Napredno estimiranje strukture i gibanja kalibriranim parom kamera

Ivan Krešo

Mentor: Siniša Šegvić

3. srpnja 2013.

Motivacija

Stereo vid

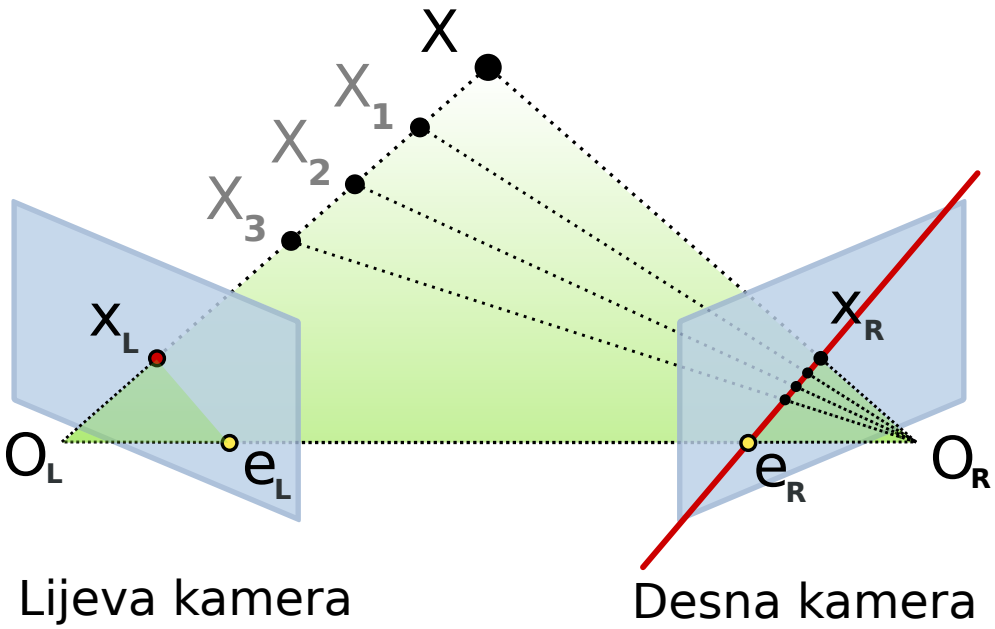
- dvije kamere omogućavaju mjerenje dubine korespondentnih točaka slike
- rekonstrukcija strukture scene
- određivanje gibanja kamere iz videa (vizualna odometrija)

Primjene u okviru projekta Vista

- položaj kamere u odnosu na ravninu ceste
- položaj vozila u odnosu na prometni trak (rub ceste)
- geometrijske indicije za položaj prometnog znaka

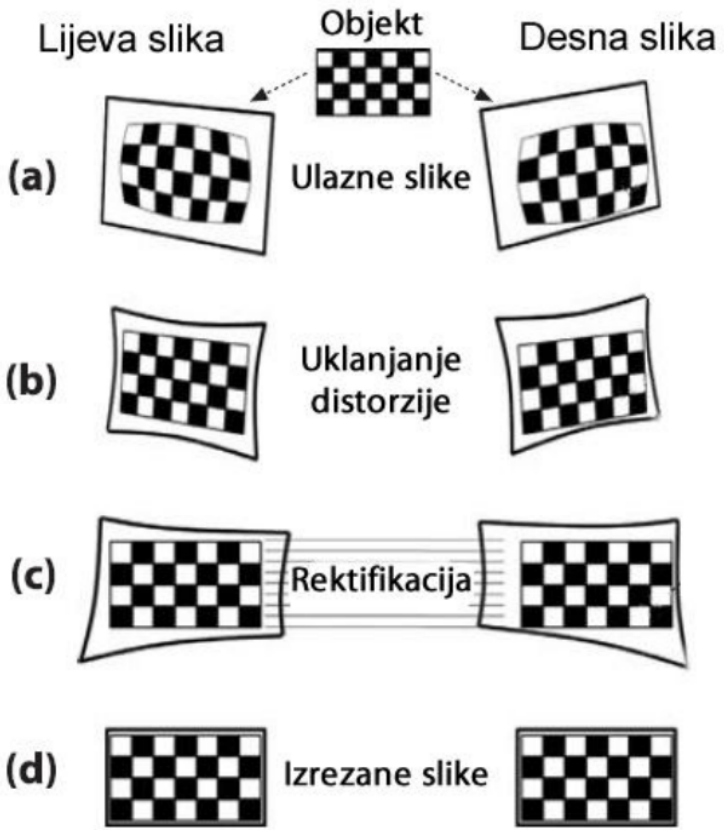
Epipolarno ograničenje

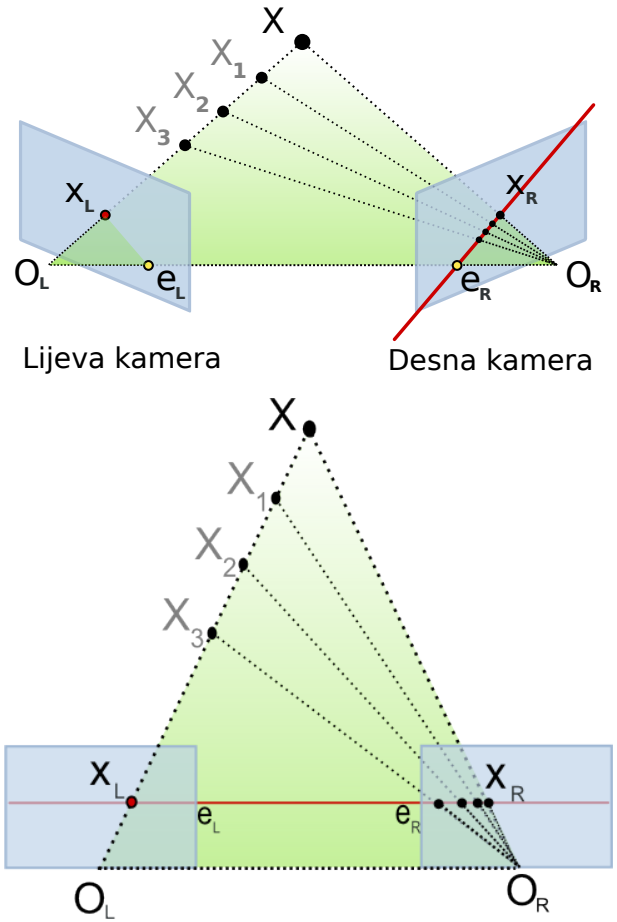
- epipolarno ograničenje - značajku x_L iz lijeve slike pretražujemo u desnoj slici na pravcu gdje ravnina određena točkama x_L, e_L, e_R siječe projekcijsku ravninu desne kamere.



