

3D REKONSTRUKCIJA SCENE NA OSNOVU KRETANJA KAMERE**3D SCENE RECONSTRUCTION BASED ON CAMERA MOTION**Jana Kudus, *Fakultet tehničkih nauka, Novi Sad***Oblast – ELEKTROTEHNIKA I RAČUNARSTVO**

Kratak sadržaj – U ovom radu opisan je proces dobijanja 3D oblaka tačaka na osnovu skupa fotografija snimljenih upotrebom drona. Dat je odnos između 3D koordinata tačke i odgovarajućih 2D projekcija kao i opis *Structure-from-Motion* (SfM) algoritma. Takođe, diskutovan je pristup za prevazilaženje problema merenja na osnovu aerosnimaka upotrebom dobijenog 3D modela.

Ključne reči: 3D rekonstrukcija, *Structure-from-Motion*, aerosnimanje

Abstract – This paper describes the process of obtaining 3D point cloud using the set of photographs taken from a drone. Relationship between 3D coordinates of the point and corresponding 2D projections as well as a description of *Structure-from-Motion* (SfM) algorithm is given. Also, an approach for overcoming the problem of measurements based on aerial photography, by using obtained 3D model, is discussed.

Keywords: 3D reconstruction, *Structure-from-Motion*, aerial photography

1. UVOD

Fotogrametrija predstavlja tehniku dobijanja pouzdanih informacija o fizičkim objektima i životnoj sredini kroz proces snimanja, interpretacije i merenja fotografske slike. Neke od primena fotogrametrije jesu kreiranje topografskih karti, ili kreiranje 3D modela kao osnove za rešavanje brojnih inženjerskih i drugih naučnih problema.

Najjednostavniji način za dobijanje informacija o 3D prostoru na osnovu 2D fotografija jeste upotrebom stereo vizije. Stereo vizija zasniva se na upotrebi dve kamere, pomerene jedna u odnosu na drugu, sa ciljem da se objekat istovremeno fotografiše iz dva različita ugla. Identifikovanjem iste ključne tačke u dve dobijene fotografije i uz pomoć triangulacije moguće je dobijanje informacije o dubini tačke i njeno projektovanje u 3D prostor.

Ovaj postupak je unapređen upotrebom većeg broja kamera, koje omogućavaju prikupljanje većeg broja informacija o sceni. Algoritmi koji se koriste za dobijanje informacija o 3D prostoru na osnovu stereo korespondencija više od dve slike nazivaju se *Multiple View Stereo* (MVS) algoritmima.

NAPOMENA:

Ovaj rad proistekao je iz master rada čiji mentor je bio docent dr Branko Brkljač.

Međutim, kako upotreba većeg broja kamera podrazumeva i veću cenu, kao i kompleksniji postupak pripreme, ovakav način dobijanja 3D informacija nije praktičan za primenu izvan laboratorijskih uslova. Pravljenjem kompromisa u kvalitetu 3D rekonstrukcije, uz pomoć *Structure from Motion* (SfM) algoritama, sličan efekat moguće je postići i upotrebom samo jedne kamere, koja se pomena oko objekta i snima fotografije u različitim vremenskim trenucima.

Fotogrametrija zasnovana na SfM algoritmima takođe može da koristi i fotografije ili video zapise snimljene iz vazduha (eng. *aerial photogrammetry*) uz pomoć aviona, helikoptera ili bespilotnih letelica. Fotogrametrijska snimanja pod uglom ili vertikalno na površinu Zemlje pronašla su primenu u kartografiji, urbanističkom planiranju, arheologiji, poljoprivredi i mnogim drugim oblastima.

2. STRUCTURE-FROM-MOTION

Structure-from-motion (SfM) algoritam predstavlja fotogrametrijsku tehniku za prikupljanje informacija o 3D prostoru na osnovu 2D slika prikupljenih iz različitih pogleda i u različitim vremenskim intervalima. SfM algoritmi kao ulaz koriste skup fotografija i određuju parametre kamere za svaku ulaznu fotografiju i skup tačaka u 3D prostoru. Dakle, kao rezultat SfM algoritama, dobija se oblak tačaka (eng. *sparse reconstruction*), na osnovu kojeg se, u zavisnosti od potrebe, može dobiti gusta rekonstrukcija (eng. *dense reconstruction*), te dobiti predstava u vidu poligonalne mreže, skupa ravni ili realističnog rendera.

