Metody Počítačového Vidění (MPV) - 3D počítačové vidění

Projektivní geometrie dvou pohledů

Ing. Zdeněk Krňoul, Ph.D.

Katedra Kybernetiky
Fakulta aplikovaných věd
Západočeská univerzita v Plzni

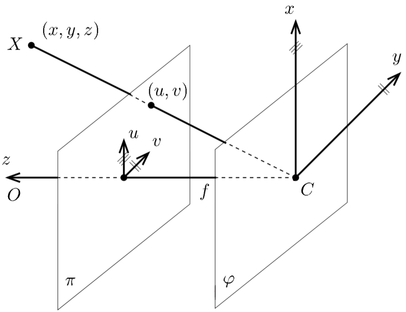
DEPARTMENT OF CYBERNETICS

Obsah přednášky

- **Perspektivní kamera**
  - model kamery
  - kalibrace kamery
  - rozklad matice projekce
- **Projektivní geometrie dvou pohledů**
  - Epipolární geometrie
  - Epipolární podmínka
  - Fundamentální matice
  - Odhad pohybu kamery
  - 3D rekonstrukce

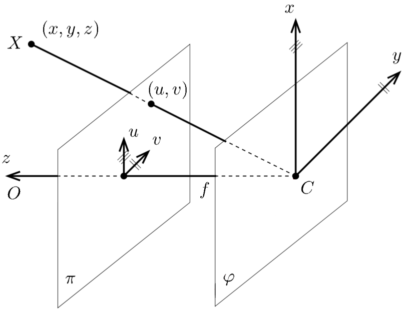
Perspektivní kamera

- obecný **model perspektivní kamery** slouží k popisu projekce $3D$ prostoru do $2D$ prostoru (obrazová rovina)
- vždy se jedná o **středovou projekci**



Pozn. speciální případ ... střed projekce leží v nekonečnu → **afinní kamera** *a jde o zobecnění tzv. paralelní projekce*

DEPARTMENT OF CYBERNETICS



- projekce bodu $X = (x, y, z)$ do obrazové roviny π souřadné osy (u, v)
- počátek souřadného systému světových souřadnic je v bode C
- vzdálenost obrazové roviny od tohoto bodu je tzv. ohnisková vzdálenost f

- model kamery je reprezentován maticí $\mathbf{P}$ o velikosti 3×4 tzv. **maticí projekce**
- libovolný bod v prostoru se transformuje do obrazové roviny pouhým násobením maticí projekce:

$$\mathbf{m} = \mathbf{P}\mathbf{X} \tag{1}$$

- a maticově zapíšeme jako:

$$\begin{bmatrix} m_1 \\ m_2 \\ m_3 \end{bmatrix} = \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \tag{2}$$

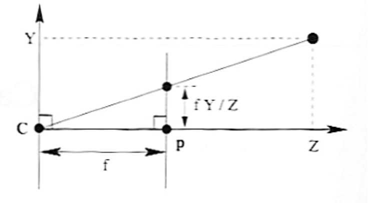
DEPARTMENT OF CYBERNETICS

Přechod z homogenní reprezentace:

- **normalizujeme** (na jedničku) třetí homogenní souřadnici $[\frac{m_1}{m_3}, \frac{m_2}{m_3}, 1]^T$ ... pak vlastně modeluje výpočet projekce takto:

$$\frac{m_1}{m_3} = \frac{fx}{z} = u, \ \frac{m_2}{m_3} = \frac{fy}{z} = v, \ \frac{m_3}{m_3} = 1 \ pro \ \ m_3 \neq 0 \qquad (3)$$

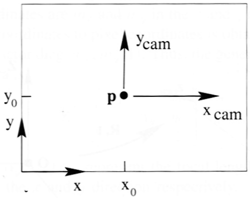
- říkáme, že vyjádříme bod v obraze
- Proč to tak je? ukázka přepočtu projekce za pomoci podobnosti trojúhelníků, náhled os (Y,Z) a obdobně platí pro (X,Z)

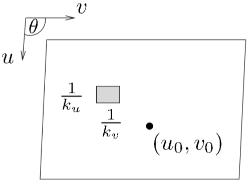


DEPARTMENT OF CYBERNETICS

Model obecné kamery (matice **P**) zahrnuje další prvky (celkem 5 vnitřních a 6 vnějších parametrů), které poskytují další stupně volnosti projekce:

- **posun počátku souřadnic** obrazové roviny do levého horního rohu
- **kompenzace nepravoúhlosti** os senzoru.