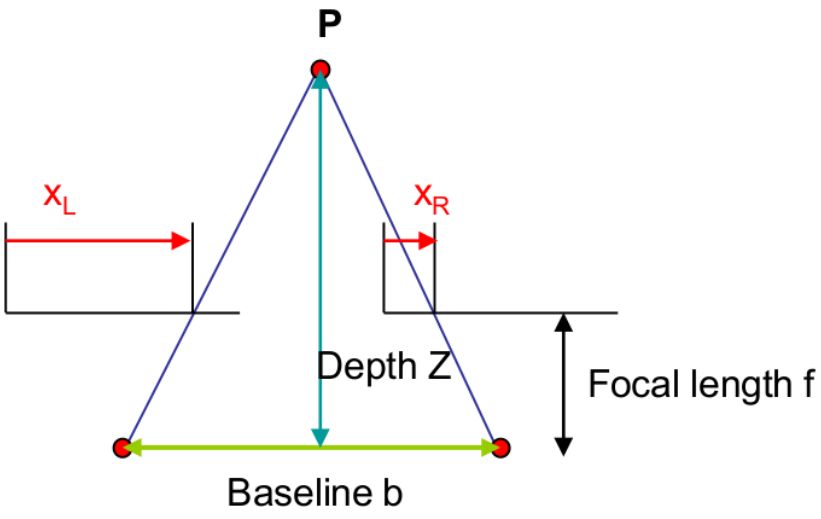
- $X \rightarrow$ promatrana točka
- $\mathrm{O_L}$ i $\mathrm{O_R} \rightarrow$ točke centara projekcije
- $\mathrm{x_L}$ i $\mathrm{x_R} \rightarrow$ projekcije točke X
- $\mathrm{e_L}$ i $\mathrm{e_R} \rightarrow$ epipolovi