Najveći broj SfM algoritama prati iste korake:

1. Detekcija i deskripcija ključnih tačaka na svakoj od ulaznih slika
2. Pronalaženje podudarajućih ključnih tačaka na različitim slikama i generisanje liste 2D koordinata ključnih tačaka
3. Određivanje 3D koordinata ključnih tačaka na osnovu odgovarajuće liste 2D koordinata
4. Dorađivanje modela podešavanjem snopa zraka (eng. *bundle adjustment*)

2.1. Detekcija i deskripcija ključnih tačaka

Detekcija ključnih tačaka predstavlja proces selekcije tačaka koje definišu sadržaj slike, nezavisnih od orijentacije, perspektive, rezolucije i osvetljenja slike.

Ključne tačke tipično se određuju analiziranjem veličine i smera promene intenziteta u lokalnim regionima slike, u cilju detekcije ivica i ćoškova. Karakteristike lokalnih regiona u okolini svake ključne tačke nazivaju se deskriptorima ključnih tačaka.

Kada su u pitanju SfM algoritmi, jedan od najčešće upotrebljivanih algoritama za detekciju i deskripciju ključnih tačaka jesu *Scale-Invariant Feature Transform* (SIFT) algoritam.

2.2. Pronalaženje podudarajućih ključnih tačaka

Podudaranje ključnih tačaka odnosi se na pronalaženja istih ključnih tačaka na paru slika, na osnovu deskriptora slike dobijenih u prethodnom koraku, pri čemu se za svaku ključnu tačku dobija lista 2D koordinata preuzetih za svaku sliku na kojoj je ona vidljiva.

Najjednostavniji način pronalaženja podudarajućih ključnih tačaka na paru slika jeste upotrebom *Brute-Force* algoritama. Ovi algoritmi funkcionišu tako što se između deskriptora neke ključne tačke sa prve slike i deskriptora svih ključnih tačaka na drugoj slici računaju udaljenosti u prostoru obeležja. Kao rezultat, pronalazi se par ključnih tačaka, koje su na osnovu izračunatih udaljenosti najbliže u prostoru obeležja.

Međutim, kako SfM algoritmi kao ulaz tipično koriste više od dve fotografije, izazov jeste pronalaženje parova fotografija koje poseduju podudarajuće ključne tačke. Ovaj postupak može da se razlikuje u zavisnosti od toga da li je skup fotografija strukturiran ili nestruktuiran.

U slučaju kada je skup ulaznih fotografija nestruktuiran, predznanje o tome na kojim slikama postoji podudaranje nije dostupno, te svaka slika mora da prođe kroz proces pronalaženja podudarajućih ključnih tačaka sa svim ostalim slikama iz ulaznog skupa, odnosno radi se potpuna pretraga, što može biti veoma računski i vremenski zahtevno. Iz tog razloga, u slučaju ulaznih skupova koji poseduju više stotina fotografija, preporučeni pristup je pronalaženje podudarajućih tačaka samo na vizuelno najslabijim fotografijama.

U slučaju kada je ulazni skup fotografija strukturiran, postoji određeno predznanje o tome na kojim fotografijama postoje podudaranja, te ga je moguće iskoristiti kako bi se ovaj korak olakšao.

2.3. Određivanje 3D koordinata ključnih tačaka

Za dobijanje podataka o 3D prostoru neophodno je određivanje fundamentalne matrice, a zatim i odgovarajućih projekcionih matrica.

Fundamentalna matrica definisana je izrazom [1]:

$$x'Fx = 0 \quad (2)$$

Gde su x i x' podudarajuće ključne tačke u dve slike.

