Vizuális helymeghatározás

Többnézetes geometria

Plósz Sándor[1]

[1]Budapesti Műszaki és Gazdaságtudományi Egyetem
Távközlési és Médiainformatikai Tanszék

2015

Tartalom

- Ismétlés
- Vizuális helymeghatározás elve
- Projektív geometria alapjai
- Kamerák reprezentálása
- Többnézetes geometria, fundamentális mátrix
- Fundamentális mátrix meghatározása
- Térépítés, trianguláció

Ismétlés

Lineáris egyenletek - SVD felbontás (1)

- Lineáris egyenletek: $\mathbf{y} = f(\mathbf{x}) = A\mathbf{x} \qquad A \in \mathbb{C}^{m \times n}$
- Minden $\mathbf{A}_{M \times N}$ mátrixra létezik egy ún. SVD (Singular Value Decomposition) felbontás

$$\mathbf{A}_{M \times N} = \mathbf{U}_{M \times M} \mathbf{\Sigma}_{M \times N} \mathbf{V}^H_{N \times N}$$

- $\mathbf{U}$ és $\mathbf{V}$ unitér (valós esetben ortogonális) mátrixok ($\mathbf{U}\mathbf{U}^H = \mathbf{U}^H\mathbf{U} = \mathbf{I}$ és $\mathbf{V}\mathbf{V}^H = \mathbf{V}^H\mathbf{V} = \mathbf{I}$)
- Σ mátrixban a főátlóban nemnegatív valós számok találhatók, máshol nulla
  - A főátló elemeit hívjuk szinguláris értékeknek (σ_n)

Ismétlés

Lineáris egyenletek - SVD felbontás (2)

- Nem négyzetes esetben tételezzük fel, hogy $A = UDV^T$ az SVD felbontás szerint
- Minimalizáljuk a $\|A\mathbf{x} - \mathbf{y}\|$ kifejezést!

$$\|A\mathbf{x} - \mathbf{y}\| = \|UDV^T\mathbf{x} - \mathbf{y}\| = \|DV^T\mathbf{x} - U^T\mathbf{y}\|$$

- Legyen $\mathbf{y}' = U^T\mathbf{y}$ és $\mathbf{x}' = V^T\mathbf{x}$!
- Ekkor a $\|D\mathbf{x}' - \mathbf{y}'\|$ kifejezést kell minimalizálnunk

$$\underset{m \times n}{\overset{X}{\begin{pmatrix} x_{11} & x_{12} & \cdots & x_{1n} \\ x_{21} & x_{22} & \cdots & \\ \vdots & \vdots & \ddots & \\ x_{m1} & & & x_{mn} \end{pmatrix}}} = \underset{m \times r}{\overset{U}{\begin{pmatrix} u_{11} & \cdots & u_{1r} \\ \vdots & \ddots & \\ u_{m1} & & u_{mr} \end{pmatrix}}} \underset{r \times r}{\overset{S}{\begin{pmatrix} s_{11} & 0 & \cdots \\ 0 & \ddots & \\ \vdots & & s_{rr} \end{pmatrix}}} \underset{r \times n}{\overset{V^{\mathsf{T}}}{\begin{pmatrix} v_{11} & \cdots & v_{1n} \\ \vdots & \ddots & \\ v_{r1} & & v_{rn} \end{pmatrix}}}$$

