

# Načrtovanje najboljšega naslednjega pogleda za gradnjo 3D-modelov predmetov iz barvnih slik

Kristian Žarn, Danijel Skočaj

Univerza v Ljubljani, Fakulteta za računalništvo in informatiko, Večna pot 113, 1000 Ljubljana, Slovenija  
E-pošta: kristian.zarn@gmail.com, danijel.skocaj@fri.uni-lj.si

**Povzetek.** Kakovost pridobljenega 3D-modela je močno odvisna od zajetih slik. Ročno zajemanje je lahko dolgotrajno opravilo, pri katerem želimo s slikami doseči ustrezno natančnost in pokritost modela. Uporabnik, ki med zajemanjem nima povratne informacije o ustreznosti slik, s težavo upošteva vse predpostavke algoritmov, kar je lahko vzrok za neuspešno rekonstrukcijo, ponovno zajemanje pa je lahko zelo drago ali celo nemogoče. V tem delu predstavimo novo metodo za načrtovanje najboljšega naslednjega pogleda, ki temelji na novi meri za oceno kakovosti 3D-modela. Na ta način uporabnik dobi namig, iz katerih pogledov naj zajame slike, da zagotovi smiselno gradnjo 3D-modela. Predlagano metodo smo preizkusili s pomočjo razvitega sistema za rekonstrukcijo, ki omogoča sprotno gradnjo modela in odziv o primernosti zajetih slik. Z evalvacijo pokažemo, da je rekonstrukcija z našo metodo načrtovanja naslednjega pogleda v večini primerov boljša od rekonstrukcije iz slik zajetih s po navidezni polkrogli enakomerno razporejenimi kamerami.

**Ključne besede:** računalniški vid, 3D-rekonstrukcija, struktura iz gibanja, najboljši naslednji pogled

## Next best view planning for building a 3D model of objects from color images

The quality of a reconstructed 3D model heavily depends on the input images. The manual image acquisition to obtain the desired accuracy and completeness of the model can be a lengthy process. A user with no feedback about the image suitability may find it hard to comply with the assumptions of the algorithms, which can result in an unsuccessful model reconstruction. An additional image acquisition can be expensive or even impossible. In this work we introduce a new method for the next best view planning. The method is based on a new quality measure giving the user a hint from which view to capture the image to ensure a sensible reconstruction. The proposed method is tested with a system that allows an online reconstruction and immediate feedback about the image adequacy. We show that the reconstruction obtained by the next best view planning is generally better than the reconstruction from images captured from evenly-spaced view points in the shape of a hemisphere.

**Keywords:** computer vision, 3D reconstruction, next best view

## 1 UVOD

Pri rekonstrukciji želimo na podlagi zajetih slik pridobiti 3D-obliko, ki se predmetu najbolj prilega. Razviti postopki in algoritmi se uspešno uporabljajo v industriji, npr. za 3D-kartiranje, spletne trgovine, 3D-tisk, posebne učinke v filmih, računalniške igre in arhiviranje kulturne dediščine. Sodobne digitalne kamere lahko zaradi visoke ločljivosti in kakovosti slik ob ustrezni uporabi proizvedejo 3D-modele visoke kakovosti, ki je v določenih pogojih primerljiva z laserskim skeniranjem [15].

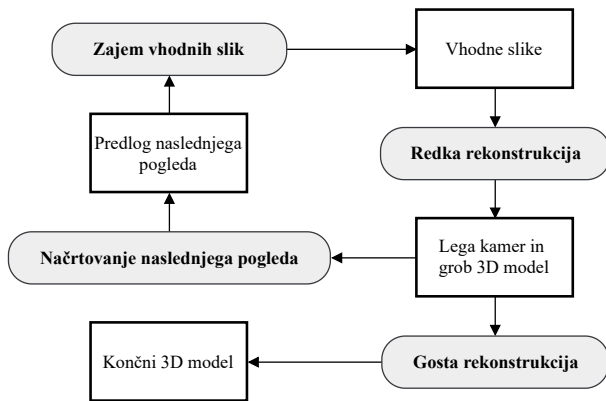
Prejet 19. december, 2019  
Odobren 26. februar, 2020

Kakovost rekonstrukcije je močno odvisna od zajetih slik. Algoritmi za rekonstrukcijo imajo nekatere predpostavke in omejitve, npr: kot in translacija med pari sosednjih pogledov morata biti primerna, rekonstruirani predmeti pa morajo imeti dovolj dobro teksturo. Uporabnik, ki med zajemanjem nima povratne informacije o ustreznosti slik, težko upošteva omenjene predpostavke. Lahko se zgodi, da zaradi neustreznih slik nekateri deli predmeta ne morejo biti rekonstruirani ali pa ni dosežena želena natančnost modelov. Ponovno zajemanje slik je lahko zelo drago ali celo nemogoče, kot npr. pri rekonstrukciji iz zračnih posnetkov in arhiviranju arheoloških najdb. Poleg tega se pri nenadzorovanemu zajemanju slik pogosto zajamejo tudi redundantne slike, ki povečajo čas procesiranja brez pomembnega prispevka h kakovosti modela.

Omenjene težave so pogoste v klasičnih pristopih, kjer je gradnja 3D-modela izvedena ločeno, po zajemanju slik. V tem delu se osredotočamo na razvoj nove metode za načrtovanje naslednjih pogledov (angl. next best view planning), ki sistematično izboljšajo kakovost rekonstrukcije. Z uporabo sodobnejših algoritmov smo razvili tudi programsko opremo, ki omogoča sprotno gradnjo grobega 3D-modela že med zajemanjem slik. Uporabili smo jo pri evalvaciji predlagane metode za načrtovanje pogledov.

Oris naše programske rešitve je prikazan na sliki 1. Diagram prikazuje odvisnost med funkcionalnostmi ter vhodnimi in izhodnimi podatki. Postopek se začne z zajemanjem slik, ki so pridobljene z IP-kamero. V našem primeru je kot IP-kamera uporabljen mobilni telefon, kar je mogoče s pomočjo dodatne mobilne aplikacije.

Iz zajetih slik se postopoma gradi redka rekonstrukcija, ki vrača lego kamer in grob 3D-model, predstavljen s trikotniško mrežo. Na podlagi te predstavitve so predlagani tudi naslednji pogledi. Ko je zajetih dovolj slik, lahko uporabnik z algoritmi goste rekonstrukcije pridobi končni 3D-model.



Slika 1: Oris programske rešitve. Sivi okvirji predstavljajo funkcionalnosti programa, beli pa njihove vhode in izhode.

V naslednjem poglavju predstavimo nekaj najbolj relevantnih sorodnih del. Kratek opis naše programske rešitve je podan v poglavju 3, podrobnejši opis naše metode za načrtovanje pogledov pa v poglavju 4. V poglavjih 5 in 6 predstavimo metodologijo evalvacije in nekaj rezultatov. Prispevek zaključimo s sklepnimi ugotovitvami ter smernicami za nadaljnje delo.

## 2 SORODNA DELA

Celoten postopek gradnje 3D-modelov iz barvnih slik v grobem delimo na redko in gosto rekonstrukcijo. Najbolj popularen postopek za prvo je imenovan struktura iz gibanja (angl. structure from motion), označen je s kratico SfM [1]. Sestavljen je lahko iz različnih algoritmov in na splošno deluje tako, da se na slikah poiščejo značilnice, nato pa se na podlagi ujemanj med slikami izračunajo njihove 3D-koordinate. Rezultat so redek oblak 3D-točk ter pozicija in orientacija kamer, ki pripadajo vhodnim slikam. Wu je v svojem delu [6] z izvirno strategijo uporabe algoritmov pokazal, da je lahko časovna zahtevnost postopka SfM skoraj linearna tudi pri rekonstrukciji z več tisoč slikami. Na takšnem postopku temeljijo številni sodobni pristopi za redko rekonstrukcijo, uporabljamo pa ga tudi v našem delu.

Če želimo na podlagi slik pridobiti 3D-model visoke ločljivosti, ki se bolj natančno prilega obliki dejanskega predmeta, potem lahko redek oblak točk nadgradimo z gosto rekonstrukcijo. V nasprotju z redko rekonstrukcijo je v tem primeru cilj pridobiti globino (in posledično 3D-lokacijo) vsakega slikovnega elementa na vsaki sliki. To storimo z algoritmi s področja večpoglednega stera (angl. multi-view stereo), ki za vhod vzamejo rezultate

redke rekonstrukcije in proizvedejo gost 3D-model. Opis in primerjava različnih pristopov rekonstrukcije na tem področju sta na voljo v preglednem članku [2].

