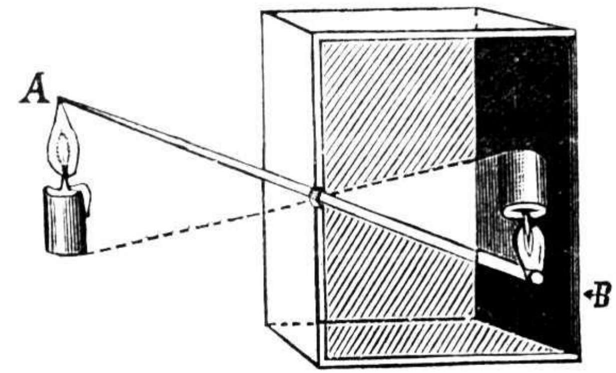
1. A projektív kamera

Kató Zoltán

Képfeldolgozás és Számítógépes Grafika tanszék
SZTE
(http://www.inf.u-szeged.hu/~kato/teaching/)

Perspektív leképezés



Camera obscura (Fizyka, 1910)

- **Kamera koordinátarendszer:**
  - **C origó: vetítési középpont, *fókuszpont***
  - **Z tengely: *optikai tengely***
  - ***Z* tengelyre merőleges *Z = f* sík a vetítési sík, *képsík***
  - ***f* a kamera *fókusztávolsága***
  - **Az optikai tengely képsíkon vett *o* döféspontja a *főpont*,**
  - **a kamera koordináta rendszer *XY* síkja pedig a *fősík*, amely párhuzamos a képsíkkal.**

Geometriai modell

Perspektív leképezés tulajdonságai

- egyenes
- méret
- párhuzam/szög
- alak
- Síkok alakja

- mélység

Photo by Robert Kosara, robert@kosara.net
http://www.kosara.net/gallery/pinholeamsterdam/pic01.html

Perspektív leképezés tulajdonságai

- ✓egyenes
- méret
- párhuzam/szög
- alak
- Síkok alakja
- mélység

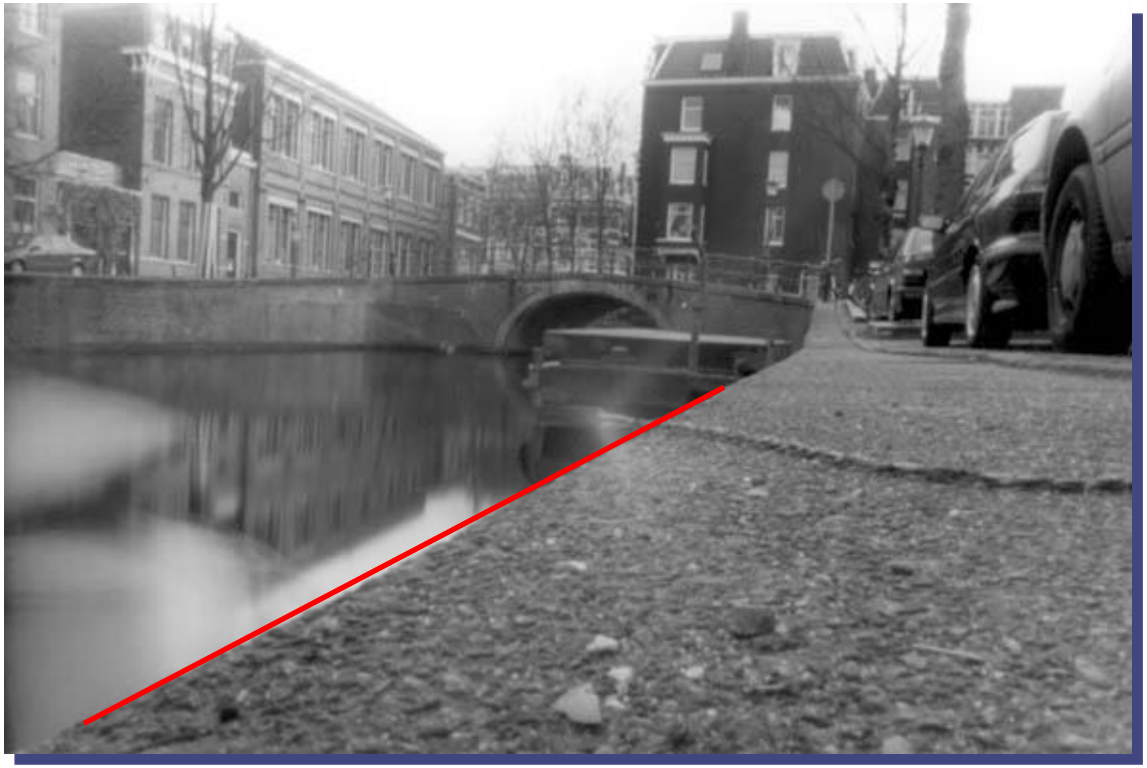


Photo by Robert Kosara, robert@kosara.net
http://www.kosara.net/gallery/pinholeamsterdam/pic01.html

Slide adopted from Zhigang Zhu Computer Vision - CSC I6716

Perspektív leképezés tulajdonságai

- ✓egyenes
- ×méret
- párhuzam/szög
- alak
- Síkok alakja
- mélység

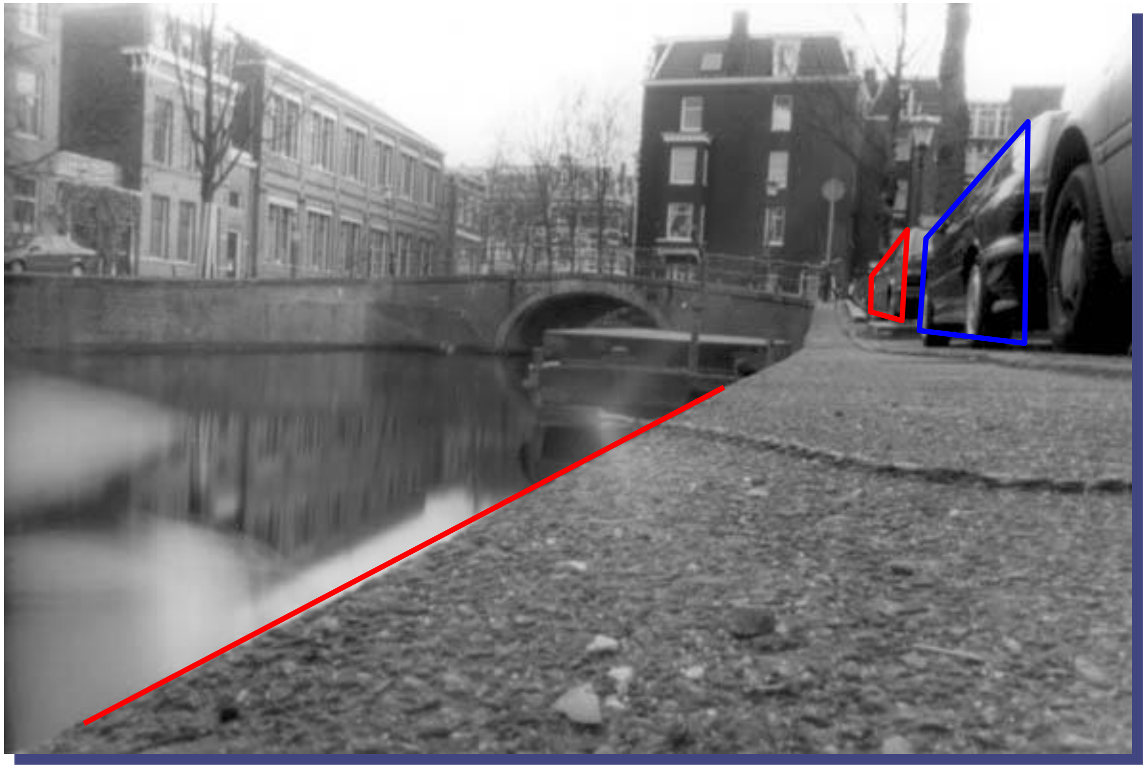


Photo by Robert Kosara, robert@kosara.net
http://www.kosara.net/gallery/pinholeamsterdam/pic01.html