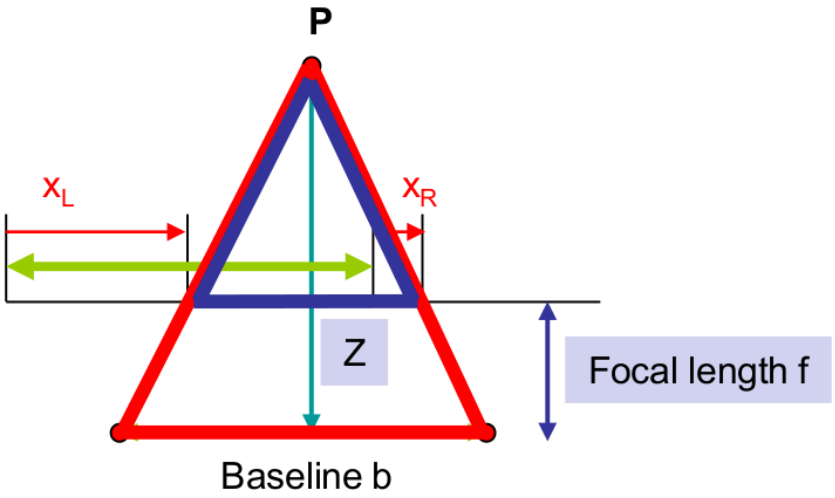
Rektificiranje stereoskopskih slika





Rektificirana triangulacija

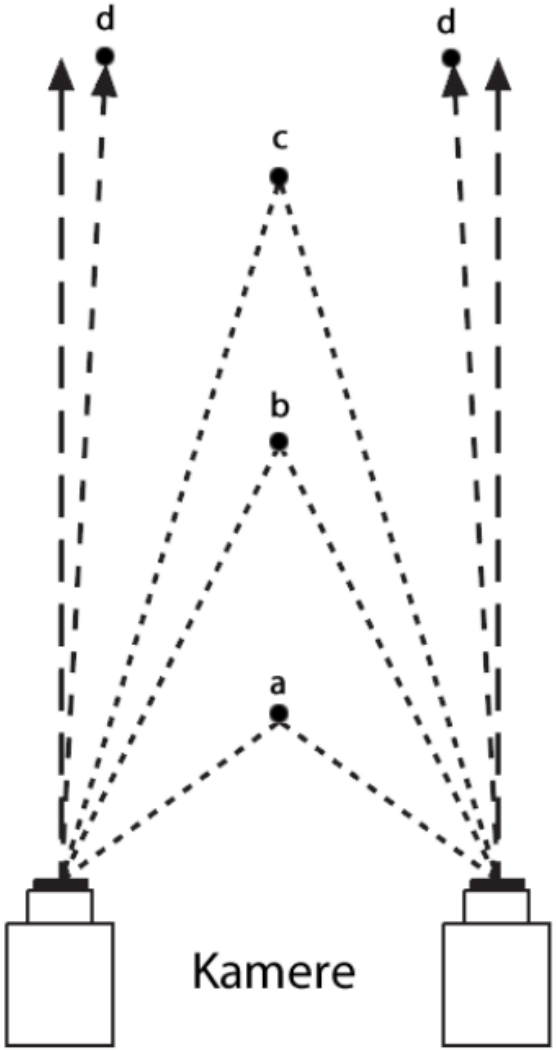


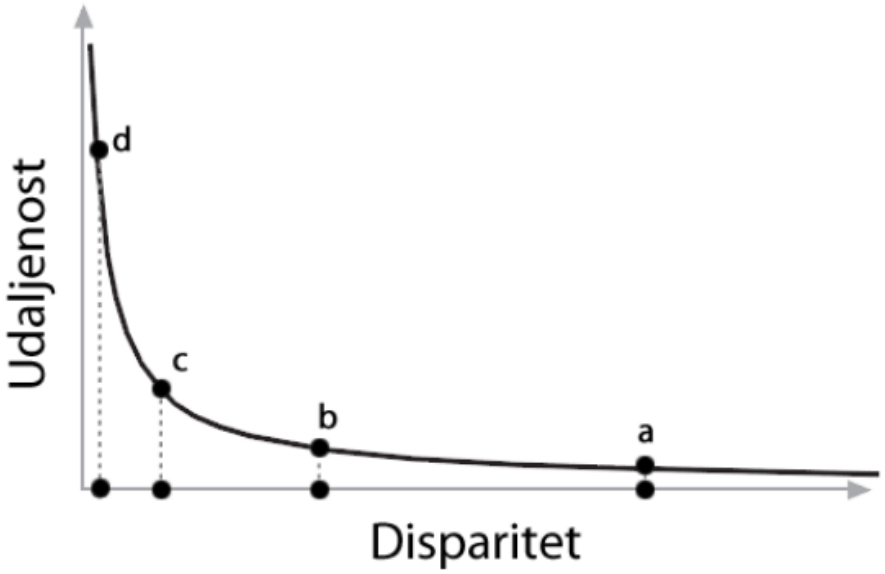


$$\frac{b}{Z} = \frac{b + x_R - x_L}{Z - f}$$

$$Z = \frac{f \cdot b}{x_L - x_R} = \frac{f \cdot b}{d}$$

Rektificirana triangulacija





Kalibriranje stereoskopskog sustava

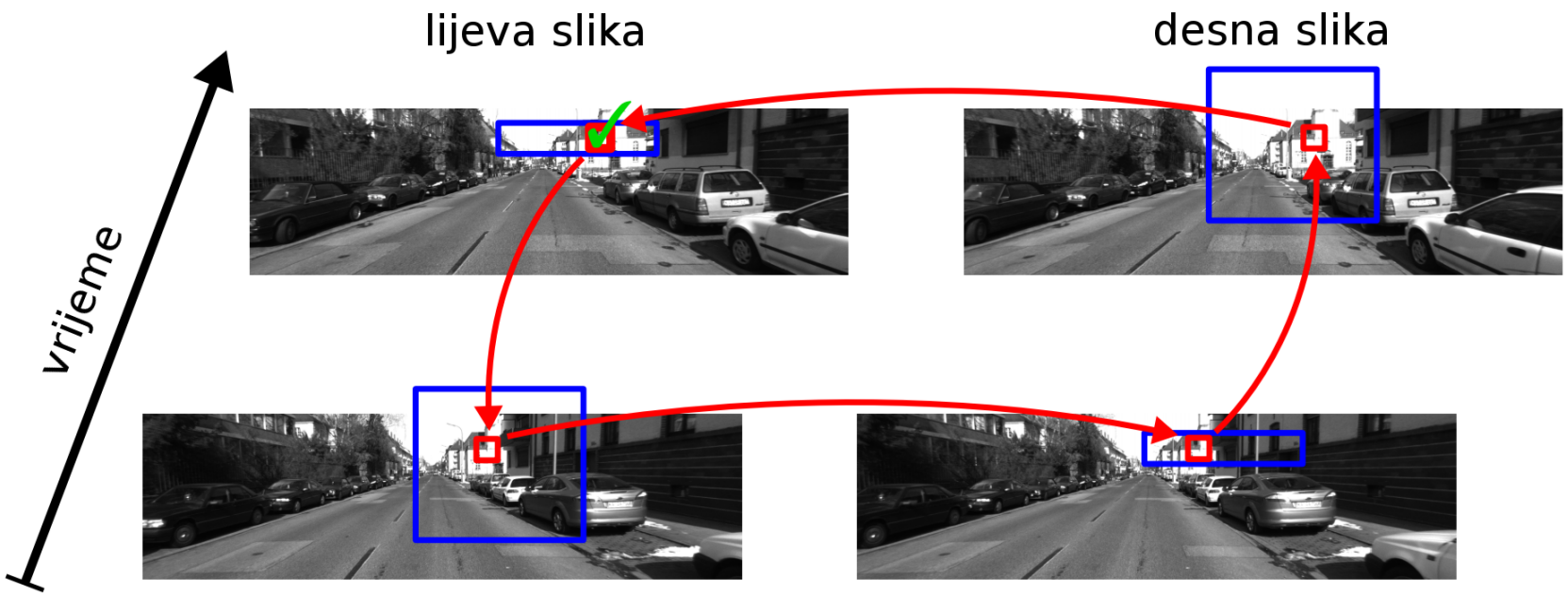
Kalibracijom stereo sustava dobivaju se:

- linearni intrinsični parametri:

$$\mathbf{K} = \begin{pmatrix} f & 0 & u_0 \\ 0 & f & v_0 \\ 0 & 0 & 1 \end{pmatrix}$$

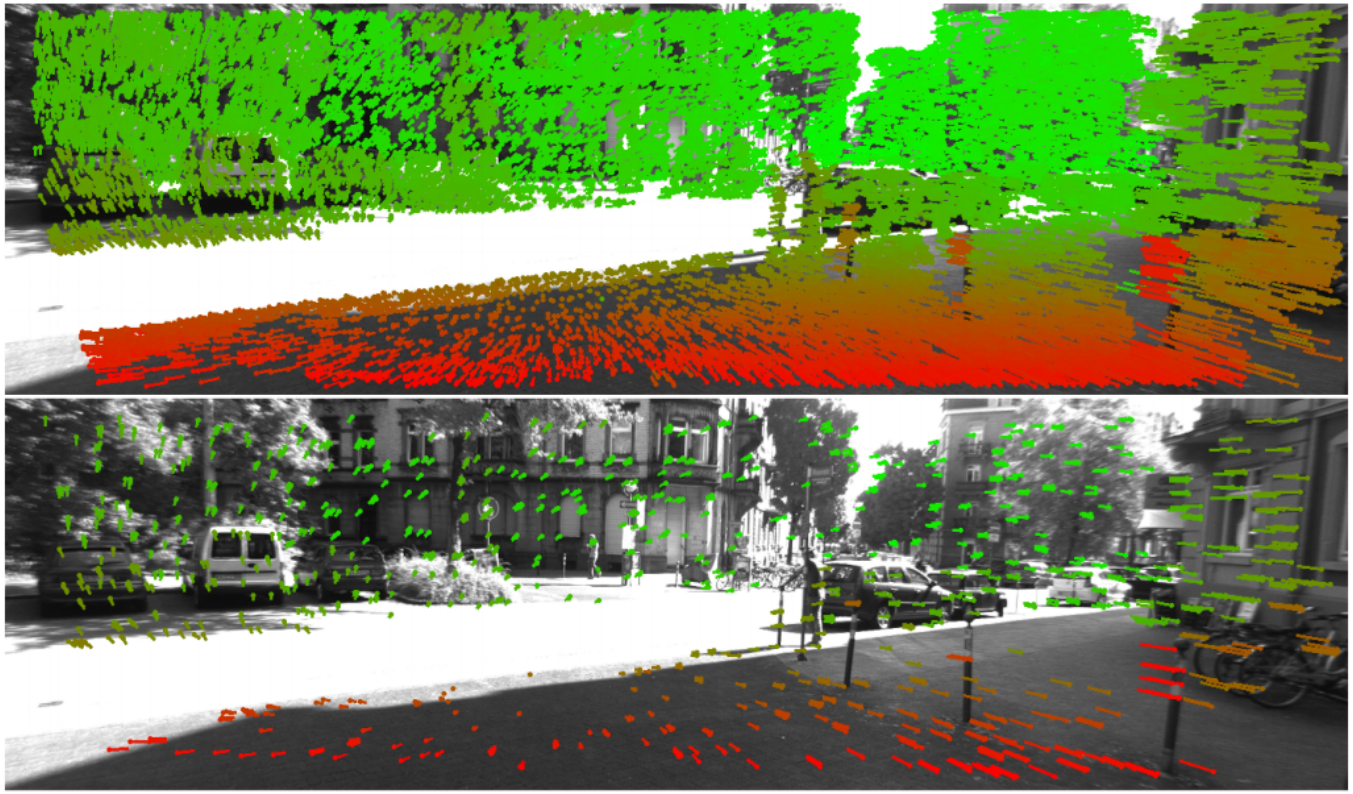
  - kamera matrice **K** opisuje pretvaranje normaliziranih koordinata projekcije u koordinate slike
- nelinearni intrinsični parametri: distorzija leće
- položaj desne kamere u odnosu na lijevu: **R**, **t**
- rektifikacijske matrice: $\mathbf{H}_l$, $\mathbf{H}_r$
- udaljenost između kamera (baseline) u rektificiranom sustavu: b

Određivanje korespondentnih značajki



- značajke se dobivaju cikličkim podudaranjem između četiri slike:
  - prethodne lijeve i desne te trenutne lijeve i desne.

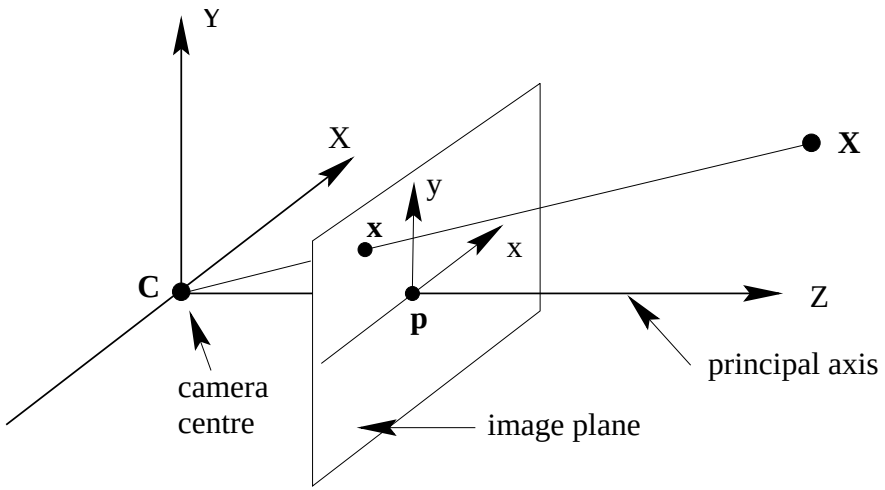
Filtriranje značajki

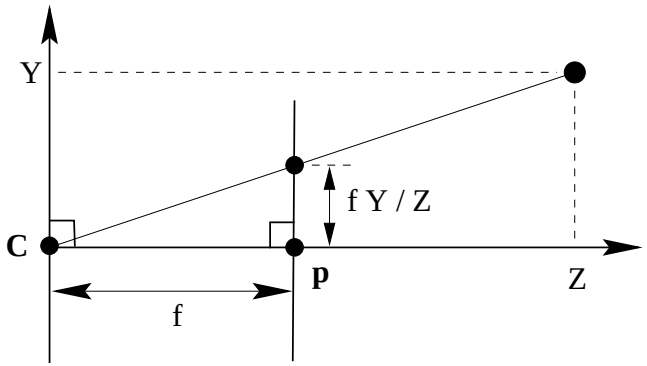


- iz pronađenih značajki gradi se optički tok (*optical flow*)
- najprije se filtriraju *outlier* značajke
- zatim se odabire relevantan skup značajki

