



MAGISTRSKA NALOGA

PETER PEER

LJUBLJANA, 2001

Gradnja globinskih panoramskih slik s postopkom mozaičenja

Peter Peer

MAGISTRSKA NALOGA
IZ
RAČUNALNIŠTVA IN INFORMATIKE

predložena
Fakulteti za računalništvo in informatiko
Univerze v Ljubljani
kot delna izpolnitev pogoja za pridobitev naslova
magister računalništva in informatike

Mentor: prof. dr. Franc Solina

April 2001

Magistrska naloga je bila izdelana pod mentorstvom prof. dr. Franca Soline in je last Fakultete za računalništvo in informatiko v Ljubljani. Za objavlanje in uporabo rezultatov magistrskega dela je potrebno soglasje zgoraj omenjene ustanove.

Tekst je oblikovan z \LaTeX urejevalnikom besedil.

Lektor: Nataša Meh Peer, prof. slov. in umet. zgod.

Gradnja globinskih panoramskih slik s postopkom mozaičenja

Peter Peer

Mentor: prof. dr. Franc Solina

POVZETEK

V nalogi je predstavljen sistem za gradnjo globinskih panoramskih slik s postopkom mozaičenja standardnih slik zajetih z eno kamero. Zaradi odmika optičnega centra kamere iz središča vrtenja sistema, lahko z eno kamero zajamemo paralakso gibanja, ki omogoča stereo rekonstrukcijo. Panoramske slike ustvarimo tako, da rotacijsko roko vrtimo za kot, ki ustreza premiku za eno kolono zajete slike. Iz geometrije sistema je razvidna enačba za izračun globine. Za izračun globine pa moramo najti korespondenčne točke na paru stereo panoram, zato moramo izračunati epipolarno geometrijo. Izkaže se, da je epipolarna geometrija zelo enostavna v primeru računanja rekonstrukcije na podlagi simetričnega para panoram. Simetričen par panoram dobimo, če vzamemo simetrične trakove levo in desno od sredine zajete slike. Epipolarne premice simetričnih parov panoram so istoležne vrstice panoram.

Poudarek naloge je predvsem v analizi rezultatov in zmogljivosti zasnovanega sistema. Sistem daje dobre rezultate pri rekonstrukciji majhnih prostorov.

KLJUČNE BESEDE

projektivna geometrija, stereo vid, rekonstrukcija, panoramska slika, globinska slika, mozaičenje, paralaksa gibanja

Panoramic depth imaging with mosaicking

Peter Peer

Supervisor: prof. dr. Franc Solina

ABSTRACT

In this work we present a panoramic depth imaging system. The system is mosaic-based which means that we use a single rotating camera and assemble the captured images in a mosaic. Due to a setoff of the camera's optical center from the rotational center of the system we are able to capture the motion parallax effect which enables the stereo reconstruction. The camera is rotating on a circular path with the step defined by an angle, equivalent to one column of the captured image. The equation for depth estimation can be easily extracted from system geometry. To find the corresponding points on stereo pair of panoramic images the epipolar geometry needs to be determined. It can be shown that the epipolar geometry is very simple whether we are doing the reconstruction based on a symmetric pair of stereo panoramic images. We get a symmetric pair of stereo panoramic images when we take symmetric stripes on the left and on the right side from the captured image center column. The epipolar lines of the symmetrical pair of panoramic views are image rows.

We focused mainly on the system analysis. The system performs well in the reconstruction of small indoor spaces.

KEY WORDS

projective geometry, stereo vision, reconstruction, panoramic image, depth image, mosaicking, motion parallax effect

Kazalo

Povzetek	vii
Abstract	viii
Slike	xii
Tabele	xiv
1 Uvod	1
Zgradba magistrske naloge	3
2 Nastanek slike	4
2.1 Uvod	4
2.2 Osnovna matrika perspektivne projekcije	4
2.3 Notranji parametri kamere	9
2.4 Zunanji parametri kamere	11
2.5 Sklep	13
3 Stereo vid	14
3.1 Uvod	14
3.2 Epipolarna omejitev	15
3.3 Računanje epipolarne geometrije	16
3.4 Delo v normaliziranem koordinatnem sistemu	18
3.5 Določanje odmika kamere na podlagi korespondence točk	19
3.5.1 Iskanje korespondenčnih točk	21
3.5.2 Algoritem osmih točk	22
3.6 Rekonstrukcija 3D točk	23
4 Panoramski pogledi in stereo rekonstrukcija	25
4.1 Uvod	25
4.2 Panoramske kamere	26
4.2.1 Katadioptrične kamere	26
4.2.2 Dioptrične kamere	27

4.2.3	Panoramski pogledi s pomočjo pomičnih delov	28
4.2.4	Mozaičenje standardnih slik v panoramske poglede	29
4.3	Analiza rezultatov panoramskih kamer	34
5	Sistem za gradnjo globinskih panoramskih slik	37
5.1	Uvod	37
5.1.1	Sestavni deli sistema	37
5.1.2	Omejitve sistema	38
5.2	Geometrija sistema	40
5.3	Ustvarjanje panoramskih slik	41
5.4	Računanje epipolarne geometrije	50
5.4.1	Projekcijska enačba za linijsko kamero	50
5.4.2	Projekcijska enačba za večperspektivno panoramsko sliko	52
5.4.3	Lastnost epipolarnih premic simetričnih parov panoram	53
5.5	3D rekonstrukcija	56
5.5.1	Iskanje korespondenčnih točk	57
5.6	Analiza sistema in rezultatov	59
5.6.1	Čas gradnje panoramskih slik	59
5.6.2	Omejitev sistema zaradi ločljivosti rotacijske roke	60
5.6.3	Omejitev iskanja na epipolarni premici	61
5.6.4	Problem prekrivanja	64
5.6.5	Zaupanje v ocenjeno globino	64
5.6.6	Rezultati	73
6	Zaključek	90
	Nadaljnje delo	91
	Dodatki	93
A	3D rotacija	93
A.1	Splošne lastnosti rotacijske matrike	93
A.2	Rotacija okoli koordinatnih osi	94
B	Računski primer vpliva zunanjih parametrov kamere na 3D točko M	95
C	Skica računskega primera rekonstrukcije 3D točke	97
D	Možne ocene globin glede na kot φ	100
E	O implementaciji	102

KAZALO	xi
F Terminološki slovarček	106
Literatura	109
Zahvala	113
Izjava	114

Slike

2.1	Nastanek slike v točkasti kameri.	5
2.2	Model točkaste kamere.	6
2.3	Goriščna ravnina \mathcal{F} je vzporedna s slikovno ravnino \mathcal{R} in je od nje oddaljena za goriščno razdaljo f	7
2.4	Vrstični vektorji matrike $\tilde{\mathbf{P}}$ podajajo goriščno ravnino in premico, definirano z optičnim centrom in točko o na slikovni ravnini.	8
2.5	Spreminjanje koordinatnega sistema slikovne ravnine.	9
2.6	Relacija med realno (pravo) in normalizirano slikovno ravnino.	11
2.7	Spreminjanje koordinatnega sistema 3D prostora.	12
3.1	Problem stereo računalniškega vida.	14
3.2	Epipolarna geometrija.	15
3.3	Dve kameri in njuni normalizirani slikovni ravnini.	18
3.4	Omejitev parametrov gibanja, ki jih dobimo iz korespondenčnih točk.	20
3.5	Princip korelacijske metode za iskanje korespondenčnih točk.	21
4.1	Svetlobni žarki, ki se sekajo v efektivnem gledišču v , se odbijejo v optični center kamere c	27
4.2	a) vhodna slika, zajeta s katadioptrično kamero (uporabljeno je sferično ogledalo), b) panoramska slika po opravljeni transformaciji vhodne slike na valj.	28
4.3	Epipolarna geometrija hiperbolične katadioptrične panoramske kamere.	29
4.4	Primer horizontalne paralakse gibanja [27].	30
4.5	Globino z točke na objektu lahko ocenimo z meritvijo kotov θ in α	30
4.6	Koncentrične panorame: a) sistem za zajemanje, b) proces ustvarjanja večperspektivnih panoram.	31
4.7	Rotacijske panorame: a) sistem za zajemanje, b) proces ustvarjanja večperspektivnih panoram.	32
4.8	Gradnja stereo (valjčnih) panoram s šivanjem trakov iz zaporednih slik.	32
4.9	Primerjava disparitete med: a) standardnim stereo parom in b) stereo parom panoram.	33

4.10	Povečan primer korespondenčnih točk v simetričnem paru panoram in ne-simetričnem paru panoram.	34
4.11	Uporaba ne-središčnih kamer.	35
5.1	Strojni deli sistema.	38
5.2	Geometrija sistema za ustvarjanje ene večperspektivne panorame. . .	40
5.3	Primer združevanja trakov zajetih slik, ki so širši od ene same kolone, v panoramski pogled.	42
5.4	Primer popačenja panoramskega pogleda zaradi dinamičnosti scene. . .	43
5.5	Radialno popačenje uporabljene kamere na mrežnem vzorcu.	44
5.6	Intenzitetno popačenje uporabljene kamere.	45
5.7	Rezultati geometrijske transformacije iz krogelnih koordinat v valjčne koordinate pri različnih naklonih kamere.	46
5.8	Zasnovan sistem zgradi v enem obhodu rotacijske roke okoli vertikalne osi 11 panoramskih slik.	48
5.9	Panoramno gradimo iz kolon zajetih slik, ki so, gledano iz optičnega centra kamere, od središčne kolone odmaknjene za kot φ	49
5.10	Projekcija 3D točke na slikovno ravnino linijske kamere.	51
5.11	Relacija med koordinatnimi sistemi.	52
5.12	Koncept projekcije točke v 3D svetovnem koordinatnem sistemu na par večperspektivnih panoramskih slik.	54
5.13	Omejitev iskanja na epipolarni premici.	63
5.14	Problem prekrivanja.	65
5.15	Graf funkcije globine l	66
5.16	Grafa odvisnosti funkcije globine l od kota θ pri različnih vrednostih kota φ	67
5.17	Direktno določevanje največje globine, v katere oceno še zaupamo. . .	71
5.18	Ustvarjene goste globinske slike.	76
5.19	Slike zaupanja v ocenjeno globino na podlagi ocene postopka normalizirane korelacije.	77
5.20	Rezultati gradnje gostih globinskih slik dela scene pri različnih vrednosti parametrov.	79
5.21	Zaupamo le v tiste ocene globine, ki so večje od praga zaupanja. . .	81
5.22	Ustvarjene redke globinske slike.	83
5.23	Skica prostora, ki smo ga rekonstruirali.	84
5.24	Tloris rekonstruirane scene (1.1).	86
5.25	Tloris rekonstruirane scene (1.2).	87
5.26	Tloris rekonstruirane scene (2.1).	88
5.27	Tloris rekonstruirane scene (2.2).	89
E.1	Uporabniški vmesniki programskih delov sistema.	102

Tabele

4.1	Primerjava različnih panoramskih kamer.	35
5.1	Nekaj osnovnih podatkov o postopku mozaičenja.	60
5.2	Ovrednotenje napake za en slikovni element pri oceni kota θ	68
5.3	Pomen napake za en slikovni element pri oceni kota θ za oceno najmanjše možne globine l_{\min} in največje možne globine l_{\max}	70
5.4	Časi gradnje gostih globinskih slik.	75
5.5	Časi gradnje redkih globinskih slik.	80
5.6	Primerjava med dejanskimi in ocenjenimi globinami.	82
D.1	Možne ocene globin l za kot $2\varphi = 3,6125^\circ$	100
D.2	Možne ocene globin l za kot $2\varphi = 29,9625^\circ$	101

Poglavje 1

Uvod

Danes se na tržišču že pojavljajo prvi komercialni izdelki, ki znajo zgraditi 3D modele resničnih objektov. Vsi poznamo pojem grafika, vendar 3D rekonstrukcija nima toliko opraviti z grafiko, kot ima opravka z računalniškim vidom. Najbolj laična razlaga razlike med grafiko in računalniškim vidom bi bila, da sta to dva pola iste stvari: pri grafiki podamo parametre in želimo dobiti sliko, pri računalniškem vidu pa ravno obratno, imamo sliko in želimo dobiti parametre (informacijo), ki kar najbolje opisujejo sliko. Seveda je to zelo pomanjkljiva definicija, saj se danes obe vedi na področju 3D modeliranja lepo dopolnjujeta, vendar ji nekaj let nazaj ne bi nihče oporekal. Osnovni namen računalniškega vida je zgraditi sistem, ki bo sposoben uporabljati vizualno informacijo iz svoje okolice za samostojno interakcijo z okolico.

Naloga 3D rekonstrukcije je torej avtomatska gradnja 3D modelov resničnih objektov. Že na samem začetku pa se metode ločijo med seboj glede na namen uporabe. Prva veja je namenjena vizualizaciji, kar pomeni, da je njen cilj narediti čimbolj vizualno ustrezen model, druga veja pa je namenjena gradnji modelov CAD, ki zahtevajo veliko natančnost in opis končnega modela s standardnim naborom pod-modelov.

Poznamo štiri pristope zajemanja 3D podatkov: s pomočjo tipala na robotski roki (ang. *Coordinate Measuring Machines*), po principu časa vrnitve (ang. *time of flight*), CAT (ang. *Computer Aided Tomography*) in s pomočjo optike.

Recimo, da množico 3D točk predstavimo s 3D matriko, kjer vsaka dimenzija predstavlja eno od osi x , y in z koordinatnega sistema. Pristop s tipalom je za laičnega opazovalca delovanja 3D skenerja najbolj razumljiva, saj opazovalec vidi dejansko polnjenje matrike. Kako? Robotska roka s tipalom se dotakne resničnega objekta in 3D točka na površini objekta se zapiše neposredno v matriko, saj poznamo vrednosti potrebnih treh koordinat. Tipalo si lahko poenostavljeno predstavljamo kot človeški prst, ob podani merski enoti in koordinatnem izhodišču pa je polnjenje matrike trivialen postopek.

Med sisteme za 3D rekonstrukcijo, ki delujejo po principu časa vrnitve, spada tudi

skener z laserskim žarkom. Ta deluje tako, da meri čas od oddaje žarka, ki ga usmeri na skeniran objekt, do trenutka, ko zazna odbiti žarek. Ker je hitrost razširjanja žarka konstantna, lahko na podlagi izmerjenega časa izračunamo opravljeno pot in s tem dobimo 3D koordinato točke na resničnem objektu.

Pristop CAT največkrat srečamo v medicinskih in industrijskih aplikacijah, saj ima to prednost, da dobimo informacijo tudi o notranjosti objekta in ne le o njegovi površini. Naprave, ki uporabljajo ta pristop, največkrat označimo z imenom skenerji za računalniško tomografijo z X-žarki.

Danes najbolj razširjeni sistemi za 3D rekonstrukcijo, tudi zaradi cene in varnosti, uporabljajo optične naprave. Podatke o 3D točkah na površini objekta pridobimo z obdelavo 2D slik, zajetih z navadnimi kamerami. Za avtomatsko gradnjo 3D modelov se danes uporabljajo skoraj izključno takšni sistemi.

Poznamo dva tipa optičnih sistemov za zajemanje 3D podatkov: Aktivni sistemi projicirajo določeno svetlobo na objekt, projicirana svetloba pa zagotavlja učinkovito in zanesljivo 3D informacijo. Vendar ne gre za kakršnokoli svetlobo, pač pa za specializirane vire, kot so recimo laserski žarek, strukturirana svetloba (vzorec, ki ga projiciramo na resničen objekt) ipd. Rezultat, zajete in tako obdelane slike, je globinska slika. Medtem ko intenzitetne slike podajajo količino svetlobe, ki pade na foto-občutljivo napravo, globinske slike podajajo oceno 3D strukture objekta, seveda le iz enega zornega kota. Svetlejša kot je točka globinske slike, bližje je ta točka v 3D prostoru. Vsaka točka globinske slike torej podaja oddaljenost znane referenčne točke od točke na objektu. Na drugi strani pa imamo pasivne sisteme, ki ne projicirajo vzorca na objekt in so zato povsem odvisni od teksture objekta. Tradicionalni postopki, za pridobivanje globine iz takšnih slik, temeljijo na algoritmu stereo (stereoskopskega) vida. Osnova stereo vida sta dve kameri, ki posnemata človeški vid. V nadaljevanju bomo delali s pasivnim sistemom.

Standardna kamera ima omejen vertikalni in horizontalni zorni kot, ki sta navadno manjša od človeškega vidnega polja. Že zaradi tega si je človek vedno prizadeval, da bi ustvaril sliko, ki bi obsegala večji zorni kot, celo večjega od človeškega vidnega polja. Takšne slike imenujemo panoramske slike.

Pod pojmom stereo rekonstrukcija razumemo gradnjo globinskih slik iz več slik. Postopek stereo rekonstrukcije temelji na relacijah med točkami in linijami na sceni in slikah le-te. Slike scene lahko nastopajo v parih, trojicah, četvericah itd., prav na podlagi števila slik pa imenujemo tudi omenjene relacije: epipolarna omejitev, trislikovna omejitev, štirislikovna omejitev itd. Zelo zaželjena lastnost slik je, da so te točke ali linije dobro vidne na vseh slikah neke scene. In takšno lastnost nam v največji meri zagotavljajo panoramske kamere. To dejstvo je predstavljalo osnovno motivacijo pri gradnji našega sistema za gradnjo globinskih slik, saj ima panoramska kamera v nasprotju s standardno kamero širše vidno polje, zato je možno rekonstruirati več globinskih točk kot s standardnim parom stereo slik.

Zgradba magistrske naloge

V nadaljevanju je najprej predstavljen postopek 3D rekonstrukcije, ki predstavlja osnovo za gradnjo globinskih slik, z vidika projektivne geometrije: V poglavju 2 je izpeljan postopek nastanka slike v kameri, kjer izpostavimo tudi pomembne relacije za postopek rekonstrukcije. Ker delamo rekonstrukcijo na podlagi dveh slik, je v poglavju 3 opisan stereo sistem. Tukaj izpeljemo epipolarno omejitev, preslikavo v normaliziran koordinatni sistem, odkim kamere na podlagi korespondence točk in samo rekonstrukcijo 3D točke.

V poglavju 4 so predstavljeni pristopi k ustvarjanju panoramskih slik, stereo panoram in globinskih panoramskih slik, s poudarkom na mozaičenju standardnih slik v panoramske poglede in pristopih, ki vodijo k 3D rekonstrukciji. Predstavljene so geometrije sistemov, načini ustvarjanja panoram, postopki rekonstrukcije in analiza rezultatov.

V poglavju 5 je predstavljen prototip sistema za gradnjo globinskih panoramskih slik s postopkom mozaičenja standardnih slik zajetih z eno kamero. Zaradi odkim optičnega centra kamere iz središča vrtenja sistema, lahko z eno kamero zajamemo paralakso gibanja, ki omogoča stereo rekonstrukcijo. Panoramske slike ustvarimo tako, da rotacijsko roko vrtimo za kot, ki ustreza premiku za eno kolono zajete slike. Iz geometrije sistema je razvidna enačba za izračun globine. Za izračun globine pa moramo najti korespondenčne točke na paru stereo panoram, zato moramo izračunati epipolarno geometrijo. Izkaže se, da je epipolarna geometrija zelo enostavna v primeru računanja rekonstrukcije na podlagi simetričnega para panoram. Simetričen par panoram dobimo, če vzamemo simetrične trakove levo in desno od sredine zajete slike. Epipolarne premice simetričnih parov panoram so istoležne vrstice panoram. Teža poglavja je predvsem v analizi rezultatov in zmogljivosti zasnovanega sistema.

Nalogo zaključimo v poglavju 6 s sklepi.

Za sklepi sledijo dodatki, med katerimi sta tudi Terminološki slovarček in Informacije o implementaciji.

Predeu gremo na naslednje poglavje pa zapišimo še dve pomembni informaciji: kot testni računalnik je služil PC Intel PII./350 MHz, vsi izračuni in izpeljave pa so bile preverjene s programom Mathematica, ki ga izdeluje podjetje Wolfram Research.

Poglavje 2

Nastanek slike

2.1 Uvod

To poglavje je povzeto po knjigi Oliverja Faugerasa: *Three-Dimensional Computer Vision: A Geometric Viewpoint* [9].

Označimo s $\tilde{\mathbf{P}}$ matriko reda 3×4 , ki predstavlja perspektivno projekcijo iz 3D svetovnega koordinatnega sveta v 2D koordinatni sistem slike. Oglejmo si postopek izpeljave matrike $\tilde{\mathbf{P}}$, ki jo imenujemo tudi matrika perspektivne projekcije ter njeno uporabo pri nastanku slike:

Na vrednost $\tilde{\mathbf{P}}$ vplivajo trije dejavniki:

- osnovna matrika perspektivne projekcije
- notranji parametri kamere in
- zunanji parametri kamere.

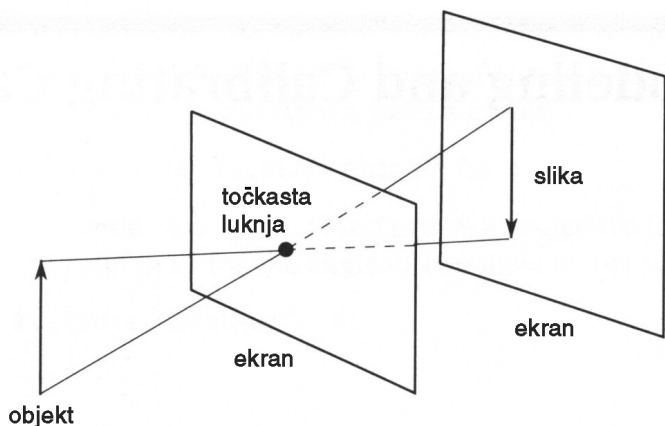
Notranje in zunanje parametre kamere dobimo ob kalibraciji sistema. S tilde (\sim) označujemo dejstvo, da delamo s homogenimi koordinatami¹.

2.2 Osnovna matrika perspektivne projekcije

Oglejmo si sistem s slike 2.1. Sistem sestavljata dva ekrana, prvi pa vsebuje točkasto luknjo v središču. Skozi to točko prehajajo svetlobni žarki, ki jih oddaja objekt ali pa so od njega odbiti. Tako na drugem ekranu dobimo invertirano sliko tega objekta.

Na podlagi tega sistema lahko zgradimo geometrijski model točkaste kamere (ang. *pinhole camera*) (slika 2.2). Model je sestavljen iz slikovne ravnine \mathcal{R} in operacije perspektivne projekcije: 3D točka C , ki predstavlja optični center kamere

¹Za homogene oziroma projektivne koordinate velja, da skaliranje ne vpliva na njih: $(X, Y, Z) = (\lambda X, \lambda Y, \lambda Z)$.



Slika 2.1: Nastanek slike v točkasti kameri.

in je od slikovne ravnine \mathcal{R} oddaljena za razdaljo f (goriščna razdalja optičnega sistema), je uporabljena pri preslikavi 3D točke M v 2D točko m tako, da poiščemo presečišče premice, ki je definirana s točkama M in C , s slikovno ravnino \mathcal{R} .

Na tem mestu pa definirajmo še en pojem za lažje razumevanje nadaljnje snovi: Optična os je premica, ki gre skozi optični center in je pravokotna na slikovno ravnino \mathcal{R} .

Iz slike 2.3 je razvidno, da relacija med koordinatnim sistemom slike in 3D prostorom (je očitna, ampak) ni linearna²:

$$-\frac{f}{z} = \frac{u}{x} = \frac{v}{y}. \quad (2.1)$$

Če uporabimo projektivne oziroma homogene koordinate, relacija postane linearna:

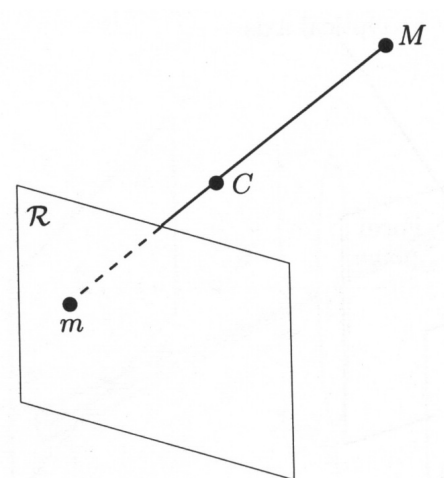
$$\begin{bmatrix} U \\ V \\ S \end{bmatrix} = \begin{bmatrix} -f & 0 & 0 & 0 \\ 0 & -f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \quad (2.2)$$

in velja:

$$u = \frac{U}{S}, \quad v = \frac{V}{S}, \quad S \neq 0. \quad (2.3)$$

Kaj pa se zgodi, če je $z = 0$? Točka M , ki leži na goriščni ravnini \mathcal{F} , nima svoje preslikave na slikovni ravnini \mathcal{R} zato, ker je premica, ki je definirana s točkama M

²Iz slike 2.3 je razvidno tudi, da je točka $m = (u, v)$ na ravnini slike tudi točka v prostoru, katere koordinate so $m_{3D} = (u, v, -f)$.



Slika 2.2: Model točkaste kamere.

in C , vzporedna s slikovno ravnino \mathcal{R} , torej je ne seka! Če govorimo projektivno, jo seka v neskončnosti; $S=0$ je enačba premice v neskončnosti slikovne ravnine \mathcal{R} .

Enačba (2.2) je zapisana s homogenimi koordinatami (U, V, S) točke m , torej je definirana do multiplikativnega faktorja natančno. Je torej projektivna, zato jo lahko zapišemo z uporabo homogenih koordinat (X, Y, Z, T) točke M :

$$\begin{bmatrix} U \\ V \\ S \end{bmatrix} = \begin{bmatrix} -f & 0 & 0 & 0 \\ 0 & -f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ T \end{bmatrix},$$

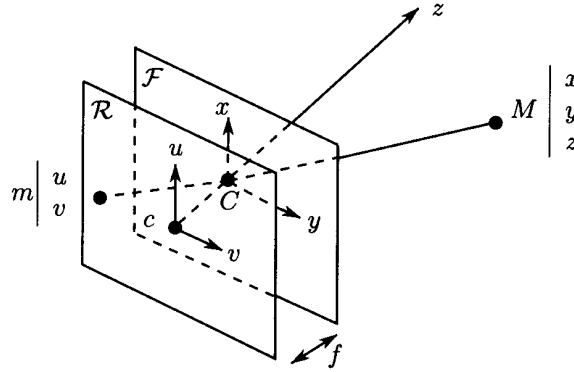
kar lahko v matrični obliki zapišemo kot:

$$\tilde{\mathbf{m}} = \tilde{\mathbf{P}}\tilde{\mathbf{M}}, \quad (2.4)$$

kjer je $\tilde{\mathbf{m}} = [U, V, S]^T$, $\tilde{\mathbf{M}} = [X, Y, Z, T]^T$, $\tilde{\mathbf{P}}$ pa je osnovna matrika perspektivne projekcije.

Na kamero lahko gledamo kot na sistem, ki izvede linearno projektivno transformacijo iz projektivnega prostora \mathcal{P}^3 na projektivno ravnino \mathcal{P}^2 . To je le eden izmed mnogih primerov, kjer se s pomočjo projektivne geometrije stvari poenostavijo: namesto, da se ukvarjamo z nelinearnimi enačbami (kot v primeru enačbe (2.1)), lahko uporabimo linearno relacijo (kot v primeru enačbe (2.2)) in izkoristimo moč linearne algebre.

Zapišimo geometrijsko interpretacijo vrstičnih vektorjev matrike $\tilde{\mathbf{P}}$:



Slika 2.3: Goriščna ravnina \mathcal{F} je vzporedna s slikovno ravnino \mathcal{R} in je od nje oddaljena za goriščno razdaljo (ang. *focal length*) f .

$$\tilde{\mathbf{P}} = \begin{bmatrix} \tilde{\mathbf{Q}}_1^T \\ \tilde{\mathbf{Q}}_2^T \\ \tilde{\mathbf{Q}}_3^T \end{bmatrix} = \begin{bmatrix} \mathbf{q}_1^T & q_{14} \\ \mathbf{q}_2^T & q_{24} \\ \mathbf{q}_3^T & q_{34} \end{bmatrix}, \quad (2.5)$$

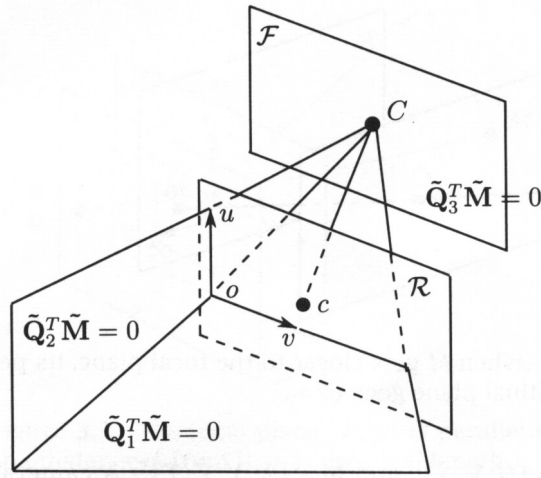
kjer je $\tilde{\mathbf{Q}}_i$, $i = 1, 2, 3$, vektor velikosti 4×1 in $\tilde{\mathbf{Q}}_i^T = [\mathbf{q}_i^T \quad q_{i4}]$. Vsak vektor predstavlja projektivno ravnino z enačbo $\tilde{\mathbf{Q}}_i^T \tilde{\mathbf{M}} = 0$. Glede na enačbo (2.2), ravnina z enačbo $\tilde{\mathbf{Q}}_3^T \tilde{\mathbf{M}} = 0$ opisuje točke na slikovni ravnini \mathcal{R} , za katere velja $S = 0$, torej točke v neskončnosti. Opisuje torej goriščno ravnino \mathcal{F} . Ravnina z enačbo $\tilde{\mathbf{Q}}_1^T \tilde{\mathbf{M}} = 0$ in ravnina z enačbo $\tilde{\mathbf{Q}}_2^T \tilde{\mathbf{M}} = 0$ opisujeta točke na slikovni ravnini \mathcal{R} , za katere velja $U = 0$ in $V = 0$. Presek teh dveh ravnin je premica, ki gre skozi optični center kamere C in točko o na slikovni ravnini \mathcal{R} (slika 2.4). Pri tem je potrebno poudariti, da v splošnem ta premica ni optična os kamere.

Na tem mestu pa podajmo še dve koristni informaciji, ki ju lahko izračunamo s pomočjo matrike perspektivne projekcije $\tilde{\mathbf{P}}$ in sta pomembni za nadaljno razlago:

- Iz slike 2.4 je razvidno, da je optični center kamere C definiran kot presek treh ravnin z enačbami $\tilde{\mathbf{Q}}_i^T \tilde{\mathbf{M}} = 0$, $i = 1, 2, 3$. C lahko torej izračunamo tako, da rešimo sistem treh linearnih enačb:

$$\tilde{\mathbf{P}} \begin{bmatrix} \mathbf{C} \\ 1 \end{bmatrix} = \mathbf{0}. \quad (2.6)$$

Če matriko $\tilde{\mathbf{P}}$, ki je reda 3×4 zapišemo kot $[\mathbf{P} \quad \tilde{\mathbf{p}}]$, pri čemer je \mathbf{P} matrika reda 3×3 , $\tilde{\mathbf{p}}$ pa vektor reda 3×1 in predpostavljamo, da je rang matrike \mathbf{P} enak tri, potem lahko optični center kamere izračunamo kot:



Slika 2.4: Vrstični vektorji matrike $\tilde{\mathbf{P}}$ podajajo goriščno ravnino in premico, definirano z optičnim centrom in točko o na slikovni ravnini.

$$\mathbf{C} = -\mathbf{P}^{-1}\tilde{\mathbf{p}}, \quad (2.7)$$

pri čemer je po enačbi (2.5) vektor $\tilde{\mathbf{p}} = [q_{14}, q_{24}, q_{34}]^T$.

Zaskrbljenost, ker računamo evklidske koordinate točke iz matrike $\tilde{\mathbf{P}}$, ki je definirana zgolj do skalarne faktorja natančno, je odveč, saj če $\tilde{\mathbf{P}}$ nadomestimo z $\lambda\tilde{\mathbf{P}}$, $\tilde{\mathbf{p}}$ z $\lambda\tilde{\mathbf{p}}$ in \mathbf{P}^{-1} z $1/\lambda\mathbf{P}^{-1}$, vidimo, da se rezultat točke C ne spremeni.

- Formalno v \mathcal{P}^2 (projektivni ravnini) ni razlike med točkami in premicami – točke in premice predstavimo na enak način. Ta lastnost je znana pod imenom *Princip dualnosti*: Na točko, predstavljeno s koordinatnim vektorjem \mathbf{x} , lahko gledamo kot na množico premic, ki gredo skozi njo. Te premice predstavimo s koordinatnim vektorjem \mathbf{u} , za katerega velja $\mathbf{u}^T\mathbf{x}=0$. Inverzno pa velja, da lahko na premico \mathbf{u} gledamo kot na množico točk \mathbf{x} , za katere velja ista enačba. Na podlagi tega velja naslednji izrek, ki ga bomo potrebovali v nadaljevanju:

IZREK 1³

Premico, ki gre skozi točki m_1 in m_2 , zapišemo kot vektorski produkt $m_1 \times m_2$.

Drugo uporabno informacijo dobimo iz enačbe 3D premice, ki je določena s slikovnim elementom m (v prostoru) in optičnim centrom C . To premico imenujemo optični žarek, ki ga definira točka m na slikovni ravnini. Po enačbi

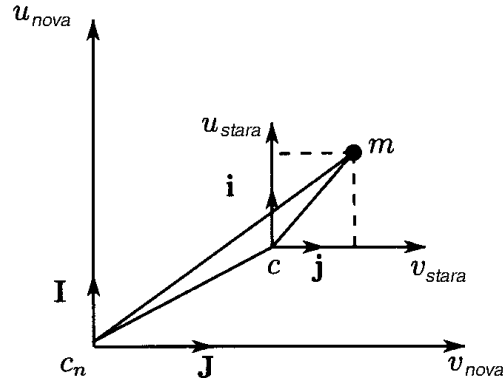
³Dokaz si lahko ogledate v [9], na strani 15.

(2.7) poznamo že točko C , ki leži na tej premici, druga točka, ki leži na tej premici pa je točka v neskončnosti $\tilde{\mathbf{D}}$, katere projektivne koordinate so $[\mathbf{D}^T, 0]^T$.⁴ Po enačbi (2.4) za \mathbf{D} velja:

$$\tilde{\mathbf{m}} = \tilde{\mathbf{P}}\tilde{\mathbf{D}} = \mathbf{P}\mathbf{D} \Rightarrow \mathbf{D} = \mathbf{P}^{-1}\tilde{\mathbf{m}}. \quad (2.8)$$

2.3 Notranji parametri kamere

Če spremenimo izhodišče koordinatnega sistema slike in uvedemo različne enote za slikovne osi k_u in k_v , dobimo tri notranje parametre kamere: koordinati starega izhodišča u_0 in v_0 ter kvocient k_u/k_v .



Slika 2.5: Spreminjanje koordinatnega sistema slikovne ravnine. Star koordinatni sistem ima izhodišče v točki c , ki jo dobimo kot presečišče optične osi s slikovno ravnino in ima na obeh oseh enake enote. Sedaj gremo iz starega koordinatnega sistema v novi koordinatni sistem, ki ima izhodišče v c_n in lahko ima različne enote na oseh (zaradi lastnosti kamere, s katero zajemamo podatke).

Iz slike 2.5 je razvidno, da za slikovni element m velja:

$$\overline{c_n m} = \overline{c_n c} + \overline{c m}. \quad (2.9)$$

$\overline{c m}$ lahko sedaj v starem koordinatnem sistemu zapišemo kot $\overline{c m} = u_{stara}\mathbf{i} + v_{stara}\mathbf{j}$, zaradi spremembe velikosti enot na novem koordinatnem sistemu pa velja $\mathbf{i} = s\mathbf{I}$ in $\mathbf{j} = s\mathbf{J}$, kjer je

⁴ \mathbf{D} je točka preseka omenjenega optičnega žarka in ravnine v neskončnosti [9] (stran 25); po principu dualnosti v \mathcal{P}^3 (projektivnem prostoru) točke in ravnine predstavimo na enak način [9] (stran 23).

$$\mathbf{s} = \begin{bmatrix} k_u & 0 \\ 0 & k_v \end{bmatrix}.$$

Če $\overrightarrow{\mathbf{c}_n \mathbf{c}}$ označimo kot vektor translacije \mathbf{t} , lahko enačbo (2.9) zapišemo kot:

$$\begin{aligned} \overrightarrow{\mathbf{c}_n \mathbf{m}} &= \overrightarrow{\mathbf{c}_n \mathbf{c}} + \overrightarrow{\mathbf{c} \mathbf{m}} = \overrightarrow{\mathbf{c}_n \mathbf{c}} + u_{stara} \mathbf{i} + v_{stara} \mathbf{j} = \\ &= \overrightarrow{\mathbf{c}_n \mathbf{c}} + u_{stara} \mathbf{s} \mathbf{I} + v_{stara} \mathbf{s} \mathbf{J} = \mathbf{t} + u_{stara} \mathbf{s} \mathbf{I} + v_{stara} \mathbf{s} \mathbf{J}, \end{aligned}$$

kar je v projektivnih koordinatah enako:

$$\tilde{\mathbf{m}}_{nova} = \tilde{\mathbf{H}} \tilde{\mathbf{m}}_{stara}, \quad (2.10)$$

kjer je $\tilde{\mathbf{H}}$ matrika reda 3×3 :

$$\tilde{\mathbf{H}} = \begin{bmatrix} \mathbf{s} & \mathbf{t} \\ \mathbf{0}_2^T & 1 \end{bmatrix}.$$

Ker po enačbi (2.4) velja:

$$\tilde{\mathbf{m}}_{stara} = \tilde{\mathbf{P}}_{stara} \tilde{\mathbf{M}},$$

sledi, da je

$$\tilde{\mathbf{P}}_{nova} = \tilde{\mathbf{H}} \tilde{\mathbf{P}}_{stara}. \quad (2.11)$$

Če označimo koordinati \mathbf{t} z (u_0, v_0) , potem velja:

$$\tilde{\mathbf{P}}_{nova} = \begin{bmatrix} -fk_u & 0 & u_0 & 0 \\ 0 & -fk_v & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}. \quad (2.12)$$

Za konec pa, zaradi nadaljnjih izpeljav, definirajmo še normaliziran koordinatni sistem kamere, ki nam omogoča normalizacijo koordinat slikovne ravnine. Ta koordinatni sistem nam pride prav pri stereo aplikacijah (poglavje 3), saj nam omogoča, da lahko ignoriramo specifične lastnosti kamer in gledamo na njo, kot da je idealna. Glede na enačbo (2.12), spremenimo slikovni koordinatni sistem tako, da lahko $\tilde{\mathbf{P}}$ zapišemo kot:

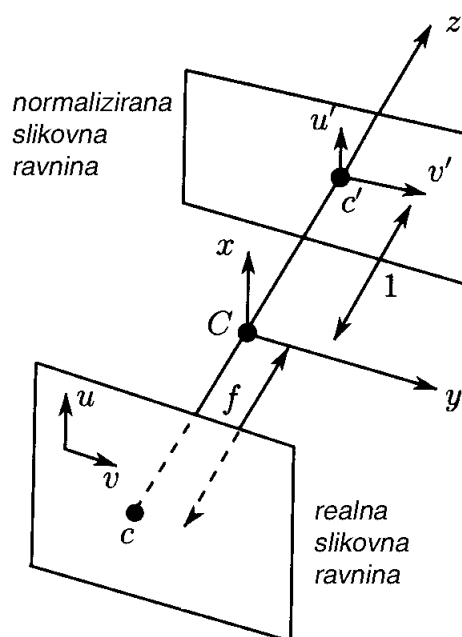
$$\tilde{\mathbf{P}}_{nova} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}. \quad (2.13)$$

Glede na enačbi (2.11) in (2.12) ima $\tilde{\mathbf{H}}$ sedaj naslednjo obliko:

$$\tilde{\mathbf{H}} = \begin{bmatrix} \frac{1}{-fk_u} & 0 & \frac{u_0}{fk_u} \\ 0 & \frac{1}{-fk_v} & \frac{v_0}{fk_v} \\ 0 & 0 & 1 \end{bmatrix}.$$

Po enačbi (2.10) tako dobimo novi slikovni koordinati:

$$\begin{aligned} u' &= \frac{u-u_0}{-fk_u} \\ v' &= \frac{v-v_0}{-fk_v} \end{aligned} \quad (2.14)$$



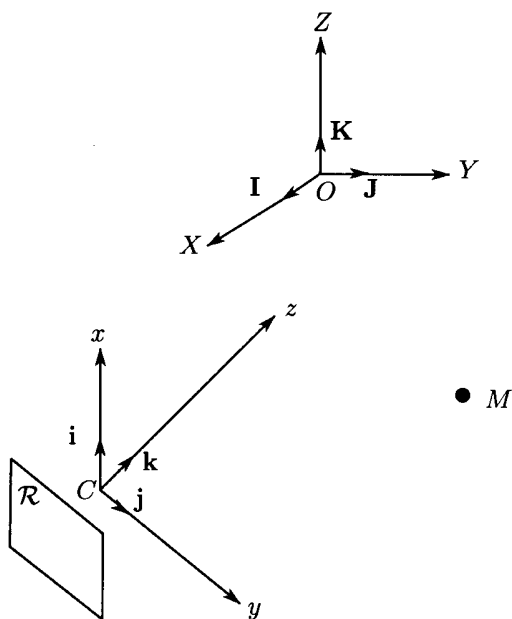
Slika 2.6: Relacija med realno (pravo) in normalizirano slikovno ravnino.

