SVEUČILIŠTE JOSIPA JURJA STROSSMAYERA U OSIJEKU ELEKTROTEHNIČKI FAKULTET

Damir Filko

Robusna lokalizacija mobilnog robota zasnovana na vizualnim obilježjima ravninskih segmenata

Doktorska disertacija

Osijek, 2013.

Doktorska disertacija izrađena je na Zavodu za industrijska postrojenja i automatizaciju Elektrotehničkog fakulteta Sveučilišta J. J. Strossmayera u Osijeku.

Mentor: Dr. sc. Robert Cupec, izvanredni profesor

Doktorska disertacija ima 173 stranice.

Disertacija br.: 25

Povjerenstvo za ocjenu doktorske disertacije:

- Dr. sc. Snježana Rimac-Drlje, redovita profesorica, Sveučilište J. J. Strossmayera u Osijeku, Elektrotehnički fakultet
- Dr. sc. Robert Cupec, izvanredni profesor, Sveučilište J. J. Strossmayera u Osijeku, Elektrotehnički fakultet
- Dr. sc. Ivan Petrović, redoviti profesor, Sveučilište u Zagrebu, Fakultet elektrotehnike i računarstva

Povjerenstvo za obranu doktorske disertacije:

- 1. Dr. sc. Snježana Rimac-Drlje, redovita profesorica, Sveučilište J. J. Strossmayera u Osijeku, Elektrotehnički fakultet
- Dr. sc. Robert Cupec, izvanredni profesor, Sveučilište J. J. Strossmayera u Osijeku, Elektrotehnički fakultet
- Dr. sc. Ivan Petrović, redoviti profesor, Sveučilište u Zagrebu, Fakultet elektrotehnike i računarstva
- 4. Dr. sc. Goran Martinović, redoviti profesor, Sveučilište J. J. Strossmayera u Osijeku, Elektrotehnički fakultet
- 5. Dr. sc. Irena Galić, docentica, Sveučilište J. J. Strossmayera u Osijeku, Elektrotehnički fakultet

Datum obrane disertacije: 19. siječnja 2013. godine

Sadržaj

Popis slika vii			
Popis tablica xi			
1. Uvoc	1	1	
2. Loka	Lokalizacija robota i kartografija4		
2.1. Pogreška lokalizacije i nesigurnost podataka		5	
2.2.	Kartografija	6	
2.3.	3D Senzori za robotsku lokalizaciju	8	
2.3.1	. Stereo vizija	8	
2.3.2	. 3D laserski daljinomjer (engl. Laser Range Finder)	9	
2.3.3	. Time-of-flight kamere	10	
2.3.4	. 3D senzori sa strukturiranom svjetlošću	11	
2.4.	Lokalizacija zasnovana na značajkama	12	
2.4.1	. Deskriptori točkastih značajki	14	
2.4.2	. Linijske i plošne značajke	16	
2.4.3	. Primjeri lokalizacije zasnovane na značajkama	18	
3. Prim	jena boje i teksture u sustavima računalnog vida	24	
3.1.	Boja	24	
3.1.1	. Refleksivni model formiranja slika	26	
3.1.2	. RGB prostor boja	30	
3.1.3	. Kromatski rgb prostor boja	30	
3.1.4	. HSV prostor boja	31	
3.1.5	. CIE 1931 XYZ prostor boja		
3.1.6	CIE L*a*b* prostor boja	34	
3.1.7	Prostor suprotstavljenih boja	35	
3.1.8	. Deskriptori boja		
3.2.	Tekstura	41	
3.2.1	. LBP porodica deskriptora teksture	43	
3.2.2	. Gabor deskriptori teksture	46	
3.2.3	. Tekston deskriptori tekstura	48	
3.2.4	. Primjene deskriptora teksture	50	
3.3.	Robotska lokalizacija zasnovana na deskriptorima vizualnih obilježja	51	
4. Susta	av za lokalizaciju mobilnog robota zasnovan na ravninskim segmentima	59	

	4.1	l.	Tria	ngulacija 2.5D slike	61
	4.2	2.	Seg	mentacija mreže trokuta na ravninske segmente	62
	4.3	3.	Prik	az i registracija ravninskih segmenata	64
	4.4	1.	Stva	aranje skupa inicijalnih parova i generiranje hipoteza	69
	4	4.4.	1.	Uvjet koplanarnosti	70
	4	4.4.2	2.	Uvjet preklapanja	70
	4	4.4.3	3.	Generiranje hipoteza	71
	4.5	5.	Eva	luacija hipoteza	75
5.]	Prin	njena	a vizualnih obilježja u lokalizaciji mobilnog robota	76
	5.1	l.	Prec	1-procesiranje RGB slike	77
		5.1.	1.	Bilateralno filtriranje	77
		5.1.2	2.	Ujednačavanje histograma	80
		5.1.3	3.	Nelinearno širenje histograma	81
	5.2	2.	Prid	ruživanje boja ravninskim segmentima	83
	5.3	3.	Gen	eriranje inicijalnog skupa parova i evaluacija hipoteza	85
		5.3.	1.	Uvjet sličnosti deskriptora vizualnih obilježja	86
		5.3.2	2.	Izmjene evaluacije hipoteza	86
	5.4	1.	Des	kriptori vizualnih obilježja	88
		5.4.	1.	Primjeri scena i ravninskih segmenata u stvarnim uvjetima	88
		5.4.2	2.	Deskriptori RGB prostora boja	91
		5.4.3	3.	Deskriptori CIE XYZ prostora boja	93
		5.4.4	4.	Deskriptori kromatskog rgb prostora boja	96
		5.4.	5.	Deskriptori CIE L*a*b* prostora boja	98
		5.4.0	5.	Deskriptori HSV prostora boja	101
		5.4.	7.	Deskriptori prostora suprotstavljenih boja	104
		5.4.8	8.	LBP porodica deskriptora teksture	106
		5.4.9	Э.	Meki histogrami	110
		5.4.	10.	Mjere sličnosti deskriptora vizualnih obilježja	112
6.]	Eksp	perin	nentalna analiza primjene vizualnih obilježja za lokalizaciju mobilnih robota.	119
	6.1	l.	Ana	liza rezultata inicijalne korespondencije	120
	6.2	2.	Ana	liza rezultata globalne lokalizacije	129
7.]	Eksp	perin	nentalni evaluacijski okvir	142
8.		Zak	ljuča	k	149

Literatura	152
Sažetak	
Abstract	
Životopis	
Prilog 1	

Popis slika

Slika 1.1. Primjer dvije geometrijski slične, a vizualno različite scene	2
Slika 2.1. a) Snimljena metrička karta, b) tlocrt kata [1]	6
Slika 2.2. Hibridna metričko-topološka karta. Svaki čvor topološke karte predstavlja lokalni	
metrički model dobiven iz panoramske slike [11]	7
Slika 2.3. Epipolarna geometrija [15]	9
Slika 2.4. Primjer "lidar" (engl. Light detection and ranging) 3D senzora [16]	. 10
Slika 2.5. Panasonic D-IMager [17]	. 10
Slika 2.6. Princip rada kamera sa strukturiranom svjetlošću [15]	. 11
Slika 2.7. Microsoft Kinect [18]	. 11
Slika 2.8. a) Infracrvena slika snimljena Kinect senzorom, b) uzorak emitirane svjetlosti na	
ravnoj površini	. 12
Slika 2.9. Obojani oblak točaka snimljen Kinectom	. 12
Slika 2.10. Princip rada lokalizacije temeljene na značajkama	. 13
Slika 2.11. Točkaste i linijske značajke	. 14
Slika 2.12. Primjer korištenja SIFT deskriptora za raspoznavanje objekata [21]	. 14
Slika 2.13. Primjer SIFT deskriptora. Dio slike označen zelenim kvadratom podijeljen je na 10	6
polja. Svakom polju je pridružen histogram orijentacija gradijenta svjetline [23]	. 15
Slika 2.14. Primjer pronalaska SURF deskriptora [24]	. 16
Slika 2.15. a) Ravne plohe detektirane korištenjem stereo vizije, b) brid između ravnih ploha	
određen njihovim presjekom [26]	. 17
Slika 2.16. Primjeri detekcije a) linijskih [12], b) plošnih značajki [27]	. 17
Slika 2.17. Primjer sustava koji koristi linijske značajke i umjetno stvorenu kartu za lokalizaci	iju
mobilnog robota [12]	. 18
Slika 2.18. Primjer sustava koji koristi ravne plohe kao značajke za lokalizaciju mobilnog rob	ota.
Panoramska slika (gore), slika dispariteta (sredina) i detektirane plohe (dolje) [27]	. 19
Slika 2.19. Primjer lokalizacije temeljene na slici koji koristi lokalne točkaste značajke [32]	. 20
Slika 2.20. Primjer sustava koji koriste svesmjernu viziju i linijske značajke za lokalizaciju	
mobilnog robota. a) Točke koje zadovoljavaju zadani kut gradijenta svjetline, b) odabrane lini	je,
c) formiranje deskriptora [35].	. 21
Slika 2.21. Primjer sustava za lokalizaciju mobilnog robota temeljena na ravninskim segmenti	ima
dobivenih iz dubinske slike [2]	. 23
Slika 3.1. Kasap svjetlosti na optičkoj prizmi	. 24
Slika 3.2. Vidijivi spektar svjetiosti [39]	. 24
Slika 3.5. V lune stanice u mieznici ijudskog oka [40]	. 25
Slika 3.4. Spektraine osjetijivosti stanica ljudskog oka [39]	. 25
Slika 3.5. Pfillijer Dayerovog obrasca na senzoru kalilere [15]	. 20
Slika 3.0. KOD prostor boja [39]	. 50
Slika 3.8. HSV prostor hoja [15]	. 31
Slika 3.0. 115 V prostor boja [15] Slika 3.0. a) CIE XVZ raspon boja [b) sPGB raspon boja [15]	. 52
Slika 3.9. a) CIE X 12 Iaspoli 00ja, 0) SKOD Iaspoli 00ja $[15]$. 55
Slika 3.10. CHE E a o prostor boja za urrazine svjetime [15]	. 54
[/8]	36
Slika 3.12 Primier a) 1D histograma Gaussove distribucije b) presieka dva 1D histograma	. 30
Slika 3.13. Primjer 1D histograma R. G i B kanala slike u boji	. 37
Slika 3.14. Primier 3D histograma RGB prostora boja	. 38
Slika 3.15. Primjeri različitih tipova tekstura u prirodi [60]	. 42
J I I LJ	

Slika 3.16. Promjenom skale se u nekim slučajevima gubi osjećaj teksture, dok se u drugima ne	e
gubi [60]	42
Slika 3.17. Način dobivanja LBP vrijednosti [62]	44
Slika 3.18. LBP vrijednost je moguće definirati nad većom okolinom s proizvoljnim brojem	
točaka [62]	44
Slika 3.19. Uniformni LBP uzorci za okolinu od 8 točaka [62]	45
Slika 3.20. Položaj, skala i orijentacija različitih Gaborovih filtera u prostorno-frekvencijskoj	
ravnini [65]	47
Slika 3.21. Prva dva retka je banka filtera s 13 izotropičnih filtera, druga dva retka banka filtera	a s
4 različita filtera nazvana MR4 [68]	48
Slika 3.22. Primjer dvije baze tekstona [68]	49
Slika 3.23. Način generiranja baze tekstona za 3D teksture [69]	49
Slika 3.24. Primjer stvaranja kompleksnijih tekstona [70]	50
Slika 3.25. Princip rada lokalizacije zasnovane na pojavnosti	52
Slika 3.26. Pionirski pokušaji lokalizacije temeljene na slici [75]	53
Slika 3.27. Okviri za testiranje koriste različite kamere i uvjete za ispitivanje sustava lokalizaci	je
temeljene na slici [78]	54
Slika 3.28. Višedimenzionalni histogrami boja se klasteriraju ne neki predefinirani broj [81]	55
Slika 3.29. Primier sustava za lokalizaciju mobilnog robota zasnovanog na slici koji koristi	
geometrijsku kartu sačinjenu od voksela [87]	57
Slika 3.30. Primier sustava za lokalizaciju mobilnog robota temeljenog na slici koji koristi	
geometrijsku kartu stvorenu od vertikalnih površina [88]	58
Slika 4.1. Sustav lokalizacije mobilnog robota zasnovan na ravninskim segmentima	59
Slika 4.2. Postunak triangulacije: prvo se određuju početne točke kao uglovi, a zatim se analizi	ra
svaki trokut zasebno i pronalazi vrh novih trokuta kao najudaljenija točka u odnosu na ravninu	
koja prolazi vrhovima analiziranog trokuta	61
Slika 4 3. Primier dualnog grafa i spajanja dva čvora [89]	62
Slika 4.4 a) Dubinska slika b) trokuti dobiveni triangulacijom c) granice između dominantnih	ິ າ
ravninskih segmenata d) konačni ravninski segmenti	63
Slika 4.5 Koordinatni sustav ravninskog segmenta $R_{\rm p}$	65
Slika 4.6. Parametri stvarnog i estimiranog ravninskog segmenta	66
Slika 4.7 Primier izgleda scena u zatvorenim prostorima [3]	72
Slika 5.1. Prošireni sustav za lokalizaciju mobilnog robota	76
Slika 5.2. a) Originalni signal b) funkcija sljčnosti c) filtrirani signal [00]	70
Slika 5.2. d) Originanii Signai, 0) funkcija snehosti, 0) mu nani Signai [70]	70
Slika 5.4. Drimier bilateralnog filtiranja slika u boji	90
Slika 5.4. Frimier viadnožavanja histograma	0U Q1
Slika 5.5. Fililijel ujedilačavalija liistografija	01
Slika 5.0. Primjer neinearnog sirenja nistograma, a) originalne, b) transformirane	82 92
Slika 5.7. Primijer slika konstenini za kanoraciju Microsoft Kinect senzora	63
Slika 5.8. Primjer iskijučivanja rubnih točaka ravninskih segmenata radi smanjenja utječaja	
pogreske reprojekcije. Gornja lijeva slika prikazuje sve točke koja bi se mogle pridružiti	
ravninskom segmentu, gornja desna slika naglasava rubove između segmenata na dubinskoj	05
slici, donja slika prikazuje konačno pridružene točke ravninskim segmentima.	85
Slika 5.9. Odabrane scene snimljene u dnevnim i nočnim uvjetima	89
Slika 5.10. Kazmatrani ravninski segmenti	90
Slika 5.11. Svi binovi RGB histograma za rezoluciju od a) 32, b) 16, c) 8 binova po dimenziji	.
prostora boja	91
Slika 5.12. Trodimenzionalni histogrami RGB prostora boja za razmatrane ravninske segmente	•
nni nazoluziti od 16 hinova na dimenziti	92

Slika 5.13. Svi binovi XYZ prostora boja za rezoluciju od a) 32, b) 16, c) 8 binova po dimenziji
Slika 5.14. Trodimenzionalni histogrami XVZ prostora boja za razmatrane ravninske segmente
pri rezoluciji od 16 binova no dimenziji
Slika 5.15. Svi binovi kromatskog rgb prostora boja za rezolucije a) 32. b) 16. c) 8 binova po
korištenoj dimenziji
Slika 5.16. Dvodimenzionalni histogrami rgb kromatskog prostora boja za razmatrane ravninske
segmente pri rezoluciji od 16 binova po dimenziji
Slika 5.17. Svi binovi CIE L*a*b* prostora boja za rezolucije a) 32, b) 16, c) 8 binova po
korištenim dimenzijama ovog prostora boja pri $L^* = 50$
Slika 5.18. Dvodimenzionalni histogrami CIE L*a*b* prostora boja za razmatrane ravninske
segmente pri rezoluciji od 16 binova po dimenziji
Slika 5.19. Svi binovi HSV prostora boja za rezolucije a) 32, b) 16, c) 8 binova po korištenoj
dimenziji, pri $V = 100$
slika 5.20. Dvodimenzionalni histogrami HSV prostora boja za razmatrane ravninske segmente
Slika 5.21. Svi binovi prostora suprotstavljenih boja za razolucija a) 32. b) 16. c) 8 binova po
korištenoi dimenziji $\mu_2 \Omega_0 \approx 220$ 104
Slika 5.22 Dvodimenzionalni histogrami prostora suprotstavljenih boja za razmatrane ravninske
segmente pri rezoluciji od 16 binova po dimenziji
Slika 5.23 Primieri izgleda <i>LBP</i> deskriptora za ravninske segmente parketa i zida snimane
pri drazmam i načnam agvistljanju. Carnij rad zu narkat a) dan b) nač a danji zid a) dan d) nač
pri dievnom i nochom osvjetijenju. Gornji red su parket a) dan, b) noc, a donji zid c) dan, d) noc. 107
Slika 5.24. Primieri izgleda LBP^{riu2}_{riu2} , za ravninske segmente parketa i zida snimane pri dnevnom
i noćnom osvjetljenju. Gornji red su parket a) dan, b) noć, a donji zid c) dan, d) noć
i noćnom osvjetljenju. Gornji red su parket a) dan, b) noć, a donji zid c) dan, d) noć
i noćnom osvjetljenju. Gornji red su parket a) dan, b) noć, a donji zid c) dan, d) noć
i noćnom osvjetljenju. Gornji red su parket a) dan, b) noć, a donji zid c) dan, d) noć
i noćnom osvjetljenju. Gornji red su parket a) dan, b) noć, a donji zid c) dan, d) noć
i noćnom osvjetljenju. Gornji red su parket a) dan, b) noć, a donji zid c) dan, d) noć
i noćnom osvjetljenju. Gornji red su parket a) dan, b) noć, a donji zid c) dan, d) noć
i noćnom osvjetljenju. Gornji red su parket a) dan, b) noć, a donji zid c) dan, d) noć
i noćnom osvjetljenju. Gornji red su parket a) dan, b) noć, a donji zid c) dan, d) noć
i noćnom osvjetljenju. Gornji red su parket a) dan, b) noć, a donji zid c) dan, d) noć
i noćnom osvjetljenju. Gornji red su parket a) dan, b) noć, a donji zid c) dan, d) noć
i noćnom osvjetljenju. Gornji red su parket a) dan, b) noć, a donji zid c) dan, d) noć
i noćnom osvjetljenju. Gornji red su parket a) dan, b) noć, a donji zid c) dan, d) noć
i noćnom osvjetljenju. Gornji red su parket a) dan, b) noć, a donji zid c) dan, d) noć
i noćnom osvjetljenju. Gornji red su parket a) dan, b) noć, a donji zid c) dan, d) noć
i noćnom osvjetljenju. Gornji red su parket a) dan, b) noć, a donji zid c) dan, d) noć
i noćnom osvjetljenju. Gornji red su parket a) dan, b) noć, a donji zid c) dan, d) noć
i noćnom osvjetljenju. Gornji red su parket a) dan, b) noć, a donji zid c) dan, d) noć
i noćnom osvjetljenju. Gornji red su parket a) dan, b) noć, a donji zid c) dan, d) noć
i noćnom osvjetljenju. Gornji red su parket a) dan, b) noć, a donji zid c) dan, d) noć
i noćnom osvjetljenju. Gornji red su parket a) dan, b) noć, a donji zid c) dan, d) noć
i noćnom osvjetljenju. Gornji red su parket a) dan, b) noć, a donji zid c) dan, d) noć

zasnovanih na RGC prostoru boja126Slika 6.6. Statistička analiza učinkovitosti inicijalne korespondencije korištenjem deskriptora127Zasnovanih na prostoru suprotstavljenih boja127Slika 6.7. Statistička analiza učinkovitosti inicijalne korespondencije korištenjem deskriptora128Zasnovanih na LBP porodici deskriptora teksture128Slika 6.8. Kumulativni histogrami mjerenja za XYZ8 deskriptor u slučaju korištenja samo 5135Slika 6.9. Kumulativni histogrami mjerenja za HSV8 deskriptor u slučaju korištenja svih137Slika 6.10. Kumulativni histogrami mjerenja za HSV8 deskriptor u slučaju korištenja samo 5138Slika 6.11. Kumulativni histogrami mjerenja za Lab32 deskriptor u slučaju korištenja svih139Slika 6.12. Kumulativni histogrami mjerenja za Lab8 deskriptor u slučaju korištenja samo 5139Slika 6.12. Kumulativni histogrami mjerenja za Lab8 deskriptor u slučaju korištenja samo 5139Slika 6.12. Kumulativni histogrami mjerenja za Lab8 deskriptor u slučaju korištenja samo 5139Slika 6.12. Kumulativni histogrami mjerenja za Lab8 deskriptor u slučaju korištenja samo 5140Slika 7.1. Pioneer 3DX s ugrađenim Microsoft Kinect senzorom143	Slika 6.5. Statistička analiza učinkovitosti inicijalne korespondencije korištenjem deskriptora	
Slika 6.6. Statistička analiza učinkovitosti inicijalne korespondencije korištenjem deskriptora 127 Slika 6.7. Statistička analiza učinkovitosti inicijalne korespondencije korištenjem deskriptora 128 Slika 6.7. Statistička analiza učinkovitosti inicijalne korespondencije korištenjem deskriptora 128 Slika 6.8. Kumulativni histogrami mjerenja za XYZ8 deskriptor u slučaju korištenja samo 5 135 Slika 6.9. Kumulativni histogrami mjerenja za HSV8 deskriptor u slučaju korištenja svih 137 Slika 6.10. Kumulativni histogrami mjerenja za HSV8 deskriptor u slučaju korištenja samo 5 137 Slika 6.11. Kumulativni histogrami mjerenja za Lab32 deskriptor u slučaju korištenja svih 138 Slika 6.12. Kumulativni histogrami mjerenja za Lab8 deskriptor u slučaju korištenja samo 5 139 Slika 6.12. Kumulativni histogrami mjerenja za Lab8 deskriptor u slučaju korištenja samo 5 140 Slika 7.1. Pioneer 3DX s ugrađenim Microsoft Kinect senzorom 143	zasnovanih na RGC prostoru boja	126
zasnovanih na prostoru suprotstavljenih boja	Slika 6.6. Statistička analiza učinkovitosti inicijalne korespondencije korištenjem deskriptora	
Slika 6.7. Statistička analiza učinkovitosti inicijalne korespondencije korištenjem deskriptora 128 zasnovanih na LBP porodici deskriptora teksture 128 Slika 6.8. Kumulativni histogrami mjerenja za XYZ8 deskriptor u slučaju korištenja samo 5 135 Slika 6.9. Kumulativni histogrami mjerenja za HSV8 deskriptor u slučaju korištenja svih 137 Slika 6.10. Kumulativni histogrami mjerenja za HSV8 deskriptor u slučaju korištenja samo 5 137 Slika 6.10. Kumulativni histogrami mjerenja za HSV8 deskriptor u slučaju korištenja samo 5 138 Slika 6.11. Kumulativni histogrami mjerenja za Lab32 deskriptor u slučaju korištenja svih 139 Slika 6.12. Kumulativni histogrami mjerenja za Lab8 deskriptor u slučaju korištenja samo 5 140 Slika 7.1. Pioneer 3DX s ugrađenim Microsoft Kinect senzorom 143	zasnovanih na prostoru suprotstavljenih boja	127
zasnovanih na LBP porodici deskriptora teksture128Slika 6.8. Kumulativni histogrami mjerenja za XYZ8 deskriptor u slučaju korištenja samo 5135binova135Slika 6.9. Kumulativni histogrami mjerenja za HSV8 deskriptor u slučaju korištenja svih137Slika 6.10. Kumulativni histogrami mjerenja za HSV8 deskriptor u slučaju korištenja samo 5138Slika 6.11. Kumulativni histogrami mjerenja za Lab32 deskriptor u slučaju korištenja svih139Slika 6.12. Kumulativni histogrami mjerenja za Lab8 deskriptor u slučaju korištenja samo 5140Slika 7.1. Pioneer 3DX s ugrađenim Microsoft Kinect senzorom143	Slika 6.7. Statistička analiza učinkovitosti inicijalne korespondencije korištenjem deskriptora	
Slika 6.8. Kumulativni histogrami mjerenja za XYZ8 deskriptor u slučaju korištenja samo 5 binova 135 Slika 6.9. Kumulativni histogrami mjerenja za HSV8 deskriptor u slučaju korištenja svih 137 Slika 6.10. Kumulativni histogrami mjerenja za HSV8 deskriptor u slučaju korištenja samo 5 137 Slika 6.10. Kumulativni histogrami mjerenja za HSV8 deskriptor u slučaju korištenja samo 5 138 Slika 6.11. Kumulativni histogrami mjerenja za Lab32 deskriptor u slučaju korištenja svih 139 Slika 6.12. Kumulativni histogrami mjerenja za Lab8 deskriptor u slučaju korištenja samo 5 140 Slika 7.1. Pioneer 3DX s ugrađenim Microsoft Kinect senzorom 143	zasnovanih na LBP porodici deskriptora teksture	128
binova	Slika 6.8. Kumulativni histogrami mjerenja za XYZ8 deskriptor u slučaju korištenja samo 5	
Slika 6.9. Kumulativni histogrami mjerenja za HSV8 deskriptor u slučaju korištenja svih 137 Slika 6.10. Kumulativni histogrami mjerenja za HSV8 deskriptor u slučaju korištenja samo 5 138 Slika 6.11. Kumulativni histogrami mjerenja za Lab32 deskriptor u slučaju korištenja svih 138 Slika 6.12. Kumulativni histogrami mjerenja za Lab8 deskriptor u slučaju korištenja samo 5 139 Slika 6.12. Kumulativni histogrami mjerenja za Lab8 deskriptor u slučaju korištenja samo 5 140 Slika 7.1. Pioneer 3DX s ugrađenim Microsoft Kinect senzorom 143	binova	135
raspoloživih binova 137 Slika 6.10. Kumulativni histogrami mjerenja za HSV8 deskriptor u slučaju korištenja samo 5 138 binova 138 Slika 6.11. Kumulativni histogrami mjerenja za Lab32 deskriptor u slučaju korištenja svih 139 Slika 6.12. Kumulativni histogrami mjerenja za Lab8 deskriptor u slučaju korištenja samo 5 140 Slika 7.1. Pioneer 3DX s ugrađenim Microsoft Kinect senzorom 143	Slika 6.9. Kumulativni histogrami mjerenja za HSV8 deskriptor u slučaju korištenja svih	
Slika 6.10. Kumulativni histogrami mjerenja za HSV8 deskriptor u slučaju korištenja samo 5 binova 138 Slika 6.11. Kumulativni histogrami mjerenja za Lab32 deskriptor u slučaju korištenja svih 139 raspoloživih binova 139 Slika 6.12. Kumulativni histogrami mjerenja za Lab8 deskriptor u slučaju korištenja samo 5 140 Slika 7.1. Pioneer 3DX s ugrađenim Microsoft Kinect senzorom 143	raspoloživih binova	137
binova	Slika 6.10. Kumulativni histogrami mjerenja za HSV8 deskriptor u slučaju korištenja samo 5	
Slika 6.11. Kumulativni histogrami mjerenja za Lab32 deskriptor u slučaju korištenja svih raspoloživih binova	binova	138
raspoloživih binova	Slika 6.11. Kumulativni histogrami mjerenja za Lab32 deskriptor u slučaju korištenja svih	
Slika 6.12. Kumulativni histogrami mjerenja za Lab8 deskriptor u slučaju korištenja samo 5 binova	raspoloživih binova	139
binova	Slika 6.12. Kumulativni histogrami mjerenja za Lab8 deskriptor u slučaju korištenja samo 5	
Slika 7.1. Pioneer 3DX s ugrađenim Microsoft Kinect senzorom143	binova	140
	Slika 7.1. Pioneer 3DX s ugrađenim Microsoft Kinect senzorom	143
Slika 7.2. Primjer sparenih scena	Slika 7.2. Primjer sparenih scena	144

Popis tablica

Tablica 5.1. Brzine izračunavanja mjera sličnosti/razlike	. 115
Tablica 6.1. Popis deskriptora i njihovih pragova koji će se koristiti za evaluaciju globalne	
lokalizacije	. 129
Tablica 6.2. Rezultati mjerenja pokusa globalne lokalizacije	. 131
Tablica 6.3. Relativne i apsolutne promjene performansi sustava lokalizacije koji koristi i	
deskriptore vizualnih obilježja u usporedbi sa sustavom koji je zasnovan samo na geometriji	. 134
Tablica 7.1. Primjer datoteke s relativnim položajima između ispitnih i referentnih scena	. 145
Tablica 7.2. Primjer datoteke s relacijama između ravninskih segmenata testnih scena i	
referentnih scena	. 145
Tablica 7.3. Algoritam za evaluaciju uspješnosti deskriptora vizualnih obilježja pokusom	
pronalaska inicijalne korespodencije	. 146
Tablica 7.4. Algoritam za evaluaciju uspješnosti deskriptora vizualnih obilježja pokusom	
globalne lokalizacije	. 147

1. Uvod

Lokalizacija mobilnog robota je jedan od fundamentalnih problema u robotici. Uspjeh mobilnog robota da izvrši neki zadatak neodvojivo je povezan sa sposobnošću da se robot uspješno lokalizira u prostoru. Stoga je izuzetno bitno riješiti taj problem točno i brzo da bi mobilni robot mogao obavljati koristan rad u prostoru gdje se kreće i u što kraćem vremenu. Posljednje desetljeće je donijelo velike razvojne korake u rješavanju problema lokalizacije sofisticiranijom implementacijom novih načina pronalaska značajki u prostoru, poboljšavanjem hardverske podrške i općenitim razvojem robotskog vida. Primjer novih i sofisticiranih sustava za lokalizaciju mobilnog robota koji su došli s napretkom znanosti su SLAM sustavi (engl. Simultaneous Localization And Mapping) koji omogućavaju istodobno lokaliziranje i stvaranje karte okoline. Prvi sustavi za preciznu lokalizaciju mobilnog robota su uglavnom koristili geometrijske značajke i to prvenstveno linijske. Znatnijim ubrzavanjem hardverske podrške i osmišljavanjem novih algoritama za pronalazak značajki su se počele uglavnom koristiti točkaste značajke i pripadajući deskriptori, kao što je SIFT, potpomognuti stereo sustavom kamera. U zadnje dvije godine na tržištu su se pojavili jeftini dubinski senzori kao što je Microsoft Kinect, koji, iako su prvenstveno namijenjeni kućnim sustavima zabave, počinju se intenzivno primjenjivati u mobilnoj robotici. Razlog Kinectove nagle popularnosti u mobilnoj robotici su izuzetne sposobnosti snimanja okoline uz pristupačnu cijenu oko 200\$ dok je prije takav sustav kamera koštao najmanje 5000\$. Stoga, korištenjem takvih sustava kamera se olakšalo i ubrzalo dobivanje preciznih trodimenzionalnih podataka iz okoline.

Postoji više sustava za lokalizaciju mobilnog robota koji koristi dubinske slike Kinect senzora kao što su [1, 2], a jedan takav sustav je razvijen i u znanstveno-istraživačkom radu na Elektrotehničkom Fakultetu u Osijeku i objavljen u [3]. Razvijeni sustav koristi isključivo podatke dubinskog djela Kinect senzora na temelju kojih generira podatke o ravninskim segmentima u prostoru gdje se robot kreće. Dobivene značajke ravninskih segmenata se potom koriste za lokalizaciju mobilnog robota u šest stupnjeva slobode kretanja u prostoru. Zbog toga što je sustav ograničen isključivo na geometrijske podatke prostora, omogućava pojavu znatne pogreške lokalizacije u slučaju globalne lokalizacije mobilnog robota zbog moguće velike geometrijske sličnosti zatvorenih prostora. Jedan od načina poboljšanja rada sustava u takvim uvjetima jest pridruživanje ravninskim segmentima deskriptore vizualnih obilježja, kao što su boja ili tekstura, koji će omogućiti razlikovanje geometrijski sličnih, ali vizualno različitih ravninskih segmenta.



Slika 1.1. Primjer dvije geometrijski slične, a vizualno različite scene

Kao što se može vidjeti na slici 1.1. radi se o prostorima koji imaju izuzetno sličan razmještaj ravninskih segmenata koji se mogu koristiti za lokalizaciju, ali u isto vrijeme ravninski segmenti kao što su pod i radijator imaju vizualno različita obilježja. Iako su geometrijski ovi prostori dosta slični, ljudska bića ih lako mogu razlikovati, dok bi mobilni robot, koji koristi samo podatke o geometriji, imao problema u razlikovanju razmatranih scena.

Korištenje vizualnih obilježja prostora za lokalizaciju mobilnog robota je do sada bilo dosta ograničeno i uglavnom isključivo povezano sa semantičko-topološkom lokalizacijom mobilnog robota zasnovanog na pojavnosti. Cilj takve lokalizacije jest grubo određivanje položaja robota u smislu raspoznavanja sobe, hodnika ili zgrade u kojoj se robot trenutno nalazi. Razlog dosadašnje ograničene primjene vizualnih obilježja jest relativno komplicirana fuzija s geometrijskim podacima prostora i velika ovisnost o promjenama u okolini s velikim naglaskom na promjene osvjetljenja. Promjena osvjetljenja okoline u stvarnim uvjetima je velika prepreka za uspješno korištenje deskriptora vizualnih obilježja u sustavima lokalizacije mobilnog robota, budući da bi mobilni robot morao pouzdano raditi bez velikog utjecaja promjene osvjetljenja u okolini. Iako postoje algoritmi očuvanja konstantnih boja, njihova uporaba je često ograničena, komplicirana i spora. Dakle, iako je ideja primjene vizualnih obilježja za lokalizaciju mobilnih robota smislena i obećavajuća, postoje i značajne prepreke u njezinoj praktičnoj realizaciji. U sklopu ove doktorske disertacije će se stoga istražiti mogućnosti primjene nekih standardnih i nestandardnih deskriptora vizualnih obilježja povezanih s trodimenzionalnim ravninskim segmentima u svrhu lokalizacije mobilnog robota. Sustav razvijen ovdje je zasnovan na onome objavljenom u [3] i [4]. Znanstveni doprinosi ove doktorske disertacije su sljedeći:

- Poboljšanje robusnosti i učinkovitosti sustava lokalizacije mobilnog robota zasnovanih na ravninskim segmentima korištenjem vizualnih obilježja.
- Analiza mogućnosti primjene postojećih deskriptora zasnovanih na vizualnim obilježjima za lokalizaciju mobilnog robota pomoću ravninskih segmenata.
- Eksperimentalni okvir i metodologija za vrednovanje algoritama sparivanja vizualnih obilježja slike i lokalizacije mobilnog robota.

Na početku ove disertacije su objašnjeni osnovni principi lokalizacije mobilnog robota pomoću karti te dosadašnji način korištenja značajki i deskriptora u sustavima lokalizacije mobilnog robota. U trećem poglavlju su objašnjeni načni formiranja i korištenja deskriptora vizualnih obilježja s posebnim naglaskom na njihove dosadašnje primjene u sustavima za lokalizaciju mobilnog robota zasnovanima na pojavnosti. U četvrtom poglavlju objašnjavaju mehanizmi osnovnog sustava lokalizacije mobilnog robota zasnovanog se na trodimenzionalnim ravninskim segmentima. Peto poglavlje objašnjava modifikacije osnovnog sustava i implementaciju mehanizama, koji omogućavaju korištenje vizualnih obilježja ravninskih segmenata čime je dobiven prošireni sustav za lokalizaciju mobilnog robota. U petom poglavlju se također na konkretnim primjerima pokazuju svojstva korištenih deskriptora vizualnih obilježja. U šestom poglavlju je statistički analizirana mogućnost korištenja implementiranih deskriptora vizualnih obilježja za inicijalno sparivanje ravninskih segmenata te se u istom poglavlju nalaze rezultati, analize i objašnjenja provedenog pokusa globalne lokalizacije mobilnog robota. Sedmo poglavlje opisuje eksperimentalni okvir za evaluaciju algoritama sparivanja vizualnih značajki slike i lokalizacije mobilnog robota razvijen u okviru eksperimentalnog dijela ovog rada, dok zadnje poglavlje predstavlja zaključak cijele disertacije.

2. Lokalizacija robota i kartografija

Širina primjene mobilnog robota za rad u određenom okruženju te njegova mogućnost izvršavanja određenih zadataka je povezana sa sposobnošću robota da odredi svoj položaj u radnoj okolini. Ovaj problem naziva se lokalizacija robota i predstavlja jedan od fundamentalnih problema u robotici te se uglavnom može podijeliti na tri podproblema:

- problem praćenja lokalnog položaja u prostoru (engl. tracking, visual odometry, incremental localization problem),
- problem određivanja globalnog položaja na karti (engl. global localization problem),
- problem otetog robota (engl. kidnapped robot problem).

Određivanje globalnog položaja je sposobnost određivanja lokacije robota u prethodno naučenoj karti radne okoline, pri čemu robot nema nikakvu dodatnu informaciju osim da se nalazi negdje na karti. Nakon što se robot lokalizirao na karti, lokalno praćenje položaja se svodi na praćenje tog položaja tijekom vremena. Postojeći algoritmi za praćenje položaja robota [5] dobro rješavaju taj problem, ali obično ne uspijevaju uspješno odrediti globalni položaj na karti na početku rada, dok su algoritmi za globalnu lokalizaciju uglavnom puno neprecizniji od onih za lokalno praćenje, ali ipak uspijevaju iz cijele karte pronaći približnu lokaciju robota. Algoritmi za globalnu lokalizaciju su zbog toga puno kompleksniji te zahtijevaju više računskih resursa od algoritama za lokalno praćenje.

Problem otetog robota se može smatrati težom varijantom globalne lokalizacije budući da je to problem gdje robot može "misliti" da je lokaliziran, a u stvari se nalazi na drugom mjestu. Stoga se ovaj problem može shvatiti kao problem ponovnog lokaliziranja u slučaju greške globalne lokalizacije.

Ovisno o načinu korištenja podataka snimljenih u okolini, metode lokalizacije se također mogu podijeliti u dvije skupine:

- lokalizacija zasnovana na značajkama (engl. feature-based),
- lokalizacija zasnovana na pojavnosti (engl. appearance-based).

2.1. Pogreška lokalizacije i nesigurnost podataka

Dok se robot giba u nekom prostoru koristi svoju vlastitu odometriju da bi estimirao svoj položaj. Senzori odometrije mjere broj okretaja kotača dok se robot kreće. Ta mjerenja potom mogu dati estimaciju pomaka robota na podu, ali zbog proklizavanja kotača i određenih izvora smetnji ta mjerenja mogu biti i neprecizna. Robot stoga mora imati određenu ideju o pogrešci svoje odometrije da bi se ona mogla uzeti u obzir prilikom estimacije položaja. Ovako ponašanje robota bi se moglo povezati sa sličnim ponašanjem čovjeka koji se kreće zatvorenih očiju. Već nakon nekoliko koraka, nesigurnost s kojom čovjek procjenjuje svoju trenutnu poziciju u odnosu na okolinu postaje vrlo velika. Kao što čovjek za nadopunu svoje odometrije ima vizualni sustav za utvrđivanje svojeg položaja u odnosu na okolinu. Tako i robot može biti opremljen percepcijskim senzorima koji će prikupljati podatke iz okoline. Ti senzori mogu uključivati viziju, kompase, inercijske sustave, GPS i sl.

Nažalost percepcijski senzori kao i odometrija mogu imati određenu pogrešku. Pogreška često proizlazi iz prirode senzora odnosno prirode prikupljenih informacija. Npr. laserski senzori su osjetljivi na refleksiju i interferenciju drugih izvora svjetlosti, vizualni sustavi mogu biti također osjetljivi na osvjetljenje, pojavu sjena u prostoru i sl. Stoga robot mora na neki način estimirati koliko su informacije koje dobiva pouzdane, odnosno odrediti kolika je njihova nesigurnost da bi napravio bolju estimaciju svog položaja.

Može se zaključiti da uzimanje u obzir pogreške izvora informacija daje realniji pogled na trenutno stanje robota u prostoru te je stoga veći potencijal za ispravnu odluku tijekom stvarnog rada robota. Gledajući problem lokalizacije robota iz probabilističke perspektive, pojmovi kao što su nesigurnosti i vjerojatnost gdje se robot nalazi dolaze gotovo prirodno. Drugim riječima, modeliranjem nesigurnosti podataka položaj robota se prikazuje kao gustoća vjerojatnosti nad prostorom stanja robota.

U svrhu praćenja i ažuriranja stanja robota kroz probabilistički model, često se koriste pristupi koji se temelje na Kalmanovim filterima. Kalmanovi filteri su rekurzivni algoritmi za obradu podataka koji omogućuju estimiranje stanja stohastičkih linearnih dinamičkih sustava koji se mogu opisati sljedećom algebarskom formulom:

$$\mathbf{x}_{\mathbf{k}} = \mathbf{F}_{\mathbf{k}} \cdot \mathbf{x}_{\mathbf{k}-1} + \mathbf{B}_{\mathbf{k}} \cdot \mathbf{u}_{\mathbf{k}} + \mathbf{w}_{\mathbf{k}}$$
(2.1)

gdje su \mathbf{x}_k i \mathbf{x}_{k-1} trenutno stvarno stanje i prethodno stanje, \mathbf{F}_k je matrica dinamike sustava, \mathbf{B}_k je matrica raspodjele upravljanja s kontrolnim vektorom \mathbf{u}_k , dok je \mathbf{w}_k šum koji se najčešće

modelira kao Gaussov bijeli šum [5]. U sustavu za lokalizaciju mobilnog robota koji se koristi u sklopu ove doktorske disertacije koristi se prošireni Kalmanov filter čiji je način korištenja detaljnije objašnjen u 4. poglavlju i prilogu 1.

Kao alternativa Kalmanovim filtrima u literaturi se često koriste čestični filtri (engl. Particle filter), također često nazvani kao sekvencijalnom Monte Carlo metodom [6, 7, 8, 9, 10].

2.2. Kartografija

Stvaranje karata je problem integriranja informacija prikupljenih pomoću senzora u konzistentni model okoline. Da bi roboti radili efikasno u nekom ljudskom okruženju oni moraju održavati ažuriranu kartu svoga okružja. Većina istraživanja u robotskoj kartografiji se odnosi na stvaranje inicijalne karte, a relativno malo istraživanja se bavi ažuriranjem karata tokom dugotrajnog rada robota u promjenjivim okruženjima kao npr. [11]. Pregledom literature za vizualnu lokalizaciju i kartografiju mobilnog robota pojavljuju se sljedeći pristupi lokalizaciji:

- metričke metode,
- topološke metode,
- hibridne metode.

Metričke metode nastoje estimirati i pratiti apsolutnu poziciju robota unutar geometrijske karte kao što se vidi na slici 2.1., dok topološki pristupi kartografiji predstavljaju okolinu kao graf čiji čvorovi predstavljaju lokacije u stvarnom prostoru.



Slika 2.1. a) Snimljena metrička karta, b) tlocrt kata [1]

Između ta dva glavna smjera postoji i hibridni pristup gdje je globalno karta prikazana kao topološka koja povezuje lokalne metričke podkarte. Potreba za ovakvim hibridnim pristupom dolazi zbog određenih prednosti i nedostataka koje imaju oba glavna pristupa kartografiji i lokalizaciji. Nedostatak kompletno metričke karte je da karte velikih prostora zahtijevaju praćenje i analiziranje velike količine podataka što rezultira velikim hardverskim zahtjevima, dok je nedostatak čiste topološke karte to što ne mogu odrediti precizni položaj robota unutar čvora. Na slici 2.2. može se vidjeti primjer hibridnog metričko-topološkog modela gdje su lokalne karte panoramske slike okoline.

Svi pristupi kartografiji i lokalizaciji zahtijevaju određeni stupanj stabilnosti, odnosno radi se o pretpostavci statičkog modela okoline. Taj uvjet često nije u potpunosti zadovoljen u stvarnim ljudskim okružjima koja uglavnom nisu sasvim statična, stoga često dolazi do problema rada robota u takvim okružjima. Određeni sustavi kao npr. u [11] dopuštaju dinamiku okoline te kao inspiraciju za svoj pristup kartografiji uzimaju principe ljudskog pamćenja.



Slika 2.2. Hibridna metričko-topološka karta. Svaki čvor topološke karte predstavlja lokalni metrički model dobiven iz panoramske slike [11].

U do sada opisanim metodama lokalizacije bilo je uvjetovano postojanje karte prostora u kojemu se robot kreće, bilo da je tu kartu napravio čovjek npr. koristeći alate za 3D modeliranje [12, 13] ili da je čovjekom upravljani robot snimio prostor u kojemu će se ubuduće samostalno kretati [1]. Postoji također još jedan sustav lokalizacije koji u isto vrijeme nastoji stvoriti kartu prostora i sebe locirati unutar nje, ti sustavi se nazivaju SLAM (engl. Simultaneous Localization And Mapping) [5]. Dakle ovi sustavi omogućavaju mobilnom robotu da samostalno istražuje i stvara kartu nepoznatog okružja dok cijelo to vrijeme prati vlastitu lokaciju unutar tog okružja. SLAM je stoga definiran kao problem stvaranja nove karte ili ažuriranjem postojeće karte te istovremenog samolociranja. Rješavanju ovog problema je izrazito doprinijelo korištenje modernih probabilističkih modela kao što su Kalmanovi filteri i čestični filteri odnosno Monte Carlo metode [5]. Također "rješavanje" SLAM problema se u pregledu literature smatra jednim od najvećih dostignuća mobilne robotike unutar posljednjih nekoliko desetljeća.

2.3. 3D Senzori za robotsku lokalizaciju

Za potrebe lokalizacije robota i stvaranje karte prostora mogu se koristiti po prirodi različiti senzori kao što su ultrazvučni, laserski, optičke kamere, time-of-flight kamere i kamere sa strukturiranom svjetlošću. Za stvaranje metričkih karata posebno su prikladni 3D senzori.

3D senzorskih tehnologija u praktičnoj primjeni ima puno budući da se uglavnom ne radi o novim tehnologijama. Gdje i kako ih se koristi zavisi o primjeni i zahtjevima koji se od njih traže, primjene mogu biti od medicinske, geodetske do onih u računalnom vidu. Ni jedna od tehnologija za 3D senzoriku nije razvijena s isključivom primjenom u robotici, već je uglavnom robotika preuzela te tehnologije iz drugih grana tehnike.

2.3.1. Stereo vizija

Robotski vid najčešće koristi optičke video kamere da bi došao do informacija o svojoj okolini, bilo da se radi o monokularnom sustavu, tj. sustavu zasnovanom na jednoj kameri ili stereo sustavu s dvije ili više kamera. Pomoću monokularnih sustava podatke o dubini možemo dobiti jedino korištenjem određenih algoritama kao što su npr. algoritmi za određivanje "strukture iz pokreta" (engl. structure from motion). Korištenjem stereo vida s dvije ili više kamera dobivamo dva ili više pogleda na istu scenu i mogućnost dobivanja informacija o dubini scene na relativno jednostavniji način u odnosu na prethodno spomenute algoritme. Tada koristeći pravila određenih geometrijskih ograničenja uvjetovanih "pinhole" modelom kamere [14] i triangulacijom možemo doći do položaja neke točke u 3D prostoru.