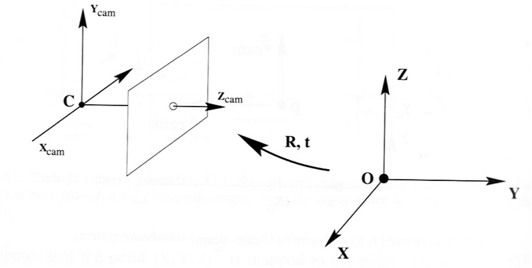


$$\mathbf{P} = \begin{bmatrix} fk_u & -fk_u \coth(\theta) & u_0 & 0 \\ 0 & fk_v / \sin(\theta) & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \text{ resp.}$$

$$\mathbf{K} = \begin{bmatrix} fk_u & -fk_u \coth(\theta) & u_0 \\ 0 & fk_v / \sin(\theta) & v_0 \\ 0 & 0 & 1 \end{bmatrix}$$ je **kalibrační matice** kamery

DEPARTMENT OF CYBERNETICS

- $+$ 6 vnějších parametrů (3 krát rotace kolem třech základních os a posun kamery, resp. souřadnice středu promítání)
- posun a otočení počátku souřadnic světových souřadnic



- $\mathbf{X} = \mathbf{R}(\mathbf{X}_{\mathbf{camera}} - \mathbf{C})$, kde $\mathbf{R}^\mathbf{T}\mathbf{R} = \mathbf{I}$ je matice rotace kamery oproti světovým souřadnicím

DEPARTMENT OF CYBERNETICS

- vše dohromady definuje obecný předpis pro perspektivní kameru,

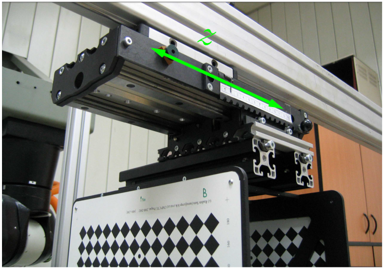
$$\mathbf{P} = \mathbf{KR}[\mathbf{I}, -\mathbf{C}] \tag{4}$$

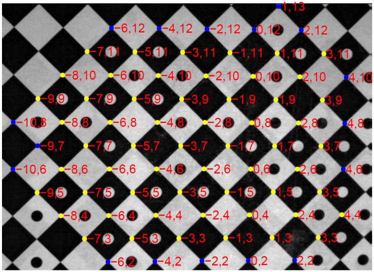
a projekce bodu je: $\mathbf{m} = \mathbf{KR}[\mathbf{I}, -\mathbf{C}]\mathbf{X}$
přepisem můžeme vyjádřit jako:

$$\mathbf{P} = \mathbf{K}[\mathbf{R}, \mathbf{t}] \; kde \; \mathbf{t} = -\mathbf{RC} \tag{5}$$

- perspektivní kamera má **11 stupňů volnosti**: 1x ohnisková vzdálenost v pixelech + 1x poměr stran pixelu + 1x zkosení os + 2x počátek obrázku + 3x posun + 3x rotace kamery= 11 DOF (degree of freedom)

Kalibrace kamery z množiny známých bodů - Camera resection





- kalibrace kamery je numerická metoda pro určení matice projekce $\mathbf{P}$
- vyžaduje pozici prostorového bodu a jeho projekci do obrazové roviny
- z několika těchto dvojic můžeme určit matici projekce
- pro každý pár $\mathbf{X}_i \leftrightarrow \mathbf{x_i}$ musí být splněna projekce $\mathbf{x}_i = \mathbf{P}\mathbf{X}_i$ pro $i = 1 : N$

