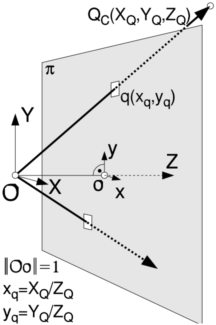
Trodimenzionalni računalni vid

geometrija jednog, dvaju i više pogleda

Siniša Šegvić
UniZg-FER D307

Uvod

Geometrijski aspekti računalnog vida: elementi za rekonstrukciju 3D modela iz 2D slika



Sadržaj:

- **Prerekviziti**: projekcijska geometrija, model stvaranja slike, umjeravanje kamere
- **Geometrija jednog pogleda**: homografije i nedogledi
- **Geometrija dvaju pogleda**: ravninske scene, epipolarna geometrija, 3D rekonstrukcija
- Primjene: autonomna navigacija, izrada 3D modela, proširena stvarnost

Projekcijska geometrija

- Omogućava **linearan** opis stvaranja slike i transformacija među koordinatnim sustavima!
- Elementi projekcijskog prostora imaju homogene koordinate:

$$\mathbf{x} = \begin{bmatrix} x_1, x_2, \ldots, x_{n+1} \end{bmatrix}^T \in \mathbb{P}^n$$
$$\mathbf{x} \sim \mathbf{y} \Longleftrightarrow \mathbf{x} = \lambda \cdot \mathbf{y}$$

- **Projekcija**: transformacija $\mathbb{P}^m \rightarrow \mathbb{P}^n$ koja čuva kolinearnost
- Projekcije su **linearna preslikavanja** (mogu se opisati homogenom matricom $\mathbf{P}_{m+1 \times n+1}$)
- Projekcijske invarijante: kolinearnost, križni omjer (udaljenosti, kutovi i paralelizam —ne)

Projekcijska geometrija

Odnos projekcijskog i Euklidskog prostora:

- Izomorfno preslikavanje $\mathbb{R}^n \rightarrow \mathbb{P}^n$:

$$\left[\begin{array}{c} x_1, x_2, \ldots, x_n \end{array}\right]^T \mapsto \left[\begin{array}{c} x_1, x_2, \ldots, x_n, 1 \end{array}\right]^T$$

- Za sve $\mathbf{y} \in \mathbb{P}^n : y_{n+1} \neq 0$ vrijedi obratno preslikavanje:

$$\left[\begin{array}{c} y_1, y_2, \ldots, y_{n+1} \end{array}\right]^T \mapsto \left[\begin{array}{c} \frac{y_1}{y_{n+1}}, \frac{y_2}{y_{n+1}}, \ldots, \frac{y_n}{y_{n+1}} \end{array}\right]^T$$

- točke $\mathbf{y} \in \mathbb{P}^n : y_{n+1} = 0$ se nazivaju točke u beskonačnosti
    - takve točke nemaju prikaz u $\mathbb{R}^n$;
    - skup svih takvih točaka čini *ravninu* (*pravac*) u beskonačnosti

Projekcijska geometrija

Elementi projekcijske ravnine:

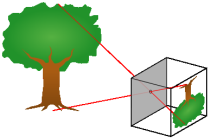
objekt	prikaz	jednadžba	transformacija
točka	$\mathbf{q} = \begin{bmatrix} x & y & t \end{bmatrix}^T$		$\mathbf{Hq}$
pravac	$\mathbf{l} = \begin{bmatrix} a & b & c \end{bmatrix}^T$	$\mathbf{l}^T\mathbf{q} = 0$	$\mathbf{H}^{-T}\mathbf{l}$
kr. 2. reda	$\mathbf{C}_{3\times 3} = \mathbf{C}_{3\times 3}^T$	$\mathbf{q}^T\mathbf{C}\mathbf{q} = 0$	$\mathbf{H}^{-T}\mathbf{C}\mathbf{H}^{-1}$

Homografija: Ravninska projekcija $\mathbb{P}^2 \to \mathbb{P}^2$; opisuje se matricom $\mathbf{H}_{3\times 3}$.

Lako se pokazuje da se presjecište dvaju pravaca da odrediti prema izrazu: $\mathbf{l}_1 \times \mathbf{l}_2 = \mathbf{q}$

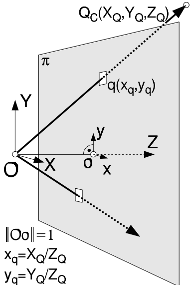
Model stvaranja slike

- **Idealna kamera**: model temeljen na kameri s rupicom
  - kamera s rupicom (stenoskop) sastoji se od kutije koja na jednom kraju ima rupicu a na drugoj senzorski element



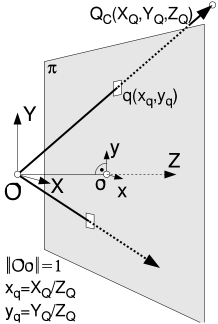
[Wikipedia]

  - zrake svjetla prolaze kroz rupicu i na senzoru stvaraju sliku rotiranu za 180°.
  - koncept radi, ali zahtijeva dugačku ekspoziciju (sati)
- U modelu, rupica je infinitezimalna a njena projekcija pada u ishodište k.s. senzora.
- udaljenost senzora postavljamo na $\|Oo\|=1$, a slikovnu ravninu crtamo ispred rupice (sliku bismo prije gledanja zakrenuli)
- stvaranje slike sada možemo opisati perspektivnom projekcijom



Model stvaranja slike

- Projekcijski, stvaranje slike opisuju **linearni izrazi**:



$$\lambda \cdot \begin{bmatrix} x_{\hat{q}} \\ y_{\hat{q}} \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} X_Q \\ Y_Q \\ Z_Q \\ 1 \end{bmatrix}$$

ili, kraće:

$$\lambda \cdot \widehat{\mathbf{q}} = \mathbf{S} \cdot \mathbf{Q_C}$$

$\lambda \in \mathbb{R}$ ... nepoznati faktor (mnoge 3D točke se preslikavaju u istu točku slike)
$\widehat{\mathbf{q}} \in \mathbb{P}^2$ ... normalizirani k.s. slike $(o, \hat{x}, \hat{y})$,
$\mathbf{Q_C} \in \mathbb{P}^3$ ... k.s. kamere (O, X, Y, Z)

$\|Oo\|$ ... udaljenost slikovne ravnine (senzora) od optičkog centra (rupice)

- Kamera s rupicom dobro modelira kamere koje daju **neizobličene** slike (bridovi scene $\mapsto$ dužine u slici)

Model stvaranja slike

Perspektivni model stvaranja slike:

- Linearno proširenje minimalnog modela s pet parametara:

$$\begin{bmatrix} x_q \\ y_q \\ 1 \end{bmatrix} = \begin{bmatrix} s_x & s_\theta & t_x \\ 0 & s_y & t_y \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} x_{\hat{q}} \\ y_{\hat{q}} \\ 1 \end{bmatrix}$$

ili, kraće:

$$\mathbf{q} = \mathbf{K} \cdot \widehat{\mathbf{q}}$$

  - s_x, s_y ...$\|Oo\|$ izražen u pikselima ($s_x \approx s_y$)
    $p_{ar} = \frac{s_y}{s_x}$; $\phi_{\mathrm{hfov}} = 2 \cdot \operatorname{arc\,tg}(\frac{w}{2s_x})$, za sliku $w \times h$
  - t_x, t_y ...pomak k.s. slike
  - $s_\theta = s_y \operatorname{ctg}(\theta)$ ...kut među osima k.s. slike ($\theta \approx \pi/2, s_\theta \approx 0$)
- **Ostali intrinsični parametri** opisuju nelinearne korekcije!