Najveći problem kod stereo vizije je određivanje korespodencija između pogleda jedne i druge kamere. Taj problem olakšava epipolarna geometrija, objašnjena na slici 2.3., jer ako obje kamere stereo vizije "vide" istu točku u prostoru tada mora vrijediti:

$$\mathbf{x}_{\mathbf{R}}^{\mathbf{T}} \cdot \mathbf{F} \cdot \mathbf{x}_{\mathbf{L}} = 0 \tag{2.2}$$

gdje je **F** fundamentalna matrica, a \mathbf{x}_{L} i \mathbf{x}_{R} su koordinate točke u koordinatnom sustavu lijeve i desne kamere.



Slika 2.3. Epipolarna geometrija [15]

Stoga promatranjem iste scene pomoću minimalno dvije kamere te pronalaženjem korespondencija među slikama nam omogućava pronalazak slike dispariteta čije vrijednosti predstavljaju razlike u koordinatama iste točke u prostoru između npr. lijeve i desne slike kod stereo vida. Iz takve slike dispariteta moguće je izračunati udaljenost točaka vidljivih na slikama u odnosu na stereo sustav kamera. Dubinske informacije dobivene na ovakav način nam omogućuju stvaranje metričkih karata.

2.3.2. 3D laserski daljinomjer (engl. Laser Range Finder)

Ovaj senzor u svakom trenutku skenira samo jednu točku u prostoru. Bazira se na mjerenju vremena koje je potrebno laserskoj zraci da prijeđe put od senzora do točke u prostoru koju mjerimo te natrag do senzora. Uz pretpostavku o konstantnoj i poznatoj brzini svjetlosti udaljenost se računa pomoću izraza:

$$s = \frac{c \cdot t}{2} \tag{2.3}$$

gdje je s pređeni put, c je brzina svjetlosti, a t proteklo vrijeme.

Korištenjem rotirajućih zrcala ovaj senzor može relativno brzo skenirati cijelu scenu. Mana im je što su skupi i osjetljivi na reflektivne površine. Na slici 2.4. može se vidjeti primjer novijeg 3D laserskog senzora za snimanje prostora.



Slika 2.4. Primjer "lidar" (engl. Light detection and ranging) 3D senzora [16]

2.3.3. Time-of-flight kamere

Ove kamere predstavljaju napredniji oblik LRF senzora za mjerenje udaljenosti s laserom opisanog u poglavlju 2.3.2. Bazira se na istom principu kao i LRF senzor tako da mjeri vrijeme putovanja svjetlosne zrake, samo što je ovdje u uređaj ugrađeno 2D polje senzora, pa u isto vrijeme mjeri podatke o dubini cijele scene. Noviji senzori obično daju dubinsku sliku veličine 320x240 točaka, dok brzina osvježavanja slike može biti i do 100Hz. Različite tehnologije ovih kamera omogućavaju različite domete, od 5-10 metara, pa do nekoliko kilometara. Relativno su skupe te su osjetljive na jakost prirodne svjetlosti. U slučaju korištenja više ovakvih senzora u isto vrijeme moguće je stvaranje interferencije, također je moguće da dođe do značajne pogreške u mjerenju prilikom skeniranja okoline s više reflektirajućih površina. Na slici 2.5. prikazan je primjer time-of-flight kamere od Panasonica.



Slika 2.5. Panasonic D-IMager [17]

2.3.4. 3D senzori sa strukturiranom svjetlošću

Ovaj tip senzora se bazira na projiciranju određenog uzorka svjetlosti na prostor koji se želi snimiti te snimanju reflektirane svjetlosti pomoću kamere. Uočavanjem pomaka u očekivanoj slici uzorka može se odrediti udaljenost od kamere kao što se vidi na slici 2.6., gdje se koristi uzorak svjetlosnih linija za detekciju pomaka.



Slika 2.6. Princip rada kamera sa strukturiranom svjetlošću [15]

Relativno im je mali domet i imaju velikih problema u slučaju skeniranja reflektirajućih površina. Primjer najnovijeg komercijalno dobavljivog 3D senzora ovog tipa je Microsoftov Kinect prikazan na slici 2.7. Originalno zamišljen kao dodatak kućnog sustava zabave, pokazao se izuzetno koristan i u druge svrhe.



Slika 2.7. Microsoft Kinect [18]

Kinect senzor se bazira na projiciranju uzorka točaka u infracrvenom spektru te potom snimanja refleksije svjetlosti na sceni pomoću infracrvene kamere. Uzorak svjetlosti na ravnoj površini i na tipičnoj sceni je vidljiv na slici 2.8. Uspoređujući položaj snimljene točke s određenim referentnim položajem odnosno modelom vrši se triangulacija i izračunava se dubina te točke. Kinect daje izuzetno gustu dubinsku sliku rezolucije 640x480 točaka pri osvježavanju slike od 30Hz. Dubinska slika je 11-bitna što rezultira s 2048 mjerljivih razina dubine. Kinect senzor ima širinu horizontalnog pogleda od 58° i vertikalnog od 45°. Prema specifikacijama proizvođača [18] preporučeno radno područje je od 0.7m do 5m. Uz Kinect

senzor je moguće koristiti službeni SDK (engl. Software Development Kit), ali zbog veće slobode korištenja sirovih podataka iz senzora u ovoj doktorskoj disertaciji se koristi neslužbeni razvoji okvir [19]. Korištenjem dobivene dubinske slike i RGB slike snimljene kamerom također ugrađenom u kućište Kinect-a, omogućeno je stvaranje oblaka obojenih točaka u 3D prostoru kao što se vidi na slici 2.9.



Slika 2.8. a) Infracrvena slika snimljena Kinect senzorom, b) uzorak emitirane svjetlosti na ravnoj površini



Slika 2.9. Obojani oblak točaka snimljen Kinectom

2.4. Lokalizacija zasnovana na značajkama

Lokalizacija zasnovana na značajkama bazira se na korištenju značajki prostora poput točaka, linija i ploha kao orijentira pomoću kojih se može odrediti položaj robota. Da bi se neka značajka mogla smatrati orijentirom, ona mora biti lako uočljiva i što više neovisna o prostornim transformacijama uvjetovanim robotskim vidom i promjenama okoline (pojava okluzije, promjena osvjetljenja i sl.). Lokalizacija zasnovana na značajkama se često povezuje s registracijom slika, pa se stoga može opisati primjerom prikazanim na slici 2.10.



Slika 2.10. Princip rada lokalizacije temeljene na značajkama

Na slici 2.10. može se vidjeti tlocrt jednog kata zgrade. Crveni kružići predstavljaju prethodne položaje robota, a ljubičasti detektirane značajke. Zelenim kružićem prikazan je trenutni položaj robota. Registracijom trenutne slike robota s prethodno snimljenom slikom te pronalaženjem odgovarajućih korespondencija između njihovih značajki moguće je odrediti trenutni relativni položaj robota u odnosu na taj prethodno snimljeni položaj. Slični princip omogućuje i lokalizaciju u odnosu na kartu. Opisivanje i raspoznavanje značajki je olakšano pomoću pridruženih deskriptora.

Deskriptori su strukture koje na temelju geometrijskih ili nekih drugih značajki kao što su boja ili tekstura opisuju neki objekt na takav konzistentan način da će u slučaju nailaska na neku drugu sliku istog objekta rezultirati približno istim deskriptorom. Deskriptori mogu služiti za raspoznavanje različitih objekata od interesa kao što su npr. objekti s kojima robot treba rukovati ili orijentiri u prostoru u kojem se robot lokalizira. U tom slučaju deskriptor za pojedini element prostora mora biti takav da što jednoznačnije opisuje taj element u karti radi što brže i točnije lokalizacije i raspoznavanja orijentira od strane mobilnog robota. Iz tog razloga deskriptori moraju biti što više neovisni o osvjetljenju scene i prostornim transformacijama koje su povezane s vizualnom percepcijom.

U posljednjem desetljeću su se najpopularnijim pokazali deskriptori vezani za točkaste značajke, odnosno deskriptori koji opisuju određenu točku i njenu okolinu u prostoru, dok su

kod geometrijskih značajki popularni deskriptori koji opisuju određenu plohu ili liniju u prostoru. Primjer točkastih i linijskih značajki može se vidjeti na slici 2.11.



Slika 2.11. Točkaste i linijske značajke

2.4.1. Deskriptori točkastih značajki

Za raspoznavanje objekata u slučaju korištenja optičkih kamera često se koriste točkaste značajke s pridruženim lokalnim deskriptorima [20]. Njihovim sparivanjem i povezivanjem sa značajkama modela omogućeno je raspoznavanje objekata i uz postojanje okluzije kao što se vidi na slici 2.12. u slučaju korištenja SIFT detektora značajki. Nacrtane konture na slici 2.12. predstavljaju originalne konture modela koje su transformirane prema položaju objekta raspoznatog na temelju prepoznatih značajki objekta.



Slika 2.12. Primjer korištenja SIFT deskriptora za raspoznavanje objekata [21]

Kao što je prethodno napomenuto, točkasti deskriptori moraju opisati određenu točku u prostoru na takav nedvosmislen način da se kasnije može prepoznati iz što više kutova

gledanja. Iz tog razloga se za opis neke točke koristi i njena okolina, bliža ili daljnja ovisno o kompleksnosti scene i geometriji prostora. Kao značajke prostora se najčešće uzimaju vrhovi, rubovi, uglovi itd., uglavnom točke definirane lokalnim ekstremima odgovarajućih filtera.

Početni oblici točkastih deskriptora su uključivali vektore intenziteta svjetline okoline razmatrane točkaste značajke koji su se uspoređivali normaliziranom korelacijom, a kasnije su se pojavili sofisticiraniji deskriptori kao što je SIFT. SIFT (engl. Scale-Invariant Feature Transform) [21] je metoda detekcije značajki koja omogućava robusnu detekciju i opisivanje određene točke u prostoru pomoću smjerova i intenziteta gradijenta u njenoj okolini. Na takav način mogu se raspoznavati pojedinačne točke u prostoru, a njihovim grupiranjem mogu se raspoznati cijeli objekti. SIFT je u biti metoda koja unutar sebe kombinira detektor točkastih značajki i deskriptora baziranog na distribuciji smjerova gradijenata svjetline unutar područja na slici oko značajke. Informacija koju daje područje slike oko značajke koristi se za normalizaciju deskriptora što ga čini neosjetljivim na promjenu orijentacije i skale. Deskriptor predstavlja 3D histogram lokacija i orijentacija gradijenata svjetline. Udio pojedinog gradijenta svjetline lokacijskom i orijentacijskom binu je otežan njegovom magnitudom. Primjer SIFT značajki može se vidjeti na slici 2.13., gdje su kvadratima različitih veličina označene okoline značajke na temelju kojih se formiraju deskriptori. Kao rezultat SIFT-a dobiva se za svaku značajku 128 elementni vektor koji predstavlja deskriptor te značajke, koji je neovisan o skali, translaciji i rotaciji, a također je djelomično neovisan i o osvjetljenju scene. Sve te karakteristike čine SIFT prikladnim za robotski vid koji je karakterističan po zahtjevima za neovisnošću o prostornim transformacijama i osvjetljenju [22]. Loše strane SIFT-a su preveliki broj značajki po slici, preveliki vektor deskriptora odnosno prevelika dimenzionalnost i relativna sporost prilikom pronalaska značajki na slici.



Slika 2.13. Primjer SIFT deskriptora. Dio slike označen zelenim kvadratom podijeljen je na 16 polja. Svakom polju je pridružen histogram orijentacija gradijenta svjetline [23].

SURF (engl. Speeded Up Robust Features) [24] je također jedan od primjera novijih deskriptora točkastih značajki. SURF je, slično kao i SIFT, kombinacija detektora značajki i deskriptora, ali SURF svoje detektiranje značajki bazira na aproksimaciji Hessijanove matrice koja u biti predstavlja kvadratnu matricu s parcijalnim derivacijama drugog reda, koja opisuje lokalnu zakrivljenost prostora. Nakon što se pronađe značajka, njezin deskriptor se dobije tako da se u njenoj okolini gleda odziv na Haarov wavelet, te se na sličan način kao i SIFT stvara histogram uzimajući u obzir orijentaciju i magnitude tako dobivenih gradijenata svjetline. Kao rezultat za jednu točku se dobije deskriptor opisan kao 64 elementni vektor. Prednosti i mane su mu slične kao i kod SIFT-a nad kojim ima blagu prednost zbog manjeg deskriptora. Pored SIFT i SURF deskriptora točkastih značajki, u literaturi se spominju još i deskriptori kao što je BRIEF, ORB, FAST, MSER i dr., njihove usporedne performanse mogu se vidjeti u [25].



Slika 2.14. Primjer pronalaska SURF deskriptora [24]

2.4.2. Linijske i plošne značajke

Deskriptori točkastih značajki imaju mnoge prednosti zbog svoje deskriptivnosti prostora, neovisnosti o pojedinim transformacijama i osvjetljenju, pa čak i zbog svoje brojnosti na scenama omogućavaju raspoznavanje objekata i u slučaju okluzije. Ali ista ta brojnost može imati i negativne konotacije budući da na scenama može biti i po nekoliko tisuća značajki, a u bazi više stotina tisuća značajki i to sve za relativno jednostavne karte. Pronalaženje korespondencije u takvim slučajevima je otežano i usporeno. Budući da su ljudska okruženja dosta strukturirana, s podovima, zidovima, stolovima, ormarima i ostalim uglavnom statičnim objektima, pojavljuje se ideja o lokalizaciji prema strukturalnim i geometrijskim značajkama kao što su plohe ili linije na rubovima ploha kao što je vidljivo na slikama 2.15. i 2.16. Prednost takvih geometrijskih značajki je u tome što za njihov opis nije potrebno puno podataka, ali to može biti i nedostatak, jer se zbog manjka podataka gubi na deskriptivnosti posebno u slučaju ako na karti postoje nekoliko stotina ili tisuća linija i ploha.

Iz tog razloga se često uz osnovne geometrijske značajke linijama i plohama nastoji dodati određene značajke koje će povećati deskriptivnost njihovih deskriptora. Osnovne geometrijske značajke kod ravnine uključuju podatke o normali na njihovu površinu, dok kod linija to može biti kut u odnosu na referentnu plohu odnosno tlo po kojem se robot kreće. Raspoznavanje linija i ploha je olakšano ako postoji pretpostavka položaja robota, odnosno ako se robot već globalno lokalizirao. Tada se uz pomoć karte vrši reprojekcija elemenata iz modela okoline koje bi robot u tom trenutku trebao detektirati na sceni koju trenutno vidi, te se potom izračunava točni položaj robota na temelju preklapanja modela i vidljivih linija ili ravnina. Ovakav način usporedbe se često koristi za provjeru hipoteze o položaju robota prilikom lokalizacije te dovodi do korekcije stvarnog položaja robota [12].



Slika 2.15. a) Ravne plohe detektirane korištenjem stereo vizije, b) brid između ravnih ploha određen njihovim presjekom [26].



a) b) Slika 2.16. Primjeri detekcije a) linijskih [12], b) plošnih značajki [27]

Osim linija i ravnih ploha postoji još mogućnost korištenja 3D primitiva odnosno jednostavnih geometrijskih tijela kao što su konveksne plohe [28], sfere, kvadri, cilindri,

stošci i sl. [29, 30, 31], kao orijentira za lokalizaciju robota. Geometrijska svojstva ovakvih primitiva potencijalno otvaraju mogućnost stvaranja raspoznatljivih deskriptora kojima bi se ostvario dovoljno točan opis scene u što manje podataka radi bržeg pronalaska korespondencije u karti. Jedan od načina povećanja deskriptivnosti ravnih ploha jest analiziranje njezinih vizualnih obilježja kao što su boje i teksture.

2.4.3. Primjeri lokalizacije zasnovane na značajkama

U literaturi veliki broj autora koristi točkaste značajke i njihove deskriptore za pronalazak korespondencija između scena, ali prije populariziranja točkastih značajki sa SIFT deskriptorom često su se koristile geometrijske značajke kao npr. detektirane linije. Jedan od ranijih takvih sustava opisan je u [12], gdje autori predstavljaju pionirski sustav lokalizacije mobilnog robota koristeći propagaciju nesigurnosti položaja robota uz korištenje proširenog Kalmanovog filtra. Radi se o monokularnom sustavu vizije na mobilnom robotu koji za lokalizaciju koristi linijske značajke. Također lokalizacija zahtjeva prethodno modeliranu geometrijsku kartu robota koja je prikazana kao žičani model okoline. Detektiranje linija odnosno pravaca na slici se provodi detekcijom rubova i korištenjem Hough prostora. Sustav provodi inkrementalnu lokalizaciju robota s tim da početni položaj robota mora biti prilično točno poznat te se potom procijenjena nesigurnost očekivanih linijskih značajki koristi za limitiranje analize slike na pojedina područja na slici u kojima se može očekivati pojavljivanje tražene značajke. Na slici 2.17. može se vidjeti primjer stvarne scene i njenog odgovarajućeg modela iz karte.



Slika 2.17. Primjer sustava koji koristi linijske značajke i umjetno stvorenu kartu za lokalizaciju mobilnog robota [12]

Također i u [13] opisana je inkrementalna lokalizacija robota korištenjem linija kao značajki. Sustav percepcije mobilnog robota se zasniva na monokularnoj viziji. Linije na slikama detektiraju se pomoću Cannyevog operatora. Sustav lokalizacije zahtjeva prethodno

ručno modeliranu geometrijsku kartu izrađenu kao žični model okoline u kojoj će se robot kretati. Korespondenciju između trenutno detektiranih linija i onih iz modela autori utvrđuju pomoću orijentacije i dužine linije. Generiranje hipoteza izvodi se pomoću strukture slične stablu koristeći korespondencije detektiranih linija. Točnost hipoteze autori utvrđuju preklapanjem trenutne scene odnosno detektirane linije i one renderirane po podacima hipoteze. Mjera sličnosti se utvrđuje pomoću Hausdorff udaljenosti. Autori su deklarirali najveću pogrešku svog sustava lokalizacije kao 0.2m i 2.5°.

Autori [27] pomoću trinokularne vizije na mobilnom robotu stvaraju kartu prostora koji u svakoj točci ima snimak prostora od 360° koji sadrži i sliku dispariteta i sliku intenziteta svjetline. Sustav prvo pronalazi rubove Cannyevim algoritmom, te zatim provodi Delaunay triangulaciju i segmentaciju pomoću širenja regija temeljeno na srednjem intenzitetu trokuta. Potom uz podršku te segmentacije i dobivenog dispariteta autori formiraju ravninske segmente. Lokalizaciju autori provode sparivanjem trenutno detektiranih ravninskih segmenata s onima iz modela na karti. Potrebno im je minimalno dvije točne korespondencije da bi se robot lokalizirao. Inicijalno sparivanje se provodi pomoću srednjeg intenziteta ravninskog segmenta što čini ovaj sustav izuzetno osjetljivim na osvjetljenje scene. Konačnu optimizaciju položaja autori izvode pomoću Levenberg-Marquardtovog algoritma. Autori su prijavili srednju pogrešku lokalizacije od 0.3m i 5°. Na slici 2.18. vide se detektirane plohe koje služe kao orijentiri u razmatranom sustavu lokalizacije.



Slika 2.18. Primjer sustava koji koristi ravne plohe kao značajke za lokalizaciju mobilnog robota. Panoramska slika (gore), slika dispariteta (sredina) i detektirane plohe (dolje) [27].

U članku [22] autori provode globalnu lokalizaciju mobilnog robota korištenjem točkastih značajki opisanih SIFT deskriptorom. Također autori pokazuju mogućnost izvođenja istovremene lokalizacije i kartografije (engl. SLAM) te zatvaranja petlje za potrebu stvaranja karte prostora. Karta koju autori koriste za lokalizaciju mobilnog robota je

trodimenzionalna i sačinjena od podmapa generiranih od nekoliko susjednih slika okoline, a lokalizacija robota se vrši samo u tri stupnja slobode. Autori koriste trinokularni sustav za dobivanje podataka o dubini. Razmatrani sustav uzima u obzir nesigurnosti za opis stanja sustava, dok se sparivanje značajki zasniva na RANSAC (engl. RANdom SAmple Consensus) pristupu.

Autori [32] koriste robotski vid samo da bi približno lokalizirali mobilnog robota sa svesmjernom kamerom, dok kao deskriptore koriste MSIFT (engl. Modified SIFT). Prilikom stvaranja karte autori su točnu poziciju robota dobivali pomoću SLAM algoritma potpomognutog LRF senzorom. Za točno raspoznavanje lokacije autori traže da broj točnih korespondencija između trenutne slike i modela iz karte pređe određeni prethodno definirani prag. Kao mjeru sličnosti MSIFT deskriptora prilikom pronalaska korespondencija su koristili euklidsku udaljenost. Na slici 2.19. može se vidjeti primjer lokalizacije, gdje postoji trenutna slika mobilnog robota i slika modela te njihove pronađene korespondencije.



Slika 2.19. Primjer lokalizacije temeljene na slici koji koristi lokalne točkaste značajke [32]

U članku [33] autori ne provode lokalizaciju mobilnog robota, već više navigaciju pomoću raspoznavanja slika i prethodno stvorene topološke karte. Autori koriste svesmjernu kameru na mobilnom robotu te za proces navigacije koriste samo robotski vid. Za opis neke slike koriste niz SIFT deskriptora te stoga za raspoznavanje trenutne slike autori zahtijevaju da broj pronađenih korespondencija s obzirom na sliku u karti zadovoljava određeni prag. Za određivanje relativne poze između pogleda koriste epipolarnu geometriju poduprtu pronađenim korespondencijama i RANSAC algoritmom. Autori za testiranje navigacije uspoređuju putanju koju robot prolazi autonomno s putanjom ručno upravljanog robota.

Autori [34] koriste RANSAC uz određena geometrijska ograničenja da bi ubrzali postupak registracije scena, što nazivaju GCRANSAC (engl. Geometrically constrained RANSAC). Koriste stereo viziju na mobilnom robotu. Razvijeni sustav se zasniva na geometrijski usmjerenom sparivanju 3D točkastih značajki radi što bržeg postizanja konzistentnijeg skupa sparenih značajki.

Omogućavanje ažuriranja hibridne karte okoline tijekom vremena je istraživano u [11]. Globalno je karta topološka sa sfernim lokalnim metričkim modelima koji sadrže točkaste značajke opisane SURF deskriptorom. Za sustav ažuriranja karte autori uzimaju inspiraciju ljudskog pamćenja, te dijele podatke na one trenutne, one u kratkoročnoj memoriji i one u dugoročnoj memoriji. Na taj način se one značajke koje nisu stabilne u vremenu brišu iz karte. Sustav mobilnog robota koristi svesmjernu kameru za prikupljanje podataka iz okoline.

Robusniju lokalizaciju mobilnog robota autori u [35] nastoje realizirati definirajući novi deskriptor vertikalnih linija svesmjerne vizije. Prvi korak predloženog sustava je detektiranje vertikalnih linija izračunavanjem gradijenata svjetline slike u vodoravnom i okomitom smjeru pomoću Sobelovog operatora nakon čega se ostavljaju samo one točke čiji je kut gradijenta svjetline unutar neke tolerancije s obzirom na vertikalni smjer. Nakon toga, realiziranjem histograma od 720 binova koji pokrivaju raspon od 360° pogleda robota utvrđuju se najvjerojatnije vertikalne linije kao lokalni maksimumi tog histograma. Deskriptor detektiranih linija se proračunava kao histogram orijentacije gradijenta svjetline (engl. Histogram of Oriented Gradients - HOG) polukrugova opisanih oko linije, kao što se može vidjeti na slici 2.20. Dakle, deskriptor se sastoji od 3 para HOG-ova, odnosno histograma. Po tvrdnjama autora, dobiveni deskriptor je neovisan o rotaciji i djelomično neovisan o promjeni osvjetljenja. Autori kao metriku sličnosti deskriptora koriste L₂ normu, a također koriste i prošireni Kalmanov filter za fuziju podataka iz odometrije robota i vizije. Autori također u članku opisuju način kalibracije sustava sa svesmjernom kamerom s obzirom na odometriju mobilnog robota.



Slika 2.20. Primjer sustava koji koriste svesmjernu viziju i linijske značajke za lokalizaciju mobilnog robota. a) Točke koje zadovoljavaju zadani kut gradijenta svjetline, b) odabrane linije, c) formiranje deskriptora [35].

Autori [36] koriste ravninske segmente za registraciju položaja i lokalizaciju mobilnog robota. Testiraju tri tipa ugrađenih senzora, time-of-flight kameru Swiss-ranger, laserske skenere Odetics lidar i SICK S300. Za registraciju položaja koriste detekciju nesigurnosti dominantnih smjerova detektiranih ravninskih segmenata. Autori napominju da bi mogli koristiti boju ravninskih segmenata za usmjeravanje pronalaska korespondencije, ali to ne koriste ni u novijim člancima [37].

Također u [38] provode 3DOF lokalizaciju mobilnog robota zasnovanu na ravninskim segmentima. Koriste Microsoft Kinect senzor, ali samo dubinske slike te uspoređuju lokalizaciju dobivenu ovim senzorom s onim pomoću LRF-a. Lokalizacija je temeljena na čestičnim filtrima. Na temelju 3D oblaka točaka autori detektiraju ravninske segmente koristeći RANSAC metodu unutar postupno povećavajućeg prozora. Potom pomoću projiciranja zraka rade projekciju ravnina u 2D prostor koji se potom koristi za lokalizaciju mobilnog robota. Autori također demonstriraju lokalizaciju robota na putanjama duljim od 4km.

Lokalizacija mobilnog robota zasnovana na 3D ravninskim segmentima i čestičnim filtrima koristeći Kinect kameru opisana je i u [2]. Predstavljeni sustav lokalizacije robota zahtjeva prethodno generiranu kartu snimljenu pomoću iste kamere te se radi o geometrijskoj karti sastavljenoj od 3D ravninskih segmenata. Sustav radi na način da grafička kartica s OpenGL grafičkim podsustavom koristi Z-Buffer grafičke kartice za generiranje pretpostavke dubinske slike svake čestice. Potom autori vrše usporedbu i proračun vjerojatnosti za svaku točku slike da se radi o točno određenoj točci u prostoru, umnožak svih dobivenih vjerojatnosti predstavlja ukupnu vjerojatnost da se radi o ispravnoj sceni. Sustav koristi poduzorkovane slike dubine za usporedbu i time omogućava brzinu mobilnog robota od 1.5m/s. Za usporedbu autori ispituju dva tipa vjerojatnosti po točci, generativni i euklidski. Autori također napominju da bi mogli koristiti i boju točaka slike za usporedbu točnosti, ali ipak ne koriste. Za svoje testiranje autori koriste više tipova mobilnih robota i ostalih mobilnih platformi. Autori navode pogrešku lokalizacije robota od 0.3 do 0.66m ovisno o tipu mobilne platforme koja je radila snimanje prostora u odnosu na stvarni položaj. Na slici 2.21. može se vidjeti primjer renderiranog prostora i dubinske slike te stvarne RGB slike i dubinske slike u tom prostoru. Boju na renderiranoj slici autori koriste samo u svrhu vizualizacije.



Slika 2.21. Primjer sustava za lokalizaciju mobilnog robota temeljena na ravninskim segmentima dobivenih iz dubinske slike [2]

Efikasna metoda za praćenje lokalnog položaja mobilnog robota koja se temelji na registraciji ravninskih segmenata snimljenih Kinect senzorom predložena je u [3]. Korištenjem metoda geometrijskih ograničenja sličnima GCRANSAC-u [34] i novim načinom generiranja hipoteza pomoću strukture slične stablu postiže se srednje vrijeme lokalizacije od oko 0.1s pri srednjoj vrijednosti pogreške od 0.31m u odnosnu na stvarni položaj. Sustav lokalizacije predstavljen u [3] je osnovni sustav korišten u ovoj doktorskoj disertaciji i detaljno će biti objašnjen u poglavlju 4.

3. Primjena boje i teksture u sustavima računalnog vida

Vizualna obilježja kao što su boja ili tekstura nekog objekta uz njegovu geometriju predstavljaju primarne izvore informacija za njegov opis. Boje i teksture se uglavnom koriste u sustavima računalnog vida u svrhu raspoznavanja slika ili njihovu segmentaciju, dok su boje korištene i za potrebe lokalizacije mobilnog robota zasnovana na pojavnosti. U ovom poglavlju će se objasniti priroda boje i teksture te njihove tipične primjene kao i primjene u sustavima za lokalizaciju mobilnog robota.

3.1. Boja

Proces percipiranja i interpretiranja boje u ljudskom mozgu je psihofiziološki 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 zabilježeno je u 17-om stoljeću kada je Sir Isaac Newton otkrio da kada sunčeva svjetlost obasja staklenu prizmu na izlazu iz nje ne dobijemo isto obojanu svjetlost već cijeli spektar bolje od ljubičaste do crvene, kao što se vidi na slici 3.1.



Slika 3.1. Rasap svjetlosti na optičkoj prizmi

Otkriveno je da je u biti vidljiva svjetlost koju ljudi mogu percipirati relativno uzak spektar frekvencija u elektromagnetskom spektru kojeg se može vidjeti na slici 3.2.



Ono što ljudi i neke životinje percipiraju kao boju je u biti reflektirana svjetlost s površine nekog objekta. Stoga je karakteriziranje svjetlosti izuzetno bitno za znanost o bojama i korištenje boje u računalnom vidu. Ako svjetlost nema boje, tada je ona akromatska i promatramo je kao intenzitet.

Foto osjetljivi dio ljudskog vizualnog sustava se zove mrežnica i sadrži vidne stanice nazvane prema njihovom obliku, čunjići i štapići, koje se mogu vidjeti na slici 3.3.



Slika 3.3. Vidne stanice u mrežnici ljudskog oka [40]

Štapići su odgovorni za osjetljivost na intenzitet svjetla 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 osjetljivost najveća na valnim duljinama 420, 534 i 564nm. Također te valne duljine približno odgovaraju plavoj, zelenoj i crvenoj boji kao što se vidi na slici 3.4.



Slika 3.4. Spektralne osjetljivosti stanica ljudskog oka [39]

Otkriveno je stoga da se ljudska percepcija boja temelji na tripletu boja RGB prostora boja te je stoga taj prostor boja postao temeljni sustav boja i za digitalnu obradu slike.
Detaljnom analizom mrežnice je utvrđeno da prosječno ljudsko oko imo oko 6 do 7 miliona čunjića i da otprilike 65% njih je osjetljivo na crvenu svjetlost, 33% na zelenu svijetlost, a samo 2% na plavu (iako su ti čunjići najosjetljiviji).

Računalni vid za prikupljanje podataka iz okoline uglavnom koristi digitalne kamere bazirane na CCD ili CMOS tehnologijama koje pretvaraju svjetlost koja pada na senzor kamere prvo u analogni signal, a nakon toga u digitalni. Senzori u digitalnim kamerama se sastoje od 3 tipa fotoosjetljivih filtera (R, G i B) raspoređenih u određenom obrascu da bi se što bolje aproksimirao rad ljudskog vizualnog sustava. Primjer obrasca filtera je npr. Bayerov filter koji možemo vidjeti na slijedećoj slici, obrazac sadrži 50% zelenih, 25% crvenih i 25% plavih ćelija.



Slika 3.5. Primjer Bayerovog obrasca na senzoru kamere [15]

3.1.1. Refleksivni model formiranja slika

Modeliranje formiranja slike na senzoru kamere se uglavnom zasniva na tzv. refleksivnom modelu, također nazvanom i Lambertovim modelom [41]. Ovaj se model zasniva na tome da slika nekog objekta ovisi o svjetlu koje ga obasjava, karakterističnoj sposobnosti reflektiranja njegove površine i o odzivnim svojstvima senzora kamere. Takav model je donekle ograničen pretpostavkom da je objekt obasjan samo jednim svjetlosnim izvorom okarakteriziranim spektralnom snagom $E(\lambda)$ koja govori o količini energije emitirane po nekoj valnoj duljini λ . Refleksivna karakteristika koja opisuje koje valne duljine i u kojoj mjeri se reflektiraju od površine je definirana sa $S(\mathbf{x}, \lambda)$, gdje \mathbf{x} predstavlja prostorne koordinate točke, dok je senzor kamere karakteriziran funkcijom $Q_k(\lambda)$ koji predstavlja spektralnu osjetljivost k-tog senzora kamere. Tada je odziv k-tog senzora kamere

$$f(\mathbf{x})_{k} = \int_{\omega} E(\lambda) \cdot S(\mathbf{x}, \lambda) \cdot Q_{k}(\lambda) d\lambda \qquad k = 1, \dots, m \qquad (3.1)$$

gdje se integral izračunava nad područjem valnih duljina ω za koje se pretpostavlja da senzor kamere ima osjetljivost. Također broj senzora k je obično 3 te svaki senzor ima osjetljivost na valne duljine koje odgovaraju crvenom, zelenom i plavom spektru boja, odnosno $k \in R, G, B$. Stoga vrijednost neke točke slike koju daje kamera je kombinacija odziva 3 senzora za zadanu prostornu koordinatu **x**. Kao što je prethodno rečeno, izraz (3.1) odgovara modelu formiranja slike za Lambertove površine nad kojima se upadna svjetlost reflektira u svim smjerovima podjednako i to neovisno o smjeru upadne svjetlosti.

Shafer [42] je predložio da se formula (3.1) proširi i doda joj se utjecaj difuzijske svjetlosti te tada imamo:

$$f(\mathbf{x})_{k} = \int_{\omega} E(\lambda) \cdot S(\mathbf{x}, \lambda) \cdot Q_{k}(\lambda) d\lambda + \int_{\omega} A(\lambda) \cdot Q_{k}(\lambda) d\lambda \qquad k = 1, \dots, m \quad (3.2)$$

gdje je $A(\lambda)$ spektralna distribucija snage difuzijske svjetlosti. Time smo dobili tzv. dikromatski refleksijski model. Smatra se da je difuzijska svjetlost manjeg intenziteta i da dolazi iz svih smjerova podjednako. Promatrajući izraz (3.2) vidimo da se, ako se derivira slika u tom modelu po prostornoj koordinati, gubi se utjecaj difuzijskog svjetla, odnosno tada je drugi član izraza na lijevoj strani jednadžbe (3.2) jednak nuli.

Kao što se može vidjeti u izrazima (3.1) i (3.2) dobivena vrijednost odziva senzora odnosno slike ovisi i o uređaju koji snima i o osvjetljenju te se stoga dobiveni RGB triplet senzorskih odziva ne može smatrati intrističnim svojstvom površine nekog objekta. U literaturi postoji niz pokušaja da se umanji utjecaj osvjetljenja na način da se estimira spektralna distribucija svjetla okoline i na neki način ispravi slika i dobije stvarna slika površine objekta, to je tzv. problem konstantne boje (engl. color constancy). Primjer korištenja algoritma konstantne boje vidi se u [43] gdje autori raspoznaju objekte na slikama u svrhu indeksiranja te pri tome nastoje smanjiti utjecaj različitih osvjetljenja. Prvo računaju logaritam po svakom kanalu prostora boja za svaku točku slike, nakon čega računaju derivacije ili Laplacijan te potom dobivene vrijednosti unose u histogram nejednakih širina binova. Testiranje autori provode na relativno maloj bazi od 55 slika modela i 24 slike za testiranje.

Postoje i određeni drugi algoritmi konstantnih boja koji eliminiraju utjecaj osvjetljenja posjedovanjem podataka o kalibriranim odzivima kamera za niz osvjetljenja kao u [44]. Autori raspoznavaju objekte na slikama, pri čemu pokušavaju umanjiti utjecaj osvjetljenja. Rade na način da imaju prethodno spremljene histograme slika u rgb kromatskom prostoru

boja pod raznim osvjetljenjima i određuju vjerojatnost pojavljivanja određene boje pod određenim osvjetljenjem. Na taj način pomoću vjerojatnosti određuju o kojem se trenutnom osvjetljenju radi na slici i na takav način ublažuju utjecaj osvjetljenja na histogram boje slike koja se pokušava raspoznati. U drugom članku [45] isti autori sugeriraju korištenje metode ujednačavanja histograma, gdje se boje transformiraju radi povećanja kontrasta te rgb kromatskog prostora boja za ublažavanje utjecaja razlika korištenih kamera i osvjetljenja na slikama te time poboljšavaju raspoznavanje objekata na slikama u raznim uvjetima. Stoga, u mnogim drugim slučajevima kada nema pristupa kalibriranim kamerama i karakterističnim spektrima osvjetljenja, algoritmi konstantnih boja uglavnom nisu primjenjivi, pa se utjecaj osvjetljenja boja pokušava umanjiti korištenjem nekih drugih metoda kao kod [45].

Smanjenje utjecaja osvjetljenja je doista kompleksan problem zbog kompleksne interakcije svjetlosti, površine i senzora koje vidimo u izrazima (3.1) i (3.2). Stoga su znanstvenici pokušali smanjiti kompleksnost interakcije u izrazima korištenjem pojednostavljenih modela od kojih je najpopularniji dijagonalni model također nazvan i Von Kries model [46]. To je model definiran dijagonalnom transformacijom odziva senzora dobivenog pod nepoznatim svjetlom:

$$\mathbf{f}^c = \mathbf{D}^{u,c} \cdot \mathbf{f}^u \tag{3.3}$$

gdje je \mathbf{f}^{u} slika pod nepoznatim osvjetljenjem, a \mathbf{f}^{c} je ista slika transformirana kao da se nalazi pod referentnim osvjetljenjem također nazvanim i kanonskim osvjetljenjem. Tada je $\mathbf{D}^{u,c}$ dijagonalna matrica transformacije koja transformira boje snimljene pod nepoznatim svjetlom u odgovarajuće boje pod referentnim svjetlom te se izraz (3.3) može prikazati matrično:

$$\begin{bmatrix} R^{c} \\ G^{c} \\ B^{c} \end{bmatrix} = \begin{bmatrix} d_{1} & 0 & 0 \\ 0 & d_{2} & 0 \\ 0 & 0 & d_{3} \end{bmatrix} \cdot \begin{bmatrix} R^{u} \\ G^{u} \\ B^{u} \end{bmatrix}$$
(3.4)

Izrazu (3.4) se može dodati difuzijski element u obliku pomaka te se tada dobiva difuzijsko-pomaknuti model [47]:

$$\begin{bmatrix} R^{c} \\ G^{c} \\ B^{c} \end{bmatrix} = \begin{bmatrix} d_{1} & 0 & 0 \\ 0 & d_{2} & 0 \\ 0 & 0 & d_{3} \end{bmatrix} \cdot \begin{bmatrix} R^{u} \\ G^{u} \\ B^{u} \end{bmatrix} + \begin{bmatrix} o_{1} \\ o_{2} \\ o_{3} \end{bmatrix}$$
(3.5)

Kombinacijom formula (3.2) i (3.5) dobivamo sljedeći dijagonalni model:

$$\begin{bmatrix} E^{c}(\lambda_{R})\\ E^{c}(\lambda_{G})\\ E^{c}(\lambda_{B}) \end{bmatrix} = \begin{bmatrix} d_{1} & 0 & 0\\ 0 & d_{2} & 0\\ 0 & 0 & d_{3} \end{bmatrix} \cdot \begin{bmatrix} E^{u}(\lambda_{R})\\ E^{u}(\lambda_{G})\\ E^{u}(\lambda_{B}) \end{bmatrix} + \begin{bmatrix} A(\lambda_{R})\\ A(\lambda_{G})\\ A(\lambda_{B}) \end{bmatrix}$$
(3.6)

Budući da je refleksivna karakteristika površine $S(\mathbf{x}, \lambda)$ jednaka i za nepoznatu i za referentnu svjetlost, slijede izrazi:

$$E^{c}\left(\lambda_{R}\right) \cdot S\left(\mathbf{x}, \lambda_{R}\right) = d_{1} \cdot E^{u}\left(\lambda_{R}\right) \cdot S\left(\mathbf{x}, \lambda_{R}\right) + A\left(\lambda_{R}\right)$$
(3.7)

$$E^{c}\left(\lambda_{G}\right) \cdot S\left(\mathbf{x}, \lambda_{G}\right) = d_{2} \cdot E^{u}\left(\lambda_{G}\right) \cdot S\left(\mathbf{x}, \lambda_{G}\right) + A\left(\lambda_{G}\right)$$
(3.8)

$$E^{c}\left(\lambda_{B}\right) \cdot S\left(\mathbf{x}, \lambda_{B}\right) = d_{3} \cdot E^{u}\left(\lambda_{B}\right) \cdot S\left(\mathbf{x}, \lambda_{B}\right) + A\left(\lambda_{B}\right)$$
(3.9)

Uzimajući u obzir dijagonalni model u izrazu (3.4) odnosno dijagonalno-pomaknuti model u izrazu (3.5) postoji pet mogućih transformacija slike s obzirom na osvjetljenje:

- Promjena intenziteta osvjetljenja s obzirom na referentno osvjetljenje (
 d₁ = d₂ = d₃ ≠ 0, o₁ = o₂ = o₃ = 0).
- Promjena boje osvjetljenja s obzirom na referentno osvjetljenje, (d₁ ≠ d₂ ≠ d₃,
 o₁ = o₂ = o₃ = 0).
- Pomak intenziteta svijetlosti gdje je $(\boldsymbol{d}_1 = \boldsymbol{d}_2 = \boldsymbol{d}_3 = 1, \boldsymbol{o}_1 = \boldsymbol{o}_2 = \boldsymbol{o}_3 \neq 0).$
- Promjena intenziteta osvjetljenja s obzorom na referentno osvjetljenje i pomak intenziteta svijetlosti (*d*₁ = *d*₂ = *d*₃ ≠ 0, *o*₁ = *o*₂ = *o*₃ ≠ 0).
- Promjena boje osvjetljenja s obzirom na referentno osvjetljenje i pomak boje osvjetljenja, (*d*₁ ≠ *d*₂ ≠ *d*₃, *o*₁ ≠ *o*₂ ≠ *o*₃).

Da bi se boje specificirale na općeniti i standardizirani način definirani su standardni prostori boja. Odabir prostora boja za rad ovisi o željenoj primjeni (prikaz, ispis, manipulacija, itd.). Budući da je primjena uporabe boja jako široka, postoji veliki izbor prostora boja. U sljedećih nekoliko podpoglavlja objasnit će se prostori boja koji se razmatraju u ovoj disertaciji.

3.1.2. RGB prostor boja

U RGB prostoru boja svaka boja je sastavljena od primarnih spektralnih komponenti crvene, zelene i plave boje. Ovaj se model bazira na Cartezijevom koordinatnom sustavu. RGB prostor boja može se predočiti slikom 3.6.



Slika 3.6. RGB prostor boja [39]

Kao što se može vidjeti na slici 3.6. primarne boje RGB prostora se nalaze na osima koordinatnog sustava i kutevima kocke, sekundarne boje cijan, magenta i žuta se nalaze na drugim kutovima kocke. Ishodište koordinatnog sustava je crna boja, a točka kocke najudaljenija od ishodišta je bijela. Zbog svoje jednostavnosti RGB prostor boja je najrašireniji u uporabi, ali radi se o prostoru koji nije perceptualno uniforman što znači da euklidska udaljenost između dvije boje u tom prostoru ne odgovaraju perceptualnoj udaljenosti tih dviju boja u ljudskom vizualnom sustavu. Također, ovaj prostor boja nije intuitivan s obzirom na ljudski način razmišljanja o bojama. S obzirom na dijagonalnopomaknuti model (3.5) ovaj prostor boja je ovisan o svim promjenama osvjetljenja što je sasvim logično ako se uzme u obzir i priroda ovog prostora boja.

3.1.3. Kromatski rgb prostor boja

Ovaj prostor boja se jednostavno izvodi iz RGB prostora boja i predstavlja kromatske komponente RGB prostora boja lišene podataka o intenzitetu, koje se računaju na sljedeći način.

$$r = \frac{R}{R+G+B} \tag{3.10}$$

$$g = \frac{G}{R+G+B} \tag{3.11}$$

$$b = \frac{B}{R+G+B} \tag{3.12}$$

Također budući da vrijedi:

$$b = 1 - r - g \tag{3.13}$$

ovaj prostor se samo promatra pomoću r i g kromatske komponente. Na sljedećoj slici može se vidjeti prostor boja koji je ovdje opisan.



Slika 3.7. Kromatski rg prostor boja

Ovaj prostor boja s obzirom na dijagonalno-pomaknuti model (3.5) je neovisan samo o promjeni intenziteta osvjetljenja.

3.1.4. HSV prostor boja

RGB prostor boja je primjeren za sustave snimanja i reprodukcije boja zbog njegove povezanosti s fizikalnim svojstvima, ali nije primjeren za interpretaciju boja budući da nije blizak ljudskom načinu razmišljanja o boji. Drugim riječima kada ljudi pokušavaju opisati neku boju uglavnom se služe terminima kao što su nijansa, zasićenje i svjetlina. Dakle, karakteristike koje se općenito najčešće koriste za razlikovanje jedne boje od druge kod ljudi su svjetlina (engl. brightness, value), nijansa (engl. hue) i zasićenje (engl. saturation). Svjetlina predstavlja akromatsko shvaćanje intenziteta svjetla. Dok se nijansa povezuje s dominantnom valnom duljinom koja prevladava u svjetlu odnosno nijansa predstavlja ono što se obično kod ljudi smatra bojom. Zasićenje predstavlja relativnu čistoću boje odnosno količinu bijele boje pomiješane s određenom nijansom. HSV prostor boja u biti predstavlja cilindričnu transformaciju RGB prostora. Prednost ovog prostora boja je što se radi o relativno jednostavnoj transformaciji RGB prostora te se stoga može brzo izvesti i to što je ovaj način opisa boja puno bliži ljudskom načinu razmišljanja. Nedostaci su slični kao i kod RGB prostora, a to je da nije perceptualno uniforman. Također postoje slični cilindrični prostori boja nazvani HSI i HSL koji se formiraju na slične načine.