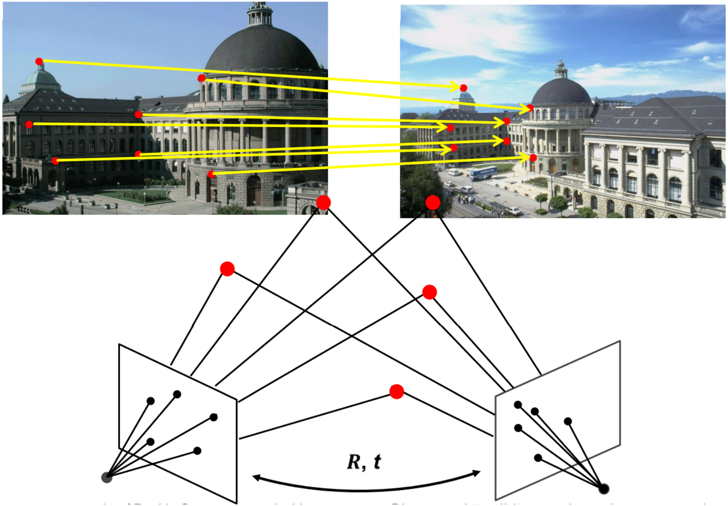
Bevezetés

Vizuális helymeghatározás lépései

- Feltérképezés:
  - *referencia* képek készítése, tartózkodási hely (C) és orientáció (R) rögzítése
  - Megalkotni minden referencia képhez a P_r referencia kamera mátrixot.
  - A kamerák belső paramétereit (K mátrix) számítással vagy kalibrációval meg kell határozni
- A tesztkép összehasonlítása referencia képekkel (valamely párosító algoritmussal), megkeresni a legjobb párt
- Összetartozó pontpárok kinyerése, ezek alapján kamera relációt leíró (fundamentális) mátrix meghatározása
- Ebből relatív hely és orientáció meghatározása
- Optimalizálás, 3d trianguláció, térépítés

Bevezetés

Kétnézetes geometria

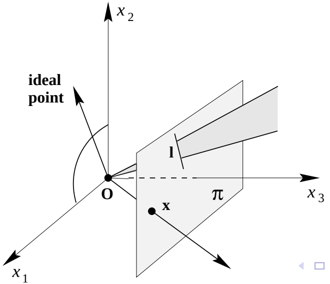


A projektív geometria alapjai

- Az n dimenziós projektív geometria n dimenziós pontokat $n+1$ dimenzióban ír le.
- Pontok leírása
  - $\pi : x_3 = 1$ egy síkot definiál, mely x_1-x_2 síkkal párhuzamos
  - Tetszőleges síkon található **x** pont leírható az O és x pontokon átmenő egyenes egyenletével:

$$ax + by + c = 0$$

- Az egyenes paraméterei: a, b, c



A projektív geometria alapjai

Projektív tér

- Egyenes konstansal való szorzással nem változik:

$$(a, b, c) \leftrightarrow (ka, kb, kc), k \neq 0$$

- $\mathbf{x} = (x, y)$ pont akkor van az egyenesen, ha teljesíti:

$$\begin{bmatrix} x & y & 1 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = 0$$

- (kx, ky, k) ugyanazt azt az $\mathbf{x}$ síkbeli pontot jelöli $\rightarrow$ pont homogén koordinátája
- Egy általános $\mathbf{x} = (x_1, x_2, x_3)$ pont a sík $(x_1/x_3, x_2/x_3)$ pontjának felel meg
- $\mathbb{P}^2$ *két dimenziós projektív tér* $\mathbb{R}^3 - (0, 0, 0)$ felett (valós számhármasok kivéve a nullvektor)

A projektív geometria alapjai

Projektív geometria tulajdonságai

- Két egyenes l, l' metszéspontja

$$\mathbf{x} = l \times l' = [l]_\times l'$$

- Hiszen $l(l \times l') = l'(l \times l') = 0$
- $l \times l'$ a két vektor vektoriális szorzata

$$a \times b = \begin{vmatrix} i & j & k \\ a_x & a_y & a_z \\ b_x & b_y & b_z \end{vmatrix} = \left(\begin{vmatrix} a_y & a_z \\ b_y & b_z \end{vmatrix}, \begin{vmatrix} a_x & a_z \\ b_x & b_z \end{vmatrix}, \begin{vmatrix} a_x & a_y \\ b_x & b_y \end{vmatrix} \right)$$

- $[l]_\times$ a vektoriális szorzás mátrix alakban

$$[a]_\times = \begin{pmatrix} 0 & -a_z & a_y \\ a_z & 0 & -a_x \\ -a_y & a_x & 0 \end{pmatrix}$$

A projektív geometria alapjai

Projektív geometria tulajdonságai

- Két párhuzamos egyenes

$$l : (a, b, c) \quad l' : (a, b, c')$$

- Metszéspontjuk:

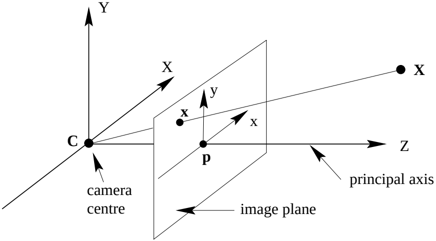
$$l \times l' = (c' - c)(b, -a, 0)^T$$

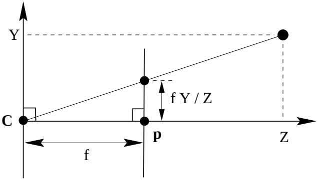
- Átváltva normál koordinátákra $(d = (c' - c))$

$$l \times l' = d(b/0, -a/0)^T$$

- A pont koordinátái végtelen nagyok $\rightarrow$ projektív térben a párhuzamosak a végtelenben találkoznak!

