1.3. Obuka mreže i verifikacija rezultata

Za obučavanje i testiranje modula za estimaciju i predikciju položaja čoveka u prostoru korišćeni su podaci dobijeni mobilnom robotskom platformom RecoRob, i snimljeni stereo kamerom Point Grey Bumblebee XB3. Eksperiment je izvršen u zatvorenom prostoru, odnosno u laboratorijskim uslovima, i dok je robot bio stacioniran u uglu prostorije, osoba se kretala po

kvadratnom šablonu nacrtanom na podu koji se nalazio ispred robota. Stvarna pozicija čoveka je merena digitalnim laserskim senzorom distance Bosch PLR 50 i služila je za evaluaciju rezultata. Na slici 7.5 prikazan je dijagram pogleda na scenu iz ptičje perspektive i na njemu kvadratnog šablona, detektovanog položaja čoveka bez filtriranja i filtriranog položaja čoveka. Ovaj dijagram predstavlja rezultate predstavljene u [25,73], koji su služili kao osnov za dalje unapređenje i primenu inteligentnog sistema za estimaciju i predikciju.

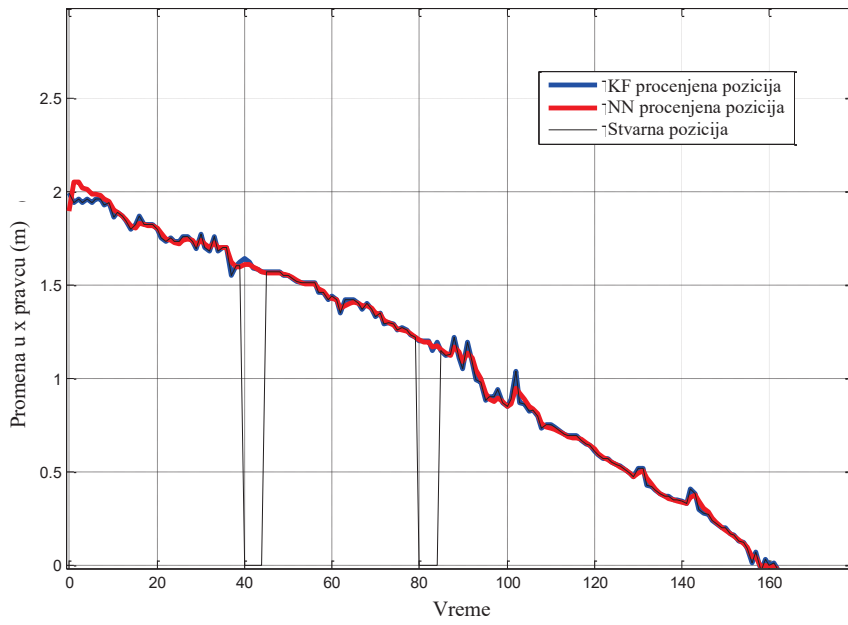


Slika 7.5 Praćenje čoveka koji hoda po šablonu nacrtanom na podu [25]

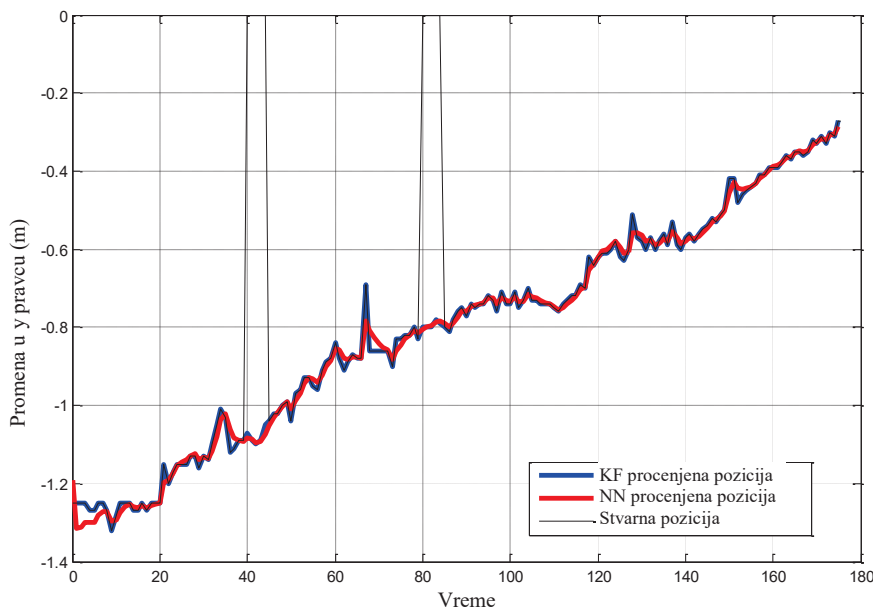
Nakon akvizicije sekvence stereo slika, formiranja mapa dispariteta, segmentacije mapa dispariteta, prepoznavanja (klasifikacije) čoveka u sceni i određivanja centara mase u svim sekvencama formirana su tri vektora promene položaja čoveka, koji su služili za obučavanje NAR neuronske mreže. Zbog zahtevnosti algoritama za obradu slika praćenje je vršeno pri snimanju 12 frejmova u sekundi.

Obučavanje nelinearne autoregresivne neuronske mreže izvršeno je primenom Levenberg-Marquardt-ove metode.

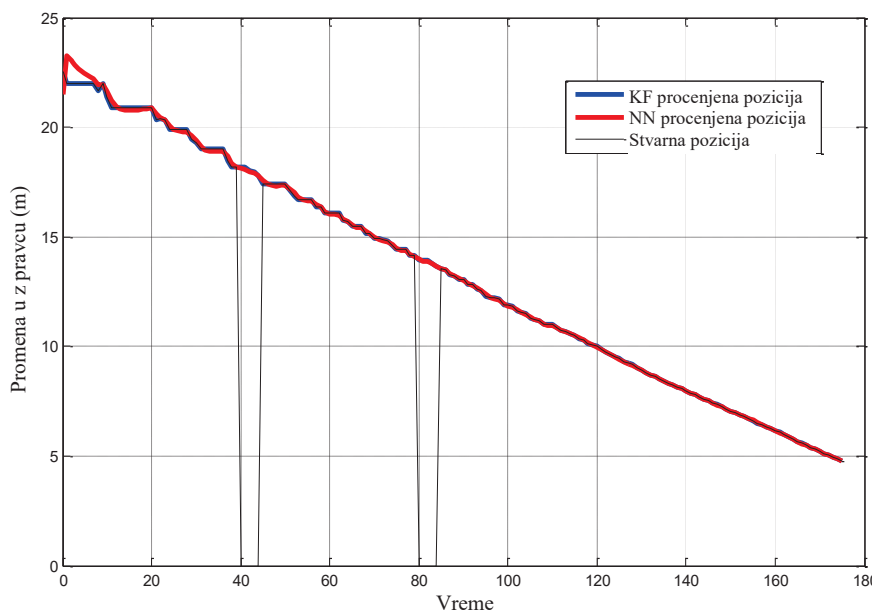
Na dijagramima prikazanim na slici 7.6 a), b) i c) prikazane su promene koordinata položaja čoveka i upoređeni su rezultati dobijeni Kalmanovim filterom i filterom zasnovanim na nelinearnoj autoregresivnoj neuronskoj mreži.



a)



b)



c)

Slika 7.6 Promena X koordinate čoveka tokom vremena a), promena Y koordinate čoveka b) i promena Z koordinate čoveka u prostoru c)

Uporednom analizom dijagrama potvrđuje se višem puta izneta tvrdnja da nelinearna autoregresivna neuronska mreža pri oceni i predviđanju vremenskih serija može da se izbori sa nelinearnostima, dok to Kalmanov filter ne može. Ova tvrdnja bi bila još očiglednija kada bi se čovek kretao po geometrijski nepravilnoj putanji, ali bi se tada pri projektovanju nelinearne autoregresivne neuronske mreže u obzir trebalo uzeti više prethodnih stanja i povećati broj neurona u skrivenom sloju.

7.2 Razvoj mobilne robotske platforme sa 3D senzorom za praćenje ljudi i prepoznavanje poza

Praćenje ljudi i prepoznavanje poza je veoma izazovno ako želimo da postignemo potpunu autonomnost robota. Većina mobilnih robota koja su namenjena za praćenje ljudi su opremljeni kamerama ili laserskim senzorima daljine kojima detektuju ljude. U ovu svrhu su korišćeni brojni senzori, uključujući ultrazvučne senzore, laserske merači razdaljine, kao i monokularne i stereo kamere. Svi ovi senzori imaju nedostatke, kao što su nepreciznost, visoka algoritamska složenost, ili jednostavno težina ili cena.

Kamere za merenje dubine projektovane da rade sa infracrvenim obrascima pojavom komercijalnih rešenja poput Microsoft Kinect i Kinect2 ili Asus Xtion senzora koji se koriste u industriji zabave postali su široko dostupne. Ove kamere su relativno precizne i obezbeđuju prikaz dubine, tj. 3D informacije o sceni direktno iz hardvera, pa time prevazilaze neka od ograničenja koja se javljaju kod konvencionalnih sistema vizije. Zbog toga je u poslednjih 5 godina sve veći broj robotskih aplikacija koje kao osnovni izvor informacija o okruženju koriste upravo ove 3D senzore.

Za razvoj pouzdanog, računski inteligentnog prepoznavanja ljudi u robotskoj viziji u ovom delu istraživanja je korišćen National Instruments Robotics Starter Kit 1.0 i Asus Xtion live 3D senzor.

7.2.1. Mobilna robotska platforma DaNI sa 3D senzorom

Za potrebe našeg tehničkog rešenja razvijen je sistem sa robotskom vizijom za prepoznavanje objekata, gde ASUS 3D senzor (Slika 7.7) omogućuje vizualne informacije o okolini robota. Zbog toga su kako ovaj senzor, tako i algoritmi koji omogućavaju praćenje ljudi i prepoznavanje gestova od izuzetnog značaja u robotskom sistemu. ASUS 3D senzor je povezan sa laptopom, koji preko Lan kabla komunicira sa robotom (Slika 7.8). Sva softverska rešenja za obradu slike, klasifikaciju i upravljanje mobilnom robotskom platformom sa robotskom vizijom realizovana su u softverskom paketu Simulink kompanije MathWorks i LabVIEW 2011 kompanije National Instruments . Osim toga u realizaciji ovog tehničkog rešenja korišćena je i ostala oprema koja je dostupna u Laboratoriji za upravljanje sistemima (LUPS) Mašinskog fakulteta Univerziteta u Nišu.



Slika 7.7. Asus Xtion PRO live 3D senzor



Slika 7.8. DaNI robot opremljen laptopom i Asus Xtion PRO LIVE kamerom i njihovi koordinatni sistemi

Asus Xtion live 3D senzor kamera se sastoji od 4 senzora: ifra-red senzor, senzor dubine, hromatski senzor i mikrofoni (zvučni senzor). Specifikacije Asus Xtion live 3D kamere i njenih senzora su prikazani u tabeli 8.

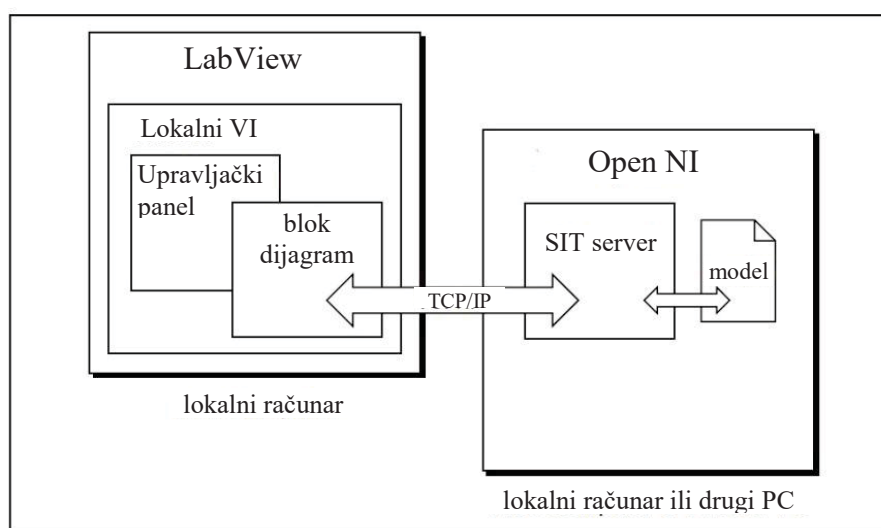
Tabela 8. Specifikacije Asus Xtion live 3D senzora

Domet kamere	Između 0.8m i 3.5m
Vidno polje	58° H, 45° V, 70° D
Senzor	RGB; Depth; Microphone*2
Rezolucija dubinske slike	VGA (640x480) : 30 fps QVGA (320x240): 60 fps
Rezolucija	SXGA (1280*1024)
Platforma	Intel X86 & AMD
Interfejs	USB2.0
Softver	OPEN NI SDK bundled

Programski jezik	C++/C# (Windows)
	C++(Linux)
	JAVA
Radno okruženje	Unutrašnjost
Dimenzije	18 x 3.5 x 5 cm

7.3. Upravljanje mobilnom robotskom platformom sa 3D senzorom

Asus Xtion PRO LIVE 3D senzor je povezan sa računarom preko USB 2.0 porta, dok je softverska veza sa računarom ostvarena preko OpenNI softvera („open source“ okruženje za 3D senzore). Podrška za 3D senzor sastoji se od IMAQ, Image, IR, Depth, Motion i Skeleton blokova, a u cilju postizanja „real-time“ akviziciju podataka sa komercijalnih 3D senzora kao što su Microsoft Kinect i Asus Xtion kamere. Za povezivanje DANI robota i Asus XtionPRO LIVE kamere, National Instruments nudi alat pod nazivom LabVIEW Simulation Interface Toolkit. Slika 7.9 prikazuje šematski dijagram veze.