Model stvaranja slike

Transformacija krutog tijela:

- homogene koordinate omogućuju linearan opis euklidskih transformacija (6 DOF):

$$\begin{bmatrix} X_Q \\ Y_Q \\ Z_Q \\ 1 \end{bmatrix} = \begin{bmatrix} \mathbf{R}_{3x3} & \mathbf{t}_{3x1} \\ 0_{1x3} & 1 \end{bmatrix} \cdot \begin{bmatrix} L_Q \\ M_Q \\ N_Q \\ 1 \end{bmatrix}$$

ili, kraće:

$$\mathbf{Q_C} = \mathbf{M} \cdot \mathbf{Q_W}$$

$\mathbf{Q_C}$ - točka izražena u k.s. kamere (O, X, Y, Z),
R - rotacija k.s. kamere oko osi k.s. svijeta,
t - translacija k.s. svijeta u k.s. kamere,
$\mathbf{Q_W}$ - točka izražena u k.s. svijeta (O_W, L, M, N)

Model stvaranja slike

Vanjski parametri: položaj kamere u k.s. svijeta
(translacija $-\mathbf{R}^\top \cdot \mathbf{t}$, rotacija $\mathbf{R}$)

- Parametrizacija rotacijske matrice (nagib,zakret,okret):

$$\mathbf{R} = \mathbf{R_z}(\gamma) \cdot \mathbf{R_y}(\beta) \cdot \mathbf{R_x}(\alpha)$$

$$\mathbf{R} = \begin{bmatrix} \cos(\beta)\ \cos(\gamma) & \cos(\gamma)\ \sin(\alpha)\ \sin(\beta) - \cos(\alpha)\ \sin(\gamma) & \cos(\alpha)\ \cos(\gamma)\ \sin(\beta) + \sin(\alpha)\ \sin(\gamma) \\ \cos(\beta)\ \sin(\gamma) & \cos(\alpha)\ \cos(\gamma) + \sin(\alpha)\ \sin(\beta)\ \sin(\gamma) & \cos(\gamma)\ \sin(\alpha) + \cos(\alpha)\ \sin(\beta)\ \sin(\gamma) \\ -\sin(\beta) & \cos(\beta)\ \sin(\alpha) & \cos(\alpha)\ \cos(\beta) \end{bmatrix}$$

Parametrizacija nije jedinstvena: zadanu matricu opisuju barem dvije trojke $(\alpha,\beta,\gamma) \in [0, \pi]^3$!

- Inverzni parametri:

$$\mathbf{M}_i = \mathbf{M}^{-1} = \begin{bmatrix} \mathbf{R_i}_{3x3} & \mathbf{t_i}_{3x1} \\ 0_{1x3} & 1 \end{bmatrix}$$

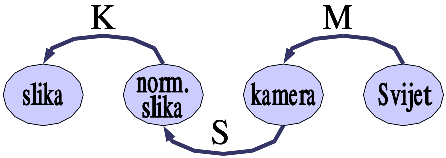
$$\mathbf{R_i} = \mathbf{R}^{-1} = \mathbf{R}^T, \quad \mathbf{t_i} = -\mathbf{R}^{-1} \cdot \mathbf{t}$$

Model stvaranja slike

Perspektivni model, sažetak:

- Pretpostavke: ``oštra'' slika i idealna (``tanka'') leća
- Svakom slikovnom elementu pridružuje se pravac kroz žarište (upotrebom intrinsičnih parametara **K**)
- Ekstrinsični parametri (**R**, **t**, 6 DOF) određuju kako te pravce interpretirati u nekom vanjskom k.s.
- Ukupna transformacija:

$$\lambda \mathbf{q} = \mathbf{K} \cdot \mathbf{S} \cdot \mathbf{M} \cdot \mathbf{Q_W} = \mathbf{P}_{3\times 4} \cdot \mathbf{Q_W}$$



Geometrija dvaju pogleda

Tri međusobno isprepletena **cilja**!

1. određivanje 3D strukture
2. određivanje pomaka kamera (5DOF)
3. određivanje korespondentnih točaka

Dva **konteksta** koji imaju mnogo toga zajedničkog

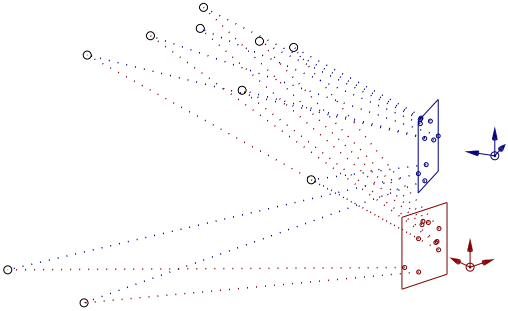
1. stereo: slike su pribavljene istovremeno
2. kretanje: slike su pribavljene sljedno u vremenu

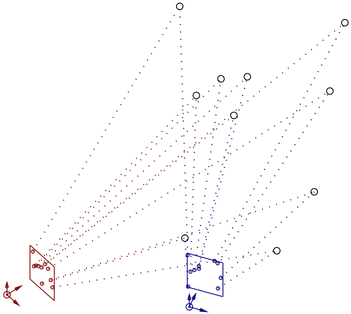
Dva **slučaja** koji se u mnogome razlikuju

1. kalibrirani slučaj: poznate kamere
   (metrička rekonstrukcija)
2. projekcijski slučaj: nepoznate kamere
   (autokalibracija ili projekcijska rekonstrukcija)

Geometrija dvaju pogleda

Primjeri:





Geometrija dvaju pogleda

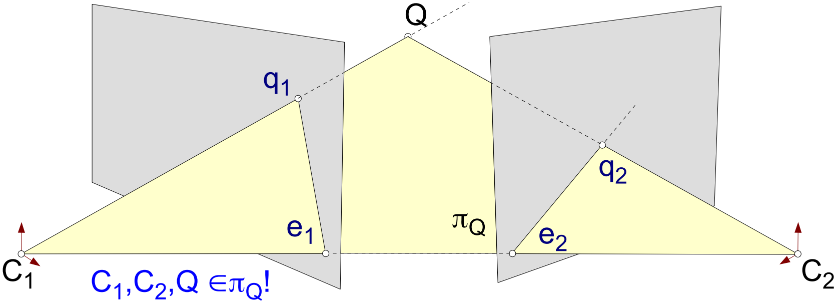
Epipolarna geometrija: odnos između dva pogleda na 3D scenu

Osnovica p(C_1, C_2): povezuje ishodišta dvaju kamera

Epipol e_i: slika ishodišta druge kamere stereo para; $e_i \in$ p(C_1,C_2)

Epipolarna ravnina π_Q: def. osnovicom i 3D točkom Q

Epipolarni pravac $p(e_i, q_i)$: presjek epipolarne i slikovne ravnine



Geometrija dvaju pogleda

Epipolarno ograničenje: geometrijski izvod

Neka je: $\mathbf{P}_1 = \begin{bmatrix} \mathbf{I} & \mathbf{0} \end{bmatrix}$, $\mathbf{P}_2 = \begin{bmatrix} \mathbf{R} & \mathbf{t} \end{bmatrix}$,

Tada u k.s. kamere C_2 vrijedi:

$\overrightarrow{\mathbf{C_2 q_2}} = \mathbf{q_2}, \quad \overrightarrow{\mathbf{C_2 C_1}} = \mathbf{t}, \quad \overrightarrow{\mathbf{C_1 q_1}} = \mathbf{R} \cdot \mathbf{q_1}$

Gornja tri vektora su koplanarna $\Rightarrow$ mješoviti produkt iščezava:

$$\mathbf{q_2}^\top \cdot (\mathbf{t} \times \mathbf{R} \cdot \mathbf{q_1}) = \mathbf{q_2}^\top \cdot [\mathbf{t}]_\times \cdot \mathbf{R} \cdot \mathbf{q_1} = 0$$

Epipolarno ograničenje: $\mathbf{q_2}^\top \cdot \mathbf{E} \cdot \mathbf{q_1} = 0$

Esencijalna matrica (homogena, 3×3, 5 DOF): $\mathbf{E} = [\mathbf{t}]_\times \mathbf{R}$

Geometrija dvaju pogleda

Epipolarno ograničenje: algebarski izvod

Neka je: $\mathbf{P}_1 = \begin{bmatrix} \mathbf{I} & \mathbf{0} \end{bmatrix}$, $\mathbf{P}_2 = \begin{bmatrix} \mathbf{R} & \mathbf{t} \end{bmatrix}$