Model kamere

Geometrijski prikaz modela kamere uz $\mathbf{K} = \begin{pmatrix} f & 0 & 0 \\ 0 & f & 0 \\ 0 & 0 & 1 \end{pmatrix}$





Model kamere

Projekcijska funkcija kamere

$$\lambda \begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = \pi(\mathbf{X}, \mathbf{R}, \mathbf{t}) = \begin{pmatrix} f & 0 & u_0 \\ 0 & f & v_0 \\ 0 & 0 & 1 \end{pmatrix} \left[\begin{bmatrix} \mathbf{R} & \mathbf{t} \end{bmatrix} \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix} - \begin{pmatrix} b \\ 0 \\ 0 \end{pmatrix} \right]$$

- 3D koordinate točke $\mathbf{X} = (x, y, z)^T$
- pomak za lijevu kameru $b = 0$, za desnu kameru b je jednak odmaku desne kamere u rektificiranom sustavu
- rotacijska i translacijska matrica $[\mathbf{Rt}]$ dimenzija 3x4
- žarišna udaljenost f
- centralna točka (u_0, v_0)
- koordinate projicirane točke u slici $(u, v, 1)^T$
- λ - informacija o dubini točke koju gubimo projiciranjem

Određivanje gibanja kamere

- Obavlja se triangulacija značajki iz para prethodnih slika:
  - N - broj značajki.
  - $\mathbf{X_i}$ - točka dobivena triangulacijom i-te značajke.
- Iterativnom gradijentnom minimizacijom izraza

$$\sum_{i=1}^{N} \|\mathbf{x_i}^{(l)} - \pi^{(l)}(\mathbf{X_i}, \mathbf{R}, \mathbf{t})\|^2 + \sum_{i=1}^{N} \|\mathbf{x_i}^{(r)} - \pi^{(r)}(\mathbf{X_i}, \mathbf{R}, \mathbf{t})\|^2$$

s obzirom na parametre $(\mathbf{R}, \mathbf{t})$ dobivamo one vrijednosti parametara za koje je pogreška projekcije minimalna.

  - $\pi^{(l)}$ - funkcija projiciranja na lijevu kameru.
  - $\pi^{(r)}$ - funkcija projiciranja na desnu kameru.
  - $\mathbf{x_i}$ - (u, v) koordinate i-te značajke u trenutnom paru slika.

Dobiveni $[\mathbf{Rt}]$ je transformacija točaka između dva trenutka u odnosu na kameru (koja je uvijek u ishodištu), a $[\mathbf{Rt}]^{-1}$ je tada gibanje kamere u odnosu na točke.

Programske biblioteke

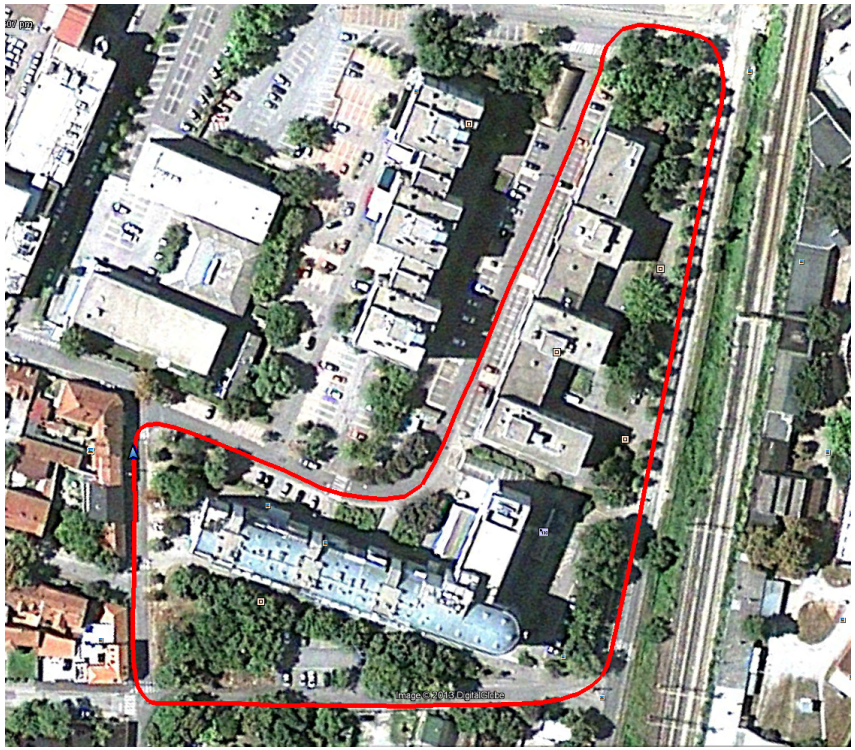
1. calib3d (stereoCalibrate, stereoRectify) - OpenCV
2. Libviso2 - KIT, Karlsruhe
3. SBA (sparse bundle adjustment) - FORTH, Heraklion
4. KLT tracker - Clemson University
5. calcOpticalFlowPyrLK - OpenCV

Eksperimentalni rezultati

1. Bumblebee dataset.
2. Sinkronizacija vizualne odometrije i GPS podataka
3. KITTI dataset.
4. Stereo model sustava kamera: testiranje Libviso biblioteke.