Slika 7.9. Konekcija OpenNI-a i LabVIEW-a

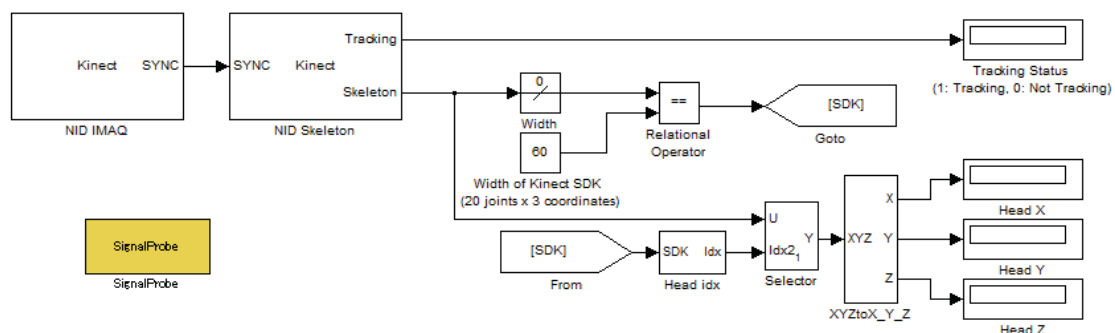
S obzirom na izrazito multidisciplinarni karakter naučnog istraživanja koji podrazumeva značajan uticaj matematičkih, mehaničkih, softverskih i senzorsko-elektronskih aspekata, razvijeni softver za upravljanje mobilnom robotskom platformom podjeljen je na dva modula.

7.3.1. Prepoznavanje i praćenje ljudi 3D senzorom

Prvi modul služi za detekciju ljudi u sceni 3D senzorom kojom se, u radnom scenariju, detektuje čovek i njegova pozicija u 3D koordinatnom sistemu. Ovaj modul se realizuje određenim algoritmima, tj. softverskim bibliotekama, koje su za ASUS Xtion Live 3D senzor identične softverskim bibliotekama koji se koriste kod Microsoft Kinect senzora. Razvijeni algoritmi su zbog toga prvenstveno namenjeni industriji zabave, ali se jednako uspešno mogu implementirati u robotski sistem.

Ovaj modul može imati ugrađen i neki filtar poput Kalmanovog filtera [8, 11, 21] ili koristiti rekurentnu neuronsku mrežu [4, 6, 14] za estimaciju i predikciju 3D položaja čoveka u sceni, čime se povećava robusnost celog sistema. 3D koordinate pozicije osobe koju prati robot šalju se do modula za upravljanje mobilnim robotom, gde se osnovu dobijenih podataka određuju se adekvatne ugaone brzine točkova.

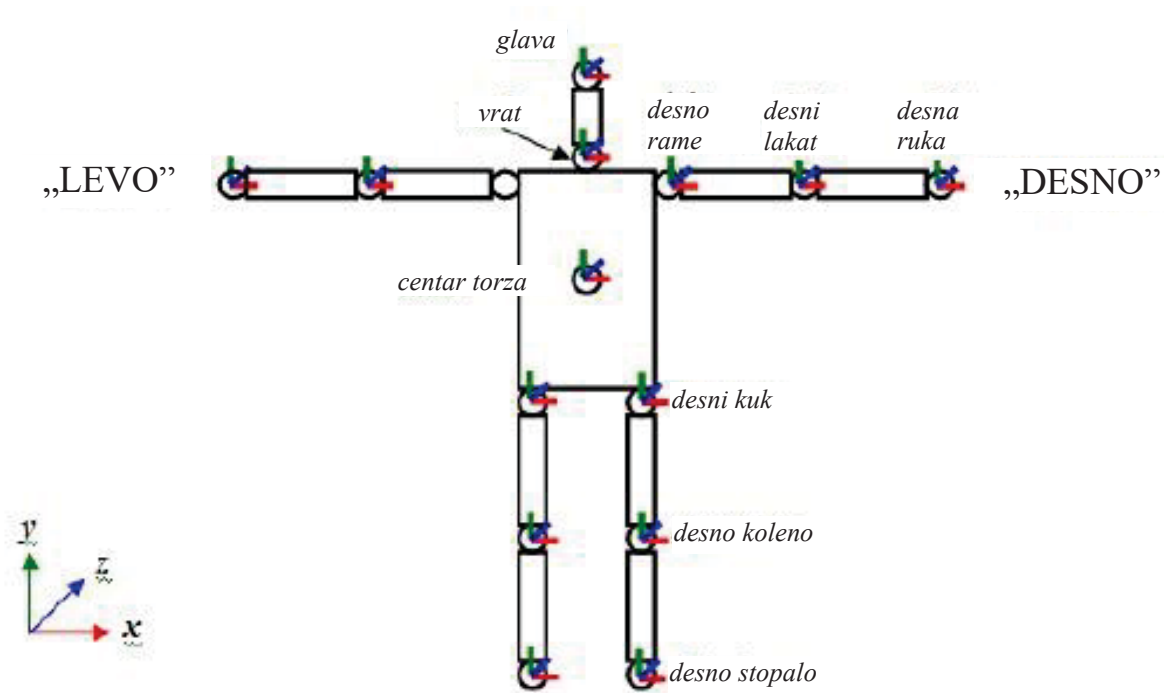
Na slici 7.10 je prikazan blok dijagram modula za detekciju ljudi koji vrši akviziciju podataka sa 3D senzora i detektuje centar mase glave čoveka.



Slika 7.10. Simulink Model za prikupljanje podataka i detekciju

Na početku blok dijagrama nalazi se "IMAQ" blok za vezu računara sa 3D senzorom, detekciju čoveka i akviziciju podataka. "Skeleton" blok na osnovu detektovanog modela čoveka daje kao izlaz tačke od interesa čoveka, odnosno prostorne koordinate zglobova, centra mase glave i centra mase torza čoveka (slika 7.11).

Ovi blokovi na osnovu mape dubine pronalaze čoveka i određuju poziciju skeleta. Skelet deli na objekte koje kasnije pojedinačno prati. Ovaj postupak ponavlja uvek kad primi novi set podataka. Skelet se pronalazi tako što se u toku programiranja kreira avatar koji odslikava ljudski kostur i u toku procesiranja mape dubine traže se podudaranja.



Slika 7.11 Tačke od interesa detektovane Skeleton blokom

7.4.2. Upravljanje mobilnom robotskom platformom na osnovu podataka sa 3D senzora u cilju praćenja ljudi

Modul za upravljanje kretanjem mobilnog robota sastoji se od dva PID kontrolera. Prvim se reguliše rastojanje od robota do čoveka, dok drugi služi za poravnanje sa smerom kretanja čoveka.

Rastojanje između čoveka i robota u koordinatnom sistemu robota je računato u svakom frejmu kao:

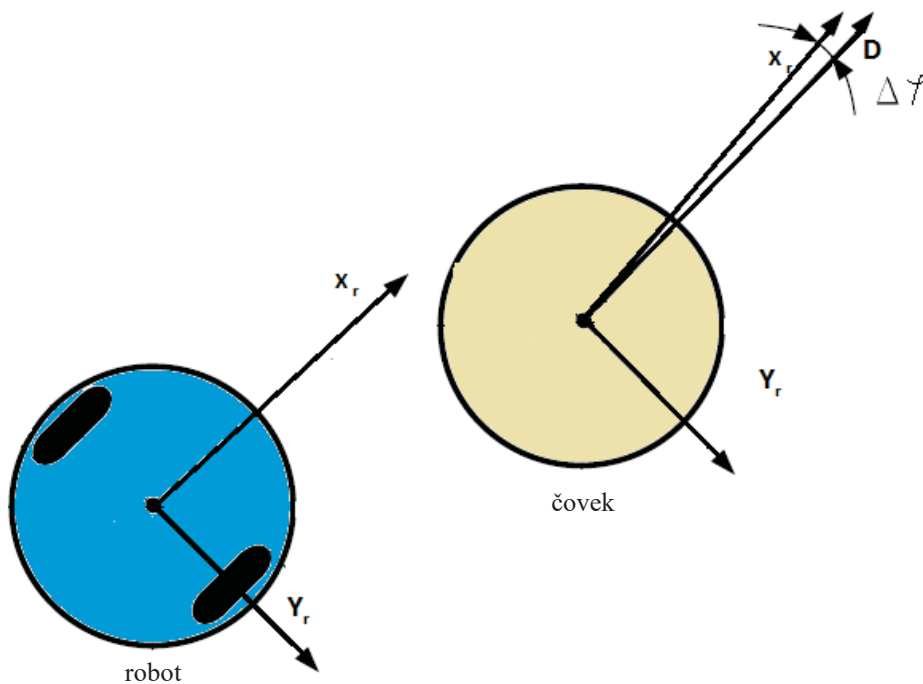
$$D = \sqrt{X_R^2 + Y_R^2 + Z_R^2}, \quad 7.3.$$

a razlika između željenog rastojanja između robota i čoveka i stvarnog rastojanja određenog prethodnim izrazom predstavlja ulaz u prvi PID kontroler.

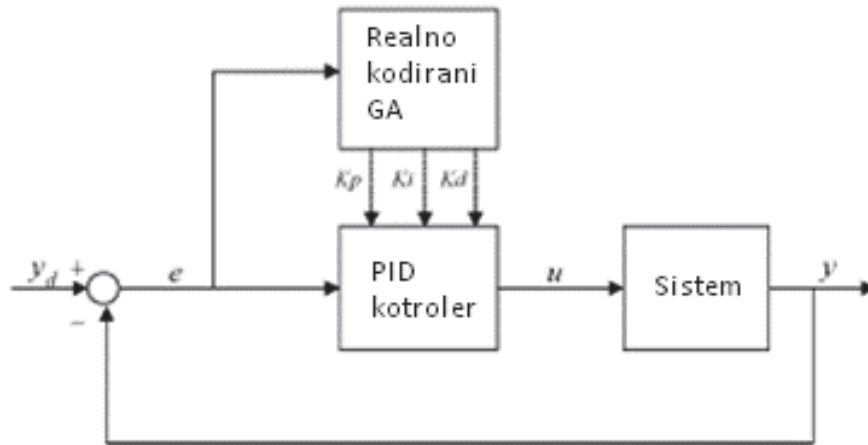
Drugi PID kontroler je korišćen da bi se robot poravnao sa smerom kretanja čoveka. Ulaz u PID kontroler predstavlja ugao između pravca rastojanja D i Y_r koordinate robota. Da bi se robot poravnao sa pravcem kretanja čoveka, ugao φ treba da bude jednak nuli (Slika 7.12).

$$\varphi = \arcsin \frac{Y_r}{\sqrt{X_r^2 + Y_r^2 + Z_r^2}}, \quad 7.4.$$

Parametri ova dva PID kontrolera mogu se odrediti nekom metaheurističkom optimizacionom metodom, a ideja se može naći u radu [71], gde se genetskim algoritmom određuju parametri dva PID kontrolera koja upravljaju složenim procesom sagorevanja. Šema podešavanja parametara PID kontrolera realno kodiranim genetskim algoritmom predstavljena je na slici 7.13. Greška e predstavlja razliku između željene vrednosti izlaza y_d i stvarne (izmerene) vrednosti izlaza y . Kako bi minimizovao grešku, funkcija cilja (fitnes funkcija) je najčešće integral (kod diskretnih sistema suma) kvadrata greške ili integral apsolutne vrednosti greške.



Slika 7.12 Šema praćenja



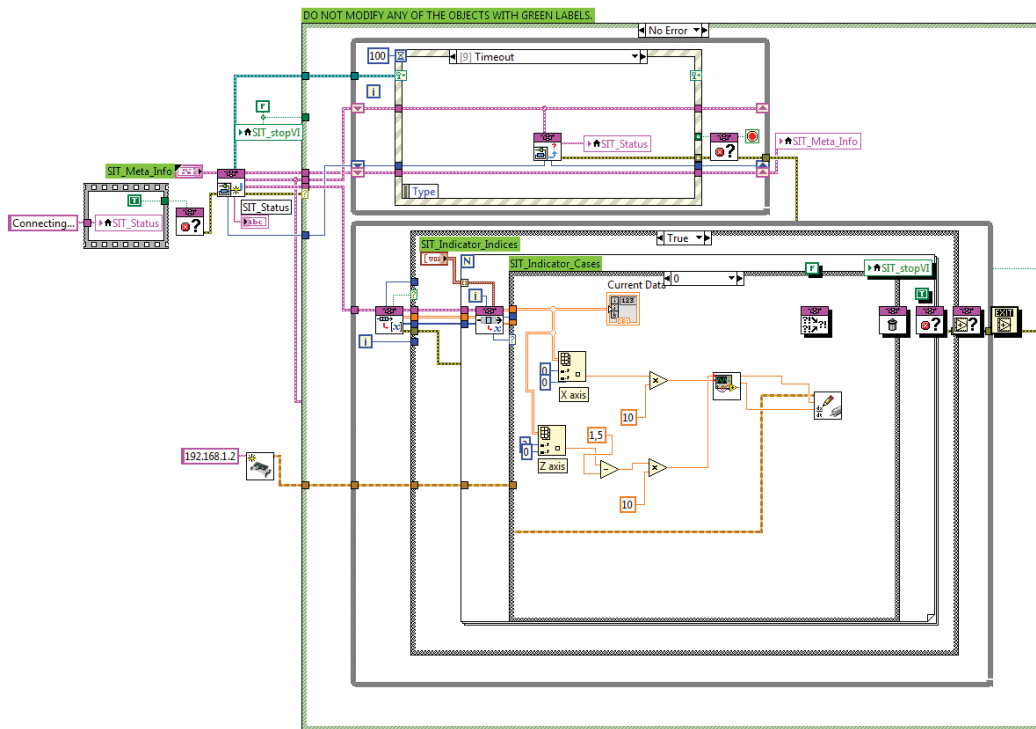
Slika 7.13 Šema genetske optimizacije parametara PID kontrolera

Vektor:

$$\Theta = [\theta_1, \theta_2, \theta_3, \theta_4, \theta_5, \theta_6] = [K_{p1}, K_{i1}, K_{d1}, K_{p2}, K_{i2}, K_{d2}], \quad 7.5.$$

čiji su elementi jednaki pojačanjima PID kontrolera predstavlja hromozom, dok su sama pojačanja $K_{p1}, K_{i1}, K_{d1}, K_{p2}, K_{i2}, K_{d2}$ zapravo geni koji čine potencijalna rešenja problema podešavanja parametara PID kontrolera primenom genetskih algoritama. Hromozom Θ sa realno kodiranim elementima evoluira iz generacije u generaciju primenom evolucionih operatora kako bi postigao bolje performanse u svakoj narednoj generaciji.

Integracijom razvijenih sistema sa sistemom za upravljanje kretanjem robota dobijen je sistem upravljanja koji daje okvir za interakciju mobilne robotske platforme sa okolinom. Mobilna robotska platforma prepoznaje i prati čoveka na osnovu algoritma prikazanog na slici 7.14.