Slika 3.8. HSV prostor boja [15]

Transformacija iz RGB prostora u HSV prostor je sljedeća:

$$H = \begin{cases} \left(\frac{G-B}{\max\left(R,G,B\right) - \min(R,G,B)}\right) \cdot 60^{\circ}, ako \ je \ \max\left(R,G,B\right) = R \\ \left(\frac{B-R}{\max\left(R,G,B\right) - \min(R,G,B)}\right) \cdot 60^{\circ}, ako \ je \ \max\left(R,G,B\right) = G \\ \left(\frac{R-G}{\max\left(R,G,B\right) - \min(R,G,B)}\right) \cdot 60^{\circ}, ako \ je \ \max\left(R,G,B\right) = B \\ S = \frac{\max\left(R,G,B\right) - \min(R,G,B)}{\max(R,G,B)} \end{cases}$$
(3.15)

$$V = \max(R, G, B) \tag{3.16}$$

HSV prostoru boja s obzirom na dijagonalno-pomaknuti model (3.5) H komponenta postaje neovisna o promjeni i pomaku intenziteta osvjetljenja. Dok je S komponenta neovisna samo u slučaju promjene intenziteta osvjetljenja.

3.1.5. CIE 1931 XYZ prostor boja

XYZ prostor boja je jedan od prvih matematički definiranih prostora boja. Nastao je 1931. godine od strane organizacije CIE (engl. International Commission on Illumination, fran. Commission Internationale de L'Éclairage). Razlog njegovog nastanka je da se bolje opiše prostor boja koje prosječno ljudsko oko može vidjeti, za razliku od RGB prostora boja, koji više ima namjenu opisati boje koje se mogu reproducirati standardnim uređajima ili snimiti kamerom. Ovaj prostor boja je nastao kao rezultat eksperimenata u kasnim 1920-ima od strane W. D. Wrighta i J. Guilda. U ovom prostoru boja X se može smatrati linearnom kombinacijom R, G i B boje, Y se može smatrati svjetlinom, dok Z komponenta odgovara donekle plavoj boji. Gamut boja odnosno raspon boja koji se može dobiti XYZ prostorom boja je prikazan xy kromatskim dijagramom na slici 3.9a, dok se na slici 3.9b, u trokutu čiji su vrhovi primarne boje (R, G i B), može vidjeti raspon boja, opisan sRGB (standardizirani RGB) prostorom boja koje se mogu reproducirati standardnim ekranima.



Slika 3.9. a) CIE XYZ raspon boja, b) sRGB raspon boja [15]

CIE XYZ prostor boja se dobiva iz RGB prostora boja (pomoću podataka D65 referentne bijele boje) na slijedeći način:

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 0,4125 & 0.3576 & 0.1804 \\ 0.2127 & 0.7152 & 0.0722 \\ 0.0193 & 0.1192 & 0.9503 \end{bmatrix} \cdot \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$
(3.17)

S obzirom na opisanu transformaciju RGB prostora boja u CIE XYZ prostor boja, novi prostor boja posjeduje slične ovisnosti o promjeni i pomaku intenziteta svijetlosti s obzirom na dijagonalno-pomaknuti model.

3.1.6. CIE L*a*b* prostor boja

CIE L*a*b* prostor boja je jedan od posljednjih objavljenih standardnih prostora boja od organizacije CIE. Objavljen je 1976. godine i zasniva se na CIE XYZ prostoru boja. Glavna značajka mu je perceptulana uniformnost, što znači da je euklidska udaljenost između dvije boje u ovom 3D prostoru boja jednaka relativnoj perceptualnoj razlici između te dvije boje u ljudskom vizualnom sustavu. Sastoji se od svjetline L* i dvije kromatske komponente a* (smještena između zelene i magente) i b* (smještena između plave i žute). Na slici 3.10. može se vidjeti sRGB raspon boja prikazan u CIE L*a*b* prostoru boja za različite vrijednosti svjetline L*.



Slika 3.10. CIE L*a*b* prostor boja za tri razine svjetline [15]

Pretvorba iz CIE XYZ u CIE L*a*b* prostor boja definirana je slijedećim izrazima:

$$L^* = 116 \cdot f\left(\frac{Y}{Y_N}\right) - 16 \tag{3.18}$$

$$a^* = 500 \cdot \left[f\left(\frac{X}{X_N}\right) - f\left(\frac{Y}{Y_N}\right) \right]$$
(3.19)

$$b^* = 200 \cdot \left[f\left(\frac{Y}{Y_N}\right) - f\left(\frac{Z}{Z_N}\right) \right]$$
(3.20)

gdje je funkcija *f*:

$$f(s) = \begin{cases} s^{1/3}, \ akos > \left(\frac{6}{29}\right)^3 \\ \frac{1}{3} \cdot \left(\frac{29}{6}\right)^2 \cdot s + \frac{4}{29}, \ ostalo \end{cases}$$
(3.21)

Pri čemu je (X_N, Y_N, Z_N) referentna bijela boja u XYZ prostoru boja. Postoji također još jedan sličan prostor boja nazvan CIE L*u*v* koji koristi istu definiciju *L** ali različite

definicije za kromatske komponente koje su u tom prostoru nazvane u* i v*. CIE L*a*b* prostor boja je ovisan o promjenama osvjetljenja, ali korištenjem samo a* i b* komponente ta ovisnost se može smanjiti.

3.1.7. Prostor suprotstavljenih boja

Ovaj prostor boja se još naziva i $O_1O_2O_3$ prostor boja te se iz RGB prostora boja dobiva na slijedeći način:

$$O_1 = R - G \tag{3.22}$$

$$O_2 = R + G - 2 \cdot B \tag{3.23}$$

$$O_3 = R + G + B \tag{3.24}$$

gdje O_1 i O_2 predstavljaju kromatske komponente, a O_3 svjetlinu. Također, ako je poznato neko osvjetljenje $c = (\alpha, \beta, \gamma)$ ovi izrazi se normiraju na sljedeći način [41]:

$$O_1 = \frac{\beta \cdot R - \alpha \cdot G}{\sqrt{\alpha^2 + \beta^2}} \tag{3.25}$$

$$O_{2} = \frac{\alpha \cdot \gamma \cdot R + \beta \cdot \gamma \cdot G \cdot (\alpha^{2} + \beta^{2}) \cdot B}{\sqrt{(\alpha^{2} + \beta^{2} + \gamma^{2}) \cdot (\alpha^{2} + \beta^{2})}}$$
(3.26)

$$O_3 = \frac{\alpha \cdot R + \beta \cdot G + \gamma \cdot B}{\sqrt{\alpha^2 + \beta^2 + \gamma^2}}$$
(3.27)

Budući da se često uzima da je c = (1,1,1), tada dobivamo konačne izraze:

$$O_1 = \frac{R - G}{\sqrt{2}} \tag{3.28}$$

$$O_2 = \frac{R + G - 2 \cdot B}{\sqrt{6}} \tag{3.29}$$

$$O_3 = \frac{R+G+B}{\sqrt{3}} \tag{3.30}$$

Ako promatramo samo kromatske komponente ovog prostora boja s obzirom na dijagonalni-pomaknuti model (3.5) možemo zaključiti da se one neovisne o pomaku intenziteta svjetlosti. Odnosno ako R, G i B komponenti dodamo iste vrijednosti, one će se za O_1 i O_2 komponentu pokratiti. Nažalost ove komponente nisu neovisne o promjeni intenziteta svjetlosti, a budući da treća komponenta O_3 sama po sebi predstavlja intenzitet svjetlosti ona ne može biti neovisna o promjeni, a ni o pomaku intenziteta svjetlosti.

3.1.8. Deskriptori boja

Korištenje boja odnosno slika za raspoznavanje objekata je počelo početkom 90-ih godina prošlog stoljeća kad je izašao značajni članak "Color Indexing" autora Swaina i Ballarda [48]. U tom članku su autori pokazali da se boja može koristiti za raspoznavanje slika objekata pri različitim kutevima i u slučaju okluzije na način da se formiraju trodimenzionalni histogrami kao deskriptori slika. Na slici 3.11. može se vidjeti primjer objekta dan u [48] u četiri položaja i odgovarajuće trodimenzionalne histograme boja.



Slika 3.11. Primjer objekta gledanog iz četiri različita kuta i odgovarajući 3D histogrami boja
[48]

Autori su kao prostor boja koristili tzv. prostor suprotstavljenih boja koji je definiran na način prikazan u poglavlju 3.1.7, iako autori sugeriraju da se njihov sustav raspoznavanja može koristiti za bilo koji prostor boja. Također su u članku [48] definirali novu mjeru sličnosti između histograma nazvanu "presjek histograma" (engl. histogram intersection) koja je definirana izrazom:

$$H(\mathbf{I}, \mathbf{M}) = \frac{\sum_{i=1}^{n} \min(I_i, M_i)}{\sum_{i=1}^{n} M_i}$$
(3.31)

gdje je I histogram slika, M histogram modela, dok su I_i i M_i vrijednosti zapisane u *i*-tom binu histograma slike odnosno modela. Po tom izrazu može se zaključiti da presjek histograma daje mjeru preklapanja dva histograma na način da zbraja najmanju količinu podataka (u ovom slučaju točaka slike) raspoređenu po binovima, koji su zajednički za oba histograma. Izraz je također normiran s obzirom na količinu podataka (točaka slike) histograma modela te je time vrijednost presjeka ograničena na interval između 0 i 1, pri čemu je 0 najmanja, a 1 najveća sličnost između dva histograma. Autori su proveli svoje eksperimente na relativno maloj bazi slika objekata bez utjecaja promjena osvjetljenja, ali uz promjene položaja i postojanje okluzije, te su utvrdili da je prosječna točnost raspoznavanja njihove metode oko 98%.

Dakle, glavni deskriptor značajke boje je statistička mjera nazvana histogram. Histogram je u biti aproksimacija funkcije gustoće vjerojatnosti slučajne varijable. Primjer izgleda jednodimenzionalnog histograma Gaussove distribucije može se vidjeti na slici 3.12a.



Slika 3.12. Primjer a) 1D histograma Gaussove distribucije, b) presjeka dva 1D histograma

Na slici 3.12b može se vidjeti presjek dva jednodimenzionalna histograma, crvenog i plavog. Presjek histograma je označen ljubičastom bojom. Jednodimenzionalni histogrami se često koriste u digitalnoj analizi slike za prikaz distribucije R, G i B komponente pojedine slike. Na slici 3.13. mogu se vidjeti tri jednodimenzionalna R, G i B histograma fotografije prikazane na lijevoj strani.



Slika 3.13. Primjer 1D histograma R, G i B kanala slike u boji

U slučaju prikazanim slikom 3.13. histogrami su za svaku dimenziju RGB prostora boja generirani i prikazani zasebno, ali histogram može biti i višedimenzionalan kako je predloženo u [48]. Na slici 3.14. prikazana su dva trodimenzionalna histograma RGB prostora fotografije sa slike 3.13. Na slici 3.14a histogram ima 256 binova po dimenziji, dok na slici 3.14b ima 7 binova po dimenziji.



Slika 3.14. Primjer 3D histograma RGB prostora boja

Pored presjeka histograma (3.31) koji su uveli Swain i Ballard [48] postoji mnogo metrika koje se mogu koristiti za određivanje sličnosti ili razlike između dva histograma. Ovdje su navedene neke od popularnijih:

• L₁ norma:

$$L_{1}(\mathbf{I},\mathbf{M}) = \sum_{i=1}^{n} |I_{i} - M_{i}|$$
(3.32)

• L₂ norma:

$$L_{2}(\mathbf{I},\mathbf{M}) = \sqrt{\sum_{i=1}^{n} (I_{i} - M_{i})^{2}}$$
(3.33)

• Bhattacharyya udaljenost:

$$B(\mathbf{I}, \mathbf{M}) = -\ln \sum_{i=1}^{n} \sqrt{I_i \cdot M_i}$$
(3.34)

• Matusita udaljenost:

$$M\left(\mathbf{I},\mathbf{M}\right) = \sqrt{\sum_{i=1}^{n} \left(\sqrt{I_i} - \sqrt{M_i}\right)^2}$$
(3.35)

38

• Divergencija:

$$D(\mathbf{I}, \mathbf{M}) = \sum_{i=1}^{n} \left((I_i - M_i) \cdot \ln \frac{I_i}{M_i} \right)$$
(3.36)

• Jeffrey divergencija:

$$d\left(\mathbf{I},\mathbf{M}\right) = \sum_{i=1}^{n} \left(I_i \cdot \log \frac{2 \cdot I_i}{I_i + M_i} + M_i \cdot \log \frac{2 \cdot M_i}{I_i + M_i} \right)$$
(3.37)

• χ^2 udaljenost:

$$\chi^{2}(\mathbf{I}, \mathbf{M}) = \sum_{i=1}^{n} \frac{(I_{i} - M_{i})^{2}}{I_{i} + M_{i}}$$
(3.38)

• Normalizirani koeficijent korelacije:

$$r(\mathbf{I}, \mathbf{M}) = \frac{\sum_{i=1}^{n} (I_i - \mu_I) \cdot (M_i - \mu_M)}{\sqrt{\sum_{i=1}^{n} (I_i - \mu_I)^2 \cdot \sqrt{\sum_{i=1}^{n} (M_i - \mu_M)^2}}}$$
(3.39)

Nakon članka [48], podaci o bojama na slici su se počeli koristiti na više područja digitalne obrade slike od pretraživanja baza slika do segmentacije slike, a kasnije i kod lokalizacije mobilnog robota zasnovanog na pojavnosti.

Primjer korištenja vizualnih obilježja za indeksiranje slika možemo vidjeti u [49] koji koristi RGB prostor boja i određene druge statističke mjere kao što su teksturiranost, gustoća rubova i sl. radi poboljšanja deskriptivnosti RGB histograma i raspoznavanja slika. Pokusi su provedeni na relativno velikoj bazi od 210000 slika.

U [50] se ispituju različiti prostori boja u svrhu prepoznavanja slika objekata. Također ispituje se ovisnost raspoznavanja objekta o njegovoj geometriji, osvijetljenosti i položaju. Autori koriste bazu od 500 slika i prostore boja; RGB, kromatski rgb, HSV te uvode određene nove bojevne sustave nazvane $l_1 l_2 l_3$, $c_1 c_2 c_3$, $m_1 m_2 m_3$ kojima se nastoji umanjiti utjecaj osvjetljenja.

U [51] koristi se CIE L*u*v* prostor boja za indeksiranje slika gdje se histogram formira na način da je svaki bin otežan ovisno o volumenu CIE L*u*v* prostora boja koji je sadržan u tom binu. Kao mjeru sličnosti dva histograma ovi autori koriste Swainov presjek histograma.

U [52] opisano je raspoznavanje slika korištenjem 2D histograma prostora suprotstavljenih boja komprimiranih na sličan način kako se komprimiraju i 2D slike. Ispituju se 3 tipa algoritma za kompresiju, te postižu kompresije i do 250 puta. Kao mjeru sličnosti komprimiranih histograma koristi se euklidska udaljenost te time postiže i 99%-tna točnost raspoznavanja slika uz korištenje malog broja dobivenih koeficijenata transformacije odnosno kompresije.

Još jedna metoda indeksiranja slika u bazi predložena je u [53], gdje se indeksiranje izvodi na temelju statističkih podataka kao što su srednja vrijednost, varijanca i nesimetrija histograma za svaki od tri kanala RGB prostora boja zasebno. Time se formira 9-elementni vektor značajki koji se koristi prilikom usporedbe slika.

Modifikacija histograma predlaže se u [54] gdje uvode pojam mekih histograma boja u službu analize slika. Autor koristi cos² funkciju za određivanje stupnja pripadnosti neke vrijednosti binovima u čijoj se okolini nalazi. Korisnost mekih histograma ispituje na primjenama kao što su lokalizacija vršnih vrijednosti aproksimacije funkcije gustoće vjerojatnosti i određivanje kuta rotacije slike.

U cilju praćenja deformabilnih objekata, kao što su ljudi na sekvencama slika, u [55] se koriste čestični filteri i RGB prostor boja. Za opis objekata koristi se RGB histogram proračunat preko kvadratnog prozora na individualnoj slici u sekvenci, dok se kao mjera sličnosti između histograma objekata koristi Bhattacharyya udaljenost.

U [56] se za raspoznavanje slika koriste histogrami gradijenta intenziteta za svaki kanal boje umjesto samog intenziteta boje. Gradijenti se proračunavaju Sobelovim filterima. Autorima je bio cilj povećati neovisnost histograma kao deskriptora slike s obzirom na promjenu položaja i deformaciju slike.

Postoje i opširnije analize deskriptora boja kao [57], gdje se analiziraju i testiraju tri tipa deskriptora bazirana na boji: histogrami, momenti boja i varijante SIFT deskriptora, u svrhu raspoznavanja klasa objekata na slikama. Testiranje je izvedeno na dvije baze slika i jednoj bazi video zapisa. Primarna je ideja autora da analiziraju različite utjecaje promjena osvjetljenja slike na performanse raspoznavanja objekata odnosno klasa objekata na slici, pri čemu su autori zaključili da neovisnost nekog deskriptora na pojedine tipove promjena osvjetljenja ne garantira uspjeh raspoznavanja, budući da postoje drugi utjecaji, kao što su promjena kuta gledanja kamere, promjena položaja svjetla s obzirom na objekt ili čak

raspoznavanja. Autori zaključuju da, iako pojedini deskriptori bazirani na histogramu dobro raspoznavaju prilikom promjene tipa osvjetljenja, općenito gledajući SIFT i njegovi derivati koji uzimaju u obzir i boju imaju bolje rezultate raspoznavanja objekata odnosno klase objekata na slikama i video zapisima.

U nekim drugim slučajevima boje mogu biti samo jedan manji dio kompleksnijeg deskriptora za opis slika. Npr. u [58] je dan pregled paradigme nazvane "vreća značajki" (engl. Bag of features) u svrhu indeksiranja slika odnosno klasifikacije i dohvata slika iz baze. Vreće značajki su u biti višedimenzionalni histogrami u kojima podaci o boji nekog objekta mogu biti par dimenzija tog deskriptora dok ostale dimenzije mogu biti podaci o teksturama, obliku i sl.

Također, u drugim slučajevima [59] boja može služiti kao dodatna informacija za segmentiranje RGB-D slike dobivene Kinect senzorom na ravne plohe. Autorov algoritam uključuje detekciju rubova od kojih stvara regije, koje potom dijeli na trokute Delaunay triangulacijom i segmentira algoritmom normaliziranih rezova pri čemu se kao mjera konzistentnosti ravnine koristi podatak o srednjoj boji trokuta. Dobivene regije segmentirane na takav način autorima sugeriraju moguću segmentaciju dubinske slike, te se pomoću RANSAC algoritma definiraju konačne ravnine pojedine RGB-D scene.

3.2. Tekstura

Teksture su karakteristične promjene u intenzitetu slike (ili boje slike) koje se tradicionalno pripisuju grubosti ili hrapavosti površine nekog objekta. U tom slučaju dobro definirana tekstura se smatra kada promjene intenziteta slike imaju i nasumične i pravilne karakteristike, te je iz tog razloga potrebno biti jako pažljiv prilikom dizajniranja određenih statističkih mjera za analizu teksture.

Dakle, tradicionalna definicija teksture jest da je ona svojstvo površine koja proizvodi lokalne promjene u intenzitetu slike. Često je tome razlog grubost površine koja nasumično rasipa svjetlost, te stoga pojačava ili smanjuje lokalnu refleksivnost pri određenim kutovima gledanja. To vrijedi za materijale kao što su npr. papir ili ljuska jajeta. Ali često grubost površine nije uzrok teksture kao što je npr. tkanina kod koje struktura materijala uzrokuje periodičnu promjenu u intenzitetu slike. Također imamo materijale kao što je drvo koje može biti obrađeno i glatko a opet teksturirano zbog intrističnih svojstava materijala. Uzroci tekstura mogu biti i određeni prostorno-vremenski poremećaji kao što su valovi na površini

vode. Stoga tradicionalno povezivanje teksture isključivo s grubosti površine nije uvijek ispravno. Na slici 3.15 se mogu vidjeti tipični primjeri tekstura u prirodi.



Slika 3.15. Primjeri različitih tipova tekstura u prirodi [60]

Nadalje, primjeri opet drugih tipova tekstura mogu biti pijesak, travnjak ili grm. Tada se teksture sastoje od mješavine zrnaca ili listova, odnosno od individualnih objekata. U tom slučaju se površine svejedno interpretiraju kao da imaju jedinstvenu teksturu, ali ako se promjeni skala tada se mogu vidjeti individualni objekti te se gubi dojam teksture, kao što se može vidjeti na slici 3.16.



Slika 3.16. Promjenom skale se u nekim slučajevima gubi osjećaj teksture, dok se u drugima ne gubi [60]

Stoga, površina se smatra teksturiranom ako je na njoj nemjerljivo mnogo elemenata teksture nazvanih tekselima (engl. texel) [60], u suprotnom se smatra skupinom objekata. Općenito gledajući, elementi teksture odnosno tekseli mogu biti razmješteni na bilo koji

način, nasumično, pravilno, usmjereno itd. Ono što se u biti želi postići uvođenjem teksela je mogućnost raspoznavanja pomoću rekonstrukcije. Odnosno ako je tekstura uspješno rekonstruirana onda je ispravno i interpretirana. Nedostatak ove metode je utjecaj nasumične komponente teksture koje mogu uzrokovati velike pogreške.

Zbog svih tih razloga interpretacija teksture se udaljava od definicije prema kojoj se radi o svojstvu površine, već se radije može govoriti o izgledu neke površine na slici. Također i ono što se smatra teksturom je sklono subjektivnoj interpretaciji. Npr. ako imamo isključivo periodični obrazac, onda se to uglavnom ne smatra teksturom. Također, ako imamo isključivo nasumični obrazac, onda se on uglavnom smatra šumom. Iz toga se zaključuje da se samo kombinacija nasumičnih i pravilnih karakteristika može smatrati teksturom. Teksture se također mogu razmatrati i analizirati na način da se podrazumijeva da je teksturirana površina globalno pravilna odnosno da makrostrukture na površini budu uniformne, a da je lokalno tekstura nasumična odnosno da su mikrostrukture na površini nasumično raspoređene.

Analiza tekstura se do sada uglavnom koristila za raspoznavanje pojedine teksture na 2D slici odnosno segmentaciju 2D slika na teksturno uniformna područja. U slijedećih nekoliko podpoglavlja objasnit će se tri popularnije metode za analizu tekstura: LBP (engl. Local Binary Pattern), tekstoni i Gabor filteri.

3.2.1. LBP porodica deskriptora teksture

LBP (engl. Local Binary Pattern) je strukturalno-statistički deskriptor teksture baziran na detektiranju lokalnih struktura na određenoj skali, tzv. lokalnih binarnih obrazaca. Ovaj tip deskriptora se prvi puta spominje u [61], gdje autori uvode lokalni opis teksture gledajući 3x3 okolinu. LBP vrijednost nad tom okolinom se definira na način da se okolne točke uspoređuju sa središnjom i ako su svjetlije ili iste svjetline pridružuje im se vrijednost '1', a ako su tamnije '0'. Na temelju tog pridruživanja izračunava se LBP vrijednost kao suma višekratnika boja 2 pomnoženih s binarnim vrijednostima dobivenih iz okoline. Primjer dobivanja LBP vrijednosti se može vidjeti na slici 3.17.



Slika 3.17. Način dobivanja LBP vrijednosti [62]

Provođenjem takve analize nad svim točkama nekog područja na slici omogućeno je dobivanje distribucije LBP vrijednosti odnosno korištenje jednodimenzionalnog histograma s 2⁸ (mogućih LBP vrijednosti) binova kao deskriptora teksture.

Kasnije se definicija LBP deskriptora proširila tako da se okolina točke može promatrati na proizvoljnoj udaljenosti s proizvoljnim brojem točaka kao što se može vidjeti na slici 3.18., dok je originalno ograničenje bilo samo za 3x3 okolinu neke točke.



Slika 3.18. LBP vrijednost je moguće definirati nad većom okolinom s proizvoljnim brojem točaka [62]

Zbog uvođenja proizvoljne udaljenosti i proizvoljnog broja točaka moguć je različiti položaj uzorkovanih točaka s obzirom na centralnu točku, pa se stoga svjetlina jedne točke uzima samo u slučaju ako uzorak pada direktno na centar točke. U svim ostalim slučajevima se uzima bilinearna interpolacija s obzirom na svjetline okolnih točaka.

LBP vrijednost za neku točku se tada računa na sljedeći način:

$$LBP_{P,R} = \sum_{p=0}^{P-1} s \left(g_p - g_c \right) \cdot 2^p$$
(3.40)

gdje je,

$$s(x) = \begin{cases} 1, x \ge 0\\ 0, x < 0 \end{cases}$$
(3.41)

a g_p i g_c su svjetlina uzorka u okolini neke točke i svjetlina te točke čiju okolinu promatramo, dok je *P* broj točaka u okolini, a *R* radijus razmatrane okoline. Time je broj LBP vrijednosti koje se mogu definirati na takav način 2^{*P*}.

Pored ove osnovne verzije LBP deskriptora, isti autori [62] su razvili dvije podvarijante LBP deskriptora nazvane $LBP_{P,R}^{riu2}$ i $LBP_{P,R}^{riu2} / VAR_{P,R}$.

Deskriptor $LBP_{P,R}^{riu2}$ je rotacijski invarijantni oblik osnovnog LBP deskriptora, koji se fokusira samo na detekciju tzv. "uniformnih" oblika. Uniformnim se smatraju samo oni oblici koji imaju neprekinuti niz '0' ili '1', odnosno smiju imati samo jedan (ili nijedan) prijelaz $0 \rightarrow 1$ i $1 \rightarrow 0$. Time je smanjen broj mogućih LBP vrijednosti na samo P+1.

Izraz za izračunavanje uniformnih LBP vrijednosti je

$$LBP_{P,R}^{riu2} = \begin{cases} \sum_{p=0}^{P-1} s(g_p - g_c), \text{ako je } U(LBP_{P,R}) \le 2\\ P+1, \text{u ostalim slučajevima} \end{cases}$$
(3.42)

gdje je,

$$U(LBP_{P,R}) = \left| s(g_{P-1} - g_c) - s(g_0 - g_c) \right| + \sum_{p=1}^{P-1} \left| s(g_p - g_c) - s(g_{p-1} - g_c) \right|$$
(3.43)

Na slici 3.19. mogu se vidjeti uniformni oblici ako se gleda samo osam točaka u okolini.

•	٠	•	•	0	•	٠	o	0	•	0	0	•	0	о	•	0	0	•	0	0	٠	0	0	0	0	0
•	0	٠	٠	1	٠	٠	2	٠	٠	3	0	٠	4	0	٠	5	0	٠	6	0	0	7	0	0	8	0
٠	•	•	٠	•	•	٠	•	٠	٠	•	•	•	•	0	•	0	0	0	о	0	0	0	0	0	0	0
				S	Slika	31	0	I Ini	forr	nni	IR	Р 117	zor	<u>.</u>	anto	lir	<u>, , , , , , , , , , , , , , , , , , , </u>	d 8 1	oč	aba	[62]					

Slika 3.19. Uniformni LBP uzorci za okolinu od 8 točaka [62]

U [62] je objašnjeno da ti uniformni oblici odgovaraju standardnim obrascima u teksturama kao što su rubovi, uglovi, završeci linija, točke i monotone regije. Također tvrde da ti uniformni oblici u analizama čine preko 90 posto dobivenih LBP uzoraka iz tekstura pa je stoga logično promatrati samo te oblike. Neki autori [63] provode analizu LBP algoritma i utvrđuju da je razlog zašto u distribuciji LBP vrijednosti dominiraju uniformni obrasci (75% po njihovoj analizi) sama matematička struktura LBP algoritma a ne karakteristika teksture. Analizu autori provode na matematičkom i statističkom modelu algoritma umjesto na stvarnim podacima te stoga na modelu potvrđuju svoju tezu.

Deskriptor $LBP_{P,R}^{riu^2} / VAR_{P,R}$, uz praćenje samo uniformnih oblika, što ga čini rotacijski invarijantnim, uzima u obzir i varijancu u okolini svake točke, što dodatno opisuje karakteristiku teksture. Time deskriptor koji koristi oba podatka postaje dvodimenzionalni histogram. Varijanca u okolini neke točke se računa na sljedeći način:

$$VAR_{P,R} = \frac{1}{P} \cdot \sum_{p=0}^{P-1} \left(g_p - \mu \right)^2$$
(3.44)

gdje je

$$\mu = \frac{1}{P} \cdot \sum_{p=0}^{P-1} g_p \tag{3.45}$$

srednja vrijednost svjetline u okolini razmatrane točke.

U novijem istraživanju [64] je LBP obitelj deskriptora pokazana kao dobar izbor za klasifikaciju tekstura pogotovo $LBP_{P,R}^{riu2}$. U slučajevima kada nema promjene osvjetljenja ili bilo kakvog drugog svjetlosnog efekta točnost klasifikacije može biti i iznad 80%. U slučaju uvođenja promjena u osvjetljenje točnost uglavnom pada i ispod 40%. Porodica LBP deskriptora teksture je implementirana i korištena za istraživanje mogućnosti primjene deskriptora vizualnih obilježja ravninskih segmenata u sklopu ove disertacije.

3.2.2. Gabor deskriptori teksture

Korištenje Gaborovih filtera za analizu tekstura je počelo još 80-ih godina prošlog stoljeća kada su znanstvenici uvidjeli da 2D gaborove funkcije strukturalno liče teksturnim primitivima odnosno tekselima.

2D Gaborovi filtri se mogu formulirati na slijedeći način:

$$g(x,y) = \frac{1}{2\pi\sigma_x\sigma_y} \cdot \exp\left[-\frac{1}{2} \cdot \left(\frac{x^2}{\sigma_x^2} + \frac{y^2}{\sigma_y^2}\right)\right] \cdot \exp\left(2\pi j u_0 x\right)$$
(3.46)

gdje σ_x i σ_y definiraju Gaussovu envelopu u x i y smjeru, u_0 je radijalna frekvencija, a $j = \sqrt{-1}$. Fourierova transformacija Gaborovog filtra je tada:

$$G(u,v) = \exp\left[-\frac{1}{2} \cdot \left(\frac{\left(u-u_0\right)^2}{\sigma_u^2} + \frac{v^2}{\sigma_v^2}\right)\right]$$
(3.47)

pri čemu su,

$$\sigma_u = \frac{1}{2\pi\sigma_x} \tag{3.48}$$

$$\sigma_{v} = \frac{1}{2\pi\sigma_{v}} \tag{3.49}$$

Na slici 3.20. može se vidjeti položaj Gaborovih filtera različitih skala i orijentacija u prostorno-frekvencijskoj ravnini.



Slika 3.20. Položaj, skala i orijentacija različitih Gaborovih filtera u prostorno-frekvencijskoj ravnini [65]

Kasnije su se Gaborovi filteri koristili za segmentaciju tekstura [66] gdje autori kao inspiraciju koriste teoriju višekanalnog filtriranja kod ljudskog vizualnog sustava. Razlog toga je otkriće da 2D Gaborovi filteri odgovaraju recepcijskom polju jednostavnih stanica u primarnom vizualnom korteksu ljudskog mozga. U članku [66] autori koriste banku od 20 Gaborovih filtera, 5 skala i 4 orijentacije, da bi dobili karakteristične odzive pojedinih tipova tekstura koje potom nastoje odvojiti u algoritmu za segmentiranje slike. Još jedan primjer korištenja Gaborovih filtera za raspoznavanje tekstura dan je u [67], gdje autori također koriste banku Gaborovih filtera od 4 skale i 6 orijentacija. Pronalaženjem odziva za sve raspoložive filtre nad nekom slikom teksture, autori proračunavaju vektor značajki, koji sadrži srednje vrijednosti i standardne devijacije koeficijenata transformacije za svaki filter. Također, proračunavanjem udaljenosti između dva takva vektora značajki, autori utvrđuju mjeru sličnosti dvije teksture. Industrijsku primjenu vidimo u [65], gdje autori ispituju mogućnosti korištenja Gaborovih filtera za analizu tekstura, sa specifičnom primjenom na detekciju defekata tkanja u tekstilnoj industriji. Detektiranje defekata se kod autora bazira na pronalaženju onog Gaborovog filtra iz banke filtera koji će imati najveći odziv za pojedinu klasu defekata. Tada taj Gaborov filter predstavlja tu klasu defekata. Autori također pružaju

mogućnost detekcije "ne naučenih" klasa defekata provođenjem multikanalne i multirezolucijske analize tkanja tijekom rada. U [64] Gaborovi filteri su pokazali dobre performanse raspoznavanja tekstura, koje su nekada iznosile i skoro 80%, ali je isto tako uočena velika osjetljivost na prostorne transformacije.

3.2.3. Tekston deskriptori tekstura

Tekstoni po definiciji predstavljaju određene mikrostrukture prisutne u slikama prirode i smatraju ih atomima ljudske podsvjesne percepcije. Raspoznavanje tekstura pomoću tekstona se bazira na tome da se iz nekog relativnog broja uzoraka za učenje izvuku određene skupine tekstona koji najbliže opisuju klase tekstura koje su predane za učenje. Potom se izračunavaju histogrami modela da se utvrdi distribucija točaka koje podupiru pojedine tekstone. Provođenjem iste analize nad slikama za testiranje utvrđuju se njihovi histogrami pomoću naučenih skupina tekstona. Potom se uspoređuje histogram testne slike s histogramima modela i utvrđuje kojoj klasi tekstura je ispitivana tekstura najsličnija.

Za potrebe dobivanja baze tekstona iz slika za učenje se pronalaze odzivi određene baze 2D filtera. Ta baza 2D filtera treba biti što bolje koncipirana radi pokrivanja što veće mogućnosti postojanja prirodnih obrazaca na slici, a da pri tome bude što je manje osjetljiva na prostorne transformacije. Na slici 3.21. vidi se dvije tako koncipirane baze.



Slika 3.21. Prva dva retka je banka filtera s 13 izotropičnih filtera, druga dva retka banka filtera s 4 različita filtera nazvana MR4 [68]

Kao što se vidi na slici 3.21., prva banka filtera koristi 13 izotropičnih i rotacijski invarijantnih filtera, dok druga koristi samo 4 filtera, dva rotacijski invarijantna i dva usmjerena pri različitim orijentacijama. Proračunom odziva za sve filtere u banci dobiva se za svaku sliku za treniranje i svaku njenu točku određeni vektor koeficijenata s onoliko elemenata vektora koliko je bilo filtera u banci. Klasteriranjem tih vektora pomoću npr. algoritma k-srednjih vrijednosti dobiva se određeni broj klastera, centri tih klastera

predstavljaju tekstone. Na slici 3.22. mogu se vidjeti 100 generiranih tekstona za 20 tipova tekstura pomoću prikazane dvije banke filtara na slici 3.21.



Slika 3.22. Primjer dvije baze tekstona [68]

Primjer takvog pristupa generiranja skupa tekstona vidimo u [69], gdje autori nastoje raspoznavati teksture s trodimenzionalnom komponentom u različitim uvjetima uvodeći pojam "3D texton". Autori koriste banku od 48 filtera i generiraju njihove odzive za svaku točku slike neke teksture koji su potom opisani vektorom značajki odziva filtera. Klasteriranjem tih vektora značajki pomoću algoritma k-srednjih vrijednosti, autori su stvarali tekstone. Da bi tekstoni što bolje opisali neku teksturu, potrebno ju je snimiti u što više uvjeta te potom ulančavanjem njihovih odziva na banku filtera dobivaju se puno robusniji vektori značajki pojedine točke.



Slika 3.23. Način generiranja baze tekstona za 3D teksture [69]

Autorima je za točno raspoznavanje teksture potrebno više slika iste teksture snimljene iz raznih kutova. Dok za raspoznavanje teksture iz jedne slike isti autori koriste Monte Carlo metodu s Markovljevim lancima. Nastavak njihovog rada vidimo u [68] gdje uvode novi način raspoznavanja teksture pomoću tekstona, koristeći manju banku dobro definiranih filtara za njihovo generiranje, nazvanu MR8. Zbog algoritma koji uzima samo najveće odzive filtara postiže se rotacijska invarijantnost. Korištenjem algoritma k-srednjih vrijednosti za klasteriranje i definiranje tekstona, te testiranjem nad standardnim bazama tekstura autori postižu točnost iznad 96%.

Također postoji i drugi način generiranja tekstona objašnjen u [70] gdje su tekstoni kombinacija određenog broja baznih funkcija kao što su Gabor i Gaussian, pri različitim skalama i orijentacijama. Time je omogućeno stvaranje kompleksnijih oblika koje karakteriziraju slike tekstura kao što su npr. zvijezda, pahuljica i sl.



Slika 3.24. Primjer stvaranja kompleksnijih tekstona [70]

3.2.4. Primjene deskriptora teksture

Analiziranje stvarnih tekstura koje nalazimo u prirodi i ljudskim okruženjima je izuzetno teško budući da postoji mnogo utjecaja koji mogu otežati detekciju pojedinih tekstura. Stoga u literaturi postoji veliki broj studija koje nastoje analizirati ponašanje raznih deskriptora za različite teksture u različitim, ali i kontroliranim uvjetima. Korištene teksture se nalaze u standardnim bazama kao što su VisTex, Outex, KTH-TIPS, UIUC i ostale.

Kroz istraživanje predstavljeno u [71] pokušalo se utvrditi da li je bolje slike tekstura klasificirati zasebno po boji ili teksturi ili združeno. Autori koriste VisTex i Outex baze za svoje ispitivanje, pri čemu ispituju veliki broj prostora boja; RGB, rgb kromatski prostor, CIE L*a*b*, HSV i I₁I₂I₃. Kao deskriptore boja koriste 2D ili 3D histograme s uglavnom 32 ili 16 binova po dimenziji prostora boja. Za deskriptore tekstura autori koriste LBP porodicu deskriptora i Gaborove filtre. Za raspoznavanje koriste klasifikator k-najbližih-susjeda. Autori na tri načina kombiniraju rezultate klasifikacije, na razini značajke, mjere sličnosti i razini klasifikatora. Zaključuju da zasebno klasificiranje ili preko boje ili preko deskriptora teksture uvijek daje bolje rezultate nego njihovo združeno djelovanje.

U [72] nastoje utvrditi koliki je utjecaj stvarnih okolnih uvjeta odnosno skale, osvjetljenja i položaja na raspoznavanje tipa materijala. Za tu svrhu autori stvaraju novu bazu za ispitivanje nazvanu KTH-TIPS koja se sastoji od 11 materijala snimljenih pod različitim skalama i osvjetljenjima. Za svoje ispitivanje koriste SVM (engl. Support Vector Machine) klasifikator i MR8 tekstone te tvrde da im točnost klasifikacije prelazi 85%. U nastavku tog istraživanja [73] autori nastoje raspoznavati materijale uz veliko generaliziranje odnosno nastoje klasificirati klase materijala umjesto individualnih materijala. Autori u ovom članku predstavljaju svoju novu bazu nazvanu KTH-TIPS2 koja uključuje 11 tipova materijala snimljenih pod različitim uvjetima, položajima i skalama. Za testiranje koriste LBP deskriptor u više skala te MR8 tekstone, dok kao klasifikator koriste k-najbližih-susjeda i SVM. Slično istraživanje autori provode i kasnije u [74].

Najdetaljniju analizu ponašanja deskriptora teksture u stvarnim uvjetima provode u [64] gdje se daje analiza 9 trenutno najboljih deskriptora za klasifikaciju tekstura u realnim uvjetima koji uključuju promjenu skale, položaja i osvjetljenja. Za svoje ispitivanje autori koriste tri baze: Outex, KTH-TIPS i UIUC. Testiraju 9 deskriptora kao što su LBP porodica deskriptora, MR8 i LM tekstoni, Gaborovi filtri, različiti histogramski deskriptori itd. Koriste algoritam k-najbližih-susjeda kao klasifikator. Općenito gledajući autori zaključuju da su MR8 prosječno najtočniji, dok su samo 50% ispitanih deskriptora imali prosječnu točnost klasifikacije iznad 60%.

3.3. Robotska lokalizacija zasnovana na deskriptorima vizualnih obilježja

Deskriptori vizualnih obilježja se prvenstveno koriste u sustavima lokalizacije mobilnog robota zasnovanih na pojavnosti. Lokalizacija zasnovana na slici se uglavnom temelji na uspoređivanju cijele slike s nekom slikom pretpostavljene lokacije i određivanju sličnosti. Ako je ta sličnost velika, tada se pretpostavlja da je velika i vjerojatnost da se robot nalazi blizu lokacije opisane slikom modela. Slikom 3.25. može se pojasniti ovakva lokalizacija.



Slika 3.25. Princip rada lokalizacije zasnovane na pojavnosti

Sustavi lokalizacije zasnovane na pojavnosti koriste uglavnom isključivo topološki prikaz karata, gdje najčešće svaki čvor predstavlja jednu sobu ili hodnik na karti te se stoga takve karte često nazivaju semantičkim. Na slici 3.25. može se vidjeti da su prostorije u razmatranoj radnoj okolini robota predstavljene slikama, tako da karta okoline robota zapravo predstavlja bazu slika. Zadatak sustava za lokalizaciju je pronaći onu sliku ili više njih iz navedene baze koje po nekim karakteristikama imaju najveću sličnost sa trenutno snimljenom slikom. Budući da se ovdje lokacije u prostoru raspoznaju pomoću cijele slike, uglavnom se koriste globalni deskriptori kao što su to histogrami boja ili nekih drugih karakteristika. Relativno rijetko se u literaturi za ovaj tip lokalizacije koriste lokalne ili točkaste značajke.

Lokalizacija zasnovana na pojavnosti se počela ozbiljno razmatrati pionirskim pokušajem [75], gdje autori nastoje topološki lokalizirati mobilnu platformu s ugrađenom svesmjernom kamerom. Karta po kojoj se robot kreće se sastoji od soba povezanih u graf, te je svaka soba opisana nizom slika u boji. Autori kao deskriptore slike koriste jednodimenzionalne histograme zasebne za svaki kanal korištenog prostora boja. Autori u članku ispituju više prostora boja: HSL, RGB i kromatski rgb prostor boja. Sustav topološke lokalizacije kod autora radi na način da uspoređuje histogram boje trenutne slike s onima koji odgovaraju onoj lokaciji na kojoj se pretpostavlja da se mobilna platforma nalazi, te s histogramima susjednih soba. Kao klasifikator autori koriste algoritam k-najbližih-susjeda, a kao mjeru sličnosti koriste Jeffrey udaljenost. Autori za lokalizaciju ne koriste odometriju budući da se radi o ručno pokretanim kolicima. Na slici 3.26a može se vidjeti jedna od fotografija koja opisuje prostoriju, a na slici 3.26b je prikazana slika korištene mobilne

platforme. Autori u članku tvrde da je postignuta točnost lokalizacije mobilne platforme od 98%.



Slika 3.26. Pionirski pokušaji lokalizacije temeljene na slici [75]

Topološka lokalizacija mobilnog robota korištenjem svesmjerne slike u boji ispituje se i u [76]. Kao dodatne informacije pored robotskog vida autori također koriste odometriju robota i diferencijalni GPS da bi ograničili pretraživanje prethodno snimljene topološke karte. Testiranje sustava lokalizacije se provodi u zatvorenim i vanjskim prostorima. Korišten je RGB prostor boja za opis zatvorenih prostora, a kromatski rgb prostor boja za vanjske prostore. Svaka lokacija u karti je opisana nizom slika u boji nad kojima se formira jednodimenzionalni histogram s 256 binova za svaku dimenziju prostora boja zasebno. Kao mjeru sličnosti dvije slike odnosno dva histograma autori su koristili L₁ normu. Autori u članku tvrde da je postignuta točnost topološke lokalizacije njihovog sustava u vanjskim prostorima od 65%, a u zatvorenim prostorima od 80%.

U članku [77] autori nastoje također topološki lokalizirati mobilnog robota, ali oni koriste višedimenzionalne histograme kao globalne deskriptore cijele slike i monokularni sustav kamere. U svojem istraživanju autori pored podataka o boji koriste različite kombinacije vizualnih podataka koje se mogu dobiti na razini jedne točke slike koje potom koriste za formiranje višedimenzionalnog histograma kao što su teksturiranost, rang i gustoća rubova. Kao prostor boja koriste kromatski rgb prostor, dok kao mjeru sličnosti koriste Jeffrey udaljenost. Autori su u članku tvrdili da su postigli točnost raspoznavanja lokacije od 83%.

Autori u [78] samo topološki lokaliziraju robota, budući da im je svrha osnovna navigacija i obavljanje specifičnih zadataka, koji ne zahtijevaju točnu lokaciju robota. Uzimaju u obzir semantički kontekst prostora gdje se robot kreće te karta predstavlja samo sobe povezane u graf. Svaku sobu autori opisuju nizom slika odnosno deskriptorom cijele scene koji je osmišljen tako da bude robustan na promjenu osvjetljenja, tip optičke kamere i neke manje fizičke promjene u sceni. Globalni deskriptor neke scene su autori zamislili kao višedimenzionalne (6-16 dimenzija) histograme koji sadrže vrijednosti različitih tipova gradijenata slike te kromatskih komponenata prostora suprotstavljenih boja. Kao klasifikator svake pojedine sobe autori koriste SVM (engl. Support Vector Machine), koji je naučen raspoznavati pojedine sobe pomoću prije spomenutih histograma. Pokuse autori provode koristeći 3264 slike označene po tome u kojoj su sobi slikane i uključuju različita doba dana/godine, različite kamere te različite robotske platforme. Korištene mobilne robotske platforme posjeduju monokularne sustave. Autori navode da su postigli točnost raspoznavanja sobe od 50% do 96% ovisno o tome koje su skupine slika korištene za učenje sustava, a koje za testiranje. Na slici 3.27. može se vidjeti ista lokacija u karti snimljena različitim kamerama i u različitim dobima dana.



Slika 3.27. Okviri za testiranje koriste različite kamere i uvjete za ispitivanje sustava lokalizacije temeljene na slici [78]

Tehnologije strojnog učenja koriste se u [79] za provođenje topološke lokalizacije monokularnog mobilnog robota koristeći SOM (engl. Self-Organizing Map) neuronske mreže kao deskriptore pojedine lokacije u karti. SOM neuronska mrežu u ovom slučaju predstavlja aproksimaciju 3D histograma u RGB prostoru boja. Prilikom faze učenja neuronske mreže potrebno joj je predati niz slika odnosno njihove histograme koji što bolje opisuju neku lokaciju. Autori u ovom članku tvrde da su postigli točnost topološke lokalizacije mobilnog robota od 86%.

U [80] koriste djelomično osmotrive Markovljeve procese odlučivanja (engl. Partialy Observable Markov Decision Process) kao evaluatore položaja mobilnog robota. Za lokalizaciju koriste samo robotski vid sa svesmjernom kamerom, dok se služe matričnom kartom (engl. occupancy grid map). Autori koriste HSL prostor boja, te uzimaju u obzir samo ograničenu skupinu boja koje prelaze neki predefinirani prag. Nad točkama slike koje zadovoljavaju tu ograničenu skupinu boja provodi se težinski obilazak za generiranje vektora značajki koje potom opisuju slike. Autori navode pogrešku njihovog sustava za lokalizaciju mobilnog robota od 0.5m.