Perspektív leképezés tulajdonságai

- ✓egyenes
- ×méret
- ×párhuzam/szög
- •alak
- •Síkok alakja
- •mélység

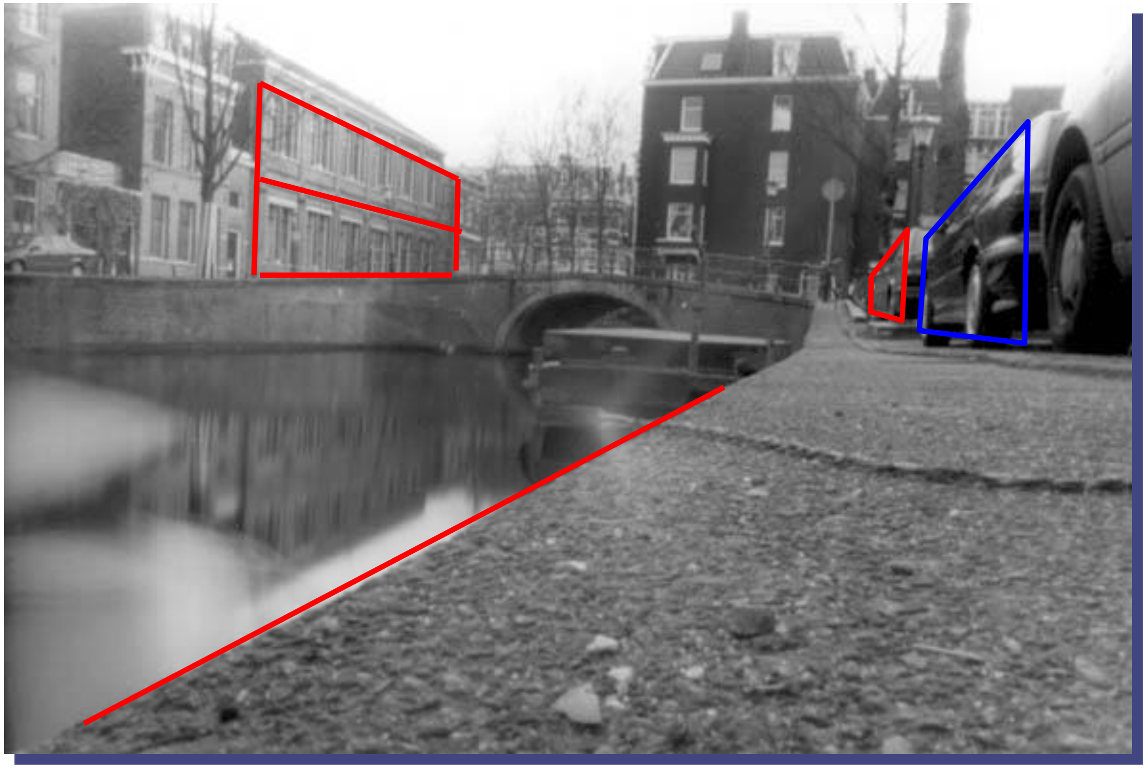


Photo by Robert Kosara, robert@kosara.net
http://www.kosara.net/gallery/pinholeamsterdam/pic01.html

Perspektív leképezés tulajdonságai

- ✓egyenes
- ×méret
- ×párhuzam/szög
- ×alak
- ●Síkok alakja
- ●mélység

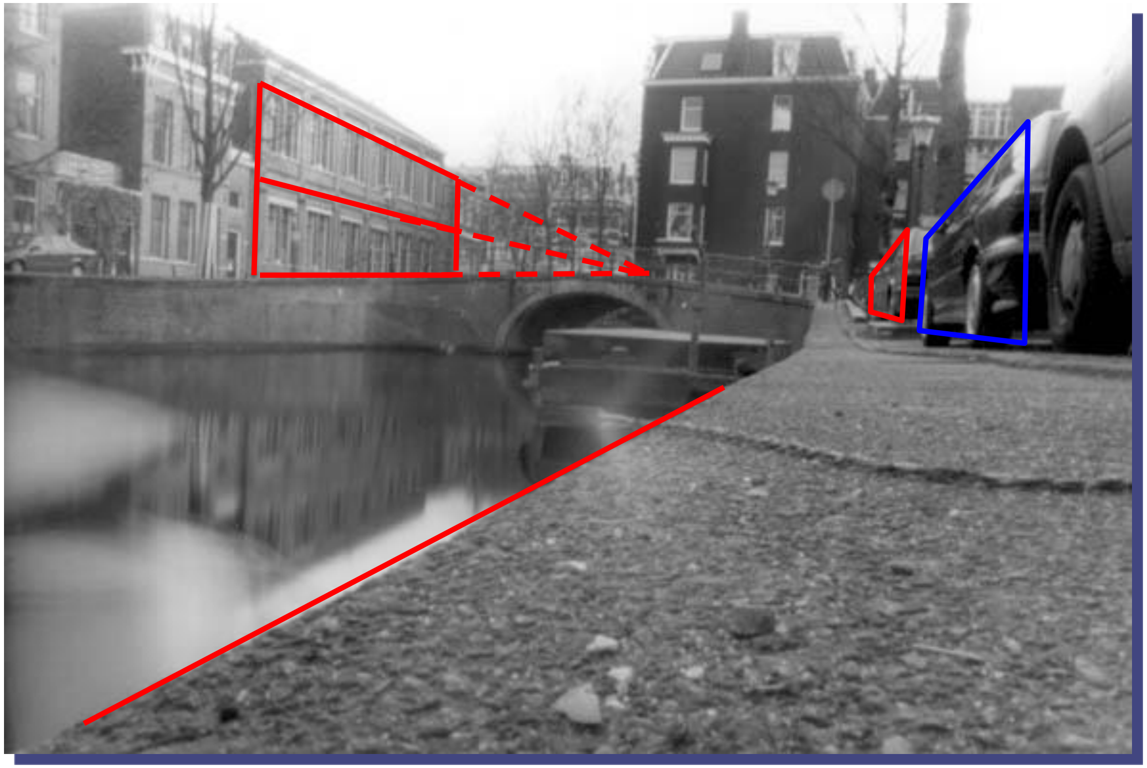


Photo by Robert Kosara, robert@kosara.net
http://www.kosara.net/gallery/pinholeamsterdam/pic01.html

Perspektív leképezés tulajdonságai

- ✓ egyenes
- × méret
- × párhuzam/szög
- × alak
- Síkok alakja
  - ✓ képsíkkal párhuzamos
- mélység
  - ✓ rekonstruálható (pl. sztereo)

