

# Ustvarjanje panoramskih globinskih slik s standardno kamero

Peter Peer, Franc Solina  
Univerza v Ljubljani, Fakulteta za računalništvo in informatiko  
Tržaška 25, 1000 Ljubljana, Slovenija  
{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. Teža članka je predvsem v analizi zmogljivosti sistema. Sistem daje dobre rezultate pri rekonstrukciji majhnih prostorov.

## 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 [8]. Takšne slike imenujemo panoramske slike.

Pod pojmom stereo rekonstrukcija razumemo gradnjo globinskih slik iz več slik. Postopek stereo rekonstrukcije temelji na relacijah med točkami in linijami na sceni in slikah le-te. Zelo zaželjena lastnost slik je, da so te točke ali linije dobro vidne na vseh slikah neke scene. In takšno lastnost nam v največji meri zagotavljajo panoramske kamere. To dejstvo je predstavljalo osnovno motivacijo pri gradnji našega sistema za gradnjo globinskih slik.

Na sliki 1 so na manjši sliki 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.

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.

## 2 Sorodna dela

Panoramske slike lahko ustvarimo ali s pomočjo posebnih panoramskih kamer ali s pomočjo standardne, ne-panoramske kamere, kjer delamo mozaike standardnih slik.

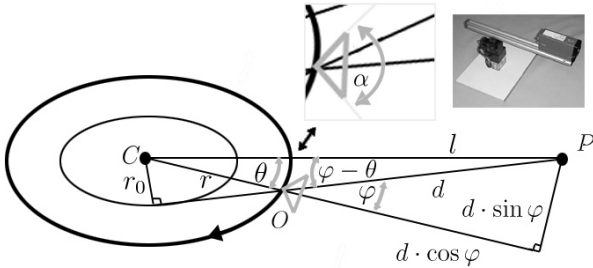
Shum in Szeliski [10] sta predstavila dva postopka za ustvarjanje mozaičnih globinskih panoramskih slik, ki temeljita na gibanju kamere po krožnici. Panorame sestavljata tako, da iz vsake zajete slike vzameta le eno kolono slikovnih elementov, ki doprinese k panorami. Te panorame imenujeta *večperspektivne 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.

Večperspektivne panorame same po sebi niso nekaj novega [10]: so poseben primer *večperspektivnih panoram za animacijo celic* (ang. *multiperspective panoramas for cel animation*) [11], so zelo podobne slikam narejenim s postopkom *večsrediščna projekcija* (ang. *multiple-center-of-projection*) [9], postopkom *mnogotera projekcija* (ang. *manifold projection*) [7] in postopkom *krožna projekcija* (ang. *circular projection*) [5, 6]. Princip je zelo podoben tudi postopku gradnje panorame z *linearno potisno kamero* (ang. *linear pushbroom camera*) [2].

V člankih, ki sta najbližja našemu delu [4, 10], 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 [4] iščejo korespondenčne točke s sledenjem značilke tako, da sledijo značilko iz kolone, ki gradi eno panoramo, do kolone, ki gradi drugo panoramo, v [10] uporabljajo nadgrajen postopek stereo s preslikavo na ravnino (ang. *plain sweep stereo*).

### 3 Geometrija sistema

Iz slike 1 je razvidna splošna geometrija našega sistema za ustvarjanje večperspektivnih panoramskih slik, ki služijo kot vhod za gradnjo globinskih panoramskih slik. Točka  $C$  predstavlja središče vrtenja, okoli katerega se vrtila kamera na oddaljenosti  $r$ , ki predstavlja polmer krožnice (pot) po kateri potuje kamera. Kamera je obrnjena navzven, torej stran od središča vrtenja. Optični center kamere se nahaja v točki  $O$ . Izbrana kolona slikovnih elementov, ki bo doprinesla k večperspektivni panorami, vsebuje preslikavo točke  $P$  na sceni. Ta je od točke  $C$  oddaljena za razdaljo  $l$ , od točke  $O$  pa za razdaljo  $d$ . S  $\theta$  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  $P$ .  $\theta$  v panoramski sliki podaja horizontalno os, ki opisuje pot gibanja kamere. S  $\varphi$  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  $\alpha$  zajete slike.