Vetítés





- Van egy $\mathbf{X}(x, y, z)$ pontunk a térben és egy **f** paraméterű kameránk a **C** pontban, mely Z tengely felé néz
- Hol látszik **X** vetülete a kameránkon (x', y')?
- Egyszerű arány: $\frac{y}{z} = \frac{y'}{f} \rightarrow y' = \frac{fy}{z}$
- Ugyanígy: $x' = \frac{fx}{z}$

$$(x, y, z)^T \mapsto (fx/z, fy/z)^T$$

Vetítés

- Kamera síkjára való vetítés:

$$y' = \frac{fy}{z} = fy\frac{1}{z} \qquad x' = \frac{fx}{z} = fy\frac{1}{z}$$

- Mivel az osztás műveletét nem tudjuk felírni mátrix alakban, ezért bevezették a homogén koordinátákat
- A hányados egy újabb koordináta

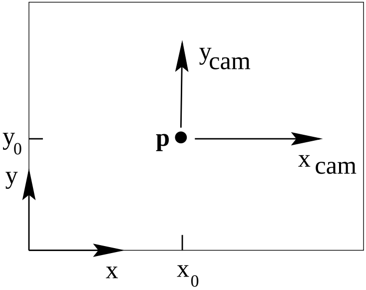
$$\begin{bmatrix} x' \\ y' \\ z' \\ 1 \end{bmatrix} \mapsto \begin{bmatrix} fx \\ fy \\ z \end{bmatrix} = \begin{bmatrix} f & & & 0 \\ & f & & 0 \\ & & 1 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$$

- Másképpen $P = \mathrm{diag}(f, f, 1)\,[I|\mathbf{0}]$
- Ahol P mátrix az ún. vetítési (projekciós) mátrix

Kamera belső paraméterei

- A valóságban f-et valamilyen fizikai mértékben adják meg (pl. mm)
- A képfeldolgozás során pixel koordináta rendszerben dolgozunk
- f-et át kell transzformálálni pixel koordinátarendszerbe
- $r = (r_x, r_y)$ a szenzor felbontása, $d = (d_x, d_y)$ a szenzor mérete [mm]
- $\alpha_x = r_x/d_x, \quad \alpha_y = r_y/d_y \quad [pixel/mm]$

$$K = \begin{bmatrix} f\alpha_x & 0 & x_0 \\ 0 & f\alpha_y & y_0 \\ 0 & 0 & 1 \end{bmatrix}$$



- x_0, y_0 a kamera origója
- K mátrixot a kamera belső (*intrinsic*) paramétereinek hívjuk

Kamera külső paraméterei

- A kamera a térben nem feltétlenül a z koordinátatengely irányába néz
- Ezt egy R forgatási mátrixszal és egy C középponttal (eltolás) jellemezhetjük
- A megfelelő külső transzformációt (*extrinsic parameters*) a következőképp kapjuk:

$$P_{ex} = \begin{bmatrix} R & -RC \\ 0 & 1 \end{bmatrix}$$

- A külső és belső paraméter mátrixok szorzata adja a teljes P kamera mátrixot:

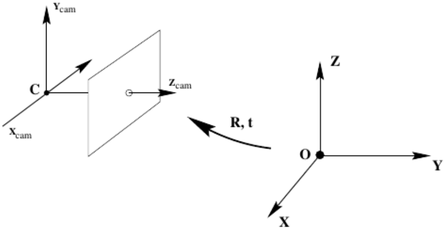
$$P = K[R| - RC] = KR[I| - C]$$

Kamera transzformációja

- Kamera projekció egyenlete:

$$P = K[R| - RC] = KR[I| - C]$$

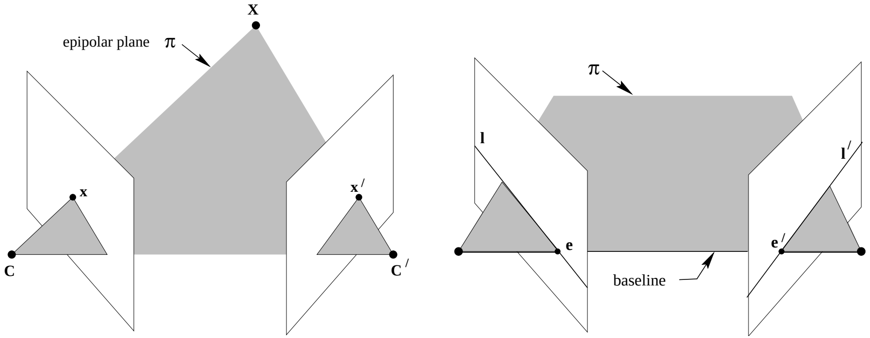
- Lépések:
  1. Eltolás $\rightarrow \mathbf{x}' = (\mathbf{x} - \mathbf{C})$
  2. Forgatás $\rightarrow \mathbf{x}'' = (\mathbf{R}\mathbf{x}')$
  3. Vetítés $\rightarrow \mathbf{x}''' = (\mathbf{K}\mathbf{x}'')$



Epipoláris geometria

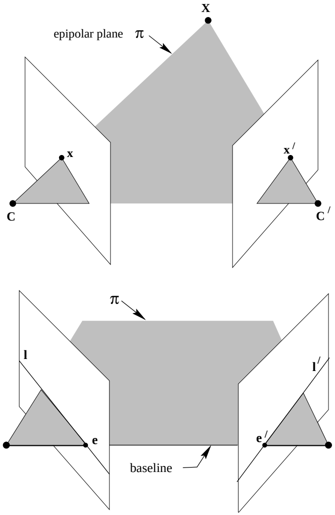
Bevezetés

- Epipoláris: kétnézetes
- Egy térbeli pont, és két vetített képe között milyen összefüggés van



Epipoláris geometria
Fogalmak

- Két kamera C és C' középponttal
- X térbeli pont, képe a két kamerán rendre x és x'
- C, C', x, x' és X pontok egy ún. π *epipoláris síkra* esnek
- A C és C' pontot összekötő egyenest *alapegyenesnek* hívjuk, a képsíkokat e és e' pontban metszik (epipólus)
- Az e és x pontok az l *epipoláris egyenesre* esnek, ahogy e' és x' az l' epipoláris egyenesre.



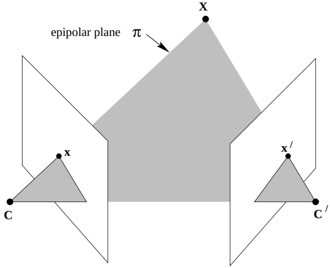
Epipoláris geometria

Epipoláris egyenlet

- Tudjuk, hogy $PX = x$, amiből X-et meghatározhatjuk a köv. egyenlet megoldásával:

$$X(\lambda) = P^{+}x + \lambda C$$

- P^{+} a P mátrix pszeudoinverze ($PP^{+} = I$)
- C a kamera középpont, (P nullvektora, hiszen $PC = 0$)
  - A megoldás egy egyenes, két kiemelt pont $P^{+}x$ ($\lambda = 0$) és C ($\lambda = \inf$)
  - Ezen pontok képei a 2. kamerán: $P'P^{+}x$ és $P'C = e'$



Epipoláris geometria

Fundamentális mátrix

- Az előzőek alapján:

$$l' = (P'C) \times (P'P^{+}x) = e' \times (P'P^{+}x) = [e']_{\times}(P'P^{+}x) = Fx$$