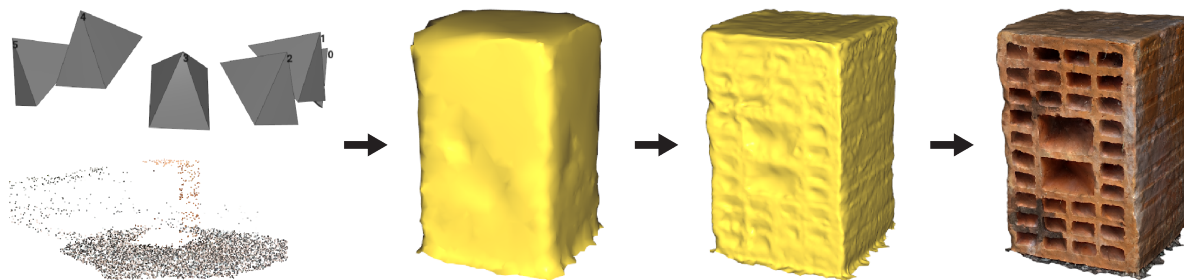
V literaturi so s pojmom načrtovanje naslednjega pogleda označeni zelo različni pristopi. V tem delu problem obravnavamo kot inkrementalni proces sistematične izboljšave natančnosti in pokritosti 3D-modela. Kriegel in drugi [3] so za gradnjo 3D-modela uporabili industrijski robotski manipulator z laserskim skenerjem. Naslednji pogledi so izbrani na podlagi robov v modelu in cenovne funkcije, ki stremi k raziskovanju. Na področju rekonstrukcije iz barvnih slik sta problem naslovila Dunn in Frahm [4]. Predlagala sta cenovno funkcijo, ki združuje negotovost v oceni strukture, projekcijo modela in videz teksture. Naslednji pogled je določen z optimizacijo te funkcije. Sproten odziv pri rekonstrukciji so naslovili tudi Hoppe in drugi [5]. V svojem delu so razvili rešitev, ki s sprotno informacijo o ustreznosti slik in kakovosti modela pomaga uporabniku pri rekonstrukciji.

## 3 GRADNJA 3D-MODELA

Kot prikazuje slika 1, je postopek gradnje 3D-modelov predmetov razdeljen na redko in gosto rekonstrukcijo. Ti komponenti naše programske opreme sta opisani v tem poglavju.

### 3.1 Redka rekonstrukcija

Za redko rekonstrukcijo smo uporabili postopek, imenovan inkrementalna struktura iz gibanja [6]. Rekonstrukcija je inicializirana na podlagi prvih dveh slik. Na slikah poiščemo značilnice SIFT [7] in njihova ujemanja. Na podlagi ujemanj izračunamo tako imenovano osnovno matriko (angl. essential matrix). Iz nje pridobimo relativno lego prvih dveh kamer [8]. Magnituda translacije iz osnovne matrike ne more biti določena, zato pravimo, da je rekonstrukcija določena do velikosti natančno. Oblak 3D-točk inicializiramo s triangulacijo [10] ujemajočih značilnic. Rekonstrukcijo nato postopoma nadgrajujemo z dodajanjem novih slik. Iz nove vhodne slike najprej pridobimo značilnice. Naslednji korak je kritičen za hitro razširitev. Namesto iskanja ujemanj z vsemi predhodnimi slikami iščemo ujemanja samo s petimi najbolj podobnimi slikami. Te kandidate za ujemanje pridobimo z uporabo tako imenovanega vizualnega slovarja [11]. Značilnice nove slike, za katere smo našli ujemanje, lahko zdaj razdelimo v dve množici. V prvi množici so tiste, ki so že del rekonstrukcije. Uporabimo jih za določanje absolutne lege kamere [9]. V drugi množici so tiste, ki še niso del rekonstrukcije. Dodamo jih s triangulacijo. Pomembno je, da po vsaki dodani sliki minimiziramo reprojekcijsko napako. S tem izboljšamo natančnost in omilimo akumulacijo napake. Na podlagi oblaka točk in lege kamer lahko rekonstruiramo površino predmeta [12]. Ker gre za hitro operacijo, jo lahko izvedemo po vsaki dodani sliki. Tako pridobimo



Slika 2: Prikazani so rezultati korakov gradnje 3D-modela. Od leve proti desni si sledijo, redka rekonstrukcija z metodo Struktura iz gibanja, rekonstrukcija površine, izboljšava ločljivosti in natančnosti ter teksturiranje.

grob 3D-model, predstavljen s trikotniško mrežo. Na sliki 2 so prikazani vmesni rezultati gradnje 3D-modela. Opisani postopek zajema prva dva koraka.

Kot osnovo pri implementaciji redke rekonstrukcije smo uporabili knjižnico TheiaSfM [17]. Ta vsebuje številne algoritme s področja strukture iz gibanja. Za hitrejše iskanje značilnic smo uporabili implementacijo na grafični kartici iz knjižnice SiftGPU [18]. Implementacijo vizualnega slovarja smo pridobili iz knjižnice COLMAP [19].

### 3.2 Gosta rekonstrukcija

Za ceno dodatnega procesiranja lahko redko rekonstrukcijo izboljšamo z algoritmi goste rekonstrukcije. Knjižnica OpenMVS [20] ponuja implementacijo algoritmov za rekonstrukcijo površine, izboljšavo ločljivosti in generiranje teksture. Te algoritme smo integrirali v našo rešitev.

Izboljšava ločljivosti 3D-modela [13] je časovno zahtevna operacija, ki jo izvedemo, ko so zajete vse slike. Ločljivost grobega modela je z delitvijo trikotnikov prilagojena ločljivosti vhodnih slik. Nad vozlišči gostejšega modela je definirana energijska funkcija. Z minimizacijo te funkcije določimo natančnejšo pozicijo vozlišč.

Končni model pridobimo s teksturiranjem [14], ki je ključnega pomena za doseg realističnega videza. Algoritem deluje v dveh korakih. V prvem koraku vsakemu trikotniku pripišemo oznako, ki določa en pogled, ki bo uporabljen za njegovo teksturiranje. Drugi korak algoritma pa je globalna uskladitev barv, tako da med trikotniki ni vidnih prehodov. Opisana algoritma zajemata zadnja dva koraka slike 2.

## 4 NAČRTOVANJE NASLEDNJEGA POGLEDA

V tem poglavju opišemo predlagano metodo za načrtovanje naslednjega pogleda, ki je glavni prispevek tega dela. Najprej predstavimo še oceno kakovosti 3D-modela, ki je temelj za delovanje naše metode.

### 4.1 Ocena kakovosti 3D-modela

Za predstavitev 3D-modelov v našem delu uporabljamo trikotniško mrežo. To je seznam točk  $M_i$  in trikotnikov  $T_i$ . Tako kot v [5], kakovost 3D-modela

definiramo s funkcijo  $Q$ , ki vsakemu trikotniku  $T_i = \{M_1, M_2, M_3\}$  priredi realno vrednost. Takšno preslikavo lahko zapišemo z enačbo

$$Q : \{\mathbb{R}^3, \mathbb{R}^3, \mathbb{R}^3\} \rightarrow \mathbb{R} \\ T_i \mapsto Q(T_i). \quad (1)$$

Želimo si, da prirejena vrednost čim bolj predstavlja kakovost končnega modela, ki je merjena z natančnostjo in pokritostjo. Tako imamo že med rekonstrukcijo na voljo informacije o tem, kateri deli modela so bolj oz. manj natančni. Tako lahko sprejmemo informirano odločitev o postavitvi naslednje kamere. Ta ideja je osnova za delovanje naše metode načrtovanja naslednjih pogledov. Hoppe in drugi v svojem delu [5] predlagajo mero GSD (angl. ground sampling distance). V našem delu predlagamo podobno, vendar nekoliko spremenjeno mero, imenovano PPA (angl. pixels per area), ki se je izkazala za bolj robustno.

Mera GSD se izračuna na naslednji način. Posamezni trikotnik  $T_i$  preslikamo nazaj na kamere  $C_j$ . Ploščino trikotnika nato delimo z maksimalno ločljivostjo preslikave in rezultat korenimo. Mero zapišemo z enačbo

$$Q_{GSD}(T_i) = \min_{C_j} \sqrt{\frac{A(T_i)}{P(T_i, C_j)}}. \quad (2)$$

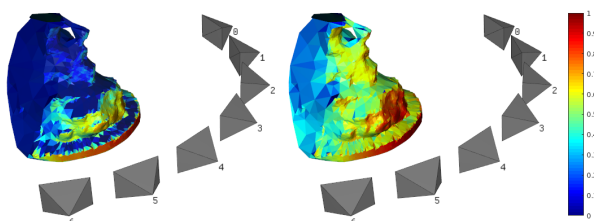
V zgornji enačbi funkcija  $P$  vrne število slikovnih elementov, ki jih trikotnik  $T_i$  pokrije na sliki kamere  $C_j$ , funkcija  $A$  pa vrne ploščino trikotnika v 3D-prostoru. Manjša vrednost funkcije  $Q_{GSD}$  pomeni večjo kakovost modela. Zaradi majhnih in manj vidnih trikotnikov ta mera proizvede tudi nekatere zelo izstopajoče vrednosti, zato odrežemo  $\alpha = 10\%$  najbolj izstopajočih, tako kot v originalnem delu [5].

Mero PPA izračunamo na podoben način. Trikotnik  $T_i$  je ponovno preslikan nazaj na kamere  $C_j$ . Število slikovnih elementov preslikav tokrat seštejemo in delimo s ploščino trikotnika v 3D-prostoru. Da omilimo vpliv večjih vrednosti, končni rezultat korenimo. Mero zapišemo z enačbo

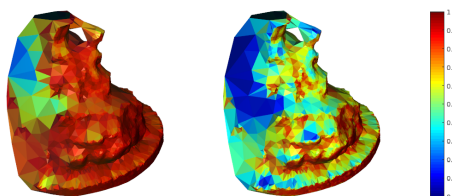
$$Q_{PPA}(T_i) = \sqrt{\frac{\sum_{C_j} P(T_i, C_j)}{A(T_i)}}. \quad (3)$$

V tem primeru večja vrednost funkcije  $Q_{PPA}$  pomeni večjo kakovost modela. Posebna normalizacija kot pri meri GSD tu ni potrebna. V zgornjih enačbah je računsko najzahtevnejša funkcija  $P$ , pri kateri moramo upoštevati medsebojno prekrivanje trikotnikov celotnega modela. Za učinkovito implementacijo smo uporabili grafično kartico.