Slično kao i u [79], autori u [81] koriste strojno učenje za topološku lokalizaciju mobilnog robota, pri čemu koriste neuronske mreže u SOM obliku kao deskriptore pojedine lokacije. Autori klasteriraju histograme RGB i HSI prostora boja na predefinirani broj klasa (16, 24, 32 i 64) kojima potom uče pojedine SOM-ove koji će predstavljati moguće lokacije u prostoru. Autori koriste svesmjernu kameru za snimanje okoline. Naglašavaju da prilikom testiranja sustava nisu uzimali u obzir promjene osvjetljenja okoline. Na slici 3.28. može se vidjeti primjer dobivenih centara klastera u RGB i HSI prostoru boja.



Slika 3.28. Višedimenzionalni histogrami boja se klasteriraju ne neki predefinirani broj [81]

Topološka lokalizacija mobilnog robota korištenjem naučene neuronske mreže kao deskriptora svake lokacije ispituje se i u [82]. Autori koriste monokularni sustav na mobilnom robotu, te ispituju sustav u zatvorenim i vanjskim prostorima. Kao značajke slike koje predaju neuronskim mrežama u fazi učenja koriste podatke o boji i detektiranim rubovima na cijeloj slici. Značajke boje autori uzimaju kao klasterirani CIE XYZ prostor boja podijeljen na 11 klastera, dok su podaci o rubovima dobiveni Robertovim operatorom. Ukupno 35 značajki ulazi u četveroslojnu neuronsku mrežu naučenu algoritmom s povratnom propagacijom pogreške. Autori u članku navode da su postigli točnost topološke lokalizacije od skoro 97% za raspoznavanje pripadnosti ispitnih slika nekoj naučenoj lokaciji.

Sustav topološke lokalizacije robota koji uzima u obzir i promjenu okruženja tijekom vremena je opisan u [83], gdje autori koriste inkrementalni SVM kao deskriptor pojedine lokacije. Inkrementalni SVM omogućava vlastito prilagođavanje novim informacijama tijekom svog rada, te time omogućava promjenu naučenih informacija o izgledu neke lokacije u slučaju da je došlo do promjene njezinog izgleda. Informacije koje sustav uči o izgledu scene uključuju derivaciju Gaussovog filtera provedenu nad slikom scene. Za testiranje sustava opisanog u članku autori koriste novostvorenu IDOL2 bazu.

U [84] opisan je sustav topološke lokalizacije mobilnog robota koji koristi klasterirane histograme kao globalne deskriptore pojedinih prostora u topološkoj karti. Autori koriste mobilnog robota sa svesmjernom kamerom te RGB prostor boja za definiranje deskriptora neke scene. Deskriptori koje autori povezuju s pojedinim lokacijama u topološkoj karti su realizirani kao klasteri 3D RGB histograma, dobiveni algoritmom k-srednjih vrijednosti. Autori formiraju deskriptore lokacija na način da generiraju 3D histograme svih slika u bazi i potom ih klasteriraju u onoliko klastera koliko imamo prostorija u topološkoj karti. Autori su formiranje deskriptora smjestili na FPGA čip ugrađen u kameri i na neki način distribuirali rad svog algoritma. Sustav se temelji na Monte Carlo lokalizaciji te autori u članku navode da su postigli prosječnu pogrešku lokalizacije od 0.61m, pri čemu za evaluaciju lokalizacije autori uspoređuju svoj sustav sa SLAM sustavom realiziranim LRF kamerom.

Okvir za testiranje algoritama za lokalizaciju mobilnog robota temeljenih na raspoznavanju slika predstavljaju autori u [85]. Oni predlažu primjenu svog razvijenog okvira za testiranje isključivo za topološku lokalizaciju robota. Testni okvir sadrži dvije baze slika nazvane INDECS i IDOL čiji su pojedini uzorci raspodijeljeni tijekom dužeg vremenskog perioda i 5 soba, te uključuju različite vremenske uvjete (sunčano, oblačno, noć) te različite druge utjecaje, kao što su promjena osvjetljenja i prisutnost ljudi. Slike u bazama su označene prema tome gdje se nalazila kamera koja ih je slikala, pri čemu su korištena 3 monokularna izvora za prikupljanje podataka u prostorima, 2 robota i jedna zasebna kamera. IDOL baza također uključuje odometriju robota, te LRF snimak prostora. Za testiranje realiziranih testnih okvira autori su kao deskriptore koristili SIFT kao lokalni deskriptore i vlastiti CRFH (engl. Compose Receptive Field Histogram) deskriptor koji u biti predstavlja višedimenzionalni histogram kojemu dimenzije predstavljaju informacije kao što su prva i druga derivacija Gaussovog filtera, magnituda gradijenta intenziteta slike i odzivi Laplacijana i Hessijana nad intenzitetom i komponentama boje slike. Kao klasifikator autori su koristili Inearni SVM.

Topološka lokalizacija monokularnog mobilnog robota provodi se u [86] s posebnim naglaskom na primjenu unutar CRESIS projekta (engl. Centre for remote sensing of ice sheets). Svaka lokacija u topološkoj karti prikazana je nizom od 1000 fotografija koje su opisane histogramom boja, svojstvenim slikama i Hu momentima. Također, dobivene histograme i svojstvene slike autori nastoje reducirati PCA (engl. Principal Component Analysis) analizom. Navedeni zbir značajki autori potom modeliraju pomoću mješavine Gaussian modela da bi opisali pojedinu lokaciju. Testove autori provode pomoću dvije topološke karte s 22 i 26 lokacija. Autori navode da su postigli točnost prepoznavanja lokacije u razvijenom sustavu od 95%.

U novije vrijeme istraživači nastoje povezati izgled scene s njenom geometrijom te na takav način bolje modelirati okolinu u kojoj se robot kreće, ali pri tome i dalje zadržati relativnu jednostavnost raspoznavanja lokacije koje nudi lokalizacija zasnovana na pojavnosti. Jedan od takvih novijih sustava je prikazan u [87] gdje autori lokaliziraju mobilni robot koristeći monokularnu viziju i matrične karte koje imaju razdiobu 3D prostora na obojane voksele. Kartu se stvara kalibriranim sustavom kamera, koji se sastoji od RGB i LRF kamere, čime je omogućeno stvaranje obojanog oblaka točaka u 3D prostoru. Točke se grupiraju u voksele dimenzija 5cm³ radi uštede prostora i olakšanog rada. Boja voksela je definirana kao srednja vrijednost boje točaka koje su obuhvaćene njegovom granicom. Lokalizaciju autori baziraju na čestičnim filtrima, te provode pokuse problema otetog robota i inkrementalne lokalizacije. Na početku lokalizacije sustav prati 1000 čestica uniformno raspoređenih po karti te za svaku česticu grafički podsustav baziran na OpenGL-u renderira pretpostavku trenutnog pogleda. Ta pretpostavka se uspoređuje sa trenutnom slikom kamere robota, uspoređujući točku po točku. Ako je zadovoljen određeni prag sličnosti točaka slike, te ako je zadovoljen predefinirani traženi postotak točaka slike, tada je stvarni položaj robota u blizini te pretpostavke odnosno čestice. Svaki kanal RGB prostora posebno se normalizira, čime se donekle umanjuje utjecaj promjene osvjetljenja. Autori navode da je postignuta prosječna pogreška razvijenog sustava lokalizacije od 0.25m do 0.59m u slučaju promjene osvjetljenja scene između modela i trenutne slike. Na slici 3.29. može se vidjeti primjer uspješne lokalizacije kada je renderirana slika približno jednaka stvarnoj.



Slika 3.29. Primjer sustava za lokalizaciju mobilnog robota zasnovanog na slici koji koristi geometrijsku kartu sačinjenu od voksela [87]

Drugi sustav koji u određenoj mjeri koristi 3D geometriju prostora uz podatke o boji predstavljen je u [88], gdje autori lokaliziraju mobilni robot sa svesmjernom kamerom koristeći Kullback-Leiber Monte Carlo lokalizaciju (engl. KLD-MCL). Autorov sustav lokalizacije koristi kartu prostora sa znatno pojednostavljenim prikazom okoline koji sadrži samo vertikalne plohe, odnosno zidove okomite u odnosu na pod po kojemu se robot kreće, s

teksturiranom površinom, koja je dobivena snimanjem prostora s robotom. Nedostatak razmatranog sustava je što robot u svakom trenutku mora vidjeti pod koji treba biti izrazito različite nijanse od zida da bi se mogao odvojiti od zida i prilikom stvaranja karte i prilikom lokalizacije robota. Na početku autorov sustav generira 65000 čestica raspoređenih po karti, pri čemu se renderiraju pretpostavke trenutnog pogleda pomoću prethodno snimljene karte prostora. Točnost svake pretpostavke provjerava se usporedbom trenutnog pogleda robota i renderirane pretpostavke. Usporedba se, kao i kod prethodnog sustava [87], vrši točka po točka te se utvrđuje postotak točaka slike koje su slične unutar neke tolerancije, ako je taj prag zadovoljen, pretpostavka je da je robot u blizini te čestice. Za točnost evaluacije čestice također se uzima u obzir izračunata udaljenost mobilnog robota od zida koja se uspoređuje s onom pretpostavljenom u čestici pomoću Mahalanobisove udaljenosti. Da bi smanjili utjecaj osvjetljenja, autori koriste kromatski rgb prostor boja. Autori provode eksperimente problema otetog robota, praćenja položaja i globalne lokalizacije. Autori navode da njihov sustav lokalizacije brže konvergira kod globalne lokalizacije od Monte Carlo sustava lokalizacije sa SIFT značajkama. Također navode da im se pogreška lokalizacije kreće između 0.6m i 1m u odnosu na stvarni položaj robota. Na slici 3.30. može se vidjeti primjer uspješne lokalizacije kada je pretpostavljena slika približno jednaka trenutnoj slici mobilnog robota.



Slika 3.30. Primjer sustava za lokalizaciju mobilnog robota temeljenog na slici koji koristi geometrijsku kartu stvorenu od vertikalnih površina [88]

4. Sustav za lokalizaciju mobilnog robota zasnovan na ravninskim segmentima

Osnovni sustav za lokalizaciju mobilnog robota koji se koristi u ovom radu je opisan u [3] i [4]. Sustav lokalizacije koristi Microsoft Kinect 3D kameru za prikupljanje podataka iz okoline. Tom sustavu je u sklopu istraživanja u ovom radu dodan novi sloj informacija u obliku vizualnih obilježja ravninskih segmenata radi postizanja bržeg i robusnijeg procesa lokalizacije mobilnog robota posebno u slučaju globalne lokalizacije. Da bi se olakšalo shvaćanje svrhe i načina korištenja vizualnih obilježja ravninskih segmenata u sustavu za lokalizaciju mobilnog robota zasnovanog na ravninskim segmentima objavljenog u [3], u ovom su poglavlju opisani svi dijelovi i faze rada navedenog sustava. Za razmatrani sustav lokalizacije koji koristi samo geometriju, bez primjene vizualnih obilježja ravninskih segmenata, u nastavku se koristi termin "osnovni sustav". Nadogradnja osnovnog sustava mehanizmima koji omogućuju primjenu vizualnih obilježja objašnjena je u 5. poglavlju.



Slika 4.1. Sustav lokalizacije mobilnog robota zasnovan na ravninskim segmentima

Sustav lokalizacije za svoj rad koristi hibridnu metričko-topološku kartu koja se sastoji od čvorova povezanih vezama, gdje svaki čvor predstavlja određenu lokaciju u prostoru opisanu lokalnim metričkim modelom. Uvjet da bi se neka lokacija našla u karti jest da se njezina slika nalazi pomaknuta za najmanje 0.5m ili da je zarotirana za najmanje 15° od susjednih čvorova. Na slici 4.1 može se vidjeti da sustav lokalizacije mobilnog robota ima 5 faza rada prije nego što može dati neku pretpostavku o položaju robota. Osnovna struktura sustava prikazana bijelim blokovima odgovara većini sustava za lokalizaciju zasnovanih na značajkama. Narandžastom bojom su naglašene specifičnosti razmatranog sustava. Prva faza

rada uključuje primanje trenutne slike okoline pomoću 3D senzora opisanog u poglavlju 2.3.4. 3D senzor predaje i dubinsku sliku i RGB sliku okoline, ali budući da je ovaj osnovni sustav za lokalizaciju mobilnog robota isključivo temeljen na geometriji, on koristi samo dubinsku sliku okoline. Kao značajke ovog razmatranog sustava za lokalizaciju se koriste ravninski segmenti dobiveni triangulacijom i segmentacijom dubinske slike 3D senzora. Korišteni mehanizmi triangulacije i segmentacije te način prikaza dobivenih ravninskih segmenata u sustavu su detaljnije objašnjeni u poglavljima 4.1., 4.2. i 4.3. Nakon pripremljenih podataka u obliku skupa ravninskih segmenata opisanih njihovim položajem, oblikom te nesigurnošću njihovog položaja, dolazi sam proces lokalizacije mobilnog robota. Problem lokalizacija se svodi na određivanje lokalnog 3D modela u karti okoline koji ima najsličniju geometrijsku strukturu geometrijskoj strukturi promatrane scene. To se izvodi tako da se skup ravninskih segmenata promatrane scene uspoređuje s ravninskim segmentima lokalnih modela koji sačinjavaju kartu. Kod inkrementalne lokalizacije usporedba se vrši sa relativno malim brojem modela u neposrednoj okolini trenutnog estimiranog položaja robota, dok se kod globalne lokalizacije usporedba vrši sa svim lokalnim modelima iz karte. U svakom slučaju, lokalizacija se izvodi registracijom dva skupa ravninskih segmenata. Registracija dva skupa ravninskih segmenata zahtijeva da se za segmente jednog skupa pronađu odgovarajući parovi među segmentima drugog skupa. Budući da se u osnovnom sustavu ne koriste nikakvi deskriptori za ravninske segmente, inicijalno sparivanje trenutno vidljivih ravninskih segmenata i onih iz prethodno snimljene karte se vrši na temelju geometrijskih ograničenja. Takvi inicijalno spareni segmenti ulaze u fazu generiranja hipoteza koja se temelji na izgradnji stablastog grafa u okviru kojeg se hipoteze ovisno o položaju mobilnog robota formiraju pomoću proširenog Kalmanovog filtera. S obzirom da među parovima segmenata na temelju kojih su se formirale hipoteze ima i neispravnih parova, među generiranim hipotezama bit će neispravnih. Na kraju se formirane hipoteze evaluiraju mjerama geometrijskog preklapanja ploha. Faze inicijalnog sparivanja ravninskih segmenata, generiranja i evaluacije hipoteza su detaljnije objašnjene u poglavljima 4.4. i 4.5. Na izlazu iz sustava dobiva se hipoteza relativnog položaja u odnosu na neki čvor prethodno snimljene hibridne metričko-topološke karte.

4.1. Triangulacija 2.5D slike

Triangulacija dubinske slike 3D senzora korištena u ovom sustavu je ostvarena Delaunay triangulacijom i algoritmom opisanim u [28]. Triangulacija je općenito proces segmentiranja nekog skupa točaka P na trokute čiji su vrhovi iz skupa P. Delaunay triangulacija skupa točaka P je takva da se ni jedna točka iz skupa P ne nalazi u opisanoj kružnici ni jednog trokuta koji je rezultat te segmentacije. Također, važno svojstvo Delaunay triangulacije je da je suma kutova nasuprot zajedničke stranice dva susjedna trokuta uvijek $\leq 180^{\circ}$.

Postupak triangulacije opisan u [28] je sljedeći. Na početku se pronalaze četiri kuta dubinske slike dobivene 3D senzorom te se pomoću njih formiraju dva trokuta s jednom od dijagonala slike kao zajedničkom stranicom. Nakon toga se kontinuirano generiraju novi trokuti na način da se za svaki postojeći trokut određuje udaljenost svake točke unutar granica tog trokuta od ravnine definirane s tri vrha tog trokuta pomoću vrijednosti dubine iz dubinske slike. Ako za svaku točku **m** unutar razmatranog trokuta vrijedi:

$$\begin{bmatrix} a_i & b_i \end{bmatrix} \cdot \mathbf{m} + c_i - d_{\mathbf{m}} \le \varepsilon_{Delaunay}$$
(4.1)

gdje su a_i , b_i i c_i parametri ravnine definirane trokutom, $d_{\mathbf{m}}$ dubina točke \mathbf{m} , a $\varepsilon_{Delaunay}$ je prag tolerancije koji definira maksimalnu udaljenost točke koja može pripadati nekom trokutu, tada se trokut više ne dijeli. Lijevi dio jednadžbe (4.1) predstavlja udaljenost točke \mathbf{m} od ravnine koju je definirao razmatrani trokut u smjeru dubine. U slučaju postojanja jedne ili više točaka u trokutu koje ne zadovoljavaju (4.1), trokut se dijeli u 3 trokuta pri čemu se kao vrh novih trokuta uzima ona točka koja ima najveću udaljenost od ravnine. Zadovoljavajući rezultati su postignuti za vrijednost praga tolerancije $\varepsilon_{Delaunay} = 4$. Na slici 4.2. se može vidjeti postupak triangulacije dubinske slike.



Slika 4.2. Postupak triangulacije: prvo se određuju početne točke kao uglovi, a zatim se analizira svaki trokut zasebno i pronalazi vrh novih trokuta kao najudaljenija točka u odnosu na ravninu koja prolazi vrhovima analiziranog trokuta
4.2. Segmentacija mreže trokuta na ravninske segmente

Kao što je u prethodnom poglavlju objašnjeno, originalna dubinska slika 3D senzora se segmentira na mrežu trokuta. Svaki tako dobiven trokut se može smatrati ravninskim segmentom unutar tolerancije koja je definirana jednadžbom (4.1). Međutim, tako segmentirana slika nije prikladna za raspoznavanje ravninskih segmenata, jer se sastoji od velikog broja trokuta koji imaju relativno male površine. Naime, mali trokuti sadrže nepreciznu informaciju o položaju robota, pa je potrebno uključiti relativno velik broj takvih trokuta u formiranje precizne hipoteze o položaju robota. Korištenje velikog broja trokuta za lokalizaciju mobilnog robota bi uzrokovalo veliku sporost rada sustava. Stoga se nad mrežom trokuta stvorenom prethodno objašnjenom triangulacijom provodi grupiranje u veće strukture koje predstavljaju dominantne ravninske segmente.

Implementirana hijerarhijska metoda grupiranja prethodno stvorenih trokuta je slična HFC metodi (engl. Hierarchical Face Clustering) opisanoj u [89]. Ova metoda primjenom pohlepnog (engl. greedy) algoritma rekurzivno grupira susjedne ravninske segmente na način da uklanja granice između njih. Konačni rezultat je hijerarhija segmenata koja se može prikazati binarnim stablom. Segmentiranje mreže trokuta započinje stvaranjem dualnog grafa nad cijelom mrežom trokuta. Svaki čvor takvog grafa na početku predstavlja jedan trokut a veza koja ga povezuje s nekim drugim čvorom predstavlja zajedničku stranicu između ta dva trokuta. Svakoj od tih veza se pridružuje određena vrijednost koja definira "jačinu" njihove granice. Počevši od najslabije granice prema najjačoj, od dva čvora povezana određenom granicom stvara se novi čvor na višoj hijerarhijskoj razini. Stari čvorovi tada postaju čvorovi djeca novostvorenom roditeljskom čvoru. Ovaj postupak se provodi sve dok se ne dobije samo jedan čvor koji tada u biti predstavlja cijelu 2.5D scenu. Kao rezultat dobiva se binarno stablo, čiji je korijen cijela 2.5D scena, a listovi trokuti dobiveni Delaunay triangulacijom. Primjer izgleda dualnog grafa i primjer spajanja dva čvora može se vidjeti na slici 4.3.



Slika 4.3. Primjer dualnog grafa i spajanja dva čvora [89]

Kao mjera jačine granice između dva segmenta korištena je standardna devijacija točaka ravninskog segmenta dobivenog spajanjem tih dvaju segmenata.

Iz dobivenog binarnog stabla se izdvajaju dominantne ravnine u 2.5D slici kao čvorovi stabla za koje vrijedi:

$$\begin{cases} \sigma_{roditelj} - \sigma_{cvor} \ge \sigma_1 \\ \sigma_{roditelj} > \sigma_2 \end{cases}$$
(4.2)

gdje je $\sigma_{roditelj}$ standardno odstupanje od idealne ravnine roditeljskog čvora, σ_{cvor} standardno odstupanje podređenog čvora, dok su σ_1 i σ_2 konstante koje definira korisnik. Zadovoljavajući rezultati su postignuti za vrijednost konstanti $\sigma_1 = 3$ i $\sigma_2 = 10$. Na slici 4.4. mogu se vidjeti sve faze obrade jedne scene dobivene 3D Kinect senzorom, od sirove dubinske slike do konačnih ravninskih segmenata.



Slika 4.4. a) Dubinska slika, b) trokuti dobiveni triangulacijom, c) granice između dominantnih ravninskih segmenata, d) konačni ravninski segmenti

4.3. Prikaz i registracija ravninskih segmenata

Svaki ravninski segment dobiven prethodno opisanim postupkom segmentacije u 2.5D prostoru odgovara ravninskom segmentu u 3D prostoru koji mora biti što jednoznačnije opisan da bi se mogla provesti registracija skupova 3D ravninskih segmenata. Ravninski segment je sačinjen od skupa približno koplanarnih 3D točaka te se njegov položaj i veličina mogu opisati centroidom $\hat{\mathbf{p}}$ i matricom kovarijance **Z**. Mjera raspršenja skupa tih točaka se može odrediti analizom glavnih komponenti (engl. Principal Component Analysis) pomoću kojih se dobiva matrica kovarijanci definirana na sljedeći način:

$$\mathbf{Z} = \frac{1}{k} \sum_{i=1}^{k} (\mathbf{p}_{i} - \hat{\mathbf{p}}) (\mathbf{p}_{i} - \hat{\mathbf{p}})^{\mathrm{T}}$$
(4.3)

gdje je $\hat{\mathbf{p}}$ srednja vrijednost točaka razmatrane ravnine odnosno centroid ravninskog segmenta, *k* je broj točaka u razmatranom skupu točaka, a \mathbf{p}_i je *i*-ta točka. Pronalaženjem svojstvenih vrijednosti:

$$\lambda_i, i = 1, 2, 3 \ (\lambda_1 \ge \lambda_2 \ge \lambda_3 \ge 0) \tag{4.4}$$

i svojstvenih vektora:

$$\boldsymbol{\Lambda}_{i}, i = 1, 2, 3 \left(\boldsymbol{\Lambda}_{1} \perp \boldsymbol{\Lambda}_{2} \perp \boldsymbol{\Lambda}_{3} \right)$$

$$(4.5)$$

matrice Z određujemo raspršenje odnosno varijancu točaka ravninskog segmenta u smjerovima svojstvenih vektora. Svojstveni vektor Λ_3 predstavlja normalu ravninskog segmenta, a odgovarajuća svojstvena vrijednost λ_3 tada predstavlja raspršenje odnosno varijancu u smjeru normale, a njezin korijen standardnu devijaciju.

Svakom ravninskom segmentu R_F pridružuje se koordinatni sustav S_F sa ishodištem u centroidu čije se osi podudaraju sa svojstvenim vektorima matrice **Z**. Stoga se u razmatranom sustavu za lokalizaciju mobilnog robota svaki ravninski segment R_F geometrijski opisuje u 3D prostoru pomoću rotacijske matrice ${}^{0}\mathbf{R}_F$, centroida \mathbf{t}_F , matrice kovarijance $\mathbf{C}_{\mathbf{q}}$ koja opisuje nesigurnost parametara ravninskog segmenta i matrice kovarijance \mathbf{C}_F . Rotacijska matrica ${}^{0}\mathbf{R}_F$ predstavlja orijentaciju koordinatnog sustava ravnine S_F u odnosu na koordinatni sustav kamere S_0 i definirana je na sljedeći način:

$${}^{0}\mathbf{R}_{F} = \begin{bmatrix} \mathbf{x}_{F} & \mathbf{y}_{F} & \mathbf{z}_{F} \end{bmatrix}$$
(4.6)

gdje su \mathbf{x}_{F} , \mathbf{y}_{F} i \mathbf{z}_{F} vektori koji predstavljaju x, y i z os koordinatnog sustava S_{F} .

Matrica kovarijance C_F opisuje raspršenje 3D točaka na površini ravninskog segmenta R_F . Matrica kovarijance C_F se izračunava iz svojstvenih vrijednosti matrice kovarijance Z na slijedeći način:

$$\mathbf{C}_{F} = {}^{0}\mathbf{R}_{F} \cdot \operatorname{diag}\left(\left[\lambda_{1} \ \lambda_{2} \ 0\right]^{\mathrm{T}}\right) \cdot {}^{0}\mathbf{R}_{F}^{\mathrm{T}}$$

$$(4.7)$$

Matrica kovarijance C_F će se prvenstveno koristiti kod utvrđivanja uvjeta preklapanja ravninskih segmenata, što je objašnjeno u poglavlju 4.4.2.

Odnos koordinatnog sustava ravninskog segmenta R_F s obzirom na koordinatni sustav kamere može se vidjeti na slici 4.5.



Slika 4.5. Koordinatni sustav ravninskog segmenta R_F

Svaki ravninski segment je definiran normalom **n** i pomaku ravnine ρ od ishodišta koordinatnog sustava kamere te za neku točku ravnine **p** tada vrijedi:

$$\boldsymbol{\rho} = \mathbf{n}^{\mathrm{T}} \cdot \mathbf{p} \tag{4.8}$$

Prethodno definirani parametri ravninskog segmenta su izračunati iz podataka dobivenih iz mjerenja 3D senzora, stoga je potrebno uzeti u obzir i nesigurnost tog mjerenja. Nesigurnost položaja ravninskog segmenta R_F je određena vektorom odstupanja:

$$\mathbf{q} = \begin{bmatrix} \mathbf{s}^{\mathrm{T}} & r \end{bmatrix}^{\mathrm{T}} = \begin{bmatrix} s_1 & s_2 & r \end{bmatrix}^{\mathrm{T}}$$
(4.9)

koji predstavlja odstupanje parametara stvarnog ravninskog segmenta **n** i ρ od parametara $\hat{\mathbf{n}}$ i $\hat{\rho}$ koji se estimiraju iz mjerenja. Parametri s_1 i s_2 predstavljaju nesigurnost normale ravninskog segmenta u smjeru \mathbf{x}_F i \mathbf{y}_F , dok *r* predstavlja nesigurnost udaljenosti ρ . Stoga se parametri stvarnog ravninskog segmenta s obzirom na njihovu nesigurnost mogu prikazati na sljedeći način:

$$\mathbf{n} = \frac{\hat{\mathbf{n}} + s_1 \cdot \mathbf{x}_F + s_2 \cdot \mathbf{y}_F}{\sqrt{1 + s_1^2 + s_2^2}}$$
(4.10)

$$\rho = r + \mathbf{n}^{\mathrm{T}} \cdot \mathbf{t}_{F} \tag{4.11}$$

Odnosi parametara stvarnog i estimiranog ravninskog segmenta mogu se vidjeti na slici 4.6.



Slika 4.6. Parametri stvarnog i estimiranog ravninskog segmenta

Budući da ne možemo u svakom trenutku znati koje je točno odstupanje estimiranog ravninskog segmenta s obzirom na stvarni ravninski segment, vektor odstupanja **q** opisan izrazom (4.9) promatra se kao slučajna varijabla s Gaussovom razdiobom i matematičkim očekivanjem 0 te ju opisujemo matricom kovarijanci. Matrica kovarijance C_q opisuje nesigurnost parametara ravninskog segmenta koje, kao što je prethodno opisano, čine jedinična normala **n** i udaljenost ρ .

$$\mathbf{C}_{\mathbf{q}} = \operatorname{diag}\left(\left[\sigma_{s_{1}}^{2} \sigma_{s_{2}}^{2} \sigma_{r}^{2}\right]^{\mathrm{T}}\right)$$
(4.12)

Nesigurnost parametara ravninskog segmenta R_F ovisi o njegovoj veličini odnosno broju točaka koje ga sačinjavaju kao i o njihovom rasporedu u prostoru. Ako se radi o malom segmentu, tada će nesigurnost njegovih parametara biti velika i obrnuto. Nadalje, ako su npr. točke segmenta raspoređene u širokom rasponu u smjeru osi \mathbf{x}_F , a u vrlo malom rasponu u smjeru osi \mathbf{y}_F , tada će nesigurnost njegove normale biti mala u smjeru \mathbf{x}_F , a velika u smjeru \mathbf{y}_F . Isto tako, nesigurnost parametara ravninskog segmenta ovisi i o njegovoj udaljenosti od kamere. Naime, što je udaljenost ravninskog segmenta od kamere veća, to će i nesigurnost pozicija točaka koje ga sačinjavaju biti veća, a time će i nesigurnost parametara razmatranog segmenta biti također veća. Iz ovog razmatranja slijede izrazi za nesigurnost parametara ravninskog segmenta

$$\sigma_{s_1}^2 = \frac{\sigma_r^2}{\lambda_1 + \sigma_r^2} \tag{4.13}$$

$$\sigma_{s_2}^2 = \frac{\sigma_r^2}{\lambda_2 + \sigma_r^2} \tag{4.14}$$

gdje σ_r^2 predstavlja reprezentativnu varijancu točaka segmenta u smjeru njegove normale. U općem slučaju, pozicije točaka nekog ravninskog segmenta su određene s različitom nesigurnošću, ovisno o svojem položaju u odnosu na kameru. Svaka točka definirana je na dubinskoj slici 3D senzora te se njihove 2.5D koordinate (*u*,*v*,*d*) također promatraju kao slučajna varijabla s Gaussovom razdiobom i matematičkim očekivanjem 0. Uz pretpostavku izotropnog i homogenog šuma mjerenja u uvd-prostoru. Matrica kovarijance 2.5D točke je:

$$\mathbf{C}_{\mathbf{p}} = \operatorname{diag}\left(\left[\sigma_{u}^{2} \sigma_{v}^{2} \sigma_{d}^{2}\right]^{\mathrm{T}}\right)$$
(4.15)

uz napomenu da je za razmatrani sustav korišteno $\sigma_u = \sigma_v = \sigma_d = 1$. Nesigurnost 3D točke u Kartezijevom prostoru, odnosno njezina matrica kovarijanci je tada:

$$\mathbf{C}_{x} = {}^{x}\mathbf{J}_{p} \cdot \mathbf{C}_{p} \cdot {}^{x}\mathbf{J}_{p}^{\mathrm{T}}$$
(4.16)

Gdje je ${}^{x}\mathbf{J}_{p}$ Jakobijan funkcije koja transformira točku iz 2.5D prostora u 3D prostor, te ona ovisi o intrističnim parametrima 3D kamere. Ako se centroid ravninskog segmenta uzme kao reprezentativna točka, tada se σ_{r}^{2} može izračunati na sljedeći način:

$$\sigma_r^2 = \mathbf{n} \cdot \mathbf{C}_x \cdot \mathbf{n}^{\mathrm{T}} \tag{4.17}$$

gdje je C_x matrica kovarijance centroida ravninskog segmenta. Varijanca σ_r^2 zapravo predstavlja projekciju matrice kovarijance C_x u smjeru normale ravninskog segmenta. Izrazi (4.13) do (4.17) su detaljno objašnjeni u [4].

Da bi mogli provesti proces registracije para ili skupa ravninskih segmenata, moramo ih svesti na isti koordinatni sustav, pri čemu je to prikladno napraviti tako da se npr. jedan ravninski segment R_F svede na koordinatni sustav drugog ravninskog segmenta $R_{F'}$. Kao što je ranije objašnjeno, svaki ravninski segment R_F je opisan u svom koordinatnom sustavu S_F kojemu je ishodište u centroidu \mathbf{t}_F , a smjerovi koordinatnih osi dobiveni PCA analizom (4.3), pri čemu je z-os u smjeru normale. Stoga da bi se transformirao neki ravninski segment R_F iz vlastitog koordinatnog sustava S_F u koordinatni sustav drugog ravninskog segmenta $S_{F'}$ moramo ga prvo transformirati u koordinatni sustav kamere koja ga je snimila S_A , zatim u koordinatni sustav kamere tog drugog ravninskog segmenta S_B te na kraju u koordinatni sustav drugog ravninskog segmenta $S_{F'}$.

Parametri ravninskog segmenta R_F opisani u koordinatnom sustavu S_F su slijedeći:

$${}^{F}\mathbf{n} = \frac{\mathbf{z} + \begin{bmatrix} \mathbf{x} & \mathbf{y} \end{bmatrix} \cdot \mathbf{s}}{\left\| \mathbf{z} + \begin{bmatrix} \mathbf{x} & \mathbf{y} \end{bmatrix} \cdot \mathbf{s} \right\|}$$
(4.18)

$${}^{F}\rho = r \tag{4.19}$$

gdje su **x**, **y** i **z** jedinični vektori koordinatnog sustava S_F izraženi u tom istom koordinatnom sustavu, tj. **x** = $\begin{bmatrix} 1 & 0 & 0 \end{bmatrix}^T$, **y** = $\begin{bmatrix} 0 & 1 & 0 \end{bmatrix}^T$, **z** = $\begin{bmatrix} 0 & 0 & 1 \end{bmatrix}^T$, a **s**= $\begin{bmatrix} s_1 & s_2 \end{bmatrix}^T$ vektor nesigurnosti normale ravninskog segmenta.

Transformiranjem u koordinatni sustav S_A kamere koja ga je snimila dobivaju se sljedeći parametri:

$${}^{A}\mathbf{n} = {}^{A}\mathbf{R}_{F} \cdot {}^{F}\mathbf{n} \tag{4.20}$$

$${}^{A}\rho = {}^{F}\rho + {}^{A}\mathbf{n}^{T} \cdot {}^{A}\mathbf{t}_{F}$$

$$(4.21)$$

što u biti odgovara formulama (4.10) i (4.11). Transformacija između dva koordinatna sustava kamere S_A i S_B je moguća samo ako poznajemo njihov relativni položaj opisan kao $\mathbf{w} = \begin{bmatrix} \boldsymbol{\phi}^T & \mathbf{t}^T \end{bmatrix}^T$, gdje vektor $\boldsymbol{\phi} = \begin{bmatrix} \alpha & \beta & \theta \end{bmatrix}^T$ predstavlja orijentaciju, a vektor \mathbf{t}^T relativni pomak. Razlika u orijentaciji ta dva sustava se može opisati i matricom ${}^B\mathbf{R}_A = \mathbf{R}(\boldsymbol{\phi})$. Transformirani parametri prvog ravninskog segmenta u koordinatnom sustavu kamere drugog ravninskog segmenta S_B su:

$${}^{B}\mathbf{n} = \mathbf{R}(\mathbf{\phi}) \cdot {}^{A}\mathbf{n} \tag{4.22}$$

$${}^{B}\rho = {}^{A}\rho + {}^{B}\mathbf{n}^{\mathrm{T}} \cdot \mathbf{t}$$

$$(4.23)$$

Konačna transformacija u koordinatni sustav $S_{F'}$ ravninskog segmenta $R_{F'}$ je sljedeća:

$$F'\mathbf{n} = {}^{B}\mathbf{R}_{F'}^{\mathrm{T}} \cdot {}^{B}\mathbf{n}$$
(4.24)

$${}^{F'}\rho = {}^{B}\rho + {}^{F'}\mathbf{n}^{\mathrm{T}} \cdot {}^{F'}\mathbf{t}_{B}$$

$$(4.25)$$

68

Ako provodimo postupak registracije para ravninskih segmenata pretpostavljamo da oni predstavljaju isti ravninski segment na promatranoj sceni sniman iz različitih pogleda, odnosno da vrijedi $F'\mathbf{n} = F'\mathbf{n}'$ i $F'\rho = F'\rho'$. Tada se iz jednadžbi (4.18) – (4.25) dobivaju konačne jednadžbe:

$$\begin{bmatrix} {}^{B}\mathbf{x}_{F'}^{\mathrm{T}} \\ \hline {}^{B}\mathbf{y}_{F'}^{\mathrm{T}} \end{bmatrix} \cdot \mathbf{R}(\mathbf{\phi}) \cdot \frac{{}^{A}\mathbf{z}_{F} + \begin{bmatrix} {}^{A}\mathbf{x}_{F} & {}^{A}\mathbf{y}_{F} \end{bmatrix} \cdot \mathbf{s}}{\sqrt{1 + \mathbf{s}^{\mathrm{T}} \cdot \mathbf{s}}} - \frac{\mathbf{s}'}{\sqrt{1 + \mathbf{s}'^{\mathrm{T}} \cdot \mathbf{s}'}} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$
(4.26)

$$r + \left({}^{A}\mathbf{t}_{F}^{\mathrm{T}} - \left({}^{B}\mathbf{t}_{F'} - \mathbf{t}\right)^{\mathrm{T}} \cdot \mathbf{R}(\boldsymbol{\phi})\right) \cdot \frac{{}^{A}\mathbf{z}_{F} + \left[{}^{A}\mathbf{x}_{F} \right] \cdot \mathbf{y}_{F} \cdot \mathbf{y}_{F}}{\sqrt{1 + \mathbf{s}^{\mathrm{T}} \cdot \mathbf{s}}} - r' = 0$$
(4.27)

Jednadžbe (4.26) i (4.27) bi trebale biti zadovoljene za svaki ispravni par ravninskih segmenata. Registracija dva skupa ravninskih segmenata se može svesti na određivanje takvog relativnog položaja kamera \mathbf{w} za koji su jednadžbe (4.26) i (4.27) zadovoljene uz minimalna odstupanja \mathbf{q} odnosno \mathbf{q}' uzimajući u obzir njihove procijenjene nesigurnosti.

Za rješavanje tog problema koristi se prošireni Kalmanov filter kojemu se korekcijski korak izvršava dodavanjem sparenih parova ravninskih segmenata. Korekcijski korak proširenog Kalmanovog filtera temelji se na jednadžbama (4.26) i (4.27). Izrazi korišteni za rad proširenog Kalmanovog filtra se mogu vidjeti u prilogu 1.

4.4. Stvaranje skupa inicijalnih parova i generiranje hipoteza

Registracija dva skupa ravninskih segmenata primjenom proširenog Kalmanovog filtra, provodi se na temelju skupa parova ravninskih segmenata, gdje je prvi element para ravninski segment detektiran na trenutno promatranoj sceni, a drugi ravninski segment iz karte okoline. Kada bi se uzimali u obzir svi mogući parovi ravninskih segmenata, broj netočnih hipoteza bi bio ogroman, što bi sustav učinilo neefikasnim. Dakle, radi postizanja zadovoljavajuće učinkovitosti sustava, potrebno je formirati što manji skup parova segmenata, ali da je u njemu sadržano što više ispravnih parova. Budući da geometrijski definirani ravninski segmenti koji se koriste u osnovnom sustavu za lokalizaciju nemaju deskriptor, inicijalni se parovi generiraju samo na osnovi geometrijskih ograničenja. Ako imamo definiran neki relativni položaj \mathbf{w} s pripadajućom nesigurnosti opisanom matricom kovarijanci \mathbf{P} , dva ravninska segmenata R_F i $R_{F'}$ se mogu smatrati ispravno sparenim ako zadovoljavaju sljedeće uvjete:

- približno su koplanarni, odnosno parametri ravninskih segmenata (normale i udaljenosti ρ) su međusobno u skladu s estimiranim relativnim položajem w,
- dovoljno se preklapaju.

Mjere koplanarnosti i preklapanja se računaju pomoću Mahalanobisove udaljenosti u prostoru parametara ravninskog segmenta. U iduća dva podpoglavlja detaljnije će se opisati ova dva uvjeta i pripadajuće mjere.

4.4.1. Uvjet koplanarnosti

Uvjet koplanarnosti se provjerava na način da se parametri ravninskog segmenta R_F transformiraju u koordinatni sustav $S_{F'}$ definiran za ravninski segment $R_{F'}$ i provjeravaju uvjeti (4.26) i (4.27). Ako se s obzirom na zadani relativni položaj w ta dva ravninska segmenta preklapaju, prethodno spomenuta dva uvjeta bi trebala biti zadovoljena za odstupanja **q** koja se nalaze unutar očekivanih granica nesigurnosti uzimajući u obzir i nesigurnost estimiranog položaja w. Mjera koplanarnosti se određuje kao Mahalanobisova udaljenost pomoću matrice kovarijanci **Q**, uvjeta (4.26) i (4.27) i vektora **e** definiranih u prilogu 1. Par ravninskih segmenata se smatra ispravnim s obzirom na uvjet koplanarnosti za zadani relativni položaj w i pripadajuću nesigurnost **P** ako zadovoljavaju:

$$\mathbf{e}^{\mathrm{T}} \cdot \mathbf{Q}^{-1} \cdot \mathbf{e} \le \varepsilon_{\parallel} \tag{4.28}$$

Za rad ovog sustava lokalizacije mobilnog robota prag uvjeta koplanarnosti ε_{\parallel} je definiran pomoću χ^2 razdiobe s 3 stupnja slobode te je postavljena na vrijednost $\varepsilon_{\parallel} = 11.34$ čime se obuhvaća 99% mjernih rezultata u slučaju točno sparenih ravninskih segmenata.

4.4.2. Uvjet preklapanja

Kod ovog uvjeta se transformira centroid \mathbf{t}_F ravninskog segmenta R_F u koordinatni sustav $S_{F'}$ ravninskog segmenta $R_{F'}$, pri čemu se također koristi matrica kovarijance \mathbf{C}_F definirana izrazom (4.7), koja daje podatke o raspršenju točaka ravninskog segmenta. Mjera preklapanja se računa kao Mahalanobisova udaljenost između centroida dva ravninska segmenta transformirana u isti koordinatni sustav pomoću estimiranog relativnog položaja **w** uzimajući u obzir njegovu nesigurnost **P**, pri čemu se uzima da je nesigurnost položaja centroida određena matricom kovarijance \mathbf{C}_F . Ova mjera preklapanja dva ravninska segmenta se može predočiti na način da je svaki segment prikazan 3D elipsoidom sa središtem u centroidu i radijusima proporcionalnim korijenima svojstvenih vrijednosti matrica $C_{F'}$ koji, pak, odražavaju distribuciju točaka tog ravninskog segmenta u prostoru, te se promatra u kolikoj se mjeri dva elipsoida preklapaju. Dva ravninska segmenta smatraju se ispravno sparenim s obzirom na uvjet preklapanja ako zadovoljavaju:

$$\widehat{\mathbf{e}}^{\mathrm{T}} \cdot \widehat{\mathbf{Q}}^{-1} \cdot \widehat{\mathbf{e}} \le \varepsilon_{\approx} \tag{4.29}$$

gdje su:

$$\widehat{\mathbf{e}} = \mathbf{R}(\mathbf{\phi}) \cdot {}^{A}\mathbf{t}_{F} + \mathbf{t} - {}^{B}\mathbf{t}_{F'}$$
(4.30)

$$\widehat{\mathbf{Q}} = \mathbf{R}(\mathbf{\phi}) \cdot \mathbf{C}_F \cdot \mathbf{R}^{\mathrm{T}}(\mathbf{\phi}) + \mathbf{C}_{F'} + \widehat{\mathbf{C}} \cdot \mathbf{P} \cdot \widehat{\mathbf{C}}^{\mathrm{T}}$$
(4.31)

$$\widehat{\mathbf{C}} = \frac{\partial \widehat{\mathbf{e}}}{\partial \mathbf{w}} \bigg|_{\mathbf{w} = \widehat{\mathbf{w}}} = \left[{}^{\mathbf{p}} \mathbf{J}_{\phi} \left(\widehat{\mathbf{\phi}}, {}^{A} \mathbf{t}_{F} \right) \middle| \mathbf{I}^{3 \times 3} \right]$$
(4.32)

gdje je ${}^{\mathbf{p}}\mathbf{J}_{\phi}$ definiran u prilogu 1. Za definiranje praga uvjeta preklapanja ε_{z} je također korištena χ^{2} razdioba s 3 stupnja slobode te je i u ovom slučaju taj prag postavljen na $\varepsilon_{z} = 11.34$ čime se obuhvaća 99% mjernih rezultata u slučaju točno sparenih ravninskih segmenata.

4.4.3. Generiranje hipoteza

Svaki stvarni prostor gdje se mobilni robot kreće može potencijalno sadržavati mnoštvo ravninskih segmenata. Korištenje svih detektiranih ravninskih segmenata na sceni za lokalizaciju mobilnog robota uglavnom nije praktično, jer bi proračunavanje svih mogućih kombinacija sparivanja s ravninskim segmentima iz karte, čak i uz geometrijska ograničenja objašnjena u podpoglavljima 4.4.1. i 4.4.2., odnijelo previše računskog vremena. Također, da bi se utvrdio relativni položaj u šest stupnjeva slobode između dvije scene nisu ni potrebne sve kombinacije parova već samo tri ispravno sparena ravninska segmenta. Pri tome ta tri korištena ravninska segmenta na sceni moraju biti u takvom međusobnom položaju da za svaki smjer u prostoru normala nekog od ta tri segmenta nije okomita na taj smjer te moraju biti dovoljno veliki da se na temelju njih može postići dovoljna točnost estimacije. Ako je zahtjev za utvrđivanjem relativnog položaja ublažen na samo pet stupnjeva slobode, tada su potrebna samo dva para točno sparenih ravninskih segmenata. Tipične scene u zatvorenim prostorima uglavnom sadrže najmanje dvije dominantne ravnine, pod po kojem se mobilni robot kreće i neki od zidova prostorije u kojoj se robot nalazi. Međutim u mnogim scenama,

pogotovo u hodnicima, postoji mogućnost da nedostaje ravnina koja će omogućiti lokalizaciju robota u smjeru zadnjeg stupnja slobode. Na slici 4.7. mogu se vidjeti primjeri tipičnih scena u zatvorenim prostorima.



a) b) Slika 4.7. Primjer izgleda scena u zatvorenim prostorima [3]

Kao što se može vidjeti na slici 4.7a postoji dovoljno dominantnih ravninskih segmenata za određivanje svih šest stupnjeva slobode. Tu mogućnost prvenstveno osiguravaju ravninski segmenti poda, lijevog stupa i desnog zida. Na slici 4.7b može se vidjeti da postoji samo ravninski segment poda i paralelni segmenti lijevog i desnog zida te da fali ravninski segment okomit na smjer gibanja mobilnog robota duž hodnika koji bi omogućio lokalizaciju mobilnog robota u tom smjeru. Stoga, može se zaključiti da svi ravninski segmenti koji se mogu detektirati na nekoj sceni nemaju jednaku važnost za određivanje relativnog položaja mobilnog robota. Drugim riječima, potrebno je na neki način vrednovati ravninske segmente prema informacijama koje daju s obzirom na osiguravanje lokalizacije u šest stupnjeva slobode. U razmatranom sustavu se to vrednovanje temelji na koeficijentu sadržaja informacije definiranom u nastavku. Raspodjela raspoložive informacije o položaju robota po svim smjerovima sadržane u skupu od N ravninskih segmenata može se opisati sljedećom matricom:

$$\mathbf{Y} = \sum_{i=1}^{N} \mathbf{n}_{i} \cdot \mathbf{n}_{i}^{\mathrm{T}} \cdot w_{i}$$
(4.33)

gdje je *N* broj detektiranih ravninskih segmenata, \mathbf{n}_i je normala i-tog ravninskog segmenta, a w_i je broj točaka sadržanih u i-tom ravninskom segmentu. Tada je mjera doprinosa u ukupnoj informaciji pojedinog ravninskog segmenta u smjeru njegove normale **n** jednaka:

$$\omega = \frac{w}{\mathbf{n}^{\mathrm{T}} \cdot \mathbf{Y} \cdot \mathbf{n}} \tag{4.34}$$

Vrijednost ω ćemo nazvati *koeficijentom sadržaja informacije*. Stoga se u razmatranom sustavu detektirani ravninski segmenti sortiraju prema vrijednosti (4.34) i u obzir za lokalizaciju se uzima samo određeni broj na vrhu tog popisa koji ćemo označiti s Q_{seg} . Redni broj ravninskog segmenta na tom popisu koristi se kod generiranja hipoteza. Zadovoljavajući rezultati lokalizacije mobilnog robota u razmatranom sustavu su postignuti za vrijednost 20 ravninskh segmenata s vrha sortiranog popisa Q_{seg} .