- $[e']_{\times}$ az e' vektorral történő keresztszorzatot jelentő mátrix
- Az F mátrixot *fundamentális mátrixnak* nevezzük
- F kapcsolatot ad a két nézet között, egy adott nézetben levetíttet ponthoz hozzárendel a másik nézetben egy egyenest.
- Mivel az x' pont az l' egyenesen fekszik, ezért minden x és x' pont között fennáll

$$x'^{T}Fx = 0$$

Epipoláris geometria

Fundamentális mátrix

- F rangja 2, és 7 szabadságfokkal bír
- Bármely két nézetet leírhatunk egy F mátrixal, úgy, hogy a vetített pontpárokra érvényes az előző egyenlet
- Bármely két kamerához egyértelműen tartozik egy fundamentális mátrix, de fordítva nem feltétlenül igaz!
- Minden kamerapárhoz, amelyek csak egy projektív transzformációban különböznek ugyanaz a fundamentális mátrix tartozik

Epipoláris geometria

Esszenciális mátrix

- Legyen $P = K[R| - RC] = KR[I| - C]$ formájú
- Vezessük be minden x pontra az ún. normalizált $\hat{x} = K^{-1}x$ koordinátákat
- Belátható, hogy minden $\hat{x} \leftrightarrow \hat{x}'$ pontpárra igaz:

$$\hat{x}'^T E\hat{x} = 0$$

- Az E mátrixot *esszenciális mátrixnak* hívjuk, és:

$$E = K'^T FK = [\mathbf{t}]_\times R.$$

- Amennyiben a két kameramátrix felépítése $P = K[I|0]$ és $P' = K[R|\mathbf{t}]$.

Epipoláris geometria
Esszenciális mátrix

- E mátrix szabadságfoka 5 $\rightarrow$ több megszorítás vonatkozik rá, mint a fundamentális mátrixra
- Belátható, hogy amennyiben egy kameramátrixhoz tartozó esszenciális mátrix ismert, akkor abból a kamera mátrixokra négy megoldás található.
- Ebből a négy megoldásból kiszűrhető az az egy, amelyik szerint az X pontok mind a kamerák nézeti irányában (s nem a kamerák mögött) találhatók.
- A pontos megoldási módszer az SVD (Singular Value Decomposition) felbontáson alapul

Fundamentális mátrix meghatározása

- Két különböző kamerával, különböző helyről rögzített képek megfelelő $\mathbf{x} \leftrightarrow \mathbf{x}'$ képpontpárjai között a köv. epipoláris megkötés áll fenn:

$$\mathbf{x}'^T F_{3x3} \mathbf{x} = 0$$

- F_{3x3} a szinguláris fundamentális mátrix (rangja 2).
- Meghatározása:
  - Kezdeti becslés
  - Hiba felírása
  - Költségfüggvény optimalizálás

Fundamentális mátrix meghatározása

Költségfüggvény

- *Szimmetrikus epipoláris hiba* költségfüggvény
- Tudjuk, hogy az egyik képen az $\mathbf{x}$ ponthoz a másik képen egy $\mathbf{l}$ egyenes tartozik, amelyekre:

$$\mathbf{l}' = F\mathbf{x}$$

- Hasonlóan az $\mathbf{x}'$ pontokra:

$$\mathbf{l} = F^T\mathbf{x}' \tag{1}$$

- A szimmetrikus epipoláris hiba így az egyenesek és pontok távolságát összegzi:

$$\sum_i d(\mathbf{x}_i', F\mathbf{x}_i)^2 + d(\mathbf{x}_i, F^T\mathbf{x}_i')^2 \tag{2}$$

- ahol $d(\mathbf{x}, \mathbf{y})$ az $\mathbf{y}$ egyenes és a $\mathbf{x}$ pont euklideszi távolságát jelenti.

Fundamentális mátrix meghatározása

Becslés

- Ha a mért képpontok normális eloszlású zajjal terheltek alkalmazható a maximum likelihood (ML) becslés
- Zajjal terhelt pontpároknak megfelelő $\hat{\mathbf{x}}_i \leftrightarrow \hat{\mathbf{x}}'_i$ pontpárok kielégítik:

$$\hat{\mathbf{x}'}_i^T F \hat{\mathbf{x}}_i = 0$$

- A *reprojekciós hiba* költségfüggvény:

$$\sum_i d(\mathbf{x}_i, \hat{\mathbf{x}}_i)^2 + d(\mathbf{x}'_i, \hat{\mathbf{x}'}_i)^2 \qquad (3)$$