Na sliki 3 sta obe meri kakovosti vizualizirani na konkretnem primeru. Za primerjavo smo na sliki 4 prikazali tudi dejansko natančnost modela. Postopek izračuna natančnosti je opisan v podpoglavju 5.2. Opazimo lahko, da se v tem primeru mera PPA vizualno bolje ujema z dejansko natančnostjo. Obsežnejša evalvacija je predstavljena v podpoglavju 6.1.



Slika 3: Vizualizacija kakovosti za konkreten 3D-model. Na levi strani je prikazana mera GSD, na desni pa PPA. Meri sta normalizirani tako, da zavzameja vrednosti na intervalu  $[0, 1]$ , kjer 1 predstavlja največjo in 0 najmanjšo kakovost. Opazimo lahko, da je bližjim trikotnikom in tistim z boljšo vidnostjo na kamerah pripisana večja kakovost.



Slika 4: Vizualizacija dejanske natančnosti modela za primerjavo z merama GSD in PPA. Leva stran prikazuje nespremenjeno natančnost. Na desni strani smo za potrebe vizualizacije odstranili 10 % največjih vrednosti. Tako omilimo vpliv velikih odstopanj in zmanjšamo razpon med vrednostmi. Rdeča barva predstavlja dobro natančnost, modra pa slabo.

Obe meri predpostavljata, da boljša vidnost trikotnikov proizvede boljše kakovost modela. Razlikujeta se v tem, da mera PPA upošteva vidnost trikotnika na vseh kamerah, medtem ko je pri meri GSD pomembna le čim večja ločljivost na eni izmed kamer. Omenjena predpostavka drži v primerih, ko ima rekonstruirani predmet dobro teksturo in je pridobljenih dovolj značilnic. V določenih primerih lahko model zaradi lukenj, odsevnih materialov, slabe teksture in šuma vsebuje trikotnike, ki se dejanski površini predmeta ne prilegajo, vendar so kljub temu dobro vidni na kamerah. V takšnih primerih opisani meri proizvedeta slabše rezultate.

## 4.2 Najboljši naslednji pogled

Problem načrtovanja naslednjega pogleda (angl. next best view planning) je definiran kot iskanje nove oz. dodatne postavitve senzorja z namenom izboljšave obstoječe rekonstrukcije oz. predstavitve prostora. Namesto barvne kamere so na splošno lahko uporabljani tudi drugi senzori (npr. globinske kamere, laserski skenerji itd.). Odvisno od konteksta lahko na načrtovanje pogledov gledamo kot na strategijo za avtonomno raziskovanje ali izbiro najboljših vhodnih slik pri rekonstrukciji iz velike množice podatkov. Tako kot v [4] v tem delu problem obravnavamo kot inkrementalni proces sistematične izboljšave natančnosti in pokritosti 3D-modela.

Naš pristop za načrtovanje naslednjega pogleda temelji na oceni kakovosti 3D-modela. Najprej definiramo cenovno funkcijo, ki za podano lego kamere oceni, kako dobra je njena postavitvev. Funkcijo zapišemo z enačbo

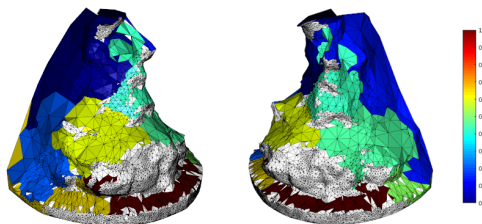
$$f_{NBV} : \mathbb{R}^{4 \times 4} \rightarrow \mathbb{R}$$

$$f_{NBV}(C) = \alpha * \mu(Q_{PPA}(v(C))) - \beta * \sigma(Q_{PPA}(v(C))). \quad (4)$$

Spremenljivka  $C$  je  $4 \times 4$  matrika lege kamere, ki vsebuje njeno pozicijo in orientacijo. Funkcija  $v$  za postavitev kamere  $C$  vrne vse vidne trikotnike, ki jih označimo s  $\mathcal{T}$ . Funkcijo  $Q$ , ki za podano množico  $\mathcal{T}$  vrne pripadajoče kakovosti, smo definirali v prejšnjem podpoglavju. V našem primeru smo uporabili mero PPA. Na podlagi pridobljenih vrednosti izračunamo dva podatka, in sicer njihovo povprečje  $\mu$  (prvi člen enačbe) in standardni odklon  $\sigma$  (drugi člen enačbe). Utemeljitev takšne definicije je naslednja. Za novo kamero si želimo, da izboljša območja slabe kakovosti (nizka vrednost  $\mu$ ), po drugi strani pa je kamera lahko pravilno lokalizirana oz. dodana k rekonstrukciji samo, če so vidni tudi bolj kakovostni deli modela (visoka vrednost  $\sigma$ ). Želimo si torej čim večji razpon med vrednostmi. Ta kompromis med vidnostjo trikotnikov z dobro in slabo kakovostjo uravnavamo s parametroma  $\alpha = 1$  in  $\beta = 5$ . Vrednost funkcije poskušamo minimizirati.

Hitrost izračuna funkcije  $f_{NBV}$  omejuje iskanje vidnih trikotnikov, pri katerem moramo upoštevati njihovo medsebojno prekrivanje. Čeprav je implementacija na grafični kartici sorazmerno hitra, je to še vedno najpočasnejši del izračuna, zato postopek načrtovanja naslednjega pogleda začnemo z gručenjem trikotnikov modela, ki opravlja dve nalogi. Na podlagi gruč generiramo bolj obvladljivo število postavitvev kamere. Te postavitve predstavljajo kandidate za naslednji pogled. Pri velikih modelih je število gruč (in posledično kandidatov) lahko še vedno preveliko, zato so kandidati dodatno filtrirani na podlagi lokalne ocene za funkcijo  $f_{NBV}$ , ki je izračunana na podlagi trikotnikov v posamezni gruči. Na koncu funkcijo  $f_{NBV}$  izračunamo le za množico najboljših kandidatov, ki je dovolj majhna za hiter izračun.

Podrobnejši opis korakov je naslednji. Na začetku model naključno razdelimo na povezane gruče, ki vsebujejo med  $c_{min}$  in  $c_{max}$  trikotnikov. Normala posameznega trikotnika v gruči od povprečja ne sme odstopati za več kot  $\phi$  stopinj. Primer gručenja je prikazan na sliki 5. Za vsako gručo ovrednotimo lokalno oceno funkcije  $f_{NBV}$  tako, da namesto trikotnikov, vidnih na kameri, uporabimo trikotnike v gruči. Ker za lokalno oceno ne potrebujemo informacije o vidnosti, je ta izračun lahko zelo hiter.



Slika 5: Primer gručenja na konkretnem primeru. Prikazana sta dva pogleda istega modela. Trikotniki iste gruče so obarvani z isto barvo. Barva predstavlja tudi lokalno oceno cenovne funkcije. Gruče na robu modela z nizko ceno (modra barva) predstavljajo dobre kandidate za postavitev kamere. Nekateri trikotniki niso del gruče (bela barva), saj njihova ocena kakovosti presega trenutno ciljno vrednost.

Za vsako gručo generiramo eno postavitev kamere, ki predstavlja kandidata za naslednji pogled. Kamera je postavljena od središča gruče  $G$  v smeri povprečja normal trikotnikov znotraj gruče. Usmerjena je proti centru gruče, oddaljenost  $d$  pa je določena z enačbo

$$d = \gamma \sqrt[3]{A(G)}. \quad (5)$$

Funkcija  $A$  vrne ploščino gruče v 3D-prostoru, korenska funkcija pa prepreči, da bi bila pri velikih gručah kamera preveč oddaljena. Dodaten parameter  $\gamma$  je odvisen od goriščne razdalje kamere. V tem delu smo uporabili le dve različni kameri (virtualno in resnično), zato smo v obeh primerih ta parameter določili empirično. V podpoglavju 6.2 uporabljamo vrednost  $\gamma = 4.0$ , v podpoglavju 6.3 pa vrednost  $\gamma = 3.0$ . Za najboljših  $n$  kandidatov (v smislu lokalne ocene) nato izračunamo cenovno funkcijo  $f_{NBV}$  in za najboljši naslednji pogled predlagamo tistega z najmanjšo ceno. Opisano načrtovanje naslednjega pogleda je povzeto v algoritmu 1. Podane so tudi konkretne vrednosti parametrov, ki smo jih določili empirično.

Izboljšava kakovosti modela je postopna. Na začetku rekonstrukcije postavimo ciljno kakovost  $q_t = 1500$ . Vsakič ko  $q_p = 85\%$  odstotkov trikotnikov doseže ciljno kakovost, parameter  $q_t$  povečamo za  $q_{inc} = 500$ . Parameter  $q_t$  se torej postopoma povečuje in zasede vrednosti 1500, 2000, 2500 itd. Tako še dodatno spodbudimo raziskovanje in preprečimo morebitno kopičenje pogledov na enem območju.