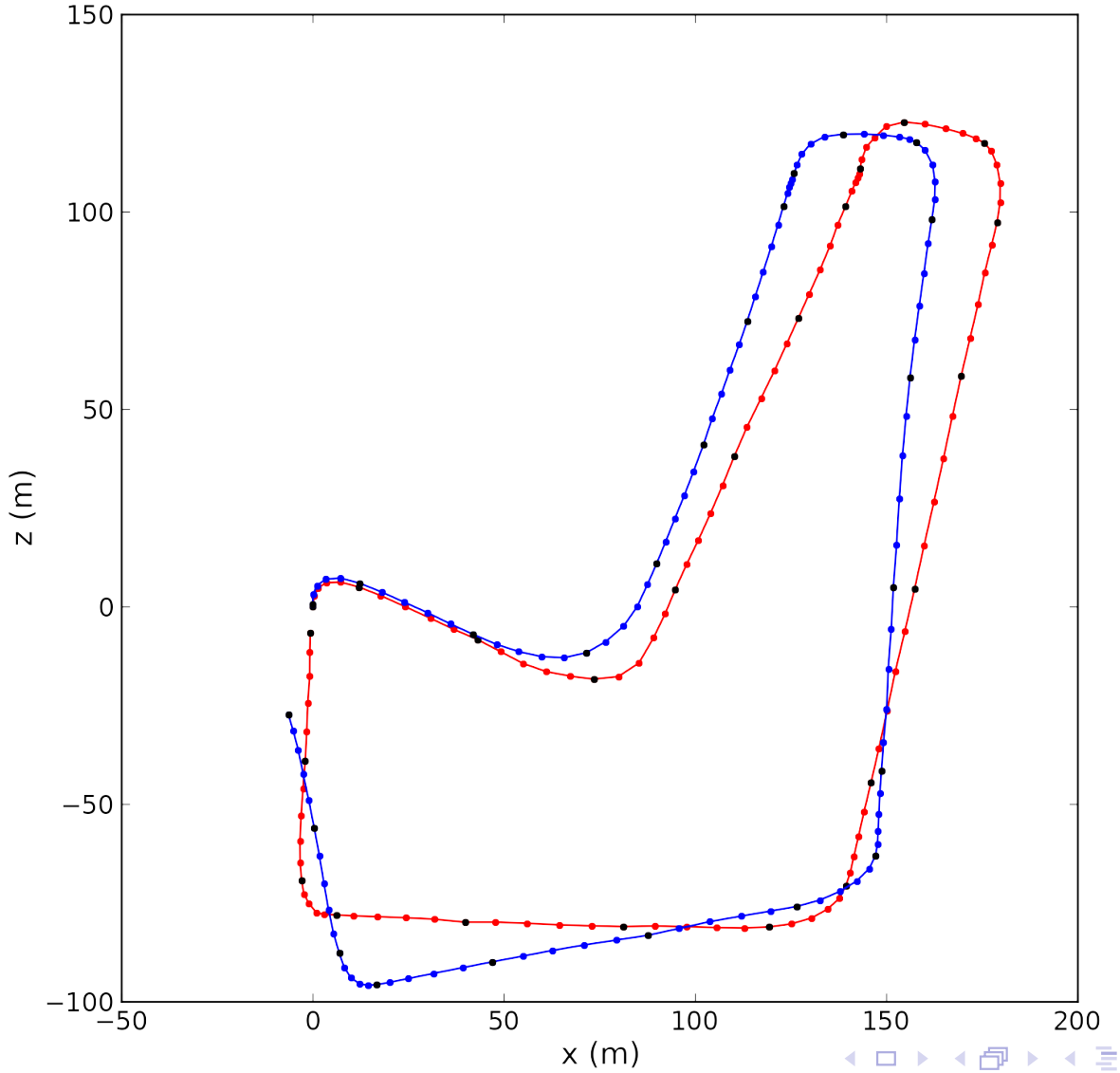
Ispitna snimka Bumblebee

- Rezolucija: 640x480
- Odmak kamere: 0.12m



Ispitna snimka Bumblebee

Rezultati vizualne odometrije (Libviso):



Interpolacija odometrije u vremenima GPS točaka

GPS podaci:

- trajanje vožnje: 111s, frekvencija dohvata: 1Hz

Vizualna odometrija:

- timestamp zapisan u pikselima
- broj okvira: 2786, nakon poduzorkovanja: 929
- ukupan timestamp vožnje: 57032 ciklusa
- 1 ciklus = $\frac{111}{57032} s \approx 2ms$

Interpolacija

Pronalazimo okvire u kojima je broj akumuliranih ciklusa najbliži trenutku GPS-a.
Položaj kamere u trenutcima GPS očitanja određujemo interpolacijom.

Sinkronizacija videa i GPS-a poklapanjem rotacije

Kut Φ između vektora pomaka kroz tri uzastopne točke:

$$\mathbf{a} \times \mathbf{b} = \begin{pmatrix} a_x \\ a_y \\ 0 \end{pmatrix} \times \begin{pmatrix} b_x \\ b_y \\ 0 \end{pmatrix} = \begin{vmatrix} i & j & k \\ a_x & a_y & 0 \\ b_x & b_y & 0 \end{vmatrix} = \begin{pmatrix} 0 \\ 0 \\ a_x b_y - b_x a_y \end{pmatrix}$$

$$\|\mathbf{a} \times \mathbf{b}\| = \|\mathbf{a}\| \|\mathbf{b}\| \sin \Phi = a_x b_y - b_x a_y$$

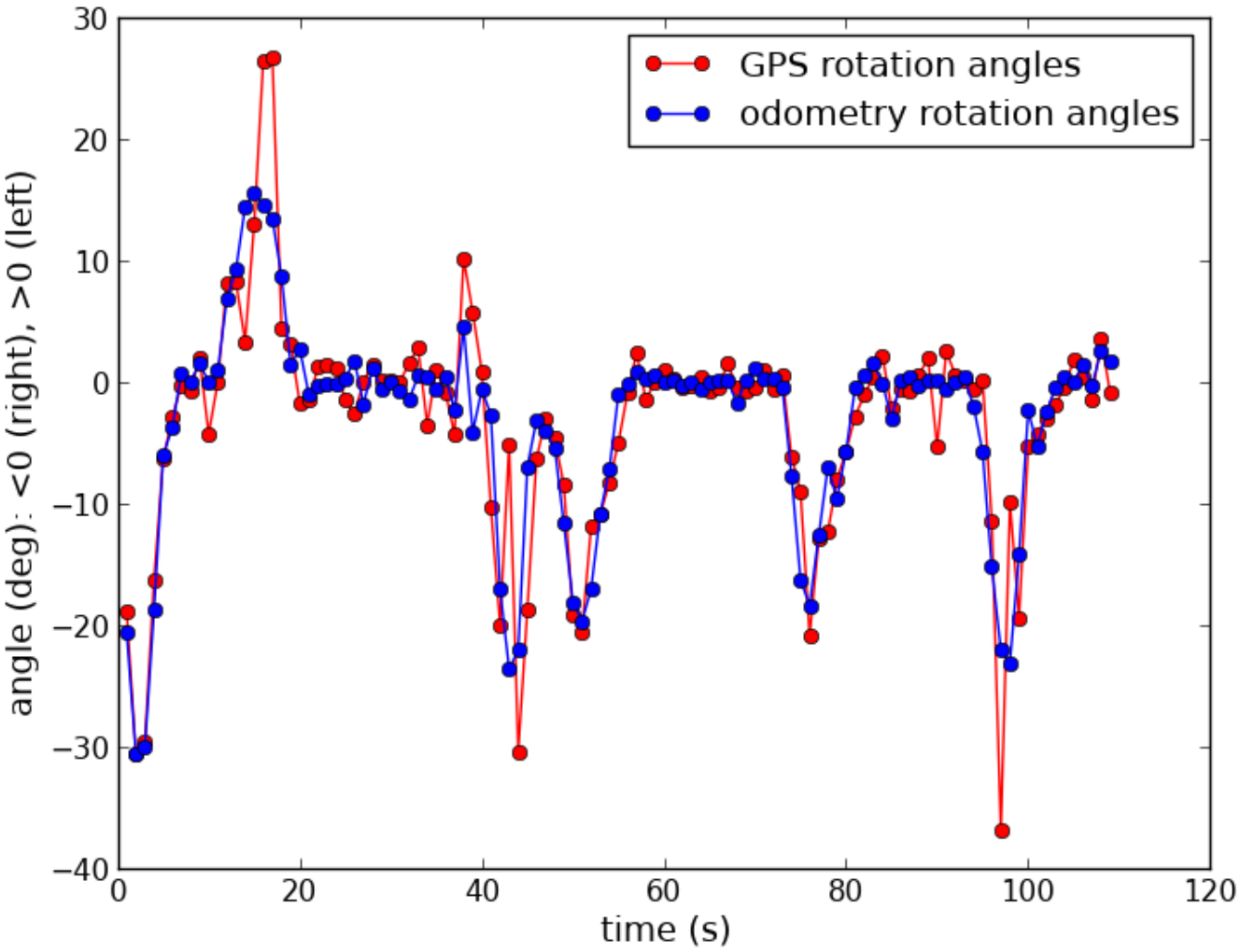
$$\Phi = \arcsin \left(\frac{a_x b_y - b_x a_y}{\|\mathbf{a}\| \|\mathbf{b}\|} \right)$$

t_s - pomak vizualne odometrije u odnosu na UTC (GPS).

