

UNIVERZA V LJUBLJANI
FAKULTETA ZA RAČUNALNIŠTVO IN INFORMATIKO

Patrik Vinovrški

**Struktura iz gibanja na osnovi
homogenih regij**

DIPLOMSKO DELO

UNIVERZITETNI ŠTUDIJSKI PROGRAM RAČUNALNIŠTVO
IN INFORMATIKA

MENTOR: doc. dr. Matej Kristan

Ljubljana, 2016

To delo je ponujeno pod licenco *Creative Commons Priznanje avtorstva-Deljenje pod enakimi pogoji 2.5 Slovenija* (ali novejšo različico). To pomeni, da se tako besedilo, slike, grafi in druge sestavine dela kot tudi rezultati diplomskega dela lahko prosto distribuirajo, reproducirajo, uporabljajo, priobčujejo javnosti in predelujejo, pod pogojem, da se jasno in vidno navede avtorja in naslov tega dela in da se v primeru spremembe, preoblikovanja ali uporabe tega dela v svojem delu, lahko distribuira predelava le pod licenco, ki je enaka tej. Podrobnosti licence so dostopne na spletni strani creativecommons.si ali na Inštitutu za intelektualno lastnino, Streliška 1, 1000 Ljubljana.



Izvorna koda diplomskega dela, njeni rezultati in v ta namen razvita programska oprema je ponujena pod licenco GNU General Public License, različica 3 (ali novejša). To pomeni, da se lahko prosto distribuira in/ali predeluje pod njenimi pogoji. Podrobnosti licence so dostopne na spletni strani <http://www.gnu.org/licenses/>.

Besedilo je oblikovano z urejevalnikom besedil L^AT_EX.

Fakulteta za računalništvo in informatiko izdaja naslednjo nalogo:

Tematika naloge:

Struktura iz gibanja je standardni problem računalniškega vida, kjer sta dani dve ali več zaporednih slik zajetih med gibanjem kamere, naloga pa je avtomatsko oceniti položaj kamere med gibanjem in hkrati rekonstrukcijo opazovane scene. Obstoječe metode za ocenjevanje strukture iz gibanja uporabljajo postopke, ki temeljijo na detekciji značilnih točk in robustnih postopkih ocenjevanja epipolarne geometrije. V nalogi obdelajte problem ocenjevanja strukture iz gibanja na podlagi homogenih regij. Predpostavite, da so homogene regije planarne in izpeljite potrebne enačbe za oceno premika kamere iz dveh pogledov. Postopek eksperimentalno ovrednotite.

IZJAVA O AVTORSTVU

diplomskega dela

Spodaj podpisani Patrik Vinovrški,

sem avtor diplomskega dela z naslovom:

Struktura iz gibanja na osnovi homogenih regij

S svojim podpisom zagotavljam, da:

- sem diplomsko delo izdelal samostojno pod mentorstvom doc. dr. Matjeja Kristana
- so elektronska oblika diplomskega dela, naslov (slov., angl.), povzetek (slov., angl.) ter ključne besede (slov., angl.) identični s tiskano obliko diplomskega dela
- soglašam z javno objavo elektronske oblike diplomskega dela v zbirki "Dela FRI".

V Ljubljani, dne 31.08.2016

Podpis avtorja:

Zahvaljujem se mentorju doc. dr. Mateju Kristanu za izjemno odzivnost in pomoč pri izdelavi diplomskega dela.

Zahvaljujem se tudi staršem, ki so mi študij omogočili in me pri tem podpirali, kot tudi ostalim, ki so kakorkoli pripomogli k mojemu dokončanju študija.

Kazalo

Povzetek

Abstract

1	Uvod	1
1.1	Sorodna dela	2
1.2	Prispevki	4
1.3	Struktura	5
2	Teoretična podlaga	7
2.1	Model kamere z luknjično odprtino	7
2.2	Epipolarna geometrija	10
2.3	Ocenjevanje parametrov	11
2.4	Algoritem MSER	14
2.5	Deskriptor SIFT	15
3	Struktura iz gibanja s homogenimi regijami	17
3.1	Primerjava regij	18
3.2	Določanje relativnega položaja kamer	22
3.3	Rekonstrukcija scene	34
4	Rezultati	35
4.1	Ocena homografije	35
4.2	Dekompozicija homografije	37
4.3	Rekonstrukcija scene	38

5 Sklep	41
5.1 Nadaljnje delo	42
Literatura	45

Seznam uporabljenih kratic

kratica	angleško	slovensko
SVD	singular value decomposition	dekompozicija singularnih vrednosti
DLT	direct linear transform	direktna linearna transformacija
RANSAC	random sample consensus	soglasje naključnega vzorca
MSER	maximally stable extremal regions	maksimalno stabilne ekstremalne regije
SIFT	scale invariant feature transform	od merila invariantna transformacija značilk

Povzetek

Naslov: Struktura iz gibanja na osnovi homogenih regij

Pridobivanje strukture iz gibanja zahteva rešitev večih problemov na področju računalniškega vida. Predstavljamo splošen cevovod za gradnjo 3D modelov iz večih kalibriranih slik, kot tudi implementacijo posameznih stopenj. Osnovne stopnje obsegajo pridobivanje značilnk, ujemanje značilnk, ocena položaja kamer in nazadnje rekonstrukcijo scene. Izbira homogenih regij kot značilnk je osnovana na predpostavki, da le te večinoma predstavljajo ravne površine upodobljene scene, kar nam omogoča predstaviti preslikavo regij med slikami s homografijo. Pri iskanju ujemanja značilnk je računsko učinkovitejše regije postaviti v normalizirano obliko ter jih predstaviti z vektorskim deskriptorjem. Predstavljena tehnika normalizacije regij omogoča večjo natančnost v stopnjah ujemanja značilnk in ocene položaja kamer. Položaj kamer ter strukturo scene pridobimo z novo metodo dekompozicije večih homografij, kar nam omogoča ustvariti rekonstrukcijo upodobljene scene. Predstavljene metode so analizirane z uporabo sintetičnih podatkov.

Ključne besede: homogene regije, normalizacija regij, homografija, ocena homografije na osnovi regij, dekompozicija homografije, ocena položaja kamere, struktura iz gibanja.

Abstract

Title: Structure from motion based on homogeneous regions

Acquiring structure from motion requires a solution to multiple problems in the field of computer vision. We present a general pipeline for the reconstruction of 3D models from multiple calibrated images and the implementation of its stages. The main stages are feature extraction, feature matching, camera pose estimation and finally scene reconstruction. The selection of homogeneous regions as features is based on the assumption that they mostly represent planar surfaces of the imaged scene, which allows us to represent mappings of such regions between images with homographies. In the search for feature matches is computationally more efficient to transform the regions into a normalized form and represent it with a vector descriptor. The subsequent region normalization technique improves the accuracy in the stages of feature matching and camera pose estimation. A novel method of multiple homography decomposition is used to obtain the camera poses and scene structure allowing us to create a reconstruction of the imaged scene. The presented methods are analyzed using synthetic data.

Keywords: homogeneous regions, region normalization, homography, region based homography estimation, homography decomposition, camera pose estimation, structure from motion.

Poglavje 1

Uvod

Slika pove več kot tisoč besed, zato se raziskovalci na področju računalniškega vida aktivno ukvarjajo s pridobivanjem informacij iz slik. Podpodročja kot so prepoznavanje predmetov [8, 16, 17, 18], 3D rekonstrukcija [9, 10, 11, 12] in določanje položaja kamere [13, 4, 15, 14], imajo praktični značaj, saj lahko z njihovo uporabo rešujemo probleme kot so navigacija robota v prostoru [19], manipulacija predmetov z robotskimi rokami [20], obogatena resničnost [21], itd. Predvsem zanimiv je problem katere informacije je moč razbrati, če imamo na razpolago dve ali več slik. Ali lahko izluščimo intrinzične ter ekstrinzične parametre kamer, ki so sliki posneli? Ali na slikah lahko prepoznamo predmete ali dele predmetov ter njihovo pozicijo v prostoru? Ali lahko s kakšno dodatno informacijo povemo kaj več o predmetih ali kamerah? V tem delu predstavljamo algoritem, kateri na navedena vprašanja odgovori pritrdilno.

Veliko od zgoraj navedenih problemov je moč rešiti z ujemanjem značilk [13, 4, 15, 14] ter ocenjevanja položaja kamere [13, 15, 14]. Standardni postopki ocenjevanja položaja kamere [13, 4, 15, 14] kot značilke uporabljajo značilne točke. Pomanjkljivost takih metod izhaja iz dejstva, da se natančne lege točk na sliki ne da določiti. Regije, ki predstavljajo ravninske površine v 3D sceni, obsegajo množico točk, zato je njihova lega natančneje določena. V nalogi izpeljemo dokaz, da preslikavi nevzporednih ravnin določajo rela-



Slika 1.1: Diagram cevovoda za rekonstrukcijo scene.

tivni položaj kamer ter oddaljenost in orientacijo opazovanih ravnin. To zakonitost, skupaj s predpostavko, da so zaznane homogene regije povprečno ravne, uporabimo za pridobivanje informacije o strukturi scene. Za celoten postopek lahko uporabimo cevovod s štirimi medsebojno neodvisnimi stopnjami, digram takega cevovoda je prikazan na Sliki 1.1. Slike iz katerih želimo narediti 3D rekonstrukcijo na prvi stopnji razdelimo na homogene regije. Na drugi stopnji nato homogene regije med slikami uparimo in ocenimo pripadajočo homografijo. S pridobljenimi uparitvami in pripadajočimi homografijami nato na tretji stopnji ocenimo relativni položaj kamer, le ta pa se skupaj z uparitvami in pripadajočimi homografijami regij uporabi v četrti stopnji za rekonstrukcijo scene.

1.1 Sorodna dela

Sorodna dela bomo razdelili na tri podsklope, saj vsaka stopnja predstavljenega cevovoda predstavlja neodvisne probleme iz področja računalniškega vida.

Prva stopnja obsega ekstrakcijo značilk, z njimi pa so se ukvarjali Lowe [8] s svojim detektorjem SIFT, Bay [16] z detektorjem SURF, Rosten [17] z detektorjem FAST, Rublee [18] z detektorjem ORB in Matas [7] z detektorjem MSER, kateri je opisan v naslednjem poglavju. Za oceno homografije obstaja nekaj različnih postopkov. Najpogosteje uporabljen je Sutherlandova metoda direktne linearne transformacije opisane v [6]. Metoda minimizira kvadratno vsoto razlik med korespondenčnimi točkami, njena natančnost pa

je močno odvisna od natančnosti pozicije točk. Izboljšavo ocene homografije sta Dubrofsky in Woodham [2] dosegla z uporabo kombinacije daljic in točk. V članku nadgradita metodo direktne linearne transformacije tako, da izpeljeta enačbo, katera ustvari kompatibilnost med korespondencami točk ter korespondencami daljic. Podobno so Zeng, Deng in Hu v [1] za oceno homografije uporabili samo daljice. Ukvarjali so se s problemom nestabilnosti ocene, ko daljice potekajo ob središču. Predstavili so metodo normalizacije daljic in se nestabilnosti ocene homografije izognili.

Problem določanja relativnega položaja kamer se v večini primerov rešuje z dekompozicijo homografije in dekompozicijo esencialne matrike. Z dekompozicijo homografije sta se ukvarjala Faugeras in Zhang. Oba iz izpeljave enačbe za homografijo pridobita rotacijsko matriko \mathbf{R} , relativni premik med kamerami \mathbf{t} in normalo ravnine \mathbf{n} . Faugeras v članku [14] v ta namen uporabi dekompozicijo singularnih vrednosti, katero uporabi na homografiji. Iz tako nastalih ortogonalnih matrik in singularne matrike izpelje osem možnih rešitev za rotacijsko matriko, premik med kamerami in normalo ravnine, nato pa predstavi omejitve, ki napačne rešitve izločijo. Podobno metodo predstavi Zhang v [15], le da njegova metoda dekompozicijo singularnih vrednosti naredi nad produktom homografije \mathbf{H} ter njeno transponirano matriko \mathbf{H}^T . Tudi Zhang izpelje osem rešitev, ter pripadajoči postopek izločanja napačnih rešitev. Najpogosteje uporabljena metoda za oceno relativnega položaja med kamerami je metoda dekompozicije esencialne matrike predstavljene v [13]. Longuet-Higgins dokaže, da je esencialna matrika \mathbf{E} enaka $\mathbf{E} = [t]_x \mathbf{R}$, kjer $[t]_x$ predstavlja relativni premik med kamerami in \mathbf{R} rotacijsko matriko. Z dekompozicijo singularnih vrednosti nato esencialno matriko razstavi ter izračuna relativni položaj kamer. Esencialno matriko Longuet-Higgins oceni s tako imenovanim osem točkovnim algoritmom. Slabost Longuet-Higginsovega osem točkovnega algoritma Hartley predstavi v [4], ter jo odpravi z normaliziranim osem točkovnim algoritmom, kjer pred izračunom esencialne matrike točke normalizira, tako da homogene slikovne koordinate točk ležijo med koordinatama $(-\sqrt{2}, -\sqrt{2})$ in $(\sqrt{2}, \sqrt{2})$.