Tada vrijedi sljedeće: $\mathbf{Q}_2 = \mathbf{R}\mathbf{Q}_1 + \mathbf{t}$

Rješavamo se zbroja: $[\mathbf{t}]_\times \mathbf{Q}_2 = [\mathbf{t}]_\times \mathbf{R}\mathbf{Q}_1 + [\mathbf{t}]_\times \mathbf{t}$

$$[\mathbf{t}]_\times = \begin{bmatrix} 0 & -t_z & t_y \\ t_z & 0 & -t_x \\ -t_y & t_x & 0 \end{bmatrix}$$, vrijedi: $[\mathbf{t}]_\times \cdot \mathbf{a} = \mathbf{t} \times \mathbf{a}$

Rješavamo se člana na lijevoj strani: $\mathbf{Q}_2^\top [\mathbf{t}]_\times \mathbf{Q}_2 = \mathbf{Q}_2^\top [\mathbf{t}]_\times \mathbf{R}\mathbf{Q}_1$

Konačno dobivamo: $0 = \mathbf{q}_2^\top \mathbf{E} \mathbf{q}_1$

Geometrija dvaju pogleda

Određivanje matrice E: algoritam s osam točaka

Raspisivanjem i reorganiziranjem epipolarnog ograničenja ($0 = \mathbf{q}_{iB}^{\top}\mathbf{E}\mathbf{q}_{iA}$) dobivamo jednadžbu:

$$\begin{bmatrix} x_{i\mathrm{B}}x_{i\mathrm{A}} & x_{i\mathrm{B}}y_{i\mathrm{A}} & x_{i\mathrm{B}} & y_{i\mathrm{B}}x_{i\mathrm{A}} & y_{i\mathrm{B}}y_{i\mathrm{A}} & y_{i\mathrm{B}} & x_{i\mathrm{A}} & y_{i\mathrm{A}} & 1 \end{bmatrix} \cdot \mathbf{e}_{9\times 1} = 0$$

Za n točaka dobivamo linearni homogeni sustav: $\mathbf{A}_{n\times 9} \cdot \mathbf{e}_{9\times 1} = \mathbf{0}_{n\times 1}$

Minimalno osam točaka za rješenje [longuet-Higgins 1981],
u praksi bismo htjeli imati mnogo više (100)

Linearna metoda postiže jako dobre rezultate (pogotovo nakon normalizacije [hartley04]), ali ima problema s outlierima i ravninama...

Dekompozicija esencijalne matrice

Promotrimo singularnu dekompoziciju matrice: $\mathbf{E} = [\mathbf{t}]_\times \mathbf{R}$:

$$\mathbf{E} = \mathbf{U} \cdot \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \mathbf{V}^\top, \quad \det(U) = \det(V) = 1$$

Tada imamo po jedan uvrnuti par za svaki smjer translacije:

- $\mathbf{t} = \pm \mathbf{U}_{:3}$
- $\mathbf{R}_a = \mathbf{U} \cdot \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \mathbf{V}^\top, \quad \mathbf{R}_b = \mathbf{U} \cdot \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \mathbf{V}^\top$

Odabir između $(\mathbf{R}_a, \mathbf{t})$, $(\mathbf{R}_a, -\mathbf{t})$, $(\mathbf{R}_b, \mathbf{t})$, $(\mathbf{R}_b, -\mathbf{t})$ provodimo testiranjem jedne ili više korespondencija.

Geometrija dvaju pogleda

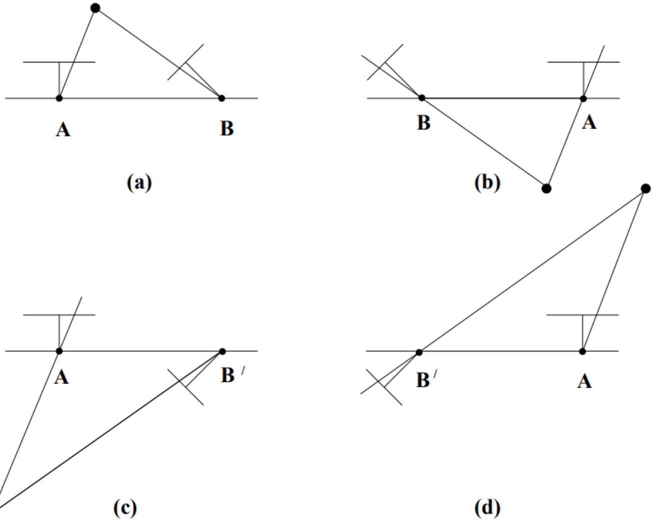


Fig. 9.12. **The four possible solutions for calibrated reconstruction from E.** *Between the left and right sides there is a baseline reversal. Between the top and bottom rows camera B rotates* 180° *about the baseline. Note, only in (a) is the reconstructed point in front of both cameras.*

[hartley04book]

Geometrija dvaju pogleda

Rekonstrukcija strukture: linearna metoda

Neka je: $\lambda_A \mathbf{q_A} = \mathbf{P_A Q}$, $\lambda_B \mathbf{q_B} = \mathbf{P_B Q}$,

Neka su poznate vrijednosti za: $\mathbf{q_A}, \mathbf{q_B}, \mathbf{P_A}$ i $\mathbf{P_B}$

Kako naći $\mathbf{Q}$?

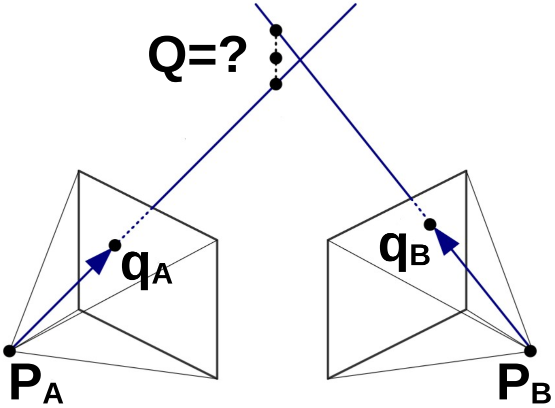
Rješenje (samo dva ograničenja su nezavisna): $[\mathbf{q_i}]_\times \mathbf{P_i Q} = 0$

Kad se to raspiše dobiva se sustav: $\mathbf{A}_{4\times4}\mathbf{Q}_{4\times1} = 0$

$$\mathbf{A}_{4\times4} = \begin{bmatrix} y_A \mathbf{P}_{\mathbf{A}[3,:]} - \mathbf{P}_{\mathbf{A}[2,:]} \\ x_A \mathbf{P}_{\mathbf{A}[3,:]} - \mathbf{P}_{\mathbf{A}[1,:]} \\ y_B \mathbf{P}_{\mathbf{B}[3,:]} - \mathbf{P}_{\mathbf{B}[2,:]} \\ x_B \mathbf{P}_{\mathbf{B}[3,:]} - \mathbf{P}_{\mathbf{B}[1,:]} \end{bmatrix}$$

Efikasnija metoda predložena je u [nister04pami].

Geometrija dvaju pogleda



[szeliski22book]

Geometrija dvaju pogleda

Rekonstrukcija strukture: linearna metoda (2)

Kao i obično, linearna metoda nije ML, pa ne postiže optimalne rezultate;

U praksi linearna metoda se koristi za inicijalizaciju, dok se bolji rezultati postižu optimizacijom.

Tipično zajedno optimiramo položaje svih 3D točaka i relativnu orijentaciju kamera (bundle adjustment) [engels06].

Geometrija dvaju pogleda

Određivanje matrice E: algoritam s pet točaka

Matrica **E** nije "bilo kakva" homogena matrica:

- samo 5 DOF (umjesto 8)
- specifična struktura: $\mathbf{E} = [\mathbf{t}]_\times \mathbf{R}$
- skup svih esencijalnih matrica zadan ``kubnim'' ograničenjem:
  $2 \cdot \mathbf{E}\mathbf{E}^T\mathbf{E} - \mathrm{trace}(\mathbf{E}\mathbf{E}^T)\mathbf{E} = 0$

Algoritam s pet točaka [Nister 2004] uzima u obzir sva ograničenja (epipolarna i interna)!

Dodatna prednost pred linearnom metodom: radi i za planarne scene!