Poznámka 1: Předpokladem je linearita projekce tak jak je zmíněna a neuvažuje se distorze obrazu daná například čočkou objektivu
Poznámka 2: Postup hledání projekční matice je velmi podobný hledání matice pro projektivní transformaci (rozdíl je pouze v rozměru matic) $\mathbf{P}$ *a* $\mathbf{H}$

DEPARTMENT OF CYBERNETICS

- pro každou dvojici $\mathbf{X}_i \leftrightarrow \mathbf{x_i}$ můžeme napsat vztah:

$$\begin{bmatrix} \mathbf{0}^T & -w_i\mathbf{X}_i^T & y_i\mathbf{X}_i^T \\ w_iX_i^T & \mathbf{0}^T & -x_i\mathbf{X}_i^T \\ -y_i\mathbf{X}^T & x_i\mathbf{X}_i^T & \mathbf{0}^T \end{bmatrix} \begin{bmatrix} \mathbf{p}_1 \\ \mathbf{p}_2 \\ \mathbf{p}_3 \end{bmatrix} = 0 \qquad (6)$$

- $\mathbf{x}_i = (x_i, y_i, w_i)$ a $\mathbf{p_1^T}$ je první řádek matice $\mathbf{P}$, podobně druhý a třetí řádek jsou složeny do sloupcového vektoru neznámých veličin o rozměru 1×12
- podobně můžeme uvažovat pouze první dva řádky soustavy rovnic ... třetí řádek je lineárně závislý na prvních dvou. Tedy:

$$\begin{bmatrix} \mathbf{0}^T & -w_i\mathbf{X}_i^T & y_i\mathbf{X}_i^T \\ w_iX_i^T & \mathbf{0}^T & -x_i\mathbf{X}_i^T \end{bmatrix} \begin{bmatrix} \mathbf{p_1} \\ \mathbf{p_2} \\ \mathbf{p_3} \end{bmatrix} = 0 \qquad (7)$$

DEPARTMENT OF CYBERNETICS

- pro množinu n známých prostorových bodů a jejich projekcí získáváme matici $\mathbf{A}$ velikosti $(2n) \times 12$

$$\mathbf{A} = \begin{bmatrix} \mathbf{0}^T & -\mathbf{X}_1^T & y_1\mathbf{X}_1^T \\ X_1^T & \mathbf{0}^T & -x_1\mathbf{X}_1^T \\ \vdots & \vdots & \vdots \\ \mathbf{0}^T & -\mathbf{X}_n^T & y_n\mathbf{X}_n^T \\ X_n^T & \mathbf{0}^T & -x_n\mathbf{X}_n^T \end{bmatrix} \tag{8}$$

- řešením této soustavy ($\mathbf{Ap} = \mathbf{0}$) získáme vektor $\mathbf{p}$ a tedy potřebné řádky matice projekce $\mathbf{P}$.
- matice $\mathbf{P}$ má 12 prvků a 11 stupňů volnosti (není modelováno měřítko, k-násobek matice je stejná projekce)
- z každého prostorového bodu získáváme dvě rovnice $\rightarrow$ teoreticky nám stačí pro DOF 11 přesně 5,5 prostorových bodů
- pak existuje jedno řešení $\rightarrow$ **pravý nulový prostor** matice $\mathbf{A}$

- však nepřesnosti měření těchto bodů $\rightarrow$ získáváme však pře-určenou soustavu rovnic ... $rank(\mathbf{A}) = 12$
- $\rightarrow$ hledáme řešení s nejmenší chybou (algebraickou nebo geometrickou)
- v základu proto použijeme SVD rozklad, nalezneme řešení s nejmenší algebraickou chybou ... $\mathbf{Ap} = \epsilon \rightarrow$ SVD minimalizuje $||\mathbf{Ap}||$ s podmínkou $||\mathbf{p}|| = 1$
- tedy $||\epsilon|| \rightarrow min$ a pokud $\mathbf{A} = \mathbf{UDV}^T$ a $\sigma_{12} \ll \sigma_{11}$
- pak řešení $\mathbf{p}$ odpovídá poslednímu sloupci matice $\mathbf{V}$ a $||\epsilon|| = \sigma_{12}$

DEPARTMENT OF CYBERNETICS

Existují jisté degenerativní konfigurace prostorových bodů, pro které nelze určit řešení a tedy matici projekce. Nejzávažnější jsou tyto:

- střed projekce kamery a prostorové body leží na "twisted cubic"
- kalibrační prostorové body leží v jedné rovině a na přímce, která prochází středem projekce aj.