$$\hat{t_s} = \underset{t_s}{\operatorname{argmin}} \sum_{t=0}^{111} (\Phi_{vo}(t + t_s) - \Phi_{gps}(t))^2$$

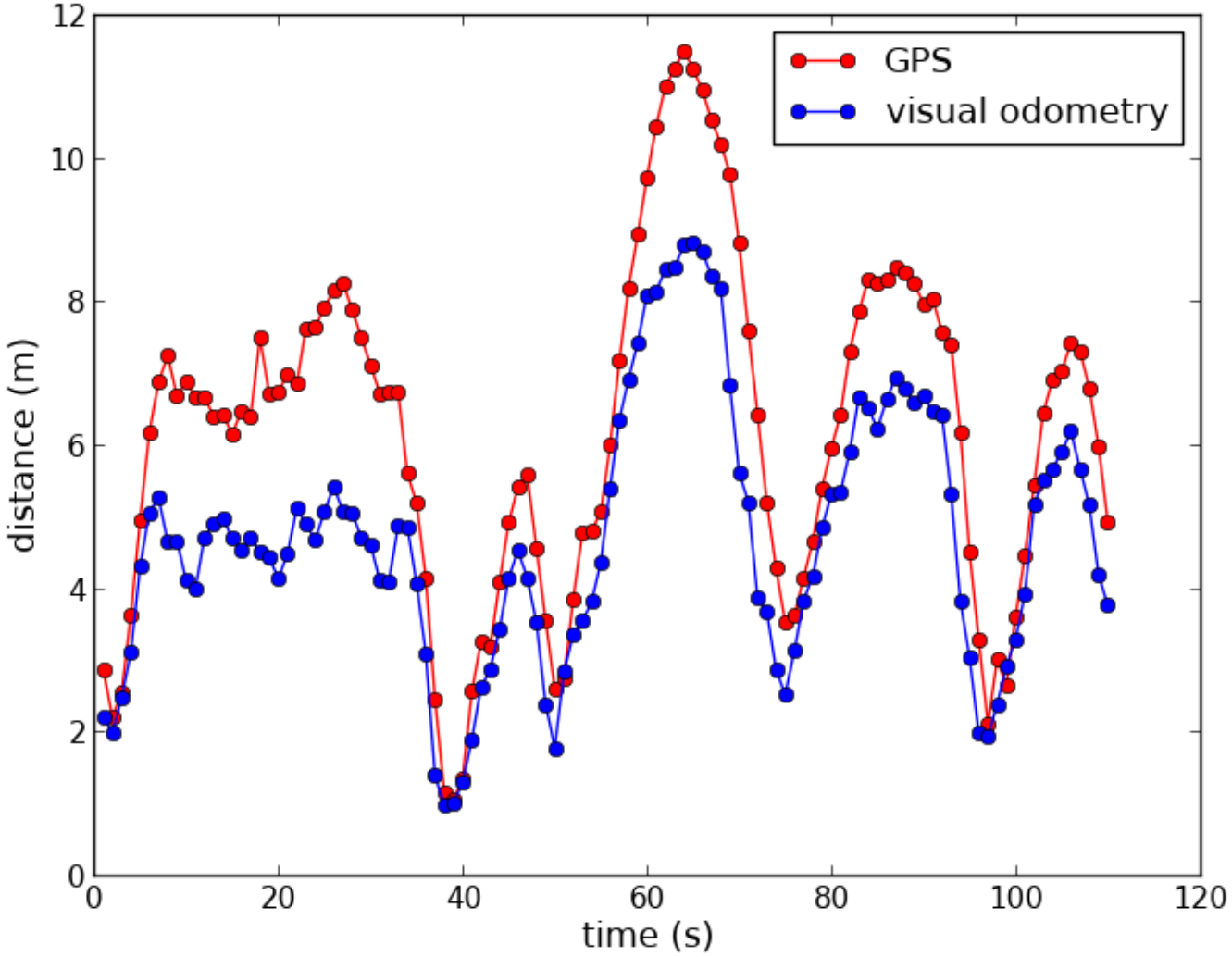
Sinkronizacija odometrije i GPS-a pomoću rotacije

- Inkrementalni zakret Φ dobiven odometrijom i GPS-om



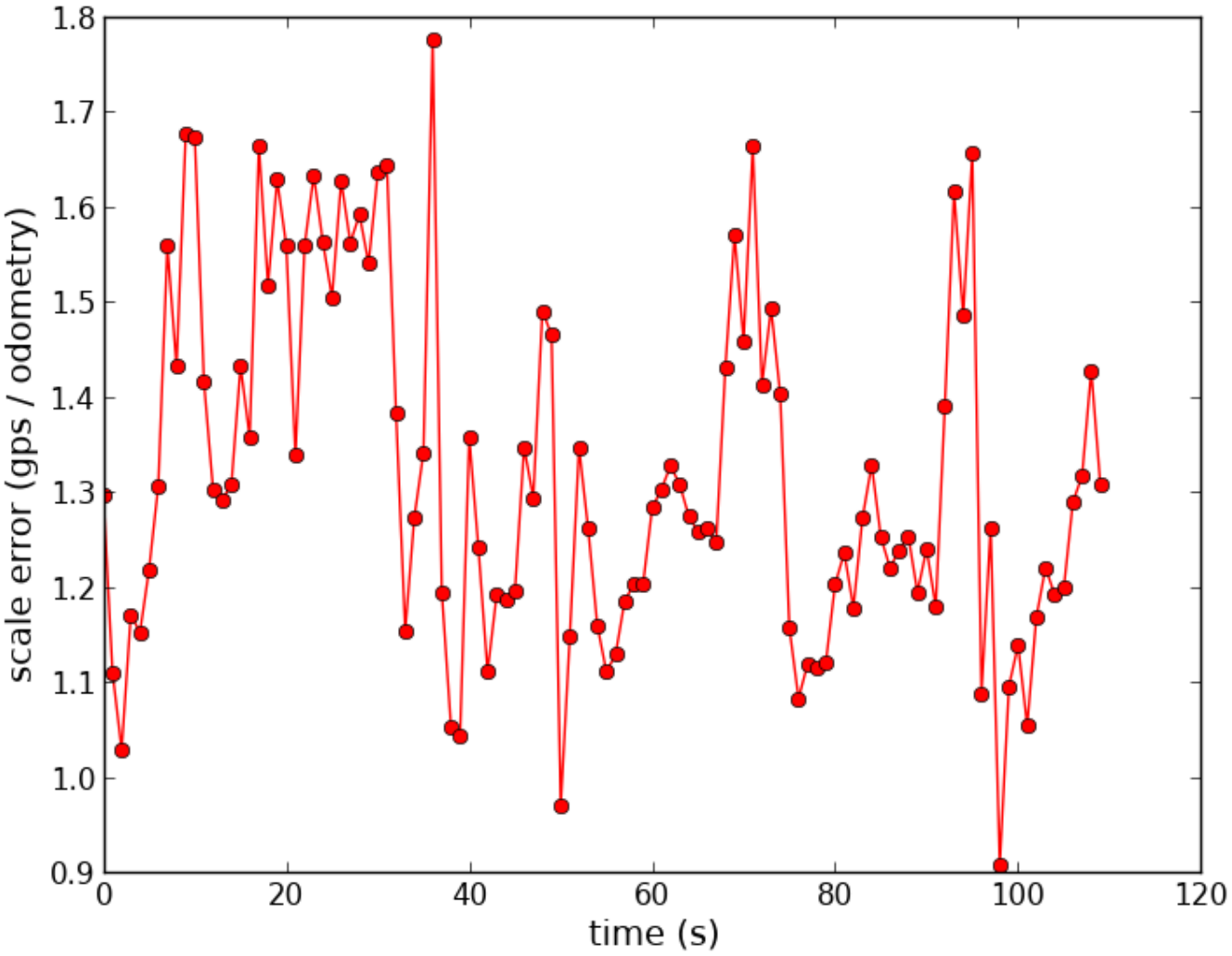
Evaluacija preciznosti translacije

Inkrementalni translacijski pomak: $\Delta s(t) = \|\mathbf{p}(t) - \mathbf{p}(t-1)\|$