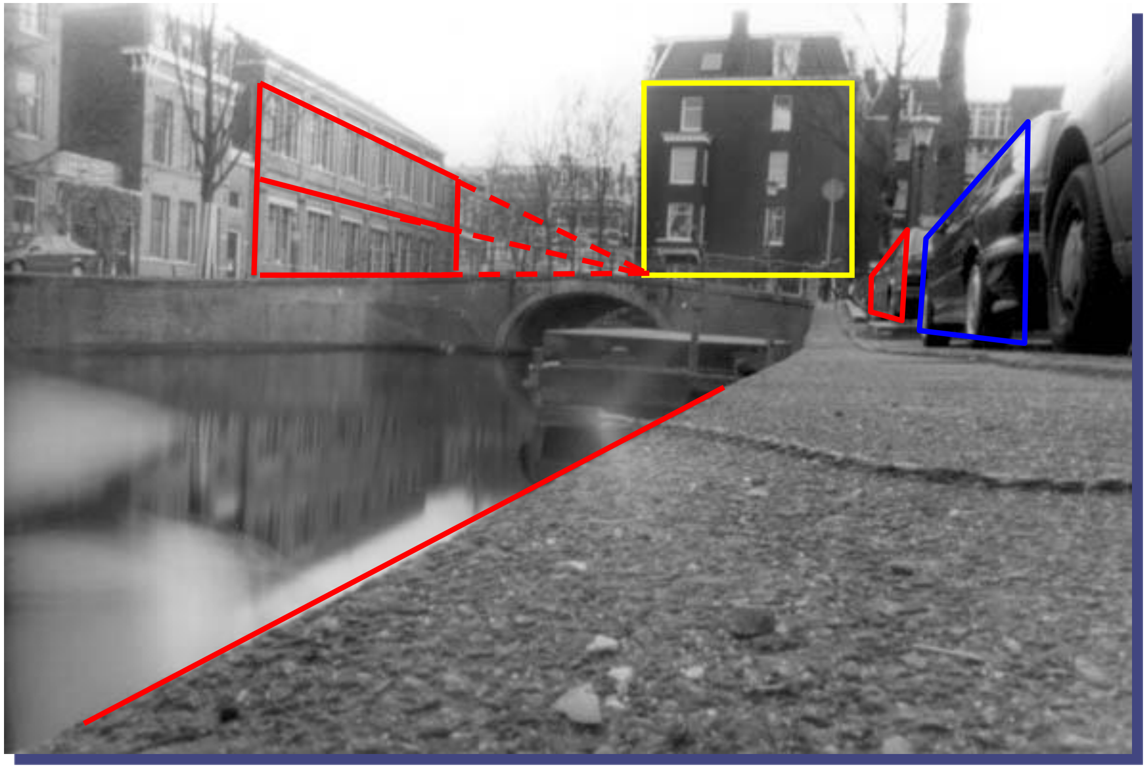
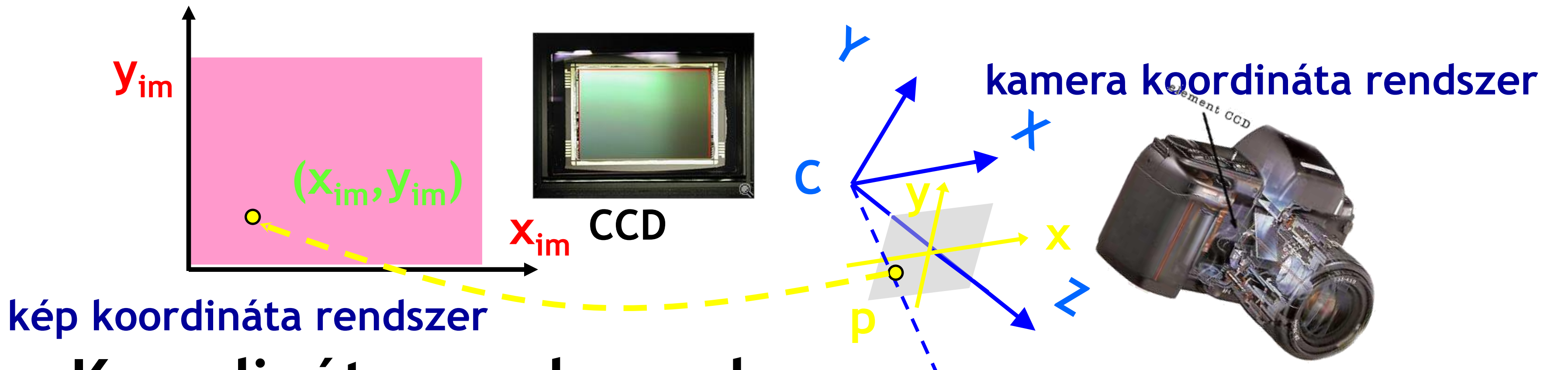
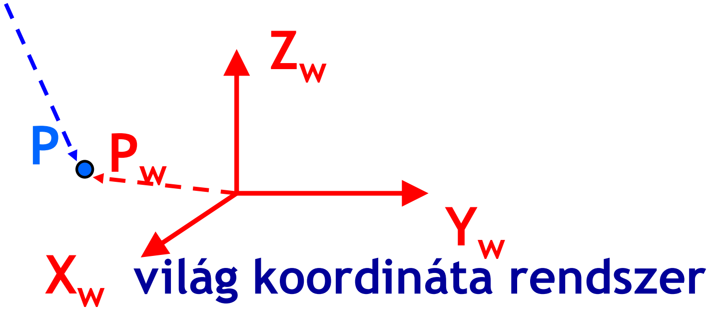


Photo by Robert Kosara, robert@kosara.net
http://www.kosara.net/gallery/pinholeamsterdam/pic01.html

Slide adopted from Zhigang Zhu Computer Vision - CSC I6716

Kamera paraméterek



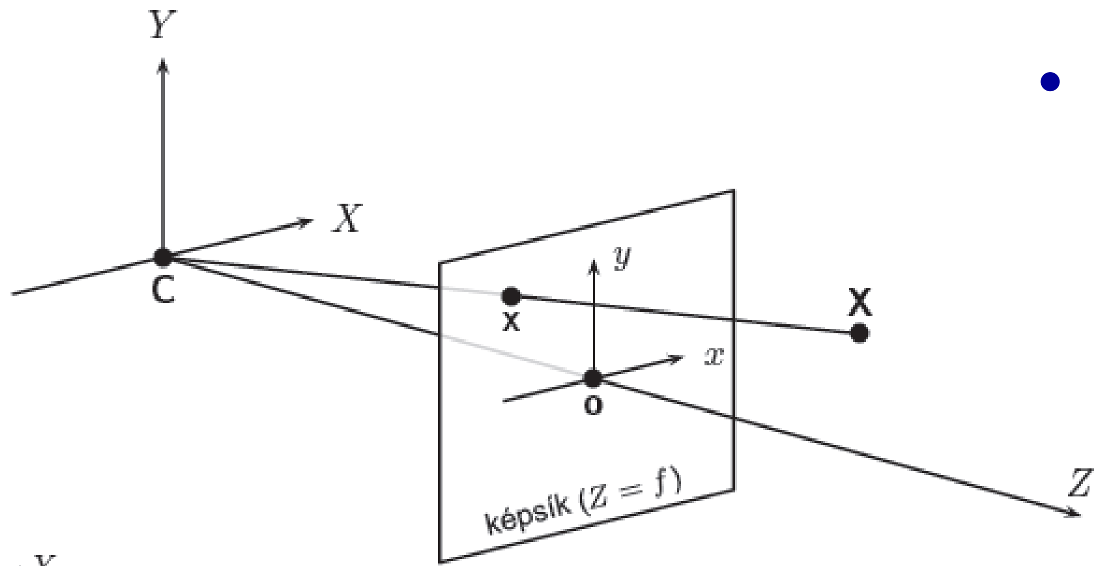


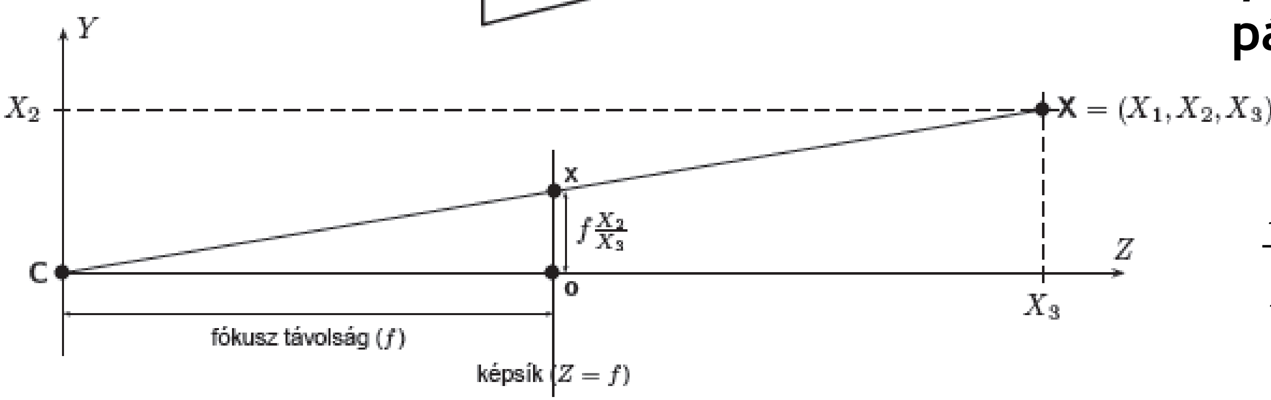
- Koordináta rendszerek
  - világ koordináták (X_w,Y_w,Z_w)
  - kamera koordináták (X,Y,Z)
  - kép koordináták (x_{im}, y_{im}) (pixel!)
    - (folytonos kép koordináták (x,y))
- Kamera paraméterek
  - Belső paraméterek : 3D→2D leképezés a kamera koordinátarendszerből a kép koordinátarendszerbe.
  - Külső paraméterek: a kamera koordinátarendszer pozíciója és orientációja a világ koordinátarendszerhez képest (3D eltolás és forgatás)

Slide adopted from Zhigang Zhu Computer Vision - CSC I6716

3D pont (X) és képe (x) közötti kapcsolat

$$\begin{pmatrix} X_1 \\ X_2 \\ X_3 \\ 1 \end{pmatrix} \mapsto \begin{pmatrix} fX_1 \\ fX_2 \\ X_3 \end{pmatrix} = \begin{bmatrix} f & & & 0 \\ & f & & 0 \\ & & 1 & 0 \end{bmatrix} \begin{pmatrix} X_1 \\ X_2 \\ X_3 \\ 1 \end{pmatrix}$$