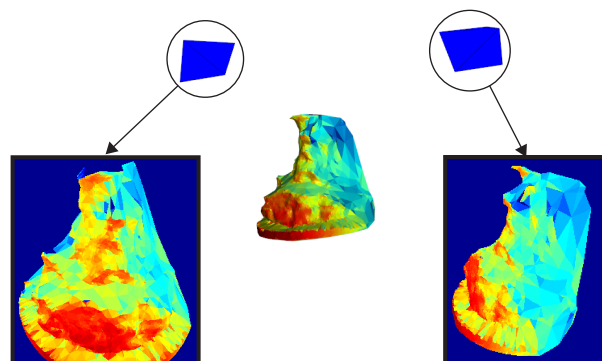
### Algoritem 1 Načrtovanje naslednjega pogleda

**Vhod:** 3D-model, predstavljen s trikotniško mrežo  $M_T$

**Izhod:** Lega predlaganega pogleda  $C_{nbv}$

- 1:  $M_Q \leftarrow$  Izračun kakovosti za trikotniško mrežo  $Q_{PPA}(M_T)$
- 2:  $G_1 \dots G_n \leftarrow$  Delitev modela na gruče ( $c_{min} = 100$ ,  $c_{max} = 300$ ,  $\phi = 100$ )
- 3:  $f_{LOC}(G_i) \leftarrow$  Lokalna ocena funkcije  $f_{NBV}$  za vse posamezne gruče  $G_i$  na podlagi trikotnikov znotraj gruče ( $\alpha = 1$ ,  $\beta = 5$ )
- 4:  $C_1 \dots C_n \leftarrow$  Za vsako gručo  $G_i$  je generirana postavitev kamere  $C_i$  (kandidati za naslednji pogled)
- 5:  $f_{NBV}(C_i) \leftarrow$  Ovrednotenje cenovne funkcije  $f_{NBV}(C_i)$  za najboljših  $n = 20$  kandidatov z najboljšo lokalno oceno  $f_{LOC}(G_i)$  ( $\alpha = 1$ ,  $\beta = 5$ )
- 6:  $C_{nbv} \leftarrow$  Izbira najboljšega kandidata

Izračun cenovne funkcije  $f_{NBV}$  na konkretnem primeru je prikazan na sliki 6. Ponovno je uporabljen isti model kot na sliki 3, le da kamere rekonstrukcije tokrat niso prikazane. Na desni strani je prikazan najboljši predlagani pogled, na levi pa šesti najboljši. Prikazana je tudi upodobitev modela z mero PPA na obeh kamerah. Vrednost cenovne funkcije  $f_{NBV}$  za levi pogled znaša 30,3, za desni pogled pa  $-111,9$ . Desni pogled ima nižjo ceno zaradi večjega razpona vrednosti mere PPA, kar je razvidno tudi iz upodobitve modela. Ta pogled je boljša izbira, saj z njim zajamemo večje število slabše ocenjenih trikotnikov in s tem spodbujamo širitev rekonstrukcije, obenem pa z vidnostjo dobro ocenjenih trikotnikov omogočimo njegovo uspešno lokalizacijo. Levi pogled je postavljen blizu obstoječih kamer in ima manjši prispevek k izboljšavi rekonstrukcije.



Slika 6: Upodobitev modela z mero PPA za dva predlagana pogleda. Vrednost cenovne funkcije  $f_{NBV}$  za levi pogled znaša 30,3, za desni primer pa  $-111,9$ . Desni pogled je bolje ocenjen zaradi manjšega povprečja in večjega standardnega odklona mere PPA na vidnih trikotnikih.

## 5 METODOLOGIJA EVALVACIJE

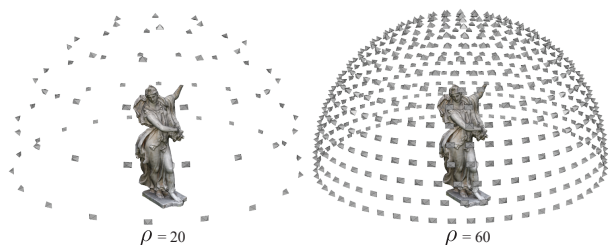
V tem poglavju predstavimo metodologijo evalvacije načrtovanja pogledov. Kvantitativno evalvacijo opravimo v simuliranem okolju, in sicer z rekonstrukcijo računalniško zgeneriranih 3D-modelov. Na ta način lahko rekonstrukcije primerjamo z referenčnimi 3D-modeli, poleg tega pa je poznana tudi natančna lega kamer pri zajemanju slik. Uporabljeni referenčni 3D-modeli so prosto dostopni na spletnem portalu Sketchfab\* in so objavljeni pod licenco CC. Rezultati so predstavljeni v poglavju 6, kjer podamo tudi nekaj primerov rekonstrukcij resničnih predmetov.

### 5.1 Enakomerna postavitve kamer

Rezultate načrtovanja naslednjega pogleda smo primerjali s t. i. osnovno rekonstrukcijo. Ta je pridobljena iz enakomerne postavitve kamer. Kamere so razporejene enakomerno po navidezni polkrogli in usmerjene proti njenemu središču. Število kamer na prvem (spodnjem) obroču določa parameter  $\rho$ , kamere na preostalih obročih pa so postavljene z enakim medsebojnim razmikom.

Primer enakomerne postavitve kamer je prikazan na sliki 7. V tem primeru sta prikazani dve postavitvi kamer, in sicer za gostoti  $\rho = \{20, 60\}$ , ki na vseh nivojih skupaj proizvedeta 72 oz. 596 kamer. Pri evalvaciji je bila uporabljena še vmesna gostota kamer, in sicer  $\rho = 40$ . Slike ločljivosti  $2048 \times 1536$  upodabljam z virtualno kamero, za katero poznamo notranje parametre. Najprej so dodane slike na spodnjem obroču, nato nadaljujemo z naslednjimi obroči proti vrhu.

Na splošno je takšna postavitve kamer zelo ugodna pri rekonstrukciji predmetov z večinoma konveksno obliko, ki jih lahko postavimo v središče polkrogle. Med drugim je takšna postavitve uporabljena tudi pri gradnji testnih množic za evalvacijo rekonstrukcijskih algoritmov [15], [16]. Za bolj pošteno primerjavo z načrtovanjem pogledov smo referenčne modele izbrali tako, da je enakomerna postavitve kamer sorazmerno ugodna za njihovo rekonstrukcijo.

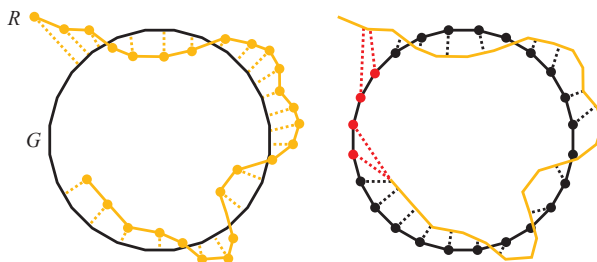


Slika 7: Prikazana je enakomerna postavitve kamer za gostoti  $\rho = \{20, 60\}$ . Skupno število kamer je 72 (levo) in 596 (desno). V sredini polkrogel je eden izmed referenčnih modelov. Rezultat rekonstrukcije referenčnega modela s takšno postavitvijo kamer je t. i. osnovna rekonstrukcija, ki jo v nadaljevanju uporabimo pri evalvaciji.

\*<https://sketchfab.com>

### 5.2 Izračun natančnosti in pokritosti modela

Za primerjavo rekonstrukcije z referenčnim modelom je v literaturi uveljavljen pristop prvotno predstavljen v delu [15]. Omenili smo že, da kakovost rekonstrukcije merimo z natančnostjo in pokritostjo modela. Tu predstavimo natančnejšo definicijo obeh pojmov. Referenčni model označimo z  $G$ , rezultat rekonstrukcije pa z  $R$ . Natančnost in pokritost sta izračunana na podlagi razdalje med  $R$  in  $G$ . Pomen te razdalje je grafično prikazan na sliki 8.



Slika 8: Leva stran prikazuje razdaljo od rekonstrukcije  $R$  do referenčnega modela  $G$ . Za vsako točko v  $R$  poiščemo najbližjo točko v  $G$ . Na podlagi teh razdalj je izračunana natančnost. Desna stran prikazuje izračun razdalje v drugo smer (od  $G$  do  $R$ ) na podlagi katere izračunamo pokritost. Na sliki je razdalja za rdeče točke prevelika, zato jih  $R$  ne pokrije. Skica je povzeta po [15].

Pri izračunu natančnosti moramo za vsako točko v  $R$  poiskati najbližjo točko v  $G$ . Modeli so v našem delu predstavljeni s trikotniško mrežo. Pri takšni predstavitvi so velikosti trikotnikov lahko zelo neenakomerne. Pri računanju razdalje zato predhodno iz originalnega modela z enakomernim vzorčenjem pridemo do oblaka točk, ki je nato uporabljen za primerjavo. S tem izničimo učinek različnih velikosti trikotnikov. Mero natančnosti predstavlja tista razdalja  $A_d$ , za katero je 90 % točk v  $R$ , do pripadajočih najbližjih točk v  $G$ , oddaljenih za največ  $A_d$ . To mero poimenujemo *nenatančnost* in si želimo, da je pri rekonstrukciji čim manjša.

Pri izračunu pokritosti iščemo razdaljo v drugi smeri, in sicer za točke v  $G$  poiščemo najbližje točke v  $R$ . Za posamezne točke referenčnega modela pravimo, da so pokrite, če je njihova razdalja do rekonstrukcije dovolj majhna. Tukaj je treba uvesti še dodatno mejno vrednost  $C_t$ , ki določa, kdaj je razdalja še sprejemljiva. Ker absolutna velikost referenčnih modelov ni znana, je to mejno vrednost treba določiti za vsak model posebej. Pri tem si pomagamo z algoritmom PCA (angl. principal component analysis). Gre za splošen algoritem, ki za vhodni oblak točk poišče novo ortogonalno bazo, tako da bazni vektorji zajemajo čim več variance v podatkih. Pri evalvaciji smo za mejno vrednost  $C_t$  izbrali 1 % dolžine najdaljšega baznega vektorja, torej mero pokritosti predstavlja delež točk  $G$ , za katere je razdalja do  $R$  manjša od  $C_t$ .