Matrica F ima 5 nenulatih elemenata, odnosno 4 stepena slobode, te se može zapisati i kao [2]:

$$F = \begin{bmatrix} 0 & 0 & a \\ 0 & 0 & b \\ c & d & e \end{bmatrix} \quad (3)$$

Iz navedenih izraza moguće je zaključiti da je poznavanjem dovoljnog broja parova tačaka moguće odrediti elemente fundamentalne matrice. Minimalni broj potrebnih korespondencija je $n = 4$, odnosno potrebno je ukupno 8 ključnih tačaka za određivanje fundamentalne matrice. Međutim, u slučaju postojanja više od četiri podudaranja, uz uvođenje pretpostavke da na slikama

postoji šum, kao i pogrešna podudaranja ključnih tačaka, rešenje nije jednoznačno, te je potrebno pronaći najverovatnije rešenje (eng. *Maximum Likelihood*, ML). ML rešenje zahteva upotrebu algoritma otpornog na pogrešna merenja, kao što je to na primer, *Random Sample Consensus* (RANSAC) algoritam. U ovom slučaju optimalna matrica F može se odrediti minimizacijom funkcije [2]:

$$\sum_i d(x_i, F^{-1}x'_i)^2 + d(x'_i, Fx_i)^2 \quad (4)$$

Određivanje koordinata 3D tačke na osnovu datih 2D koordinata naziva se triangulacijom. Kako bi triangulacija bila moguća, nakon određivanja fundamentalne matrice F , potrebno je odrediti par projekcionih matrica dve kamere. Uz pretpostavku da je tačka x'_t , koja se nalazi na slici t sa odgovarajućom projekcionom matricom P_t , projekcija nepoznate tačke u 3D prostoru X , uslov koji ona mora da zadovolji dat je sledećim izrazom [2]:

$$x'_t \times P_t X = S_{x'_t} P_t X = 0 \quad (5)$$

Gde je $S_{x'_t}$ kososimetrična matrica koordinata tačke x'_t .

Međutim, uzimajući u obzir nesavršenost merenja, ovaj uslov nije moguće zadovoljiti za sve projekcije istovremeno. Iz tog razloga, traži se približno rešenje X kroz optimizaciju podešavanjem snopa zraka.

2.4. Podešavanje snopa zraka

Optimalno rešenje triangulacije podrazumevalo bi pronalazak zraka koji prolaze kroz podudarne tačke na različitim fotografijama, a zatim pronalazak njihovog preseka u 3D prostoru koji definiše tačku X . Međutim, uzimajući u obzir nesavršenost merenja, potrebno je u obzir uzeti stohastičke osobine podudarnih tačaka u cilju pronalaženja optimalne tačke preseka snopa zraka u 3D. Prilikom rešavanja problema rekonstrukcije, rešava se naredni problem: na osnovu koordinata podudarnih ključnih tačaka dobijenih iz skupa ulaznih slika (koordinate x'_j), odrediti skup matrica kamere P^i i 3D tačaka X_j , takvih da je zadovoljena jednačina: $P^i X_j = x'_j$. Međutim, zbog šuma prisutnog u merenjima, ova jednačina nikada neće biti u potpunosti zadovoljena za sve tačke, te se traži približno ML rešenje koje istovremeno estimira matrice kamere i 3D tačke minimizovanjem greške projekcije [2]:

$$\min_{\hat{P}^i, \hat{X}_j} \sum_{ij} d(\hat{P}^i \hat{X}_j, x'_j)^2 \quad (6)$$

3. MERENJA NA OSNOVU AEROSNIMAKA

Za određivanje fizičkih dimenzija piksela u slučaju aerosnimaka neophodno je poznavanje visine sa koje je snimak formiran u odnosu na teren. Dron visinu može da beleži na dva načina:

1. Upotrebom GPS
2. Upotrebom barometra

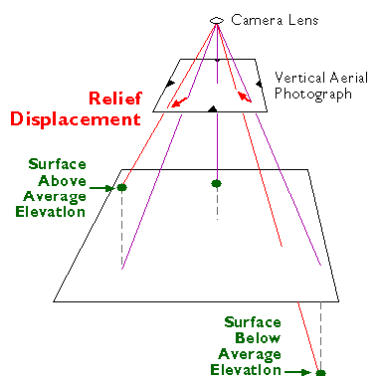
Međutim, visina izmerena upotrebom ovih senzora ne predstavlja visinu u odnosu na teren. Visina dobijena upotrebom GPS predstavlja visinu u odnosu na referentni elipsoid, dok visina dobijena upotrebom barometra odgovara visini u odnosu na početnu tačku snimanja.