Iz slike 2.6 je razvidno, da ravnina, ki je vzporedna s slikovno ravnino in je za enoto oddaljena od optičnega centra kamere, skupaj z optičnim centrom definira t.i. normalizirano kamero.

Vidimo lahko, da je normalizirana slikovna ravnina na drugi strani optičnega centra C , torej je tukaj slika, v primerjavi z originalno, invertirana.

2.4 Zunanji parametri kamere

S spremembo položaja in orientacije kamere v prostoru dobimo šest zunanjih parametrov kamere, tri za rotacijo \mathbf{R} (matrika velikosti 3×3) (dodatek A) in tri za translacijo \mathbf{T} (vektor velikosti 3×1).



Slika 2.7: Spreminjanje koordinatnega sistema 3D prostora.

Kot je razvidno iz slike 2.7, gremo iz starega koordinatnega sistema z izhodiščem v optičnem centru C v nov koordinatni sistem z izhodiščem v točki O z rotacijo \mathbf{R} in translacijo $\mathbf{T} = \overrightarrow{CO}$. Glede na notacijo slike in po podobnem principu kot v primeru notranjih parametrov, velja:

$$\overrightarrow{CM} = \overrightarrow{CO} + \overrightarrow{OM}.$$

Če izrazimo \overrightarrow{OM} v novem koordinatnem sistemu, dobimo:

$$\overrightarrow{OM} = x_{nova}\mathbf{I} + y_{nova}\mathbf{J} + z_{nova}\mathbf{K},$$

ob upoštevanju rotacije \mathbf{R} iz starega koordinatnega sistema v novi koordinatni sistem, pa velja:

$$\mathbf{I} = \mathbf{R}\mathbf{i}, \quad \mathbf{J} = \mathbf{R}\mathbf{j}, \quad \mathbf{K} = \mathbf{R}\mathbf{k}.$$

Če sedaj zapišemo vektor \overrightarrow{CO} v starem koordinatnem sistemu kot $\mathbf{T} = [t_x, t_y, t_z]^T$, velja:

$$\begin{aligned}
\overline{\mathbf{CM}} &= \overline{\mathbf{CO}} + \overline{\mathbf{OM}} = \mathbf{T} + \overline{\mathbf{OM}} = \mathbf{T} + x_{nova}\mathbf{I} + y_{nova}\mathbf{J} + z_{nova}\mathbf{K} = \\
&= \mathbf{T} + x_{nova}\mathbf{Ri} + y_{nova}\mathbf{Rj} + z_{nova}\mathbf{Rk} \\
&= \mathbf{T} + \mathbf{R} \begin{bmatrix} x_{nova} \\ y_{nova} \\ z_{nova} \end{bmatrix}.
\end{aligned}$$

Na podlagi zgornje enačbe v homogenih koordinatah dobimo:

$$\tilde{\mathbf{M}}_{stara} = \tilde{\mathbf{K}}\tilde{\mathbf{M}}_{nova}, \quad (2.15)$$

kjer je $\tilde{\mathbf{K}}^5$ matrika reda 4×4 :

$$\tilde{\mathbf{K}} = \begin{bmatrix} \mathbf{R} & \mathbf{T} \\ \mathbf{0}_3^T & 1 \end{bmatrix}.$$

Računski primer vpliva zunanjih parametrov kamere na 3D točko je podan v dodatku B.

Po enačbi (2.4) in enačbi (2.15) velja:

$$\tilde{\mathbf{m}} = \tilde{\mathbf{P}}_{stara}\tilde{\mathbf{M}}_{stara} = \tilde{\mathbf{P}}_{stara}\tilde{\mathbf{K}}\tilde{\mathbf{M}}_{nova},$$

torej velja, da je

$$\tilde{\mathbf{P}}_{nova} = \tilde{\mathbf{P}}_{stara}\tilde{\mathbf{K}}.$$

2.5 Sklep

Zapišimo enačbo, ki pove, kako se spreminja matrika perspektivne projekcije $\tilde{\mathbf{P}}$, ko prestavimo koordinatni sistem slikovne ravnine in/ali 3D prostora:

$$\tilde{\mathbf{P}}_{nova} = \tilde{\mathbf{H}}\tilde{\mathbf{P}}_{stara}\tilde{\mathbf{K}}.$$

⁵ $\tilde{\mathbf{K}}$ je matrika, katere determinanta je različna od nič in definira linearno transformacijo (kolineacijo) iz \mathcal{P}^3 v \mathcal{P}^3 . [9] (strani 10 in 26)

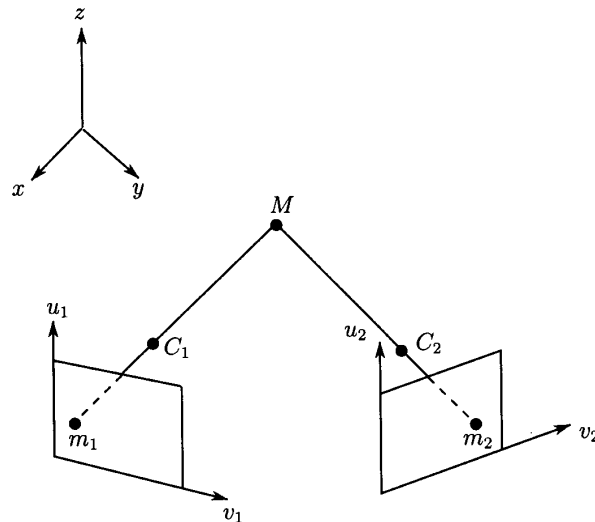
Poglavje 3

Stereo vid

3.1 Uvod

To poglavje je po večini povzeto po knjigi Oliverja Faugerasa: *Three-Dimensional Computer Vision: A Geometric Viewpoint* [9].

Stereo vid predstavlja zelo pomemben praktičen problem računalniškega vida: rekonstrukcija 3D koordinat točk na sceni iz dveh (ali več) slik, zajetih s kamerama, za kateri lahko poznamo relativno lokacijo in orientacijo v prostoru¹ (kalibrirani kameri [9] (stran 51), nekalibrirani kameri [1]).



Slika 3.1: Problem stereo računalniškega vida.

Osnovni problem stereo vida je prikazan na sliki 3.1. Točkasti kameri tvorita sliki

¹Poznamo lahko matriko $\tilde{\mathbf{P}}$, torej tudi notranje in zunanje parametre obeh kamer.

\mathcal{R}_1 in \mathcal{R}_2 (slikovni ravnini), v katerih se fizična točka M preslika v točki m_1 in m_2 . Opraviti imamo s tremi koordinatnimi sistemi: dvema slikovnim koordinatnim sistemoma $((u_1, v_1)$ in $(u_2, v_2))$ in 3D koordinatnim sistemom $((x, y, z))$. Glede na formirani sliki \mathcal{R}_1 in \mathcal{R}_2 , želimo rešiti naslednja dva problema [9]:

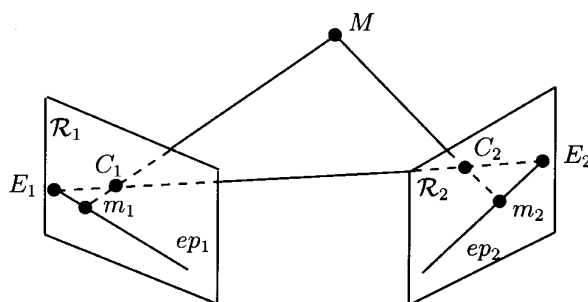
Problem korespondence - Za točko m_1 v slikovni ravnini \mathcal{R}_1 najdi korespondenčno točko m_2 v slikovni ravnini \mathcal{R}_2 . Korespondenca pomeni, da točki m_1 in m_2 predstavljata sliko iste fizične točke M .

Problem rekonstrukcije - Na podlagi dane točke m_1 in točke m_2 izračunaj 3D koordinato fizične točke M .

Točka m_1 iz slikovne ravnine \mathcal{R}_1 ima možnost biti v korespondenci s katerokoli točko m_2 iz slikovne ravnine \mathcal{R}_2 . Zato uporabimo določene omejitve, s katerimi zmanjšamo število potencialnih ujemanj za dano točko m_1 . Za nadaljne razumevanje snovi si bomo v nadaljevanju ogledali epipolarno omejitve.

Rešitev problema rekonstrukcije pa lahko po sliki 3.1 zapišemo kot presečišče premic, ki sta definirani s točkama C_1 in m_1 ter C_2 in m_2 .² Seveda je rezultat močno odvisen od natančnosti kalibracijskega postopka sistema kamer, torej kako natančno poznamo vrednosti notranjih parametrov kamer, lokacijo 3D točk C_1 in C_2 (optičnih centrov) in slikovnih ravnin \mathcal{R}_1 in \mathcal{R}_2 v 3D koordinatnem prostoru. O rekonstrukciji 3D točk bomo več spregovorili v poglavju 3.6.

3.2 Epipolarna omejitve



Slika 3.2: Epipolarna geometrija.

Iz slike 3.2 je razvidno, da vse možne 3D fizične točke M , ki ustvarijo dano slikovno točko m_1 , ležijo na neskončnem poltraku, ki je definiran s točkama m_1 in

²Kako vemo, da ti dve premici nista mimobežni? Dokaz se skriva v epipolarni omejitvi, predstavljeni v naslednjem poglavju; bralec naj bo pozoren na pojem epipolarna ravnina točke M . [9] (izrek na strani 24)

C_1 . Posledica tega je, da vse možne korespondenčne točke m_2 iz slikovne ravnine \mathcal{R}_2 točki m_1 ležijo na sliki tam, kamor se ta poltrak preslika. Na sliki \mathcal{R}_2 se ta preslikan poltrak imenuje epipolarna premica ep_2 točke m_1 , ki gre skozi točko E_2 . E_2 se imenuje epipol in predstavlja presečišče premice, ki je definirana s točkama C_1 in C_2 ter slikovne ravnine \mathcal{R}_2 . Na točko E_2 slikovne ravnine \mathcal{R}_2 lahko gledamo tudi kot na preslikavo optičnega centra C_1 prve kamere na slikovno ravnino \mathcal{R}_2 .

Ustrezna omejitev tako pravi, da ob dani točki m_1 iz slikovne ravnine \mathcal{R}_1 , možne korespondenčne točke na slikovni ravnini \mathcal{R}_2 ležijo na premici. Tako smo zmanjšali preiskovalni prostor iz dveh dimenzij na eno. Epipolarna omejitev je seveda simetrična. Epipolarni premici ep_1 in ep_2 dobimo kot rezultat presečišča slikovnih ravnin \mathcal{R}_1 in \mathcal{R}_2 z ravnino, ki je definirana s točkami C_1 , C_2 in M . To ravnino imenujemo epipolarna ravnina točke M .

Velikokrat zaradi enostavnosti obravnave sistema predpostavljamo, da sta slikovni ravnini \mathcal{R}_1 in \mathcal{R}_2 vzporedni s premico, ki je definirana s točkama C_1 in C_2 . Na ta način gresta oba epipola v neskončnost, epipolarni premici ep_1 in ep_2 pa sta vzporedni. Splošen primer izračuna epipolarne premice je podan v naslednjem poglavju, tukaj pa si oglejmo, kako je le-ta določena v tem posebnem primeru: Kot rečeno sta presečišči epipolarne ravnine točke M in slikovnih ravnin \mathcal{R}_1 in \mathcal{R}_2 epipolarni premici ep_1 in ep_2 , ki sta vzporedni. Torej je epipolarna premica v tem primeru določena že zgolj z eno točko (enačba (3.1)) in omenjenim dejstvom vzporednosti s premico, ki je definirana s točkama C_1 in C_2 . Pri delu v normaliziranem koordinatnem sistemu (poglavje 3.4) pa vemo, da je translacija dovoljena le po osi x in osi y , ne pa tudi po osi z (sicer ne velja več pravilo vzporednosti s premico, ki je definirana s točkama C_1 in C_2), rotacija pa je (iz istega vzroka) omejena le na os z .

3.3 Računanje epipolarne geometrije

Po enačbi (2.4) velja:

$$\tilde{\mathbf{m}}_1 = \tilde{\mathbf{P}}_1 \tilde{\mathbf{M}} \text{ in } \tilde{\mathbf{m}}_2 = \tilde{\mathbf{P}}_2 \tilde{\mathbf{M}} .$$

Kot sledi iz enačbe (2.6), dobimo koordinati obeh optičnih centrov C_i ($i = 1, 2$) v 3D prostoru tako, da rešimo naslednja sistema treh linearnih enačb:

$$\tilde{\mathbf{P}}_i \tilde{\mathbf{M}} = \mathbf{0} , \quad i = 1, 2 .$$

Epipole E_i izračunamo po enačbi:

$$\tilde{\mathbf{e}}_i = \tilde{\mathbf{P}}_i \begin{bmatrix} \mathbf{C}_j \\ 1 \end{bmatrix} , \quad i, j = 1, 2 , \quad i \neq j ,$$

saj je epipol E_i perspektivna preslikava optičnega centra kamere j na slikovno ravnino i . Z $\tilde{\mathbf{e}}_i$ v tem primeru označimo epipol E_i , zapisan v homogenih koordinatah projektivne ravnine \mathcal{P}^2 .

Sedaj pa si oglejmo, kako lahko izračunamo epipolarno premico ep_2 , če poznamo slikovno točko m_1 . Premica je definirana z dvema točkama:

- Epipolom E_2 , za katerega po enačbi (2.4) in enačbi (2.7) velja:

$$\tilde{\mathbf{e}}_2 = \tilde{\mathbf{P}}_2 \tilde{\mathbf{C}}_1 = \tilde{\mathbf{P}}_2 \begin{bmatrix} -\mathbf{P}_1^{-1} \tilde{\mathbf{p}}_1 \\ 1 \end{bmatrix} .$$

- Drugo točko ($\tilde{\mathbf{m}}_2$) dobimo s pomočjo točke v neskončnosti $\tilde{\mathbf{D}}$ optičnega žarka (ta je določen s slikovnim elementom m_1 (v prostoru) in optičnim centrom C_1), torej točko, ki jo dobimo s presekom optičnega žarka z ravnino v neskončnosti (stran 9). Glede na enačbi (2.4) in (2.8), dobimo točko m_2 , ki je preslikava točke D_1 na drugo slikovno ravnino, po naslednji enačbi:

$$\tilde{\mathbf{m}}_2 = \tilde{\mathbf{P}}_2 \tilde{\mathbf{M}} = \tilde{\mathbf{P}}_2 \tilde{\mathbf{D}}_1 = \mathbf{P}_2 \mathbf{D}_1 = \mathbf{P}_2 \mathbf{P}_1^{-1} \tilde{\mathbf{m}}_1 . \quad (3.1)$$

Projektivna predstavitev ep_2 je sedaj, po izreku 1 (stran 8), enaka vektorskemu produktu $\tilde{\mathbf{e}}_2 \times \tilde{\mathbf{m}}_2$.

Če vektorju $\tilde{\mathbf{e}}_2$ priredimo matriko $\tilde{\mathbf{E}}_2$, za katero velja:

$$\tilde{\mathbf{E}}_2 \mathbf{y} = \tilde{\mathbf{e}}_2 \times \mathbf{y} , \text{ za vsak vektor } \mathbf{y} ,$$

potem lahko $\tilde{\mathbf{e}}_2 \times \tilde{\mathbf{m}}_2$ po enačbi (3.1) zapišemo kot:

$$\tilde{\mathbf{e}}_2 \times \tilde{\mathbf{m}}_2 = \tilde{\mathbf{E}}_2 \tilde{\mathbf{m}}_2 = \tilde{\mathbf{E}}_2 \mathbf{P}_2 \mathbf{P}_1^{-1} \tilde{\mathbf{m}}_1 = \mathbf{F} \tilde{\mathbf{m}}_1 . \quad (3.2)$$

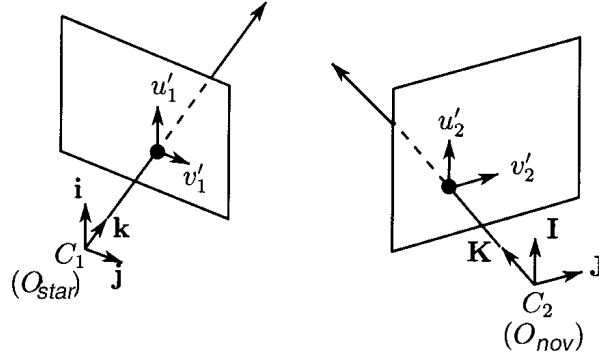
Elementi matrike $\tilde{\mathbf{E}}_2$ so pri tem razvidni iz naslednje enačbe:

$$\tilde{\mathbf{e}}_2 = \begin{bmatrix} a \\ b \\ c \end{bmatrix} \Rightarrow \tilde{\mathbf{E}}_2 = \begin{bmatrix} 0 & -c & b \\ c & 0 & -a \\ -b & a & 0 \end{bmatrix} . \quad (3.3)$$

Vsak slikovni element m_2 , ki leži na epipolarni premici ep_2 , katero določa slikovni element m_1 , tako zadovoljuje enačbo³ (epipolarna omejitev):

$$\tilde{\mathbf{m}}_2^T \mathbf{F} \tilde{\mathbf{m}}_1 = 0 , \quad (3.4)$$

ki se imenuje Longuet-Higginsova enačba in jo bomo s pridom uporabili v nadaljevanju.



Slika 3.3: Dve kameri in njuni normalizirani slikovni ravnini.

3.4 Delo v normaliziranem koordinatnem sistemu

Če uporabimo normaliziran model kamere (poglavje 2.3) in predpostavimo relacije iz slike 3.3, se opis epipolarne geometrije poenostavi. Uporabimo notacijo, ki smo jo uporabili v poglavju 2.4: $\mathbf{T} = \overrightarrow{C_1 C_2}$, $\mathbf{R}[\mathbf{i}, \mathbf{j}, \mathbf{k}] = [\mathbf{I}, \mathbf{J}, \mathbf{K}]$, matrika $\tilde{\mathbf{K}}$ pa je enaka⁴:

$$\tilde{\mathbf{K}} = \begin{bmatrix} \mathbf{R} & \mathbf{T} \\ \mathbf{0}_3^T & 1 \end{bmatrix}.$$

$\tilde{\mathbf{P}}_1$ zapišemo kot:

$$\tilde{\mathbf{P}}_1 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}.$$

Matrika $\tilde{\mathbf{P}}_2$ v novem koordinatnem sistemu C_2XYZ je enaka $\tilde{\mathbf{P}}_1$.⁵ Matrika $\tilde{\mathbf{P}}_2$ izražena v koordinatnem sistemu C_1xyz pa je enaka⁶:

³Natančnejša razlaga je podana v poglavju 3.5.

⁴Če bi smer translacije in rotacije obrnili ($\mathbf{T}_2 = \overrightarrow{C_2 C_1}$, $\mathbf{R}_2[\mathbf{I}, \mathbf{J}, \mathbf{K}] = [\mathbf{i}, \mathbf{j}, \mathbf{k}]$), bi za matriko $\tilde{\mathbf{K}}_2$ veljalo:

$$\tilde{\mathbf{K}}_2 = \begin{bmatrix} \mathbf{R}_2 & \mathbf{T}_2 \\ \mathbf{0}_3^T & 1 \end{bmatrix} = \begin{bmatrix} \mathbf{R}^T & -\mathbf{R}^T \mathbf{T} \\ \mathbf{0}_3^T & 1 \end{bmatrix} = \tilde{\mathbf{K}}^{-1}.$$

⁵Smo v normaliziranem koordinatnem sistemu; notranji parametri kamere ostajajo enaki, spremembe zunanjih parametrov pa tukaj ne upoštevamo, kar pomeni, da je optični center kamere C_2 v izhodišču 3D koordinatnega sistema.

⁶Podobno velja tudi za matriko $\tilde{\mathbf{P}}_1$, ki jo izrazimo v koordinatnem sistemu C_2xyz :

$$\tilde{\mathbf{P}}_1 = \tilde{\mathbf{P}}_2 \tilde{\mathbf{K}} = [\mathbf{R} \quad \mathbf{T}].$$

$$\tilde{\mathbf{P}}_1 = \tilde{\mathbf{P}}_2 \tilde{\mathbf{K}} \Rightarrow \tilde{\mathbf{P}}_2 = \tilde{\mathbf{P}}_1 \tilde{\mathbf{K}}^{-1} = \begin{bmatrix} \mathbf{R}^T & -\mathbf{R}^T \mathbf{T} \end{bmatrix} .$$

Sedaj lahko izračunamo koordinati epipolov E_1 in E_2 :

$$\tilde{\mathbf{e}}_1 = \tilde{\mathbf{P}}_1 \tilde{\mathbf{C}}_2 = \tilde{\mathbf{P}}_1 \begin{bmatrix} \overrightarrow{\mathbf{C}_1 \mathbf{C}_2} \\ 1 \end{bmatrix} = \tilde{\mathbf{P}}_1 \begin{bmatrix} \mathbf{T} \\ 1 \end{bmatrix} = \mathbf{T} ,$$

$$\tilde{\mathbf{e}}_2 = \tilde{\mathbf{P}}_2 \tilde{\mathbf{C}}_1 = \tilde{\mathbf{P}}_2 \begin{bmatrix} \mathbf{0}_3 \\ 1 \end{bmatrix} = -\mathbf{R}^T \mathbf{T} .$$

Če slikovni element predstavimo z $\tilde{\mathbf{m}}_1 = [u'_1, v'_1, 1]^T$, je epipolarna premica ep_2 , glede na zgornje enačbe in enačbo (3.2), sedaj predstavljena z enačbo:

$$\tilde{\mathbf{e}}_2 \times \tilde{\mathbf{m}}_2 = \tilde{\mathbf{e}}_2 \times \mathbf{P}_2 \mathbf{P}_1^{-1} \tilde{\mathbf{m}}_1 = -\mathbf{R}^T \mathbf{T} \times \mathbf{R}^T \tilde{\mathbf{m}}_1 ,$$

kar je projektivno⁷ enako:

$$\mathbf{R}^T (\mathbf{T} \times \tilde{\mathbf{m}}_1) .$$

Podobno velja tudi za $\tilde{\mathbf{m}}_2 = [u'_2, v'_2, 1]^T$, kjer je epipolarna premica ep_1 predstavljena z enačbo:

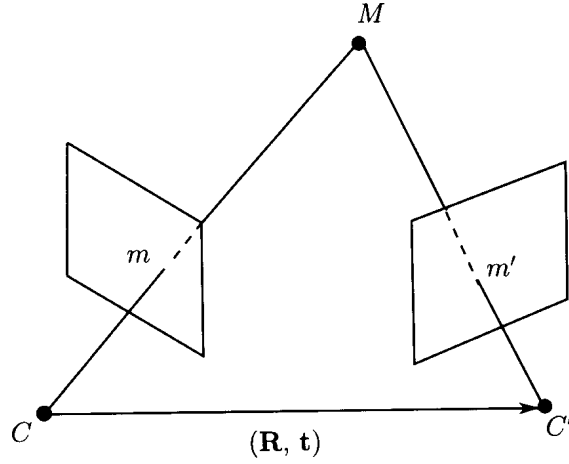
$$\tilde{\mathbf{e}}_1 \times \tilde{\mathbf{m}}_1 = \tilde{\mathbf{e}}_1 \times \mathbf{P}_1 \mathbf{P}_2^{-1} \tilde{\mathbf{m}}_2 = \mathbf{T} \times \mathbf{R} \tilde{\mathbf{m}}_2 . \quad (3.5)$$

3.5 Določanje odmika kamere na podlagi korespondence točk

Predpostavimo, da imamo opravka z normaliziranim koordinatnim sistemom, kar torej pomeni, da moramo kalibrirati kamero in izračunati njene notranje parametre. Ti definirajo transformacijo slikovnih koordinat v normalizirane slikovne koordinate. S transformacijo dobimo normalno kamero z enotsko goriščno razdaljo f in koordinatni sistem $Cxyz$ z izhodiščem v optičnem centru kamere (slika 2.6). Na premik kamere lahko tako gledamo kot na premik tega koordinatnega sistema. Odmik D_i iz koordinatnega sistema ene kamere v koordinatni sistem druge kamere, kar je ekvivalentno premikanju kamere iz lokacije in orientacije ob času t_i v lokacijo in orientacijo ob času t_{i+1} , opišemo kot rotacijo (\mathbf{R}) okoli osi, ki gre skozi C_i , kateri sledi translacija $\mathbf{t} = \overrightarrow{C_i C_{i+1}}$.

Kot je razvidno iz slike 3.4, se 3D točka M preslika v korespondenčni slikovni točki m in m' . Sedaj nas zanima, kakšna omejitev velja za \mathbf{R} in \mathbf{t} ? Iz slike je razvidno, da sta točki m in m' v korespondenci le v primeru, če so prostorski vektorji $\mathbf{m} = \overrightarrow{C\mathbf{m}}$, $\mathbf{m}' = \overrightarrow{C'\mathbf{m}'}$ in $\mathbf{t} = \overrightarrow{CC'}$ koplanarni. Če želimo to omejitev zapisati v obliki

⁷Uprabljammo homogene koordinate.



Slika 3.4: Omejitev parametrov gibanja, ki jih dobimo iz korespondenčnih točk.

enačbe, moramo te vektorje najprej prestaviti v skupni koordinatni sistem. Za vektor \mathbf{m}' , ki je izražen v koordinatnem sistemu druge kamere, velja, da je enak vektorju $\mathbf{R}\mathbf{m}'$ v koordinatnem sistemu prve kamere (poglavje 2.4). V koordinatnem sistemu prve kamere lahko torej omejitev zapišemo kot⁸:

$$\mathbf{m} \cdot (\mathbf{t} \times \mathbf{R}\mathbf{m}') = 0 .$$

Na podlagi zgornje enačbe in enačbe (3.5) lahko epipolarno omejitev v koordinatnem sistemu prve kamere zapišemo kot:

$$\tilde{\mathbf{m}}_1 \cdot (\mathbf{t} \times \mathbf{R}\tilde{\mathbf{m}}_2) = 0 .$$

Če vektorju \mathbf{t} priredimo matriko \mathbf{T} , za katero velja isto kot za matriko v enačbi (3.3), potem lahko omejitev zapišemo kot (enačba (3.4)):

$$\tilde{\mathbf{m}}_1^T \mathbf{T} \mathbf{R} \tilde{\mathbf{m}}_2 = \tilde{\mathbf{m}}_1^T \mathbf{E} \tilde{\mathbf{m}}_2 = 0 . \quad (3.6)$$

Osnovna matrika \mathbf{E} (ang. *essential matrix*)⁹ predstavlja osnovo linearne metode za izračun odmika D_i , ki se imenuje *Algoritem osmih točk* (ang. *The eight-point algorithm*) [9]. Ta ob znanih korespondenčnih točkah na izhodu poda \mathbf{R} in \mathbf{t} ¹⁰.

⁸Enačbo dobimo na podlagi dveh pravil: rezultat vektorskega produkta je vektor, ki je pravokoten na ravnino, katero tvorita ta dva vektorja, skalarni produkt dveh pravokotnih vektorjev pa je enak nič.

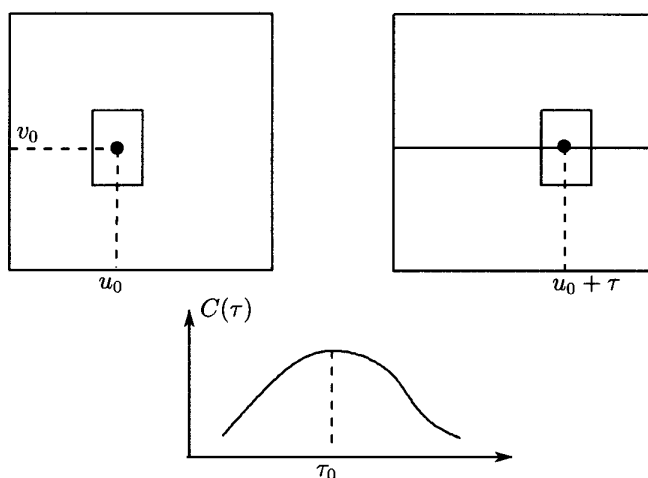
⁹Poznamo tudi pojem *temeljna matrika* (ang. *fundamental matrix*), katere vloga je identična vlogi osnovne matrike \mathbf{E} , le da temeljna matrika deluje v enotah slikovnih elementov, osnovna matrika pa v enotah prostora. [1]

¹⁰V nadaljevanju predpostavljamo, da je $\mathbf{t} \neq \mathbf{0}$ ([9], stran 249).

Preden pa si ogledamo podrobno to metodo, pa na kratko naprej opišimo, kako poiščemo ustrezne korespondenčne točke.

3.5.1 Iskanje korespondenčnih točk

Veliko je raziskav korelacijske tehnike intenzitetnih področij slik za komercialne aplikacije stereofotogrametrije, hkrati pa le-te predstavljajo ene najstarejših metod računalniškega vida.



Slika 3.5: Princip korelacijske metode za iskanje korespondenčnih točk.

Princip korelacijske metode za iskanje korespondenčnih točk je razviden iz slike 3.5. Da bi našli koordinati slikovnega elementa na drugi sliki, ki ustrežata koordinati (u_0, v_0) slikovnega elementa na prvi sliki, računamo korelacijo $C_{12}(\tau)$ pravokotnega okna¹¹ velikosti $(2P+1) \times (2N+1)$, katerega središče je (u_0, v_0) , z drugo intenzitetno sliko vzdolž vrstice ($v_2 = v_0$). Pri tem seveda upoštevamo izračunano epipolarno premico. Krivulja $C_{12}(\tau)$ ima navadno en maksimum, ki ga dosežemo v točki τ_0 . Dispariteto¹² slikovnega elementa (u_0, v_0) tako izenačimo s τ_0 .

Ena izmed pomankljivosti te metode je predpostavka, da so epipolarne premice vrstice slike [9] (stran 191). Če vhodne slike ne ustrezajo tej zahtevi, moramo slike predhodno procesirati s postopkom izravnave (ang. *rectification*) [9] (stran 188).

Ker za razumevanje nadaljnje snovi tega poglavja sama predstavitev enačbe korelacijske metode ni potrebna, bomo enačbo zapisali v poglavju 5.5.1, ko bomo govorili o sistemu za gradnjo globinskih panoramskih slik.

¹¹To okno imenujemo korelacijsko okno.

¹²Definicija disparitete v primeru vzporednih kamer: ob danem slikovnem elementu m_1 s koordinatama (u_1, v_1) v prvi slikovni ravnini in korespondenčnim slikovnim elementom m_2 s koordinatama (u_2, v_2) v drugi slikovni ravnini, je dispariteta definirana kot razlika $v_2 - v_1$.

3.5.2 Algoritem osmih točk

Na podlagi dejstva, da je enačba (3.6) linearna in homogena glede na koeficiente matrike \mathbf{E} , sta Longuet in Higgins razvila metodo, s pomočjo katere lahko najdemo te koeficiente, če imamo na voljo osem takšnih enačb.

Metoda je sestavljena iz dveh korakov:

- ocena matrike \mathbf{E}
- izračun \mathbf{t} in \mathbf{R} iz \mathbf{E} .

Ocena matrike \mathbf{E}

Zapišimo \mathbf{X} kot vektor velikosti 9×1 :

$$\mathbf{X} = [\mathbf{e}_1^T, \mathbf{e}_2^T, \mathbf{e}_3^T]^T,$$

kjer je \mathbf{e}_i i -ta vrstica matrike \mathbf{E} .

Sedaj lahko enačbo (3.6) zapišemo kot:

$$\mathbf{a}^T \mathbf{X} = 0,$$

kjer je \mathbf{a} vektor velikosti 9×1 ¹³:

$$\mathbf{a} = [x_1 \tilde{\mathbf{m}}_2^T \quad y_1 \tilde{\mathbf{m}}_2^T \quad \tilde{\mathbf{m}}_2^T]^T.$$

Če imamo n korespondenčnih parov točk, nam vsak par da enačbo tipa $\mathbf{a}^T \mathbf{X} = 0$.

Ker je \mathbf{E} , podobno kot \mathbf{t} , definirana le do multiplikativnega faktorja natančno, imamo opravka le z osmimi neznankami. Torej je najmanjša vrednost za n , s katero lahko ocenimo \mathbf{X} , $n = 8$ (in od tod ime algoritma).

Enačbe tipa $\mathbf{a}^T \mathbf{X} = 0$ lahko združimo:

$$\mathbf{A}_n \mathbf{X} = \mathbf{0}, \tag{3.7}$$

kjer je \mathbf{A}_n matrika velikosti $n \times 9$:

$$\mathbf{A}_n = \begin{bmatrix} \mathbf{a}_1^T \\ \vdots \\ \mathbf{a}_n^T \end{bmatrix}.$$

Če predpostavljamo, da obstaja v 3D prostoru množica točk M_1, \dots, M_8 , za katere velja, da je \mathbf{A}_8 reda 8, potem lahko iz enačbe (3.7) dobimo vrednosti koeficientov \mathbf{X} . Tako dobimo oceno matrike \mathbf{E} .

¹³Ker imamo opravka z normaliziranim koordinatnim sistemom, je $\tilde{\mathbf{m}}_1 = [x_1, y_1, 1]^T$ in $\tilde{\mathbf{m}}_2 = [x_2, y_2, 1]^T$.

Izračun \mathbf{t} in \mathbf{R} iz \mathbf{E}

Naj bodo \mathbf{c}_i in \mathbf{r}_i , $i = 1, 2, 3$, vektorji kolon matrik \mathbf{E} in \mathbf{R} . Iz enačbe $\mathbf{E} = \mathbf{TR}$ (po enačbi (3.6)) sledi, da je $\mathbf{c}_i = t \times \mathbf{r}_i$ za vsak i . Iz tega sledi, da so vsi trije kolonski vektorji \mathbf{E} pravokotni na \mathbf{t} , torej je \mathbf{t} vzporeden z vektorskim produktom dveh \mathbf{c}_i . To dejstvo določi \mathbf{t} do predznaka natančno:

$$\mathbf{t} = \pm \mathbf{c}_1 \times \mathbf{c}_2 .$$

Če poznamo predznak ene od komponent \mathbf{t} (na primer t_3), poznamo tudi predznak \mathbf{t} .

Iz relacije $\mathbf{c}_i = t \times \mathbf{r}_i$ lahko sedaj izrazimo:

$$\mathbf{c}_i \times \mathbf{c}_{i+1} = (\mathbf{r}_{i+2} \cdot \mathbf{t}) \mathbf{t} , \quad i = 1, 2, 3 ,$$

kjer je $\mathbf{r}_{i+2} = \mathbf{r}_{((i+1) \bmod 3)+1}$. Matrika kofaktorjev \mathbf{E}^* matrike \mathbf{E} je dana kot:

$$\begin{aligned} \mathbf{E}^* &= [\mathbf{c}_2 \times \mathbf{c}_3, \mathbf{c}_3 \times \mathbf{c}_1, \mathbf{c}_1 \times \mathbf{c}_2]^\top = \\ &= [(\mathbf{r}_1 \cdot \mathbf{t}) \mathbf{t}, (\mathbf{r}_2 \cdot \mathbf{t}) \mathbf{t}, (\mathbf{r}_3 \cdot \mathbf{t}) \mathbf{t}]^\top = \\ &= (\mathbf{t}(\mathbf{R}^\top \mathbf{t})^\top)^\top = (\mathbf{t} \mathbf{t}^\top \mathbf{R})^\top . \end{aligned}$$

Na podlagi naslednjega izreka

IZREK 2¹⁴

Za matriko $\mathbf{E} = \mathbf{TR}$ velja, da je matrika $\mathbf{E} \mathbf{E}^\top$ odvisna le od translacijskega vektorja \mathbf{t} .

velja:

$$\mathbf{TE} = \mathbf{T}^2 \mathbf{R} = \mathbf{t} \mathbf{t}^\top \mathbf{R} - (\mathbf{t} \cdot \mathbf{t}) \mathbf{R} ,$$

od tod pa sledi:

$$(\mathbf{t} \cdot \mathbf{t}) \mathbf{R} = \mathbf{E}^{*\top} - \mathbf{TE} ,$$

kar nam da \mathbf{R} v odvisnosti od \mathbf{E} in \mathbf{t} .

3.6 Rekonstrukcija 3D točk

Oglejmo si linearno metodo rekonstrukcije:

Predvidevamo, da imamo dve sliki iz dveh kamer, na katerih smo našli korepondenčni točki m_1 in m_2 , ki imata koordinati (u_1, v_1) in (u_2, v_2) .

Če s p_{ij} označimo elemente (znane) matrike $\hat{\mathbf{P}}$, dobimo iz enačbe (2.4) naslednje enačbe:

¹⁴Dokaz si lahko ogledate v [9], na strani 250.

$$\begin{aligned} U &= p_{11}X + p_{12}Y + p_{13}Z + p_{14}T \\ V &= p_{21}X + p_{22}Y + p_{23}Z + p_{24}T \\ S &= p_{31}X + p_{32}Y + p_{33}Z + p_{34}T . \end{aligned}$$

Za prehod iz projektivnih koordinat v evklidske koordinate upoštevamo enačbo (2.3) in izberemo $X = x$, $Y = y$, $Z = z$ in $T = 1$. Dobimo:

$$\begin{aligned} u(p_{31}x + p_{32}y + p_{33}z + p_{34}) &= p_{11}x + p_{12}y + p_{13}z + p_{14} \\ v(p_{31}x + p_{32}y + p_{33}z + p_{34}) &= p_{21}x + p_{22}y + p_{23}z + p_{24} , \end{aligned}$$

kar lahko zapišemo kot:

$$\begin{aligned} (p_{11} - up_{31})x + (p_{12} - up_{32})y + (p_{13} - up_{33})z + p_{14} - up_{34} &= 0 \\ (p_{21} - vp_{31})x + (p_{22} - vp_{32})y + (p_{23} - vp_{33})z + p_{24} - vp_{34} &= 0 . \end{aligned}$$

V dobljenih enačbah poznamo koeficiente p_{ij} ter slikovne koordinate (u, v) . Ko v enačbi vstavimo vrednosti p_{ij_1} , p_{ij_2} ¹⁵, (u_1, v_1) in (u_2, v_2) , dobimo sistem štirih linearnih enačb s tremi neznankami. Rešitev tega sistema nam da iskane 3D koordinate (x, y, z) točke, katere rekonstrukcijo smo iskali.

Posplošitev postopka na n slik je očitna.

V dodatku C je podana skica računskega primera celotnega postopka rekonstrukcije 3D točke.

¹⁵Matriki $\tilde{\mathbf{P}}_1$ in $\tilde{\mathbf{P}}_2$ morata biti izraženi v istem koordinatnem sistemu (stran 18).

Poglavje 4

Panoramski pogledi in stereo rekonstrukcija

4.1 Uvod

Standardna kamera ima omejen vertikalni in horizontalni zorni kot, ki sta navadno manjša od človekovega vidnega polja. Že zaradi tega si je človek vedno prizadeval, da bi ustvaril sliko, ki bi obsegala večji zorni kot, celo večjega od človekovega vidnega polja [24, 25, 34]. Takšne slike imenujemo panoramske slike (slika 4.2b).

Pod pojmom *stereo rekonstrukcija* razumemo gradnjo globinskih slik iz več slik. Postopek stereo rekonstrukcije temelji na relacijah med točkami in linijami na sceni in slikah le-te. Slike scene lahko nastopajo v parih, trojicah, četvericah itd., prav na podlagi števila slik pa imenujemo tudi omenjene relacije: epipolarna omejitve, trislikovna omejitve, štirislikovna omejitve¹ itd. [11]. Zelo zaželjena lastnost slik je, da so te točke ali linije dobro vidne na vseh slikah neke scene. In takšno lastnost nam v največji meri zagotavljajo panoramske kamere.