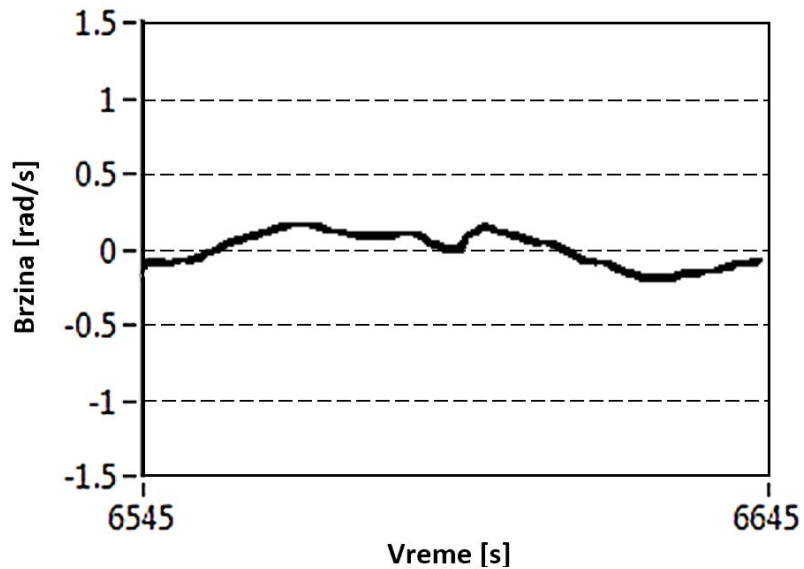
Slika 7.14. LabView program za upravljanje DANI robotom u cilju praćenja čoveka

Osim ovog pristupa, umesto PID kontrolera ovde se mogu naći fazi kontroleri, a podešavanje njihovih parametara se takođe može vršiti genetskim algoritmima [99,119]. Kod sistema upravljanja sa dva fazi kontrolera se za svaki može podešavati veći broj parametara, poput parametara normalizacije, denormalizacije, broja i oblika funkcija pripadnosti, kao i sami parametri funkcija pripadnosti. Kod ovako kompleksnih kontrolera metaheurističke metode podešavanja parametara su jedine koje mogu dati optimalno ili skoro optimalno rešenje. Funkcija cilja kod optimalnog podešavanja parametara kontrolera ovakve mobilne robotske platforme, ali i uopšte kompleksnih sistema, bez obzira na to koje kontrolere koristimo jeste minimizacija neke od karakteristika odziva sistema, poput greške ustaljenog stanja, vremena uspona, preskoka, vremena smirenja i sl. Moguća je i višekriterijumska optimizacija koja bi uzimala u obzir veći broj ovih karakteristika.

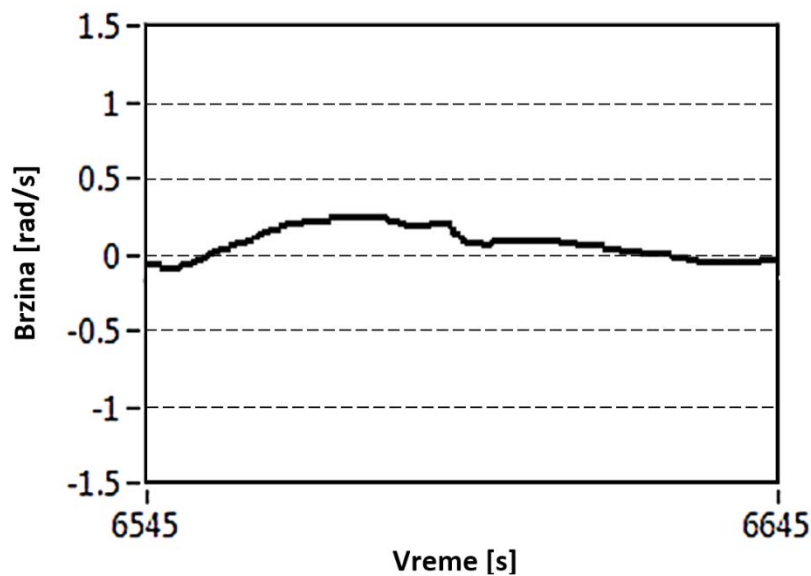
7.3.3. Verifikacija rezultata praćenja čoveka mobilnom robotskom platformom sa 3D senzorom

U ovom istraživanju je efektivnost mobilnog robotskog sistema testirana za radni scenario gde je cilj bio da mobilni robotski sistem da prepozna čoveka u svom okruženju i da ga prati, održavajući rastojanje između mobilne robotske platforme i čoveka od 1.5 m.

Promene ugaonih brzina na levog i desnog točku su prikazane na dijagramima na slici 7.15 i slici 7.16, dok je promena rastojanja između robota i čoveka prikazana na dijagramu na slici 7.17. Osim promene rastojanja između robota i čoveka na slici 18 je prikazana i ocenjena vrednost ove veličine proširenom Kalmanovim filterom.



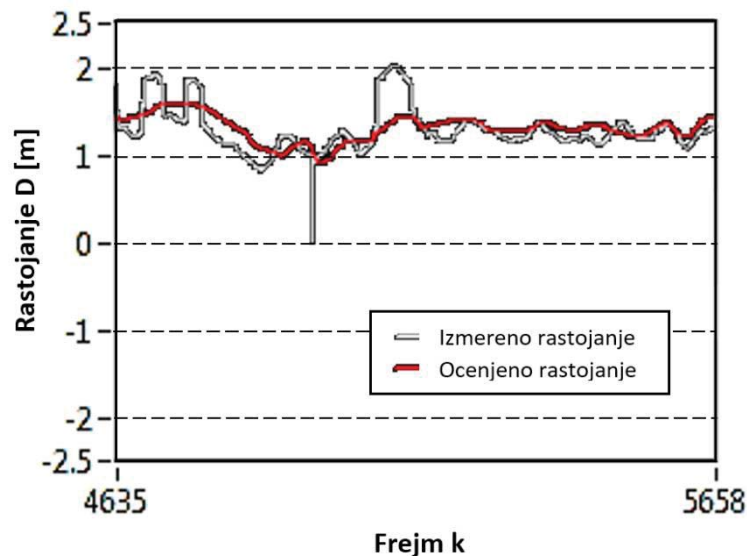
Slika 7.15 Brzina levog točka DaNI mobilnog robota



Slika 7.16 Brzina desnog točka DaNI mobilnog robota

Primenom savremenih metoda i algoritama, razvijen je upravljački sistem inteligentnog robotskog sistema sa senzorom stereo vizije koji ima sposobnosti detekcije, prepoznavanja i

prostornog lociranja. U ovom tehničkom rešenju postignuto je autonomno i pouzdano prepoznavanje čoveka i njegovo praćenje. Robusnost robotske vizije protiv promenljivih uslova osvetljenja postignuta je uz pomoć dodatnog modula sa Proširenim Kalmanovim filterom, omogućujući nesmetan rad sistema i slučajevima kada senzor vizije ne daje potrebne podatke usled poremećaja u sistemu. Eksperimentalni rezultati (slika 7.17) pokazuju da ovakva vrsta pristupa prilikom praćenja ljudi u robotici daje dobre rezultate.



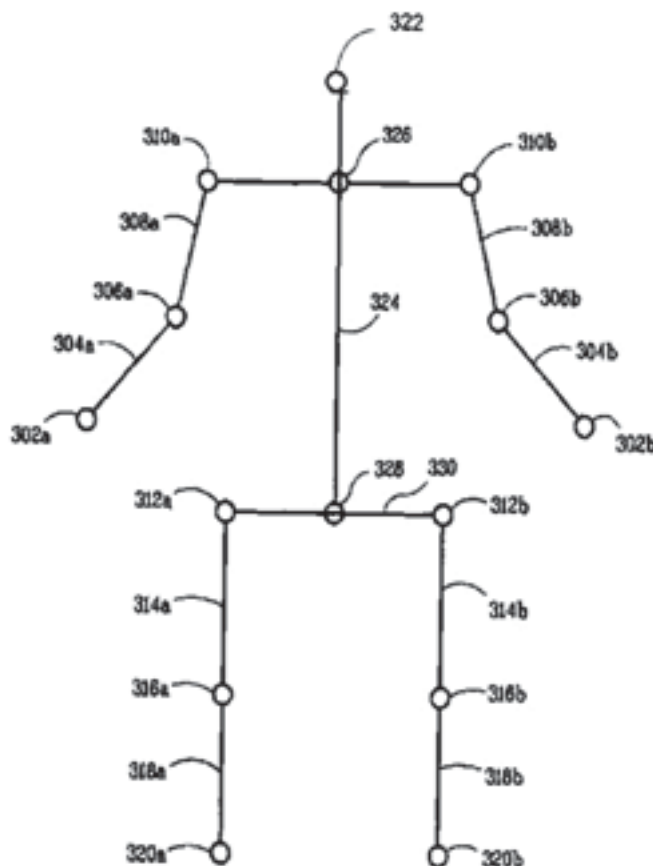
Slika 7.17 Distanca dobijena sistemom robotske vizije i distanca dobijena iz modula za praćenje

7.4. Prepoznavanje gestova 3D senzorom

Kao što je već napomenuto, algoritmi za obradu podataka dobijenih sa ASUS Xtion Live 3D senzora i Microsoft Kinect-a su identični i nalaze se u odgovarajućim softverskim bibliotekama. Microsoft je izbacivanjem Kinect 3D senzora načinio veliki iskorak i značajno unapredio praćenje ljudskih pokreta. U odnosu na konkurentske algoritme, njihov algoritam ima značajne prednosti i ovaj algoritam je ugrađen u njihov SDK (engl. Software Development Kit) tako da se može koristiti, ali ga ne možete izmeniti .

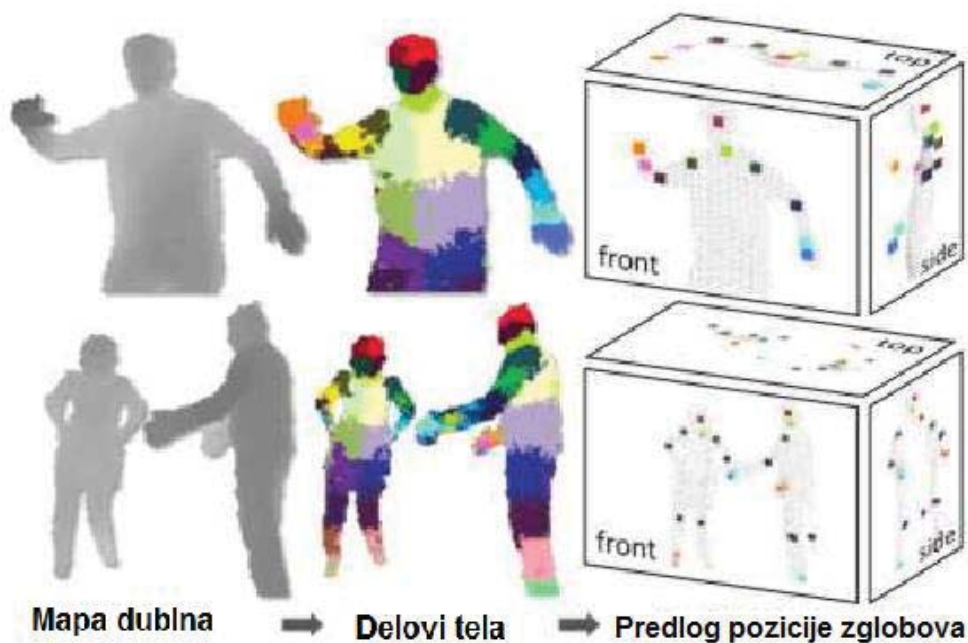
Microsoft-ovi inženjeri su osmislili izuzetno efikasan algoritam za transformaciju dubinske mape u skeletni model osobe koja koristi Kinect, odnosno osobe koja svojim pokretima zadaje komande uređaju, a sve zarad lakše i preciznije detekcije i obrade

ljudskih pokreta [120]. Prvo se iz dubinske mape oduzme pozadina, tj. u obzir se uzimaju pikseli koji pripadaju isključivo jednoj osobi. Ovo je mnogo lakše uraditi iz slike koja je dobijena uz pomoć (IR) dubinskog senzora, nego iz slike koja potiče od obične RGB kamere. Nakon ovog koraka, svakom pikselu se dodeli deo tela kojem pripada, na osnovu čega se kasnije formira skeletni model osobe. Glavna karakteristika ovog modela je što je ljudsko telo podeljeno na segmente, i tačke spajanja tih segmenata, tj. zglobove (body joints). Na primer, ruka se može posmatrati kao spoj tri segmenta: deo od lakta prema šaci, deo od lakta do ramena, i deo od ramena do vrata, uz dva zgloba lakat i rame. Na taj način, konačno, dobijamo model iz koga se veoma efikasno može odrediti pokret koji je upravo načinjen i to sa velikim brojem frejmova u sekundi (do 200 fps). Spoj ovog algoritma, koji je potekao iz Microsoft Research laboratorija, i PrimeSense patenta je ono što čini ovakve 3D senzore relativno jeftinim i pouzdanim uređajima ove vrste na tržištu danas. Na osnovu dubinske mape formira se model skeleta prikazan na slici 7.18



Slika 7.18 Slika skelete sa zglobovima [120]

Svaki deo tela je kodiran posebnom bojom i jedan deo njih se može preslikati u tačno jedan spoj (zglob) od interesa, dok ostali mogu poslužiti da se popune praznine ili se mogu iskoristiti za procenu lokacija drugih spojeva. To se lako uočava sa slike 7.19 - uzmimo za primer deo tela koji predstavlja koleno (svi pikseli u okolini su tamno ljubičasti), i nije teško uočiti njegov par iste boje u vidu zgloba na skeletnom modelu (skroz desno na slici). Sa druge strane, donji deo stomaka koji je predstavljen lila bojom, nema svog parnjaka u 3D modelu, jer jednostavno ne predstavlja nijedan "zglob" koji dozvoljava susednim segmentima da vrše pokrete.