V tretji sklop sorodnih del spadajo algoritmi za pridobivanje strukture iz gibanja. Formherz in Bischel [9] si pomagata s pridobivanjem oblike iz obrisov. V članku predstavita metodo, s katero najprej pridobita obdajajoči prostor predmeta. Pridobljene točke so natančne le za konveksno ovojnico predmeta, zato na konkavnih delih uporabita integracijo komplementarnih namigov o obliki. Predmetno osredotočeno metodo [10] Fua in Leclerc uporabita za rekonstrukcijo oblike in reflektivnosti površin predmeta. Na začetnem približku površine predmeta poiščeta minimum objektivne funkcije, ki kombinira podatke večih slik. Objektivna funkcija je vsota uteženih vrednosti različnih metod rekonstrukcij, kjer so uteži določene za vsako točko površine in so večje za metode, ki so za obdelovani del slike natančnejše. Minimizacijo kriterijske funkcije uporabljata tudi Rockwood in Winget v [11]. Kriterijsko funkcijo predstavlja energija med predmetom na slikah ter mrežo točk površine predmeta. Z iterativnim algoritmom simuliranega ohlajanja nato energijo minimizirata in pridobita končno mrežo točk površine predmeta. Zhang in Seitz [12] v njunem algoritmu poskušata iterativno prilegati mrežo točk površine predmeta iz splošnega začetnega približka kot je krogla ali kocka. V ta namen mrežo točk opremita s teksturami in nato začetni približek mreže točk iterativno deformirata na podlagi slikovnih in geometričnih sil. Vsaka iteracija zavzema izračun variance barv za točke na mreži, nato premik točk, tako da je varianca čim manjša, deformacijo in delitev mreže, da se prilega premaknjenim točkam ter poenostavitev modela.

1.2 Prispevki

Predstavljen postopek spada v področje pridobivanja strukture iz gibanja. Rezultat postopka je gosto povezana mreža točk strukture scene ter relativni položaj kamer. S predstavljenim algoritmom normalizacije homogenih regij izboljšamo natančnost ocene homografije, kot tudi omogočimo direktno medsebojno primerjavo regij. Ker so homogene regije normalizirane z uporabo projektivne transformacije, so invariantne na vse affine in projektivne tran-

formacije. Zahteva po enotni velikosti normaliziranih regij odstrani potrebo po izdelavi piramide slik različnih velikosti. Dokazana izpeljava homografije omogoča analitično dekompozicijo le-te na intrinzične ter ekstrinzične parametre kamere in informacijo o legi ravnine, ki jo predstavlja regija. Pridobljene informacije so zadostne za metrično 3D rekonstrukcijo scene.

1.3 Struktura

V tej nalogi se najprej v Poglavlju 2 srečamo s teoretično podlago, ki bo služila kot osnova za izpeljavo enačb homografije v Poglavlju 3, ter nekaj algoritmi, s katerimi si bomo pomagali pridobiti osnovne konstrukte, nad katerimi bodo prikazani predstavljeni algoritmi. V Poglavlju 3 sledi predstavitev postopka rekonstrukcije scene. Postopek ima več stopenj zato najprej predstavimo stopnjo primerjanja, kjer je predstavljen nov iterativni algoritem za normalizacijo regij ter pripadajoči lokalno invariantni deskriptor. Sledi stopnja določanja položaja kamere. Tu najprej izpeljemo definicijo homografije, nato pa z analitično dekompozicijo enačb homografije sestavimo poenostavljen sistem enačb, ki ga rešimo z opisanim iterativnim algoritmom na osnovi Gauss-Newtonove metode. V Poglavlju 4 so predstavljeni rezultati vmesnih algoritmov, kot tudi rezultati celotnega postopka. Sklepno Poglavlje 5 povzame vsebino naloge ter rezultate. Splošnost algoritma ter njegove možne uporabe so povzete v Podpoglavju 5.1.

Poglavje 2

Teoretična podlaga

2.1 Model kamere z luknjično odprtino

Hartley in Zisserman [3] sta v odmevni knjigi predstavila nekaj osnovnih pojmov s katerimi se pogosto srečamo na področju računalniškega vida. V knjigi izpeljeta homogeni koordinatni sistem ter z uporabo preprostega modela kamere z luknjično odprtino prikažeta projekcijo 3D koordinatnega sistema na slikovni koordinatni sistem. Pričujoče poglavje temelji na [3] in tudi uporablja oznake vpeljane v [3].

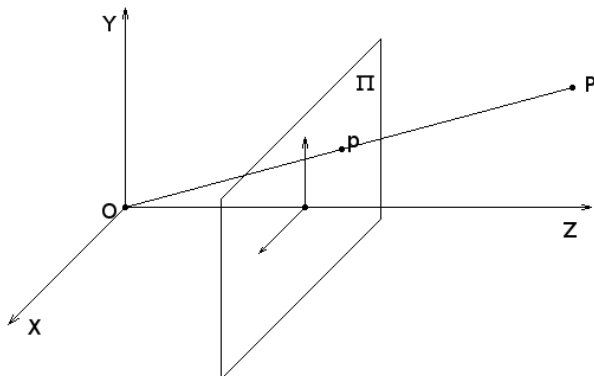
Homogeni koordinatni sistem dobimo tako, da poljubnemu evklidskem prostoru \mathbb{R}^n dodamo še eno koordinato in jo enačimo z 1. S tako spremembo ne pridobimo ali vzamemo informacije, zato je pretvorba med prostori trivialna. V kolikor zadnja koordinata ni enaka 1, pa sta dodala naslednje pravilo.

Izrek 2.1 *V homogenem koordinatnem sistemu za vsako koordinato $\mathbf{X} = [x_1, \dots, x_n, 1]^T$ velja*

$$[x_1, \dots, x_n, 1]^T = [kx_1, \dots, kx_n, k]^T, \quad (2.1)$$

kjer k predstavlja poljuben realni faktor različen od 0.

Prednost homogenega koordinatnega sistema je v temu, da lahko sedaj točko v neskončnosti izrazimo tako, da zadnjo koordinato enačimo z 0. Linearne



Slika 2.1: Model kamere z luknjično odprtino.

transformacije homogenega prostora lahko nato razdelimo na dve osnovni skupini. Skupina afinih transformacij ohranja končnost oziroma neskončnost točk. Projektivne transformacije pa končne točke preslikajo v neskončnost ter obratno. Iz osnov geometrije poznamo zakonitost, da se dve vzporedni premici nikoli ne sekata oziroma njuno sečišče leži v neskončnosti. Na slikah pa lahko opazimo, da lahko sečišču takih premic, recimo železniški tir, določimo točno koordinato. Iz tega sledi da je preslikava realnega prostora na sliko projektivna transformacija iz \mathbb{R}^3 na \mathbb{P}^2 . Najbolj preprost opis take transformacije je model kamere z luknjično odprtino. Na Sliki 2.1 vidimo, da ga sestavljajo slikovna ravnina Π ter goriščna točka \mathbf{O} . Točka na slikovni ravnini \mathbf{p} pa predstavlja sečišče žarka, ki poteka skozi točko \mathbf{X} in goriščno točko, s slikovno ravnino. V matrični obliki lahko transformacijo koordinat točke predmeta $\mathbf{P} = [X, Y, Z, T]^T$ na slikovne koordinate $\mathbf{p} = [x, y, w]^T$ zapišemo

$$\begin{bmatrix} x \\ y \\ w \end{bmatrix} = \mathbf{P}_{3 \times 4} \begin{bmatrix} X \\ Y \\ Z \\ T \end{bmatrix}, \quad (2.2)$$

kjer matrika $\mathbf{P}_{3 \times 4}$ predstavlja projekcijsko matriko kamere, njeni členi pa

predstavljajo intrinzične in ekstrinzične parametre projekcije. Intrinzične parametre lahko združimo v matriko

$$\mathbf{K} = \begin{bmatrix} \alpha_x & \gamma & u_0 \\ 0 & \alpha_y & v_0 \\ 0 & 0 & 1 \end{bmatrix}, \quad (2.3)$$

kjer $\alpha_x = fm_x$ in $\alpha_y = fm_y$ predstavljata goriščno razdaljo v enotah pikslov, f predstavlja goriščno razdaljo v enotah razdalje ter m_x in m_y faktor razmerja med piksli in razdaljo. Parameter γ pove faktor nagnjenosti med x in y osjo, parametra u_0 in v_0 pa predstavljata principalno točko. Med ekstrinzične parametre spada rotacija koordinatnega sistema \mathbf{R} in translacijski vektor goriščne točke $\mathbf{T} = [c_x, c_y, c_z]^T$. Projekcijsko matriko kamere lahko zapišemo kot $\mathbf{P} = \mathbf{K}[\mathbf{R}\mathbf{T}]$. Sedaj z uporabo enačbe (2.2) preslikavo točke $\mathbf{P} = [x, y, z, 1]^T$ na slikovni koordinati točke $\mathbf{p} = [u, v, 1]^T$ izrazimo z naslednjo enačbo:

$$\mathbf{R} = \begin{bmatrix} i_x & i_y & i_z \\ j_x & j_y & j_z \\ k_x & k_y & k_z \end{bmatrix}, \quad (2.4)$$

$$\begin{bmatrix} u' \\ v' \\ w' \end{bmatrix} = \begin{bmatrix} \alpha_x & \gamma & u_0 \\ 0 & \alpha_y & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} i_x & i_y & i_z & c_x \\ j_x & j_y & j_z & c_y \\ k_x & k_y & k_z & c_z \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}, \quad (2.5)$$

$$u = \frac{u'}{w'}, \quad (2.6)$$

$$v = \frac{v'}{w'}. \quad (2.7)$$

V kolikor vse točke ležijo na ravnini, lahko enačbo (2.2) poenostavimo na

$$\begin{bmatrix} x \\ y \\ w \end{bmatrix} = \mathbf{H}_{3 \times 3} \begin{bmatrix} X \\ Y \\ T \end{bmatrix}, \quad (2.8)$$

kjer matrika $\mathbf{H}_{3 \times 3}$ predstavlja projekcijsko matriko ravnine. V kolikor vzamemo dva različna pogleda, kjer točka $\mathbf{p}_1 = [x_1, y_1, w_1]^T$ predstavlja projekcijo točke $\mathbf{P} = [X, Y, Z, T]$ na prvo slikovno ravnino ter točka $\mathbf{p}_2 = [x_2, y_2, w_2]^T$ projekcijo iste točke na drugo ravnino lahko povezavo med njima zapišemo kot:

$$\begin{bmatrix} x_1 \\ y_1 \\ w_1 \end{bmatrix} = \mathbf{H}_1 \begin{bmatrix} X \\ Y \\ T \end{bmatrix}, \quad (2.9)$$

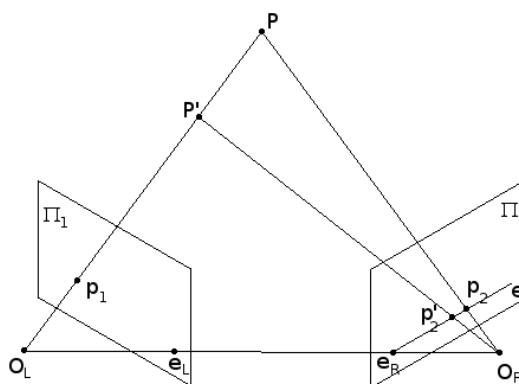
$$\begin{bmatrix} x_2 \\ y_2 \\ w_2 \end{bmatrix} = \mathbf{H}_2 \begin{bmatrix} X \\ Y \\ T \end{bmatrix}, \quad (2.10)$$

$$\begin{bmatrix} x_2 \\ y_2 \\ w_2 \end{bmatrix} = \mathbf{H}_2 \mathbf{H}_1^{-1} \begin{bmatrix} x_1 \\ y_1 \\ z_1 \end{bmatrix} = \mathbf{H} \begin{bmatrix} x_1 \\ y_1 \\ z_1 \end{bmatrix}. \quad (2.11)$$

Projekcijski matriki ravnin \mathbf{H}_1 in \mathbf{H}_2 združimo v matriko \mathbf{H} , ki opisuje preslikavo projekcije ravnine iz prve slikovne ravnine na drugo slikovno ravnino. Le tej pravimo homografija.