Fundamentális mátrix meghatározása

8 pontos módszer

- $\mathbf{x}'^T F_{3x3} \mathbf{x} = 0$ egyenlet az F lineáris függvénye
- F mátrixra lineáris egyenletrendszer írható fel, pontonként egy egyenlettel
- Legyen $\mathbf{x}_i = (x, y, 1)$ és $\mathbf{x}'_i = (x', y', 1)$, ekkor $\forall(\mathbf{x}_i, \mathbf{x}'_i)$-re:

$$x'xf_{11} + x'yf_{12} + x'f_{13} + y'xf_{21} + y'yf_{22} + y'f_{23} + xf_{31} + yf_{32} + f_{33} = 0$$

- Legyen $\mathbf{f}^T = (f_{11}, f_{12}, f_{13}, f_{21}, f_{22}, f_{23}, f_{31}, f_{32}, f_{33})$

$$\mathbf{Af} = \begin{bmatrix} x'_1x_1 & x'_1y_1 & x'_1 & y'_1x_1 & y'_1y_1 & y'_1 & x_1 & y_1 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x'_nx_n & x'_ny_n & x'_n & y'_nx_n & y'_ny_n & y'_n & x_n & y_n & 1 \end{bmatrix} \mathbf{f} = 0$$

- F nem feltétlenül lesz szinguláris!

Fundamentális mátrix meghatározása

8 pontos módszer

- Legtöbbször több mint 8 pontunk van
- Legkisebb négyzetes megoldás elve: legkisebb szingularitáshoz tartozó szinguláris vektor
- Ha A SVD felbontása $A = UDV^T$, akkor a megoldás V_i^T amennyiben $\sigma_i = \min_k \sigma_k$
- F szinguláris felbontása $F = U\mathrm{diag}(r, s, 0)V^T$ formájú, hiszen a rangja 2
- Célszerű $F = U\mathrm{diag}(r, s, t)V^T$ helyett az $F' = U\mathrm{diag}(r, s, 0)V^T$ mátrix használata

Fundamentális mátrix meghatározása

7 pontos módszer

- F fundamentális mátrix szabadsági foka 7
- A mátrix 9 oszloppal rendelkezik, ha 7 pontból építjük fel magtere két dimenziós lesz
- Legyen e kétdimenziós térnek a bázisa $\mathbf{f}_1$ és $\mathbf{f}_2$, amelyeket az A mátrix SVD felbontásából kaphatunk

$$F = F_1 + \alpha F_2 \tag{4}$$

- F mátrix az $A\mathbf{f} = 0$ homogén egyenletnek megfelelő mátrix, α ismeretlen szorzó. A köv. egyenletet kell megoldani:

$$\det(F_1 + \alpha F_2) = 0 \tag{5}$$

- Ez már zárt alakban megolható!
- Előnye: 7 pont elég, és F szinguláris lesz

Fundamentális mátrix meghatározása

Optimális megoldás

- Kezdeti becslésre jó a normalizált 8 pontos és 7 pontos algoritmus
- ML becsléshez reprojekciós hiba költségfüggvény minimalizálása kell
- A legkézenfekvőbb megoldás a tér projektív rekonstrukciója
  - F-nek megfelelő P és P' kameramátrixok meghatározása
  - Pontok triangulációja
  - Optimalizáció
- Segítségével a kamerák relatív helye meghatározható $\rightarrow$ skálázás erejéig

Esszenciális mátrix meghatározása

- Ha F megvan $\rightarrow E = K^T F K$
- Ha $P = K[I|0] \rightarrow P' = K[R|t]$ hogy kapjuk meg?
- Legyen az esszenciális mátrix SVD felbontása $E = U\mathrm{diag}(1,1,0)V^T$. Definiáljuk:

$$W = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad Z = \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

- Ha $P = K[I|0]$, akkor P' kameramátrixra négy megoldás adódik:

$$P'_1 = \begin{bmatrix} UWV^T & +\mathbf{u}_3 \end{bmatrix} \qquad P'_3 = \begin{bmatrix} UW^TV^T & +\mathbf{u}_3 \end{bmatrix}$$
$$P'_2 = \begin{bmatrix} UWV^T & -\mathbf{u}_3 \end{bmatrix} \qquad P'_4 = \begin{bmatrix} UW^TV^T & -\mathbf{u}_3 \end{bmatrix}$$

- Csak egy megoldás helyes (pontok mindkét kamera előtt)

