

SVEUČILIŠTE U ZAGREBU  
FAKULTET ELEKTROTEHNIKE I RAČUNARSTVA

ZAVRŠNI RAD br. 5118

# **Gusta stereoskopska rekonstrukcija poluglobalnim podudaranjem**

Nikola Bunjevac

Zagreb, svibanj 2017.

*Umjesto ove stranice umetnite izvornik Vašeg rada.  
Da bi ste uklonili ovu stranicu obrišite naredbu \izvornik.*



# SADRŽAJ

<b>1. Uvod</b>	<b>1</b>
<b>2. Stereoskopska rekonstrukcija</b>	<b>3</b>
2.1. Epipolarna geometrija . . . . .	3
2.2. Rektifikacija . . . . .	3
<b>3. Lokalne metode</b>	<b>5</b>
3.1. Opći algoritam . . . . .	5
3.2. SSD . . . . .	6
3.3. ZSAD . . . . .	6
3.4. Census . . . . .	6
3.5. Korespondencija neovisna o uzorkovanju (Birchfield-Tomasi) . . .	6
3.6. Usporedba lokalnih i globalnih metoda . . . . .	6
<b>4. Poluglobalno podudaranje</b>	<b>7</b>
<b>5. Rezultati</b>	<b>8</b>
<b>6. Zaključak</b>	<b>9</b>
<b>A. Upute za korištenje</b>	<b>10</b>

# 1. Uvod

Računalni vid iznimno je zanimljivo interdisciplinarno područje koje se svrstava kao grana umjetne inteligencije. Ono se bavi omogućavanjem računalima da shvate i interpretiraju podatke iz digitalnih slika i videozapisa. Vid je jedno od najznačajnijih osjetila jer pomoću njega primamo mnoštvo korisnih informacija. Pomoću vida se orijentiramo u prostoru, prepoznajemo druge ljude i njihova lica, čitamo, prikupljamo informacije itd. Kada bi računala mogla interpretirati vizualne podražaje poput ljudi, to bi omogućilo velik napredak u umjetnoj inteligenciji. Nažalost, to je još uvijek neriješen problem.

Cijelo područje je nastalo šezdesetih godina prošlog stoljeća kada se mislilo da će se taj problem relativno brzo riješiti. Naime, profesor je kao zadatak studentu zadao da na robota stavi kameru i da robot opisuje ono što vidi. Vrlo brzo se pokazalo kako je problem puno složeniji nego što se mislilo. S druge strane, od tada su ostvareni veliki napretci u tom području kao i općenito u umjetnoj inteligenciji.

U ovom radu ćemo opisati neke metode za stvaranje rekonstrukcije prostora iz slika dviju kamera. Takvi sustavi se nazivaju stereo sustavi, a takva vrsta vida stereo vid. Oni funkcioniraju analogno ljudskom vidu koji se još naziva binokularni vid zbog toga što ljudi gledaju pomoću dva oka. To ljudima omogućava stvaranje vrlo kvalitetnog dojma o 3D osobinama prostora kojeg promatraju. Mogu zaključiti koliko je nešto udaljeno, odnos veličina raznih predmeta itd.

Gusta stereoskopska rekonstrukcija pokušava iz dvaju slika rekonstruirati gusti oblak točaka koji odgovara prostoru koji se promatra. Postupak se odvija u nekoliko koraka. Prvo je potrebno obraditi slike kako bi se otklonile fizičke nesavršenosti sustava poput nesavršenosti leće i položaja senzora. Zatim je potrebno piksele transformirati kako bi svi korespondentni pikseli ležali na istom pravcu, odnosno epipolarnoj liniji. Nakon toga se može krenuti u rekonstrukciju scene.

Postoje dvije glavne podjele metoda guste stereoskopske rekonstrukcije, a to su

1. lokalne metode i
2. globalne metode.

Lokalne metode se temelje na promatranju lokalne okoline svakog pojedinog piksela koju ćemo nazvati prozorčić. Postupak rekonstrukcije tada se svodi na traženje najslbličnijeg prozorčića na drugoj slici. Važno je naglasiti kako se traženje vrši samo duž epipolarne linije. Za razliku od lokalnih, globalne metode rekonstrukciju rade pomoću svih piksela slike što intuitivno dovodi do zaključka da je postupak složeniji i zahtijeva više računalnih resursa.

Metoda poluglobalnog podudaranja, koju ćemo detaljnije obrađivati u ovom radu, može se svrstati negdje između. Naime, ona kao ulaz prima izračunate korespondencije lokalnih metoda, a zatim radi optimizaciju nad pikselima duž iste linije u osam ili šesnaest smjerova. Takva optimizacije dovodi do kvalitetnijih rezultata u dijelovima gdje dolazi do nagle promjene u dubini scene.