Na tem mestu opišimo dva pojma, ki opisujeta dva osnovna tipa panoramskih kamer: *klasična* in *ne-klasična panoramska kamera* [31]. Za klasično panoramsko kamero velja, da se vsi svetlobni žarki, ki tvorijo sliko, sekajo v eni točki. Za ne-klasično kamero pa velja, da se ti žarki ne sekajo v eni točki, pač pa se lahko sekajo na premici, lahko predstavljajo tangente na neko krožnico ipd. V klasični kameri nastane slika s pomočjo perspektivne projekcije oziroma *središčne projekcije* (ang. *central projection*), zato se velikokrat kot sinonim klasični kameri uporablja pojem *središčna kamera* (ang. *central camera*), kot sinonim ne-klasični kameri pa pojem *ne-središčna kamera* (ang. *non-central camera*).

V nadaljevanju bomo pokazali, da nekaterih metod za zajemanje panoramskih slik ne moremo modelirati s središčno projekcijo [3, 20, 27, 31]. Veliko standardnih algoritmov iz področja računalniškega vida (na primer epipolarna omejitve) pa

¹Le za omenjene (tri) omejitve pa obstaja linearna rešitev rekonstrukcije.

zahteva uporabo središčne projekcije. V prihodnosti lahko pričakujemo raziskave o teoriji ne-središčnih kamer, katerih cilj bo poiskati posplošitev teh algoritmov, torej njihovo uporabo tudi v sistemih z ne-središčnimi kamerami [2, 3].

Panoramske slike lahko ustvarimo s pomočjo panoramskih kamer, ali s pomočjo standardne, ne-panoramske kamere, kjer delamo mozaike standardnih slik. Osredotočili se bomo predvsem na pristope, ki vodijo k 3D rekonstrukciji, še posebej pa na mozaike, ki nas še posebej zanimajo. Preden pa si ogledamo posamezne pristope, si oglejmo še nekaj pojmov, ki so pomembni za nadaljne razumevanje snovi [31]: Sinonimi za središčno kamero so tudi *kamera s središčem projekcije* in *kamera z enim glediščem* (ang. *single (effective) viewpoint*). V primeru standardne kamere je gledišče v optičnem centru kamere (slika 2.2). Glede na zorni kot kamere, delimo le-te na: *usmerjene*, *panoramske* in *vsesmerno*. Kamera je vsesmerna, če ima celotno vidno polje, kar pomeni, da slika zajeta s takšno kamero pokriva celotno površino krogle (ang. *view-sphere*). Kamera je usmerjena, če je njeno vidno polje ustrezna podmnožica polkrogle na površini krogle. Kamera je panoramska, če njeno vidno polje vsebuje vsaj eno krožnico na površini krogle. Glede na konstrukcijo kamere ločimo: *dioptrične* in *katadioptrične kamere*. Sestavni del dioptričnih kamer so leče, katadioptričnih kamer pa vsaj eno ogledalo, lahko pa tudi leče. Navadnim kameram bomo rekli *standardne kamere*, ki so središčne, usmerjene in dioptrične kamere, kar lahko z drugimi besedami opišemo kot točkaste kamere, katerih vidno polje je del polkrogle.

V naslednjem poglavju so predstavljene panoramske kamere s poudarkom na 3D rekonstrukciji. Sledi analiza rezultatov panoramskih kamer.

4.2 Panoramske kamere

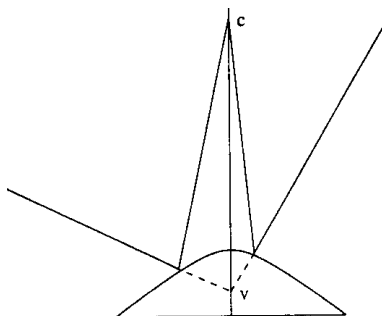
Kriterij razdelitve panoramskih kamer v spodnja poglavja je način ustvarjanja panoram: ustvarjanje panoram s pomočjo ogledal(a), leče, pomičnih delov in mozaičenja.

4.2.1 Katadioptrične kamere

Katadioptrična kamera je sestavljena iz ogledala in kamere, ki zajema sliko, odbito od tega ogledala [31]. Najenostavnejša katadioptrična kamera je sestavljena iz kamere in ravninskega ogledala, seveda pa takšna sestava ne omogoča zajemanje panoramskih slik. Že delno ravninska ogledala pa povečajo zorni kot kamere.

Nekatera ogledala izpolnjujejo pomemben pogoj: imajo eno *efektivno gledišče* (ang. *single effective viewpoint*). Svetlobni žarki, ki tvorijo sliko, se namreč odbijajo od ogledala proti kameri, ki te žarke obdela. Če se ti žarki ne bi odbili, bi se sekali v točki, ki ji pravimo efektivno gledišče (slika 4.1). Ta lastnost nam omogoča modeliranje sistema s središčno projekcijo.

Vendar pa ta omejitev velja le za določene oblike ogledal in kar je še pomembnejše, le za določene postavitve kamere glede na ogledalo [31]. Na primer za sferično



Slika 4.1: Svetlobni žarki, ki se sekajo v efektivnem gledišču v , se odbijejo v optični center kamere c .

ogledalo je ta omejitev izpolnjena le, če se kamera nahaja v središču sfere. Podoben problem nastane pri uporabi stožčastega ogledala, saj bi v tem primeru kamera morala biti pritrjena na vrhu stožca. Edina ogledala, ki s praktičnega vidika zadovoljujejo to omejitev, so: ravninska, eliptična, parabolična in hiperbolična ogledala, vendar le zadnja dva povečata zorni kot kamere. Na sliki 4.2 je prikazan primer panoramske slike, zgrajene s pomočjo katadioptrične kamere².

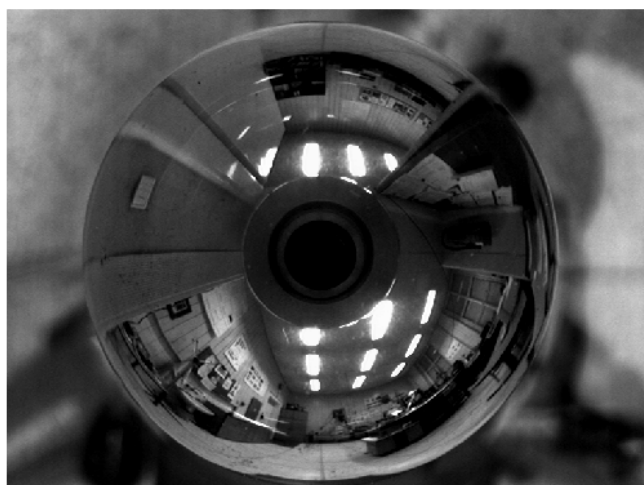
Vsa ostala ogledala nimajo omenjene lastnosti, zato imamo pri njih opravka z ne-središčnimi kamerami.

Da si bomo lažje predstavljali kompleksnost stereo rekonstrukcije v katadioptričnem sistemu, si oglejmo primer: Svoboda s sodelavci [31, 32] je predstavil epipolarno geometrijo za primer hiperbolične katadioptrične panoramske kamere, torej za sistem z dvema standardnima kamerama in dvema hiperboličnima ogledaloma. Epipolarna ravnina π seka ogledali v obliki stožčastega preseka (slika 4.3), ki se preslikajo v stožce na slikovni ravnini s pomočjo središčne projekcije. Za vsako točko \mathbf{u}_1 na eni sliki leži korespondenčna točka na drugi sliki nekje na epipolarnem stožcu.

4.2.2 Dioptrične kamere

Dioptrične kamere uporabljajo posebne leče, ki povečajo zorni kot kamere. Primer takšnih leč so ribje oči (ang. *fish-eye*). Kamero s širokokotnimi lečami lahko uporabimo za 3D rekonstrukcijo, ko upoštevamo parametre izkrivljanja [37]. Seveda pa velja, da takšne kamere niso prave panoramske kamere, saj je njihov zorni kot še vedno dosti manjši od 360° . Bralec si lahko več o širokokotnih lečah in panoramskih pogledih prebere v [4].

²Slika 4.2 je vzeta iz baze katadioptričnih slik Laboratorija za računalniški vid, Fakulteta za računalništvo in informatiko, Univerza v Ljubljani in laboratorija CMP (Center for Machine Perception), Fakulteta za elektrotehniko, Tehniška univerza v Pragi. Slika prikazuje laboratorij CMP. [14]



a)



b)

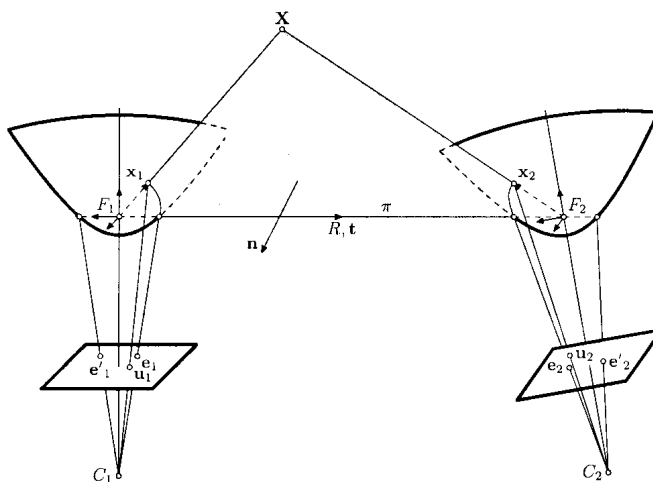
Slika 4.2: a) vhodna slika, zajeta s katadioptrično kamero (uporabljeno je sferično ogledalo), b) panoramska slika po opravljeni transformaciji vhodne slike na valj.

4.2.3 Panoramski pogledi s pomočjo pomičnih delov

Če senzor (kamero) premikamo po neki poti, lahko sestavimo panoramski pogled. Tudi postopki mozaičenja standardnih slik v panoramske poglede spadajo v to skupino postopkov. Ker pa nas postopki mozaičenja še posebej zanimajo, jih bomo obravnavali v posebnem poglavju.

V [5] je opisan postopek gradnje panoramskih slik s pomočjo *linijske kamere* (ang. *line-scan camera*), ki se vrti okoli vertikalne osi. 3D rekonstrukcija je možna na podlagi uporabe dveh takšnih kamer, ki sta postavljeni ena nad drugo in imata skupno os vrtenja.

Linijska kamera je 1D tabela CCD senzorjev. Sinonim za njo je tudi *trakovna*



Slika 4.3: Epipolarna geometrija hiperbolične katadioptrične panoramske kamere [31, 32].

kamera³ (ang. *slit camera*). Če premikamo takšno tabelo po ravni liniji s konstantno hitrostjo dobimo efekt *linearne potisne kamere* (ang. *linear pushbroom camera*) [10]. Ta kamera je poenostavitev modela potisne kamere, ki se uporablja za satelitsko zajemanje slik. Takšna kamera je seveda ne-središčna, saj se svetlobni žarki, ki tvorijo sliko, ne sekajo v eni točki. Takšne kamere torej ne moremo točno aproksimirati z modelom točkaste kamere.

4.2.4 Mozaičenje standardnih slik v panoramske poglede

Če želimo ustvariti prave mozaične panoramske poglede, moramo kamero seveda pomikati po neki zaključeni poti, ki je največkrat krožnica. Sama ideja mozaičenja slik v panoramske poglede je že zelo stara⁴: princip gradnje panoram iz le dela zajete slike, pri čemer kamero vrtimo okoli optičnega centra, so integrirali v fotoaparati že v 19. stoletju.

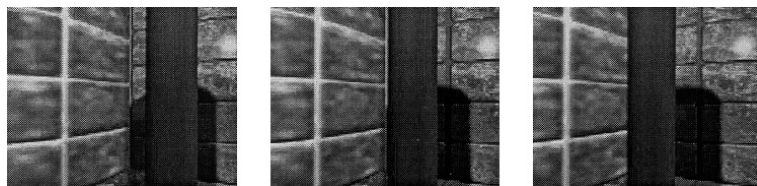
Eden izmed najbolj znanih komercialnih paketov za ustvarjanje mozaičnih panoramskih pogledov je QTVR (QuickTime Virtual Reality)⁵. Ta deluje na principu šivanja standardnih slik, ki so zajete s pomočjo rotirajoče se kamere [8]. Peleg s sodelavci [19, 22] je predstavil metodo ustvarjanja mozaičnih panoramskih pogle-

³Včasih se pojem trakovna kamera uporablja tudi v primeru, ko iz slike, zajete s standardno kamero, uporabimo le določen trak slikovnih elementov (tipično nekaj kolon slikovnih elementov), vendar v tem primeru pojem linijska kamera ne predstavlja sinonima tako definirani trakovni kameri.

⁴Bill McBride: Panoramic Cameras Time Line 1843-1994, <http://www.panphoto.com/TimeLine.html>.

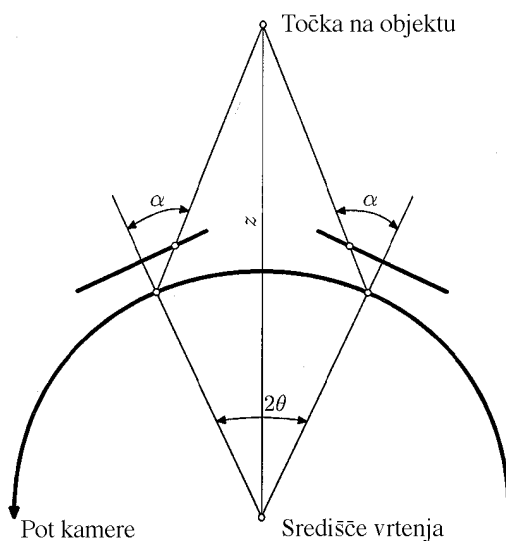
⁵<http://www.qtvrworld.com/>

dov iz standardnih slik, zajetih s poljubno gibajočo se kamero. Podobno metodo sta predlagala tudi Szeliski in Shum [35], ki pravtako ne omejuje gibanja oziroma postopka zajemanja slik, predpostavlja pa, da ni velike paralakse gibanja⁶. Seveda pa omenjene metode služijo zgolj za vizualizacijo, saj ni prisotnih poskusov po rekonstrukciji scene.



Slika 4.4: Primer horizontalne paralakse gibanja [27].

Ishiguro s sodelavci [13] je predlagal metodo, ki omogoča rekonstrukcijo scene. Uporabili so standardno kamero, ki rotira na krožnici, katere polmer ne sme biti nič. Optične osi kamere na vsaki lokaciji te krožnice se sekajo v njenem središču. Globina objekta se oceni na podlagi projekcije v dve sliki, ki sta bili zajeti na različnih lokacijah kamere na krožnici (slika 4.5).



Slika 4.5: Globino z točke na objektu lahko ocenimo z meritvijo kotov θ in α [13].

Peleg in Ben-Ezra [20, 21, 23] sta predstavila metodo za ustvarjanje stereo panoram. Stereo panorame pa so ustvarjene brez dejanskega računanja 3D strukture

⁶Paralaksa gibanja je efekt, ki povzroči spremembo lokacije opazovanega objekta, ker se spremeni lokacija človeka oziroma instrumenta, ki opazuje objekt (slika 4.4).

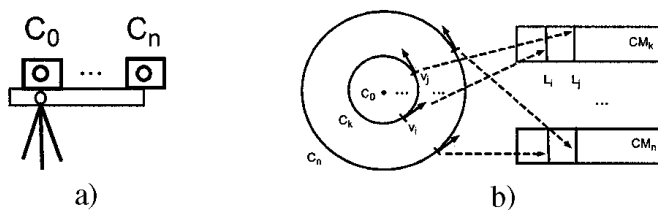
— efekt globine se ustvari le v možganih opazovalca.

Shum in Szeliski [27] sta predstavila dva postopka za ustvarjanje globinskih panoramskih slik in sicer oba postopka izrabljata standardne postopke za 3D rekonstrukcijo. Oba postopka temeljita na gibanju kamere po krožnici. Panorame sestavljata tako, da iz vsake zajete slike vzameta le eno ali nekaj kolon slikovnih elementov (kar je ekvivalentno uporabi linijske oziroma trakovne kamere), ki doprinesejo k panorami. Te panorame imenujeta *večperspektivne panorame* (ang. *multiperspective panorama*). Pomembna lastnost (dveh ali več) večperspektivnih panoram je, da hranijo informacijo o paralaksi gibanja, saj so njihove kolone zajete iz različnih perspektiv. Za samo stereo rekonstrukcijo nato uporabita tako dobljene panorame.

Večperspektivne panorame same po sebi niso nekaj novega [27]: so poseben primer *večperspektivnih panoram za animacijo celic* (ang. *multiperspective panoramas for cel animation*) [38], so zelo podobne slikam narejenim s postopkom *večsrediščna projekcija* (ang. *multiple-center-of-projection*) [26], postopkom *mnogotera projekcija* (ang. *manifold projection*) [19, 22] in postopkom *krožna projekcija* (ang. *circular projection*) [20, 21, 23]. Princip je zelo podoben tudi postopku gradnje panorame z linearno potisno kamero [10].

Oglejmo si sistema za zajemanje panoram obeh pristopov Shuma in Szeliskega [27]:

- Prvi pristop temelji na uporabi več standardnih kamer, ki so pritrjene na rotacijsko mizo (roko) (slika 4.6a). Ekvivalenten pristop lahko temelji tudi na eni kameri, ki pa jo moramo ustrezno premikati, preden rotiramo mizo. Večperspektivna panorama nastane iz središčnih trakov (kolon) zajetih slik na vsaki lokaciji rotacijske mizice, kot je to razvidno iz slike 4.6b. Tako dobljene panorame avtorja imenujeta *koncentrične panorame*.

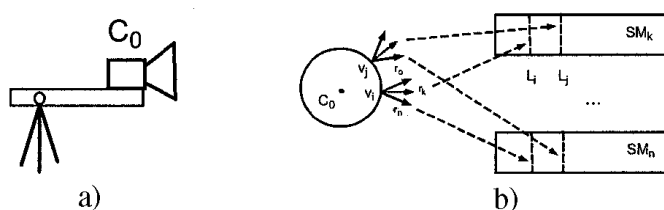


Slika 4.6: Koncentrične panorame: a) sistem za zajemanje, b) proces ustvarjanja večperspektivnih panoram. [27]

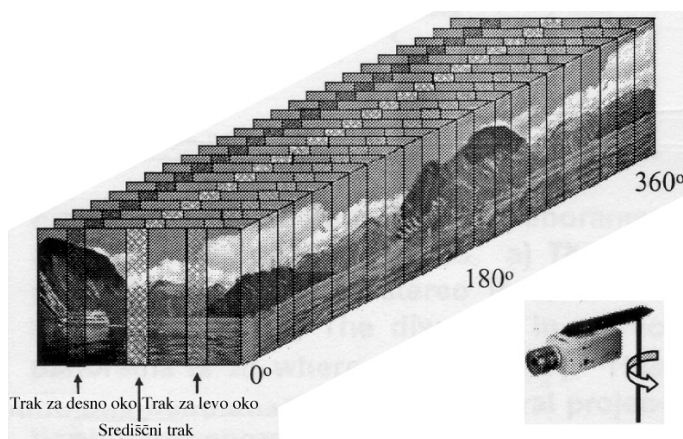
- Drugi pristop temelji na vrtenju standardne kamere, ki gleda navzven, torej stran od središča vrtenja (slika 4.7a). V tem primeru večperspektivne panorame nastanejo iz različnih trakov (kolon) zajetih slik, kot je to razvidno iz slike 4.7b in slike 4.8. Tako dobljene panorame avtorja imenujeta *rotacijske panorame* (ang. *swing panorama*). Peleg in Ben-Ezra [20, 21, 23] imenujeta opisan

postopek ustvarjanja panoram *krožna projekcija* (ang. *circular projection*). Tukaj bi lahko uporabili tudi bolj splošen pristop: uporabimo več kamer, pritrjenih na rotacijsko roko, ki gledajo vsaka v svojo smer [27].

Podoben pristop k zajemanju slik sta uporabila tudi Huang in Pajdla [12], le da je tokrat kamera obrnjena navznoter, torej proti središču vrtenja. Njun cilj je bil izpeljati enačbo epipolarne krivulje, ki bi na podlagi dane točke (slikovnega elementa) v eni panorami, izrisala ustrezno krivuljo na drugi panorami.



Slika 4.7: Rotacijske panorame: a) sistem za zajemanje, b) proces ustvarjanja večperspektivnih panoram. [27]



Slika 4.8: Gradnja stereo (valjčnih) panoram s šivanjem trakov iz zaporednih slik: če uporabimo trakove iz središča slike, dobimo navadno panoramo, če uporabimo trakove iz leve strani slike, dobimo panoramo za desno oko in če uporabimo trakove iz desne strani slike, dobimo panoramo za levo oko. Efekt globine se ustvari v možganih opazovalca. [20, 21, 23]

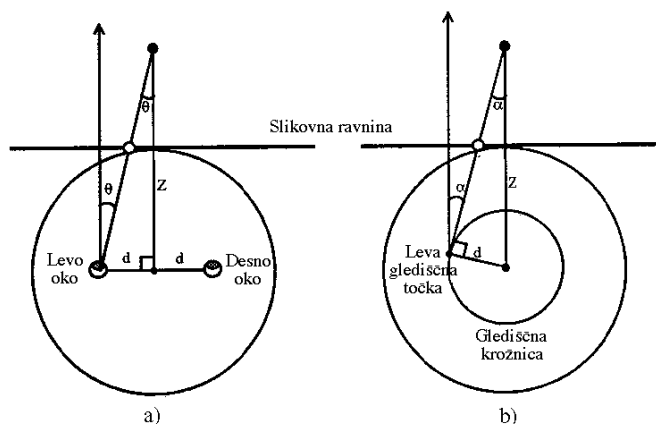
Na sliki 4.9 je predstavljena primerjava disparitete med standardnim stereo parom in stereo parom (valjčnih) panoram [20]. Medtem ko je dispariteta standardnega sterea 2θ , kjer je

$$\theta = \arctan\left(\frac{d}{Z}\right),$$

je dispariteta stereo panoram 2α , kjer je

$$\alpha = \arcsin\left(\frac{d}{Z}\right).$$

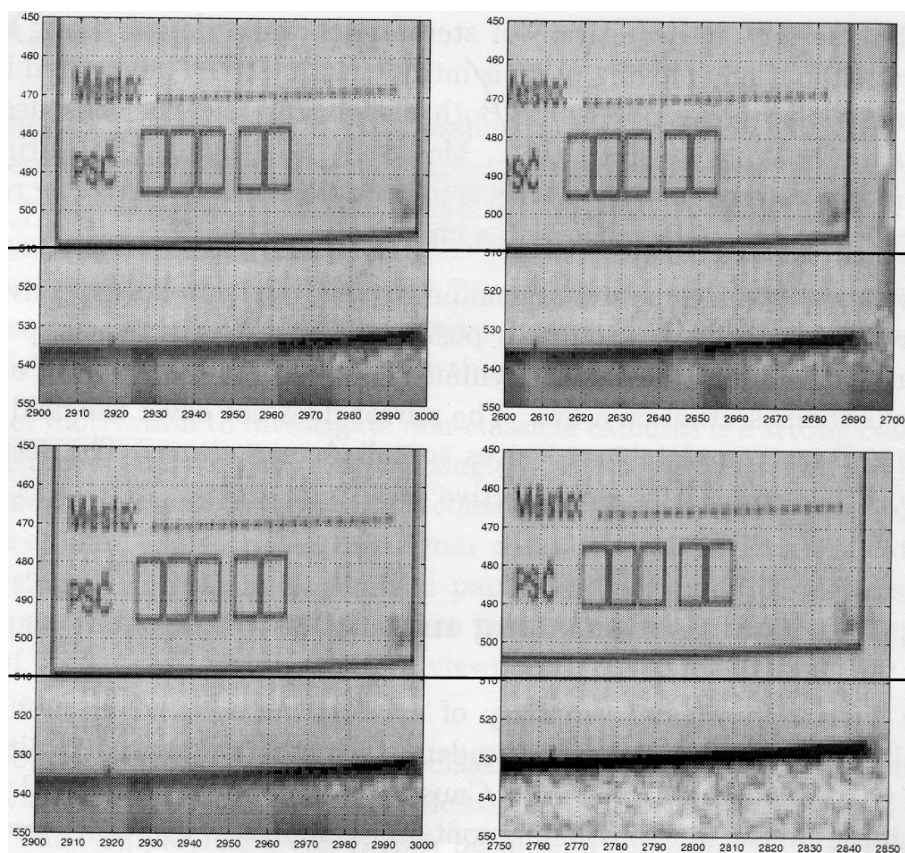
Ta dispariteta se, za opazovalca v središču valja, s središčno projekcijo panorame na ravnino ohranja, s tem pa se ohranja efekt globine.



Slika 4.9: Primerjava disparitete med: a) standardnim stereo parom in b) stereo parom panoram [20].

Omejitvi, ki sta si ju zadala pri svojem delu, torej gibanje po krožnici in uporaba le dela zajete slike, pa omogočata enostavno iskanje korespondence (uporabimo lahko tradicionalne postopke), saj se izkaže, da so epipolarne premice kar horizontalne premice oziroma vrstice zajete panorame [12, 27]. Ta trditev velja za *simetrične pare stereo (rotacijskih) panoram*. Simetričen par panoram dobimo, če vzamemo simetrične trakove levo in desno od sredine zajete slike (torej na enaki razdalji od sredine zajete slike), ali če pritrdimo dve kameri na enaki razdalji od središča vrtenja, vendar na nasprotnih koncih rotacijske roke (slika 4.10). Ti dve panorami imata prav zaradi horizontalne epipolarne geometrije lepo vidna horizontalna paralaksa gibanja, vertikalne paralakse gibanja pa ni [28]. Tudi takšne panorame imenujemo rotacijske panorame [27]; Huang in Pajdla [12] jih imenujeta koncentrične panorame. V primeru ne-simetričnih parov panoram se globina izračuna s pomočjo postopka *stereo s preslikavo na valj* (ang. *cylinder sweep stereo*) [27]. Cilj njune rekonstrukcije je povezava globinske slike s panoramskimi pogledi, ki bi omogočali scenarij tipa “poglej okoli in malo se prestavi” — ustvarjene globinske slike omogočajo upodabljanje na osnovi slik, torej interpolacijo in ekstrapolacijo pogledov.

Za oba pristopa velja, da uporabljata ne-središčne kamere: v prvem primeru so svetlobni žarki, ki tvorijo (panoramsko) sliko, tangentni na krožnico, ki predstavlja pot kamere (slika 4.11a), v drugem primeru pa so le-ti tangentni na *glediščno krožnico* (ang. *viewing circle*) (slika 4.11b).



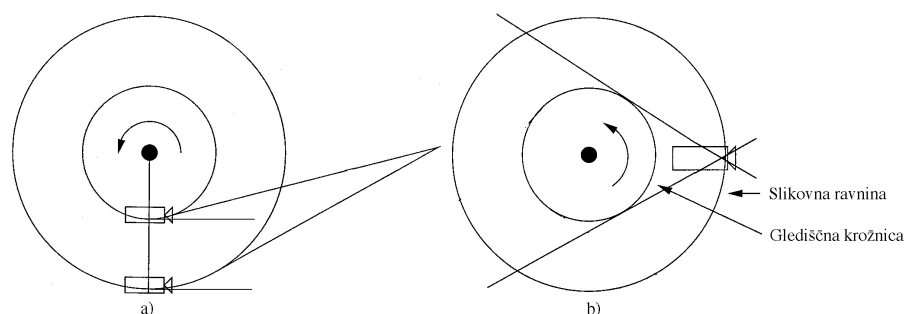
Slika 4.10: Povečan primer korespondenčnih točk v simetričnem paru panoram (zgornja vrstica) in ne-simetričnem paru panoram (spodnja vrstica) [12].

Seveda lahko tudi v primeru ne-simetričnih panoram računamo globino s pomočjo preiskovanja epipolarne premice, ki pa v tem primeru ni vrstica panorame. Torej se postopek rekonstrukcije v primeru ne-simetričnih panoram z vidika časovne kompleksnosti poveča.

4.3 Analiza rezultatov panoramskih kamer

V tabeli 4.1 je prikazana primerjava različnih panoramskih kamer. Pristopi so primerjani glede na število standardnih slik, ki so potrebne za gradnjo panoramske slike, glede na ločljivost dobljene panorame in glede na vprašanje delovanja v realnem času.

Zaključimo lahko, da imajo le panoramske kamere z enim ogledalom eno gledišče [31]. Vsi pristopi omogočajo 3D rekonstrukcijo. Mozaične pristope lahko označimo



Slika 4.11: Uporaba ne-središčnih kamer: a) svetlobni žarki, ki tvorijo (panoramsko) sliko, so tangenti na krožnico, ki predstavlja pot kamere, b) svetlobni žarki, ki tvorijo (panoramsko) sliko, so tangenti na glediščno krožnico. [2, 3, 12, 27]

Princip gradnje panoramske slike	Število slik	Ločljivost panorame	Realni čas	Literatura
katadioptrična kamera	1	majhna	da	[31, 32]
dioptrična kamera	1	majhna	da	[4, 37]
pomični deli	veliko	velika	ne	[5, 10]
mozaičenje	veliko	velika	ne	[8, 12, 13, 19, 20, 21, 22, 23] [24, 25, 26, 27, 28, 34, 35, 38]

Tabela 4.1: Primerjava različnih panoramskih kamer glede na število standardnih slik, ki so potrebne za gradnjo panoramske slike, glede na ločljivost dobljene panorame in glede na vprašanje delovanja v realnem času.

kot ne-središčne (pristop nima le enega središča projekcije), pristope z dolgotrajnim časom gradnje panoram in kot pristope, ki dajejo visoko ločljive rezultate. Kot takšni niso primerni za zajemanje dinamičnih scen in posledično manj primerni za rekonstrukcijo. Podobno velja tudi za pristope s pomičnimi deli, ki so hitrejši, vendar še vedno ne delujejo v realnem času. Dioptrične pristope s širokokotnimi lečami lahko označimo kot ne-središčne, kot pristope, ki delujejo v realnem času in kot pristope, ki dajejo rezultate majhne ločljivosti. Kamere s širokokotnimi lečami so primerne za hitro zajemanje panoramskih slik in njihovo procesiranje, na primer za iskanje ovir, ali za lokalizacijo mobilnega robota, niso pa najbolj primerne za rekonstrukcijo. Podobno velja tudi za pristope z več ogledali. Pristopi z enim ogledalom delujejo v realnem času in dajejo rezultate z majhno ločljivostjo. Le pristopi s hiperboličnim in paraboličnim ogledalom so središčni. Ti pristopi so primerni za rekonstrukcijo dinamičnih scen v majhni ločljivosti in za oceno gibanja. Zanje tudi velja, da so to

edini pristopi, za katere lahko enostavno posplošimo epipolarno geometrijo [11, 31].

Za konec pa omenimo le še idejo, ki sta jo predstavila Swaminathan in Nayar [33]: Bistven problem pri mozaičenju standardnih slik v panoramske poglede je torej, da postopek ne deluje v realnem času. Omenjena avtorja sta zasnovala sistem, ki vsebuje štiri kamere s horizontalnim zornim kotom okoli 115° . Širokokotne kamere so združene v en strojni del tako, da je optična os ene kamere pravokotna na optično os sosednje kamere, vse optične osi pa ležijo na isti ravnini. Takšen strojni del imenujeta *panoramska polikamera* (ang. *panoramic polycamera*), ki omogoča mozaičenje le štirih slik v panoramski pogled z vidnim poljem 360° v realnem času. Seveda je potrebno v postopku mozaičenja odpraviti tudi popačenje, ki ga povzročijo širokokotne leče. Ker smo še vedno omejeni z velikostjo kamer, v tem primeru ne moremo sestaviti polikamere z enim optičnim centrom, kar pomeni, da lahko delamo lepe mozaike scen, kjer so objekti vsaj za neko minimalno razdaljo oddaljeni od središča polikamere. Avtorja se nista dotaknila problema rekonstrukcije.

Poglavje 5

Sistem za gradnjo globinskih panoramskih slik

5.1 Uvod

Če želimo ustvariti večperspektivne panorame, moramo omejiti gibanje kamere na krožnico v ravnini. Večperspektivno panoramo zgradimo tako, da vedno izberemo isto kolono iz standardnih perspektivnih slik, ki jih zajamemo na posameznih lokacijah na krožnici. Večperspektivne panorame se razlikujejo od standardnih panoram po lastnosti, da večperspektivne panorame (dve ali več) hranijo informacijo o paralaksi gibanja, saj so njihove kolone zajete iz različnih perspektiv (slike 4.6, 4.7 in 4.8). Tako so postavljeni temelji, ki omogočajo 3D rekonstrukcijo (poglavje 4).

5.1.1 Sestavni deli sistema

V osnovi lahko razdelimo sestavne dele sistema na strojne in programske dele.

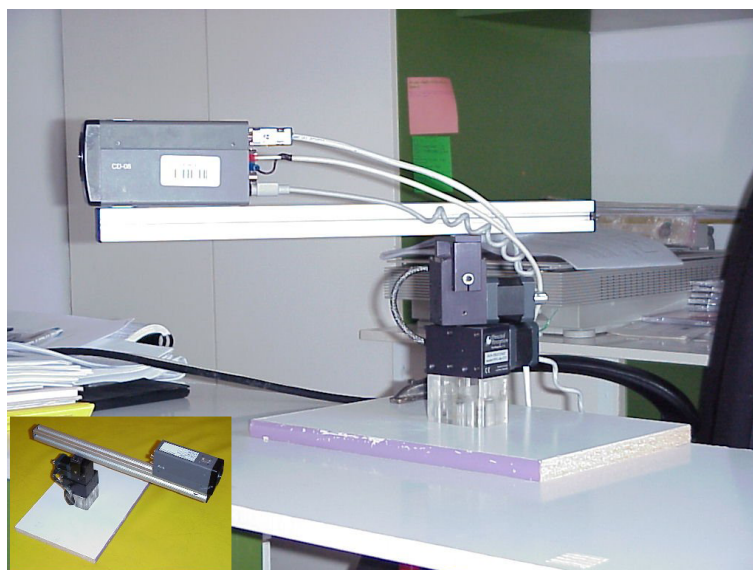
Strojni deli sistema so (slika 5.1):

- rotacijska roka¹, ki se vrti v horizontalni smeri,
- na njo je pritrjen distančnik, ki omogoča premik optičnega centra kamere iz središča vrtenja in
- na ta distančnik je pritrjena ena sama standardna (barvna) kamera², ki gleda navzven, torej stran od središča vrtenja.

Rotacijska roka je z računalnikom povezana preko posebne kontrolne enote, ki z računalnikom komunicira preko komunikacijskega kanala COM1. Komunikacijski kanal COM2 je rezerviran za komunikacijo računalnika in kamere, ki je povezana tudi z zajemalnikom slik preko kanala S-VIDEO.

¹Model PTU-46-17.5 podjetja Directed Perception, Inc.

²Model CD-08 podjetja Chugai Boyeki Co., Ltd.



Slika 5.1: Strojni deli sistema.

Programski deli sistema so:

- modul za krmiljenje rotacijske roke,
- modul za komunikacijo s kamero,
- modul za zajemanje žive slike in
- modul za ustvarjanje (globinskih) panoramskih slik.

Več o programskih delih sistema si lahko bralec prebere v dodatku E.

Opraviti imamo torej s postopkom mozaičenja standardnih slik v (stereo) panoramske poglede in računanjem stereo rekonstrukcije na podlagi vrtenja ene kamere, katere optični center ni v središču vrtenja (slika 4.7).

5.1.2 Omejitve sistema

Osnovne omejitve izvirajo iz omenjene (strojne) sestave našega sistema:

- Kamero lahko okoli vertikalne osi vrtenja obrnemo v vsako stran za 159° .
- Ločljivost rotacijske roke je 185,1428 kotnih sekund na en korak rotacijske roke. To pomeni, da en korak rotacijske roke ustreza zasuku za $0,051428\bar{5}^\circ$.

Ob tem podatku se pojavi vprašanje ponovljivosti rezultatov. V primeru, da je kamera pritrjena v središču vrtenja, se izkaže, da kamera pride na isto mesto,

če opravimo zaporedje različno velikih premikov³, ki predstavljajo zaokroženo pot. To velja vsaj za slike do vključno ločljivosti 640×480 slikovnih elementov. V primeru uporabe našega distančnika ta lastnost velja le v primeru, če je premik rotacijske roke majhen⁴. Ker bomo v nadaljevanju uporabljali premik za 4 korake rotacijske roke, imamo zagotovljeno ponovljivost rezultatov za naš sistem. Razlog za problem pri večjem premiku rotacijske roke je zagotovo nepravilna utežitev tovora rotacijske roke. Problem pa bi najverjetneje bil rešljiv tudi z ustrezno nastavitvijo hitrosti in/ali pospeška premika rotacijske roke.

- Opcija ZOOM na kameri je nastavljena tako, da je horizontalni zorni kot kamere $\alpha = 34^\circ$ in vertikalni zorni kot kamere $\beta = 25^\circ$. Nastavitve kamere se med postopkom zajemanja slik ne spreminjajo.
- Zaradi oblike rotacijske roke, distančnik, ki je pritrjen na roko, ne more biti poljubne (željene) dolžine. Problem je v maksimalni dovoljeni teži tovora, ki je lahko pritrjen na rotacijsko roko (1,81 kg) in uteženju tovora. Teža distančnika, ki omogoča odmik optičnega centra kamere iz središča vrtenja za maksimalno 40 cm, je 0,4 kg. Vendar, ker tovor pri maksimalnem odmiku ni bil pravilno utežen, zajemanje s takšnim odmikom ni priporočljivo; trpi rotacijska roka, tresenje distančnika pa je zelo očitno. Če bi želeli zajemati z odmikom 40 cm, bi po pravilih na primer morali uporabiti distančnik dvakratne dolžine uporabljenega distančnika, kar bi pomenilo težo 0,8 kg. Ker je teža kamere 0,5 kg, moramo na drugi strani distančnika pritrčiti utež enake teže. Skupaj tako dosežemo maksimalno dovoljeno težo tovora, pri čemer je distančnik pritrjen na rotacijsko roko na sredini. V nadaljevanju tega poglavja pa bomo pokazali, da je prav z odmikom optičnega centra kamere od središča vrtenja pogojena maksimalna (smiselna) globina, v katero oceno še lahko verjamemo z veliko stopnjo zaupanja.

Na tem mestu pa zapišimo še dva problema strojnih delov sistema, ki nista bila tako očitna:

- Optika v kameri je rahlo premaknjena okoli optične osi (os, ki je pravokotna na slikovno ravnino).
- Zajemalnik slik prikaže ob levem in zgornjem robu zajete slike temen rob, katerega širina je odvisna od ločljivosti zajete slike. S tem se pojavi vprašanje: Kje je središčna kolona zajete slike? Odgovor na to vprašanje je zelo pomemben, saj, kot bomo videli v nadaljevanju tega poglavja, predstavlja osnovo za

³ 5×100 korakov rotacijske roke v desno, 3×100 korakov rotacijske roke navzdol, 2×300 korakov rotacijske roke v desno, 5×100 korakov rotacijske roke navzgor, 4×50 korakov rotacijske roke navzdol, 2×500 korakov rotacijske roke v levo in 2×50 korakov rotacijske roke v levo.

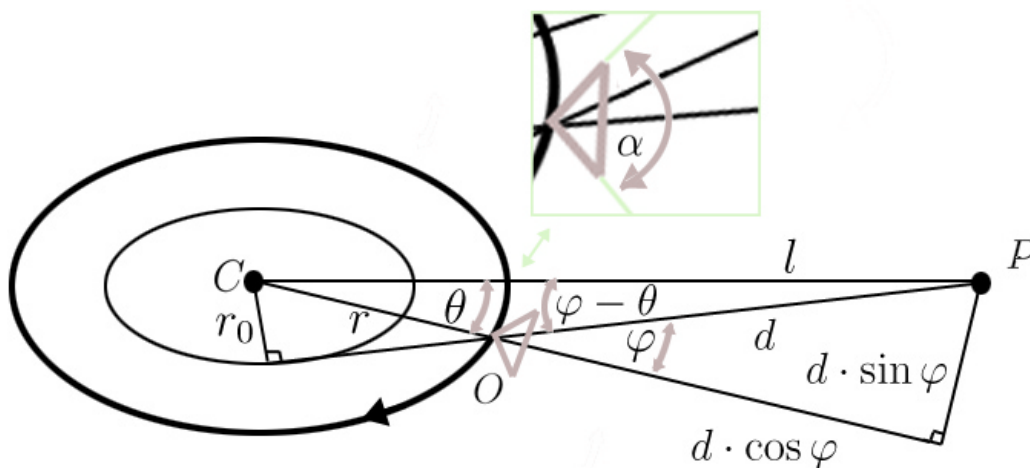
⁴ 20×4 koraki rotacijske roke v desno in 20×4 koraki rotacijske roke v levo.

ustvarjanje simetričnih parov panoram. Za idealni primer velja, da kamerina optična os seka zajeto sliko v središčni koloni, vendar pa to v praksi ni nujno.