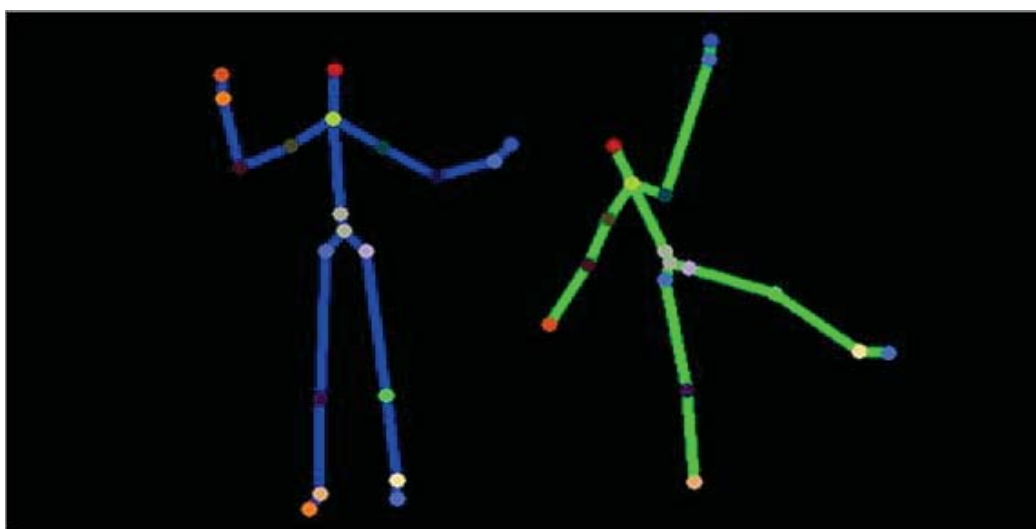
Slika 7.19 Proces formiranja zglobova

Očigledno, sve ovo ne bi imalo smisla kad ne bi postojao algoritam da se svakom pikselu dodeli pripadajući deo tela. Za svaki piksel x iz mape dubine I imamo koordinate u prostoru, jer smo to određivali uz pomoć IR dubinskog senzora, a ujedno i podatak o dubini tog piksela u prostoru (u odnosu na senzor). Ako funkcija $d(x, I)$ vraća podatak o dubini piksela x iz mape dubina I , onda će, za određene pomeraje u i v , funkcija:

$$f_{\theta}(I, \mathbf{x}) = d_I \left(\mathbf{x} + \frac{\mathbf{u}}{d_I(\mathbf{x})} \right) - d_I \left(\mathbf{x} + \frac{\mathbf{v}}{d_I(\mathbf{x})} \right) \quad 7.6.$$

predstavljati razliku između dubine piksela predstavljenim pomerajima $\Theta=(u, v)$ u odnosu na piksel x koji pokušavamo da klasifikujemo (da mu dodelimo pripadajući deo tela). Svaki pomeraj se normalizuje faktorom $1/d(x)$ da bi njegova vrednost bila nezavisna u odnosu na udaljenost piksela koji se posmatra. Normalizacijom, gornja formula je dubinski invarijantna i ne zavisi od udaljenosti posmatranog piksela u stvarnom prostoru. Što je piksel dalji od kamere, to će se pomeraj normalizovati na manju vrednost, ali će predstavljati isti razmak od posmatrane tačke u stvarnom 3D prostoru, i obrnuto. Ako pomeraj piksela pada na izolovanu pozadinu, ili izlazi izvan posmatrane slike, njegova dubina dobija neku veliku pozitivnu konstantu.

Inženjeri iz Microsoft-a su za potrebe ovog algoritma snimili čak 500 000 frejmova ljudi u pokretu u nekoliko stotina sesija trčanja, skijanja, igranja, upravljanja menijem i još nekoliko karakterističnih situacija u kojima će igrači učestvovati dok upravljaju nekim fikcionim karakterom u igri. I sve to u raznim varijetetima odeće, teksturama na odeći, itd. Na osnovu tih podataka, formiran je veliki broj stabala odlučivanja koja će uz gornju formulu izdejsstvovati finalni rezultat algoritma. Svaki piksel se propušta kroz određeni broj stabala odlučivanja (šumu stabala) gde će u zavisnosti od parametra stići do poslednjeg čvora (lista) svakog stabla. Kao rezultat celog algoritam dobijamo sliku 7.20 koja je dobijena iscrtavanjem skeleta.



Slika 7.20 Skelet iscrtan na osnovu podataka dobijenih od Kinect uređaja

Pri prijemu podataka o poziciji skeleta izvlačimo informacije o osobi čije pokrete pratimo. Kinect uređaj ima mogućnost praćenja više osoba, pa ako se iza ili pored korisnika nalazi neka osoba može doći do greške jer Kinect neće znati čije pokrete da prati. Sve informacije koje su nam potrebne od uređaja su zapravo pozicije određenih delova tela (Joints i JointsType). U koliko je potrebno jednostavno upravljanje mobilnom robotskom platformom gestovima, što podrazumeva komande tipa "prati me" (mahanje šakom od robota ka sebi), "stani" (pomeranje dlana ka robotu) ili "idi levo/desno" (odmahivanje levom rukom u levo, odnosno desnom rukom u desno), dovoljno je pratiti samo poziciju desne i leve šake. Joint klasa nam pruža informacije o poziciji u koordinatnom sistemu čije koordinate se kreću od -2 do 2 za sve tri ose pa te vrednosti treba skalirati na svetske(realne) koordinate. Za skaliranje se može koristiti biblioteka Coding4Fun.Kinect.Wpf koja pruža gotove metode za ova prevođenja.

U zavisnosti od same aplikacije i broja gestova koje robotska platforma sa 3D senzorom treba da prepozna, u cilju prepoznavanja gestova mogu se koristiti različiti klasifikatori. Klasifikator na osnovu podataka o poziciji, tj. promeni pozicije leve ili desne šake, detektuju odgovarajuće gestove koji služe za upravljanje robotom.

8. Zaključci i pravci daljih istraživanja

Izložena istraživanja sprovedena u toku izrade ove doktorske disertacije po sadržaju obuhvataju više aktuelnih naučnoistraživačkih pravaca koje se odnose na razvoj robotskih mobilnih platformi sa monokularnom kamerom, stereovizijskom kamerom, termovizijskom kamerom i 3D senzorom sa struktuiranom svetlošću, razvoj neuro-fazi-genetskih sistema za prepoznavanje i klasifikaciju objekata na osnovu slika dobijenih akvizicijom sa senzora robotske vizije, razvoj neuro-fazi-genetskih algoritama za detekciju, prostornu lokalizaciju i praćenje ljudi pomoću mobilne robotske platforme sa senzorom vizije, razvoj algoritama za optimizaciju sistema upravljanja baziranih na genetskim algoritmima i srodnim optimizacionim tehnikama iz domena veštačke inteligencije, uporednu analizu različitih inteligentnih algoritama za rešavanje problema detekcije i prepoznavanja objekata i ljudi, uporednu analizu različitih inteligentnih algoritama za rešavanje problema praćenja pokretnih objekata i ljudi, kao i razvoj novog integrisanog inteligentnog sistema upravljanja mobilnim robotom na osnovu neuro-fazi-genetskog prepoznavanja objekata i praćenja ljudi u robotskoj viziji.

Primenom savremenih metoda i algoritama iz domena veštačke inteligencije, moguće je razviti upravljački sistem inteligentnog robotskog sistema sa senzorom vizije koji ima sposobnosti detekcije, prepoznavanja i prostornog lociranja objekata i ljudi kao i odgovarajuću interakciju ovog robotskog sistema sa njima. Implementacijom više tehnika veštačke inteligencije, kao što su neuronske mreže, fazi logika i genetski algoritmi, u proces razvoja inteligentnog upravljačkog sistema moguće je obezbediti autonomno i pouzdano prepoznavanje objekata i ljudi i njihovo praćenje koje je superiorno u odnosu na postojeća rešenja. Razvojem hibridnog upravljačkog sistema inteligentnog mobilnog robota na bazi neuronskih mreža, fazi logike i genetskih algoritama moguće je ostvariti inteligentnu interakciju mobilne robotske platforme sa okolinom.

Doprinosi ove teze ogledaju se u razvoju većeg broja novih metodologija projektovanja sistema vizije, zasnovanih na tehnikama veštačke inteligencije, kao osnovnog dela sistema upravljanja kod mobilnih robotskih platformi za prepoznavanje objekata i praćenja ljudi.

Razvijeni su i veći broj algoritama zasnovanih na tehnikama računarske inteligencije i mašinskog učenja koji se mogu implementirati u sisteme robotske vizije i najviše nivoa upravljanja mobilnim robotom.

Razvijeni algoritmi vizije osnovni su element sistema upravljanja predstavljenim robotskim platformama i uspešno se mogu koristiti kako kod prepoznavanja objekata u cilju planiranja putanje manipulatora, tako i kod mobilnih robotskih platformi za prepoznavanje i praćenje ljudi. Predstavljene metodologije su zapravo osnova hijerarhijski najvišeg nivoa upravljanja zasnovanog na viziji mobilnim robotima i manipulatorima.

U okviru teze predstavljeni su sledeći doprinosi:

- Metodologija razvoja mobilne robotske platforme s monokularnom kamerom za detekciju i praćenje objekata;
- Nova metodologija razvoja sistema vizije i sistema upravljanja mobilne robotske platforme sa monokularnom kamerom zasnovane na neuro-fazi sistemu za prepoznavanje i praćenje objekata sa razvojem pratećih algoritama;
- Nova metodologija razvoja sistema vizije robotske platforme sa stereo kamerom zasnovane na tehnikama mašinskog učenja za prepoznavanje objekata i razvoj pratećih algoritama;
- Predlog novog inteligentnog adaptivnog algoritma segmentacije;
- Nova metodologija detekcije i praćenja ljudi mobilnom robotskom platformom sa stereo kamerom i razvoj pratećih algoritama zasnovanih na neuronskim mrežama i genetskim algoritmima;
- Predlog nove metodologije za optimalan izbor vektora ulaza u klasifikator.
- Nova metodologija detekcije i praćenja ljudi mobilnom robotskom platformom sa termovizijskom kamerom zasnovana na tehnikama računarske inteligencije i razvoj odgovarajućih novih algoritama segmentacije i klasifikacije;
- Nova metodologija estimacije i predikcije položaja čoveka u cilju praćenja zasnovana na nelinearnoj autoregresivnoj neuronskoj mreži;
- Razvoj nove mobilne robotske platforme i odgovarajućih upravljačkih algoritama za praćenje ljudi, prepoznavanje poza i gestova čoveka;
- Metodologija razvoja hijerarhijski najvišeg nivoa upravljanja zasnovanog na viziji kod mobilnih robota.

Pravci daljeg istraživanja podrazumevaju dodatnu optimizaciju algoritama i sistematizaciju predloženih rešenja. Prva ideja koja se javila a još uvek nije sprovedena u delo jeste optimalni izbor veličine bloka za formiranje mape dispariteta. Od ovog parametra zavisili kasnija segmentacija mape dispariteta i sami rezultati klasifikacije.

Algoritam segmentacije s povratnom spregom bi se dodatno mogao unaprediti ekspertnim sistemom preporuke koji bi preporučivao lokalne minimume koje je potrebno objediniti kako bi se formirala što vernija binarna segmentirana slika koja bi poslužila kao osnova za 3D rekonstrukciju.

Predložena metodologija optimalnog izbora ulaznog vektora klasifikatora može se dodatno unaprediti poboljšanim odabirom trening skupa, validacionog i test skupa. Kako je predstavljeno u odeljku 4.2.4, odabir trening skupa značajno može da utiče na rezultate klasifikacije, pa je realno upoređenje performansi klasifikatora moguće tek nakon većeg broja slučajnog odabira trening, validacionog i test skupa.

Kako se vidi, u ovoj tezi su predstavljeni različiti algoritmi klasifikacije, ali bi dodatno istraživanje moglo da podrazumeva odgovarajuću sistematizaciju, tj. primenu svih spomenutih klasifikatora na sve probleme, njihovu uporednu analizu performansi, čime bi se mogao preporučiti najbolji algoritam klasifikacije za svaku aplikaciju ponaosob.

Sistem vizije je kod mobilne robotske platforme sa termovizijskom kamerom razvijan isključivo za upotrebu u zatvorenom prostoru, gde su temperaturni uslovi relativno konstantni, ali nije obrađivan problem upotrebe na otvorenom, što je jedan od potencijalnih pravaca daljeg istraživanja koji može doneti veliki broj potencijlnih izazova.

Sistem praćenja ljudi zasnovan na nelinearnoj autoregresivnoj neuronskoj mreži još uvek nije implementiran u sistem za praćenje ljudi kod mobilne robotske platforme sa 3D sensorom sa struktuiranom svetlošću, što je sledeći korak u razvoju celovitog inteligentnog sistema upravljanja ovom mobilnom robotskom platformom. Kako su najveći problemi vezani za prepoznavanje i praćenje ljudi kod ove platforme prevaziđeni, ostaje da se razvijaju algoritmi za prepoznavanje akcija, poza i gestova, što značajno može unaprediti performanse ove robotske platforme.

