

Univerza v Ljubljani  
Fakulteta za računalništvo in informatiko

**Matej Artač**

**MODELIRANJE OKOLJA S PANORAMSKIMI  
PREDSTAVITVAMI ZA LOKALIZACIJO IN  
NAVIGACIJO MOBILNEGA ROBOTA**

DOKTORSKA DISERTACIJA

Ljubljana, 2006

Mentor: **prof. dr. Aleš Leonardis**



# Povzetek

Računalniški vid omogoča vizualno učenje okolja mobilnega robota in ocenjevanje robotove trenutne lokacije. Nekatere metode vizualnega učenja predstavijo okolje v obliki zaporedja vizualnih vtisov, zajetih v učnih lokacijah. V fazi lokalizacije trenutni vizualni vtis primerjajo s shranjenimi in tako na podlagi videza določijo trenutno robotovo lokacijo. Te metode omogočajo učinkovito delovanje v poljubnem okolju, vendar omogočajo lokalizacijo le v bližini učnih lokacij. V disertaciji bomo predstavili novo metodo lokalizacije na podlagi videza, ki to omejitev odpravi.

Rešitev, ki jo predlagamo, v postopek lokalizacije vključuje metodo s področja računalniške grafike. Izrisovanje novih slik z elementi referenčnih slik je metoda, namenjena vizualizaciji pogledov v navideznih lokacijah. Pri tem uporablja predstavitev v obliki prostorsko-časovnega volumna referenčnih slik. V disertaciji bomo opisali delovanje izrisovanja novih slik z navzkrižnimi režami. Za postopek bomo uvedli zgoščeno predstavitev, ki s pomočjo metode glavnih komponent zmanjša prostorsko zahtevnost predstavitve.

Predlagani postopek lokalizacije prejme sliko robotovega trenutnega pogleda, nato pa izmed množice slik v navideznih lokacijah izbere najbolj podobno sliko. Shranjene predstavitve ne uporablja neposredno pri primerjavi pogledov, ampak kot vir za izrisovanje približkov slik v navidezni lokaciji znanih koordinat. Njegovo delovanje temelji na preprosti geometrični predpostavki o planarnosti prizora, zato ne potrebuje zahtevnih izračunov tridimenzionalne rekonstrukcije prostora.

Metodo bomo utemeljili s poskusi na umetnih in resničnih slikah v okoljih poljubne oblike. Pokazali bomo primernost njenega delovanja na mobilnih robotih in njeno uspešnost pri ocenjevanju lokacij, ki jih robot pred tem še ni obiskal.



# Abstract

Computer vision enables the mobile robots to perform visual learning of the environment and the estimation of their current location. Some methods for visual learning utilise representations that consist of a series of visual impressions, captured during the exploration of the environment. In the localisation phase they compare the momentary visual impression with those stored in the representation. Thus they estimate their current location by matching the appearance. These methods perform efficiently in an arbitrary environment. However, their location estimation is limited to the vicinity of the actual training locations. In this dissertation we propose a novel method for appearance-based localisation that overcomes this limitation.

The solution we propose is a synergy of a mobile robot localisation technique with a computer graphics technique. The mobile robot acquires a set of images we organise into a spatio-temporal image volume. An image-based rendering technique called the X-slits rendering uses this representation as a source for generating novel views from virtual viewpoints. We compress the representation using the principal component analysis in order to reduce the memory requirements for the representation.

The proposed localisation method acquires an image from the robot's current viewpoint. It then selects the most similar image from a set of virtual viewpoints. In this way we replace the selection of pre-stored views with the rendering of the image approximations at the views with known coordinates. In order to do this, the method does not require any complex 3D reconstruction computations. Instead, it performs on a simple geometrical assumption that the scene is planar.

We demonstrate the performance of the proposed method with our experiments in artificial as well as real environments of arbitrary shape. We show that the method estimates the robot's locations well despite the fact that they are arbitrary and not a part of a training set. We thus show that the proposed method is suitable for the mobile robot localisation tasks.



# Kazalo

Povzetek . . . . .	iii
Abstract . . . . .	v
Kazalo slik . . . . .	x
Kazalo algoritmov . . . . .	xi
<b>1 Uvod</b>	<b>1</b>
1.1 Vid pri učenju in prepoznavanju okolja . . . . .	1
1.2 Načini predstavitve okolja s podatki iz vizualnih senzorjev . . . . .	2
1.2.1 Geometrijske predstavitve . . . . .	3
1.2.2 Predstavitve na podlagi videza . . . . .	6
1.3 Opis problema . . . . .	9
1.4 Sorodna dela . . . . .	10
1.5 Naš pristop . . . . .	12
1.6 Zgradba disertacije . . . . .	14
1.7 Prispevki k znanosti . . . . .	15
<b>2 Metode na podlagi videza</b>	<b>17</b>
2.1 Uvod . . . . .	17
2.2 Globalne značilnice . . . . .	17
2.2.1 Organizacija znanja . . . . .	18
2.2.2 Metoda glavnih komponent . . . . .	18
2.3 Lokalizacija mobilnega robota z globalnimi značilnicami . . . . .	20
2.4 Povzetek poglavja . . . . .	22
<b>3 Predstavitev okolja s prostorsko-časovnim volumnom slik</b>	<b>23</b>
3.1 Uvod . . . . .	23
3.2 Izrisovanje z navzkrižnimi režami za sestavljanje navideznih pogledov	24
3.3 Izrisovanje iz slik s posplošeno geometrijo . . . . .	30
3.4 Izrisovanje navideznih slik iz zgoščenih volumnov . . . . .	34

---

3.5	Lastnosti predstavitve okolice s prostorsko-časovnimi volumni . . . . .	36
3.6	Povzetek poglavja . . . . .	41
<b>4</b>	<b>Lokalizacija mobilnega robota</b>	<b>43</b>
4.1	Uvod . . . . .	43
4.2	Primerjanje slik . . . . .	43
4.3	Lokalizacija . . . . .	45
4.4	Povzetek poglavja . . . . .	47
<b>5</b>	<b>Preskusi</b>	<b>49</b>
5.1	Uvod . . . . .	49
5.1.1	Testni nabori slik . . . . .	49
5.2	Metode na podlagi videza . . . . .	54
5.2.1	Lokalizacija v eni dimenziji . . . . .	54
5.2.2	Lokalizacija v dveh dimenzijah . . . . .	58
5.2.3	Lokalizacija izven dosega učnih slik . . . . .	58
5.3	Izrisovanje navideznih slik . . . . .	60
5.3.1	Primeri navideznih slik . . . . .	60
5.3.2	Kakovost navideznih slik . . . . .	63
5.4	Lokalizacija . . . . .	68
5.4.1	Iskanje najustrežnejšega navideznega pogleda . . . . .	68
5.4.2	Lokalizacija v eni dimenziji . . . . .	70
5.4.3	Lokalizacija v dveh dimenzijah . . . . .	72
5.4.4	Vplivi elementov na različnih globinah prizora na lokalizacijo .	73
5.4.5	Lokalizacija z mobilnim robotom . . . . .	81
5.5	Povzetek poglavja . . . . .	86
<b>6</b>	<b>Zaključek</b>	<b>87</b>
6.1	Obnova disertacije . . . . .	87
6.2	Prispevki k znanosti . . . . .	90
6.3	Razprava . . . . .	91
<b>A</b>	<b>Notacija</b>	<b>93</b>
	<b>Zahvala</b>	<b>97</b>
	<b>Izjava</b>	<b>99</b>

# Slike

1.1	Diagram različnih pristopov k vizualni lokalizaciji . . . . .	11
1.2	Shema poteka predlagane metode . . . . .	13
3.1	Prostorsko-časovni slikovni volumen . . . . .	25
3.2	Geometrija nastanka slike s perspektivno kamero . . . . .	26
3.3	Shema izrisovanja z navzkrižnimi režami . . . . .	27
3.4	Sprememba navpične skale za oddaljenost $d_e$ . . . . .	28
3.5	Prikaz slik iz navideznih pogledov . . . . .	30
3.6	Vpliv parametra $d_e$ na navidezne slike . . . . .	30
3.7	Odpravljanje vpliva odklona od smeri pogleda . . . . .	31
3.8	Poravnava slike odklonjenega pogleda . . . . .	32
3.9	Primeri navideznih slik s posplošeno geometrijo . . . . .	33
3.10	Referenčni žarki pri sestavljanju navideznega pogleda . . . . .	37
3.11	Območja pri predstavitvi s prostorsko-časovnimi volumni . . . . .	38
3.12	Primeri navideznih slik glede na območja dosega . . . . .	39
3.13	Zmanjševanje in povečevanje slike pri izrisovanju iz slik . . . . .	40
4.1	Skica robotovega okolja pri lokalizaciji . . . . .	46
5.1	Prikaz umetnega prizora — zunanje okolje . . . . .	51
5.2	Tloris umetnega okolja, podobnega pisarni . . . . .	52
5.3	Umetno okolje, podobno pisarni . . . . .	52
5.4	Slike nadzorovanega resničnega prizora . . . . .	54
5.5	Klasična enodimenzionalna lokalizacija, perspektivna kamera . . . . .	56
5.6	Klasična enodimenzionalna lokalizacija, panoramska kamera . . . . .	57
5.7	Klasična lokalizacija v dveh dimenzijah, panoramska kamera . . . . .	59
5.8	Poskusi lokalizacije izven območja učnih slik . . . . .	61
5.9	Vpliv parametra $d_e$ na navidezno sliko . . . . .	62

---

5.10	Prikaz izrisovanja navideznih slik iz nabora umetnega prizora . . . . .	63
5.11	Navidezni pogled iz nabora nadzorovanega resničnega prizora . . . . .	64
5.12	Numerična primerjava navidezne slike z resničnimi. . . . .	64
5.13	Koordinate resnične slike, ki je najbolj podobna navideznim; vpliv $d_e$ . . . . .	65
5.14	Primerjava metod . . . . .	67
5.15	Primerjava slikovnih razdalj med resnično in navideznimi slikami . . . . .	70
5.16	Relativne lokalizacijske napake . . . . .	72
5.17	Shema lokacij in lokalizacija za nabor nadzorovanega resničnega prizora . . . . .	73
5.18	Histogram lokalizacijske napake . . . . .	74
5.19	Primer lokalizacije v okolju, podobnem pisarni . . . . .	75
5.20	Uspešna lokalizacija kljub delnem zakrivanju prizora . . . . .	76
5.21	Lokalizacija iz referenčnega nabora 0 . . . . .	77
5.22	Slike pravih in najboljših navideznih pogledov . . . . .	78
5.23	Lokalizacija iz referenčnih naborov 560 in 840 . . . . .	79
5.24	Lokalizacija iz zamaskiranega referenčnega nabora 560 . . . . .	80
5.25	Referenčne in testna slika pri lokalizaciji z mobilnim robotom . . . . .	82
5.26	Lokalizacija z mobilnim robotom . . . . .	83
5.27	Testna in navidezna slika pri lokalizaciji z mobilnim robotom . . . . .	83
5.28	Lokalizacija z dvema pravokotnima naboroma referenčnih slik . . . . .	85
5.29	Primerjava rezultatov navadnega in predlaganega postopka lokalizacije . . . . .	86

# Algoritmi

1	Lokalizacija z metodo glavnih komponent . . . . .	22
2	Izrisovanje z navzkrižnimi režami . . . . .	29
3	Izrisovanje navidezne slike pod ohlapnejšimi pogoji . . . . .	34
4	Izrisovanje navidezne slike iz zgoščene predstavitve . . . . .	36
5	Izrisovanje navidezne slike, ki vrača masko z definiranimi elementi . .	45
6	Lokalizacija neznanega pogleda iz slike . . . . .	47



# Poglavje 1

## Uvod

### 1.1 Vid pri učenju in prepoznavanju okolja

Ljudje imamo nabor čutil, ki nam omogočajo, da zaznavamo dražljaje iz okolja. Z učenjem in izkušnjami smo sposobni iz dražljajev izluščiti informacije in ugotavljati lastnosti okolja. Izmed vseh čutil je vid najpomembnejši, saj nam omogoča kompleksne zaznave okolja in dostop do širokega razpona informacij. Njegovo izgubo lahko sicer do neke mere nadomestimo s kombinacijo drugih čutov, vendar je sam po sebi še vedno najbolj učinkovit pri funkciji spoznavanja okolja, njegove analize in kasnejšega prepoznavanja že znanega okolja.

Med bivanjem v nekem okolju ali gibanjem po njem prejmemo različne vtise, ki nam pomagajo, da si neko okolje zapomnimo. Vtisi so lahko v obliki posameznih predmetov, ki smo jih videli v okolju, splošnih lastnosti okolja ali zaporedja, v katerem so nastali med našim gibanjem skozi okolje. Če nam kasneje nekdo pokaže fotografijo iz istega okolja, lahko v splošnem ugotovimo, v katerem okolju je fotografija nastala in kje znotraj okolja je stal fotograf. Naša predstava o okolju nam omogoča, da spoznamo isti prostor ob drugem obisku, čeprav smo vanj vstopili na drugi strani kot ob prvem obisku.

Mobilni roboti pogosto opravljajo naloge, ki zahtevajo gibanje po nekem okolju ali prostoru. Postopek navigacije skrbi za to, da robot poišče pot iz svojega izhodišča na cilj, se na poti izogne oviram ter pot uspešno zaključi. Za izvedljivost te naloge mora sistem, ki vodi robota, hraniti podatke o okolju ter iz njih znati ugotoviti robotovo trenutno lokacijo.

Ko robota postavimo v nek prostor in ga vključimo ali ga med njegovim delovanjem prenesemo v drug prostor, sistem nima pravih podatkov o robotovi trenutni

lokaciji. Na morebitne pretekle ocene lokacije se pri tem ne more zanašati, ampak mora lokacijo oceniti nanovo. V žargonu takšni situaciji rečemo, da je bil robot ugrabljen (ang. *kidnapped robot problem*). *Lokalizacija* je postopek, ki na vhod dobi meritve trenutnega robotovega okolja, nato pa na podlagi shranjenih podatkov o okolju oceni robotovo trenutno lokacijo. Ta postopek ni pomemben samo na začetku, ko želimo določiti robotovo izhodišče, ampak tudi med samo navigacijo. Postopek navigacije namreč z občasno lokalizacijo preverja pravilnost zahtevanih akcij ter izvaja morebitne popravke.

Za meritve okolice in določanje njenih lastnosti so roboti opremljeni z različnimi senzorji. Ti imajo različne lastnosti in postopke meritev. Aktivni senzorji, kot so zvočni ali laserski merilci razdalje, v okolico sevajo energijo in merijo intenzitete ali zamike povratka energije, s tem pa ugotavljajo razdaljo do predmetov v okolici. Po drugi strani so kamere pasivni senzorji, ki zgolj merijo energijo iz okolice. Njihove meritve so slike, ki jih strojno obdelujemo s postopki računalniškega vida.

Vid pri lokalizaciji mobilnih robotov zavzema vse pomembnejšo vlogo, čeprav še vedno ne prednjači. Razlog za to vidimo v relativno kompleksni uporabi slik kot rezultatov meritev, sploh v primerjavi s senzorji, ki merijo razdaljo do predmetov v okolici. Med robotiki so tako najbolj priljubljene naprave, ki z laserjem merijo razdaljo. Takšen senzor lahko že takoj, ko ga prvič vklopimo, vrača podatke o oddaljenosti do okoliških predmetov. S tem izvemo obliko okolice in do neke mere umeščenost robota v zemljevid prostora. Te informacije lahko razberemo tudi iz slik, vendar na precej bolj kompleksen način.

Uporaba vida na mobilnem robotu tako predstavlja izziv, vendar je hkrati privlačna alternativa ustaljenim tehnikam zbiranja podatkov o okolju. Laserska tipala dobivajo podatke v eni dimenziji, kamera pa običajno v dveh dimenzijah. Poleg tega je lahko sama intenziteta (svetlost, barva) na sliki vir precej bogatejšemu naboru podatkov, kot jih lahko zajamejo ostali senzorji. Poleg tega ponuja računalniški vid dobro osnovo za druge naloge (npr. prepoznavanje predmetov, sledenje uporabnika, določanje lokacije drugih robotov ipd.).

## 1.2 Načini predstavitve okolja s podatki iz vizualnih senzorjev

Ko robot pride v neko novo okolje ali prostor, ga mora spoznati, ga raziskati in se naučiti njegovih lastnosti. Pri tem dobi množico vizualnih vtisov, podkrepjenih s

časovnimi, prostorskimi in drugimi dodatnimi informacijami, ki jih nato organizira v predstavitev okolja. Ta predstavitev nam omogoča, da kasnejše nove vtise uporabi za to, da iz njih sklepa, v katerem okolju in kje znotraj okolja se nahaja. Dobljena predstavitev služi kot zemljevid raziskanega okolja, opremljen s podatki, ki jih je med raziskovanjem in učenjem dobil iz vizualnih senzorjev.

Vizualni senzorji se med seboj razlikujejo po lastnostih, kot je geometrija nastanka slike, obseg valovnih dolžin ipd. Najpogostejši imajo obliko običajne perspektivne kamere, pri katerih goriščna razdalja določa zorni kot zajemanja vizualnih podatkov. Pri sistemih, ki potrebujejo zelo širok zorni kot, je na objektiv kamere nameščena posebna leča, imenovana ribje oko (ang. *fish-eye lens*). Vedno pogosteje pa na mobilnega robota raziskovalci nameščajo tudi *panoramski senzor*, sestavljen iz kombinacije kamere in ustrezno oblikovanega zrcala, na primer perspektivno kamero in hiperbolično zrcalo ali ortografsko kamero in parabolično zrcalo [53]. Panoramska slika vsebuje slikovno informacijo, ki prihaja iz vseh smeri okrog robota. To med drugim pomeni, da dobimo med vrtenjem robota okrog navpične osi zaporedje slik, ki se med seboj razlikujejo le po zamiku slikovnih stolpcev.

Ne glede na izbrano vrsto senzorja dobimo meritev v obliki matrike slikovnih elementov, ki predstavlja sliko nekega pogleda. Ko robot okolico raziskuje, prejema poleg slik še podatke, kot so koordinate pogledov, kjer so slike nastale, oznake prostora ali delov prostora ipd. Naša naloga je, da te podatke skupaj s slikami obdelamo in organiziramo tako, da bomo kasneje iz novih slik lahko sklepali o neznanih koordinatah. V preteklosti so raziskovalci predstavili različne načine obdelave slikovnih in drugih podatkov ter vrste predstavitev za lokalizacijo mobilnih robotov iz slik [8]. Pri obravnavi slik v grobem ločimo med *geometrijskimi pristopi* in *pristopi*, ki temeljijo na *videzu* okolice.

### 1.2.1 Geometrijske predstavitve

Prostor lahko opišemo kot skupek elementov, umeščenih v trirazsežni okvir. Vsak element ima svoje *metrične* koordinate, lokacijo robota pa določimo glede na *geometrijske odnose* med robotom in elementi v okolju. Osnova za te predstavitve je merjenje razdalj od robota do elementov v okolici, kar na svoj način posnema zvočne, infra-rdeče in laserske senzorje. Ti senzorji so pogojevali nastanek predstavitev z mrežo celic, ki hranijo verjetnost prehodnosti območja, ki ga pokriva posamezna celica [37, 55].

Vizualni senzorji merijo svetlobno energijo, ki prihaja iz prostora. Slika nastane

kot projekcija točk iz trirazsežnega prostora v ravnino dveh razsežnosti, pri čemer izgubimo podatke o globini (oddaljenosti od kamere) do točk, preslikanih na slikovno ravnino. Iz posamezne slike tako v splošnem ni mogoče izvesti rekonstrukcije točk v trirazsežnem prostoru, ampak potrebujemo dve ali več slik istega predmeta iz različnih zornih kotov. Pri tem moramo biti sposobni na vsaki sliki določiti tiste točke predmeta, ki jih vsebujejo tudi ostale slike. Običajno se osredotočimo na elemente posamezne slike, ki jih odlikujejo posebne lastnosti, in jih zato lahko odkrijemo, opišemo in med seboj primerjamo. To so lahko posamezne točke, okolice točk ali večja območja, ki jih imenujemo *lokalne značilnice*.

Lokalne značilnice na sliki iščemo s postopki ali operatorji, imenovanimi detektorji značilnic. Ti nam določijo koordinate ali območja na sliki, ki so potencialno zanimiva. Ta območja potem numerično označimo z opisnikom, s katerimi lahko značilnice med seboj primerjamo. Dve značilnici, ki imata podobno vrednost opisnika (se ujemata), imenujemo korespondenčni par. Če je vsaka od njih na svoji sliki, je zelo verjetno, da značilnici predstavljata isti element prostora.

Razvoj detektorjev značilnic je odvisen od namena aplikacije, ki bo značilnice uporabljala. Nekatera okolja imajo takšne značilnice, ki jih je preprosto odkriti. Na slikah pisarniškega in urbanega okolja najdemo številne elemente, kot so robovi in vogali [56], pravokotne oblike, ki predstavljajo vrata in okna [54, 17], meje med tlemi in steno [13], elemente na stropu [12] ali druge navpične elemente v prostoru [14]. Kompleksnejši detektorji so sposobni odkrivati značilnice v bolj splošnih okoljih [36]. Eden od njih je operator za značilnice, neodvisne od povečevanja, zmanjševanja in vrtenja (SIFT, ang. *scale-invariant feature transform*) [31], drugi operator pa na slikah išče stabilna območja (MSER, ang. *maximally stable extremal region*) [33].

Z detektorji lokalnih značilnic torej lahko na slikah predmeta iz različnih zornih kotov poiščemo in povežemo točke, ki se preslikajo na vse slike. Izračun razdalje (3D rekonstrukcija) nato izvedemo s triangulacijo. To lahko naredimo z umerjenim parom kamer (stereo vid), ki točko projicirata vsaka na svojo slikovno ravnino, pri čemer poznamo tako parametre nastanka slike v posamezni kameri kot medsebojno oddaljenost in relativni zasuk kamer [40]. Točke prizora lahko rekonstruiramo tudi, če imamo zaporedje vsaj treh slik prizora iz različnih pogledov, pri čemer nam kamere ni treba umeriti [23, 16].

Običajno dobimo diskreten nabor točk v prostoru, ki lahko dopolnijo podatke iz senzorjev oddaljenosti pri predstavitvi z verjetnostno mrežo prehodnosti [32] ali mrežo zgradijo samostojno [57]. Bolj razširjene rešitve pa raje uporabljajo trirazsežni model, sestavljen iz teh točk [28, 47, 12]. Ti postopki so namenjeni delovanju v

zaprtih prostorih, pisarnah in industrijskih halah, redkeje pa v zunanjem okolju [27, 46]. Geometrijske pristope uporabljajo tudi postopki, ki jih označujemo s kratico SLAM (ang. *simultaneous localisation and mapping*) [29]. Ti postopki vzdržujejo stanje robota in okolice, sestavljene iz zgodovine robotovih premikov po prostoru in opažanj v obliki zaznanih točk v okolici robota (ang. *landmarks*). Po vsakem novem premiku robot prejme nova opažanja o okolici, s katerimi dopolni stanje. To stanje je torej zemljevid, ki ga postopki uporabljajo za to, da hkrati ugotovijo robotovo trenutno lokacijo ter zemljevid dopolnijo z novimi podatki. Predpogoj za delovanje teh postopkov je sposobnost iz vsake lokacije zanesljivo ugotoviti razdalje do točk v okolju. Sistemi, ki merijo razdalje s pomočjo para kamer (stereo SLAM), morajo zato izločiti šum in delovati z meritvami, ki imajo veliko praznega prostora (npr. [48, 9]). Možne so tudi rešitve z eno samo kamero na robotu (enooki SLAM, ang. *bearings-only SLAM*), ki obravnavajo dodatno stopnjo negotovosti pri meritvah [50, 12].

Metode, ki merijo razdalje do elementov v prostoru, odlikuje sposobnost gradnje celostne predstavitve okolja z obiskom relativno majhnega števila učnih lokacij v prostoru. Elementi, ki so vidni iz neke učne lokacije, so namreč lahko vidni tudi iz večje množice drugih lokacij. Zato je pri tej predstavitvi pomembno le, katere elemente spravimo vanjo, ne pa tudi, iz katerih (učnih) lokacij smo te elemente odkrili. Posledično se tudi pri lokalizaciji osredotočimo na iskanje ujemanj vidnih elementov iz trenutne lokacije namesto na določanje lokacij, ki so prispevale k tem elementom.

Model, zgrajen iz množice točk v dvo- ali tridimenzionalnem prostoru je blizu naši predstavi o zemljevidih ter opisu prostora. Poleg tega je enostaven za uporabo v sistemu z mobilnimi roboti, saj ti na nizkem nivoju operirajo z metričnimi enotami prostora. Razdalje do elementov v modelu preprosto prevedemo v razdalje do ovir med gibanjem robota, hkrati pa so dobra osnova za načrtovanje poti ali izgradnje topoloških predstavitev.

Rešitve, namenjene nekemu določenemu okolju, delujejo dobro, saj v njem vedno lahko najdemo dovolj izbranih značilnic, vendar si v praksi želimo večjo fleksibilnost in splošnost njihovega delovanja. Za te detektorje in njihove izpeljanke obstajajo študije o njihovi stabilnosti ob različnih spremembah slik [36], vendar menimo, da so pogoji, pri katerih so avtorji preskušali različne detektorje, premalo rigorozni in realni, hkrati pa podajajo samo relativno primerjavo med detektorji, same uporabnosti pa ne izrazijo.

Poleg težav s splošnostjo uporabe geometrijskih metod v različnih okoljih je tu vprašanje, kako so te sposobne delovati v velikih okoljih, ki zahtevajo zelo velik

model (skalabilnost). Pričakujemo namreč lahko težave s hranjenjem velike količine značilnic ter preveliko obsežnost iskanja ujemanja s trenutnim pogledom. V manj strukturiranem okolju, na primer v naravi, je velika možnost napačnega prepoznavanja nekega področja okolja zaradi podobnosti z drugim področjem (ang. *aliasing*). Poleg tega je 3D rekonstrukcija precej zahteven problem, ki pri slabo pogojenih vhodnih podatkih lahko odpove.

### 1.2.2 Predstavitve na podlagi videza

Slika nekega prizora ni samo skupek elementov, ki jih obravnavamo neodvisno med seboj, ampak nosi veliko informacij tudi kot celota z zabeleženim *videzom* prizora. Na videz vplivajo oblika elementov prizora, njihove odbojne lastnosti ter osvetlitev. S spreminjanjem lokacije in smeri pogleda se videz spreminja, saj s projekcijo na slikovno ravnino posredno beleži geometrijske odnose med kamero in elementi prizora. Osnova za metode, ki uporabljajo predstavitev okolja na podlagi videza, je na eni strani relativna edinstvenost videza posamezne lokacije in na drugi strani podobnost videza pogledov, ki so si dovolj blizu, ob predpostavki neizrojenih pogledov.

Za predstavitev videza lahko uporabimo lokalne značilnice, ki smo jih opisali že v predhodnem razdelku. Njihova vloga je, da s svojo prisotnostjo na sliki predstavljajo lastnosti pogleda. Tudi tu iščemo korespondenčne pare med lokalnimi značilnicami dveh slik. Vsak odkriti par lahko prispeva glas za primerjano sliko, zmagovalka glasovanja pa predstavlja sliko z najbolj podobnim videzom [25].

Drugi avtorji so preskočili iskanje značilnic in vsako sliko že vnaprej razdelili na območja, v katerih metoda izračuna opisnike lokalnih značilnic. Delovanje metode so prikazali z lokalizacijo vozila, ki se vozi po gozdnih in poljskih poteh ob različnih vremenskih pogojih [6]. Ker ob različnih obiskih iste lokacije na poti ni bilo lateralnih premikov, je njihova metoda delovala uspešno. V splošnem pa lahko pri uporabi lokalnih metod za opisovanje videza pričakujemo že opisane težave, saj takšni pristopi uporabljajo iste vrste lokalnih značilnic kot geometrijske metode.

Bolj razširjene metode za obravnavo videza pogledov pa uporabljajo *globalne značilnice*. Te namesto iskanja posebnih elementov slike upoštevajo *celotno sliko*, v predstavitev pa shranijo njen opis, ki ima obliko točke v nekem prostoru značilnic. Predstavitev okolja sestavlja množica točk preslikav učnih slik v tem prostoru značilnic. Kot globalne značilnice lahko uporabimo barvni histogram [5], ki hrani nabor in pojavnost barvnih intenzitet na sliki, izgubi pa podatke o strukturi slike. Večjo mero opisa strukture dobimo s koeficienti po razvoju slike v Fourierovo vr-

sto [20], čeprav s tem še vedno ne ohranimo same prostorske razporeditve elementov. Namesto razgradnje posamezne slike na numerične elemente, ki predstavljajo določen aspekt slike, pa lahko uporabimo intenzitetne vrednosti slike same, vektorje katerih obravnavamo statistično preko celotnega nabora učnih slik. Statistično orodje, ki ga uporabljamo za to, je metoda glavnih komponent (PCA, ang. *principal component analysis*) [41, 58, 38].

Z metodo glavnih komponent iz prostora učnih slik dobimo podprostor lastnih vektorjev, razvrščenih po padajoči stopnji variance podatkov v smeri pripadajoče komponente. Ker manj informativne komponente lahko zavržemo, dobimo kompakten podprostor, v katerem hranimo projekcije učnih slik. Takšna predstavitev zaseda le manjši delež prostora, ki bi ga zasedali izvorni podatki, pri čemer je izguba natančnosti predstavitve minimalna. Raziskovalci so sprva na ta način predstavili nabor slik obrazov in ga uporabili za razpoznavanje oseb iz njihovih slik [58]. Kasneje so na podoben način prepoznavali vrsto in orientacijo predmetov [38]. Na področju robotike je metoda določala položaj in odmik robotske roke s kamero glede na statični predmet [39]. Ta problem pa lahko prevedemo na problem razpoznavanja lokacij iz pogledov, kar dokazujejo metode za lokalizacijo mobilnega robota [21, 3]. Dobljeni podprostor s projekcijami učnih slik odseva prostorske odnose učnih lokacij v prostoru ter odnose glede na videz pogledov [21]. Primerjavo med slikami lahko izvedemo neposredno z merjenjem metrične razdalje parov projekcij slik, to razdaljo pa lahko izrazimo tudi v obliki verjetnosti, da smo v neki lokaciji [26].

Prednosti metode glavnih komponent se kažejo v fleksibilnosti tako pri gradnji kot pri primerjanju pogledov (razpoznavi). Če učne slike nabiramo postopno ali jih je preprosto preveč za izvedljivost razgradnje na glavne komponente, lahko uporabimo postopno metodo gradnje podprostora [15, 3, 52]. Pri razpoznavi pa lahko izkoristimo možnost projekcije podmnožice elementov slik v podprostor značilnic [30], pri čemer nam vhodnih slik ni treba vnaprej razdeliti na podpodročja [43]. To nam omogoča selektivno izločanje motečih delov na slikah in razpoznavo, ki je odporna na delno zakrivanje prizora [21].

S predstavitvijo, ki temelji na videzu okolja, smo precej bolj neodvisni od izbranega okolja in elementov v njem kot pri geometrijskih metodah. To velja zlasti za metode, ki za opis videza uporabljajo globalne značilnice. Za elemente v okolju ni nujno, da so strukturirani, njihova oblika je lahko poljubna, pomembno je le, da posamezen pogled vključuje dovolj elementov, ki omogočajo razlikovanje med ostalimi pogledi.

Če se osredotočimo na uporabo metode glavnih komponent, je uspešnost iska-

nja ujemanja neke neznane slike s slikami iz učnega nabora odvisna od tega, v kakšni meri se vrednosti (intenzitete) slikovnih elementov neznane slike skladajo z vrednostmi istoležnih slikovnih elementov najbolj podobne učne slike. Najboljše ujemanje dobimo s sliko, ki je identična primerjani sliki. Temu se lahko v največji meri približamo, če sta obe sliki zajeti iz točno istega pogleda. V praksi takšnega ujemanja ne moremo doseči, poleg tega v naprednejših aplikacijah z avtonomnimi roboti ne želimo ponovno obiskovati točno istih lokacij, ki smo jih prešli med učenjem. Realno torej lahko pričakujemo, da se bo lokacija pogleda, kjer nastane neznana slika, vedno vsaj malo razlikovala od lokacije najbližjega učnega pogleda. Če bi se znašli v okolju, kjer imajo vsi predmeti samo tanke obrise, bi to pomenilo precej težavno iskanje ujemanja. Vendar večina predmetov, tako naravnih kot umetnih, na slikah zavzema dovolj veliko površino, da lahko pričakujemo dobro ujemanje tudi med slikami pogledov, katerih lokacije so med seboj nekoliko oddaljene.

Domet posamezne učne slike ali območje okrog pripadajoče lokacije, znotraj katerega še lahko najdemo ujemanje z drugim pogledom, torej ni trivialno majhen, ampak obsega neko območje. To območje je odvisno od vrste kamere, lastnosti okolja in smeri premika med paroma pogledov. Za perspektivno kamero velja, da so spremembe videza najmanjše pri premikih v smeri pogleda in največje pri premikih pravokotno na smer pogleda. Domet je pri tem obratno sorazmeren hitrosti spreminjanja videza. Pri panoramski kameri smer gibanja na domet vpliva manj, vendar je ta večji na sredini prostora kot v bližini predmetov.

Kot bomo opisali v 2. poglavju in s preskusi pokazali v 5. poglavju, je domet učnih pogledov pri metodah, ki uporabljajo metodo glavnih komponent, omejen. Posledica tega je, da moramo za dobro predstavitev okolja zajeti učne poglede na gosti mreži. Gostota mreže je sicer odvisna od že prej opisanega dometa posamezne učne lokacije na nekem območju. Nadalje zaradi hranjenja vtisov, ki ne vsebujejo geometrijskih odnosov elementov na sliki, lokacije neznanega pogleda ne moremo določiti natančno, ampak dobimo le diskretne ocene, vezane na učne lokacije. Višjo natančnost lahko dosežemo le pod pogojem, da se neznana lokacija nahaja znotraj dometa več učnih lokacij, saj lahko z interpolacijo v prostoru značilnic izračunamo približek projekcij vmesnih pogledov.

Slabost metod, ki temeljijo na predstavitvi videza in uporabljajo pri tem globalne značilnice, je torej ta, da znajo modelirati videz samo tistih delov okolja, ki jih je robot med raziskovanjem dejansko obiskal. Za zanesljivo in uspešno lokalizacijo s poljubne lokacije zato zahtevajo velik nabor učnih podatkov, ki z lokacijami pokrije celotno okolje.

## 1.3 Opis problema

Rešiti želimo problem učenja in lokalizacije mobilnega robota s pomočjo računalniškega vida. Obstoječe rešitve za lokalizacijo v grobem uvrščamo k pristopom, ki temeljijo na geometrijskih metodah, in pristopom, ki temeljijo na obravnavi videza pogledov. Vsaka vrsta pristopov ima določene prednosti, zaradi njihovih omejitev pa želimo vpeljati izboljšave. Nekateri avtorji so predstavili rešitve, pri katerih se metode obeh pristopov med seboj dopolnjujejo [13, 24], sami pa bomo predlagali metodo, ki delno združuje oba postopka.

Teoretično se geometrijske metode zdijo boljše od tistih, ki temeljijo na videzu, ker z njimi iz peščice pogledov lahko pokrijemo celo vidno območje v nekem prostoru ter dobimo mere razdalj do elementov prizora. Med navigacijo in lokalizacijo smo potem neodvisni od konkretnih učnih lokacij, saj elementi, vidni v neznanih lokacijah, omogočajo določanje razdalj in umeščanje robota v model prostora. Vendar je njihova uporaba preveč vezana na določeno okolje, ki jo pogojuje izbira lokalnih značilnic. Metode z globalnimi značilnicami imajo manj težav pri izbiri okolja, saj so sposobne delovati s precej širšim naborom elementov danega okolja. Vendar je njihova slabost ta, da smo pri lokalizaciji omejeni na neposredno bližino učnih lokacij, posledično pa za zajemanje in hranjenje dobre predstavitve potrebujemo veliko časa in sorazmerno več prostega pomnilnika.

Naše pretekle izkušnje z metodami, ki uporabljajo videz za predstavitev in razpoznavo okolja [3, 21, 1], so pokazale njihovo dobro delovanje. Z njimi je zato smiselno delati tudi v nadaljevanju, vendar v takšnem okviru, ki bi potreboval le manjši delež učnih slik za delovanje v nekem okolju. Razliko v številu učnih slik bomo zapolnili z uporabo prvin geometrijskih metod.

V lokacijah, ki jih med raziskovanjem nismo obiskali, videza in lastnosti pogledov ne poznamo. Te lokacije so s stališča učnega ali referenčnega nabora *navidezne*, saj v modelu okolja neposredno ne nastopajo. V tridimenzionalnem modelu okolja je to razlikovanje manj pomembno, saj je model tak, da lahko dobimo navidezno meritev z izračunom razdalj in kotov do elementov v modelu iz poljubne lokacije znotraj pokritega okolja. Po drugi strani pa model, ki hrani videz učnih lokacij, sam po sebi ni namenjen iskanju ujemanja učnih slik s takšnimi slikami, ki se od učnih občutno razlikujejo. Pri tem velja poudariti, da se dve sliki s stališča teh metod lahko povsem razlikujeta že, če sta pogleda premaknjena oz. eden prikazuje isti prizor bolj od blizu kot drugi. Podobnost takih slik bi lahko ugotovili šele, če bi poznali transformacijo videza, ki se je zgodila zaradi spremembe lokacije pogleda.

Problem, ki ga želimo rešiti, je torej vpeljava zmožnosti lokalizacije mobilnega robota z metodo, ki modelira in uporablja videz okolja tudi v lokacijah, ki jih med raziskovanjem nismo obiskali.

Pogledi v teh lokacijah vsebujejo slike elementov, ki jih vsebujejo tudi slike v predstavitvi. Ti pogledi se po videzu razlikujejo od vsakega učnega pogleda, vendar med njimi obstajajo geometrijske preslikave. Za lokalizacijo neznanega pogleda bi moralo biti dovolj, da ta pogled vsebuje elemente, ki so vidni iz pogledov v predstavitvi, iz geometrijske preslikave med pogledi pa bi lahko sklepali na oceno lokacije neznanega pogleda.