Slika 1: Geometrija sistema za ustvarjanje ene večperspektivne panorame. Na manjši sliki so prikazani strojni deli sistema.

### 4 Epipolarna geometrija

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. Zaradi pomanjkanja prostora bomo v tem poglavju zapisali le še trditev, da so epipolarne premice simetričnih parov panoram istoležne vrstice panoram. Dokaz si lahko bralec ogleda v [3].

### 5 Stereo rekonstrukcija

Vrnimo se k sliki 1. Na osnovi trigonometričnih relacij moramo zapisati enačbo za oceno globine  $l$

točke  $P$  na sceni. Po sinusnem izreku velja:

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

Iz zgornje enačbe sledi, da lahko globino  $l$  ocenimo le, če poznamo tri parametre:  $r$ ,  $\varphi$  in  $\theta$ .  $r$  je dan. Kot  $\varphi$  izračunamo glede na horizontalni zorni kot kamere  $\alpha$  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 kolonom, ki gradita simetrični par panoram, prav tako v slikovnih elementih. Za izračun kota  $\theta$  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  $\theta_0$ , potem za kot  $\theta$  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 [1].

## 6 Analiza zmogljivosti sistema

### 6.1 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 čim bolj 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 omejit: 1) Če je  $\theta_0$  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. 2) Teoretično velja, da ocena globine navzgor ni omejena, vendar iz enačbe (1) sledi, da imenovalec ne sme biti enak 0. To dejstvo lahko zapišemo kot:  $\theta_{\max} = n \cdot \frac{\theta_0}{2}$ , pri  $n = \varphi \text{ div } \frac{\theta_0}{2}$  in  $\varphi \bmod \frac{\theta_0}{2} \neq 0$ . 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.

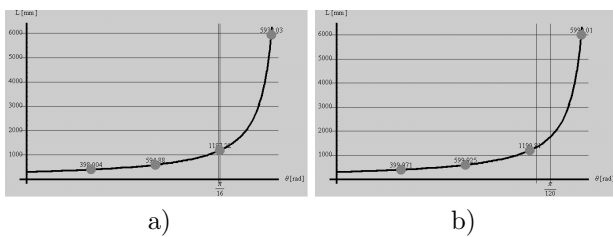
Tako na primer za naš sistem velja, da od 1501 slikovnih elementov, kolikor znaša širina panorame, ki se nahajajo na epipolarni premici, pregledamo le  $n = 149$  slikovnih elementov pri  $2\varphi = 29,9625^\circ$  in le  $n = 18$  slikovnih elementov pri  $2\varphi = 3,6125^\circ$ .

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

## 6.2 Pomen napake za en slikovni element pri oceni kota $\theta$

Grafa na sliki 2 prikazujeta odvisnost funkcije globine  $l$  od kota  $\theta$  pri različnih vrednostih kota  $\varphi$ . Iz grafov sledi, da funkcija globine  $l$  narašča počasneje pri večjem kotu  $\varphi$ . Ta lastnost zmanjša napako v oceni globine  $l$  pri večjem kotu  $\varphi$ , vendar pa to zmanjšanje napake pride še toliko bolj do izraza, če vemo, da je abscisna os  $\theta$  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  $\varphi$  veliko manjši. To pa pomeni, da je tudi napaka za en slikovni element pri oceni kota  $\theta$  (enačba (3)) veliko manjša pri večjem kotu  $\varphi$ , saj prav premik za kot  $\theta_0$  opisuje premik za en slikovni element v panoramski sliki.