Zbog nepostojanja deskriptora za geometrijske ravninske segmente, standardna RANSAC metoda za sparivanje ne bi bila efikasna zbog velikog broja krivo sparenih ravninskih segmenata. Primjer rješenja takvih problema je korištenje modifikacije nazvane GCRANSAC [34] (engl. Geometrically Constrained RANSAC). Osnovna ideja GCRANSAC-a je da se hipoteze generiraju na način da svaki par uključen u hipotezu sužava izbor za sljedeći par postrožavajući geometrijska ograničenja koje sljedeći par mora zadovoljiti. U razmatranom sustavu lokalizacije mobilnog robota se koristi izmijenjena metoda, gdje se i dalje uzimaju u obzir geometrijska ograničenja, ali se umjesto nasumično odabranog konzistentnog para ravninskih segmenata, oni odabiru sekvencijalno uzimajući u obzir koeficijent sadržaja informacije razmatranog para ravninskih segmenata. Dakle proces generiranja hipoteza o položaju robota radi na takav način da u hipotezu sekvencijalno uvodi parove koji su geometrijski konzistentni s njom dok nesigurnost estimirane orijentacije ne padne ispod neke prethodno definirane razine. Time je postupak generiranja hipoteze osmišljen da prvo generira hipoteze koje su vjerojatnije što znatno može ubrzati proces lokalizacije mobilnog robota.

Postupak generiranja hipoteze je podijeljen u dvije faze. Prva faza koristi strukturu sličnu stablu za određivanje pet stupnjeva slobode gibanja mobilnog robota (sve tri komponente orijentacije i dva smjera translacije), dok se u drugoj fazi određuje posljednji stupanj slobode gibanja odnosno preostali smjer translacije.

Ulazni podaci prve faze generiranja hipoteze uključuju početnu estimaciju relativnog položaja s pripadajućom nesigurnosti te sortirani popis Q_{sort} sparenih ravninskih segmenata (R_F, R_F) koji zadovoljavaju uvjete (4.28) i (4.29). Popis je sortiran rastućim redoslijedom s obzirom na združenu mjeru koeficijenta sadržaja informacije para ravninskih segmenata. Ta mjera se za neki par ravninskih segmenata izračunava tako da se zbroje njihovi redni brojevi u odgovarajućim popisima Q_{seg} . Kao što je ranije napomenuto, proces generiranja hipoteza

koristi stablastu strukturu, gdje svaki čvor stabla, osim korijena stabla, predstavlja par ravninskih segmenata iz popisa Q_{sort} . Put od nekog čvora stabla N_V do korijena stabla predstavlja jednu hipotezu estimacije relativnog položaja s pripadajućom nesigurnosti koja je dodijeljena čvoru N_V . U početnom stanju prve faze imamo stablo s korijenom koji predstavlja početnu estimaciju relativnog položaja s pripadajućom nesigurnosti. S tim početnim stanjem se paralelno razvijaju hipoteze tako da se obilazi sortirani popis Q_{sort} i skida par s vrha popisa te se svakom čvoru stabla N_V , dodaje novi čvor koji sadrži taj skinuti par ravninskih segmenata (R_F , R_F). Pri tome se mora naglasiti da se parovi ravninskih segmenata koji se skidaju sa sortiranog popisa povezuju samo s onim čvorovima stabla N_V za koje vrijedi:

- ni jedan član novog para ravninskih segmenta (*R_F*, *R_{F'}*) se ne nalazi kao član ni jednog drugog para ravninskih segmenata u čvorovima na putu od čvora *N_V* prema korijenu stabla,
- novi par ravninskih segmenata (R_F , R_F) zadovoljava uvjete (4.28) i (4.29) za relativni položaj s pripadajućom nesigurnosti koji je pridružen čvoru N_V .

Novododanom čvoru je pridružena estimacija relativnog položaja s pripadajućom nesigurnosti dobivena korekcijom estimiranog položaja pridruženog čvoru N_V na temelju informacije o relativnom položaju ravninskih segmenata R_F , $R_{F'}$. Korekcija položaja se izračunava primjenom proširenog Kalmanovog filtra opisanog u prilogu 1. Ako je procijenjena nesigurnost orijentacije novog dodanog čvora manja od neke unaprijed zadane, ta grana stabla se zaključava i odlazi u drugu fazu, dok se ostale grane nastavljaju proširivati parovima iz sortiranog popisa Q_{sort} . Postupak generiranja hipoteza se nastavlja do nekog unaprijed definiranog broja hipoteza.

Druga faza generiranja hipoteze se provodi samo za one hipoteze koje su završile prvu fazu. Svodi se na pronalazak onog para ravninskih segmenata (R_F , R_F) iz sortiranog popisa Q_{sort} koji će omogućiti estimiranje posljednjeg stupnja slobode odnosno preostalog smjera translacije koji ćemo nazvati deficitarnim smjerom i označiti vektorom Λ_d . Pri tome se formira skup T_d svih parova iz sortiranog popisa Q_{sort} koji sadrže ravninske segmente postavljene pod dovoljno velikim kutem u odnosu na deficitarni smjer Λ_d razmatrane hipoteze. Među parovima iz skupa T_d traži se onaj koji omogućuje najbolje poklapanje preostalih parova ravninskih segmenata iz skupa T_d . Druga faza generiranja hipoteze se izvršava samo za one hipoteze koje nisu slične već nekoj obrađenoj hipotezi, dok se sama sličnost dviju hipoteza izračunava na temelju sličnosti njihove orijentacije te translacije u smjeru okomitom na deficitarni smjer.

4.5. Evaluacija hipoteza

Postupak opisan u prethodnom poglavlju će rezultirati skupom hipoteza, gdje će za svaku hipotezu biti definiran indeks μ lokalnog 3D modela iz karte s kojim je provedena registracija, relativni položaj **w** u odnosu na taj model s pripadajućom nesigurnosti te popis geometrijski konzistentnih parova ravninskih segmenata $\Omega(\mu, \mathbf{w}) \subseteq Q_{sort}$ za koje vrijede uvjeti koplanarnosti i preklapanja opisani s (4.28) i (4.29) s obzirom na položaj **w**. Kvaliteta neke hipoteze se tada određuje prema:

$$\mathfrak{I}(\mu, \mathbf{w}) = \sum_{R_F \in F} \max_{(R_F, R_{F'}) \in \Omega(\mu, \mathbf{w})} \left\{ P\left(d\left(R_F, R_{F'}; \mathbf{w}\right)\right) \cdot \min\left\{|R_F|, |R_{F'}|\right\} \right\}$$
(4.35)

gdje |.| označava broj točaka ravninskog segmenta, *d* je Mahalanobisova udaljenost koja odgovara lijevoj strani izraza (4.28), a P(x) predstavlja vjerojatnost da za ispravni par ravninskih segmenata (R_F , R_F) Mahalanobisova udaljenost na lijevoj strani jednadžbe (4.28) bude veća ili jednaka vrijednosti *x* uz pretpostavku χ^2 razdiobe s tri stupnja slobode. Hipoteza koja ima najveću vrijednost dobivenu izrazom (4.35) smatra se najboljom estimacijom trenutnog stvarnog položaja mobilnog robota.

5. Primjena vizualnih obilježja u lokalizaciji mobilnog robota

Osnovni sustav opisan u 4. poglavlju uzima u obzir samo dubinsku sliku dobivenu Kinect senzorom te radi samo s geometrijskim podacima dobivenih iz točaka sadržanih u toj slici. Da bi sustav za lokalizaciju mobilnog robota uzimao u obzir i podatke o vizualnim obilježjima potrebno ga je izmijeniti. Na slici 5.1. mogu se vidjeti promjene u fazama rada sustava opisanog u poglavlju 4. koje su razvijene i implementirane u sklopu ove disertacije.



Slika 5.1. Prošireni sustav za lokalizaciju mobilnog robota

Dijelovi obojani crvenom bojom na slici 5.1. su dijelovi koji su prošireni radi omogućavanja korištenja podataka o vizualnim obilježjima. Da bi se smanjili utjecaji promjene osvjetljenja okoline i mjernog šuma na RGB kameru potrebno je ulaznu RGB sliku pred-procesirati čime se postiže značajno poboljšanje performansi deskriptora vizualnih obilježja. Kada se koriste deskriptori zasnovani na prostorima boja sa slika se šum uklanja pomoću bilateralnog filtriranja opisanog u poglavlju 5.1.1. Potom se još slika modificira nelinearnim širenjem histograma po svakom ulaznom kanalu RGB slike na način opisan u poglavlju 5.1.3. Kada se koristi deskriptori zasnovani na značajkama teksture, tada se RGB slika transformira u sliku intenziteta svjetline te joj se ujednači histogram radi naglašavanja kontrasta na način opisan u poglavlju 5.1.2. Nakon stvaranja ravninskih segmenata segmentacijom opisanom u poglavlju 4.2., svakom detektiranom ravninskom segmentu se dodjeljuje generirani deskriptor vizualnih obilježja. Primjere generiranih deskriptora vizualnih obilježja može se vidjeti u poglavlju 5.4. Budući da Kinect senzor nije originalno dizajniran za korištenje u mobilnoj robotici, moguće je postojanje veće pogreške prilikom povezivanja točka dubinske slike s njima odgovarajućim točkama na RGB slici. Stoga je potrebno učiniti

sustav robusnim na pogreške kalibracije tako da se ne uzimaju u obzir točke na rubovima segmenata, što rezultira smanjenim brojem pridruženih RGB točaka nekom ravninskom segmentu, ali i smanjenjem pogreške izračuna sličnosti deskriptora vizualnih obilježja. Ovaj postupak izuzimanja rubnih točaka ravninskih segmenata je opisan u poglavlju 5.2. Uvođenje dodatnih uvjeta prilikom pronalaženja inicijalne korespondencije radi uzimanja u obzir podataka o vizualnim obilježjima je objašnjeno u poglavlju 5.3. U istom poglavlju je objašnjena i izmjena u načinu evaluacije generiranih hipoteza radi njihovog boljeg vrednovanja, koje sada uključuje i podatke o vizualnim obilježjima.

5.1. Pred-procesiranje RGB slike

U ovom će se poglavlju objasniti korišteni načini pred-procesiranja slika radi uklanjanja šuma i smanjenja utjecaja promjene osvjetljenja u slučaju korištenja sva tri kanala ulazne RGB slike te načni naglašavanja kontrasta slike intenziteta svjetline radi boljeg razlikovanja tekstura.

5.1.1. Bilateralno filtriranje

Pod filtriranjem slike se u tradicionalnom smislu podrazumijeva postupak kojim se od neke slike stvara nova, filtrirana slika, pri čemu je vrijednost neke točke filtrirane slike dobivena kao funkcija nad vrijednostima neke lokalne okoline odgovarajuće točke ulazne slike. Kao primjer može se uzeti Gaussov nisko-propusni filtar koji težinski ujednačuje vrijednosti neke okoline. Kod njega se težine smanjuju udaljavanjem od centra razmatrane okoline. Time se šum uklanja zamućivanjem budući da se smatra da je slika originalno lokalno "glatka", a superponirani šum su diskontinuiteti koje treba "zagladiti". Nažalost se time i rubovi na slici koji se također mogu smatrati diskonituitetima zaglađuju odnosno zamućuju što uglavnom nije poželjni rezultat filtriranja.

Stoga je ideja bilateralnog filtriranja objašnjenog u [90] kombinirati tradicionalno filtriranje u prostoru s filtriranjem intervala intenziteta slike u svrhu povećanja efikasnosti uklanjanja šuma uz očuvanje rubova. Možemo promatrati nisko-propusni prostorni filter koji generira sljedeći odziv:

$$\mathbf{h}(\mathbf{x}) = \left(k_d(\mathbf{x})\right)^{-1} \cdot \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \mathbf{f}(\boldsymbol{\xi}) \cdot c(\boldsymbol{\xi} - \mathbf{x}) d\boldsymbol{\xi}$$
(5.1)

77

gdje je $c(\xi - \mathbf{x})$ mjera geometrijske bliskosti centralne točke \mathbf{x} razmatrane okoline i neke točke ξ te okoline, $\mathbf{f}(\xi)$ vrijednost intenziteta točke ξ , a k_d je normalizacijska konstanta. Da bi se održala istosmjerna komponenta signala, k_d je:

$$k_d\left(\mathbf{x}\right) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} c\left(\boldsymbol{\xi} - \mathbf{x}\right) d\boldsymbol{\xi}$$
(5.2)

Filter intervala intenziteta slike se može definirati na sličan način:

$$\mathbf{h}(\mathbf{x}) = \left(k_r(\mathbf{x})\right)^{-1} \cdot \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \mathbf{f}(\boldsymbol{\xi}) \cdot s(\mathbf{f}(\boldsymbol{\xi}) - \mathbf{f}(\mathbf{x})) d\boldsymbol{\xi}$$
(5.3)

gdje je $s(\mathbf{f}(\boldsymbol{\xi})-\mathbf{f}(\mathbf{x}))$ mjera bliskosti intenziteta slike centralne točke \mathbf{x} razmatrane okoline i neke točke $\boldsymbol{\xi}$ te okoline, a k_r normalizacijska konstanta. Funkcija *s* ima veću vrijednost za točke čiji su intenziteti jako blizu, a malu za po intenzitetu jako različite točke. Normalizacijska konstanta k_r je tada:

$$k_r(\mathbf{x}) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} s(\mathbf{f}(\boldsymbol{\xi}) - \mathbf{f}(\mathbf{x})) d\boldsymbol{\xi}$$
(5.4)

Ako se kombiniraju ova dva filtriranja po prostoru i intervalu intenziteta dobiva se bilateralno filtriranje koje je tada definirano:

$$\mathbf{h}(\mathbf{x}) = \left(k\left(\mathbf{x}\right)\right)^{-1} \cdot \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \mathbf{f}\left(\boldsymbol{\xi}\right) \cdot c\left(\boldsymbol{\xi} - \mathbf{x}\right) \cdot s\left(\mathbf{f}\left(\boldsymbol{\xi}\right) - \mathbf{f}\left(\mathbf{x}\right)\right) d\boldsymbol{\xi}$$
(5.5)

gdje je normalizacijska konstanta k tada:

$$k(\mathbf{x}) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} c(\boldsymbol{\xi} - \mathbf{x}) s(\mathbf{f}(\boldsymbol{\xi}) - \mathbf{f}(\mathbf{x})) d\boldsymbol{\xi}$$
(5.6)

Ovo filtriranje mijenja vrijednost intenziteta u točci \mathbf{x} u ovisnosti o udaljenosti točaka i sličnosti s njihovim intenzitetima u okolini. Princip rada bilateralnog filtriranja može se objasniti slikom 5.2. koja prikazuje okolinu točke na granici između svijetlog i tamnog intenziteta.



Slika 5.2. a) Originalni signal, b) funkcija sličnosti, c) filtrirani signal [90]

Kao što se može vidjeti na slici 5.2a svijetli i tamni dio razmatrane okoline je prožet šumom. Ako se npr. postavi centralna točka razmatrane okoline da je za dvije točke unutar svjetlijeg djela može se na slici 5.2b vidjeti funkciju sličnosti za zadanu okolinu. Dakle, može se vidjeti da će filter jače otežavati samo one svijetle točke u okolini zadane točke, dok će tamne praktički ignorirati. Ako provedemo zadano filtriranje za sve točke razmatrane okoline na slici 5.2a, možemo na slici 5.2c vidjeti traženi odziv gdje su površine filtrirane a rub održan. Na slici 5.3. vidi se efekt bilateralnog filtriranja na slici intenziteta svjetline.



Slika 5.3. Primjer bilateralnog filtiranja crno-bijele slike [90]

Filtiranje slika intenziteta svjetline je relativno jednostavno, jer se radi samo o jednom kanalu podataka te filtriranje proizvodi nove vrijednosti piksela samo u tom kanalu. Drugim riječima, kada se koriste samo sive boje rezultat filtriranja mogu biti samo neke druge opet sive boje. Filtiranje slika u boji je kompliciranije, jer iako se radi samo o tri slike intenziteta za, R, G i B kanal zasebno, filtriranje sva tri kanala uzrokuje pojavu artefakata zbog međusobne korelacije između boja, budući da već npr. između samo dva kanala slike u boji postoji niz drugih boja i nijansi koji se dobivaju njihovom kombinacijom. Stoga filtriranje slika u boji zasebnim filtriranjem kanala ne postiže željeni efekt. Stoga se bilateralno filtiranje slika u boji provodi u CIE L*a*b* prostoru boja, gdje euklidski pomak između neke dvije boje odgovara njihovom perceptualnom pomaku u ljudskom vizualnom sustavu. Drugim

riječima, na takav način se filtriranjem samo perceptualno relevantne boje izjednačuju te se time samo preceptualno relevantni rubovi održavaju. Na slici 5.4. mogu se vidjeti prednosti korištenja ovog tipa filtriranja na primjerima slika koje se koriste za lokalizaciju mobilnog robota u ovoj disertaciji.



Slika 5.4. Primjer bilateralnog filtiranja slike u boji

Kao što se vidi na slici 5.4., utjecaj šuma je znatno umanjen a rubovi očuvani, nedostatak je jedino smanjenje osjećaja teksturiranosti na ravnim površinama, ali nam za analizu boja to nije bitno. U ovoj disertaciji korištena je implementacija bilateralnog filtra u sklopu OpenCV okvira [91].

5.1.2. Ujednačavanje histograma

Ujednačavanje histograma je transformacija intenziteta svjetline slike pomoću kumulativne distribucije intenziteta slike [39]. Vjerojatnost pojavljivanja pojedinog intenziteta na slici se može aproksimirati sljedećim izrazom:

$$p_r(r_k) = \frac{n_k}{MN}, k = 0, 1, 2..., L-1$$
 (5.7)

gdje je *MN* ukupni broj točaka na slici, n_k je broj točaka na slici koje imaju vrijednost intenziteta r_k , dok je L ukupan broj intenziteta koji mogu biti na slici (kod slika intenziteta svjetline to je 256). Prikaz podataka $p_r(r_k)$ naspram r_k se u biti naziva histogram. Kumulativna distribucija intenziteta slike može se tada definirati na sljedeći način:

$$s_{k} = T(r_{k}) = (L-1) \cdot \sum_{j=0}^{k} p_{k}(r_{j}) = \frac{(L-1)}{MN} \cdot \sum_{j=0}^{k} n_{j}, k = 0, 1, 2, ..., L-1$$
(5.8)

Time je dobivena funkcija koja transformira svaku točku ulazne slike s intenzitetom r_k u odgovarajuću točku s intenzitetom s_k u konačnoj slici. Na slici 5.5 može se vidjeti primjer ujednačavanja histograma za sliku intenziteta svjetline koja se koristila za lokalizaciju mobilnog robota.



Slika 5.5. Primjer ujednačavanja histograma

Na slici se može primijetiti da se predstavljenom transformacijom naglašavaju kontrasti i time teksturiranost ravnih površina.

5.1.3. Nelinearno širenje histograma

Nelinearno širenje histograma se uglavnom izvodi da bi se izjednačila svjetlina na slici te se stoga često naziva prilagodba osvjetljenja (engl. gamma adjustment). Drugim riječima,

ideja je da se svjetla područja potamne, a tamna posvijetle. Za sliku koja koristi RGB prostor boja ta se transformacija najčešće izvodi pomoći sljedećeg izraza:

$$I_i = 255 \cdot \left(\frac{orig_i - lo_i}{hi_i - lo_i}\right)^{\gamma_i}, i = R, G, B$$
(5.9)

gdje je I_i nova vrijednost intenziteta i-tog kanala slike, $orig_i$ je originalna vrijednost intenziteta i-tog kanala slika, dok su lo_i i hi_i je najniža i najviša vrijednost intenziteta i-tog kanala slike. Varijabla γ_i predstavlja vrijednost osvjetljenja za i-ti kanal slike. Vrijednost osvjetljenja se računa kao logaritamska vrijednost, jer se nastoji imitirati odziv ljudskog vizualnog sustava. Osvijetljenost se može izračunati na sljedeći način:

$$\gamma_{i} = \log_{\frac{md_{i} - lo_{i}}{hi_{i} - lo_{i}}} \frac{1}{2}, i = R, G, B$$
(5.10)

gdje je md_i srednja vrijednost i-tog kanala slike. Na slici 5.6. može se vidjeti utjecaj ove transformacije na slike po danu i po noći za približno isto mjesto u prostoru.



Slika 5.6. Primjer nelinearnog širenja histograma, a) originalne, b) transformirane

Kao što se može vidjeti na slici 5.6., ova transformacija ima više utjecaja na slike snimane po noći. Korištenje ove transformacije se u ovoj disertaciji pokazalo korisnim za djelomično smanjenje utjecaja promjene osvjetljenja okoline.

5.2. Pridruživanje boja ravninskim segmentima

Kada je Microsoft Kinect senzor izašao na tržište nije bio namijenjen primjeni u mobilnoj robotici, već kućnom sustavu zabave zajedno s Microsoft Xbox 360 igraćom konzolom. Zahtjevi za preciznošću kalibracije Kinect senzora su stoga s obzirom na njenu originalnu primjenu smanjeni, što dovodi do relevantne pogreške prilikom reprojekcije neke točke s dubinske slike na RGB sliku Kinect senzora, odnosno prilikom pridruživanja svakoj točci dubinske slike odgovarajuću točku na RGB slici.

Da bi se izvela reprojekcija točke s dubinske slike na RGB sliku potrebno je poznavati intristične podatke RGB kamere i dubinske kamere te rotacijsku matricu i translacijski vektor između koordinatnih sustava tih kamera. Ovi podaci se dobivaju kalibracijom kamera koja je ovdje provedena standardnim kvadratnim uzorkom istovremeno snimljenim RGB kamerom i infracrvenom kamerom. Infracrvena kamera čini pasivnu komponentu dubinskog dijela sustava Kinect senzora. Na slici 5.7. može se vidjeti jedan primjer para slika korištenih za kalibraciju Kinect senzora.



Slika 5.7. Primjer slika korištenih za kalibraciju Microsoft Kinect senzora

Postupak reprojekcije točke s dubinske slike na RGB sliku uključuje sljedeću transformaciju iz koordinatnog sustava dubinske slike u koordinatni sustav 3D prostora kamere:

$$\mathbf{p}_{D} = \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} (u - u_{D}) \cdot z / f_{u_{D}} \\ (v - v_{D}) \cdot z / f_{v_{D}} \\ z \end{bmatrix}$$
(5.11)

gdje su *u* i *v* koordinate na dubinskoj slici, dok su u_D , v_D , f_{u_D} i f_{v_D} intristični parametri dubinske kamere. Koordinata *z* iz vrijednosti dubine se dobiva izrazom [19]:

$$z = k_3 \cdot tg\left(\frac{d}{k_2} + k_1\right) \tag{5.12}$$

gdje je *d* vrijednost dubine na dubinskoj slici, a konstante su k_1 =1.1863, k_2 =2842.5 i k_3 =0.1236. Nakon toga se točka iz koodinatnog sustava dubinske kamere transformira u koordinatni sustav RGB kamere uz poznavanje rotacijske matrice **R** i translacijskog vektora **t** na sljedeći način:

$$\mathbf{p}_{RGB} = \begin{bmatrix} x_{RGB} \\ y_{RGB} \\ z_{RGB} \end{bmatrix} = \mathbf{R} \cdot \mathbf{p}_{D} + \mathbf{t}$$
(5.13)

Potom je transformacija u koordinate RGB slike sljedeća:

$$u' = \frac{x_{RGB} \cdot f_{u_{RGB}}}{z_{RGB}} + u_{RGB}, \qquad (5.14)$$

$$v' = \frac{y_{RGB} \cdot f_{v_{RGB}}}{z_{RGB}} + v_{RGB}$$
(5.15)

gdje su *u*' i *v*' koordinate točke na RGB slici, a u_{RGB} , v_{RGB} , $f_{u_{RGB}}$ i $f_{v_{RGB}}$ intristični parametri RGB kamere. Zbog nesavršene kalibracije kamera Kinect senzora unosi se pogreška u ove prethodno opisane transformacije što u konačnici rezultira da se nekom ravninskom segmentu pridruže točke susjednog segmenta na RGB slici koji mu ne pripadaju. Stoga, da bi se smanjio utjecaj pogrešno pridijeljenih RGB točaka, potrebno je isključiti točke u okolini ruba između ravninskih segmenata prilikom formiranja deskriptora. Time je smanjen broj točaka RGB slike koji mogu biti pridijeljene nekom ravninskom segmentu, ali je isto tako smanjen utjecaj pogreške reprojekcije. Optimalna širina pojasa između dva ravninska segmenta unutar kojeg se ne uzimaju u obzir točke RGB slike je eksperimentalno procijenjena na 3 piksela. Označavanje točaka koji se mogu razmatrati prilikom formiranja deskriptora se jednostavno može provesti formiranjem binarne slike gdje su iscrtane granice određene debljine između ravninskih segmenata. Primjer segmentirane RGB slike sa i bez naglašenih granica dan je na slici 5.8.



Slika 5.8. Primjer isključivanja rubnih točaka ravninskih segmenata radi smanjenja utjecaja pogreške reprojekcije. Gornja lijeva slika prikazuje sve točke koja bi se mogle pridružiti ravninskom segmentu, gornja desna slika naglašava rubove između segmenata na dubinskoj slici, donja slika prikazuje konačno pridružene točke ravninskim segmentima.

Točke RGB slika na slici 5.8. koje imaju superponiranu neku boju su one koje su pridružene nekom ravninskom segmentu dok one koje nemaju ni jednu boju superponiranu se ili nalaze na granici između segmenata ili za njih nema podataka o dubini. Prikazana binarna slika je generirana za dubinsku sliku i reprojicirana na RGB sliku. Na slici se također može primijetiti da RGB kamera Kinect senzora ima širi horizontalni i vertikalni kut gledanja od dubinske kamere. Na slici 5.8. se jasno može vidjeti da isključenje rubnih točaka dovodi do smanjenja pridruživanja krivih točaka pojedinim ravninskim segmentima, što rezultira konzistentnijim deskriptorima temeljenim na vizualnim obilježjima ravninskih segmenata.

5.3. Generiranje inicijalnog skupa parova i evaluacija hipoteza

U ovom poglavlju će se objasniti izmjena izraza za evaluaciju hipoteza te novi uvjet koji ravninski segmenti moraju zadovoljiti da bi bili spareni na temelju sličnosti deskriptora vizualnih obilježja.

5.3.1. Uvjet sličnosti deskriptora vizualnih obilježja

Postupak generiranja inicijalnog skupa parova na temelju geometrijskih ograničenja je objašnjen u poglavlju 4.4. Ovdje će se objasniti na koji način uklopiti informaciju o vizualnim obilježjima ravninskih segmenata radi robusnijeg postupka generiranja inicijalnog skupa parova. Kao što je objašnjeno u poglavlju 4.4., da bi se par ravninskih segmenata $(R_F, R_{F'})$ smatrao uspješno sparenim, oni moraju zadovoljavati uvjete koplanarnosti i preklapanja objašnjene u poglavljima 4.4.1. i 4.4.2. Drugim riječima, moraju biti približno koplanarni i moraju se preklapati prema nekoj unaprijed definiranoj mjeri. Pored ta dva prethodno definirana uvjeta, ovdje se uvodi treći uvjet koji zahtjeva da se deskriptori vizualnih obilježja definirani za dva ravninska segmenta $(R_F, R_{F'})$ preklapaju unutar prethodno definirane mjere. Taj uvjet se može zapisati na sljedeći način:

$$h(R_F, R_{F'}) \ge \varepsilon_s \tag{5.16}$$

gdje je $h(R_F,R_F)$ funkcija koja računa vrijednost sličnosti između deskriptora vizualnih obilježja dva ravninska segmenta, dok je ε_s zadani prag sličnosti. Bez smanjenja općenitosti, pretpostavit ćemo da je funkcija *h* normirana tako da poprima vrijednosti iz intervala [0, 1]. Prag sličnosti se definira provedenim eksperimentalnim ispitivanjem deskriptora vizualnih obilježja radi utvrđivanja granice koja osigurava najmanje 90%-tnu točnost raspoznavanja parova ravninskih segmenata za zadani deskriptor. Ako se npr. za mjeru sličnosti u izrazu (5.16) koristi presjek histograma, tada se ona može zapisati na sljedeći način:

$$\sum_{i=1}^{N} \min\left(\frac{R_{Fi}}{n_{R_{F}}}, \frac{R_{F'i}}{n_{R_{F'}}}\right) \ge \varepsilon_{s}$$
(5.17)

gdje su R_{Fi} i $R_{F'i}$ vrijednosti i-tog bina histograma para ravninskih segmenata, N je broj binova u histogramima, a n_{R_F} i $n_{R_{F'}}$ brojevi točaka sadržanih u ravninskim segmentima. Dakle, u skladu s postupkom generiranja hipoteza objašnjenim u poglavlju 4.4.3. koristi se lista inicijalno sparenih parova Q_{sort} koji sada moraju zadovoljiti uvjete (4.28), (4.29) i (5.16) odnosno (5.17).

5.3.2. Izmjene evaluacije hipoteza

Smanjivanjem broja inicijalno sparenih parova ravninskih segmenata koji ulaze u postupak stvaranja hipoteza očekuje se znatno smanjenje broja generiranih hipoteza i time

ubrzavanje faze evaluacije hipoteza. Faza evaluacije hipoteza zasnovana na geometrijskim ograničenjima je opisana u poglavlju 4.5., pa će se ovdje samo objasniti modifikacija izraza za evaluaciju koji će uzimati u obzir i podatke deskriptora vizualnih obilježja ravninskih segmenata. Kao što se može vidjeti u izrazu (4.35), mjera kvalitete podudaranja hipoteze je dobivena kao vjerojatnost Mahalanobisove udaljenosti prema χ^2 -razdiobi s 3 stupnjeva slobode otežana s manjom površinom para ravninskih segmenata. Uvođenje podataka o vizualnim obilježjima postižemo dodatnim otežavanjem s iznosom mjere sličnosti deskriptora vizualnih obilježja para ravninskih segmenata, time se izraz (4.35) modificira na slijedeći način:

$$\Im(\mathbf{w},\mu) = \sum_{R_{F} \in F} \max_{(R_{F},R_{F'}) \in \Omega(\mathbf{w},\mu)} \left\{ \left(c + h(R_{F},R_{F'}) \right) \cdot P(d(R_{F},R_{F'};\mathbf{w})) \cdot \min\left\{ |R_{F}|,|R_{F'}| \right\} \right\} (5.18)$$

gdje je $h(R_{F'}M_{F'})$ korištena mjera sličnosti deskriptora boje, a *c* je konstanta. U slučaju korištenja presjeka histograma kao mjere sličnosti izraz (5.18) postaje sljedeći:

$$\mathfrak{I}(\mathbf{w},\mu) = \sum_{R_{F}\in F} \max_{(R_{F},R_{F'})\in\Omega(\mathbf{w},\mu)} \left\{ \left(c + \sum_{i=1}^{N} \min\left(\frac{R_{Fi}}{n_{R_{F}}},\frac{R_{F'i}}{n_{R_{F'}}}\right) \right) \cdot P\left(d\left(R_{F},R_{F'};\mathbf{w}\right)\right) \cdot \min\left\{ |R_{F}|,|R_{F'}|\right\} \right\}$$
(5.19)

Budući da h() u ovom slučaju poprima vrijednosti iz intervala [0, 1], težina koju unose deskriptori vizualnih obilježja se kreće između c i c+1. Razlog zadržavanja težine koju unose deskriptori vizualnih obilježja na minimalno c jest da mjera sličnosti deskriptora pomogne usmjeriti evaluaciju hipoteza tamo gdje postoji sličnost, ali da zbog manjka sličnosti vizualnih obilježja ipak ne umanji previše doprinos sličnosti geometrije. Manjak sličnosti se često može dogoditi zbog znatne promjene osvjetljenja okoline, koja će, bez obzira na mjere smanjenja ovisnosti o osvjetljenju, u stvarnim uvjetima dovesti do smanjenja sličnosti deskriptora boje, dok će geometrija prostora ostati ista. Stoga, da bi sustav lokalizacije robota u slučaju zakazivanja podsustava za analizu vizualnih obilježja ipak mogao donijeti ispravnu odluku temeljem samo geometrije, postavlja se minimalna težina sličnosti deskriptora na c. Za pokuse provedene u ovoj disertaciji koristila se vrijednost c = 1.

5.4. Deskriptori vizualnih obilježja

U ovom poglavlju su opisani i vizualizirani deskriptori tekstura i boja. Namjera je prikazati na primjerima konkretnih ravninskih segmenata "izgled" i funkcioniranje implementiranih deskriptora boja za različite prostore boja i deskriptora teksture. Ideja je prikazati sličnosti i razlike deskriptora vizualno istih, sličnih i različitih ravninskih segmenata. Također će se u ovom poglavlju objasniti modifikacija standardnih histograma koje koriste neizrazite granice između binova, dok će se u zadnjem podpoglavlju objasniti razlog odabiranja presjeka histograma kao mjere sličnosti. Budući da je karakteristike deskriptora boja najlakše objasniti na primjerima konkretnih ravninskih segmenata, prvo će biti prikazane odabrane scene snimljene u dnevnim i noćnim uvjetima te pripadajući ravninski segmenti.

5.4.1. Primjeri scena i ravninskih segmenata u stvarnim uvjetima

Za potrebe ovog ispitivanja i vizualizacije izabrano je 4 scene koje opisuju četiri lokacije na Elektrotehničkom fakultetu u Osijeku snimane u dnevnim i noćnim uvjetima. Na slici 5.9. mogu se vidjeti 4 odabrane scene, gdje lijevi stupac predstavljaja scene snimljene po danu, a desni stupac iste scene snimljene po noći. Slike koje su prikazane na slici 5.9. prethodno su obrađene na isti način na koji se to izvodi i u razmatranom sustavu za lokalizaciju mobilnog robota, odnosno primijenjeno je nelinearno širenje histograma slike na način opisan u poglavlju 5.1.3. te je provedeno bilateralno filtriranje radi uklanjanja šuma. Za demonstraciju deskriptora teksture, slike su obrađene na način opisan u poglavlju 5.1.2., ali radi uštede prostora one neće biti prikazane. Kao što se može primijetiti, parovi slika prikazuju približno istu lokaciju u prostoru snimljenu u različitim uvjetima. Razlika u pomaku i uvjetima je onakva kakva se može očekivati u stvarnom radu mobilnog robota.

Na slici 5.10. mogu se vidjeti naglašeni ravninski segmenti koji sudjeluju u ovom ispitivanju. Tako se može vidjeti da je na prvoj sceni sa slike 5.9. odabran pod sačinjen od parketa i lijevi jednoliko obojeni zid, na drugoj sceni je odabran crveni pod, na trećoj crnobijeli pod, a na četvrtoj pod sačinjen od pločica sive i blago crvene boje. Odabrani ravninski segmenti reflektiraju vizualno distinktne klase ravninskih segmenata koje uglavnom prevladavaju u testnoj bazi snimaka koja je korištena u ovoj doktorskoj disertaciji.



Slika 5.9. Odabrane scene snimljene u dnevnim i noćnim uvjetima

Naglašene točke na scenama na slici 5.10. predstavljaju isključivo one točke koje su sudjelovale u formiranju deskriptora boja. Drugim riječima, samo one točke koje su pripadale ravninskim segmentima nakon procesa segmentacije, isključivanja graničnih točaka i projiciranja točaka dubinske slike 3D senzora na njegovu RGB sliku, kao što je opisano u poglavlju 5.2. Kao što se može primijetiti na slici 5.10., segmenti snimljeni po noći sadržavaju uglavnom puno više točaka nego oni po danu. Razlog tome je znatna interferencija dnevnog svjetla s aktivnim dijelom Kinect senzora, dok umjetna rasvjeta u noćnim snimkama nema dovoljni intenzitet da uzrokuje značajniju interferenciju, što rezultira većim vidljivim površinama.



Slika 5.10. Razmatrani ravninski segmenti

5.4.2. Deskriptori RGB prostora boja

Kao što je objašnjeno u poglavlju 3.1.2., RGB prostor boja se prvenstveno koristi u sustavima snimanja i reprodukcije slike u boji, stoga on opisuje cijeli spektar boja koji se može reproducirati na standardnim uređajima za prikaz slike. Budući da se RGB prostor boja sastoji od tri kromatske komponente, kao deskriptor ovog prostora boja se koristi trodimenzionalni histogram. Trodimenzionalni histogram se može prikazati kao skupina kuglica razmještenih u trodimenzionalnim prostoru, gdje svaka kuglica predstavlja jedan bin tog histograma. Radijus kuglice predstavlja zastupljenost boja koje obuhvaća taj bin unutar histograma. Na slici 5.11. mogu se vidjeti trodimenzionalni histogrami za cjelokupni spektar boja RGB prostora za tri rezolucije histograma.



Slika 5.11. Svi binovi RGB histograma za rezoluciju od a) 32, b) 16, c) 8 binova po dimenziji prostora boja

Boja pojedine kuglice na slici 5.11. je definirana kao srednja vrijednost boja koje taj bin obuhvaća, dok je radijus svake kuglice jednak za potrebe vizualizacije RGB prostora boja. Prikazani histogrami bi u biti odgovarali histogramu slike koja sadrži jedan piksel za svaku moguću nijansu boje u RGB prostoru. Na slici 5.12. može se za odabrane ravninske segmente vidjeti izgled trodimenzionalnog histograma RGB prostora boja za rezoluciju od 16 binova po dimenziji prostora, odnosno maksimalni broj binova za taj histogram je 4096. Slika 5.12. prikazuje normalizirane histograme ravninskih segmenata s obzirom na ukupan broj točaka koji ih je stvorio. Budući da radijus kuglica prikazanih u prostoru govori koliko je zastupljena boja tog bina na razmatranom ravninskim segmentu, suma svih radijusa prikazanih kuglica bi onda trebala biti 1. Da bi vizualizacija bila uočljivija veličina kuglica je u korištenom 3D vizualizacijskom alatu automatski skalirana s obzirom na najveću kuglicu da bi se mogla bolje uočiti distribucija boja.



Slika 5.12. Trodimenzionalni histogrami RGB prostora boja za razmatrane ravninske segmente pri rezoluciji od 16 binova po dimenziji

Na slici 5.12. su prikazani histogrami za segmente poredane na isti način kao i na slici 5.10., odnosno poredani su s lijeva na desno i od gore prema dolje slijedećim redoslijedom: pod parket dan, pod parket noć, zid dan, zid noć, crveni pod dan, crveni pod noć, crno-bijeli pod dan, crno-bijeli pod noć, pod s pločicama dan, pod s pločicama noć. Boja pojedine kuglice na histogramu je definirana pomoću koordinate centra bina. Promatrajući dobivene histograme razmatranih ravninskih segmenata možemo općenito primijetiti da su histogrami segmenata snimljenih u dnevnom svjetlu koncentrirani u manji broj binova nego oni snimljeni po noći. Razlog tome je djelomice to što segmenti snimljeni u noćnim uvjetima imaju puno više točaka pa obuhvaćaju i dijelove plohe koje nisu obuhvaćene danju. Dok je veći razlog to što ravninski segmenti snimljeni po noći nisu ravnomjerno osvijetljeni te zbog pojave sjena dolazi do raspršivanja boja s obzirom na dijagonalu RGB prostora boja odnosno dužinom tzv. linije sive boje koja je vidljiva na slici 3.12. Radikalniju promjenu histograma jedino ima ravninski segment koji opisuje crno-bijeli pod gdje se zbog refleksije umjetnog svjetla naglašavaju njegovi bijeli dijelovi. Konačni zaključak je taj da između ravninskih segmenata snimljenih po danu i onih snimljenih po noći postoji znatna razlika u distribuciji boja po binovima, posebno u smjeru dijagonale. Nadalje, u disertaciji će deskriptori zasnovani na RGB prostoru boja biti označeni s "RGBkk", gdje kk definira broj binova po dimenziji RGB prostora boja.

5.4.3. Deskriptori CIE XYZ prostora boja

Slično kao i kod RGB prostora boja, za deskriptor XYZ prostora boja se također uzima trodimenzionalni histogram. Primjer izgleda histograma cijelog spektra boja ovog prostora boja za tri rezolucije se može vidjeti na slici 5.13. Slično kao kod RGB primjera i ovdje kuglice predstavljaju binove histograma, a radijus kuglice zastupljenost te boje u histogramu. Boja pojedine kuglice je srednja vrijednost boja obuhvaćenih granicama bina, odnosno odgovara boji centra bina.



Slika 5.13. Svi binovi XYZ prostora boja za rezoluciju od a) 32, b) 16, c) 8 binova po dimenziji prostora boja

Dakle, prikaz primjera deskriptora XYZ prostora boja za odabrane ravninske segmente je identičan onome za RGB prostor boja. Boja neke kuglice bina se određuje transformacijom koordinata centra bina iz XYZ u RGB prostor boja. Položaj same kuglice se zadržava u koordinatama XYZ prostora boja. Na slici 5.14. može se vidjeti prikaz trodimenzionalnih histograma XYZ prostora boja za odabrane ravninske segmente u slučaju kada imamo 16 binova po dimenziji ovog prostora boja.

Pogledom na dobivene slike histograma može se doći do istog zaključka kao i kod RGB prostora boja, a to je da se promjenom osvjetljenja na sceni uvelike mijenja distribucija boje u histogramu, što može dovesti do poteškoća prilikom raspoznavanja istih segmenata snimljenih pri različitim svjetlosnim uvjetima. Nadalje, u disertaciji će deskriptori zasnovani na CIE XYZ prostoru boja biti označeni s "XYZ*kk*", gdje *kk* definira broj binova po dimenziji CIE XYZ prostora boja.



Slika 5.14. Trodimenzionalni histogrami XYZ prostora boja za razmatrane ravninske segmente pri rezoluciji od 16 binova po dimenziji

5.4.4. Deskriptori kromatskog rgb prostora boja

U poglavlju 3.1.3. opisano je kako se kromatski rgb prostor izvodi iz RGB prostora na način da se R, G i B komponenta normira sumiranim intenzitetima sve tri komponente RGB prostora boja. Na takav način, po dijagonalnom modelu opisanom u poglavlju 3.1.1., gubi se ovisnost o promjeni intenziteta svijetlosti. Budući da se jedna dimenzija ovog prostora boja uvijek može izvesti iz ostale dvije, deskriptor ovog prostora boja se formira s obzirom samo na r i g komponentu. Stoga je deskriptor ovog prostora boja dvodimenzionalni histogram čiji je spektar vrijednosti binova vidljiv na slici 5.15.



Slika 5.15. Svi binovi kromatskog rgb prostora boja za rezolucije a) 32, b) 16, c) 8 binova po korištenoj dimenziji

Na slici 5.15. obojani su samo oni binovi histograma koji imaju smisla u ovom prostoru boja, odnosno samo oni binovi kojima je $r+g \le 1$, jer bi se u suprotnom mogle dobiti vrijednosti koje nisu podržane u originalnom RGB prostoru boja. Time se je efektivna površina histograma smanjila za pola. Zbog prirode ovog prostora boja i činjenice da gubimo informaciju o intenzitetu svjetlosti po dimenziji RGB prostora boja, pojedina boja koja je pridijeljena nekom binu u ovoj vizualizaciji je stoga dobivena relacijama:

1

$$R = r \cdot 255 \tag{5.20}$$

$$G = g \cdot 255 \tag{5.21}$$

$$B = (1 - r - g) \cdot 255 \tag{5.22}$$

Također boja nekog bina je definirana kao srednja vrijednost boja sadržanih unutar granica tog bina, odnosno kao boja centra bina. Dvodimenzionalni histogram ovog prostora boja prikazuje se kao stupčasti dijagram, gdje visina stupca odgovara zastupljenosti boja obuhvaćenih granicama tog bina za neki razmatrani ravninski segment. Budući da iz samog položaja bina histograma ne možemo dobiti točnu boju, zbog odsustva informacije o

intenzitetu, te, kao što je rečeno, ona se računa prema izrazima (5.20), (5.21) i (5.22), ovaj prikaz histograma može služiti samo za vizualizaciju i ne mora predstavljati stvarnu boju točaka sadržanih u binu. Na slici 5.16. se za 10 razmatranih ravninskih segmenata može vidjeti prikaz dvodimenzionalnog histograma kromatskog rgb prostora boja kada imamo 16 binova po dimenziji prostora boja.