9. Literatura

- [1] Ćirić, Ivan T; Ćojbašić, Žarko M; Nikolić, Vlastimir D; Igić, Tomislav S; Turšnek, Branko AJ; (2014), „Intelligent optimal control of thermal vision-based Person-Following Robot Platform“, *Thermal Science*, 18,3,957-966,2014, (M23)
- [2] Ćirić, Ivan; Ćojbašić, Žarko; Nikolić, Vlastimir; Živković, Predrag; Petković, Dušan; Tomić, Mladen; Tomić, Miša; „Thermal Vision Integration In Mobile Robot Vision System“, *Annals of the Faculty of Engineering Hunedoara-International Journal of Engineering*, 12,2,2014, (M51)
- [3] Ivan Ćirić, Žarko Ćojbašić, Miša Tomić, Milan Pavlović, Vukašin Pavlović, Ivan Pavlović, Vlastimir Nikolić (2012) Intelligent Control of DaNi Robot Based on Robot Vision and Object Recognition *Facta Universitatis, Series: Automatic Control And Robotics*, Vol.11, No 2, 2012 pp. 129 – 140 (M52)
- [4] Petrović, E., Ćojbašić, Ž., Ristić-Durrant, D., Nikolić, V., Ćirić, I., & Matić, S. (2013). Kalman Filter and Narx Neural Network for Robot Vision Based Human Tracking. *Facta Universitatis, Series: Automatic Control And Robotics*, Vol. 12, No 1, pp. 43-51. (M52)
- [5] Ivan Ćirić, Žarko Ćojbašić, Vlastimir Nikolic, Milica Ćirić, Predrag Živković, Ivan Pavlović, Emina Petrovic, INTELLIGENT CONTROL OF MOBILE ROBOT FOR OBJECT RECOGNITION AND GRASPING, 12th International Conference on Accomplishments in Electrical and Mechanical Engineering and Information Technology DEMI 2015, Banjaluka, BiH (2015) (M33)
- [6] Ivan Ćirić, Žarko Ćojbašić, Vlastimir Nikolić, Milica Ćirić, Mladen Tomić, Emina Petrović, Miloš Simonović, NEURAL NETWORK PREDICTION OF PERSON POSITION FOR HUMAN FOLLOWING MOBILE ROBOT PLATFORM, 12th International Conference on Accomplishments in Electrical and Mechanical Engineering and Information Technology DEMI 2015, Banjaluka, BiH (2015) (M33)

- [7] Ivan Ćirić, Žarko Čojbašić, Miloš Simonović, Milica Ćirić, Ivan Pavlović, Emina Petrović, Vlastimir Nikolić ,(2014) INTELLIGENT SYSTEM FOR OBJECT RECOGNITION IN STEREO VISION BASED ROBOTIC APPLICATIONS, ISSN 1313-4264, Sozopol, Bulgaria, pp.143-149. (M33)
- [8] Emina Petrović, Ivan Ćirić, Vlastimir Nikolić,(2014) ,HUMAN TRACKING BY A DANI MOBILE ROBOT BASED ON EXTENDED KALMAN FILTER, ISSN 1313-4264, Sozopol, Bulgaria, pp.136-142. (M33)
- [9] Ivan Ćirić , Žarko Čojbašić , Vlastimir Nikolić , Dragan Antić, Computationally Intelligent System for Thermal Vision People Detection and Tracking in Robotic Applications, Proceedings of the 11th International Conference on Telecommunications in Modern Satellite, Cable and Broadcasting Services - TELSIXS 2013, Niš, Serbia, 2013.
- [10] Ćirić, I., Čojbašić, Ž., Nikolić, V., Igić, T., Intelligent Control System for Thermal Vision-Based Person-Following Robot Platform, Proceedings, 16th Symposium on Thermal Science and Engineering SIMTERM, Sokobanja, Serbia, 2013, pp. 640-648
- [11] Petrović E., Tomić M., Nikolić V., Čojbašić Ž., Pavlović V., Ćirić I., (2013), Human Tracking With a Person Following Robot Based on Extended Kalman Filter, 2nd International Conference (Mechanical Engineering in the 21st century), Niš, Serbia. (M33)
- [12] Ćirić, Ž. Čojbašić, V. Nikolić, P. Živković, D. Petković, M. Tomić, M. Tomić Thermal Vision Integration in Mobile Robot Vision System, 11 International Conference on Accomplishments in Electrical and Mechanical Engineering and Information Technology DEMI 2013, Banjaluka, BiH (2013) (M33)
- [13] Ćirić, Ž. Čojbašić, M. Tomić, M. Pavlović, V. Pavlović, Computationally Intelligent Object Recognition for DaNI Robot Vision, Proceedings of XI International SAUM Conference, Niš, Serbia, 2012, pp.132 – 135, (M33)
- [14] Ž. Čojbašić, V. Nikolić, D. Ristić-Durant, I. Ćirić, E. Petrović, S. Matić, Neural Networks Based Human Tracking for Robot Vision, Proceedings of XI International SAUM Conference, Niš, Serbia, 2012, pp.164 – 167, (M33)
- [15] V. Nikolić, Ž. Čojbašić, D. Ristić-Durant, E. Petrović, S. Matić, I. Ćirić, Kalman Filter For Robot Vision-Based Human Tracking, Proceedings of 29th DANUBIA-ADRIA Symposium on Advances in Experimental Mechanics, Belgrade, Serbia, 2012., pp. 178-181 (M33)

- [16] Žarko Čojbašić, Danijela Ristić-Durrant, Ivan Ćirić, Sorin M. Grigorescu, Axel Gräser, Vlastimir Nikolić (2013) Reliable Computationally Intelligent Object Recognition and Human Tracking in Robotic Vision, Zbornik 57. konferencije ETRAN, Zlatibor, 3-6. juna 2013, str. RO3.6.1-5 (M63)
- [17] Milica Ćirić, Ivan Ćirić, Milan Gocić, Application of Naive Bayes Classifier in Stereo-Vision Based Object Recognition, The Second National Conference on Information Theory and Complex Systems, TINKOS 2014, Niš, Serbia, pp 39-40 (M63)
- [18] Žarko Čojbašić, Vlastimir Nikolić, Ivan Ćirić, Ivan Pavlović, Neuro-fazi klasifikator objekata kod robotske vizije složenog rehabilitacionog robotskog sistema FRIEND. (M85)
- [19] Ivan Ćirić, Žarko Čojbašić, Vlastimir Nikolić, Ivan Pavlović, Emina Petrović, Milica Ćirić, Miloš Simonović, Mobilna robotska platforma sa termovizijskom kamerom za detekciju i praćenje ljudi (M81)
- [20] Vlastimir Nikolić, Žarko Čojbašić, Ivan Ćirić, Miša Tomić, Emina Petrović, Andrija Milojević, Inteligentno upravljanje mobilnom robotskom platformom zasnovano na robotskoj viziji i prepoznavanju objekata (M85)
- [21] Vlastimir Nikolić, Žarko Čojbašić, Emina Petrović, Ivan Ćirić, Miša Tomić, Miloš Simonović, Vukašin Pavlović, Sistem za detekciju i praćenje ljudi mobilnim robotom zasnovan na robotskoj viziji i proširenom kalmanovom filteru (M85)
- [22] Danijela Ristic. Feedback Structures in Image Processing. PhD thesis, Bremen University, Institute of Automation, Bremen, Germany, 2007.
- [23] GONZALEZ, Rafael C.; WOODS, Richard E. Digital image processing 3rd edition. 2007.
- [24] GONZALEZ, Rafael C.; WOODS, Richard Eugene; EDDINS, Steven L. Digital image processing using MATLAB. Pearson Education India, 2004.
- [25] Adrian Leu, Robust Real-time Vision-based Human Detection and Tracking, PhD thesis, Bremen University, Institute of Automation, Bremen, Germany, 2014.
- [26] Szeliski R (2011) Computer Vision: Algorithms and Applications. Springer Verlag London
- [27] Qidwai, U., Chen, C. H., "Digital image processing: An Algorithmic Approach with MATLAB", CRC Press, 2010.
- [28] Wikipedia, en.wikipedia.org

- [29] Damir Filko, „Robusna lokalizacija mobilnog robota zasnovana na vizualnim obilježjima ravninskih segmenata“, doktorska disertacija, 2013
- [30] Hartley R, Zisserman A (2003) Multiple View Geometry in computer vision. Cambridge University Press
- [31] Marques, Oge. Practical image and video processing using MATLAB. John Wiley & Sons, 2011.
- [32] LU, Tong, et al. Video Preprocessing. In: *Video Text Detection*. Springer London, 2014. p. 19-47.
- [33] H.S. Koppula, R.Gupta, A.Saxena, Learning human activities and object affordance from RGB-D videos, *IJRR*, 2013, 32(8), 951-970
- [34] J. Sinapov, T. Bergquist, C. Schenck, U. Ohiri, S. Griffith, A. Stoytchev, Interactive object recognition using proprioceptive and auditory feedback, *IJRR*, 2011, 30(10), 1250-1262
- [35] B. Rasolzadeh, M.Bjorkman, K. Huebner, D. Kragic, An Active Vision System for detecting, Fixating and Manipulating objects in the Real World, *IJRR*, 2010, 29(2-3), 133-154
- [36] W.H.Li, L. Kleeman, Segmentation and modeling of visually symmetric objects by robot actions, *IJRR*, 2011, 30(9), 1124-1142
- [37] C. Papazov, S. Haddadin, S. Parusel, K. Krieger, D. Burschka, Rigid 3D geometry matching for grasping of known objects in cluttered scenes, *IJRR*, 2012, 31(4), 538-553
- [38] Kaijen Hsiao, L.P. Kaelbling, T. Lozano-Perez, Robust Grasping under Object Pose Uncertainty, *Auton Robot* (2011), 31:253-268
- [39] K.Parsa, J. Angeles, A.K. Misra, Rigid-body pose and twist estimation using an accelerometer array, *Archive of Applied Mechanics* 74(2004) 223-236
- [40] Z.M. Bi, L. Wang, Advances in 3D data acquisition and processing for industrial applications, *Robotics and Computer-Integrated Manufacturing*, 2010, 26:403-413
- [41] Paolo Bellandi, Giovanna Sansoni, Angelo Vertuan, Development and characterization of a multi-camera 2D-vision system for enhanced performance of a drink serving robotic cell, *Robotics and Computer-Integrated Manufacturing*, Volume 28, Issue 1, February 2012, Pages 35-49, ISSN 0736-5845
- [42] Ashutosh Saxena, Justin Driemeyer, Justin Kearns, and Andrew Y. Ng. Robotic Grasping of Novel Objects, *Neural Information Processing Systems (NIPS 19)*, 2007.

- [43] Ivan Volosyak, Oleg Ivlev, and Axel Graeser. Rehabilitation robot FRIEND II - the general concept and current implementation. In Proc. of the IEEE 9th Int. Conf.on Rehabilitation Robotics ICORR 2005, pages 540-544, Chicago, USA, 2005
- [44] Sorin M. Grigorescu, Ristić-Durrant Danijela, Vuppala K. Sai, Gräser Axel, Closed-Loop Control in Image Processing for Improvement of Object Recognition, Proceedings of the 17th World Congress The International Federation of Automatic Control, Seoul, Korea, July 6-11, 2008
- [45] Grigorescu, Sorin M; Ristić-Durrant, Danijela; Gräser, Axel; „ROVIS: Robust machine vision for service robotic system FRIEND, "Intelligent Robots and Systems, 2009. IROS 2009. IEEE/RSJ International Conference on" „,3574-3581, 2009, IEEE
- [46] Grigorescu Sorin, Robust Machine Vision for Service Robotics, 2010., phd teza
- [47] Natarajan, S.K.; Ristic-Durrant, D.; Leu, A.; Graser, A., "Robust stereo-vision based 3D modelling of real-world objects for assistive robotic applications," Intelligent Robots and Systems (IROS), 2011 IEEE/RSJ International Conference on , vol., no., pp.786,792, 25-30 Sept. 2011
- [48] Ristić-Durrant D., Grigorescu S.M., Gräser A., Čojbašić Ž., Nikolić V. (2011), Robust Stereo-Vision Based 3D Object Reconstruction for the Assistive Robot FRIEND, Advances in Electrical and Computer Engineering, Issue 4, Year 2011, 15 – 22, DOI: 10.4316/AECE.2011.04003.
- [49] Gräser A., Heyer T., Fotoohi L., Lange U., Kampe H., Enjarini B., Heyer S., Fragkopoulos C., Ristić-Durrant D., A Supportive FRIEND at Work: Robotic Workplace Assistance for the Disabled, Robotics & Automation Magazine, IEEE, vol.20, no.4, pp.148-59, Dec. 2013.
- [50] Gräser A., Kuzmicheva O., Ristić-Durrant D., Natarajan S., Fragkopoulos C., Vision-based Control of Assistive Robot FRIEND: Practical Experiences and Design Conclusions, at – Automatisierungstechnik, Vol. 60, Nr. 5 (2012) , p. 297-308, 2012.
- [51] CUPEC, Robert; NYARKO, Emmanuel Karlo; GRBIC, Ratko. Detection of Dominant Planar Surfaces in Disparity Images Based on Random Sampling. *Tehnicki Vjesnik*, 2011, 18.3: 301-308.
- [52] SEZGIN, Mehmet, et al. Survey over image thresholding techniques and quantitative performance evaluation. *Journal of Electronic imaging*, 2004, 13.1: 146-168.

- [53] MARSHALL, David; LUKACS, Gabor; MARTIN, Ralph. Robust segmentation of primitives from range data in the presence of geometric degeneracy. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 2001, 23.3: 304-314.
- [54] Toshev, Alexander, Ben Taskar, and Kostas Daniilidis. "Object detection via boundary structure segmentation." *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on*. IEEE, 2010.
- [55] LEU, Adrian; AITEANU, Dorin; GRÄSER, Axel. High speed stereo vision based automotive collision warning system. In: *Applied Computational Intelligence in Engineering and Information Technology*. Springer Berlin Heidelberg, 2012. p. 187-199.
- [56] MIKOLAJCZYK, Krystian; SCHMID, Cordelia. A performance evaluation of local descriptors. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 2005, 27.10: 1615-1630.
- [57] LOWE, David G. Object recognition from local scale-invariant features. In *Computer vision, 1999. The proceedings of the seventh IEEE international conference on (Vol. 2, pp. 1150-1157)*. 1999.
- [58] O'HARA, Stephen; DRAPER, Bruce A. Introduction to the bag of features paradigm for image classification and retrieval. *arXiv preprint arXiv:1101.3354*, 2011.
- [59] VIOLA, Paul; JONES, Michael. Rapid object detection using a boosted cascade of simple features. In: *Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on*. IEEE, 2001. p. I-511-I-518 vol. 1.
- [60] Abbasgholipour, M., Omid, M., Keyhani, A., & Mohtasebi, S. S. (2011). Color image segmentation with genetic algorithm in a raisin sorting system based on machine vision in variable conditions. *Expert Systems with Applications*, 38(4), 3671-3678.
- [61] Mullen, Rob J., Dorothy N. Monekosso, and Paolo Remagnino. "Ant algorithms for image feature extraction." *Expert Systems with Applications* 40.11 (2013): 4315-4332.
- [62] Portela, N. M., Cavalcanti, G. D., & Ren, T. I. (2014). Semi-supervised clustering for MR brain image segmentation. *Expert Systems with Applications*, 41(4), 1492-1497.
- [63] Zhao, F., Fan, J., & Liu, H. (2014). Optimal-selection-based suppressed fuzzy c-means clustering algorithm with self-tuning non local spatial information for image segmentation. *Expert systems with applications*, 41(9), 4083-4093.