Dakle, za određivanje stvarne visine drona u odnosu na teren potrebno je poznavanje digitalnog modela elevacije (eng. *Digital Elevation Model*, DEM) ili digitalnog modela površine (eng. *Digital Surface Model*, DSM). DEM i DSM sadrže informacije o stvarnoj visini terena na svakoj od posmatranih lokacija, pri čemu DEM predstavlja visinu površine Zemlje ne uzimajući u obzir rastinje, građevine i druge objekte, dok DSM modeluje površinu Zemlje, kao i sve objekte na njoj.

Dobijanje DSM moguće je na osnovu višestrukih slika terena, korišćenjem prethodno opisanih SfM algoritama. Tačnije, ovo predstavlja prvi korak kreiranja ortomozaika – aerosnimka na osnovu kojeg je moguće vršiti merenja.

3.1. Kreiranje ortomozaika

Fizička veličina piksela zavisi od visine sa koje su oni formirani. Dakle, razlike u visini terena utiču na razmeru fotografija. Dodatno, efekat koji se javlja kao posledica razlike u visini terena u odnosu na centar aerosnimka naziva se reljefnim pomerajem. Kao što se može uočiti na slici 1, objekti koji poseduju veću visinu su na aerosnimku pomereni od centra slike, dok su objekti koji su na nižoj visini pomereni ka centru slike. Iz ovog razloga aerosnimci nisu pogodni za merenje udaljenosti ili dimenzija objekata bez prethodne obrade.



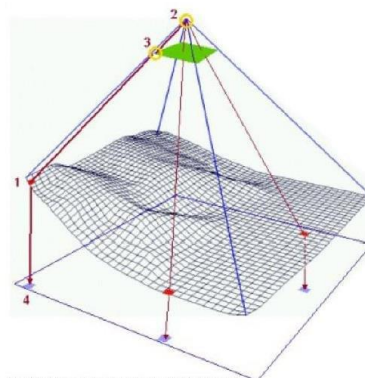
Slika 1. Reljefni pomeraj [3]

Ortofotografija, stvarna ortofotografija i ortomozaik predstavljaju geometrijski korigovane aerosnimke, te se za razliku od aerosnimka mogu koristiti za merenje.

Ortofotografija i stvarna ortofotografija razlikuju se po tome što se prilikom procesa formiranja ortofotografije koristi DEM, dok se pri formiranju stvarne ortofotografije koristi DSM. Međutim, kako su tradicionalni pristupi formiranja DSM bili veoma skupi, formiranje stvarne ortofotografije nije bilo široko zastupljeno. Način za prevazilaženje ovog problema jeste korišćenje SfM tehnika za 3D rekonstrukciju kojima se na osnovu generisanog DSM omogućava ortorektifikacija snimaka i generisanje ortomozaika. Primena SfM na skup slika čije geometrijsko korigovanje je potrebno izvršiti, proces je učinilo značajno jeftinijim i dostupnijim.

Na slici 2 ilustrovan je proces ortorektifikacije jedne slike prilikom formiranja ortomozaika. Za svaki korigovani piksel, označen sa „(4)“ u diskretnoj mreži, odgovarajuća tačka DSM (1) projektuje se na odgovarajuću lokaciju na ravni slike (3) na osnovu poznatih parametara kamere (2). Iz tog razloga opisana operacija se još naziva i projekcijom unazad (eng. *backprojection*). Vrednosti

intenziteta novog piksela (4) moguće je dobiti na osnovu interpolacije vrednosti intenziteta piksela u okolini pozicije (3), jer u opštem slučaju piksel (3) jednoznačno ne odgovara ni jednoj diskretnoj vrednosti u novoj slici.