Slika 5.16. Dvodimenzionalni histogrami rgb kromatskog prostora boja za razmatrane ravninske segmente pri rezoluciji od 16 binova po dimenziji
Promatrajući prikazane histograme na slici 5.15. može se doći do zaključka da ovaj prostor boja ne diskriminira jasno pojedine tipova razmatranih ravninskih segmenata. Razlog zašto je to tako se krije u međusobnim odnosima RGB koordinata piksela razmatranih ravninskih segmenata. Ako se ponovo promotri slika 5.12. koja prikazuje trodimenzionalne RGB histograme razmatranih ravninskih segmenata može se primijetiti da su većina većih kuglica jako blizu glavne dijagonale uz manja odstupanja prema crvenoj osi RGB prostora boja za segmente koji su malo crveniji. Sve to sugerira da će, kada se normiraju RGB vrijednosti i dobiju koordinate u kromatskom rgb prostoru boja, one biti relativno blizu jedna drugoj, što je vidljivo i u prikazanim histogramima. Većina stupaca koji se ističu se nalaze oko centralnog bina histograma kromatskog rgb prostora boja, koji odgovara bojama s približno istim omjerima R, G i B koordinate RGB prostora boja, manji stupići su uglavnom pomaknuti prema crvenoj komponenti kromatskog rgb prostora boja zbog prvenstveno crvenih ravninskih segmenata u slučajevima parketa, crvenog poda i poda s pločicama. Dobra strana ovog prostora boja koja je vidljiva na prikazanim histogramima je relativna neosjetljivost na promjenu intenziteta osvjetljenja scene. Nadalje, u disertaciji će deskriptori zasnovani na kromatskom grb prostoru boja biti označeni s "RGCkk", gdje kk definira broj binova po korištenoj dimenziji kromatskog rgb prostora boja.

5.4.5. Deskriptori CIE L*a*b* prostora boja

CIE L*a*b* je drugi matematički definiran prostor boja objavljen od strane organizacije CIE koji je korišten u ovoj disertaciji. Ovaj prostor boja je izveden iz CIE XYZ prostora boja na način objašnjen u poglavlju 3.1.6., te se sastoji od tri dimenzije, kromatskih komponenti a* i b* te svjetline L*. Kod ovog deskriptora imamo mogućnost umanjiti utjecaj osvjetljenja te se stoga za formiranje deskriptora ovog prostora boja koriste samo a* i b* komponente. Time stvaramo dvodimenzionalne histograme čiji se primjeri spektra binova za različite rezolucije mogu vidjeti na slici 5.17.



Slika 5.17. Svi binovi CIE L*a*b* prostora boja za rezolucije a) 32, b) 16, c) 8 binova po korištenim dimenzijama ovog prostora boja pri L* = 50

Budući da prilikom stvaranja deskriptora ovog prostora boja gubimo jednu dimenziju, za potrebe vizualizacije se mora pretpostaviti određena razine svjetline L*. Stoga prikazani spektar binova na slici 5.17. predstavlja presjek L*a*b* prostora boja po a*-b* ravnini za polovicu maksimalne vrijednosti svjetline L*, pa ovaj prikaz može služiti samo za potrebe vizualizacije i ne mora predstavljati snimljenu boju u stvarnom prostoru. Dakle, boja pridijeljena nekom binu je dobivena obrnutnom transformacijom koordinata centra bina u RGB prostor boja preko XYZ prostora boja uz pretpostavku svjetline $L^*=50$. Također zbog svoje prirode i prostora koji opisuje, ovaj prostor boja ima smanjenu efektivnu iskoristivost prostora dvodimenzionalnog histograma.

Budući da se ovdje radi o dvodimenzionalnom histogramu, kao i kod kromatskog rgb prostora boja, prikazujemo ga kao stupčasti dijagram gdje visina stupca odgovara zastupljenosti boja obuhvaćenih granicama pojedinog bina. Na slici 5.18. mogu se vidjeti dvodimenzionalni histogrami za L*a*b* prostor boja pri rezoluciji od 16 binova po korištenoj dimenziji ovog prostora boja.



Slika 5.18. Dvodimenzionalni histogrami CIE L*a*b* prostora boja za razmatrane ravninske segmente pri rezoluciji od 16 binova po dimenziji

Slično kao i kod dvodimenzionalnog histograma za kromatski rgb prostor boja i ovdje se stupci uglavnom kreću oko neke centralne vrijednosti koja odgovara sivim bojama. Ipak, ovaj prostor boja daje veću deskriptovnost za razlikovanje razmatranih primjera ravninskih segmenata. Po prikazanim histogramima se također može zaključiti da ovaj prostor boja pokazuje znatnu ovisnost o promjeni osvjetljenja okoline što je posebno vidljivo za zidni segment i crno-bijeli pod gdje se za noćnu scenu znatno promijenila distribucija boja na histogramu. Nadalje, u disertaciji će deskriptori zasnovani na CIE L*a*b* prostoru boja biti označeni s "Lab*kk*", gdje *kk* definira broj binova po korištenoj dimenziji CIE L*a*b* prostora boja.

5.4.6. Deskriptori HSV prostora boja

HSV prostor boja je detaljno objašnjen u poglavlju 3.1.4., pa će se stoga ovdje samo objasniti deskriptor koji se koristio za ovaj prostor boja. Budući da je cilj koristiti deskriptore boja koji su što više neovisni o osvjetljenju prisutnom u okolini, kod korištenih prostora boja ne koriste se one dimenzije prostora boja koje su izrazito osjetljive na osvjetljenje ili predstavljaju intenzitet osvjetljenja same po sebi. Stoga, slično kao i kod L*a*b* prostora boja, u HSV prostoru boja ćemo za formiranje deskriptora koristiti samo H i S komponente, dakle komponente koje opisuju nijansu i zasićenje pronađene boje na slici. Time naravno gubimo informaciju o svjetlini pojedine nijanse na analiziranoj slici te stoga iz ovog deskriptora slično kao i kod ostalih dvodimenzionalnih histograma nije moguće rekonstruirati točnu boju koja je bila na slici. Budući da je HSV prostor boja opisan polarnim koordinatama, za njegovu vizualizaciju ćemo koristiti polarni dijagram čiji se spektri binova za tri rezolucije histograma mogu vidjeti na slici 5.19.



Slika 5.19. Svi binovi HSV prostora boja za rezolucije a) 32, b) 16, c) 8 binova po korištenoj dimenziji, pri V = 100

Kao što se može vidjeti na slici 5.19. svaki krug predstavlja jedan bin dvodimenzionalnog histograma. Boja krugova je definirana srednjom vrijednošću H i S koordinata točaka prostora boja koje su obuhvaćene granicama bina, odnosno vrijednostima centra bina. Radijus odnosno površina prikazanih obojanih krugova predstavlja zastupljenost boja obuhvaćenih tim binom u histogramu. Slično kao i kod primjera RGB histograma, prikazani histogrami bi mogli predstavljati histogram slike u kojoj je svaka moguća boja zastupljena jednim pikselom.

Na slici 5.20. mogu se vidjeti histogrami HSV prostora boja sa 16 binova po korištenoj dimenziji prostora boja za razmatranih 10 segmenata. Promatranjem dobivenih histograma može se primijetiti, slično kao i kod RGB prostora boja, da segmenti snimani u noćnim uvjetima imaju raspršenije distribucije. Jedini izuzetak je segment crno-bijelog poda koji počne jako naglašavati svoje bijele dijelove. Također se može zaključiti da HSV prostor boja, za razliku od RGB prostora boja, puno bolje zadržava zastupljenost istih binova kada se uspoređuju isti segmenti snimani u različitim uvjetima, što sugerira da ovaj prostor boja jasno pokazuje puno bolja svojstva neovisnosti o promjenama osvjetljenja okoline. Nadalje, u disertaciji će deskriptori zasnovani na HSV prostoru boja biti označeni s "HSV*kk*", gdje *kk* definira broj binova po korištenoj dimenziji HSV prostora boja.



Slika 5.20. Dvodimenzionalni histogrami HSV prostora boja za razmatrane ravninske segmente pri rezoluciji od 16 binova po dimenziji

5.4.7. Deskriptori prostora suprotstavljenih boja

Prostor suprotstavljenih boja se izvodi iz RGB prostora boja na način objašnjen u poglavlju 3.1.7. Ovaj prostor boja se sastoji od tri dimenzije O_1 , O_2 i O_3 pri čemu ova posljednja komponenta predstavlja svjetlinu, koja zbog namjere da se smanji utjecaj osvjetljenja neće biti korištena. Stoga je deskriptor ovog prostora boja također dvodimenzionalni histogram opisan dimenzijama O_1 i O_2 . Također, ovaj je prostor boja svojim transformacijama za ove dvije korištene dimenzije postigao neovisnost o pomaku intenziteta svjetlosti, što dodatno potvrđuje korisnost ovog prostora boja. Na slici 5.21. može se vidjeti spektar vrijednosti binova za tri rezolucije histograma.



Slika 5.21. Svi binovi prostora suprotstavljenih boja za rezolucije a) 32, b) 16, c) 8 binova po korištenoj dimenziji, uz $O_3 \approx 220$

Kao što se može primijetiti na slici 5.21., ovaj prostor boja izrazito podsjeća na HSV prostor boja. Također, kao i kod kromatskog rgb prostora boja i L*a*b* prostora boja i ovdje su obojani samo oni binovi čije koordinate centra daju neku postojeću boju u RGB prostoru boja. Ovaj prostor boja također ima određeni broj binova koji su neiskorišteni, jer obrnutom transformacijom od one opisane u poglavlju 3.1.7. ne opisuju neku boju u RGB prostoru boja. Također, budući da se korištenjem samo dvije komponente ovog prostora boja izgubila informacija o svjetlini, boje za binove prikazane na slici su dobivene za polovicu maksimalne vrijednosti svjetline. Stoga, zbog te pretpostavke o svjetlini, prikazani binovi u biti predstavljaju presjek ovog prostora boja u ravnini O₁O₂ za polovicu svjetline te time pridružene boje mogu samo služiti za potrebe vizualizacije i ne moraju predstavljati stvarne boje točaka korištenih na RGB slici. Prostor suprotstavljenih boja se također izvodi iz RGB prostora boja, ali njegova transformacija s obzirom na dijagonalni model osigurava neovisnost s obzirom na pomak intenziteta svijetlosti, pa se očekuju bolje performanse od RGB prostora boja. Na slici 5.22. se za razmatranih 10 ravninskih segmenata mogu vidjeti stupčasti prikazi

dvodimenzionalnih histograma za prostor suprotstavljenih boja s rezolucijom od 16 binova po korištenoj dimenziji prostora boja.



Slika 5.22. Dvodimenzionalni histogrami prostora suprotstavljenih boja za razmatrane ravninske segmente pri rezoluciji od 16 binova po dimenziji

Promatranjem dobivenih histograma na slici 5.22., može se primijetiti da se i ovdje, slično kao i kod kromatskog rgb i L*a*b* prostora boja, glavni stupci kreću oko neke centralne točke koja odgovara nijansama sive boje, ali ovdje, slično kao i kod L*a*b*

prostora boja, pojedini tipovi razmatranih ravninskih segmenata jasno se mogu diskriminirati od ostalih. Time je ovaj prostor boja sličan L*a*b* i određenoj mjeri HSV prostoru boja u svojoj diskriminaciji pojedinih segmenata. Nadalje, u disertaciji će deskriptori zasnovani na prostoru suprotstavljenih boja biti označeni s "Oppkk", gdje kk definira broj binova po korištenoj dimenziji prostora suprotstavljenih boja.

5.4.8. LBP porodica deskriptora teksture

U sklopu ove disertacije također su implementirani deskriptori teksture zasnovani na porodici LBP deskriptora teksture. Ova porodica deskriptora je zbog svoje jednostavnosti implementacije, brzine stvaranja i prijavljenih rezultata [62], jedina izabrana za implementaciju. Pojedini deskriptori, koji imaju generalno bolje prijavljene rezultate raspoznavanja teksture kao što su MR8 tekstoni [68, 64], zahtijevaju puno kompliciraniju implementaciju, koja bi previše usporila rad sustava za lokalizaciju mobilnog robota. Stoga su ovdje implementirana tri deskriptora LBP porodice: $LBP_{P,R}$, $LBP_{P,R}^{riu^2}$ i $LBP_{P,R}^{riu^2} / VAR_{P,R}$. Kao što je objašnjeno u poglavlju 3.2.1., LBP deskriptori tekstura su strukturalno-statističkog tipa. Strukturalno iz razloga što svaka LBP kombinacija u okolini neke točke predstavlja jednu moguću strukturu, dok statističku komponentu predstavlja formiranje histograma na osnovu broja tih pojedinih strukturnih elemenata. Formiranje ovih deskriptora teksture je provedeno nad slikom intenziteta svjetline koja je prethodno obrađena na način opisan u poglavlju 5.1.2., odnosno ulazna slika intenziteta je obrađena ujednačavanjem histograma radi naglašavanja pojedinih obrazaca teksture. Također, budući da ovi deskriptori teksture promatraju obrasce u nekoj okolini r, broj raspoloživih točaka nekog ravninskog segmenta se dodatno smanjuje na način opisan u poglavlju 5.2. Povećanje granica između ravninskih segmenta odgovara radijusu r okoline točaka koje se promatraju. Način korištenja ovih deskriptora teksture u razmatranom sustavu nije tipičan. Tipičan način korištenja deskriptora teksture podrazumijeva fazu učenja s prethodno klasificiranim i grupiranim slikama tekstura po njihovim tipovima i formiranje općeg deskriptora odnosno klasifikatora tih tipova tekstura. Nasuprot tome, proces lokalizacije mobilnog robota koji se ovdje razmatra podrazumijeva automatsko generiranje karte okoline bez ručne klasifikacije tekstura koja bi omogućila njihovo klasteriranje u opći oblik pojedinog tipa tekstura. Odnosno, u radu ovog sustava ne postoji faza učenja sustava izuzev stvaranja karte. Stoga se ovdje LBP deskriptori pojedinog ravninskg segmenta formiraju pomoću jednog histograma kojeg formiraju točke samo jednog uzorka tog ravninskog segmenta te se zbog toga mogu očekivati slabije performanse od onih prijavljenih u literaturi.

 $LBP_{P,R}$ deskriptor je formiran kao histogram kojem broj binova odgovara vrijednosti $n_B = P^2/2$, gdje je *P* broj točaka koje se promatraju u okolini. Korištene konfiguracije ovog deskriptora uključuju *P*=8,16,24 za okolinu *R*=1,2,3, što će stoga rezultirati histogramima s 32, 128 i 288 binova. Distribuciju *LBP*_{16,2} obrazaca za prethodno korištene segmente parketa i zida snimane po danu odnosno noću možemo vidjeti na slici 5.23.



Slika 5.23. Primjeri izgleda $LBP_{16,2}$ deskriptora, za ravninske segmente parketa i zida snimane pri dnevnom i noćnom osvjetljenju. Gornji red su parket a) dan, b) noć, a donji zid c) dan, d) noć.

Na slici 5.23. je vidljiva velika sličnost između histograma parketa i zida koji imaju vidno različite teksture te iz tog razloga nisu prikazani ostali histogrami ravninskih segmenata, koji su služili kao primjeri kod prostora boja.

Deskriptor $LBP_{P,R}^{riu^2}$ je po definiciji neovisan o rotaciji pogleda, budući da se u njemu promatraju samo tzv. "uniformni" obrasci. Po literaturi "uniformni" obrasci bi trebali činiti većinu čak i do 90% u standardiziranim tipovima teksturnih obrazaca [62]. Na slici 5.24. mogu se vidjeti distribucije $LBP_{16,2}^{riu^2}$ obrazaca za parket i zid snimane po danu odnosno noću. Broj binova histograma ovog deskriptora je definiran kao $n_B = P + 2$, što odgovara broju "uniformnih" obrazaca za zadani broj točaka koje se promatra u okolini uvećan za još jedan bin koji predstavlja ostale "neuniformne" obrasce. Na slici 5.24. se može primijetiti da ovdje uniformni obrasci čine samo oko 50% svih tipova obrazaca.

Analizirajući histograme na slici 5.24. može se primijetiti da je broj obrazaca kod parketa koji predstavljaju rubove (obrasci koji imaju 2 do 14 točaka u okolini s većom svjetlinom od centra) uglavnom uniformno raspoređen, dok su kod zida oni uglavnom grupirani u one obrasce koje odgovaraju područjima približno konstantnog gradijenta svjetline (obrasci koji imaju približno jednaki broj točaka u okolini koje su svjetlije ili tamnije od centra).



Slika 5.24. Primjeri izgleda $LBP_{16,2}^{riu^2}$, za ravninske segmente parketa i zida snimane pri dnevnom i noćnom osvjetljenju. Gornji red su parket a) dan, b) noć, a donji zid c) dan, d) noć

Treći tip LBP deskriptora koji je implementiran uključuje kombinaciju $LBP_{P,R}^{riu^2}/VAR_{P,R}$. Ovdje se radi o dvodimenzionalnim histogramu kao deskriptoru koji se formira tako da se za svaku točku nađe $LBP_{P,R}^{riu^2}$ obrazac te vrijednost varijance za razmatranu okolinu prema izrazu (3.44). U dimenziji $LBP_{P,R}^{riu^2}$ raspored binova je isti kao i za prethodni deskriptor, dok je dimenzija $VAR_{P,R}$ podijeljena na ukupno 17 binova čijih 16 dijeli varijance

do vrijednosti 200 dok zadnji 17. bin prihvaća sve ostale vrijednosti varijance veće od 200. U članku [62] je preporučeno da se broj i raspored binova po $VAR_{P,R}$ dimenziji odredi tako da svaki bin ima jednaku vjerojatnost odabira, pri čemu se granice trebaju odrediti analizom iz prethodno klasteriranih podataka o tipovima tekstura. Budući da mi ovdje ne klasteriramo podatke po tipu tekstura već segmente nastojimo raspoznati kao individualne objekte, binovi su po $VAR_{P,R}$ dimenziji raspoređeni linearno na 16 binova. Na slici 5.25. može se vidjeti distribucija točaka po binovima $LBP_{16,2}^{riu2} / VAR_{16,2}$ deskriptora za ravninske segmente parketa i poda.



Slika 5.25. Primjeri izgleda $LBP_{16,2}^{riu^2} / VAR_{16,2}$ za ravninske segmente parketa i zida snimane pri dnevnom i noćnom osvjetljenju. Gornji red su parket a) dan, b) noć, a donji zid c) dan, d) noć

5.4.9. Meki histogrami

Klasični način stvaranja histograma, gdje se svaka točka uvrštava samo u onaj bin unutar čijih granica se nalazi, može se smatrati neprikladnim s obzirom na točke koje su bliske granicama binova histograma. Drugim riječima, ako se neka točka ne nalazi u središtu bina, već je pomaknuta za određeni iznos, prikladnije bi bilo da se ta točka uvrsti u nekoj mjeri i u susjedne binove prema kojima naginje. Stoga je ideja mekih (engl. soft) histograma da za svaku točku koja se uvrštava u neki histogram proračunava njezina udaljenost od središta susjednih binova i u ovisnosti o toj udaljenosti proračunavaju težine s kojima se ta točka dodaje tim susjednim binovima. Pri tome ukupni zbroj težina za neku točku mora biti 1. Određivanje težina odnosno pripadnosti točke binovima pomoću linearne funkcije za jednodimenzionalni histogram može se vidjeti na slici 5.26.



Slika 5.26. Određivanje težine kod jednodimenzionalnih mekih histograma

Na slici 5.26. može se vidjeti okolina jednog bina za jednodimenzionalni histogram, pri čemu su sa većim crticama određene granice binova, a sa manjim središta binova koji također imaju vrijednosti A, B i C. Linearna funkcija koja definira težinu pripadnosti točke nekom binu ima trokutasti oblik. Funkcija pripadnosti ima središte u centru bina, nakon čega u oba smjera pada do centara susjednih binova, dakle nagibom od $-1/\Delta$, gdje je Δ širina bina histograma. Ako promatramo neku točku Z, vrijednost težina koje će se dodati binovima u njezinom susjedstvu možemo dobiti prema sljedećim izrazima:

$$w_A = 1 - \frac{|Z - A|}{\Delta} \tag{5.23}$$

$$w_B = 1 - \frac{|Z - B|}{\Delta} \tag{5.24}$$

Ako se pogleda slika 5.26, može se reći da su za točku Z težine $w_A = b$, $w_B = a + b$. Isti način razmišljanja se može primijeniti i za višedimenzionalne histograme. U tom slučaju će se

težina koja se dodaje nekom binu dobiti umnoškom dobivenih težina za svaku dimenziju, kao što se može vidjeti za dvodimenzionalni slučaj u sljedećem izrazu:

$$w_{B} = \left(1 - \frac{|x - x_{B}|}{\Delta}\right) \cdot \left(1 - \frac{|y - y_{B}|}{\Delta}\right)$$
(5.25)

gdje je w_B težina koja se dodaje binu B, x i y su koordinate trenutno razmatrane točke, a x_B i y_B koordinate centra bina B, Δ je širina bina u x i y smjeru. U slučaju jednodimenzionalnih histograma najveći broj binova koji se moraju ažurirati je uvijek 2, onaj čiji je centar najbliži razmatranoj točci i onaj drugi najbliži. Ako razmatrana točka ima jednaku koordinatu kao centar bina, onda se ažurira samo taj bin. Kada koristimo višedimenzionalne histograma, najveći broj binova koje moramo ažurirati u nekom trenutku raste eksponencijalno s brojem dimenzija. Stoga kada koristimo dvodimenzionalne histograme, najveći broj binova koje trebamo ažurirati je 4, kao što se može vidjeti na slici 5.27.



Slika 5.27. Najveći broj binova koji se trebaju ažurirati u slučaju dvodimenzionalnih mekih histograma je 4

Na slici 5.27. može se vidjeti koje binove je potrebno ažurirati kod dvodimenzionalnih histograma na primjeru tri točke X, Y i Z. Crvenom bojom je na slici 5.27. označen slučaj kada se mora ažurirati 4 bina, a zelenom i plavom bojom kada se ažurira samo 2 bina. Na slici 5.28. može se vidjeti raspored vrijednosti težina koje će biti dodane binu za točke koje se nalaze unutar tog istog bina.



Slika 5.28. Raspodjela težina unutar jednog bina dvodimenzionalnog mekog histograma

Slično se može dobiti i za trodimenzionalne histograme, samo se u tom slučaju zbog povećanja dimenzije povećava i najveći broj binova koje je potrebno ažurirati, te u formuli za težinu imamo tri umnoška. Zbog sličnosti s jednodimenzionalnim i dvodimenzionalnim primjerima mekih histograma, trodimenzionalne histograme nećemo ovdje zasebno objašnjavati. Deskriptori zasnovani na prostorima boja koji budu formirani kao histogrami mekog tipa će u svom nazivu pored naziva prostora boja i broja korištenih binova po dimenziji imati oznaku 'S'.

5.4.10. Mjere sličnosti deskriptora vizualnih obilježja

U poglavlju 3.1.8. je spomenut niz mjera kojima se može utvrditi sličnost ili razlika distribucija opisanih histogramima. Smisao korištenja mjera ovdje je dakle utvrđivanje koliko su korišteni deskriptori vizualnih obilježja ravninskih segmenata slični odnosno različiti unutar neke skale. Stoga je cilj odabrati onu mjeru koja će što bolje istaknuti ispravne parove ravninskih segmenata, dakle one koji predstavljaju istu površinu u okolini robota, od onih netočnih. Pored što boljeg isticanja ispravnih parova, kod izbora dobre mjere se također mora paziti o njenoj numeričkoj stabilnosti, a zbog zahtjeva mobilne robotike za što bržom obradom podataka, mora se voditi računa i o računskoj kompleksnosti mjere odnosno vremenu potrebnom za njezin proračun. Na slici 5.29. može se vidjeti razmještaj razmatranih ravninskih segmenata na matrici sparivanja radi lakše analize vrijednosti na slici 5.30. Na slici 5.29. crvenim je simbolima označeno sparivanje segmenta samog sa sobom te će one uvijek biti maksimalne vrijednosti za pojedinu mjeru ako se radi o mjeri sličnosti ili 0 ako se radi o mjeri razlike. Plavim simbolima su označeni točni parovi ravninskih segmenata, oni se nalaze

i iznad i ispod glavne dijagonale budući da usporedba segmenata može ići u oba smjera. Ostale ćelije u grafu su netočni parovi.



Slika 5.29. Razmještaj ravninskih segmenata u matrici sparivanja

Na osnovi grafa na slici 5.29. možemo na slici 5.30. promatrati matrice sparivanja za devet mjera sličnosti odnosno razlike nabrojanih u poglavlju 3.1.8. Ovdje su se mjere proračunavale za HSV16 deskriptor i prethodno razmatranih 10 ravninskih segmenata. HSV deskriptor je odabran za ovo ispitivanje zbog prikazanog najboljeg razlikovanja pojedinih ravninskih segmenata u poglavlju 5.4.6.

U vezi slike 5.30. je potrebno naglasiti da, budući da su neke mjere mjere sličnosti, a neke mjere razlike, vrijednosti koje one prikazuju za točne parove(na dijagonali) bi trebale biti maksimalne ako se radi o mjeri sličnosti odnosno minimalne ako se radi o mjeri razlike. Također većina mjera daje iste vrijednosti bez obzira o smjeru usporedbe odnosno d(I,M) = d(M,I), ali za neke mjere kao što su to divergencija ili normalizirani koeficijent korelacije to ne vrijedi. Stoga su zbog lakše usporedbe mjera proračunate i prikazane vrijednosti za oba smjera usporedbe za sve razmatrane mjere uključujući i usporedbu segmenta sa samim sobom. Kao što se može vidjeti na slici 5.30., uglavnom sve mjere donekle ističu ispravne parove ravninskih segmenata u odnosu na neispravne, jedina je iznimka L₂ udaljenost, koja za segmente zida pokazuje dosta malu razliku u udaljenosti između točnih i netočnih parova. Također, sve mjere imaju problema s jasnim razlikovanjem između vizualno sličnih, ali pogrešnih parova, kao što su usporedbe parketnog poda s crvenim podom i podom napravljenim od pločica. Zajednička crvena komponenta kod tih tipova



segmenata uzrokuje veliku mjeru sličnosti odnosno malu mjeru razlike kada se međusobno

uspoređuju.

Slika 5.30. Matrice sparivanja za 9 mjera sličnosti/razlike koje uključuju: a) presjek histograma, b) L₁ udaljenost, c) L₂ udaljenost, d) Bhattacharyya udaljenost, e) Matusita udaljenost, f) χ^2 udaljenost, g) divergenciju, h) Jeffrey divergenciju, i) normalizirani koeficijent korelacije

Kada govorimo o numeričkoj stabilnosti neke mjere, ovdje uglavnom mislimo na pojavu nedopuštenih operacija kao što je dijeljenje s nulom ili logaritam od nule. Budući da se uspoređuje veliki broj binova histograma, jako često prilikom usporedbe segmenata se može dogoditi da na nekom binu histograma kod jednog segmenta bude vrijednost nula a kod drugog ne, ili da oboje budu nula. Takve događaje za pojedine mjere treba detektirati što može dovesti do komplikacija prilikom implementacije te pogrešaka prilikom njihovog izračuna. To se posebno odnosi na sve mjere koje koriste logaritme i dijeljenje ili oboje. U tablici 5.1. može se vidjeti za deskriptore HSV16 i RGB16 prosječno vrijeme izračuna pojedinih mjera za 1000 usporedbi histograma. Mjerenje vremena i ovo ispitivanje je izvedeno u programskom jeziku Python, što znači da sama izvedba programa nije optimalna po brzini izvođenja, ali može dati indikaciju o brzini s kojom se može izračunati vrijednost pojedine mjere.

HSV	HSV16	RGB16
115 V	(ms)	(ms)
Presjek histograma	0.0324	0. 0953
L ₁ udaljenost	0.0332	0.1023
L ₂ udaljenost	0.0342	0.1105
Bhattacharyya udaljenost	0.0376	0.1316
Matusita udaljenost	0.0416	0.1759
χ^2 udaljenost	0.0988	0.2017
Divergencija	2.0955	56.411
Jeffrey divergencija	2.1234	56.265
Normalizirani koeficijent korelacije	0.0899	0.1922

Tablica 5.1. Brzine izračunavanja mjera sličnosti/razlike

Kao što se može primijetiti u tablici 5.1., presjek histograma ima prosječno najkraće vrijeme izračuna, dok L_1 i L_2 udaljenosti imaju približno slične, ali ipak malo duže vrijeme izračuna. Ostale mjere imaju još duže prosječno vrijeme izračuna, a posebno dugo vrijeme izračuna imaju divergencija i Jeffrey divergencija. Razlog za njihovo duže vrijeme izračuna je zbog očuvanja numeričke stabilnost odnosno da bi se spriječilo pojavljivanje logaritma nule na čemu se gubi puno vremena. Provedena mjerenja vremena su izvedena na Intel Core i5-2430M procesoru i Windows 7 32-bit operacijskom sustavu.

Zbog predočenih razloga, kao što su jednostavnost implementacije, brzina i numerička stabilnost, presjek histograma je odabran kao optimalna mjera za implementaciju i korištenje za usporedbu deskriptora vizualnih obilježja ravninskih segmenata u razmatranom sustavu za lokalizaciju mobilnog robota. Stoga će se u idućih nekoliko paragrafa analizirati presjek histograma. Na slici 5.31. mogu se vidjeti vrijednosti presjeka histograma za šest konfiguracija deskriptora HSV prostora boja u slučaju kada koristimo sve raspoložive binove za raspoznavanje razmatranih ravninskih segmenata.



Slika 5.31. Matrice sparivanja za HSV prostor boja pri rezolucijama od 32 a) i d),16 b) i e), 8 binova c) i f). Gornji red su standardni histogrami, a donji mekog tipa.

Na slici 5.31. može se razaznati da se smanjivanjem broja binova po dimenziji razmatranog prostora boja povećavaju iznosi mjere presjeka histograma za točne parove, ali i za one koji nisu točni. Razlog tome je, naravno, što će u tom slučaju binovi obuhvaćati veći dio prostora boja te će se time povećati svojstvo generaliziranja deskriptora, ali će se pri tome smanjiti i svojstvo diskriminacije. Drugim riječima, histogrami s manjim brojem binova će biti manje osjetljivi na promjene osvjetljenja okoline, ali će se smanjiti razlike između pojedinih tipova ravninskih segmenata. To je posebno vidljivo međusobnim sparivanjem segmenata koji uključuju pod sačinjen od pločica i segmenata koji uključuju crveni pod. Kada je rezolucija 32 bina po dimenziji, raspoznavanje segmenata s pločicama je na granici diskriminacije s obzirom na ostale moguće parove. Smanjivanjem broja binova po dimenziji se iznos presjeka histograma za točan par samo blago povećava, dok za netočne taj iznos raste puno brže. Da bi se donekle umanjio utjecaj smanjenja diskriminacije može se koristiti meke histograme. U donjem redu na slici 5.31. mogu se vidjeti vrijednosti presjeka histograma u slučaju korištenja mekih histograma za tri rezolucije HSV deskriptora boje. Općenito se može zaključiti da korištenje mekih histograma ne daje znatne razlike u iznosu presjeka histograma velikih rezolucija. Kod histograma manjih rezolucija korištenje mekih histograma dolazi više do izražaja što ima i smisla budući da okrupnjavanjem binova oni sadržavaju više prostora odnosno točaka, pa time te točke imaju i više utjecaja na okolne binove. Također se može primijetiti da za primjer kada se uspoređuju segmenti s pločicama i segmenti s crvenim podom, smanjivanjem broja binova i korištenjem mekih histograma znatnije se smanjuju iznosi presjeka histograma za netočne parove, a više ističu oni točni.

Uspoređivanje ravninskih segmenata deskriptorima boja može potencijalno biti dug proces zbog velikog broja binova koje u histogramu mogu imati neku vrijednost različitu od nule. Ako se pogleda vizualizacija histograma pojedinih prostora boja prikazanih u prethodnim poglavljima, može se primijetiti da je većina histograma prazna te uz relativno mali broj binova koji se ističu imamo puno binova s malim vrijednostima. Pri tome, binovi koji se ističu mogu se smatrati glavnim nositeljima informacija o boji pojedinog ravninskog segmenta, dok se ostali binovi tada ne smatraju relevantnima za razmatrani ravninski segment te ih se mogu zanemariti i u presjeku histograma. Stoga je otvorena mogućnost, da se prilikom presjeka histograma boje dva ravninska segmenta, uzimaju samo n najvećih binova. Na slici 5.32. može se vidjeti iznose presjeka histograma za HSV prostor boja za prethodno već razmatranih 10 ravninskih segmenata u slučaju parcijalnog presjeka histograma. Na slici 5.32. su prikazani iznosi parcijalnog presjeka histograma za samo 5 najvećih binova histograma koji se uspoređuju. Korištene su iste konfiguracije deskriptora kao i na slici 5.31. Na slici 5.32. može se vidjeti da se promatranjem samo pet najvećih binova histograma efikasno izbacuje iz razmatranja većina potencijalnih parova ravninskih segmenata, pogotovo u slučaju većih rezolucija histograma. Na žalost, također, u slučaju korištenja većih rezolucija, se time izbacuju i neki točni parovi za crno-bijeli pod i pod sačinjen od pločica. Smanjivanjem rezolucije se vraća točan par za pod od pločica, ali onaj za crno-bijeli pod se ne vraća. Razlog zašto ravninski segment crno-bijelog poda snimljen po danu i po noći nemaju ni jedan zajednički bin u najvećih pet je vidljiv na slici 5.20. koja vizualizira histograme za HSV prostor boja. Na slici 5.20. može se vidjeti da se promjenom osvjetljenja za ravninske segmente tipa crno-bijelog poda deskriptor drastično mijenja što onemogućava raspoznavanje korištenjem HSV prostora boja za ovakve tipove ravninskih segmenata pomoću parcijalnog presjeka histograma.



rezolucijama od 32 a) i d),16 b) i e), 8 binova c) i f). Gornji red su standardni histogrami, a donji mekog tipa.

6. Eksperimentalna analiza primjene vizualnih obilježja za lokalizaciju mobilnih robota

U ovom poglavlju bit će provedena analiza eksperimentalnih rezultata korištenja vizualnih obilježja ravninskih segmenata u svrhu lokalizacije mobilnog robota. Pokusi analizirani u ovom poglavlju provedeni su pomoću hibridne metričko-topološke karte koja je sačinjena od 854 čvora, odnosno lokacija podijeljenih na tri kata Elektrotehničkog fakulteta u Osijeku. Svaki čvor u karti predstavlja 3D model dijela prostora u kojem se robot kretao tokom postupka izgradnje karte. Dio prostora prikazan jednim čvorom predstavlja neposrednu okolinu robota snimljenu u trenutku kada je robot bio u određenom položaju na karti. Lokalni modeli prikazani čvorovima u karti međusobno su udaljeni za najmanje 500 mm ili se orijentacije robota iz kojih su snimljeni razlikuju za kut od 15°. Slike lokalnih modela koje su služile za formiranje karata tri kata snimljene su u noćnim uvjetima rada pod umjetnim osvjetljenjem radi postizanja što bolje dubinske slike. Testne sekvence slika korištene u ovom poglavlju sadržavale su 1237 slika u slučaju pokusa globalne lokalizacije i 1046 slika u slučaju pokusa inicijalne korespondencije. Testne slike su snimane u različitim vremenskim uvjetima, tipovima osvjetljenja i godišnjim dobima te su također raspoređene na sva tri kata iste zgrade Elektrotehničkog fakulteta kao i karta. Svaka referentna i testna slika je sadržavala RGB i dubinsku sliku te podatke odometrije robota.

Osnovna ideja provedenih analiza je dobiti performanse deskriptora vizualnih obilježja u slučaju raspoznavanja parova ravninskih segmenata te točnost globalne lokalizacije mobilnog robota pomoću referentnih podataka inkrementalne lokalizacije. Postupak inkrementalne lokalizacije provodi se pomoću prethodno stvorene hibridne karte i približnog trenutnog položaja mobilnog robota u odnosu na neki čvor karte. Tada sustav za lokalizaciju mobilnog robota uz postavljanje relativno male početne nesigurnosti zna u odnosu na koji lokalni model u karti da pronađe svoj relativni položaj, što čini izuzetno brzo i precizno [4]. Nakon tog početnog stadija, mobilni robot, poznavajući svoj originalni relativni položaj u odnosu na čvor u karti i uz pomoć vlastite odometrije, uspješno se lokalizira relativno u odnosnu na ostale lokalne čvorove u karti koji mu u pojedinom trenutku budu najbliži. Radi utvrđivanja točnosti inkrementalne lokalizacije, na Fakultetu Elektrotehnike i Računarstva u Zagrebu provedeno je testiranje razmatranog sustava [4], u kojem su rezultati inkrementalne lokalizacije postignuti razmatranim sustavom uspoređeni s podacima dobivenim pomoću laserskog mjerača udaljenosti i Monte-Carlo metode. Tim testiranjem je utvrđeno da razmatrani sustav lokalizacije postiže srednju vrijednost pogreške od 0.031m i 0.91°. Postignuta točnost se može smatrati dovoljnom da se podaci o lokaciji mobilnog robota prilikom provođenja eksperimenta inkrementalne lokalizacije mogu smatrati tzv. "ground truth" odnosno referentnim podacima za evaluaciju globalne lokalizacije. Dakle, pod pretpostavkom da je inkrementalnom lokalizacijom za svaku testnu sliku određen točan položaj robota na karti, parovi ravninskih segmenata na temelju kojih je taj položaj izračunat se također mogu smatrati ispravnim. Nadalje, u ovom poglavlju će se točnom lokacijom smatrati lokacija koja će u koordinatnom sustavu karte odstupati od one dobivene referentnim podacima inkrementalne lokalizacije za najviše 0.2m i 2°. Svi pokusi provedeni u ovom poglavlju izvršeni su pomoću Intel Core i5-2430M procesora na 2.4 GHz i Windows 7 32-bit operacijskog sustava.

U prvom podpoglavlju su prikazani i objašnjeni rezultati inicijalne korespondencije za implementirane deskriptore vizualnih obilježja. Na temelju te analize odabrana su dva optimalna deskriptora za svaki korišteni prostor boja te tri optimalna deskriptora zasnovana na značajkama teksture, čije će se performanse ispitati u kontekstu poboljšavanja rješenja problema globalne lokalizacije. Stoga će se u zadnjem podpoglavlju 6. poglavlja prikazati i analizirati rezultati eksperimentalnog ispitivanja problema globalne lokalizacije razmatranog sustava za lokalizaciju mobilnog robota.

6.1. Analiza rezultata inicijalne korespondencije

Kao što je objašnjeno u četvrtom poglavlju, proces lokalizacije započinje korakom inicijalne korespondencije, tj. stvaranjem parova ravninskih segmenata na temelju kojih se, u sljedećem koraku, generiraju hipoteze. U poglavlju 5.3.1. je objašnjeno da se u toj početnoj fazi procesa lokalizacije sličnost deskriptora vizualnih obilježja ravninskih segmenata koristi kao mjera selekcije prilikom sastavljanja popisa parova Q_{sort} , koji se potom koriste za generiranje hipoteza o trenutnom položaju mobilnog robota. U toj se inicijalnoj fazi rada sustava na popis Q_{sort} stavljaju samo oni parovi ravninskih segmenata kojima se deskriptori boja preklapaju u određenoj mjeri. Odabrana mjera odnosno prag sličnosti deskriptora vizualnih obilježja mora osigurati što veću propusnost točnih parova, a što manju propusnost onih neispravnih parova koji mogu uvesti smetnju i time otežati i usporiti rad sustava za lokalizaciju mobilnog robota. Utvrđivanje odgovarajućeg praga sličnosti nije trivijalno zbog evidentne sličnosti u deskriptorima boja za različite ravninske segmente, kao što možemo

vidjeti na matricama sparivanja u poglavlju 5.4.10., pa je za očekivati da mnogi neispravni parovi imaju veću vrijednost presjeka histograma od nekih ispravnih parova. Dakle, ne postoji vrijednost praga koja će sve točne parove odijeliti od netočnih, pa se mora naći odgovarajući kompromis, tj. optimalni prag.

Utvrđivanje praga sličnosti je ovdje provedeno analizom dva pokusa. Prvi pokus je proveden tako da se za svaki ravninski segment testnih slika odredi sličnost deskriptora naspram svakog ravninskog segmenta referentnih slika korištenjem referentnih podataka dobivenih inkrementalnom lokalizacijom koji omogućavaju određivanje ispravnih i neispravnih parova ravninskih segmenata. Podešavanje praga sličnosti deskriptora vizualnih obilježja je izvedeno tako da se postigne propusnosti od najmanje 90% točnih parova ravninskih segmenata. Naime, odbacivanje više od 10% točnih parova može značajno smanjiti funkcionalnost sustava onemogućujući stvaranje točnih hipoteza. Drugi pokus omogućava podešavanje vrijednosti praga sličnosti tako da se zadovolji određeni stupanj ukupne propusnosti svih mogućih parova ravninskih segmenata. Zajedno ta dva pokusa omogućavaju podešavanje optimalnog praga sličnosti koji će osigurati što veću propusnost točnih parova u proces generiranja hipoteza, a ograničiti broj neispravnih parova koje bi sustav morao inače uzeti u obzir, što će u konačnici rezultirati bržim i robusnijim sustavom za lokalizaciju mobilnog robota. Pouzdanost deskriptora vizualnih obilježja se pored samog izvora informacija na kojemu je zasnovan (boja, tekstura) može povezati i s brojem točaka koji se koristio za njegovo formiranje. Deskriptor vizualnih obilježja koji se formira za mali broj točaka RGB slike je ovisan o razmještaju tih točaka na ravninskom segmentu, pa se zbog toga ne može smatrati dovoljno pouzdanim. Stoga, za potrebe statističke analize i određivanja praga, u obzir su dolazili samo oni parovi ravninskih segmenata gdje oba člana u paru zadovoljavaju neke minimalne zahtjeve o broju pridruženih RGB točaka. U provedenim pokusima je zahtijevano da taj minimalni broj pridruženih točaka RGB slike bude 20.

Za analizu pokusa provedenog za utvrđivanje praga propusnosti točnih parova ravninskih segmenata korišteno je 242.899 parova segmenata od kojih su 11.426 točna para, dok je za utvrđivanje praga ukupne propusnosti korišteno 193.644.918 parova ravninskih segmenata. Na slici 6.1. može se vidjeti statistička analiza propusnosti točnih parova i ukupne propusnosti parova ravninskih segmenata za RGB deskriptore s obzirom na visinu praga. Analiziran je slučaj kada se uspoređuju svi binovi histograma te slučaj kada se uspoređuje samo 5 najpopunjenijih binova.



Slika 6.1. Statistička analiza učinkovitosti inicijalne korespondencije korištenjem deskriptora zasnovanih na RGB prostoru boja

Na dijagramima slike 6.1. može se očekivano vidjeti da propusnost točnih parova ravninskih segmenata monotono pada s porastom praga. U slučaju korištenja svih binova histograma za proračun presjeka histograma taj pad je približno linearan za skoro sve razmatrane deskriptore. U istom slučaju, također može se primijetiti da meki histogrami imaju veći stupanj propusnosti točnih parova od standardnih histograma. Kada se koristi samo 5 binova za proračun presjeka može se općenito primijetiti da deskriptori ovog prostora boja imaju značajno manji stupanj propusnosti točnih parova ravninskih segmenata, što je posebno istaknuto za deskriptore većih rezolucija. Također, u istom slučaju korištenja mekih histograma nema znatnije razlike u odnosu na standardne histograme. Postotak svih parova koji su prošli prag je primjetno manji u slučaju korištenja 5 binova nego u slučaju kada se koriste svi binovi. Konačno se može zaključiti da meki deskriptori i deskriptori manjih rezolucija imaju veću propusnost točnih parova zbog većih svojstava generaliziranja što uzrokuje i veću ukupnu propusnost svih mogućih parova. U svrhu testiranja sustava lokalizacije opisanog u poglavlju 6.2. kod ovog prostora boja su izabrani deskriptori "RGB32S" za slučaj korištenja svih binova i "RGB8" u slučaju korištenja 5 binova. "RGB32S" deskriptor postiže propusnost točnih parova od 93.37% uz ukupnu propusnost od 51.38% pri pragu 0, dok deskriptor "RGB8" postiže propusnost točnih parova od 85.46% uz ukupnu propusnost 40.61% također pri pragu 0.

Na slici 6.2. vidi se statistička analiza propusnosti točnih parova i ukupne propusnosti parova ravninskih segmenata za XYZ deskriptore s obzirom na visinu praga.



Slika 6.2. Statistička analiza učinkovitosti inicijalne korespondencije korištenjem deskriptora zasnovanih na CIE XYZ prostoru boja

Na dijagramima slike 6.2. može se primijetiti slične odnose kao i kod RGB prostora boja. Deskriptori mekog tipa i deskriptori manjih rezolucija imaju veću propusnost točnih parova, ali isto tako i veću ukupnu propusnost. Kod ovog prostora boja su za ispitivanje sustava lokalizacije korišteni deskriptori "XYZ8S" u slučaju korištenja svih binova i "XYZ8" u slučaju korištenja samo pet binova. Deskriptor "XYZ8S" ostvaruje propusnost točnih parova od 92.14% i ukupnu propusnost od 56.11% pri pragu od 0.1, dok deskriptor "XYZ8" postiže propusnost točnih parova od 89.19% i ukupnu propusnost od 49.58% pri pragu od 0.

Na slici 6.3. može se vidjeti statistička analiza propusnosti točnih parova i ukupne propusnosti parova ravninskih segmenata za HSV deskriptore s obzirom na visinu praga.



Slika 6.3. Statistička analiza učinkovitosti inicijalne korespondencije korištenjem deskriptora zasnovanih na HSV prostoru boja

HSV deskriptori su dvodimenzionalni deskriptori i zbog gubitaka dimenzije intenziteta svjetlosti postižu dobre rezultate poglavito u slučajevima promjene osvjetljenja scene. Slično kao i kod ostalih prostora boja, i ovdje u slučaju korištenja svih raspoloživih binova deskriptori mekog tipa i manjih rezolucija imaju bolje performanse. U slučaju korištenja samo pet binova deskriptori manjih rezolucija imaju znatno bolje performanse od onih s većim rezolucijama, a između mekih i standardnih histograma gotovo da i nema razlike. Za potrebe testiranja sustava lokalizacije kod ovog prostora boja je odabran "HSV8" i u slučaju korištenja svih binova i za samo pet binova. Deskriptor "HSV8" u slučaju korištenja svih binova od 90.61% uz ukupnu propusnost od 47.69% pri pragu od 0.1, dok u slučaju korištenja pet binova postiže propusnost točnih parova od 92.46% i ukupnu propusnost od 47.29% pri pragu 0.