U sljedećem poglavlju ćemo ukratko objasniti postupke kalibracije stereo sustava te rektifikacije parova slika kao predradnja postupku rekonstrukcije.

U trećem poglavlju ćemo obraditi nekoliko lokalnih metoda guste stereoskopske rekonstrukcije. Lokalne metode nude relativno dobre rezultate uz prihvatljivu performansu.

Zatim ćemo definirati algoritam poluglobalnog podudaranja, razraditi njegove hiperparametre te vidjeti kakva poboljšanja nudi u odnosu na lokalne metode.

U petom poglavlju ćemo vidjeti rezultate eksperimenata na standardnim skupovima slika, usporedbe obrađenih metoda i komentare.

U Dodatku A iza zaključka se nalaze kratke upute za korištenje implementiranih programa i skripti za testiranje ispitnih slijedova.

## 2. Stereoskopska rekonstrukcija

Prije same rekonstrukcije, potrebno je obraditi slike dobivene iz para kamera. Za to je potrebno znati geometrijske parametre sustava dvaju kamera. Kao model kamere uzet ćemo projekcijsku ravninu koja je udaljena od izvora/kamere za neku udaljenost  $d$ . Tako se sve točke u prostoru ispred kamere i ravnine projiciraju na tu zamišljenu ravninu.

### 2.1. Epipolarna geometrija

Geometrijske parametre sustava kamera možemo podijeliti u dvije skupine:

1. intrinzični parametri
2. ekstrinzični parametri.

Intrinzični parametri su svojstveni svakoj kameri pojedinačno, dok su im ekstrinzični parametri zajednički. Neka svojstva koja utječu na intrinzične parametre su nesavršenost leće (radijalna distorzija), pomak senzora od centra leće (tangencijalna distorzija) i slične fizičke nesavršenosti koje je ponekad nemoguće izbjeći. Stoga priskaćemo programskom otklanjanju tih nedostataka. Ekstrinzični parametri dovode do transformacija kako bismo slike doveli u istu ravninu projekcije te postigli da svi pikseli duž horizontalnog pravca jedne slike odgovaraju istom pravcu (na istoj visini) na drugoj slici. Taj pravac se naziva epipolarna linija. Razlog takvih transformacija je olakšavanje traženja korespondentnih točaka koje sada treba tražiti samo duž epipolarne linije čime je problem sveden na jednu dimenziju. Vidimo da uz malo pretprocesiranja puno olakšavamo problem rekonstrukcije.

## 2.2. Rektifikacija

Postupak rektifikacije odgovara transformaciji slika kako bismo dobili slike koje zadovoljavaju gore navedena svojstva, a to su:

1. Pikseli koji odgovaraju točkama u prostoru na istoj visini nalaze se duž iste epipolarne linije
2. Nema distorzije uzrokovane lećom
3. Projekcijske ravnine leže u istoj ravnini

Za rektifikaciju su potrebni ekstrinzični parametri sustava kamera, a za izračun ekstrinzičnih parametara su potrebni intrinzični. Nakon provedenih transformacija intrinzični parametri obaju kamera su jednaki (npr. imaju istu žarišnu udaljenost što je lako predočiti kada znamo da projekcije sada leže u istoj ravnini te je svaka točka u prostoru ispred jednako udaljena od oba projekcijska platna).

Ovaj postupak se obično provodi uz pomoć šahovske ploče koja je pogodna za to zbog svojih svojstava. Potrebno je parom kamera uslikati dovoljan broj slika na kojima se u potpunosti vidi šahovnica. Nakon toga se mogu odrediti ranije navedeni parametri. Postupak se odvija u nekoliko koraka, od kojih je prvi traženje šahovskih polja. Nakon toga se pomoću kuteva iz uglova polja može odrediti distorzija i ostali parametri.

Navedene postupke implementira i nudi biblioteka OpenCV <sup>1</sup>. Korisno je istaknuti kako je u korištenim skupovima slika postupak kalibracije i rektifikacije već proveden.

---

<sup>1</sup><http://opencv.org/>



## 3. Lokalne metode

Metode guste stereoskopske rekonstrukcije se mogu podijeliti na lokalne i globalne. U ovom poglavlju ćemo načelno opisati takve metode te detaljnije obraditi neke koje su korištene u ovom radu.

Lokalne metode se temelje na principu prozorčića ili okna fiksne širine oko piksela. Označimo širinu prozora s  $w$ .

Postavlja se pitanje što napraviti na rubovima slike gdje nije moguće izračunati sve vrijednosti. Nekoliko je mogućih rješenja:

- računati samo piksele koje je moguće obuhvatiti unutar slike
- proširiti sliku crnim okvirom debljine  $w/2$
- uopće ne računati korespondenciju za rub slike.