Oba problema prideta manj do izraza pri manjši ločljivosti zajetih slik.

5.2 Geometrija sistema

Iz slike 5.2 je razvidna splošna geometrija našega sistema za ustvarjanje večperspektivnih panoramskih slik, ki služijo kot vhod za gradnjo globinskih panoramskih slik. Točka C predstavlja središče vrtenja, okoli katerega se vrtila kamera na oddaljenosti r , ki predstavlja polmer krožnice (pot), po kateri potuje kamera. Kamera je obrnjena navzven, torej stran od središča vrtenja. Optični center kamere se nahaja v točki O . Izbrana kolona slikovnih elementov, ki bo doprinesla k večperspektivni panorami, vsebuje preslikavo točke P na sceni. Ta je od točke C oddaljena za razdaljo l , od točke O pa za razdaljo d . S θ je označen trenutni kot zasuka med daljicama, ki sta definirani s središčem vrtenja C in optičnim centrom O ter središčem vrtenja C in točko na sceni P . θ v panoramski sliki podaja horizontalno os, ki opisuje pot gibanja kamere. S φ pa je označen kot med daljicama, ki sta definirani z optičnim centrom O in sredinsko kolono zajete slike ter optičnim centrom O in izbrano kolono zajete slike, ki bo doprinesla k večperspektivni panorami. Na ta kot lahko gledamo tudi kot na zmanjšanje horizontalnega zornega kota α zajete slike.



Slika 5.2: Geometrija sistema za ustvarjanje ene večperspektivne panorame.

Geometrijo zajemanja večperspektivne panoramske slike lahko torej opišemo s parom parametrov (r, φ) .

Sistem na sliki 5.2 je očitno ne-središčni, saj se svetlobni žarki ($\varphi \neq 0$), ki tvorijo panoramsko sliko, ne sekajo v eni točki, pač pa predstavljajo tangente na krožnico s

polmerom r_0 , ki predstavlja polmer glediščne krožnice. Torej imamo opravka s panoramskimi slikami, ki nastanejo s projekcijo iz več projekcijskih središč, to pa pomeni, da bo določena točka na sceni vidna v panoramski sliki le iz enega projekcijskega središča.

5.3 Ustvarjanje panoramskih slik

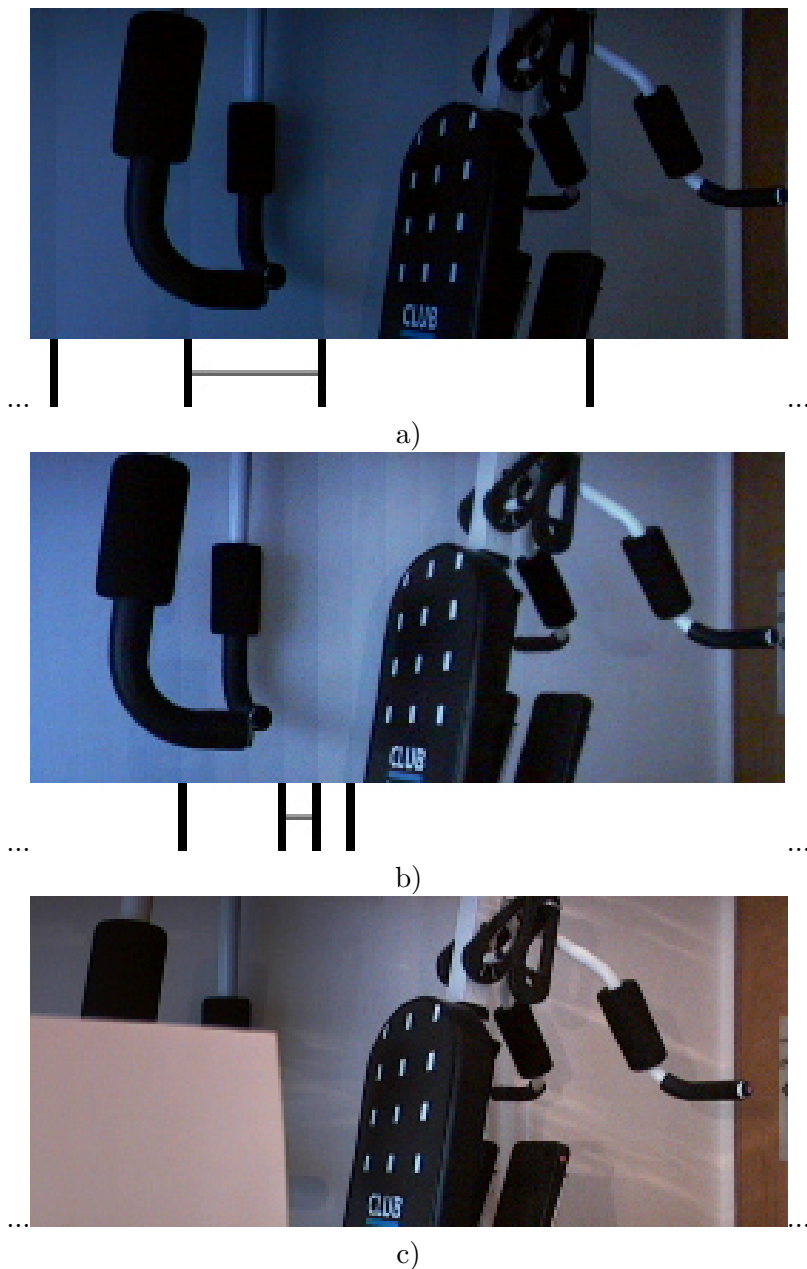
Postopek ustvarjanja panoramskih slik, ki sledi iz geometrije sistema, je lepo razviden iz slike 4.7 in slike 4.8. Podoben je postopkom, ki so jih uporabljali Shum in Szeliski [27], Peleg in Ben-Ezra [20, 21, 23] ter Huang in Pajdla [12].

Panoramske poglede ustvarimo tako, da obračamo kamero in zlagamo zajete slike v panoramsko sliko. Večperspektivne panorame ustvarimo iz različnih trakov (kolon) zajetih slik. Na predlog Shuma in Szeliskega [27] jih poimenujmo *rotacijske panorame* (ang. *swing panoramas*).

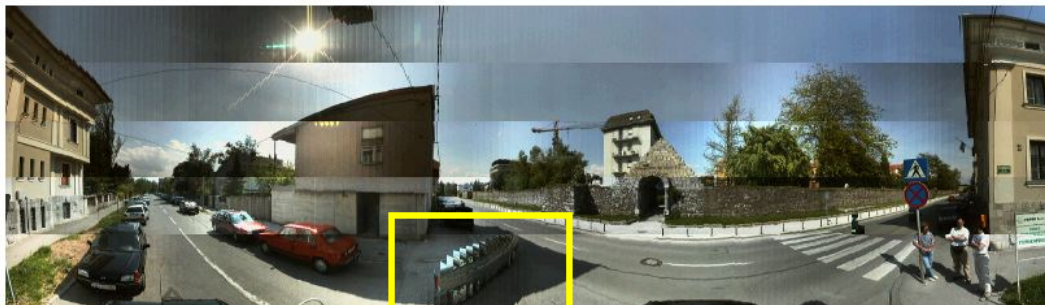
Pri zajemanju slik naletimo na več problemov, ki se jih moramo zavedati pri nadaljni gradnji panoramskih pogledov [24]:

- Intenzitetni zamik korespondenčnih točk. — V idealnem primeru naj bi imele slikovne vrednosti korespondenčnih točk na sosednjih slikah enako vrednost. Toda zaradi različnih vplivov, kot sta na primer nihanje svetlobe v prostoru in samodejna uravnavna svetlosti slike v sami kameri⁵, pride v realnosti do intenzitetnega (barvnega) zamika korespondenčnih točk. Tako v primeru, ko gradimo panoramski pogled iz več kot le ene kolone zajete slike, v ustvarjeni panorami vidimo šive, ki združuje trakove zajetih slik v panoramo (slika 5.3).
- Dinamičnost scen. — V času, ki poteče med zajetjem dveh zaporednih slik, lahko pride do spremembe scene. Objekti na sceni lahko spremenijo lokacijo, obliko, se zasučejo, lahko vstopijo na sceno, ali izginejo z nje. V takšnih primerih se izgubi del korespondence med zaporednimi slikami, kar povzroči popačen panoramski pogled (slika 5.4).
- Radialno popačenje. — Tovrstno popačenje je odvisno od vrste leč kamere in goriščne razdalje le-teh. Radialno popačenje leč na uporabljeni kameri je razvidno iz slike 5.5.
- Intenzitetno popačenje kamere. — Intenziteta slikovnih točk je odvisna od položaja slikovnih točk na sliki. Na sliki 5.6 vidimo 3D prikaz intenzitete slikovnih točk sivinske slike homogene bele površine (bel list papirja), ki nam da lep vpogled v naravo intenzitetnega popačenja uporabljene kamere. Intenziteta slikovnih točk pada z oddaljenostjo od optičnega centra slike. Popačenje

⁵Kamerino samodejno uravnavanje parametrov, ki vplivajo na svetlost slike, je ključnega pomena za kvaliteto dobljene panoramske slike, saj nam omogoča zajemanje detajlov na sceni, na kateri imamo različne svetlobne razmere.



Slika 5.3: Primer združevanja trakov zajetih slik, ki so širši od ene same kolone, v panoramski pogled: a) trak širine 47 kolon, b) trak širine 12 kolon. V obeh primerih so lepo vidni šivi, ki združuje trakove zajetih slik v panoramo. V primeru c), kjer gradimo panoramski pogled iz le ene kolone zajete slike, pa šivi niso vidni.



Slika 5.4: Primer popačenja panoramskega pogleda zaradi dinamičnosti scene: s pravokotnikom je označeno popačenje, ki ga je povzročil avtomobil, ki se je med snemanjem peljal mimo [24].

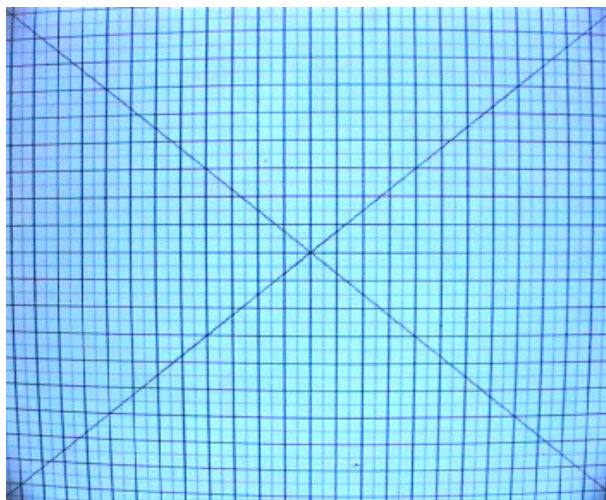
je največje na vogalih in ga lahko na sliki opazimo kot zakrivljenost površine navzdol.

Intenzitetno popačenje kamere je lahko tudi eden izmed razlogov za prvi omenjen problem, torej problem intenzitetnega zamika korespondenčnih točk.

Kot lahko vidimo, problema intenzitetnega zamika korespondenčnih točk v primeru gradnje panoram iz le ene kolone ni. Pravzaprav šivov ni, v paru panoram, namenjenih za rekonstrukcijo, pa je viden vpliv samodejne uravnave svetlosti slike v sami kameri (glej območje vrat na zgornjem paru panoram na sliki 5.8; korespondenčni točki se opazno razlikujeta v barvi). Ker postopek mozaičenja ni primeren za rekonstrukcijo dinamičnih scen, se moramo omejiti na statične scene. V uporabljeni literaturi avtorji ne izpostavljajo problema radialnega popačenja, čeprav je problem očiten. To dejstvo si lahko razlagamo z naslednjo lastnostjo: Velja namreč, da imata simetrični koloni, ki sta enako oddaljeni od središčne kolone zajete slike, enako radialno popačenje, ki pa je seveda usmerjeno obratno. V nadaljevanju bomo delali s simetričnimi pari panoram, zato nam ta lastnost rešuje problem radialnega popačenja, saj so korespondenčne točke⁶ v simetričnem paru panoram vedno enako oddaljene od središčne kolone zajete slike, ki prispeva koloni k panoramama. In to zato, ker je tudi radialno popačenje simetrično. Ta lastnost ne velja za ne-simetrične pare panoram, vendar v uporabljeni literaturi avtorji kljub vsemu problema ne izpostavljajo. Podobno lahko rečemo tudi za problem intenzitetnega popačenja kamere: Namreč tudi to popačenje je (približno) simetrično in kot tako ne vpliva bistveno na gradnjo simetričnih parov panoram.

Iz zgornjega odstavka sledi, da je priporočljivo graditi panorame za rekonstrukcijo iz le ene kolone zajete slike, česar se držijo vsi avtorji uporabljenih literatur,

⁶O epipolarni geometriji bomo spregovorili v naslednjem poglavju, kjer bomo pokazali, da so epipolarne premice simetričnih parov panoram res istoležne vrstice panoram.



Slika 5.5: Radialno popačenje uporabljene kamere na mrežnem vzorcu.

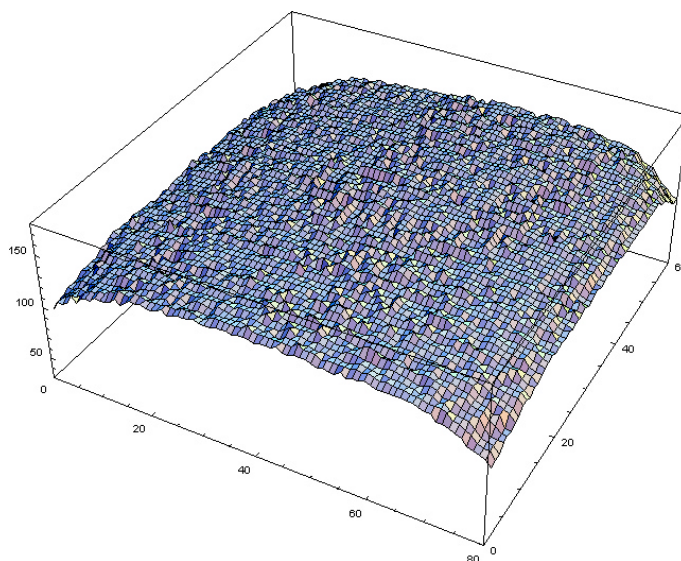
vendar pa to ni nujno. Pri tem pa moramo paziti, da iz zajete slike res vzamemo trakove tako, da se bo isti del (trak) scene pojavil v obeh panoramah, ki sta namenjeni za rekonstrukcijo, s šivi na istem mestu⁷. Bistvena razlika uporabe le ene kolone je, da dobimo gladke panorame, torej brez vidnih šivov. Bolj uniformno pa je tudi radialno popačenje znotraj kolone kot pa znotraj širšega traku. Pri izbiri širine traku moramo vedno narediti kompromis med kvaliteto (gladkostjo) slike in časom, potrebnim za generiranje panoramske slike.

Če za gradnjo panoramskih slik, pri čemer vrtimo rotacijsko roko tudi okoli njene horizontalne osi, uporabimo le ozko področje blizu sredine kamerinega horizonta, potem transformacija iz krogelnih koordinat v valjčne koordinate ni potrebna, ker je izkrivljanje na tem delu dovolj majhno in simetrično [24, 25] (slika 5.7). Podobno predpostavljamo tudi v primeru gradnje naših panoram (slika 5.8) pri naklonu kamere 0° (slika 5.7). Če pa bi kamero v našem primeru vrteli tudi okoli horizontalne osi rotacijske roke in ne bi uporabljali zgolj ozkega področja blizu sredine zajete slike, bi bil vertikalni zorni kot panorame veliko večji kot sicer in bi bila zato geometrijska transformacija nujno potrebna, če bi želeli dobiti pravokotno (valjčno) panoramsko sliko.

V uporabljeni literaturi avtorji za gradnjo panoram uporabljajo ozko področje blizu kamerinega horizonta in ne izpostavljajo potrebe po transformaciji. Dejstvo pa je, da nam postopek rekonstrukcije zaradi zanemarjanja takšnih in podobnih izkrivljanj ne da optimalnih možnih rezultatov.

Pri svojem delu smo uporabili metodo gradnje (globinskih) panoramskih slik pri

⁷To dosežemo le v primeru, če je oddaljenost traku od središčnega traku zajete slike deljiva s širino traku, saj predstavljamo rotacijsko roko za širino traku.



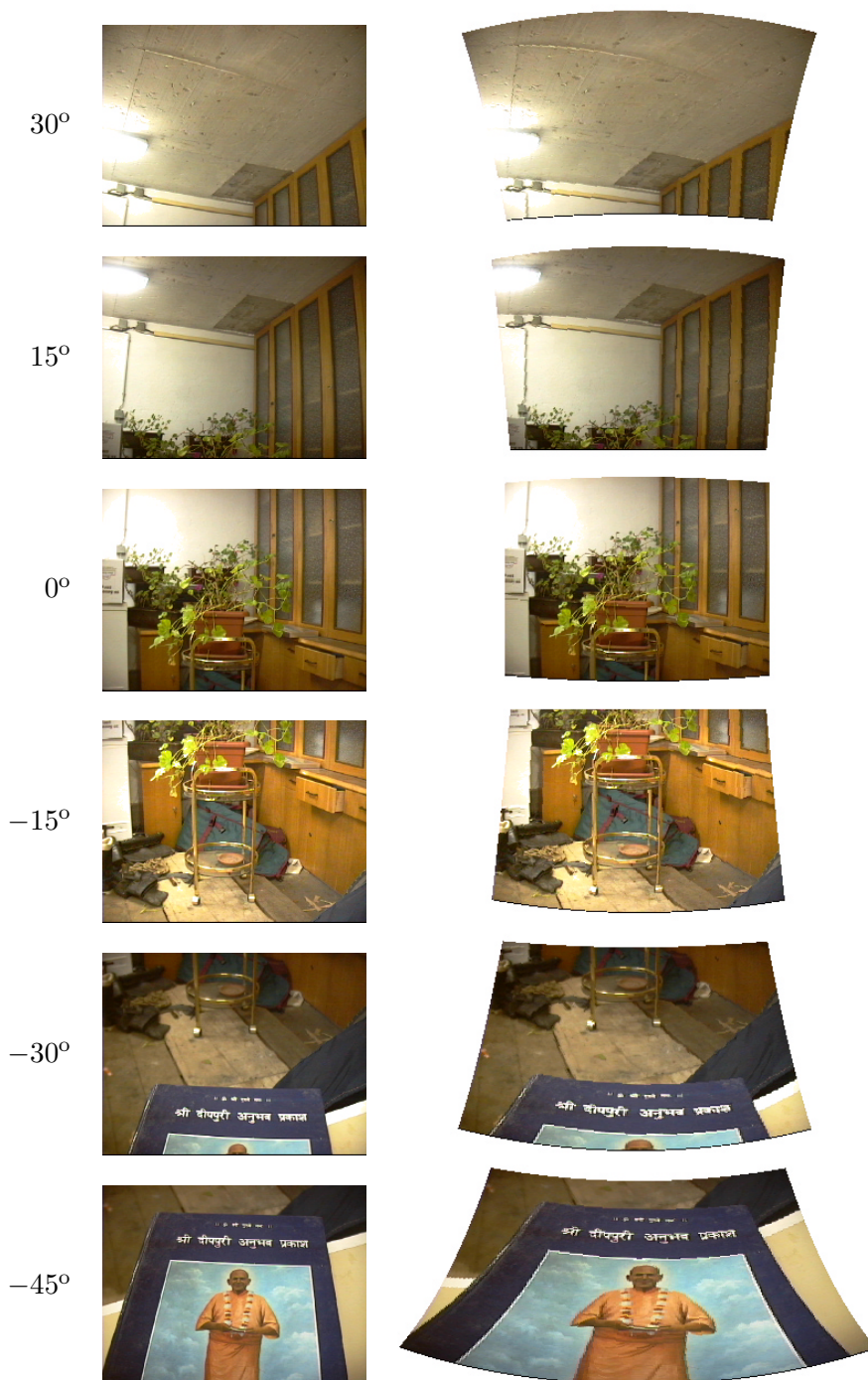
Slika 5.6: 3D prikaz intenzitete slikovnih točk sivinske slike homogene bele površine (bel list papirja), ki nam da lep vpogled v naravo intenzitetnega popačenja uporabljene kamere. [24]

naklonu kamere 0° (slika 5.7) in iz ene same kolone zajete slike. Tako smo se izognili potrebi po uporabi transformacije in zgradili gladke panorame, torej panorame brez vidnih šivov.

Če želimo izvesti samodejno poravnavo slik direktno iz smeri, v katero gleda kamera, moramo poznati horizontalni zorni kot kamere α in vertikalni zorni kot kamere β , saj le tako dobimo informacijo o ločljivosti kotne stopinje, torej o tem, koliko kolon oziroma vrstic opisuje kot ene stopinje. Za naš sistem je pomemben predvsem horizontalni zorni kot, saj rotacijsko roko vrtimo le okoli vertikalne osi. Za izračun teh dveh parametrov smo uporabili algoritem opisan v [24, 25], ki deluje s kamerami, kjer so goriščna razdalja leč in ostali interni parametri kamere neznani. Algoritem se naslanja na zanesljivost rotacijske roke. Osnovna ideja algoritma je izračunati premik med dvema slikama dx (v slikovnih elementih) medtem, ko je bila kamera obrnjena za nek določen kot $d\gamma$ v horizontalni smeri. Ker poznamo natančen kot, za katerega smo obrnili kamero, lahko izračunamo horizontalni zorni kot kamere:

$$\alpha = \frac{W}{dx} \cdot d\gamma, \quad (5.1)$$

kjer je W širina slike v slikovnih elementih. Sedaj lahko izračunamo ločljivost kotne stopinje x_0 :



Slika 5.7: Rezultati geometrijske transformacije iz krogelnih koordinat v valjčne koordinate pri različnih naklonih kamere. Če za gradnjo panoramskih slik uporabljamo le ozko področje blizu kamerinega horizonta, potem transformacija iz krogelnih koordinat v valjčne koordinate ni potrebna, ker je izkrivljanje na tem delu dovolj majhno in simetrično. [24, 25]

$$x_0 = \frac{W}{\alpha}, \quad (5.2)$$

kar nam omogoča izračun širine traku W_s , ki bo doprinesel k panorami ob premiku rotacijske roke za kot θ_0 :

$$W_s = x_0 \cdot \theta_0. \quad (5.3)$$

Iz te enačbe lahko izračunamo tudi potreben premik, ki ga moramo narediti, da je trak širok le eno kolono zajete slike.

Kot že omenjeno, je horizontalni zorni kot uporabljene kamere znašal $\alpha = 34^\circ$, vertikalni zorni kot pa $\beta = 25^\circ$. Med zajemanjem slik se ta dva parametra nista spreminjala.

Vrnimo se za hip nazaj na sliko 5.2, kjer je prikazana geometrija sistema za ustvarjanje ene večperspektivne panorame. V ilustriranem primeru gradimo panoramo iz kolon zajetih slik, ki so, gledano iz optičnega centra kamere, od središčne kolone odmaknjene za kot φ . Naš sistem zgradi v enem obhodu rotacijske roke okoli vertikalne osi 11 panoramskih slik (slika 5.8):

- panoramo iz središčne kolone in
- pet simetričnih parov panoram s koti $2\varphi = 29,9625^\circ, 26,5625^\circ, 18,9125^\circ, 11,2625^\circ$ in $3,6125^\circ$.

Zapisani koti so rezultat izbire kolon, ki doprinesejo k posamezni panorami (slika 5.9). Po enačbi (5.1) velja:

$$2\varphi = \frac{\alpha}{W} \cdot W_{2\varphi}, \quad (5.4)$$

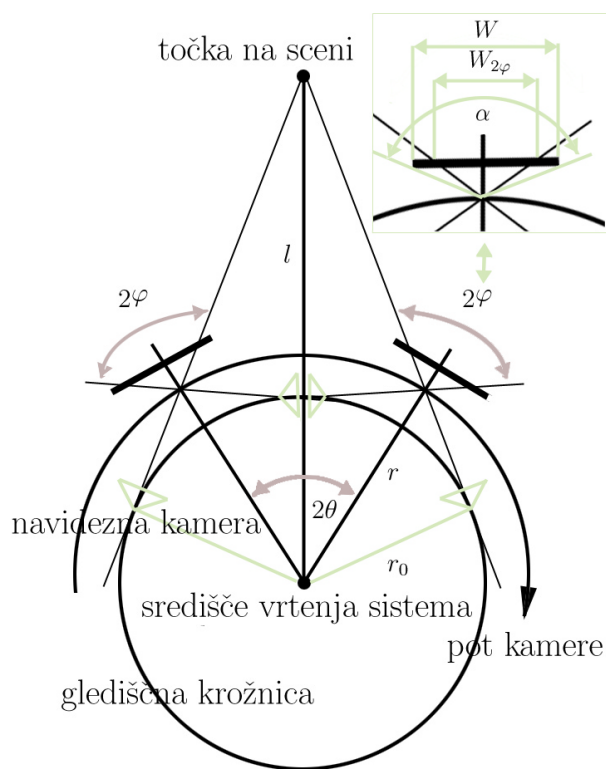
pri čemer je α horizontalni zorni kot kamere, W je širina zajete slike v slikovnih elementih, $W_{2\varphi}$ pa širina slike med kolonama, ki gradita simetrični par panoram, prav tako v slikovnih elementih. Iz slike 5.9 je razvidno, da lahko na naš sistem gledamo tudi kot na sistem dveh kamer. V našem primeru navidezni kameri potujeta po glediščni krožnici polmera r_0 (glej tudi sliko 5.2):

$$r_0 = r \cdot \sin \varphi. \quad (5.5)$$

Optična os navidezne kamere je vedno tangenta na glediščno krožnico in velja, da gradimo panoramo iz središčnih kolon slik, ki jih zajame navidezna kamera. Če opazujemo neko točko na sceni, vidimo, da navidezni kameri, ki obe vidita to točko (zunanji navidezni kameri na sliki 5.9), tvorita tradicionalni stereo sistem konvergirajočih kamer.



Slika 5.8: Zasnovan sistem zgradi v enem obhodu rotacijske roke okoli vertikalne osi 11 panoramskih slik (glej tekst). Pri vsakem kotu $\varphi \neq 0$ predstavlja zgornja slika panoramo za levo oko, spodnja slika panoramo za desno oko, skupaj pa panorami predstavljata simetričen par stereo panoram.



Slika 5.9: Panoramo gradimo iz kolon zajetih slik, ki so, gledano iz optičnega centra kamere, od središčne kolone odmaknjene za kot φ . Kolona na levi strani zajete slike prispeva k panorami za desno oko (oziroma k panorami, ki jo gradi desna navidezna kamera), kolona na desni strani zajete slike pa prispeva k panorami za levo oko (oziroma k panorami, ki jo gradi leva navidezna kamera). Tako dobimo simetričen par (stereo) panoram. Kot vidimo lahko na naš sistem gledamo tudi kot na sistem dveh kamer.

Ker smo kot bazno ločljivost izbrali velikost slike 160×120 slikovnih elementov, so tudi koti izračunani za to ločljivost⁸. To ločljivost smo izbrali kot bazno zaradi kompromisa med velikostjo dobljene panorame in časom, ki je potreben za ustvarjanje le-te (glej poglavje 5.6.2).

Za konec tega poglavja povzemimo glavne korake gradnje panoramskih pogledov z mozaičenjem slik našega sistema:

1. umerimo kamero, torej določimo horizontalni zorni kot kamere α ,

⁸V primeru, da bi izbrali večjo ločljivost, na primer 640×480 , bi bili koti malo drugačni: na primer prvi kot bi namesto vrednosti $29,9625^\circ$ imel vrednost $29,803125^\circ$. Razlika je posledica diskretizacije slike na slikovne elemente.

2. zajamemo slike z znanim premikom θ_0 rotacijske roke med dvema zaporedno zajetima slikama,
3. na podlagi znanega premika rotacijske roke θ_0 in horizontalnega zornega kota kamere α združimo zajete slike v panorame, pri čemer gradimo panorame iz kolon zajetih slik, ki so, gledano iz optičnega centra kamere, od središčne kolone odmaknjene za določen kot φ .

5.4 Računanje epipolarne geometrije

V tem poglavju bomo pokazali, da so epipolarne premice simetričnih parov panoram res istoležne vrstice panoram [12, 27].

Dokaz temelji na vrednosti r_0 iz slike 5.2, ki predstavlja polmer glediščne krožnice in jo izračunamo iz enačbe (5.5). Na r_0 lahko gledamo kot na polmer zajete panoramske slike, saj dobimo enako sliko tudi v primeru, če vrtimo linijsko kamero po krožnici radija r_0 , pri čemer je njena optična os tangenta na to krožnico, ki opisuje pot kamere.



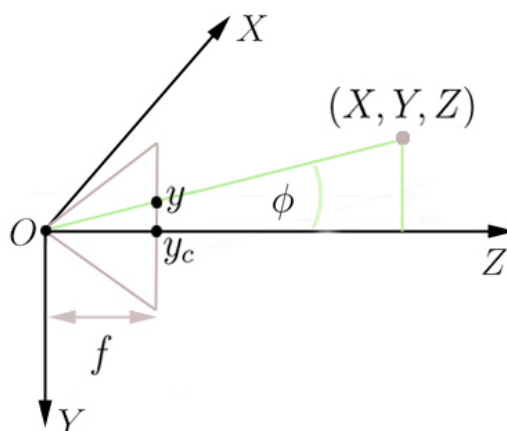
Dokaz bomo izpeljali v treh korakih: najprej moramo izpeljati projekcijsko enačbo za linijsko kamero, potem moramo zapisati projekcijsko enačbo za večperspektivno panoramsko sliko in šele v tretjem koraku bomo lahko dokazali lastnost epipolarnih premic simetričnih parov panoram. V prvem delu nas zanima, kako se neka točka v prostoru preslika na ravnino kamere (poglavje 2.2), ki je tokrat ravnina⁹ velikosti $n \times 1$. V drugem delu bomo podali ekvivalentne opise točk v prostoru v evklidskih koordinatah 3D koordinatnega sistema, valjčnih koordinatah 3D koordinatnega sistema, kotnih koordinatah 2D koordinatnega sistema panoramske slike in slikovnih koordinatah 2D koordinatnega sistema panoramske slike ter tako zapisali enačbo za projekcijo točk v prostoru na ravnino (valj) večperspektivne panorame. Prav iz lastnosti kotnih koordinat koordinatnega sistema panoramske slike pa bomo v tretjem delu pokazali lastnost epipolarnih premic simetričnih parov panoram.

5.4.1 Projekcijska enačba za linijsko kamero

Predpostavimo, da imamo opraviti z idealno kamero, kar nam pomeni, da je optična os kamere pravokotna na sliko linijske kamere in da jo optična os seka na sredini. Potem lahko koordinato y slike linijske kamere zapišemo s pomočjo kotne koordinate ϕ kot (slika 5.10):

$$\phi = \arctan \frac{y - y_c}{f},$$

⁹Slikovna ravnina linijske kamere je premica, vendar bomo zaradi konsistentnosti s standardno kamero, na katero najprej pomislimo, ko slišimo besedo kamera, pojem ravnina uporabljali tudi v primeru linijske kamere.



Slika 5.10: Projekcija 3D točke na slikovno ravnino linijske kamere.

kjer je y_c koordinata sredine slike, f pa goriščna razdalja kamere. Če projeciramo neko 3D točko, ki ima koordinate (X, Y, Z) , na slikovno ravnino linijske kamere, dobimo projekcijsko enačbo:

$$y = f \tan \phi + y_c. \quad (5.6)$$

Ker pa iz slike 5.10 sledi, da je projekcija možna le v primeru, da je koordinata $X=0$, lahko zgornjo enačbo zapišemo kot:

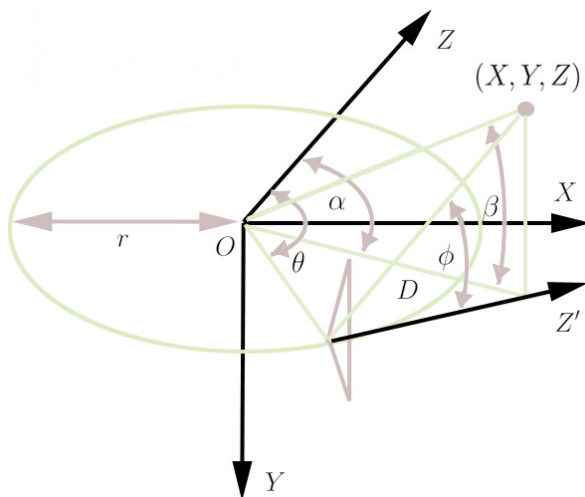
$$y = f \frac{Y}{Z} + y_c.$$

Na podlagi zgornje enačbe lahko zapišemo enačbo optičnega žarka (poglavje 2.2) v 3D prostoru (izhodišče prostora je v optičnem centru O linijske kamere), ki je določen s slikovnim elementom $m = (0, y)$ (če slikovni element predstavimo v prostoru pa imamo $m_{3D} = (0, y - y_c, f)$) in optičnim centrom O (na enačbo lahko gledamo tudi kot na vektor iz optičnega centra O do projicirane 3D točke):

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \lambda \begin{bmatrix} 0 \\ y - y_c \\ f \end{bmatrix}, \quad (5.7)$$

kjer je $\lambda \in \mathbb{R}$ skalar, saj delamo perspektivno preslikavo, za njo pa vemo, da jo najlažje obravnavamo s projektivnega vidika [9] (stran 7), kar pa pomeni, da imamo opravka s homogenimi koordinatami točke m , ki je definirana do multiplikativnega faktorja natančno. Prav z enačbo optičnega žarka pa si bomo pomagali v poglavju 5.4.3, kjer bomo dokazali lastnost epipolarnih premic simetričnih parov panoram.

5.4.2 Projekcijska enačba za večperspektivno panoramsko sliko



Slika 5.11: Relacija med 3D svetovnim koordinatnim sistemom, kjer točko opišemo s trojico (X, Y, Z) , 3D valjčnim koordinatnim sistemom, kjer točko opišemo s trojico (α, β, D) in 2D kotnim koordinatnim sistemom panoramske slike, kjer točko opišemo s parom (θ, ϕ) .

Sedaj definirajmo 3D svetovni koordinatni sistem z izhodiščem O v središču vrtenja našega sistema (slika 5.11). Če želimo neko točko v 3D svetovnem koordinatnem sistemu s koordinatami (X, Y, Z) zapisati v 3D valjčnem koordinatnem sistemu s koordinatami (α, β, D) , velja:

$$\begin{pmatrix} \alpha \\ \beta \\ D \end{pmatrix} = \begin{pmatrix} 2A\pi - B \arccos \frac{Z}{\sqrt{X^2 + Z^2}} \\ \arctan \frac{Y}{\sqrt{X^2 + Z^2}} \\ \sqrt{X^2 + Z^2} \end{pmatrix},$$

kjer sta parametra A in B določena glede na vrednost izraza:

$$E = \arccos \frac{X}{\sqrt{X^2 + Z^2}}$$

in sicer:

$$\begin{cases} A = 0, B = -1, & \text{če } E \leq \frac{\pi}{2} \\ A = 1, B = 1, & \text{sicer.} \end{cases}$$

Če definiramo 2D kotni koordinatni sistem panoramske slike, kot je to razvidno iz slike 5.11, potem lahko točko v 3D valjčnem koordinatnem sistemu s koordinatami (α, β, D) zapišemo v 2D kotnem koordinatnem sistemu s koordinatami (θ, ϕ) kot:

$$\begin{pmatrix} \theta \\ \phi \end{pmatrix} = \begin{pmatrix} \alpha - C \arccos \frac{r}{D} \\ \arctan \frac{D \tan \beta}{\sqrt{D^2 - r^2}} \end{pmatrix},$$

kjer je parameter $C = 1$ v primeru, da večperspektivno panoramo gradimo iz kolone, ki je na levi strani sredinske kolone zajete slike (panorama za desno oko) in $C = -1$ v primeru, da večperspektivno panoramo gradimo iz kolone, ki je na desni strani sredinske kolone zajete slike (panorama za levo oko).

Tako na primer velja: če gradimo panoramo za levo oko ($C = -1$) in projiciramo nanjo neko 3D točko s koordinatami (X, Y, Z) , pri tem pa velja $E < \frac{\pi}{2}$ ($A = 0$, $B = -1$), potem to točko v 2D kotnem koordinatnem sistemu panoramske slike opišemo s parom:

$$\begin{pmatrix} \theta \\ \phi \end{pmatrix} = \begin{pmatrix} \arccos \frac{Z}{\sqrt{X^2 + Z^2}} + \arccos \frac{r}{\sqrt{X^2 + Z^2}} \\ \arctan \frac{Y}{\sqrt{X^2 + Z^2 - r^2}} \end{pmatrix}. \quad (5.8)$$

Na podlagi dejstva, da kot θ definira horizontalno os panoramske slike, ki opisuje kot zasuka okoli vertikalne osi našega sistema (zaželjena je možnost zasuka za 2π), in enačbe 5.6, lahko sedaj zapišemo slikovni koordinati (x, y) 2D koordinatnega sistema panoramske slike v odvisnosti od kotnih koordinat (θ, ϕ) 2D koordinatnega sistema panoramske slike kot:

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \frac{\theta}{2\pi} W \\ f \tan \phi + y_c \end{pmatrix}, \quad (5.9)$$

kjer z W označimo širino panoramske slike.

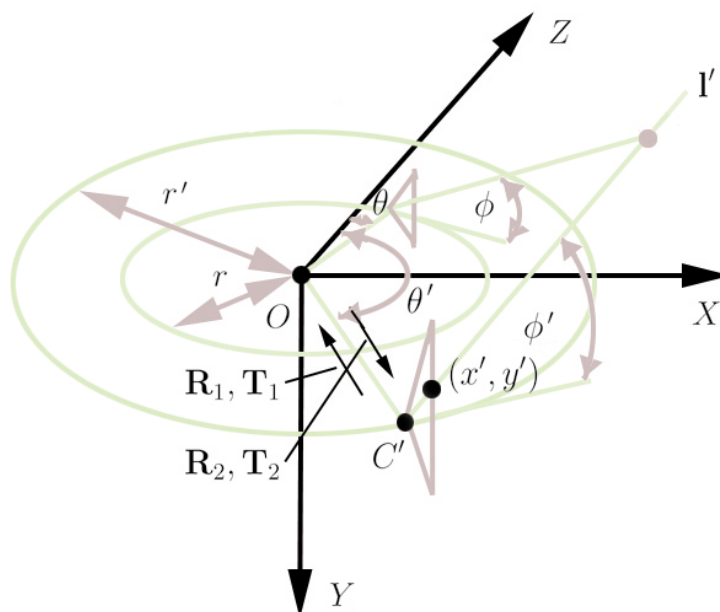
5.4.3 Lastnost epipolarnih premic simetričnih parov panoram

Osnovna ideja zadnjega dela dokaza sledi iz naslednje trditve:

Če imamo dano neko slikovno točko na eni večperspektivni panoramski sliki, lahko izrazimo optični žarek (enačba 5.7) v 3D svetovnem koordinatnem sistemu. Če optični žarek izražen v svetovnem koordinatnem sistemu projiciramo na drugo večperspektivno panoramsko sliko in pri tem spreminjamo vrednost λ (glej razlago enačbe 5.7), dobimo epipolarno premico za dano slikovno točko na prvi večperspektivni panoramski sliki.

Kot opombo na tem mestu zapišimo še, da je izhodišče svetovnega koordinatnega sistema v središču vrtenja našega sistema.

Pri spodnjem dokazu bomo izhajali iz poljubnega para večperspektivnih panoram, na koncu pa bomo v dobljeno enačbo vstavili ustrezne relacije za simetričen par panoram.



Slika 5.12: Koncept projekcije točke v 3D svetovnem koordinatnem sistemu na par večperspektivnih panoramskih slik: optični žarek, ki je določen z optičnim centrom C' kamere in dano 2D točko s koordinatama (x', y') na panoramski sliki, nam opiše epipolarno premico za dano 2D točko na drugi panorami.

Dokaz: [12]

Koncept projekcije je razviden iz slike 5.12. Zanima nas epipolarna premica na večperspektivni panoramski sliki, ki jo dobimo z navidezno kamero, ki potuje po glediščni krožnici s polmerom r , pri čemer je dana točka na večperspektivni panoramski sliki, ki jo dobimo z navidezno kamero, ki potuje po glediščni krožnici s polmerom r' .¹⁰ Dano točko označimo s koordinatama (x', y') . Enačba (5.7) opisuje optični žarek (3D premico) glede na 3D koordinatni sistem kamere z izhodiščem v optičnem centru kamere C' . Lokacijo C' na glediščni krožnici lahko podamo s kotom θ' , ki je, kot smo že omenili (enačba 5.9), odvisna od dane koordinate x' :

$$\theta' = \frac{x'}{W} 2\pi,$$

kjer je W širina panoramske slike.