2.2 Epipolarna geometrija

V [3] Hartley definira epipolarno geometrijo kot intrinzično projektivno geometrijo med dvema pogledoma. Dalje nadaljuje, da je epipolarna geometrija neodvisna od strukture scene in je odvisna samo od internih parametrov ter relativnega položaja kamer. Slika 2.2 prikazuje epipolarno geometrijo dveh kamer. Točki \mathbf{e}_L in \mathbf{e}_R sta epipolarni točki. Predstavljata preslikavo goriščne točke \mathbf{O}_R na ravnino kamere Π_1 ter preslikavo goriščne točke \mathbf{O}_L na ravnino kamere Π_2 . Vidimo, da točke \mathbf{e}_L , \mathbf{e}_R , \mathbf{O}_L , \mathbf{O}_R vedno ležijo na isti premici. Epipolarna premica e je preslikava premice, ki poteka skozi točki \mathbf{O}_L in \mathbf{p}_2 , na ravnino kamere Π_2 . Pove nam vse možne točke (naprimer točka \mathbf{p}'_2) na kateri se lahko nahaja preslikava točke \mathbf{P} na ravnini Π_2 . Z znanimi koordinatami goriščnih in epipolarnih točk, kot tudi koordinatami projekcije točke \mathbf{P}



Slika 2.2: Epipolarna geometrija dveh kamer.

na ravnini Π_1 ter Π_2 , torej točki \mathbf{p}_1 in \mathbf{p}_2 , lahko izračunamo koordinate točke \mathbf{P} . Ta postopek uporabljamo za pridobitev 3D rekonstrukcije iz dveh ali več slik. Opazimo tudi, da je točko \mathbf{O}_R lahko izrazimo z vektorjem $t_1 \overrightarrow{\mathbf{O}_L \mathbf{e}_L}$, prav tako točko \mathbf{O}_L lahko izrazimo z vektorjem $t_2 \overrightarrow{\mathbf{O}_R \mathbf{e}_R}$.

2.3 Ocenjevanje parametrov

2.3.1 Direktna linearna transformacija

Direktna linearna transformacija je algoritem, ki ga najpogosteje uporabljamo za oceno homografije. Težava nastane pri določanju členov homografije, saj je le ta definirana nad prej predstavljenim projekcijskim prostorom in jo je zaradi tega možno pomnožiti s poljubnim skalarjem in rezultat še vedno predstavlja enako linearno transformacijo:

$$\mathbf{x}_i \propto \mathbf{A} \mathbf{y}_i, \quad (2.12)$$

kjer je \mathbf{x}_i vektor točke v prvem projekcijskem prostoru, \mathbf{y}_i vektor točke v drugem projekcijskem prostoru ter \mathbf{A} matrika linearne transformacije med projekcijskima prostoroma. Tak linearen sistem enačb ni mogoče rešiti enolično,

zato je Sutherland v članku [6] predstavil postopek, po katerem sistem enačb prepisemo in tako dobimo linearen sistem, ki ga je mogoče enolično rešiti. Ravno temu prepisu enačb pravimo direktna linearna transformacija. Na primeru homografije za enačbo (2.11) prepis enačb ustvari naslednji dve enačbi:

$$u_2 = \frac{h_{11}u_1 + h_{12}v_1 + h_{13}}{h_{31}u_1 + h_{32}v_1 + h_{33}}, \quad (2.13)$$

$$v_2 = \frac{h_{21}u_1 + h_{22}v_1 + h_{23}}{h_{31}u_1 + h_{32}v_1 + h_{33}}, \quad (2.14)$$

kjer členi h_{ij} predstavljajo pripadajoče člene homografije \mathbf{H} , u_1 in v_1 sta koordinati točke na prvi slikovni ravnini ter u_2 in v_2 koordinati točke na drugi slikovni ravnini. Enačbi nato lineariziramo ter tako za vsak ujemaajoči par točk na dveh slikovnih ravninah dobimo:

$$h_{11}u_1 + h_{12}v_1 + h_{13} - h_{31}u_1u_2 - h_{32}v_1u_2 - h_{33}u_2 = 0, \quad (2.15)$$

$$h_{21}u_1 + h_{22}v_1 + h_{23} - h_{31}u_1v_2 - h_{32}v_1v_2 - h_{33}v_2 = 0. \quad (2.16)$$

Faktorju h_{33} določimo vrednost 1 in nato z uporabo štirih ali več točk sistem enačb rešimo.

2.3.2 Algoritem Random sample consensus (RANSAC)

Algoritem RANdom SAMple Consensus (RANSAC) sta Fischler in Bolles leta 1981 predstavila v [5]. Ukvarjala sta se z problemom določanja lokacije in orientacije kamere, katera je posnela določene značilke z znano lokacijo v prostoru. V splošnem je to problem prileganja točk na določen matematični model, katerega parametre želimo izračunati. Predhodno so take probleme reševali z metodami osnovanimi na optimizaciji parametrov z najmanjšo kvadratno vstoto napake in nato izločili točko, ki je najdlje od ustreznega modela. Postopek so ponavljali, dokler je bila povprečna napaka dovolj velika. Ugotovila sta, da točka, ki leži zelo daleč od zelenega modela, močno vpliva na parametre modela in zato metode, ki upoštevajo vse točke, za množice točk, ki vsebujejo veliko napačnih podatkov, niso primerne. Razvila sta metodo

RANSAC, ki poskuša iz naključno izbrane najmanjše množice podatkov, ki so potrebni za določitev modela, dobljenemu modelu prilagoditi čim več podatkov iz preostale množice.

Algoritem 1 Algoritem RANSAC.

- 1: **ponavlja**
 - 2: Izberi naključno toliko točk kot je potrebno za določitev parametrov modela
 - 3: Izračunaj parametre modela
 - 4: Preveri koliko točk izmed preostale množice se modelu prilega
 - 5: **dokler** Odstotek prilegajočih točk manjši od zahtevane meje ali število iteracij presega mejo N
 - 6: Ponovno izračunaj model samo s prilegajočimi točkami
-

Rezultat algoritma sta množici prilegajočih ter neprilegajočih se točk. Največje število ponovitev N izberemo tako, da je verjetnost prilegajoče se točke v naključno izbrani začetni množici večja od želenega odstotka (ponavadi $\geq 99\%$).

$$N = \frac{\log(1 - p)}{\log(1 - (1 - v)^m)}, \quad (2.17)$$

kjer p predstavlja želeni odstotek, v verjetnost, da se točka ne prilega modelu ter m najmanjše število točk za določitev modela.

2.3.3 Gauss-Newtonova metoda

Gauss-Newtonova metoda [22] je poimenovana po matematikoma Carlu Friederichu Gaussu ter Isaacu Newtonu. Metoda se uporablja za iskanje minimuma oziroma maksimuma določene funkcije. Predvsem pa je uporabna za prileganje parametrov nelinearnih sistemov enačb z najmanjšo kvadratno vsoto napake. Metoda je izpeljanka Newtonove metode, katere osnova temelji na Taylorjevi ekspanziji $f_T(x)$ funkcije $f(x)$ okoli točke $x = x_n + \Delta x$:

$$f_T(x) = f_T(x_n + \Delta x) \approx f(x_n) + f'(x_n)\Delta x + f''(x_n)\Delta x^2. \quad (2.18)$$

Prednost Gauss-Newtonove metode pred Newtonovo metodo je, da ne potrebuje drugih odvodov funkcij, posledica tega pa je slabost, da jo lahko uporabimo samo za iskanje najmanjše kvadratne vsote vrednosti funkcij. Metoda iterativno izboljšuje začetni približek, dokler ne doseže najmanjše kvadratne vsote funkcij. Sistem m nelinearnih enačb predstavimo s funkcijami $f_i(\mathbf{X}) = 0$, kjer i predstavlja indeks funkcije za pripadajočo enačbo ter \mathbf{X} pa vrednosti spremenljivk. Za n spremenljivk, ki jih želimo oceniti, vzamemo začetni približek $X_0 = [x_1, \dots, x_n]^T$ ter izračunamo vrednosti funkcij za približek $\mathbf{r}(\mathbf{X}) = [f_1(\mathbf{X}), \dots, f_m(\mathbf{X})]^T$. V kolikor imamo večje število enačb kot spremenljivk, $m \geq n$, naslednji približek X_{i+1} izračunamo z

$$X_{i+1} = X_i - (J_r^T J_r)^{-1} J_r^T \mathbf{r}(X_i), \quad (2.19)$$

kjer J_r predstavlja jakobijevo matriko. Jakobijeva matrika je matrika prvih delnih odvodov funkcij, njene člene pa izračunamo z enačbo

$$(R_r)_{ij} = \frac{\partial r_i(\mathbf{X})}{\partial x_j}, \quad (2.20)$$

kjer i in j predtavljata indeks vrstice in stolpca matrike. V kolikor imamo enako število enačb in spremenljivk, $m = n$, se enačba (2.19) poenostavi na

$$X_{i+1} = X_i - J_r^{-1} \mathbf{r}(X_i). \quad (2.21)$$

2.4 Algoritem MSER

Na področju računalniškega vida se pri raznih problemih, kot so 3D rekonstrukcija, zaznavanje lege kamere, detekcija objektov, itd., pojavlja skupna točka. Med različnimi skupinami slik je potrebno najdeti ujemanja oziroma tako imenovane korenspondence. Korenspondence lahko predstavljajo točke, okolico okoli točke ali zaključene regije. Ker so slike na katerih iščemo korespondece lahko posnete z različnimi kamerami in pod različnimi pogoji osvetlitve, moramo le to pri iskanju korespondenc upoštevati. Za blažitev takih razlik je nastal nov razred algoritmov [8, 7], kateri temeljijo na lokalnih invariantnih deskriptorjih. Med take algoritme spada tudi algoritem MSER [7].

Osnovo algoritma predstavljajo ekstremalne regije, katerih značilne lastnosti predstavljajo zaprtost množice pod perspektivnimi transformacijami slikovnih koordinat ter zaprtost množice pod monotonimi spremembami intenzitet na sliki.

2.5 Deskriptor SIFT

V predhodnem poglavju 2.4 omenjeni lokalni invariantni deskriptorji omogočajo bolj robustno določanje korespondenc med slikami posnetimi z različnimi kamerami in pod različnimi pogoji. Eden izmed takih deskriptorjev je deskriptor SIFT, ki ga je predstavil Lowe v [8]. Deskriptor SIFT je vektorski deskriptor, sestavljen tako, da sta si dva vektorja blizu, ko se intezitete na sliki spreminjajo s podobno orientacijo in magnitudo. Deskriptor izračunamo tako, da vzamemo regijo okoli interesne točke ter jo razdelimo v $n \times n$ kvadratov enake velikosti. Nato za vsak kvadratni del regije izračunamo histogram orientacij intezitet za pripadajoče slikovne točke.

Poglavje 3

Struktura iz gibanja s homogenimi regijami

V tem poglavju se bomo ukvarjali z vprašanjem kako pridobiti informacijo o strukturi scene iz dveh kalibriranih pogledov. Problem pridobivanja strukture scene razdelimo na manjše podprobleme, ki jih rešujemo v posameznih stopnjah predstavljenega postopka. Hartley [3] v splošnem postopek razdeli na štiri stopnje. Na prvi stopnji moramo iz slik pridobiti značilke. Značilke morajo vsebovati vsaj nekaj informacije o preslikavi strukture scene na slike. Lahko so recimo točke, daljice, krivulje ali množice točk. Nato moramo na drugi stopnji značilke na različnih slikah medsebojno upariti. Ker značilke nosijo nekaj informacije o preslikavi, lahko na tretji stopnji iz več uparjenih značilk pridobimo informacijo o relativnem položaju kamer, ki so sliki posneli. Z uporabo epipolarne geometrije, informacije o položaju kamer in položajem značilk, nato na četrti stopnji izračunamo lego značilk v prostoru in tako dobimo rekonstrukcijo strukture scene. Predstavljen postopek za značilke uporablja homogene regije, saj lahko za le te sklepamo da so povprečno ravninske. V prvi fazi zato za detekcijo homogenih regij uporabimo algoritem MSER [7].