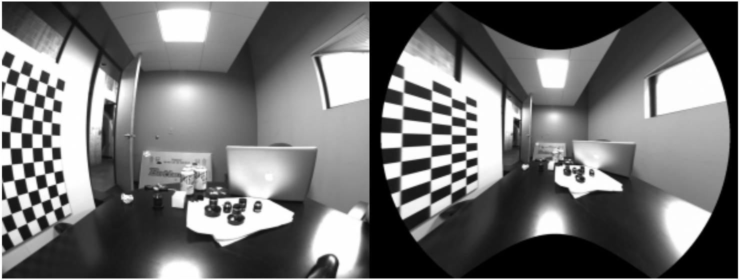
Radiální zkreslení - Radial distorsion

- Všechny předchozí vztahy platí pro případ, že projekce je ideální středové promítání
- skutečný přístroj (fotoaparát nebo kamera) obsahuje čočku, která způsobuje více či méně jev, že prostorové přímky nejsou promítány na přímky v obraze - tzv. **radiální zkreslení**
- tento jev narůstá důležitosti s klesající ohniskovou vzdáleností a cenou objektivu

DEPARTMENT OF CYBERNETICS

- **korekci zkreslení** souřadnic obrázku můžeme přespat jako:
  $\hat{x} = x_c + L(r)(x - x_c)$
  $\hat{y} = y_c + L(r)(y - y_c)$
  - (x, y) je bod v obraze který je podřízen radiálnímu zkreslení
  - (x_c, y_c) střed radiálního zkreslení
  - r je radiální vzdálenost od středu radiálního zkreslení $\sqrt{x^2 + y^2}$
  - $L(r)$ je funkce popisující zkreslení, parametrem je vzdálenost od středu
- aproximaci funkce $L(r)$ můžeme zvolit:
  $L(r) = 1 + \kappa_1 r + \kappa_2 r^2 + \kappa_3 r^3 + \dots$
- parametry popisující radiální zkreslení jsou pak
  $(\kappa_1, \kappa_2, \kappa_3, \dots), x_c, y_c$

Pozn. střed radiálního zkreslení může být zvolen "principal point"... projekce středu projekce do obrazu



- určení funkce $L(r)$ je často provedenou současně s výpočtem projekční matice (je zahrnuta do minimalizačního procesu)
- chybová veličina pak určuje odchylku skutečných bodů kalibračního obrazce v obraze od bodů popsaných lineární transformací

Pozn. obdobně mohou být parametry radiálního zkreslení určeny během výpočtu homografie

DEPARTMENT OF CYBERNETICS

Rozklad matice projekce $\mathbf{P}$

- metodou kalibrace kamery získáváme přímo projekční matici $\mathbf{P}$ jako celek (tedy matici 3×4)
- matici můžeme použít pro případnou 3D rekonstrukci a není bezprostředně nutné znát jednotlivé vnitřní a vnější parametry kamery
- pokud však tyto parametry potřebujeme určit, musíme získanou projekční matici rozložit do zmíněného maticového součinu
- tedy víme, že $\mathbf{P} = \mathbf{KR}[\mathbf{I}, -\mathbf{C}]$

DEPARTMENT OF CYBERNETICS

- dále matici budeme značit jako:
- $\mathbf{P} = [\mathbf{Q}, \mathbf{q}] \ \rightarrow \ \mathbf{KR}[\mathbf{I}, -\mathbf{C}]$, kde
- $\mathbf{Q} = \mathbf{KR} \in \mathbb{R}^{3,3}$ je čtvercová matice, pro perspektivní kameru má plnou hodnost
- $\mathbf{q} \in \mathbb{R}^{3,1}$ střed systému světových souřadnic
- $\mathbf{K} \in \mathbb{R}^{3,3}$ je čtvercová matice horní trojúhelníková
- $\mathbf{R} \in \mathbb{R}^{3,3}$ je čtvercová matice rotace, je ortogonální ($\mathbf{R}^{-\mathbf{1}} = \mathbf{R}^{\mathbf{T}}$ a tedy $\mathbf{R}^{\mathbf{T}}\mathbf{R} = \mathbf{I}$)
  pozn. vektory sloupců takové matice mají jednotkovou normu a jsou na sebe kolmé