Po enačbi (5.7) in enačbi (5.6) označimo optični žarek, ki je določen z optičnim centrom C' kamere in dano 2D točko s koordinatama (x', y') na panoramski sliki, z

¹⁰Dokaz v obratnem primeru je simetričen.

oznako¹¹ $\tilde{\mathbf{I}}'$:

$$\tilde{\mathbf{I}}' = \lambda \begin{bmatrix} 0 \\ y' - y'_c \\ f \end{bmatrix} = \lambda \begin{bmatrix} 0 \\ f \tan \phi \\ f \end{bmatrix}.$$

Sedaj moramo ta žarek opisati v svetovnem koordinatnem sistemu, katerega izhodišče je v središču vrtenja O našega sistema. Transformacijo žarka $\tilde{\mathbf{I}}'$ iz koordinatnega sistema kamere v svetovni koordinatni sistem opišemo z rotacijsko matriko \mathbf{R} in translacijskim vektorjem \mathbf{T} (poglavje 2.4). Na sliki 5.12 sta označena dva para (\mathbf{R}, \mathbf{T}) : $(\mathbf{R}_1, \mathbf{T}_1)$ in $(\mathbf{R}_2, \mathbf{T}_2)$, med katerima veljata relaciji:

$$\mathbf{R}_2 = \mathbf{R}_1^T \text{ in } \mathbf{T}_2 = -\mathbf{R}_1^T \mathbf{T}_1$$

Iz slike 5.12 lahko enostavno izrazimo \mathbf{R}_2 in \mathbf{T}_2 (poglavje 3.4). Tako matrika¹² \mathbf{R}_2 kot vektor \mathbf{T}_2 sta odvisna od parametra θ' . Rotacijsko matriko \mathbf{R}_2 določimo na osnovi vrtenja okoli osi y (dodatek A) in sicer je kot med kamerama¹³ enak $\theta' - 90^\circ$, translacijski vektor \mathbf{T}_2 pa izrazimo s premikom v ravnini xz . Tako dobimo:

$$\mathbf{R}_2 = \begin{bmatrix} \sin \theta' & 0 & -\cos \theta' \\ 0 & 1 & 0 \\ \cos \theta' & 0 & \sin \theta' \end{bmatrix},$$

$$\mathbf{T}_2 = \begin{bmatrix} r' \sin \theta' \\ 0 \\ r' \cos \theta' \end{bmatrix}.$$

Če z $\tilde{\mathbf{I}}$ označimo optični žarek $\tilde{\mathbf{I}}'$ v svetovnem koordinatnem sistemu in predpostavimo, da se nahajamo v normaliziranem koordinatnem sistemu ($f = 1$; slika 2.6) (poglavje 3.4), velja:

$$\tilde{\mathbf{I}} = \tilde{\mathbf{P}} \begin{bmatrix} \tilde{\mathbf{I}}' \\ 1 \end{bmatrix} = [\mathbf{R}_2 \quad \mathbf{T}_2] \begin{bmatrix} 0 \\ \lambda \tan \phi \\ \lambda \\ 1 \end{bmatrix} =$$

¹¹S tildo (\sim) označujemo dejstvo, da delamo s homogenimi koordinatami (poglavje 2.1).

¹²V primeru, da bi bil optični žarek določen z optičnim centrom C kamere in dano točko na večperspektivni panoramski sliki, ki jo dobimo z navidezno kamero, ki potuje po glediščni krožnici s polmerom r (slika 5.12), bi za rotacijsko matriko veljalo:

$$\mathbf{R}_{2,r}(\theta') = -\mathbf{R}_{2,r}(\theta').$$

¹³Optična os kamere v svetovnem koordinatnem sistemu je usmerjena v smeri osi z .

$$\begin{bmatrix} r' \sin \theta' - \lambda \cos \theta' \\ \lambda \tan \phi' \\ r' \cos \theta' + \lambda \sin \theta' \end{bmatrix} = \begin{bmatrix} X \\ Y \\ Z \end{bmatrix},$$

kjer je $\tilde{\mathbf{P}}$ matrika perspektivne projekcije izražena v 3D koordinatnem sistemu linijske kamere z izhodiščem v C' .

Če sedaj spreminjamo vrednost λ , dobimo točke na optičnem žarku $\tilde{\mathbf{I}}$ v svetovnem koordinatnem sistemu. Projekcija teh točk na drugo večperspektivno panoramsko sliko, ki jo dobimo z navidezno kamero, ki potuje po glediščni krožnici s polmerom r (panorama za levo oko), opisuje epipolarno premico, ki jo lahko v 2D kotnem koordinatnem sistemu zapišemo po enačbi (5.8) kot:

$$\begin{pmatrix} \theta \\ \phi \end{pmatrix} = \begin{pmatrix} \arccos \frac{r' \cos \theta' + \lambda \sin \theta'}{\sqrt{r'^2 + \lambda^2}} + \arccos \frac{r}{\sqrt{r'^2 + \lambda^2}} \\ \arctan \frac{\lambda \tan \phi'}{\sqrt{r'^2 + \lambda^2 - r^2}} \end{pmatrix}.$$

Zgornja enačba torej podaja epipolarno premico za poljuben par večperspektivnih panoram, nas pa zanima enačba epipolarne premice za specifičen primer simetričnega para večperspektivnih panoram, za katerega velja $r' = r$. Ker je slikovna koordinata y (vrstica panoramske slike) 2D koordinatnega sistema panoramske slike povezana s kotno koordinato ϕ 2D koordinatnega sistema panoramske slike (enačba 5.9), uporabimo omenjeno zvezo $r' = r$ v izrazu za ϕ zgornje enačbe:

$$\phi = \arctan \frac{\lambda \tan \phi'}{\sqrt{r'^2 + \lambda^2 - r^2}} \stackrel{r'=r}{\implies} \phi = \phi'.$$

□

Rezultat, ki smo ga dobili nam pove, da sta korespondenčni točki v istoležnih vrsticah simetričnih večperspektivnih panoram (efekt vertikalne paralakse gibanja ni prisoten), torej da so epipolarne premice res vrstice panoram.

To pomeni, da nam pri implementaciji rekonstrukcije iz simetričnih parov večperspektivnih panoram ni treba računati epipolarne geometrije, s čimer se kompleksnost postopka bistveno zmanjša. Po drugi strani pa nam dejstvo, da so epipolarne premice vrstice slik, omogoča, da za stereo rekonstrukcijo uporabimo tradicionalne postopke.

5.5 3D rekonstrukcija

V tem poglavju bomo izpeljali postopek za oceno globine za najdene korespondenčne točke.

Vrnimo se k sliki 5.2. Na osnovi trigonometričnih relacij moramo zapisati enačbo za oceno globine L točke P na sceni. Po sinusnem izreku velja:

$$\frac{r}{\sin(\varphi - \theta)} = \frac{d}{\sin \theta} = \frac{l}{\sin(180^\circ - \varphi)},$$

iz te enačbe pa lahko izrazimo enačbo za oceno globine l :

$$l = \frac{r \cdot \sin(180^\circ - \varphi)}{\sin(\varphi - \theta)} = \frac{r \cdot \sin \varphi}{\sin(\varphi - \theta)}. \quad (5.10)$$

Iz zgornje enačbe sledi, da lahko globino l ocenimo le, če poznamo tri parametre: r , φ in θ . r je dan, kot φ izračunamo glede na horizontalni zorni kot kamere α po enačbi (5.4) (slika 5.9), za izračun kota θ pa moramo najprej najti korespondenčni točki na panoramskih slikah. Naš sistem deluje tako, da kamero po krožni poti premikamo za kot, ki ustreza eni koloni zajete slike. Če ta kot označimo s θ_0 , potem za kot θ velja:

$$\theta = dx \cdot \frac{\theta_0}{2}, \quad (5.11)$$

kjer je dx absolutna vrednost razlike med slikovnjima koordinatama korespondenčnih točk na horizontalni osi x panoramskih slik¹⁴.

5.5.1 Iskanje korespondenčnih točk

Soočeni smo z naslednjim problemom: na eni večperspektivni panorami imamo dano slikovno točko, za katero iščemo korespondenčno točko na drugi večperspektivni panorami in vemo, da se korespondenčna točka nahaja nekje na epipolarni premici. Ker obdelujemo simetričen par večperspektivnih panoram, velja, da so epipolarne premice vrstice panoram, torej lahko za iskanje korespondenčnih točk uporabimo korelacijsko metodo¹⁵ (poglavje 3.5.1, slika 3.5).

Okoli dane slikovne točke postavimo korelacijsko okno velikosti $(2n + 1) \times (2m + 1)$ slikovnih elementov tako, da je dani slikovni element v sredini tega okna. Po epipolarni premici na drugi panorami pa premikamo okno enake velikosti in za vsako točko na tej premici izračunamo korelacijo tako definiranih bližnjih okolice slikovnih elementov. Če označimo slikovno točko na prvi panorami z $\mathbf{m}_1 = (x_1, y_1)$, slikovno točko na drugi panorami pa $\mathbf{m}_2 = (x_2, y_2)$, lahko mero za korelacijo zapišemo kot [9]:

$$\text{korelacija}(\mathbf{m}_1, \mathbf{m}_2) = \sum_{i=-n}^n \sum_{j=-m}^m f(I_1(x_1 + i, y_1 + j), I_2(x_2 + i, y_2 + j)),$$

¹⁴Podobno je definirana tudi dispariteta v primeru vzporednih kamer. (poglavje 3.5.1)

¹⁵Na korelacijsko metodo lahko gledamo tudi kot na postopek ujemanja vzorcev (ang. *template matching*).

kjer je $I_k(x, y)$ intenzitetna (barvna) vrednost slikovnega elementa s koordinatama (x, y) na sliki I_k ($k = 1, 2$), f pa funkcija za oceno podobnosti. Za $f = f(u, v)$ se največkrat uporabi *križna korelacija* (ang. *cross-correlation*) f_x ali *vsota kvadratov razlik* (ang. *Sum of Squared Differences*) f_{SSD} :

$$f_x(u, v) = u \cdot v,$$

$$f_{SSD}(u, v) = -(u - v)^2 = -u^2 + 2 \cdot u \cdot v - v^2.$$

Križna korelacija deluje dobro le, če je energija, ki je definirana kot vsota kvadratov vrednosti slikovnih elementov v korelacijskem oknu, na drugi sliki konstantna na epipolarni premici [1, 36]. Vsota kvadratov razlik je boljša alternativa, saj za razliko od križne korelacije ne daje višjih vrednosti pri izrazito svetlih, ali izrazito temnih delih prizora. Iz zgornjih izrazov je razvidno tudi, da je križna korelacija poseben primer vsote kvadratov razlik, saj se lahko prepričamo, da vsota kvadratov razlik vrne enako vrednost kot križna korelacija pod pogoji konstantne energije na epipolarni premici.

V korespondenco postavimo tisti dve točki, katerih korelacija je maksimalna.

Slabost obeh opisanih mer je, da vračata vrednosti praktično na poljubnih intervalih: korelacija f_x vrne rezultat na intervalu $[0, \infty]$, korelacija f_{SSD} pa vrne rezultat na intervalu $[-\infty, 0]$. Izboljšana korelacijska metoda, ki daje rezultate na intervalu $[-1, 1]$, kjer 1 pove, da sta korelacijski okni identični, -1 pa, da sta popolnoma različni, se imenuje *normalizirana korelacija* (ang. *normalized correlation*):

normalizirana_korelacija($\mathbf{m}_1, \mathbf{m}_2$) =

$$\frac{\sum_{i=-n}^n \sum_{j=-m}^m [(I_1(x_1+i, y_1+j) - \overline{I_1(x_1, y_1)})(I_2(x_2+i, y_2+j) - \overline{I_2(x_2, y_2)})]}{(2n+1)(2m+1)\sigma(I_1)\sigma(I_2)},$$

kjer je

$$\overline{I_k(x, y)} = \frac{\sum_{i=-n}^n \sum_{j=-m}^m I_k(x+i, y+j)}{(2n+1)(2m+1)}$$

povprečje intenzitetnih vrednosti,

$$\sigma(I_k) = \sqrt{\frac{\sum_{i=-n}^n \sum_{j=-m}^m (I_k(x+i, y+j) - \overline{I_k(x, y)})^2}{(2n+1)(2m+1)}}$$

pa standardni odklon na sliki I_k , oboje izračunano znotraj korelacijskega okna, ki ima središče v slikovni točki (x, y) .

Če opazujemo enačbe za korelacijo, opazimo, da enačbe niso simetrične glede na I_1 in I_2 , torej v primeru, da imamo dano točko \mathbf{m}_2 in iščemo korespondenčno točko \mathbf{m}_1 . Najboljša ocena dobljena po enačbi za katerokoli korelacijo, nam ne opisuje

nujno tudi dejanske korespondenčne točke, zato je dobro, če bi lahko dobljeni rezultat dodatno potrdili in s tem povečali zaupanje vanj. Postopek, ki nam to omogoča, je znan pod imenom *vzratna korelacija* (ang. *back correlation*) in zaupa le v tiste rezultate, za katere velja:

$$(\text{normalizirana_})\text{korelacija}(\mathbf{m}_1, \mathbf{m}_2) = (\text{normalizirana_})\text{korelacija}(\mathbf{m}_2, \mathbf{m}_1),$$

ostale rezultate pa enostavno zavrže kot pogreške. To pomeni, da v točkah, kjer je prišlo do pogreška, globine ne moremo oceniti po enačbi (5.10). Na globinskih slikah v takšnih primerih ustrezno točko največkrat predstavimo s črno barvo [9] (str. 192). Tako, kot rezultat rekonstrukcije, dobimo globinsko sliko z manjkajočo informacijo o globini posameznih delov te slike, ki pa jih lahko zapolnimo s postopkom interpolacije. Pri tem moramo seveda uporabiti nek pameten postopek interpolacije, ki bo kar najboljše zapolnil vrzeli v globinski sliki, hkrati pa bo znal oceniti, kdaj je interpolacija nesmiselna (na primer glede na velikost vrzeli).

5.6 Analiza sistema in rezultatov

5.6.1 Čas gradnje panoramskih slik

Pri gradnji panoramskih slik z našim sistemom, predstavlja časovno ozko grlo, premikanje rotacijske roke, pri tem pa moramo paziti tudi na to, da mora biti rotacijska roka, preden zajamemo sliko, povsem umirjena. Čas, potreben za gradnjo panoram, je torej predolg, da bi lahko trdili, da so sploh kakšne možnosti, da sistem deluje v realnem času. Tudi v prihodnosti najverjetneje nikoli ne bo, pa čeprav lahko rotacijska roka opravi zasuk za 300° v eni sekundi.

V tabeli 5.1 je prikazanih nekaj osnovnih podatkov o postopku mozaičenja za naš sistem ter sisteme opisane v [24, 25]¹⁶, [13] in [27]. Ob upoštevanju naslednjih dejstev, ki so razvidni iz tabele 5.1:

- z večanjem števila zajetih slik, ki prispevajo k panorami, se podaljša čas potreben za gradnjo panorame,
- z večanjem ločljivosti zajetih slik se prav tako podaljša čas potreben za gradnjo panorame,
- podobno velja tudi za število panoram, ki jih ustvarimo v enem zasuku sistema okoli svoje vertikalne osi,

¹⁶Sistem gradi panoramo z vrtenjem kamere okoli vertikalne in horizontalne osi rotacijske roke. Tipično je panoramska slika zgrajena iz petih premikov okoli horizontalne osi. Zaradi tega dejstva bi v tabeli 5.1 morala na mestu podatka o koraku rotacijske roke biti vrednost 2° , vendar je zaradi lažje primerjave z drugimi sistemi ta vrednost petkrat manjša, saj se čas gradnje panorame ne razlikuje bistveno, če namesto petih premikov okoli horizontalne osi raje zajemamo slike s petkrat manjšim korakom rotacijske roke.

	[24, 25]	[13]	[27]	Naš sistem
št. slik	700	900	1280	1501
ločljivost [slikovnih pik]	80×60	342×256	240×320	160×120
št. panoram	1	2	?	11
θ_0 [°]	0,4	0,4	0,28125	0,2
r [cm]	0	20	?	30
čas [min.]	3	10	?	15,5

Tabela 5.1: Nekaj osnovnih podatkov o postopku mozaičenja za naš sistem ter sisteme opisane v [24, 25], [13] in [27] (glej tekst). V zadnji vrstici je podan približen čas, ki je potreben za zasuk kamere okoli vertikalne osi sistema in hkratno gradnjo panoram(e) ob zgoraj podanih parametrih sistema. Vprašaj (?) v polju pomeni, da ta podatek v literaturi ni bil naveden, oziroma da se ga ni dalo izračunati. Pri sistemu opisanem v [27], lahko opazimo, da je kamera pritrjena v pokončni legi. Kot opombo navedimo, da v primeru našega sistema in sistema, ki je opisan v [24, 25], panoramska slika ne opisuje kot 360° (poglavje 5.1.2).

- večji je korak rotacijske roke θ_0 , prej sistem opiše pot okoli svoje vertikalne osi in tako prej zgradi panoramsko sliko in
- večji je odmik optičnega centra kamere od središča vrtenja sistema r , več časa potrebuje sistem, da se na dani lokaciji povsem umiri,

lahko zaključimo, da je čas, potreben za gradnjo ene panoramske slike v vseh (treh) sistemih približno enak.

5.6.2 Omejitev sistema zaradi ločljivosti rotacijske roke

Glede na to, da smo pri vrtenju roke omejeni z njeno ločljivostjo (poglavje 5.1.2), lahko naredimo analizo, kakšna je lahko maksimalna velikost zajete slike, glede na ločljivost roke. Ker poznamo horizontalni zorni kot kamere $\alpha = 34^\circ$, velja, da lahko širino slike W izračunamo iz enačbe (5.2) kot:

$$W = \alpha \cdot x_0,$$

kjer je x_0 ločljivost kotne stopinje, ki jo lahko izračunamo iz enačbe (5.3), saj vemo, da je minimalen možen premik rotacijske roke ($\theta_0 = 0,0514285^\circ$), vsaka zajeta slika pa doprinese k panorami le eno kolono ($W_s = 1$):

$$x_0 = \frac{W_s}{\theta_0} = 19,444450448 \text{ slikovnih pik na stopinjo.}$$

Torej znaša maksimalna možna širina slike W 661,11131523 slikovnih pik, kar pa dejansko v praksi, zaradi zaokroževanja in standardnih ločljivosti slike, pomeni, da je lahko maksimalna širina 640 slikovnih pik. Za naš sistem to pomeni, da lahko zajemamo slike do velikosti 640×480 slikovnih pik, glede na tabelo 5.1 pa velja, da je čas, ki je potreben za zasuk sistema okoli svoje vertikalne osi in hkratno gradnjo panoram, približno štirikrat večji od navedenega. Pri tem zajamemo 6000 slik.

Ločljivost 160×120 slikovnih pik smo izbrali kot bazno zaradi kompromisa med velikostjo dobljene panorame in časom, ki je potreben za njeno gradnjo. Z zmanjšanjem ločljivosti bi seveda postopek še pospešili, vendar pa bi bila tudi globinska slika ustrezno manjše ločljivosti.

5.6.3 Omejitev iskanja na epipolarni premici

Izhajamo iz simetričnega para panoram in želimo zgraditi globinsko panoramsko sliko. To lahko zgradimo za levo oko in/ali za desno oko. V nadaljevanju bomo gradili globinske slike za levo oko, dejstvo pa je, da globinsko sliko za desno oko dobimo s simetričnim postopkom, torej izhajamo iz slike za desno oko in iščemo korespondenčne točke na sliki za levo oko. Glede na to, da je širina panoramske slike veliko večja od širine standardne slike¹⁷, kar pomeni, da je iskanje korespondenčne točke potrebno izvršiti vzdolž zelo dolge epipolarne premice, želimo ta preiskovalni prostor čimbolj omejiti in tako pospešiti delovanje sistema.

Če izhajamo iz enačbe (5.10), lahko zapišemo dve ugotovitvi, ki nam preiskovalni prostor močno omejita:

- Teoretično je najmanjša možna ocena globine enaka $l_{\min} = r$, kar velja za primer, ko je $\theta = 0^\circ$. Seveda bi to pomenilo, da je neka točka na sceni hkrati vidna v koloni, ki prispeva k panorami za levo oko in hkrati v koloni, ki prispeva k panorami za desno oko. To pa je praktično nemogoče. Glede na smer vrtenja sistema velja, da je vedno neka točka na sceni, če opazujemo horizontalno os panoramske slike, najprej vidna na panorami za levo oko in šele nato na panorami za desno oko. Torej moramo počakati, da se točka iz kolone, ki prispeva k panorami za levo oko, s časom prestavi v kolono, ki prispeva k panorami za desno oko. Če je θ_0 kot, za katerega premikamo kamero, je $2\theta_{\min} = \theta_0$. Da pride točka, ki se nahaja v desni koloni in prispeva k panorami za levo oko, v levo kolono, moramo torej narediti obrat za vsaj en osnovni premik.

Na podlagi zapsanega velja, da korespondenčno točko na desni panorami iščemo od vključno horizontalne slikovne koordinate $x + \frac{2\theta_{\min}}{\theta_0} = x + 1$ naprej, pri čemer je x horizontalna slikovna koordinata točke na panorami za levo oko, za katero iščemo korespondenčno točko. Vrednost $+1$ torej dobimo na podlagi dejstva, da premik za θ_0 opisuje premik za eno kolono zajete slike.

¹⁷V našem primeru velja, da je razmerje med širino panoramske slike in širino standardne slike enako $1501/160$.

Za naš sistem tako na primer velja, da je najmanjša možna ocena globine l_{\min} , pri razdalji $r = 300$ mm, odvisna od kota φ (stran 47) (glej sliki 5.9 in 5.2 ter tabelo 5.1):

$$\begin{aligned} l_{\min}(2\varphi = 29,9625^\circ) &= 302 \text{ mm} \\ &\dots \\ l_{\min}(2\varphi = 3,6125^\circ) &= 318 \text{ mm.} \end{aligned}$$

- Teoretično velja, da ocena globine navzgor ni omejena, vendar iz enačbe (5.10) sledi, da imenovalec ne sme biti enak 0. Praktično to pomeni, da mora biti za oceno največje možne globine l_{\max} razlika $(\varphi - \theta_{\max})$ enaka vrednosti na intervalu $(0, \frac{\theta_0}{2})$. To dejstvo lahko zapišemo elegantneje kot: $\theta_{\max} = n \cdot \frac{\theta_0}{2}$, pri $n = \varphi \operatorname{div} \frac{\theta_0}{2}$ in $\varphi \bmod \frac{\theta_0}{2} \neq 0$.¹⁸

Če po analogiji z zgornjim primerom določanja začetne točke preiskovalnega prostora na epipolarni premici določimo končno točko preiskovalnega prostora na epipolarni premici, velja, da korespondenčno točko na desni panorami iščemo do vključno horizontalne slikovne koordinate $x + \frac{2\theta_{\max}}{\theta_0} = x + n$, pri čemer je x horizontalna slikovna koordinata točke na panorami za levo oko, za katero iščemo korespondenčno točko.

Tudi tukaj velja, da je največja možna ocena globine l_{\max} , pri razdalji $r = 300$ mm, odvisna od kota φ :

$$\begin{aligned} l_{\max}(2\varphi = 29,9625^\circ) &= 54687 \text{ mm} \\ &\dots \\ l_{\max}(2\varphi = 3,6125^\circ) &= 86686 \text{ mm.} \end{aligned}$$

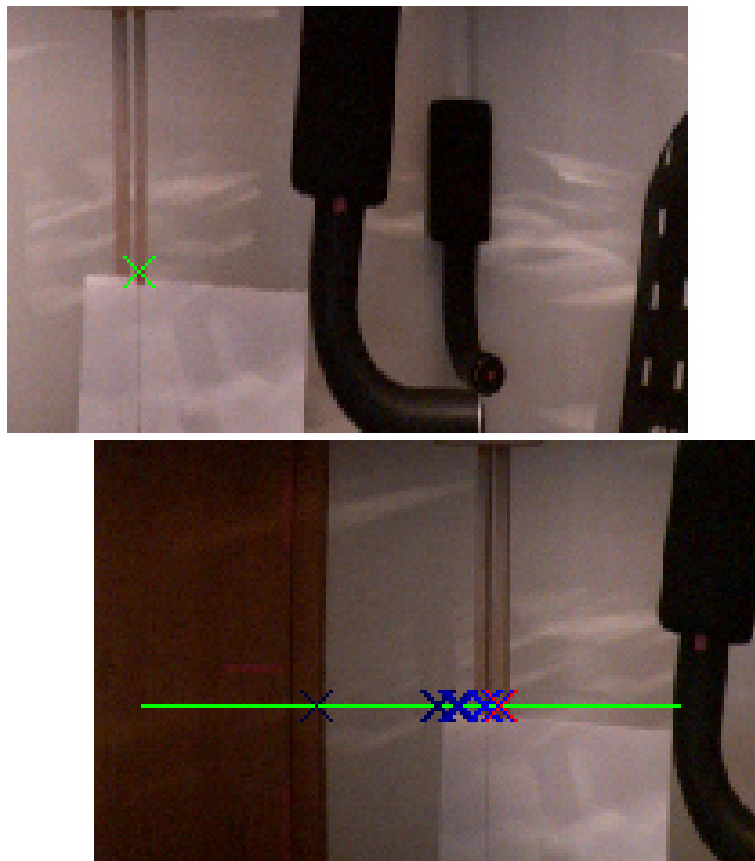
V nadaljevanju bomo pokazali, da v takšno oceno sicer ne moremo zaupati, vendar pa smo tukaj pokazali, da lahko tudi navzgor učinkovito omejimo preiskovalni prostor.

Tako na primer za naš sistem velja, da od omenjenih 1501 slikovnih elementov, kolikor znaša širina panorame, ki se nahajajo na epipolarni premici, pregledamo le $n = 149$ slikovnih elementov pri $2\varphi = 29,9625^\circ$ (slika 5.13) in le $n = 18$ slikovnih elementov pri $2\varphi = 3,6125^\circ$ (dodatek D).

Iz povedanega lahko sklepamo, da je postopek rekonstrukcije hitrejši, če je kot φ manjši, vendar pa bomo v poglavju 5.6.5 videli, da ima manjši kot φ tudi negativno lastnost.

Seveda pa lahko pri omejevanju preiskovalnega prostora uporabimo tudi druge lastnosti našega sistema, na primer, da zaradi tega, ker naš sistem ne zgradi panoramskih slik, ki bi opisovale kot 360° , ne moremo delati rekonstrukcije točk, ki

¹⁸Na srečo izbrane vrednosti za φ ne morejo dati rezultata 0.



Slika 5.13: Omejitev iskanja na epipolarni premici v primeru, ko je $2\varphi = 29,9625^\circ$. Na panorami za levo oko (zgornja slika) je z zelenim križcem označena točka, za katero iščemo korespondenčno točko. Na panorami za desno oko (spodnja slika) pa je z zeleno barvo označen del epipolarne premice, na kateri se mora nahajati korespondenčna točka. Najdena korespondenčna točka je označena z rdečim križcem. Z modrimi križci je označenih nekaj (maksimalno 10) prej najboljše ocenjenih točk, ki so predstavljale trenutno najboljšo korespondenčno točko, preden smo našli točko z maksimalno korelacijo.

ležijo na pasu širine 2φ na koncu panorame za levo oko, saj korespondenčnih točk na panorami za desno oko ni.

5.6.4 Problem prekrivanja

Problem prekrivanja je lepo razviden iz slike 5.14 (primerjaj sliki znotraj rdečih pravokotnikov). Vidimo lahko, da na eni sliki luč v ospredju prekriva predmete v ozadju, na drugi sliki pa te predmete vidimo v celoti. V takšnih primerih torej korespondenčne točke ne moremo najti, ker na drugi sliki ni vidna. Ker pa vedno najdemo najboljšo možno korespondenco, se lahko pri oceni globine zelo uštejemo, zato moramo na ustrezen način takšne primere izločiti iz nadaljne obdelave. To naredimo z že omenjenim postopkom vzvratne korelacije (poglavje 5.5.1) in primere, v katerih z vzvratno korelacijo ne najdemo ustrezne korespondence, izločimo.

Na sliki 5.14 je prikazano tudi omejevanje preiskovalnega prostora pri postopku vzvratne korelacije. Opazimo lahko, da tudi pri postopku iskanja vzvratne korespondenčne točke uporabimo enake lastnosti sistema za omejevanje preiskovalnega prostora, kot smo to storili pri iskanju korespondenčne točke (poglavje 5.6.3). Na podlagi prikazanega rezultata vzvratne korelacije, sistem za točko, za katero smo iskali korespondenco, ni ocenil globine, saj z vzvratno korelacijo nismo našli te točke, pač pa neko drugo. Opazimo lahko tudi, da se je sistem že pri iskanju korespondenčne točke zmotil (prava korespondenčna točka je bila ocenjena kot druga najboljša), pri iskanju vzvratne korespondenčne točke pa je našel pravo korespondenčno točko.

Seveda pa lahko v literaturi zasledimo tudi alternativne postopke za reševanje problema prekrivanja, ki pa največkrat zahtevajo tudi drugačne postopke za iskanje korespondenčnih točk [13, 27].

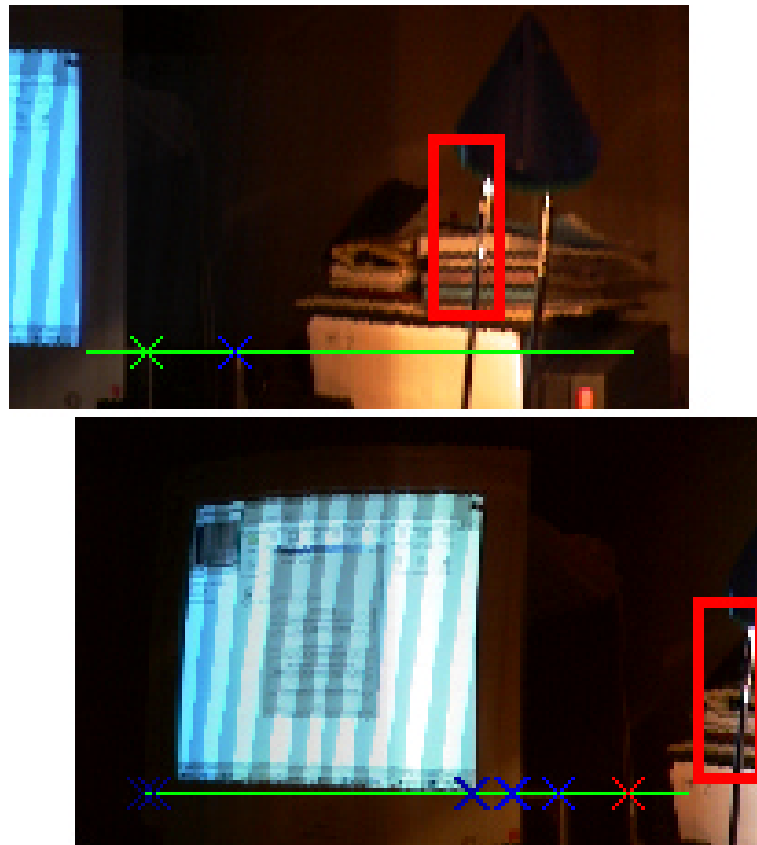
5.6.5 Zaupanje v ocenjeno globino

Na sliki 5.15 je podan graf funkcije globine l (enačba (5.10)) v odvisnosti od kota θ , na katerega lahko gledamo tudi kot na dispariteto našega sistema. Dosedanje ugotovitve o oceni globine l lahko povzamemo v dveh točkah:

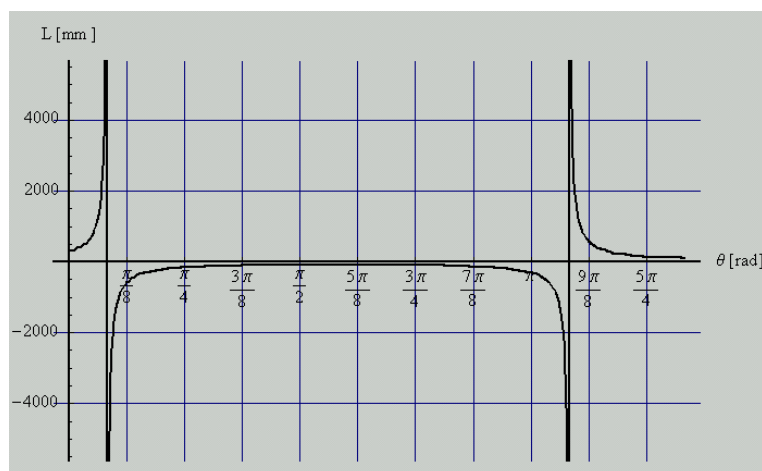
- bistven del grafa je prikazan na intervalu $\theta \in [0, \varphi]$, saj lahko le za te vrednosti teoretično korektno izračunamo globino l za neko točko na sceni in
- glede na kot θ_0 , ki predstavlja kot zasuka kamere med dvema zaporednima slikama, ki prispevata k simetričnemu paru panoram, lahko določimo najmanjšo možno oceno globine l_{\min} in največjo možno oceno globine l_{\max} (poglavje 5.6.3),

na podlagi grafa pa lahko zapišemo še dve dodatni trditvi:

- z večanjem kota θ se večja pomen napake za en slikovni element pri oceni velikosti tega kota (enačba 5.11) in



Slika 5.14: Problem prekrivanja je lepo opazen znotraj rdečih pravokotnikov. Glede križcev in premic velja vse, kar smo zapisali pri opisu slike 5.13, le da je tukaj prikazano tudi omejevanje preiskovalnega prostora pri postopku vzvratne korelacije: sedaj za točko, ki je označena z rdečim križcem na panorami za desno oko (spodnja slika), iščemo ustrezno vzvratno korespondenčno točko na panorami za levo oko (zgornja slika). Iskana točka se mora nahajati na delu epipolarne premice, ki je na panorami za levo oko označen z zeleno barvo. Z modrim križcem pa na panorami za levo oko označimo najdeno vzvratno korespondenčno točko.



Slika 5.15: Graf funkcije globine l (enačba (5.10)) v odvisnosti od kota θ pri kotu $2\varphi = 29,9625^\circ$ in razdalji $r = 30$ cm.

- glede na vse bolj strmo naraščanje funkcije l , ko se kot θ približuje kotu φ , je smiselno določiti manjšo največjo možno oceno globine l_{maks} , saj je napaka za en slikovni element pri oceni kota θ na tem delu ogromna.

V nadaljevanju bomo predstavili argumente, ki nas bodo prepričali v veljavnost obeh trditev.

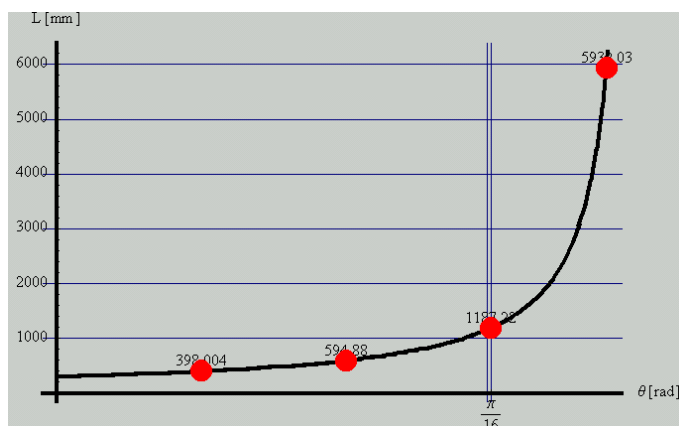
Pomen napake za en slikovni element pri oceni kota θ

Na graf na sliki 5.15 lahko gledamo tudi kot na graf odvisnosti napake v oceni globine l od napake v oceni kota θ , ki ga ocenimo na podlagi najdenih korespondenčnih točk. To pomeni, da lahko naredimo analizo velikosti napake v oceni globine l glede na vrednost kota θ . Preden pa ilustriramo pomen napake za en slikovni element pri oceni kota θ , si oglejmo še grafa na sliki 5.16. Grafa prikazujeta odvisnost funkcije globine l od kota θ pri različnih vrednostih φ . Iz grafov sledi, da funkcija globine l narašča počasneje pri večjem kotu φ . Ta lastnost zmanjša napako v oceni globine l pri večjem kotu φ , vendar pa to zmanjšanje napake pride še toliko bolj do izraza, če vemo, da je abscisna os θ diskretizirana na intervale velikosti $\frac{\theta_0}{2}$, kot je to razvidno iz grafov. Če primerjamo interval širine $\frac{\theta_0}{2}$ na obeh grafih glede na širino definicijskega območja $\theta \in [0, \varphi]$, vidimo, da je interval širine $\frac{\theta_0}{2}$ pri večjem kotu φ veliko manjši. To pa pomeni, da je tudi napaka za en slikovni element pri oceni kota θ veliko manjša pri večjem kotu φ , saj, kot smo že zapisali, prav premik za kot θ_0 opisuje premik za en slikovni element v panoramski sliki.

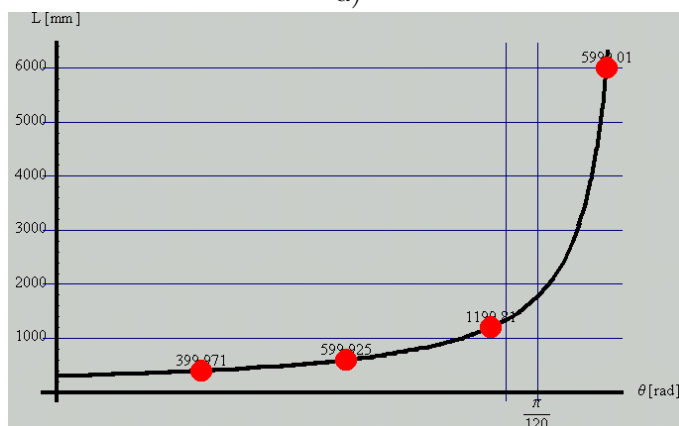
Zaradi diskretizacije osi θ (slika 5.16) na intervale širine $\frac{\theta_0}{2}$ (v našem primeru velja: $\theta_0 = 0,2^\circ$) velja, da je število možnih ocen globine sorazmerno s kotom φ : za

kot $2\varphi = 29,9625^\circ$ velja, da lahko izračunamo največ ($\varphi \text{ div } \frac{\theta_0}{2} =$)149 ocen globine, za kot $2\varphi = 3,6125^\circ$ pa velja, da lahko izračunamo največ le 18 ocen globine (dodatek D).

S tem smo lepo opisali negativno lastnost majhnega kota φ , ki smo jo omenili v poglavju 5.6.3.



a)



b)

Slika 5.16: Grafa odvisnosti funkcije globine l od kota θ pri razdalji $r = 30$ cm in različnih vrednostih kota φ : a) $2\varphi = 29,9625^\circ$, b) $2\varphi = 3,6125^\circ$. Za lažjo primerjavo hitrosti naraščanja funkcije globine l so na grafih prikazane štiri ocene globine: $\theta \in \{\frac{\varphi}{4}, \frac{\varphi}{2}, \frac{3\varphi}{4}, \frac{19\varphi}{20}\}$. Za lažjo primerjavo velikosti napake za en slikovni element pri oceni kota θ pa je med vertikalnima črtama pri tretji točki prikazan interval širine $\frac{\theta_0}{2} = 0,1^\circ$.

Za konec tega poglavja ilustrirajmo pomen napake za en slikovni element pri oceni kota θ : Zanima nas, kakšna je ta napaka za kot θ , ki je na začetku defini-

cijskega območja $\theta \in [0, \varphi]$ in kakšna je ta napaka za kot θ , ki je na koncu tega definicijskega območja. V ta namen izberimo kota $\theta_1 = \frac{\varphi}{4}$ in $\theta_2 = \frac{7\varphi}{8}$. Zanima pa nas seveda tudi narava napake za različni kot φ , zato bomo za kot φ izbrali, zdaj že standardni vrednosti, $2\varphi = 29,9625^\circ$ in $2\varphi = 3,6125^\circ$. Rezultati v tabeli 5.2 podajajo ovrednotenje napake v oceni globine l za omenjene vrednosti parametrov θ in φ .