Zaradi diskretizacije osi  $\theta$  (slika 2) 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  $\varphi$ : za kot  $2\varphi = 29,9625^\circ$  velja, da lahko izračunamo največ ( $\varphi \div \frac{\theta_0}{2} =$ ) 149 ocen globine, za kot  $2\varphi = 3,6125^\circ$  pa velja, da lahko izračunamo največ le 18 ocen globine. S tem smo lepo opisali negativno lastnost majhnega kota  $\varphi$ .



Slika 2: Grafa odvisnosti funkcije globine  $l$  od kota  $\theta$  pri razdalji  $r = 30$  cm in različnih vrednostih kota  $\varphi$ : a)  $2\varphi = 29,9625^\circ$ , b)  $2\varphi = 3,6125^\circ$ . Za lažjo primerjavo velikosti napake za en slikovni element pri oceni kota  $\theta$  je med vertikalnima črtama pri tretji točki prikazan interval širine  $\frac{\theta_0}{2} = 0,1^\circ$ .

Vidimo, da je pri majhnem kotu  $\varphi$  napaka veliko večja kot pri večjem kotu  $\varphi$ . Velja, da je hitrost rekonstrukcije v obratnem sorazmerju z natančnostjo le-te.

## 6.3 Določitev največje globine, v katere oceno še zaupamo

V poglavju 6.1 smo ocenili najmanjšo možno globino  $l_{\min}$  in največjo možno globino  $l_{\max}$ , nismo pa zapisali, kakšen je pomen napake za en slikovni element pri oceni kota  $\theta$  za ti dve oceni. Oglejmo si torej vrednost te napake ( $\Delta l$ ) za ti dve oceni:  $\Delta l_{\min}$  izračunamo kot absolutno vrednost razlike med  $l_{\min}$  in globino  $l$ , za katero velja, da je kot  $\theta$  za kot  $\frac{\theta_0}{2}$  večji od kota  $\theta_{\min}$ :

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

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

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

kjer je spremenljivka  $n$  pozitivno celo število, ki jo dobimo po enačbi:  $n = \varphi \div \frac{\theta_0}{2}$ . Iz tabele 1 so razvidne vrednosti napak glede na kot  $\varphi$ .

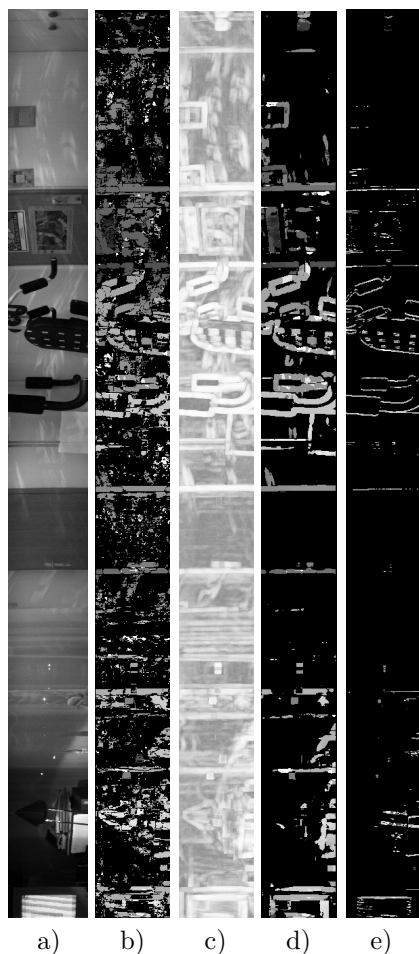
	$2\varphi = 29,9625^\circ$	$2\varphi = 3,6125^\circ$
$\Delta l_{\min}$	2 mm	19 mm
$\Delta l_{\max}$	30172 mm	81587 mm

Tabela 1: Pomen napake ( $\Delta l$ ) za en slikovni element pri oceni kota  $\theta$  za oceno najmanjše možne globine  $l_{\min}$  in največje možne globine  $l_{\max}$  glede na kot  $\varphi$ .