Geometrija dvaju pogleda

Određivanje matrice E: algoritam s pet točaka (2)

Prvo se pronađu donja 4 nul-vektora linearnog sustava
$\mathbf{A}_{5\times 9} \cdot \mathbf{e}_{9\times 1} = \mathbf{0}_{5\times 1}$

Zatim se njihova linearna kombinacija uvrsti u kubno ograničenje, te se traže koeficijenti za koje je ograničenje zadovoljeno

Ovo rezultira polinomom 11-og stupnja koji ima 1-11 nultočaka (svaka nultočka odgovara zasebnom rješenju)

Na kraju je iz skupa dobivenih hipoteza potrebno naći točnu: to se radi uvrštanjem šeste točke u epipolarno ograničenje.

Algoritam s pet točaka je metoda izbora u minimalnom kontekstu (5+1 točka), dok se za velik broj točaka bolji rezultati postižu kondicioniranim linearnim metodama

Geometrija dvaju pogleda

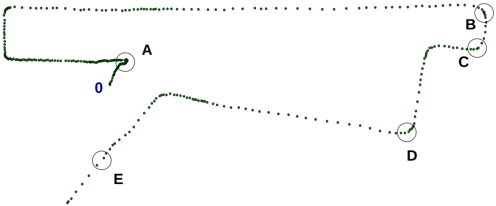
Geometrija više pogleda: proširenje tehnika za dva pogleda

Primjer: monokularna vizualna odometrija [nister04cvpr2]

- ideja: estimirati gibanje kamere konkateniranjem estimiranih pomaka u kratkim intervalima videa
- položaje kamere moguće je odrediti:
  1. iz esencijalne matrice u odnosu na prvu sliku intervala
  2. iz položaja trijanguliranih 3D točaka (PnP)
- intervali su potrebni zbog rastuće nesigurnosti praćenja
- mjerilo u odnosu na prethodni interval određeno na temelju 3D rekonstrukcije točaka koje su vidljive u oba intervala (glavni nedostatak pristupa!)
- za najbolje rezultate optimirati reprojekcijsku pogrešku u posljednjih n slika [engels06]

Geometrija dvaju pogleda

Rezultati kartiranja puta od oko 1100 m (sveučilišni kampus):



- **oblik trajektorije** dobro rekonstruiran, znatna odstupanja u lokalnom mjerilu
- karta predstavljena grafom gdje su rubne slike čvorovi, a lokalne 3D rekonstrukcije lukovi [segvic09cviu]
- lukovi sadrže pomake kamere, 2D izglede praćenih značajki te trijangulirane 3D položaje
- vozilo je uspjelo autonomno ponoviti put uz 5 intervencija.

Geometrija jednog pogleda

Geometriju jednog pogleda definira projekcijska matrica $\mathbf{P}_{3\times 4}$:

$$\lambda \mathbf{q} = \mathbf{K} \cdot \mathbf{S} \cdot \mathbf{M} \cdot \mathbf{Q}_{\mathsf{W}} = \mathbf{K} \cdot \begin{bmatrix} \mathbf{R} & \mathbf{t} \end{bmatrix} \cdot \mathbf{Q}_{\mathsf{W}} = \mathbf{P}_{3\times 4} \cdot \mathbf{Q}_{\mathsf{W}},$$

Projekcijska matrica u sebi sadrži i vanjske ($\mathbf{R}$, $\mathbf{t}$) i unutrašnje ($\mathbf{K}$) parametre kamere!

Nulprostor projekcijske matrice je položaj kamere: $\mathrm{Nul}(\mathbf{P}) = (-\mathbf{R}^\top \mathbf{t}, 1)$

Ako je kamera umjerena, $\mathbf{P} = \mathbf{S} \cdot \mathbf{M}$
Ako je još i k.s. svijeta poravnat s k.s. kamere: $\mathbf{P} = \mathbf{S} = \begin{bmatrix} \mathbf{I} & \mathbf{0} \end{bmatrix}$

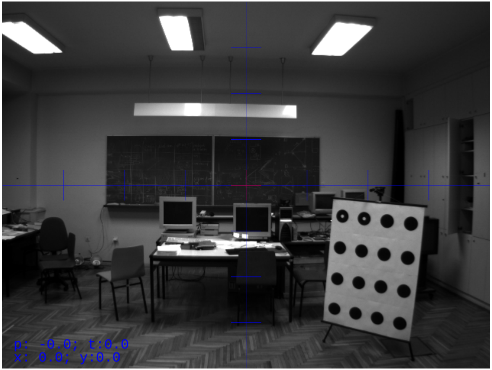
3D iz jednog pogleda samo uz pomoć apriornog znanja

- poznati elementi ravninske scene
- paralelne glavne osi strukture scene

Geometrija jednog pogleda

Homografija: ravninsko projekcijsko preslikavanje

- Koja je veza između ravnine i njezine slike?
- Možemo li iz te veze odrediti položaj kamere (**P**)?



Geometrija jednog pogleda

Homografija: veza između ravnine i njene slike

Neka je k.s. svijeta (O, X, Y, Z) orijentiran tako da je Z okomit na ravninu uzorka (analogija: nogometno igralište iz helikoptera):

$$\lambda \cdot \begin{bmatrix} x_q \\ y_q \\ 1 \end{bmatrix} = \mathbf{K} \begin{bmatrix} \mathbf{R} & \mathbf{t} \end{bmatrix} \begin{bmatrix} X_Q \\ Y_Q \\ 0 \\ 1 \end{bmatrix} = \mathbf{K} \begin{bmatrix} \mathbf{R}_{:,1} & \mathbf{R}_{:,2} & \mathbf{t} \end{bmatrix} \begin{bmatrix} X_Q \\ Y_Q \\ 1 \end{bmatrix}$$

$$\lambda \cdot \mathbf{q} = \mathbf{H} \cdot \mathbf{q_R},$$

$$\mathbf{q} = \begin{bmatrix} x_q & y_q & 1 \end{bmatrix}^\top, \quad \mathbf{q_R} = \begin{bmatrix} X_Q & Y_Q & 1 \end{bmatrix}^\top$$

Iz **H** možemo dobiti i **R** i **t**!

Geometrija jednog pogleda

Homografija iz korespondencija između poznatog objekta i slike

- Nepoznati faktor izbacujemo iz igre vektorskim produktom:
  $\mathbf{H} \cdot \mathbf{q}_R = \lambda \cdot \mathbf{q} \Longleftrightarrow \mathbf{q} \times (\mathbf{H} \cdot \mathbf{q}_R) = \mathbf{0}$

$$\mathbf{q} \times (\mathbf{H} \cdot \mathbf{q}_R) = \begin{bmatrix} x_q \\ y_q \\ 1 \end{bmatrix} \times \begin{bmatrix} \mathbf{H}_{1:}\mathbf{q}_R \\ \mathbf{H}_{2:}\mathbf{q}_R \\ \mathbf{H}_{3:}\mathbf{q}_R \end{bmatrix} = \begin{bmatrix} y_q \cdot \mathbf{H}_{3:}\mathbf{q}_R - 1 \cdot \mathbf{H}_{2:}\mathbf{q}_R \\ 1 \cdot \mathbf{H}_{1:}\mathbf{q}_R - x_q \cdot \mathbf{H}_{3:}\mathbf{q}_R \\ x_q \cdot \mathbf{H}_{2:}\mathbf{q}_R - y_q \cdot \mathbf{H}_{1:}\mathbf{q}_R \end{bmatrix}$$

$$= \begin{bmatrix} \mathbf{0}_3^T & -\mathbf{q}_R^T & y_q\mathbf{q}_R^T \\ \mathbf{q}_R^T & \mathbf{0}_3^T & -x_q\mathbf{q}_R^T \\ -y_q\mathbf{q}_R^T & x_q\mathbf{q}_R^T & \mathbf{0}_3^T \end{bmatrix} \cdot \begin{bmatrix} \mathbf{H}_{1,:}^T \\ \mathbf{H}_{2,:}^T \\ \mathbf{H}_{3,:}^T \end{bmatrix} = \mathbf{0}_3$$

- Imamo dvije nezavisne jednadžbe $\forall$ par točaka!
- Za sve parove, dobivamo linearni sustav: $\mathbf{A}_{2n\times 9} \cdot \mathbf{h}_9 = 0_{2n}$

Geometrija jednog pogleda

Homografija: kako riješiti sustav s prethodne stranice?