## 6 REZULTATI

Rezultati evalvacije našega dela so razdeljeni v tri večje sklope. Predlagano mero PPA najprej primerjamo z obstoječo mero. Nato predstavimo rezultate evalvacije načrtovanja naslednjega pogleda v simuliranem okolju, poglavje pa zaključimo z nekaj primeri rekonstrukcij resničnih predmetov.

### 6.1 Evalvacija mere za kakovost

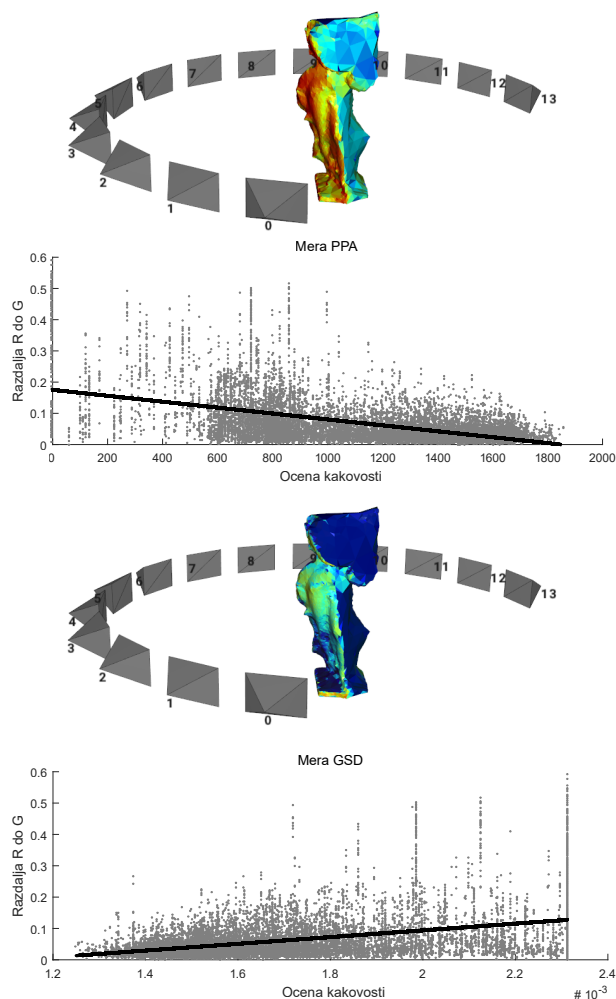
Za meri PPA in GSD smo preverili, kako dobro odražata dejansko natančnost rekonstrukcije. To smo storili z izračunom Pearsonovega koeficienta korelacije, ki je definiran z enačbo

$$r_{xy} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}} \quad (6)$$

Gre za najpogosteje uporabljeno mero linearne povezanosti dveh številskih spremenljivk. Koeficient lahko zavzame vrednosti na intervalu  $[-1, 1]$ , kjer vrednosti  $\pm 1$  predstavljata popolno linearno odvisnost, vrednost 0 pa pomeni, da spremenljivki nista odvisni.

Pri evalvaciji smo uporabili 4 referenčne modele, in sicer fontano, nagrobnik, hidrant in kip. Za vsak model smo gradili rekonstrukcijo z enakomerno postavitvijo kamer gostote  $\rho = 20$ . Po vsaki dodani sliki smo za vmesni model izračunali razdalje od točk rekonstrukcije do najbližjih točk referenčnega modela (razdalja R do G) ter obe meri kakovosti. Na sliki 9 je prikazan primer korelacije med natančnostjo in oceno kakovosti za meri PPA in GSD. Primer prikazuje vmesni model rekonstrukcije. Vsaka točka na grafu sovpada z eno izmed točk modela ter prikazuje njeno razdaljo do reference (os y) in oceno kakovosti (os x). Na obeh grafih je razvidno, da je napaka večja pri točkah s slabše ocenjeno kakovostjo. V tem primeru je vrednost koeficienta korelacije za mero PPA  $-0,52$ , za mero GSD pa  $0,44$ . Predznak se razlikuje, ker pri meri PPA večja vrednost ponazarja manjšo razdaljo do reference (pri meri GSD pa obratno). Za primerjavo nas zanima le njeno absolutna vrednost.

Za vsak testni model smo med gradnjo izračunali 69 koeficientov korelacije (od 72 slik so 3 uporabljene za inicializacijo) in na koncu njihove vrednosti povprečili. Rezultati so zbrani v tabeli 1. Prikazane so absolutne vrednosti. Koeficient korelacije je za mero PPA v povprečju 1,6-krat večji od mere GSD. Čeprav dejanske vrednosti predstavljajo nizko do srednjo linearno odvisnost spremenljivk, so rezultati nove mere občutno boljši od obstoječe. Pri interpretaciji rezultatov moramo upoštevati še, da odvisnost spremenljivk v resnici ni povsem linearna. Na neki točki namreč novi pogledi ne prispevajo pomembno k izboljšavi natančnosti in jo lahko v določenih primerih celo poslabšajo. Naj še omenimo, da je v tem primeru ocena kakovosti izračunana zgolj na podlagi vidnosti trikotnikov. Z vpeljavo dodatnih informacij (npr. videz texture, postavitev kamer itd.) je tu še nekaj prostora za izboljšave.



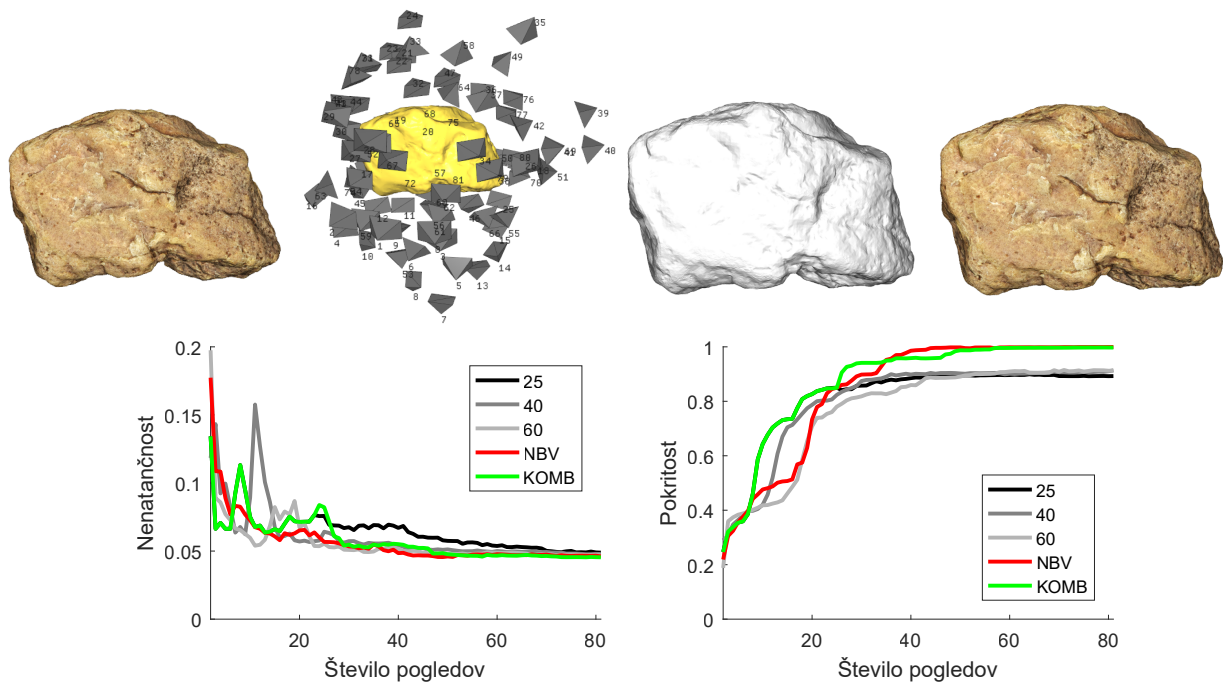
Slika 9: Grafa prikazujeta natančnost v odvisnosti od ocene kakovosti za meri PPA (zgoraj) in GSD (spodaj). Črna premica predstavlja linearni model, ki se podatkom najbolj prilaga (minimizira vsoto kvadratov napake). Grafa sta ustvarjena na podlagi prikazanih vmesnih modelov rekonstrukcije.

	Fontana	Nagrobnik	Hidrant	Kip
PPA	0,265	0,321	0,407	0,336
GSD	0,144	0,240	0,246	0,218

Tabela 1: Absolutne vrednosti koeficientov korelacije med oceno kakovosti in razdaljo do referenčnega modela. Posamezna vrednost predstavlja povprečje koeficientov za vse vmesne modele med rekonstrukcijo.

### 6.2 Evalvacija načrtovanja pogledov

V tem poglavju predstavimo rezultate evalvacije načrtovanja naslednjega pogleda. Gradnjo modela z načrtovanjem pogledov (v nadaljevanju rekonstrukcija NBV) primerjamo z osnovno rekonstrukcijo pri različnih vrednostih parametra  $\rho$ . Poleg tega smo preverili tudi delovanje kombiniranega pristopa, pri katerem smo za prvih 20 slik uporabili enakomerno postavitev kamer (gostote  $\rho = 20$ ), nadaljnje slike pa smo zajeli z



Slika 10: Evalvacija rekonstrukcije modela kamna. Na zgornjem delu slike si od leve proti desni sledijo referenčni model, postavitev kamer pri rekonstrukciji NBV ter končni rekonstruirani model brez teksture in z njo. Spodnja grafa prikazujeta nenatančnost in pokritost v odvisnosti od števila dodanih pogledov (enako strukturo imajo tudi preostale slike tega podpoglavja).