3.1 Primerjava regij

Dve regiji lahko medsebojno primerjamo na več načinov. Lahko poiščemo korespondenco med točkami na eni ter drugi regiji, lahko poiščemo korespondenco med daljicami robov regije ali primerjamo celotno obliko regije oziroma njen obris. Problem se pojavlja pri primerjavi oblik med regijami, saj se pri projektivni transformaciji le te deformirajo in se jih ne da direktno primerjati. Rešitev predstavlja normalizacija regije z uporabo projektivne transformacije. Ob predpostavki, da točke regije ležijo na ravnini v prostoru lahko preslikavo regije v normalno obliko izrazimo s homografijo.

Izrek 3.1 *Normalizirano obliko regije R_n definiramo kot homografijo regije R na enotno krožnico, kjer je površina regije znotraj krožnice največja. Med točkami regije X_i ter točkami normalizirane regije X_{ni} določimo preslikavo H_n tako, da velja:*

$$X_{ni} = H_n X_i, X_i \in R, \quad (3.1)$$

$$\|X_{ni}\| \leq 1, \quad (3.2)$$

$$\max \text{Area}(X_{n0} \dots X_{nk}), X_{n0} \dots X_{nk} \in R_n. \quad (3.3)$$

Dve tako normalizirani regiji sta direktno primerljivi, saj lahko točke prve regije X_{1n} , njeno normalizirano preslikavo H_{n1} , točke druge regije X_{2n} in njeno normalizirano preslikavo H_{n2} povežemo s skupno preslikavo $H_{n2}^{-1}H_{n1}$ na sledeči način:

$$X_{1n} = H_{n1} X_1, \quad (3.4)$$

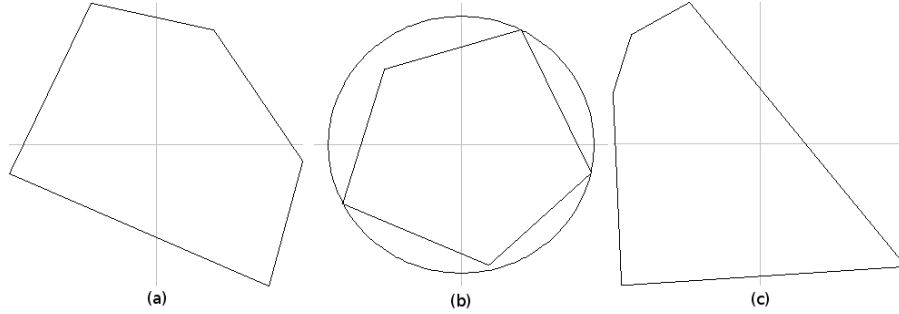
$$X_{2n} = H_{n2} X_2, \quad (3.5)$$

$$X_{1n} = X_{2n}, \quad (3.6)$$

$$H_{n2} X_2 = H_{n1} X_1, \quad (3.7)$$

$$X_2 = H_{n2}^{-1} H_{n1} X_1. \quad (3.8)$$

Dve tako normalizirani regiji predstavljata projektivno transformacijo vsake regije na skupno slikovno ravnino, zato je normalna oblika regije invariantna



Slika 3.1: Primeri normaliziranih regij.

na transformacije rotacije, premika, nagiba, spremembe velikosti in projekcije. Primer transformacije na skupno ravnino vidimo na Sliki 3.1, kjer del (a) predstavlja obris regije levega pogleda, del (c) obris regije desnega pogleda ter del (b) njuno normalno obliko.

3.1.1 Algoritem za iterativno normalizacijo regije

Iz same definicije (3.1) in (3.3) vidimo, da na površino regije \mathbf{R} vplivajo samo ogljišča konveksne ovojnice $\mathbf{X}_k \in \mathbf{K}$ okoli točk $\mathbf{X}_i \in \mathbf{R}$;

$$\text{Area}(\mathbf{X}_{k0}, \dots, \mathbf{X}_{kn}) = \text{Area}(\mathbf{X}_{k0}, \dots, \mathbf{X}_{kn}, \mathbf{X}_i), \mathbf{X}_i \notin \mathbf{K}. \quad (3.9)$$

Zato lahko pri normalizaciji uporabimo samo točke na konveksni ovojnici. Za normalizacijo regije lahko uporabimo numerično fizikalno aproksimacijo na osnovi vzmeti.

Izrek 3.2 *Fizikalni sistem definiramo tako, da so vsa ogljišča konveksne ovojnice $\mathbf{X}_i \in \mathbf{K}$ medsebojno povezana z vzmetmi enake dolžine, nato pa iščemo postavitev ogljišč \mathbf{K}^* , ki predstavlja minimalno energijo sistema:*

$$\mathbf{K}^* = \underset{\mathbf{K}}{\text{argmin}} E(\mathbf{K}), \quad (3.10)$$

$$\mathbf{E}(\mathbf{K}) = \sum \|F(\mathbf{X}_i)\|, \quad (3.11)$$

$$F(\mathbf{X}_i) = \sum F(\mathbf{X}_i, \mathbf{X}_j) + F_{\text{ext}}, \mathbf{X}_i, \mathbf{X}_j \in \mathbf{K} \wedge i \neq j, \quad (3.12)$$

$$F(\mathbf{X}_i, \mathbf{X}_j) = -k(\|\mathbf{X}_i - \mathbf{X}_j\| - l). \quad (3.13)$$

Celotno energijo sistema predstavlja vsota sil $F(X_i)$, ki deluje na posamezne točke, in jo predstavimo z energijsko funkcijo $E(\mathbf{K})$. Sile, ki delujejo na točko, so sile vzmeti med točkami $F(\mathbf{X}_i, \mathbf{X}_j)$, katere izračunamo z uporabo koeficienta trdnosti vzmeti k in dolžino vzmeti l , in zunanje sile, ki jih predstavimo s faktorjem F_{ext} .

V primeru, da je F_{ext} enaka 0, ima sistem najmanjšo energijo, ko konveksna ovojnica predstavlja pravilni N-kotnik. V kolikor pravih N-kotnikov na sliki ne moremo zaznati, je potrebno sistem omejiti tako, da dovolimo premik točk samo v smeri, katero dovoljuje projektivna transformacija. To lahko dosežemo tako, da izračunamo homografijo med prvotnimi pozicijami X_{i-1} točk ter njihovimi zelenimi premiki X_i :

$$\mathbf{X}_i = \mathbf{H}_i \mathbf{X}_{i-1}, \quad (3.14)$$

$$\mathbf{X}_i = \mathbf{X}_{i-1} + \mathbf{F}(\mathbf{X}_{i-1}), \quad (3.15)$$

kjer je $\mathbf{F}(\mathbf{X}_{i-1})$ sila, ki na točko deluje. Za reševanje zgornjega problema lahko uporabimo direktni linearni transform. Po nekaj iteracijah (odvisno od zelene natančnosti) se sistem stabilizira, dosežena je najmanjša energija sistema. Pogoji (3.10) je dosežen, potrebno je še doseči pogoji (3.2). Doseg pogoja (3.2) je trivialen, saj je potrebno konveksno ovojnico le premakniti tako, da je težišče ovojnice v izhodišču koordinatnega sistema, ter da točka, ki je najdlje od težišča, leži na krožnici. Opisani postopek je povzet v Algoritmu 2.

Regija z največjo površino znotraj krožnice predstavlja tudi najmanjšo elipso okoli množice točk s štirimi prostostnimi stopnjami. Splošno najmanjšo elipso s petimi prostorskimi stopnjami lahko dobimo, če v Algoritmu 2 namesto izračuna homografije uporabimo splošnejšo obliko linearne transformacije. Enačbam direktne linearne transformacije (2.13) in (2.14) dodamo še člen $x_1 y_1$:

$$x_2 = \frac{f_{11}x_1 + f_{12}y_1 + f_{13}x_1y_1 + f_{14}}{f_{31}x_1 + f_{32}y_1 + f_{33}x_1y_1 + f_{34}}, \quad (3.16)$$

$$y_2 = \frac{f_{21}x_1 + f_{22}y_1 + f_{23}x_1y_1 + f_{24}}{f_{31}x_1 + f_{32}y_1 + f_{33}x_1y_1 + f_{34}}. \quad (3.17)$$

Algoritem 2 Algoritem za normalizacijo regij.

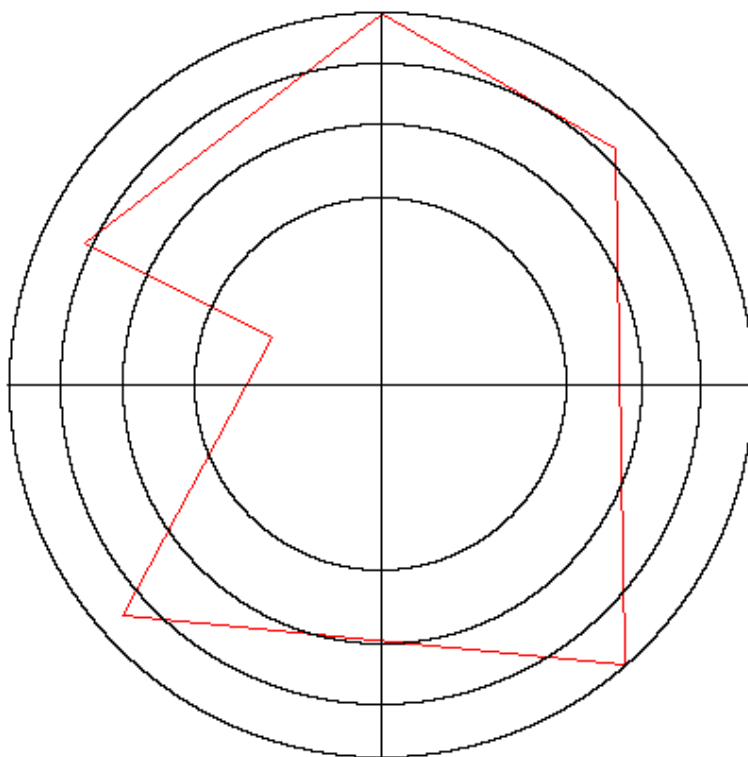
- 1: **ponavlja**j
 - 2: Za vsako točko na konveksni ovojnici izračunajte silo katera na točko deluje.
 - 3: Izračunajte začasno konveksno ovojnico iz trenutne konveksne ovojnice, tako da vsako točko premaknete za vektor pripadajoče sile
 - 4: Izračunajte homografijo med trenutno konveksno ovojnico ter začasno konveksno ovojnico z metodo DLT.
 - 5: Preslikaj trenutno konveksno ovojnico kot to zahteva homografija
 - 6: **dokler** Izračunana homografija ni blizu identitete
 - 7: Točke konveksne ovojnice premakni tako, da težišče ovojnice leži v središču enotnega kroga ter točka najdlje od središča leži na krožnici.
-

Iz ocenjenih faktorjev f_{ij} , $i \in \{1, 2, 3\} \wedge j \in \{1, 2, 3, 4\}$ lahko sedaj izračunamo parametre najmanjše elipse.

3.1.2 Primerjava normaliziranih regij

Za optimalno primerjavo normaliziranih regij je najbolje uporabiti eno izmed oblik vektorskih deskriptorjev. V večini primerov se v te namene uporablja deskriptor SIFT [8], vendar za potrebe normalizacije po postopku 2 predhodno že izračunamo obris regije. Če obris regije dalje razdelimo na daljice, dosežemo manjše število operacij za izračun konveksne ovojnice, saj je potrebno upoštevati le začetne in končne točke daljic. Ker algoritem MSER [7] vrne regije invariantne na spremembo intenzitete svetlobe, je tudi nepotrebno da bi regijo iz slike transformirali in iz nje ustvarili deskriptor SIFT. Zato je bolje uporabiti deskriptor podoben deskriptorju SIFT, le da histogram gradientov intenzitet zamenjamo s histogramom orientacije daljic. Opazimo lahko, da je prostor zunaj enotne krožnice vedno prazen, zato lahko namesto kvadratnih regij uporabimo krožne izseke enakih površin. Višine posameznih izsekov h_i dobimo s formulo

$$h_i = \sqrt{\frac{1}{N}}, \quad (3.18)$$



Slika 3.2: Krožni deskriptor z normalizirano regijo.

kjer je kot krožnega izseka $\alpha = 2\pi/M$ in N število delitev po višini. Skupaj s številom košev histograma O lahko dolžino deskriptorja izračunamo z $M \times N \times O$. Na Sliki 3.2 imamo primer deskriptorja razdeljenega na 4×4 krožnih izsekov ter normalizirano regijo v rdečem. Za vsako točko regije, ki leži na notni krožnici ustvarimo svoj deskriptor, tako da leži na točki $(0,1)$. Koše histograma orientacij daljic nato napolnemo z dolžinami daljic, ki dani krožni izsek prečkajo. Vsak histogram nato normaliziramo, tako da je skupna dolžina daljic v največjem košu enaka 1.