Slika 6.4. prikazuje statističku analizu propusnosti točnih parova i ukupne propusnosti parova ravninskih segmenata za L*a*b* deskriptore s obzirom na visinu praga.



Slika 6.4. Statistička analiza učinkovitosti inicijalne korespondencije korištenjem deskriptora zasnovanih na CIE L*a*b* prostoru boja

L*a*b* deskriptori su također dvodimenzionalni, ali zbog svoje specifične trodimenzionalne konstrukcije i njihove projekcije na a*-b* ravninu relativno je mali broj binova histograma raspoloživo za isticanje razlike između boja posebno za deskriptore manjih rezolucija kao što je vidljivo na slici 5.17. Stoga ovi deskriptori imaju propusnost točnih parova iznad 90% i za velike iznose pragova, ali zbog toga je i ukupna propusnost na jednako visokoj razini. Korištenje svih raspoloživih binova ili samo pet nema značajnije razlike zbog već prije spomenutog malog broja raspoloženih binova. Značajniju razliku ima jedino deskriptor "Lab32S" koji smanjivanjem broja korištenih binova gubi bitne informacije. Kod ovog prostora boja su izabrani deskriptori "Lab32" u slučaju korištenja svih raspoloživih binova je odabran deskriptor "Lab8". Deskriptor "Lab32" postiže propusnost točnih parova od 91.7% uz ukupnu propusnost od 47.97% i prag od 0.2, dok deskriptor "Lab8" postiže propusnost točnih parova od 92.11% uz ukupnu propusnost od 55.67% i prag od 0.3.

Na grafovima slike 6.5. može se vidjeti statistička analiza propusnosti točnih parova i ukupne propusnosti parova ravninskih segmenata za RGC deskriptore s obzirom na visinu praga.



Slika 6.5. Statistička analiza učinkovitosti inicijalne korespondencije korištenjem deskriptora zasnovanih na RGC prostoru boja

Slično kao i ostali deskriptori opisani dvodimenzionalnim histogramima, i ovdje se postiže veća mogućnost generalizacije izbacivanjem jedne dimenzije i transformacijama kojima se donekle osigurava neovisnost o promjeni intenziteta osvjetljenja. Time se osigurava veći postotak propusnosti točnih parova, ali isto tako veća ukupna propusnost svih mogućih parova. Suprotno od drugih, kod ovog prostora boja korištenje mekih histograma postižu se bolji rezultati u slučajevima većih rezolucija nego kod manjih. Ovdje također nema većih razlika između korištenja svih binova ili samo pet, što sugerira da je distribucija točaka koncentrirana te da je i inače iskorišten mali broj raspoloživih binova. Kod ovog prostora boja odabran je deskriptor "RGC16S" u slučaju korištenja svih binova i deskriptor "RGC16" u slučaju korištenja samo pet binova. Deskriptor "RGC16S" postiže propusnost točnih parova od 92.46% uz ukupnu propusnost od 53.40% pri pragu od 0.4, dok deskriptor "RGC16" postiže propusnost točnih parova od 90.41% uz ukupnu propusnost od 54.71% pri pragu od 0.2.

Na slici 6.6. može se vidjeti statistička analiza propusnosti točnih parova i ukupne propusnosti parova ravninskih segmenata za deskriptore prostora suprotstavljenih boja s obzirom na visinu praga.



Slika 6.6. Statistička analiza učinkovitosti inicijalne korespondencije korištenjem deskriptora zasnovanih na prostoru suprotstavljenih boja

Ovaj prostor boja, slično kao i L*a*b* prostor boja, pokazuje velike razlike između korištenja mekih i standardnih histograma. Korištenje mekih histograma i histograma manjih rezolucija bitno povećava svojstvo generaliziranja, što uzrokuje veliku ukupnu propusnost svih mogućih parova ravninskih segmenata. Ovdje također nema znatne razlike u propusnosti točnih parova i ukupne propusnosti između korištenja svih binova i samo pet binova. Ovdje je izabran deskriptor "Opp16S" i u slučaju korištenja svih binova i za slučaj korištenja samo pet binova. Taj deskriptor postiže propusnost točnih parova od 92.04% uz ukupnu propusnost 47.58% pri pragu od 0.5 kada se koriste svi binovi, a u slučaju korištenja samo pet binova postiže propusnost točnih parova od 94.42% i ukupnu propusnost od 57.12% pri pragu od 0.4.

Slika 6.7. prikazuje statističku analizu propusnosti točnih parova i ukupne propusnosti parova ravninskih segmenata za LBP porodicu deskriptora s obzirom na visinu praga.



Slika 6.7. Statistička analiza učinkovitosti inicijalne korespondencije korištenjem deskriptora zasnovanih na LBP porodici deskriptora teksture

Za stvaranje deskriptora iz LBP obitelji deskriptora svaka točka nekog ravninskog segmenta se analizira promatranjem njene okoline. Uz to što zbog greške reprojekcije, koja je objašnjena u poglavlju 5.2., smanjujemo broj točaka koji možemo analizirati, zbog zahtjeva za obradom okoline svake točke taj broj točaka se ovdje još dodatno smanjuje. Povećanje područja isključivanja točaka uz granicu ravninskih segmenata odgovara iznosu radijusa okoline koja se promatra za svaku točku, što u konačnici rezultira s manjim brojem ravninskih segmenata koji mogu zadovoljavati unaprijed zadanu normu od 20 pridruženih točaka. Na grafovima na slici 6.7. može se razaznati da deskriptori tipa $LBP_{P,R}^{riu_2}$ imaju najveću propusnost točnih parova, ali isto tako i najveću ukupnu propusnost svih mogućih parova ravninskih segmenata, dok deskriptori tipa $LBP_{P,R}^{riu^2} / VAR_{P,R}$ imaju najveći pad propusnosti točnih parova i ukupne propusnosti s porastom vrijednosti praga. Budući da je pad ukupne propusnosti mnogo veći od pada propusnosti točnih parova, deskriptori tipa $LBP_{P,R}^{riu^2} / VAR_{P,R}$ se nameću kao optimalni izbor iz ove porodice deskriptora. Za potrebe dodatnog testiranja u sustavu za lokalizaciju mobilnog robota je odabran optimalni deskriptor od svakog tipa iz ove porodice deskriptora teksture, pa su stoga odabrani: $LBP_{16,2}$, $LBP_{24,3}^{riu_2}$ i $LBP_{24,3}^{riu_2}/VAR_{24,3}$. Deskriptor LBP₁₆₂ postiže propusnost točnih parova od 90.23% i ukupnu propusnost od 61.10% pri pragu od 0.4, deskriptor *LBP*^{*riu*2}_{24,3} postiže propusnost točnih parova od 94.19% i ukupnu propusnost od 72.44% pri pragu od 0.5, dok deskriptor LBP^{riu2}/VAR_{24.3} postiže propusnost točnih parova od 92.53% uz pripadajuću ukupnu propusnost od 59.8% pri pragu od 0.2.

U tablici 6.1. može se vidjeti konačan popis odabranih deskriptora s pripadajućim pragovima koji će se koristiti kod inicijalnog sparivanja ravninskih segmenata u pokusu

provedenom u poglavlju 6.2. Pragovi su djelomično promijenjeni u odnosu na one napisane u prethodnom tekstu radi postizanja najboljeg omjera propusnosti točnih parova i ukupne propusnosti koji odabrani deskriptori mogu postići. Maksimalni broj binova također spomenut u tablici 6.1. govori kolika je maksimalna dimenzionalnost deskriptora, koja ovisi o korištenom prostoru boja i broju binova po korištenoj dimenziji prostora boja. Stvarni broj korištenih binova za presjek histograma će uglavnom biti mnogo manji, budući da će se koristiti samo oni binovi koji imaju vrijednost iznad 0.

Tablica 6.1. Popis deskriptora i njihovih pragova koji će se koristiti za evaluaciju globalne lokalizacije

Deskriptor	Maksimalni broj binova	Prag
RGB32S	32768	0.02
RGB8	5	0.02
XYZ8S	512	0.1
XYZ8	5	0.02
HSV8	64	0.1
HSV8	5	0.02
Lab32	1024	0.2
Lab8	5	0.3
RGC16S	256	0.4
RGC16	5	0.2
Opp16S	256	0.5
Opp16S	5	0.45
LBP16	128	0.4
LBP RIU 24	26	0.55
LBP RIU VAR 24	442	0.21

6.2. Analiza rezultata globalne lokalizacije

Cilj globalne lokalizacije je da mobilni robot na temelju trenutno snimljene RGB-D slike bez ikakvog znanja o prethodnom položaju uspije odrediti svoj položaj na karti. U nekim slučajevima neposredna okolina robota ne pruža dovoljno informacija da se robot točno lokalizira, što znači da robot neće istog trenutka naći svoj točan položaj, već da će se kroz određeno vrijeme istraživanja svoje okoline uspjeti točno lokalizirati. Mobilni robot prilikom globalne lokalizacije obilazi sve čvorove u karti i za svaki čvor generira niz hipoteza na način opisan u poglavlju 4.4.3. uz veliku početnu nesigurnost svojeg položaja. Za očekivati je da najviše rangirana hipoteza koju će sustav generirati nije uvijek točna, ali isto tako možemo očekivati da će se točna hipoteza nalaziti u prvih *N* hipoteza sortiranih po kriteriju objašnjenom u poglavlju 4.5. U tom slučaju potrebno je samo pratiti i ažurirati *N* najviše rangiranih hipoteza. Sličan način rada je i u sustavima zasnovanima na Monte Carlo metodama odnosno čestičnim filtrima. Određivanje točnog broja *N*, koji govori koliko

hipoteza bi trebali pratiti da bi se robot u konačnici točno lokalizirao, se uglavnom može samo eksperimentalnim putem utvrditi. Što je broj N veći, veći su i zahtjevi za računalnim vremenom potrebnim za praćenje N hipoteza. Stoga se učinkovitost algoritma globalne lokalizacije može mjeriti upravo najmanjim brojem N unutar kojega se s velikom sigurnošću nalazi točna hipoteza.

Usporedbom položaja svake generirane hipoteze s njoj odgovarajućim referentnim položajem dobivenim inkrementalnom lokalizacijom može se ocijeniti točnost hipoteze. Hipoteza se smatra točnom prema kriteriju navedenom u uvodnom dijelu 6. poglavlja. Među generiranim hipotezama može biti više točnih. Svrha primjene deskriptora vizualnih obilježja ravninskih segmenata je, kao što je objašnjeno u poglavlju 5.3., da se smanji broj hipoteza među kojima se nalazi točna hipoteza te time ubrza rad sustava. Dakle, kao mjera uspješnosti globalne lokalizacije uzet je redni broj prve točne hipoteze u popisu hipoteza sortiranom prema kriteriju (4.35) kod osnovnog sustava odnosno sortiranom prema kriteriju (5.19) kod proširenog sustava. Taj redni broj se u nastavku naziva *indeks* hipoteze. U nastavku je prikazana usporedbena evaluacija osnovnog sustava za globalnu lokalizaciju, koji koristi samo geometriju ravninskih segmenata, i proširenog sustava za lokalizaciju mobilnog robota, koji uzima u obzir i vizualna obilježja ravninskih segmenata. Primarni interes evaluacije je procijeniti poboljšanje sustava postignuto upotrebom vizualnih obilježja ravninskih segmenata s obzirom na vremensko ubrzanje rada sustava te ostvarenje što manjeg indeksa prve točne hipoteze.

Korištenje proširenog sustava za lokalizaciju mobilnog robota je ispitano za deskriptore vizualnih obilježja i pripadajuće pragove navedene u tablici 6.1. Također, pokusi globalne lokalizacije za korištene deskriptore provedeni su za dva slučaja:

- prilikom inicijalnog sparivanja parova ravninskih segmenata proračunava se vrijednost presjeka njihovih deskriptora i uspoređuje se s iznosom postavljenog praga bez obzira koliko je točaka RGB slike bilo korišteno za formiranje njihovih deskriptora,
- u slučaju ako je broj točaka korišten za formiranje deskriptora manji od 20 za bilo koji element para ravninskih segmenata koji se pokušavaju inicijalno spariti, taj par se bez uspoređivanja propušta dalje na inicijalno sparivanje zasnovano na geometrijskim ograničenjima. Ovi slučajevi su označeni s (*) poslije naziva deskriptora u tablicama 6.2. i 6.3.

U slučaju ako se deskriptor vizualnih obilježja formira za manje od 20 točaka, takav deskriptor se ne može smatrati dovoljno pouzdanim, mada, bez obzira na nedostatak informacije o vizualnim obilježjima, takav ravninski segment može nositi presudnu geometrijsku informaciju potrebnu za lokaliziranje robota s obzirom na zadnji stupanj slobode na način objašnjen u poglavlju 4.4.3. S druge strane, propuštanje svih parova ravninskih segmenata čiji barem jedan segment ima manje od 20 točaka korištenih za formiranje deskriptora znatno povećava ukupni broj generiranih hipoteza, što zahtjeva osjetno više računskog vremena nego kada se ti parovi ne obrađuju. U tablici 6.2. mogu se vidjeti rezultati mjerenja provedenih pokusa globalne lokalizacije.

	Maks.		Prva	Nema	Nema	Srednje	Srednje	Srednji broj	Srednja vrijednost
Deskriptor	broj	Hip.	hip.	točne	zadnji	vrijeme	vrijeme	generiranih	indeksa prve točne
	binova		točna	hip.	DOF	lokalizacije[s]	obrade[s]	hipoteza	hipoteze
Geometrija	/	1162	470	75	0	1,062	0,050	3636,6	39,3
RGB32S	32768	1115	578	95	27	3,336	0,211	1834,8	14,8
RGB32S (*)	32768	1147	577	89	1	3,309	0,211	2564,0	19,5
RGB8	5	1096	528	116	25	0,427	0,071	1652,6	17,2
RGB8 (*)	5	1138	537	97	2	0,895	0,071	2275,3	22,2
XYZ8S	512	1128	515	83	26	0,915	0,097	2226,2	25,1
XYZ8S (*)	512	1150	516	87	0	1,221	0,097	2808,9	29,4
XYZ8	5	1122	528	90	25	0,539	0,072	1942,8	23,2
XYZ8 (*)	5	1150	529	86	1	0,947	0,072	2546,1	25,5
HSV8	64	1144	588	69	24	0,680	0,073	2075,0	19,0
HSV8 (*)	64	1162	576	75	0	1,049	0,073	2699,9	20,2
HSV8	5	1127	546	84	26	0,497	0,073	1908,3	19,3
HSV8 (*)	5	1153	546	84	0	0,898	0,073	2518,6	20,8
Lab32	1024	1134	569	77	26	0,671	0,073	1946,5	19,1
Lab32 (*)	1024	1156	563	80	1	1,042	0,073	2528,8	21,3
Lab8	5	1129	556	84	24	0,636	0,070	2126,7	21,4
Lab8 (*)	5	1154	543	81	2	0,996	0,070	2671,1	26,6
RGC16S	256	1113	546	95	29	0,717	0,081	2119,3	22,6
RGC16S (*)	256	1148	551	89	0	1,046	0,081	2612,2	25,3
RGC16	5	1122	513	87	28	0,643	0,071	2208,4	30,1
RGC16 (*)	5	1153	525	84	0	0,940	0,071	2666,0	36,3
Opp16S	256	1126	560	85	26	0,613	0,081	1875,5	20,1
Opp16S (*)	256	1149	559	88	0	0,954	0,081	2369,7	23,0
Opp16S	5	1126	534	84	27	0,565	0,081	1980,2	23,7
Opp16S (*)	5	1149	531	88	0	0,889	0,081	2452,9	23,0
LBP16	128	1099	500	106	32	0,810	0,175	2480,4	29,8
LBP16 (*)	128	1151	518	85	1	1,253	0,175	3422,2	33,3
LBP RIU 24	26	1042	468	155	40	0,604	0,223	2367,5	30,9
LBP RIU 24 (*)	26	1133	482	104	0	1,099	0,223	3422,3	40,3
LBP RIU VAR 24	442	1089	504	101	47	0,722	0,252	1974,1	27,2
LBP RIU VAR 24 (*)	442	1156	518	81	0	1,271	0,252	3228,8	38,3

Tablica 6.2. Rezultati mjerenja pokusa globalne lokalizacije

Prethodna tablica sadrži slijedeće stupce:

 "Deskriptor" – Navodi koji se deskriptor vizualnih obilježja koristio ili se radilo o sustavu koji koristi samo geometrijske podatke. Svaki naziv deskriptora vizualnih obilježja sadrži akronim korištenog prostora boja i broj binova po korištenoj dimenziji prostora boja. Ako se radi o deskriptoru s histogramom mekog tipa, ima dodanu oznaku 'S', dok se oznaka (*) koristi u slučaju ako se u tom pokusu propuštalo parove kojima su deskriptori formirani s manje od 20 točaka. U slučaju deskriptora teksture navodi se korištena varijanta LBP deskriptora te broj razmatranih točaka u okolini.

- "Maks. broj binova" Navodi koji je maksimalni broj binova koji bi deskriptor mogao imati što sugerira o njegovoj kompleksnosti i hardverskim zahtjevima.
- "Hip." Broj regularnih scena, odnosno onih za koje je sustav uspio generirati barem jednu točnu hipotezu prema zadanom kriteriju (200mm i 2°) i pronaći ju na sortiranom popisu.
- "Prva hip. točna" Broj scena za koje je točna hipoteza bila na samom vrhu sortiranog popisa hipoteza.
- "Nema točne hip." Predstavlja broj onih scena za koje sustav nije uspio proizvesti ni jednu točnu hipotezu, a scena ima dovoljno informacija u svim razmatranim stupnjevima slobode. Razlog za postojanje takvih slučajeva jest da Kalmanov filter s obzirom na zadanu početnu nesigurnost nije mogao generirati ni jednu ispravnu pretpostavku položaja i orijentacije s obzirom na zadano maksimalno odstupanje od 200mm i 2° od točnog položaja.
- "Nema zadnji DOF" Broj onih scena za koje sustav nije mogao proizvesti ni jednu hipotezu budući da scena nema dovoljno informacija za procjenu zadnjeg stupnja slobode. Zajedno s prethodnim stupcem čini ukupan broj scena za koje sustav nije mogao naći nikakvu točnu pretpostavku lokacije.
- "Srednje vrijeme lokalizacije[s]" Srednje vrijeme potrebno da se prođe kroz fazu rada sustava koji provodi samo algoritam lokalizacije opisan u poglavljima 4.4 i 4.5.
- "Srednje vrijeme obrade[s]" Srednje vrijeme generiranja 3D modela scene koje uključuje: triangulaciju, segmentaciju na ravninske segmente, predobradu RGB slike koje zahtjeva razmatrani deskriptor, stvaranje maske za točke slike koje se mogu iskoristiti s obzirom na granice segmenata i sam proces stvaranja deskriptora vizualnih obilježja za svaki detektirani ravninski segment na sceni. Kada se koriste samo geometrijska ograničenja za proces lokalizacije onda je ovdje razmatrano samo vrijeme triangulacije i segmentacije.
- "Srednji broj generiranih hipoteza" Srednja vrijednost broja hipoteza koje je sustav generirao u procesu lokalizacije.

 "Srednja vrijednost indeksa prve točne hipoteze" – Navodi koja ja bila srednja vrijednost indeksa prve točne hipoteze na sortiranom popisu hipoteza.

Na temelju rezultata prikazanih tablicom 6.2., može se zaključiti da većina deskriptora vizualnih obilježja pozitivno pridonosi radu sustava, što je vidljivo u smanjenom broju generiranih hipoteza, što rezultira i vremenski kraćom fazom lokalizacije mobilnog robota. Također se općenito može primijetiti da se vrijeme obrade scene znatno povećava zbog svih dodatnih radnji koje se moraju izvršiti kada se koriste deskriptori vizualnih obilježja, ali bez obzira na to, konačno vrijeme rada sustava je uglavnom manje zbog ušteda u fazi lokalizacije. Smanjenje vremena faze lokalizacije je vidljivo manje u slučajevima kada se ne ispituje sličnost parova gdje je barem jedan formiran s manje od 20 točaka, već ih se propušta dalje u proces geometrijskog sparivanja, ali time se smanjuje i broj scena za koje se ne može generirati ni jedna hipoteza što je također vidljivo u tablici. Na tablici 6.3. je lakše uočiti relativne i apsolutne promjene performansi sustava koji koristi vizualna obilježja ravninskih segmenata u usporedbi s sustavom koji koristi samo informacije o geometriji.

Stupci koji su u tablici 6.3. navedeni kao relativne promjene opisuju kolika je postotna promjena nekog mjerenja za sustav koji koristi deskriptore vizualnih obilježja kada se usporedi s odgovarajućim mjerenjem sustava zasnovanog samo na geometriji. Tablica 6.3. također sadrži dva stupca koji opisuju apsolutne promjene. Stupac tablice 6.3. koji je označen " Δ Hip.(abs.) [%]" opisuje apsolutnu promjenu broja scena za koje se uspjela barem jedna točna hipoteza generirati u odnosu na ukupan broj testnih scena. Drugi stupac koji je naveden kao apsolutna promjena je "Prva hip. točna(abs.) [%]" te opisuje apsolutnu promjenu broja scena kojoj je točna hipoteza bila na samom vrhu sortiranog popisa hipoteza u odnosu na ukupan broj testnih scena. S brojem 5 unutar zagrada označeni su oni slučajevi kada se koristi samo 5 binova za presjek histograma. U tablici 6.3. zelenom bojom su označene vrijednosti koji predstavljaju poboljšanje performansi u odnosu na sustav zasnovan samo na geometriji, dok su crvenom bojom označena pogoršanja performansi.
Deskriptor	∆Hip.(rel.) [%]	ΔHip.(aps.) [%]	Prva hip. točna (rel.) [%]	Prva hip. točna (abs.) [%]	Relativna promjena vremena lokalizacije [%]	Relativna promjena vremena obrade [%]	Relativna promjena ukupnog vremena [%]	Relativna promjena broja generiranih hipoteza [%]	Relativna promjena indeksa prve točne hipoteze [%]
Geometrija	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
RGB32S	-4,04	-3,60	22,98	8,26	214,07	323,75	218,98	-49,55	-62,35
RGB32S (*)	-1,29	-1,15	22,77	8,19	211,55	323,75	216,56	-29,49	-50,39
RGB8 (5)	-5,68	-5,05	12,34	4,44	-59,81	43,19	-55,20	-54,56	-56,27
RGB8 (5*)	-2,07	-1,84	14,26	5,13	-15,72	43,19	-13,08	-37,43	-43,67
XYZ8S	-2,93	-2,60	9,57	3,44	-13,80	95,23	-8,92	-38,78	-36,19
XYZ8S (*)	-1,03	-0,92	9,79	3,52	14,94	95,23	18,53	-22,76	-25,21
XYZ8 (5)	-3,44	-3,06	12,34	4,44	-49,21	44,98	-45,00	-46,58	-41,01
XYZ8 (5*)	-1,03	-0,92	12,55	4,51	-10,79	44,98	-8,29	-29,99	-35,09
HSV8	-1,55	-1,38	25,11	9,03	-35,98	46,71	-32,29	-42,94	-51,76
HSV8 (*)	0,00	0,00	22,55	8,11	-1,21	46,71	0,93	-25,76	-48,56
HSV8 (5)	-3,01	-2,68	16,17	5,81	-53,23	46,71	-48,76	-47,53	-50,80
HSV8 (5*)	-0,77	-0,69	16,17	5,81	-15,46	46,71	-12,68	-30,74	-47,12
Lab32	-2,41	-2,14	21,06	7,57	-36,81	46,81	-33,07	-46,48	-51,52
Lab32 (*)	-0,52	-0,46	19,79	7,12	-1,86	46,81	0,32	-30,46	-45,94
Lab8 (5)	-2,84	-2,52	18,30	6,58	-40,15	40,21	-36,56	-41,52	-45,71
Lab8 (5*)	-0,69	-0,61	15,53	5,59	-6,21	40,21	-4,14	-26,55	-32,45
RGC16S	-4,22	-3,75	16,17	5,81	-32,52	62,95	-28,25	-41,72	-42,62
RGC16S (*)	-1,20	-1,07	17,23	6,20	-1,55	62,95	1,33	-28,17	-35,63
RGC16 (5)	-3,44	-3,06	9,15	3,29	-39,48	41,86	-35,85	-39,27	-23,50
RGC16 (5*)	-0,77	-0,69	11,70	4,21	-11,48	41,86	-9,09	-26,69	-7,73
Opp16S	-3,10	-2,75	19,15	6,89	-42,33	62,71	-37,63	-48,43	-48,81
Opp16S (*)	-1,12	-0,99	18,94	6,81	-10,19	62,71	-6,93	-34,84	-41,55
Opp16S (5)	-3,10	-2,75	13,62	4,90	-46,81	62,71	-41,92	-45,55	-39,65
Opp16S (5*)	-1,12	-0,99	12,98	4,67	-16,26	62,71	-12,73	-32,55	-41,56
LBP16	-5,42	-4,82	6,38	2,30	-23,77	252,92	-11,40	-31,79	-24,34
LBP16 (*)	-0,95	-0,84	10,21	3,67	17,98	252,40	28,47	-5,90	-15,24
LBP RIU 24	-10,33	-9,18	-0,43	-0,15	-43,14	349,26	-25,59	-34,90	-21,44
LBP RIU 24 (*)	-2,50	-2,22	2,55	0,92	3,45	349,08	18,90	-5,89	2,56
LBP RIU VAR 24	-6,28	-5,59	7,23	2,60	-32,06	405,95	-12,47	-45,72	-30,97
LBP RIU VAR 24 (*)	-0,52	-0,46	10,21	3,67	19,65	405,99	36,93	-11,21	-2,73

Tablica 6.3. Relativne i apsolutne promjene performansi sustava lokalizacije koji koristi i deskriptore vizualnih obilježja u usporedbi sa sustavom koji je zasnovan samo na geometriji

Deskriptori zasnovani na RGB prostoru boja su se pokazali relativno prikladnim deskriptorom za razlikovanje ravninskih segmenata, što je dovelo do značajnog povećanja broja scena čija je točna hipoteza bila na vrhu sortiranog popisa hipoteza, posebno u slučaju korištenja RGB32S deskriptora gdje je povećanje iznosilo iznad 22%. Nažalost, zbog povećane kompleksnosti RGB32S deskriptora, budući da je mekog tipa i ima potencijalno mnogo binova, vrijeme lokalizacije, a time i ukupno vrijeme rada sustava se povećava i za više od 200%, iako se broj generiranih hipoteza drastično smanjuje. Korištenjem manjeg RGB8 deskriptora vremena lokalizacije i ukupna vremena rada sustava su mnogo manja naročito u slučaju kad se ispituju sličnosti svih parova bez obzira na broj točaka od kojih su formirani. Međutim, manji deskriptor isto tako ima i manju moć razlikovanja što rezultira

manjim poboljšanjem prosječnog položaja prvog točnog indeksa u odnosu na veći RGB32S deskriptor. Također, korištenje ovih deskriptora u svim slučajevima uzrokuje smanjenje broja scena za koje sustav generira barem jednu točnu hipotezu. Smanjenje je od 1.29% kod RGB32S(*) do 5.68% kod RGB8 deskriptora.

Korištenje XYZ prostora boja u deskriptoru pokazuje nešto skromnije povećanje brzine rada sustava naspram RGB prostora boja. U slučajevima kada se propuštaju parovi ravninskih segmenata s manje od 20 točaka za deskriptor XYZ8S, vrijeme faze lokalizacije i ukupno vrijeme je lošije i od sustava zasnovanog samo na geometriji. Općenito ovi deskriptori uspiju za veći broj scena generirati barem jednu točnu hipotezu za razliku od RGB deskriptora, pa je ovdje smanjenje broja scena u najboljem slučaju oko 1.03%, a u najgorem 3.44%. Budući da deskriptor XYZ8 ima značajnije smanjenje broja generiranih hipoteza, a time i vremena faze lokalizacije uz u određenoj mjeri prihvatljivi gubitak scena za koje uspije generirati točne hipoteze, prikazat će se na slici 6.8. usporedno kumulativni histogrami mjerenja sustava koji koristi taj deskriptor naspram sustava koji je zasnovan samo na geometriji.



Slika 6.8. Kumulativni histogrami mjerenja za XYZ8 deskriptor u slučaju korištenja samo 5 binova

Kumulativni histogram daje podatak o postotku mjerenja koja su manja ili jednaka određenoj vrijednosti mjerenja, x-os prikazuje vrijednosti mjerenja, a y-os postotak mjerenja. Gornja dva dijagrama slike 6.8. pokazuju kumulative histograme indeksa prve točne hipoteze na sortiranom popisu hipoteza, lijevi dijagram prikazuje cijeli raspon izmjerenih indeksa, dok desni dijagram prikazuje uvećani dio lijevog dijagrama za prvih 100 indeksa. Na dijagramima slike 6.8. može se vidjeti da korištenjem XYZ8 deskriptora zajedno s geometrijskim značajkama dovodi do značajnijeg smanjenja vremena faze lokalizacije sustava, dok se indeks prve točne hipoteze na sortiranom popisu također smanjuje. Vrhunac krivulja na slici 6.8. je normiran s obzirom na broj scena za koje je sustav zasnovan samo na geometriji uspio generirati barem jednu točnu hipotezu.

Deskriptori zasnovani na HSV prostoru boja pokazuju svoja dobra svojstva razlikovanja ravninskih segmenata, što rezultira znatnim povećanjem broja scena za koje je prva hipoteza točna čak i za 25.11% u slučaju korištenja svih binova HSV8 deskriptora. U slučaju korištenja ovih deskriptora smanjenje broja scena za koje sustav ne uspije generirati ni jednu točnu hipotezu je prosječno manji nego za prethodna dva prostora boja. U slučaju propuštanja parova ravninskih segmenata koji sadržavaju manje od 20 točaka i korištenja svih binova HSV8 deskriptora postiže se jednak broj scena za koje se ne uspije generirati ni jedna točna hipoteza kao i kada se koristi samo geometrija. Za isti slučaj postiže se 22.55% više scena za koje je prva hipoteza točna, dok je prosječni indeks prve točne hipoteze manji za 48.56% uz prosječno povećanje ukupnog vremena rada sustava za 0.93% u odnosu na sustav zasnovan samo na geometriji, time se ovaj deskriptor može smatrati najboljim od ispitanih ako se želi zadržati ista razina scena za koju se ne može generirati ni jedna hipoteza. Na dijagramima slike 6.9. može se vidjeti kumulativni histogrami mjerenja sustava u slučaju korištenja samo geometrije i kombinacije s HSV8 deskriptorom kada se koriste svi njegovi binovi.



Slika 6.9. Kumulativni histogrami mjerenja za HSV8 deskriptor u slučaju korištenja svih raspoloživih binova

Ako se koristi samo 5 najvećih binova ovog deskriptora smanjuje se broj scena za koje sustav uspije generirati barem jednu točnu hipotezu, ali se pri tome postiže i značajno ubrzanje faze lokalizacije sustava koja je i u najsporijem slučaju brža za 15.46% u odnosu na sustav koji koristi samo geometriju. Na slici 6.10. mogu se vidjeti dijagrami kumulativnih histograma za sustav s HSV8 deskriptorom kada koristimo samo 5 najvećih binova za proračun sličnosti.



Slika 6.10. Kumulativni histogrami mjerenja za HSV8 deskriptor u slučaju korištenja samo 5 binova

Na dijagramima slike 6.10. može se primijetiti da, kada se koristi samo 5 binova HSV8 deskriptora i kada propuštamo parove ravninskih segmenata koji imaju manje od 20 točaka, postiže se značajna ušteda vremena u fazi lokalizacije uz vrlo malo povećanje broja scena za koje se ne može generirati ni jedna točna hipoteza od 0.77%. Stoga se ovaj deskriptor i u slučaju korištenja samo 5 binova može smatrati optimalnim izborom ako je dopušteno barem malo pogoršanje u smislu broja scena za koje se ne može generirati ni jedna točna za koje se ne može generirati ni jedna točaka, postiže se ne može generirati ni jedna točna hipoteza od 0.77%.

Odabrani deskriptori zasnovani na CIE L*a*b* prostoru boja postižu usporedive rezultate s onima zasnovanima na HSV prostoru boja te imaju prosječno samo malo veće povećanje broja scena za koje se ne može generirati ni jedna točna hipoteza. Stoga odabrani deskriptori zasnovani na ovom prostoru boja također spadaju u one dobre izbore za korištenje u sustavima za lokalizaciju mobilnog robota. Na dijagramima slike 6.11. može se vidjeti kumulativni histogrami mjerenja za Lab32 deskriptor.



Slika 6.11. Kumulativni histogrami mjerenja za Lab32 deskriptor u slučaju korištenja svih raspoloživih binova

Na slici 6.11. može se primijetiti da su performanse Lab32 deskriptora dosta slične onima HSV8 kada se koriste svi binovi, ali budući da HSV8 deskriptor ima jednak broj scena za koje se ne može generirati barem jedna točna hipoteza kao i sustav zasnovan samo na geometriji, tada on uzima primat u odnosu na Lab32 deskriptor. Deskriptor Lab8 ima malo lošije performanse od Lab32 deskriptora odnosno ima manji broj scena za koje je prva hipoteza točna i veći broj scena za koje se ne može generirati ni jedna točna hipoteza. Međutim, po istim tim mjerenjima Lab8 deskriptor postiže usporedne rezultate s HSV8 deskriptorom uz malo manju uštedu vremena u fazi lokalizacije što ga čini također zanimljivim. Na dijagramima slike 6.12. mogu se vidjeti kumulativni histogrami mjerenja za Lab8 deskriptor.



Slika 6.12. Kumulativni histogrami mjerenja za Lab8 deskriptor u slučaju korištenja samo 5 binova

Kromatski rgb prostor boja se pokazao lošijim izborom za stvaranje deskriptora, jer, kada se uspoređuje s ostalim prostorima boja, deskriptori zasnovani na ovom prostoru boja uglavnom imaju veći broj scena za koje se ne može generirati barem jedna točna hipoteza izuzev RGB prostora boja od kojeg je samo malo bolji te manji broj onih scena za koje je prva hipoteza točna izuzev XYZ prostora boja od kojeg je u tom mjerenju malo bolji. Također, u usporedbi s ostalim prostorima boja, imaju prosječno najslabije poboljšanje indeksa prve točne hipoteze na sortiranom popisu hipoteza te su ti rezultati usporedivi s korištenim deskriptorima iz LBP porodice deskriptora teksture.

Deskriptori zasnovani na prostoru suprotstavljenih boja se, ukupno gledajući, mogu opisati kao prosjek svih korištenih deskriptora vizualnih obilježja zasnovanih na prostorima boja. Kod korištenja Opp16S deskriptora nema nikakve razlike u promjeni broja scena za koje se ne može generirati ni jedna točna hipoteza bez obzira da li se koriste svi binovi ili samo 5 najvećih, dok broj scena za koje je prva hipoteza točna manji ako se koristi samo 5 binova. Također, i ovdje u slučaju kada se propuštaju parovi koji sadrže manje od 20 točaka značajno se povećava vrijeme faze lokalizacije nego kada se i za te parove uspoređuje sličnost

deskriptora. Poboljšanje prosječnog indeksa prve točne hipoteze malo je manje, ali svejedno dosta slično kao kod HSV i Lab deskriptora.

Odabrani deskriptori zasnovani na LBP porodici deskriptora teksture su po svojim performansama uglavnom puno lošiji od deskriptora zasnovanih na prostorima boja. To se najviše odnosi na $LBP_{PR}^{riu^2}$ deskriptore koji sustavu za lokalizaciju mobilnog robota unose pogoršanje u skoro sva mjerenja od kojih se najviše ističe povećanje broja scena za koje se ne može generirati ni jedna točna hipoteza. Pogoršanje iznosi i 10% ako se ne gleda da li deskriptori imaju manje od 20 pridruženih točaka prilikom usporedbe deskriptora. Dodavanje podataka o varijanci i stvaranjem $LBP_{P,R}^{riu^2}/VAR_{P,R}$ deskriptora ne poboljšava puno performanse, već uglavnom samo povećava vrijeme potrebno na stvaranje deskriptora, što dodatno opterećuje i usporava sustav. Optimalni izbor kod korištenih deskriptora teksture se pokazao *LBP_{P,R}* deskriptor, budući da je prosječno najmanje vremena bilo utrošeno za njegovo stvaranje te ima relativno malo povećanje broja scena za koje se ne može generirati ni jedna točna hipoteza uz postizanje umjerenog poboljšanja srednjeg indeksa prve točne hipoteze. Ovaj deskriptor ima i umjereno povećanje broja scena za koje je prva hipoteza točna kada se propuštaju parovi ravninskih segmenata koji imaju manje od 20 točaka. Općenito gledajući korišteni deskriptori tekstura ne pomažu dovoljno radu sustava za lokalizaciju mobilnog robota, već ga uglavnom samo opterećuju, te ih stoga za ovu primjenu treba izbjegavati, budući da su raspoloživi deskriptori zasnovani na prostorima boja puno brži i precizniji, što je vidljivo i u rezultatima mjerenja.

7. Eksperimentalni evaluacijski okvir

Smisao eksperimentalnog evaluacijskog okvira jest pružiti znanstvenoj zajednici odgovarajuće pripremljene podatke i kriterij za evaluaciju sustava za lokalizaciju mobilnog robota zasnovanog na vizualnim obilježjima ravninskih segmenata u zatvorenim prostorima. Evaluacijski okvir koji će ovdje biti opisan je proizašao iz istraživanja u sklopu ove doktorske disertacije te ima za cilj omogućavanje lakog proširenja spoznaja i zaključaka o mogućnostima korištenja deskriptora vizualnih obilježja u svrhu lokalizacije mobilnog robota.

Postoje već slični evaluacijski okviri za topološku lokalizaciju mobilnog robota [85], ali ti okviri omogućuju samo približnu i semantičku lokalizaciju mobilnog robota zasnovanu na globalnim deskriptorima cijelih slika. Takvi evaluacijski okviri su također monomodalni odnosno ograničeni pristupom informacija iz okoline samo putem RGB kamera. Neki drugi evaluacijski okviri [92, 93], s druge strane, služe za evaluaciju SLAM sustava mobilnog robota što je problem koji se ne razmatra u sklopu ove disertacije. Okvir [92, 93] također koristi Microsoft Kinect kameru za prikupljanje informacija iz okoline, ali zbog fokusa na SLAM sustave ima drugačiji skup zahtjeva te snimanje nije vršeno pri različitim osvjetljenjima radi mogućnosti ispitivanja deskriptora vizualnih obilježja.

Razvijeni evaluacijski okvir se sastoji od:

- dvije skupine segmentiranih RGB-D slika, ispitne i referentne,
- datoteke koje sadrže točne podatke o relativnom položaju ispitnih slika naspram referentnih slika,
- datoteke koja sadrži relacije između ravninskih segmenata ispitnih i referentnih skupova slika,
- kriterija za evaluaciju deskriptora,
- podataka o kalibraciji Kinect senzora te
- primjera algoritma za korištenje pripremljenih podataka

Referentni i ispit skup podataka sadrže za svaku scenu RGB i dubinsku sliku snimljenu Kinect senzorom ugrađenim na mobilnu platformu Pioneer 3DX (slika 7.1.), podatke ugrađene odometrije mobilnog robota te mape segmentacije RGB i dubinske slike. Slike referentnog skupa su snimane u noćnim uvjetima s umjetnim osvjetljenjem radi postizanja što bolje dubinske slike, dok su sekvence ispitnih slika snimane u različitim vremenskim uvjetima, tipovima osvjetljenja i godišnjim dobima. Sve slike su snimane u zgradi Elektrotehničkog fakulteta u Osijeku.



Slika 7.1. Pioneer 3DX s ugrađenim Microsoft Kinect senzorom

Mape segmentacije pridružene RGB-D slikama su u osnovi tekstualne datoteke u koje su zapisana dvodimenzionalna polja jednakih dimenzija kao RGB i dubinska slika. Svaki element tog dvodimenzionalnog polja predstavlja identifikacijski broj ravninskog segmenta kojemu ta točka RGB odnosno dubinske slika pripada. U slučaju ako je vrijednost nekog elementa polja 0 to znači da ta točka ne pripada ni jednom ravninskom segmentu, jer nije imala definiranu vrijednost dubine u dubinskoj slici ili se nalazi na granici između ravninskih segmenata, pa je isključena iz daljnje obrade radi postizanja robusnosti na pogreške u kalibraciji sustava RGB-D kamera, kao što je objašnjeno u poglavlju 5.2. Takav način označavanja točaka na mapi segmentacije dubinske i RGB slike omogućava korisniku jednostavan i brz način formiranja deskriptora vizualnih obilježja ravninskog segmenta. Učinkovitost deskriptora nekog vizualnog obilježja će se moći izračunati na sličan način kako je to riješeno u sklopu ove doktorske disertacije, pomoću postotka propusnosti točnih parova ravninskih segmenata, odnosno pomoću izraza:

$$p_c = \frac{N_c}{N_{ref}} \cdot 100\% \tag{7.1}$$

gdje p_c predstavlja postotak točnih parova koje je deskriptor uspio prepoznati, dok je N_c broj uspješno prepoznatih točnih parova, a N_{ref} je ukupni broj točnih parova. Pomoću sljedećeg izraza se dobiva ukupna propusnost:

$$p_{tot} = \frac{N_{tot}}{n \cdot m} \cdot 100\% \tag{7.2}$$

gdje je p_{tot} ukupna propusnost koja predstavlja postotak koliko parova ravninskih segmenata je prošlo evaluaciju deskriptora (N_{tot}) naspram ukupnog broja svih mogućih parova ravninskih segmenata između ispitnih i referentnih scena ($n \cdot m$).



Slika 7.2. Primjer sparenih scena

Na slici 7.2. može se vidjeti jedan primjer sparenih scena. Ispitna scena je prikazana lijevo, a referentna scena desno. Slika 7.2. također prikazuje mape segmentacija dubinske i RGB slike za obje scene te točno sparene ravninske segmente koji su omogućili uspješan pronalazak relativnog položaja. U tablici 7.1. može se vidjeti primjer izgleda tekstualne datoteke u kojoj će biti upisani relativni položaji između ispitnih scena i njima bliskih lokalnih modela. Relativni položaji između ispitnih i referentnih scena su dobiveni globalnom lokalizacijom uz korištenje referentnih podataka inkrementalne lokalizacije. Provođenjem globalne lokalizacije za slike ispitnih sekvenci uz kartu stvorenu od referentnih scena te sortiranjem generiranih hipoteza za svaku scenu prema kriteriju (4.35) omogućeno je pomoću referentnih podataka inkrementalne lokalizacije pronalaženje točnih hipoteza o položaju svih referentnih modela u okolini razmatrane ispitne scene. Pri tome apsolutni položaji tih pronađenih hipoteza u sortiranom popisu moraju biti unutar okoline od 200mm od položaja dobivenog referentnim podacima inkrementalne lokalizacije i s razlikom u rotaciji od najviše 2°. Dakle, provedbom globalne lokalizacije dobivamo popis svih lokalnih modela u karti čije hipoteze svojim relativnim položajem zadovoljavaju navedenu okolinu dobivenu referentnim podacima.

SceneID	ModelID	Х	Y	Alpha
0	3	1361.073375	204.230409	8.619176
0	6	-489.033456	175.862630	8.307426
0	5	303.224910	160.349176	8.736810
1	6	-100.536369	196.392627	1.907758
1	7	-573.430685	233.788719	2.562933
2	7	-93.155392	253.799602	2.317635

Tablica 7.1. Primjer datoteke s relativnim položajima između ispitnih i referentnih scena

Struktura tekstualne datoteke prikazane u tablici 7.1. je stupčana. Prvi redak u datoteci je zaglavlje koje opisuje koje informacije nosi pojedeni stupac. Ostali redci u datoteci predstavljaju sparene ispitne i njima bliske referentne scene. Ispitna i referentne scene su identificirane preko prva dva stupca datoteke koji predstavljaju identifikacijske brojeve, dok zadnja tri stupca predstavljaju relativne 3DOF koordinate u milimetrima i stupnjevima koje predstavljaju položaj ispitne scene u odnosu na referentnu.

Relacije između ravninskih segmenata ispitnih i referentnih scena su dobivene na isti način pomoću pokusa globalne lokalizacije uz korištenje referentnih podataka inkrementalne lokalizacije. Dobivanjem svih točnih hipoteza lokalnih modela iz karte bliskih ispitnoj sceni omogućeno je da se za svaki ravninski segment ispitnih slika dobije popis svih točno sparenih ravninskih segmenata referentnih slika. Relacije između ravninskih segmenata ispitnih i referentnih scena su zapisane u tekstualnoj datoteci čiji se primjer može vidjeti u tablici 7.2. Datoteka prikazana tablicom 7.2. ima četiri stupca koji govore koji ravninski segment testne slike ("SegmentID") je sparen s kojim ravninskim segmentom referentne slike ("ModelSegmentID"). Identifikacijski brojevi segmenata odgovaraju onim identifikacijskim brojevima koji su korišteni u segmentacijskim mapama RGB i dubinskih slika za ispitne i referente scene.

Tablica 7.2. Primjer datoteke s relacijama između ravninskih segmenata testnih scena i referentnih scena

SceneID	SegmentID	ModelID	ModelSegmentID
0	1	3	1
0	2	3	5
0	3	3	4
0	5	3	9
0	6	3	8
0	10	3	11
0	13	3	12

Uz navedene datoteke, u evaluacijskom okviru će se također nalaziti i datoteka s kalibracijskim podacima korištenog Kinect senzora, koji uključuju intristične podatke RGB i dubinske kamere te podatke o rotacijskoj matrici i translacijskom vektoru između koordinatnih sustava RGB i dubinske kamere. Ovi podaci nisu potrebni za provođenje pokusa za evaluaciju deskriptora vizualnih obilježja, budući da uz svaku referentnu i ispitnu RGB-D sliku dolaze segmentacijske mape, ali ti podaci omogućuju korisniku da koristi vlastiti algoritam za segmentaciju na ravninske segmente ako to želi.

Pored podataka o ispitnim i referentnim scenama te njihovim međusobnim odnosima kao i odnosima njihovih ravninskih segmenata, uz evaluacijskih okvir dani su primjeri algoritama koji predstavljaju upute kako koristiti te navedene podatke. Ako se želi ispitati učinkovitost deskriptora vizualnih obilježja pomoću evaluacijskih kriterija propusnosti točnih parova ravninskih segmenata i ukupne propusnosti svih parova, potrebno je provesti algoritam opisan pseudo kodom u tablici 7.3.