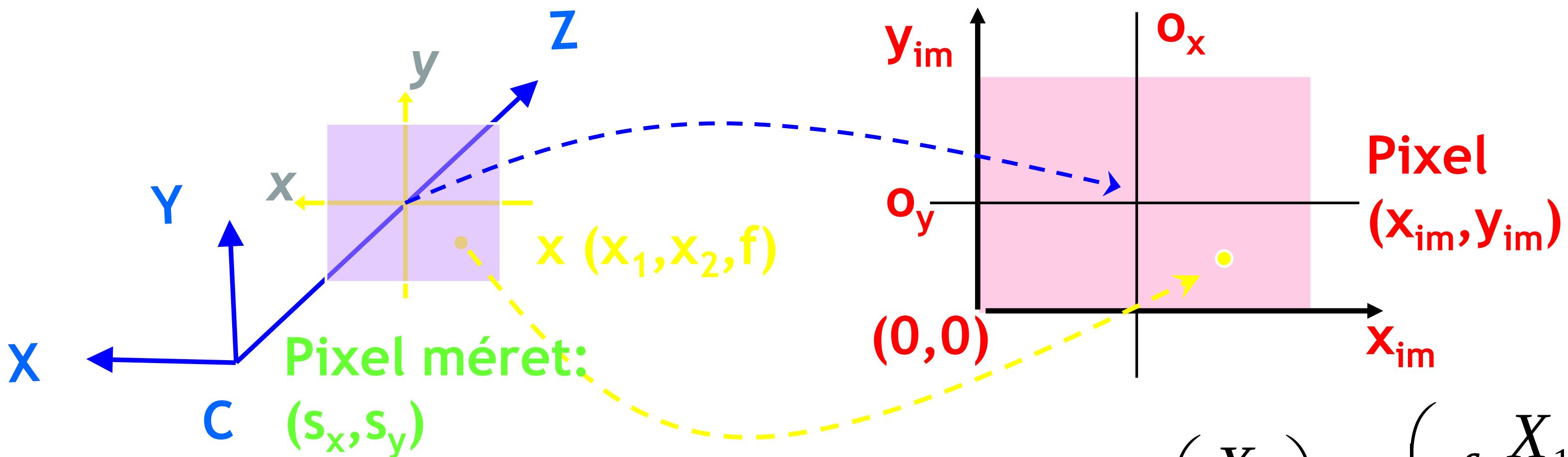




- **Nem lineáris leképezés az Euklideszi térben!**
- **Homogén koordináták → projektív tér, ahol lineáris leképezés lesz.**
- **A képpont koordinátáit egy olyan 2D euklideszi koordináta rendszerben kapjuk, melynek**
  - **origója az *o* főpont,**
  - ***x* és *y* tengelyei pedig a kamera koordináta rendszer megfelelő *X* és *Y* tengelyeivel párhuzamosak.**

$$\frac{X_2}{X_3} = \frac{x_2}{f} \Rightarrow x_2 = f\,\frac{X_2}{X_3}$$

x a kép koordinátarendszerben



$$\begin{pmatrix} X_1 \\ X_2 \\ X_3 \end{pmatrix} \mapsto \begin{pmatrix} f\dfrac{X_1}{X_3} \\ f\dfrac{X_2}{X_3} \end{pmatrix}$$

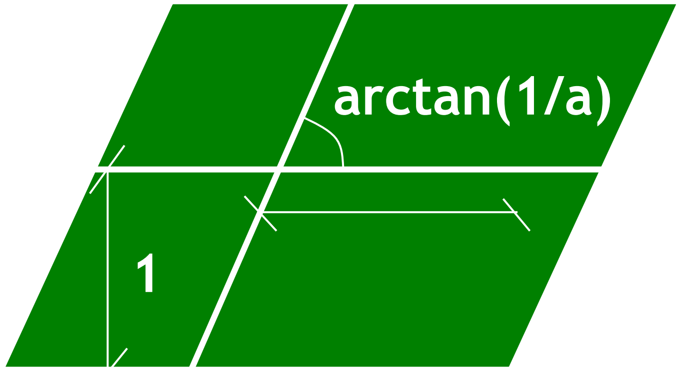
- **Belső paraméterek**
  - (o_x, o_y): főpont (kép középpontja)
  - (s_x, s_y): pixel méret
  - f: fókusz távolság

$$\begin{pmatrix} X_1 \\ X_2 \\ X_3 \\ 1 \end{pmatrix} \mapsto \begin{pmatrix} f_x X_1 + X_3 o_x \\ f_y X_2 + X_3 o_y \\ X_3 \end{pmatrix} = \begin{bmatrix} fs_x & & o_x & 0 \\ & fs_y & o_y & 0 \\ & & 1 & 0 \end{bmatrix} \begin{pmatrix} X_1 \\ X_2 \\ X_3 \\ 1 \end{pmatrix}$$

Slide adopted from Zhigang Zhu Computer Vision - CSC I6716

Kalibrációs mátrix

$$\mathbf{K} = \begin{bmatrix} f_x & a & o_x \\ & f_y & o_y \\ & & 1 \end{bmatrix}$$



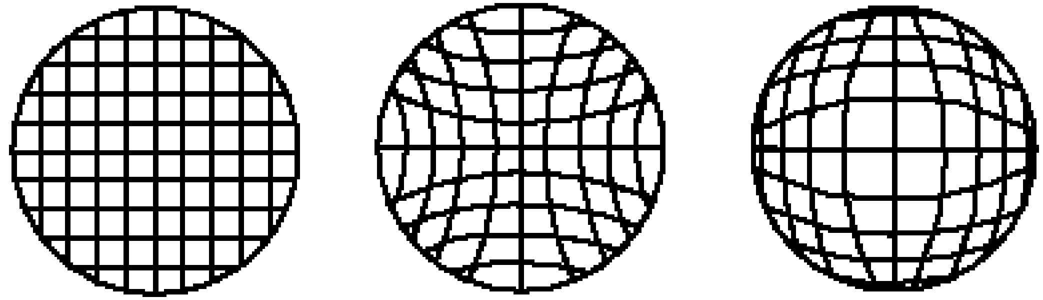
a ➔ nyírás
CCD/CMOS esetén a=0

$$\mathbf{P} = [\mathbf{K}|\mathbf{0}] = \mathbf{K}[\mathbf{I}|\mathbf{0}]$$

- P a 3×4 homogén kamera mátrix
- 3×3 felső trianguláris K mátrix a *kamera kalibrációs mátrix*
  - 5 belső paraméter: f_x, f_y, a, o_x, o_y
  - Általános pixelek: paralelogramma
  - Téglalap: a=0
  - Négyzetes pixelek: $f_x = f_y$, a=0

Kató Zoltán: Ipari Képfeldolgozás

Optikai torzítások

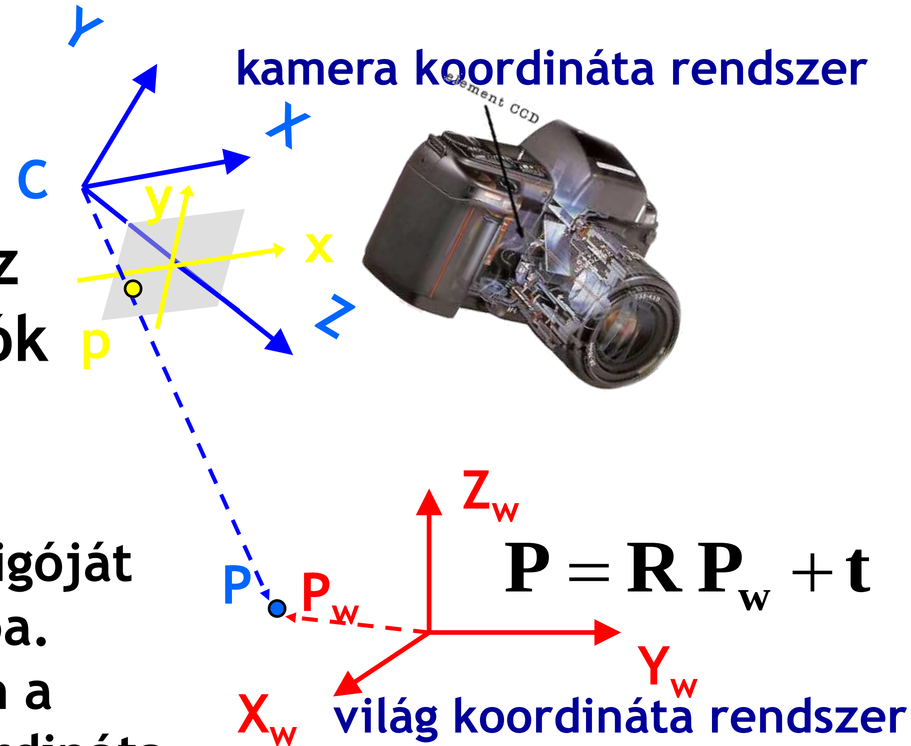


- **Minden optikának van geometriai torzítása**
  - A perspektív vetítés képletei torzításmentes esetben érvényesek
  - Pontos mérésekhez / 3D rekonstrukcióhoz szükséges ezen torzítások korrekciója
    - Jobb minőségű (és így drágább) optikák torzítását minimalizálják
    - Nagy látószögű optikák mindig erősebben torzítanak
- **Tipikusan radiális torzítás, melyet korrigálhatunk**
  - (x_d , y_d) torzított koordináták
  - $r^2 = x_d^2+y_d^2$
  - k_1 , k_2: torzítási együtthatók (paraméterek)