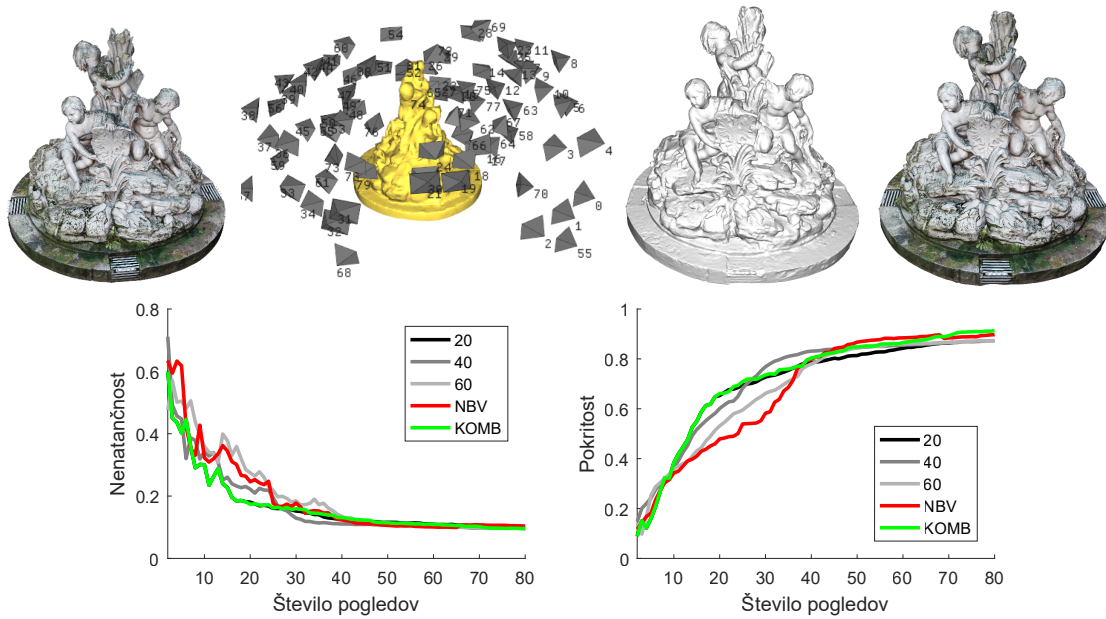
načrtovanjem pogledov (označeno kot rekonstrukcija KOMB). Zanima nas predvsem, kako se med gradnjo spreminjata nenatančnost in pokritost modela. Za primerjavo zgradimo 3 osnovne rekonstrukcije modelov z gostotami postavitve  $\rho = 20, 40, 60$ , ki predstavljajo redko, srednje gosto in gosto enakomerno postavitev kamer. V nekaterih primerih je vrednost  $\rho = 20$  premajhna za uspešno rekonstrukcijo, zato jo v takšnih primerih povečamo na  $\rho = 25$ . Za inicializacijo rekonstrukcije NBV smo uporabili prve tri poglede enakomerne postavitve, vsi naslednji pogledi pa so bili izbrani z našim algoritmom za načrtovanje naslednjega pogleda. V primeru, da najbolje ocenjeni pogled ni uspešno dodan (npr. neuspešna lokalizacija), se poskuša dodati pogled z drugo najboljšo oceno itd. V rezultatih poročamo o povprečnem indeksu izbranega pogleda, kjer indeks 0 pomeni, da so bile ustrezne vse slike, indeks 1 pa, da je bila za vsako izbrano sliko ena izpuščena itn.

Za začetek si pogledjmo primer, ki je prikazan na sliki 10. Z njim želimo prikazati večjo fleksibilnost uporabe načrtovanja pogledov v primerjavi s fiksno postavitvijo. Čeprav gre za sorazmerno preprost model z večinoma konveksno obliko, enakomerna postavitev ne more doseči polne pokritosti, saj spodnji del modela na kamerah ni viden v celoti. Najmanjša gostota, za katero je osnovna rekonstrukcija v tem primeru uspešna, je 25, poleg tega pa smo uporabili še gostoti 40 in 60. Nenatančnost je majhna pri vseh načinih rekonstrukcije (nekoliko slabša je  $\rho = 25$ , vendar ne bistveno), bolj

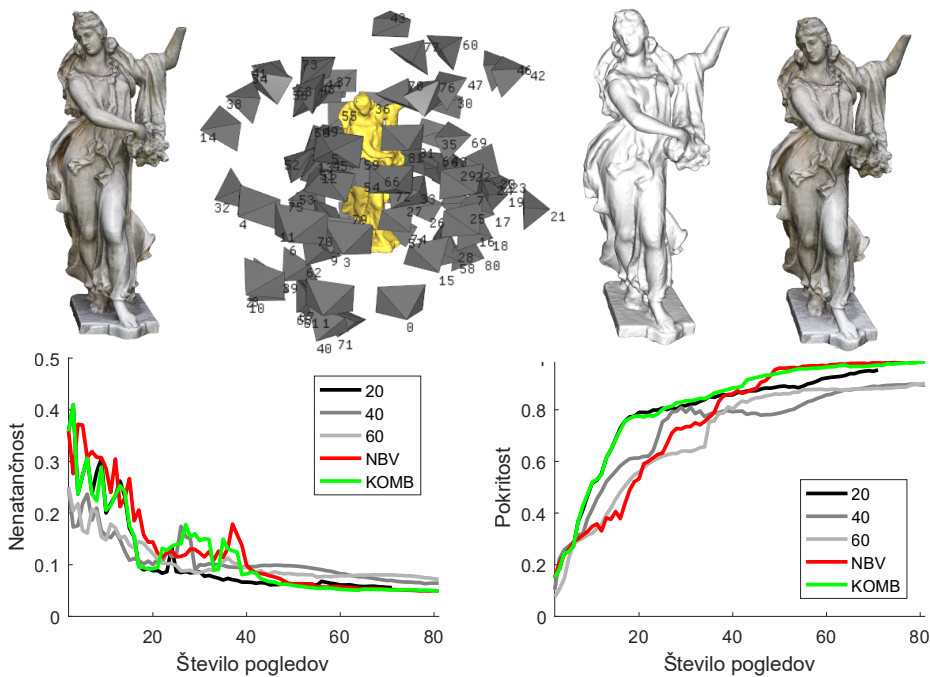
zanimiva je pokritost. Najprej lahko opazimo, da gostota kamer vpliva na hitrost naraščanja pokritosti, kar je v skladu s pričakovanji. Pri gostejši postavitvi moramo namreč dodati več pogledov, da pokrijemo model. Pokritost rekonstrukcije NBV na intervalu med 5. in 20. pogledom sicer narašča nekoliko počasneje (primerljivo z  $\rho = 60$ ), pozneje pa se približa polni pokritosti, saj so z uporabo načrtovanja pogledov kamere dodane tudi na spodnji strani modela. Povprečni indeks izbranega naslednjega pogleda je 0,15, kar pomeni, da je bil v veliki večini uspešno dodan prvi predlagani pogled. Rekonstrukcija KOMB v tem primeru združuje prednosti obeh pristopov in daje najboljše rezultate. Pokritost je skoraj celoten čas rekonstrukcije boljša kot pri preostalih pristopih. Izjema je vmesni del okoli 40. pogleda, kar je verjetno posledica naključnosti gručenja pri določanju naslednjega pogleda.

Naslednja modela sta bila izbrana zaradi njune nekoliko bolj kompleksne oblike. Z njima želimo prikazati robustnost naše metode za načrtovanje naslednjega pogleda. Kljub kompleksnejšim oblikam je enakomerna postavitev v teh primerih precej ugodna, saj so vidni vsi deli modelov na vsaj eni izmed kamer, poleg tega so vsi pogledi uspešno dodani tudi pri redki postavitvi kamer. Referenčna modela s pripadajočimi rezultati sta prikazana na slikah 11 in 12. Rezultati evalvacije so v teh dveh primerih podobni. Nekoliko bolj izstopa model kipa, pri katerem je končna nenatančnost rekonstrukcije NBV boljša od osnovne rekonstrukcije, čeprav je za





Slika 11: Evalvacija rekonstrukcije modela fontane.

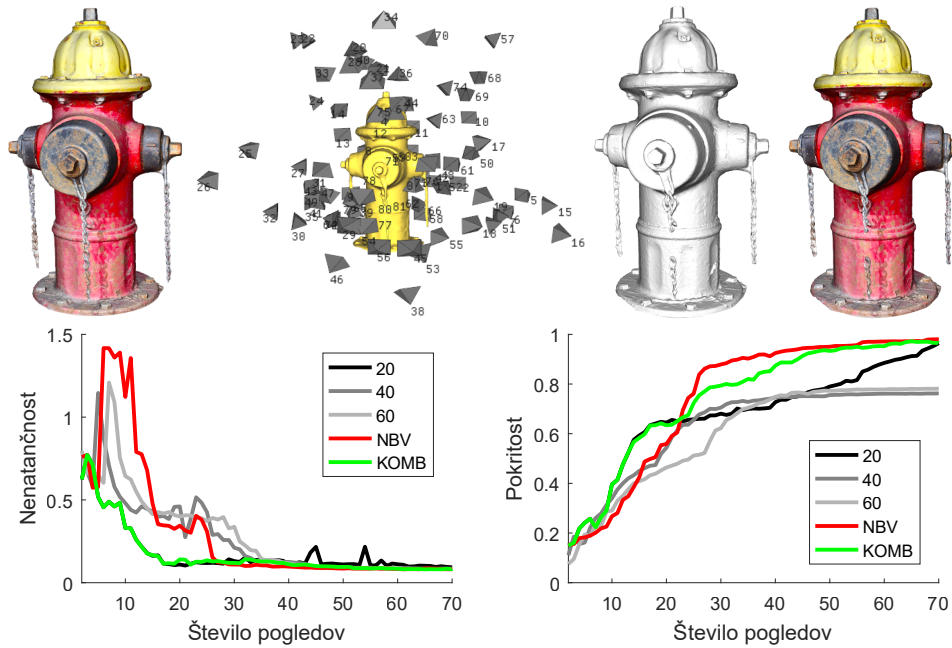


Slika 12: Evalvacija rekonstrukcije modela kipa.