- Pro určení vnitřních parametrů můžeme například použít **QR rozklad**, resp. variantu RQ
- z teorie je RQ rozklad je dekompozice nějaké matice A tak, že platí $A = RQ$ za podmínky, že R je horní trojúhelníková matice a Q je ortogonální matice (*Nezaměnit s naším označením pro matice* $\mathbf{Q}$ *a* $\mathbf{R}$*!*)
- jednou z možností takového rozkladu je použití Givensových rotací
- postupně uvažujeme násobení matice $\mathbf{Q}$ (*ta naše co vznikla z matice* $\mathbf{P}$) zprava maticemi $\mathbf{R}_1$, $\mathbf{R}_2$ a $\mathbf{R}_3$ tak, aby platilo:

$$\mathbf{K} = \mathbf{Q}\mathbf{R_1}\mathbf{R_2}\mathbf{R_3}, \quad \mathbf{R_1} = \begin{bmatrix} c & s & 0 \\ -s & c & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

kde $c^2 + s^2 = 1$ apod. $\mathbf{R_2}$ a $\mathbf{R_3}$

DEPARTMENT OF CYBERNETICS

Rotace kamery oproti světovým souřadnicím:

- $\mathbf{K} = \mathbf{Q}\mathbf{R_1}\mathbf{R_2}\mathbf{R_3}, \quad \mathbf{R_1} = \begin{bmatrix} c & s & 0 \\ -s & c & 0 \\ 0 & 0 & 1 \end{bmatrix}$
- chceme tedy na levé straně dostat horní trojúhelníkovou matici
- postupně tedy, hledám nejprve úhel reprezentovaný maticí $\mathbf{R}_1$ $\rightarrow$ na pozici $\mathbf{Q}_{32}$ byl nulový. Pak hledám druhý úhel matice $\mathbf{R}_2$ tak, aby prvek na pozici $\mathbf{Q}_{31}$ byl nulový
- nakonec najdu třetí úhel v matici $\mathbf{R}_3$, aby i třetí prvek pod diagonálou byl nulový, tedy prvek $\mathbf{Q}_{21}$
- máme tedy 3 vnější parametry ... rotaci kamery ve světových souřadnicích

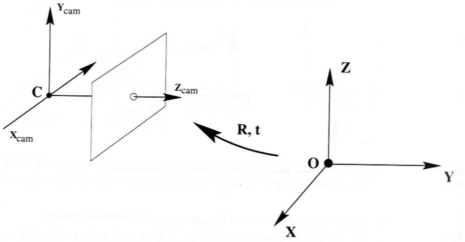
Ohnisková vzdálenost, posun počátku obrázku a kolmost os obrázku:

- přenásobením původní matice $\mathbf{Q}$ zprava všemi třemi získanými maticemi získávám kalibrační matici $\mathbf{K}$ a tedy potažmo i vnitřní parametry kamery: ohnisko, velikost pixelu, posun počátku obrázku a úhel os obrázku
- tedy $\mathbf{K} = \mathbf{Q}\mathbf{R_1}\mathbf{R_2}\mathbf{R_3}$

Optický střed:

- poslední 3 vnější parametry jsou pro posun kamery od počátku světových souřadnic
- optický střed má tu vlastnost, že projekce tohoto bodu (C) je nulová
- pak $\mathbf{PC} = \mathbf{0}$ a $\mathbf{C}$ jsou prostorové souřadnice středu projekce
- můžeme odvodit vztah pro určení optického středu jako:

$$\mathbf{0} = \mathbf{PC} = [\mathbf{Q}, \mathbf{q}] \begin{bmatrix} \mathbf{C} \\ 1 \end{bmatrix} = \mathbf{QC} + \mathbf{q} \Rightarrow \mathbf{C} = -\mathbf{Q}^{-1}\mathbf{q}$$

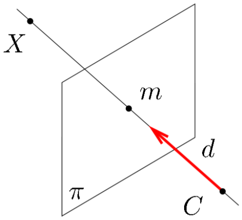


DEPARTMENT OF CYBERNETICS

Uveďme ještě další vlastnosti z odvozené matice projekce, které nejsou vnitřní ani vnější parametry kamery:

- optický paprsek
- optická rovina
- vztah optická rovina a optický paprsek
- ...