	$\theta - \frac{\theta_0}{2}$	θ	$\theta + \frac{\theta_0}{2}$
l [mm]	394,5	398	401,5
Δl [mm]	3,5		
(napaka)	3,5		

a) $\theta = \theta_1 = \frac{\varphi}{4}$, $2\varphi = 29,9625^\circ$

	$\theta - \frac{\theta_0}{2}$	θ	$\theta + \frac{\theta_0}{2}$
l [mm]	372,5	400	431,8
Δl [mm]	27,5		
(napaka)	31,8		

b) $\theta = \theta_1 = \frac{\varphi}{4}$, $2\varphi = 3,6125^\circ$

	$\theta - \frac{\theta_0}{2}$	θ	$\theta + \frac{\theta_0}{2}$
l [mm]	2252,9	2373,2	2507
Δl [mm]	120,3		
(napaka)	133,8		

c) $\theta = \theta_2 = \frac{7\varphi}{8}$, $2\varphi = 29,9625^\circ$

	$\theta - \frac{\theta_0}{2}$	θ	$\theta + \frac{\theta_0}{2}$
l [mm]	1663	2399,6	4307,4
Δl [mm]	736,6		
(napaka)	1907,8		

d) $\theta = \theta_2 = \frac{7\varphi}{8}$, $2\varphi = 3,6125^\circ$

Tabela 5.2: Ovrednotenje napake za en slikovni element pri oceni kota θ in posledično po enačbi (5.10) pri oceni globine l , za različne vrednosti kotov θ in φ , pri čemer je razdalja r enaka 30 cm, kot θ_0 pa je enak $0,2^\circ$. (glej tekst)

Iz rezultatov lahko vidimo, da je pri majhnem kotu φ napaka veliko večja kot pri večjem kotu φ . Dejstvo je, da želimo pri rekonstrukciji dobiti verodostojno oceno globine (vsaj do neke mere natančno), hkrati pa želimo, da je postopek rekonstrukcije hiter. To pa sta želji, kateri lahko v našem sistemu zadovoljimo le s kompromisom

med natančnostjo in hitrostjo rekonstrukcije. Če bi želeli doseči največjo možno natančnost, potem bi uporabili največjo možno velikost kota φ , vendar bi s tem morali preiskovati tudi največji del epipolarne premice, kar bi postopek rekonstrukcije upočasnilo. Podobno lahko sklepamo tudi v primeru, če bi želeli doseči večjo hitrost sistema. Velja torej, da je hitrost rekonstrukcije v obratnem sorazmerju z njeno natančnostjo. Več o časovni analizi in analizi natančnosti gradnje globinskih panoramskih slik bomo zapisali v poglavju 5.6.6, na tem mestu pa omenimo le še dejstvo, ki je pravtako razvidno iz rezultatov, da je napaka v primeru kota θ_2 prevelika. To pa pomeni, da lahko določimo bolj smiselno maksimalno globino l_{maks} , v katere oceno še zaupamo.

Rezultati so podani za primer, ko v panoramsko sliko mozaičimo standardne slike ločljivosti 160×120 slikovnih elementov ($\theta_0 = 0, 2^\circ$) in je oddaljenost r optičnega centra kamere od središča vrtenja sistema enaka 30 cm. S spreminjanjem teh dveh parametrov se seveda spreminja tudi velikost napake:

- Z večanjem ločljivosti zajete slike manjšamo kot θ_0 , torej manjšamo kot zasuk kamere med dvema zaporednima slikama, ki prispevata k panorami. Za približno¹⁹ enak faktor kot povečamo (zmanjšamo) ločljivost zajete slike, se zmanjša (poveča) tudi vrednost napake Δl , ob tem pa velja, da postopek rekonstrukcije traja za približno enak faktor več (manj) časa. Na večanje ločljivosti lahko gledamo tudi kot na postopek *pod-točkovne natančnosti* (ang. *sub-pixel accuracy*), saj je na primer slikovni element na sliki velikosti 160×120 slikovnih elementov na štirikrat večji sliki velikosti 640×480 slikovnih elementov predstavljen s 16. slikovnimi elementi (tabela velikosti 4×4).
- Za enak faktor kot povečamo (zmanjšamo) razdaljo r , se poveča (zmanjša) tudi ocena globine l in vrednost napake Δl . Zaradi spreminjanja razdalje r se postopek rekonstrukcije ne izvrši nič počasneje oziroma hitreje. Z večanjem razdalje r se veča tudi največja možna ocena globine l_{maks} . V praksi to pomeni, da lahko z večjo razdaljo r rekonstruiramo večje prostore. Dejstvo je, da je naš sistem uporaben le za rekonstrukcijo (majhnih) prostorov in ne poljubnih scen, saj praktično ne moremo oceniti globine zelo oddaljenih točk na sceni, ker je vrednost napake Δl za te točke prevelika. (glej poglavje 5.1.2)

Določitev največje globine, v katere oceno še zaupamo

V poglavju 5.6.3 smo ocenili najmanjšo možno globino l_{min} in največjo možno globino l_{maks} , zapisali pa nismo, kakšen je pomen napake za en slikovni element pri oceni kota θ za ti dve oceni. Oglejmo si torej vrednost te napake (Δl) za ti dve oceni: Δl_{min} izračunamo kot absolutno vrednost razlike med l_{min} in globino l , za katero velja, da je kot θ za kot $\frac{\theta_0}{2}$ večji od kota θ_{min} :

¹⁹Zaradi uporabe funkcije \sin v enačbi za oceno globine l (enačba (5.10)).

$$\Delta l_{\min} = |l_{\min}(\theta_{\min}) - l(\theta_{\min} + \frac{\theta_0}{2})| = |l_{\min}(\frac{\theta_0}{2}) - l(\theta_0)|.$$

Δl_{\max} pa izračunamo kot absolutno vrednost razlike med l_{\max} in globino l , za katero velja, da je kot θ za kot $\frac{\theta_0}{2}$ manjši od kota θ_{\max} :

$$\Delta l_{\max} = |l_{\max}(\theta_{\max}) - l(\theta_{\max} - \frac{\theta_0}{2})| = |l_{\max}(n\frac{\theta_0}{2}) - l((n-1)\frac{\theta_0}{2})|,$$

kjer je spremenljivka n pozitivno celo število, ki jo dobimo po enačbi: $n = \varphi \operatorname{div} \frac{\theta_0}{2}$. Iz tabele 5.3 so razvidne vrednosti napak glede na kot φ .

	$2\varphi = 29,9625^\circ$	$2\varphi = 3,6125^\circ$
Δl_{\min} [mm]	2	19
Δl_{\max} [mm]	30172	81587

Tabela 5.3: Pomen napake (Δl) za en slikovni element pri oceni kota θ za oceno najmanjše možne globine l_{\min} in največje možne globine l_{\max} glede na kot φ . (glej tekst)

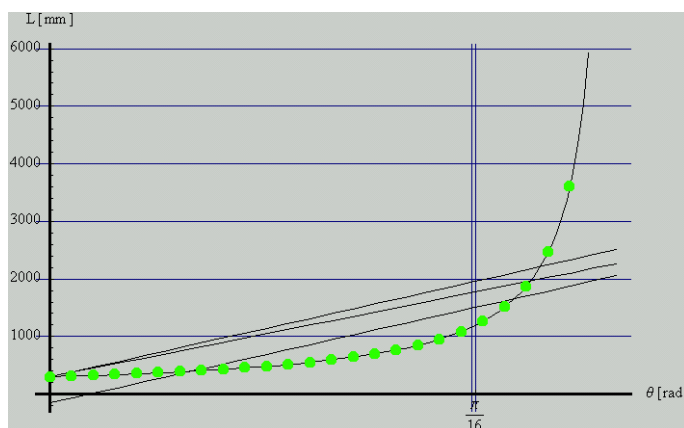
Dobljeni rezultati potrjujejo že zapisane ugotovitve (poglavje 5.6.5), zapišemo pa lahko še dve dodatni ugotovitvi:

- Vrednost napake Δl_{\max} je nesprejemljivo visoka in to neglede na kot φ .
- Kot φ je vedno odvisen od horizontalnega zornega kota α (enačba (5.4)), le-ta pa na standardnih kamerah dosega vrednosti do približno 40° . To pomeni, da smo v praksi pri oceni globine omejeni s kotom α , saj v najboljšem primeru velja: $\varphi_{\max} = \frac{\alpha}{2}$. Torej je naš sistem res uporaben le v primeru rekonstrukcije majhnih prostorov.

Na podlagi prve ugotovitve moramo smiselno zmanjšati oceno za največjo možno globino l_{\max} . To pomeni, da moramo definirati zgornjo mejo dovoljene napake (Δl) za en slikovni element pri oceni kota θ in s tem posledično določimo največjo globino, v katere oceno še zaupamo. Zgornjo mejo dovoljene napake lahko določimo na dva načina: enostavno podamo njeno vrednost, ali pa zapišemo njeno velikost v odstotkih glede na ocenjeno globino. Če dejanska napaka presega dano vrednost, v ocenjeno globino ne zaupamo. Ilustrirajmo določitev zgornje meje dovoljene napake: Če za napako podamo vrednost 10 cm in si ogledamo primer a) s slike 5.16 ter dodatek D, vidimo, da takšno napako dosežemo pri vrednosti $l = 213,5$ cm (upoštevamo diskretizacijo abscisne osi na intervale širine $\frac{\theta_0}{2}$). Enako velja tudi v primeru, če za napako podamo vrednost 5%. Tako bi na primer določili največjo globino, v katere oceno še zaupamo, kot $l_{\max} = 213,5$ cm. Splošen algoritem za določanje največje globine, če za napako podamo vrednost Δl_{dana} , lahko v psevdo kodi zapišemo kot:

za $i = 1$ do $((\varphi \operatorname{div} \frac{\theta_0}{2}) - 1)$ delaj
 $\theta = i \cdot \frac{\theta_0}{2}$
 izračunaj $l_1 = l(\theta)$
 izračunaj $l_2 = l(\theta + \frac{\theta_0}{2})$
 če $(l_2 - l_1 > \Delta l_{\text{dana}})$ potem
 $l_{\text{maks}} = l_1$
 izhod iz zanke

Lahko pa poskušamo določiti največjo globino tudi direktno: s pomočjo linearne interpolacije točk na krivulji funkcije globine l (enačba (5.10)) določimo največjo globino kot presečišče dobljene premice in krivulje (slika 5.17).



Slika 5.17: Direktno določevanje največje globine, v katere oceno še zaupamo: s pomočjo linearne interpolacije točk na krivulji funkcije globine l določimo največjo globino kot presečišče dobljene premice in krivulje. 25 točk je enakomerno porazdeljenih na intervalu $[0, \frac{19\varphi}{20}]$ abscisne osi s korakom 0,01 rad ($2\varphi = 29,9625^\circ$). (Zakaj do $\frac{19\varphi}{20}$? Ker gredo naprej točke zelo hitro proti neskončnosti.) Prikazani so trije primeri interpolacije: najnižja premica predstavlja interpolacijo s splošno enačbo premice $y = k_1x + n$, sledi interpolacija z enačbo premice $y = k_2x + 300$ ($r=300$ mm), najvišja premica pa predstavlja interpolacijo z enačbo premice $y = k_1x + 300$.

Razmišljanje

Za konec tega poglavja o zaupanju v ocenjeno globino zapišimo dve ideji oziroma razmišljanji:

Dodatno omejevanje preiskovalnega prostora na epipolarni premici. — V prejšnjem poglavju smo smiselno zmanjšali oceno za največjo možno globino l_{maks} , sedaj pa nas zanima, ali lahko s tem podatkom še dodatno omejimo preiskovalni prostor na epipolarni premici. Recimo, da opazujemo neko točko

na sceni, katere globina je večja od l_{maks} . Če iščemo korespondenčno točko po epipolarni premici in sedaj preiskovalni prostor na njej omejimo navzgor s kotom θ_{maks} (glej poglavje 5.6.3), katerega izrazimo iz enačbe (5.10), potem sistem postavi v korespondenco dve točki, ki imata največjo vrednost korelacije, nista pa dejansko korespondenčni, saj dejanske korespondenčne točke ni v preiskovalnem prostoru. Torej velja, da je dejanska globina $l > l_{\text{maks}}$, sistem pa vrne rezultat $l \leq l_{\text{maks}}$. Torej z največjo možno globino l_{maks} ne moremo dodatno omejiti preiskovalnega prostora na epipolarni premici. Ali je to res? Pravzaprav bi lahko tako dodatno omejili preiskovalni prostor, če bi upoštevali tudi oceno, ki jo vrne normalizirana korelacija: če bi bila ta ocena večja od nekega praga, potem lahko zaupamo v ocenjeno globino, sicer pa ne. Vprašanje pa je, za koliko nam bi to zmanjšalo preiskovalni prostor in kakšni so negativni učinki tega zmanjšanja? Ilustrirajmo odgovor na že omenjeni največji globini $l_{\text{maks}} = 213,5$ cm (poglavje 5.6.5), pri vrednostih: $r = 30$ cm, $2\varphi = 29,9625^\circ$, $\theta_0 = 0,2^\circ$. Če izrazimo kot θ_{maks} kot večkratnik kota $\frac{\theta_0}{2}$, dobimo, da je $\theta_{\text{maks}} = 12,9^\circ$. Sedaj zapišimo kot φ kot večkratnik kota $\frac{\theta_0}{2}$ in ga označimo s $\theta_{\text{maksstara}}$. Za kot $\theta_{\text{maksstara}}$ velja, da je manjši od kota φ . Razlika kotov $2\theta_{\text{maksstara}} - 2\theta_{\text{maks}}$ je 4° , kar glede na kot θ_0 predstavlja 20 slikovnih elementov. Torej se preiskovalni prostor zmanjša iz 149 slikovnih elementov (poglavje 5.6.3) na 129 slikovnih elementov. V tem primeru dosežemo torej dokaj majhno zmanjšanje preiskovalnega prostora, negativni učinek zmanjšanja pa je, da lahko normalizirana korelacija vrne dobro oceno tudi za točki na sceni, ki dejansko nista v korelaciji, torej se ocena globine kljub očitni napaki zapiše v globinsko sliko. Poleg tega pa je nastavljanje praga dokaj nevhvaležno opravilo.

V primeru, če bi l_{maks} nastavili na 101,5 cm, bi zmanjšanje bilo nekoliko večje, skupaj za 43 slikovnih elementov, vendar pa lahko delamo rekonstrukcijo le majhnih prostorov oziroma le predmetov v naši neposredni bližini. Dodatno omejevanje preiskovalnega prostora iz praktičnega vidika torej ni priporočljivo, saj:

- preiskovalni prostor ni veliko manjši
- normalizirana korelacija lahko dobro oceni tudi nekorrespondenčni točki
- obstaja problem nastavitve praga.

Kot pa bomo videli v poglavju 5.6.6, lahko zaupanje v ocenjeno globino spremenimo tudi na podlagi uteževanja le-te z oceno normalizirane korelacije.

Rekonstrukcija na osnovi premikanja središča vrtenja našega sistema in gradnje panoramskih slik iz središčnih stolpcev. — Če želimo z našim sistemom delati rekonstrukcije večjih prostorov, moramo povečati dispariteto, ki je pri nas omejena z največjim kotom φ , ki ga lahko dosežemo glede na

horizontalni zorni kot α . Ker je horizontalni zorni kot α omejen (lastnost standardne kamere), lahko dispariteto povečamo le tako, da premikamo središče vrtenja našega sistema. Takšen način rekonstrukcije je opisan v [13]: Sistem rekonstruira sceno tako, da najprej ustvari panoramsko sliko enako kot naš sistem, vendar iz središčnih kolon, nato pa se sistem premakne²⁰ na novo lokacijo in tam zgradi še eno panoramsko sliko. Ti dve panoramski sliki predstavljata vhod v postopek rekonstrukcije. Dobra lastnost tako zasnovanega sistema je, da je z njim možno oceniti večjo globino, slabih lastnosti pa je več:

- na premici, ki povezuje točki središč vrtenja sistema, je dispariteta enaka nič (0), torej globine ne moremo oceniti povsod, hkrati pa se zaupanje v oceno globine spreminja glede na oddaljenost točke na sceni od te premice (bliže je točka premici, manjše je zaupanje),
- zaradi premikanja sistema je čas od začetka zajemanja panoramskih slik do zaključka rekonstrukcije neprimerno daljši kot v primeru našega sistema²¹,
- epipolarne “premice” so sinusne krivulje.

Cilj sistema, opisanega v [13], je bil sestaviti globalni načrt prostora, po katerem se premika sistem, ki je pritrjen na robota. Globinske slike sistem gradi tako, da na slikah išče vertikalne robove s Sobelovim filtrom in za njih išče korespondenčne točke po epipolarnih krivuljah. Tako zgradi lokalne redke globinske slike, ki jih nato združi v globalni načrt prostora, ki je pravitako redek in izpostavlja vertikalne robove v prostoru.

5.6.6 Rezultati

Rezultati našega sistema, predstavljeni v tem poglavju, bodo osvetlili proces rekonstrukcije s časovnega vidika, natančnost rekonstrukcije, slikovni material pa bo potrdil marsikatero ugotovitev, ki smo jo zapisali do sedaj. Osnovne informacije o implementaciji sistema si lahko bralec ogleda v dodatku E, preden si ogledamo

²⁰Sistem pravzaprav najprej oceni globino kot to naredimo mi, le da ne uporablja korelacije za iskanje korespondenčnih točk na epipolarni premici, pač pa uporablja korelacijo za sledenje značilkam (vertikalnim robovom) na zajeti sliki in tako generira redko globinsko sliko. Ta postopek je časovno zelo kompleksen. Nato naj bi sistem na podlagi teh ocen globin sam določil smer in velikost premika. Ker pa v gibanje robota, na katerega je pritrjen sistem, ne zaupamo povsem (na primer zaradi zdrsa koles), avtorji sistem premikajo ročno, saj moramo poznati, za korektno oceno globine, resnično razdaljo med središčema vrtenja.

²¹Na opisan sistem bi lahko gledali tudi kot na dva naša sistema, ki sta eden od drugege oddaljena za neko razdaljo. V tem primeru bi bilo zajemanje panoramskih slik enako hitro kot pri našem sistemu, vendar pa bi imeli opravka z dinamično sceno, zaradi hkratnega premikanja nasproti stoječega sistema. Takšne panorame pa, kot omenjeno (poglavje 5.3), niso zaželjene kot vhod v postopek rekonstrukcije.

rezultate, pa izpostavimo tiste lastnosti sistema, na katere bomo v nadaljevanju posebej pozorni:

- Zanimajo nas rezultati ob iskanju korespondenčnih točk brez in s postopkom vzvratne korespondence.
- Zanimajo nas rezultati ob iskanju korespondenčnih točk pri kotu $2\varphi = 29,9625^\circ$ in kotu $2\varphi = 3,6125^\circ$.
- Zanimajo nas rezultati gradnje gostih in redkih globinskih slik.
- Zanima nas, kako lahko povečamo zaupanje v ocenjeno globino na podlagi uteževanja globine z oceno, ki nam jo vrne normalizirana korelacija.
- Zanima nas, kako in kdaj je smiselno zapolni vrzeli v gosti globinski sliki, ki nastanejo zaradi uporabe postopka vzvratne korespondence.

V nadaljevanju bomo korespondenčne točke iskali s postopkom normalizirane korelacije, saj je dal boljše rezultate kot postopek vsote kvadratov razlik.

Glede na ugotovitev, da smo zelo omejeni z velikostjo prostora, ki ga lahko rekonstruiramo, smo za testni prostor izbrali sobo velikosti 307×295 cm, kjer je bila največja razdalja v horizontalni smeri od središča vrtenja sistema enaka 229 cm (slika 5.23). Sistem vrne oceno globine v milimetrih, da pa lahko razdaljo zapišemo v globinsko sliko (namenjeni vizualizaciji rezultatov), moramo to oceno preslikati na interval $[0,255]$, ki predstavlja zalogo vrednosti 8-bitne sivinske slike. Glede na to, da je maksimalna razdalja v sobi 229 cm, ocenjeno globino najprej zaokrožimo na centimeter natančno in jo zapišemo v globinsko sliko tako, da so točke, ki so ocenjene kot bližnje, predstavljene s svetlo sivino, točke, ki so ocenjene kot daljne, pa so predstavljene s temno sivino. Pri tem je upoštevano tudi dejstvo, da nobena globina ne more biti manjša od razdalje r . S črno barvo so označene točke, ki nimajo pripisane ocene globine zaradi uporabe postopka vzvratne korespondence, ali gradnje redke globinske slike, z rdečo barvo pa so označene točke, katerih ocena je presegala vrednost 255 cm.

Naš sistem dela z barvnimi slikami, torej delamo rekonstrukcijo na podlagi informacije v treh barvnih kanalih (barvni prostor RGB), kar pomeni, da bi bil, v primeru obdelave sivinskih slik, naš sistem v postopku rekonstrukcije hitrejši.

Časovna analiza gradnje gostih globinskih slik

Gradimo goste globinske slike, kar pomeni, da poskušamo za vsako točko na panorami za levo oko najti ustrezno korespondenčno točko na panorami za desno oko. V tabeli 5.4 so predstavljeni časi gradnje globinskih slik v štirih primerih:

- uporabimo postopek vzvratne korespondence, kot $2\varphi = 29,9625^\circ$

- ne uporabimo postopka vzvratne korespondence, kot $2\varphi = 29,9625^\circ$
- uporabimo postopek vzvratne korespondence, kot $2\varphi = 3,6125^\circ$
- ne uporabimo postopka vzvratne korespondence, kot $2\varphi = 3,6125^\circ$.

Na sliki 5.18 so predstavljene ustrezne globinske slike. Če hkrati z rekonstrukcijo v drugo sliko zapisujemo tudi ocene postopka normalizirane korelacije²², dobimo sliko zaupanja v ocenjene globine (slika 5.19). Zaradi tega se čas rekonstrukcije ne spremeni bistveno.

	čas rekonstrukcije [ur/min./sek.]
uporabimo vzvratno korespondenco $2\varphi = 29,9625^\circ$	6/42/20
ne uporabimo vzvratne korespondence $2\varphi = 29,9625^\circ$	3/21/56
uporabimo vzvratno korespondenco $2\varphi = 3,6125^\circ$	0/52/56
ne uporabimo vzvratne korespondence $2\varphi = 3,6125^\circ$	0/29/6

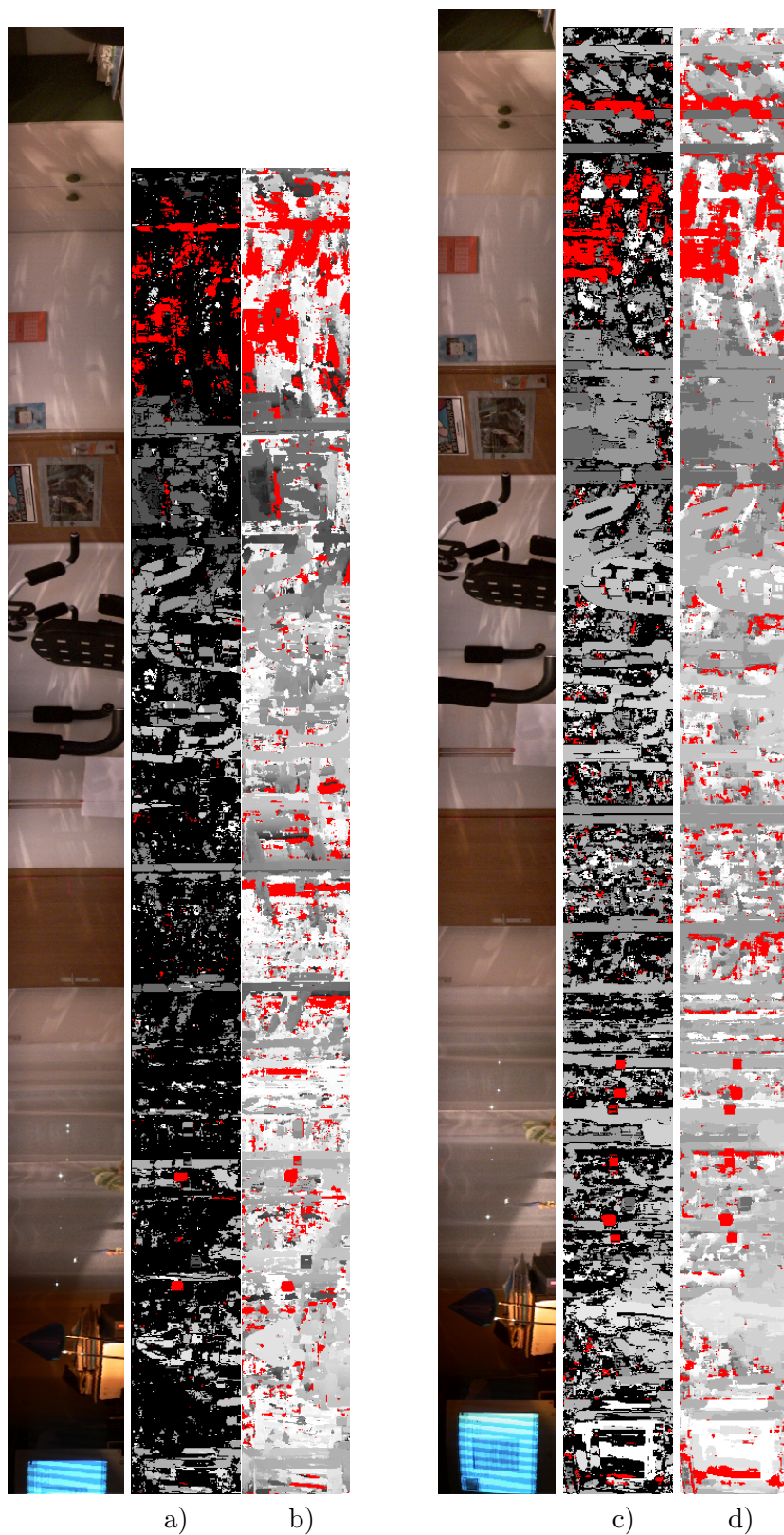
Tabela 5.4: Časi gradnje gostih globinskih slik v štirih primerih.

Časi rekonstrukcije iz tabele 5.4 so med seboj tesno povezani:

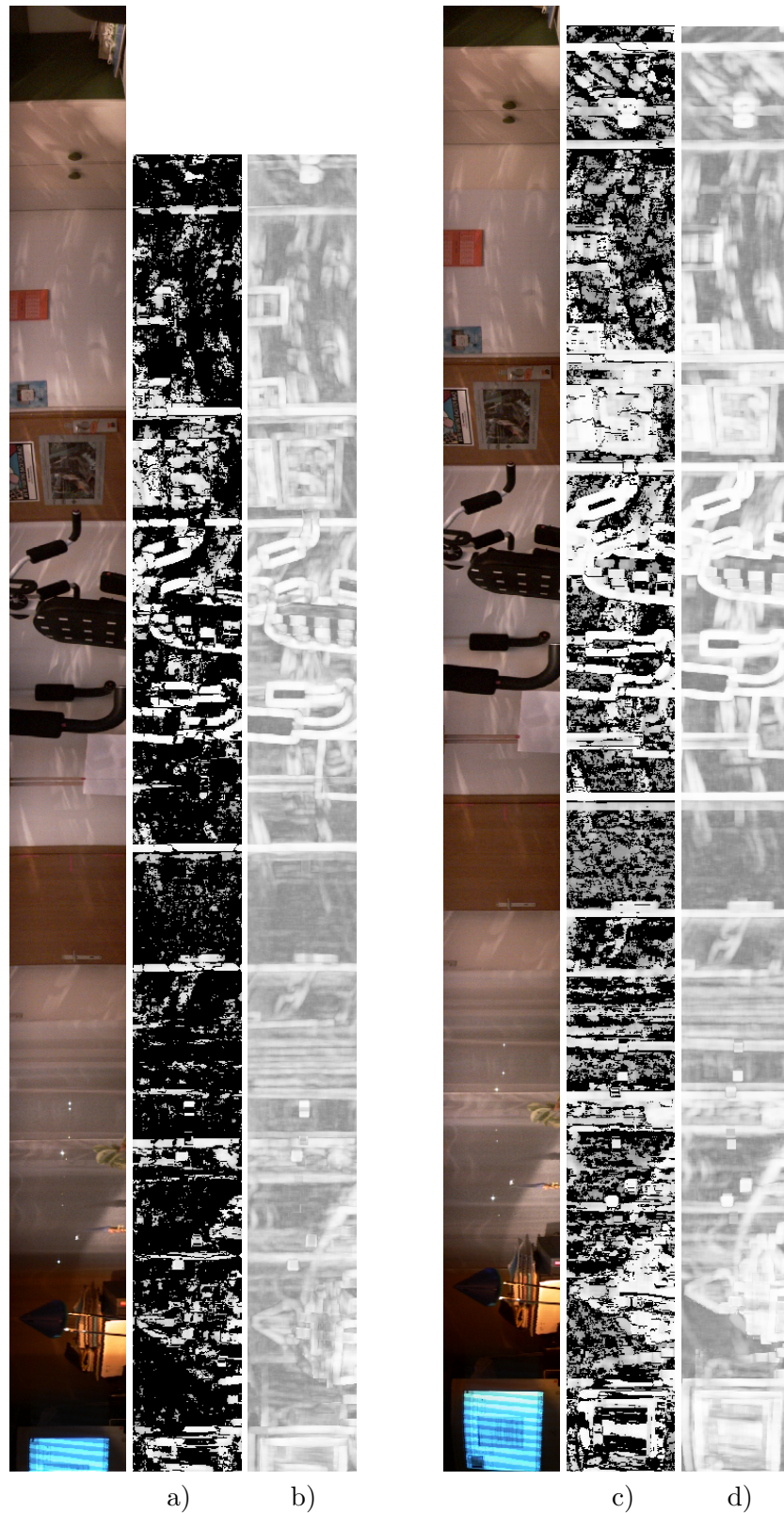
- če ne uporabimo postopka vzvratne korespondence, se čas rekonstrukcije približno prepolovi,
- na podlagi zapisane ugotovitve, da na epipolarni premici pregledamo le 149 slikovnih elementov pri $2\varphi = 29,9625^\circ$ in le 18 slikovnih elementov pri $2\varphi = 3,6125^\circ$ (poglavje 5.6.3), pa se čas rekonstrukcije zmanjša za približno osemkrat.

Na ustvarjenih globinskih slikah (slika 5.18) lahko opazimo, da so robovi na sceni dobro izpostavljeni in, kot bomo videli v nadaljevanju, je tudi ocena njihove globine dobra, hkrati pa lahko opazimo probleme pri oceni globine, ki so povezani z uniformnimi deli na sceni. Na teh delih (na primer na beli steni) zelo težko najdemo pravo korespondenco, zato so vrzeli v gosti globinski sliki, ki nastanejo zaradi uporabe postopka vzvratne korespondence, pričakovane. Na panoramskih slikah je prisoten tudi ekran monitorja, ki zaradi osveževanja slike predstavlja dinamičen del

²²Ocena postopka normalizirane korelacije je vedno na intervalu $[-1,1]$ (poglavje 5.5.1), zato moramo to oceno preslikati na interval $[0,255]$, ki predstavlja zalogo vrednosti 8-bitne sivinske slike. Velja, da bolj zaupamo v tisto oceno, ki je na sliki bolj svetla.



Slika 5.18: Ustvarjene goste globinske slike: a) uporabimo postopek vzratne korespondence, kot $2\varphi = 29,9625^\circ$, b) ne uporabimo postopka vzratne korespondence, kot $2\varphi = 29,9625^\circ$, c) uporabimo postopek vzratne korespondence, kot $2\varphi = 3,6125^\circ$, d) ne uporabimo postopka vzratne korespondence, kot $2\varphi = 3,6125^\circ$.



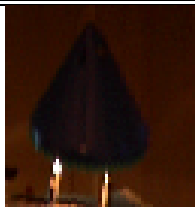








Slika 5.19: Slike zaupanja v ocenjeno globino (slika 5.18) na podlagi ocene postopka normalizirane korelacije. Velja, da bolj zaupamo v tisto oceno, ki je na sliki bolj svetla. a) Uporabimo postopek vzratne korespondence, kot $2\varphi = 29,9625^\circ$, b) ne uporabimo postopka vzratne korespondence, kot $2\varphi = 29,9625^\circ$, c) uporabimo postopek vzratne korespondence, kot $2\varphi = 3,6125^\circ$, d) ne uporabimo postopka vzratne korespondence, kot $2\varphi = 3,6125^\circ$.

scene. Torej na ustreznem delu globinske slike ne moremo pričakovati korektnih ocen globine.

Dokazali smo, da je epipolarna premica vrstica panorame, na kateri naj bi bila korespondenčna točka. Splošna praksa pri preiskovanju epipolarne premice pa je, da preiščemo tudi njeno bližnjo okolico. To v splošnem naredimo zaradi prisotnosti šuma pri zajemanju slik, pri izračunu same epipolarne geometrije in pri impliciranju epipolarne geometrije na slike. Glede na to, da imamo v našem primeru opraviti pravzaprav z idealno epipolarno geometrijo (vrstice panoram), vertikalna ločljivost panoram pa ni velika (118 slikovnih elementov), iščemo korespondenčno točko zgolj po ustrezni vrstici panorame. Sicer bi bil čas rekonstrukcije za ustrezen faktor daljši. Mogoče bi zahtevali, ki jo postavljamo s postopkom vzvratne korelacije, lahko malo omilili: namesto, da moramo z vzvratno korelacijo najti izbrano točko, za katero smo iskali korespondenčno točko, dopustimo neko majhno odstopanje od te izbrane točke. Seveda to odstopanje predstavlja dvorezen meč: dobra lastnost uporabe odstopanja je, da ima več točk na globinski sliki pripisano oceno globine, slaba lastnost pa je, da s tem zagotovo povečamo napako v ustvarjeni globinski sliki. Na sliki 5.20 je prikazana primerjava rezultatov gradnje globinskih slik dela scene pri različnih vrednosti parametrov, ki med drugim določajo: bližnjo okolico epipolarne premice, odstopanje od izbrane točke in izbiro korelacijske metode (vsota kvadratov razlik ali normalizirana korelacija). Rezultati potrjujejo zapisane ugotovitve:

- Postopek vsote kvadratov razlik daje slabše rezultate kot postopek normalizirane korelacije, deluje pa hitreje.
- Zaradi preiskovanja okolice epipolarne premice se čas rekonstrukcije podaljša za ustrezen faktor, rezultat pa je boljši.
- Z dovoljenim odstopanjem od izbrane točke dobimo gostejšo globinsko sliko, ki pa vsebuje večjo napako.
- Razmerja med zapisanimi časi rekonstrukcije so pričakovana.

Vrzeli v gosti globinski sliki, ki nastanejo zaradi uporabe postopka vzvratne korespondence, lahko poskušamo smiselno zapolniti z interpolacijo: Za neko točko na globinski sliki, ki nima pripisane ocene globine, poiščemo n najbližjih točk, ki imajo pripisano oceno globine. Nato interpoliramo globino omenjene točke na podlagi ocenjenih globin najdenih najbližjih točk tako, da vsaka točka prispeva k interpolirani oceni globine delež svoje ocene globine, ki je sorazmeren z oddaljenostjo med najdeno točko in točko, za katero interpoliramo globino. Na ta način je smiselno zapolniti majhne vrzeli, velikih vrzeli pa načeloma ni smiselno zapolniti, saj bi bila napaka tam najverjetneje prevelika. Torej se moremo najprej odločiti, do kakšne velikosti vrzeli sploh zapolnjujemo. Ta podatek najlažje podamo v obliki parametra, ki določa površino vrzeli (v slikovnih elementih). Da izračunamo površino vrzeli lahko uporabimo algoritem rast regij (ang. *region growing*) [30].

	UseSSD=1	UseSSD=0
 <p>del scene (63×67 pik)</p>	 <p>$t = 3$ min. in 12 sek. UseBackCorr=0, search_win_size=0</p>	 <p>$t = 5$ min. in 50 sek. UseBackCorr=0, search_win_size=0</p>
	 <p>$t = 6$ min. in 20 sek. UseBackCorr=1, search_win_size=0, DevOfBCPFromSelPoint=0</p>	 <p>$t = 11$ min. in 34 sek. UseBackCorr=1, search_win_size=0, DevOfBCPFromSelPoint=0</p>
	 <p>$t = 6$ min. in 21 sek. UseBackCorr=1, search_win_size=0, DevOfBCPFromSelPoint=3</p>	 <p>$t = 11$ min. in 37 sek. UseBackCorr=1, search_win_size=0, DevOfBCPFromSelPoint=3</p>
	 <p>$t = 15$ min. in 52 sek. UseBackCorr=0, search_win_size=3</p>	 <p>$t = 28$ min. in 56 sek. UseBackCorr=0, search_win_size=3</p>

Slika 5.20: Rezultati gradnje gostih globinskih slik dela scene pri različnih vrednosti parametrov, ki določajo: izbiro korelacijske metode (vsota kvadratov razlik (UseSSD=1) ali normalizirana korelacija (UseSSD=0)), uporabo postopka vzratne korelacije (postopek uporabimo, če UseBackCorr=1), bližnjo okolico epipolarne pre-mice (search_win_size — pove odmik v slikovnih elementih od vrstice panorame, ki ga še preiskujemo) in odstopanje od izbrane točke (DevOfBCPFromSelPoint — pove največje dovoljeno odstopanje v slikovnih elementih). (glej tekst) S spremenljivko t je označen čas rekonstrukcije. ($2\varphi = 29,9625^\circ$, $r = 30$ cm)

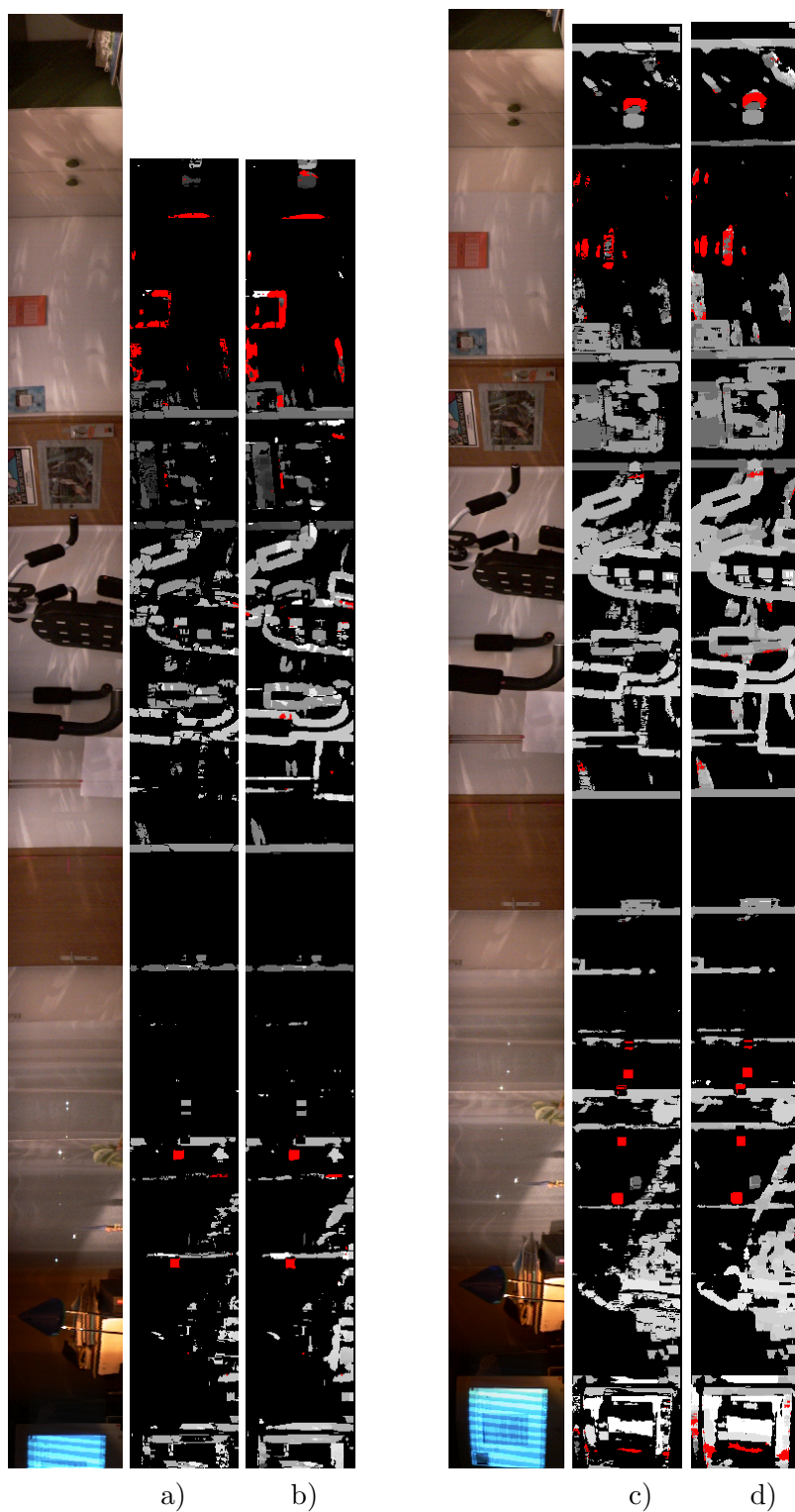
Zaradi uniformnih delov scene pa so vprašljive tudi določene ocene globine na teh delih, saj je naše zaupanje v te ocene majhno. Ta problem lahko rešujemo s pomočjo slike zaupanja v ocenjene globine, ki jo dobimo na podlagi ocen postopka normalizirane korelacije (slika 5.19). Slaba lastnost te rešitve je, da moramo nastaviti prag, ki pove, v katero oceno postopka normalizirane korelacije še zaupamo, hkrati pa kot rezultat pravzaprav dobimo redko globinsko sliko. Redke globinske slike pa lahko delamo veliko učinkoviteje, kot bomo to opisali v naslednjem poglavju. Rezultati v primeru, ko prag zaupanja nastavimo na 230, so prikazani na sliki 5.21. Iz rezultatov je razvidno, da so, neglede na to, ali smo uporabili postopek vzvratne korelacije ali ne, globinske slike za enak kot φ skoraj identične. To pomeni, da bi lahko namesto postopka vzvratne korelacije uporabljali ocene postopka normalizirane korelacije, saj dosežemo zelo podoben učinek. Čas rekonstrukcije pa ostaja na nivoju gradnje globinskih slik, ko ne uporabimo postopka vzvratne korelacije (tabela 5.4). Torej, dobimo podobne rezultate, čas rekonstrukcije pa se približno razpolovi, saj postopek upragavljanja podaljša celoten postopek le za nekaj sekund.