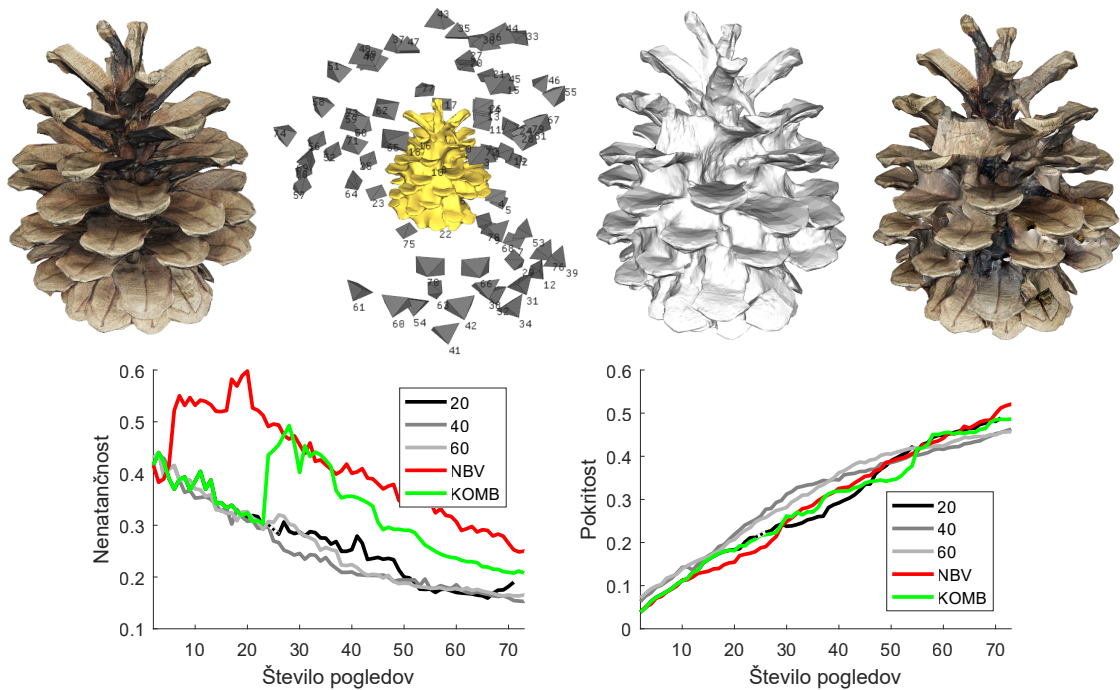
prvih 40 pogledov slabša. Enako situacijo opazimo tudi pri pokritosti, ki se v tem primeru približa 99 %. Obema primeroma je skupno, da pokritost med 10. in 40. pogledom nekoliko zaostaja za redkejšo osnovno rekonstrukcijo  $\rho = 20$  in je bolj primerljiva z gostejšo  $\rho = 60$ . Z uporabo kombiniranega pristopa je ta zaostanek izničen, končna pokritost pa je v obeh primerih (NBV in KOMB) boljša od osnovne rekonstrukcije. Povprečni indeks izbranega naslednjega pogleda je 1,87 pri fontani

in 1,48 pri kipu. Vrednosti so zaradi kompleksnejših modelov nekoliko večje kot v prejšnjem primeru.

Naslednji primer je prikazan na sliki 13. Kljub sorazmerno kompleksni geometriji modela je rekonstrukcija NBV bolj učinkovita od osnovne. Nenatančnost je na začetku sicer večja, vendar okoli 30. pogleda vrednost pade in postane primerljiva z osnovno rekonstrukcijo. Pokritost je od 20. pogleda dalje precej boljša in na koncu preseže 99 %. Pri kombiniranem pristopu začne



Slika 13: Evalvacija rekonstrukcije modela hidranta.



Slika 14: Evalvacija rekonstrukcije modela storža.

ob preklopu na načrtovanje pogledov pokritost hitro naraščati. To naraščanje je sicer počasnejše od NBV. Predvidevamo, da do te razlike pride zaradi naslednjega razloga: Ob preklopu so novi pogledi lahko pravilno lokalizirani le v okolici prvega obroča kamer, kar ima v tem primeru vidne posledice. Končna pokritost je kljub temu primerljiva. Povprečni indeks izbranega naslednjega pogleda je v tem primeru dober in znaša 0,1.

Za konec si pogledimo še primer, ki je za našo metodo načrtovanja pogledov zelo neugoden. Gre za kompleksen model s tanko geometrijo in majhnimi ravnimi površinami. Prikazan je na sliki 14. Čeprav je pokritost NBV primerljiva z osnovno, ta narašča počasi in pri 80. pogledu ne preseže 60 %. Nenatančnost je med rekonstrukcijo NBV ves čas precej večja od osnovne. Tudi kombiniran pristop v tem primeru ne izboljša

rezultatov. Povprečni indeks izbranega naslednjega pogleda je visok in znaša 3,48. Razlog za to sta ponovno velika kompleksnost modela in posledično neuspešna lokalizacija nekaterih kamer.

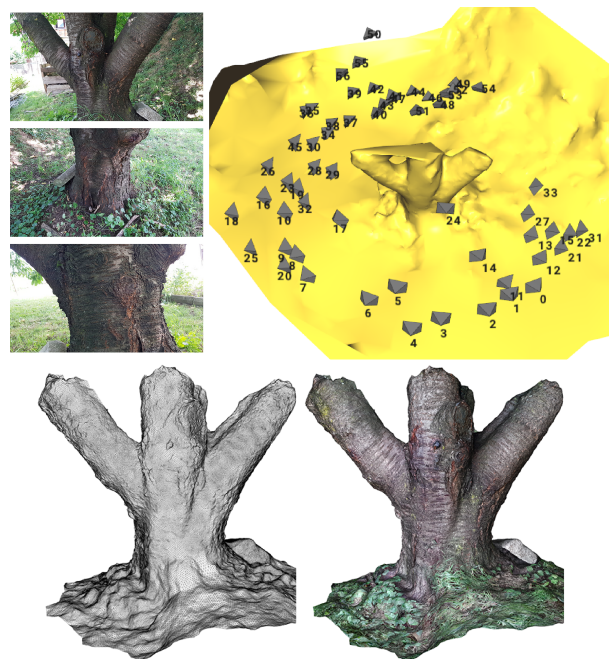
### 6.3 Primeri rekonstrukcij resničnih predmetov

V tem poglavju predstavimo še nekaj rekonstrukcij resničnih predmetov. Za inicializacijo rekonstrukcije smo po lastni presoji izbrali nekaj pogledov in s pomočjo programske opreme označili območje interesa. Pri dodajanju novih slik smo upoštevali predlagane poglede. Lokalizacija predlaganega pogleda je včasih neuspešna ali pa natančna postavitev kamere zaradi fizičnih omejitev prostora ni mogoča. V takšnih primerih smo izbrali naslednji predlagani pogled, v rezultatih pa poročamo o povprečnem indeksu izbranega pogleda ( $I_{nbv}$ ). Poleg tega poročamo še o absolutni ( $\epsilon_{abs}$ ) in relativni ( $\epsilon_{rel}$ ) napaki, ki jo naredimo pri postavitvi kamere. Absolutna napaka položaja kamere je izračunana na podlagi meritev resničnega predmeta, relativno napako pa izračunamo glede na razdaljo dveh najbolj oddaljenih kamer. Napako orientacije ( $\epsilon_{or}$ ) predstavlja kot med optičnima osema predlagane in resnične postavitve kamere. Ločljivost zajetih slik znaša  $2560 \times 1440$ , kamera pa je predhodno kalibrirana. Ker referenčni modeli tokrat niso na voljo, lahko kakovost rekonstrukcije ocenimo zgolj vizualno.

Prvi primer prikazuje rekonstrukcijo debla drevesa, za katero smo uporabili 57 slik. Na zgornjem delu slike 15 so prikazani postavitve kamer in nekaj primerov zajetih slik. Podatki o napakah pri postavitvi kamere so naslednji:  $I_{nbv} = 0,8$ ,  $\epsilon_{abs} = 20$  cm,  $\epsilon_{rel} = 6,5$  % in  $\epsilon_{or} = 13,6$  °C. Končni 3D-model, ki je rezultat goste rekonstrukcije po odstranitvi odvečnih trikotnikov, je prikazan na spodnjem delu slike 15. Geometrija debla se vizualno dobro ujema z njegovim resničnim videzom. Rekonstrukcija travnate površine in preostalih rastlin v okolici sicer ni naš cilj, opazimo pa lahko, da so ta področja bolj problematična in manj natančno rekonstruirana.