Optický paprsek:



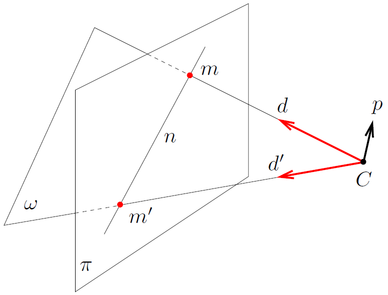
- optický paprsek je vektor, který směřuje z optického středu (C) směrem k prostorovému bodu (X)
- v obrazové rovině pak určuje bod m
- bod v prostoru je dán jako:

$$\mathbf{X} = \mathbf{C} + \lambda\mathbf{d} = \mathbf{C} + \lambda\mathbf{Q}^{-1}\mathbf{m} \Rightarrow \mathbf{d} = \mathbf{Q}^{-1}\mathbf{m} \tag{9}$$

DEPARTMENT OF CYBERNETICS

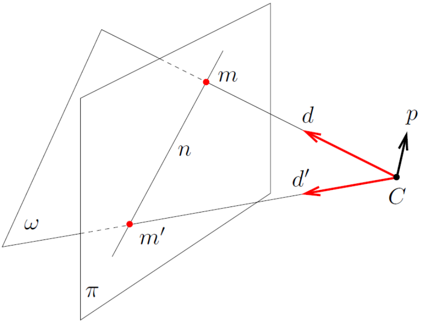
Optická rovina:

- optická rovina je prostorová rovina, procházející optickým středem a určující přímku v obrazové rovině.
- potom optický paprsek daný bodem m je $\mathbf{d} = \mathbf{Q}^{-1}\mathbf{m}$
- druhý paprsek jako $\mathbf{d}' = \mathbf{Q}^{-1}\mathbf{m}'$



DEPARTMENT OF CYBERNETICS

- po dosazení získáme vztah pro normálový vektor optické roviny jako:
  $\mathbf{p} = \mathbf{d} \times \mathbf{d}' = \mathbf{Q}^\mathsf{T}(\mathbf{m} \times \mathbf{m}') = \mathbf{Q}^\mathsf{T}\mathbf{n}$



Poznámka: optický paprsek z tohoto pohledu je si možné představit také jako průsečík dvou optický rovin

DEPARTMENT OF CYBERNETICS

Shrnutí

- $\mathbf{P} = [\mathbf{Q}, \mathbf{q}] = \begin{bmatrix} \mathbf{q}_1^T & q_{14} \\ \mathbf{q}_2^T & q_{24} \\ \mathbf{q}_3^T & q_{34} \end{bmatrix} = \mathbf{KR}[\mathbf{I}, -\mathbf{C}]$
- $\mathbf{K} = \begin{bmatrix} fk_u & -fk_u \coth(\theta) & u_0 \\ 0 & fk_v / \sin(\theta) & v_0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} \alpha_x & s & x_0 \\ 0 & \alpha_y & y_0 \\ 0 & 0 & 1 \end{bmatrix}$
- $\mathbf{R}$ ... orientace kamery, ortogonální matice 3×3
- $\mathbf{C} = \mathit{rnull}(\mathbf{P})$ ... optický střed
- $\mathbf{d} = \mathbf{Q}^{-1}\mathbf{m}$ ... optický paprsek
- $\mathit{det}(\mathbf{Q})\mathbf{q}_3$ ... optická osa
- $\mathbf{Qq_3}$ ... principal point
- $\mathbf{p} = \mathbf{Q^T n}$ ... optická rovina (n je přímka v obrazové rovině)