Trianguláció

Bevezetés

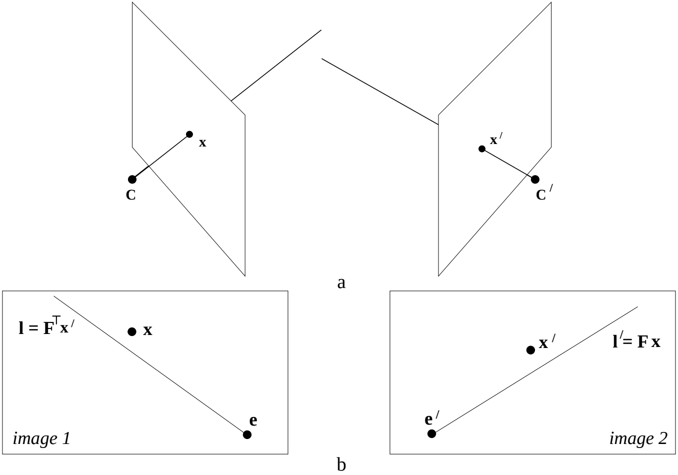
- Legyen P és P' a két kamera mátrixunk
- Legyen $\mathbf{X}$ térbeli pont, mely képe $\mathbf{x} = P\mathbf{X}$ és $\mathbf{x}' = P'\mathbf{X}$
- A trianguláció ennek megfordítása, tehát ismerve $\mathbf{x} \leftrightarrow \mathbf{x}'$ pontpárokat, melyik az az $\mathbf{X}$ pont, amelyre $\mathbf{x} = P\mathbf{X}$ és $\mathbf{x}' = P'\mathbf{X}$?
- A feladat a következő τ trianguláció megtalálása:

$$\mathbf{X} = \tau(\mathbf{x}, \mathbf{x}', P, P')$$

- Csak zajmentes esetben lenne egyértelmű

Trianguláció

Zaj szemléltetése



Trianguláció

Egyenletek

- A következő egyenleteket írhatjuk fel:

$$\mathbf{x} \times (P\mathbf{X}) = 0$$
$$\mathbf{x}' \times (P'\mathbf{X}) = 0$$

- Az első egyenletet felbonthatjuk az alábbi formában ($\mathbf{p}^{iT}$ a P mátrix i-dik sorát jelöli)

$$x(\mathbf{p}^{3T}\mathbf{X}) - (\mathbf{p}^{1T}\mathbf{X}) = 0$$
$$y(\mathbf{p}^{3T}\mathbf{X}) - (\mathbf{p}^{2T}\mathbf{X}) = 0$$
$$x(\mathbf{p}^{2T}\mathbf{X}) - y(\mathbf{p}^{1T}\mathbf{X}) = 0$$

- Lineáris egyenletrendszer

Trianguláció

Megoldás

- A másik keresztszorzatot hasonlóan felbontva összesen hat egyenletet kapunk, amelyeket összevonhatunk $A\mathbf{X} = 0$ alakra

$$A = \begin{bmatrix} x\mathbf{p}^{3T} - \mathbf{p}^{1T} \\ y\mathbf{p}^{3T} - \mathbf{p}^{2T} \\ x\mathbf{p}^{2T} - \mathbf{p}^{1T} \\ x'\mathbf{p}'^{3T} - \mathbf{p}'^{1T} \\ y'\mathbf{p}'^{3T} - \mathbf{p}'^{2T} \\ x'\mathbf{p}'^{2T} - \mathbf{p}'^{1T} \end{bmatrix}$$

- Megoldása a legkisebb négyzetes hiba minimalizálásával (a legkisebb szinguláris értékhez tartozó szinguláris vektor)
- A mátrix szinguláris felbontásával (SVD) kaphatjuk meg