3.2 Določanje relativnega položaja kamer

Najpogosteje uporabljen pristop za določanje relativnega položaja kamer je Hartleyev normalizirani osem točkovni algoritem [4]. Alternativno Hartleye-

vem algoritmu predstavlja metoda dekompozicije homografije. Pri dekompoziciji homografije predstavlja problem to, da je enolično določena s štirimi točkami, kar močno omejuje njeno natančnost. Faugeras [14] in Zhang [15] homografijo \mathbf{H} izpeljeta kot $\mathbf{H} = \mathbf{R} + \mathbf{t}\mathbf{n}^T$, kjer je \mathbf{R} rotacijska matrika, \mathbf{t} vektor premika med kamerami in \mathbf{n} normala ravnine. Če želimo pri dekompoziciji homografije doseči večjo natančnost, moramo uporabiti dve ali več homografij in zato enačbo homografije izpeljati drugače. Iz nove izpeljave homografije lahko naredimo analitično dekompozicijo in tako določimo relativni položaj kamer.

3.2.1 Izpeljava enačbe homografije

Pri izpeljavi enačbe homografije bomo uporabili model kamere z luknjičasto odprtino ter predpostavko, da so intrinzični parametri kamere γ , u_0 in v_0 enaki 0 ter α_x in α_y enaka $1/f_1$ v primeru prve kamere in $1/f_2$ v primeru druge kamere. Za izračun slikovnih koordinat točke $[u_1, v_1]$ na prvi sliki v enačbah (2.5), (2.6) in (2.7) člene rotacijske matrike nadomestimo z identiteto, člene vektorja premika pa enačimo z $[0, 0, 0]^T$. Pridobimo naslednji enačbi:

$$u_1 = \frac{x}{zf}, \quad (3.19)$$

$$v_1 = \frac{y}{zf}. \quad (3.20)$$

Ob predpostavki, da točke ležijo na ravnini, lahko uporabimo enačbo

$$n_x x + n_y y + n_z z + d = 0, \quad (3.21)$$

kjer vektor $[n_x, n_y, n_z]^T$ predstavlja normalo na ravnino, d pa oddaljenost ravnine od izhodišča. Za slikovne koordinate točke $[u_2, v_2]$ ekstrinzične parametre rotacije nadomestimo s pripadajočimi enotnimi vektorji $\mathbf{i} = [i_x, i_y, i_z]^T$, $\mathbf{j} = [j_x, j_y, j_z]^T$ in $\mathbf{k} = [k_x, k_y, k_z]^T$, vektor translacije $\mathbf{c} = [c_x, c_y, c_z]^T$ pa predstavimo k točki $\mathbf{P} = [x, y, z, 1]^T$. Sedaj sta koordinati izraženi z naslednjima enačbama:

$$u_2 = \frac{i_x x + i_y y + i_z z - \mathbf{c}^T \mathbf{i}}{f_2(k_x x + k_y y + k_z z - \mathbf{c}^T \mathbf{k})}, \quad (3.22)$$

$$v_2 = \frac{j_x x + j_y y + j_z z - \mathbf{c}^T \mathbf{j}}{f_2(k_x x + k_y y + k_z z - \mathbf{c}^T \mathbf{k})}. \quad (3.23)$$

Iz enačb (3.19), (3.20) in (3.21) izpostavimo člene x , y in z ter jih vstavimo v enačbi (3.22) ter (3.23) nato dobimo preslikavo med točkami prvega ter drugega slikovnega koodrinatnega sistema:

$$x = -\frac{b_{x1} df_1}{b_{x1} f_1 n_x + b_{y1} f_1 n_y + n_z}, \quad (3.24)$$

$$y = -\frac{b_{y1} df_1}{b_{x1} f_1 n_x + b_{y1} f_1 n_y + n_z}, \quad (3.25)$$

$$z = -\frac{d}{b_{x1} f_1 n_x + b_{y1} f_1 n_y + n_z}, \quad (3.26)$$

$$b_{x2} = \frac{b_{x1} f_1 (di_x + n_x \mathbf{c}^T \mathbf{i}) + b_{y1} f_1 (di_y + n_y \mathbf{c}^T \mathbf{i}) + di_z + n_z \mathbf{c}^T \mathbf{i}}{b_{x1} f_1 f_2 (dk_x + n_x \mathbf{c}^T \mathbf{k}) + b_{y1} f_1 f_2 (dk_y + n_y \mathbf{c}^T \mathbf{k}) + f_2 (dk_z + n_z \mathbf{c}^T \mathbf{k})}, \quad (3.27)$$

$$b_{y2} = \frac{b_{x1} f_1 (dj_x + n_x \mathbf{c}^T \mathbf{j}) + b_{y1} f_1 (dj_y + n_y \mathbf{c}^T \mathbf{j}) + dj_z + n_z \mathbf{c}^T \mathbf{j}}{b_{x1} f_1 f_2 (dk_x + n_x \mathbf{c}^T \mathbf{k}) + b_{y1} f_1 f_2 (dk_y + n_y \mathbf{c}^T \mathbf{k}) + f_2 (dk_z + n_z \mathbf{c}^T \mathbf{k})}. \quad (3.28)$$

Vidimo, da imata enačbi (3.27) in (3.28) enako strukturo kot enačbi (2.13) in (2.14) ter zato predstavljata prepis homografije z metodo direktne linearne transformacije. Izpostaviti je potrebno tudi prosti člen, torej tisti, ki ni vezan na spremenljivki u_1 ali v_1 , saj je le ta lahko 0 samo v primeru, ko opazovana ravnina poteka skozi goriščno točko druge kamere in so zato točke na drugi sliki kolinearne.

3.2.2 Dekompozicija homografije

V prejšnjem podpoglavju smo dokazali da členi enačb (3.27) in (3.28) predstavljajo člene homografije \mathbf{H}

$$\mathbf{H} = \begin{bmatrix} h_{11} & h_{12} & h_{31} \\ h_{21} & h_{22} & h_{32} \\ h_{31} & h_{32} & h_{33} \end{bmatrix}. \quad (3.29)$$

Člene lahko enačimo in tako pridobimo naslednji sistem enačb;

$$f_1 (di_x + n_x \mathbf{c}^T \mathbf{i}) = h_{11} h_{33}, \quad (3.30)$$

$$f_1(di_y + n_y \mathbf{c}^T \mathbf{i}) = h_{12} h_{33}, \quad (3.31)$$

$$di_z + n_z \mathbf{c}^T \mathbf{i} = h_{13} h_{33}, \quad (3.32)$$

$$f_1(dj_x + n_x \mathbf{c}^T \mathbf{j}) = h_{21} h_{33}, \quad (3.33)$$

$$f_1(dj_y + n_y \mathbf{c}^T \mathbf{j}) = h_{22} h_{33}, \quad (3.34)$$

$$dj_z + n_z \mathbf{c}^T \mathbf{j} = h_{23} h_{33}, \quad (3.35)$$

$$f_1 f_2(dk_x + n_x \mathbf{c}^T \mathbf{k}) = h_{31} h_{33}, \quad (3.36)$$

$$f_1 f_2(dk_y + n_y \mathbf{c}^T \mathbf{k}) = h_{32} h_{33}, \quad (3.37)$$

$$f_2(dk_z + n_z \mathbf{c}^T \mathbf{k}) = h_{33}. \quad (3.38)$$

Sistem sestavlja 18 nelinearno povezanih spremenljivk. Ker so $i_x, i_y, i_z, j_x, j_y, j_z, k_x, k_y, k_z$ komponente enotnih vektorjev veljajo tudi naslednje omejitve:

$$i_x^2 + i_y^2 + i_z^2 = 1, \quad (3.39)$$

$$j_x^2 + j_y^2 + j_z^2 = 1, \quad (3.40)$$

$$i_x j_x + i_y j_y + i_z j_z = 0, \quad (3.41)$$

$$k_x = i_y j_z - i_z j_y, \quad (3.42)$$

$$k_y = i_z j_x - i_x j_z, \quad (3.43)$$

$$k_z = i_x j_y - i_y j_x. \quad (3.44)$$

Vidimo, da sistem enačb (3.30)–(3.38) skupaj z zgoraj navedenimi omejitvami (3.39)–(3.44) imamo premalo enačb za rešitev sistema, zato potrebujemo najmanj dve homografiji. Upoštevaajoč epipolarno geometrijo lahko goriščni točki kamer $\mathbf{O}_L = [0, 0, 0]^T$ in $\mathbf{O}_R = [c_x, c_y, c_z]^T$ izrazimo z vektorjema $t_1 \overrightarrow{\mathbf{O}_L \mathbf{e}_L}$ in $t_2 \overrightarrow{\mathbf{O}_R \mathbf{e}_R}$, zato pridobimo tudi naslednji dve omejitvi:

$$t_1 [e_{Lx}, e_{Ly}, \frac{1}{f_1}]^T = [c_x, c_y, c_z]^T, \quad (3.45)$$

$$-t_2 [e_{Rx} i_x + e_{Ry} j_x + \frac{k_x}{f_2}, e_{Rx} i_y + e_{Ry} j_y + \frac{k_y}{f_2}, e_{Rx} i_z + e_{Ry} j_z + \frac{k_z}{f_2}]^T = [c_x, c_y, c_z]^T. \quad (3.46)$$

Iz sistema enačb lahko opazimo, da so faktorji d , c_x , c_y , c_z medsebojno linearno odvisni, kar nam omogoča oceniti le relativno razdaljo ravnine, glede na relativni premik kamere po epipolarni osi. Dve homografiji si delita faktorje i_x , i_y , i_z , j_x , j_y , j_z , k_x , k_y , k_z , c_x , c_y , c_z , f_1 , f_2 . Člene prve homografije označimo s h_1^{ij} , kjer so i in j indeksi pozicije člena v matriki homografije, podobno označimo tudi člene druge homografije h_2^{ij} . Faktorje ravnine, katere preslikavo določa prva homografija označimo z n_{x1} , n_{y1} , n_{z1} , d_1 , faktorje ravnine druge homografije pa z n_{x2} , n_{y2} , n_{z2} , d_2 . Sedaj lahko izpeljemo pet osnovnih enačb, ki jih bomo uporabili za dekompozicijo homografije.

Dokaz. Dokaz drži le ob predpostavki, da faktorji n_{x1} , $\mathbf{c}^T \mathbf{i}$, $\mathbf{c}^T \mathbf{j}$, $\mathbf{c}^T \mathbf{k}$ niso enaki 0. Za izpeljavo prve enačbe najprej uporabimo enačbo (3.23). Za vsako izmed homografij izpostavimo faktor n_{xi} , $i \in \{1, 2\}$, tako nastali enačbi nato medsebojno delimo ter izpostavimo faktor n_{x2} :

$$n_{x1} = \frac{h_1^{11}h_1^{33} - f_1d_1i_x}{f_1\mathbf{c}^T\mathbf{i}}, \quad (3.47)$$

$$n_{x2} = \frac{h_2^{11}h_2^{33} - f_1d_2i_x}{f_1\mathbf{c}^T\mathbf{i}}, \quad (3.48)$$

$$n_{x2} = n_{x1} \frac{h_2^{11}h_2^{33} - f_1d_2i_x}{h_1^{11}h_1^{33} - f_1d_1i_x}. \quad (3.49)$$

Novonastalo izpeljavo faktorja n_{x2} iz enačbe (3.49) lahko sedaj zamenjamo v enačbah (3.33) in (3.36), iz njih pa izpostavimo faktor n_{x1} :

$$n_{x1} = \frac{h_1^{21}h_1^{33} - f_1d_1j_x}{f_1\mathbf{c}^T\mathbf{j}}, \quad (3.50)$$

$$n_{x1} = \frac{(h_2^{21}h_2^{33} - f_1d_2j_x)(h_1^{11}h_1^{33} - f_1d_1i_x)}{f_1\mathbf{c}^T\mathbf{j}(h_2^{11}h_2^{33} - f_1d_2i_x)}, \quad (3.51)$$

$$n_{x1} = \frac{h_1^{31}h_1^{33} - f_1f_2d_1k_x}{f_1f_2\mathbf{c}^T\mathbf{k}}, \quad (3.52)$$

$$n_{x1} = \frac{(h_2^{31}h_2^{33} - f_1f_2d_2k_x)(h_1^{11}h_1^{33} - f_1d_1i_x)}{f_1f_2\mathbf{c}^T\mathbf{k}(h_2^{11}h_2^{33} - f_1d_2i_x)}. \quad (3.53)$$

Sedaj enačbi (3.50) in (3.51) združimo ter enako storimo z enačbami (3.52) in (3.53). Iz novo nastalih enačb izpostavimo faktor d_2 . Ko le ti združimo,

vidimo da bi moral faktor d_1 biti enak $\frac{h_1^{11}h_1^{33}}{f_1i_x}$. Iz enačbe (3.30) ter predpostavke da n_{x1} , $\mathbf{c}^T\mathbf{i}$, h_2^{33} niso enaki 0 vidimo, da to ne more držati, kar pomeni da velja:

$$f_2k_x(h_2^{11}h_1^{21} - h_1^{11}h_2^{21}) + i_x(h_2^{21}h_1^{31} - h_1^{21}h_2^{31}) + j_x(h_1^{11}h_2^{31} - h_2^{11}h_1^{31}) = 0. \quad (3.54)$$

□

Dokaz. Pravilnost enačbe (3.55) dokažemo po enakem postopku kot smo to naredili za enačbo (3.54), le da enačbe (3.30), (3.33) in (3.36) zamenjamo z enačbami (3.31), (3.34) in (3.37), ter namesto faktorjev n_{x1} in n_{x2} izpostavimo faktorja n_{y1} in n_{y2} . V tem dokazu velja podobna predpostavka, to je, da faktorji n_{y1} , $\mathbf{c}^T\mathbf{i}$, $\mathbf{c}^T\mathbf{j}$, $\mathbf{c}^T\mathbf{k}$ niso enaki 0. Rezultat predstavlja naslednja enačba:

$$f_2k_y(h_2^{12}h_1^{22} - h_1^{12}h_2^{22}) + i_y(h_2^{22}h_1^{32} - h_1^{22}h_2^{32}) + j_y(h_1^{12}h_2^{32} - h_2^{12}h_1^{32}) = 0. \quad (3.55)$$

□

Dokaz. Tudi za dokaz enačbe (3.56) uporabimo enak postopek kot pri dokazu enačbe (3.54). Tukaj namesto enačb (3.30), (3.33) in (3.36) uporabimo enačbe (3.32), (3.35) in (3.38) ter namesto faktorjev n_{x1} in n_{x2} izpostavimo faktorja n_{z1} in n_{z2} . Tudi v tem primeru mora veljati, da n_{z1} , $\mathbf{c}^T\mathbf{i}$, $\mathbf{c}^T\mathbf{j}$, $\mathbf{c}^T\mathbf{k}$ niso enaki 0. Rezultat je naslednja enačba:

$$f_2k_z(h_2^{13}h_1^{23} - h_1^{13}h_2^{23}) + i_z(h_1^{23} - h_2^{23}) + j_z(h_2^{13} - h_1^{13}) = 0. \quad (3.56)$$

□

Dokaz. Dokaz enačbe (3.57) sloni na preslikavi točk $p_1 = (0, 0)$, $p_2 = (0, 1)$, $p_3 = (1, 0)$, $p_4 = (1, 1)$ z obema homografijama. Preslikave točk s prvo homografijo označimo s p_1^i , preslikave točk z drugo homografijo pa s p_2^i , kjer i predstavlja indeks prej naštetih točk. Sečišče premice, ki poteka skozi točki p_1^1 in p_2^2 , s premico, ki poteka skozi točki p_1^3 in p_1^4 označimo s točko t_1 , sečišče premice, ki poteka skozi točki p_2^1 in p_2^2 , s premico, ki poteka skozi točki p_2^3 in p_2^4 označimo s točko t_2 . Vidimo da ima normalizirani vektor $\overrightarrow{t_1t_2}$ člene enake $[\frac{c_zj_y - c_yj_z}{c_yi_z - c_zi_y}, \frac{c_yi_z - c_zi_y}{c_zj_y - c_yj_z}] = [v_1, \frac{1}{v_1}]$. Vzamemo lahko koordinato x normaliziranega

vektorja ter na njej uporabimo enačbo (3.45). Pridobimo naslednjo enačbo:

$$j_y + i_y v_1 - e_{Ly} f_1(j_z + i_z v_1) = 0. \quad (3.57)$$

Ekvivalentno izpeljavo predstavlja tudi naslednja enačba:

$$\frac{c_z j_y - c_y j_z}{c_y i_z - c_z i_y} = \frac{h_2^{11} h_1^{31} - h_1^{11} h_2^{31}}{h_2^{21} h_1^{31} - h_1^{21} h_2^{31}}. \quad (3.58)$$

□

Dokaz. Dokazovanje enačbe (3.59) sledi enakemu postopku kot dokazovanje enačbe (3.57) le, da točko t_1 predstavlja presečišče premice skozi točki p_1^1 in p_1^3 ter premice skozi točki p_1^2 in p_1^4 , točko t_2 pa presečišče premice skozi točki p_2^1 in p_2^3 ter premice skozi točki p_2^2 in p_2^4 . Normalizirani vektor $\overrightarrow{t_1 t_2}$, ki ga s to izpeljavo dobimo, pa ima člene enake $[\frac{c_z j_x - c_x j_z}{c_x i_z - c_z i_x}, \frac{c_x i_z - c_z i_x}{c_z j_x - c_x j_z}] = [v_2, \frac{1}{v_2}]$. Ponovno na koordinati x uporabimo enačbo (3.45) in tako dobimo:

$$j_x + i_x v_2 - e_{Lx} f_1(j_z + i_z v_2) = 0. \quad (3.59)$$

Ekvivalentna enačba za to izpeljavo pa je naslednja:

$$\frac{c_z j_x - c_x j_z}{c_x i_z - c_z i_x} = \frac{h_2^{12} h_1^{32} - h_1^{12} h_2^{32}}{h_2^{22} h_1^{32} - h_1^{22} h_2^{32}}. \quad (3.60)$$

□

Izrek 3.3 *Sistem enačb dveh homografij lahko poenostavimo na naslednje enačbe:*

$$f_2 k_x (h_2^{11} h_1^{21} - h_1^{11} h_2^{21}) + i_x (h_2^{21} h_1^{31} - h_1^{21} h_2^{31}) + j_x (h_1^{11} h_2^{31} - h_2^{11} h_1^{31}) = 0, \quad (3.61)$$

$$f_2 k_y (h_2^{12} h_1^{22} - h_1^{12} h_2^{22}) + i_y (h_2^{22} h_1^{32} - h_1^{22} h_2^{32}) + j_y (h_1^{12} h_2^{32} - h_2^{12} h_1^{32}) = 0, \quad (3.62)$$

$$f_2 k_z (h_2^{13} h_1^{23} - h_1^{13} h_2^{23}) + i_z (h_1^{23} - h_2^{23}) + j_z (h_2^{13} - h_1^{13}) = 0, \quad (3.63)$$

$$j_y + i_y v_1 - e_{Ly} f_1(j_z + i_z v_1) = 0, \quad (3.64)$$

$$j_x + i_x v_2 - e_{Lx} f_1(j_z + i_z v_2) = 0. \quad (3.65)$$

3.2.3 Izrojeni sistemi

Enačbe izreka 3.3 predstavljajo rešitev le v primeru, ko velja, da $n_{x1}, n_{y1}, n_{z1}, \mathbf{c}^T \mathbf{i}, \mathbf{c}^T \mathbf{j}, \mathbf{c}^T \mathbf{k}$ niso enaki 0. V tem podpoglavju so predstavljene nadomestne rešitve za enačbe izreka 3.3, ko navedeni pogoji ne veljajo, in kako take pogoje zaznamo.

Goriščni točki obeh kamer v isti točki

V kolikor se dva pogleda razlikujeta le z rotacijo okoli izhodišča, vidimo, da se sistem enačb iz definicije homografije poenostavi tako, da izpadejo vsi faktorji, ki se nanašajo na opazovano ravnino. Ostanejo le faktorji $i_x, i_y, i_z, j_x, j_y, j_z, k_x, k_y, k_z, f_1$ ter f_2 , zato jih lahko določimo le z uporabo ene homografije. Posebej je potrebno omeniti, da člen k_z ne sme biti enak 0, saj v nasprotnem primeru rešitve ne moremo dobiti. Rešiti je potrebno naslednji sistem:

$$\frac{f_1}{f_2} i_x = h_{11}(i_x j_y - i_y j_x), \quad (3.66)$$

$$\frac{f_1}{f_2} i_y = h_{12}(i_x j_y - i_y j_x), \quad (3.67)$$

$$\frac{1}{f_2} i_z = h_{13}(i_x j_y - i_y j_x), \quad (3.68)$$

$$\frac{f_1}{f_2} j_x = h_{21}(i_x j_y - i_y j_x), \quad (3.69)$$

$$\frac{f_1}{f_2} j_y = h_{22}(i_x j_y - i_y j_x), \quad (3.70)$$

$$\frac{1}{f_2} j_z = h_{23}(i_x j_y - i_y j_x), \quad (3.71)$$

$$f_1(i_y j_z - i_z j_y) = h_{31}(i_x j_y - i_y j_x), \quad (3.72)$$

$$f_1(i_z j_x - i_x j_z) = h_{11}(i_x j_y - i_y j_x). \quad (3.73)$$

Ker velja $f_1 > 0$ in $f_2 > 0$ lahko v primeru, da je kateri izmed faktorjev $h_{11}, h_{12}, h_{13}, h_{21}, h_{22}, h_{23}$ enak 0, sklepamo, da so pripadajoči faktorji $i_x, i_y, i_z, j_x, j_y, j_z$ tudi enaki 0.

Brez premika po kateri izmed osi

Možnost, da je kateri izmed faktorjev c_x , c_y ali c_z enak 0 opazimo iz epipolarne omejitve v enačbi (3.45). V primeru, da je vsaj eden izmed faktorjev e_{Lx} in e_{Ly} enak 0, je tudi pripadajoči premik po osi x in y enak 0. V kolikor pa je eden izmed njiju neskončen ali nedefiniran, pa velja, da je premik po osi z enak 0. V takem primeru je onemogočena uporaba enačb (3.49) in (3.50). Lahko pa izkoristimo eno izmed enakosti v enačbi (3.46) in tako pridobimo novo enačbo, ki nadomesti enačbi (3.49) in (3.50). Sedaj faktor f_1 ni prisoten v sistemu enačb, zato ga moramo izračunati analitično.

Projekcija premika kamere na vektor \vec{i} enaka 0

V primeru, da je e_{Rx} enak 0, velja, da je $\mathbf{c}^T \mathbf{i}$ tudi enaka 0. Enačbe (3.30), (3.31) in (3.32) se sedaj poenostavijo na:

$$f_1 di_x = h_{11} h_{33}, \quad (3.74)$$

$$f_1 di_y = h_{12} h_{33}, \quad (3.75)$$

$$di_z = h_{13} h_{33}. \quad (3.76)$$

V kolikor je kateri izmed faktorjev h_{11} , h_{12} ali h_{13} enak 0 so tudi pripadajoči i_x , i_y ali i_z enaki 0. Ker velja $\mathbf{c}^T \mathbf{i} = 0$, lahko uporabimo enakosti iz enačbe (3.45) ter pridobimo enačbo

$$f_1 e_x i_x + f_1 e_y i_y + i_z = 0. \quad (3.77)$$

Z uporabo enačb (3.74), (3.75), (3.76), (3.77) ter preostankom enačb je sistem mogoče analitično rešiti.

Projekcija premika kamere na vektor \vec{j} enaka 0

Podobno kot v prejšnjem primeru, velja, da je $\mathbf{c}^T \mathbf{j}$ enaka 0, če je e_{Ry} enak 0. Namesto enačb (3.30), (3.31), (3.32) in enakosti $\mathbf{c}^T \mathbf{i} = 0$ sedaj sistem sestavimo z uporabo enačb (3.33), (3.34), (3.35) in enakosti $\mathbf{c}^T \mathbf{j} = 0$ ter ga analitično rešimo.

Projekcija premika kamere na vektor \vec{k} enaka 0

Tudi v primeru zaznavanja, da je $\mathbf{c}^T \mathbf{k}$ enaka 0, uporabimo epipolarni točki e_{Rx} in e_{Ry} , le da mora biti v tem primeru vsaj ena izmed njiju v neskončnosti. Postopek reševanja je podoben kot v prejšnjih dveh primerih, le da sedaj uporabimo enačbe (3.36), (3.37), (3.38) in enakost $\mathbf{c}^T \mathbf{k} = 0$.

Faktor $n_x 1$ enak 0

Da je faktor $n_x 1$ enak 0 ugotovimo tako, ker velja $h_2^{11} h_1^{21} - h_1^{11} h_2^{21} = 0$, $h_2^{21} h_1^{31} - h_1^{21} h_2^{31} = 0$, $h_1^{11} h_2^{31} - h_2^{11} h_1^{31} = 0$ in $\mathbf{c}^T \mathbf{i} \neq 0$. V tem primeru uporabimo poenostavljene enačbe (3.30), (3.33) in (3.36) ter iz njih izločimo faktorje f_1 , $\mathbf{c}^T \mathbf{i}$, n_z in d . Pridobimo lahko naslednje tri enačbe

$$j_x h_1^{11} - i_x h_1^{21} = 0, \quad (3.78)$$

$$f_2 k_x h_1^{11} - i_x h_1^{31} = 0, \quad (3.79)$$

$$f_2 k_x h_1^{21} - j_x h_1^{31} = 0. \quad (3.80)$$

Zgoraj navedeni pogoji onemogočajo uporabo enačb (3.61) in (3.64), zato ju nadomestimo z enačbami (3.78) in (3.79) v primeru, da h_1^{11} ni enak 0, (3.78) in (3.80) v primeru, da h_1^{21} ni enak 0 ali (3.79) in (3.80) v primeru, da h_1^{31} ni enak 0.

Faktor $n_y 1$ enak 0

Podobno kot v prejšnjem primeru, lahko tudi za faktor $n_y 1$ ugotovimo ali je 0. Veljati mora $h_2^{12} h_1^{22} - h_1^{12} h_2^{22} = 0$, $h_2^{22} h_1^{32} - h_1^{22} h_2^{32} = 0$, $h_1^{12} h_2^{32} - h_2^{12} h_1^{32} = 0$ in $\mathbf{c}^T \mathbf{i} \neq 0$. Navedeni pogoji nam onemogočajo uporabo enačb (3.62) in (3.65), zato ponovimo enak postopek kot pri prejšnjem primeru, le da namesto enačb (3.30), (3.33) in (3.36) uporabimo enačbe (3.31), (3.34) in (3.37), nato pa sestavimo nadomestka enačbam (3.62) in (3.65).

Faktor n_z1 enak 0

Če velja $h_2^{13}h_1^{23} - h_1^{13}h_2^{23} = 0$, $h_1^{23} - h_2^{23} = 0$, $h_2^{13} - h_1^{13} = 0$ in $\mathbf{c}^T \mathbf{i} \neq 0$, velja tudi, da je faktor n_z1 enak 0. Ponovno lahko uporabimo enak postopek kot pri prejšnjem primeru, tokrat enačbe (3.30), (3.33) in (3.36) nadomestijo enačbe (3.32), (3.35) in (3.38), rezultat pa nadomesti enačbo (3.63).

Klasični stereo sistem

Klasični stereo sistem sestavljata dve kameri, kateri imata enako orientacijo, ter sta zamaknjeni samo po osi x . Iz pravil epipolarne geometrije sledi, da sta e_{Ly} in e_{Ry} enaka 0, e_{Lx} in e_{Rx} pa v neskončnosti.

3.2.4 Iterativni algoritem za dekompozicijo dveh homografij

Iterativni algoritem za dekompozicijo dveh homografij temelji na Gauss-Newtonovi metodi. Jakobijevo matriko sestavimo z uporabo izreka 3.3 ter omejitev (3.39), (3.40), (3.41), (3.42), (3.43) in (3.44). Tako bomo ocenili faktorje $i_x, i_y, i_z, j_x, j_y, j_z, f_1$ in f_2 . Sedaj za začetne vrednosti potrebujemo le dovolj dober približek. Tega lahko ocenimo s Hartleyevim normaliziranim osem točkovnim algoritmom [4] ali katero izmed metod dekompozicije homografije. Gradient v naključno določenem začetnem približku ima eksperimentalno ocenjeno 40 odstotno verjetnost, da kaže v smeri globalnega minimuma. Če predpostavimo, da začetni približek, ki ga ocenimo z eno izmed prej navedenih metod, spada med preostalih 60 odstotkov, je bolje izbrati nekaj naključnih začetnih vrednosti ter preveriti koliko se globalnemu minimumu približajo. Uporabimo lahko enačbo (2.17) za izračun števila iteracij pri algoritmu RANSAC [5] in izračunamo, da ima 20 naključno izbranih vrednosti le okoli stotinko odstotka, da ne doseže globalnega minimuma. Sedaj moramo oceniti katera izmed naključnih začetnih vrednosti se najbolj približa globalnemu minimumu. To naredimo tako, da izračunamo še preostale faktorje in z uporabo enačbah iz definicije homografije izračunamo

novi homografiji. Novi homografiji nato primerjamo s starimi, kjer vzamemo kombinacijo faktorjev, ki se starim najboljše prilega. Postopek je povzet v Algoritmu 3

Algoritem 3 Iterativna dekompozicija dveh homografij.

- 1: **ponavlja**
 - 2: Izberi naključni začetni približek
 - 3: Z Gauss-Newtonovo metodo oceni faktorje $i_x, i_y, i_z, j_x, j_y, j_z, k_x, k_y, k_z, f_1$ in f_2 .
 - 4: Analitično izračunaj faktorje $n_{x1}, n_{y1}, n_{z1}, d_1, n_{x2}, n_{y2}, n_{z2}$ in d_2 .
 - 5: Izračunaj novi homografiji ter ju primerjaj s starimi.
 - 6: **dokler** Razlika med novimi in starimi homografijami ni dovolj majhna ali preseženo določeno število iteracij
 - 7: Kot rezultat vzamemo faktorje, kateri se danima homografijama najboljše približata
-

Konvergenca algoritma

Eksperimentalno je ugotovljeno, da algoritem lahko konvergira k enemu izmed osmih globalnih minimumov. Tej predstavljajo naslednje možnosti:

- vsi faktorji so pravilni,
- $i_x, i_y, i_z, j_x, j_y, j_z$ preslikani preko izhodišča, eden ali oba izmed faktorjev f_1 in f_2 negativna,
- $i_x, i_y, i_z, j_x, j_y, j_z$ preslikani preko z osi, eden ali oba izmed faktorjev f_1 in f_2 negativna,
- $i_x, i_y, i_z, j_x, j_y, j_z$ preslikani preko izhodišča ter z osi, eden ali oba izmed faktorjev f_1 in f_2 negativna,
- zgoraj navedene možnosti rotirane okoli vektorja (c_x, c_y, c_z) za π .

Vidimo, da ne glede h kateremu izmed globalnih minimumov algoritem konvergira, preostale lahko trivialno izračunamo. Pogoji, da morata biti faktorja f_1 in f_2 pozitivna, izloči šest od osmih naštetih globalnih minimumov. Pravilno rotacijo okoli vektorja (c_x, c_y, c_z) pa lahko nato določimo s pogojem, da morajo normale ploskev ravnin vedno gledati proti normali ravnine kamere.

Izračun preostalih spremenljivk

Pri pravilno ocenjeni rotacijski matriki ter goriščnima razdaljama f_1 in f_2 je izračun preostalih faktorjev trivialen. Najprej izračunamo premik druge kamere. To storimo z uporabo enačbe (3.45) tako, da izberemo poljubni faktor t_1 . Vrednosti nato vstavimo v enačbe homografije, ter zračunamo še normali na ravnini ter njuno oddaljenost od izhodišča.

3.3 Rekonstrukcija scene

Postopek za iterativno dekompozicijo homografij Algoritem 3 hkrati že izračuna lego homogene regije v prostoru, kot tudi vseh točk na njej. Za rekonstrukcijo scene je potrebno med pari regij določiti take, ki ustrezajo globalnemu položaju kamere in jo nato oceniti. Potrebujemo robustno metodo, ki bo nepravilne pare izločila, kot je naprimer RANSAC [6]. Za izgradnjo modela potrebujemo dve homografiji, torej potrebujemo dve ujemanji regij. Preostala ujemanja nato lahko preverimo v kombinaciji z enim od ujemanj s katerim smo model zgradili. Iz globalne ocene položaja kamere ter potrjenih parov regij lahko sedaj trivialno sestavimo globinsko sliko tako, da poiščemo sečišče ravnine z vektorjem, ki poteka med goriščno točko ter točko na slikovni ravnini.

Poglavje 4

Rezultati

Začetne faze predstavljenega postopka je mogoče testirati neodvisno od preostalnih. Na točnost homogenih regij, ki jih vrne algoritem MSER [7], ne moremo vplivati, zato za testiranje ocene homografije uporabimo analitično določene projekcije točk. Prav tako je tudi testiranje uspešnosti dekompozicije odvisno od natančnosti homografij zato tudi v tem primeru uporabimo analitično določene projekcije točk. Za oceno uspešnosti rekonstrukcije pa moramo uporabiti celoten postopek za pridobivanje strukture iz gibanja. Testiramo lahko na dva načina in sicer iz realnih slik, katerim smo z natančnejšimi postopki določili rekonstruiran model, ali z uporabo upodobljenih slik računalniških 3D modelov. Uporaba upodobljenih slik omogoča večjo svobodo pri izbiri scene, kot tudi natančje določene lege točk scene.

4.1 Ocena homografije

Za predstavljeno metodo za oceno homografije z uporabo normaliziranih regij ne potrebujemo regij samih, pač pa nekaj točk, ki jih regija vsebuje. Točke na regiji dobimo z uporabo enačbe (3.21), kjer vrednosti parametrov n_x , n_y , n_z in d določimo naključno. Za projekcijo točk na slikovni ravnini uporabimo enačbo kamere z luknjičasto odprtino (2.5), kjer intrinzične in ekstrinzične parametre določimo naključno, ter pravilno homografijo izračunamo

število točk	Normalizacija regij			DLT		
	$\sigma = 1$	$\sigma = 2$	$\sigma = 3$	$\sigma = 1$	$\sigma = 2$	$\sigma = 3$
10	0,180px	0,350px	0,43px	0,160px	0,220px	0,250px
20	0,090px	0,019px	0,16px	0,050px	0,012px	0,095px
30	0,047px	0,086px	0,12px	0,036px	0,031px	0,087px
40	0,085px	0,099px	0,12px	0,065px	0,035px	0,051px
50	0,590px	0,085px	0,16px	0,033px	0,033px	0,093px

Tabela 4.1: Primerjava ocene homografije z uporabo normaliziranih regij in DLT metode za različne vrednosti Gaussovega šuma σ .

z uporabo enačb (2.15) in (2.16). Pozicijam projeciranih točk nato dodamo Gaussov šum ter ocenimo homografijo. Razliko med pravilno ter ocenjeno homografijo dobimo tako, da ponovno izberemo nekaj naključnih točk na regiji, izračunamo projekcijo točk na prvo slikovno ravnino, nato točke preslikamo z obema homografijama ter izmerimo povprečno razdaljo med točkami. Metodo smo preverili proti klasični metodi DLT za oceno homografije na različnem številu naključno določenih korespondenc ter napako ocenili s povprečno razdaljo pri preslikavi testnih korespondenc.

V Tabeli 4.1 so prikazani rezultati primerjave ocene homografije z metodo normalizacije regij ter metodo DLT. Metoda normalizacije regij uporablja metodo DLT kot osnovo za določitev vmesne preslikave, zato pri enakih korespondencah ne more doseči boljšega rezultata. Zaradi preslikovanja točk v vsaki iteraciji se dodatno pojavlja tudi računsko napaka. Normalizacija regij lahko boljši rezultat doseže pri preslikavi obrisa predmeta, na kateremu točnih korespondenc ni mogoče najti, oblika predmeta pa ostaja podobna.

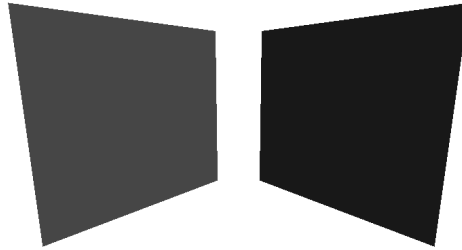
f_1/f_2	Iterativna dekompozicija			Hartleyev algoritem [4]			Zhangova dekompozicija [15]		
	$\sigma = 1$	$\sigma = 2$	$\sigma = 3$	$\sigma = 1$	$\sigma = 2$	$\sigma = 3$	$\sigma = 1$	$\sigma = 2$	$\sigma = 3$
1	0,297	0,257	0,283	0,665	0,723	0,645	1,251	1,218	1,194
2	0,273	0,238	0,281	0,828	0,793	0,843	2,198	2,104	2,098
3	0,284	0,244	0,290	0,834	0,828	0,762	3,258	3,438	3,298
4	0,268	0,306	0,342	0,874	0,807	0,855	4,498	4,540	4,711
5	0,306	0,301	0,321	0,854	0,830	0,916	6,149	5,890	6,208

Tabela 4.2: Primerjava ocene položaja kamere med iterativno dekompozicijo, Hartleyevim algoritmom in Zhangovo dekompozicijo za različne faktorje goriščnih razdalj f_1/f_2 ter za različne vrednosti Gaussovega šuma σ .

4.2 Dekompozicija homografije

Pri preverjanju iterativnega algoritma za dekompozicijo homografije Algoritem 3 lahko uporabimo postopek iz testa ocene homografije. Za intrinzične parametre prve kamere uporabimo enaki naključni goriščni razdalji za x in y os, ostale parametre pa enačimo z 0. Intrinsicni parametri druge kamere se od prve razlikujejo le v goriščni razdalji. Ekstrinzične parametre druge kamere določimo naključno, prav tako normalni in oddaljenost ravnin. Nato naključno izberemo nekaj točk na obeh ravninah, njihovim pozicijam dodamo Gaussov šum ter izračunamo homografiji za pripadajoči ravnini. Osredotočili smo se na natančnost ocene pri različnem razmerju goriščnih razdalj. Rezultate pa smo primerjali z metodama Hartleyevega normaliziranega osem točkovnega algoritma [4] ter Zhangovo dekompozicijo homografije [15].

Rezultati primerjave med metodami iterativne dekompozicije, Hartleyevega algoritma [4] in Zhangove dekompozicije [15] so prikazani v Tabeli 4.2. Vrednosti prikazujejo povprečno razdaljo enotnih vektorjev rotacijske matrike ter enotnega vektorja translacije. Brez znanih intrinzičnih parametrov kamere se najbolje obnese metoda iterativne dekompozicije, saj jih deloma pridobi sama. Opazimo lahko, da je Zhangova dekompozicija [15] močno

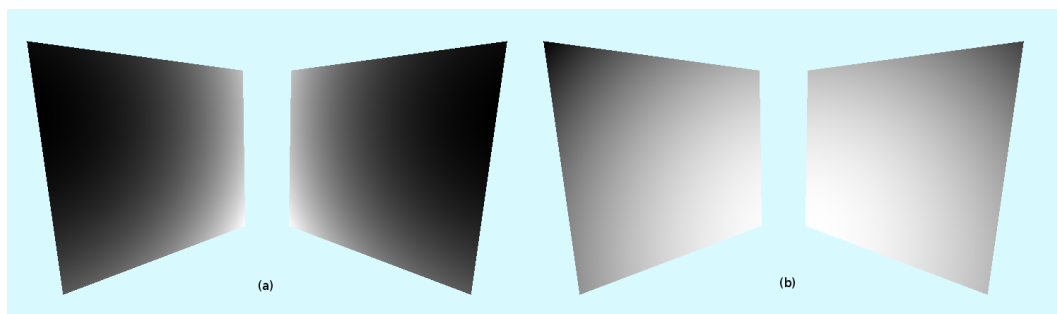


Slika 4.1: Enostavni model z dvema ravninama.

odvisna od razmerja med goriščnima razdaljama, Hartleyev algoritem [4] pa na razmerje ni občutljiv. Ocenjevanje homografije, na kateri slonita metodi dekompozicije, kot tudi sama metoda ocene fundamentalne matrike sta na Gaussov šum razmeroma neobčutljivi, saj obe šum izničita z iskanjem najmanjše kvadratne vsote.

4.3 Rekonstrukcija scene

Točnost rekonstrukcije smo preverili z uporabo upodobljenih slik 3D modelov. Za vsako upodobitev 3D modela zabeležimo parametre kamere ter ustvarimo globinsko sliko. Na upodobljenih slikah izvedemo postopek rekonstrukcije ter primerjamo globinski sliki. Zanimiva sta predvsem podatka o povprečni oddaljenosti med globinskima slikama ter odstotek rekonstruirane scene. Na Sliki 4.1 imamo primer uporabljenega enostavnega modela z dvema ravninama. V opisanem cevovodu se odstotek rekonstruirane scene določi že na prvi stopnji, saj kasnejše stopnje velikosti regij ne spreminjajo, in je zato odvisen od uporabljenega algoritma za ekstrakcijo regij. Povprečna oddaljenost med globinskima slikama pa je odvisna od natančnosti ocene homografij ter dejanske ravnosti homogene regije. Razliki med globinskima slikama vidimo na Sliki 4.2, kjer so v svetlo modro pobarvana področja brez določene globine. Za model na Sliki 4.1 je odstotek rekonstruirane scene 99,99%, povprečna razdalja pa 0,304 enote pri goriščni razdalji 0,0026 enote za vidno



Slika 4.2: Globinska slika modela (a) ter rekonstruirana globinska slika (b).

polje 90° . Največji vpliv na pravilnost rekonstrukcije sta imela intrinzična parametra goriščnih razdalj.

Poglavje 5

Sklep

Pridobivanje strukture iz več pogledov je med težjimi problemi na področju računalniškega vida. Nedoločenost položaja kamere kot tudi njenih intrinzičnih parametrov predstavlja prvo oviro pri gradnji 3D modelov. Pomagamo si s homogenimi regijami, saj so le te namreč velikokrat ravnine v prostoru. Ravno zato pa homogene regije nosijo več informacij o prostoru in pogledu kot posamezne nepovezane točke. Slabost ravnin je, da jih je težje upariti med različnimi pogledi kot lahko to storimo za točke. V ta namen vpeljemo postopek za normalizacijo ravnin, kateri sloni na osnovi projekтивne transformacije. Normalizirane ravnine so tako odporne na vse affine kot tudi projekтивne transformacije ter jih lahko predstavimo z vektorskim deskriptorjem za hitrejše primerjanje. V podpoglavju 3.2.1 izpeljemo enačbe s katerimi rešimo problem ocene položaja kamere ter določimo nekaj intrinzičnih parametrov z uporabo dveh ali več homogenih regij. Primerjava med predstavljenim Algoritmom 3 ter Hartleyevim normaliziranim osem točkovnim algoritmom [4] pokaže, da je iterativna dekompozicija homografije približno dvakrat bolj natančna pri določanju položaja kamere. Goriščni razdalji se že pri majhnem šumu precej oddaljita, zato ju je potrebno iz enačb odstraniti ali bolje stabilizirati, saj imata velik vpliv na končno rekonstrukcijo.

5.1 Nadaljnje delo

Posamezne stopnje uporabljenega cevovoda predstavljajo probleme, s katerimi se pogosto soočamo na področju računalniškega vida. Le tem lahko dodamo novo stopnjo ter tako spremenimo namembnost cevovoda. V tem pod poglavju predstavljamo možne uporabe predstavljenih algoritmov z manjšimi dopolnitvami.

5.1.1 3D rekonstrukcija

Ker algoritem razstavi homografijo regije na orientacijo in relativne položaje obeh kamer, normalo ravnine regije ter njeno relativno oddaljenost, lahko za vsako točko na regiji določimo relativno lego v prostoru. V kolikor poznamo absolutno razdaljo med kamerami ali velikost regije, lahko vse razdalje določimo absolutno.

5.1.2 Detekcija objektov

Za detekcijo objektov algoritem uporabimo tako, da pri učenju objektov sestavimo 3D model za posamezen objekt ter za vsako regijo objekta shranimo pripadajoč krožni deskriptor normalizirane regije in pripadajočo homografijo. Regije na sliki na kateri želimo objekt razpoznati nato normaliziramo in primerjamo s shranjenimi normaliziranimi regijami. Kot dodatno potrditev detekcije lahko uporabimo tudi prostorsko informacijo iz 3D modela ter tako izboljšamo rezultat.

5.1.3 Lociranje kamere

V kolikor imamo že predhodno absolutno 3D rekonstrukcijo prostora, lahko kamero tudi lociramo. Namreč poznamo faktorje f_1 , n_{x1} , n_{y1} , n_{z1} , d_1 , n_{x2} , n_{y2} , n_{z2} in d_2 , rotacijsko matriko in faktor f_2 nato ocenimo in ju uporabimo za izračun faktorja t_1 v enačbi (3.45), tako da rešitev ustreza vsem pogojem.

5.1.4 Obogatena resničnost

Pri obogateni resničnosti je najbolj zahteven problem pravilna umestitev navideznih konstruktov v prostor. Z uporabo dveh kamer ali pa s premikanjem po že rekonstruiranem prostoru, pa lahko algoritem uporabimo za določanje normal ravin ter njihovo razdaljo v prostoru in nanje nato umestimo navidezne konstrukte.

Literatura

- [1] H. Zeng, X. Deng, Z. Hu. “A new normalized method on line-based homography estimation”, *Pattern Recognition Letters*, št. 29, zv. 9, str. 1236–1244, 2008.
- [2] E. Dubrofsky, R. J. Woodham. “Combining line and point correspondences for homography estimation”, na ISVC '08: International Symposium on Visual Computing, 2008.
- [3] R. Hartley, A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, drugi izvod, 2003.
- [4] R. Hartley. “In Defense of the Eight-Point Algorithm”, *IEEE Transaction on Pattern Recognition and Machine Intelligence*, št. 19, zv. 6, str. 580–593, 1997.
- [5] M. A. Fischler, R. C. Bolles. “Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography”, *Communications of the ACM*, št. 24, zv. 6, str. 381–395, 1981.
- [6] I. E. Sutherland. “Three-dimensional data input by tablet”, *Proceedings of the IEEE*, št. 62, zv. 4, str. 453–461, 1974.
- [7] J. Matas, O. Chum, M. Urban, T. Pajdla. “Robust Wide Baseline Stereo from Maximally Stable Extremal Regions”, *Proceedings of British Machine Vision Conference*, str. 384–396, 2002.

-
- [8] D. G. Lowe. “Object recognition from local scale-invariant features”, *Proceedings of the International Conference on Computer Vision*, str. 1150–1157, 1999.
- [9] T. Formherz, M. Bichsel. “Shape from Multiple Cues: Integrating Local Brightness Information”, *Proceedings of the Fourth International Conference for Young Computer Scientists, ICYCS 95*, 1995.
- [10] P. Fua, Y. G. Leclerc. “Object-Centered Surface Reconstruction: Combining Multi-Image Stereo and Shading”, *International Journal of Computer Vision*, št. 16, zv. 1, str. 35–56, 1995.
- [11] A. P. Rockwood, J. Winget. “Three-dimensional object reconstruction from two-dimensional images”, *Computer-Aided Design*, št. 29, zv. 4, str. 279–285, 1997.
- [12] L. Zhang, S. M. Seitz. “Image-based multiresolution shape recovery by surface deformation”, *Proc. SPIE 4309, Videometrics and Optical Methods for 3D Shape Measurement*, št. 4309, str. 51–60, 2000.
- [13] H.C. Longuet-Higgins. “A computer algorithm for reconstructing a scene from two projections”, *Nature*, št. 293, str. 133–135, 1981.
- [14] O. Faugeras, F. Lustman. “Motion and structure from motion in a piecewise planar environment”, *International Journal of Pattern Recognition and Artificial Intelligence*, št. 2, zv. 3, str. 485–508, 1988.
- [15] Z. Zhang, A. R. Hanson. “3D Reconstruction based on homography mapping”, *ARPA Proceedings of the 1996 Image Understanding Workshop*, str. 1007–1012, 1996.
- [16] H. Bay, T. Tuytelaars, L. Van Gool. “SURF: Speeded Up Robust Features”, *Computer Vision – ECCV 2006*, št. 3951, str. 404–417, 2006.

-
- [17] E. Rosten, T. Drummond. “Fusing points and lines for high performance tracking”, *IEEE International Conference on Computer Vision*, št. 2, str. 1508–1511, 2005.
- [18] E. Rublee, V. Rabaud, K. Konolige, G. Bradski. “ORB: An efficient alternative to SIFT or SURF”, *IEEE International Conference on Computer Vision*, 2011.
- [19] G.N. Desouza, A. C. Kak. “Vision for mobile robot navigation: a survey”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, št. 24, zv. 2, str. 237–267, 2002.
- [20] D. Kragic, M. Björkman, H. I. Christensen, J. Eklundh. “Vision for robotic object manipulation in domestic settings”, *Robotics and Autonomous Systems*, št. 52, zv. 1, str. 85–100, 2005.
- [21] P. Jain, J. Manweiler, R. R. Choudhury. “OverLay: Practical Mobile Augmented Reality”, *MobiSys '15 Proceedings of the 13th Annual International Conference on Mobile Systems, Applications, and Services*, str. 331–344, 2015.
- [22] Å. Björck. *Numerical Methods for Least Squares Problems*. SIAM, 1996.