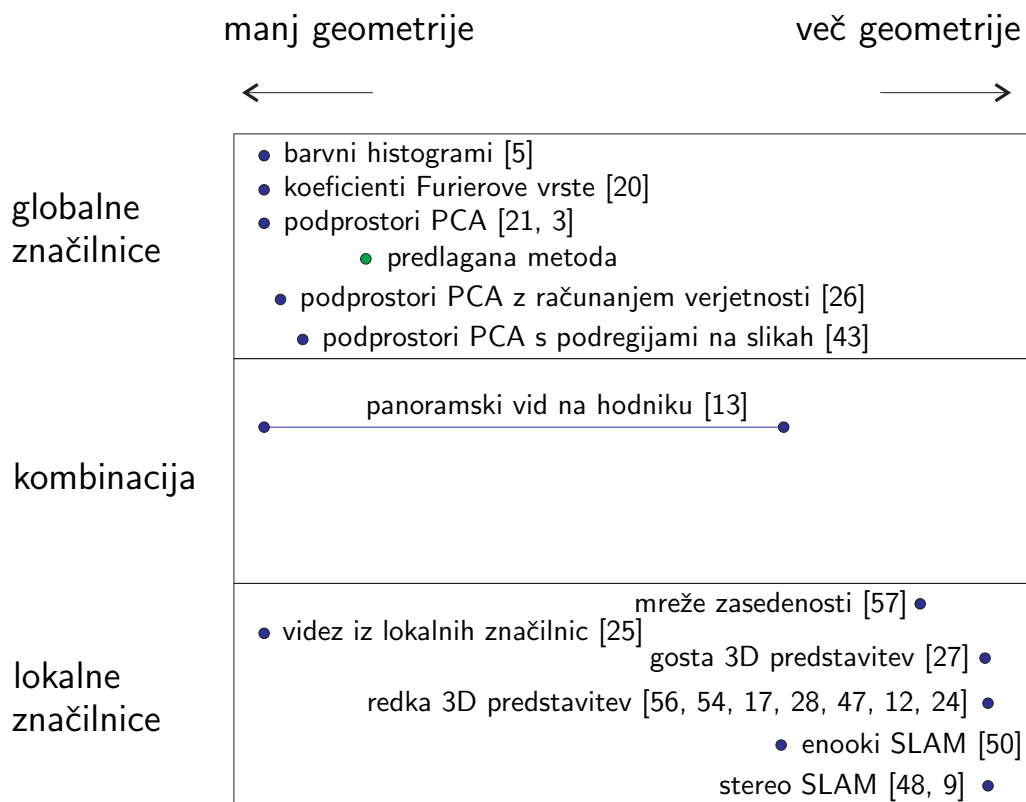
Področje, ki se ukvarja z izrisovanjem novih in drugačnih pogledov iz osnovnih slikovnih gradnikov, je računalniška grafika. Del tega področja vključujejo postopki, ki iz obstoječih slikovnih podatkov zgradijo novo sliko, ki simulira premike od referenčnih (znanih) lokacij v novo (navidezno) lokacijo. Ta nova slika nam lahko služi kot *približek* predstavitve, ki bi jo dobili z obiskom te lokacije, in nam omogoča enakovredno primerjavo s slikami resničnih pogledov.

## 1.4 Sorodna dela

Delo, ki ga bomo opisali v nalogi, predstavlja združitev rezultatov raziskav na področju shranjevanja videza lokacij kot predstavitve okolja in raziskav vizualizacije navideznih pogledov iz skupka slik. Diagram na sliki 1.1 grafično ponazarja umešitev metode, ki jo bomo opisali v disertaciji (**predlagana metoda**), v primerjavi z izborom sorodnih metod. Metode so v diagramu razporejene v eni razsežnosti glede na vrsto značilnic, ki jih uporabljajo, v drugi razsežnosti pa po načinu, kako te značilnice uporabljajo. Levi del diagrama tako ustreza pristopom, ki delujejo na podlagi videza, desni del pa geometrijskim pristopom.

Raziskave, ki so blizu opisanemu pristopu, okolje predstavijo s slikovnimi informacijami iz gosto razporejenih učnih pogledov. Te slike kasneje služijo za primerjavo s slikami pogledov iz neznanih lokacij [20, 60, 26, 21, 3]. Te postopke smo omenili že pri obravnavi metod, ki temeljijo na videzu okolja.

Relativno blizu našim raziskavam so metode, ki zajemajo zaporedje pogledov iz bočno usmerjene kamere ter zgradijo panoramsko predstavitev poti [62, 44], namenjeno kasnejšemu primerjanju globalnih in lokalnih značilnic trenutnega pogleda s sestavljeno panoramo [62] ali iskanju ujemanj v zaporedju slik in njihovih podslik [44]. Obe rešitvi so avtorji uporabili pri vožnji in lokalizaciji po mestnih ulicah. Ker običajno med zajemanjem pride do tresenja, je moč slike poravnati [65] in za-



Slika 1.1: Diagram pristopov k učenju okolice in lokalizacije s podatki iz vizualnih senzorjev; SLAM: sočasna gradnja zemljevida in lokalizacija; PCA: metoda glavnih komponent [41].

poredje regularizirati [45], kar odpravi morebitne razmike med zaporednimi pogledi.

Izrisovanje navideznih pogledov iz slik izhaja iz postopkov za vizualizacijo podatkov in okolja. Pomemben vidik njihovih raziskav (npr. projekt BENOGO<sup>1</sup>) je prepričati uporabnika, da se sprehaja po nekem okolju, ki ga resnici ni obiskal. Iz teh raziskav izhajajo postopki, ki za izrisovanje okolja ne uporabljajo teksturiranih poligonov, ampak sestavijo nove slike iz elementov obstoječih slik [49]. Eden takšnih postopkov deluje z izrisovanjem z navzkrižnimi režami [66], ki ga je moč uporabljati v programih za fotorealistično navidezno resničnost [4].

Pretežni namen raziskav in aplikacij, ki vključujejo izrisovanje navideznih pogledov, pa ostaja vizualizacija iz fizično neobiskanih lokacij [19]. Tudi aplikacije, ki vključujejo mobilne robote, želijo z ustvarjanjem navideznih pogledov le pomagati operaterju [42], za lokalizacijo pa uporabljajo bolj konvencionalne pristope. Edini

<sup>1</sup><http://www.benogo.dk/>

postopek, za katerega vemo, da uporablja lokalizacijo s pomočjo videza iz navideznih pogledov, je predstavil Yagi s sodelavci [59]. Njihova metoda med raziskovanjem po ravni poti zajema enodimenzionalne panoramske slike v višini horizonta. Zaporedje pogledov združi v časovno-prostorsko ravnino. Lokacijo neznanega pogleda, ki je ravno tako enodimenzionalen in v višini horizonta, nato poišče tako, da z minimizacijo določi par krivulj, ki iz shranjene ravnine izreže ustrezen navidezni pogled.

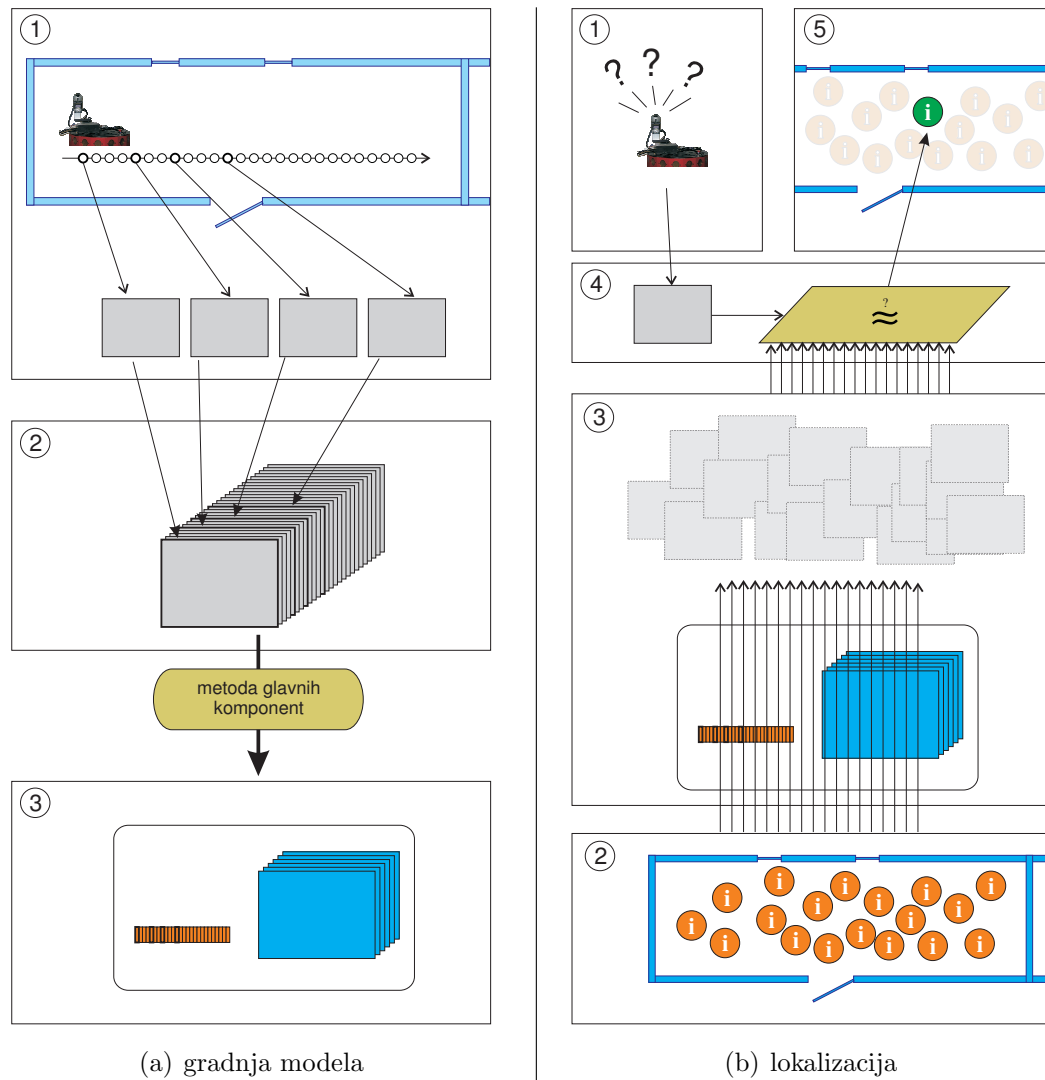
Med sorodne pristope lahko uvrstimo tudi metodo, ki za referenco uporablja 360° panoramo, obogateno s 3D informacijo, pri lokalizaciji pa uporablja videz in ravninske zaplate [7].

## 1.5 Naš pristop

V disertaciji bomo opisali svojo metodo za lokalizacijo mobilnega robota. Metoda združuje pristope za gradnjo predstavitve okolja in lokalizacijo, ki temeljijo na modeliranju videza [3], in postopke za izrisovanje slik navideznih pogledov in zbranih referenčnih slik [49], s čimer se metoda približa geometrijskim pristopom. Del predstavljenih izsledkov smo objavili v [2]. Slika 1.2 povzema naš pristop v obliki diagrama.

V fazi učenja, ki jo prikazuje slika 1.2 (a), robot pridobi podatke o okolju in sistem zgradi njegov model (predstavitev okolja). Podatke dobi tako, da zajame zaporedje bočnih pogledov na enodimenzionalni, ravni poti skozi okolje. Te slike sestavljajo prostorsko-časovni volumen referenčnih slik. Dobljena predstavitev je že v osnovi manjša od predstavitve običajnih metod, ki temeljijo na videzu, saj bi za njih morali zajemanje nadaljevati po vzporednih poteh, dokler ne dobimo dvorazsežne mreže lokacij. Kot prikazuje slika 1.2 (a), pa pri naši metodi volumen preoblikujemo v še kompaktnjšo obliko s postopkom sprotne, postopne gradnje podprostora metode glavnih komponent [3, 1, 52].

Namen zgrajenega modela je njegova uporaba pri lokalizaciji mobilnega robota. Delovanje te faze prikazuje slika 1.2 (b). Robot se nahaja v neznani lokaciji, v kateri lahko zajame sliko pogleda. Naloga postopka je, da s pomočjo predstavitve okolja in na podlagi podane slike ugotovi robotovo trenutno lokacijo. Pri tem podatki iz predstavitve ne nastopajo neposredno pri primerjavi s podano sliko, ampak so vir slikovnih elementov, ki jih postopek izrisovanja z navzkrižnimi režami [66] vključi v slike navideznih pogledov. Navidezni pogled ima znane koordinate lokacije, izrisana slika pa je približek slike iz pogleda te lokacije. Lokalizacija poteka tako, da postopek tvori hipoteze za trenutno lokacijo robota. Za vsako hipotezo izriše navidezno sliko,



Slika 1.2: Shema poteka metode, opisane v disertaciji. (a) Faza učenja: (1) raziskovanje okolja in zajemanje slik pogledov, (2) predstavitev s prostorsko-časovnim volumnom referenčnih slik, (3) zgoščena predstavitev. (b) Faza lokalizacije: (1) robot na trenutni, neznani lokaciji zajame sliko pogleda, (2) izbira hipotez za lokacijo, (3) izrisovanje navideznih slik za hipoteze, (4) primerjanje navideznih slik s sliko robotovega pogleda, (5) zmagovalna hipoteza kot rezultat lokalizacije.

ki jo primerja s sliko iz robotovega pogleda. Izmed navideznih slik izbere najbolj podobno podani, pripadajočo hipotezo pa vrne kot oceno robotove trenutne lokacije.

V primerjavi s sorodno metodo [59] pri našem postopku v svojo predstavitev vključimo celotno višino (sicer perspektivnih) slik pogledov. Tako shranimo precej več vizualnih informacij, kar zmanjša možnost dvoumnih ali napačnih rezultatov (ang. *aliasing*). S tem sicer povečamo porabo prostora, vendar tudi to lahko nadziramo z uporabo ustrezno velikega podprostora metode glavnih komponent [38, 30]. Prednost naše metode je tudi, da za izrisovanje navideznih slik potrebujemo le tri parametre. Nenazadnje pred lokalizacijo ne potrebujemo niti grobega poznavanja izhodišča robota, kar bomo pokazali s preskusi.

Naš pristop uspešno nadgradi koncept vizualne lokalizacije, ki temelji na videzu, in s precej manjšo količino učnih slik doseže večje območje pokrivanja okolja. To smo dosegli s preprostim geometrijskim formalizmom za izrisovanje navideznih slik. Postopek pri tem ne potrebuje množice podatkov o razdalji do posameznih elementov, vidnih na prizoru (kot npr. v [7]), ampak zadošča preprosta predpostavka o okolju, ki ga modeliramo z ravnino, vzporedno slikovni ravnini. O učinkovitosti in uspešnosti pristopa govorijo tudi rezultati preskusov v poglavju 5.

## 1.6 Zgradba disertacije

V prejšnjih razdelkih tega poglavju smo podali uvod v disertacijo, ki predstavi splošno ozadje naših raziskav in področje našega dela. Podali smo pregled nekaterih metod za lokalizacijo mobilnih robotov s pomočjo računalniškega vida, ki so sorodne naši metodi in iz katerih izhaja motivacija za naše raziskave.

V 2. poglavju bomo predstavili postopek učenja in razpoznavne na podlagi videza. Predstavili bomo vlogo globalnih značilnic pri gradnji predstavitve in razpoznavi. Opisali bomo zgoščevanje predstavitve z metodo glavnih komponent in uporabo zgoščene predstavitve pri lokalizaciji mobilnega robota.

V 3. poglavju bomo opisali predstavitev okolja s prostorsko-časovnim volumnom slik in njeno uporabo. Opisali bomo postopek izrisovanja z navzkrižnimi režami, ki za podane vrednosti parametra koordinat navideznega pogleda in oddaljenosti do prizora iz predstavitve izriše približek slike tega pogleda. Predstavili bomo, kako za izrisovanje navideznih pogledov uporabiti predstavitev okolja, zgoščeno z metodo glavnih komponent. Na koncu bomo podali lastnosti predstavitve in navideznih slik, ki jih dobimo z opisanim postopkom.

V 4. poglavju bomo opisali postopek lokalizacije mobilnega robota. Bistvo te

lokalizacije je primerjanje slike neznanega pogleda z navideznimi slikami iz predstavitve, zato bomo najprej uvedli mero za primerjavo slike z navidezno sliko. Nato bomo opisali postopek preiskovanja in ocenjevanja hipotez za pogled, ki z optimizacijo najde najboljšo hipotezo.

V 5. poglavju bomo predstavili opise in rezultate empiričnih preskusov metode. S preskusom običajne metode za lokalizacijo na podlagi videza bomo še v praktičnem smislu motivirali uvedbo novega postopka. Nato bomo preverili delovanje postopka za izrisovanje navideznih slik in ugotovili vpliv vrednosti parametrov na rezultat izrisovanja. Uporabo navideznih slik bomo utemeljili za lokalizacijo neznanih pogledov ter prikazali uspešnost postopka za lokalizacijo mobilnega robota.

V 6. poglavju bomo podali povzetek disertacije, sklepne misli in razpravo, poudarili prispevke k znanosti in našeli odprte teme.

## 1.7 Prispevki k znanosti

V disertaciji bomo opisali razvoj in utemeljitev novega postopka za lokalizacijo mobilnega robota, ki bo združeval pristope, ki temeljijo na obravnavi videza, s prvimi geometrijskih pristopov. Opisani postopek hrani predstavitev okolja in obravnava nove poglede na enak način kot običajni postopki lokalizacije mobilnega robota na podlagi videza. Njegova novost je, da predstavitev vsebuje le delež učnih slik, ki jih potrebujejo običajni postopki, naš postopek pa iz nje sestavi približke poljubnih pogledov v okolju.

Naš postopek lokalizacije bo vključeval pristop s področja računalniške grafike. Z njim bo postopek zmožen slikovne podatke iz predstavitve izkoristiti za tvorbo novih pogledov, ki niso neposredno shranjeni v predstavitvi. Predstavitev tako ne bo več samo množica točk v podprostoru značilnic, ki so razvrščene samo glede na medsebojno vektorsko podobnost.

Pri delovanju robota v velikih okoljih lahko pričakujemo, da bo prostorska zahtevnost predstavitve narasla. Zato bomo uporabili metodo glavnih komponent, s katero bomo zmanjšali porabo pomnilnika za predstavitev. Prirejeni postopek izrisovanja navideznih pogledov bo lahko neposredno uporabljal zgoščeno predstavitev.



# Poglavje 2

## Metode na podlagi videza

### 2.1 Uvod

Postopki učenja in razpoznavanja na podlagi videza predstavljajo preprosto, vendar močno orodje računalniškega vida. Delo z globalnimi slikovnimi značilnicami omogoča reševanje širokega razpona nalog. Ena od takšnih nalog vključuje učenje videza okolja, ki ga raziskuje mobilni robot, in lokalizacija mobilnega robota v raziskanem okolju. V nadaljevanju bomo opisali delovanje teh postopkov.

### 2.2 Globalne značilnice

Sliko na najnižjem nivoju predstavimo s strukturo matrike. Element matrike je skalarna vrednost, ki predstavlja intenziteto žarka, ki osvetli pripadajoče slikovno tipalo v kameri. Pomen posamezne slike je odvisen od rezultatov njene obdelave na višjih nivojih, vendar je določanje njene vsebine zapleten postopek. Po drugi strani je določanje numeričnih lastnosti slike preprosto izvedljivo. Namen globalnih značilnic je ravno ta, da določa numerične lastnosti slik. Z njimi lahko slike označimo, jih organiziramo v predstavitev znanja in jih primerjamo z drugimi slikami.

Najosnovnejša oblika globalne značilnice slik je slikovni vektor slike. Intenzitetne vrednosti (barva) slikovnih elementov določa posamezno sliko, podobne slike pa imajo podobne vrednosti istoležnih slikovnih elementov. Dve sliki tako lahko med seboj primerjamo z evklidsko razdaljo njunih pripadajočih slikovnih vektorjev.

### 2.2.1 Organizacija znanja

Metoda ima svojo učno fazo, v kateri dobi seznam slik s pripisanim pomenom. Ta seznam je podatkovna zbirka znanja, ki povezuje primerke slik z vrednostjo oznak ali atributov. V fazi uporabe znanja potem po vrednosti oznak ali atributov povprašujemo s primerkom slike neznanih vrednosti. Kot odgovor dobimo vrednosti oznak ali atributov, ki so pripisane najbolj podobni shranjeni sliki.

Pri delu z globalnimi značilnicami slik znanje organiziramo tako, da v prostor značilnic preslikamo učne slike, ob vsaki značilnici pa shranimo še dodatne podatke. Dobimo predstavitev, ki naše znanje modelira glede na videz učnih slik, nad modelom pa imamo še strukturo, ki modelu daje pomen. V splošnem je torej pomen od organizacije podprostora značilnic ločen. To nam na eni strani zagotavlja delovanje ne glede na vsebino učnih slik pod predpostavko, da iz različnega videza slik sledijo različne vrednosti atributov. Po drugi strani pa je težko posplošiti videz do te mere, da bi lahko sklepali na vrednosti atributov značilnic, ki jih nimamo v predstavitvi.

### 2.2.2 Metoda glavnih komponent

Metoda glavnih komponent (ang. *principal component analysis*, PCA) [18, 51] je statistična metoda, ki za množico točk v visokodimenzionalnem prostoru določi novo bazo. Ta baza točke dekorelira, njeno razsežnost pa lahko zmanjšamo tako, da dobimo kompaktnjšo predstavitev vhodnih podatkov, pri tem pa naredimo najmanjšo možno napako. V nadaljevanju bomo na kratko opisali delovanje metode na slikovnih podatkih. Pri uporabi matematičnih izrazov bomo sproti opisali pomen posameznih oznak, njihov povzetek pa navajamo v dodatku A.

Imejmo  $n$  slik  $\mathbf{A}_i \in \mathbb{R}^R \times \mathbb{R}^C$ ,  $i = 1 \dots n$ , s  $C$  slikovnimi elementi v širino in  $R$  slikovnimi elementi v višino, ki predstavljajo nabor učnih slik. Slikovno matriko  $\mathbf{A}_i$  lahko enolično predstavimo s slikovnim vektorjem  $\mathbf{a}_i \in \mathbb{R}^{RC}$ . Posamezen element slike  $\mathbf{A}_i$  naslavljamo s številko njegovega stolpca  $c$  in vrstice  $r$ . Ta element je v slikovnem vektorju  $\mathbf{a}_i$  skalar, ki ga označimo kot  $a_{(r,c),i}$ . Zveza  $(r, c)$  v indeksu predstavlja linearni naslov elementa v slikovnem vektorju, ki ga iz naslova vrstice  $r$  in stolpca  $c$  slikovne matrike dobimo z izrazom  $cR + r$ .

Slikovne vektorje  $\mathbf{a}_i$ ,  $i = 1 \dots n$ , ki jim odštejemo srednjo vrednost  $\bar{\mathbf{a}} = \frac{1}{n} \sum_{i=1}^n \mathbf{a}_i$ , zložimo v matriko  $\mathbf{B} = [\mathbf{a}_1 - \bar{\mathbf{a}}, \mathbf{a}_2 - \bar{\mathbf{a}}, \dots, \mathbf{a}_n - \bar{\mathbf{a}}]$ . Stolpci matrike  $\mathbf{B}$  tako predstavljajo množico točk, katere središče je v izhodišču visokodimenzijskega prostora. Nato izračunamo kovariančno matriko  $\mathbf{K} \in \mathbb{R}^{RC} \times \mathbb{R}^{RC}$  matrike  $\mathbf{B}$ :

$$\mathbf{K} = \frac{1}{n} \mathbf{B} \mathbf{B}^\top, \quad (2.1)$$

matriko  $\mathbf{K}$  pa razcepimo na matriko lastnih vektorjev  $\hat{\mathbf{U}} \in \mathbb{R}^{RC} \times \mathbb{R}^n$  in matriko lastnih vrednosti  $\mathbf{\Lambda} \in \mathbb{R}^n \times \mathbb{R}^n$  tako, da velja,

$$\mathbf{K} \hat{\mathbf{U}} = \hat{\mathbf{U}} \mathbf{\Lambda}. \quad (2.2)$$

Matriko  $\hat{\mathbf{U}} = [\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n]$  sestavljajo lastni vektorji  $\mathbf{u}_i \in \mathbb{R}^{RC}$  kovariančne matrice  $\mathbf{K}$ . Lastni vektorji razpenjajo novo bazo prostora, pri čemer ima  $\mathbf{u}_1$  smer največje variance vhodnih podatkov, nanj pravokotni vektor  $\mathbf{u}_2$  kaže v smeri naslednje največje variance, vsak naslednji vektor pa je pravokoten na vse predhodne in kaže v smeri največje variance preostalih smeri. Mero variance posamezne smeri predstavljajo lastne vrednosti  $\boldsymbol{\lambda} = \text{diag}(\mathbf{\Lambda})$ ,  $\boldsymbol{\lambda} = [\lambda_1, \lambda_2, \dots, \lambda_n]$ , za katere velja,  $\lambda_i \geq \lambda_j$ ,  $1 \leq i < j \leq n$ .

Število dimenzij globalnih značilnic nato učinkovito zmanjšamo tako, da zavržemo nekaj zadnjih komponent koordinatnega sistema. Ostane nam  $k \ll n$  lastnih vektorjev, ki razpenjajo podprostor značilnic, ta pa hrani delež variance prvotnega podprostora, ki ga izrazimo z energijo lastnega spektra  $E_k$ :

$$E_k = \frac{\sum_{j=1}^k \lambda_j}{\sum_{j=1}^n \lambda_j}. \quad (2.3)$$

Običajno si izberemo nek delež lastnega spektra  $\tau$ , ki ga želimo obdržati z zmanjšanim podprostorom. V tem primeru izberemo  $k = \arg \min_j \{E_j \mid E_j \geq \tau\}$ .

Rezultat opisanih postopkov je podprostor značilnic, ki ga določa nova in zmanjšana matrika  $\mathbf{U} \in \mathbb{R}^{RC} \times \mathbb{R}^k$  baznih vektorjev  $\mathbf{U} = [\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k]$  in katerega izhodišče je v točki  $\bar{\mathbf{a}}$ . V ta podprostor projiciramo učne slike  $\mathbf{a}_i$  in dobimo točke  $\mathbf{w}_i$ :

$$\mathbf{w}_i = \mathbf{U}^\top (\mathbf{a}_i - \bar{\mathbf{a}}). \quad (2.4)$$

Dobljeni vektorji  $\mathbf{w}_i$  so globalne značilnice, ki v podprostoru značilnic predstavljajo učne slike  $\mathbf{A}_i$ . Teh nam zdaj ni več treba hraniti, saj lahko njihove približke rekonstruiramo:

$$\mathbf{a}_i \approx \mathbf{U} \mathbf{w}_i + \bar{\mathbf{a}}. \quad (2.5)$$

Rekonstrukcijska napaka  $r_i = \|\mathbf{U} \mathbf{w}_i + \bar{\mathbf{a}} - \mathbf{a}_i\|$ , ki jo dobimo zaradi zmanjšanja podprostora, je zaradi konstrukcije podprostora najmanjša možna za podano množico

$\mathbf{a}_i$ ,  $i = 1, \dots, n$ . Hkrati pa zgoščena predstavitev omogoča več učnih podatkov pri dani količini pomnilnika.

Pri implementaciji je računsko učinkoviteje namesto izraza (2.1) uporabiti notranji produkt,

$$\mathbf{K}' = \frac{1}{n} \hat{\mathbf{B}}^\top \hat{\mathbf{B}}, \quad (2.6)$$

katerega lastne vrednosti so identične lastnim vrednostim kovariančne matrike  $\mathbf{K}$ , iz lastnih vektorjev  $\mathbf{u}'_i$ ,  $i = 1, \dots, n$  matrike  $\mathbf{K}'$  pa dobimo lastne vektorje  $\mathbf{u}_i$  z izrazom,

$$\mathbf{u}_i = \frac{\mathbf{B}\mathbf{u}'_i}{\sqrt{n\lambda_i}}. \quad (2.7)$$

Že zgrajeno predstavitev lahko vedno dopolnimo, če dobimo nove učne podatke [52, 1]. Postopna metoda gradnje, ki jo uporabimo pri tem, pa lahko pri veliki količini učnih podatkov omogoči sicer neizvedljiv izračun predstavitve.

V času razpoznavne sistem dobi sliko  $\mathbf{Q}$ , za katero mora izmed učnega nabora poiskati najbolj podobno sliko. Namesto neposrednega primerjanja slik najprej vektorsko obliko  $\mathbf{q}$  slike  $\mathbf{Q}$  projiciramo v podprostor značilnic,

$$\mathbf{w}_\mathbf{q} = \mathbf{U}^\top(\mathbf{q} - \bar{\mathbf{a}}), \quad (2.8)$$

nato pa poiščemo najbližjo projekcijo učnih slik,

$$i = \arg \min_j \|\mathbf{w}_\mathbf{q} - \mathbf{w}_j\|. \quad (2.9)$$

Če imajo torej slike, ki predstavljajo podoben pomen, podoben videz, bodo njihove projekcije v podprostoru spremenljivk blizu, indeks  $i$ , ki ga vrne izraz (2.9), pa bo videz slike  $\mathbf{q}$  povezal s pomenom učne slike  $\mathbf{a}_i$ .

## 2.3 Lokalizacija mobilnega robota z globalnimi značilnicami

Postopek, opisan v predhodnem razdelku, predstavlja osnovo za rešitve metod razpoznavanja iz slik. V zbirki slik obrazov ljudi je tako vsaki sliki pripisano ime osebe, ki jo slika predstavlja. Iz slike obraza lahko prepoznamo osebo, ki jo slika prikazuje [58]. Tudi predmeti se med seboj razlikujejo po videzu, zato se lahko njihovih

videzov naučimo in jih kasneje prepoznavamo. Videz predmeta se glede na njegovo orientacijo spreminja, zato se lahko sistem nauči videza vsakega predmeta v različnih orientacijah. Pri razpoznavi tako ne izvemo le, kateri predmet prikazuje slika, ampak tudi njegovo orientacijo [38].

S stališča postopka za učenje in prepoznavanje iz videza je tudi okolje, v katerem se giblje mobilni robot, neke vrste predmet, katerega videz se spreminja glede na gibanje robota.

Med delovanjem mobilnega robota mora sistem vedeti, v katerem prostoru se robot nahaja in kje v tem prostoru je njegova trenutna lokacija. Videz okolice se z gibanjem robota spreminja, pri čemer je videz pogledov iz lokacij podoben v njihovi ozki okolici, med seboj oddaljene lokacije pa imajo različen videz pogledov. To pomeni, da lahko postopke za razpoznavanje na podlagi videza uporabimo tudi za določanje lokacije robota [3, 21].

Lokacijo mobilnega robota označimo s koordinatami znotraj prostora in orientacijo. Med raziskovanjem okolja robot zajema slike pogledov in hrani koordinate lokacij, ki jih je izmeril. Sistem s paketno ali postopno [1, 52] metodo glavnih komponent zgradi podprostor globalnih značilnic zajetih (učnih) slik. Tudi v tem primeru dobimo podprostor značilnic, ki so razporejene glede na podobnost, medtem ko moramo same metrične odnose pripadajočih lokacij zanje hraniti posebej.

V fazi lokalizacije robot zajame sliko iz trenutnega pogleda. To potem projicira v podprostor značilnic učnih slik in poišče najbližjo učno projekcijo. Koordinate, pripisane odkriti projekciji, predstavljajo oceno lokacije trenutnega pogleda. Algoritem 1 opisuje ta postopek.

Kot vidimo iz algoritma 1, je dobljena ocena diskretna in vezana na učne lokacije. Boljši približek lahko dobimo, če med točkami  $\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_n$  izračunamo interpolacijo glede na parametre  $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n$  [21]. Ta interpolacija je možna samo, če so lokacije učnih pogledov dovolj goste, da so si slike sosednjih pogledov podobne.

Posledica tega, da postopek za oceno lokacije izbira le med učnimi lokacijami, je tudi ta, da moramo za uspešno lokalizacijo robota v neki lokaciji že med raziskovanjem poskrbeti, da bo robot to lokacijo obiskal in sliko pogleda vključil v predstavitev. Iz same predstavitve namreč ni mogoče sklepati na to, kakšne bi bile globalne značilnice v lokaciji, ki je ni v učni množici. Hkrati ni mogoče pravilno oceniti lokacije neznanega pogleda, za katerega je slika premalo podobna katerikoli učni sliki. V danem okviru je namreč mera podobnosti vezana na nizkonivojske vrednosti slikovnih elementov. Podobnost je tako lahko majhna za par slik, ki sta si podobni na visokem nivoju, medtem ko se na nizkem nivoju razlikujeta po zamiku med prika-

---

**Algoritem 1** Lokalizacija mobilnega robota z metodo glavnih komponent.

---

**Vhod:** *Predstavitev učnih slik:* matrika baznih vektorjev  $\mathbf{U}$  podprostora značilnic, srednja vrednost  $\bar{\mathbf{a}}$  globalne značilnice  $\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_n$  učnih slik, lokacije pogledov učnih slik  $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n$ ; slika  $\mathbf{Q}$ .

**Izhod:** ocena lokacije  $\mathbf{x}_q$  vhodne slike  $\mathbf{Q}$ .

- 1:  $\mathbf{q} \leftarrow$  vektorska oblika  $\mathbf{Q}$
  - 2:  $\mathbf{w}_q \leftarrow \mathbf{U}^\top(\mathbf{q} - \bar{\mathbf{a}})$  // Preslikava  $\mathbf{q}$  v podprostor učnih značilnic
  - 3:  $d_{\min} \leftarrow \infty$  // Trenutna najmanjša razdalja
  - 4:  $i_{\min} \leftarrow -1$  // Indeks trenutne najboljše lokacije
  - 5: **za vse**  $i \in \{1, 2, \dots, n\}$  **izvedi**
  - 6:    $d \leftarrow \|\mathbf{w}_q - \mathbf{w}_i\|$  // Slikovna razdalja med učno sliko z indeksom  $i$  in sliko  $\mathbf{Q}$
  - 7:   **če** velja  $d < d_{\min}$ , **potem**
  - 8:      $i_{\min} \leftarrow i$
  - 9:      $d_{\min} \leftarrow d$
  - 10:   **konec če**
  - 11: **konec za**
  - 12:  $\mathbf{x}_q \leftarrow \mathbf{x}_{i_{\min}}$
- 

zanimi elementi ali njihovi povečavi (zasuk kamere, približevanje ali oddaljevanje). Konkretno primere navedenih težav bomo s preskusi prikazali v 5. poglavju.

## 2.4 Povzetek poglavja

Globalne značilnice numerično opisujejo videz slike, kar omogoča obravnavo slik in njihovo organizacijo v obliki znanja. Značilnice v predstavitvi služijo kot povezava med videzom in naučenimi podatki. Metoda glavnih komponent omogoča zmanjšanje porabe pomnilnika za dano predstavitev za ceno minimalne izgube natančnosti predstavitve. S takšno predstavitvijo lahko modeliramo videz okolja iz množice učnih pogledov. Lokalizacijo mobilnega robota v okolju izvajamo z iskanjem sliki trenutnega pogleda najbolj podobne učne slike iz naučene predstavitve.

## Poglavje 3

# Predstavitev okolja s prostorsko-časovnim volumnom slik

### 3.1 Uvod

Mobilni robot, med raziskovanjem okolja zbira podatke o tem okolju, ki jih dobi iz svojih senzorjev. To zaporedje podatkov moramo obdelati tako, da dobimo primerno predstavitev raziskanega okolja. Predstavitev okolje modelira in omogoča, da lahko po njej kasneje poizvedujemo ali opravljamo analize na višjih nivojih. Pri naših raziskavah želimo s pomočjo predstavitve okolja rešiti problem lokalizacije. Za podano sliko želimo ugotoviti lokacijo pogleda, iz katerega je slika nastala.

Preden začnemo reševati ta problem, najprej opišimo njemu obratni problem: za podano lokacijo pogleda dobiti približek slike, ki bi jo dobili iz te lokacije. To je problem, ki ga rešuje področje računalniške grafike. Aplikacije navidezne resničnosti težijo k temu, da na čim bolj realističen način uporabniku prikažejo neko okolje, po katerem se uporabnik na videz giblje. To okolje je v splošnem lahko resnično ali izmišljeno, vendar mora biti uporabniku prikazano dovolj prepričljivo. Da to dosežejo, morajo biti sposobne pokazati okolja in prizore pravih oblik, elementi prizorov pa morajo biti v takšnih geometrijskih odnosih, da odražajo pogled iz uporabnikovega gledišča v okolju. Razvoj računalniške grafike je prinesel postopke, ki znajo iz zbirke slik izveči slikovne elemente in jih sestaviti v novo sliko [49]. Nova slika tako vsebuje elemente, ki jih uporabnik prepozna, od samega postopka pa je odvisno, kako pravilna je ta slika. Ljudje pri ogledovanju slik dopuščamo

manjša odstopanja in napake, medtem ko nas večje napake (napačna razmerja ali odnosi elementov, ponavljajoči se ali manjkajoči deli) motijo. Postopki morajo zato uporabniku slike izrisovati čim bolj pravilno, kar pomeni, da je približek slike nekega pogleda zelo podoben pravi sliki iz tega pogleda.

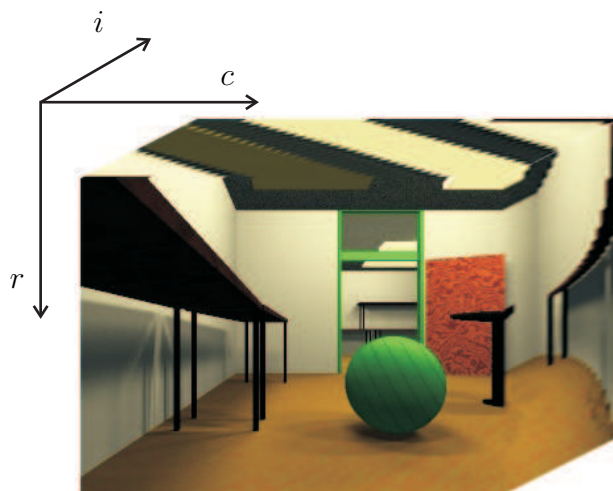
Na prvi pogled opisano področje z lokalizacijo mobilnega robota nima veliko skupnega. Robot zajame sliko, nato pa po vrsti prikliče vsako sliko, za katero ima v svoji predstavitvi shranjeno lokacijo. Vendar je ta postopek zelo podoben poskusom vizualizacije nečesa znanega z namenom identificiranja neznanega. Če torej postopek neposrednega priklica slike iz predstavitve nadomestimo s postopkom izrisovanja približkov slike, s tem pridobimo možnost ocenjevanja lokacij iz *vezne množice lokacij*.