$$x = x_d\left(1+k_1 r^2 + k_2 r^4\right)$$
$$y = y_d\left(1+k_1 r^2 + k_2 r^4\right)$$

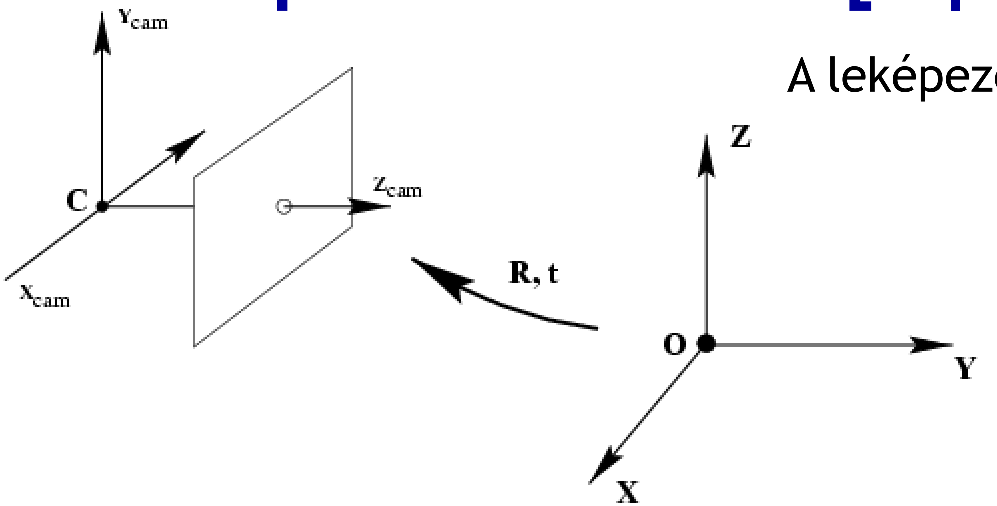
Külső paraméterek: [R|t]

- A két koordináta rendszer között egy 3D merevtest transzformáció hat, mely az alábbi elemi transzformációk kompozíciója:
  1. Egy t eltolás, amely a világ koordináta rendszer origóját átviszi a kamera középpontba. Az eltolás nagysága pontosan a kamera középpont világ koordináta rendszerbeli inhomogén koordinátája lesz: $-\tilde{\mathbf{C}}$
  2. Ezután következik egy R forgatás, amely a világ koordináta tengelyeket illeszti a kamera koordináta tengelyekre.



$$\mathbf{R}^{-1} = \mathbf{R}^T, vagyis\ \mathbf{R}\mathbf{R}^T = \mathbf{R}^T\mathbf{R} = \mathbf{I}$$

Külső paraméterek: [R|t]



A leképezés homogén koordinátákkal:

$$\mathbf{X}_{\mathrm{cam}} = \begin{bmatrix} \mathbf{R} & -\mathbf{R}\tilde{\mathbf{C}} \\ 0 & 1 \end{bmatrix} \begin{pmatrix} X_1 \\ X_2 \\ X_3 \\ 1 \end{pmatrix}$$

- **Az általános kamera mátrixot, amely a világ koordináta rendszerből képez le a kép koordináta rendszerbe, az alábbi alakokban írhatjuk:**

$$\mathbf{P} = \mathbf{K}\mathbf{R}[\mathbf{I}|\mathbf{t}] = \mathbf{K}[\mathbf{R}|\mathbf{R}\mathbf{t}] = \mathbf{K}[\mathbf{R}|-\mathbf{R}\tilde{\mathbf{C}}] = \mathbf{K}\mathbf{R}[\mathbf{I}|-\tilde{\mathbf{C}}]$$

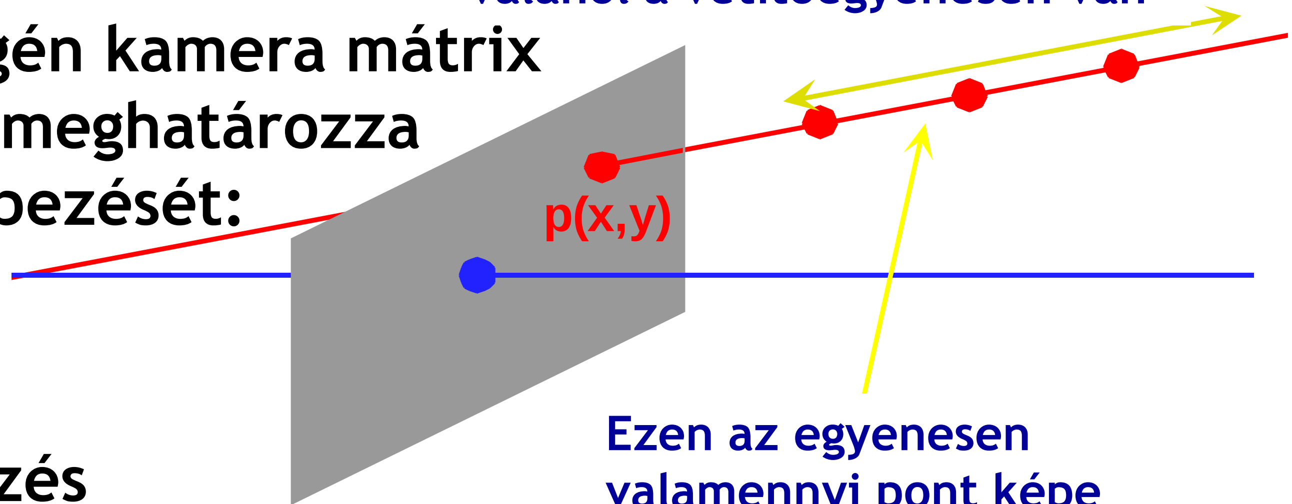
- **Az (R,t) transzformáció 3 forgatási és 3 eltolási paramétere a kamera külső paramétere**
  - **kizárólag a kamera világ koordináta rendszerben vett helyzetétől és irányától függenek**
- **A P általános kamera mátrix szabadségi foka 5+6=11**

Vetítés és visszavetítés

A visszavetített 3D pont
valahol a vetítőegyenesen van

- A P 3x4 homogén kamera mátrix egyértelműen meghatározza a kamera leképezését:

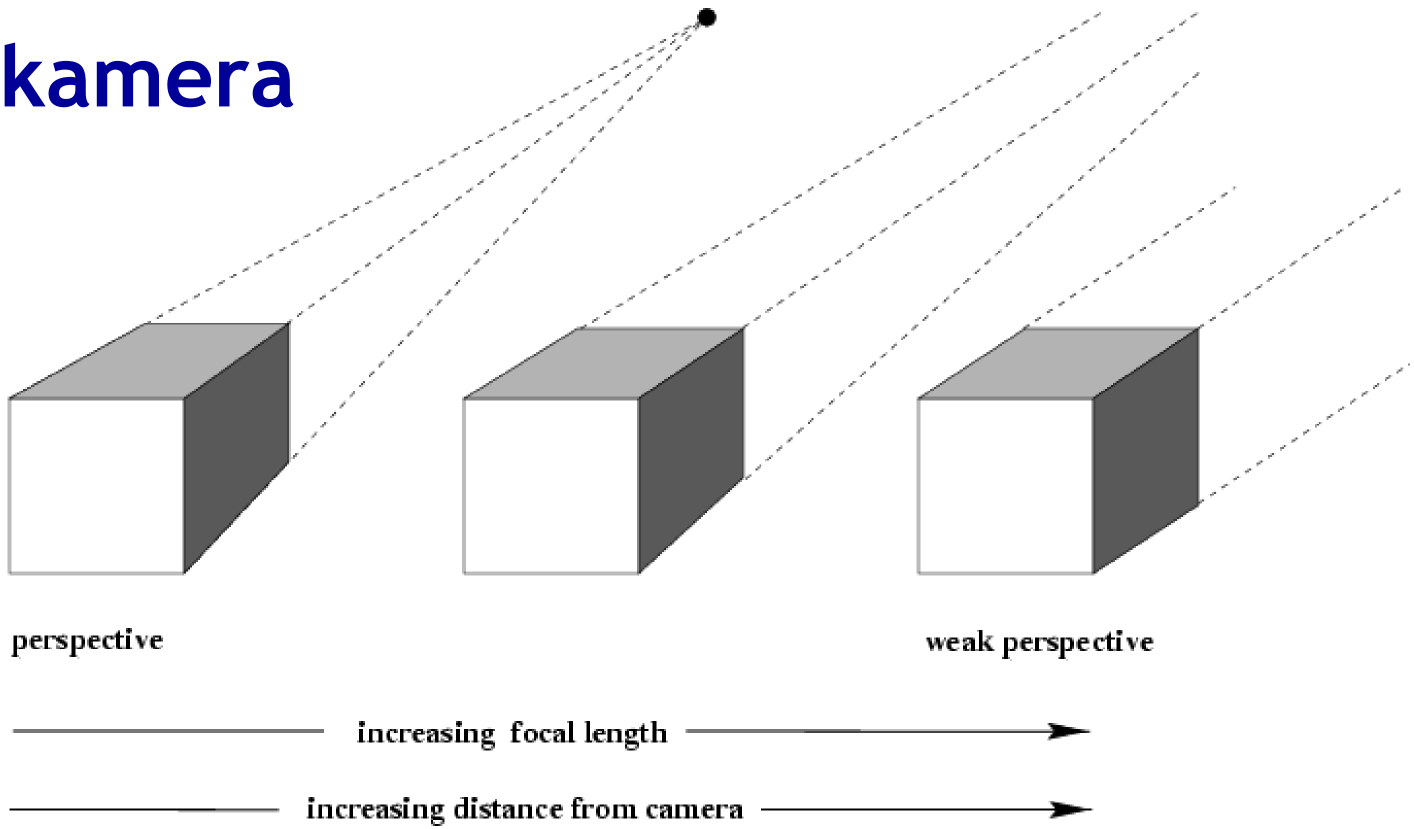
$$\mathbf{x} = \mathbf{P}\mathbf{X}$$

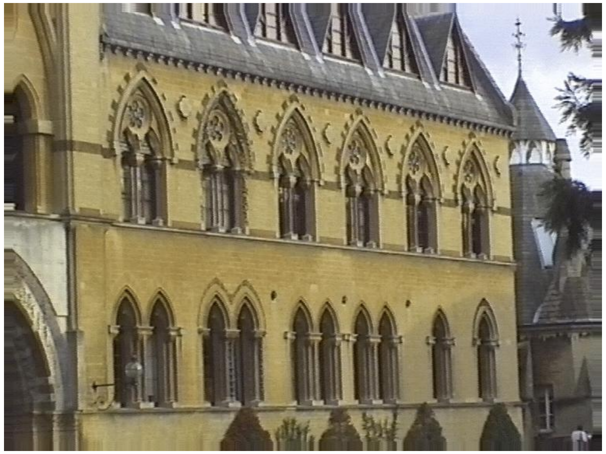


Ezen az egyenesen
valamennyi pont képe
ugyanaz a képpont lesz

- A fenti leképezés inverze, vagyis egy képpont visszavetítése azonban csak egy 3D egyenes erejéig határozható meg
  - ez az egyenes az adott képponthoz tartozó vetítősugár
  - A 3D pont meghatározásához több információ kell (pl. sztereó képpár)

Affin kamera





Affin kamera

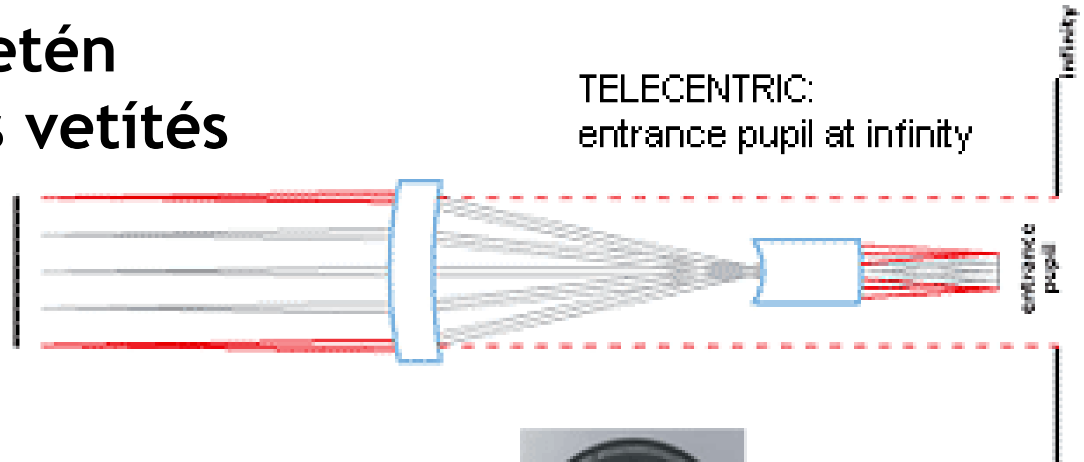
- **Affin kamera=olyan kamera, melynek fősíkja a végtelen távoli sík Π_∞**
- **Párhuzamos egyenesek képe párhuzamos**
- **Nincs vetítési középpont, nincs főpont csak vetítési irány $P_A d=0$ (pont a Π_∞ -n)**
  - **Párhuzamos vetítés**
- **Kamera mátrix utolsó sora (0,0,0,1) lesz**

$$\mathbf{P}_A = \begin{bmatrix} f_x & a & \\ & f_y & \\ & & 1 \end{bmatrix} \begin{bmatrix} \mathbf{r}^{1\mathrm{T}} & t_1 \\ \mathbf{r}^{2\mathrm{T}} & t_2 \\ \mathbf{0}^{\mathrm{T}} & 1 \end{bmatrix} = \begin{bmatrix} m_{11} & m_{12} & m_{13} & t_1 \\ m_{21} & m_{22} & m_{23} & t_2 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

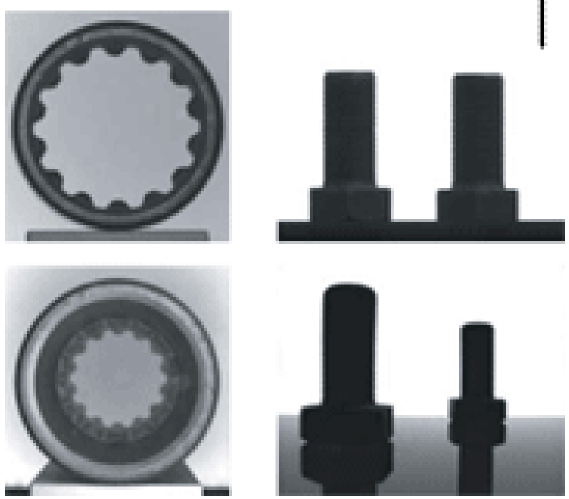
  - **M_{2x3} rangja 2**
  - **(t_1,t_2) a világ koordináta rendszer origójának képe**
  - **8 szabadsági fok**

Telecentrikus optika

- Telecentrikus optika affin kameraként működik egy (optika függő) tárgytávolság-intervallumban.
- Négyzetes pixelek esetén skálázott ortografikus vetítés
  - 6 szabadági fok
    - 3 forgatási szög
    - 2 eltolás (t_1,t_2)
    - skálafaktor k
  - $\mathbf{r}^{1T}$ és $\mathbf{r}^{2T}$ ortogonális és normájuk egyenlő
    - A 3D forgatómátrix első 2 sora



$$\mathbf{P}_A = \begin{bmatrix} k & & \\ & k & \\ & & 1 \end{bmatrix} \begin{bmatrix} \mathbf{r}^{1\mathrm{T}} & t_1 \\ \mathbf{r}^{2\mathrm{T}} & t_2 \\ \mathbf{0}^{\mathrm{T}} & 1 \end{bmatrix} = \begin{bmatrix} \mathbf{r}^{1\mathrm{T}} & t_1 \\ \mathbf{r}^{2\mathrm{T}} & t_2 \\ \mathbf{0}^{\mathrm{T}} & 1/k \end{bmatrix}$$