- [64] Tao, W., Jin, H., & Liu, L. (2007). Object segmentation using ant colony optimization algorithm and fuzzy entropy. *Pattern Recognition Letters*, 28(7), 788-796.
- [65] GÜLER, Inan; ÜBEYLI, Elif Derya. Adaptive neuro-fuzzy inference system for classification of EEG signals using wavelet coefficients. *Journal of neuroscience methods*, 2005, 148.2: 113-121.
- [66] NAZMY, T. M.; EL-MESSIRY, H.; AL-BOKHITY, B. Adaptive neuro-fuzzy inference system for classification of ECG signals. In: *Informatics and Systems (INFOS)*, 2010 The 7th International Conference on. IEEE, 2010. p. 1-6.
- [67] SHARMA, Minakshi. Artificial Neural Network Fuzzy Inference System (ANFIS) For Brain Tumor Detection. arXiv preprint arXiv:1212.0059, 2012.
- [68] LAHMIRI, Salim; BOUKADOUM, Mounir. Comparison of ANFIS and SVM for the classification of brain MRI Pathologies. In: *Circuits and Systems (MWSCAS)*, 2011 IEEE 54th International Midwest Symposium on. IEEE, 2011. p. 1-4.
- [69] BHARDWAJ, Anant; SIDDHU, Kapil Kumar. An Approach to Medical Image Classification Using Neuro Fuzzy Logic and ANFIS Classifier, *International Journal of Computer Trends and Technology- volume4 Issue3- 2013*, pp. 236-240
- [70] HOSSEINI, Monireh Sheikh; ZEKRI, Maryam. Review of medical image classification using the adaptive neuro-fuzzy inference system. *Journal of medical signals and sensors*, 2012, 2.1: 49.
- [71] Čojbašić, Ž. M., Nikolić, V. D., Ćirić, I. T., & Čojbašić, Lj. R. (2011). Computationally intelligent modeling and control of fluidized bed combustion process. *Thermal science*, 15(2), 321-338.
- [72] GASCUEÑA, José Manuel; FERNÁNDEZ-CABALLERO, Antonio. Agent-oriented modeling and development of a person-following mobile robot. *Expert Systems with Applications*, 2011, 38.4: 4280-4290.
- [73] PETROVIĆ, Emina, et al. V, "Stereo Vision-Based Human Tracking for Robotic Follower,". *International Journal of Advanced Robotic Systems*, 2013.
- [74] Hu, C., Ma, X., Dai, X., & Qian, K. (2010). Reliable people tracking approach for mobile robot in indoor environments. *Robotics and Computer-Integrated Manufacturing*, 26(2), 174-179.
- [75] Alcantarilla, P. F., Stasse, O., Druon, S., Bergasa, L. M., & Dellaert, F. (2013). How to localize humanoids with a single camera?. *Autonomous Robots*, 34(1-2), 47-71.

- [76] Muñoz-Salinas, R., Aguirre, E., & García-Silvente, M. (2007). People detection and tracking using stereo vision and color. *Image and Vision Computing*, 25(6), 995-1007.
- [77] Benezeth, Y., Jodoin, P. M., Emile, B., Laurent, H., & Rosenberger, C. (2010). Human detection with a multi-sensors stereovision system. In *Image and Signal Processing* (pp. 228-235). Springer Berlin Heidelberg.
- [78] Bertozzi, M., Binelli, E., Broggi, A., & Rose, M. D. (2005, June). Stereo vision-based approaches for pedestrian detection. In *Computer Vision and Pattern Recognition-Workshops, 2005. CVPR Workshops. IEEE Computer Society Conference on* (pp. 16-16). IEEE.
- [79] Kwak, S., Ko, B., & Byun, H. (2007). Salient human detection for robot vision. *Pattern Analysis and Applications*, 10(4), 291-299.
- [80] Wang, Y., & Morrison, G. (2009, February). Automatic human detecting and tracking using stereo vision technique. In *IS&T/SPIE Electronic Imaging* (pp. 723803-723803). International Society for Optics and Photonics
- [81] Muñoz-Salinas, R., Aguirre, E., García-Silvente, M., & Gonzalez, A. (2005). People detection and tracking through stereo vision for human-robot interaction. In *MICAI 2005: Advances in Artificial Intelligence* (pp. 337-346). Springer Berlin Heidelberg.
- [82] Tang, C. Y., Chen, Z., & HUNG, Y. P. (2000). Automatic detection and tracking of human heads using an active stereo vision system. *International Journal of Pattern Recognition and Artificial Intelligence*, 14(02), 137-166.
- [83] PINTO, Andry Maykol; ROCHA, Luís F.; MOREIRA, A. Paulo. Object recognition using laser range finder and machine learning techniques. *Robotics and Computer-Integrated Manufacturing*, 2013, 29.1: 12-22.
- [84] HE, B. W.; LONG, Zhou Xiao; LI, You Fu. The research of an automatic object reconstruction method based on limit visible region of the laser-scanning vision system. *Robotics and Computer-Integrated Manufacturing*, 2010, 26.6: 711-719.
- [85] Cherubini, Andrea, and François Chaumette. "Visual navigation of a mobile robot with laser-based collision avoidance." *The International Journal of Robotics Research* 32.2 (2013): 189-205.
- [86] Bellotto, N., & Hu, H. (2010). Computationally efficient solutions for tracking people with a mobile robot: an experimental evaluation of Bayesian filters. *Autonomous Robots*, 28(4), 425-438.

- [87] KOHAVI, Ron; JOHN, George H. Wrappers for feature subset selection. *Artificial intelligence*, 1997, 97.1: 273-324.
- [88] GUYON, Isabelle; ELISSEEFF, André. An introduction to variable and feature selection. *The Journal of Machine Learning Research*, 2003, 3: 1157-1182.
- [89] PECK, Charles C.; DHAWAN, Atam P.; MEYER, Claudia M. Genetic algorithm based input selection for a neural network function approximator with applications to SSME health monitoring. In: *Neural Networks, 1993., IEEE International Conference on. IEEE*, 1993. p. 1115-1122.
- [90] D'HEYGERE, Tom; GOETHALS, Peter; DE PAUW, Niels. Use of Genetic Algorithms to select Input Variables in Artificial Neural Network Models for the Prediction of Benthic Macroinvertebrates. In: *Integrated Assessment and Decision Support, Proceedings of the First Biennial Meeting of the International Environmental Modelling and Software Society/Rizzoli, AE and Jakeman, AJ,(eds.). Volume 2, 2002.-ISBN 88-900787-0-7. 2002. p. 136-141.*
- [91] CATENI, Silvia; COLLA, Valentina; VANNUCCI, Marco. A Genetic Algorithm-based approach for selecting input variables and setting relevant network parameters of a SOM-based classifier. In: *International Journal of Simulation Systems, Science & Technology. UKSim 4th European Modelling Symposium on Mathematical modelling and computer simulation. 2011.*
- [92] Antonio Fernández-Caballero, José Carlos Castillo, Javier Martínez-Cantos, Rafael Martínez-Tomás, "Optical flow or image subtraction in human detection from infrared camera on mobile robot", *Robotics and Autonomous Systems*, 58, 1273–1281, 2010.
- [93] Andr'e Treptow, Grzegorz Cielniak, Tom Duckett, "Real-time people tracking for mobile robots using thermal vision", *Robotics and Autonomous Systems*, 54, pp. 729–739, 2006.
- [94] St-Laurent, Louis, D. Prévost, and Xavier Maldague. "Thermal imaging for enhanced foreground-background segmentation." *Int. Conf. on Quantitative InfraRed Thermography*. 2006.
- [95] BRUNNER, Christopher, et al. Selective combination of visual and thermal imaging for resilient localization in adverse conditions: Day and night, smoke and fire. *Journal of Field Robotics*, 2013, 30.4: 641-666.
- [96] K. Senthil Kumar, G. Kavitha, R. Subramanian and G. Ramesh, "Visual and Thermal Image Fusion for UAV Based Target Tracking" in "MATLAB - A Ubiquitous Tool

- for the Practical Engineer", 2011., book edited by Clara M. Ionescu, ISBN 978-953-307-907-3
- [97] Dudić, S. P., Ignjatović, I. M., Šešlija, D. D., Blagojević, V. A., & Stojiljković, M. M. Leakage quantification of compressed air on pipes using thermovision. *Thermal Science*, 16 (2012). suppl. 2, pp. 555-565
- [98] Jin Guofeng, Zhang Wei, Yang Zhengwei, Huang Zhiyong¹, Song Yuanjia, Wang Dongdong, Tian Gan¹ "Image Segmentation of Thermal Waving Inspection based on Particle Swarm Optimization Fuzzy Clustering Algorithm" *Measurement Science Review*. Volume 12, Issue 6, Pages 296–301
- [99] V. Nikolić, D. Antić, Ž. Čojbašić, I. Ćirić, S. Nikolić, S. Perić, Evolutionary Computation in Hybrid Aeroturbine Fuzzy Control, Proceedings of XI International SAUM Conference, Niš, Serbia, 2012, pp.176 – 184, invited paper
- [100] C. Cortes and V. Vapnik, Support-Vector Networks, *Machine Learning*, 20 (1995), 3, pp.273-297
- [101] Soyguder, S. (2011). Intelligent control based on wavelet decomposition and neural network for predicting of human trajectories with a novel vision-based robotic. *Expert Systems with Applications*, 38(11), 13994-14000.
- [102] Elfring, J., Van De Molengraft, R., & Steinbuch, M. (2014). Learning intentions for improved human motion prediction. *Robotics and Autonomous Systems*, 62(4), 591-602.
- [103] Malgireddy, M. R., Nwogu, I., & Govindaraju, V. (2013). Language-motivated approaches to action recognition. *The Journal of Machine Learning Research*, 14(1), 2189-2212.
- [104] Wang, Y., Tran, D., Liao, Z., & Forsyth, D. (2012). Discriminative hierarchical part-based models for human parsing and action recognition. *The Journal of Machine Learning Research*, 13(1), 3075-3102.
- [105] Chaaoui, A. A., Climent-Pérez, P., & Flórez-Revuelta, F. (2012). A review on vision techniques applied to human behaviour analysis for ambient-assisted living. *Expert Systems with Applications*, 39(12), 10873-10888.
- [106] Pellegrini, S., & Iocchi, L. (2006). Human posture tracking and classification through stereo vision. In *Proc. of Intl. Conf. on Computer Vision Theory & Applications*.
- [107] Cooper, H., Ong, E. J., Pugeault, N., & Bowden, R. (2012). Sign language recognition using sub-units. *The Journal of Machine Learning Research*, 13(1), 2205-2231.

- [108] Nayak, S., Duncan, K., Sarkar, S., & Loeding, B. (2012). Finding recurrent patterns from continuous sign language sentences for automated extraction of signs. *The Journal of Machine Learning Research*, 13(1), 2589-2615.
- [109] Lui, Y. M. (2012). Human gesture recognition on product manifolds. *The Journal of Machine Learning Research*, 13(1), 3297-3321.
- [110] Burger, B., Ferrané, I., Lerasle, F., & Infantes, G. (2012). Two-handed gesture recognition and fusion with speech to command a robot. *Autonomous Robots*, 32(2), 129-147.
- [111] Ćirić, Ivan T., et al. "Air quality estimation by computational intelligence methodologies." *Thermal Science* 16.suppl. 2 (2012): 493-504.
- [112] DIACONESCU, Eugen. The use of NARX neural networks to predict chaotic time series. *WSEAS Transactions on Computer Research*, 2008, 3.3: 182-191.
- [113] OANCEA, Bogdan; CIUCU, Ştefan Cristian. Time series forecasting using neural networks. arXiv preprint arXiv:1401.1333, 2014.
- [114] RAMESH, K.; SHUKOR, SR Abd; AZIZ, N. Nonlinear model predictive control of a distillation column using NARX model. *Computer Aided Chemical Engineering*, 2009, 27: 1575-1580.
- [115] AKKAYA, Ramazan; AYDOGDU, Omer; CANAN, Suleyman. An ANN based NARX GPS/DR System for Mobile Robot Positioning and Obstacle Avoidance. *Journal of Automation and Control*, 2013, 1.1: 6-13.
- [116] JAEGER, Herbert. Tutorial on training recurrent neural networks, covering BPPT, RTRL, EKF and the "echo state network" approach. *GMD-Forschungszentrum Informationstechnik*, 2002.
- [117] DOULAMIS, Nikolaos; DOULAMIS, Anastasios; VARVARIGOU, Theodora. Adaptable neural networks for modeling recursive non-linear systems. In: *Digital Signal Processing*, 2002. DSP 2002. 2002 14th International Conference on. IEEE, 2002. p. 1191-1194.
- [118] YEREDOR, Arie. Yule-Walker equations applied to Hessians of the characteristic function for improved AR estimation. In: *Acoustics, Speech and Signal Processing*, 2007. ICASSP 2007. IEEE International Conference on. IEEE, 2007. p. III-1025-III-1028.

- [119] Živković, P. M., Nikolić, V. D., Ilić, G. S., Čojbašić, Ž. M., & Ćirić, I. T. (2012). Hybrid soft computing control strategies for improving the energy capture of a wind farm. *Thermal Science*, 16(suppl. 2), 483-491.
- [120] Kinect for Windows Sensor Components and Specifications
<https://msdn.microsoft.com/en-us/library/jj131033.aspx>

BIOGRAFIJA

Ivan Ćirić, asistent Mašinskog fakulteta Univerziteta u Nišu, rođen je 06.06.1980. godine u Prokuplju.

Završio je osnovnu školu „Rodoljub Čolaković” u Nišu i Gimnaziju „Svetozar Marković” u Nišu sa odličnim uspehom.

Školske 1999/2000. otpočeo je studije na Mašinskom fakultetu Univerziteta u Nišu, a iste je završio 2004. godine. Diplomirao je na profilu Mehatronika i upravljanje sa prosečnom ocenom 9,16 u toku studija i ocenom 10 na diplomskom radu.

Na magistarskim studijama koje je upisao 2004. godine na Mašinskom fakultetu Univerziteta u Nišu na studijskom programu „automatsko upravljanje“ položio je sa sve ispite sa prosečnom ocenom 10 i 2010. godine odbranio magistarski rad pod nazivom „Neuro-fazi-genetsko modeliranje i upravljanje procesom sagorevanja“.

Od 2005. godine angažovan je od strane Mašinskog fakulteta Univerziteta u Nišu na realizaciji naučno-istraživačkih i nastavnih projekata i izvođenju nastavno-obrazovnog procesa na Katedri za mehatroniku i upravljanje, najpre kao stipendista Ministarstva nauke i tehnološkog razvoja Republike Srbije, a od 2009. godine kao asistent. Kao asistent, učestvovao je u izvođenju nastavno-obrazovnog procesa na više od 20 predmeta.