Svaki pristup ima svoje mane i prednosti. U vlastitoj implementaciji je korišteno treće rješenje gdje se uopće ne računa korespondencija, radi jednostavnosti. Razlika u pogrešci je relativno mala, a u skupu slika KITTI <sup>1</sup> su rubovi uglavnom i isključeni iz provjere točnih dispariteta.

### 3.1. Opći algoritam

Sada ćemo opisati općeniti algoritam za lokalno podudaranje.

za svaki piksel  $(x, y)$  lijeve slike:

za svaki disparitet  $0 \leq d \leq \text{maksimalni\_disparitet}$ :

$\text{cijena}[y][x] = f(x, y, d)$

pri čemu je funkcija  $f(x, y, d)$  funkcija korespondencije koja određuje cijenu razlike odgovarajućih piksela lijeve i desne slike. Svaki piksel lijeve slike s pozicijom  $(x, y)$  uspoređuje se s pikselom desne slike na poziciji  $(x - d, y)$ , gdje je  $d$  disparitet. Vidimo kako se odgovarajući pikseli nalaze na istoj  $y$  koordinati što

---

<sup>1</sup>[http://www.cvlibs.net/datasets/kitti/eval\\_scene\\_flow.php?benchmark=stereo](http://www.cvlibs.net/datasets/kitti/eval_scene_flow.php?benchmark=stereo)

je posljedica rektifikacije slika. Stoga je potrebno razmatrati samo piksele duž epipolarne linije.

## 3.2. SSD

Prva metoda koju ćemo opisati je SSD (eng. *Sum of Squared Differences*) koja se temelji na zbroju kvadrata razlika intenziteta piksela. Matematički ju možemo opisati sljedećim izrazom:

$$C_{SSD}(x, y, d) = \sum_i (I_R(x - d, y) - I_L(x, y))^2$$

gdje su  $(x, y)$  koordinate piksela lijeve slike,  $d$  trenutni disparitet, a  $I_L$  lijeva slika te  $I_R$  desna slika. Zbroj ide po svim pikselima unutar prozora promatranog piksela. Najbolja podudarnost kod SSD-a postiže se kada je vrijednost funkcije cijene jednaka 0. Lako je vidjeti zašto je tomu tako kada bismo usporedili dva prozora s istim intenzitetima piksela. Tada će vrijednost funkcije biti 0 jer oduzimamo iste vrijednosti. Iz toga slijedi da što se intenziteti više razlikuju, to će i cijene biti veće. Također, kvadrat doprinosi tome da vrijednosti ne budu negativne tako da možemo raditi samo s pozitivnim brojevima.

## 3.3. ZSAD

Sljedeća metoda je ZSAD (eng. *Zero Sum of Absolute Differences*). Matematički ju možemo definirati ovako:

$$C_{ZSAD}(x, y, d) = \sum_i [(M_L - I_L(x, y)) - (M_R - I_R(x - d, y))]$$

gdje je  $I_L$  lijeva slika,  $I_R$  desna slika, a  $M_L$  i  $M_R$  srednje vrijednosti intenziteta piksela u trenutnom prozoru lijeve, odnosno desne slike.

### 3.4. Census

### 3.5. Korespondencija neovisna o uzorkovanju (Birchfield-Tomasi)

### 3.6. Usporedba lokalnih i globalnih metoda

U ovom odjeljku dana je usporedba lokalnih i globalnih metoda podudaranja.

**Tablica 3.1:** Svojstva lokalnih i globalnih metoda

Lokalne metode	Globalne metode
cijene temeljene na prozorčićima oko svakog pojedinog piksela	cijene temeljene na svim pikselima slike
odabir pojedinog dispariteta po principu minimalne agregirane cijene	minimizacija globalne funkcije kroz vrijednosti svih piksela slike
značajno jednostavnije za implementaciju	teže za implementaciju
u pravilu brže	u pravilu sporije

U sljedećem poglavlju ćemo uvesti algoritam poluglobalnog podudaranja koji kao cijene prima rezultate lokalne korespondencije, a optimizaciju radi nad svim pikselima u određenim smjerovima. Time ga, kao što mu i naziv govori, možemo svrstati negdje između jer koristi lokalne korespondencije, a optimira (točnije, aproksimira) globalno.

## 4. Poluglobalno podudaranje

## 5. Rezultati

## 6. Zaključak

Zaključak.

# Dodatak A

## Upute za korištenje

## **Gusta stereoskopska rekonstrukcija poluglobalnim podudaranjem**

### **Sažetak**

Sažetak na hrvatskom jeziku.

**Ključne riječi:** računalni vid, stereoskopska rekonstrukcija, poluglobalno podudaranje, disparitet, korespondencija

## **Dense stereoscopic reconstruction with semi-global matching**

### **Abstract**

Abstract.

**Keywords:** computer vision, stereoscopic reconstruction, semi-global matching, disparity, correspondence