KAMERA KALIBRÁCIÓ

Kalibráció

- Kamera a mátrixával egyértelműen reprezentálható
- A kamera mátrix meghatározása a *kamera kalibráció*
  - A rendelkezésre álló információtól függően sokféle algoritmus
- Kalibrációs minta alapú eljárásokat tárgyaljuk
- Rendelkezésre állnak pontmegfeleltetések a valós 3D pontok (*kontrolpont*ok) és 2D képeik között:
  - Ismert 3D méretű kalibrációs mintát lefényképezünk
  - a könnyen azonosítható sarokpontokat keressük a képen
    - ezért általában sakktábla-szerű kalibrációs mintákat használunk
  - Ezekből kapjuk a 3D-2D pontpárokat

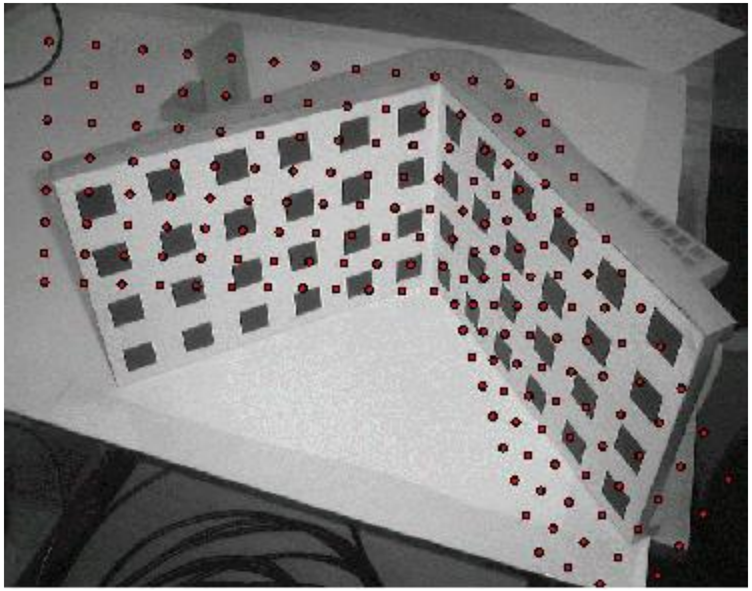
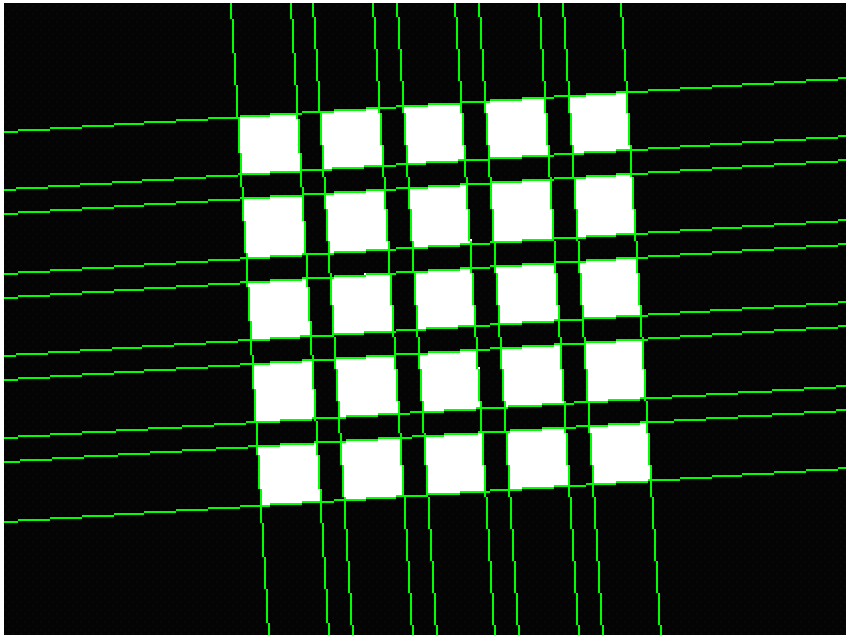


Image from Stanford CS223B Computer Vision course (Thrun, Szeliski, Dahlkamp)

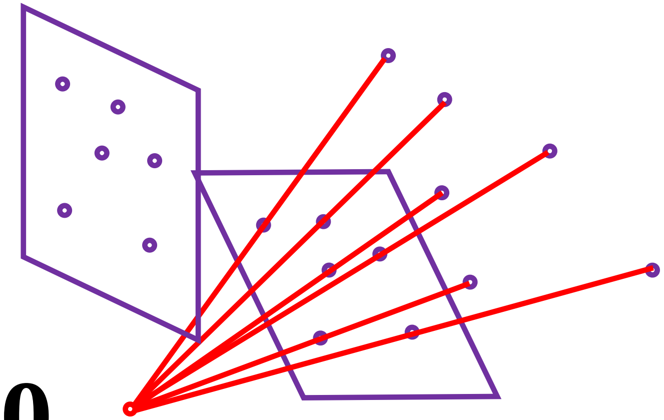
Kató Zoltán: Ipari Képfeldolgozás

Kalibrációs minta és pontosság



- Adott egy vagy több kép egy kalibrációs mintáról
- Határozzuk meg a kamera
  - Belső paramétereit
  - Külső paramétereit, avagy
  - Valamennyi paraméterét
- Milyen pontos lesz a kalibráció?
  - A kalibrációs minta tervezése
    - A kontrolpontok a képen detektálható (sarok)pontok, amelyekben ismerjük a 3D méretet.
    - Egyenletes eloszlásúak és nem esnek egy síkba
    - A minta konstrukciója egy vagy két nagyságrenddel pontosabb legyen, mint az elérni kívánt kalibrációs pontosság
    - pl. 0,1mm pontosság eléréséhez 0,01mm pontosságú minta kell
  - Hogyan nyerjük ki a kontrolpontoknak megfelelő képpontokat?
    - <u>Sarokdetektálás</u>: működik optikai torzítás esetén is, de bizonytalanabb
    - <u>Egyenes illesztés</u>: torzításmentes képet igényel, de szubpixeles pontosságú pontokat ad

Algebrai alapmegoldás



- Adott N pontpár: $\mathbf{x_i} \leftrightarrow \mathbf{X_i}$
- Valamennyi pontpárra teljesül:

$$\mathbf{x_i} = \mathbf{P}\mathbf{X_i} \quad \Rightarrow \quad \mathbf{x_i} \times \mathbf{P}\mathbf{X_i} = \mathbf{0}$$

- Kifejtve és az ismeretlen P-re rendezve:

$$\begin{bmatrix} \mathbf{0}^\top & -w_i\mathbf{X}_i{}^\top & y_i\mathbf{X}_i{}^\top \\ w_i\mathbf{X}_i{}^\top & \mathbf{0}^\top & -x_i\mathbf{X}_i{}^\top \\ -y_i\mathbf{X}_i{}^\top & x_i\mathbf{X}_i{}^\top & \mathbf{0}^\top \end{bmatrix} \begin{pmatrix} \mathbf{P}^1 \\ \mathbf{P}^2 \\ \mathbf{P}^3 \end{pmatrix} = 0$$

- Mivel csak 2 sor lesz független (homogén koord. miatt):

$$\begin{bmatrix} \mathbf{0}^\top & -w_i\mathbf{X}_i{}^\top & y_i\mathbf{X}_i{}^\top \\ w_i\mathbf{X}_i{}^\top & \mathbf{0}^\top & -x_i\mathbf{X}_i{}^\top \end{bmatrix} \begin{pmatrix} \mathbf{P}^1 \\ \mathbf{P}^2 \\ \mathbf{P}^3 \end{pmatrix} = 0$$