- **Prijedlog**: minimizirati L2 normu algebarskog reziduala
  $\widehat{\mathbf{h}}_{\mathbf{alg}} = \arg\min_{|\mathbf{h}_9|=1} ||\mathbf{A}_{2n\times 9} \cdot \mathbf{h}_9||$
  - Prednost: može se izračunati eksplicitno (*in closed form*) i brzo SVD dekompozicijom!
  - ali...
- **Problem 1**: dobiveno rješenje ne maksimira vjerodostojnost!
  - ML rješenje: $\widehat{\mathbf{h}}_{\mathbf{geom}} = \arg\min_{\mathbf{h}_9} \sum_i \mathrm{d}_{\mathbb{R}^2}(\mathbf{H} \cdot \mathbf{q}_{Ri}, \mathbf{q}_i)^2$
  - nužan iterativan postupak (gradijentna optimizacija)
  - iterativna optimizacija traži dobru inicijalizaciju...
- **Problem 2**: postupak je osjetljiv na *outliere*!
  - Outlieri se mogu detektirati Monte Carlo analizom (RANSAC), ili optimizacijom robustne norme

Geometrija jednog pogleda

Homografija: kako riješiti sustav s pret-prethodne stranice?

1. Odbaciti outliere Monte Carlo analizom (RANSAC)
2. Odrediti $\widehat{\mathbf{h}}_{\mathbf{alg}}$ kao inicijalno rješenje (SVD)
3. Poboljšati rješenje iterativnom metodom (optimirati ML kriterijsku funkciju)
4. Pokušati proširiti skup korespondencija vođenim uparivanjem, iterirati.

Geometrija jednog pogleda

Nedogled (vanishing point): točka slike u kojoj se sijeku projekcije paralelnih 3D pravaca.

Uz pomoć nedogleda možemo odrediti 3D smjer pravaca u k.s. kamere.

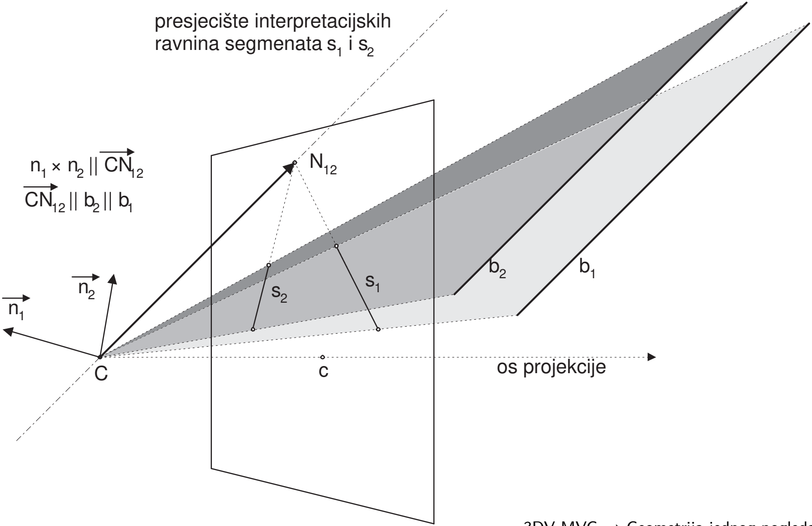
Projekcijsko objašnjenje nedogleda:
većina točaka u beskonačnosti $Q_n \in \mathbb{P}^3$ se preslikavaju u konačne točke ravnine $q_n \in \mathbb{P}^2$

$$\lambda \begin{bmatrix} x_n \\ y_n \\ 1 \end{bmatrix} = \begin{bmatrix} P'_{3\times3} & P''_{3\times1} \end{bmatrix} \cdot \begin{bmatrix} k_n \\ l_n \\ m_n \\ 0 \end{bmatrix} = P'_{3\times3} \cdot \begin{bmatrix} k_n \\ l_n \\ m_n \end{bmatrix}$$

Geometrija jednog pogleda

Nedogled: Euklidsko objašnjenje

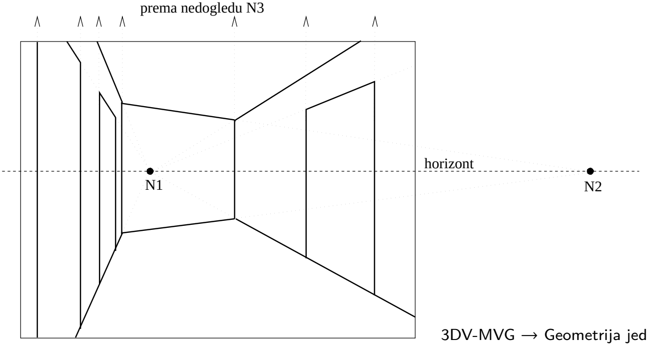
Nedogled kao presjecište interpretacijskih ravnina:



Geometrija jednog pogleda

Projekcija preslikava paralelne 3D pravce u 2D pravce koji se sijeku u **nedogledu**

- svakom 3D smjeru odgovara točno jedan nedogled!
- skup konvergentnih 2D pravaca može prokazati glavnu os strukture scene:
  1. pretp: 2D pravci odgovaraju paralelnim bridovima scene
  2. 3D smjer bridova je određen položajem nedogleda

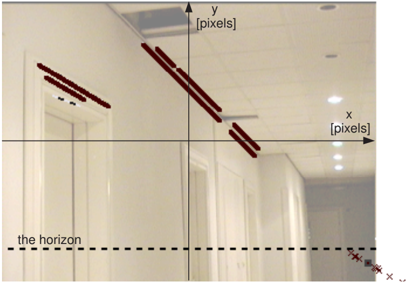


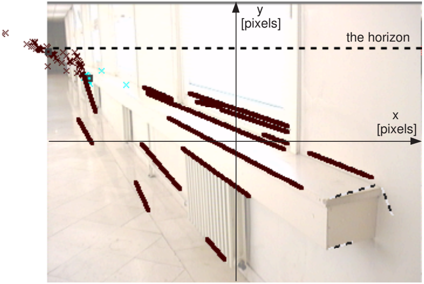
Geometrija jednog pogleda

Nedogled: rezultati detekcije

Geometrija jednog pogleda

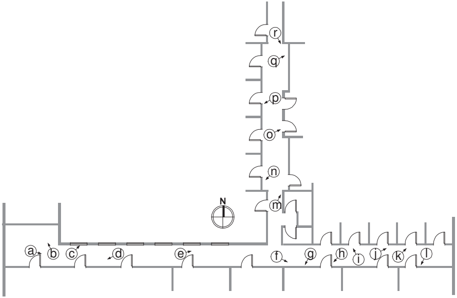
Nedogled: rezultati detekcije

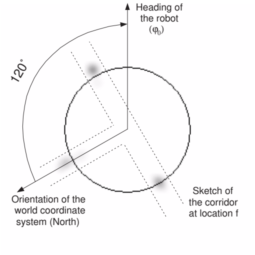


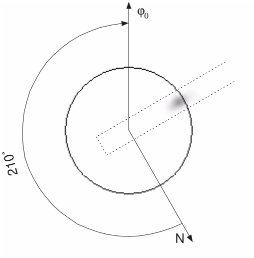


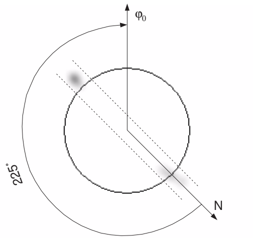
Geometrija jednog pogleda

Nedogled: rezultati lokalizacije na lokacijama **f**, **l** i **n**









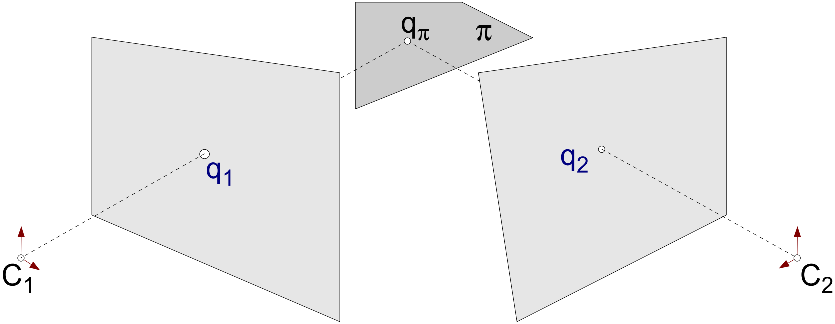
Dva pogleda na ravninsku scenu

Ravninske scene: homografija homografijine inverzije

$\mathbf{q}_1 = \mathbf{H}_1 \cdot \mathbf{q}_\pi$

$\mathbf{q}_2 = \mathbf{H}_2 \cdot \mathbf{q}_\pi$

$\Rightarrow \mathbf{q}_2 = \mathbf{H}_2 \cdot \mathbf{H}_1^{-1} \cdot \mathbf{q}_1 = \mathbf{H}_{21} \cdot \mathbf{q}_1$