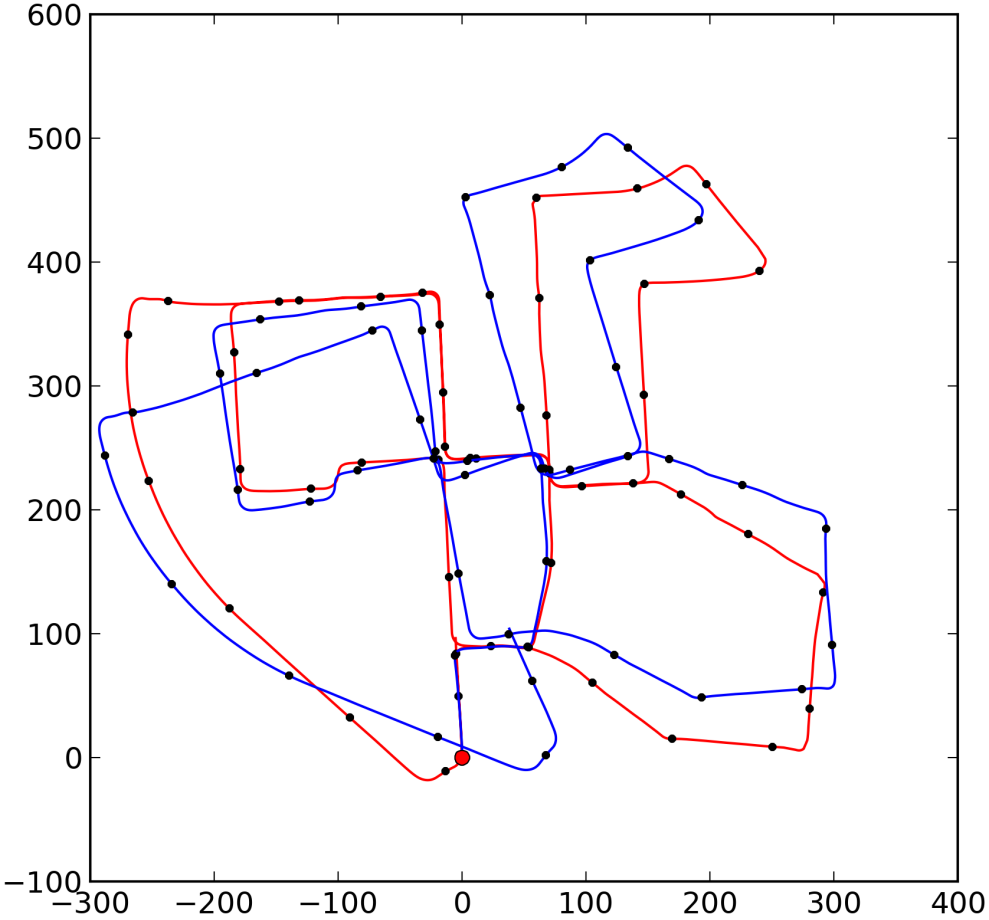
Evaluacija odometrije i GPS-a

- Pogreška skale: $E(t) = \frac{\|\mathbf{p}_{vo}(t) - \mathbf{p}_{vo}(t-1)\|}{\|\mathbf{p}_{GPS}(t) - \mathbf{p}_{GPS}(t-1)\|}$



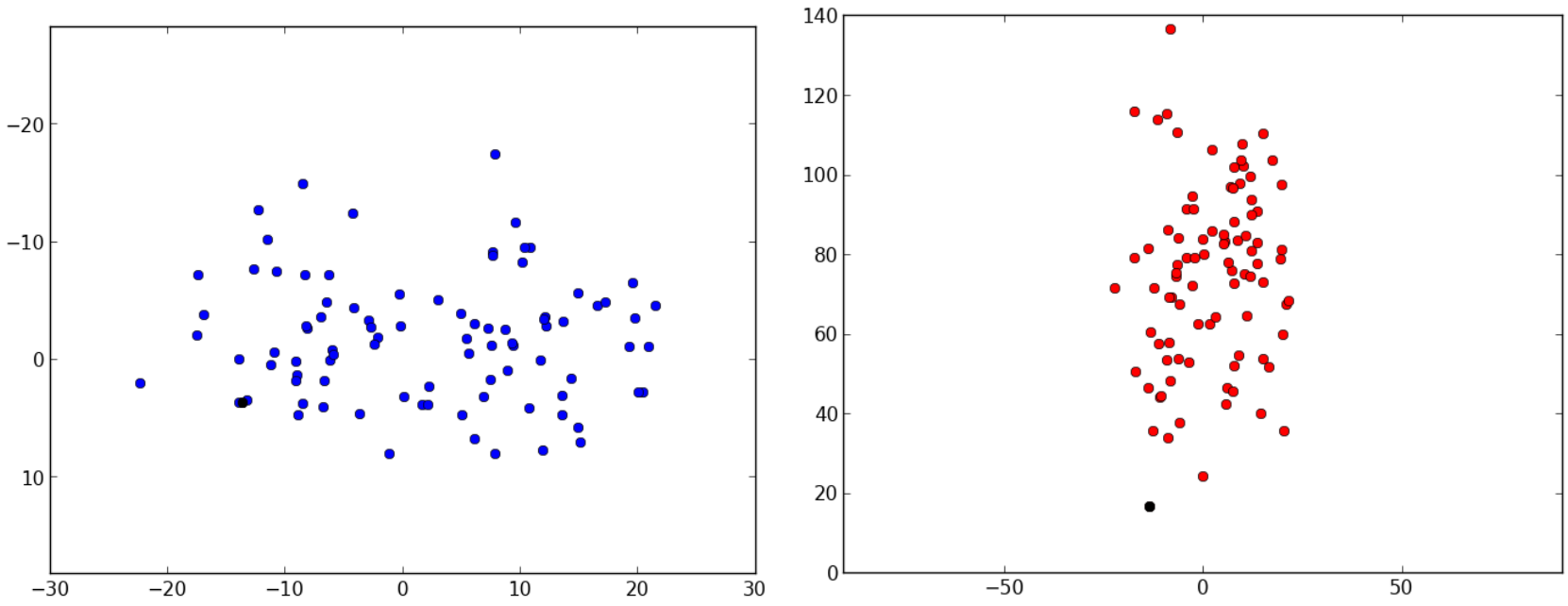
Ispitna snimka KITTI

- Rezolucija: 1241x376
- Razmak: 0.537m



Umjetni ispitni skup podataka

- Modelira linearne intrinsične i ekstrinsične parametre
- Ulaz: ekstrinsični parametri kretanja kamere
- Izlaz: projekcije generiranih 3D točaka u točke lijeve i desne slike - ulazni podaci za vizualnu odometriju