Naš namen je torej v postopek lokalizacije vgraditi postopek s področja računalniškega vida. S stališča lokalizacije mobilnega robota mora postopek uporabiti predstavitev iz podatkov, ki jih lahko pridobi mobilni robot. Parametri izrisa morajo biti takšni, da njihove vrednosti odsevajo vsebino navidezne slike. Te zahteve izpolnjuje postopek za izrisovanje iz slik (ang. *image-based rendering, IBR*), ki ga bomo opisali v nadaljevanju.

## 3.2 Izrisovanje z navzkrižnimi režami za sestavljanje navideznih pogledov

Pri naših raziskavah smo izbrali postopek izrisovanja navideznih pogledov, ki temelji na postopku izrisovanja z navzkrižnimi režami (ang. *X-slits rendering*) [66]. Ta postopek uporablja preprost geometrični formalizem, ki ne zahteva poznavanja oblike prizora ali izračuna 3D rekonstrukcije iz slik. Njegov izhod je slika, ki na realističen način modelira premike po prostoru navideznih lokacij. V nadaljevanju bomo predstavili geometrične lastnosti in delovanje postopka, ki se nanaša na navadno perspektivno kamero. Splošnosti postopka s tem ne zmanjšujemo, saj ga lahko z ustreznimi spremembami uporabimo tudi z drugačno vrsto kamer, na primer s kamero, ki uporablja zelo širokokoten objektiv (ribje oko, ang. *fish-eye lens*), panoramsko zrcalo [4] ali katerokoli drugo vrsto kamere s poznano geometrijo.

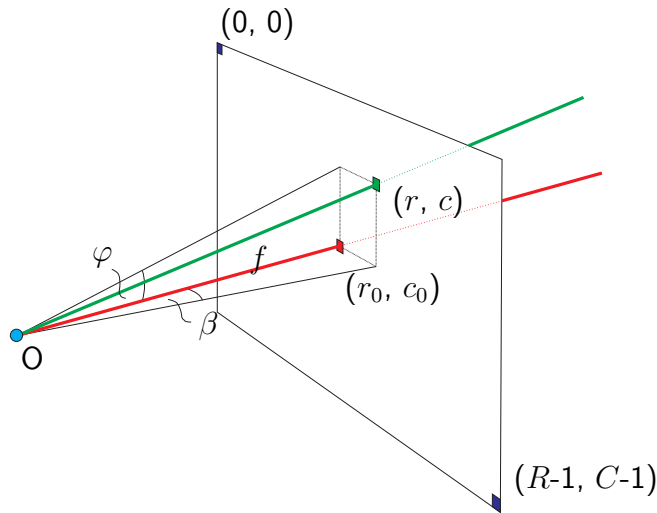
Vir slikovnih elementov za izrisovanje z navzkrižnimi režami je zaporedje slik, zloženih v *prostorsko-časovni slikovni volumen*. Slikovni volumen je kvader, katerega prvi dve razsežnosti predstavljata slikovne koordinate, tretja razsežnost pa čas oz. zaporedje slik. Sestavljen je iz diskretnih slikovnih elementov, ki predstavljajo za-



Slika 3.1: Prostorsko-časovni slikovni volumen sestavljajo slikovni elementi, razporejeni v prostoru vrstic  $r$ , stolpcev  $c$  in zaporednih številik slik (časa)  $i$ .

pis *žarkov* na slikovni ravnini kamere. Na sliki 3.1 vidimo primer slikovnega volumna in njegove razsežnosti. V nadaljevanju bomo ta volumen uporabili za ilustracijo opisanih postopkov. Prerezi volumna z ravnino, pravokotno na os časovne razsežnosti, nam dajo referenčne slike, ki sestavljajo volumen. Z različnimi drugačnimi prerezi dobimo nove, navidezne poglede na prizor.

Označimo časovno zaporedje  $n$  slik širine  $C$  stolpcev in višine  $R$  vrstic, ki tvorijo volumen, z  $\mathbf{A}_1, \mathbf{A}_2, \dots, \mathbf{A}_n$ . Vsaka od slik  $\mathbf{A}_i \in \mathbb{R}^R \times \mathbb{R}^C$ ,  $i = 1 \dots n$ , je matrika, sestavljena iz slikovnih elementov  $a_{r,c,i}$ , ki jih naslavljamo s koordinatami vrstice  $r$  in stolpca  $c$ . Levi zgornji slikovni element ima tako koordinati  $(0, 0)$ , desni spodnji element pa  $(R - 1, C - 1)$ . Slikovni elementi nosijo podatke o barvi ali intenziteti svetlobnega žarka, ki na mestu slikovnega elementa seka slikovno ravnino. Zvezo med žarki in slikovnimi elementi določa geometrija nastanka slike. Pri perspektivni kameri, ki jo ponazarja slika 3.2, se vsi žarki sekajo v točki goriščnici. Razdaljo med goriščnico in slikovno ravnino označimo z goriščno razdaljo  $f$ . Žarek, ki slikovno ravnino seka pod pravim kotom, določa središče slike. Pripadajoči slikovni element ima slikovne koordinate  $(r_0, c_0)$ . Žarek, ki prispeva k nekemu poljubnemu slikovnemu elementu  $(r, c)$ , opišemo s premico, usmerjeno pod kotom  $\beta$  v vodoravni smeri (azimut) in kotom  $\varphi$  v navpični smeri. Zvezo med koordinatami slikovnega elementa in smerjo žarka določata naslednja izraza:



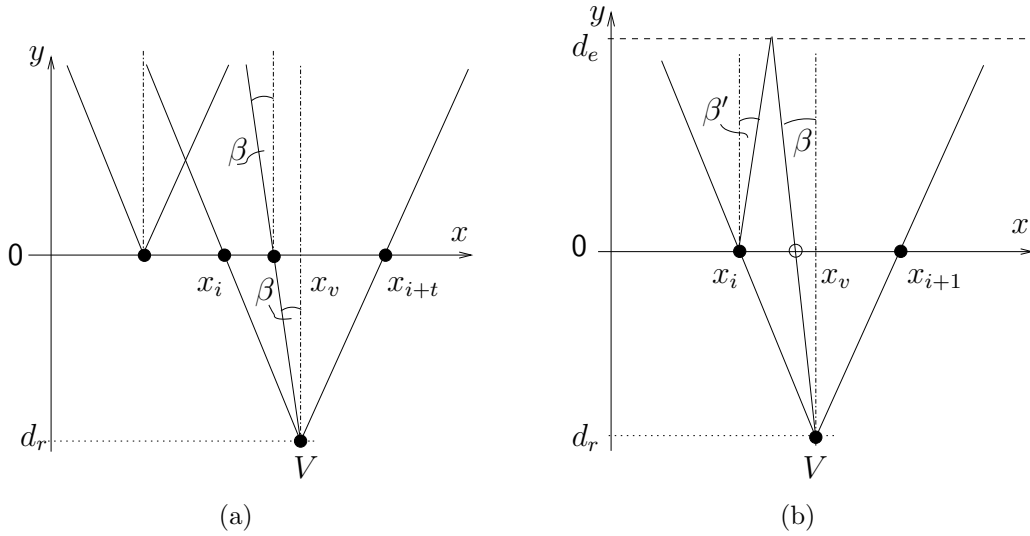
Slika 3.2: Geometrija nastanka slike na slikovni ravnini perspektivne kamere.

$$\beta = \tan\left(\frac{c - c_0}{f}\right), \quad (3.1)$$

$$\varphi = \tan\left(\frac{r - r_0}{f}\right). \quad (3.2)$$

Slike za referenčni slikovni volumen zajamemo iz zaporedja kolinearnih lokacij z relativno oddaljenostjo  $x_1, x_2, \dots, x_n$ , vendar s pogledi, usmerjenimi pravokotno na pot. S tem si namreč zagotovimo, da bomo med gibanjem zajeli največ informacij o okolici, saj se njen videz v smeri pravokotno na gibanje spreminja v največji meri. To naredimo tako, da kamero na vozilu ali robotu usmerimo bočno na smer gibanja. V nadaljevanju bomo torej predpostavili, da ležijo vse referenčne lokacije na skupni premici in da je pogled kamere usmerjen pod kotom  $90^\circ$  od smeri gibanja. Odstopanja od te predpostavke bomo obravnavali v razdelku 3.3.

Postopek izrisovanja z navzkrižnimi režami deluje tako, da iz referenčnih slik izbere slikovne stolpce in jih vstavi v ciljno sliko. Novo sliko torej sestavi iz elementov referenčnih slik. Ker postopek izbira celotne stolpce slikovnih elementov, je pri smereh žarkov pomemben le njihov vodoravni kot  $\beta$ . Naj bo koordinatni sistem ravnine, po kateri se gibljemo, postavljen tako, da njegova os  $x$  sovpada s premico nosilko referenčnih lokacij. Referenčno sliko  $\mathbf{A}_i$  smo zajeli v točki  $(x_i, 0)$ , smer pogleda pa je vzporedna z osjo  $y$  koordinatnega sistema. Elementi prizora, ki jih prikazujejo referenčne slike, so na polravnini  $y > 0$ . Točka  $V$  naj bo navidezna lokacija s koordinato  $x_v$  in oddaljenostjo  $d_r$  od premice. Sestaviti želimo sliko  $\mathbf{V}$ ,



Slika 3.3: (a) Izrisovanje z navzkrižnimi režami in (b) normalizacija pri globini  $d_e$ .

ki predstavlja približek pogleda iz točke  $V$ , pri čemer je pogled usmerjen vzporedno s pogledi referenčnih slik. Slika 3.3 (a) ponazarja potek sestavljanja navideznega pogleda. Za vsak žarek  $z$  vodoravnim kotom  $\beta$  v navideznem pogledu vzorčimo ustrezen žarek  $\beta$  referenčnega pogleda v točki presečišča žarka s premico poti.

Nabor referenčnih lokacij je diskreten, zato na večini presečišč žarkov s premico poti nimamo referenčne slike. Izbrati moramo približek, žarek  $\beta'$ , iz referenčnega pogleda  $x_i$ , ki je najbližji presečišču žarka  $\beta$  in premice poti. To naredimo tako, da na razdalji  $d_e$  od te premice določimo navidezno ravnino, nato pa približek  $\beta'$  izberemo kot projekcijo presečišča žarka  $\beta$  z navidezno ravnino na sliki v lokaciji  $x_i$ . Slika 3.3 (b) ponazarja takšen primer.

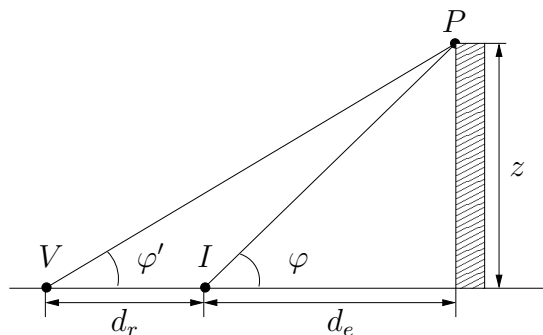
Formalno si za prizor ali del prizora izberemo razdaljo  $d_e$ , ki jo imenujemo *normalizacijska globina*. Za vsak žarek  $\beta$  navidezne slike poiščemo referenčni pogled s koordinato  $x_i$  tako, da velja

$$i = \arg \min_j |x_j - (x_v - d_r \tan(\beta))|, \quad (3.3)$$

nato pa iz pogleda  $x_i$  izberemo žarek  $\beta'$ , ki predstavlja približek žarka  $\beta$ :

$$\beta' = \arctan \left( \frac{(d_e - d_r) \tan(\beta) - (x_v - x_i)}{d_e} \right). \quad (3.4)$$

Izraz (3.4) poskrbi, da navidezno sliko sestavljajo stolpci, ki predstavljajo najboljše približke žarkov za navidezno sliko. Upoštevati moramo še spremembe v



Slika 3.4: Sprememba navpične skale za oddaljenost  $d_e$ .

oddaljenosti od prizora, ki so odvisne od vrednosti  $d_e$  in  $d_r$ . To naredimo tako, da žarke z navpičnim kotom  $\varphi$  (vrstice) preslikamo v žarke z navpičnim kotom  $\varphi'$ . Skico te spremembe prikazuje slika 3.4. Preslikavo izračunamo kot

$$\varphi' = \arctan\left(\frac{d_e \tan(\varphi)}{d_e - d_r}\right). \quad (3.5)$$

Slika se torej pri približevanju k prizoru v navpični smeri povečuje, pri oddaljevanju pa zmanjšuje.

Celoten postopek izrisovanja navideznih slik opisuje algoritem 2. Glavna zanka iterira preko vseh ciljnih stolpcev navidezne slike  $\mathbf{V}$ . Koordinate slikovnih elementov referenčnih slik določimo v vrsticah 5 in 8, preslikavo v navidezno sliko pa izvedemo v vrstici 10. Koordinate slikovnih elementov so sicer iz množice naravnih števil, zato moramo pri preslikavi uporabiti interpolacijo po slikovnih elementih referenčne slike  $\mathbf{A}_i$ . Če vrednosti  $c'$  in  $r'$  zaokrožujemo, dobimo interpolacijo, ki poišče najbližjega sosedu.

Algoritem ima v vrstici 9 vgrajen pogoj, ki pred preslikavo vrednosti referenčnega slikovnega elementa v navidezno sliko preveri, če so koordinate izvornega slikovnega elementa v mejah referenčne slike. Ta pogoj ni vedno izpolnjen, kar povzroči, da določeni slikovni elementi navidezne slike niso definirani. Tipična situacija, ki pripelje do nedefiniranih območij navidezne slike, je pri vrednostih  $d_r < 0$ , ko bi potrebovali več referenčnih slikovnih elementov po višini, kot jih imamo na razpolago (npr. slika 3.5).

Postopek izrisovanja z navzkrižnimi režami torej zahteva tri parametre za izris navideznega pogleda. Parametra  $x_v$  in  $d_r$  predstavljata koordinati navidezne lokacije. Tretji parameter,  $d_e$ , pa predstavlja okolje oz. prizor, ki ga prikazujejo referenčne slike v volumnu. Razumemo ga lahko kot oddaljenost do prizora. Če

---

**Algoritem 2** Izrisovanje z navzkrižnimi režami.

---

**Vhod:** *Prostorsko-časovni volumen referenčnih slik:* slike  $\mathbf{A}_1, \mathbf{A}_2, \dots, \mathbf{A}_n$ , koordinate  $x_1, x_2, \dots, x_n$ ; *parametri kamere:* goriščnica  $f$ , središče slike  $(c_0, r_0)$ ; *parametri navideznega pogleda:* koordinati  $x_v$  in  $d_r$ , normalizacijska razdalja  $d_e$ .

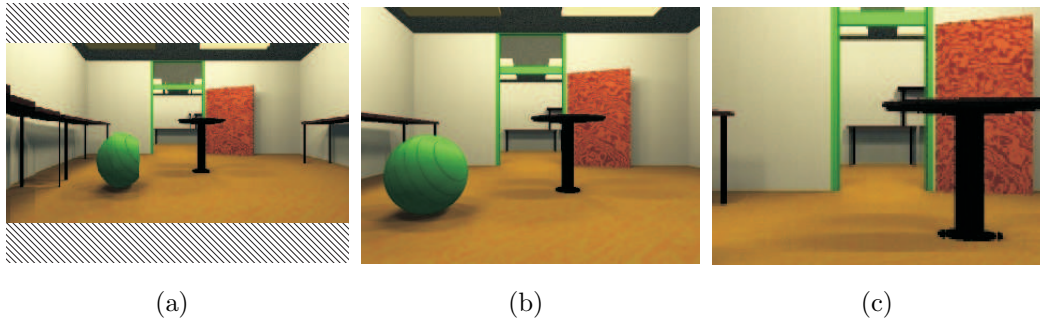
**Izhod:** navidezna slika  $\mathbf{V}$ .

- 1:  $k \leftarrow 1 - \frac{d_r}{d_e}$  //koeficient povečave/pomanjšave
  - 2: **za vse stolpce  $c$  navidezne slike izvedi**
  - 3:  $\beta \leftarrow \arctan\left(\frac{c-c_0}{f}\right)$  //Pretvori v kote.
  - 4:  $i \leftarrow \arg \min_j |x_j - (x_v - d_r \tan(\beta))|$  //Poišči referenčno sliko.
  - 5:  $\beta' \leftarrow \arctan\left(k \tan(\beta) - \frac{(x_i - x_v)}{d_e}\right)$  //Določi približek žarka.
  - 6:  $c' \leftarrow f \tan(\beta') + c_0$  //Izberi stolpec približka žarka.
  - 7: **za vse vrstice  $r$  navidezne slike izvedi**
  - 8:  $r' \leftarrow k(r - r_0) + r_0$  //Navpično povečevanje ali zmanjševanje.
  - 9: **če sta koordinati  $(r', c')$  znotraj meja slike  $\mathbf{A}_i$ , potem**
  - 10:  $v_{r,c} \leftarrow a_{r',c',i}$  //Preslikava slikovnih elementov v sliko  $\mathbf{V}$ .
  - 11: **konec če**
  - 12: **konec za**
  - 13: **konec za**
- 

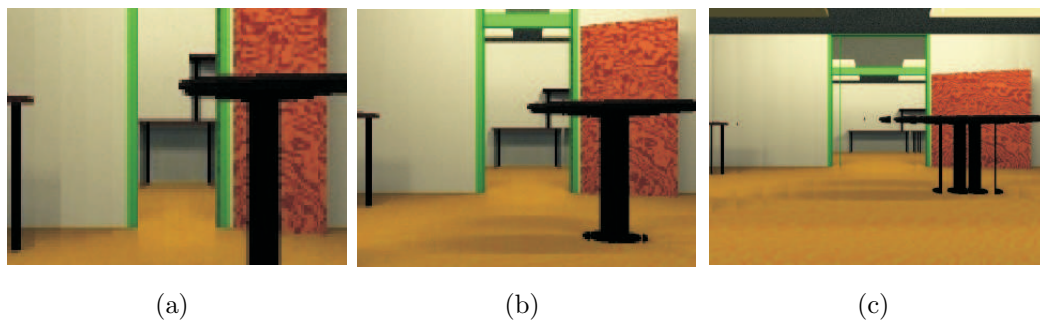
parameter nastavimo na okolju ustrezno vrednost, dobimo slike, kot so tiste na sliki 3.5. Na sliki 3.5 (a) vidimo rezultat pri  $d_r < 0$ , slika 3.5 (b) prikazuje primer pri  $d_r = 0$  (referenčna slika), slika 3.5 (c) pa pri  $d_r > 0$ . Pri vzvratnem premikanju navideznega pogleda (slika 3.5 (a)) dobimo sliko z nedefiniranimi elementi v zgornjem in spodnjem pasu slike. Ko se iz tega navideznega pogleda pomaknemo spet naprej, lahko na slikah 3.5 (b) in 3.5 (c) vidimo, da postopek realistično modelira spreminjanje zakrivanja delov prizora v vodoravni smeri.

Slika 3.6 ilustrira vpliv spreminjanja vrednosti parametra  $d_e$  na rezultat izrisovanja navidezne slike. Če parametru nastavimo previsoko vrednost, lahko nekateri elementi prizora manjkajo, drugi elementi pa so po širini pretirano skrčeni, kot vidimo na sliki 3.6 (a). Obraten učinek dosežemo s previsoko vrednostjo parametra: na sliki 3.6 (c) vidimo, da se določeni elementi ponovijo večkrat, drugi pa postanejo nesorazmerno široki. Običajno nastavimo vrednost  $d_e$  na povprečno oddaljenost do prizora ali pa oddaljenost do tistega elementa prizora, ki pokriva največji del slike. Na sliki 3.6 (b) je ta vrednost enaka razdalji do predelne stene z vrati.

Pri posamezni navidezni lokaciji je torej vrednost tega parametra  $d_e$  konstantna za celotno navidezno sliko. Ker nam ni treba izračunavati globinskega reliefa



Slika 3.5: (a) Navidezni korak nazaj, (b) referenčna slika, (c) navidezni korak naprej. Nedefinirano območje na sliki (a) je označeno s šrafuro.

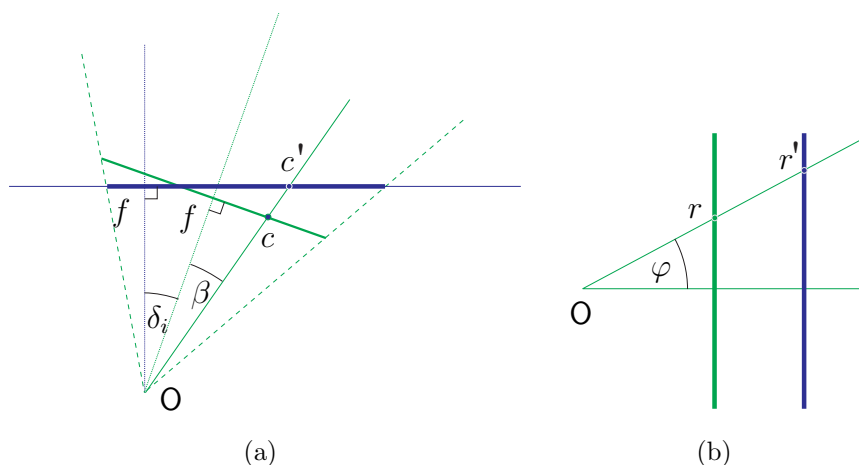


Slika 3.6: Prikaz vpliva vrednosti parametra  $d_e$ , če je vrednost (a) prevelika, (b) ustrezna, (c) premajhna.

(ang. *depth map*) [16], je delo z opisanim postopkom precej poenostavljeno. Cena te poenostavitve je večje število napak na navideznih slikah prizorov, ki imajo elemente razporejene na različnih globinah. V tem primeru nastavljanje vrednosti  $d_e$  lahko primerjamo z nastavljanjem ostrine slike na kameri s slabo globinsko ostrino. Elementi na razdalji, podani z  $d_e$ , bodo na navidezni sliki prikazani pravilno, pri ostalih pa bo prikaz vseboval napake.

### 3.3 Izrisovanje iz slik s posplošeno geometrijo

Postopek, ki smo ga opisali, predpostavlja, da so vse referenčne lokacije med seboj kolinearne, vsi pogledi pa so usmerjeni pravokotno na njihovo premico nosilko. V praksi je takšne pogoje zelo težko doseči, zato želimo, da bi metoda delovala pod ohlapnejšimi pogoji. Metoda, ki odpravi posledice tresenja, neenakomernega zajemanja slik ter tudi sicer regularizira časovno-prostorski volumen, je opisan v [45]. V nadaljevanju bomo opisali postopek, ki upošteva poznavanje lokacij in orientacij

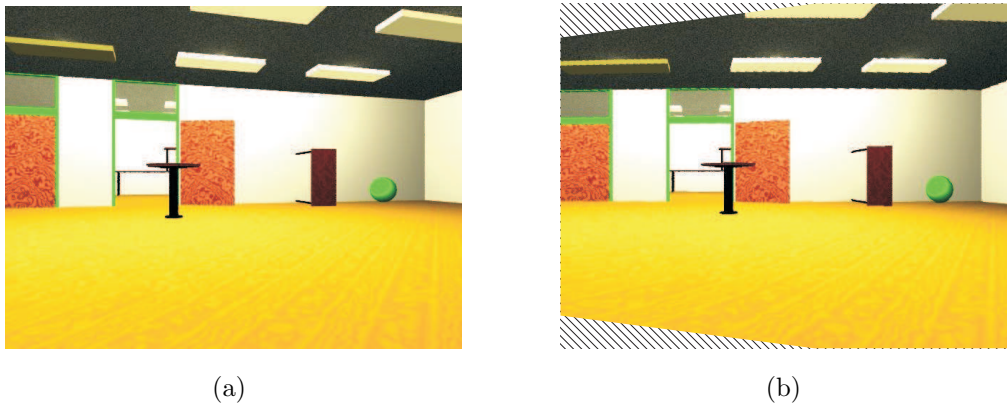


Slika 3.7: Odprava vpliva odklona od smeri pogleda za kot  $\delta_i$ : (a) tloris, (b) stranski prerez.

referenčnih pogledov, pri čemer bomo razširili geometrijske osnove, predstavljene v predhodnih razdelkih. Predpostavka, da so relativne lokacije in orientacije referenčnih slik znane, je dovolj realna, saj mobilni robot s spremljanjem vrtenja koles (odometrom) med gibanjem beleži svoje relativne koordinate.

Naj bo  $\delta_i$  odklon smeri referenčnega pogleda z zaporedno številko  $i$  od pravega kota. To situacijo prikazujejo zelene premice na sliki 3.7 (a). Odebeljena prečna daljica zelene barve označuje slikovno ravnino, ki ni vzporedna z navidezno ravnino na razdalji  $d_e$ . Zato ta kamera žarke beleži drugače kot kamera z vzporedno slikovno ravnino (odebeljena modra daljica). Smeri zabeleženih žarkov so tako premaknjene za kot  $\delta_i$ , na sliki 3.7 (b) pa vidimo tudi, da žarke, ki bi jih na vzporedni slikovni ravnini beležili slikovni elementi vrstice  $r'$ , predstavljajo slikovni elementi v vrstici  $r$ . Primer slike, ki nastane iz takšnega odklonjenega pogleda, prikazuje slika 3.8 (a).

Preden lahko v svojo predstavitev vključimo sliko drugače usmerjenega pogleda, jo moramo spremeniti tako, da bodo slikovni elementi ležali na osnovni slikovni ravnini. To naredimo tako, da vsakemu žarku z vodoravnim kotom  $\beta$  slike pripišemo nov vodoravni kot  $\beta' = \beta + \delta_i$ . Kot vidimo na sliki 3.7 (a), bo žarek, ki prispeva k slikovnemu elementu v stolpcu  $c$  izvirne slike, na popravljeni sliki prispeval k slikovnemu elementu v stolpcu  $c'$ . Na popravljeni slikovni ravnini so presečišča žarkov drugače porazdeljena kot na izvirni slikovni ravnini. Vendar s stališča vodoravnih kotov zadostuje, da si za vsak stolpec slikovnih elementov označimo vodoravni kot  $\beta'$  pripadajočega žarka, medtem ko velja  $c' = c$ . Upoštevati moramo le preslikave



Slika 3.8: (a) Slika pogleda, ki ni pravokoten na normalizacijsko ravnino pri  $d_e$ , (b) poravnana slika; šrafirano področje predstavlja nedefinirane slikovne elemente.

slikovnih elementov v navpični smeri, kot to prikazuje slika 3.7 (b).

V navpični smeri obravnavamo slikovne elemente po stolpcu  $c$  (vsi žarki pod vodoravnim  $\beta$ ). Oddaljenost središča kamere do slikovne ravnine v tem stolpcu je  $\frac{f}{\cos(\beta)}$  (razdalja do zelene navpičnice s slike 3.7 (b)). Ko slikovno ravnino za kot  $\delta_i$  zavrtimo v vzporedno lego, postane razdalja do stolpca velika  $\frac{f}{\cos(\beta+\delta_i)}$  (razdalja do modre navpičnice s slike 3.7 (b)). Ti dve razdalji omogočata izračun preslikave slikovnega elementa iz stolpca  $c$  in vrstice  $r$  zajete slike v element stolpca  $c' = c$  in vrstice  $r'$  popravljene slike:

$$r' = r \frac{\cos(\beta)}{\cos(\beta + \delta_i)}. \quad (3.6)$$

Tako spremenjene slike lahko uporabimo neposredno za sestavljanje navideznih pogledov z izrisovanjem iz slik. Po videzu in vsebini se od pravih poravnanih slik razlikujejo v tem, da so žarki porazdeljeni drugače, del slike pa lahko že v osnovi manjka zaradi odsotnosti informacije na vrhu in dnu posameznih stolpcev. Na sliki 3.8 (b) vidimo rezultat opisane preslikave, ki jo izvedemo na sliki 3.8 (a).

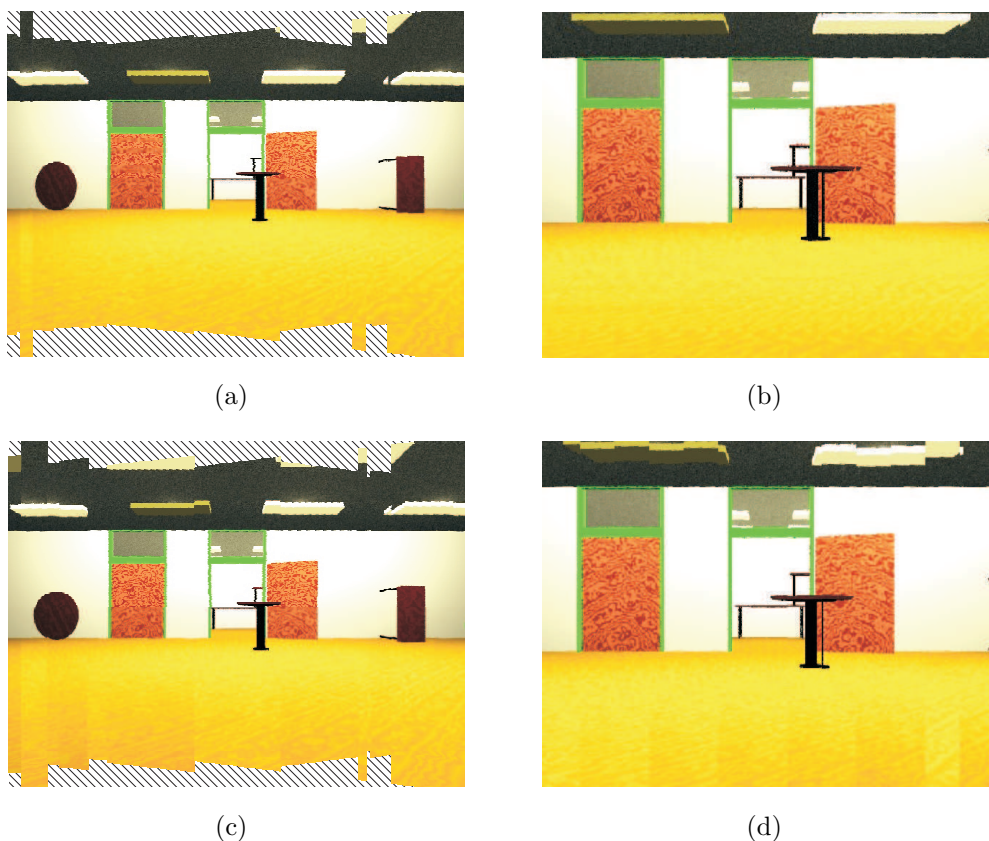
Poleg odklonov pri orientaciji referenčnih pogledov moramo upoštevati tudi, da lahko ležijo izven premice poti. Razdaljo med referenčno lokacijo  $i$  in premico označimo z  $dy_i$ , spremenjeno oddaljenost do ravnine pri  $d_e$  pa vključimo v izraza (3.4) in (3.5). Dobimo naslednja izraza:

$$\beta' = \arctan \left( \frac{(d_e - d_r) \tan(\beta) - (x_v - x_i)}{d_e - dy} \right), \quad (3.7)$$

$$\varphi' = \arctan\left(\frac{(d_e - dy) \tan(\varphi)}{d_e - d_r}\right). \quad (3.8)$$

Z opisanimi spremembami izvirnih izrazov lahko sedaj dopolnimo algoritem 2 in dobimo postopek, ki ga povzema algoritem 3. Koordinate referenčnih slik na vhodu imajo zdaj dve dimenziji in kot odklona, kar torej omogoča raziskovanje okolja v ravnini. Ker imajo referenčne lokacije lahko različne oddaljenosti od normalizacijske ravnine, moramo koeficient povečevanja ali zmanjševanja (vrstica 4) izračunavati individualno glede na izbrani referenčni pogled. Podobno v vrstici 7 vidimo, da povečevanje ali zmanjševanje slike v navpični smeri ni več enako preko celotne slike.

Slika 3.9 prikazuje primere rezultatov opisanega postopka. Če so referenčne lokacije kolinearne, vendar njihove smeri pogledov odstopajo od pogleda pravokotno na



Slika 3.9: Primeri navideznih slik, izrisanih iz prostorsko-časovnega volumna slik pogledov, ki (a)-(b) so usmerjene pod kotom, različnim od  $90^\circ$  in (c)-(d) tudi odstopajo od premice poti referenčnih pogledov. Sliki (a) in (c) sta nastali pri  $d_r < 0$ , (b) in (d) pa pri  $d_r > 0$ . Šrafirana območja predstavljajo nedefinirane slikovne elemente.

---

**Algoritem 3** Izrisovanje navidezne slike pod ohlapnejšimi pogoji.

---

**Vhod:** *Prostorsko-časovni volumen referenčnih slik:* slike  $\mathbf{A}_1, \mathbf{A}_2, \dots, \mathbf{A}_n$ , koordinate  $x_1, x_2, \dots, x_n$ , odmiki od premice  $dy_1, dy_2, \dots, dy_n$ , odkloni pogledov  $\delta_1, \delta_2, \dots, \delta_n$ ; *parametri kamere:* goriščnica  $f$ , središče slike  $(c_0, r_0)$ ; *parametri navideznega pogleda:* koordinati  $x_v$  in  $d_r$ , normalizacijska razdalja  $d_e$ .

**Izhod:** navidezna slika  $\mathbf{V}$ .

- 1: **za vse** stolpce  $c$  navidezne slike **izvedi**
  - 2:  $\beta \leftarrow \arctan\left(\frac{c-c_0}{f}\right)$  // Pretvori v kote.
  - 3:  $i \leftarrow \arg \min_j |x_j - (x_v - d_r \tan(\beta))|$  // Poišči referenčno sliko.
  - 4:  $k \leftarrow \frac{d_e - d_r}{d_e - dy_i}$  // koeficient povečave/pomanjšave
  - 5:  $\beta' \leftarrow \arctan\left(k \tan(\beta) - \frac{(x_i - x_v)}{d_e - dy_i}\right)$  // Določi približek žarka.
  - 6:  $c' \leftarrow f \tan(\beta' - \delta_i) + c_0$  // Izberi stolpec približka žarka.
  - 7:  $k_v \leftarrow k \frac{\cos(\beta')}{\cos(\beta' - \delta_i)}$  // koeficient navpične povečave/pomanjšave
  - 8: **za vse** vrstice  $r$  navidezne slike **izvedi**
  - 9:  $r' \leftarrow k_v(r - r_0) + r_0$  // Navpično povečevanje ali zmanjševanje.
  - 10: če sta koordinati  $(r', c')$  znotraj meja slike  $\mathbf{A}_i$ , **potem**
  - 11:  $v_{r,c} \leftarrow a_{r',c',i}$  // Preslikava slikovnih elementov v sliko  $\mathbf{V}$ .
  - 12: **konec če**
  - 13: **konec za**
  - 14: **konec za**
- 

premico poti, dobimo navidezne slike, kot sta tisti na slikah 3.9 (a) in (b). V najbolj splošnem primeru pa dobimo rezultate, kot jih prikazujeta primera s slike 3.9 (c) in (d). Na spremembe v orientaciji referenčnih pogledov nakazuje oblika meje med območji definiranih in nedefiniranih slikovnih elementov (sliki 3.9 (a) in (c)). Opazimo tudi manjše nepravilnosti elementov, ki od referenčnih pogledov niso oddaljeni za vrednost  $d_e$ .

### 3.4 Izrisovanje navideznih slik iz zgoščenih volumnov

Slikovni volumni v pomnilniku zasedajo veliko prostora. Hkrati je izkoristek shranjenih podatkov relativno majhen, saj do posameznega slikovnega elementa dostopamo redko. Zato potrebujemo postopek, ki bi zmanjšal porabo pomnilnika, povečal njegov izkoristek in hkrati še vedno omogočal sestavljanje navideznih pogledov.

Za to nalogo smo izbrali metodo glavnih komponent, ki smo jo opisali že v poglavju 2. Tudi tukaj imamo za izhodišče zaporedje  $n$  slik  $\mathbf{A}_i \in \mathbb{R}^R \times \mathbb{R}^C$ ,  $i = 1 \dots n$ , širine  $C$  in višine  $R$  elementov, ki sestavljajo slikovni volumen  $\mathcal{A} = \{\mathbf{A}_1, \mathbf{A}_2, \dots, \mathbf{A}_n\}$ . Elemente slikovne matrike  $\mathbf{A}_i$  zložimo v slikovni vektor  $\mathbf{a}_i$ . Matriko navidezne slike, ki jo želimo iz volumna dobiti, označimo z  $\mathbf{V}$ .

Z metodo glavnih komponent za prostorsko-časovni volumen referenčnih slik  $\mathcal{A}$  dobimo podprostor, razpet med lastnimi vektorji  $\mathbf{U} = [\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k]$ . Podprostor ima svoje središče v srednji vrednosti  $\bar{\mathbf{a}}$  slikovnih vektorjev  $\mathbf{a}_i$ ,  $i = 1 \dots n$ . V ta podprostor projiciramo slikovne vektorje  $\mathbf{a}_i$  in dobimo vektorje njihovih globalnih značilnic  $\mathbf{w}_i = [w_{1i}, w_{2i}, \dots, w_{ki}]^\top$ ,  $i = 1, \dots, n$ . Našo zgoščeno predstavitev sedaj sestavlja matrika baznih vektorjev  $\mathbf{U}$ , vektor srednje vrednosti  $\bar{\mathbf{a}}$  in vektorji projekcij  $\mathbf{w}_i$ ,  $i = 1, \dots, n$ . Izvirnih slik  $\mathbf{A}_i$ ,  $i = 1, \dots, n$ , ne potrebujemo več, saj lahko njihove približke rekonstruiramo z izrazom

$$\mathbf{a}_i \approx \sum_{j=1}^k w_{ji} \mathbf{u}_j + \bar{\mathbf{a}} = \mathbf{U} \mathbf{w}_i + \bar{\mathbf{a}}. \quad (3.9)$$

Ko imamo referenčne slike predstavljene v podprostoru metode glavnih komponent, lahko navidezne poglede sestavimo podobno, kot smo to opisali v razdelku 3.2. Za vsak stolpec navidezne slike  $\mathbf{V}$  najprej izberemo zaporedno številko  $i$  referenčnega pogleda z izrazom (3.3). Nato izvedemo rekonstrukcijo izbrane referenčne slike  $\mathbf{A}_i$  s pomočjo izraza (3.9). Od tu naprej ravnamo enako, kot če bi imeli izvirne referenčne slike.