Ivan Ćirić je do sada publikovao više od 80 naučnih rezultata, a učestvovao je u realizaciji pet nacionalnih naučno-istraživačkih projekata i osam međunarodnih nastavnih/bilateralnih projekata. Tokom svoje profesionalne karijere, više puta je bio na studijskom boravku i stručnom usavršavanju u okviru DAAD, CEEPUS i TEMPUS programa.

Član je asocijacije Srbije za sisteme, automatsko upravljanje i SAUM.

Oženjen je, otac jednog deteta i živi u Nišu.

ИЗВЕШТАЈ О ОЦЕНИ ДОКТОРСКЕ ДИСЕРТАЦИЈЕ

ПОДАЦИ О КАНДИДАТУ

Презиме, име једног родитеља и име	Ђирић Топлица Иван
Датум и место рођења	06.06.1980., Прокупље, Србија

Основне студије

Универзитет	Универзитет у Нишу
Факултет	Машински факултет
Студијски програм	Мехатроника
Звање	Дипломирани инжењер машинства
Година уписа	1999.
Година завршетка	2004.
Просечна оцена	9,16

Магистарске студије, магистарске студије

Универзитет	Универзитет у Нишу
Факултет	Машински факултет
Студијски програм	Аутоматско управљање
Звање	Магистар машинских наука
Година уписа	2004.
Година завршетка	2010.
Просечна оцена	10,00
Научна област	Аутоматско управљање
Наслов завршног рада	Неуро-фази-генетско моделирање и управљање процесом сагоревања

Докторске студије

Универзитет	-
Факултет	-
Студијски програм	-
Година уписа	-
Остварен број ЕСПБ бодова	-
Просечна оцена	-

НАСЛОВ ТЕМЕ ДОКТОРСКЕ ДИСЕРТАЦИЈЕ

Наслов теме докторске дисертације	Интелигентно управљање мобилним роботима на основу неуро-фази-генетског препознавања објеката и праћења људи у роботској визији
Име и презиме ментора, звање	др Жарко Ђојбашић, редовни професор
Број и датум добијања сагласности за тему докторске дисертације	8/20-01-001/13-019 29.01.2013.

ПРЕГЛЕД ДОКТОРСКЕ ДИСЕРТАЦИЈЕ

Број страна	170
Број поглавља	8
Број слика (шема, графикона)	98
Број табела	8
Број прилога	-

УНИВЕРЗИТЕТ У НИШУ			
Примљено	18.09.2015		
Орг. јед.	Број	Портал	Вредност
	612-577	115	

**ПРИКАЗ НАУЧНИХ И СТРУЧНИХ РАДОВА КАНДИДАТА
који садрже резултате истраживања у оквиру докторске дисертације**

Р. бр.	Аутор-и, наслов, часопис, година, број волумена, странице	Категорија
1	<p>Ćirić, Ivan T; Ćojbašić, Žarko M; Nikolić, Vlastimir D; Igić, Tomislav S; Turšnek, Branko AJ; (2014) "Intelligent optimal control of thermal vision-based Person-Following Robot Platform", Thermal Science, 18,3,957-966,2014,</p> <p><i>Кратак опис садржине (до 100 речи)</i></p> <p>У овом раду представљен је развој интелигентног система управљања мобилном роботском платформом која као основни извор информација о окоolini користи термовизијску камеру. Проблем детекције човека у окружењу мобилног робота реализује се сегментационим алгоритмом чији су параметри оптимално подешени генетским алгоритмом и предложен је класификатор заснован на методи потпорних вектора.</p>	M23
2	<p>Ćirić Ivan, Ćojbašić Žarko, Miša Tomić, Milan Pavlović, Vukašin Pavlović, Ivan Pavlović, Vlastimir Nikolić (2012) "Intelligent Control of DaNi Robot Based on Robot Vision and Object Recognition", Facta Universitatis, Series: Automatic Control And Robotics, Vol.11, No 2, 2012 pp. 129 – 140</p> <p><i>Кратак опис садржине (до 100 речи)</i></p> <p>Представљено је једно решење проблема праћења објеката у роботској визији. Развијена је мобилна роботска платформа са монокуларном камером, као и одговарајући алгоритми софтверска решења потребна за препознавање и праћење објеката. Развијени алгоритми засновани су на алатима из домена рачунарске интелигенције, и експериментални резултати показују њихову применљивост у решавању представљених проблема.</p>	M52
3	<p>Pavlović, I., Ćirić, I., Djekić, P., Nikolić, V., Pavlović, R., Ćojbašić, Ž., & Radenković, G. (2013). Rheological model optimization using advanced evolutionary computation for the analysis of the influence of recycled rubber on rubber blend dynamical behaviour. Meccanica, 48(10), 2467-2477.</p> <p><i>Кратак опис садржине (до 100 речи)</i></p> <p>Решавању проблема подешавања параметара реолошког модела је у овом раду приступљено на неконвенционалан начин, применом метахеуристичког оптимизационог алгоритма – реално кодираног генетског алгоритма. Посебна пажња посвећена је дефинисању функције циља оптимизације и самом одабиру параметара оптимизационог алгоритма. Развој реолошких модела са оптимално подешеним параметрима омогућава једноставнију анализу утицаја присуства рециклиране гуме у гуменој смеси на динамичке карактеристике саме смесе. Развијена методологија генетске оптимизације комплексних система је у предметној тези примењена на проблеме сегментације слике.</p>	M21
4	<p>Ćirić, I. T., Ćojbašić, Ž. M., Nikolić, V. D., Živković, P. M., & Tomić, M. A. (2012). Air quality estimation by computational intelligence methodologies. Thermal Science, 16(suppl. 2), 493-504.</p> <p><i>Кратак опис садржине (до 100 речи)</i></p> <p>На основу вишемесечног мерења атмосферских параметара и параметара који дефинишу квалитета ваздуха формирана је конвенционална неуронска мрежа са простирањем сигнала унапред и нелинеарна ауторегресивна неуронска мрежа за предвиђање квалитета ваздуха. У овом раду је представљено и упоређење ове две мреже за предикцију временске серије и наведен је низ предности нелинеарне ауторегресивне неуронске мреже у односу на конвенционалну неуронску мрежу при естимацији и предикцији временских серија. Предикција временских серија се као проблем јавља и у тези кандидата, те је у циљу решавања овог проблема имплементиран алгоритам развијен у описаном раду</p>	M23
5	<p>Živković, P. M., Nikolić, V. D., Ilić, G. S., Ćojbašić, Ž. M., Ćirić, I. T. (2012). Hybrid soft computing control strategies for improving the energy capture of a wind farm. Thermal Science, 16(suppl. 2), 483-491.</p> <p><i>Кратак опис садржине (до 100 речи)</i></p>	M23

	<p>Аутори су у овом раду представили фази-управљачки систем за управљање пољем ветротурбина, који има низ предности у односу на конвенционалне управљачке алгоритме, али доноси и изазове који се првенствено огледају у оптималном подешавању фази-управљачке структуре. Како би превазишли овај проблем, аутори предлажу примену реално кодираних генетских алгоритама, и представљају симулационе резултате који верификују овакав приступ решавању проблема управљању пољем ветротурбина. Примена фази контролера чији су параметри подешени реално кодираним генетским алгоритима су разматрани и у разматраној докторској дисертацији.</p>	
6	<p>Ćirić, Ivan; Ćojbašić, Žarko; Nikolić, Vlastimir; Živković, Predrag; Petković, Dušan; Tomić, Mladen; Tomić, Miša; (2014) “Thermal Vision Integration In Mobile Robot Vision System., Annals of the Faculty of Engineering Hunedoara-International Journal of Engineering”, 12,2,2014,</p> <p><i>Кратак опис садржине (до 100 речи)</i></p> <p>У циљу праћења људи у роботској визији предложена је интеграција термовизијске камере у систем визије. Овакав приступ захтева развој интелигентног алгоритма обраде слике, па је представљена фази сегментација термовизијске слике, као и неуронска мрежа за класификацију сегментисаних објеката.</p>	M51
7	<p>Petrović, E., Ćojbašić, Ž., Ristić-Durrant, D., Nikolić, V., Ćirić, I., Matić, S. (2013). “Kalman Filter and Narx Neural Network for Robot Vision Based Human Tracking” Facta Universitatis, Series: Automatic Control and Robotics, Vol. 12, No 1, pp. 43-51.</p> <p><i>Кратак опис садржине (до 100 речи)</i></p> <p>Праћење људи у роботској визији може се унапредити имплементацијом одговарајућих филтера који би помогли у оцени положаја људи, као и предвиђању положаја људи у наредном фрејму. У ту сврху се у овом раду користе Калманов филтер и ауторегресивна неуронска мрежа, и упоређени су њихови резултати.</p>	M52
8	<p>Ćirić Ivan , Žarko Ćojbašić, Vlastimir Nikolic, Milica Ćirić, Predrag Živković, Ivan Pavlović, Emina Petrovic, (2015) “Intelligent Control of Mobile Robot for Object Recognition And Grasping”, 12th International Conference on Accomplishments in Electrical and Mechanical Engineering and Information Technology DEMI 2015, Banjaluka, BiH</p> <p><i>Кратак опис садржине (до 100 речи)</i></p> <p>У овом раду разматран је проблем прихватања објеката од стране роботске руке, где робот као основни извор информација користи стерео камеру. Представљени су алгоритми сегментације мапе диспаратета која се формира на основу стерео слике, као и интелигентни алгоритми класификације објеката. Ови подаци користе се касније за планирање путање робота и само управљање моторима који покрећу зглобове роботске руке.</p>	M33
9	<p>Ćirić Ivan, Žarko Ćojbašić, Vlastimir Nikolić, Milica Ćirić, Mladen Tomić, Emina Petrović, Miloš Simonović, (2015) “Neural Network Prediction of Person Position for Human Following Mobile Robot Platform”, 12th International Conference on Accomplishments in Electrical and Mechanical Engineering and Information Technology DEMI 2015, Banjaluka, BiH</p> <p><i>Кратак опис садржине (до 100 речи)</i></p> <p>Мобилна роботска платформа која прати људе задатке праћења може да оствари много успешније уколико у сваком тренутку предвиђа наредну позицију човека. На овај начин се заправо смањује област претраге слике наредног фрејма у коме ће се наћи човек, што значајно убрзава извршење самог алгоритма препознавања и праћења људи стерео камером. Како се проблем своди на предвиђање временских серија, у овом раду је предложено решење засновано на имплементацији неуронске мреже за ово предвиђање.</p>	M33
10	<p>Ćirić Ivan, Žarko Ćojbašić, Miloš Simonović, Milica Ćirić, Ivan Pavlović, Emina Petrović, Vlastimir Nikolić, (2014) “Intelligent System for Object Recognition in Stereo Vision Based Robotic Applications”, ISSN 1313-4264, Sozopol, Bulgaria, pp.143-149.</p> <p><i>Кратак опис садржине (до 100 речи)</i></p>	M33

	<p>Препознавање објеката у роботској визији применом стерео камере није нов проблем, али савремена достигнућа у области вештачке интелигенције отварају нове могућности имплементације нових алгоритама у циљу повећања поузданости и робусности роботског система. У раду су представљени алгоритми сегментације једне слике применом алгоритма сегментације с повратном спрегом као и сегментације мапе диспаратета, док се сегментисани бинарни објекти класификују класификаторима заснованим на неуронској мрежи и класификатором заснованим на неуро-фази структури. Експериментални резултати представљени у раду омогућили су упоређење перформанси ова два класификатора.</p>	
11	<p>Ćirić Ivan, Žarko Ćojbašić, Vlastimir Nikolić, Dragan Antić, (2013) "Computationally Intelligent System for Thermal Vision People Detection and Tracking in Robotic Applications", Proceedings of the 11th International Conference on Telecommunications in Modern Satellite, Cable and Broadcasting Services - TELSIXS 2013, Niš, Serbia, 2013.</p> <p><i>Кратак опис садржине (до 100 речи)</i></p> <p>У овом раду представљени су напредни алгоритми засновани на вештачкој интелигенцији и машинском учењу за детекцију, локализацију и праћење људи мобилном роботском платформом са термовизијском камером. Интелигентног алгоритми обраде слике подразумевали су пре-процесирање, фази сегментацију термовизијске слике и алгоритам за разумевање термовизијске слике. Како би се одредили бинарни сегментисани објекти који представљају људе на слици, предложен је класификатор заснован на неуро-фази структури.</p>	M33
12	<p>Ćirić, I., Ćojbašić, Ž., Nikolić, V., Igić, T., (2013) "Intelligent Control System for Thermal Vision-Based Person-Following Robot Platform", Proceedings, 16th Symposium on Thermal Science and Engineering SIMTERM, Sokobanja, Serbia, 2013, pp. 640-648</p> <p><i>Кратак опис садржине (до 100 речи)</i></p> <p>Идеја примене генетских алгоритама за оптимално одређивање интервала прага сегментације слике први пут је представљена у овом раду. Овакав вид интелигентне сегментације примењен је на мобилну роботску платформу за праћење људи која информације о околини добија помоћу термовизијске камере. У раду су представљени и развијени алгоритми класификације којим се детектује човек у окружењу, и одређује се просторна локација човека у сцени, што представља основ за развој система управљања мобилном роботском платформом.</p>	M33
13	<p>Ćirić, Ž. Ćojbašić, M. Tomić, M. Pavlović, V. Pavlović, (2012) "Computationally Intelligent Object Recognition for DaNI Robot Vision", Proceedings of XI International SAUM Conference, Niš, Serbia, 2012, pp.132 – 135</p> <p><i>Кратак опис садржине (до 100 речи)</i></p> <p>Овај рад анализира примену интелигентних алгоритама развијених применом алата вештачке интелигенције и машинског учења у циљу решавања проблема детекције и праћења објеката у роботској визији. Развијена мобилна роботска платформа користи монокуларну камеру као извор информација, а представљени су и одговарајући алгоритми и софтверска решења потребна за препознавање и праћење објеката.</p>	M33
14	<p>Ivan Ćirić, Žarko Ćojbašić, Vlastimir Nikolić, Ivan Pavlović, Emina Petrović, Milica Ćirić, Miloš Simonović, "Mobilna robotska platforma sa termovizijskom kamerom za detekciju i praćenje ljudi"</p> <p><i>Кратак опис садржине (до 100 речи)</i></p> <p>Ово техничко решење обједињује различите идеје праћења људи мобилном роботском платформом са термовизијском камером које су представљене у већем броју радова аутора техничког решења. Представљен је развој мобилне роботске платформе, интеграција термовизијске камере у систем визије, као и развој и имплементација одговарајућих управљачких алгоритама и алгоритама за обраду слике. Сви имплементирани алгоритми су хибридни и настали комбиновањем већег броја интелигентних алгоритама.</p>	M81

ИСПУЊЕНОСТ УСЛОВА ЗА ОДБРАНУ ДОКТОРСКЕ ДИСЕРТАЦИЈЕ

Кандидат испуњава услове за оцену и одбрану докторске дисертације који су предвиђени Законом о високом образовању, Статутом Универзитета и Статутом Факултета.