Dobljeni rezultati potrjujejo že zapisane ugotovitve (poglavje 6.2), zapišemo pa lahko še dve dodatni ugotovitvi: 1) Vrednost napake  $\Delta l_{\max}$  je nesprejemljivo visoka in to neglede na kot  $\varphi$ . Torej moramo smiselno zmanjšati oceno za največjo možno globino  $l_{\max}$ . To pomeni, da moramo definirati zgornjo mejo dovoljene napake ( $\Delta l$ ) za en slikovni element pri oceni kota  $\theta$  in s tem posledično določimo največjo globino, v katere oceno še zaupamo. 2) Kot  $\varphi$  je vedno odvisen od horizontalnega zornega kota  $\alpha$  (enačba (2)), le-ta pa na standardnih kamerah dosega vrednosti do približno  $40^\circ$ . To pomeni, da smo v praksi pri oceni globine omejeni s kotom  $\alpha$ , saj v najboljšem primeru velja:  $\varphi_{\max} = \frac{\alpha}{2}$ . Torej je naš sistem res uporaben le v primeru rekonstrukcije majhnih prostorov.

## 6.4 Rezultati

Na sliki 3 so prikazani primeri rezultatov našega sistema. V primeru b) smo gradili gosto 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,



Slika 3: Primer rezultatov za kot  $2\varphi = 29,9625^\circ$ , ko delamo rekonstrukcijo za levo oko: a) panorama za levo oko, b) gosta globinska slika / uporabimo vzvratno korelacijo / čas rekonstrukcije: 6 ur, 42 min., 20 sek., c) informacija o zaupanju v ocenjeno globino, d) gosta 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 sek. Ločljivost slik:  $1501 \times 118$  slikovnih pik.

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 [1, 4]. Če uporabimo manjši kot  $2\varphi = 3,6125^\circ$  so časi rekonstrukcije manjši tudi do 8-krat.

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 [4].

## 7 Sklepi

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; zaupanje v ocenjeno globino se spreminja: večja je strmina krivulje funkcije  $l$ , manjše je zaupanje v ocenjeno globino; gostih globinskih slik ni smiselno graditi, že zaradi samega časa, potrebne za rekonstrukcijo; sistem je uporaben v primeru rekonstrukcije majhnih prostorov.

## Literatura

- [1] O. Faugeras, *Three-Dimensional Computer Vision: A Geometric Viewpoint*, MIT Press, Cambridge, Massachusetts, London, England, 1993.
- [2] R. Gupta, R. I. Hartley, Linear pushbroom cameras, *IEEE Trans. PAMI*, 19(9):963–975, September 1997.
- [3] F. Huang, T. Pajdla, Epipolar geometry in concentric panoramas, *Technical Report CTU-CMP-2000-07*, Czech Technical University, 2000.
- [4] H. Ishiguro, M. Yamamoto, S. Tsuji, Omnidirectional stereo, *IEEE Trans. PAMI*, 14(2):257–262, February 1992.
- [5] S. Peleg, M. Ben-Ezra, Stereo panorama with a single camera, *IEEE Conf. CVPR*, pp. 395–401, Fort Collins, USA, 1999.
- [6] S. Peleg, Y. Pritch, M. Ben-Ezra, Cameras for stereo panoramic imaging, *IEEE Conf. CVPR*, pp. 208–214, Hilton Head Island, USA, 2000.
- [7] S. Peleg, B. Rousso, A. Rav-Acha, A. Zomet, Mosaicing on adaptive manifolds, *IEEE Trans. PAMI*, 22(10):1144–1154, October 2000.
- [8] B. Prihavec, F. Solina, User interface for video observation over the internet, *Journal of Network and Computer Applications*, 21:219–237, 1998.
- [9] P. Rademacher, G. Bishop, Multiple-center-of-projection images, *Computer Graphics (ACM SIGGRAPH)*, pp. 199–206, Orlando, USA, 1998.
- [10] H. Y. Shum, R. Szeliski, Stereo reconstruction from multiperspective panoramas, *IEEE ICCV*, pp. 14–21, Kerkyra, Greece, 1999.
- [11] 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.