Časovna analiza gradnje redkih globinskih slik

V prejšnjem poglavju smo že omenili en način gradnje redkih globinskih slik, ki pa s časovnega vidika ni bil ugoden. Redko globinsko sliko zgradimo veliko hitreje, če na panoramskih slikah najprej najdemo neke značilke ter nato ocenimo globine le za te značilke. Pri tem uporabljamo enak postopek iskanja korespondenčnih točk kot v primeru gradnje gostih globinskih slik. Sedaj moramo pri iskanju korespondenčne točke, za neko značilko na panorami za levo oko, preiskati le tiste slikovne elemente na epipolarni premici, ki so bili označeni kot značilka tudi na panorami za desno oko. Značilke lahko na primer predstavljajo vertikalni robovi na sceni, ki jih hitro dobimo, če panoramo filtriramo s Sobelovim filtrom za iskanje vertikalnih robov

	čas rekonstrukcije [min./sek.]
uporabimo vzvratno korespondenco $2\varphi = 29,9625^\circ$	1/10
ne uporabimo vzvratne korespondence $2\varphi = 29,9625^\circ$	0/38
uporabimo vzvratno korespondenco $2\varphi = 3,6125^\circ$	0/33
ne uporabimo vzvratne korespondence $2\varphi = 3,6125^\circ$	0/21

Tabela 5.5: Časi gradnje redkih globinskih slik v štirih primerih. Čas, ki je potreben za iskanje značilk na sceni, je vštet v čas rekonstrukcije.



Slika 5.21: Rezultati po obdelavi primerov s slike 5.18, ko zaupamo le v tiste ocene globine, predstavljene na ustreznih primerih na sliki 5.19, ki so večje od praga zaupanja 230. a) Uporabimo postopek vzratne korespondence, kot $2\varphi = 29,9625^\circ$, b) ne uporabimo postopka vzratne korespondence, kot $2\varphi = 29,9625^\circ$, c) uporabimo postopek vzratne korespondence, kot $2\varphi = 3,6125^\circ$, d) ne uporabimo postopka vzratne korespondence, kot $2\varphi = 3,6125^\circ$.

[9, 13]. Slabost redkih globinskih slik je, da je globina ocenjena le za določene izstopajoče točke (značilke) na sceni, vendar navadno v te ocene zelo zaupamo, hkrati pa je postopek rekonstrukcije veliko hitrejši. — Primerjaj rezultate podane v tabeli 5.4 in tabeli 5.5. Na sliki 5.22 so predstavljene ustrezne globinske slike.

Glede na čase rekonstrukcije, ki so podani v tabeli 5.5, lahko sklepamo, da lahko dosežemo delovanje rekonstrukcije v realnem času, če bi delali z 8-bitnimi sivinskimi panoramami manjše ločljivosti in/ali gradili redko globinsko sliko le dela panorame. Takšen rezultat bi načeloma lahko uporabili pri navigaciji robota v prostoru [13]. Prej pa se moramo seveda prepričati v kakovost dobljenih globinskih slik.

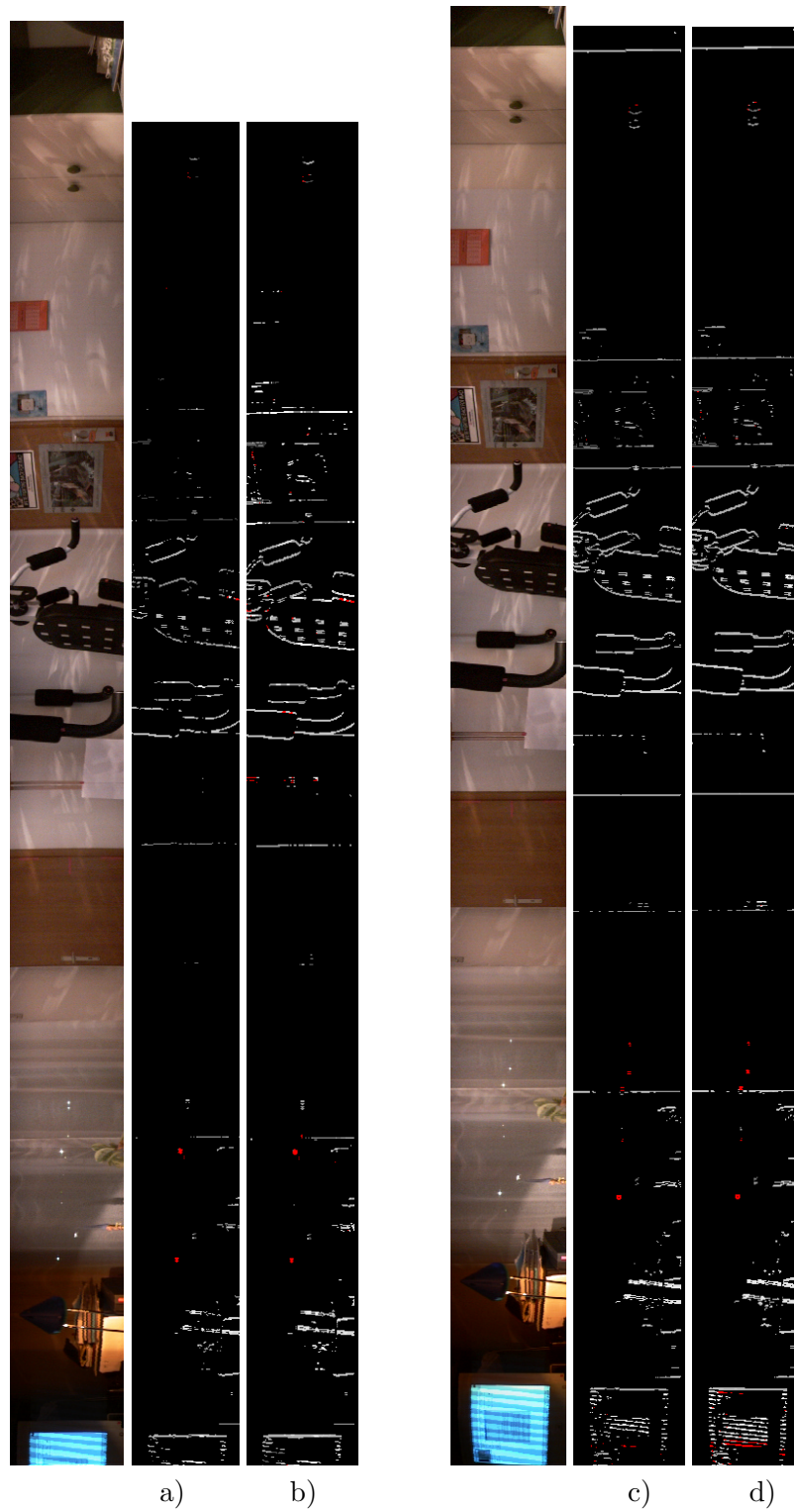
Primerjava rezultatov z dejanskimi globinami

oznaka globine na sliki 5.23	dejanska globina d [cm]	ocenjena globina l [cm] pri $2\varphi =$		razlika $l - d$ [cm (%)] pri $2\varphi =$	
		$3,6125^\circ$	$29,9625^\circ$	$3,6125^\circ$	$29,9625^\circ$
1	111,5	89,4	108,9	-22,1 (-19,8)	-2,6 (-2,3)
2	95,5	76,7	89,3	-18,8 (-19,6)	-6,2 (-6,5)
3	64	53,8	59,6	-10,2 (-15,9)	-4,4 (-6,9)
4	83,5	76,7	78,3	-6,8 (-8,1)	-5,2 (-6,2)
5	92	89,4	89,3	-2,6 (-2,8)	-2,7 (-2,9)
6	86,5	76,7	82,7	-9,8 (-11,3)	-3,8 (-4,4)
7	153	133,4	159,8	-19,6 (-12,8)	6,8 (4,4)
8	130,5	133,4	135,5	2,9 (2,2)	5 (3,8)
9	88	76,7	87,5	-11,3(-12,8)	-0,5 (-0,6)
10	92	89,4	89,3	-2,6 (-2,8)	-2,7 (-2,9)
11	234,5	176,9	213,5	-57,6 (-24,6)	-21 (-9)
12	198	176,9	179,1	-21,1 (-10,7)	-18,9 (-9,5)
13	177	176,9	213,5	-0,1 (-0,1)	36,5 (20,6)

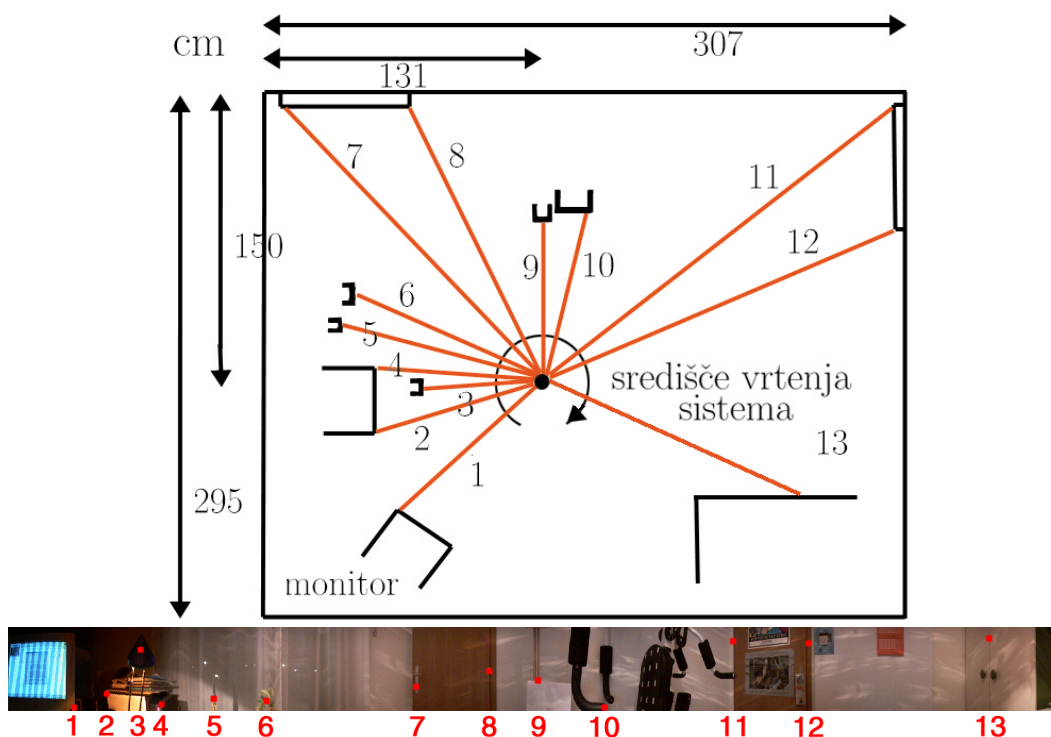
Tabela 5.6: Primerjava med dejanskimi in ocenjenimi globinami (enačba (5.10)) nekaterih značilnk na panoramah za različna kota φ in pri razdalji $r = 30$ cm. Za vse značilke je bila najdena ustrezna korespondenčna značilka.

Na podlagi dosedanjih rezultatov (tabeli 5.4 in 5.5, sliki 5.18 in 5.22) lahko zaključimo, da gostih globinskih slik največkrat ni smiselno graditi, že zaradi samega časa, potrebnega za rekonstrukcijo. V tabeli 5.6 je zato predstavljena primerjava med dejanskimi in ocenjenimi globinami nekaterih značilnk na panoramah. Primerjavo delamo za značilke, ki so različno oddaljene od središča vrtenja sistema (od 64 cm pa do 234,5 cm) (slika 5.23) in za različna kota φ . Iz rezultatov sledi:

- Za večji kot φ so ocene globine korektnješe. Razlog najdemo v diskretizaciji



Slika 5.22: Ustvarjene redke globinske slike: a) uporabimo postopek vzvratne korespondence, kot $2\varphi = 29,9625^\circ$, b) ne uporabimo postopka vzvratne korespondence, kot $2\varphi = 29,9625^\circ$, c) uporabimo postopek vzvratne korespondence, kot $2\varphi = 3,6125^\circ$, d) ne uporabimo postopka vzvratne korespondence, kot $2\varphi = 3,6125^\circ$.



Slika 5.23: Na zgornji sliki je skica prostora, ki smo ga rekonstruirali. Na spodnji sliki so podane oznake od 1 do 13, ki predstavljajo značilke na sceni. Na te oznake se sklicujemo v tabeli 5.6.

osi θ (slika 5.16) na intervale širine $\frac{\theta_0}{2}$ (v našem primeru velja: $\theta_0 = 0,2^\circ$), kar pomeni, da je število možnih ocen globine sorazmerno s kotom φ : za kot $2\varphi = 29,9625^\circ$ velja, da lahko izračunamo največ ($\varphi \text{ div } \frac{\theta_0}{2} =$) 149 ocen globine, za kot $2\varphi = 3,6125^\circ$ pa velja, da lahko izračunamo največ le 18 ocen globine (dodatek D).

- Velikosti ocenjenih globin se za podobne dejanske globine ponavljajo. Razlog je enak kot pri prejšnji točki.
- Če lahko na podlagi predstavljenih rezultatov sklepamo o natančnosti, bi lahko zapisali, da sistem daje smiselne oziroma uporabne ocene globine za značilke do oddaljenosti dveh metrov (2 m). To velja za primer, ko je φ_{maks} enak polovici največjega možnega horizontalnega zornega kota kamere α (stran 70). Seveda je to splošen sklep, ki bi ga morali ustrezno preoblikovati glede na zahteve nekega praktičnega sistema oziroma aplikacije, ki bi uporabljal(a) naš postopek rekonstrukcije.

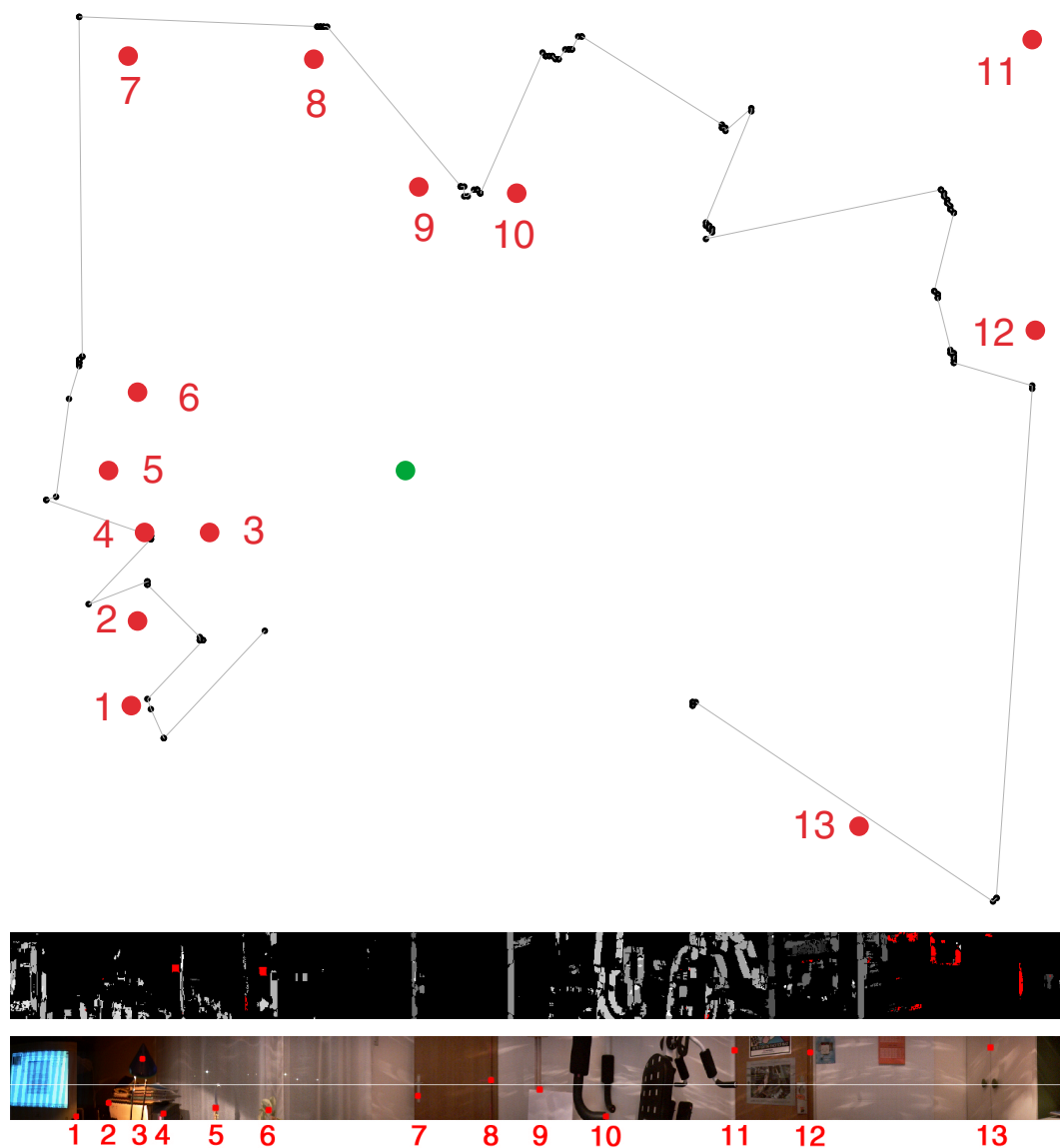
Za konec pa naredimo še nekaj rekonstrukcij prostora na podlagi ustvarjenih globinskih slik. Tako bomo lahko ocenili kakovost dobljenih globinskih slik in posledično kakovost samega sistema. Za slike 5.24, 5.25, 5.26 in 5.27 veljajo naslednje lastnosti²³:

- Z rdečimi točkami so označene značilke na sceni, za katere smo podali dejanske globine v tabeli 5.6. V nadaljevanju bomo upoštevali dejstvo, da te točke ne ležijo na isti vrstici globinske slike (slika 5.23, spodaj).
- Z zeleno točko je označeno središče sistema.
- S črnimi točkami so označene rekonstruirane točke na sceni.
- Linije med črnimi točkami podajajo povezave med zaporedno rekonstruiranimi točkami.
- Rezultat postopka (3D) rekonstrukcije je tloris scene.

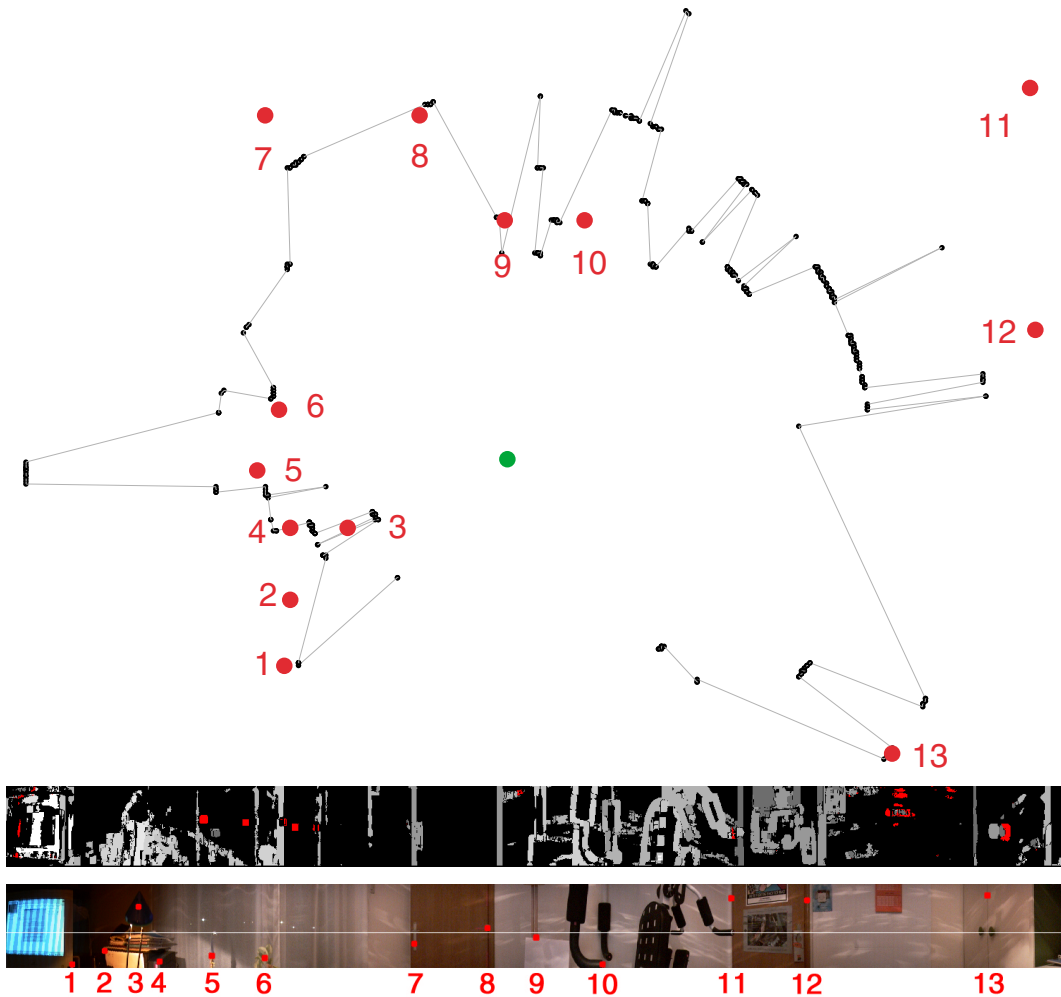
Na sliki 5.24 delamo rekonstrukcijo na podlagi (68.) vrstice globinske slike iz slike 5.21a) ($2\varphi = 29,9625^\circ$), na sliki 5.25 pa delamo rekonstrukcijo na podlagi (68.) vrstice globinske slike iz slike 5.21c) ($2\varphi = 3,6125^\circ$). Za obe sliki (sliki 5.24 in 5.25) torej velja, da so črne točke rekonstruirane na podlagi ocenjenih globin, ki so v globinski sliki zapisane v isti vrstici, medtem ko značilke na sceni, ki smo jih označili z rdečimi točkami, niso nujno vidne v tej vrstici. Iz slike 5.27 so lepo vidne dve lastnosti sistema: rekonstruirane točke se nahajajo na koncentričnih krožnicah s središčem v središču sistema, razdalja med krožnicami pa narašča, ko se oddaljujemo od središča. Slika torej lepo ilustrira dejstvo, da lahko pri majhnem kotu φ ocenimo le malo globin in dejstvo, da se napaka za slikovni element pri oceni kota θ večja, ko se oddaljujemo od središča vrtenja sistema.

V primeru s slike 5.22 smo najprej poiskali vertikalne robove na slikah, zato v nadaljevanju predpostavimo, da so vertikalni robovi (bolj ali manj) na isti globini. Na sliki 5.26 delamo rekonstrukcijo na podlagi cele globinske slike 5.22a) ($2\varphi = 29,9625^\circ$) tako, da najprej povprečimo ocene globin vsake kolone in nato prikažemo povprečno globino na tlorisu prostora, na sliki 5.27 pa delamo rekonstrukcijo na podlagi cele globinske slike 5.22c) ($2\varphi = 3,6125^\circ$). Pri tem smo postavili dodatno omejitev, da mora vsaka kolona vsebovati vsaj štiri točke, ki imajo pripisano oceno globine, sicer točke s povprečno globino ne prikažemo v tlorisu prostora.

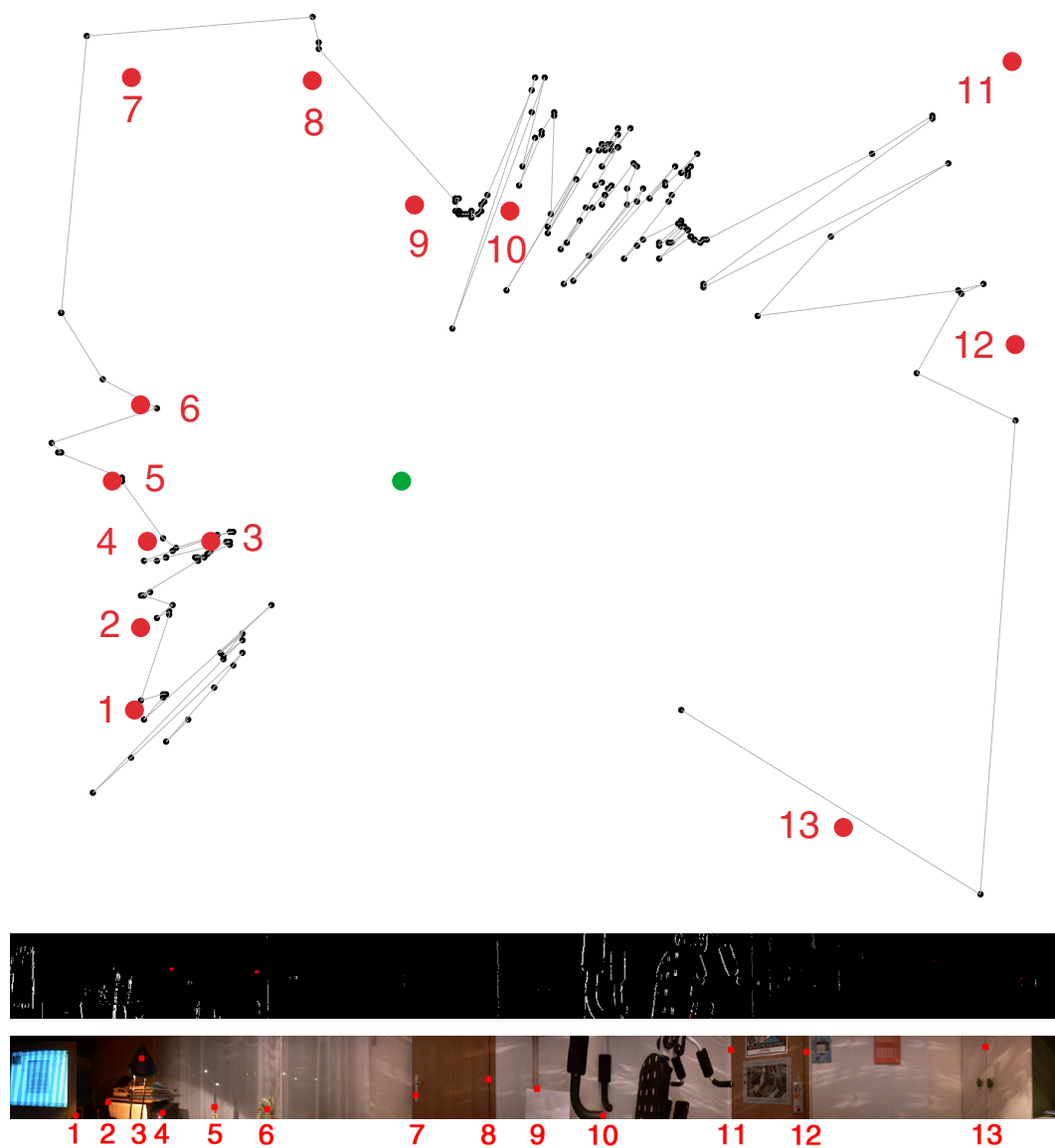
²³Zakaj na teh slikah ni točnega obrisa prostora? Razlogov je več: dejanske razdalje (tabela 5.6) niso povsem točne, saj so bile izmerjene ročno; kot zasuka sistema, glede na kot φ , ko zajamemo prvo kolono panorame, ni povsem točen, saj je bil izmerjen ročno; dejanske razdalje (tabela 5.6) niso podane na isti vrstici panorame.



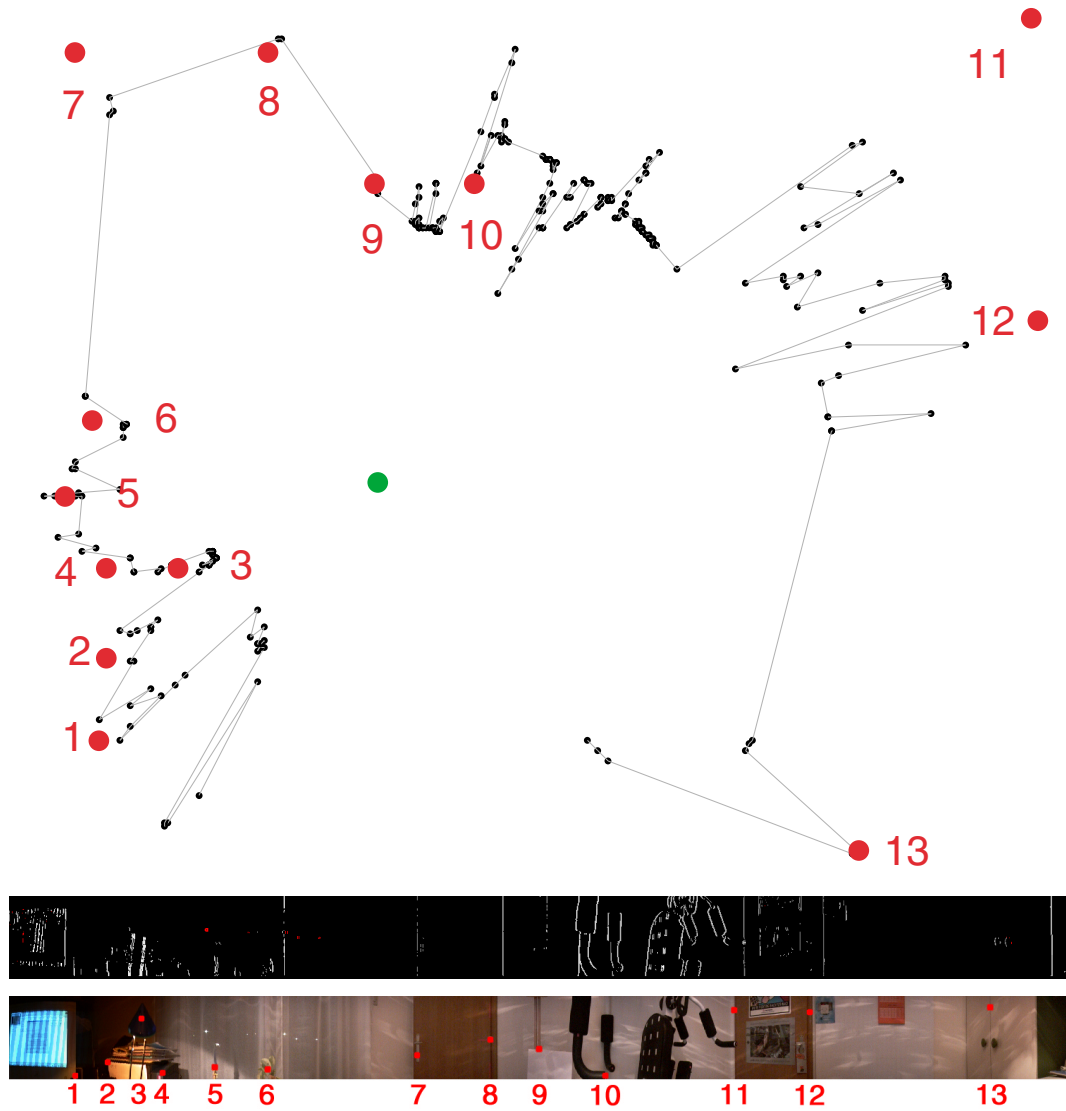
Slika 5.24: Na zgornji sliki je tloris rekonstruirane scene na podlagi (68.) vrstice globinske slike iz slike 5.21a) ($2\varphi = 29,9625^\circ$), ki je predstavljena na srednji sliki. Za orientacijo so na spodnji sliki podane oznake od 1 do 13, ki predstavljajo značilke na sceni in rekonstruirana vrstica. Na te oznake se sklicujemo v tabeli 5.6. Značilke na sceni, ki smo jih označili z rdečimi točkami, niso nujno vidne v tej vrstici. (glej tekst)



Slika 5.25: Na zgornji sliki je tloris rekonstruirane scene na podlagi (68.) vrstice globinske slike iz slike 5.21c) ($2\varphi = 3,6125^\circ$), ki je predstavljena na srednji sliki. Za orientacijo so na spodnji sliki podane oznake od 1 do 13, ki predstavljajo značilke na sceni in rekonstruirana vrstica. Na te oznake se sklicujemo v tabeli 5.6. Značilke na sceni, ki smo jih označili z rdečimi točkami, niso nujno vidne v tej vrstici. (glej tekst)



Slika 5.26: Na zgornji sliki je tloris rekonstruirane scene na podlagi povprečne globine vsake kolone globinske slike iz slike 5.22a) ($2\varphi = 29,9625^\circ$), ki je predstavljena na srednji sliki. Za orientacijo so na spodnji sliki podane oznake od 1 do 13, ki predstavljajo značilke na sceni. Na te oznake se sklicujemo v tabeli 5.6. (glej tekst)



Slika 5.27: Na zgornji sliki je tloris rekonstruirane scene na podlagi povprečne globine vsake kolone globinske slike iz slike 5.22c) ($2\varphi = 3,6125^\circ$), ki je predstavljena na srednji sliki. Za orientacijo so na spodnji sliki podane oznake od 1 do 13, ki predstavljajo značilke na sceni. Na te oznake se sklicujemo v tabeli 5.6. (glej tekst)

Poglavje 6

Zaključek

V nalogi smo najprej predstavili postopek nastanka slike v kameri in postopek stereo vida. S tem smo podali temelje za ustvarjanje panoramskih (globinskih) slik. Implementirali smo sistem za gradnjo globinskih panoramskih slik s postopkom mozaičenja in pri tem uporabili le eno standardno kamero. Naredili smo analizo zasnovanega sistema in rezultatov ter s tem razkrili naravo sistema. Povzemimo osnovne ugotovitve:

- postopek ustvarjanja panoramskih slik je dolgotrajen in ni izvršljiv v realnem času,
- epipolarne premice simetričnih parov panoram so istoležne vrstice panoram,
- na podlagi enačbe za oceno globine l lahko omejimo preiskovalni prostor na epipolarni premici,
- normalizirana korelacija daje pri iskanju korespondenčnih točk boljše rezultate kot postopek vsote kvadratov razlik,
- ker je število možnih ocen globine sorazmerno z velikostjo kota φ , krivulja funkcije globine l pa vse bolj strmo narašča, se zaupanje v ocenjeno globino spreminja: večja je strmina krivulje, manjše je zaupanje v ocenjeno globino,
- gostih globinskih slik največkrat ni smiselno graditi, že zaradi samega časa, potrebnega za rekonstrukcijo,
- sistem daje dobre (dovolj natančne) rezultate pri rekonstrukciji majhnih prostorov in bi jih recimo lahko uporabili pri navigaciji mobilnega robota.

Če izhajamo iz sorodnih del lahko zapišemo, da je bistven prispevek pričujoče naloge k obravnavani tematiki v analizi zasnovanega sistema, pri čemer uporabljamo korelacijsko metodo za iskanje korespondenčnih točk po epipolarni premici. Nihče

izmed avtorjev sorodnih del namreč ni podal tako izčrpne analize sistema (če sploh¹) in ni uporabil korelacijske metode za iskanje korespondenčnih točk po epipolarni premici. Med drugim, nihče izmed avtorjev ni omenil možnosti omejitve preiskovalnega prostora na epipolarni premici.

Za sam konec pa spregovorimo še nekaj besed o nadaljnjem delu.

Nadaljnje delo

V nalogi smo že omenili več postopkov, ki bi lahko služili kot alternativa uporabljenim postopkom, recimo *stereo s preslikavo na valj* (stran 33) in *stereo s sledenjem značilk* (stran 73), vendar pa ti postopki ne bi bistveno spremenili rezultatov sistema². Zato je naše nadaljnje razmišljanje usmerjeno predvsem v razvoj aplikacije za navigacijo robota v prostoru. S tega vidika sta bili omenjeni tudi dve najbolj zanimivi ideji: a) premikanje središča vrtenja našega sistema (stran 73), pri čemer bi lahko gradnjo panoramskih slik pospešili z uporabo transformacije iz krogelnih koordinat v valjčne koordinate (stran 44), ali pa vsaj z gradnjo panoram iz širših trakov zajetih slik in b) uporaba *panoramske polikamere* (stran 36). Zapisali smo tudi osnovno idejo za reševanje problema delovanja rekonstrukcije v realnem času: delo z 8-bitnimi sivinskimi panoramami manjše ločljivosti in/ali gradnja redkih globinskih slik le dela panorame (stran 82). Bolj pereč je seveda problem postopka gradnje panoram, ki ne deluje v realnem času, vendar pa lahko tudi pomembnost tega problema (zelo) zmanjšamo: Namesto, da panorame gradimo iz samo ene kolone zajete slike, bi lahko uporabili širši trak in tako zelo pospešili gradnjo panoram (stran 44). S tem se znotraj traku spreminja kot φ , a epipolarna geometrija ostaja nespremenjena. V primeru, da bi želeli povečati tudi vertikalni zorni kot panorame, bi morali vrteti rotacijsko roko tudi okoli njene horizontalne osi³, s tem pa bi bili primorani uporabiti transformacijo iz krogelnih koordinat v valjčne koordinate. — Idej za nadgradnjo sistema je torej še veliko.

¹Peleg s sodelavci [19, 20, 21, 22, 23] gradi stereo panorame, kjer se efekt globine ustvari v možganih opazovalca. Torej ne dela rekonstrukcije s pomočjo računalnika in tako sploh ne poda željene analize zmogljivosti sistema. Ishiguro s sodelavci [13] navaja le, da je sistem primeren za rekonstrukcijo majhnih prostorov. Huang in Pajdla [12] navajata dejstvo, da se iz simetričnih parov panoram lahko dela stereo rekonstrukcija. Shum in Szeliski [27] sta naredila analizo vertikalne in horizontalne paralakse gibanja, torej podobno kot Huang in Pajdla [12], ki sta izpeljevala epipolarno geometrijo v parih panoram. Le Shum in Szeliski [27] sta pokazala tudi primere ustvarjenih panoramskih globinskih slik.

²Rezultati opisani v [13, 27] niso nič boljši, tako iz časovnega kot iz kakovostnega vidika, od naših, kvečjemu slabši [13].

³Pri tem se morata horizontalna os rotacijske roke in vertikalna os rotacijske roke sekati v optičnem centru kamere. Za našo rotacijsko roko to ne drži.

Dodatek A

3D rotacija

A.1 Splošne lastnosti rotacijske matrike

Rotacijo 3D vektorja \mathbf{v} v 3D vektor \mathbf{v}' lahko predstavimo z linearno transformacijo, ki je definirana z matriko \mathbf{R} , velikosti 3×3 :

$$\mathbf{v}' = \mathbf{R}\mathbf{v} ,$$

za katero velja:

$$\mathbf{R}\mathbf{R}^T = \mathbf{R}^T\mathbf{R} = \mathbf{I} \text{ in } \det(\mathbf{R}) = 1 ,$$

kjer je \mathbf{I} matrika identitete. Prva enačba nam pove, da je inverz matrike \mathbf{R} enak transponirani matriki \mathbf{R} , druga enačba pa nam pove, da transformacija ohranja relativno orientacijo referenčnega okvirja (osi koordinatnega sistema) [9] (strani 20, 21, 29 in 143).