Dva pogleda na ravninsku scenu

Ravninske scene: geometrijska interpretacija kalibrirane homografije:

$$\mathbf{P}_1 = \begin{bmatrix} \mathbf{I} & \mathbf{0} \end{bmatrix}, \mathbf{P}_2 = \begin{bmatrix} \mathbf{R} & \mathbf{t} \end{bmatrix}, \pi \equiv \mathbf{n}^\top \mathbf{Q}_1 = d$$

$$\begin{aligned}
\Rightarrow \mathbf{Q}_2 &= \mathbf{R}\mathbf{Q}_1 + \mathbf{t} \\
&= \mathbf{R}\mathbf{Q}_1 + \mathbf{t} \cdot \frac{\mathbf{n}^\top \mathbf{Q}_1}{d} \\
&= (\mathbf{R} + \frac{\mathbf{t} \cdot \mathbf{n}^\top}{d}) \cdot \mathbf{Q}_1
\end{aligned}$$

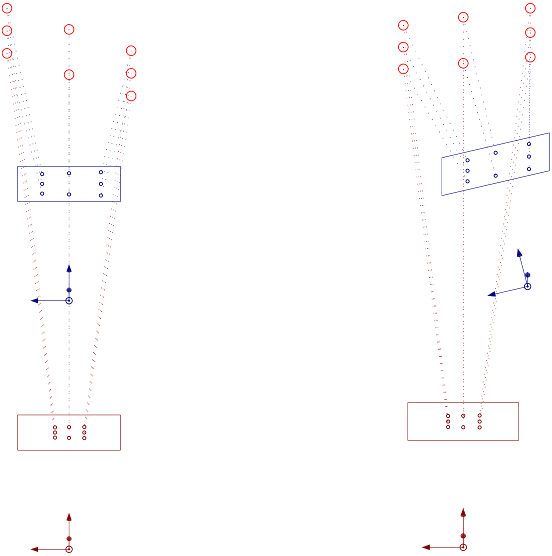
$$\Rightarrow \mathbf{H}_{21} = \mathbf{R} + \frac{\mathbf{t} \cdot \mathbf{n}^\top}{d}$$

Dekompozicija homografije: ako je zadan $\mathbf{H}_{21}$, odrediti $\mathbf{R}$, $\frac{\mathbf{t}}{d}$, $\mathbf{n}$
(obavezna kalibracija, rješenje nije uvijek jednoznačno [faugeras88ijprai])

Dva pogleda na ravninsku scenu

Nejednoznačnost dekompozicije homografije

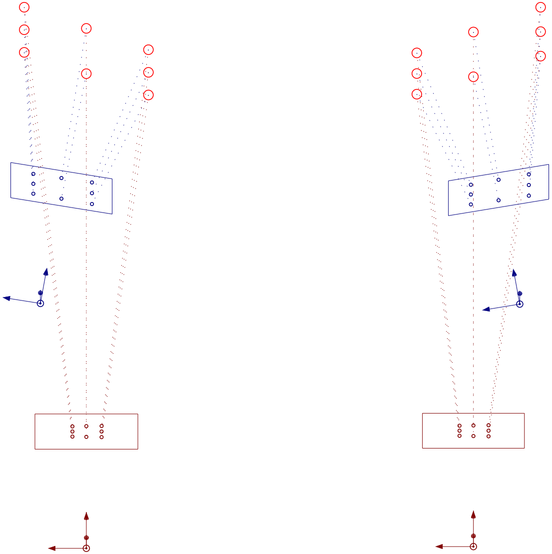
θ=00°:



Dva pogleda na ravninsku scenu

Nejednoznačnost dekompozicije homografije

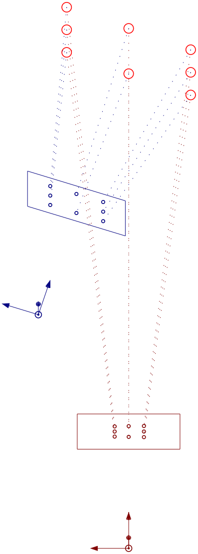
θ=10°:

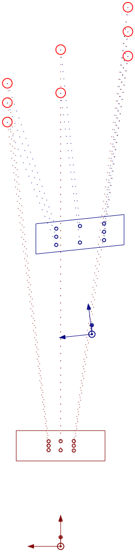


Dva pogleda na ravninsku scenu

Nejednoznačnost dekompozicije homografije

θ=20°:

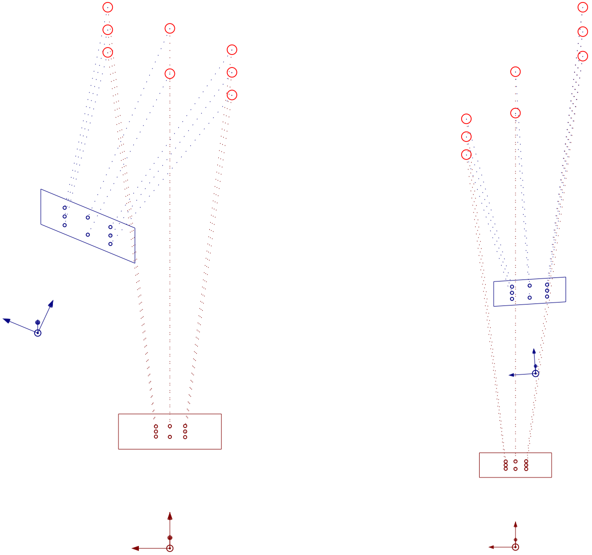




Dva pogleda na ravninsku scenu

Nejednoznačnost dekompozicije homografije

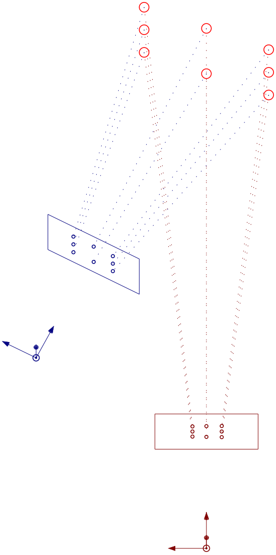
θ=30°:

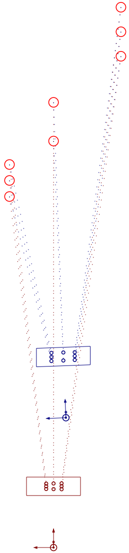


Dva pogleda na ravninsku scenu

Nejednoznačnost dekompozicije homografije

θ=40°:

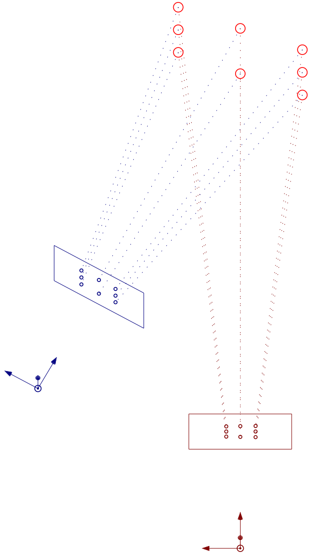


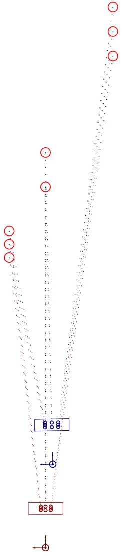


Dva pogleda na ravninsku scenu

Nejednoznačnost dekompozicije homografije

θ=50°:

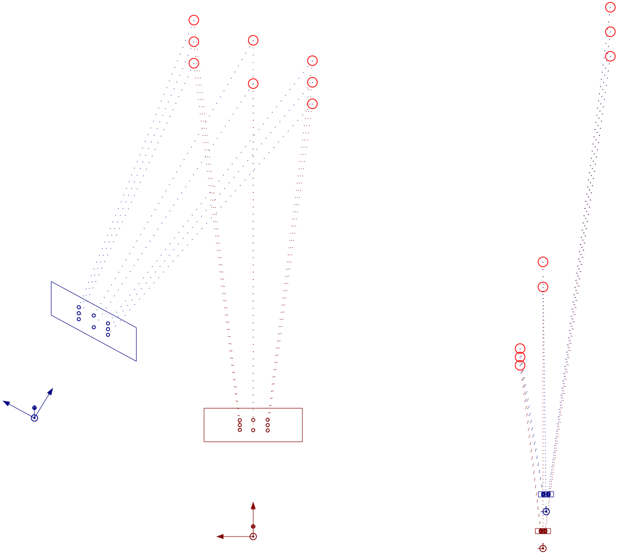




Dva pogleda na ravninsku scenu

Nejednoznačnost dekompozicije homografije

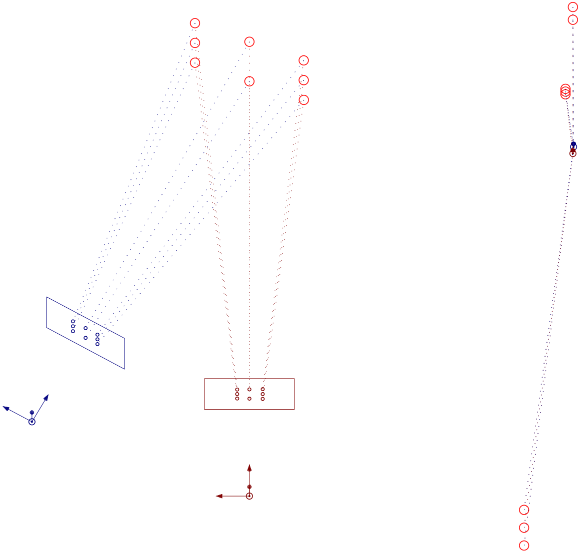
θ=60°:



Dva pogleda na ravninsku scenu

Nejednoznačnost dekompozicije homografije

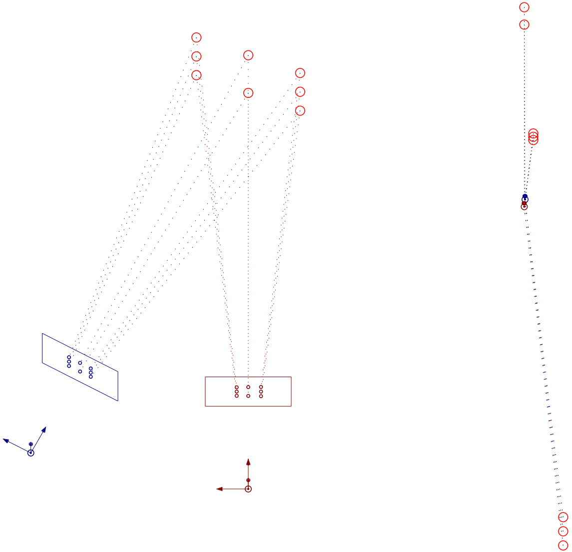
θ=70°:



Dva pogleda na ravninsku scenu

Nejednoznačnost dekompozicije homografije

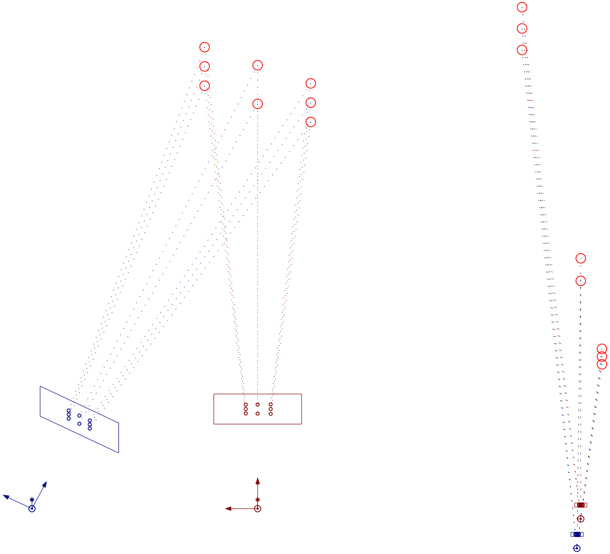
θ=80°:



Dva pogleda na ravninsku scenu

Nejednoznačnost dekompozicije homografije

θ=90°:



Stvarne kamere

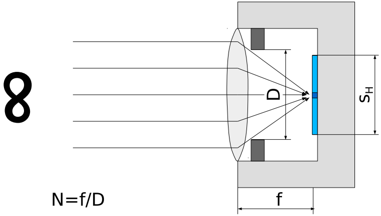
Stvarne kamere:

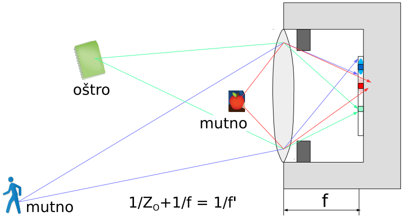
- svojstva senzora (pretvara fotone u naboj):
  - veličina $s_W \times s_H$ (npr. 2/3" $\Rightarrow$ 8.98 mm $\times$ 6.71 mm)
  - kvantna efikasnost (ICX285: 60% na 550 nm)
  - tehnologija (CMOS, CCD), rezolucija, bit/piksel, boja (siva, Bayer ili 3CCD), preplitanje, ...
- ostala svojstva stvarnih kamera:
  - izvedba zatvarača (rolling, global), ležište leće (C, F, ...), frekvencija (fps), sučelje, mogućnost okidanja, ...
- svojstva leće (prikuplja svjetlo na senzor):
  - udaljenost središta leće od senzora f (=$\|Oo\|$!)
  - format slike (veličina senzora)
  - dijametar otvora (D) odnosno F-broj (N=f/D)
  - standard ležišta (C, F, ...)

Stvarne kamere

Geometrija stvarne kamere:

- širina vidnog polja ovisi o veličini senzora i udaljenosti leće
- $\phi_{\mathrm{hfov}} = 2 \cdot \mathrm{arc\,tg}(\frac{s_w/2}{\mathrm{f}}),\ \mathrm{f} = \|Oo\|$
- leća prikuplja konus zraka ``ispucanih'' s objekta
- geometrijski, leća se "trudi" biti što manje vidljiva (izobličenja su neželjena)
- žarišna daljina leće $\mathrm{f}' \neq \mathrm{f}$ određuje gdje su objekti oštri
- što je D veći, vidno polje je ``pliće''



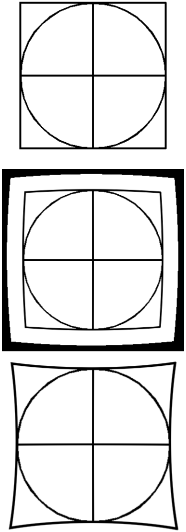


Stvarne kamere

Model radijalnog izobličenja:

- Stvarne kamere dobro aproksimiraju projekciju, ali izobličenja postoje:
  - neravnomjerni efektivni presjek leće, $S_{\mathrm{eff}} = S \cdot \cos(\varphi)$
  - **radijalno** i tangencijalno izobličenje
  - kromatska aberacija
- Model radijalnog izobličenja i korekcije:

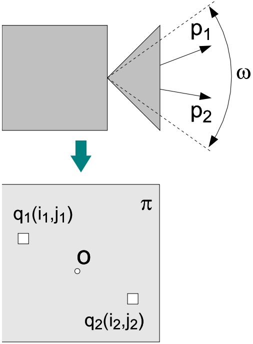
$$
\begin{aligned}
\hat{x}_d &= f_{ud}(\hat{x}_u) \\
\hat{x}_u &= f_{du}(\hat{x}_d) \\
f_{ab}(\hat{x}_a) &= \hat{x}_a \cdot (1 + k_1^{ab} \cdot \hat{r}_a^2 + k_2^{ab} \cdot \hat{r}_a^4) \\
\hat{r}_a^2 &= \hat{x}_a^2 + \hat{y}_a^2
\end{aligned}
$$



Umjeravanje perspektivne kamere

Kamera kao crna kutija:

- Stvaranje slike je **preslikavanje** $f: \mathbb{R}^3 \rightarrow \mathbb{R}^2$
  - $\mathcal{D}(f)$ - Euklidski 3D prostor $\mathbb{R}^3$
  - $\mathcal{K}(f)$ - slikovna ravnina π
- Zbog pravocrtnog širenja svjetlosti:
  $\forall y \in \pi \; \exists p \in \mathcal{P} : f(x) = y, \forall x \in p$
  - $\mathcal{P}$ - skup pravaca iz $\mathcal{D}(f)$
  - $p \in \mathcal{P}$ definira zraku koja podražava odgovarajući senzorski element
- Problem umjeravanja (kalibracije):
  - **inverzno** preslikavanje $\kappa : \pi \rightarrow \mathcal{P}$?
  - ocijeniti preciznost postupka



$\mathcal{P}$ ... skup pravaca scene

$p_1, p_2 \in \mathcal{P}$; $q_1, q_2 \in \pi$;

kalibracija $\equiv \kappa : \pi \rightarrow \mathcal{P}$

$\kappa(q_1) = p_1$; $\kappa(q_2) = p_2$

Umjeravanje perspektivne kamere

Umjeravanje s ravninskim uzorkom [Zhang00]:

- Koristi se više pogleda na ravninski uzorak (pravokutna mreža prepoznatljivih značajki)
- Dosjetljivom linearnom metodom određuju se:
  - linearni intrinsični parametri kamere ($\mathbf{K}$)
  - vanjski parametri za svaki pogled ($\mathbf{R}, \mathbf{t}$)
- Parametri radijalnog izobličenja određuju se optimizacijom
  - kriterij optimiranja je ukupna **projekcijska pogreška**!
  - dimenzionalnost funkcije cilja: $n_{\text{pogleda}} \cdot n_{\text{znacajki}} \cdot 2 \cdot 2$
  - optimiraju se **svi** intrinsični parametri i parametri pomaka kamere za svaki pogled: $9 + n_{\text{pogleda}} \cdot 6$
    (dotjeruju se rezultati linearne analize)
  - postupak poznat pod imenom *bundle adjustment*

Umjeravanje perspektivne kamere

Još o nelinearnoj optimizaciji:

- U postupak ulaze:
  - liste uparenih točaka $\{\mathbf{Q}_{Rij}\}, \{\mathbf{q}_{ij}\}, i = 1..n_p, j = 1..n_{zi}$
  - inicijalne procjene parametara pomaka: $\{\mathbf{R}_i\},\{\mathbf{t}_i\}, i = 1..n_p$
  - inicijalna procjena intrinsičnih parametara: $\mathbf{K}$, $(\mathbf{k}_1^{ud}, \mathbf{k}_2^{ud})$, $(\mathbf{k}_1^{du}, \mathbf{k}_2^{du})$,
- Minimizira se norma funkcije $\mathrm{fcn} : \mathbb{R}^{9+6n_p} \longrightarrow \mathbb{R}^{4n_p\bar{n}_{zi}}$

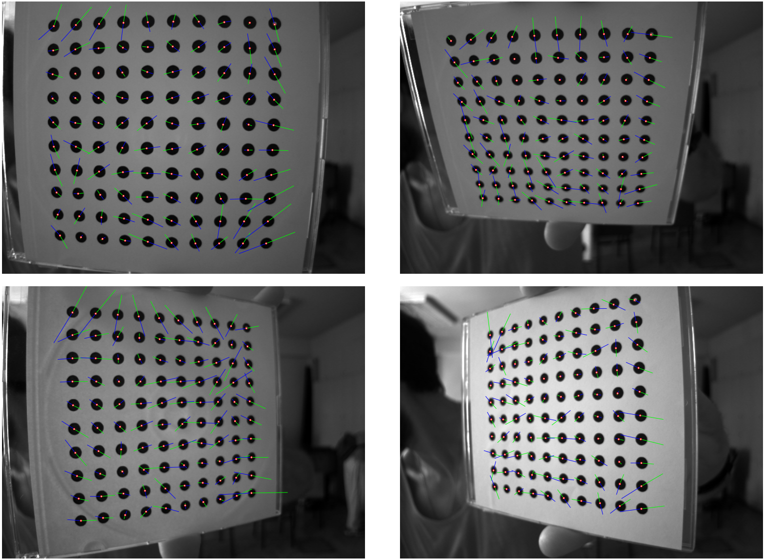
$$\mathrm{fcn}(\{\mathbf{R}_i\}, \{\mathbf{t}_i\}, \mathbf{K}, \{\mathbf{k}_l^k\}) = \{\delta_{ij}^d, \delta_{ij}^u\}, i = 1..n_p, j = 1..n_{zi}$$

$$\delta_{ij}^d = \mathrm{d}_{\mathbb{R}^2}(\mathbf{K}\mathbf{f}_{ud}(\mathbf{S}\,\mathbf{M}_i\mathbf{Q}_{Rij}), \mathbf{q}_{ij})^2$$

$$\delta_{ij}^u = \mathrm{d}_{\mathbb{R}^2}(\mathbf{K}\mathbf{S}\,\mathbf{M}_i\mathbf{Q}_{Rij}, \mathbf{K}\mathbf{f}_{du}(\mathbf{K}^{-1}\mathbf{q}_{ij}))^2$$

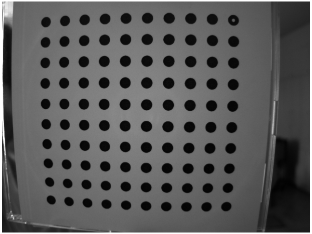
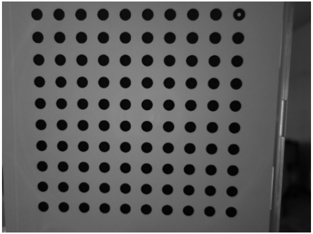
UMJERAVANJE PERSPEKTIVNE KAMERE

Projekcijska pogreška (×100)



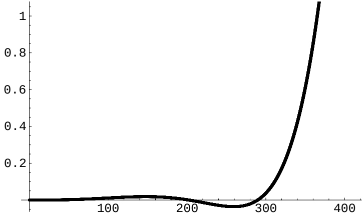
Umjeravanje perspektivne kamere

- Korekcija radijalnog izobličenja:

 $\longrightarrow$ 

- Preciznost parametara radijalnog izobličenja

$$f_{\varepsilon} = r - f_{du}(f_{ud}(r))$$



Umjeravanje perspektivne kamere

Zaključak:

- Optimizacijski postupak pronalazi optimalne vrijednosti unutrašnjih i vanjskih (za svaki pogled) parametara kamere: $\mathbf{K}, \{\mathbf{k}_l^k\}, \{\mathbf{R}_i\}, \{\mathbf{t}_i\}$
- ukupna projekcijska pogreška modela $\|\mathrm{fcn}()\|_2$ obuhvaća:
  - pogrešku δ_S^d u izobličenoj slici (kalibriramo $\mathbf{f}_{ud}$)
  - pogrešku δ_S^u u idealnoj slici (kalibriramo $\mathbf{f}_{du}$)
- Rješenje **ocjenjujemo** parametrima projekcijske pogreške u slikovnoj ravnini δ_S^d:
  - standardna devijacija pojedinačnih odstupanja
  - najveća pojedinačna projekcijska pogreška

Umjeravanje perspektivne kamere

Primjene 3D vida: 3D modeliranje, proširena stvarnost, autonomna navigacija, detekcija objekata, ...

Velika kočnica: procesna moć današnjih računala.