Ugotovimo lahko, da z rekonstrukcijo celotnih referenčnih slik opravimo več dela, kot je potrebno, saj običajno iz vsake referenčne slike potrebujemo le majhno število stolpcev. Količino potrebnih izračunov lahko občutno zmanjšamo, če rekonstruiramo le tiste slikovne elemente referenčnih slik, ki jih dejansko uporabimo pri sestavljanju navidezne slike. Izraz (3.9) deluje po posameznih komponentah vektorjev, zato lahko izrazimo samo skalarje, ki nastopajo v teh vektorjih:

$$a_{(r,c)i} \approx \sum_{j=1}^k w_{ji} u_{(r,c)j} + \bar{a}_{(r,c)}. \quad (3.10)$$

Zdaj lahko popravimo algoritem 3 tako, da bo namesto prostorsko-časovnega volumna referenčnih slik na vhod dobil zgoščeno predstavitev volumna. Vrstico 11 algoritma nadomestimo z rekonstrukcijo slikovnega elementa iz izraza (3.10), pri čemer naredimo ustrezno preslikavo indeksov. Algoritem 4 prikazuje opisane spremembe.

**Algoritem 4** Izrisovanje navidezne slike iz zgoščene predstavitve.

**Vhod:** *Zgoščen prostorsko-časovni volumen referenčnih slik:* srednja vrednost  $\bar{\mathbf{a}}$ , lastni vektorji  $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k$ , projekcije referenčnih slik  $\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_n$ , koordinate  $x_1, x_2, \dots, x_n$ , odmiki od premice  $dy_1, dy_2, \dots, dy_n$ , odkloni pogledov  $\delta_1, \delta_2, \dots, \delta_n$ ; *parametri kamere:* goriščnica  $f$ , središče slike  $(c_0, r_0)$ ; *parametri navideznega pogleda:* koordinati  $x_v$  in  $d_r$ , normalizacijska razdalja  $d_e$ .

**Izhod:** navidezna slika  $\mathbf{V}$ .

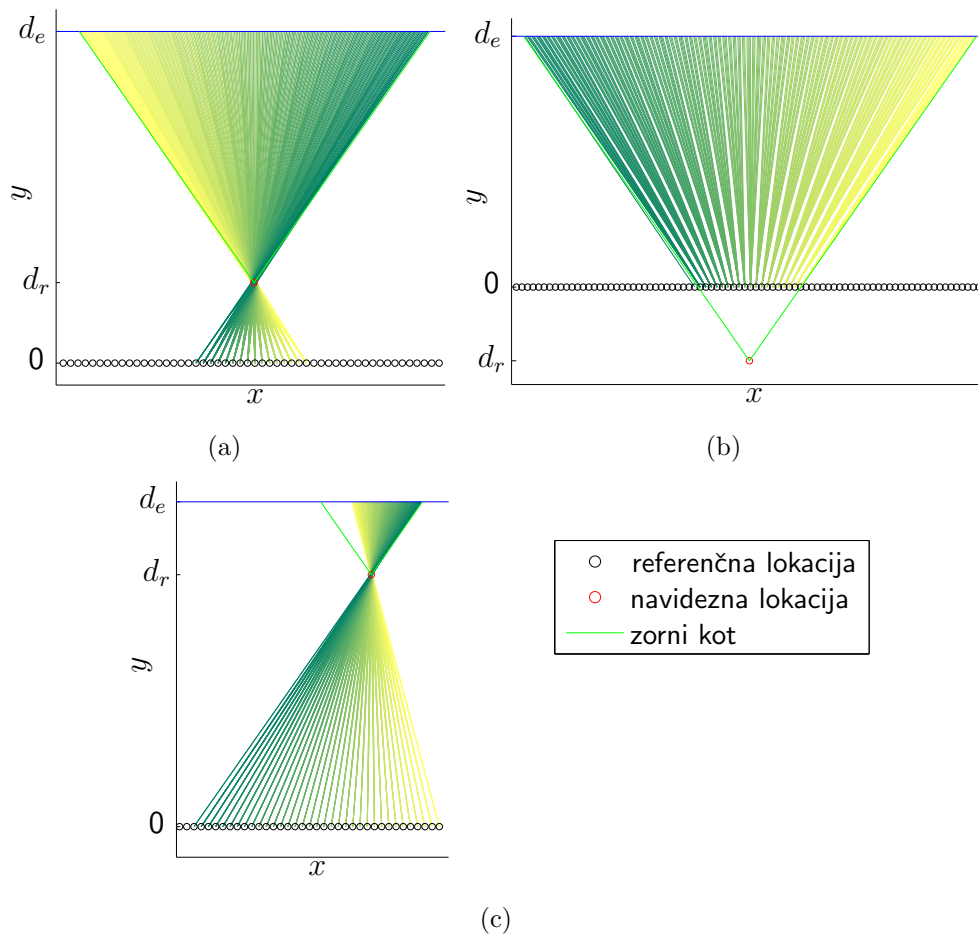
- 1: **za vse** stolpce  $c$  navidezne slike **izvedi**
- 2:   izračun  $i$  in  $c'$  // *algoritem 3, vrstice 3–7*
- 3:   **za vse** vrstice  $r$  navidezne slike **izvedi**
- 4:    izračun  $r'$  // *algoritem 3, vrstica 9*
- 5:    **če** sta koordinati  $(r', c')$  znotraj meja slik, **potem**
- 6:       $v_{r,c} \leftarrow \sum_{j=1}^k w_{ji} u_{(r',c')j} + \bar{a}_{(r',c')}$  // *Rekonstrukcija in preslikava slik. elementov.*
- 7:    **konec če**
- 8:    **konec za**
- 9: **konec za**

### 3.5 Lastnosti predstavitve okolice s prostorsko-časovnimi volumni

V predhodnih razdelkih smo predstavili teoretične in geometrijske osnove predstavitve nekega prostora ali okolja s prostorsko-časovnim volumnom slik. V nadaljevanju bomo opisali praktični pomen in lastnosti takšne predstavitve.

Slikovni elementi v prostorsko-časovnem volumnu hranijo zapis posameznih žarkov, ki padajo na slikovno ravnino. Določen slikovni element (stolpec) torej lahko uporabimo pri sestavljanju samo tistih navideznih pogledov, katerih izhodišče je kolinearno s tem žarkom in katerih zorni kot vključuje smer tega žarka. Zaradi diskretnosti nabora referenčnih pogledov in posledično načina njihovega izbora z izrazom (3.3) je ta omejitev nekoliko splošena na množico točk med simetralama med žarkom in vsakim od sosednjih žarkov. Temu navkljub velja, da mora biti navidezni pogled usmerjen v isto smer kot referenčni pogledi, če želimo dobiti navidezno sliko z največ definiranimi slikovnimi elementi. Navidezne poglede, ki so usmerjeni drugače, pa lahko izrišemo le delno, torej tisti del zornega kota, ki je v preseku z zornim kotom referenčnih slik.

Slika 3.10 prikazuje, katere referenčne slike in njihove žarke uporabimo za izris navideznega pogleda. Referenčni pogledi imajo na prikazanem primeru zorni kot

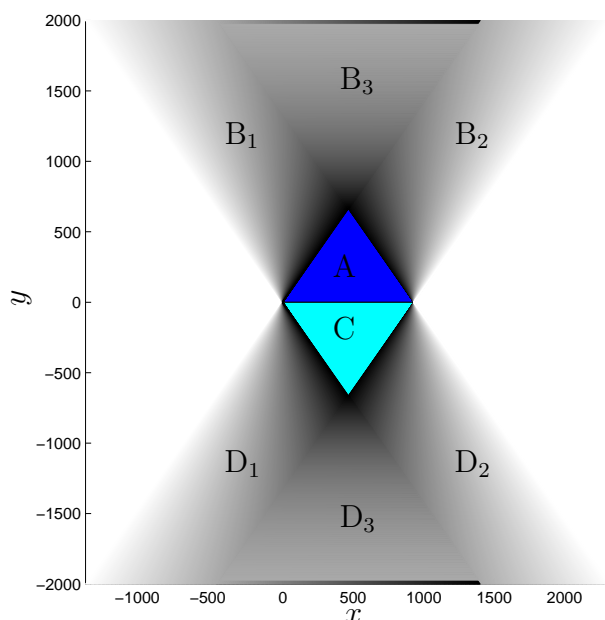


Slika 3.10: Referenčni žarki, ki jih uporabimo pri izrisovanju navideznega pogleda, pri čemer je navidezna lokacija (a) bliže prizoru, (b) bolj oddaljen od prizora in (c) ob robu znanega območja.

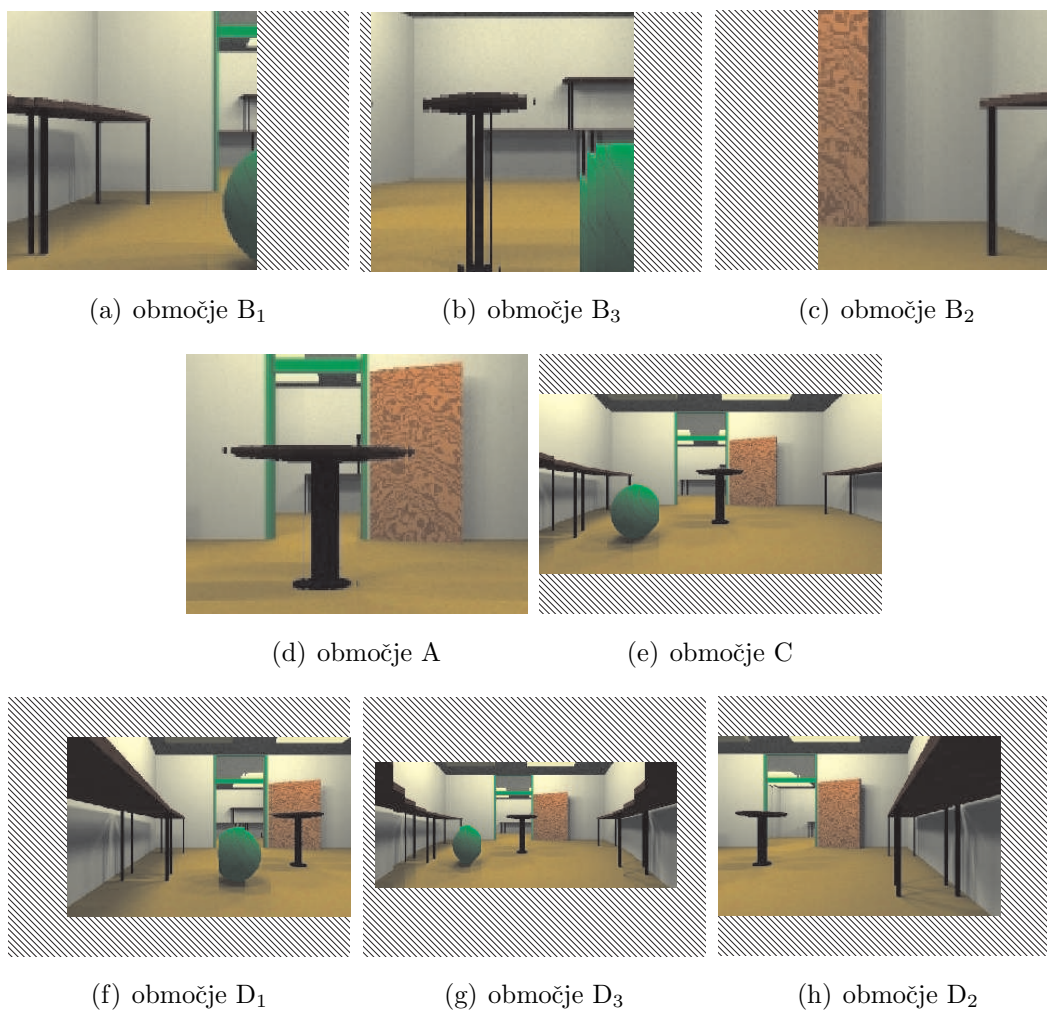
$70^\circ$ , zato tudi navidezni pogled izrišemo s tem zornim kotom. Navidezna lokacija je na sliki 3.10 (a) bliže prizoru, ki ga predstavljajo referenčni pogledi (približevanje), na sliki 3.10 (b) pa je ta lokacija bolj oddaljena od prizora kot so referenčni pogledi (oddaljevanje). Shema grafično prikazuje, kako algoritem 3 v vrstici 3 za nek ciljni stolpec navidezne slike izbere referenčno sliko. Vrstni red izbranih referenčnih slik bo pri  $d_r > 0$  obratno sorazmeren, sicer pa premo sorazmeren z vrstnim redom izrisovanja stolpcev navidezne slike. Če pri več zaporednih stolpcih navidezne slike izberemo isto referenčno sliko, bo vrstni red stolpcev izbrane referenčne slike v vrstici 6 algoritma 3 vedno premo sorazmeren z vrstnim redom izrisovanja stolpcev navidezne slike.

Ko pomaknemo navidezno lokacijo proti eni od skrajnih referenčnih lokacij, dobimo situacijo s slike 3.10 (c). Vidimo, da na desni strani nimamo več referenčnih slik, ki bi hranile slikovne elemente za zapolnitev levega dela navidezne slike. Uporabimo lahko sicer preostale elemente iz skrajne referenčne slike, vendar je to relativno slabši približek.

Zorni kot referenčnih slik določa območja v ravnini, kjer lahko dobimo celotni izris navidezne slike, pa tudi to, kakšne deleže slike lahko dobimo izven tega območja. Shema na sliki 3.11 prikazuje vsa območja za vzorčni primer predstavitve s prostorsko-časovnim volumnom. Žarki prvega in zadnjega stolpca obeh skrajnih referenčnih slik razdelita ravnino na 8 območij. Slika 3.12 prikazuje po en primer navideznega pogleda iz vsakega od teh območij. Najmanjše trikotno območje, razpeto med te meje in referenčno potjo, je območje A (slika 3.12 (d)) in njemu zrcalno območje C (slika 3.12 (e)). Ti območji predstavljata množico lokacij, kjer lahko izrišemo navidezne slike za celoten zorni kot. Ko se pomaknemo v širša območja, v njihovih lokacijah nimamo več vseh potrebnih žarkov, zato je navidezni pogled tam le delno zgrajen. V območju  $B_1$  manjkajo žarki za desni del slike (slika 3.12 (a)), v območju  $B_2$  pa ne moremo izrisati levega dela navidezne slike (slika 3.12 (c), slika 3.10 (c)). Podobno velja za območji  $D_2$  in  $D_1$  (sliki 3.12 (h) in (f)). Točke iz območja  $B_3$  (slika 3.12 (b)) in njemu zrcalnega območja  $D_3$  (slika 3.12 (g)) pa so že



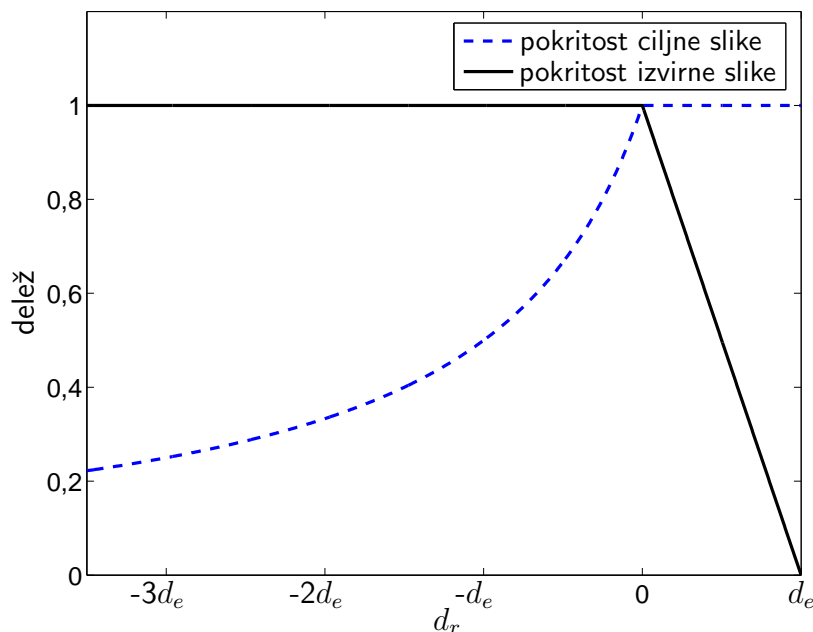
Slika 3.11: Območja dosega pri predstavitvi okolice s prostorsko-časovnimi volumni.



Slika 3.12: Primeri navideznih slik glede na območja iz slike 3.11. Šrafirana območja predstavljajo nedefinirane slikovne elemente.

tako oddaljene od referenčnih lokacij, da pri navidezni pogledih manjka tako levi kot desni del slike. Delež pokrite navidezne slike je sicer enak pri vseh navidezni pogledih, ki imajo isto vrednost  $y$ , vendar se njegov začetni stolpec spreminja s koordinato  $x$  navidezne lokacije. Slika 3.11 z intenziteto sivine ponazarja deleže pokritosti navidezne slike v posameznih lokacijah: temnejša območja pomenijo boljšo pokritost.

Pri opisu predstavljenih območij smo izhajali iz izrazov (3.3) in (3.4), ki izbirata referenčne slike in slikovne stolpce na teh slikah. Pri gibanju navidezne lokacije v smeri  $y$  pa moramo upoštevati še spremembe, ki nastanejo na nivoju vrstic zaradi izraza (3.5), torej spremembe slike v navpični smeri. Slika 3.13 povzema te



Slika 3.13: Ponazoritev deležev navidezne slike, ki jo pokrivajo slikovni elementi, in deležev referenčne slike, ki jo uporabimo za izrisovanje.

spremembe.

Ko navidezno lokacijo pomikamo v smeri  $y$  stran od prizora, torej pri negativnih vrednostih  $d_r$ , se elementi na sliki zmanjšujejo. Pri tem izrabljamo celotno višino izvernih referenčnih slik, vendar se povečuje zgornji in spodnji pas, kjer so slikovni elementi navidezne slike nedefinirani, sama slika pa s tem izgublja podrobnosti. Te spremembe ponazarja krivulja **pokritost ciljne slike** na sliki 3.13. Ko pa se gibljemo na območju  $d_r > 0$ , pa nimamo več nedefiniranih elementov na sliki (pokritost ciljne slike je 100%), saj povečujemo podsliko na velikost cele slike. Informacij torej ne izgubljam, vendar hkrati ne pridobivamo na podrobnostih, sicer prisotnih na resničnih slikah v navidezni lokaciji. Krivulja **pokritost izvirne slike** grafa na sliki 3.13 tako ponazarja, kakšen delež referenčne slike izberemo pri povečavi na celotno višino navidezne slike.

Glede na ugotovitve te analize lahko sklepamo, kako velik doseg ima konkretna predstavitev s prostorsko-časovnim volumnom. Z daljšo potjo, ki jo opravimo med zajemanjem referenčnih slik, dobimo večji obseg znanega okolja. Če zmanjšamo zorni kot navideznega pogleda, lahko podaljšamo območji A in C s slike 3.11, vendar oženje pogleda lahko slabo vpliva na uspešnost primerjanja slik in razpoznavne loka-

cije. Dobro razpoznavo lahko pričakujemo v območju pred referenčnimi pogledi, saj z interpolacijo pokrijemo vse slikovne elemente slike, ki jo primerjamo. Zato je boljše, če so referenčni elementi bolj oddaljeni od prizora. Na območju za referenčnimi lokacijami lahko primerjamo le osrednje vrstice neznane slike, saj v zgornjem in spodnjem delu navidezne slike informacije ni. Vseeno lahko do oddaljenosti v obsegu polovice  $d_e$  pričakujemo zadostni delež prisotnih informacij na navidezni sliki.

Seveda pa v praksi na področje pokrivanja vplivajo elementi v samem prostoru. Predmeti, ki so bližje referenčnim slikam od razdalje  $d_e$ , zakrivajo dele prizora. Njihov vpliv postane moteč, če se nahajajo med referenčnimi lokacijami in navidezno lokacijo, saj se njihovi deli zaradi zamenjave vrstnega referenčnih slik (slika 3.10 (b)) pojavljajo na nepričakovanih delih. Primer takih elementov vidimo na sliki 3.12 (b), kjer se deli žoge z leve strani prizora pojavljajo na desnem delu slike. Vpliv teh zakrivanj bi lahko zmanjšali tako, da z robustno metodo primerjanja izločimo moteče elemente, ali pa uporabimo za območje za to oviro svojo predstavitev, ki ovire ne vključuje.

## 3.6 Povzetek poglavja

Opisana predstavitev okolja ima obliko prostorsko-časovnega volumna referenčnih slik, ki prikazujejo bočni pogled na okolje med vožnjo po premici. Postopek izrisovanja z navzkrižnimi režami iz prereza volumna glede na podane parametre navideznega pogleda sestavi približek slike iz tega pogleda. Razširjeni postopek izrisovanja z navzkrižnimi režami na vhod lahko prejme tudi predstavitve, pri katerih referenčni pogledi niso vsi na premici in niso enako usmerjeni. S pomočjo metode glavnih komponent dobimo zgoščeno predstavitev, izrisovanje pa nato poteka med sprotno rekonstrukcijo slikovnih elementov, ki jih postopek potrebuje. Konkretna predstavitev okolja razdeli množico navideznih pogledov na več področij glede na to, kakšen delež slikovnih stolpcev bodo pripadajoče navidezne slike imeli definirane. Svoj vpliv na kakovost navidezne slike ima tudi oddaljenost navideznega pogleda od referenčnih pogledov.



# Poglavje 4

## Lokalizacija mobilnega robota

### 4.1 Uvod

Lokalizacija mobilnega robota je postopek, ki iz slike trenutnega robotovega pogleda s pomočjo predstavitve prostora ugotovi, kje se robot nahaja. Predstavitev prostora nastane v učni fazi, ki vključuje raziskovanje okolja, zajemanje slik ter organizacijo zajetih podatkov. Njen namen je kasnejše primerjanje z novimi slikami in določanje odnosov neznanih lokacij do modela.

V prejšnjem poglavju smo opisali predstavitev okolja, ki jo zgradimo tako, da robota peljemo po ravni poti in zajemamo slike na robotovem levem ali desnem boku. Ob vsaki sliki robot shrani še relativne koordinate in orientacijo robota, to so podatki o odometriji. Slike podajamo postopni gradnji podprostora metode glavnih komponent, ki nam vrne zgoščeno obliko prostorsko-časovnega volumna referenčnih slik. Naš model okolja je torej predstavnik modelov, ki okolje predstavijo z množico pogledov, namenjenih primerjanju videza.

Naš postopek modela ne uporabljamo neposredno za primerjanje slik iz robotovega trenutnega pogleda, saj vsebuje premalo učnih pogledov. Model je namenjen izrisovanju slik v navideznih pogledih. V nadaljevanju bomo opisali, kako s tem postopkom dobimo oceno robotove trenutne lokacije.

### 4.2 Primerjanje slik

Pri nalogi lokalizacije običajno zajamemo sliko  $\mathbf{Q} \in \mathbb{R}^R \times \mathbb{R}^C$  pogleda prizora v lokaciji, kjer se trenutno nahajamo. Sliko  $\mathbf{Q}$  primerjamo s slikami, katerih lokacije poznamo, lokacijo najbolj podobne izmed njih pa določimo za oceno lokacije na-

stanka slike  $\mathbf{Q}$ . Potrebujemo torej mero podobnosti, ki jo želimo maksimizirati, ali, dualno, mero različnosti, katere najnižjo vrednost želimo najti.

Običajna mera za podobnost (različnost) dveh slik je razdalja med vektorjema slik, ki ju primerjamo. Kvadrat te razdalje je *vsota kvadratov razlik* in jo med slikovnima vektorjema  $\mathbf{v} \in \mathbb{R}^{RC}$  in  $\mathbf{q} \in \mathbb{R}^{RC}$  izrazimo kot

$$d_{\text{ssd}}(\mathbf{q}, \mathbf{v}) = \sum_{j=1}^{RC} (q_j - v_j)^2. \quad (4.1)$$

V praksi se pogosto zgodi, da sliki, ki ju primerjamo, nista popolni in da za neko podmnožico slikovnih elementov nimamo podatkov. Manjkajoči deli slik so posledica odsotnosti vseh potrebnih žarkov pri izrisu navideznih slik. Poleg tega lahko z dodatnim postopkom na sliki zaznamo zakrite ali nezanesljive dele, ki jih izločimo iz slike. Izračun razdalje v obliki izraza (4.1) tako ni mogoč, saj pri elementih, kjer eden od faktorjev manjka, nimamo vrednosti za primerjavo, če pa bi manjkajoče vrednosti zapolnili z neko drugo vrednostjo, pa nam lahko ti manjkajoči deli napačno vplivajo na oceno.

Za slikovna vektorja  $\mathbf{v}$  in  $\mathbf{q}$  uvedemo binarno masko  $\mathbf{m} \in \{0, 1\}^{RC}$  tako, da velja

$$m_j = \begin{cases} 1 & \text{podatka } v_j \text{ in } q_j \text{ sta oba definirana in zanesljiva,} \\ 0 & \text{sicer.} \end{cases} \quad (4.2)$$

Vrednosti elementom maske dodelimo med izrisovanjem navidezne slike. V ta namen smo dopolnili postopek tako, kot ga opisuje algoritem 5. Maska  $\mathbf{m}$ , ki jo dobimo iz matrike  $\mathbf{M} \in \{0, 1\}^R \times \{0, 1\}^C$  kot rezultat algoritma, torej z vrednostjo 1 označuje tiste pare istoležnih elementov, ki jih lahko uporabimo pri primerjavi, z 0 pa takšne, pri katerih je nedefiniran vsaj eden od njih. Razdaljo iz izraza (4.1) nato spremenimo tako, da upošteva le pare definiranih elementov, hkrati pa z normalizacijo kompenzira za manjkajoče elemente:

$$d_{\text{nssd}}(\mathbf{q}, \mathbf{v}; \mathbf{m}) = \frac{RC}{\sum_{j=1}^{RC} m_j} \sum_{j=1}^{RC} m_j (q_j - v_j)^2. \quad (4.3)$$

To mero smo izbrali zaradi njene preprostosti in hitrosti izračunavanja. Dobro se je izkazala tudi pri preskusih. Seveda bi jo bilo moč nadomestiti z uteženo navzkrižno korelacijo, normirano navzkrižno korelacijo ali katero drugo mero, pri čemer ne bi vplivali na opisan postopek kot celoto.

---

**Algoritem 5** Izrisovanje navidezne slike, ki vrača masko z definiranimi elementi.

---

**Vhod:** *Zgoščen prostorsko-časovni volumen referenčnih slik:* srednja vrednost  $\bar{a}$ , lastni vektorji  $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k$ , projekcije referenčnih slik  $\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_n$ , koordinate  $x_1, x_2, \dots, x_n$ , odmiki od premice  $dy_1, dy_2, \dots, dy_n$ , odkloni pogledov  $\delta_1, \delta_2, \dots, \delta_n$ ; *parametri kamere:* goriščnica  $f$ , središče slike  $(c_0, r_0)$ ; *parametri navideznega pogleda:* koordinati  $x_v$  in  $d_r$ , normalizacijska razdalja  $d_e$ .

**Izhod:** navidezna slika  $\mathbf{V}$ , binarna maska  $\mathbf{M}$ .

- 1: **za vse** stolpce  $c$  navidezne slike **izvedi**
  - 2:     izračun  $i$  in  $c'$  // *algoritem 3, vrstice 3–7*
  - 3:     **za vse** vrstice  $r$  navidezne slike **izvedi**
  - 4:         izračun  $r'$  // *algoritem 3, vrstica 9*
  - 5:         **če** sta koordinati  $(r', c')$  znotraj meja slik, **potem**
  - 6:              $v_{r,c} \leftarrow \sum_{j=1}^k w_{ji} u_{(r',c')j} + \bar{a}_{(r',c')}$  // *Rekonstr. in preslikava slikovnih elementov.*
  - 7:              $m_{r,c} \leftarrow 1$  // *Element vključen v primerjavo.*
  - 8:         **sicer**
  - 9:              $m_{r,c} \leftarrow 0$  // *Element izključen iz primerjave.*
  - 10:         **konec če**
  - 11:     **konec za**
  - 12: **konec za**
- 

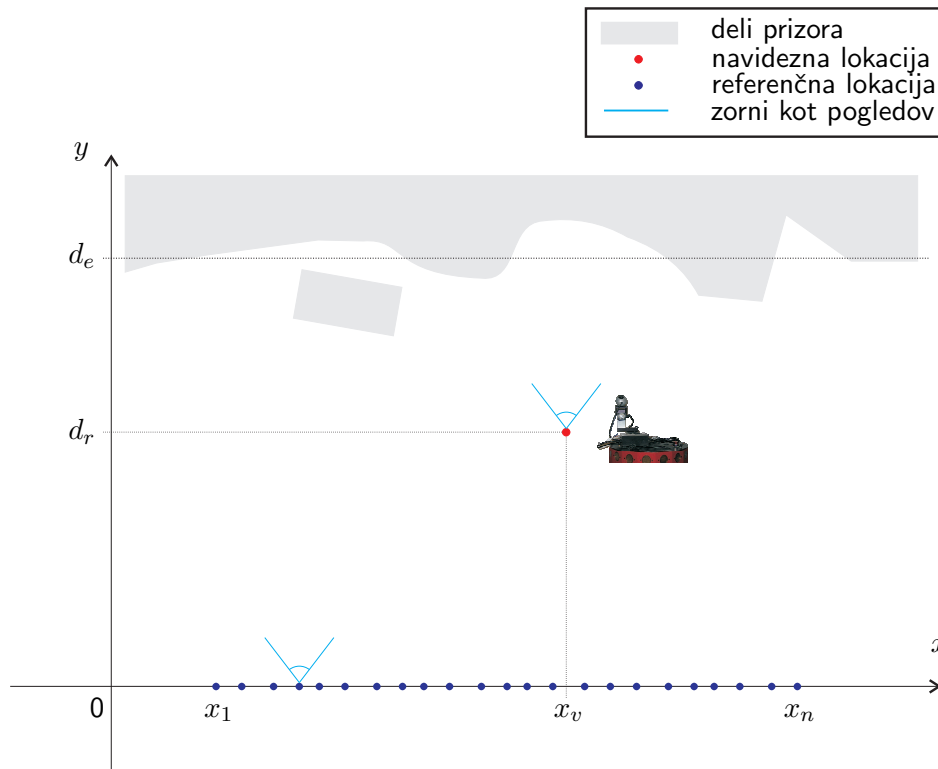
## 4.3 Lokalizacija

Zdaj, ko imamo mero za primerjanje navidezne slike s katerokoli drugo sliko, lahko opišemo postopek lokalizacije. Naša predstavitev okolice ima obliko prostorsko-časovnega volumna referenčnih slik, ki sama neposredno ne nastopa pri postopku primerjave znanih pogledov s podanim neznanim pogledom. Predstavitev uporabimo za izrisovanje navideznih slik v lokacijah kandidatkah, vsako od teh slik pa nato z mero (4.3) primerjamo z resnično sliko iz neznanega pogleda. Lokacija navideznega pogleda, ki doseže najnižjo vrednost mere, postane ocena za neznano lokacijo.

Če želimo iz neznanne slike z opisano metodo ugotoviti lokacijo robota, se mora smer pripadajočega pogleda delno ali v celoti pokrivati s smerjo referenčnih pogledov predstavitve. V nadaljevanju bomo predpostavili, da lahko s pomočjo kompasa ali drugega postopka dovolj natančno ugotovimo trenutno orientacijo robota, pogled (kamero) pa usmerimo tako, kot so usmerjeni referenčni pogledi. Z lokalizacijo želimo ugotoviti koordinati  $x$  in  $y$  trenutne lokacije. Brez izgube splošnosti

lahko predpostavimo, da je  $x$  oddaljenost od prve referenčne točke do projekcije neznane lokacije na premico poti,  $y$  pa razdalja med neznano lokacijo in premico poti. Slika 4.1 predstavlja tloris vzorčnega okolja, v katerem se robot nahaja, koordinatni sistem pa je poravnan z lokacijami referenčnih pogledov. Koordinati  $x$  in  $y$  lahko neposredno prevedemo v parametra  $x_v$  in  $d_r$  s slike 4.1. Poleg teh parametrov za izrisovanje navideznega pogleda potrebujemo še parameter globine prizora  $d_e$ . Vsi trije skupaj tvorijo trirazsežni preiskovalni prostor, katerega točke ocenjujemo z mero razdalje (4.3) med navidezno sliko, ki nastane v tej točki, in sliko neznanega pogleda. Algoritem 6 povzema postopek preiskovanja tega prostora.

Opisano iskanje rešitve pri problemu lokalizacije je torej enakovredno reševanju optimizacijskega problema v treh razsežnostih za mero ocene (4.3) pri podani sliki neznanega pogleda in prostorsko-časovnem volumnu referenčnih slik. Preiskovalni prostor je lahko razmeroma obsežen, zato je proces dejanske lokalizacije odvisen od izbrane strategije pregledovanja možnih rešitev. Najpreprostejšo strategijo smo na-



Slika 4.1: Skica robotovega okolja pri lokalizaciji, koordinatni sistem predstavitev in parametri navideznega pogleda.

---

**Algoritem 6** Lokalizacija neznanega pogleda iz slike.

---

**Vhod:** Prostorsko-časovna predstavitev okolja; slikovni vektor  $\mathbf{q}$  pogleda neznanе lokacije.

**Izhod:** ocena koordinat  $x_v$ ,  $d_r$  in parametra  $d_e$ .

- 1: izberi  $\mathbf{H} \in \mathbb{R}^{h \times 3}$  // Določi preiskovalni prostor s  $h$  hipotezami.
  - 2:  $s \leftarrow \infty$  // začetna vrednost ocene hipoteze
  - 3:  $[x_v, d_e, d_r] \leftarrow [?, ?, ?]$  // Najboljša hipoteza je še nedefinirana.
  - 4: **za vse**  $[x'_v, d'_e, d'_r] \in \mathbf{H}$  **izvedi**
  - 5: Izriši navidezno sliko  $\mathbf{V}$  in masko  $\mathbf{M}$  za parametre  $[x'_v, d'_e, d'_r]$  // Algoritem 5.
  - 6:  $s' \leftarrow d_{\text{nssd}}(\mathbf{q}, \mathbf{v}; \mathbf{m})$  // Ocenimo trenutno hipotezo z mero (4.3).
  - 7: **če** velja  $s' < s$ , **potem**
  - 8:  $[x_v, d_e, d_r] \leftarrow [x'_v, d'_e, d'_r]$  // Shranimo trenutno najboljšo hipotezo.
  - 9:  $s \leftarrow s'$  // Shranimo oceno trenutno najboljše hipoteze.
  - 10: **konec če**
  - 11: **konec za**
- 

vedli z algoritmom 6 v obliki prebora točk iz neke množice hipotez. Ta ima lahko strukturo trirazsežne diskretne mreže, naključnega vzorčenja prostora ali kakšnega drugega vzorčenja. Če je množica hipotez dovolj gosta, bomo zagotovo našli najboljšo rešitev, vendar bomo za to porabili veliko časa. Boljše strategije sproti prilagajajo množico hipotez v odvisnosti od predhodnega uspeha. Primer takšnega preiskovanja je simulirano ohlajanje [35], ki preiskovalni prostor pregleduje naključno in gravitira proti vedno boljši rešitvi. Seveda lahko uporabimo katerokoli drugo splošno znano strategijo za numerično optimizacijo.

Tako obsežno pregledovanje preiskovalnega prostora je smiselno samo na začetku, ko lokacija robota še ni poznana, torej pri lokalizaciji po t.i. ugrabitvi robota. Med nadaljnjo navigacijo zadošča pregledovanje le majhnega dela preiskovalnega prostora v okolici lokacije, ki jo predvidimo po akciji (premiki) robota. V tej fazi bi lahko robotu sledili s Kalmanovim filtrom [22] ali s pomočjo filtriranja delcev (ang. *particle filtering*) [10].

## 4.4 Povzetek poglavja

Lokalizacija na podlagi videza temelji na iskanju ujemanj parov slik. Zato uvedemo numerično mero za primerjanje dveh slik, pri katerih so lahko posamezni slikovni elementi nedefinirani ali nezanesljivi. Lokalizacija na vhod prejme sliko iz pogleda,

ki mu mora določiti koordinate. S tvorjenjem hipotez pogledov in primerjanjem pripadajočih navideznih slik s podano sliko poiščemo hipotezo, ki je podani sliki najbolj podobna. Koordinate te hipoteze so rezultat lokalizacije.

# Poglavje 5

## Preskusi

### 5.1 Uvod

Lokalizacijo s pomočjo navideznih slik smo predstavili s teoretičnega in tehničnega vidika. V tem poglavju bomo predstavili še praktično delovanje predstavljene metode, uspešnost pri reševanju te naloge in rezultate.

Struktura poglavja bo odsevala strukturo disertacije. Najprej bomo predstavili, kakšne nabore podatkov smo pri preskusih uporabili in zakaj. Sledi motivacija v obliki preskusov lokalizacije mobilnega robota z običajnim postopkom, ki deluje na podlagi videza. Metodo, opisano v disertaciji, bomo nato s preskusi utemeljevali postopno. Preverili bomo primernost predstavljene metode za izrisovanje navideznih slik, v zadnjem delu poglavja pa predstavili še reševanje naloge lokalizacije. Pri tem bomo najprej opisali rezultate preskusov pod bolj nadzorovanimi pogoji, nato pa z vedno manj nadzora nad postopkom zajemanja slik. Ugotovili bomo, kako predmeti, ki delno zakrivajo prizore, vplivajo na lokalizacijo, nazadnje pa predstavili še lokalizacijo z mobilnim robotom in digitalno kamero.

#### 5.1.1 Testni nabori slik

Med razvojem neke metode si pripravimo podatke za ocenjevanje delovanja in uspešnosti metode. Z vnaprej pripravljenimi podatki na eni strani zagotovimo ponovljivost določenih situacij znotraj posameznega nabora, po drugi strani pa se z raznolikostjo naborov ugotovimo pripravljenost postopkov na različne situacije.

Podatki, ki jih bomo predstavili, imajo obliko slik in podatkov o pogledih, kjer so slike nastale, ter o kameri, ki je slike zajela. Začeli bomo s slikami, ki so nastali v umetnih, simuliranih okoljih, nadaljevali pa z nabori slik iz resničnih okolij.