Slika 2. Proces formiranja ortomozaika [4]

U poređenju sa aerosnimkom, na ortorektifikovanom snimku nije prisutna distorzija prouzrokovana različitim visinama terena, čime je uklonjen i problem različitog skaliranja objekata koji se nalaze na različitim visinama, te je omogućeno tačno merenje na osnovu slika.

4. FORMIRANJE SKUPA PODATAKA ZA FOTOGRAMETRIJSKU 3D REKONSTRUKCIJU

Prikupljanje skupa fotografija je korak koji ima najznačajniji uticaj na krajnji ishod 3D rekonstrukcije. Neadekvatno podešeno osvetljenje, okruženje i kamera mogu dovesti do nedostajućih ili pogrešno izmerenih vrednosti, što rezultuje manje tačnim modelom. Međutim, primene na otvorenom ne dozvoljavaju potpuno kontrolisanje uslova. Ipak, prilikom formiranja skupa podataka, važno je u obzir uzeti preporuke, te ih poštovati u meri u kojoj to dozvoljava data primena.

4.1. Preporuke za formiranje skupa podataka za fotogrametrijsku 3D rekonstrukciju

Pri formiranju skupa podataka, pre svega je važno utvrditi da je li je objekat dobar kandidat za fotogrametrijsku 3D rekonstrukciju. Određeni tipovi površina, kao što su metalne ili mokre površine, koje imaju jaku refleksiju, otežavaju pronalaženje podudarnih tačaka na različitim fotografijama. Isto važi i za transparentne površine, kao što su staklo ili tečnosti, ili za objekte koji se kreću. Takođe, objekti sa uniformnim površinama nisu naročito dobri kandidati za fotogrametrijsku 3D rekonstrukciju zbog manjka ključnih tačaka. Međutim, projektovanjem različitih šablona na objekat, moguće je postići veštačko formiranje potrebnih ključnih tačaka.

Broj fotografija koji se nalazi u skupu može značajno da varira u zavisnosti od veličine, oblika i tipa objekta koji je predmet rekonstrukcije. Međutim, kako će posedovanje većeg broja fotografija na kojima je vidljiv isti deo objekta iz više različitih uglova povećati verovatnoću da objekat bude tačno reprezentovan, preporuka je formiranje što većeg broja fotografija. Takođe, preporučeno je minimalno preklapanje susednih ili uzastopno načinjenih fotografija od najmanje 60% - 70%.

Kada je objekat previše taman, pronalaženje ključnih tačaka je otežano, a samim tim i 3D rekonstrukcija. Takođe, senke mogu prouzrokovati pojavu tamnih tačaka na modelu ukoliko su vidljive na većini fotografija. Prema tome, veoma važan uticaj na konačan ishod rekonstrukcije ima adekvatno osvetljenje, koje bi trebalo da je nepromenljivo tokom čitavog procesa akvizicije skupa podataka, kao i što manja promena i prisustvo senki, što je teško kontrolisati u primenama na otvorenom. Iz tog razloga, pri formiranju skupa podataka, preporučuje se izbegavanje sunčanih dana zbog izraženijih senki, ali i delimično oblačnih dana zbog čestih promena osvetljenja. Takođe, preporučuje se izbegavanje kiše i snega, zbog mogućeg izazivanja refleksije površina, kao i vetra koji može prouzrokovati pomeranja u sceni. Dakle, preporučuje se formiranje skupa fotografija po oblačnom danu bez kiše.

4.2. Posmatrani skup podataka

U ovom radu, posmatran je skup aerosnimaka dobijen upotrebom *DJI P4 Multispectral* drona. Skup podataka sastoji se iz 59 slika. Slike su snimljene sa preklapanjem od 70 %, pri čemu je prosečna visina letenja u odnosu na početnu tačku iznosila 90 m. Na slici 3 prikazano je nekoliko susednih (uzastopno načinjenih) fotografija iz korišćenog skupa.