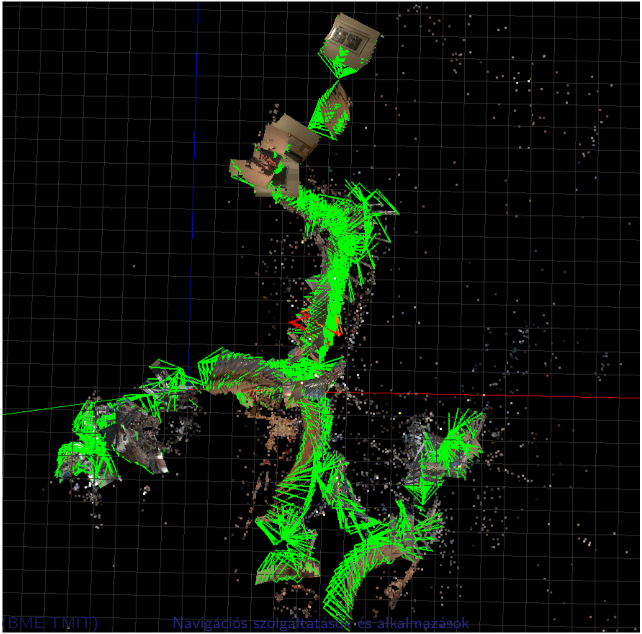
Optimalizáció

- a $\mathbf{x}_i \leftrightarrow \mathbf{x}'_i$ pontpárok ismeretében az $\mathbf{X}_i$ három dimenziós pontok meghatározhatók triangulációval
- A feladat innentől a következő egyenletrendszer reprojekciós hibájának nemlineáris optimalizációja $\mathbf{X}_i$ és P' függvényében:

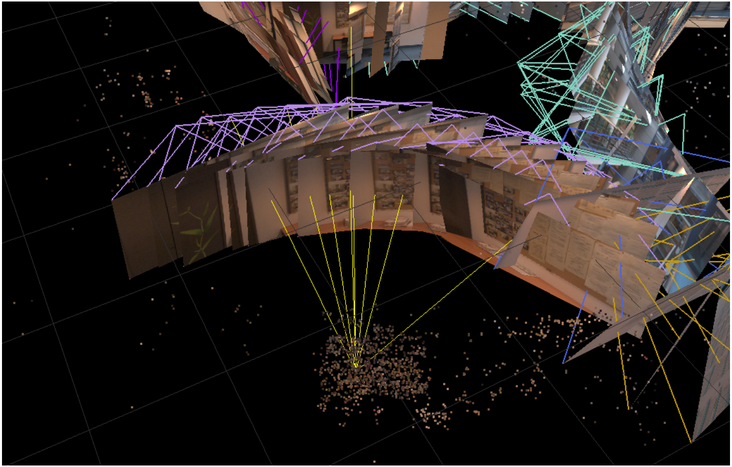
$$\begin{aligned} \mathbf{x}_i &= P\mathbf{X}_i \\ \mathbf{x}'_i &= P'\mathbf{X}_i \end{aligned} \tag{6}$$

- Összesen $3n + 12$ változó, ahol n a 3d pontpárok száma (ritka Levenberg-Marquardt minimalizáció)

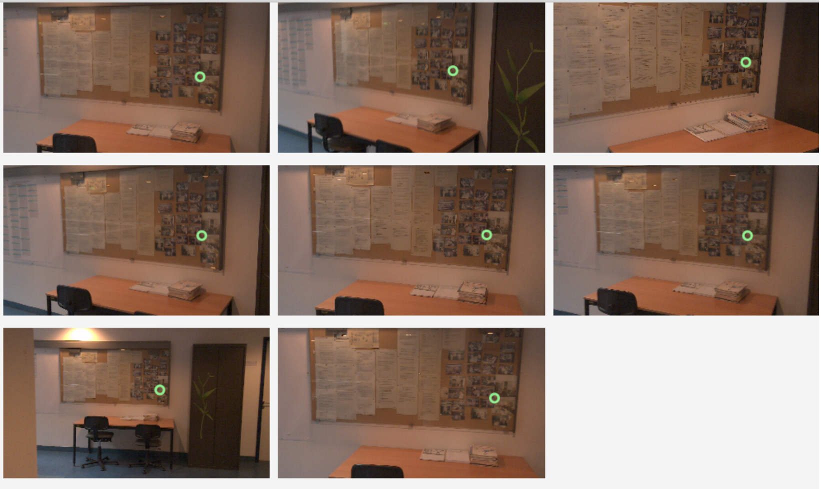
Térmodell



Térmodell



Térmodell



Forrás

- Richard Hartley and Andrew Zisserman, "Multiple View Geometry in Computer Vision", 2nd edition, 2004