- N pontpárra tehát 2N egyenletet kapunk
  - A egy 2Nx12 mátrix.
  - P a lineáris egyenletrendszer megoldása lesz

$$\mathbf{Ap} = \mathbf{0}$$

Kató Zoltán: Ipari Képfeldolgozás

Algebrai alapmegoldás

$$\mathbf{Ap} = \mathbf{0}$$

- P 12 elemű, de csak 11 szabadsági foka van ➔ 11 lineárisan független egyenletre van szükség
  - 5½ pontpár szükséges
  - A gyakorlatban azonban a pontmegfeleltetések hibával terheltek ➔ az egyenleteink csak közelítően teljesülnek és így nem lesz egzakt megoldás.
- Túlhatározott egyenletrendszer: N≥6
  - az egyenletrendszer ||Ap|| algebrai hibáját minimalizáljuk
  - azzal a további feltétellel, hogy||p=1||
    - Kizárja a p=0 triviális megoldást
    - Rögzíti P skáláját (homogén mátrix csak egy skálafaktor erejéig meghatározott)
  - a megoldást A SVD felbontásával kapjuk

Redukált mérési mátrix

- Ha sok a pontmegfeleltetés, akkor A-nak nagyon sok sora lesz (2N)
- Valójában elegendő egy 12X12 mátrixal dolgozni (redukált mérési mátrix), amit A SVD felbontásával kapunk:

$$\|\mathbf{Ap}\| = \mathbf{p}^{\mathrm{T}}\mathbf{A}^{\mathrm{T}}\mathbf{Ap} = \|\tilde{\mathbf{A}}\mathbf{p}\|$$

$$\mathbf{A}^{\mathrm{T}}\mathbf{A} = (\mathbf{VDU}^{T})(\mathbf{UDV}^{T}) = (\mathbf{VD})(\mathbf{DV}^{T}) = \tilde{\mathbf{A}}^{T}\tilde{\mathbf{A}}$$

- Az egyenletrendszer megoldása szempontjából tehát az eredeti A és az alabbi redukált mérési mátrix egyenértékű, hiszen ugyanaz lesz az algebrai hiba

$$\mathbf{A}_{2N\times 12} = (\mathbf{UDV}^{T}) \Rightarrow \tilde{\mathbf{A}}_{12\times 12} = (\mathbf{DV}^{T})$$

Kató Zoltán: Ipari Képfeldolgozás

Normalizálás

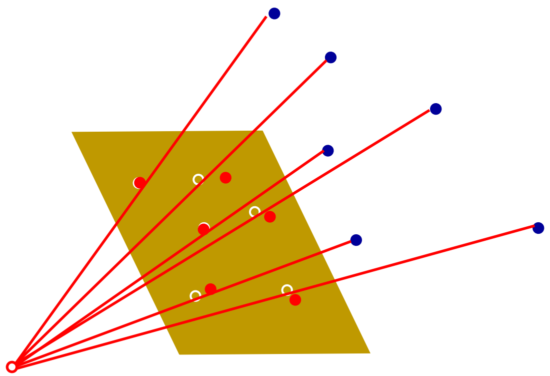
- **Az egyenletek numerikus stabilitásához elengedhetetlenül szükséges a pontkoordináták normalizálása**
  - egyenletrendszer együtthatói nagyságrendekkel eltérhetnek egymástól
- **összetett transzformációval érhetjük el:**
  1. Számoljuk ki a pontjaink középpontját, majd az így kapott vektorral toljunk el valamennyi pontot. Ezzel az origó a centroidba került.
  2. Ezután alkalmazzunk egy olyan skálázást, mellyel a pontok origótól vett átlagos távolsága a képpontok esetén $\sqrt{2}$, míg a 3D kalibrációs minta pontjai esetén $\sqrt{3}$ lesz.
- **A kapott P' mátrixot denormalizálva kapjuk a P megoldást:**

$$\mathbf{P} = \mathbf{N}_{2D}^{-1}\mathbf{P}'\mathbf{N}_{3D}$$

Geometriai hiba

- A megoldás során az algebrai hibát minimalizáltuk ➔ a kapott megoldásnak nem feltétlenül lesz legkisebb a geometriai hibája
  - A kalibrációs minta 3D pontkoordinátáit hibátlannak tekintve, a képen mért geometriai hiba:

$$\sum_i \left\| \mathbf{x}_i - \mathbf{P}\mathbf{X}_i \right\|^2$$



  - Vagyis a kinyert és a P által vetített pontok négyzetes távolságainak összege
  - a fenti hibát minimalizáló P mátrix lesz a kamera mátrix *legnagyobb valószínűség* (angolul *Maximum Likelihood*) értelemben vett becslése.

További kényszerfeltételek

- A kalibrációs mátrix bizonyos elemei ismertek lehetnek:
  - Nincs nyírás (a=0)
  - Négyzetes pixelek ($f_x = f_y$, a=0)
  - Főpont ismert
  - Teljes kalibrációs mátrix ismert

$$\mathbf{K} = \begin{bmatrix} f_x & a & o_x \\ & f_y & o_y \\ & & 1 \end{bmatrix}$$

- Algebrai megoldás:
  - A megoldásként kapott P mátrixot bontsuk fel összetevőire:

$$\mathbf{P} = \mathbf{KR}\left[\mathbf{I} \middle| -\tilde{\mathbf{C}}\right] = \left[\mathbf{M} \middle| \mathbf{p}_4\right] \Rightarrow \mathbf{KR} = RQ(\mathbf{M}) \;\; \acute{e}s \;\; \tilde{\mathbf{C}} = -\mathbf{M}^{-1}\mathbf{p}_4$$

  - majd az összetevőkben a kívánt értékeket beállítva megkapjuk a kívánt tulajdonságú kamera mátrixot.
  - jól működik, ha a kiindulási mátrixunkban már jó közelítéssel teljesülnek a kívánt feltételek
  - Sajnos a gyakorlatban ez ritkán teljesül!

További kényszerfeltételek

- A gyakorlatban célszerűbb az egyes paraméterekre vonatkozó kényszereket a geometriai hibához hozzáadott költségként megfogalmazni
  - Például négyzetes pixelekre vonatkozó feltételt ($f_x = f_y$, a=0):
$$\sum_i \left\| \mathbf{x}_i - \mathbf{P}\mathbf{X}_i \right\|^2 + wa^2 + w(f_x - f_y)^2$$
  - Ahol w súly értéke kezdetben kicsi, majd minden iteráció után növeljük
  - Ezzel elérjük, hogy a feltételeket fokozatosan érvényesítjük és az így kapott kamera mátrix elemei már numerikusan jól közelítik azt.
  - Ezután a megfelelő elemeket az előzőleg bemutatott faktorizációs módszerrel már biztonságosan beállíthatjuk a kívánt értékekre, ezzel pontosan előállítva a kívánt tulajdonságú kamera mátrixot.

Gold Standard algoritmus

Cél

Adott N≥6 3D -2D pontmegfeleltetés $\{\mathbf{X}_i \leftrightarrow \mathbf{x}_i'\}$, határozzuk meg P *maximális valószínűség* szerinti becslését.

Algoritmus

(i) **Lineáris (algebrai) megoldás**

(a) Pont koordináták normalizáláss: $\tilde{\mathbf{X}}_i = \mathbf{U}\mathbf{X}_i \quad \tilde{\mathbf{x}}_i = \mathbf{T}\mathbf{x}_i$

(b) **Ap=0 lineáris egyenletrendszer megoldása SVD felbontással**

(ii) **Geometriai hiba minimalizálása:** az előző lépésben kapott megoldással inicializálva minimalizáljuk (pl. Levenberg-Marquardt algoritmussal)

- A geometriai hibát: $\sum_i \left\| \tilde{\mathbf{x}}_i - \tilde{\mathbf{P}}\tilde{\mathbf{X}}_i \right\|^2$
- vagy a geometriai hibát együtt a súlyozott kényszerekkel, pl.:

$$\sum_i \left\| \tilde{\mathbf{x}}_i - \tilde{\mathbf{P}}\tilde{\mathbf{X}}_i \right\|^2 + wa^2 + w(f_x - f_y)^2$$

(iii) **Denormalizálás:** $\mathbf{P} = \mathbf{T}^{-1}\tilde{\mathbf{P}}\mathbf{U}$

Kató Zoltán: Ipari Képfeldolgozás

Kalibrált kamera pozíciója

- Mozgó kamerák (pl. mobil robot) esetén leginkább előforduló eset, amikor egy
  - ismert K kalibrációs mátrixú kamera
  - pozícióját és
  - orientációját

  akarjuk meghatározni a világ koordináta rendszerben
- 6 szabadsági fok
  - 3 forgatási szög
  - 3 eltolási paraméter
  - Minimum 3 pontmegfeleltetés szükséges