Matrika \mathbf{R} , za katero velja prva enačba oziroma lastnost, se imenuje ortogonalna. Lastnost ortogonalnosti si lažje predstavljamo, če \mathbf{v} in \mathbf{v}' izrazimo v dveh medsebojno ortogonalnih koordinatnih sistemih, ki sta definirana z enotskimi vektorji $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ in $\mathbf{e}'_1, \mathbf{e}'_2, \mathbf{e}'_3$. Iz zamišljenega sledi, da je element r_{ij} matrike \mathbf{R} enak kosinusu kota, ki ga definirata enotski vektorj \mathbf{e}_i in \mathbf{e}'_j :

$$r_{ij} = \mathbf{e}_i^T \mathbf{e}'_j .$$

Na podlagi enačbe $\mathbf{R}\mathbf{R}^T = \mathbf{R}^T\mathbf{R} = \mathbf{I}$ in zgornje enačbe, velja:

$$\sum_{j=1}^3 r_{ij} r_{kj} = \sum_{j=1}^3 r_{ji} r_{jk} = \begin{cases} 0, & i \neq k \\ 1, & i = k \end{cases} ,$$

torej vrstice (in kolone) matrike \mathbf{R} so medsebojno ortogonalni vektorji enotske dolžine [9] (stran 143):

$$\mathbf{r}_i^T \mathbf{r}_j = \begin{cases} 0, & i \neq j \\ 1, & i = j \end{cases}, \quad i, j = 1, 2, 3, \quad i \leq j,$$

kjer smo z \mathbf{r}_i ($i = 1, 2, 3$) označili vrstične (kolonske) vektorje matrike \mathbf{R} .

Devet elementov matrike \mathbf{R} torej ustreza zgornji relaciji, ki tvori šest neodvisnih omejitev. To nam zmanjša število prostostnih stopenj (parametrov) 3D rotacije na $9-6=3$ [36]. Obstaja več parametrizacij matrike \mathbf{R} : rotacija okoli koordinatnih osi, rotacija na osnovi osi in kota, rotacija na osnovi Eulerjevih kotov, rotacija na osnovi kvaternionov itd. V nadaljevanju pa si oglejmo prvo.

A.2 Rotacija okoli koordinatnih osi

3D rotacijo lahko izrazimo kot zaporedje treh rotacij okoli koordinatnih osi \mathbf{e}_1 , \mathbf{e}_2 in \mathbf{e}_3 : okoli koordinatne osi \mathbf{e}_1 (x os) za kot α^1 , okoli koordinatne osi \mathbf{e}_2 (y os) za kot β in okoli koordinatne osi \mathbf{e}_3 (z os) za kot γ . Ti koti tako predstavljajo tri parametre matrike \mathbf{R} , vsaka rotacija pa je predstavljena z rotacijsko matriko \mathbf{R}_j , ki rotira okoli osi \mathbf{e}_j :

$$\begin{aligned} \mathbf{R}_1(\alpha) &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \alpha & -\sin \alpha \\ 0 & \sin \alpha & \cos \alpha \end{bmatrix} \\ \mathbf{R}_2(\beta) &= \begin{bmatrix} \cos \beta & 0 & \sin \beta \\ 0 & 1 & 0 \\ -\sin \beta & 0 & \cos \beta \end{bmatrix} \\ \mathbf{R}_3(\gamma) &= \begin{bmatrix} \cos \gamma & -\sin \gamma & 0 \\ \sin \gamma & \cos \gamma & 0 \\ 0 & 0 & 1 \end{bmatrix} \end{aligned}.$$

Matriko \mathbf{R} , ki opisuje celotno rotacijo, dobimo z zmnožkom matrik \mathbf{R}_j ²:

$$\mathbf{R} = \mathbf{R}_1 \mathbf{R}_2 \mathbf{R}_3 = \begin{bmatrix} \cos \beta \cos \gamma & -\cos \beta \sin \gamma & \sin \beta \\ \sin \alpha \sin \beta \cos \gamma + \cos \alpha \sin \gamma & -\sin \alpha \sin \beta \sin \gamma + \cos \alpha \cos \gamma & -\sin \alpha \cos \beta \\ -\cos \alpha \sin \beta \cos \gamma + \sin \alpha \sin \gamma & \cos \alpha \sin \beta \sin \gamma + \sin \alpha \cos \gamma & \cos \alpha \cos \gamma \end{bmatrix}.$$

Izračun kotov α , β in γ je na podlagi prejšnjega poglavja enostaven, vendar pa se moramo v primeru, da smo matriko \mathbf{R} dobili kot izhod neke numerične metode, vedno prepričati, da je dobljena matrika \mathbf{R} res ortogonalna.

¹V katero smer vrtimo? Če "primemo" koordinatno os za puščico, potem moramo zavrteti koordinatni sistem okoli te osi za dani kot v smeri urinega kazalca.

²Pri tem je seveda vrstni red množenja pomemben, saj nam drugačno zaporedje da drugačen rezultat v odvisnosti od kotov α , β in γ . Torej tukaj obstaja šest različnih načinov sestavljanja matrike \mathbf{R} .

Dodatek B

Računski primer vpliva zunanjih parametrov kamere na 3D točko M

S spremembo položaja in orientacije kamere v prostoru dobimo torej šest zunanjih parametrov kamere, tri za rotacijo \mathbf{R} in tri za translacijo \mathbf{T} (poglavje 2.4). Vzemimo ilustrativen primer, v katerem za tri parametre rotacije (dodatek A) velja: $\alpha=180^\circ$, $\beta = \gamma = 0^\circ$, vektor translacije pa naj bo določen s $\mathbf{T}=[1,1,1]^T$. Zanima nas, kakšne so nove koordinate 3D točke M , ki je v starem koordinatnem sistemu imela koordinate $M_{stara}=(10,10,10)$.

Rešitev:

Po enačbi (2.15) velja:

$$\tilde{\mathbf{M}}_{nova} = \tilde{\mathbf{K}}^{-1} \tilde{\mathbf{M}}_{stara} ,$$

kjer je

$$\tilde{\mathbf{K}} = \begin{bmatrix} \mathbf{R} & \mathbf{T} \\ \mathbf{0}_3^T & 1 \end{bmatrix} ,$$

$$\mathbf{T} = [1, 1, 1]^T \text{ in}$$

$$\mathbf{R} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix} .$$

Ko izračunamo

$$\tilde{\mathbf{K}}^{-1} = \begin{bmatrix} 1 & 0 & 0 & -1 \\ 0 & -1 & 0 & 1 \\ 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix},$$

dobimo $M_{nova}=(9,-9,-9)!$

Dodatek C

Skica računskega primera rekonstrukcije 3D točke

Podane imamo naslednje podatke:

- Delamo z normaliziranimi kamerama (poglavje 2.3).
- Izhodišče 3D prostora je v optičnem centru prve kamere.
- Točka m_2 na normalizirani slikovni ravnini \mathcal{R}_2 (poglavje 3.4) druge kamere:

$$m_2 = \left(\frac{1}{8}, 0\right).$$

- Translacijski vektor med dvema koordinatnima sistemoma kamer:

$$\mathbf{t} = \left[5, 1, \frac{13}{2}\right]^T.$$

- Rotacijska matrika med dvema koordinatnima sistemoma kamer, kjer so parametri $\alpha = \gamma = 0^\circ$, $\beta = 270^\circ$ (dodatek A):

$$\mathbf{R} = \begin{bmatrix} 0 & 0 & -1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}.$$

Sedaj nas zanimajo 3D koordinate točke M , ki se preslika v koordinatnem sistemu druge kamere v točko m_2 .

Rešitev:

Po enačbi (3.6), kjer je \mathbf{T} po enačbi (3.3) enak:

$$\mathbf{T} = \begin{bmatrix} 0 & -\frac{13}{2} & 1 \\ \frac{13}{2} & 0 & -5 \\ -1 & 5 & 0 \end{bmatrix},$$

velja:

$$\begin{aligned} \tilde{\mathbf{m}}_1^T \mathbf{T} \mathbf{R} \tilde{\mathbf{m}}_2 &= \\ \tilde{\mathbf{m}}_1^T \begin{bmatrix} 0 & -\frac{13}{2} & 1 \\ \frac{13}{2} & 0 & -5 \\ -1 & 5 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 & -1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} \frac{1}{8} \\ 0 \\ 1 \end{bmatrix} &= \\ \tilde{\mathbf{m}}_1^T \begin{bmatrix} \frac{1}{8} \\ -\frac{57}{8} \\ 1 \end{bmatrix} &= 0. \end{aligned}$$

Točko $m_1 = (u, v)$ na slikovni ravnini prve kamere zapišemo v projektivnih koordinatah kot $(u, v, 1)$, pri čemer za u in v po zgornji enačbi velja:

$$v = \frac{1}{57}u + \frac{8}{57}.$$

Možni kandidati za točko m_1 na slikovni ravnini prve kamere tako izpolnjujejo zgornji pogoj (epipolarna omejitev). S pomočjo korelacije dobimo korespondenčno točko (poglavje 3.5.1) m_1 , ker pa korelacijska metoda presega namen te seminarske naloge, zapišimo $m_1 = (\frac{1}{7}, \frac{1}{7})^1$.

Ker se nahajamo v normaliziranem koordinatnem sistemu, vemo, da je:

$$\tilde{\mathbf{P}}_1 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}.$$

Vemo tudi, da je matrika $\tilde{\mathbf{P}}_2$ v novem koordinatnem sistemu C_2XYZ enaka matriki $\tilde{\mathbf{P}}_1$. Matrika $\tilde{\mathbf{P}}_2$ izražena v koordinatnem sistemu C_1xyz pa je tako enaka:

$$\tilde{\mathbf{P}}_1 = \tilde{\mathbf{P}}_2 \tilde{\mathbf{K}} \Rightarrow \tilde{\mathbf{P}}_2 = \tilde{\mathbf{P}}_1 \tilde{\mathbf{K}}^{-1} = \begin{bmatrix} \mathbf{R}^T & -\mathbf{R}^T \mathbf{t} \end{bmatrix},$$

kjer je $\tilde{\mathbf{K}}$ matrika kolineacije (stran 13):

$$\tilde{\mathbf{K}} = \begin{bmatrix} \mathbf{R} & \mathbf{t} \\ \mathbf{0}_3^T & 1 \end{bmatrix}.$$

¹Ta točka je dobljena iz realne situacije: preslikava 3D točke $M = (1, 1, 7)$.

$\tilde{\mathbf{P}}_2$ je torej enaka:

$$\tilde{\mathbf{P}}_2 = \begin{bmatrix} 0 & 0 & 1 & -\frac{13}{2} \\ 0 & 1 & 0 & -1 \\ -1 & 0 & 0 & 5 \end{bmatrix}.$$

Sedaj poznamo slikovne koordinate točk m_1 in m_2 ter vrednosti elementov matrik $\tilde{\mathbf{P}}_1$ in $\tilde{\mathbf{P}}_2$, zato lahko izračunamo 3D koordinate točke M (poglavje 3.6) tako, da te podatke vstavimo v enačbi:

$$\begin{aligned} (p_{11} - up_{31})x + (p_{12} - up_{32})y + (p_{13} - up_{33})z + p_{14} - up_{34} &= 0 \\ (p_{21} - vp_{31})x + (p_{22} - vp_{32})y + (p_{23} - vp_{33})z + p_{24} - vp_{34} &= 0. \end{aligned}$$

Rešiti moramo torej sistem štirih linearnih enačb s tremi neznankami:

$$\begin{aligned} -\frac{57}{8} + \frac{1}{8}x + z &= 0 \\ -1 + y &= 0 \\ x - \frac{1}{7}z &= 0 \\ y - \frac{1}{7}z &= 0, \end{aligned}$$

ki nam dajo rešitev:

$$\underline{\underline{M = (1, 1, 7)}}.$$

Dodatek D

Možne ocene globin glede na kot φ

θ [°]	l [cm]	θ [°]	l [cm]
0,1	31,8	0,2	33,7
0,3	36	0,4	38,5
0,5	41,5	0,6	44,9
0,7	49	0,8	53,8
0,9	59,8	1	67,2
1,1	76,7	1,2	89,4
1,3	107	1,4	133,4
1,5	176,9	1,6	262,7
1,7	509,9	1,8	8668,6

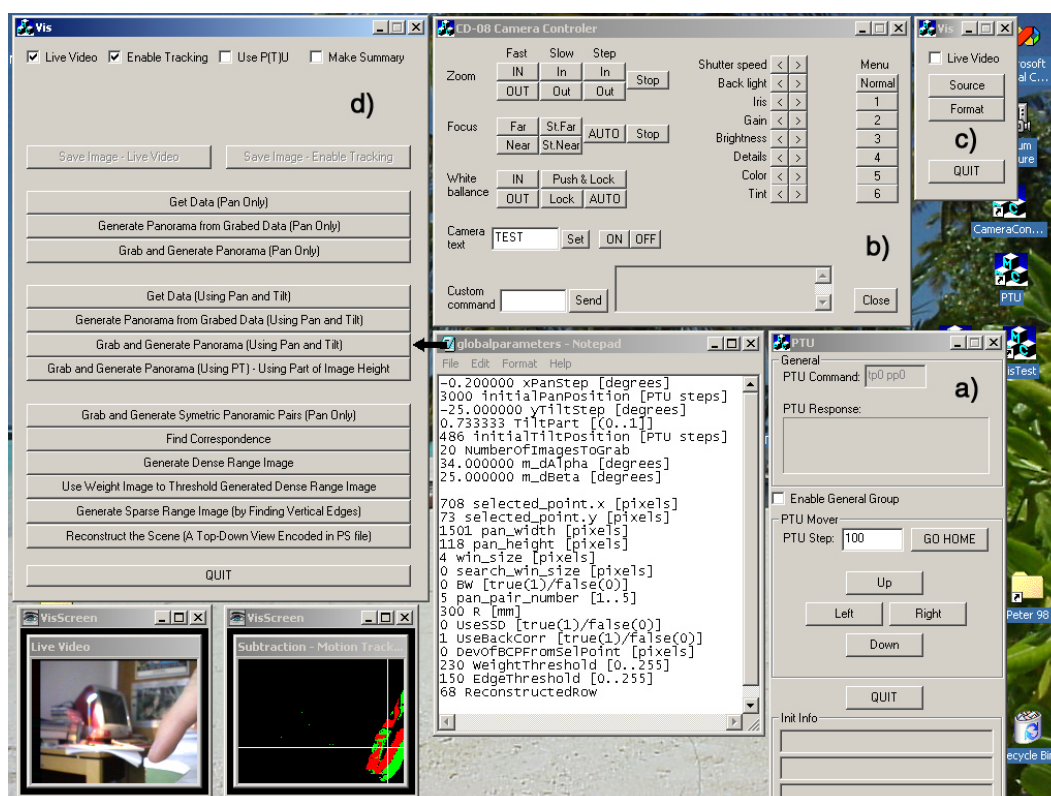
Tabela D.1: Možne ocene globin l (enačba (5.10)) za kot $2\varphi = 3,6125^\circ$, ob razdalji $r = 30$ cm. Pri $\theta_0 = 0,2$ velja, da lahko izračunamo največ $\varphi \operatorname{div} \frac{\theta_0}{2} = 18$ ocen globin.

θ [°]	l [cm]	θ [°]	l [cm]	θ [°]	l [cm]	θ [°]	l [cm]	θ [°]	l [cm]
0,1	30,2	0,2	30,4	0,3	30,6	0,4	30,8	0,5	31
0,6	31,2	0,7	31,4	0,8	31,7	0,9	31,9	1	32,1
1,1	32,3	1,2	32,6	1,3	32,8	1,4	33	1,5	33,3
1,6	33,5	1,7	33,8	1,8	34	1,9	34,3	2	34,5
2,1	34,8	2,2	35,1	2,3	35,3	2,4	35,6	2,5	35,9
2,6	36,2	2,7	36,5	2,8	36,8	2,9	37,1	3	37,4
3,1	37,7	3,2	38	3,3	38,3	3,4	38,6	3,5	39
3,6	39,3	3,7	39,6	3,8	40	3,9	40,3	4	40,7
4,1	41,1	4,2	41,5	4,3	41,8	4,4	42,2	4,5	42,6
4,6	43	4,7	43,5	4,8	43,9	4,9	44,3	5	44,7
5,1	45,2	5,2	45,6	5,3	46,1	5,4	46,6	5,5	47,1
5,6	47,6	5,7	48,1	5,8	48,6	5,9	49,1	6	49,7
6,1	50,2	6,2	50,8	6,3	51,4	6,4	52	6,5	52,6
6,6	53,2	6,7	53,8	6,8	54,5	6,9	55,2	7	55,9
7,1	56,6	7,2	57,3	7,3	58	7,4	58,8	7,5	59,6
7,6	60,4	7,7	61,2	7,8	62	7,9	62,9	8	63,8
8,1	64,7	8,2	65,7	8,3	66,7	8,4	67,7	8,5	68,7
8,6	69,8	8,7	70,9	8,8	72	8,9	73,2	9	74,4
9,1	75,7	9,2	77	9,3	78,3	9,4	79,7	9,5	81,2
9,6	82,7	9,7	84,3	9,8	85,9	9,9	87,6	10	89,3
10,1	91,1	10,2	93	10,3	95	10,4	97,1	10,5	99,3
10,6	101,5	10,7	103,9	10,8	106,4	10,9	109	11	111,7
11,1	114,6	11,2	117,6	11,3	120,8	11,4	124,2	11,5	127,7
11,6	131,5	11,7	135,5	11,8	139,7	11,9	144,3	12	149,1
12,1	154,3	12,2	159,8	12,3	165,8	12,4	172,2	12,5	179,1
12,6	186,7	12,7	194,8	12,8	203,8	12,9	213,5	13	224,3
13,1	236,2	13,2	249,5	13,3	264,3	13,4	281	13,5	300,
13,6	321,7	13,7	346,8	13,8	376,2	13,9	411	14	452,8
14,1	504,2	14,2	568,8	14,3	652,2	14,4	764,5	14,5	923,3
14,6	1165,5	14,7	1579,9	14,8	2451,5	14,9	5468,7		

Tabela D.2: Možne ocene globin l (enačba (5.10)) za kot $2\varphi = 29,9625^\circ$, ob razdalji $r = 30$ cm. Pri $\theta_0 = 0,2$ velja, da lahko izračunamo največ $\varphi \div \frac{\theta_0}{2} = 149$ ocen globin.

Dodatek E

O implementaciji



Slika E.1: Uporabniški vmesniki programskih delov sistema: a) modul za krmiljenje rotacijske roke, b) modul za komunikacijo s kamero, c) modul za zajemanje žive slike in d) modul za ustvarjanje (globinskih) panoramskih slik.

Zasnovan sistem je implementiran v programskem jeziku C++, v okolju Mi-

Microsoft Visual C++. Kot tak izkorišča prednosti objektnega načina programiranja. Za lažjo berljivost kode so uporabljena dolga imena spremenljivk v angleščini, zamiki in komentarji v angleščini. V okolje je integrirana funkcionalnost paketa Microsoft Vision Software Development Kit (Ms VisSDK), ki v prvi vrsti omogoča povezavo aplikacije z zajemalnikom slik. Programski deli sistema so:

- modul za krmiljenje rotacijske roke (slika E.1a)),
- modul za komunikacijo s kamero (slika E.1b)),
- modul za zajemanje žive slike (slika E.1c)) in
- modul za ustvarjanje (globinskih) panoramskih slik (slika E.1d)).

Zadnji modul deluje na osnovi parametrov, ki so na sliki E.1 predstavljeni v oknu z imenom *globalparameters*. Podajmo kratek opis parametrov, kjer na začetku podamo vrednost parametra v primeru opisanih rezultatov¹, ime parametra in mersko enoto ali definicijsko območje parametra:

`-0.200000 xPanStep [°]` — Parameter podaja premik kamere po horizontali med dvema zaporednima slikama, ki prispevata k panoramam. Kot je podan v stopinjah. Negativni predznak pove, da se sistem vrti v negativni matematični smeri.

`initialPanPosition [korak rotacijske roke]` — Parameter podaja začetno horizontalno lokacijo rotacijske roke. Enota je osnovni korak rotacijske roke.

`yTiltStep [°]` — Parameter podaja premik kamere po vertikali, če gradimo panorame iz več obhodov okoli vertikalne osi rotacijske roke. Kot je podan v stopinjah. Negativni predznak pove, da se sistem giba od zgoraj navzdol.

`TiltPart [(0..1)]` — Parameter podaja velikost zajete slike po vertikali, ki prispeva k panorami, če gradimo panorame iz več obhodov okoli vertikalne osi rotacijske roke. Če je vrednost parametra enaka 1, se ohranja celotna višina zajete slike.

`initialTiltPosition [korak rotacijske roke]` — Parameter podaja začetno horizontalno lokacijo rotacijske roke. Enota je osnovni korak rotacijske roke.

`NumberOfImagesToGrab` — Parameter podaja število zaporedno zajetih slik, iz katerih zgradimo panoramo. Parameter uporabimo v primeru, da ne želimo zgraditi panorame največje možne velikosti.

`34.000000 m_dAlpha [°]` — Parameter podaja horizontalni zorni kot kamere v stopinjah.

¹Vrednost je podana le v primeru, če je pomembna za postopek rekonstrukcije.

- `m_dBeta` [°] — Parameter podaja vertikalni zorni kot kamere v stopinjah.
- `selected_point.x`, `selected_point.y` [slikovni element] — Parametra podajata horizontalno in vertikalno koordinato slikovnega elementa na panorami za levo oko, za katerega iščemo korespondenčno točko na panorami za desno oko.
- 1501 `pan_width` [slikovni element] — Parameter podaja širino panorame v slikovnih elementih.
- 118 `pan_height` [slikovni element] — Parameter podaja višino panorame v slikovnih elementih.
- 4 `win_size` [slikovni element] — Parameter podaja velikost korelacijskega okna v slikovnih elementih: $(2n + 1) \times (2n + 1)$.
- 0 `search_win_size` [slikovni element] — Parameter podaja velikost iskalnega okna okoli epipolarne premice: korespondenčne točke iščemo tudi za do vključno n slikovnih elementov nad in pod epipolarno premico.
- 0 `BW` [0/1] — Parameter podaja ukaz sistemu, kako naj obravnava vhodne panoramske slike v primeru iskanja korespondenčnih točk. Slike so zajete v barvnem prostoru RGB in če želimo, da sistem uporabi vso razpoložljivo barvno informacijo, nastavimo ta parameter na 0. Sicer sistem naredi najprej transformacijo v 8-bitni (sivinski) barvni prostor.
- 5 `pan_pair_number` [1..5] — S tem parametrom izberemo simetričen par panoram, na osnovi katerega bomo gradili globinske slike:
 1 — $2\varphi=3,6125^\circ$, 2 — $2\varphi=11,2625^\circ$, 3 — $2\varphi=18,9125^\circ$,
 4 — $2\varphi=26,5625^\circ$, 5 — $2\varphi=29,9625^\circ$.
- 300 `R` [mm] — Parameter pove razdaljo optičnega centra kamere od središča vrtenja v milimetrih.
- 0 `UseSSD` [0/1] — S tem parametrom povemo sistemu, katero korelacijsko metodo bomo uporabili pri iskanju korespondenčnih točk: 0 — normalizirana korelacija, 1 — vsota kvadratov razlik.
- 1 `UseBackCorr` [0/1] — S tem parametrom povemo sistemu, ali bomo uporabili postopek vzvratne korelacije: 0 — postopka ne uporabimo, 1 — postopek uporabimo.
- 0 `DevOfBCPFFromSelPoint` [slikovni element] — Mogoče bi zahtevo, ki jo postavljamo s postopkom vzvratne korelacije, lahko malo omilili: namesto da moramo z vzvratno korelacijo najti izbrano točko, za katero smo iskali korespondenčno točko, dopustimo neko majhno odstopanje od te izbrane točke. Parameter podaja največje dovoljeno odstopanje od izbrane točke v slikovnih elementih.

- 230 `WeightThreshold` [0..255] — Zaradi uniformnih delov scene so vprašljive določene ocene globine na teh delih, saj je naše zaupanje v te ocene majhno. Ta problem lahko rešujemo s pomočjo slike zaupanja v ocenjene globine, ki jo dobimo na podlagi ocen postopka normalizirane korelacije. Parameter podaja prag, ki pove, v katero oceno postopka normalizirane korelacije še zaupamo. Ocena postopka normalizirane korelacije je vedno na intervalu [-1,1] (realna števila), ki jo preslikamo na interval [0,255] (cela števila).
- 150 `EdgeThreshold` [0..255] — Redko globinsko sliko zgradimo veliko hitreje, če na panoramskih slikah najprej najdemo neke značilke ter nato ocenimo globine le za te značilke. Značilke lahko na primer predstavljajo vertikalni robovi na sceni, ki jih hitro dobimo, če panoramo filtriramo s Sobelovim filtrom za iskanje vertikalnih robov. Parameter podaja zahtevano moč oziroma izrazitost značilke.
- 68 `ReconstructedRow` [[win_size..pan_height-win_size)] — S tem parametrom povemo sistemu, katero vrstico globinske slike naj rekonstruira; rezultat postopka (3D) rekonstrukcije je tloris scene, ki je zapisan v datoteko tipa PostScript. Na sliki so označene tudi točke iz tabele 5.6 in središče vrtenja našega sistema. Za te točke moramo vedeti, da načeloma ne ležijo v vrstici podani s tem parametrom.

Iz slike E.1d) je razvidno tudi, da modul za ustvarjanje (globinskih) panoramskih slik vsebuje tudi funkcijo sledenja gibanju na sceni, kjer lahko delamo vizualne in numerične povzetke dogajanja. Ta del sistema predstavlja prototip (komercialnega) nadzornega in varnostnega sistema SecurityAgent[©], ki v sebi združuje funkcionalnost našega sistema in katadioptričnega sistema.

Dodatek F

Terminološki slovarček

Dioptrična kamera je kamera, ki vsebuje leče.

Dispariteta je kotna razlika med dvema pogledoma na isto točko na sceni, ki je vidna na dveh slikah (slika 4.9). Človeški možgani interpretirajo dispariteto kot globino.

Epipolarna geometrija kamer je geometrija stereo sistema, ki omogoča enostavnejše iskanje korespondenčnih (sovpadajočih) točk.

Glediščna krožnica podaja pot kamere, katere optična os je vedno tangenta na to krožnico. Glediščna krožnica opisuje ne-središčni sistem, ki ga dobimo v primeru ustvarjanja večperspektivnih panoram.

Globinska slika podaja oddaljenost posamezne točke objekta od točke opazovanja.

Horizontalni in vertikalni zorni kot kamere določata vidno polje kamere.

Katadioptrična kamera je kamera, ki vsebuje vsaj eno ogledalo, lahko pa vsebuje tudi leče.

Klasična panoramska kamera; glej *središčna panoramska kamera*.

Koncentrična panorama je panorama, ki temelji na uporabi več standardnih kamer, ki so pritrjene na rotacijsko mizo (roko) (slika 4.6a); vrtenje standardnih kamer po krožnicah. Večperspektivna panorama nastane iz središčnih trakov (kolon) zajetih slik na vsaki lokaciji rotacijske mizice, kot je to razvidno iz slike 4.6b.

Linearna potisna kamera (ang. *linear pushbroom camera*) je 1D tabela CCD senzorjev (linijska kamera), ki jo premikamo po ravni liniji s konstantno hitrostjo.

Linijska kamera (ang. *line-scan camera*); glej *trakovna kamera*.

Ne-klasična panoramska kamera; glej *ne-središčna panoramska kamera*.

Ne-središčna panoramska kamera je panoramska kamera, za katero velja, da se svetlobni žarki, ki tvorijo sliko, ne sekajo v eni točki, pač pa se lahko sekajo na premici, lahko predstavljajo tangente na neko krožnico ipd. Sinonim: ne-klasična panoramska kamera.

Optična os kamere je definirana z optičnim centrom kamere in dejanskim izhodiščem slikovnega koordinatnega sistema. V idealnem primeru velja, da je optična os kamere pravokotna na sliko kamere in da jo optična os seka na sredini.

Optični center kamere je točka, v kateri se sekajo vsi svetlobni žarki, ki tvorijo sliko v točkasti kameri.

Panorama je slika, dobljena s panoramsko kamero.

Panoramska kamera je kamera za katero velja, da je njeno vidno polje vsebuje vsaj eno krožnico na površini krogle (ang. *view-sphere*); za dodatno razlago glej *vesemerna kamera*. Če ta pogoj malce omilimo, lahko rečemo, da za panoramsko kamero velja, da (zelo) poveča vidno polje standardne kamere.

Paralaksa gibanja je efekt, ki povzroči spremembo lokacije opazovanega objekta, ker se spremeni lokacija človeka oziroma instrumenta, ki opazuje objekt (slika 4.4).

Perspektivna projekcija; glej *središčna projekcija*.

Rotacijska panorama (ang. *swing panorama*) je panorama, ki temelji na vrtenju standardne kamere (po krožnici), ki gleda navzven, torej stran od središča vrtenja (slika 4.7a). Večperspektivne panorame nastanejo iz različnih trakov (kolon) zajetih slik, kot je to razvidno iz slike 4.7b in slike 4.8.

Simetričen par (rotacijskih) panoram dobimo, če vzamemo simetrične trakove levo in desno od sredine zajete slike (torej na enaki razdalji od sredine zajete slike), ali če pritrdimo dve kameri na enaki razdalji od središča vrtenja, vendar na nasprotnih koncih rotacijske roke (slika 4.10). Ti dve panorami imata zaradi horizontalne epipolarne geometrije lepo vidno horizontalno paralakso gibanja, vertikalne paralakse gibanja pa ni. Glej *rotacijska panorama*.

Središčna panoramska kamera je panoramska kamera, za katero velja, da se vsi svetlobni žarki, ki tvorijo sliko, sekajo v eni točki. Sinonimi: klasična panoramska kamera, panoramska kamera s središčem projekcije, panoramska kamera z enim samim glediščem (ang. *single (effective) viewpoint*).

Središčna projekcija je projekcija 3D scene na 2D slikovno ravnino skozi optični center kamere. Sinonim: perspektivna projekcija.

Standardna kamera je središčna, usmerjena in dioptrična kamera, kar lahko z drugimi besedami opišemo kot točkasta kamera, katere vidno polje je del polkrogle.

Stereo sistem omogoča 3D rekonstrukcijo scen iz dveh ali več gledišč.

Točkasta kamera je standardna kamera, za katero velja, da se vsi svetlobni žarki, ki tvorijo sliko, sekajo v eni točki, ki se imenuje optični center kamere.

Trakovna kamera (ang. *slit camera*) je 1D tabela CCD senzorjev. Sinonim: linijska kamera. Včasih se pojem trakovna kamera uporablja tudi v primeru, ko iz slike, zajete s standardno kamero, uporabimo le določen trak slikovnih elementov (tipično nekaj kolon slikovnih elementov), vendar v tem primeru pojem linijska kamera ne predstavlja sinonima tako definirani trakovni kameri.

Usmerjena kamera je kamera za katero velja, da je njeno vidno polje ustrezna podmnožica polkrogle na površini krogle (ang. *view-sphere*); za dodatno razlago glej *vsesmerna kamera*.

Večperspektivna panorama (ang. *multiperspective panorama*) je sestavljena tako, da iz vsake zajete slike vzamemo le eno ali nekaj kolon slikovnih elementov (kar je ekvivalentno uporabi linijske oziroma trakovne kamere), ki doprinesejo k panorami. Pomembna lastnost (dveh ali več) večperspektivnih panoram je, da hranijo informacijo o paralaksi gibanja, saj so kolone le-teh zajete iz različnih perspektiv (slike 4.6, 4.7 in 4.8).

Vsesmerna kamera je kamera, ki ima celotno vidno polje, kar pomeni, da slika zajeta s takšno kamero pokriva celotno površino krogle (ang. *view-sphere*).

Literatura

- [1] M. Artač, *Projekcijska rekonstrukcija prizora iz zaporedja slik*, diplomska naloga, Fakulteta za računalništvo in informatiko, Univerza v Ljubljani, Ljubljana, Slovenija, 2000.
- [2] H. Bakstein, T. Pajdla, Non-central cameras: A review, Research report CTU-CMP-2000-14, Center for Machine Perception, Czech Technical University, Praga, Češka, 2000.
- [3] H. Bakstein, T. Pajdla, An overview of non-central cameras, v B. Likar, urednik, zbornik *Computer Vision Winter Workshop*, str. 223-233, Bled, Slovenija, 2001.
- [4] A. Basu, S. Licardie, Alternative models for fish-eye lenses, *Pattern Recognition Letters*, letnik 16, št. 4, str. 433-441, 1995.
- [5] R. Benosman, T. Maniere, J. Devars, Panoramic stereovision sensor, v A. K. Jain, S. Venkatesh, B. C. Lovell, uredniki, *IEEE 14th International Conference on Pattern Recognition*, str. 767-769, Brisbane, Avstralija, 1998.
- [6] S. Birchfield, An Introduction to Projective Geometry (for computer vision), <http://robotics.stanford.edu/~birch/projective/>, 1998.
- [7] J. N. Bronštejn, K. A. Semendjajev, *Matematični priročnik: za inženirje in slušatelje tehniških visokih šol*, Tehniška založba Slovenije, v Ljubljani, 1988.
- [8] S. Chen, Quicktime VR — an image-based approach to virtual environment navigation, zbornik *ACM SIGGRAPH*, str. 29-38, Los Angeles, ZDA, 1995.
- [9] O. Faugeras, *Three-Dimensional Computer Vision: A Geometric Viewpoint*, MIT Press, Cambridge, Massachusetts, London, England, 1993.
- [10] R. Gupta, R. I. Hartley, Linear pushbroom cameras, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, letnik 19, št. 9, str. 963-975, 1997.
- [11] R. Hartley, A. Zisserman, *Multiple View Geometry in Computer Vision*, Cambridge University Press, Cambridge, Velika Britanija, 2000.

- [12] F. Huang, T. Pajdla, Epipolar Geometry in Concentric Panoramas, Research report CTU-CMP-2000-07, Center for Machine Perception, Czech Technical University, Praga, Češka, 2000.
- [13] H. Ishiguro, M. Yamamoto, S. Tsuji, Omni-directional stereo, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, letnik 14, št. 2, str. 257-262, 1992.
- [14] M. Jogan, T. Pajdla, Image-based map of the lab 130, Research report CTU-CMP-1999-23, Center for Machine Perception, Czech Technical University, Praga, Češka, 1999.
- [15] E. Kramar, *Linearna algebra — zapiski iz predavanj*, Fakulteta za računalništvo in informatiko, Univerza v Ljubljani, Ljubljana, Slovenija, 1995.
- [16] P. Peer, *Primer uporabe projekтивne geometrije v računalniškem vidu: samodajna izdelava 3D modelov*, tehnično poročilo, Fakulteta za računalništvo in informatiko, Univerza v Ljubljani, Ljubljana, Slovenija, 2000.
- [17] P. Peer, *Panoramski pogledi in 3D rekonstrukcija*, tehnično poročilo, Fakulteta za računalništvo in informatiko, Univerza v Ljubljani, Ljubljana, Slovenija, 2001.
- [18] P. Peer, *Sistem za gradnjo globinskih panoramskih slik*, tehnično poročilo, Fakulteta za računalništvo in informatiko, Univerza v Ljubljani, Ljubljana, Slovenija, 2001.
- [19] S. Peleg, J. Herman, Panoramic mosaics by manifold projection, zbornik *IEEE Conference on Computer Vision and Pattern Recognition*, str. 338-343, San Juan, Puerto Rico, 1997.
- [20] S. Peleg, M. Ben-Ezra, Stereo panorama with a single camera, zbornik *IEEE Conference on Computer Vision and Pattern Recognition*, str. 395-401, Fort Collins, ZDA, 1999.
- [21] S. Peleg, Y. Pritch, M. Ben-Ezra, Cameras for Stereo Panoramic Imaging, zbornik *IEEE Conference on Computer Vision and Pattern Recognition*, str. 208-214, Hilton Head Island, ZDA, 2000.
- [22] S. Peleg, B. Rousso, A. Rav-Acha, A. Zomet, Mosaicing on Adaptive Manifolds, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, letnik 22, št. 10, str. 1144-1154, 2000.
- [23] S. Peleg, M. Ben-Ezra, Y. Pritch, Omnistereo: Panoramic Stereo Imaging, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, letnik 23, št. 3, str. 279-290, 2001.

- [24] B. Prihavec, *Sistem za prenašanje žive slike in aktivno opazovanje preko interneta*, magistrska naloga, Fakulteta za računalništvo in informatiko, Univerza v Ljubljani, Ljubljana, Slovenija, 1999.
- [25] B. Prihavec, F. Solina, User interface for video observation over the internet, *Journal of Network and Computer Applications*, letnik 21, str. 219–237, 1998.
- [26] P. Rademacher, G. Bishop, Multiple-center-of-projection images, zbornik *Computer Graphics (ACM SIGGRAPH)*, str. 199–206, Orlando, ZDA, 1998.
- [27] H.-Y. Shum, R. Szeliski, Stereo reconstruction from multiperspective panoramas, zbornik *IEEE 7th International Conference on Computer Vision*, str. 14–21, Kerkyra, Grčija, 1999.
- [28] H.-Y. Shum, A. Kalai, S. M. Seitz, Omnivergent stereo, zbornik *IEEE 7th International Conference on Computer Vision*, str. 22–29, Kerkyra, Grčija, 1999.
- [29] D. Skočaj, *Avtomatsko modeliranje 3-dimenzionalnih večbarvnih predmetov z uporabo globinskega senzorja*, magistrska naloga, Fakulteta za računalništvo in informatiko, Univerza v Ljubljani, Ljubljana, Slovenija, 1999.
- [30] M. Sonka, V. Hlavac, R. Boyle, *Image Processing, Analysis and Machine Vision*, Chapman & Hall Computing, London, Glasgow, New York, Tokyo, Melbourne, Madras, 1993.
- [31] T. Svoboda, T. Pajdla, Panoramic cameras for 3D computation, v T. Svoboda, urednik, zbornik *Czech Pattern Recognition Workshop*, str. 63–70, Praga, Češka, 2000.
- [32] T. Svoboda, *Central Panoramic Camera Design, Geometry, Egomotion*, doktorska disertacija, Center for Machine Perception, Czech Technical University, Praga, Češka, 2000.
- [33] R. Swaminathan, S. K. Nayar, Nonmetric calibration of wide-angle lenses and polycameras, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, letnik 22, št. 10, str. 1172–1178, 2000.
- [34] R. Szeliski, Video Mosaics for Virtual Environments, *IEEE Computer Graphics and Applications*, letnik 16, št. 2, str. 22–30, 1996.
- [35] R. Szeliski, H.-Y. Shum, Creating full view panoramic image mosaics and texture-mapped models, zbornik *Computer Graphics (ACM SIGGRAPH)*, str. 251–258, Los Angeles, ZDA, 1997.
- [36] E. Trucco, A. Verri, *Introductory Techniques for 3-D Computer Vision*, Prentice-Hall, 1998.

- [37] J. Weng, P. Cohen, M. Herniou, Camera calibration with distortion models and accuracy evaluation, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, letnik 14, št. 10, str. 965-980, 1992.
- [38] D. Wood, A. Finkelstein, J. Hughes, C. Thayer, D. Salesin, Multiperspective panoramas for cel animation, zbornik *Computer Graphics (ACM SIGGRAPH)*, str. 243-250, Los Angeles, ZDA, 1997.

Zahvala

Prof. dr. Franc Solina mi stoji ob strani že od začetka mojega študija. Že leta 1993, ko sem se vpisal na fakulteto, mi je omogočil delo v Laboratoriju za računalniški vid in me tako pravzaprav usmeril na pot, po kateri hodim danes. Med študijem se je moje zanimanje za področje umetne inteligence in računalniškega vida samo še stopnjevalo, rezultat tega pa je bila tudi diplomska naloga pod mentorstvom prof. Soline. Nato sem pod mentorstvom prof. Soline dobil staž mladega raziskovalca, ki ga zaključujem s pričujočim magistrskim delom. — Prof. Solina, najlepša hvala.

Seveda so me na moji poti spremljali tudi sodelavci Laboratorija za računalniški vid, ki so bili vedno dosegljivi za vprašanja in razprave. Posebna zahvala gre tudi nekdanjemu sodelavcu mag. Boru Prihavcu.

Velika zahvala je namenjena tudi moji družini: ženi Nataši, hčerki Petji in sinu Jašu. Seveda, tudi psički Pegi. Hvala za spodbudo, razumevanje, potrpežljivost in popestritev redkih trenutkov sprostitve v zadnjih nekaj mesecih.

Hvala tudi staršema in starim staršem ter ostalim sorodnikom, ki so ves čas šolanja zaupali vame in mi izražali vso podporo.

V spomin očetu Urošu Peeru in staremu očetu Heribertu Kopsetu, ki sta me naučila poštenosti in vztrajnosti.

A življenje teče dalje, v katerega moramo zreti optimistično...

; -)

Izjava

Izjavljam, da sem magistrsko nalogo izdelal samostojno pod vodstvom mentorja prof. dr. Franca Soline. Izkazano pomoč drugih sodelavcev sem v celoti navedel v zahvali.

V Ljubljani, 23. aprila 2001

Peter Peer