ДА НЕ

Образложење

Кандидат је магистар техничких наука, има одобрену тему докторске дисертације, објавио је већи број научних и стручних радова и поднео докторску дисертацију одговарајуће садржине, обима и квалитета, у складу са одобреном темом докторске дисертације.

ВРЕДНОВАЊЕ ПОЈЕДИНИХ ДЕЛОВА ДОКТОРСКЕ ДИСЕРТАЦИЈЕ

Кратак опис појединих делова дисертације *(до 500 речи)*

Разматрана теза састоји се из 8 поглавља. Претходе им резиме на српском и енглеском језику, садржај и листе табела и слика. На крају се налази списак 120 референци и биографија аутора.

У уводу је представљен проблем истраживања, представљени су основни циљеви истраживања, као и досадашње публикације аутора. Истакнути су најзначајнији сегменти дисертације и дат је преглед садржине осталих поглавља. Кандидат је у свом истраживању користио 4 различита сензора визије и то монокуларну камеру, стерео камеру, термовизијску камеру и 3Д сензор са структурираном светлошћу. Део истраживања спроведен је на Институту за аутоматику Универзитета у Бремену, где је кандидат вршио истраживања користећи рехабилитациони робот и мобилну роботску платформу за праћење људи.

У другом поглављу су представљене основе компјутерске, односно роботске визије које подразумевају аквизицију и обраду слике. У трећем поглављу представљени су напредни алгоритми обраде слике и стерео слике у циљу препознавања објеката. Ови алгоритми подразумевају формирање мапе диспаратитета и сегментационе алгоритме који су били основа за даље унапређење алатима из домена рачунарске интелигенције.

У четвртном поглављу представљене су две роботске платформе код којих је основни задатак био развити алгоритме за препознавање објеката. У ту сврху развијени су АНФИС класификатор и класификатор заснован на неуронским мрежама. Осим тога, у овом поглављу су предложена унапређења постојећег алгоритма класификације с повратном спрегом која подразумевају интелигентну адаптивну сегментацију.

Пето поглавље бави се развојем интелигентног система за детекцију људи заснованог на неуронској мрежи за класификацију. Овде је представљен и нови концепт примене генетског алгоритма за одабир оптималног вектора карактеристика као улаза у класификатор.

У шестом поглављу се као алтернатива најчешће коришћеним сензорима визије предлаже примена термовизијске камере. Развијена су два алгоритма сегментације од којих се један базира на фази логици, док се код другог параметри подешавају генетским алгоритмима. Представљен је и метод потпорних вектора за класификацију сегментисаних објеката на људе и оне који то нису.

Седмо поглавље се бави проблемом праћења људи и препознавања њихових гестова. Најпре је представљена нелинеарна ауто-регресивна неуронска мрежа за предикцију и естимацију положаја човека која се може применити на експерименталној роботској платформи представљеној у петој глави. Потом је представљена роботска платформа са 3Д сензором, код које је препознавање људи решено хардверски, па је било само потребно развити систем праћења, управљања и препознавања гестова човека како би се омогућила успешна интеракција робота и човека.

У осмом поглављу су представљени закључци и правци даљих истраживања.

ВРЕДНОВАЊЕ РЕЗУЛТАТА ДОКТОРСКЕ ДИСЕРТАЦИЈЕ

Ниво остваривања постављених циљева из пријаве докторске дисертације *(до 200 речи)*

Циљеви постављени у пријави докторске дисертације су остварени, уз поштовање предложеног оквирног садржаја дисертације.

У оквиру дисертације, приказан је већи број сродних методологија развоја система визије и одговарајућих управљачких алгоритама код мобилних роботских платформи са монокуларном камером, стереовизијском камером, термовизијском камером и 3Д сензором са структурираном светлошћу. Представљена истраживања су по садржају обухватала више актуелних научноистраживачких праваца од роботске визије и дигиталне обраде слике, преко вештачке интелигенције и машинског учења, до филтрирања, предикције и естимације временских серија.

Резултати истраживања представљени у разматраној тези потврђују да је применом савремених метода и алгоритама из домена вештачке интелигенције могуће развити управљачки систем интелигентног роботског система са сензором визије који има способности детекције, препознавања и просторног лоцирања објеката и људи као и одговарајућу интеракцију овог роботског система са њима. Имплементацијом више техника вештачке интелигенције, као што су неуронске мреже, фази логика и генетски алгоритми, у процес развоја интелигентног управљачког система могуће је обезбедити аутономно и поуздано препознавање објеката и људи и њихово праћење које је супериорно у односу на постојећа решења.

Вредновање значаја и научног доприноса резултата дисертације (до 200 речи)

Обрађивана тема докторске дисертације је веома значајна и актуелна, како у научном смислу, тако и смислу практичне применљивости. Поднета докторска дисертација представља оригиналан и вредан научни и стручни допринос кандидата. Научни допринос разматраног рукописа и објављених радова се пре свега огледа у следећем:

- Дефинисана је нова методологија развоја система визије и система управљања мобилне роботске платформе са монокуларном камером засноване на неуро-фази систему за препознавање и праћење објеката и развијени су пратећих алгоритми и софтвер;
- Дефинисана је нова методологија развоја система визије роботске платформе са стерео камером засноване на техникама машинског учења за препознавање објеката и предложен је нови интелигентни адаптивни алгоритам сегментације;
- Представљена је нова методологија детекције и праћења људи мобилном роботском платформом са стерео камером, као и развој пратећих алгоритама и софтвера заснованих на неуронским мрежама и генетским алгоритмима;
- Предложена је нова методологија за оптималан избор вектора улаза у класификатор;
- Дефинисана је нова методологија детекције и праћења људи мобилном роботском платформом са термовизијском камером заснована на техникама рачунарске интелигенције и представљен је развој одговарајућих нових алгоритама сегментације и класификације слике добијене термовизијском камером;
- Развијени су алгоритми и одговарајућа софтверска решења за естимацију и предикције положаја човека у циљу праћења, који су засновани на нелинеарној ауторегресивној неуронској мрежи;
- Представљена је методологија развоја новог система визије код мобилне роботске платформе са 3Д сензором са структурираном светлешћу, као и развој одговарајућих управљачких алгоритама за праћење људи, препознавање поза и гестова човека.

Оцена самосталности научног рада кандидата (до 100 речи)

Кандидат мр Иван Ћирић показао је значајно теоријско и практично знање, као и висок ниво самосталности, систематичности и креативности у бављењу научно-истраживачким радом. Кандидат је приказао детаљну, свеобухватну и квалитетну анализу постојеће научне литературе из области теме докторске дисертације. Познавање литературе и стечена знања из више области је искористио да на креативан начин осмисли, формулише и примени софистицирани научни приступ развоја управљачких алгоритама мобилних роботских платформи са сензорима визије који је заснован на комбиновању различитих метода из домена вештачке интелигенције и машинског учења. Креирао је одговарајуће алгоритме и софтверска решења која омогућавају примену предложених методологија.

Неки од научних резултата представљених у разматраној тези презентирани су у оквиру већег броја научних радова који су штампани у међународним и домаћим часописима и представљени су на међународним и домаћим конференцијама, те штампани у зборницима радова. Осим тога, потребно је истаћи да су се неки од резултата представљених у овој тези нашли у 4 техничка решења, што још једном истиче практичну применљивост предложених решења.

ЗАКЉУЧАК (до 100 речи)

На основу прегледа поднете радне верзије докторске дисертације и увидом у публиковане научне радове кандидата, чланови Комисије за оцену и одбрану докторске дисертације закључују следеће:

- Поднети рукопис одговара теми докторске дисертације одобреној од стране Наставно научног већа Машинског факултета у Нишу и Научно стручног већа Универзитета у Нишу.

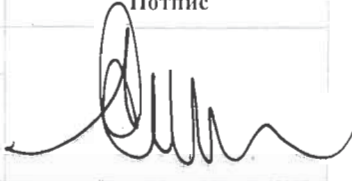
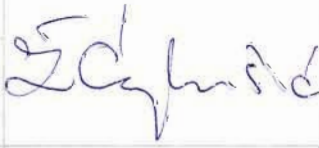



- Докторска дисертација представља оригиналан и вредан научни и стручни допринос веома актуелној и значајној проблематици управљања мобилним роботима заснованом на роботској визији.
- Научни допринос и оригиналност дисертације показани су објављивањем већег броја радова и техничких решења.
- Докторска дисертација је адекватно конципирана и технички квалитетно урађена.
- Резултати приказаног научног рада имају висок степен општости и применљивости.
- Кандидат поседује висок ниво теоријских и практичних знања из више области потребних за решавање комплексних проблема интелигентног управљања мобилним роботима заснованог на роботској визији а и добро је упознат са досадашњим научним достигнућима.
- Кандидат је показао висок ниво самосталности и систематичности у бављењу научно-истраживачким радом, као и креативан приступ формулацији и решавању разматраних проблема.

Имајући у виду напред наведено, Комисија предлаже Наставно научног већу Машинског факултета у Нишу да се поднети рукопис кандидата **мр Ивана Ђирића**, дипломираног инжењера машинства, под називом:

„ИНТЕЛИГЕНТНО УПРАВЉАЊЕ МОБИЛНИМ РОБОТИМА НА ОСНОВУ НЕУРО-ФАЗИ-ГЕНЕТСКОГ ПРЕПОЗНАВАЊА ОБЈЕКТА И ПРАЋЕЊА ЉУДИ У РОБОТСКОЈ ВИЗИЈИ“

прихвати као докторска дисертација, а кандидат позове на усмену јавну одбрану.

КОМИСИЈА

Број одлуке ННВ о именовану Комисије		612-508-6/2015	
Датум именовања Комисије		25.08.2015	
Р. бр.	Име и презиме, звање		Потпис
1.	др Драган Антић, редовни професор	председник	
	Аутоматика (Научна област)	Универзитет у Нишу, Електронски факултет (Установа у којој је запослен)	
2.	др Жарко Ђојбашић, редовни професор	ментор, члан	
	Аутоматско управљање и роботика (Научна област)	Универзитет у Нишу, Машински факултет (Установа у којој је запослен)	
3.	др Властимир Николић, редовни професор	члан	
	Аутоматско управљање и роботика (Научна област)	Универзитет у Нишу, Машински факултет (Установа у којој је запослен)	
4.	др Данијела Ристић-Durrant, доцент	члан	
	Аутоматско управљање и роботика (Научна област)	Универзитет у Нишу, Машински факултет (Установа у којој је запослен)	
5.	др Стеван Станковски, редовни професор	члан	
	Мехатроника, роботика и аутоматизација (Научна област)	Универзитет у Новом Саду, Факултет техничких наука (Установа у којој је запослен)	

Датум и место:
У Нишу,
18.09.2015.



Универзитет у Нишу

Изјава 1.

ИЗЈАВА О АУТОРСТВУ

Изјављујем да је докторска дисертација, под насловом

Интелигентно управљање мобилним роботима на основу неуро-фази-генетског
препознавања објеката и праћења људи у роботској визији

која је одбрањена на Машинском факултету Универзитета у Нишу:

- резултат сопственог истраживачког рада;
- да ову дисертацију, ни у целини, нити у деловима, нисам пријављивао/ла на другим факултетима, нити универзитетима;
- да нисам повредио/ла ауторска права, нити злоупотребио/ла интелектуалну својину других лица.

Дозвољавам да се објаве моји лични подаци, који су у вези са ауторством и добијањем академског звања доктора наука, као што су име и презиме, година и место рођења и датум одбране рада, и то у каталогу Библиотеке, Дигиталном репозиторијуму Универзитета у Нишу, као и у публикацијама Универзитета у Нишу.

У Нишу, 16.12.2015.

Аутор дисертације: Иван Т Ћирић

Потпис аутора дисертације:

Иван Т Ћирић



Универзитет у Нишу

Изјава 2.

**ИЗЈАВА О ИСТОВЕТНОСТИ ШТАМПАНОГ И ЕЛЕКТРОНСКОГ ОБЛИКА
ДОКТОРСКЕ ДИСЕРТАЦИЈЕ**

Име и презиме аутора: Иван Т Ђирић

Наслов дисертације: Интелигентно управљање мобилним роботима на основу неуро-
фази-генетског препознавања објеката и праћења људи у роботској визији

Ментор: др Жарко Ђојбашић, ред. проф.

Изјављујем да је штампани облик моје докторске дисертације истоветан електронском облику, који сам предао/ла за уношење у Дигитални репозиторијум Универзитета у Нишу.

У Нишу, 16.12.2015.

Потпис аутора дисертације:

Иван Т Ђирић



Универзитет у Нишу

Изјава 3:

ИЗЈАВА О КОРИШЋЕЊУ

Овлашћујем Универзитетску библиотеку „Никола Тесла“ да, у Дигитални репозиторијум Универзитета у Нишу, унесе моју докторску дисертацију, под насловом:

Интелигентно управљање мобилним роботима на основу неуро-фази-генетског препознавања објеката и праћења људи у роботској визији

Дисертацију са свим прилозима предао/ла сам у електронском облику, погодном за трајно архивирање.

Моју докторску дисертацију, унету у Дигитални репозиторијум Универзитета у Нишу, могу користити сви који поштују одредбе садржане у одабраном типу лиценце Креативне заједнице (Creative Commons), за коју сам се одлучио/ла.

1. Ауторство (CC BY)
2. Ауторство – некомерцијално (CC BY-NC)
3. Ауторство – некомерцијално – без прераде (CC BY-NC-ND)
4. Ауторство – некомерцијално – делити под истим условима (CC BY-NC-SA)
5. Ауторство – без прераде (CC BY-ND)
6. Ауторство – делити под истим условима (CC BY-SA)

У Нишу, 16.12.2015.

Аутор дисертације: Иван Т Ћирић

Потпис аутора дисертације:

