

Gradnja globinskih panoramskih slik s postopkom mozaičenja

Peter Peer, Franc Solina

Univerza v Ljubljani, Fakulteta za računalništvo in informatiko, Laboratorij za računalniški vid,
Tržaška 25, 1000 Ljubljana, Slovenija

E-pošta: {peter.peer, franc.solina}@fri.uni-lj.si

Povzetek. V članku 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 lepo 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 enako oddaljeni koloni levo in desno od sredine zajete slike. Epipolarne premice simetričnih parov panoram so istoležne vrstice panoram. Teža članka je predvsem v analizi zmožljivosti sistema. Sistem daje dobre rezultate pri rekonstrukciji majhnih prostorov.

Ključne besede: stereo vid, rekonstrukcija, panoramska slika, globinska slika, mozaičenje, paralaksa gibanja, epipolarna geometrija

Panoramic depth imaging with mosaicing

Extended abstract. In this article 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 (Fig. 1) 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 θ_0 , equivalent to one column of the captured image (Fig. 2). The equation for depth estimation can be easily extracted from system geometry (Eq. 1). To find the corresponding points on a stereo pair of panoramic images the epipolar geometry needs to be determined. It can be shown that the epipolar geometry is very simple if 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 columns on the left and on the right side from the captured image center column. Epipolar lines of the symmetrical pair of panoramic images are image rows. We focused mainly on the system analysis. Some results of the stereo reconstruction procedure are given in Fig. 4. An example of quality evaluation of generated depth image is presented in Fig. 5. The system performs well in the reconstruction of small indoor spaces.

Key words: stereo vision, reconstruction, panoramic image, depth image, mosaicing, motion parallax effect, epipolar geometry

1 Uvod

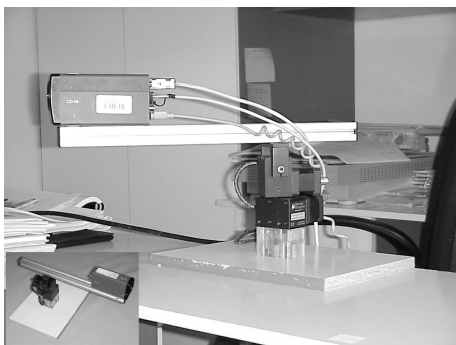
Standardna kamera ima omejen vertikalni in horizontalni zorni kot, ki sta navadno manjša od človeškega vidnega polja. Že zaradi tega si je človek vedno prizadeval, da bi ustvaril sliko, ki bi obsegala večji zorni kot, celo večjega od človeškega vidnega polja [10]. Takšne slike imenujemo panoramske slike.

Pod pojmom stereo rekonstrukcija razumemo gradnjo globinskih slik iz več slik. Globinska slika hrani informacijo o oddaljenosti do točk na sceni. Postopek stereo rekonstrukcije temelji na relacijah med točkami in linijami na sceni in slikah le-te. Če želimo dobiti linearno rešitev rekonstrukcije, lahko slike scene nastopajo v parih, trojicah ali četvericah, prav na podlagi števila slik pa imenujemo tudi omenjene relacije: epipolarna omejitve, trislikovna omejitve ali štirislikovna omejitve. Zelo zaželeno 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.

Če bi uporabili dve standardni kameri sočasno in poskušali zgraditi panoramski sliki, bi imeli probleme že pri zajemanju panoramskih slik zaradi gibanja nasprotno kamere. Zato smo se odločili, da uporabimo le eno

kamero. Naš končni cilj je uporaba sistema za samodejno navigacijo mobilnega robota v prostoru.

Na sliki 1 so prikazani strojni deli našega sistema: rotacijska roka, 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) kamera, ki gleda navzven, torej stran od središča vrtenja. Panoramske slike ustvarimo tako, da rotacijsko roko vrtimo za kot, ki ustreza premiku za eno kolono zajete slike.



Slika 1. Strojni deli sistema
Figure 1. Hardware part of our system

Izkaže se, da je epipolarna geometrija našega sistema zelo enostavna v primeru računanja rekonstrukcije na podlagi simetričnega para panoram. Simetričen par panoram dobimo, če vzamemo enako oddaljeni koloni levo in desno od sredine zajete slike.

V naslednjem poglavju bomo predstavili podobna dela ter izpostavili prispevek pričujočega članka k obravnavani tematiki. Sledil bo opis geometrije sistema, epipolarne geometrije in postopka rekonstrukcije. Poudarek članka bo na analizi zmogljivosti sistema, ki bo opisana v predzadnjem poglavju. Članek bomo zaključili s sklepi in smernicami za nadaljnje delo.

2 Sorodna dela

Panoramske slike lahko ustvarimo ali s pomočjo posebnih panoramskih kamer ali s pomočjo standardne, nepanoramske kamere, kjer delamo mozaike standardnih slik. Če želimo ustvariti prave mozaične panoramske poglede, moramo kamero seveda pomikati po neki zaključeni poti, ki je največkrat krožnica.

Eden izmed najbolj znanih komercialnih paketov za ustvarjanje mozaičnih panoramskih pogledov je QTVR (QuickTime Virtual Reality). Ta deluje na principu šivanja standardnih slik, ki so zajete s pomočjo rotirajoče se kamere [1]. Peleg s sodelavci [9] je predstavil metodo ustvarjanja mozaičnih panoramskih pogledov iz standardnih slik, zajetih s poljubno gibajočo se kamero. Podobno metodo sta predlagala tudi Szeliski in Shum [13], ki prav tako ne omejuje gibanja oziroma postopka zajemanja slik,

predpostavlja pa, da ni velike paralakse gibanja. Seveda pa omenjene metode služijo zgolj za vizualizacijo, saj ni prisotnih poskusov po rekonstrukciji scene.

Ishiguro s sodelavci [5] je predlagal metodo, ki omogoča rekonstrukcijo scene. Uporabili so standardno kamero, ki rotira na krožnici. Sceno rekonstruirajo tako, da najprej ustvarijo panoramsko sliko iz središčnih kolon zajetih slike, nato sistem prestavijo na novo lokacijo in tam zgradijo še eno panoramsko sliko. Ti dve panoramski sliki predstavljata vhod v postopek rekonstrukcije. Globino objekta so sprva ocenili na podlagi projekcije v dveh panoramskih slikah, ki sta bili zajeti na različnih lokacijah kamere na krožnici (torej sta bili zgrajeni iz različnih kolon zajetih slik), ker pa je bil njihov glavni cilj ustvariti globalni zemljevid prostora tako, da se sistem, ki je pritrjen na robota, premika po prostoru, globine niso ocenjevali na ta način.

Peleg in Ben-Ezra [7, 8] sta predstavila metodo za ustvarjanje stereo panoram. Stereo panorame pa so ustvarjene brez dejanskega računanja 3D strukture — efekt globine se ustvari le v možganih opazovalca.

Shum in Szeliski [12] sta predstavila dva postopka za ustvarjanje globinskih panoramskih slik in sicer oba postopka izrabljata standardne postopke za stereo rekonstrukcijo. Oba postopka temeljita na gibanju kamere po krožnici. Panorame sestavljata tako, da iz vsake zajete slike vzameta le eno kolono slikovnih elementov, ki doprinese k panorami. Te panorame imenujeta *večperspektivne panorame* (ang. *multiperspective panorama*). Pomembna lastnost (dveh ali več) večperspektivnih panoram je, da hranijo informacijo o paralaksu gibanja, saj so kolone le-teh zajete iz različnih perspektiv. Za samo stereo rekonstrukcijo nato uporabita tako dobljene panorame.

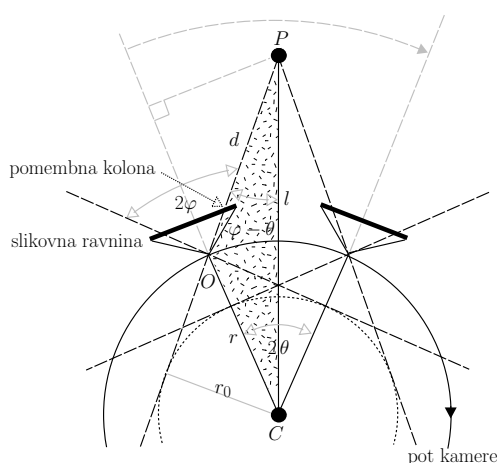
Večperspektivne panorame same po sebi niso nekaj novega [12]: so poseben primer *večperspektivnih panoram za animacijo celic* (ang. *multiperspective panoramas for cel animation*) [14], so zelo podobne slikam narejenim s postopkom *večsrediščna projekcija* (ang. *multiple-center-of-projection*) [11], postopkom *mnogotereta projekcija* (ang. *manifold projection*) [9] in postopkom *krožna projekcija* (ang. *circular projection*) [7, 8]. Princip je zelo podoben tudi postopku gradnje panorame z *linearno potisno kamero* (ang. *linear push-broom camera*) [3].

V člankih, ki sta najbližja našemu delu [5, 12], smo pogrešali predvsem dve stvari: analizo zmogljivosti sistema in uporabo standardnega postopka korelacije za iskanje korespondenčnih točk, zato smo temu posvetili bistveno pozornost. Medtem ko v [5] iščejo korespondenčne točke s sledenjem značilke tako, da sledijo značilko iz kolone, ki gradi eno panoramo, do kolone, ki gradi drugo panoramo, v [12] uporabljajo nadgrajen postopek stereo s preslikavo na ravnino (ang. *plane sweep stereo*).

3 Geometrija sistema

Začnimo to poglavje z opisom postopka nastanka stereo para panoramskih slik. Iz vsake zajete slike na kamerini krožni poti vzamemo le dve koloni, ki sta enako oddaljeni od sredine zajete slike. Kolono na desni strani zajete slike nato mozaičimo v panoramo za levo oko, kolono na levi strani zajete slike pa mozaičimo v panoramo za desno oko. Panoramo torej gradimo iz le ene kolone zajete slike. Na ta način dobimo simetričen par stereo panoramskih slik.

Iz slike 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 vrti 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 . 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. 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 kamere α zajete slike.



Slika 2. Geometrija sistema za ustvarjanje ene večperspektivne panorame. Optična os kamere je v horizontalnem položaju
Figure 2. Geometry of our system for constructing a multiperspective panorama. The optical axis of the camera is kept horizontal

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

Sistem na sliki 2 je očitno ne-središčni, saj se svetlobni žarki, ki tvorijo panoramsko sliko, ne sekajo v eni točki, pač pa predstavljajo tangente ($\varphi \neq 0$) na valj s polmerom r_0 , ki predstavlja polmer glediščnega valja. 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.

4 Epipolarna geometrija

Iskanje korespondenčnih točk na dveh slikah je zelo težek problem. V osnovi je lahko korespondenčna točka kjerkoli na drugi sliki. Zato želimo preiskovalni prostor čimbolj omejiti. Z epipolarno omejitvijo zmanjšamo preiskovalni prostor iz 2D na 1D, torej na epipolarno premico [2]. V poglavju 6.2 pa bomo pokazali, da lahko v našem sistemu učinkovito omejimo tudi preiskovalni prostor na epipolarni premici.

V tem poglavju bomo zapisali le postopek dokaza, da so epipolarne premice simetričnih parov panoram istotežne vrstice panoram. Bralec si lahko dokaz ogleda v [4].

Dokaz temelji na vrednosti r_0 iz slike 2, ki predstavlja polmer glediščnega valja in jo izračunamo po enačbi:

$$r_0 = r \cdot \sin \varphi.$$

Dokaz izpeljemo 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 lahko dokažemo lastnost epipolarnih premic simetričnih parov panoram. V prvem delu nas zanima, kako se neka točka v prostoru preslika na ravnino kamere [2], ki je tokrat ravnina velikosti $n \times 1$. V drugem delu podamo 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 zapišemo enačbo za projekcijo točk v prostoru na ravnino (valj) večperspektivne panorame. Prav iz lastnosti kotnih koordinat koordinatnega sistema panoramske slike pa v tretjem delu pokažemo 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, ki je definiran z dano točko in optičnim centrom kamere, v 3D svetovnem koordinatnem sistemu. Če optični žarek izražen v svetovnem koordinatnem sistemu projeciramo na drugo večperspektivno panoramsko sliko, dobimo epipolarno premico za dano slikovno točko na prvi večperspektivni panoramski sliki.

5 Stereo rekonstrukcija

Vrnimo se k sliki 2. Na osnovi trigonometričnih relacij moramo zapisati enačbo za oceno globine l točke P na sceni. Po sinusnem izreku velja (opazujemo pobarvan trikotnik na sliki 2):

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

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

$$l = \frac{r \cdot \sin(180^\circ - \varphi)}{\sin(\varphi - \theta)} = \frac{r \cdot \sin \varphi}{\sin(\varphi - \theta)}. \quad (1)$$

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:

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

pri čemer je W širina zajete slike v slikovnih elementih, $W_{2\varphi}$ pa širina slike med kolonoma, ki gradita simetrični par panoram, prav tako v slikovnih elementih. 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:

$$\theta = dx \cdot \frac{\theta_0}{2}, \quad (3)$$

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

Za iskanje korespondenčnih točk uporabimo postopek normalizirane korelacije, za povečanje zaupanja v dobljeno oceno globine pa postopek vzvratne korelacije, s katerim rešujemo tudi problem prekrivanja [2].

Postopek normalizirane korelacije deluje na osnovi iskanja podobnosti med deli scene na dveh slikah. Osnovna ideja postopka je: poišči tisti del scene na drugi sliki, ki je najbolj podoben danemu delu scene na prvi sliki. Postopek uporablja okenski operator, znotraj katerega se meri podobnost med deli scene s tehniko korelacije.

6 Analiza zmogljivosti sistema

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 rotacijska roka, preden zajamemo sliko, biti povsem umirjena. Čas, potreben za gradnjo panoram, je torej veliko prevelik, da

bi lahko trdili, da so sploh kakšne možnosti, da sistem deluje v realnem času.

Za naš sistem velja, da zgradimo v enem obhodu okoli vertikalne osi sistema 11 panoram (pet simetričnih parov in panoramsko sliko iz sredinskih kolon zajetih slik). Zajamemo 1501 slik, ločljivosti 160×120 slikovnih pik, razdalja $r = 30$ cm, kot premika rotacijske roke $\theta_0 = 0,2^\circ$ in za to porabimo dobrih 15 minut na računalniku PC Intel PII./350 MHz.

6.2 Omejitev iskanja na epipolarni premici

Glede na to, da je širina panoramske slike veliko večja od širine standardne slike, kar pomeni, da je iskanje korespondenčne točke potrebno izvršiti vzdolž zelo dolge epipolarne premice, želimo ta preiskovalni prostor čimbolj omejiti. Tako povečamo zaupanje v ocenjeno globino in pospešimo delovanje sistema.

Če izhajamo iz enačbe (1), lahko zapišemo dve ugotovitvi, ki nam preiskovalni prostor močno omejita:

1. 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 zapisanega 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.

Za naš sistem tako na primer velja, da je najmanjša možna ocena globine l_{\min} odvisna od kota φ . Pri $r = 30$ cm velja:

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

2. Teoretično velja, da ocena globine navzgor ni omejena, vendar iz enačbe (1) sledi, da imenovalec ne sme biti enak 0. Praktično to pomeni, da mora biti za oceno največje možne globine l_{maks} razlika $(\varphi - \theta_{\text{maks}})$ enaka vrednosti na intervalu $(0, \frac{\theta_0}{2})$. To dejstvo lahko zapišemo kot: $\theta_{\text{maks}} = n \cdot \frac{\theta_0}{2}$, pri $n = \varphi \operatorname{div} \frac{\theta_0}{2}$ in $\varphi \bmod \frac{\theta_0}{2} \neq 0$.

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 $x + \frac{2\theta_{\text{maks}}}{\theta_0} = x + n$. x je horizontalna slikovna koordinata točke na panorami za levo oko, za katero iščemo korespondenčno točko.

Tudi tukaj velja, da je največja možna ocena globine l_{maks} odvisna od kota φ . Pri $r = 30$ cm velja:

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

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

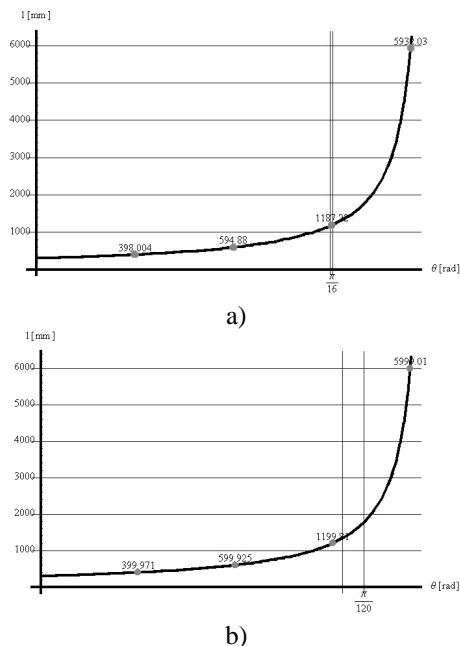
Tako na primer za naš sistem velja, da od omenjenih 1501 slikovnih elementov, kolikor znaša širina panorame, ki se nahajajo na epipolarni premici, pregledamo le $n = 149$ slikovnih elementov pri $2\varphi = 29,9625^\circ$ in le $n = 18$ slikovnih elementov pri $2\varphi = 3,6125^\circ$.

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

6.3 Pomen napake za en slikovni element pri oceni kota θ

Najprej definirajmo pojem napake za en slikovni element: slike so diskretne, zato bi radi vedeli, kakšna je napaka v oceni globine, če zgrešimo pravo korespondenčno točko za le en slikovni element. In to informacijo bi radi imeli za različne vrednosti kota φ .

Praden ilustriramo pomen napake za en slikovni element pri oceni kota θ , si oglejmo še grafa na sliki 3. Grafa prikazujeta odvisnost funkcije globine l od kota θ pri različnih vrednostih kota 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,



Slika 3. 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 grafi 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 $\frac{\theta_0}{2} = 0,1^\circ$

Figure 3. Graphs showing dependence of depth function l from the angle θ while radius $r = 30$ cm and using different values of the angle φ : a) $2\varphi = 29.9625^\circ$, b) $2\varphi = 3.6125^\circ$. To ease the comparison of the error for a pixel in estimation of the angle θ , we showed the interval of width $\frac{\theta_0}{2} = 0.1^\circ$ between the vertical lines around the third point

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 (enačba (3)) veliko manjša pri večjem kotu φ , saj prav premik za kot θ_0 opisuje premik za en slikovni element v panoramski sliki.

Zaradi diskretizacije osi θ (slika 3) 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 $\varphi \operatorname{div} \frac{\theta_0}{2} = 149$ različnih ocen globine, za kot $2\varphi = 3,6125^\circ$ pa velja, da lahko izračunamo največ le 18 različnih ocen globine. S tem smo lepo opisali negativno lastnost majhnega kota φ .

Sedaj ilustrirajmo pomen napake za en slikovni element pri oceni kota θ : Zanima nas, kakšna je ta napaka za kot θ , ki je na začetku definicijskega 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 $\theta_1 = \frac{\varphi}{4}$

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

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

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

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

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

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

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

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

Tabela 1. Ovrednotenje napake za en slikovni element pri oceni kota θ (enačba 3) in posledično po enačbi (1) 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 $0,2^\circ$

Table 1. The meaning of the error for a pixel in estimation of the angle θ , where $r = 30$ cm and $\theta_0 = 0.2^\circ$ (eqs. (1) and (3))

in $\theta_2 = \frac{7\varphi}{8}$. Zanima pa nas seveda tudi narava napake za različne 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 1 podajajo ovrednotenje napake v oceni globine l za različne vrednosti kotov θ in φ .

Iz rezultatov v tabeli 1 lahko vidimo, da je pri majhnem kotu φ napaka veliko večja kot pri večjem kotu φ . Opazimo lahko tudi, da se vrednost napake veča, ko se vrednost kota θ približuje vrednosti kota φ . To velja glede na vrednost kota 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. Namreč, 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, da je hitrost rekonstrukcije v obratnem sorazmerju z natančnostjo le-te.

S spreminjanjem parametrov θ_0 in r se seveda spreminja tudi velikost napake:

- Z večanjem ločljivosti zajete slike manjšamo kot θ_0 , torej manjšamo kot zasuka kamere med dvema zaporednima slikama, ki prispevata k panorami. Za približno 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*).
- 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.

6.4 Določitev največje globine, v katere oceno še zaupamo

V poglavju 6.2 smo ocenili najmanjšo možno globino l_{min} in največjo možno globino l_{maks} , nismo pa zapisali, 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} :

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

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

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

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

Dobljeni rezultati potrjujejo že zapisane ugotovitve (poglavje 6.3), zapišemo pa lahko še dve dodatni ugotovitvi:

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

Tabela 2. 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
 Table 2. The meaning of error (Δl) for one pixel in estimation of the angle θ for the minimal possible depth estimation l_{\min} and the maximal possible depth estimation l_{\max} regarding the angle φ

1. Vrednost napake Δl_{\max} je nesprejemljivo visoka in to neglede na kot φ . Torej 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.
2. Kot φ je vedno odvisen od horizontalnega zornega kota α (enačba (2)), 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.

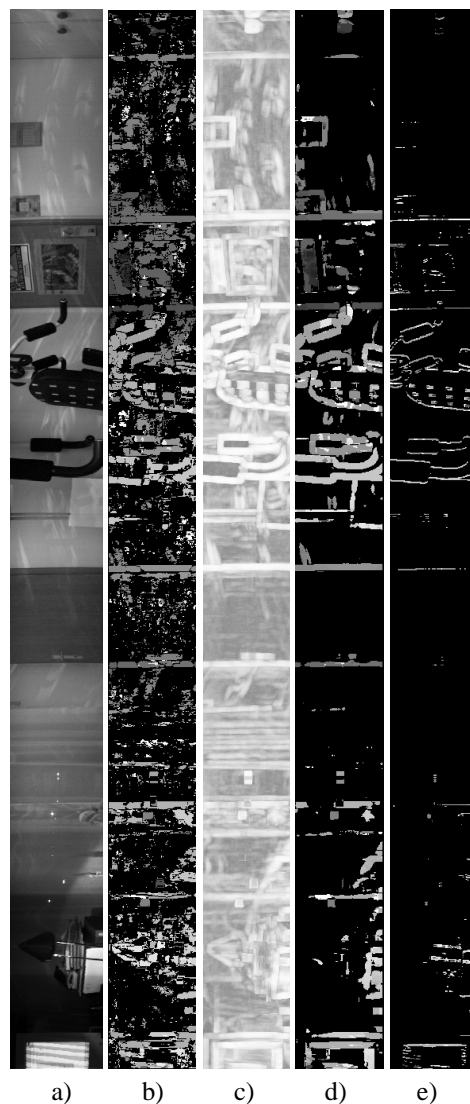
6.5 Rezultati

Na sliki 4 so prikazani primeri rezultatov našega sistema (priključenega na računalnik PC Intel PII./350 MHz). V primeru b) smo gradili gsto globinsko sliko, kar pomeni, da smo poskušali za vsako točko na panorami za levo oko najti ustrezno korespondenčno točko na panorami za desno oko. S črno barvo so podane točke na sceni, ki nimajo pripisane globine, sicer pa velja, da je bližje tista točka, ki je svetlejša.

V primeru d) uporabimo informacijo o zaupanju v ocenjeno globino (primer c)), ki jo dobimo na podlagi ocene postopka normalizirane korelacije. Tako iz goste globinske slike izločimo vse tiste ocene globin, ki nimajo pripisane dovolj velike ocene zaupanja.

V primeru e) pa gradimo redko globinsko sliko tako, da iščemo korespondenčne točke le za značilke na vhodnih slikah. Značilke lahko na primer predstavljajo vertikalni robovi na sceni, ki jih hitro dobimo, če panoramo filtriramo s Sobelovim filtrom za iskanje vertikalnih robov [2, 5]. Če uporabimo manjši kot $2\varphi = 3.6125^\circ$, so časi rekonstrukcije manjši tudi do 8-krat.

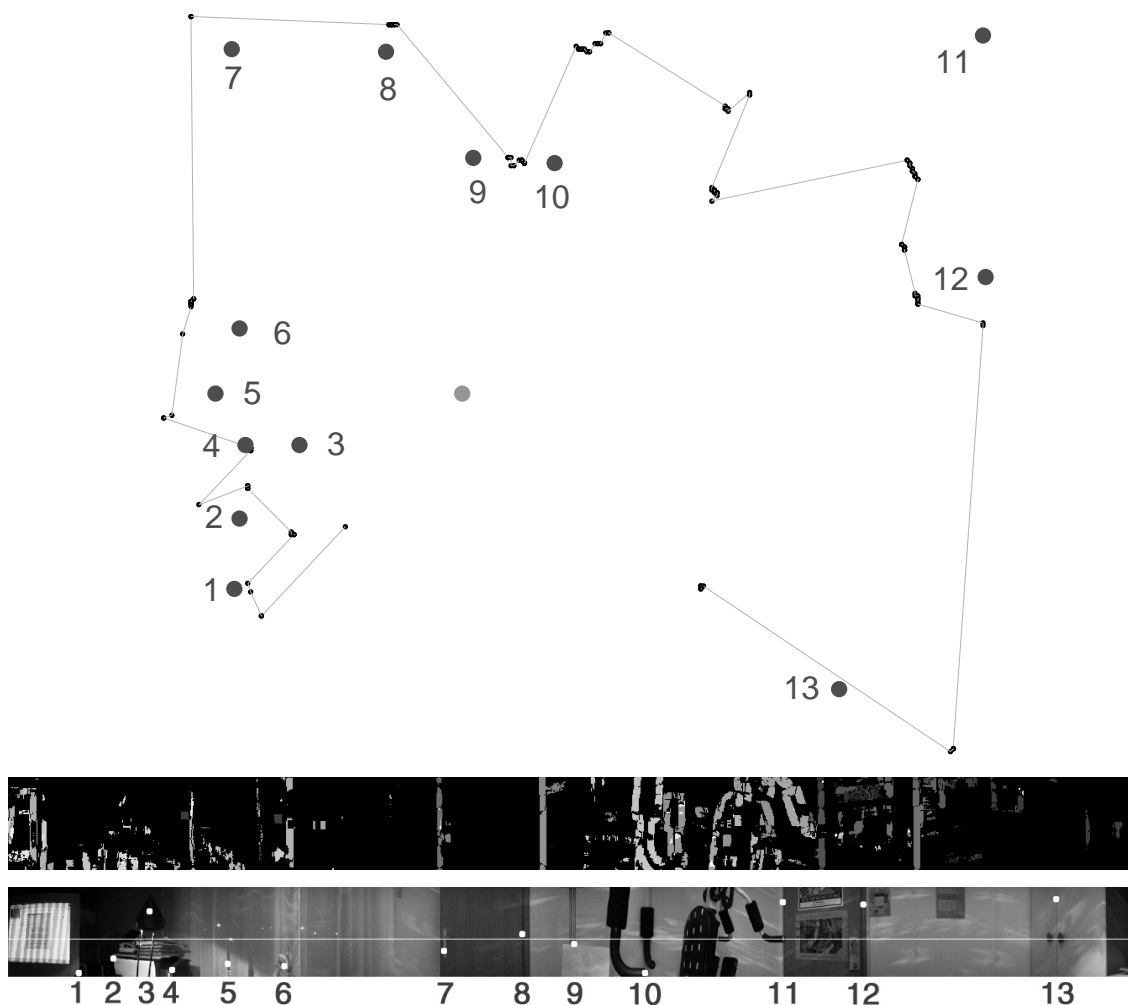
Vsi predstavljeni rezultati so bili ustvarjeni pod enakimi pogoji: velikost korelacijskega okna je bila $(2n+1) \times (2n+1)$, $n = 4$, korespondenčno točko pa smo iskali le na ustreznem delu vrstice panoramske slike, ki je bila določena z epipolarno geometrijo.



Slika 4. Primer rezultatov za kot $2\varphi = 29,9625^\circ$, ko delamo rekonstrukcijo za levo oko: a) panorama za levo oko, b) gsta globinska slika / uporabimo vzvratno korelacijo / čas rekonstrukcije: 6 ur, 42 min., 20 sek., c) informacija o zaupanju v ocenjeno globino, d) gsta globinska slika po uteževanju / ne uporabimo vzvratne korelacije / čas rekonstrukcije: 3 ur, 21 min., 56 sek., e) redka globinska slika / ne uporabimo vzvratne korelacije / čas rekonstrukcije: 38 sekund

Figure 4. Some results of stereo reconstruction when creating the depth image for the left eye while angle $2\varphi = 29.9625^\circ$: a) left eye panorama, b) dense depth image / using back-correlation / reconstruction time: 6 hours, 42 min., 20 sec., c) confidence of estimated depth, d) dense depth image after weighting / without back-correlation / reconstruction time: 3 hours, 21 min., 56 sec., e) sparse depth image / without back-correlation / reconstruction time: 38 seconds

Seveda pa je na podlagi rezultatov s slike 4 težko oceniti kakovost dobljenih globinskih slik. Zato naredimo primer rekonstrukcije prostora na podlagi ustvarjene globinske slike. Tako bomo lahko ocenili kakovost dobljene globinske slike in posledično kakovost samega sistema. Rezultat postopka (3D) rekonstrukcije je tloris scene. Za sliko 5 veljajo naslednje lastnosti:



Slika 5. Na zgornji sliki je tloris rekonstruirane scene na podlagi 68. vrstice globinske slike, ki je predstavljena na srednji sliki. Uporabili smo vzvratno korelacijo in uteževanje, kot 2φ pa je znašal $29,9625^\circ$. Za orientacijo so na spodnji sliki podane oznake od 1 do 13, ki predstavljajo (ročno) izmerjene dejanske globine značilk na sceni in rekonstruirana vrstica. Značilke na sceni, ki smo jih označili z velikimi točkami, niso nujno vidne v tej vrstici

Figure 5. On top is a ground-plan showing the results of the reconstruction process based on the 68th row of the depth image. We used back-correlation and weighting for angle $2\varphi = 29.9625^\circ$. The corresponding depth image is shown on the middle picture. For orientation, the reconstructed row and the features on the scene for which we measured the actual depth by hand are shown on the bottom picture. The features on the scene marked with big dots and associated numbers are not necessarily visible in this row

- Z velikimi točkami so označene značilke na sceni, za katere smo (ročno) izmerili dejanske globine.
- Z veliko točko blizu sredine rekonstrukcije je označeno središče sistema.
- Z majhnimi točkami so označene rekonstruirane točke na sceni.
- Linije med majhnimi točkami podajajo povezave med zaporedno rekonstruiranimi točkami.

Na sliki 5 delamo rekonstrukcijo na podlagi 68. vrstice globinske slike. Uporabili smo vzvratno korelacijo in uteževanje, kot 2φ pa je znašal $29,9625^\circ$. Za sliko 5 velja, da so majhne 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 velikimi točkami, niso nujno vidne v tej vrstici.

7 Sklepi in nadaljnje delo

Naredili smo analizo sistema ter s tem razkrili naravo sistema. Zapišimo 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 učinkovito omejimo preiskovalni prostor na epipolarni premici; ker je število možnih ocen globine premosorazmerno z vrednostjo kota φ , krivulja funkcije globine l pa vse bolj strmo narašča, se zaupanje v ocenjeno globino spreminja: večja je strmina krivulje

funkcije l , manjše je zaupanje v ocenjeno globino; gostih globinskih slik ni smiselno graditi, že zaradi samega časa, potrebnega za rekonstrukcijo.

Bistvena sklepa pa bi lahko zapisali tako:

1. Sistem je uporaben v primeru rekonstrukcije majhnih prostorov.
2. Glede na podane čase rekonstrukcije, lahko sklepamo, da lahko dosežemo delovanje rekonstrukcije v realnem času, če bi delali z 8-bitnimi sivinskimi panoramami manjše ločljivosti, gradili redko globinsko sliko le dela panorame in/ali enostavno uporabili hitrejši računalnik. Takšen rezultat bi lahko uporabili pri navigaciji robota v prostoru [5].

Pereč je seveda problem postopka gradnje panoram, ki ne deluje v realnem času, vendar pa lahko dosežemo tudi skrajšanje tega časa: Namesto, da panorame gradimo iz samo ene kolone zajete slike, bi lahko uporabili širši trak in tako zelo pospešili gradnjo panoram [10]. S tem se znotraj traku spreminja kot φ , zato moramo odgovoriti na vprašanje: "Kako takšen način gradnje vpliva na proces rekonstrukcije?"

Naše nadaljnje razmišljanje je usmerjeno predvsem v razvoj aplikacije za samodejno navigacijo mobilnega robota v prostoru.

8 Literatura

- [1] S. Chen, Quicktime VR — an image-based approach to virtual environment navigation, *ACM SIGGRAPH*, pp. 29–38, Los Angeles, USA, 1995.
- [2] O. Faugeras, *Three-Dimensional Computer Vision: A Geometric Viewpoint*, MIT Press, Cambridge, Massachusetts, London, England, 1993.
- [3] R. Gupta, R. I. Hartley, Linear pushbroom cameras, *IEEE Trans. PAMI*, 19(9):963–975, September 1997.
- [4] F. Huang, T. Pajdla, Epipolar geometry in concentric panoramas, *Technical Report CTU-CMP-2000-07*, Czech Technical University, 2000.
- [5] H. Ishiguro, M. Yamamoto, S. Tsuji, Omni-directional stereo, *IEEE Trans. PAMI*, 14(2):257–262, February 1992.
- [6] P. Peer, F. Solina, Panoramic Depth Imaging: Single Standard Camera Approach, *International Journal of Computer Vision*, sprejeto (v oktobru 2001) za objavo.
- [7] S. Peleg, M. Ben-Ezra, Stereo panorama with a single camera, *IEEE Conf. CVPR*, pp. 395–401, Fort Collins, USA, 1999.
- [8] S. Peleg, Y. Pritch, M. Ben-Ezra, Cameras for stereo panoramic imaging, *IEEE Conf. CVPR*, pp. 208–214, Hilton Head Island, USA, 2000.
- [9] S. Peleg, B. Rousso, A. Rav-Acha, A. Zomet, Mosaicing on adaptive manifolds, *IEEE Trans. PAMI*, 22(10):1144–1154, October 2000.
- [10] B. Pihavec, F. Solina, User interface for video observation over the internet, *Journal of Network and Computer Applications*, 21:219–237, 1998.
- [11] P. Rademacher, G. Bishop, Multiple-center-of-projection images, *Computer Graphics (ACM SIGGRAPH)*, pp. 199–206, Orlando, USA, 1998.
- [12] H. Y. Shum, R. Szeliski, Stereo reconstruction from multiperspective panoramas, *IEEE ICCV*, pp. 14–21, Kerkyra, Greece, 1999.
- [13] R. Szeliski, H. Y. Shum, Creating full view panoramic image mosaics and texture-mapped models, *Computer Graphics (ACM SIGGRAPH)*, pp. 251–258, Los Angeles, USA, 1997.
- [14] D. Wood, A. Finkelstein, J. Hughes, C. Thayer, D. Salesin, Multiperspective panoramas for cel animation, *Computer Graphics (ACM SIGGRAPH)*, pp. 243–250, Los Angeles, USA, 1997.

Peter Peer je mladi raziskovalec na Fakulteti za računalništvo in informatiko, Univerze v Ljubljani. Dodiplomski študij je zaključil leta 1998, magistrski študij pa leta 2001, oboje na Fakulteti za računalništvo in informatiko. Raziskovalno se ukvarja z računalniškim vidom.

je redni profesor na Fakulteti za računalništvo in informatiko, Univerze v Ljubljani in vodja Laboratorija za računalniški vid na isti fakulteti. Diplomiral in magistriral je v letih 1979 in 1982 na Fakulteti za elektrotehniko, Univerze v Ljubljani. Doktoriral je leta 1987 na University of Pennsylvania, ZDA. Raziskovalno se ukvarja z računalniškim vidom.