## Nabori umetnih slik

Umetne slike nastanejo s pomočjo programov, ki izrisujejo predmete, prizore in okolja iz poligonov, tekstur ali s postopkom sledenja žarkov. Ti programi skušajo posnemati resnična okolja in prizore ter lahko dosegajo visoko stopnjo podobnosti z njimi. Prednost takšnih naborov je, da lahko umetna okolja, lokacije in postopek nastanka posamezne slike zelo natančno nadziramo, kar olajša zgodnje faze raziskav. Pri zajemanju slik namreč izločimo moteče elemente, ki bi lahko vplivali na rezultate, na primer nenatančnost meritev lokacij. Strojni postopki, ki nabore slik sprejemajo, ne ločijo med umetnimi in resničnimi slikami, sploh če umetna okolja vsebujejo dovolj elementov resničnih okolij. Pri spoznavanju metode in njenem prvem preskušanju pa so nepogrešljivi, saj skupaj s slikami dobimo tudi točne podatke, ki so nujni pri ocenjevanju uspešnosti in primerjanju izračunov z resničnimi podatki.

Umetne slike smo pripravili z grafičnim pogonom Source engine. Grafični pogoni za interaktivne sisteme, ki okolje izrisujejo v realnem času, so dosegli visoko stopnjo slikovne zvestobe izrisa. Source engine je tržni izdelek podjetja Valve Corporation<sup>1</sup>, ki uporabnikom omogoča ustvarjanje lastnih umetnih okolij in programiranje delovanja in obnašanja tega okolja. S tem orodjem smo pripravili dva nabora slik.

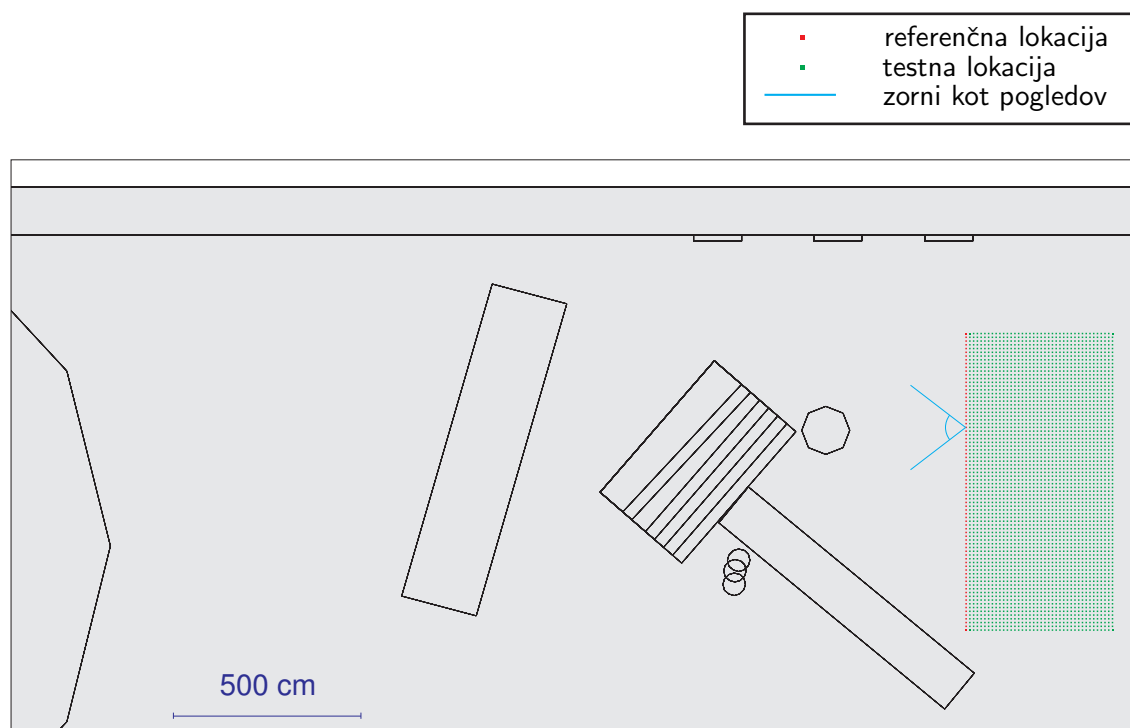
**Umetni prizor — zunanje okolje.** Pripravili smo navidezno okolje, ki ga prikazuje slika 5.1. Iz lokacij, ki opisujejo mrežo velikosti 80x40 mest, smo zajeli nabor slik. Razdalja med sosednjima mestoma je bila 10 cm, kamera pa je bila v vseh lokacijah enako usmerjena. Tloris na sliki 5.1 (a) prikazuje razsežnost okolja ter postavitev referenčnih in testnih lokacij. Slika 5.1 (b) prikazuje okolje samo skupaj z mrežo lokacij. Za referenčni nabor slik smo izbrali zaporedje 80 slik, ki so elementom prizora najbližji.

**Umetni prizor — okolje, podobno pisarni.** V praksi ne moremo pričakovati, da bodo vsi prizori imeli konfiguracijo, ki bi popolnoma ustrezala geometrijski poenostavitvi izrisovanja navideznih slik z ravnino na razdalji  $d_e$  od lokacij referenčnih pogledov. Za proučevanje vplivov elementov na različnih globinah prizora smo za pogon Source engine pripravili prostor, ki je podoben prostornejši pisarni. Slika 5.2 prikazuje tloris tega prostora, slika 5.3 pa enega izmed označenih pogledov.

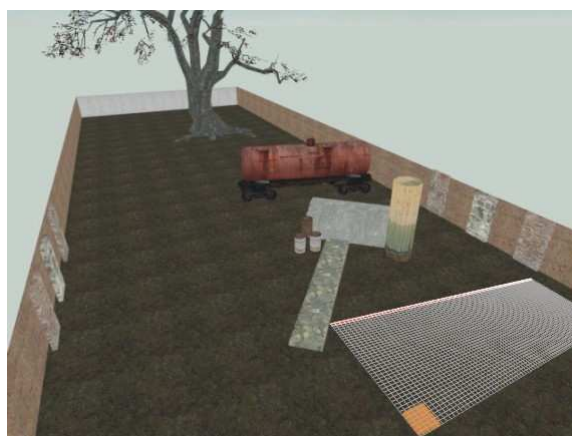
Prostor smo pripravili tako, da vsebuje dve plasti predmetov, ki delno zakrivajo pogled iz skrajne lokacije v prostoru na drugo stran prostora. Pisalne mize z zasloni tako predstavljajo manjšo, škatle na sredini prostora pa večjo stopnjo zakrivanja prizora. Na zadnjo steno prostora smo namestili takšne elemente, ki oponašajo

---

<sup>1</sup><http://www.half-life.com/>

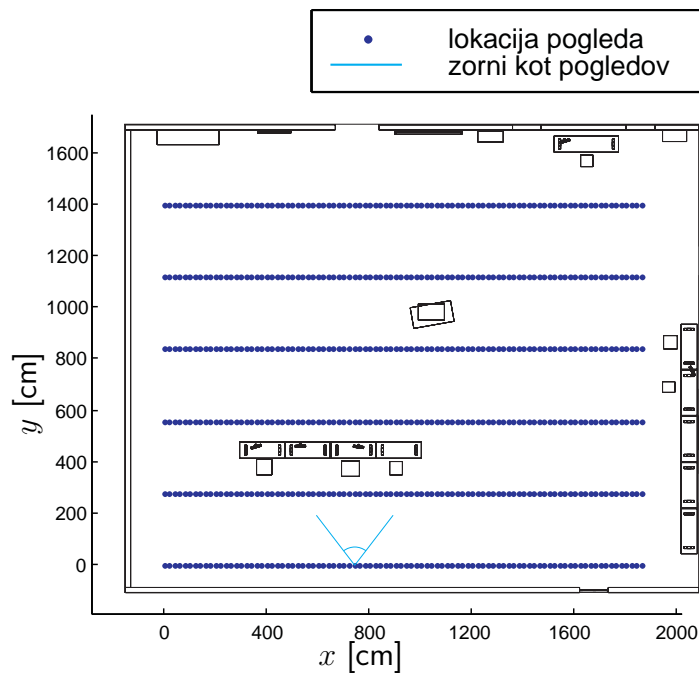


(a)



(b)

Slika 5.1: Prikaz umetnega prizora — zunanje okolje: (a) tloris prizora, (b) pogled na prizor z označenimi referenčnimi in testnimi lokacijami.



Slika 5.2: Tloris umetnega okolja, podobnega pisarni, z označenimi lokacijami pogledov in zornim kotom, ki je enak za vse poglede.



Slika 5.3: Pogled na umetno okolje, podobno pisarni, katerega tloris je na sliki 5.2.

naše delovno okolje, hkrati pa omogočajo nedvoumno razpoznavo iz slik pogledov poljubne oddaljenosti.

Tloris na sliki 5.2 vključuje lokacije, na katerih smo zajeli slike pogledov. Smer pogledov je v vsaki lokaciji vzporeden z osjo  $y$  in usmerjen proti naraščajočim vrednostim koordinate  $y$ . Razdalja med zaporednima lokacijama je 20 cm, korak v smeri  $y$  pa 120 cm.

### Nabori resničnih slik

Glavni namen preskusov je prikazati delovanje metode v resničnem okolju. Pri zajemanju resničnih slik imamo precej manj nadzora, hkrati pa so pogoji enaki kot med delovanjem sistema v živo. Resnična okolja so lahko precej kompleksnejša od umetnih, poveča pa se tudi negotovost pri ugotavljanju znanih lokacij ter hkrati zmanjša natančnost meritev.

Za zajemanje teh slik smo uporabili različno opremo:

- digitalni fotoaparati Canon PowerShot S60 na tračnicah,
- digitalna kamera Flea na robotu ATRVMini.

Tračnice, na katere smo položili digitalni fotoaparati, nam omogočajo relativno natančno določanje lokacije zajetih slik in stalnost usmeritve kamere. Z njimi smo dosegli nabor slik, ki je po stopnji nadzora nad zajemom blizu umetnim naborom slik. S tem postopkom smo zajeli množico slik, ki prikazuje *novoletno jelko*. Tak prizor smo izbrali zaradi njegove kompleksnosti, ki je posledica pretežno nepravilnih (ne-geometrijskih) oblik, gostote elementov, njihovega medsebojnega prekrivanja in pojavitev na različnih globinah prizora. Slika 5.4 (a) prikazuje primer slike iz nabora, ki smo ga uporabili za referenčne slike, slika 5.4 (b) pa primer testne slike.

**Umerjanje kamere.** Postopek za izrisovanje navideznih slik potrebuje podatke o smeri žarka, ki je prispeval k vrednosti posameznega slikovnega elementa na sliki. Teh podatkov pri delu s kamerami v resničnem okolju ne dobimo, lahko pa jih izračunamo, če poznamo velikost slikovnih elementov in goriščno razdaljo kamere. Te notranje parametre kamere dobimo z umerjanjem (kalibracijo) kamere.

Notranje parametre dobimo lahko samodejno med 3D rekonstrukcijo iz zaporedja slik [16] ali s postopkom, ki parametre izračuna iz slik znanega in vnaprej pripravljenega vzorca [61]. Mi smo uporabili orodje<sup>2</sup>, ki uporabi slike šahovnice podanih

---

<sup>2</sup>Camera Calibration Toolbox for Matlab, <http://www.vision.caltech.edu/bouguetj/calib.doc/>



Slika 5.4: Primer slik nadzorovanega resničnega prizora iz (a) referenčnega nabora in (b) testnega nabora, oddaljenega od prizora.

razsežnosti, vrne pa oceno koordinat središča slike, goriščne razdalje in parametre radialnega popačenja slike.

## 5.2 Metode na podlagi videza

Pomemben del motivacije za razvoj metode, opisane v disertaciji, so metode, ki za predstavitev okolja in lokalizacijo mobilnega robota uporabljajo videz iz pogledov. V nadaljevanju bomo ponazorili delovanje metode, ki za modeliranje videza uporablja prostor globalnih značilnic, zgrajen z metodo glavnih komponent. Podobnost med parom slik meri z razdaljo med projekcijama slikovnih vektorjev v podprostoru značilnic. Pri preskusih bomo uporabili slike, zajete z mobilnim robotom.

### 5.2.1 Lokalizacija v eni dimenziji

Če se robot med raziskovanjem giblje v neki smeri, ki je ne spreminja, dobimo enodimenzionalno linearno zaporedje. Takšno zaporedje bi lahko uporabili za lokalizacijo na hodnikih [34, 13] ali cestah [6]. Robot mora pri sledenju učne poti voziti na enakem odmiku od robov poti kot med učenjem. Rezultat lokalizacije je potem odmik trenutnega pogleda od izhodišča vzdolž poti.

Enodimenzionalno lokalizacijo bomo prikazali z uporabo dveh zaporedij, zajetih med vožnjo robota po ravni poti preko laboratorija. Za prvo zaporedje smo za zajem slik uporabili običajno perspektivno kamero, za drugo zaporedje pa panoramski senzor.

### Perspektivna kamera

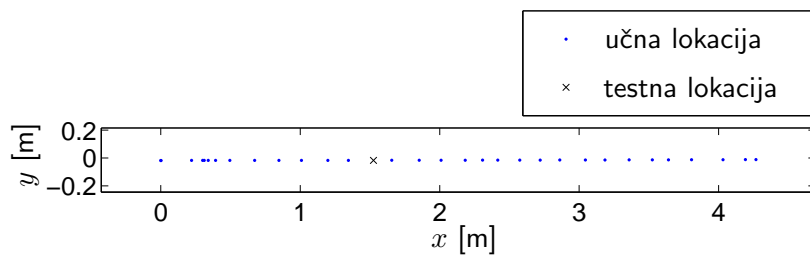
Kamero smo na robota namestili tako, da je bil njen pogled usmerjen bočno na smer vožnje robota. V tej smeri se videz pogleda med ravnim gibanjem spreminja v največji meri. Poleg tega je takšno zaporedje osnova za prostorsko-časovno predstavitev, ki jo uporablja lokalizacija z navideznimi pogledi, to pa nam bo kasneje omogočalo primerjavo delovanja obeh metod.

Postavitev učnih lokacij vidimo na sliki 5.5 (a). Iz učnih slik smo zgradili podprostor globalnih značilnic, zgoščen na 30% prvotnega prostora v pomnilniku (85% energije variance). Testno sliko, ki ni del učne množice, smo projicirali v ta podprostor in izračunali razdalje od dobljene projekcije do učnih projekcij. Te razdalje prikazuje slika 5.5 (b) v odvisnosti od lokacije učnih slik. Točka v najnižji vrednosti krivulje predstavlja učno sliko, ki je testni najbolj podobna. Rezultat lokalizacije pri danem testnem pogledu je vrednost  $x = 1,66$  m, kar je pri danih učnih lokacijah najboljši približek za testno sliko v lokaciji  $x_t = 1,53$  m. Vrednost slikovnih razdalj z oddaljenostjo od testne lokacije strmo naraste, saj se videz hitro spreminja. Slika 5.5 (c) prikazuje učne slike, označene na sliki 5.5 (b), slika 5.5 (d) pa prikazuje testno sliko.

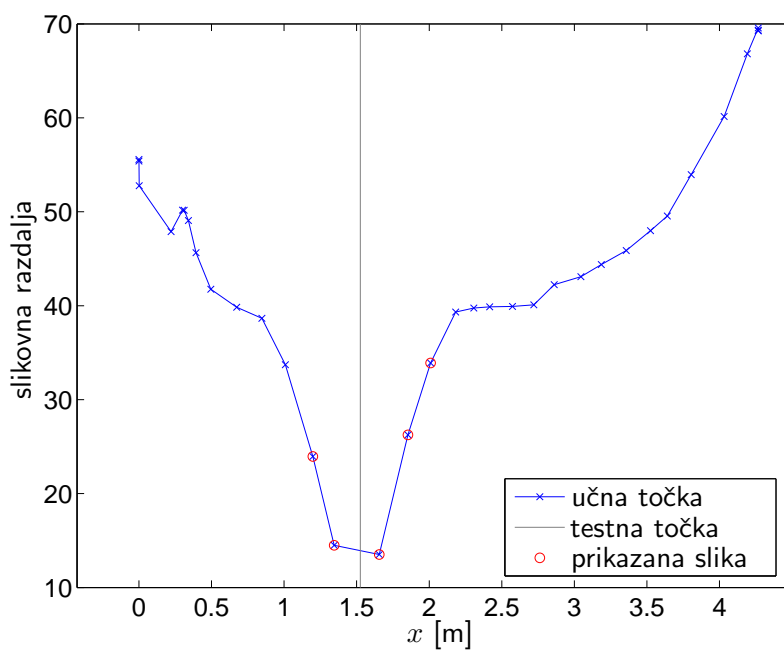
### Panoramska kamera

S panoramsko kamero na robotu smo ponovili nalogo, opisano v predhodnem razdelku. Slika 5.6 (a) prikazuje ravno pot z učnimi slikami ter testno lokacijo, ki ni iz množice učnih slik. Tudi v tem primeru smo obdržali 85% energije variance, kar zmanjša obseg podatkov na 16% prvotnega obsega. Testno sliko smo projicirali v podprostor značilnic in dobili krivuljo razdalj do učnih projekcij, prikazano na sliki 5.6 (b). Podobno kot prej dobimo minimum v neposredni bližini testne slike, krivulja pa je bolj gladka. Razdalja do pogleda, v katerem sliko še lahko označimo za podobno testni sliki, je daljša za okoli 20 cm. Razlog za to je, da se deli slike, ki prikazujejo prizor pred in za robotom, spreminjajo počasneje kot elementi ob straneh robota.

S preprostim preskusom, opisanim v predhodnem in tem razdelku, smo prikazali delovanje prepoznavanja lokacij iz zaporedja pogledov. Krivulja funkcije slikovnih razdalj do testne slike v okolici minimuma nakazuje, kakšen doseg imajo učne slike na danem območju. Panoramska kamera nam vrne slike, katerih videz se med gibanjem spreminja na bolj zvezen način, kar lahko zmanjša gostoto učnih slik (za ceno manjše ločljivosti koordinat lokalizacije).



(a)



(b)

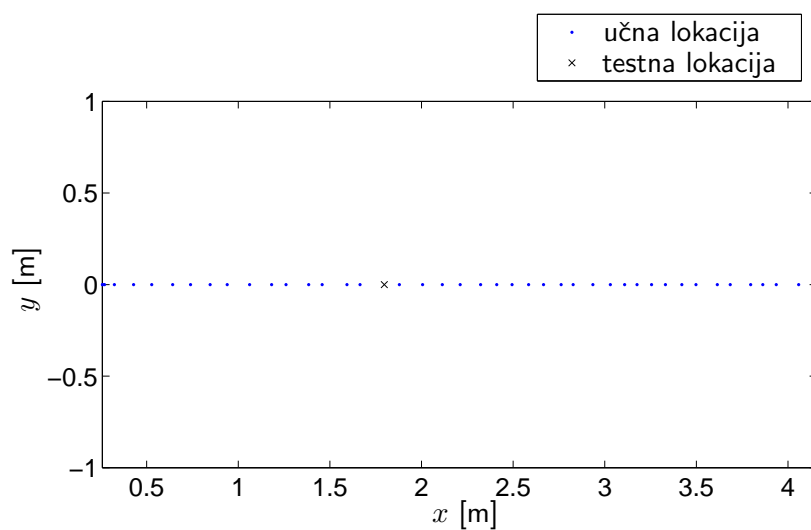


(c)

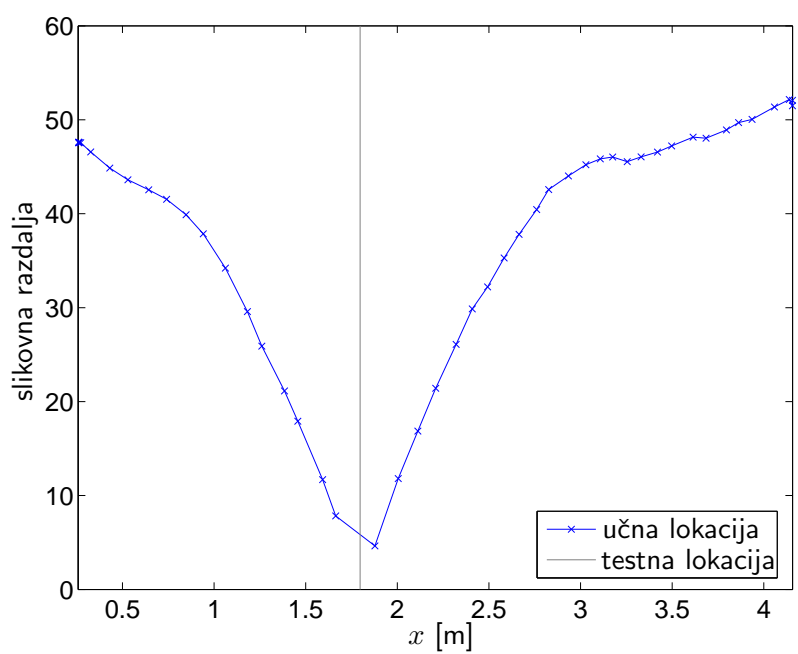


(d)

Slika 5.5: Lokalizacija v eni dimenziji s perspektivno kamero: (a) odometrija učnih in testne lokacije, (b) slikovne razdalje med testno in učnimi slikami, (c) slike učnih slik, označenih na sliki (b), (d) testna slika.



(a)



(b)

Slika 5.6: Lokalizacija v eni dimenziji s panoramsko kamero: (a) odometrija učnih in testne lokacije, (b) slikovne razdalje med testno in učnimi slikami.

## 5.2.2 Lokalizacija v dveh dimenzijah

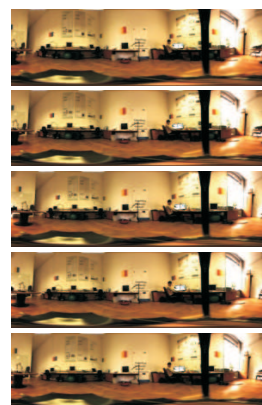
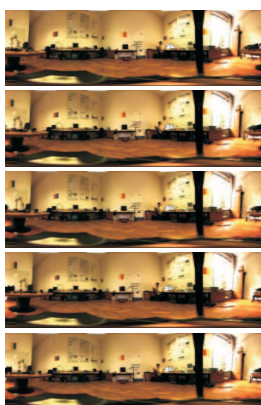
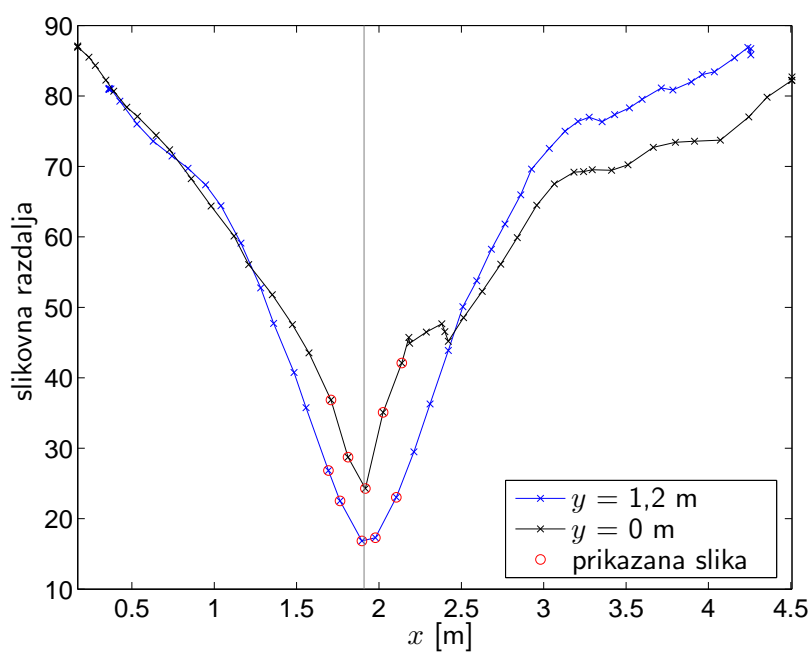
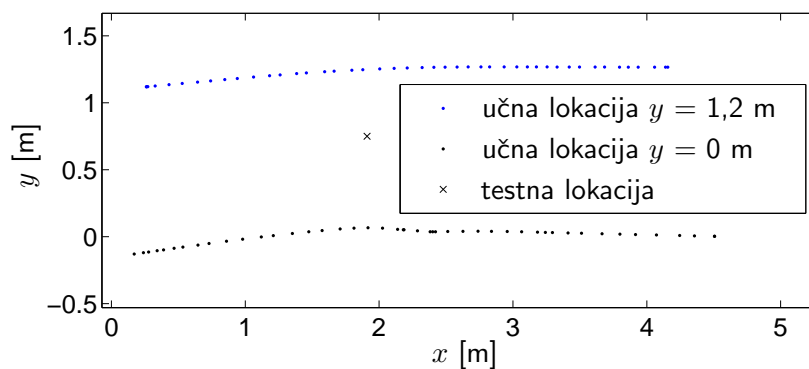
V podprostoru globalnih značilnic so projekcije porazdeljene glede na videz in vsebino slik. Strukturo prostora, ki ga podprostor predstavlja, in medsebojne odnose značilnic določajo podatki, ki jih dodatno shranimo s projekcijami učnih slik. Če je torej enodimenzionalna lokalizacija bolj podobna določanju zaporedne številke učne slike, ki je najbolj podobna testni sliki, lahko dvodimenzionalno lokalizacijo dosežemo z ustreznim označevanjem učnih značilnic.

Na sliki 5.7 (a) vidimo postavitev naslednjega preskusa, kjer imamo panoramske učne slike porazdeljene v ravnini. Testno lokacijo smo izbrali na območju znotraj ovojnice učnih lokacij, pripadajoči pogled pa prikazuje slika 5.7 (d). Glede na krivuljo iz slike 5.6 (b) lahko pričakujemo, da bomo med učnimi slikami našli dovolj slik, podobnih testni sliki. To preverimo tako, da projiciramo testno sliko v podprostor učnih značilnic in izračunamo razdalje do njih.

Slika 5.7 (b) prikazuje dobljene razdalje. Prikaz na sliki je oblikovan tako, da odseva strukturo učnih podatkov: ravninski koordinati  $x$ ,  $y$  in vrednost funkcije. Slike pogledov, označenih na krivulji  $y = 0$  m, prikazuje slika 5.7 (c), poglede, označene na krivulji  $y = 1,2$  m, pa slika 5.7 (e). Pri iskanju ocene lokacije testne slike imamo sedaj več možnosti. Prva možnost je, da krivuljo obravnavamo na enak način kot v predhodnih razdelkih, da torej izberemo zaporedno številko učne točke z najnižjo vrednostjo funkcije. Koordinate lokacije, shranjene skupaj s točko, nam dajo oceno lokacije testne slike. Za primer, prikazan na sliki 5.7 bi tako dobili koordinate učne lokacije, ki je izmed vseh tudi fizično najbližja testni lokaciji, lokalizacijska napaka pa bi bila enaka tej minimalni oddaljenosti. Druga možnost je, da obravnavamo dve ločeni zaporedji učnih slik, kot sta prikazani na sliki 5.7 (b). Vsaka krivulja nam da oceno lokacije iz svojega minimuma, končno oceno pa bi dobili z linearno kombinacijo med njima v razmerju vrednosti funkcij v minimumih. Takšna ocena bi bila lahko natančnejša od predhodne. Poleg tega lahko uporabimo zvezo učnih lokacij s slike 5.7 (b), da v podprostoru značilnic z interpolacijo izračunamo gostejšo mrežo značilnic, ki služijo kot učni primeri pri lokalizaciji.

## 5.2.3 Lokalizacija izven dosega učnih slik

V predhodnih razdelkih smo prikazali primere lokalizacije, pri katerih so bile testne lokacije znotraj območja, ki ga pokrivajo učne slike. Pri enodimenzionalnih primerih je bila testna lokacija v neposredni bližini dveh učnih lokacij, v dvodimenzionalnem primeru pa znotraj konveksne ovojnice učnih lokacij.



Slika 5.7: Lokalizacija v dveh dimenzijah s panoramsko kamero: (a) odometrija učnih in testne lokacije, (b) slikovne razdalje med testno in učnimi slikami, (c) učne slike  $y = 0$  m, (d) testna slika, (e) učne slike  $y = 1,2$  m.

Na sliki 5.8 pa prikazujemo rezultate poskusov lokalizacije, pri čemer smo testne lokacije izbrali iz vzporedne poti, oddaljene 3 m od poti, ki vsebuje učne lokacije. Lokalizacija v dveh dimenzijah tukaj ni mogoča, saj so testne lokacije preveč oddaljene od najbližjih učnih lokacij. Prikazali bomo le koordinate  $x$  najbolj podobnih učnih slik. Za rezultate na sliki 5.8 (a) smo uporabili panoramske slike. Krivulja ocena lokacije prikazuje koordinate najbolj podobnih učnih slik za posamezno testno sliko, odometrija pa pripadajoče izmerjene koordinate. Vidimo, da lokalizacija v veliki večini primerov vrne napačne koordinate. Še slabše rezultate dobimo, če uporabimo perspektivno kamero, kot vidimo s slike 5.8 (b). Krivulja ocena lokacije v tem primeru prikazuje pretežno poljubne vrednosti.

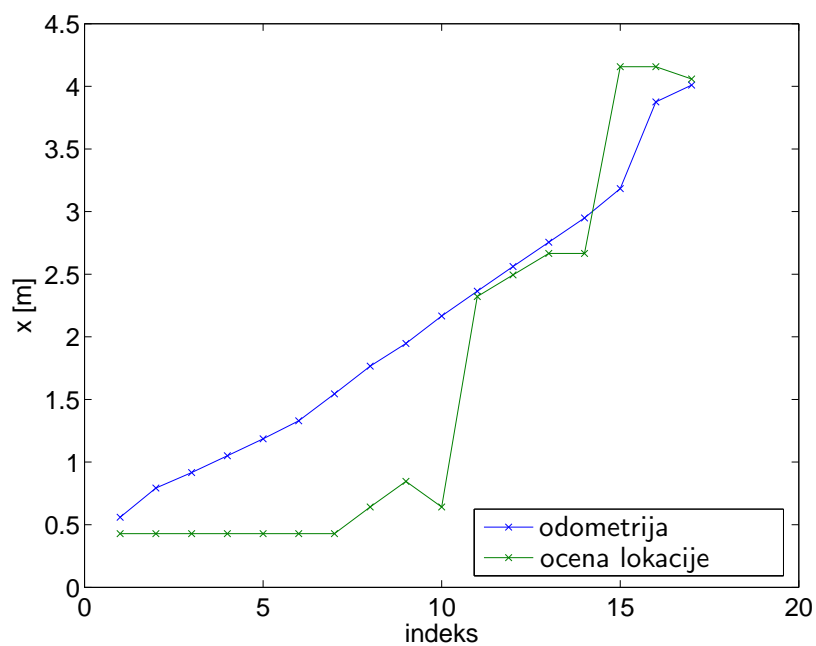
Opisani preskusi ponazarjajo slabosti in omejitve običajnih metod lokalizacije na podlagi videza. V nadaljevanju bomo preskusili delovanje postopka, opisanega v disertaciji, in ugotovili, v kakšni meri odpravi omenjene omejitve.

## 5.3 Izrisovanje navideznih slik

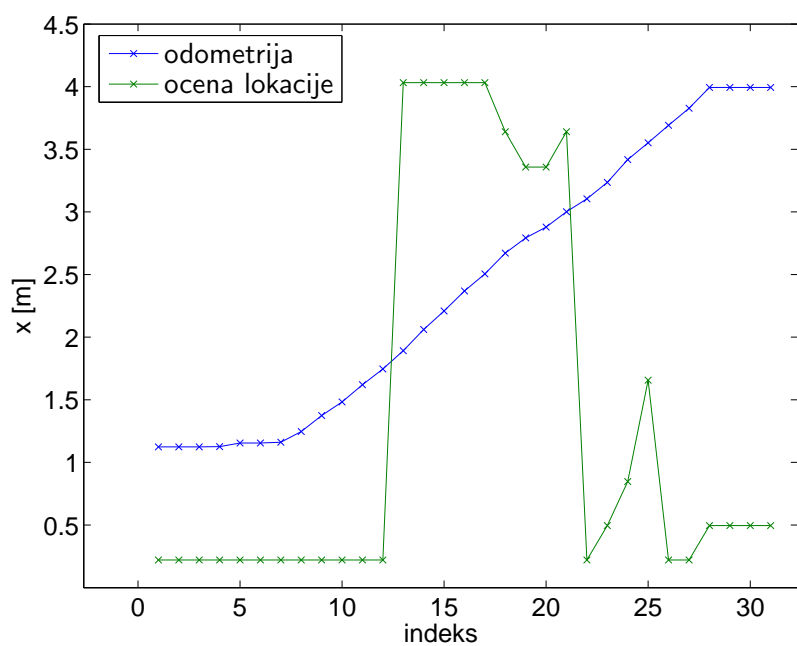
Postopki za izrisovanje navideznih slik so v osnovi namenjeni vizualizaciji navideznih pogledov v sistemih za navidezno resničnost ali kot pomoč operaterjem pri upravljanju mobilnih robotov. Njihov rezultat je torej slika, katere pravilnost lahko preverimo z ogledom. V nadaljevanju bomo zato prikazali nekaj navideznih slik iz naborov referenčnih slik, ki smo jih predstavili. Bolj kot sam videz navideznih slik pa nas bo zanimalo numerično ocenjevanje njihove kakovosti.

### 5.3.1 Primeri navideznih slik

Iz umetnega nabora slik v ločljivosti  $640 \times 480$  slikovnih elementov smo izbrali referenčne slike in jih uporabili kot prostorsko-časovni volumen. Izbrali smo si koordinati za nek navidezni pogled, ki simulira vzvratni premik od premice referenčnih lokacij, in za parameter oddaljenosti do prizora  $d_e$  določili najprej nizko, nato pa zelo visoko vrednost. V prvem primeru smo dobili rezultat, ki ga prikazuje slika 5.9 (a). V zgornjem in spodnjem delu vidimo področje, kjer slikovni elementi niso definirani. Sama slika vsebuje veliko napak v obliki večkratnih ponovitev istih elementov. Slika 5.9 (b) prikazuje rezultat pri vrednosti  $d_e$ , ki presega razdaljo do najbolj oddaljenega elementa prizora. Področje nedefiniranih slikovnih elementov je zelo majhno, kar kaže na to, da je premik od referenčnih slik relativno majhen v primerjavi z vrednostjo  $d_e$ . Učinek je obraten kot pri premajhnih vrednostih  $d_e$ , kar prikazujeta povečana

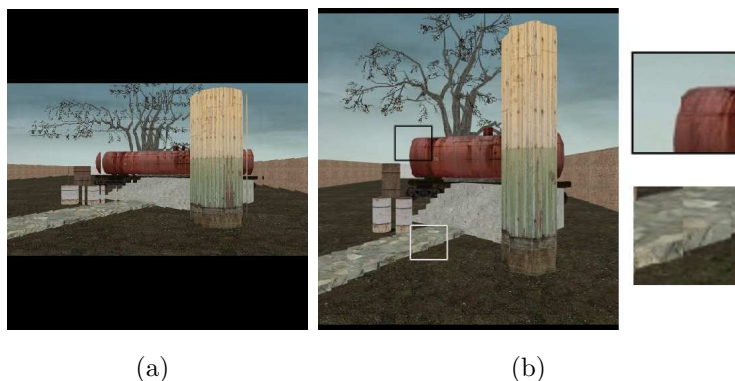


(a)



(b)

Slika 5.8: Rezultati lokalizacije izven območja, pokritega z uĉnimi slikami (a) za panoramske slike, (b) za perspektivne slike, pri ĉemer naša metoda predstavlja lokalizacijo z navideznimi slikami.

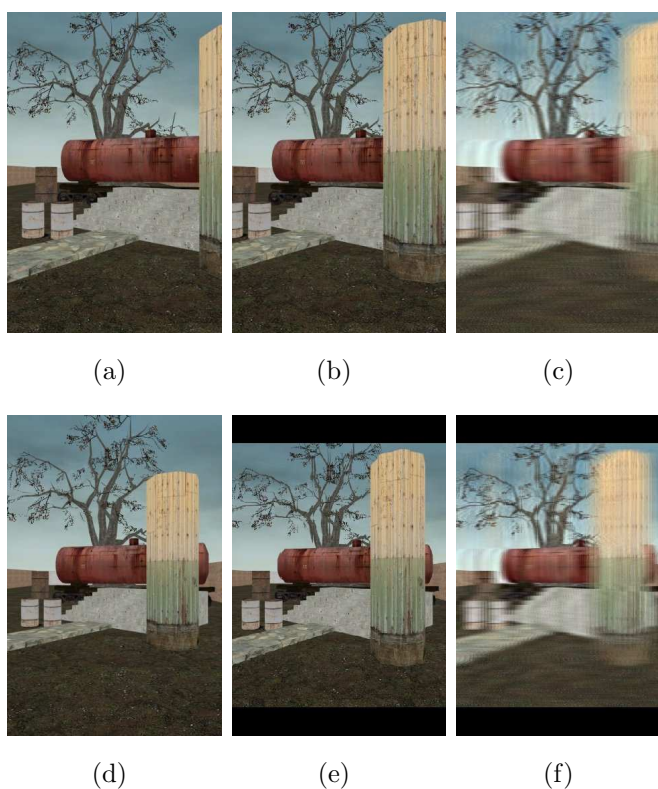


Slika 5.9: Vpliv parametra  $d_e$  na navidezno sliko, če je njegova vrednost (a) prenizka, (b) preveč oddaljena (označene povečave so prikazane na desni strani slike).

dela na desni strani slike 5.9 (b): nezveznosti in manjkajoči deli elementov prizora.

Iz prostorsko-časovnega volumna referenčnih slik smo s postopno gradnjo podprostora metode glavnih komponent zgradili zgoščeno predstavitev. Dobljena predstavitev ohranja 80% variance izvornih podatkov, v pomnilniku pa zasede 25% prostora. Slika 5.10 (c) prikazuje rekonstrukcijo referenčne slike, prikazane na sliki 5.10 (b), iz zgoščene predstavitve. Zaradi zgostitve rekonstrukcija kaže posledice zamegljevanja v času, vendar iz nje lahko razločimo posamezne predmete. Slika 5.10 (e) prikazuje navidezno sliko, zgrajeno iz (nezgoščene) prostorsko-časovnega volumna referenčnih slik. Koordinate navideznega pogleda, kjer je nastala ta slika, ustreza pogledu, v katerem dobimo resnični pogled, ki ga prikazuje slika 5.10 (d). Navidezna slika kaže na dober približek, saj izrisovanje modelira zakrivljanje delov prizora s predmeti v ospredju. Slika 5.10 (b) prikazuje najbližji referenčni pogled, slika 5.10 (a) pa referenčni pogled, iz katerega steber manj zakriva elemente za seboj. Če postopku izrisovanja navideznih pogledov podamo parametre, ki smo jih za pogled iz slike 5.10 (e), pri čemer uporabimo zgoščeno predstavitev, dobimo rezultat na sliki 5.10 (f).