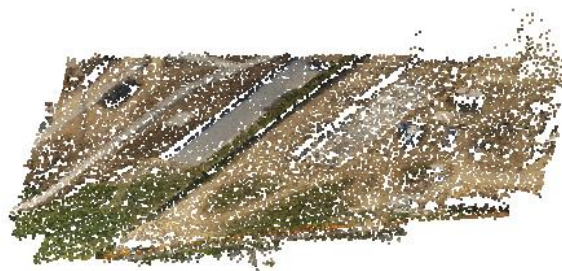
Slika 3. Primer susednih fotografija iz načinjenog skupa

5. REZULTAT

Kako je fotogrametrija pronašla primenu u mnogim oblastima, danas postoji veliki broj dostupnih alata za 3D rekonstrukciju na osnovu slika. Neki od najpopularnijih alata su: *COLMAP*, *Meshroom*, *MicMac*, *Multi-View Environment*, *OpenMVG*, *PIX4D*, *Agisoft Metashape*, itd.

U ovom radu poređeni su rezultati 3D rekonstrukcije oblaka tačaka dobijeni korišćenjem nekoliko navedenih biblioteka. Radi prikaza, na slici 4 i slici 5 formiran je 3D oblak tačaka na osnovu skupa fotografija opisanog u poglavlju 4.2. primenom softvera [5]. Tom prilikom izabran je *Frame camera* model, kako bi bilo omogućeno modelovanje distorzije ulaznih podataka. Kako unutrašnji parametri kamere nisu prethodno uneti, korišćena je automatska estimacija na osnovu ulaznog skupa podataka. Nakon toga, izvršeno je formiranje 3D oblaka tačaka na osnovu početnih podešavanja. Dobijeni 3D oblak tačaka prikazan je na slici 4.

Na slici 5 prikazan je 3D model dobijen dodatnom interpolacijom oblaka tačaka na slici 4.



Slika 4. 3D oblak tačaka rekonstruisan primenom [5]



Slika 5. 3D model dobijen upotrebom [5]

6. ZAKLJUČAK

Izborom adekvatnih detektora i deskriptora ključnih tačaka moguće je uticati na kvalitet 3D rekonstrukcije, kao i na potrebno vreme za obradu snimaka. U tom smislu istraživanja su usmerena ka postizanju rekonstrukcije u realnom vremenu, koja je prikladnija za primene koje zahtevaju brzo dobijanje informacija o karakteristikama scene i orijentaciju na terenu. U datim situacijama SfM tehnike 3D rekonstrukcije obezbeđuju način za efikasno generisanje 2D ortomozaika i ispravnu interpretaciju i merenja na osnovu aerosnimaka, za razliku od alata, npr. [6], koji ne podrazumevaju proces ortorektifikacije i ne obezbeđuju potreban kvalitet fotointerpretacije.

7. LITERATURA

- [1] W. Förstner, B. Wrobel, "Photogrammetric Computer Vision: Statistics, Geometry, Orientation and Reconstruction", Springer, 2016
- [2] R. Hartley, A. Zisserman, "Multiple View Geometry in Computer Vision", Cambridge UP, 2004
- [3] <https://www.e-education.psu.edu/natureofgeoinfo/book/export/html/1815> (pristupljeno 01.2022)
- [4] L. Abrahamczyk, J. Schwatz, "Forecast Engineering: From Past Design to Future Decision", Bauhaus-Universität Weimar, Fakultät Bauingenieurwesen, Institut für Konstruktiven Ingenieurbau, 2017
- [5] <https://www.agisoft.com/> (pristupljeno 01.2022)
- [6] <https://verticalphotoplacer.github.io/VerticalPhotoPlacer/> (pristupljeno 01.2022)

Kratka biografija:



Jana Kudus rođena je u Somboru 1998. god. Diplomski rad na Fakultetu tehničkih nauka iz oblasti Biomedicinsko inženjerstvo – Prepoznavanje oblika odbranla je 2020.god. kontakt: janakudus1998@gmail.com