Tablica 7.3. Algoritam za evaluaciju uspješnosti deskriptora vizualnih obilježja pokusom pronalaska inicijalne korespodencije

Ulazni podaci: Popis parova ravninskih segmenata testnih i referentnih scena generiran od strane korisnika, gdje je svaki segment testne scene predstavljen s identifikatorima SceneID i SegmentID, a svaki segment referentne scene s ModelID i ModelSegmentID, referentna datoteka s relacijama između segmenata
Izlazni podaci : broj točno sparenih ravninskih segmenata N_c , broj obrađenih parova N_{tot}
Algoritam:
Za svaki par s popisa:
Pronaći odgovarajuće linije u referentnoj datoteci koje odgovaraju segmentu scene (SceneID, SegmentID)
Za svaku pronađenu liniju:
Provjeri da li pročitane vrijednosti (ModelID, ModelSegmentID) odgovaraju onima za razmatrani par
Ako da:
Povećaj brojač točno sparenih ravninskih segmenata za 1
Prekini : za svaku pronađenu liniju
Kraj : za svaku pronađenu liniju
Povećaj brojač obrađenih parova za 1
Kraj : za svaki par s popisa

Kao što se može vidjeti u tablici 7.3., za izvršavanje navedenog pseudo koda je potrebno imati popis parova ravninskih segmenata koji su spareni po nekom korisničkom kriteriju. Svaki par ravninskih segmenata je opisan s identifikacijskim brojem trenutno razmatrane ispitne scene "SceneID", identifikacijskim brojem ravninskog segmenta scene "SegmentID" te odgovarajućim vrijednostima za razmatranu referentnu scenu "ModelID" i spareni ravninski segment "ModelSegmentID". Algoritam za svaki ravninski segment ispitne scene parova na popisu pronalazi linije referentne datoteke koji odgovaraju identifikacijskom

broju scene i segmenta te provjerava da li identifikacijski brojevi modela i ravninskog segmenta modela odgovaraju onima koji opisuju trenutno razmatran par. Ako odgovaraju poveća se broj točno sparenih ravninskih segmenata N_c za 1 i nastavi dalje obrađivati ostale parove. Prije prelaska na idući par poveća se broj obrađenih parova za 1. Ako se algoritam izvrši za sve moguće parove testna scena – referentna scena dobiti će se ukupni brojevi točno sparenih ravninskih segmenata N_c i ukupni broj segmenata N_{tot} koji su prošli korisnikov kriterij. Referentni ukupni broj točnih parova N_{ref} može se znati po broju linija u referentnoj datoteci umanjen za 1 zbog zaglavlja datoteke, dok ukupan broj parova koji su mogli proći zadani kriterij korisnik može odrediti kao umnožak ukupnog broja segmenata testne scene n i ukupnog broja segmenata referentne scene m. Tada korisnik pomoću izraza (7.1) i (7.2) može odrediti učinkovitost vlastitih deskriptora vizualnih obilježja.

Učinkovitost deskriptora vizualnih obilježja ravninskih segmenata se može prikazati i performansama pri rješavanju problema globalne lokalizacije na sličan način kako je provedeno i u ovoj doktorskoj disertaciji. Za provođenje tog pokusa koristi se datoteka relativnih položaja ispitnih i referentnih scena prikazana u tablici 7.1. Pseudo kod koji je tada potrebno izvršiti se vidi u tablici 7.4.

Tablica 7.4. Algoritam za evaluaciju uspješnosti deskriptora vizualnih obilježja pokusom globalne lokalizacije

Ulazni podaci: Sortirani popis generiranih hipoteza za trenutnu ispitnu scenu pri čemu je svaka hipoteza opisana
relativnim 3DOF položajem i identifikacijskim brojem referentne scene "ModelID" na koju se taj
položaj odnosi, referentna datoteka s relativnim položajima ispitnih i referentnih scena, dopuštena
pogreška lokalizacije (200mm i 2°), identifikacijski broj trenutno razmatrane ispitne scene
"SceneID"
Izlazni podaci: indeks prve točne hipoteze za svaku ispitnu scenu
Algoritam:
Pronaći odgovarajuće linije u referentnoj datoteci koje odgovaraju trenutnoj ispitnoj sceni (SceneID)
Za svaku hipotezu na sortiranom popisu:
Za svaku pronađenu liniju u referentnoj datoteci:
Provjeri da li pročitana vrijednost(ModelID) odgovara onoj zapisanoj u trenutno razmatranoj hipotezi
Ako da:
Provjeri da li je relativni položaj hipoteze unutar zadane pogreške od one u zapisane u liniji(200mm i 2°)
Ako je:
Indeks = redni broj trenutno razmatrane hipoteze na sortiranom popisu
Prekini : za svaku hipotezu
Kraj : za svaku pronađenu liniju
Kraj : za svaku hipotezu

Kao što se može vidjeti u tablici 7.4. algoritam opisan pseudo kodom omogućava pronalazak prve točne hipoteze na sortiranom popisu generiranih hipoteza. Hipoteze na popisu su sortirane od najbolje prema najlošijoj prema nekom korisnikovom evaluacijskom

kriteriju. Svaka hipoteza je opisana identifikacijskim brojem referentne scene za koju je generirana i relativnim položajem u tri stupnja slobode u odnosu na trenutno razmatranu ispitnu scenu. Algoritmu je također potrebno predati identifikacijski broj trenutno razmatrane ispitne scene i datoteku s referentnim podacima o relativnim odnosima između ispitnih i referentnih scena. Korisnik također mora zadati dopuštenu pogrešku položaja lokalizacije u mm i pogrešku rotacije u ° ili može koristiti vrijednosti koje su korištene u ovoj disertaciji, a koje iznose 200mm i 2°. Prikazani algoritam u tablici 7.4 vraća u varijabli Indeks redni broj prve točne hipoteze u sortiranom popisu koja je pronađena na način da algoritam obilazi cijeli popis i za svaku hipotezu pokuša pronaći onaj točni relativni položaj definiran referentnim podacima u datoteci. Točan relativni položaj je onaj koji zadovoljava zadanu pogrešku lokalizacije za odgovarajući par identifikacijskih brojeva (SceneID, ModelID). Ako se algoritam prikazan u tablici 7.4. primijeni na sve testne scene tako da se za svaku testnu scenu odredi indeks prve točne hipoteze, može se odrediti prosječni indeks prve točne hipoteze na sortiranom popisu generiranih hipoteza koji potom može služiti kao mjera učinkovitosti primjene deskriptora vizualnih obilježja ravninskih segmenata u sustavima za lokalizaciju mobilnog robota. Nadalje, pobrojavanjem svih testnih scena za koje je indeks prve točne hipoteze jednak 1 može se dobiti još jedan pokazatelj učinkovitosti koji se analizira u ovom radu, a prikazan je u tablici 6.2. u stupcu pod naslovom "Prva hip. točna". Praćenje ostalih mjera učinkovitosti kao što su broj generiranih hipoteza, vrijeme faze lokalizacije i ukupno vrijeme rada sustava korisnik treba sam realizirati u sklopu svog sustava za lokalizaciju mobilnog robota.

8. Zaključak

U ovoj doktorskoj disertaciji su analizirane mogućnosti korištenja deskriptora vizualnih obilježja trodimenzionalnih ravninskih segmenata za lokalizaciju mobilnog robota u stvarnim uvjetima zatvorenih prostora. Promjene osnovnog sustava za lokalizaciju mobilnog robota zasnovanog na ravninskim segmentima su uključivale uvođenje koraka pred-procesiranja, novog uvjeta kod inicijalnog sparivanja ravninskih segmenata te novog kriterija za evaluaciju hipoteza koji koriste sličnost deskriptora vizualnih obilježja radi bolje procjene točnosti hipoteze. Novim uvjetom kod inicijalnog sparivanja se postiže točnije sparivanje ravninskih segmenata odnosno odbacivanje netočnih parova, što rezultira manjim brojem generiranih hipoteza i kraćim vremenom lokalizacije. Korak pred-procesiranja služi za uklanjanje utjecaja šuma i smanjivanje utjecaja promjene osvjetljenja. Kod korištenja deskriptora teksture ovaj je korak uključivao naglašavanje kontrasta radi isticanja teksturiranih površina.

Za potrebe opisivanja vizualnih obilježja implementirani su deskriptori šest prostora boja, RGB, CIE XYZ, CIE L*a*b*, HSV, kromatski rgb prostor i prostor suprotstavljenih boja te tri deskriptora teksture zasnovana na LBP porodici deskriptora teksture. Također se svojstvo generaliziranja deskriptora i smanjenja utjecaja promjene osvjetljenja pokušalo povećati uvođenjem mekih histograma za deskriptore zasnovane na prostorima boja. Svi implementirani deskriptori su inicijalno ispitani za tri rezolucije, deskriptori boja za rezolucije od 32, 16 i 8 binova po dimenziji prostora boja, a za LBP deskriptore su se promatrale okoline na udaljenosti 1 s 8 točaka, udaljenosti 2 s 16 točaka i udaljenosti 3 s 24 točke u okolini. Kao mjera sličnosti deskriptora je odabran presjek histograma kao optimalna mjera koja je jednostavna, brza i numerički stabilna.

Analiza inicijalnog sparivanja ravninskih segmenta putem njihovih deskriptora vizualnih obilježja je služila za dolaženje do spoznaja o mogućnostima diskriminacije samih deskriptora i njihove primjene za raspoznavanje ravninskih segmenata. Analizom inicijalnog sparivanja su se također odredili pragovi sličnosti za pojedine deskriptore koji se moraju na odgovarajući način podesiti da bi se osigurali zahtjevi za određenom propusnošću točnih parova i ukupne propusnosti svih mogućih parova u ostale faza rada sustava za lokalizaciju. Općenito se iz analize inicijalne korespondencije može zaključiti da dvodimenzionalni tipovi deskriptora boja pokazuju bolje sposobnosti raspoznavanja točnih parova od trodimenzionalnih, ali isto tako imaju i veći stupanj propusnosti svih mogućih parova. Korištenje mekih tipova histograma rezultira povećanjima propusnosti točnih parova

isključivo u slučajevima kada se koristi puni presjek histograma, dok u slučajevima parcijalnog presjeka kod nekih prostora boja razlike između standardnih i mekih tipova gotovo da ni nema. Parcijalni presjek histograma sa samo 5 najvećih binova rezultira skoro identičnim rezultatima kao i puni presjek za većinu dvodimenzionalnih histograma, što sugerira da se uglavnom radi o kompaktnim distribucijama u histogramima posebno onim manjih rezolucija. Općenito se može zaključiti da niti za jedan od ispitivanih deskriptora ne postoji vrijednost praga koja će sve točne parove odijeliti od netočnih, pa se mora naći odgovarajući kompromis, tj. optimalni prag. Rezultati analize inicijalne korespondencije su potencijalno optimalni deskriptori za uporabu s ravninskim segmentima u svrhu lokalizacije mobilnog robota.

Performanse deskriptora zasnovanih na vizualnim obilježjima su se također ispitale u kontekstu globalne lokalizacije mobilnog robota. Ispitivana su po dva deskriptora za svaki implementirani prostor boja, jedan kada se koristi puni presjek histograma i drugi kada se koristi parcijalni presjek te tri konfiguracije LBP deskriptora teksture. Svaki odabrani deskriptor je imao i svoj pripadajući prag sličnosti dobiven prethodnom analizom inicijalnog sparivanja koji par segmenata mora dostići u fazi inicijalnog sparivanja ravninskih segmenata da bi bio propušten u fazu generiranja hipoteza. Analiza pokusa globalne lokalizacije obuhvaćala je promatranje vremena trajanja faze lokalizacije, ukupnog vremena rada sustava, broja generiranih hipoteza te indeksa prve točne hipoteze i njegove razdiobe za osnovni i prošireni sustav za lokalizaciju. Kao referentni podaci za određivanje indeksa prve točne hipoteze na sortiranom popisu hipoteza generiranih u okviru postupka globalne lokalizacije korišteni su podaci inkrementalne lokalizacije za iste sekvence slika. Pokusi su provedeni na način da se svi parovi ravninskih segmenata kojima barem jedan segment ima manje od 20 pridruženih točaka propuštaju bez evaluacije deskriptora vizualnih obilježja kao i na način da im se neovisno o tom broju procjeni sličnost deskriptora vizualnih obilježja i usporedi sa zahtijevanim pragom.

Analizom dobivenih podataka globalne lokalizacije je utvrđeno da korištenje većine deskriptora vizualnih obilježja donosi primjetne uštede vremena u fazi lokalizacije ako se dopusti evaluacija deskriptora vizualnih obilježja svih parova ravninskih segmenata neovisno o broju točaka koje sadrže iz razloga što se izbacuju parovi nedovoljne sličnosti i time smanjuje broj generiranih hipoteza. Ali u tim slučajevima također dolazi do povećanja broja scena za koje se ne može generirati ni jedna točna hipoteza, najčešće zbog gubitka informacije potrebne za određivanje zadnjeg stupnja slobode. Naime, često su segmenti koji nose informaciju o zadnjem stupnju slobode vrlo mali i kao takvi sadrže nedovoljan broj točaka za pouzdano formiranje deskriptora vizualnih obilježja, pa dolazi do njihovog odbacivanja u slučaju da se ti deskriptori koriste kod inicijalnog sparivanja. U slučaju propuštanja parova ravninskih segmenata kojima barem jedan element para ima manje od 20 točaka ubrzanja su manja. Ipak, u nekim slučajevima, kao npr. kod HSV8 deskriptora s parcijalnim presjekom, postižu se primjetna ubrzanja bez značajnijeg povećanja broja scena za koje se ne može generirati ni jedna točna hipoteza. Broj scena za koje je prva hipoteza točna je u slučajevima korištenja pojedinih deskriptora vizualnih obilježja znatno veći u odnosu na sustav zasnovan samo na geometriji. Stoga je i prosječni indeks prve točne hipoteze u sortiranom popisu također znatno manji, što u konačnici rezultira znatno manjim brojem hipoteza koje robot mora uzimati u obzir prilikom globalne lokalizacije. Najboljim deskriptorima prostora boja su se pokazali HSV i Lab deskriptori, dok su se korišteni RGC deskriptori pokazali najlošijim. Deskriptori tekstura zasnovani na LBP porodici pokazuju znatno lošije performanse od deskriptora boja za primjenu u sustavima lokalizacije mobilnog robota. Deskriptori tekstura su čak u mnogim slučajevima pokazali znatno pogoršanje rada sustava po broju scena za koje se ne može generirani ni jedna točna hipoteza u odnosu sustav zasnovan samo na geometriji.

Zadnja stavka realizirana u ovoj disertaciji uključuje eksperimentalni okvir za evaluaciju algoritama sparivanja vizualnih obilježja ravninskih segmenata i lokalizacije mobilnog robota. Razvijeni eksperimentalni okvir će omogućiti znanstvenoj zajednici jednostavnu i konzistentnu evaluaciju razvijenih deskriptora vizualnih obilježja za primjenu u sustavima lokalizacije mobilnog robota zasnovanih na ravninskim segmentima. Okvir će sadržavati sve referentne i ispitne sekvence RGB-D slika te njihove prostorne relacije korištene i u ispitivanju u sklopu ove disertacije. Sve scene u evaluacijskom okviru će biti segmentirane na ravninske segmente s danim relacijama između segmenata ispitnih i referentnih scena. Predloženi kriteriji evaluacije deskriptora vizualnih obilježja uključuju propusnost točnih parova ravninskih segmenata i ukupnu propusnost svih mogućih parova. Također je dan pseudo kod algoritma za korištenje podataka evaluacijskog okvira radi računanja istih mjera uspješnosti deskriptora vizualnih obilježja ravninskih segmenata koji su korišteni u ovoj disertaciji.

Literatura

- [1] Henry, P., Krainin, M., Herbst, E., Ren, X., Fox, D., "RGB-D Mapping: Using Depth Cameras for Dense 3D Modeling of Indoor Environments", Proceedings of International Symposium on Experimental Robotics, Delhi, India, 2010.
- [2] Fallon, M. F., Johannsson, H., Leonard, J., "Efficient Scene Simulation for Robust Monte Carlo Localization using an RGB-D Camera", Proceedings of ICRA 2012, St Paul, Minnesota, USA. May 2012.
- [3] Cupec, R., Nyarko, K. E., Filko, D., Petrović, I., "Fast Pose Tracking Based on Ranked 3D Planar Patch Correspondences", Proceedings of 10th IFAC Symposium on Robot Control(SYROCO 2012), 05-07. rujna 2012., Dubrovnik
- [4] Nyarko, E. K., "Planar surface segments as features for mobile robot localization in 3D indoor environments", Technical report, Elektrotehnički fakultet, Osijek, 2012.
- [5] Thrun, S., Burgard, W., Fox, D., "Probabilistic robotics", Intelligent Robotics and Autonomous Agents series, The MIT Press, 2005.
- [6] Fox, D., Burgard, W., Thrun, S., "Markov localization for mobile robots in dynamic environments", Journal of Artificial Intelligence Research, no.11, pp. 391-427, 1999.
- [7] Fox, D., Burgard, W., Dellaert ,F., Thrun, S., "Monte Carlo localization: Efficient position estimation for mobile robots", Proceedings of the National Conference on Artificial Intelligence(AAAI), Orlando, FL, 1999.
- [8] Dellaert, F., Fox, D., Burgard, W., Thrun, S., "Monte Carlo Localization for Mobile Robots", Proceedings of the 1999 International Conference on Robotics and Automation, pp. 1322-1328, 1999.
- [9] Thrun, S., Fox, D., Burgard, W., Dellaert, F., "Robust Monte Carlo localization for mobile robots", Artificial Intelligence, no. 128, pp 99-141, 2001.
- [10] Jensfelt, P., Wijk, O., Austin, D.J., Andersson, M., "Experiments on augmenting condensation for mobile robot localization", Proceedings of the International Conference on Robotics and Automation (ICRA), San Francisco, CA, USA, pp. 2518–2524, 2000.
- [11] Dayoub, F., Duckett, T., Cielniak, G., "An Adaptive Spherical View Representations for Navigation in Changing Environments", Proceedings of the 4th European Conference on Mobile Robots, pp. 1-6, Dubrovnik, Croatia, 2009.

- [12] Kosaka, A., Kak, A., "Fast vision-guided mobile robot navigation using model-based reasoning and prediction of uncertainties", CVGIP: Image Understanding, vol. 56, no. 3, pp. 271-329, 1992.
- [13] Aider, O. A., Hoppenot, P., Colle, E., "A model-based method for indoor mobile robot localization using monocular vision and straight-line correspondences", Robotics and Autonomous Systems, vol. 52, no. 2-3, pp. 229-246, 2005.
- [14] R. Hartley, A. Zisserman, "Multiple View Geometry in Computer Vision", Cambridge University Press, Cambridge, 2003.
- [15] Wikipedia, en.wikipedia.org
- [16] Leica, http://www.leica-geosystems.com/en/5574.htm
- [17] Panasonic D-IMager, http://pewa.panasonic.com/components/built-in-sensors/3d-imagesensors/d-imager/
- [18] Microsoft Xbox Kinect, http://www.xbox.com/en-US/KINECT
- [19] Libfreenect, http://openkinect.org/
- [20] Mikolajczyk, K., Schmid, C., "Performance evaluation of local descriptors", IEEE Transactions on pattern analysis and machine intelligence, vol. 27, no. 10, pp. 1615-1630, 2005.
- [21] Lowe, D. G., "Object recognition from local scale-invariant features", Proceedings of International Conference on Computer Vision, Corfu, Greece, pp. 1150-1157., 1999.
- [22] Se, S., Lowe, D. G., Little, J. J., "Vision-based global localization and mapping for mobile robots", IEEE Transactions on Robotics, vol. 21, no. 3, pp. 364-375, 2005.
- [23] Vidović, I., "Lokalni deskriptori temeljeni na histogramu orijentacije gradijenta svjetline slike", Diplomski rad, Elektrotehnički fakultet, 2011.
- [24] Bay, H., Ess, A., Tuytelaars, T., Gool, L. V., "SURF: Speeded Up Robust Features", Computer Vision and Image Understanding (CVIU), vol. 110, no. 3, pp. 346-359, 2008.
- [25] Computer vision talks blog, http://computer-vision-talks.com/2011/01/ comparison-offeature-descriptors/
- [26] Cupec, R., Nyarko, E. K., Grbić, R., "Detection of dominant planar surfaces in disparity images based on random sampling", Tehnički Vjesnik, vol. 18, no. 3, pp. 301-308., 2011.

- [27] Cobzas, D., Zhang, H., "Mobile Robot Localization using Planar Patches and a Stereo Panoramic Model", Vision Interface, pp. 94-99., 2001.
- [28] Cupec, R., Nyarko, E. K., Filko, D., "Fast 2.5D Mesh Segmentation to Approximately Convex Surfaces", Proceedings of the 5th European Conference on Mobile Robots, pp. 127-132, Örebro, Švedska, 2011.
- [29] Marshall, D., Lukacs, G., Martin, R., "Robust segmentation of primitives from range data in the presence of geometric degeneracy", IEEE Transactions in Pattern Analysis and Machine Intelligence, vol. 23, no. 3, pp. 304–314, 2001.
- [30] Attene, M., Spagnuolo, M., Falcidieno, B., "Hierarchical Mesh Segmentation based on Fitting Primitives", The Visual Computer, vol.22, no. 3, pp. 181-193, 2006.
- [31] Lafarge, F., Keriven, R., Brédif, M., "Insertion of 3-D-Primitives in Mesh-Based Representations: Towards Compact Models Preserving the Details", IEEE Transactions on Image Processing, vol. 19, no. 7, 2010.
- [32] Valgren, C., Lilienthal, A., Duckett, T., "Incremental Topological Mapping Using Omnidirectional Vision", Proceedings of IROS2006, pp. 3441-3447, 2006.
- [33] Booij, O., Terwijn, B., Zivkovic, Z., Krose, B., "Navigation using an appearance based topological map", Proceedings of ICRA2007, pp. 3927-3932, 2007.
- [34] Cupec, R., Nyarko, E. K., Kitanov, A., Petović, I., "RANSAC-Based Stereo Image Registration with Geometrically Constrained Hypothesis Generation", AUTOMATIKA: časopis za automatiku, mjerenje, elektroniku, računarstvo i komunikacije, vol.50, no.3-4, pp. 195 - 204, Prosinac 2009.
- [35] Scaramuzza, D., Siegwart, R., Martinelli, A., "A Robust Descriptor for Tracking Vertical Lines in Omnidirectional Images and Its Use in Mobile Robotics", The International Journal of Robotics Research, vol. 28, no. 2, pp. 149-171, 2009.
- [36] Pathak, K., Birk, A., Vaskevicius, N., Poppinga, J., "Fast Registration Based on Noisy Planes with Unknown Correspondences for 3D Mapping", IEEE Transactions on Robotics, vol. 26, no. 3, pp. 424 – 441, 2010.
- [37] Pathak, K., Birk, A., Vaskevicius, N., Pfingsthorn, M., Schwertfeger, S., Poppinga., J., "Online three-dimensional SLAM by registration of large planar surface segments and closed-form pose-graph relaxation". Journal of Field Robotics, vol. 27, no. 1, 52-84, January 2010.

- [38] Biswas, J., Veloso, M., "Depth Camera Based Indoor Mobile Robot Localization and Navigation", Proceedings of ICRA 2012, pp. 1697-1702, St Paul, Minnesota, 2012.
- [39] Gonzalez, R. C., Woods, R. E., "Digital Image processing", Pearson Prentice Hall, 2008.
- [40] Qidwai, U., Chen, C. H., "Digital image processing: An Algorithmic Approach with MATLAB", CRC Press, 2010.
- [41] Gevers, T., Gijsenj, A., Weijer, J. V. D., Geusebroek, J. M., "Color in Computer Vision", Wiley–IS&T Series in Imaging Science and Technology, John Wiley & Sons Ltd, 2012.
- [42] Shafer, M., "Using color to separate reflection components", Color Research and Applications, vol. 10, no. 4, pp. 210–218, 1985.
- [43] Funt, B.V., Finlayson, G.D., "Color Constant Color Indexing", IEEE Transaction on Pattern Analysis and Machine Intelligence, vol. 17, no. 5, pp. 522-529, 1995.
- [44] Finlayson, G. D., Hordley, S., Hubel, P. M., "Illuminant estimation for object recognition", Color Research & Application, vol. 27, no 4., pp. 260-270., 2002.
- [45] Finlayson, G., Hordley, S., Schaefer, G., Tian, G., "Illuminant and device invariant colour using histogram equalization", Pattern Recognition, no. 38, pp. 179 – 190, 2005.
- [46] Kries, J. V., "Influence of adaptation on the effects produced by luminous stimuli", Sources of Color Vision, MIT Press, Cambridge, 1970.
- [47] Finlayson, G. D., Hordley, S. D., Xu, R., "Convex programming colour constancy with a diagonal-offset model", IEEE International Conference on Image Processing, pp. 948-951, 2005.
- [48] Swain, M. J., Ballard, D. H., "Color Indexing", International Journal of Computer Vision, vol. 7, no.1, pp. 11-32, 1991.
- [49] Pass, G. Zabih, R., "Comparing Images Using Joint Histograms", ACM Journal of Multimedia Systems, vol. 7, no. 3, pp.234, 1999.
- [50] Gevers, T., Smeulders, A.W. M., "Color-based object recognition", Pattern Recognition, no. 32, pp. 453-464, 1999.
- [51] Park, D.S., Park, J.S., Kim, T.Y., Han, J.H., "Image Indexing Using Weighted Color Histogram", Proceedings of the IEEE 10th International Conference on Image Analysis and Processing (ICIAP '99), 1999.

- [52] Berens, J., Finlayson, G.D., Qiu, G., "Image indexing using compressed colour histograms", IEEE Proceedings of Vision Image and Signal Processing, pp. 349-355, 2000.
- [53] Al-Jarrah, M., Al-Omari, F., "An efficient statistical content-based color image indexing and retrieval", Proceedings of ICITA 2002, 2002.
- [54] Forseen, P. E., "Image analysis using soft histograms", Proceedings of the SSAB Symposium on Image Analysis, pp. 109-112, 2001.
- [55] Czyz, J., Ristic, B., Macq, B., "A particle filter for joint detection and tracking of color objects", Image Vision Computing, vol.27, no.7, 2007.
- [56] Domke, J., Aloimonos, Y., "Deformation and viewpoint invariant color histograms", Proceedings of BMVC (British Machine Vision Conference), 2006.
- [57] Sande, K. V. D., Gevers, T., Snoek, C., "Evaluating Color Descriptors for Object and Scene Recognition", IEEE Transaction on Pattern Analysis and Machine Intelligence, vol. 32, no. 9, pp. 1582-1596, 2010.
- [58] O'Hara, S., Draper, B. A., "Introduction to the Bag of Features Paradigm for Image Classification and Retrieval", CoRR abs/1101.3354, 2011.
- [59] Taylor, C. J., Cowley, A., "Segmentation and Analysis of RGB-D data", Proceedings of RSS 2011 Workshop on RGB-D Cameras, 2011.
- [60] Mirmehdi, M., Xie, X., Suri, J., "Handbook of texture analysis", Imperial College Press, 2009.
- [61] Ojala, T., Pietikäinen, M., Harwood, D., "Performance evaluation of texture measures with classification based on Kullback discrimination of distributions", Proceedings of the 12th IAPR International Conference on Pattern Recognition (ICPR 1994), vol. 1, pp. 582 - 585, 1994.
- [62] Ojala, T., Pietikainen, M., Maenpaa, T., "Multiresolution Gray-Scale and Rotation Invariant Texture Classification with Local Binary Patterns", IEEE Transactions in Pattern Analysis and Machine Intelligence, vol. 24, no. 7, pp. 971-987, 2002.
- [63] Bianconi, F., Fernández, A., "On the Occurrence Probability of Local Binary Patterns: A Theoretical Study", Journal of Math. Imaging Vision, vol. 40, no. 3, pp. 259-268, 2011.

- [64] Kandaswamy, U., Schuckers, S.A., Adjeroh, D., "Comparison of Texture Analysis Schemes Under Nonideal Conditions", IEEE Transactions on Image Processing, vol.20, no. 8, pp. 2260-2275, 2011.
- [65] Kumar, A., Pang, G. K. H., "Defect Detection in Textured Materials Using Gabor Filters", IEEE Transactions on industry applications, vol. 38, no. 2, pp. 425-440, 2002.
- [66] Jain, A. K., Farrokhnia, F., "Unsupervised texture segmentation using Gabor filters", Pattern Recognition, vol. 24, no. 12, pp. 1167-1186, 1991.
- [67] Manjunath, B. S., Ma, W. Y., "Texture Features for Browsing and Retrieval of Image Data", IEEE Transaction in Pattern Analysis and Machine Intelligence, vol. 18, no. 8, pp. 837-842, 1996.
- [68] Varma, M., Zisserman, A., "Classifying Images of Materials: Achieving Viewpoint and Illumination Independence", Proceedings of the 7th European Conference on Computer Vision, pp.255-271, Copenhagen, Denmark , 2002.
- [69] Leung, T., Malik, J., "Representing and Recognizing the Visual Appearance of Materials using Three-dimensional Textons", International Journal of Computer Vision, vol. 43, no. 1, pp. 29-44, 2001.
- [70] Zhu, S. C., Guo, C.E., Wang, Y., Xu, Z., "What are Textons?", International Journal of Computer Vision, vol. 62, no. 1-2, pp. 121-143, 2005.
- [71] Maenpaa, T., Pietikainen, M., "Classification with color and texture: jointly or separately", Pattern Recognition, no. 37, pp. 1629 – 1640, 2004.
- [72] Hayman, E., Caputo, B., Fritz, M., Eklundh, J. O., "On the Significance of Real-World Conditions for Material Classification", Proceedings of European Conference on Computer Vision, vol. 4, pp. 253-266, 2004.
- [73] Caputo, B., Hayman, E., Mallikarjuna, P., "Class-Specific Material Categorisation", Proceedings of the Tenth IEEE International Conference on Computer Vision (ICCV '05), vol. 2, Washington, USA, pp. 1597-1604, 2005.
- [74] Caputo, B., Hayman, E., Fritz, M., Eklundh, J. O., "Classifying materials in the real world", Image Vision Computing, vol. 28, no. 1, pp. 150-163, 2010.
- [75] Ulrich, I., Nourbakhsh, I. R., "Appearance-Based Place Recognition for Topological Localization", Proceedings of ICRA 2000:, pp.1023-1029, 2001.

- [76] Blear, P., Allen, P., "Topological Mobile Robot Localization Using Fast Vision Techniques", Proceedings of ICRA 2002, pp. 1031-1036, 2002.
- [77] Zhou, C., Wei, Y., Tan, T., "Mobile Robot Self-Localization Based on Global Visual Appearance Features", Proceedings of ICRA2003, pp.1271-1276, 2003.
- [78] Pronobis, A., Caputo, B., Jensfelt, P., Christensen, H.I., "A Discriminative Approach to Robust Visual Place Recognition", Proceedings of IROS 2006, pp.3829-3836, 2006.
- [79] Rano, I., Lazkano, E., Sierra, B., "On The Application Of Colour Histograms For Mobile Robot Localisation", Proceedings of ECMR2005, pp.189-193, 2005.
- [80] Zingaretti, P., Frontoni, E., "Appearence Based Robotics", IEEE Robotics & Automation Magazine, vol. 13, no. 1, pp. 59-68, 2006.
- [81] Werner, F., Sitte, J., Maire, F. D., "Automatic Place Determination using Colour Histograms and Self-Organising Maps", Proceedings 13th International Conference on Advanced Robotics, pp. 111-116, Jeju, Korea Republic, 2007.
- [82] Jafar, F.A., Suzuki, Y., Tateno, Y., Yokota, K., Matsuoka, T., "Localization Method for Autonomous Mobile Robot using Visual Features in Environment", International Journal of Intelligent Information Technology Application, vol. 4, no. 2, pp.169-176, 2009.
- [83] Luo, J., Pronobis, A., Caputo, B., Jensfelt, P., "Incremental Learning for Place Recognition in Dynamic Environments", In Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS07), pp. 721-728, San Diego, CA, USA, 2007.
- [84] Werner, F. Sitte, J. Maire, F. D., "Visual Topological Mapping and Localisation using Colour Histograms", Proceedings International Conference on Control, Automation, Robotics and Vision (ICARCV08), pp. 341-346, Hanoi, Vietnam, 2008.
- [85] Pronobis, A., Caputo, B., Jensfelt, P., Christensen, H.I., "A realistic benchmark for visual indoor place recognition", Robotics and Autonomous Systems, vol. 58, no. 1, pp. 81-96, Siječanj 2010.
- [86] Akers, E. L., Agah, A., "Topological localization using appearance-based recognition", Journal of Automation, Mobile Robotics & Intelligent Systems, vol. 4, no. 1, pp.68-84, 2010.

- [87] Mason, J., Ricco, S., Parr, R., "Textured occupancy grids for monocular localization without features", Proceedings of ICRA2011, pp. 5800-5806, 2011.
- [88] Fu, Y., Tully, S., Kantor, G., Choset, H., "Monte Carlo Localization using 3D Texture Maps", Proceedings of IROS2011, pp.482-487, 2011.
- [89] Garland, M., Willmott, A., Heckbert, P. S., "Hierarchical face clustering on polygonal surfaces", Proceedings of the 2001 symposium on Interactive 3D graphics (I3D '01), pp. 49-58, 2001.
- [90] Tomasi, C., Manduchi, R., "Bilateral Filtering for Gray and Color Images", Proceedings of the Sixth International Conference on Computer Vision (ICCV '98), Washington, DC, USA, pp. 839-847, 1998.
- [91] OpenCV, http://opencv.willowgarage.com
- [92] Sturm, J., Magnenat, S., Engelhard, N., Pomerleau, F., Colas, F., Burgard, W., Cremers, D., Siegwart, R., "Towards a benchmark for RGB-D SLAM evaluation ", Proc. of the RGB-D Workshop on Advanced Reasoning with Depth Cameras at Robotics: Science and Systems Conf., 2011.
- [93] Sturm, J., Engelhard, N., Endres, F., Burgard, W., Cremers, D., "A Benchmark for the Evaluation of RGB-D SLAM Systems", Proc. of the International Conference on Intelligent Robot Systems, 2012.

Sažetak

U ovoj doktorskoj disertaciji se istražuje mogućnost primjene deskriptora vizualnih obilježja kao što su boja i tekstura u sustavu za lokalizaciju mobilnog robota zasnovanom na ravninskim segmentima. U disertaciji su objašnjene teoretske osnove i primjeri sustava lokalizacije mobilnog robota zasnovanih na značajkama. Također su objašnjena fizikalna i matematička svojstva vizualnih obilježja te njihove dosadašnje tipične primjene i primjene u sustavima za lokalizaciju mobilnog robota zasnovanog na pojavnosti. Opisan je rad osnovnog sustava za lokalizaciju mobilnog robota koji detektira ravninske segmente na dubinskoj slici snimljenoj 3D kamerom, sparuje ih s ravninskim segmentima prethodno izrađenih 3D modela okoline i na temelju tih parova generira niz hipoteza o položaju robota na hibridnoj metričkotopološkoj karti sastavljenoj od navedenih modela. Generirane hipoteze se evaluiraju na temelju određenog kriterija tako da najviše rangirana hipoteza ima najveću vjerojatnost ispravnosti. Fokus ovog rada su modifikacije navedenog sustava lokalizacije koje omogućavaju korištenje vizualnih obilježja ravninskih segmenata. Modifikacije sustava uključuju novi uvjet sličnosti deskriptora vizualnih obilježja u fazi inicijalnog sparivanja segmenata te modificiranu evaluaciju generiranih hipoteza. ravninskih Svojstva implementiranih deskriptora vizualnih obilježja su ispitivana na konkretnim primjerima RGB-D slika snimljenih Microsoft Kinect senzorom montiranim na mobilnog robota. Provedena je analiza primjenljivosti vizualnih obilježja za redukciju broja parova ravninskih segmenata od kojih će se generirati hipoteze kao i za poboljšanje kriterija evaluacije generiranih hipoteza. Pokus globalne lokalizacije je omogućen korištenjem referentnih podataka inkrementalne lokalizacije provedenim za iste ispitne sekvence slika. Analiza pokusa globalne lokalizacije pokazuje povećanje brzine rada sustava te povećanje točnosti lokalizacije mobilnog robota u slučaju korištenja deskriptora vizualnih obilježja ravninskih segmenata naspram osnovnog sustava zasnovanog samo na geometrijskim svojstvima ravninskih segmenata. Izrađen je eksperimentalni evaluacijski okvir za ispitivanje učinkovitosti sustava lokalizacije zasnovanih na vizualnim obilježjima ravninskih segmenata.

Abstract

This PhD thesis explores the possibilities of applying visual characteristics such as color and texture in a mobile robot localization system based on planar surface segments. Basic principles and applications of feature based mobile robot localization systems are given. Furthermore, the physical and mathematical properties of these visual characteristics are explained. Typical applications involving color and texture along with applications in appearance based localization are also provided. This thesis describes the principles and properties of a basic mobile robot localization system which detects planar surface segments in a depth image acquired by a 3D camera and matches them with previously captured 3D models of the environment. Based on these matches the system generates hypotheses of the robot's pose on a hybrid metric-topological map composed of the aforementioned 3D models. The generated hypotheses are evaluated according to a criterion so that the highest ranked hypotheses are the most probable ones. Modifications of the basic localization system which enable utilizing visual characteristics are the main focus of this thesis. The implemented modifications include a new criterion for visual descriptor similarity and a modified hypothesis evaluation criterion. The characteristics of the implemented visual descriptors are examined on actual examples of RGB-D images captured by a Microsoft Kinect sensor mounted on a mobile robot. Analysis is performed on the applicability of visual descriptors in reducing the number of pairs of planar surface segments which are used for hypothesis generation and in improving the evaluation criterion. A global localization experiment is enabled by utilizing reference data generated by incremental localization applied to the same test series of images. The analysis of this global localization experiment has revealed an increase in speed and localization precision when utilizing visual descriptors compared to only utilizing geometrical constraints of planar surface segments. An experimental evaluation benchmark is developed for mobile robot localization systems based on visual descriptors of planar surface segments.

Životopis

Damir Filko je rođen 23. lipnja 1982. godine u Osijeku. Godine 2006. završio je na Elektrotehničkom fakultetu dodiplomski studij i stekao zvanje dipl. ing. elektrotehnike, obranom diplomskog rada pod nazivom "Precizno mjerenje temperature". Slijedeće godine nakon civilnog služenja vojnog roka zapošljava kao asistent na Zavodu za Automatiku i Procesno računarstvo. Tijekom rada na fakultetu držao je nastavu iz predmeta Operacijski sustavi, Računalni sustavi koji toleriraju kvarove, Programiranje, Programiranje 1, Algoritmi i strukture podataka, Objektno orijentirano programiranje, Meko računarstvo i Industrijska informatika na stručnim i sveučilišnim prediplomskim i diplomskim studijima. Također je bio sumentor na završnim i diplomskim radovima. Glavna područja njegovog znanstvenog i stručnog interesa uključuju medicinske aplikacije, računalni i robotski vid. Bio je istraživač i suradnik na nekoliko MZOŠ RH projekata. Autor i suautor je brojnih članaka u zbornicima međunarodnih konferencija te znanstvenim i stručnim časopisima.

Prilog 1

Kalmanov filter za estimaciju položaja robota korištenjem ravninskih segmenata

Kalmanov filter se u ovom radu koristi kao postupak rekurzivne estimacije relativnog položaja robota $\hat{\mathbf{w}}$ u odnosu na lokalni 3D model okoline. Uz pripadajuću nesigurnost opisanu matricom kovarijance \mathbf{P} , sastoji se od predikcijskog i korekcijskog koraka. Predikcijski korak predstavlja estimaciju trenutnog položaja robota na temelju njegovog prethodnog položaja i podataka dobivenih odometrijom ukoliko se radi o inkrementalnoj lokalizaciji ili praćenju lokalnog položaja robota. Kod globalne lokalizacije nema predikcijskog koraka, već se pretpostavlja da je početni relativni položaj 0 i pripadajuća nesigurnost jako velika. Korekcijski se korak implementiranog Kalmanovog filtera zasniva na jednadžbama (4.26) i (4.27). Radi se o korekciji položaja $\hat{\mathbf{w}}$ i pripadajuće matrice kovarijance \mathbf{P} opisanoj sljedećim izrazima:

$$\mathbf{K} = \mathbf{P} \cdot \mathbf{C}^{\mathrm{T}} \cdot \mathbf{Q}^{-1},$$
$$\hat{\mathbf{w}}' = \hat{\mathbf{w}} - \mathbf{K} \cdot \mathbf{e},$$
$$\mathbf{P}' = (\mathbf{I} - \mathbf{K} \cdot \mathbf{C}) \cdot \mathbf{P},$$

gdje je e pogreška estimacije dana izrazom:

$$\mathbf{e} = \mathbf{h} \left(\mathbf{0}, \mathbf{0}, \hat{\mathbf{w}}; {}^{A}\mathbf{R}_{F}, {}^{A}\mathbf{t}_{F}, {}^{B}\mathbf{R}_{F'}, {}^{B}\mathbf{t}_{F'} \right) = \left[\frac{\left[{}^{B}\mathbf{x}_{F'} \right]^{T} \cdot \mathbf{R} \left(\hat{\mathbf{\phi}} \right) \cdot {}^{A}\mathbf{z}_{F}}{\left({}^{A}\mathbf{t}_{F}^{T} - \left({}^{B}\mathbf{t}_{F'} - \mathbf{t} \right)^{T} \cdot \mathbf{R} \left(\hat{\mathbf{\phi}} \right) \right) \cdot {}^{A}\mathbf{z}_{F}} \right]$$

a **Q** matrica dana izrazom:

$$\mathbf{Q} = \mathbf{E} \cdot \mathbf{C}_{\mathbf{q}} \cdot \mathbf{E}^{\mathrm{T}} + \mathbf{E}' \cdot \mathbf{C}_{\mathbf{q}'} \cdot \mathbf{E}'^{\mathrm{T}} + \mathbf{C} \cdot \mathbf{P} \cdot \mathbf{C}^{\mathrm{T}}$$

C, E i E' su Jacobieve matrice opisane izrazima:

$$\mathbf{C} = \frac{\partial \mathbf{h} \left(\mathbf{0}, \mathbf{0}, \hat{\mathbf{w}}; {}^{A}\mathbf{R}_{F}, {}^{A}\mathbf{t}_{F}, {}^{B}\mathbf{R}_{F'}, {}^{B}\mathbf{t}_{F'} \right)}{\partial \mathbf{w}} \bigg|_{\mathbf{w} = \hat{\mathbf{w}}} = \begin{bmatrix} \begin{bmatrix} {}^{B}\mathbf{x}_{F'} & {}^{B}\mathbf{y}_{F'} \end{bmatrix}^{\mathrm{T}} \cdot {}^{\mathbf{p}}\mathbf{J}_{\phi} \left(\hat{\phi}, {}^{A}\mathbf{z}_{F} \right) & \mathbf{0}^{2\times3} \\ - \left({}^{B}\mathbf{t}_{F'} - \hat{\mathbf{t}} \right)^{\mathrm{T}} \cdot {}^{\mathbf{p}}\mathbf{J}_{\phi} \left(\hat{\phi}, {}^{A}\mathbf{z}_{F} \right) & {}^{A}\mathbf{z}_{F}^{\mathrm{T}} \cdot \mathbf{R}^{\mathrm{T}} \left(\hat{\phi} \right) \end{bmatrix} \\ \mathbf{E} = \frac{\partial \mathbf{h} \left(\mathbf{q}, \mathbf{q}', \hat{\mathbf{w}}; {}^{A}\mathbf{R}_{F}, {}^{A}\mathbf{t}_{F}, {}^{B}\mathbf{R}_{F'}, {}^{B}\mathbf{t}_{F'} \right)}{\partial \mathbf{q}} \bigg|_{\mathbf{q} = \mathbf{0}} = \begin{bmatrix} \begin{bmatrix} {}^{B}\mathbf{x}_{F'} & {}^{B}\mathbf{y}_{F'} \end{bmatrix}^{\mathrm{T}} \cdot \mathbf{R} \left(\hat{\phi} \right) \cdot \begin{bmatrix} {}^{A}\mathbf{x}_{F} & {}^{A}\mathbf{y}_{F} \end{bmatrix} & \mathbf{0} \\ ({}^{A}\mathbf{t}_{F}^{\mathrm{T}} - \left({}^{B}\mathbf{t}_{F'} - \hat{\mathbf{t}} \right)^{\mathrm{T}} \cdot \mathbf{R} \left(\hat{\phi} \right) \cdot \begin{bmatrix} {}^{A}\mathbf{x}_{F} & {}^{A}\mathbf{y}_{F} \end{bmatrix} & \mathbf{0} \\ ({}^{A}\mathbf{t}_{F}^{\mathrm{T}} - \left({}^{B}\mathbf{t}_{F'} - \hat{\mathbf{t}} \right)^{\mathrm{T}} \cdot \mathbf{R} \left(\hat{\phi} \right) \right) \cdot \begin{bmatrix} {}^{A}\mathbf{x}_{F} & {}^{A}\mathbf{y}_{F} \end{bmatrix} & \mathbf{0} \\ \mathbf{E}' = \frac{\partial \mathbf{h} \left(\mathbf{q}, \mathbf{q}', \hat{\mathbf{w}}; {}^{A}\mathbf{R}_{F}, {}^{A}\mathbf{t}_{F}, {}^{B}\mathbf{R}_{F'}, {}^{B}\mathbf{t}_{F'} \right) \\ \partial \mathbf{q}'} \bigg|_{\mathbf{q} = \mathbf{0}} \\ \partial \mathbf{q}' \end{bmatrix}_{\mathbf{q} = \mathbf{0}} = -\mathbf{I}^{3\times3},$$

163

koje vrše propagaciju nesigurnosti, dok je

$${}^{\mathbf{p}}\mathbf{J}_{\boldsymbol{\phi}}(\boldsymbol{\phi},\mathbf{p}) = \frac{\partial \left(\mathbf{R}(\boldsymbol{\Psi})\cdot\mathbf{p}\right)}{\partial \boldsymbol{\Psi}} \bigg|_{\boldsymbol{\Psi}=\boldsymbol{\phi}}.$$

Primjenom ovog koraka dobiva se nova procjena odnosno korigirani relativni položaj $\hat{\mathbf{w}}'$ uz manju pripadajuću nesigurnost **P**'. Stoga, svaki spareni par ravninskih segmenata odnosno par ravninskih segmenata koji zadovolje jednadžbe (4.26) i (4.27) predstavljaju jedno mjerenje pomoću kojega se korigira estimirani relativni položaj $\hat{\mathbf{w}}$ i pripadajuća nesigurnost P.