Za naslednji primer smo izbrali nekoliko manjši predmet, in sicer gorske čevlje. Pri zajemanju slik je bila rekonstrukcija površine 4-krat neuspešna. Ker je načrtovanje pogledov v takšnem primeru nemogoče, smo 4 slike dodali po lastni presoji. Skupno smo zajeli 31 slik. Podatki o napakah pri postavitvi kamere so naslednji:  $I_{nbv} = 0,48$ ,  $\epsilon_{abs} = 18$  cm,  $\epsilon_{rel} = 12,8$  % in  $\epsilon_{or} = 15,3$  °C. Končni 3D-model je prikazan na sliki 16. Na slabše teksturiranih predelih čevlja sicer lahko opazimo nekaj nepravilnosti, na splošno pa se geometrija vizualno dobro ujema z resničnim predmetom.

Rezultati rekonstrukcije zadnjega primera so prikazani na sliki 17. Čeprav je oblika predmeta navidezno preprosta, je zaradi ostrih robov in nekoliko manj ugodne teksture žoge ta primer bolj težaven za rekonstrukcijo. Zajeli smo 43 slik, od tega smo 3 slike morali dodati po lastni presoji, saj predlagani pogledi niso bili uspešno lokalizirani. Podatki o napakah pri postavitvi kamere so

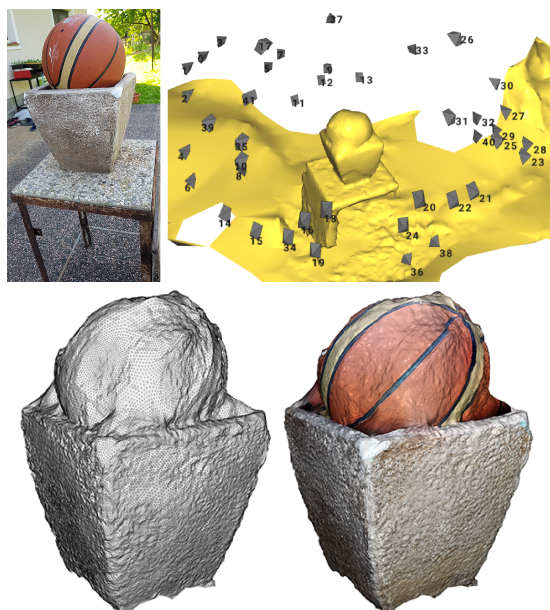


Slika 15: Nekaj primerov slik, uporabljenih pri rekonstrukciji (levo zgoraj), postavitev kamer (desno zgoraj) in rezultat goste rekonstrukcije debla (spodaj).



Slika 16: Nekaj primerov slik, uporabljenih pri rekonstrukciji (levo zgoraj), postavitev kamer (desno zgoraj) in rezultat goste rekonstrukcije gorskih čevljev (spodaj).

naslednji:  $I_{nbv} = 1,1$ ,  $\epsilon_{abs} = 15$  cm,  $\epsilon_{rel} = 7,7$  % in  $\epsilon_{or} = 15,7$  °C. Rekonstrukcija spodnjega dela (lonca) je zelo uspešna in se vizualno dobro ujema z resničnim predmetom. Več nepravilnosti lahko opazimo na žogi. Rekonstruirana površina je pregroba, oblika pa ni povsem sferična, kar je najbolj razvidno na zgornjem delu. Razlog za te nepravilnosti je manj ugodna tekstura.



Slika 17: Primer slike, uporabljene pri rekonstrukciji (levo zgoraj), postavitev kamer (desno zgoraj) ter rezultat goste rekonstrukcije lonca in žoge (spodaj).

## 7 ZAKLJUČEK

Prispevek našega dela sta zasnova in implementacija nove metode za načrtovanje najboljšega naslednjega pogleda, ki temelji na lastni meri za oceno kakovosti 3D-modela. Za delovanje potrebuje le trenutni 3D-model, predstavljen s trikotniško mrežo, in informacije o legi obstoječih kamer, zato je metoda dovolj splošna in je lahko uporabljena tudi kot gradnik pri avtonomni rekonstrukciji z različnimi roboti (npr. kvadrokopter, robotski manipulator itd.). Programska oprema, razvita v sklopu tega dela, je prosto dostopna\*. V delu [21] je na voljo še podrobnejši opis programske opreme in uporabljenih metod za rekonstrukcijo.

Rezultati evalvacije so pokazali uspešno delovanje našega sistema. Predlagana mera za oceno kakovosti na testiranih modelih deluje bolje od obstoječe mere. Pokazali smo tudi, da je kakovost rekonstrukcije z načrtovanjem naslednjega pogleda v večini primerov boljša od rekonstrukcije z enakomerno postavljenimi kamerami. V nasprotju s fiksno postavitvijo kamer naša metoda nima posebnih predpostavk o obliki in velikosti predmeta. Uspešno delovanje rekonstrukcije smo prikazali tudi na resničnih predmetih.

V nadaljnjem delu se bomo posvetili še nekaterim izboljšavam. Oceno kakovosti bi lahko izboljšali s pomočjo strojnega učenja. Pri načrtovanju naslednjih pogledov pa bi bilo v cenovno funkcijo smiselno vplejati dodatne informacije o videzu teksture in vidnosti rekonstruiranih značilnic. Tako bi bil proces gradnje 3D-modelov še bolj zanesljiv.

\*Dostopno na: <https://github.com/KristianZarn/Reconstruction>

## LITERATURA

- [1] M. Pollefeys, L. Van Gool, M. Vergauwen, F. Verbiest, K. Cornelis, J. Tops, R. Koch, Visual modeling with a hand-held camera, in: *IJCV* 59 (3) (2004) 207–232.
- [2] Y. Furukawa, C. Hernández, Multi-view stereo: A tutorial, *Foundations and Trends® in CGV* 9 (1-2) (2015) 1–148.
- [3] S. Kriegel, C. Rink, T. Bodenmüller, M. Suppa, Efficient next-best-scan planning for autonomous 3D surface reconstruction of unknown objects, *Journal of Real-Time Image Processing* 10 (4) (2015) 611–631.
- [4] E. Dunn, J.-M. Frahm, Next best view planning for active model improvement, in: *BMVC*, 2009, pp. 1–11.
- [5] C. Hoppe, M. Klopschitz, M. Rumpler, A. Wendel, S. Kluckner, H. Bischof, G. Reitmayr, Online feedback for structure-from-motion image acquisition, in: *BMVC*, Vol. 2, 2012, p. 6.
- [6] C. Wu, Towards linear-time incremental structure from motion, in: *3D Vision-3DV 2013*, IEEE, 2013, pp. 127–134.
- [7] D. G. Lowe, Distinctive image features from scale-invariant keypoints, in: *IJCV* 60 (2) (2004) 91–110.
- [8] H. Stewenius, C. Engels, D. Nistér, Recent developments on direct relative orientation, *ISPRS Journal of Photogrammetry and Remote Sensing* 60 (4) (2006) 284–294.
- [9] L. Kneip, D. Scaramuzza, R. Siegwart, A novel parametrization of the perspective-three-point problem for a direct computation of absolute camera position and orientation, *CVPR* 2011, pp. 2969–2976.
- [10] R. Szeliski, *Computer vision: algorithms and applications*, Springer Science & Business Media, 2010.
- [11] R. Arandjelović, A. Zisserman, Dislocation: Scalable descriptor distinctiveness for location recognition, in: *ACCV*, Springer, 2014, pp. 188–204.
- [12] M. Jancosek, T. Pajdla, Exploiting visibility information in surface reconstruction to preserve weakly supported surfaces, *International scholarly research notices* 2014.
- [13] H.-H. Vu, P. Labatut, J.-P. Pons, R. Keriven, High accuracy and visibility-consistent dense multiview stereo, *IEEE transactions on pattern analysis and machine intelligence* 34 (5) (2012) 889–901.
- [14] M. Waechter, N. Moehrle, M. Goesele, Let there be color! large-scale texturing of 3D reconstructions, in: *ECCV*, Springer, 2014, pp. 836–850.
- [15] S. M. Seitz, B. Curless, J. Diebel, D. Scharstein, R. Szeliski, A comparison and evaluation of multi-view stereo reconstruction algorithms, *IEEE*, 2006, pp. 519–528.
- [16] R. Jensen, et al. Large scale multi-view stereopsis evaluation, in: *CVPR*, 2014.
- [17] C. Sweeney, Theia multiview geometry library: Tutorial & reference, <http://theia-sfm.org>.
- [18] W. Changchang, SiftGPU: a GPU implementation of scale invariant feature transform (SIFT), <https://github.com/pitzer/SiftGPU> (2007).
- [19] J. L. Schönberger, J.-M. Frahm, Structure-from-motion revisited, in: *CVPR*, 2016.
- [20] Openmvs: open multi-view stereo reconstruction library, <https://github.com/cdcseacave/openMVS>.
- [21] Kristian Žarn, Gradnja 3D modelov predmetov iz barvnih slik z načrtovanjem najboljšega naslednjega pogleda, magistrsko delo, Univerza v Ljubljani, Fakulteta za računalništvo in informatiko, 2019.

**Kristian Žarn** je leta 2015 diplomiral in leta 2019 magistriral na Fakulteti za računalništvo in informatiko Univerze v Ljubljani. Trenutno je zaposlen v slovenskem IT-podjetju v raziskovalnem oddelku.

**Danijel Skočaj** je izredni profesor na Fakulteti za računalništvo in informatiko Univerze v Ljubljani. Je predstojnik Laboratorija za umetne vizualne spoznavne sisteme. Njegovi glavni raziskovalni interesi segajo na področja računalniškega vida, razpoznavanja vzorcev, strojnega učenja in spoznavne robotike.