Slika 5.11 prikazuje rezultat izrisovanja navidezne slike iz referenčnega nabora *novoletna jelka*. Prikazana navidezna slika je sestavljena iz stolpcev vsake od referenčnih slik in zato omogoča precej širši pogled na prizor kot je to mogoče iz referenčnih pogledov. Na njej je zelo malo napak in je v celoti videti prepričljiva.



Slika 5.10: Izrisovanje navideznih slik iz nabora umetnega prizora: (a) referenčni pogled 1, (b) referenčni pogled 2, (c) rekonstrukcija referenčnega pogleda 2 iz zgoščene predstavitev, (d) pogled, ki ni iz referenčnega nabora, (e) navidezni pogled, (f) navidezni pogled iz zgoščene predstavitev.

### 5.3.2 Kakovost navideznih slik

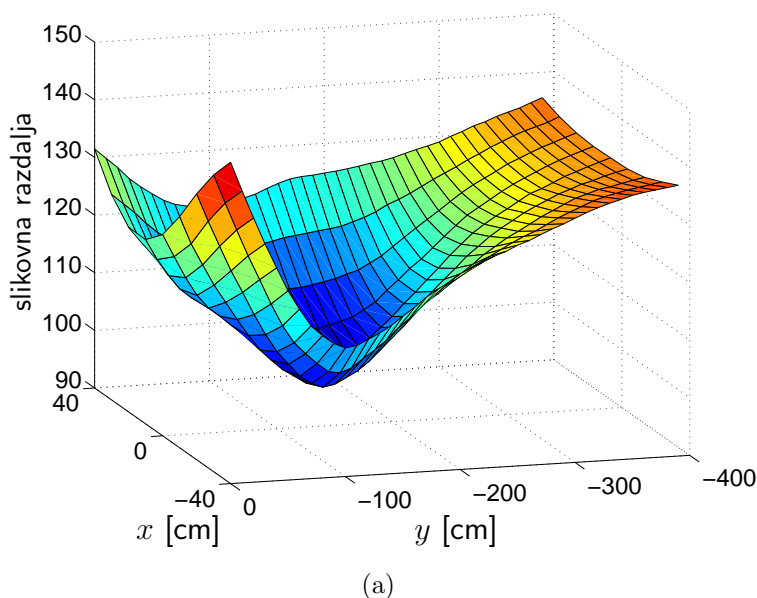
Kakovost navidezne slike numerično primerjamo tako, da dobljeno navidezno sliko primerjamo z resnično sliko iz lokacije, ki jo določajo parametri navidezne slike.

Ob merjenju slikovne razdalje med navidezno in pravo sliko v neki točki dobimo vrednost, ki sama po sebi ni dovolj informativna. Pravi pomen dobi šele, ko izračunamo slikovne razdalje od navidezne slike do pravih slik iz širše okolice izbrane navidezne lokacije. Preskus bomo torej izvedli tako, da bomo eno navidezno sliko primerjali z naborom resničnih slik.

Za ta preskus smo izbrali nabor umetnih slik, ustvarjenih z grafičnim pogonom Source engine. Izbrali smo si naključne vrednosti parametrov  $x_v$  in  $d_r$  ter določili neko smiselno vrednost parametru  $d_e$ . Dobljeno navidezno sliko smo primerjali z resničnimi slikami in dobili funkcijo razdalje, ki jo prikazuje slika 5.12. Funkcija



Slika 5.11: Navidezni pogled iz nabora nadzorovanega resničnega prizora.



Slika 5.12: Slikovne razdalje med resničnimi in navidezno sliko pri izbrani vrednosti  $d_e$  in  $d_r$ .

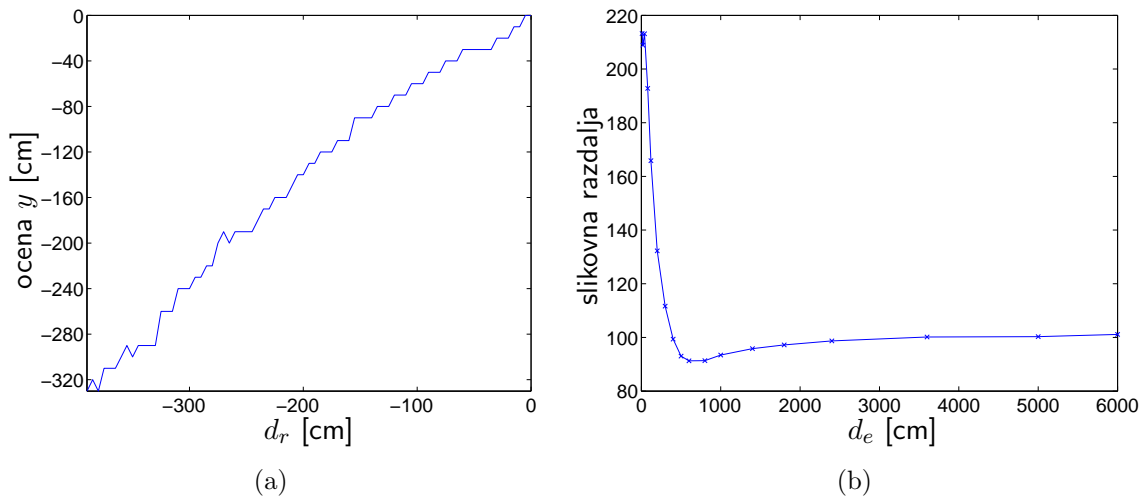
je odvisna od koordinat  $(x, y)$  resničnih slik in ima lepo obliko z izrazit minimum. Iz tega lahko sklepamo, da ima navidezna slika največjo podobnost z eno izmed resničnih slik in da mora ta resnična slika biti tudi vizualno podobna navidezni sliki.

Koordinati  $x$  in  $y$  točke najnižje vrednosti funkcije s slike 5.12 bi morali biti v neposredni zvezi parametrov  $x_v$  in  $d_r$  navidezne slike. Ugotovimo, da koordinata  $x$  dejansko ustreza podanemu parametru  $x_v$  navidezne slike, koordinata  $y$  pa od podane vrednosti  $d_r$  odstopa za 45 cm.

Vprašanje, ki si ga tu postavimo, je, kako natančno parameter  $d_r$  določa koordi-

nato  $y$  najbolj podobnega resničnega pogleda. Za odgovor na to vprašanje smo pri fiksnem  $x_v$  spreminjali  $d_r$ , ustvarili zaporedje navideznih slik, ter za vsako od njih poiskali najbolj podobno resnično sliko. Na sliki 5.13 (a) lahko vidimo koordinate  $y$  tako najdenih resničnih slik. Oblika krivulje je približek linearne funkcije, kar pomeni, da  $d_r$  dejansko deluje po pričakovanju. Vendar pa dobljena funkcija nima takšnega naklona, da bi dobili  $d_r = y$ . Razlog za to je, da smo parameter  $d_e$  nastavili poljubno in torej pri podanem okolju neustrezno.

Parameter  $d_e$  se navezuje na specifičnost okolja, v katerem želimo izrisati navidezne poglede. Če mu nastavimo pravilno vrednost, bomo dobili navidezne slike, ki so pravim slikam podobne v največji meri, saj bo na takšni sliki najmanj popačenj, ponovljenih ali izpuščenih elementov prizora. Predvidevamo torej lahko, da bo tudi v numeričnem smislu navidezna slika s pravilno nastavljenjo vrednostjo tega parametra najbolj podobna resnični sliki. To hipotezo smo preverili podobno kot prej, torej ob fiksnih  $x_v$  in  $d_r$  smo spreminjali vrednost  $d_e$  in za dobljene navidezne slike poiskali vrednost slikovne razdalje do najbolj podobne resnične slike. Krivuljo te funkcije vidimo na sliki 5.13 (b), rezultat pa potrjuje hipotezo, da obstaja takšna vrednost  $d_e$ , pri kateri je dobljena navidezna slika v največji meri podobna najbolj podobni resnični sliki, hkrati pa ta vrednost ustreza razdalji do najbolj dominantnega predmeta v okolju. Ta predmet je torej takšen, da ga mora postopek izrisovanja navideznih slik izrisati čim bolj pravilno, v nasprotnem primeru (npr., ko je pravilno izrisan



Slika 5.13: (a) Koordinata  $y$  najbolj podobne resnične slike navidezni sliki pri parametru  $d_r$ ; (b) slikovna razdalja do najbližje resnične slike v odvisnosti od  $d_e$  navideznih slik.

drugače oddaljen predmet) pa napake pri izrisovanju povišajo vrednost mere slikovne razdalje. Stopnja dominantnosti predmeta na sliki je sorazmerna deležu slike, ki jo predmet zavzema, ter stopnji izstopanja intenzitete predmeta v primerjavi z ostalimi predmeti na sliki (svetlejši ali temnejši od okolice).

Opisani preskusi so na primeru pokazali, da z izrisovanjem navideznih slik dobimo slike, ki so primerljive resničnim slikam in katerih parametri ustrezajo lokacijam in lastnostim dejanskega prizora. Podobno lahko ugotovimo, da je mera slikovne razdalje ustrezna za nadaljnjo uporabo metode za lokalizacijo.

### Primerjava s podobnimi metodami

V predhodnem preskusu smo na primeru ene navidezne slike pokazali ustreznost tako postopka za izrisovanje navideznih slik kot same mere za primerjanje slik. V nadaljevanju bomo preskus razširili tako, da bo vključeval celoten nabor lokacij, iz katerih imamo znane poglede. S tem bomo posplošili rezultate iz posamezne izbrane navidezne lokacije na celotno pokrito okolje.

Drugi namen naslednjega preskusa je utemeljitev izbranega postopka za izrisovanje navideznih slik. Izbrani postopek je sicer preprost, vendar vseeno bolj zapleten kot povečevanje ali zmanjševanje le ene referenčne slike. Poleg tega uporabljamo metodo glavnih komponent za zgoščevanje predstavitve volumna, kar vpliva na kakovost navidezne slike. S preskusom bomo torej opravičili dodatno kompleksnost sestavljanja navideznih slik.

Preskus smo izvedli tako, da smo ustvarili navidezne poglede pri različnih parametrih  $x_v$  in  $d_r$ , pri čemer je bila vrednost  $d_e$  pri vseh navideznih lokacijah enaka. Za vsako navidezno sliko smo poiskali najbolj podobno resnično sliko in shranili njuno medsebojno razdaljo. Navidezne slike smo ustvarili z naslednjimi postopki:

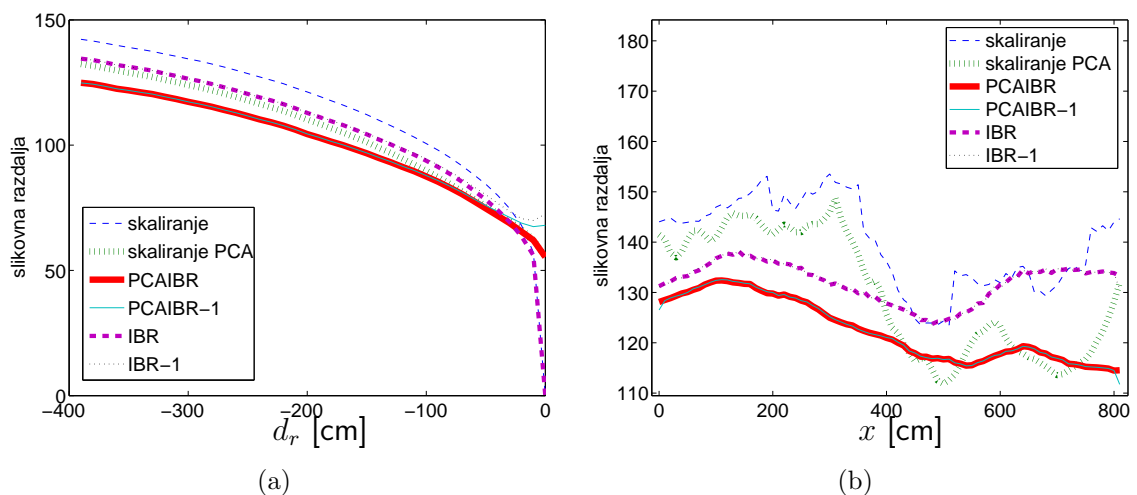
- skaliranje izbere najbližjo referenčno sliko  $x_i = x_v$  in jo poveča ali zmanjša v odvisnosti od parametrov  $d_r$  in  $d_e$ . To doseže tako, da uporabi postopek izrisovanja navideznih slik, pri čemer je prostorsko-časovno volumen referenčnih slik sestavljen samo iz slike  $\mathbf{A}_i$ , dobljena slika pa je povečana ali pomanjšana za  $\frac{d_e}{d_e - d_r}$ ,
- IBR je postopek izrisovanja iz slik, kot smo ga opisali v disertaciji,
- IBR-1 je enak postopek kot IBR, le da smo iz referenčne množice slik odstranili najbližjo referenčno sliko ( $x = x_v$ ) lokaciji trenutno izrisane navidezne slike. Pomen tega postopka je, da pokažemo, da celoten postopek še vedno

deluje, četudi manjka navidezni lokaciji najbližja referenčna slika, brez katere pa postopek skaliranje ne more delovati.

Vsakega od postopkov smo ponovili z uporabo metode glavnih komponent in dobili še postopke skaliranje PCA, PCAIBR in PCAIBR-1. Uporabili smo zgostitev, ki ohrani 80% variacije izvirnih podatkov, in dobili predstavitev, ki je zmanjšala zahtevnost po pomnilniku na 25% izvirne velikosti. Slika 5.14 prikazuje primerjavo delovanja naštetih metod. Vidimo, da dobimo najbolj podobne navidezne slike z izrisovanjem iz slik, ki za referenčni volumen uporablja predstavitev z metodo glavnih komponent (PCAIBR). Prednost tega postopka se povečuje, ko se oddaljujemo od poti z referenčnimi lokacijami.

Najslabše se obnese metoda, ki povečuje ali zmanjšuje referenčno sliko. Postopek skaliranje PCA vrača nižje slikovne razdalje in deluje torej bolje od postopka skaliranje. V sredini sta postopka IBR in IBR-1, ki dajeta identične rezultate povsod razen pri nizkih vrednostih  $d_r$ . Najboljša sta postopka PCAIBR in PCAIBR-1 ki, tako kot predhodnika, delujeta enako razen pri nizkih vrednostih  $d_r$ . Prednost postopkov, ki uporabljajo metodo glavnih komponent, pred ekvivalentnimi metodami brez zgoščene predstavitve lahko razložimo tako, da zaradi glajenja postanejo napake numerično manj vplivne pri izračunu slikovne razdalje.

Postopek skaliranje (PCA) ima pred postopkom (PCA)IBR-1 prednost pri vrednostih  $|d_r|$ , ki so manjše od ločljivosti referenčnih lokacij. Ta prednost je zanemarljiva,



Slika 5.14: Primerjava uspešnosti delovanja različnih metod: (a) povprečna vrednost v odvisnosti od  $d_r$  in (b) pri  $d_r = -350$  cm.

saj nastane zaradi referenčne slike, ki smo jo pri (PCA)IBR-1 namenoma izpustili. V praksi je namreč bolj verjetno, da pri podanem  $x_v$  referenčne slike ne bo, zaradi česar postopek skaliranje v večini primerov sploh ni mogoč.

Zaključimo lahko, da predlagana metoda učinkovito ustvari vizualno zveste navidezne poglede v poljubni navidezni lokaciji in tudi pri večji oddaljenosti od referenčnih lokacij. Postopek nam torej omogoča pripravo hipotez, ki jih primerjamo s sliko neznane lokacije pri lokalizaciji.

## 5.4 Lokalizacija

V preskusih, ki smo jih izvajali do sedaj, smo želeli utemeljiti izbor postopka za izrisovanje navideznih pogledov. To smo uresničili tako, da smo podano navidezno sliko primerjali z množico resničnih slik. Prostor smo imeli pokrit s slikami znanih resničnih pogledov, zato je bil takšen pristop izvedljiv. Pri lokalizaciji pa imamo situacijo obrnjeno: dobimo eno samo resnično sliko, njene lokacije pogleda ne poznamo, lahko pa ustvarimo množico navideznih slik, ki jih z njo primerjamo. Tako deluje postopek lokalizacije, opisan v disertaciji.

V nadaljevanju bomo najprej pokazali, da dobimo z našo metodo takšen preiskovalni prostor, po katerem lahko izvajamo optimizacijo in katere rezultat je ocena trenutne lokacije iz slike pogleda. Nato bomo opisali preskuse lokalizacije v nadzorovanih pogojih, najprej v eni in nato v dveh dimenzijah. Sledil bo prikaz vplivov različnih predmetov v okolju, ki delno zakrivajo prizor. Nazadnje bomo pokazali rezultate lokalizacije mobilnega robota.

### 5.4.1 Iskanje najustreznejšega navideznega pogleda

Robot se običajno giblje v ravnini med lokacijami, ki imajo koordinati  $x$  in  $y$ . Ko želi na neki lokaciji izvedeti vrednosti teh parametrov, zajame sliko trenutnega pogleda, z njo pa izvede povpraševanje pri postopku za lokalizacijo.

V predhodnem razdelku smo pokazali, da imata koordinati  $x$  in  $y$  resničnega pogleda neposredno zvezo s koordinatama  $x_v$  in  $d_r$  navideznega pogleda, če je vrednost parametra  $d_e$  pravilno nastavljena. Oceno koordinat  $x$  in  $y$  lahko torej dobimo tako, da poiščemo vrednosti  $x_v$  in  $d_r$  navideznega pogleda, iz katerega je slika najbolj podobna podani sliki. Prostor parametrov navideznih slik tako postane preiskovalni prostor, v katerem iščemo najnižjo vrednost cenilne funkcije  $d_{\text{nssd}}$  iz izraza (4.3).

Parameter  $d_e$  bi v praksi lahko ugotovili s senzorjem oddaljenosti ali s pomočjo

triangularizacije para lokalnih značilnic, vendar želimo opisano metodo uporabiti neodvisno od ostalih naprav in postopkov. Zato ima tudi parameter  $d_e$  neznano vrednost, vrednosti vseh treh parametrov pa skušamo ugotoviti z minimizacijo slikovnih razdalj. Z naslednjim preskusom želimo ugotoviti, ali je opisana metoda primerna za iskanje ocene lokacije neznanega pogleda.

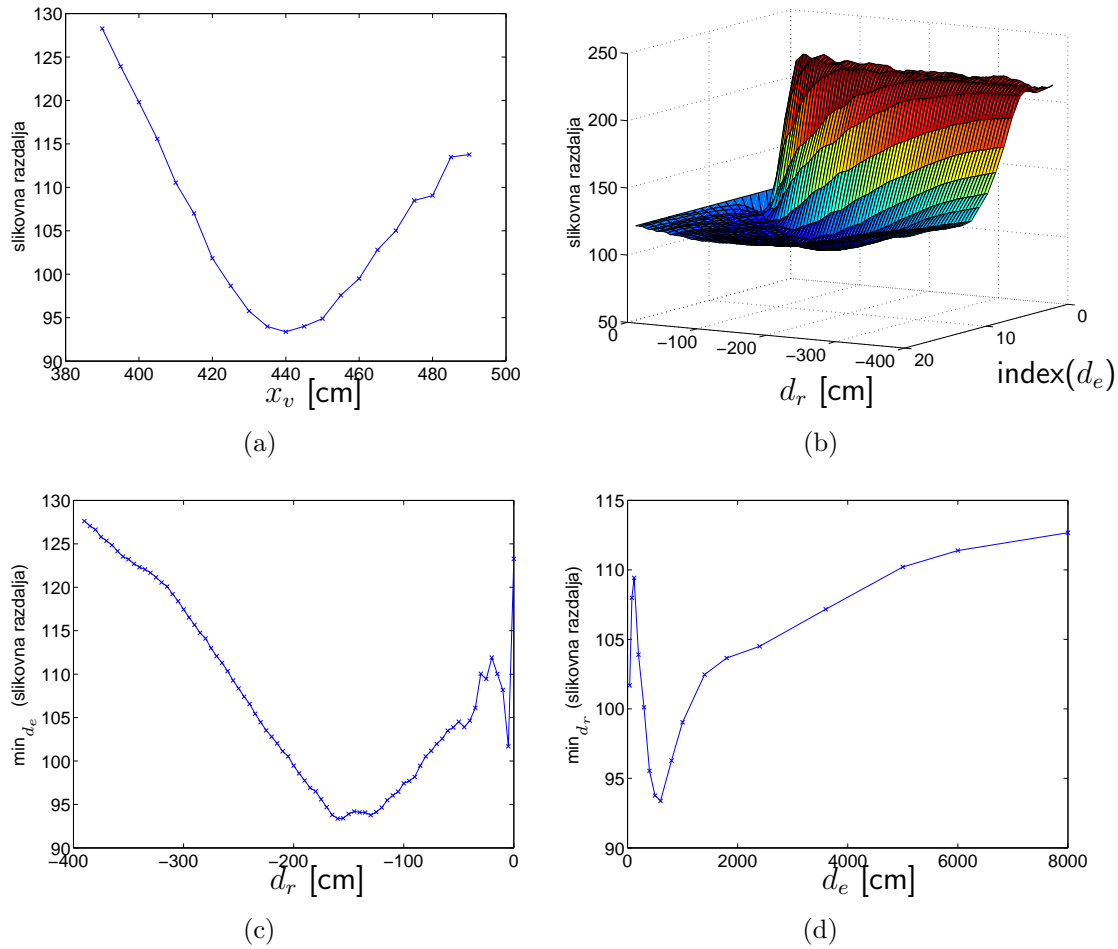
Vzeli smo nabor slik, ustvarjenih z grafičnim pogonom Source engine, in naključno izbrali sliko s koordinatami  $x = 440$  cm in  $y = 160$  cm. V tej lokaciji in v njeni okolici smo določili množico navideznih pogledov, v vsakem od njih smo ustvarili navidezno sliko za množico vrednosti parametra  $d_e$  ter izračunali  $d_{\text{nssd}}$  med njimi in izbrano resnično sliko. Dobili smo tridimenzionalno strukturo z dobljenimi vrednostmi slikovnih razdalj.

Na sliki 5.15 (a) lahko vidimo najmanjše vrednosti slikovnih razdalj v odvisnosti od parametra  $x_v$ . Funkcija ima izrazit globalni minimum pri  $x_v = 440$  cm, torej v vrednost  $x$  izbrane resnične slike. Če potem pogledamo prerez preiskovalnega prostora pri  $x_v = 440$  cm, dobimo površino, ki jo prikazuje slika 5.15 (b).

Sliki 5.15 (c) in (d) prikazujeta najmanjše vrednosti površine iz slike 5.15 (b) za parametra  $d_r$  in  $d_e$ . Funkcija se pri zelo majhnih vrednostih  $d_e$  obnaša dokaj nepredvidljivo, kar je posledica velikosti slik 10% izvirne velikosti ali manj. Drugod je funkcija oblikovana lepo, globalni minimum pa ima v  $d_r = -160$  cm, kar je vrednost  $y$  izbrane resnične točke. Vrednost  $d_e$ , ki jo izbere minimizacija, ustreza oddaljenosti od referenčne poti do cisterne na sliki 5.1. Ta element se je izkazal kot najbolj dominanten, saj je najbolj kontrasten glede na ostale elemente prizora in hkrati zavzema dovolj veliko površino slike.

Neničelna vrednost minimuma slikovne razdalje nakazuje, da je slika v navideznem pogledu, ki sovpada z resničnim pogledom, le približek prave slike. Slikovno razdaljo povečujejo napake pri modeliranju preslikav v navpični smeri predmetov, ki niso na razdalji  $d_e$  od referenčnih pogledov (npr. steber na slikah 5.10 (d) nima enake višine kot tisti na sliki 5.10 (e)), ponavljajoči se ali manjkajoči elementi na teh predmetih, izguba podrobnosti pri približevanju ali normalizacija mere slikovne razdalje zaradi nedefiniranih slikovnih elementov pri oddaljevanju. Nadalje na vrednost slikovne razdalje vpliva izguba določenih podrobnosti zaradi zgostitve predstavitve. Vseeno bomo pri primerjanju dveh neustreznih pogledov dobili višje vrednosti slikovne razdalje, saj bodo razlike med slikama večje zaradi slabega prekrivanja elementov ali celo povsem različne vsebine slike.

Preskus je na poljubno izbranem pogledu iz nabora pokazal, da s predstavljeno metodo lahko poiščemo oceno lokacije neznanega pogleda. Poleg koordinat pogleda



Slika 5.15: Slikovne razdalje v preiskovalnem prostoru: (a) najmanjše razdalje v odvisnosti od  $x_v$ , (b) funkcija v odvisnosti od  $d_r$  in  $d_e$  pri  $x_v = 440$  cm, (c) najmanjše razdalje v odvisnosti od  $d_r$  in (d)  $d_e$ .

smo dobili tudi vrednost parametra  $d_e$ , ki je nosila vrednost razdalje do enega od elementov prizora. Za izbrani primer smo videli, da je oblika preiskovalnega prostora ustrezna za numerično optimizacijo. Če to velja za vse ostale poglede iz testnega nabora, pa bomo preverili v nadaljevanju.

### 5.4.2 Lokalizacija v eni dimenziji

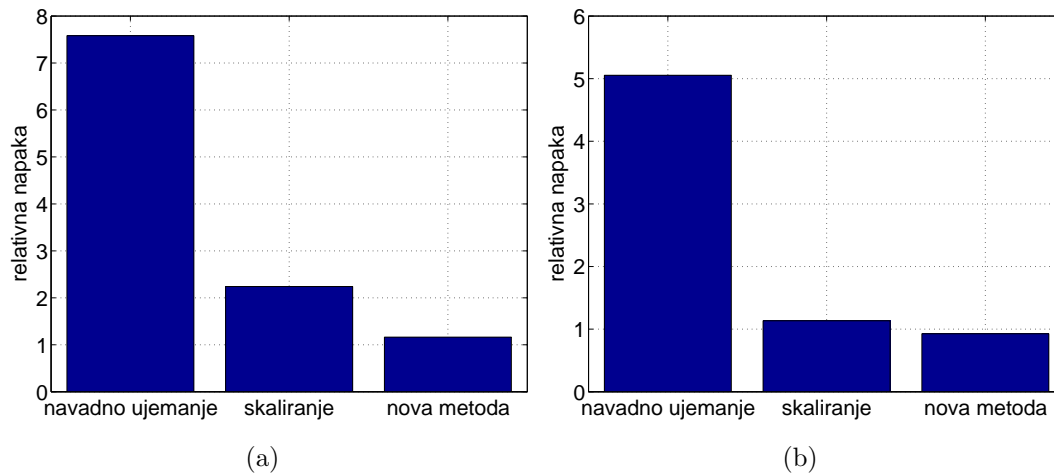
V prvi fazi preskusov lokalizacije se bomo osredotočili na sposobnost predlagane metode, da določi koordinato  $x$ , torej na lokalizacijo v eni dimenziji. Tako pri umetnem kot pri naravnem naboru slik smo za učne slike uporabili le polovico nabora

zajetih slik, in sicer tako, da smo izpustili slike s sodo zaporedno številko. Ker smo slike zajeli v obliki mreže, je tako polovica testnih slik obdržala najbližjo referenčno sliko na lokaciji, katere koordinata  $x$  je enaka koordinati  $x$  testne slike, druga polovica testnih slik pa se je znašla med dvema zaporednima učnima slikama. S takšnim izborom učnih lokacij smo se približali realni situaciji, v kateri običajno neznana slika ne deli koordinate  $x$  s katero od učnih lokacij.

Tako z umetnim kot naravnim naborom slik smo za vsako testno sliko poiskali oceno parametra  $x_v$  in odšteli njegovo vrednost od vrednosti  $x$  testne slike. Absolutna vrednost te razlike predstavlja *lokalizacijsko napako* za parameter  $x_v$  testne slike. To lokalizacijsko napako smo delili s polovico razdalje med sosednjima učnima lokacijama in dobili relativno lokalizacijsko napako. Vrednost te napake, ki je enaka 1 ali manj, pomeni, da je ocena lokacije znotraj ločljivosti učnih lokacij.

Takšno obliko meritve smo izbrali, ker smo želeli delovanje predlagane metode primerjati z delovanjem dveh preprostejših metod. Prva takšna metoda je neposredna primerjava testnih slik z učnimi, ne da bi testne slike kakorkoli spreminjali (*navadno ujemanje*). S to metodo lahko ugotovimo, kako se videz prizora spremeni, ko povečamo razdaljo med najbližjo učno in testno sliko. Druga metoda (*skaliranje*) iz vsake testne slike ustvari več povečanih ali pomanjšanih različic te slike, nato vsako od njih primerja z vsemi učnimi slikami, lokacijo v smeri  $x$  pa nam določi par z najmanjšo slikovno razdaljo. Takšen postopek je podoben običajnemu iskanju ujemanj na podlagi videza [3], le da z različnimi velikostmi izvirne slike skuša simulirati oddaljevanje ali približevanje slikam primerjanih pogledov.

Slika 5.16 prikazuje povprečje vseh relativnih napak pri lokalizaciji za parameter  $x_v$ . Graf na sliki 5.16 (a) prikazuje rezultate za umetni prizor, tisti na sliki 5.16 (b) pa rezultate za nabor *novoletna jelka*. Metoda *navadno ujemanje* potrjuje hipotezo, da pri večji oddaljenosti v smeri  $y$  razlike v videzu med testnimi in učnimi slikami presežejo možnost njihove neposredne primerjave. Podobno kot v razdelku 5.2.3 tudi tukaj dobimo v povprečju zelo visoke napake, zaradi katerih *navadno ujemanje* v splošnem ni uporabno. Metoda *skaliranje* se je obnesla bistveno bolje in je v primeru nabora „novoletna jelka“ že blizu mejne relativne napake 1. Metoda je v osnovi poseben primer predstavljenе metode, ki ima en sam referenčni pogled. En sam referenčni pogled pa pomeni, da ne omogoča iskanja ujemanja za različne vrednosti parametra  $x_v$  in jo zato omejuje ločljivost učnih lokacij. Najboljše rezultate dosežemo s predlagano metodo (*naša metoda*).

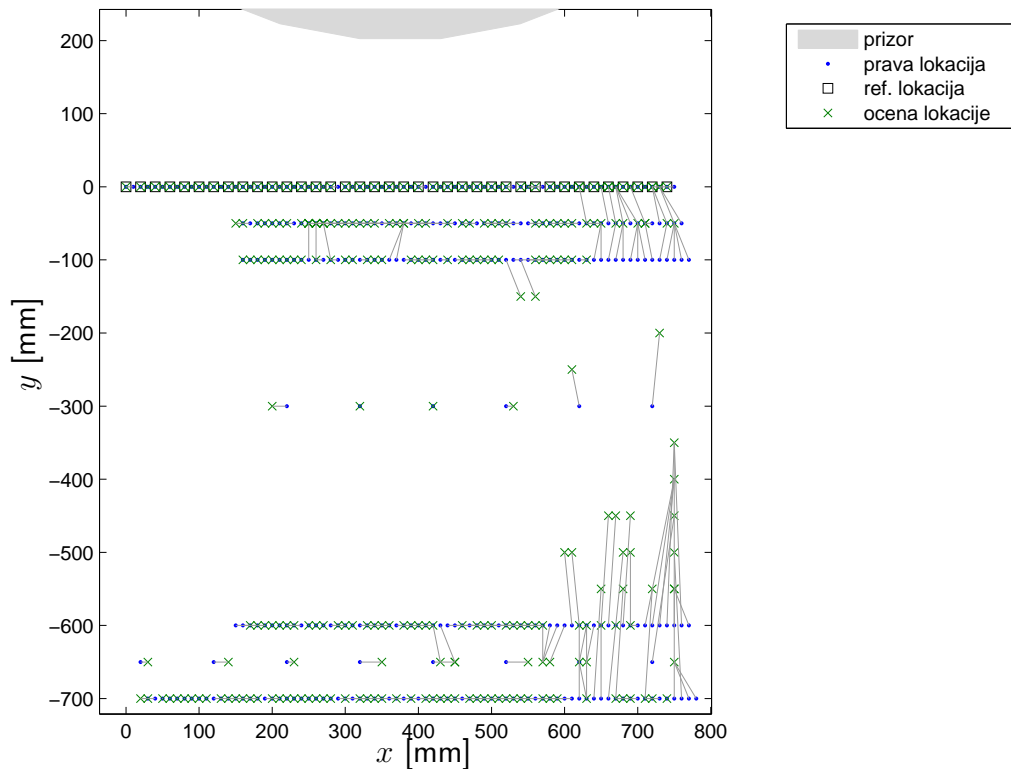


Slika 5.16: Relativne lokalizacijske napake pri (a) naboru umetnin slik in (b) nadzorovanem naboru resničnih slik s prizorom novoletne jelke.

### 5.4.3 Lokalizacija v dveh dimenzijah

Preskusi, opisani v predhodnem razdelku, prikazujejo prednost naše metode pri lokalizaciji v eni dimenziji pred metodami, ki so iz referenčnih slik, zajetih v eni dimenziji, sposobne samo takšne lokalizacije. Metoda, opisana v disertaciji, pa je zmožna isto predstavitev uporabiti za ocenjevanje koordinat neznane lokacije tudi v ravnini. V nadaljevanju bomo zato ocenili uspešnost delovanja metode v dveh dimenzijah.

Pri predhodnem preskusu smo pri naboru „novoletna jelka“ za vsako testno lokacijo dobili oceno tako parametra  $x_v$  (koordinata  $x$ ) kot parametra  $d_r$  (koordinata  $y$ ). Slika 5.17 prikazuje te rezultate pri vsaki testni lokaciji. Na sliki vidimo shemo prostora z modrimi točkami, ki označujejo testne lokacije, črni kvadratici pa ustrezajo učnim lokacijam v vrstici  $y = 0$  mm. Od vsake modre točke vodi daljica do križca, ki označuje ugotovljeno lokacijo za testno sliko, ki je nastala v lokaciji, označeni z modro točko. Povprečna lokalizacijska napaka, ki jo je naredila naša metoda, je 16,9 mm. Slika 5.18 prikazuje histogram lokalizacijske napake. Vidimo, da smo pri lokalizaciji deleža 73% vseh testnih slik dobili napako, manjšo od 25 mm, v 20% primerov pa je napaka velika 25–75 mm. Na območju, velikem  $780 \times 700$  mm, je to dobra ocena neznanih lokacij, zlasti glede na to, da teh lokacij robot prej še ni obiskal, do ocene pa pridemo brez izračunov 3D rekonstrukcije. V 7% testnih lokacij smo dobili oceno, ki je od prave lokacije oddaljena 75–250 mm. Na sliki 5.17 vidimo, da k tej napaki prispeva ocena koordinate  $y$  testnih pogledov v desnem spodnjem



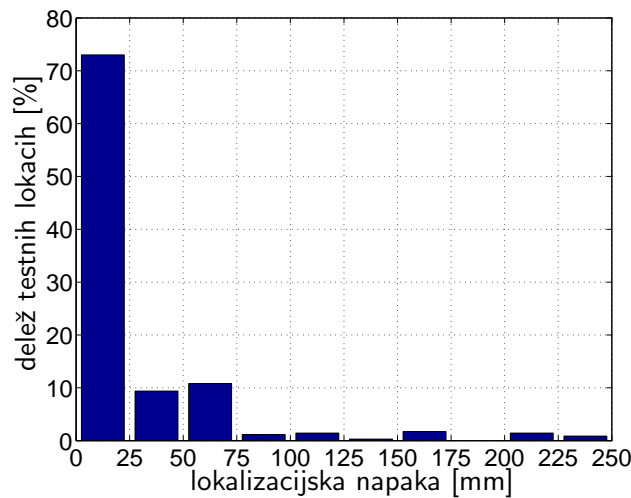
Slika 5.17: Shema lokacij in njihove lokalizacije za nabor nadzorovanega resničnega prizora. Učne slike so pri  $y = 0$  mm.

kotu diagrama. Razlog za napake je, da so te lokacije iz območja  $D_2$ , kjer je na navidezni sliki desni del slike nedefiniran (sliki 3.11 in 3.12 (h)).

#### 5.4.4 Vplivi elementov na različnih globinah prizora na lokalizacijo

V praksi ne moremo pričakovati, da bodo vsi prizori imeli konfiguracijo, ki bi popolnoma ustrezala geometrijski poenostavitvi izrisovanja navidezni sliki z ravnino na razdalji  $d_e$  od lokacij referenčnih pogledov. Za proučevanje vplivov elementov na različnih globinah prizora bomo uporabili umetni nabor, ki prikazuje okolje, podobno pisarni (slika 5.2 in slika 5.3).

V naslednji skupini preskusov bomo ugotovili, kako razgibana struktura prizora vpliva na lokalizacijo, kakšne so posledice predmetov, ki se znajdejo med referenčnimi in navidezno sliko ter proučili morebitno izboljšanje delovanja ob hipotetični zmožnosti odkrivanja motečih elementov na referenčnih slikah.