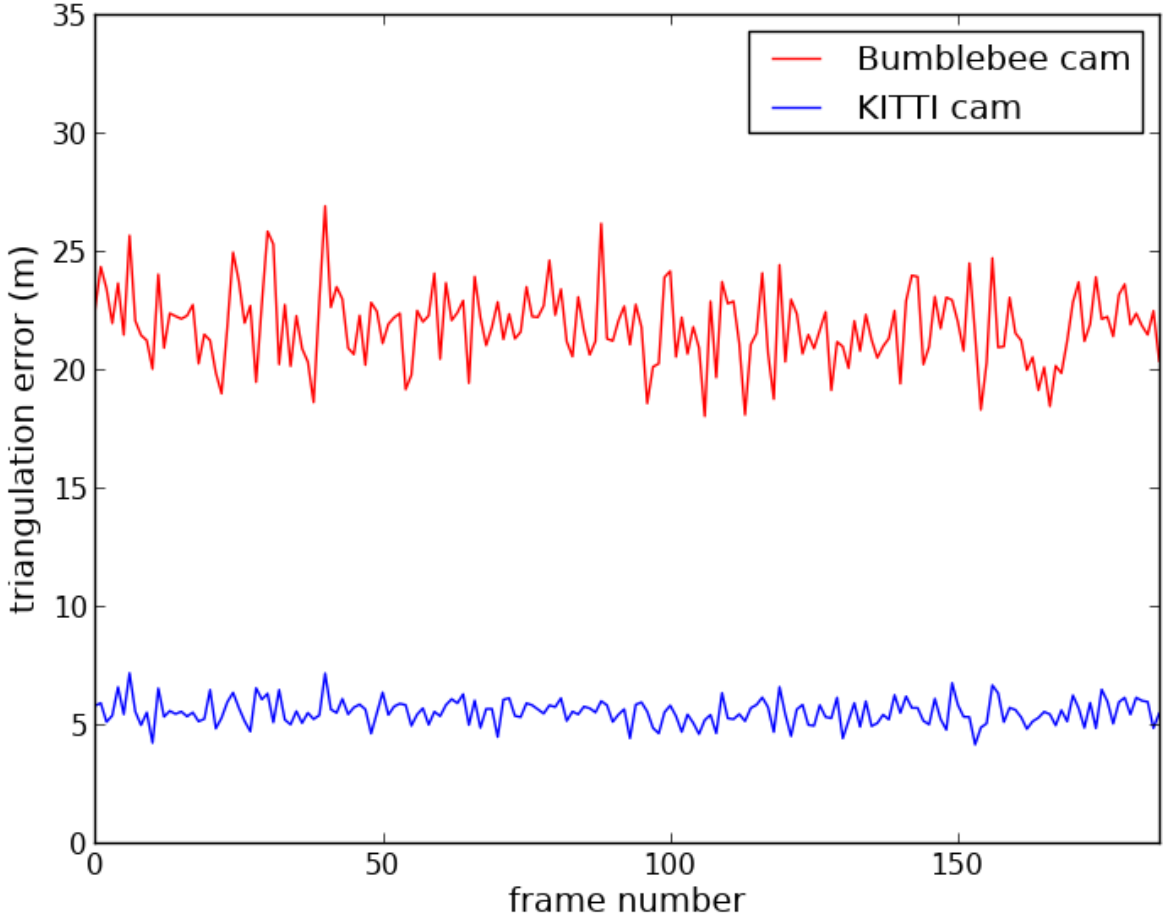


(e) XY osi (f) XZ osi

Slika: Distribucija generiranih točaka u 3D prostoru

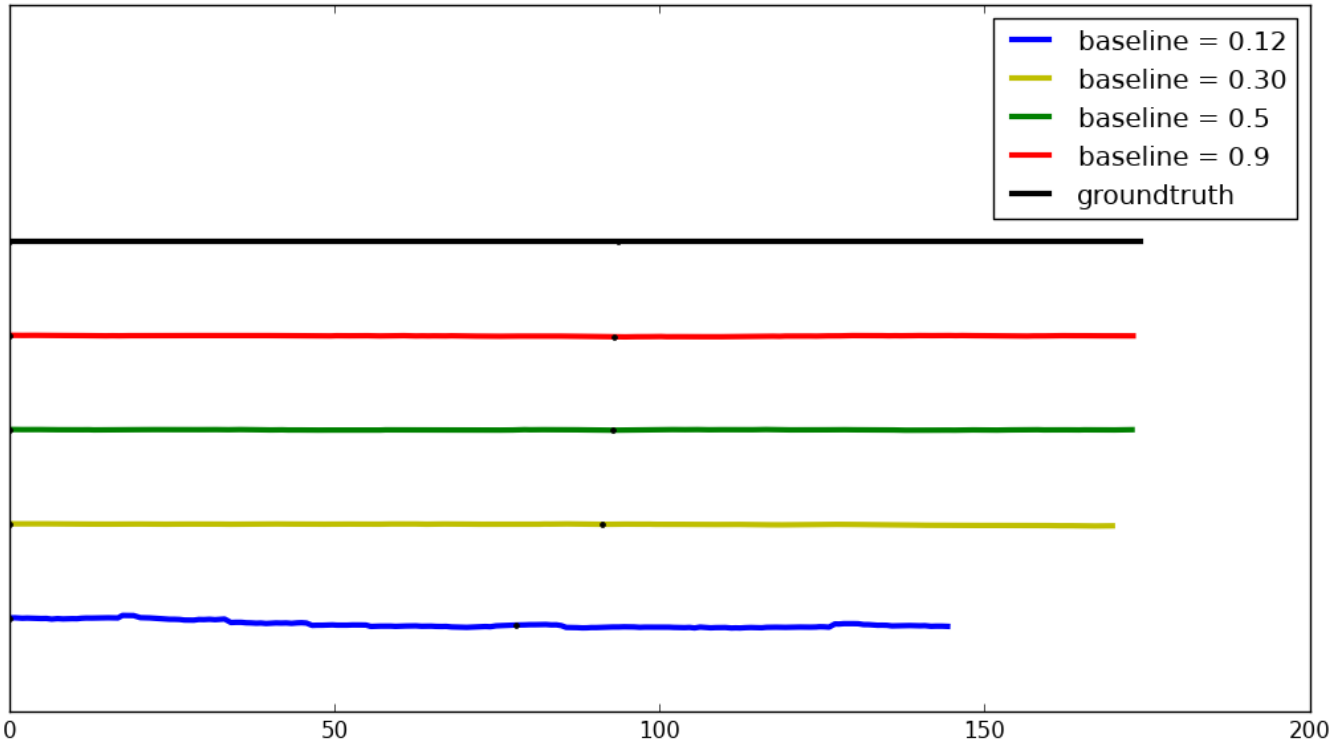
Umjetni ispitni skup podataka

- usporedba uprosječene pogreške triangulacije za parametre kamera Bumblebee i KITTI



Umjetni ispitni skup podataka

Trajektorije dobivene za različite razmake kamera (Bumblebee)



Zaključak i budući ciljevi

- praćenje kroz više okvira
  - proširiti SBA biblioteku za rad sa stereo sustavom i povezati je s vizualnom odometrijom
- optimizacija: ukloniti proces rektifikacije čitave slike:
  - rektificirati odvojen skup značajki iz lijeve slike i desne slike te ih zatim podudarati
  - podudarati značajke pretraživanjem po epipolarnim pravcima
    - rektificirati samo pronađene značajke
- upotrijebiti estimaciju strukture scene i gibanja kamere za ciljeve projekta Vista:
  - praćenje oznaka na kolniku - razdjelna linija i rubovi ceste
  - lane departure warning
  - detekcija i filtriranje prometnih znakova

Hvala na pažnji!