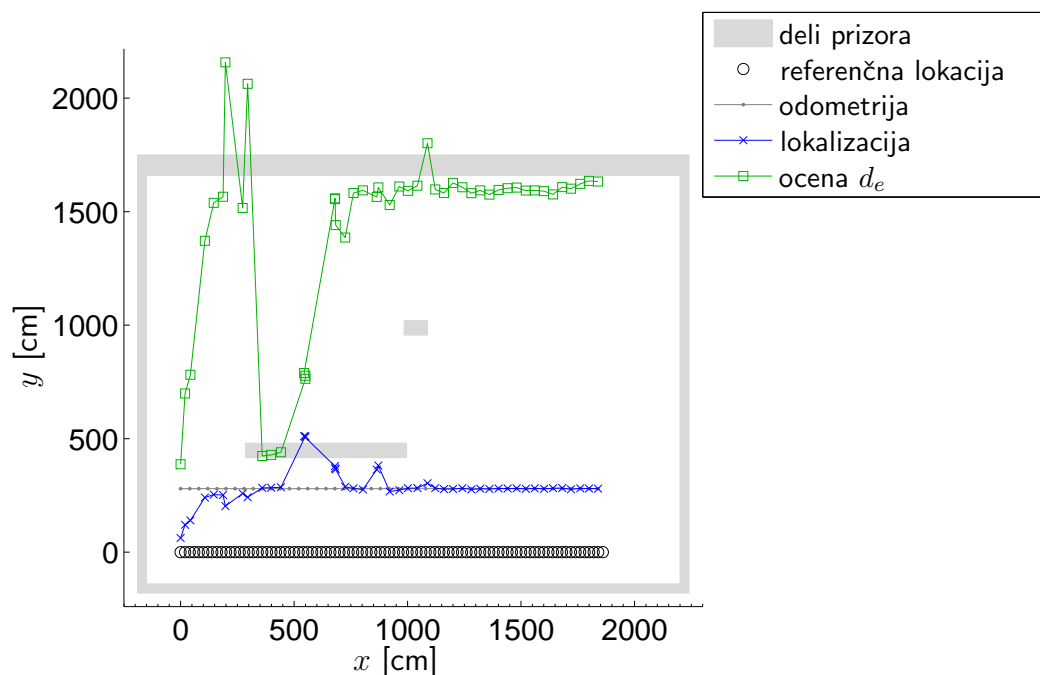


Slika 5.18: Histogram lokalizacijske napake pri lokalizaciji nabora nadzorovanega resničnega prizora s slike 5.17.

### Struktura prostora

Minimizacija slikovne razdalje med navidezno in podano sliko po parametrih  $x_v$ ,  $d_e$  in  $d_r$  nam vrne oceno lokacije neznanega pogleda s koordinatama  $x_v$  in  $d_r$  ter normalizacijsko razdaljo  $d_e$  do elementa prizora. Slika 5.19 prikazuje rezultate lokalizacije, pri čemer smo uporabili referenčne lokacije pri  $y = 0$  mm in testne lokacije pri  $y = 240$  mm. Krivulja lokalizacija prikazuje ugotovljeni koordinati  $x_v$  in  $d_r$  testnih slik, krivulja ocena  $d_e$  pa dobljene vrednosti parametra  $d_e$  pri pripadajočem  $x_v$ . Testni nabor lokacij smo izbrali tako, da je korak v smeri  $y$  velik s stališča vidnih sprememb elementov, hkrati pa je dovolj majhen, da med referenčnimi in testnimi lokacijami še ni nobenih elementov prizora.

Iz krivulj na sliki 5.19 se precej dobro vidi struktura prostora. Ocena parametrov  $d_e$  in  $d_r$ , ki se nanašata na oddaljenost v smeri  $y$ , je pri prvih treh testnih lokacijah pomaknjena zaradi stene prostora na levi strani. Sledi skupina lokacij (do približno  $x = 800$  cm), pri katerih minimizacija skuša prilagoditi parametre glede na elemente z miz, ki so v neposredni bližini testnih pogledov, in elemente na oddaljeni steni. Na območju, kjer je vrednost  $x$  okrog 400 cm, so prevladali bližnji elementi z mize, na kar nakazuje razlika med oceno  $d_e$  in  $d_r$ . Pri točkah v okolici  $x = 500$  cm je prilagoditev manj uspešna, saj so se parametri prilagodili tako, da bližnji elementi ne zakrivajo dominantnejših oddaljenih elementov. To sicer poslabša točnost lokalizacije, vendar v prikazanem primeru rezultati ostanejo v bližini pravih vrednosti.

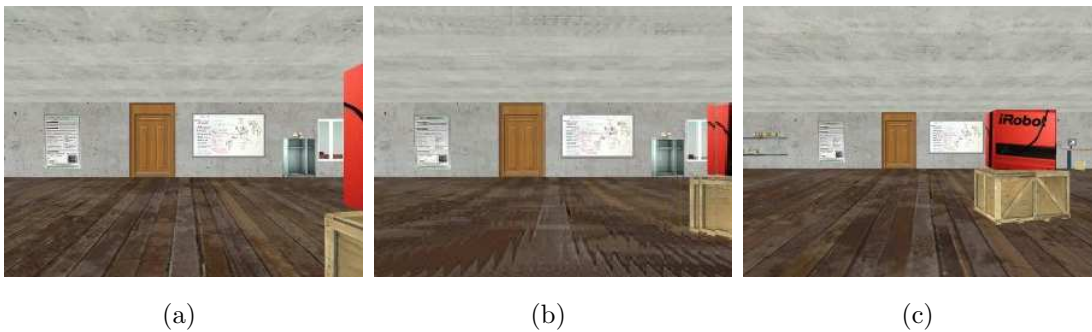


Slika 5.19: Lokalizacija v umetnem prostoru, podobnem pisarni z elementi, ki niso enako oddaljeni od referenčnih lokacij.

Preostanek krivulje se lepo prilega dejanskim meritvam. Na območju, kjer med referenčnimi lokacijami in oddaljeno steno ni nobenih predmetov, je to sicer pričakovano. Tudi škatle na sredini prostora ne povzročijo večjih odklonov pri lokalizaciji od izmerjenih lokacij, kar kaže na njihov manjši vpliv v primerjavi z ostalimi elementi prizora.

### Elementi, ki zakrivajo dele prizora

Pri spreminjanju lokacije navideznega pogleda lahko elementi prizora na različne načine zakrivajo druge dele prizora. Običajno so zakrivajoči elementi na isti strani tako referenčnih kot navideznega pogleda. V predhodnem razdelku smo videli, da jih večino metoda uspešno razreši. Drug tovrsten primer prikazuje slika 5.20. Slika 5.20 (a) predstavlja testni pogled, slika 5.20 (c) pa njemu najbližji referenčni pogled. Na desni strani slike 5.20 (a) vidimo omaro in okno, elementa, ki ju na sliki 5.20 (c) zakriva škatla, postavljena na sredino prostora. Lokalizacija za sliko 5.20 (a) vrne navidezni pogled, kot ga prikazuje slika 5.20 (b). Dobljeni približek pravilno modelira spremembo v pogledu, ki na desni strani vsebuje tudi ne-



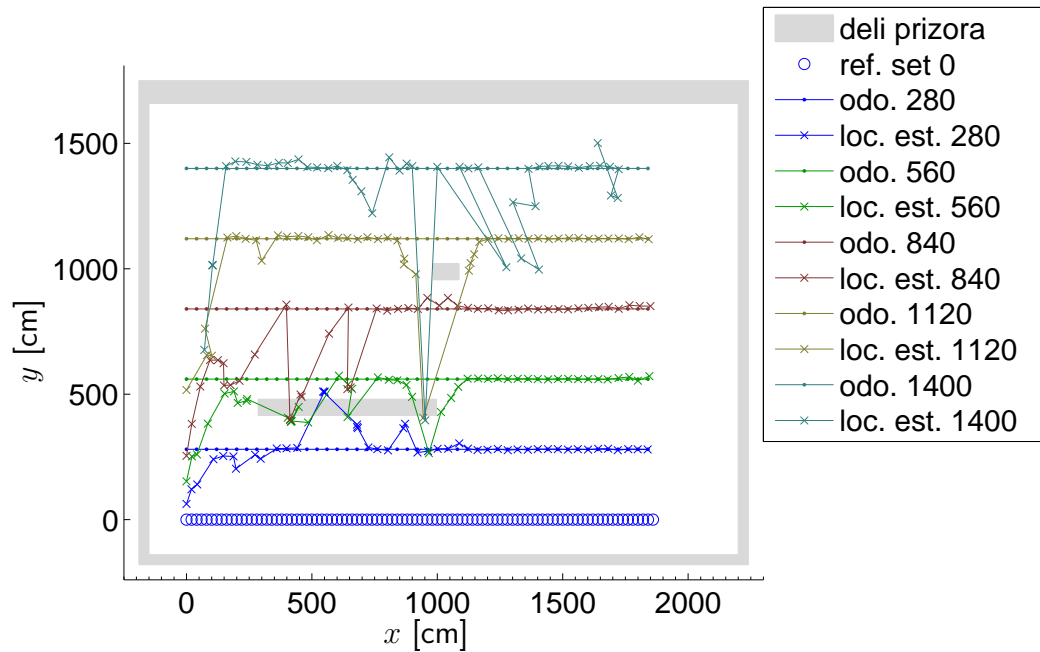
Slika 5.20: Uspešna lokalizacija kljub delnem zakrivanju prizora: (a) testna slika, (b) najuspešnejša navidezna slika, (c) najbližji referenčni pogled.

zakrit pogled na okno in omaro.

Drugačne vrste zakrivanja dobimo, ko razdalja od referenčne do navidezne lokacije preseže razdaljo do najbližjega elementa prizora. V tem primeru se lahko zgodi, da bomo poskušali izrisati elemente, ki so pred referenčnimi pogledi zakriti. Žarki, ki med referenčno in navidezno lokacijo sekajo nek predmet, bodo na navidezni sliki prispevali dele tega predmeta. To običajno povzroči vizualno nepričakovane učinke, ki ustrezno vplivajo na razpoznavo iz videza.

Za ugotavljanje vplivov opisanih zakrivanj smo prej opisani preskus razširili tako, da smo tudi preostale zajete slike vključili v testne nabore slik. Referenčni nabor je ostal enak kot prej, torej slike pri  $y = 0$  cm. Slika 5.21 prikazuje rezultate tega preskusa za lokalizirane vrednosti parametrov  $x_v$  in  $d_r$  testnih slik. Podobno kot prej krivulja loc. est.  $y$  predstavlja lokalizacijo slik pogledov, katerih izmerjene lokacije predstavlja *odo. y*, pri čemer je  $y \in \{280 \text{ cm}, 560 \text{ cm}, 840 \text{ cm}, 1120 \text{ cm}, 1400 \text{ cm}\}$ .

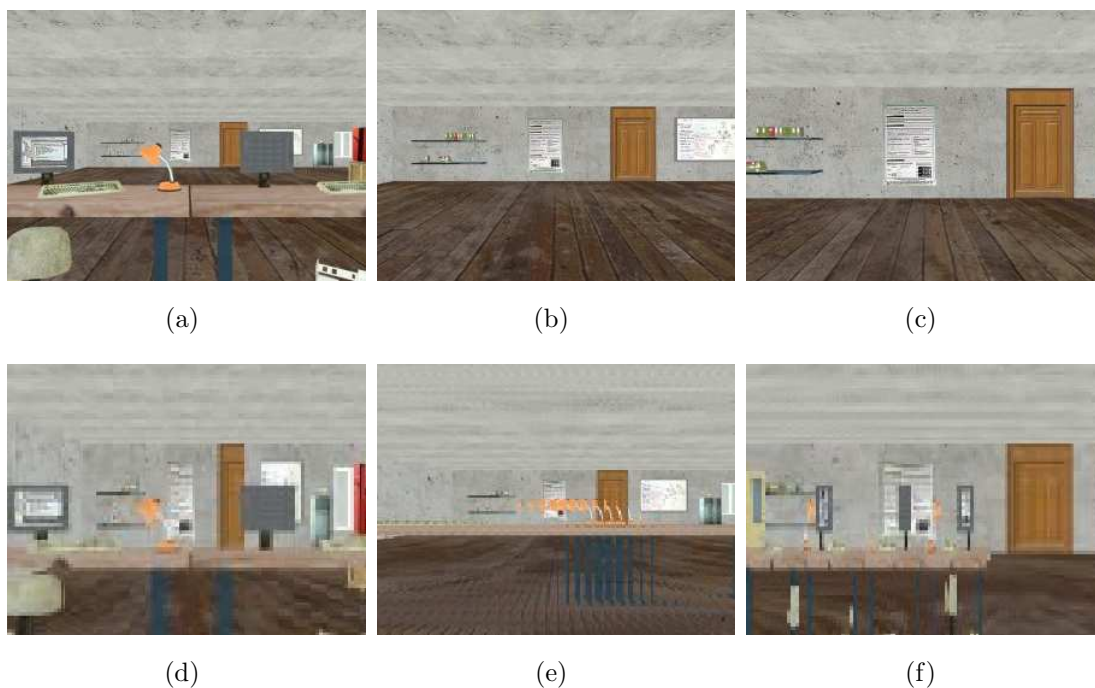
Iz rezultatov lahko analiziramo vpliv miz v levem sprednjem delu prostora. Že prej smo ugotovili, da je neposredno pred temi elementi ( $y = 280$  cm) lokalizacija manj natančna, vendar še vedno relativno dobra. Na sliki 5.22 (a) vidimo primer precej zapletenega prizora, na sliki 5.22 (d) pa navidezni pogled, ki ga je sistem določil za najbolj podobnega tistemu s slike 5.22 (a). Vrednosti parametrov tega pogleda so prilagojeni predmetom na mizi, kot to prikazuje slika 5.21. V navidezni lokaciji na drugi strani miz ( $y \in \{560 \text{ cm}, 840 \text{ cm}\}$  s slike 5.21) je vpliv zakritih delov tako velik, da je tam lokalizacija loc. est. 560 in loc. est. 840 povsem nezanesljiva. Na sliki 5.22 (e) je primer navidezne slike, ki je zaradi prekritih delov glede na mero slikovne razdalje bolj podobna sliki 5.22 (b) kot navidezna slika v dejanski lokaciji testnega pogleda. Kljub temu pri bolj oddaljenih navidezni lokacijah



Slika 5.21: Lokalizacija iz slik umetnega prostora, podobnega pisarni, referenčni nabor  $y = 0$  cm.

( $y \in \{1120 \text{ cm}, 1400 \text{ cm}\}$ ) vpliva zakrivanj iz podatkov o lokalizaciji *ne zaznamo več*. Na sliki 5.21 to razberemo iz krivulj loc. est. 1120 in loc. est. 1400, ki se prilegata vsaka svoji krivulji odometrije, torej odo. 1120 in odo. 1400. Vpliv nekaterih elementov prizora, ki se nahajajo med referenčnimi in testnimi lokacijami, se torej z oddaljenostjo zmanjšuje, ker se na navideznih slikah pojavljajo v vedno manjši meri. Primer takšnega navideznega pogleda je na sliki 5.22 (f), ki smo ga dobili za testno sliko 5.22 (c).

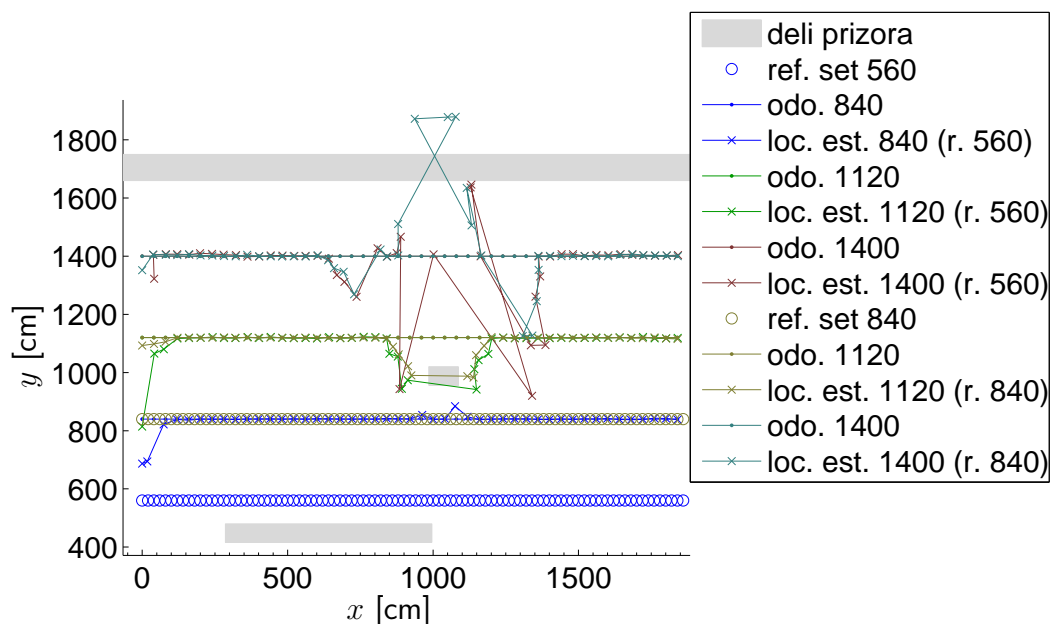
Škatle na sredini prostora predstavljajo drugačno vrsto elementov, ki delno zakrivajo prizor. Njihov vpliv se zelo jasno vidi na krivuljah lokalizacije testnih slik pri  $y \in \{1120 \text{ cm}, 1400 \text{ cm}\}$ . Padajoče vrednosti  $d_r$  na levi in naraščajoče vrednosti na desni strani ovire nakazujejo na to, da minimizacija skuša izločiti moteče elemente, ki se pojavijo na navideznih slikah, na vmesnem območju pa ni mogoče iz danega referenčnega nabora sestaviti takšnega navideznega pogleda, ki bi ustrezal pravim pogledom brez zakrivajočih elementov. Ti elementi so moteči, ker povečajo slikovno razdaljo, ko primerjamo resnični sliko z navidezno sliko iz pravega pogleda, zato sistem izbere pogled z nižjo slikovno razdaljo, ki pa je običajno napačen ali vsaj netočen.



Slika 5.22: Prikaz vpliva zakritih delov prizora na lokalizacijo: (a)–(c) testne slike, (d)–(f) navidezni pogledi v dobljeni oceni lokacije za pripadajoče testne slike.

Naslednjo skupino rezultatov prikazuje slika 5.23. Navidezne slike za lokalizacijo smo sestavili iz dveh referenčnih naborov, ki ne vsebujeta elementov miz (vendar oba še vedno prikazujeta škatle). Krivulja loc. est.  $y$  ( $r. y_r$ ) predstavlja rezultate testnih slik lokacij  $odo. y$ , pri katerih smo uporabili referenčni nabor, prikazan s točkami ref. set  $y_r$ . Iz slike lahko ugotovimo, da dobimo enakovredno zanesljive rezultate za loc. est. 1120 ( $r. 560$ ) in loc. est. 1120 ( $r. 840$ ), primerljiva pa sta tudi rezultata loc. est. 1400 ( $r. 560$ ) in loc. est. 1400 ( $r. 840$ ). Manjše odstopanje je le na levi strani, kjer pri referenčnem naboru, ki je bolj oddaljen, leva stena močnejše vpliva na oceno parametrov v smeri  $y$ . Iz tega lahko sklepamo, da pri lokalizaciji v tem delu prostora zadošča referenčni nabor pri  $y = 560$  cm. Po eni strani ta nabor pokrije večji del prostora, hkrati pa se pri bolj oddaljenih lokacijah obnese enakovredno. V splošnem bi lahko primerjali odziv lokalizacije istega nabora testnih slik z dvema referenčnima naboroma, nato pa enega od njih izločili, če bi v obeh primerih dobili podoben rezultat.

Opazimo lahko tudi, da se črka „V“, ki jo opisujejo krivulje s slike 5.21 zaradi škatel, pokriva z obliko krivulj istega območja na sliki 5.23, pri čemer je kót, ki ga na tem delu opiše krivulja, približno enak zornemu kotu pogledov. Ta oblika je



Slika 5.23: Lokalizacija iz slik umetnega prostora, podobnega pisarni, referenčna nabora  $y \in \{560 \text{ cm}, 840 \text{ cm}\}$ .

posledica izvora in smeri žarkov, ki sekajo predmet na sredini prostora, postopek lokalizacije pa se jim poskuša izogniti, ker njihovo vključevanje v hipotezo poveča slikovno razdaljo.

Rezultati opisanih preskusov so pokazali, da moremo predstavljeno metodo uporabiti tudi v okolju, ki ni ploske, ravninske oblike, ampak vsebuje predmete, razpostavljene po okolju. Lahko se sicer zgodi, da zaradi zakrivanja ključnih delov prizora lokalizacija poskuša zakrivajoče elemente zaobiti in s tem vrne manj točne ali zanesljive rezultate. Vendar pa nekateri zakrivajoči elementi „izzvenijo“ z globino lokacij pogledov, kar omogoča pokrivanje okolja tudi na tistih območjih, ki niso nujno v čistem vidnem območju najbližjih referenčnih pogledov. Problematična področja pa vedno lahko pokrijemo z dodatnim referenčnim naborom slik.

### Maskiranje predmetov, ki zakrivajo pogled na prizor

Vpliv motečih elementov na referenčnih slikah lahko torej odpravimo tako, da zajamemo referenčni nabor slik, ki motečih predmetov ne vključuje. Vedno pa ne želimo dodatno obremeniti pomnilnika in uporabljati več referenčnih naborov. Zato bomo v nadaljevanju poskusili ugotoviti, ali nam pri tem lahko pomaga poznavanje

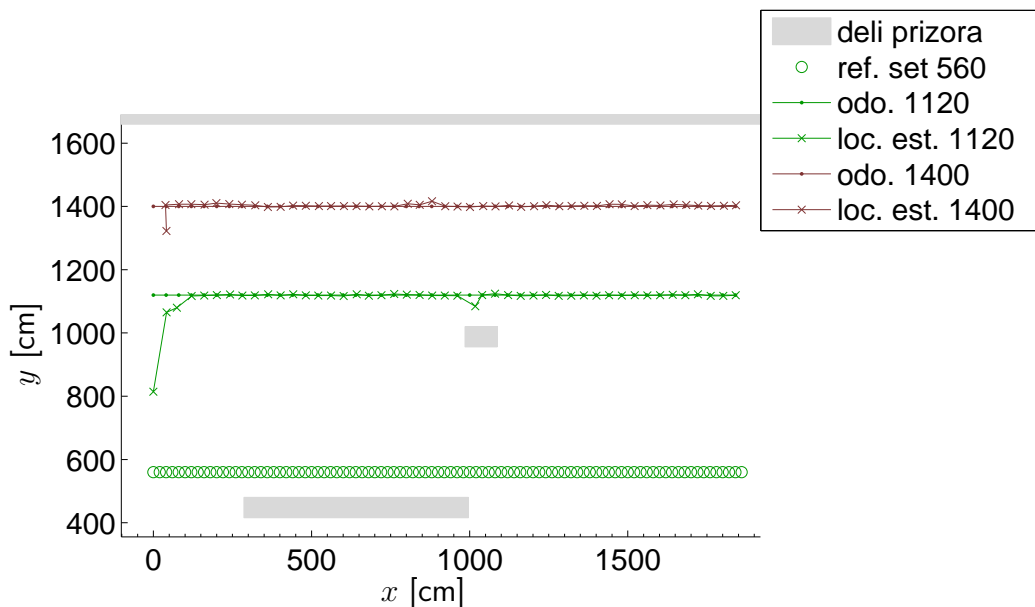
področij na referenčnih slikah, ki vsebujejo moteče elemente.

Elemente, ki na referenčnih slikah prikazujejo moteče predmete, označimo na binarni maski. Ta maska potem določa, katere slikovne elemente uporabimo pri primerjavi dveh slik in katere iz primerjave izključimo. Mera (4.3) za primerjavo para slik že v osnovi uporablja masko za izbor uporabljenih slikovnih elementov, ki jo moramo torej samo še dopolniti z označevanjem motečih elementov.

Za iskanje teh elementov bi lahko uporabili analizo prostorsko-časovnih prerezov referenčnega volumna [63] ali sledili lokalnim značilnicam v omenjenem volumnu. Omenjena postopka bi maske zgradila vnaprej, v okviru učne faze. Med lokalizacijo samo pa bi lahko masko določali iterativno z naključno podmnožico slikovnih elementov, izvedli primerjavo ter ocenili uspešnosti njenega ujemanja.

Konkretnega postopka na tem mestu nismo razvili, ampak smo ročno označili slikovne elemente referenčnih pogledov, ki vsebujejo potencialno problematične predmete. Referenčni pogledi, ki smo jih za preskus izbrali in tako maskirali, so pri  $y = 560$  cm.

Rezultati preskusa na sliki 5.24 prikazujejo uspešnost lokalizacije z maskiranjem. Postopek oceni koordinate neznanih lokacij z zanemarljivim odstopanjem od izmerjenih koordinat tudi na območju na sredini, kjer brez maskiranja (slika 5.23) dobimo



Slika 5.24: Lokalizacija po maskiranju problematičnih elementov referenčnega nabora  $y = 560$  cm iz slike 5.23.

napačne lokacije. S tem preskusom smo pokazali, da je moč predstavljeno metodo uporabiti tudi na način, ki je odporen na zakrite dele prizora.

### 5.4.5 Lokalizacija z mobilnim robotom

Z naslednjimi poskusi želimo ugotoviti in pokazati delovanje opisane metode v resničnem okolju, z digitalno kamero na mobilnem robotu. Robot zna ocenjevati svoje relativno gibanje s pomočjo odometra, kar uporabimo za določanje lokacij referenčnih slik. Digitalno kamero smo umerili in s tem dobili oceno goriščne razdalje  $f$ , središče slike  $c_0$  in  $r_0$  ter posredno zorni kot kamere. Koordinatne osi ravnine  $x$ - $y$  smo pretvorili tako, da je vsota kvadratov razdalj referenčnih lokacij do osi  $x$  najmanjša, prva referenčna točka pa ima koordinato  $x = 0$  m. S tem smo dobili množico lokacij učnih pogledov, ki so najbližje premici  $y = 0$  m.

Izvedli smo dva preskusa. Za prvi preskus smo uporabili en prostorsko-časovni volumen referenčnih slik, pri drugem pa smo dodali še njemu ortogonalni prostorsko-časovni volumen za določanje koordinat testnih lokacij. S preskusi želimo ugotoviti odpornost lokalizacije na moteče elemente (nenatančnost umerjanja kamere, napake pri odometriji) in prikazati, da so medsebojni geometrijski odnosi ugotovljenih lokacij testnih slik skladni z njihovimi dejanskimi geometrijskimi odnosi.

#### Lokalizacija vzporedne poti

V tem preskusu bomo ugotovili delovanje metode pri lokalizaciji iz enega prostorsko-časovnega volumna referenčnih slik. V fazi raziskovanja smo najprej peljali robota po dokaj ravni poti, dolgi približno 3 m vzdolž stene laboratorija. Tako dobljene slike smo uporabili za referenčni prostorsko-časovni volumen. Slika 5.25 (a) prikazuje pogled, sestavljen iz referenčnih slik, iz katerega vidimo, da vsebuje večino elementov prostora.

Robota smo nato peljali po približno vzporedni poti, oddaljeni 3 m od predhodne poti. Na njej smo zajeli testne slike, primer ene od njih pa prikazuje slika 5.25 (b). Tudi za te slike smo shranili meritve odometrije, relativne glede na prvo testno sliko, za primerjavo z rezultati lokalizacije.

Pri lokalizaciji smo vsako testno sliko obravnavali neodvisno od lokalizacije in meritev ostalih testnih lokacij. To pomeni, da je bil robot „ugrabljen“ v vsaki testni lokaciji. Za hitrejše določanje ocene lokacije smo preiskovalni prostor razdelili na redko diskretno mrežo točk. Najbolj uspešna od njih je postala izhodišče za

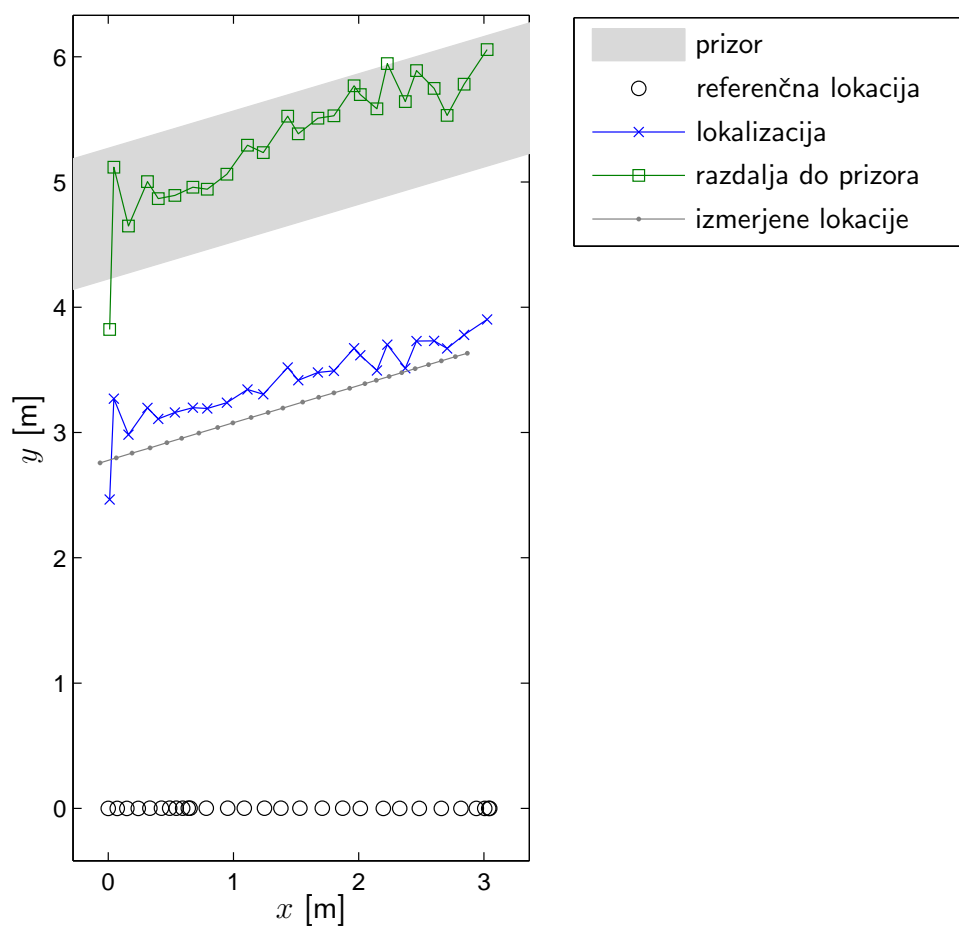


Slika 5.25: (a) Pogled, sestavljen iz referenčnih slik; (b) primer testne slike.

optimizacijo s simuliranim ohlajanjem [35]. Na koncu smo dobljeni rezultat izboljšali z determinističnim pokoordinatnim spustom do minimuma.

Slika 5.26 prikazuje rezultate lokalizacije za vse testne slike. Shema za vsako testno lokacijo prikazuje z lokalizacijo ugotovljene parametre glede na referenčne lokacije, ki so tudi sami na shemi. Koordinata  $x$  modrih točk označuje vrednost parametra  $x_v$  pripadajoče testne lokacije, koordinata  $y$  oznake pa vrednost parametra  $d_r$ . Vrednost tretjega parametra,  $d_e$ , prikazujejo zeleni kvadratici v točkah  $x = x_v$  in  $y = d_e$ . Sosedne oznake so povezane z daljicami, celotna krivulja pa prikazuje zaporedje zajemanja testnih lokacij. Shema prikazuje tudi točke iz odometrije testnih lokacij za primerjavo z rezultati lokalizacije. Vidimo lahko, da si ugotovljene vrednosti  $x_v$  testnih pogledov sledijo v takem vrstnem redu, kot smo jih izmerili z odometrijo. Ocena parametra  $d_r$  je manj natančna in kaže na eni strani na sistematično napako, ki je posledica netočnosti pri umerjanju kamere, ter dodatni šum, ki nastane zaradi manjše točnosti pri odometriji referenčnih pogledov. Slika 5.27 (a) prikazuje eno izmed testnih slik, uporabljenih pri lokalizaciji, postopek pa je vrnil lokacijo navideznega pogleda, iz katerega izriše sliko 5.27 (b).

Preskus je pokazal, da s predstavljeno metodo ne moremo dobiti natančne metrične ocene lokacije mobilnega robota. Vendar smo dobili oceno koordinat posamezne lokacije, ki je pred tem nismo obiskali, in pri oceni napako, enakovredno dvojni razdalji dveh zaporednih referenčnih lokacij. Lokalizacija zaporednih testnih pogledov je v skladu z zaporedjem izmerjenih lokacij kljub temu, da je izvajanje lokalizacije v vsaki testni lokaciji neodvisno od rezultatov lokalizacije v ostalih testnih lokacijah. To je dovolj, da za robota ugotovimo, kje se je znotraj prostora nahajal na začetku in kako se je po prostoru gibal. Na podlagi teh ugotovitev ocenjujemo,



Slika 5.26: Rezultati lokalizacije relativno glede na referenčne lokacije.



Slika 5.27: (a) Testna slika; (b) navidezni pogled, ki se ujema s testno sliko.

da je metoda, opisana v disertaciji, primerna za uporabo pri lokalizaciji mobilnega robota.

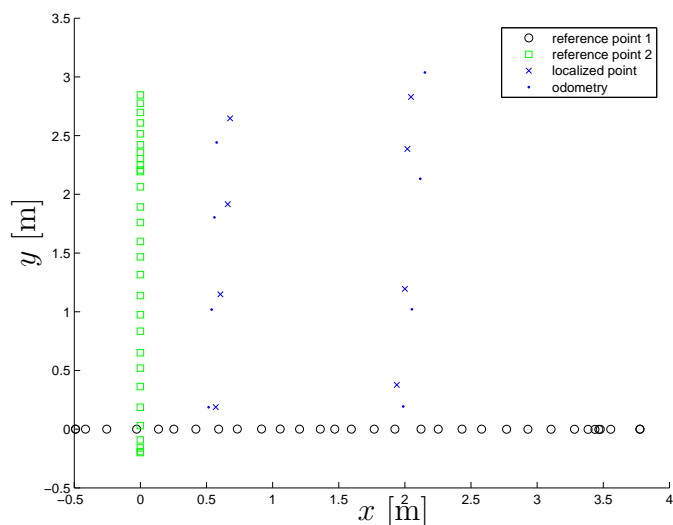
### Lokalizacija z ortogonalnimi referenčnimi nabori slik

Z dosedanjimi preskusi smo pokazali, da predstavljena metoda zmore lokalizirati poglede z neznane lokacije, ki se nahaja v isti ravnini kot referenčni pogledi in katerih smer pogleda je enaka kot smer referenčnih pogledov. Skladnost ugotovljenih parametrov  $x_v$  z izmerjenimi vrednostmi je zelo visoka, zato so lokalizacijske napake v tej smeri majhne. Večje napake dobimo pri oceni oddaljenosti v smeri  $y$ , kjer ocene parametra  $d_r$  lahko odstopajo od resnične vrednosti.

Če želimo višjo točnost lokalizacije v obeh razsežnostih, lahko uporabimo dva pravokotno usmerjena nabora referenčnih pogledov. Ortogonalnost teh referenčnih naborov ni samo v smereh vožnje in pogledov, ampak tudi s stališča vsebine in nabora žarkov. Vsakega od teh naborov uporabimo ločeno, nato pa za oceno koordinat neznane lokacije vzamemo vrednost parametra oddaljenosti vzdolž posameznega referenčnega nabora. Ker sta si takšna nabora slik ortogonalna, enodimenzionalno funkcionalnost lokalizacije posameznega referenčnega nabora združimo v dvodimenzionalno lokalizacijo.

Za ta preskus smo vzeli referenčni nabor pogledov, opisan v predhodnem preskusu, dodatno pa smo peljali robota po pravokotni poti in dobili ortogonalni nabor referenčnih slik. Nato smo zajeli poglede v nekaj testnih lokacijah. Ker se mora smer testnega pogleda ujemati s smerjo referenčnih pogledov, smo iz vsake lokacije zajeli po tri slike, pri čemer smo kamero po navpični osi vrteli s korakom  $90^\circ$ . Tako smo v vsaki testni lokaciji dobili pogled, usmerjen v levo, naravnost in v desno, dva od teh pogledov pa se pokrivata s po enim referenčnim naborom. V postopek lokalizacije iz vsakega od referenčnih naborov smo vključili vse zajete slike, izmed treh pa smo v vsaki testni lokaciji izbrali le tistega, ki je dosegel najnižjo slikovno razdaljo. Dobljeno vrednost parametra  $x_v$  iz prvega referenčnega nabora smo dodelili koordinati  $x$ , vrednost  $x_v$  drugega referenčnega nabora pa koordinati  $y$  testnih lokacij.

Na sliki 5.28 vidimo postavitev tega preskusa skupaj z rezultati lokalizacije. Prvi referenčni nabor slik (krogci) je usmerjen vzporedno z osjo  $y$  in v smeri naraščanja koordinate, drugi referenčni nabor (kvadratki) pa v smeri naraščajočih vrednosti  $x$ . Točke na sliki označujejo izmerjene koordinate iz odometrije robota. Modre oznake prikazujejo vrednosti koordinat, ki smo jih dobili pri lokalizaciji, relativno glede



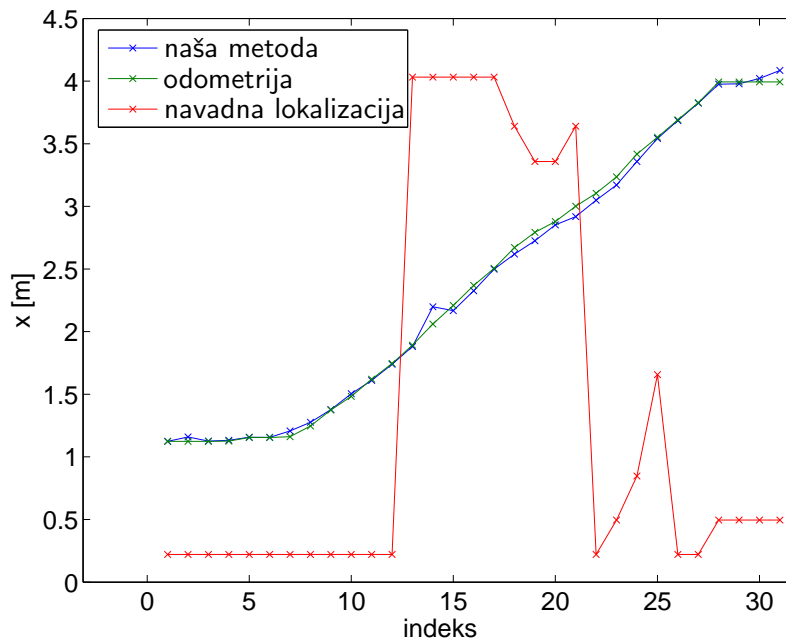
Slika 5.28: Lokalizacija z dvema pravokotnima naboroma referenčnih slik.

na referenčne lokacije. Ugotovljene lokacije so glede geometrijske razporeditve po ravnini skladne z izmerjenimi. V celoti se sicer ne pokrivajo, vendar je to posledica napak pri meritvah odometrije.

### Primerjava z navadnimi metodami na podlagi videza

V razdelku 5.2.3 smo ponazorili, kakšne rezultate dobimo z običajnim postopkom za lokalizacijo na podlagi videza, če so lokacije testnih pogledov preveč oddaljene od lokacij učnih pogledov. Videz testnih slik se je preveč razlikoval od videza učnih slik, zato smo dobili veliko lokalizacijsko napako.

Učne slike opisanega preskusa smo uporabili kot prostorsko-časovni volumen referenčnih slik za lokalizacijo pogledov testnih slik s pristopom, opisanim v disertaciji. Krivulja *naša metoda* na sliki 5.29 prikazuje rezultate te lokalizacije in s tem dopolni graf s slike 5.8. Krivulja se prilega krivulji *odometrija*, ki prikazuje prave lokacije testnih pogledov. S tem smo pokazali, da testne slike vsebujejo elemente prizora, vidne iz učnih slik, vendar običajne metode lokalizacije na podlagi videza sprememb v videzu niso sposobne modelirati.



Slika 5.29: Rezultati lokalizacije izven območja, pokritega z učnimi slikami, ki jih dobimo z običajnim postopkom lokalizacije na podlagi videza (krivulja navadna lokalizacija) in s postopkom, opisanim v disertaciji (krivulja naša metoda). Krivulja odometrija predstavlja točne vrednosti.

## 5.5 Povzetek poglavja

Za preskuse metode smo pripravili širok nabor slik in podatkov, ki smo jih dobili tako na visoko nadzorovan umetni način kot z zajemom resničnih slik s srednjo in majhno stopnjo nadzora nad postopkom. Za motivacijo smo ponazorili delovanje običajnih metod lokalizacije na podlagi videza in prikazali nekatere njihove slabosti. Numerično smo ocenili kakovost izrisovanja navideznih slik z navzkrižnimi režami. Preverili smo primernost navideznih pogledov za iskanje lokacije neznanega pogleda. Raziskali smo, kako različni deli prizora, ki delno zakrivajo pogled na okolje, vplivajo na lokalizacijo. Najpomembnejši del preskusov predstavlja lokalizacija mobilnega robota.

# Poglavje 6

## Zaključek

### 6.1 Obnova disertacije

Uporaba računalniškega vida za mobilne robote pridobiva na veljavi. V panogi robotike poznamo številne rešitve, ki so tako namenske in torej delujejo le v določenem vnaprej pripravljenem okolju, kot splošnejše rešitve. Slike, ki jih dobimo iz vizualnih senzorjev, nosijo bogat nabor podatkov. Možnosti za pridobivanje informacij iz njih je veliko, ustrezno številni pa so tudi postopki za obravnavo slikovnih podatkov. Poglavitni nalogi, ki jo rešujemo z računalniškim vidom na mobilnem robotu, sta, pridobiti opis okolja, ki ga robot razišče, ter uporabiti sliko trenutnega pogleda za umeščanje robota v okolje glede na pridobljeni opis.

Za obravnavo slikovnih podatkov, namenjeno predstavitev raziskovalnega okolja, sta se uveljavili dve vrsti postopkov, geometrijski postopki ter postopki, ki temeljijo na videzu. Geometrijski postopki rekonstruirajo model raziskanega okolja, pri lokalizaciji pa se opirajo na metrične razdalje med elementi in robotom. Dobljena predstavitev je podobna zemljevidom, njeno metričnost pa lahko izkoristimo neposredno pri lokalizaciji in navigaciji robota. Slabost geometrijskih pristopov je, da temeljijo na inverznih izračunih rekonstrukcije, ki so lahko slabo pogojeni. Njihova šibka točka je tudi odvisnost od detektorjev lokalnih značilnic, ki ne delujejo v vseh okoljih ali pogojih.

Postopki, ki temeljijo na videzu okolja, zgradijo predstavitev v obliki zaporedja vtisov iz obiskanih lokacij. Ti vtisi imajo lahko obliko množic lokalnih značilnic, pogosteje pa shranimo opise celotnih slik ali njihovih globalnih lastnosti v neki zgoščeni obliki. Postopki vsebine slik ne interpretirajo, zato se izognejo dodatnim izračunom, ki bi lahko vračali napačne rezultate, hkrati pa zagotavljajo delovanje v praktično

poljubnem okolju. Strukturo preiskanega okolja shranimo v obliki dodatnih podatkov o koordinatah in orientaciji lokacij, v katerih so zajeti vtisi. Lokalizacija potem poteka tako, da izmed shranjenih vtisov poiščemo tistega, ki se najbolj ujema s trenutnim vtisom, dodatni podatki najdenega vtisa pa potem podajo oceno trenutne lokacije. Zmožnost lokalizacije je torej omejena na ožje območje lokacij, obiskanih med raziskovanjem. Iz tega sledi, da za dobro predstavitev potrebujemo obsežno in relativno gosto pokrito okolje. To na eni strani zahteva veliko prostora v pomnilniku za hranjenje predstavitev, ki ga je moč nadzorovati z uporabo tehnik zgoščevanja podatkov, kot je metoda glavnih komponent. Po drugi strani pa moramo v raziskovanje vložiti čas in skrbno raziskati okolje, kar je lahko v praksi problematično. V nadaljevanju smo se zato osredotočili na reševanje problema, kako za lokalizacijo mobilnega robota uporabiti predstavitev, ki temelji na videzu, vendar predstavlja le majhen delež predstavitev, potrebne za pokrivanje okolja.

Če robota vodimo strogo po premici in v znanih intervalih zajemamo slike pogledov, dobimo zaporedje slik, ki je poseben primer predstavitev okolja na podlagi videza in ki omogoča lokalizacijo v eni dimenziji na isti premici. To velja tudi, če je kamera usmerjena pravokotno na premico. V slednjem primeru lahko dobljeno predstavitev obravnavamo tudi kot prostorsko-časovni volumen slik, ta pa predstavlja referenčno osnovo za izrisovanje novih pogledov na okolje. Postopek, imenovan izrisovanje z navzkrižnimi režami, slikovne elemente slik iz referenčnih pogledov zloži v približek slike navideznega pogleda iz lokacije, ki je med raziskovanjem nismo obiskali. Pri tem izkorišča geometrijo nastanka slike in način, kako se svetlobni žarki preslikujejo na slikovno ravnino. Žarki, ki bi prispevali k elementom navidezne slike pogleda v podani navidezni lokaciji, določajo referenčne slike, ki intenzitete teh žarkov hranijo. Pri slikah iz običajne perspektivne kamere postopek uporablja normalizacijsko ravnino, ki omogoča reprojekcijo žarkov ter iskanje približkov za tiste žarke, ki jih zaradi diskretnosti nabora referenčnih slik nimamo shranjenih v predstavitvi.

Predstavitev, ki bi jo lahko uporabili za lokalizacijo na podlagi videza v eni dimenziji, torej lahko uporabimo za sestavljanje približkov slik pogledov iz lokacij, izhajajočih iz zvezne množice, ki opišejo ravnino. To nam občutno razširi možnosti uporabe predstavitev za lokalizacijo na podlagi videza. Vendar je omejitev, da morajo biti referenčni pogledi kolinearni, v praksi težko doseči, zato jo omilimo in omogočimo delovanje izrisovanja tudi, če referenčni pogledi od premice odstopajo in so usmerjeni pod različnimi koti. To dosežemo tako, da slikovne elemente referenčne slike preslikamo na slikovno ravnino, ki je obrnjena za negativni kot odklona od

pravega kota, izraze, ki določajo sestavljanje navideznega pogleda, pa dopolnimo z dodatno spremenljivko oddaljenosti referenčne lokacije od premice.

Zaporedje, ki predstavlja prostorsko-časovni volumen referenčnih slik, lahko sedaj zajame tudi mobilni robot, ki se ne giblje strogo po premici. Iz dobljene predstavitve lahko izrišemo približke poljubnih pogledov v ravnini, te pa nato uporabimo v ustreznem postopku lokalizacije na podlagi videza. V primerjavi z običajnimi pristopi lokalizacije na podlagi videza sedaj v predstavitvi potrebujemo le majhen delež podatkov. Vseeno pa bomo za velika okolja potrebovali sorazmerno veliko predstavitev v pomnilniku. Ko želimo nadzorovati to velikost, lahko uporabimo postopek za zgoščevanje predstavitev s postopno izgradnjo podprostora metode glavnih komponent. Dobljeni podprostor ima to lastnost, da ob zmanjšanju porabe pomnilnika ohranja informacije v največji možni meri. Poleg tega lahko rekonstruiramo slike po delih, kar nam omogoča neposredno sestavljanje navideznih slik iz zgoščene predstavitve.

Ko imamo predstavitev zgrajeno, lahko izvajamo lokalizacijo. To poteka tako, da robot v trenutni neznani lokaciji zajame pogled in dobljeno sliko poda postopku za lokalizacijo. Postopek iz predstavitve v lokaciji hipotezi izriše navidezno sliko, ki jo primerja s podano sliko. Primerjavo opravi z mero vsote kvadratov razlik, ki je prilagojena temu, da navidezna slika nima vedno definiranih vseh slikovnih elementov. Sledi primerjanje podane slike s slikami iz nadaljnje množice hipotez lokacij, dokler se postopek ne ustavi in vrne tisto hipotezo, ki da najboljšo oceno. Koordinate te hipoteze so rezultat postopka lokalizacije.

Opisano lokalizacijo lahko teoretično izvajamo najuspešneje v dveh trikotnih območjih pred in za premico referenčne poti. Na teh območjih namreč lahko izrišemo sliko po vsej širini navidezne slike. Izven teh območij ta sposobnost postopoma pada, s tem pa tudi pričakovana zmožnost lokalizacije. Pri oddaljevanju od premice poti moramo upoštevati tudi zmanjševanje pokritosti slike oz. povečevanje dela referenčnih slik. Ne glede na to opisani postopek ob manjši porabi pomnilnika in časa raziskovanja omogoča lokalizacijo mobilnega robota na večjem območju kot običajni postopki lokalizacije na podlagi videza.

Rezultati preskusov so pokazali, da izrisovanje z navzkrižnimi režami vrača navidezne slike, ki so primerne za ugotavljanje podobnosti z resničnimi slikami. Parametra, ki določata lokacijo navideznega pogleda, pravilno določata videz izrisane slike, pravilnost pa je odvisna tudi od parametra oddaljenosti do normalizacijske ravnine. Kljub predpostavki, da okolje modelira ravnina, dobimo dobre približke tudi v okolju, ki ima elemente na različnih globinah prizora. Izrisovanje približkov

slik iz poljubnega pogleda v zveznem prostoru torej lahko nadomesti pregledovanje pravih slik iz diskretnega in omejenega nabora učnih slik. Rezultati preskusov lokalizacije so pokazali, da je izbira postopka ustrezna in da uspešno določa oceno lokacij iz slik pogledov prej neobiskanih lokacij. Čeprav je prostorsko-časovni volumen referenčnih slik enodimenzionalna predstavitev, pa same slike skupaj z geometrijo izrisovanja navideznih slik modelirajo tudi globino prizora ter oddaljenost od referenčnih lokacij. Postopek je uspešno lokaliziral mobilnega robota tudi v lokacijah, kjer običajen postopek lokalizacije na podlagi videza odpove.

## 6.2 Prispevki k znanosti

Z delom, opisanim v disertaciji, smo k znanosti prispevali naslednje:

- Razvili smo postopek za lokalizacijo mobilnega robota, ki temelji na pristopu učenja in prepoznavanja okolja **na podlagi videza**, hkrati pa sloni na **prvinah geometrijskih pristopov**. Postopek ima s pristopi učenja in razpoznavanja na podlagi videza skupno obliko predstavitve, ki hrani zaporedje vizualnih vtisov, in način razpoznave s primerjanjem trenutnega vtisa z znanimi vtisi. Razlika v predstavitvi pri naši metodi je, da **ne hrani eksplicitno videza vseh pogledov v prostoru**, za katere pričakujemo, da jih bomo med lokalizacijo in navigacijo potrebovali, ampak hranimo **le osnovo**, iz katere lahko s **pomočjo geometrije dobimo približek poljubnega pogleda**, če ta vsebuje elemente, vidne iz shranjenih pogledov.
- Geometrijski del našega postopka skrbi za to, da iz predstavitve zgradi slike iz pogledov, ki jih med raziskovanjem nismo obiskali. To smo dosegli z uporabo izrisovanja iz slik, ki je metoda s področja **računalniške grafike**. Tako smo uporabo metode s področja **vizualizacije** uspešno **razširili tudi na področje lokalizacije mobilnih robotov**. Izbrani postopek izrisovanja iz slik ne zahteva izračunov oblike prostora in razdalj do posameznih elementov, kot velja pri ostalih geometrijskih pristopih, zato je **računsko nezahteven**.
- Uvedli smo **zgoščeno predstavitev** okolja, ki jo zgradimo z metodo glavnih komponent. Postopek izrisovanja navideznih slik smo priredili tako, da lahko neposredno uporabljata tudi takšno predstavitev. S tem smo omogočili **nadzor nad porabo pomnilnika** in s tem delovanje postopka v velikih okoljih.

Predlagane postopke smo utemeljili in njihovo delovanje pokazali na umetnih slikah, resničnih slikah pod nadzorovanimi pogoji ter **na mobilnem robotu**.

## 6.3 Razprava

V disertaciji smo predstavili in preskusili postopek, ki uspešno izvaja lokalizacijo na podlagi videza slike. Lokalizacija pri tem ni omejena na prej obiskane lokacije, ampak omogoča oceno lokacij iz poljubnega pogleda, ki prikazuje elemente na učnih slikah. Postopek uporablja predstavitev v obliki zaporedja slik, združenih v prostorsko-časovni volumen učnih slik, iz katere tvori nove, navidezne slike.

Postopek za pravilno delovanje določa omejitve, da morajo neznani pogledi, iz slike katerih želimo oceniti lokacijo, biti usmerjeni vzporedno z referenčnimi pogledi v predstavitvi. Ta omejitev je posledica izvedbe metode s perspektivno kamero, ki ima svoj zorni kot omejen. Drugače usmerjena kamera tako zajame sliko, ki ima drugačno vsebino in videz. Pravilno usmerjenost kamere lahko zagotovimo s pomočjo vgrajenega kompasa. Če ta ni del opreme robota, lahko robot pravilno usmerjenost pogledov ugotovi tako, da zajame zaporedje delno prekrivajočih se pogledov med vrtenjem okrog lastne osi. Za vsako sliko bi potem dobili neko oceno lokacije ter slikovno razdaljo v minimumu preiskovalnega prostora. Slika, pri kateri je ta vrednost najnižja, nato predstavlja smer pogleda, ki se izmed vseh pokriva najboljše s smerjo referenčnih pogledov.

Gradnja predstavitve s panoramsko kamero in uporaba ustreznega postopka za izrisovanje iz slik [4, 11] bi zmanjšala omejitve glede usmerjenosti robota med lokalizacijo. Neznano sliko bi morali zgolj poravnati tako, da bi se elementi na njej pokrili z elementi iz navideznih pogledov. S tem bi torej pridobili večjo fleksibilnost, vendar za ceno večjega števila parametrov izrisa navidezne slike.

Pri nadaljnjih raziskavah bo smiselno obravnavati, kako združiti več prostorsko-časovnih volumnov referenčnih slik v skupno predstavitev. Okolja z večjo kompleksnostjo bo tako nujno raziskati v več vzporednih prehodih, da bo predstavitev vsebovala vsa področja okolja, ki jih zaradi zakritih delov en sam prehod ne bi vseboval. Pri takšnem raziskovanju bomo verjetno shranili več prostorsko-časovnih volumnov referenčnih slik, kot bi jih potrebovali za uspešno delovanje. V času, ko robot ne raziskuje ali navigira po okolici, lahko sistem sam preveri, katere poglede v predstavitvi lahko dobro predstavi z ostalimi referenčnimi pogledi, in obdrži le tiste prostorsko-časovne volumne, ki jim referenčnih slik ni mogoče predstaviti z ostalimi volumni.

Pri opisanih preskusih smo lokalizacijo vsakega testnega pogleda izvajali neodvisno od lokalizacije ostalih testnih pogledov, čeprav je njihovo zaporedje določalo gibanje robota. Vodenje robota in njegovo gibanje zmanjša preiskovalni prostor za lokalizacijo po premiku iz znanje lokacije. Ker imajo hipoteze za lokacijo robota obliko točk v preiskovalnem prostoru, bi gibanju robota lahko sledili s pomočjo Kalmanovega filtra [22] ali s filtriranjem delcev (ang. *particle filtering*) [10].

Problematiko zakritih delov prizora smo v disertaciji že obravnavali. V zapletenejših okoljih bi njihovo odkrivanje pripomoglo k zanesljivejši lokalizaciji. Uporabljena predstavitev je primerna za analizo in odkrivanje plasti na različnih globinah prizora [64], kar lahko izvajamo med fazo učenja. Iz dobljenih podatkov bi lahko sklepali, kateri elementi so med referenčnimi pogledi in podanim učnim pogledom, ter morebitne moteče elemente izločili iz navidezne slike. Drugače bi morali obravnavati elemente, ki se na resničnih slikah pojavijo nanovo v obliki dinamičnih elementov. Te bi morali odkrivati in izločati med samo lokalizacijo. Primerna oblika izločanja bi bila, da med preverjanjem hipotez namesto celotne razpoložljive navidezne slike za primerjanje uporabimo le naključno izbrano podmnožico obstoječih slikovnih elementov. Vsako hipotezo bi torej morali oceniti večkrat, ob različnem izboru slikovnih elementov, ter obdržati tisto, ki prispeva najnižjo slikovno razdaljo.

Naši dosedanji preskusi so torej pokazali uspešnost delovanja predlaganega pristopa. V tej razpravi smo ugotovili možnosti njegove uporabe ter nadaljnjega razvoja.

# Dodatek A

## Notacija

Pri zapisu matematičnih izrazov v disertaciji uporabljamo naslednja pravila zapisa:

- Skalarje pišemo v ležečem zapisu (npr.  $i$ ,  $n$ ,  $R$ ,  $C$ ).
- Vektorje označujemo z malimi tiskanimi črkami v krepki pisavi (npr.  $\mathbf{a}$ ).
- Elemente vektorja  $\mathbf{a}$  dolžine  $R$  označujemo kot skalar  $a_r$  z indeksom  $r$ , ki predstavlja zaporedno številko elementa, da velja,

$$\mathbf{a} = \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_{R-1} \end{bmatrix} = [a_0, a_1, \dots, a_{R-1}]^\top. \quad (\text{A.1})$$

- Vektor, ki ima v zaporedju  $n$  vektorjev zaporedno številko  $i$ , označujemo z indeksom, npr.  $\mathbf{a}_i$ .
- Element vektorja  $\mathbf{a}_i$  dolžine  $R$  z zaporedno številko  $r$  označujemo z dvema indeksoma  $a_{r,i}$ , da velja  $\mathbf{a}_i = [a_{0,i}, a_{1,i}, \dots, a_{R-1,i}]^\top$ .
- Srednjo vrednost zaporedja vektorjev  $\mathbf{a}_i$ ,  $i = 1, \dots, n$  označujemo s prečno črto nad črko vektorja,  $\bar{\mathbf{a}}$ .
- Matrike označujemo z velikimi tiskanimi črkami v krepki pisavi (npr.  $\mathbf{A}$ ,  $\mathbf{V}$ ).
- Element matrike  $\mathbf{A}$ , velike  $R$  vrstic in  $C$  stolpcev, ki je v vrstici  $r$  in stolpcu

$c$  matrike, označujemo z indeksi vrstice in stolpca,  $a_{r,c}$ , da velja,

$$\mathbf{A} = \begin{bmatrix} a_{0,0} & a_{0,1} & \cdots & a_{0,C-1} \\ a_{1,0} & a_{1,1} & \cdots & a_{1,C-1} \\ \vdots & \vdots & \ddots & \vdots \\ a_{R-1,0} & a_{R-1,1} & \cdots & a_{R-1,C-1} \end{bmatrix} \quad (\text{A.2})$$

- Matrika, ki ima v zaporedju  $n$  matrik zaporedno številko  $i$ , označujemo z indeksom, npr.  $\mathbf{A}_i$ .
- Element v vrstici  $r$  in stolpcu  $c$  matrike  $\mathbf{A}_i$  označujemo kot  $a_{r,c,i}$ .
- Elemente matrike  $\mathbf{A}$  lahko predstavimo v obliki vektorja  $\mathbf{a}$ , ki ga označimo z isto črko v mali tiskani obliki. Element iz vrstice  $r$  in stolpca  $c$  matrike  $\mathbf{A}$  predstavlja element  $a_{(r,c)}$  vektorja  $\mathbf{a}$ . Primer preslikave iz matrike  $\mathbf{A}$  velikosti  $R \times C$  v vektor  $\mathbf{a}$ :

$$\mathbf{a} = [a_{(0,0)}, a_{(1,0)}, \dots, a_{(R-1,0)}, a_{(0,1)}, \dots, a_{(R-1,C-1)}]^\top. \quad (\text{A.3})$$

- Element vektorske predstavitve  $\mathbf{a}_i$  matrike  $\mathbf{A}_i$ , ki se nahaja v vrstici  $r$  in stolpcu  $c$  matrike  $\mathbf{A}_i$ , označimo z  $a_{(r,c),i}$ .

Določene oznake, ki jih uporabljamo v besedilu, imajo v večini primerov enak pomen:

- $R$  — število vrstic v slikah.
- $C$  — število stolpcev v slikah.
- $n$  — število učnih ali referenčnih slik.
- $k$  — število dimenzij podprostora značilnic.
- $\mathbf{A}_i \in \mathbb{R}^R \times \mathbb{R}^C$  — učna ali referenčna slika z zaporedno številko  $i$ .
- $\mathbf{a}_i \in \mathbb{R}^{RC}$  — vektorska oblika učne ali referenčne slike  $\mathbf{A}_i$ .
- $\mathbf{V} \in \mathbb{R}^R \times \mathbb{R}^C$  — navidezna slika.
- $\mathbf{v} \in \mathbb{R}^{RC}$  — vektorska oblika slike  $\mathbf{V}$ .
- $\mathbf{Q} \in \mathbb{R}^R \times \mathbb{R}^C$  — slika, s katero povprašujemo (slika pogleda za lokalizacijo).

- $\mathbf{q} \in \mathbb{R}^{RC}$  — vektorska oblika učne ali referenčne slike  $\mathbf{Q}$ .
- $\mathbf{U} \in \mathbb{R}^{RC} \times \mathbb{R}^k$  — matrika lastnih vektorjev  $\mathbf{U} = [\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k]$ .
- $\mathbf{w}_i \in \mathbb{R}^k$  — preslikava slike  $\mathbf{A}_i$  v prostoru globalnih značilnic.
- $\mathbf{w}_{\mathbf{q}} \in \mathbb{R}^k$  — preslikava slike  $\mathbf{Q}$  v prostoru globalnih značilnic.



# Zahvala

Zahvaljujem se prof. dr. Alešu Leonardisu za mentorstvo, vodstvo, ideje in pomoč pri delu in raziskavah za doktorsko disertacijo. Za izkazano strokovno pomoč se zahvaljujem tudi dr. Danijelu Skočaju in Matjažu Joganu, za sodelovanje na mednarodnem nivoju pa Hyneku Baksteinu s Tehnične Univerze v Pragi.

Izvedba dela ne bi bila mogoča brez delovnega in ustvarjalnega vzdušja v Laboratoriju za umetne vizualne spoznavne sisteme, za kar gre zahvala sočlanom in mojim sodelovcem. Za potrpežljivost in podporo se zahvaljujem tudi svojim prijateljem in staršem.



# Izjava

Izjavljam, da sem doktorsko disertacijo izdelal samostojno pod vodstvom mentorja prof. dr. Aleša Leonardisa. Izkazano pomoč drugih sodelavcev sem v celoti navedel v zahvali.



# Literatura

- [1] Matej Artač. *Lokalizacija in osnovna navigacija mobilnega robota s pomočjo vizualnega učenja in razpoznavanja*. Magistrsko delo, Univerza v Ljubljani, Fakulteta za računalništvo in informatiko, marec 2003.
- [2] Matej Artač, Matjaž Jogan, Hynek Bakstein, in Aleš Leonardis. Panoramic volumes for robot localization. V *IEEE/RSJ International Conference on Intelligent Robots and Systems*, strani 3776–3782, Edmonton, Alberta, Canada, avgust 2005.
- [3] Matej Artač, Matjaž Jogan, in Aleš Leonardis. Mobile robot localization using an incremental eigenspace model. V *IEEE International Conference on Robotics and Automation, maj 11–15, 2002, Washington, D. C.*, letnik 1, strani 1025–1030. IEEE, 2002.
- [4] Hynek Bakstein in Tomáš Pajdla. Rendering novel views from a set of omnidirectional mosaic images. V *Proceedings of Omnivis 2003: Workshop on Omnidirectional Vision and Camera Networks*, Los Alamitos, USA, junij 2003. IEEE Computer Society Press.
- [5] Paul Blaer in Peter Allen. Topological mobile robot localization using fast vision techniques. V *IEEE International Conference on Robotics and Automation, maj 11–15, 2002, Washington, D. C.*, strani 1031–1037. IEEE, 2002.
- [6] David M. Bradley, Rashmi Patel, Nicolas Vandapel, in Scott M. Thayer. Real-time image-based topological localization in large outdoor environments. V *IEEE/RSJ International Conference on Intelligent Robots and Systems*, strani 3062–3069, Edmonton, Alberta, Canada, avgust 2005.
- [7] Dana Cobzas in Hong Zhang. Cylindrical panoramic image-based model for robot localization. V *Proc. of IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2001)*, strani 1924–1930, oktober 2001.

- 
- [8] Guilherme N. DeSouza in Avinash C. Kak. Vision for mobile robot navigation: A survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(2):237–267, februar 2002.
- [9] James Diebel, Kjell Reuterswård, Sebastian Thrun, James Davis, in Rakesh Gupta. Simultaneous localization and mapping with active stereo vision. V *Proceedings of 2004 IEEE/RSJ International Conference on Intelligent Robots and Systems*, strani 3426–3443, 2004.
- [10] Arnaud Doucet in Nando de Freitas, uredniki. *Sequential Monte Carlo Methods in Practice*. Springer-Verlag, New York, 2001.
- [11] Doron Feldman in Daphna Weinshall. Realtime IBR with omnidirectional crossed-slits projection. V *Tenth IEEE International Conference on Computer Vision*, strani 839–845, Beijing, China, oktober 2005.
- [12] John Folkesson, Patric Jensfelt, in Henrik I. Christensen. Graphical SLAM using vision and the measurement subspace. V *IEEE/RSJ International Conference on Intelligent Robots and Systems*, strani 3383–3388, Edmonton, Alberta, Canada, avgust 2005.
- [13] José Gaspar, Niall Winters, in José Santos-Victor. Vision-based navigation and environmental representations with an omnidirectional camera. *IEEE Transaction on Robotics and Automation*, 16(6):890–898, december 2000.
- [14] Toon Goedemé, Tinne Tuytelaars, in Luc Van Gool. Fast wide baseline matching for visual navigation. V *2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, letnik 1, strani 24–29, 2004.
- [15] Peter M. Hall, David Marshall, in Ralph R. Martin. Incremental eigenanalysis for classification. V *British Machine Vision Conference*, letnik 1, strani 286–295, september 1998.
- [16] Richard I. Hartley in Andrew Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, ISBN: 0521540518, druga izdaja, 2004.
- [17] Jean-Bernard Hayet, Frédéric Lerasle, in Michel Devy. A visual landmark framework for indoor mobile robot navigation. V *2002 IEEE International Conference on Robotics and Automation*, letnik 4, strani 3942–3947, Washington, DC, maj 2002.

- [18] H. Hotelling. Analysis of a complex of statistical variables into principal components. *Journal of Educational Psychology*, 24:417–441, 1933.
- [19] Hiroshi Ishiguro, Kim C. Ng, Richard Capella, in Mohan M. Trivedi. Omnidirectional image-based modeling: three approaches to approximated plenoptic representations. *Machine Vision and Applications*, 14(2):94–102, 2003.
- [20] Hiroshi Ishiguro in Saburo Tsuji. Image-based memory of environment. V *Proc. IEEE/RSJ Int. Conf. Intelligent Robots and Systems*, strani 634–639, 1996.
- [21] Matjaž Jogan in Aleš Leonardis. Robust localization using an omnidirectional appearance-based subspace model of environment. *Robotics and Autonomous Systems, Elsevier Science*, 45(1):51–72, 2003.
- [22] Rudolf Kalman. A new approach to linear filtering and prediction problems. *Transactions of the ASME, Journal of Basic Engineering*, 82:35–45, 1960.
- [23] Reinhard Koch, Marc Pollefeys, in Luc J. Van Gool. Realistic surface reconstruction of 3D scenes from uncalibrated image sequences. *Journal of Visualization and Computer Animation*, 11(3):115–127, 2000.
- [24] Jana Košecká, Fayin Li, in Xiaolong Yang. Global localization and relative positioning based on scale-invariant keypoints. *Robotics and Autonomous Systems*, 52(1):27–38, julij 2005.
- [25] Jana Košecká in Xiaolong Yang. Location recognition and global localization based on scale-invariant keypoints. V H. Bischof A. Leonardis, urednik, *Statistical Learning in Computer Vision, ECCV 2004 Workshop, Prague, Czech Republic*, strani 49–58, maj 2004.
- [26] Ben Kröse in Roland Bunschoten. Probabilistic localization by appearance models and active vision. V *In Proceedings of the IEEE International Conference on Robotics and Automation*, strani 2255–2260, 1999.
- [27] Simon Lacroix, Anthony Mallet, David Bonnafous, Gérard Bauzil, Sara Fleury, Matthieu Herrb, in Raja Chatila. Autonomous rover navigation on unknown terrains. *International Journal of Robotic Research*, 21(10–11):917–942, 2002.
- [28] John J. Leonard in Hugh F. Durrant-Whyte. Mobile robot localization by tracking geometric beacons. *IEEE Transactions on Robotics and Automation*, 7(3):376–382, 1991.

- [29] John J. Leonard in Hugh F. Durrant-Whyte. Simultaneous map building and localization for an autonomous mobile robot. V *IEEE/RSJ International Workshop on Intelligent Robots and Systems*, strani 1442–1447, Osaka, Japan, november 1991.
- [30] Aleš Leonardis in Horst Bischof. Robust recognition using eigenimages. *Computer Vision and Image Understanding*, 78(1):99–118, 2000.
- [31] David G. Lowe. Object recognition from local scale-invariant features. V *Proc. of the International Conference on Computer Vision ICCV, Corfu*, strani 1150–1157, 1999.
- [32] Martin C. Martin in Hans Moravec. Robot evidence grids. Tehnično poročilo CMU-RI-TR-96-06, Robotics Institute, Carnegie Mellon University, Pittsburgh, PA, marec 1996.
- [33] Jiří Matas, Ondřej Chum, Martin Urban, in Tomáš Pajdla. Robust wide baseline stereo from maximally stable extremal regions. V *Proceedings of the British Machine Vision Conference*, letnik 1, strani 384–393, London, 2002.
- [34] Yoshio Matsumoto, Masayuki Inaba, in Hirochika Inoue. View-based approach to robot navigation. *Journal of Robotics Society of Japan*, 20(5), 2002.
- [35] Nicholas Metropolis, Arianna W. Rosenbluth, Marshall N. Rosenbluth, avgusta H. Teller, in Edward Teller. Equation of state calculations by fast computing machines. *Journal of Chemical Physics*, 21, 6:1087–1092, 1953.
- [36] Krystian Mikolajczyk in Cordelia Schmid. A performance evaluation of local descriptors. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27(10):1615–1630, oktober 2005.
- [37] Hans Moravec. The stanford cart and the CMU Rover. V *Proceedings of the IEEE*, letnik 71, strani 872–884, 1983.
- [38] Hiroshi Murase in Shree K. Nayar. Visual learning and recognition of 3-D objects from appearance. *IJCV*, 14(1):5–24, januar 1995.
- [39] Shree K. Nayar, Sameer A. Nene, in Hiroshi Murase. Subspace methods for robot vision. *IEEE Trans. on Robotics and Automation*, 12(5):750–758, oktober 1996.

- [40] David Nistér in Frederik Schaffalitzky. What do four points in two calibrated images tell us about the epipoles? V *8th European Conference on Computer Vision*, strani 41–57, Prague, Czech Republic, maj 2004.
- [41] Erkki Oja. *Subspace methods of pattern recognition*. Research Studies Press, 1983.
- [42] Kuniaki Otsuka, Kanji Tanaka, Nobuhiro Okada, in Eiji Kondo. Image based rendering for a mobile robot using visual landmarks. V *IEEE/RSJ International Conference on Intelligent Robots and Systems*, strani 3783–3788, Edmonton, Alberta, Canada, avgust 2005.
- [43] Lucas Paletta, Simone Frintrop, in Joachim Hertzberg. Robust localization using context in omnidirectional imaging. V *IEEE International Conference on Robotics and Automation*, maj 21–26, Seoul, Korea, strani 2072–2077. IEEE, 2001.
- [44] Lucas Paletta, Christophe Le Gal, Georg Ogris, Roland Wack, in Gerhard Paar. Appearance based positioning in urban environments using kalman filtering of PCA sensors. V *28th OAGM/AAPR Workshop*, strani 23–31, junij 2004.
- [45] Alex Rav-Acha, Yael Shor, in Shmuel Peleg. Mosaicing with parallax using time warping. V *Second IEEE Workshop on Image and Video Registration (IVR'04)*, strani 164–164, 2004.
- [46] Eric Royer, Jonathan Bom, Michel Dhome, Benoit Thuilot, Maxime Lhuillier, in François Marmoton. Outdoor autonomous navigation using monocular vision. V *IEEE/RSJ International Conference on Intelligent Robots and Systems*, strani 3395–3400, Edmonton, Alberta, Canada, avgust 2005.
- [47] Stephen Se, David Lowe, in Jim Little. Mobile robot localization and mapping with uncertainty using scale-invariant visual landmarks. *International Journal of Robotic Research*, 8(21):735–760, avgust 2002.
- [48] Juan Manuel Sáez in Francisco Escolano. Entropy minimization SLAM using stereo vision. V *Proceedings of the 2005 IEEE International Conference on Robotics and Automation*, strani 36–43, april 2005.
- [49] Heung-Yeung Shum in Sing Bing Kang. A review of image-based rendering techniques. V *IEEE/SPIE Visual Communications and Image Processing (VCIP) 2000*, strani 2–13, junij 2000.

- [50] Robert Sim. Stable exploration for bearings-only SLAM. V *Proceedings of the 2005 IEEE International Conference on Robotics and Automation*, strani 2411–2416, april 2005.
- [51] H. Sirovich in M. Kirby. Low-dimensional procedure for the characterization of human faces. *J. Opt. Soc. America A*, 4(3):519–524, 1987.
- [52] Danijel Skočaj. *Robustni pristopi k vizualnemu učenju in razpoznavanju na osnovi podprostorov*. PhD thesis, Univerza v Ljubljani, Fakulteta za računalništvo in informatiko, 2003.
- [53] Tomáš Svoboda. *Central Panoramic Cameras Design, Geometry, Egomotion*. PhD thesis, Faculty of Electrical Engineering of Czech Technical University, 1999.
- [54] Charles Thorpe, Justin David Carlson, David Duggins, Jay Gowdy, Robert MacLachlan, Christoph Mertz, Arne Suppe, in Chieh-Chih Wang. Safe robot driving in cluttered environments. V *Proceedings of the 11th International Symposium of Robotics Research*, oktober 2003.
- [55] Sebastian Thrun, Arno Bücken, Wolfram Burgard, Dieter Fox, Thorsten Fröhlinghaus, Daniel Henning, Thomas Hofmann, Michael Krell, in Timo Schmidt. Map learning and high-speed navigation in RHINO. V D. Kortenkamp, R.P. Bonasso, in R. Murphy, uredniki, *AI-based Mobile Robots: Case Studies of Successful Robot Systems*. MIT Press, 1998.
- [56] Carlo Tomasi in Takeo Kanade. Detection and tracking of point features. Tehnično poročilo CMU-CS-91-132, Carnegie Mellon University, april 1991.
- [57] Vladimir Tucakov, Michael Sahota, Don Murray, Alan Mackworth, Jim Little, Stewart Kingdon, Cullen Jennings, in Rod Barman. Spinoza: A stereoscopic visually guided mobile robot. V *Proceedings of the Thirtieth Hawaii International Conference on System Sciences*, letnik 5, strani 188–197, januar 1997.
- [58] Matthew Turk in Alex Pentland. Face recognition using eigenfaces. V *Proc. Computer Vision and Pattern Recognition, CVPR-91*, strani 586–591, 1991.
- [59] Yasushi Yagi, Kousuke Imai, Kentaro Tsuji, in Masahiko Yachida. Iconic memory-based omnidirectional route panorama navigation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27(1):78–87, januar 2005.

- 
- [60] Yasushi Yagi, Yoshimitsu Nishizawa, in Masahiko Yachida. Map-based navigation for a mobile robot with omnidirectional image sensor COPIS. *IEEE Trans. on Robotics and Automation*, 11(5):634–648, oktober 1995.
- [61] Zhengyou Zhang. Flexible camera calibration by viewing a plane from unknown orientations. V *The Proceedings of the Seventh IEEE International Conference on Computer Vision*, letnik 1, strani 666–673, 1999.
- [62] Jiang Yu Zheng in Saburo Tsuji. Panoramic representation of scenes for route understanding. V *10th International Conference on Pattern Recognition*, strani 161–167, junij 1990.
- [63] Zhigang Zhu in Allen R. Hanson. 3D LAMP: a new layered panoramic representation. V *8th IEEE International Conference on Computer Vision*, letnik 2, strani 723–730, Vancouver, B.C., Canada, julij 2001.
- [64] Zhigang Zhu in Allen R. Hanson. LAMP: 3D layered, adaptive-resolution and multi-perspective panorama — a new scene representation. *Computer Vision and Image Understanding*, 96:294–326, december 2004.
- [65] Zhigang Zhu, Guangyou Xu, in Xueyin Lin. Panoramic EPI generation and analysis of video from a moving platform with vibration. V *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, strani 531–537, Fort Collins, Colorado, junij 1999.
- [66] Assaf Zomet, Doron Feldman, Shmuel Peleg, in Daphna Weinshall. Mosaicing new views: The crossed-slits projection. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 25(6):741–754, junij 2003.