DEPARTMENT OF CYBERNETICS

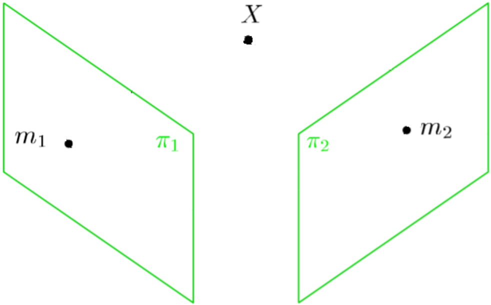
Projektivní geometrie dvou pohledů

- projektivní geometrie může být dále rozšířena v případě, že máme dva pohledy na stejnou scénu z různých směrů
- dva pohledy mohou být získány souběžně v jeden okamžik (dva přístroje)
- nebo jedním přístrojem postupně za sebou (pohyb přístroje před objektem)
- podobně pohyb objektu před přístrojem (stejné)
- tyto úlohy jsou duální a vedou na stejné řešení

DEPARTMENT OF CYBERNETICS

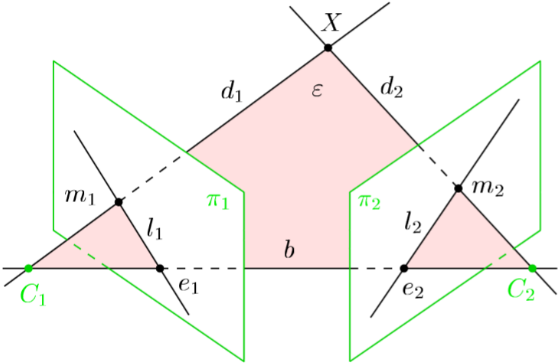
- první a druhý pohled je popsán maticí projekce $\mathbf{P}$ resp. $\mathbf{P}'$
- označení $'$ použije pro druhý pohled
- dále platí ... $\mathbf{x} = \mathbf{PX}$ a $\mathbf{x}' = \mathbf{P}'\mathbf{X}$
- body v obraze $\mathbf{x}$ a $\mathbf{x}'$ jsou tzv. sobě korespondující body protože pocházejí projekcí od stejného prostorového bodu $\mathbf{X}$
- projektivní geometrie dvou pohledů umožňuje řešit následující skupiny problémů:
  - geometrie korespondence - pro daný bod $\mathbf{x}$ nás zajímá, v jaké části v druhém obrazu je jeho korespondent $\mathbf{x}'$
  - geometrie pohybu kamery - známe množinu sobě korespondujících obrazových bodů a zajímá nás, jaký pohyb je kamery (tedy pohyb mezi snímkem jedna a dva)
  - geometrie trojrozměrné scény - opět známe sobě korespondující obrazové body a zajímá nás jejich pozice ve 3D

Epipolární geometrie



- epipolární geometrie popisuje vzájemný vztah dvou pohledů na scénu
- všechny vztahy a výpočty epipolární geometrie jsou nezávislé na geometrii scény
- závisí pouze na vnitřních parametrech daných dvou pohledů a na jejich vzájemné pozici

DEPARTMENT OF CYBERNETICS



- b je **báze**, spojnice středů kamer $b = C_2 - C_1$
- epipól $e_i \in \pi_i$ je obraz středů v obrazové rovině, $\mathbf{e_1} = \mathbf{P_1 C_2}$ a $\mathbf{e_2} = \mathbf{P_2 C_1}$
- $l_i \in \pi_i$ je obraz roviny (**epipolární roviny**) spojující prostorové body $\epsilon = (C_2, X, C_1)$
- l_i je **epilopární přímka**

DEPARTMENT OF CYBERNETICS

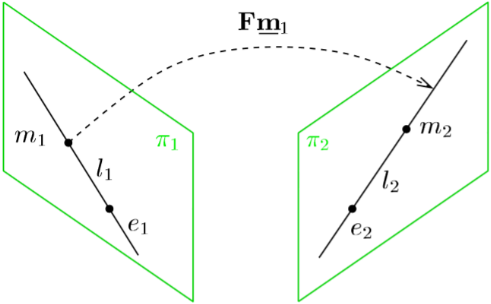
Epipolární podmínka

- vychází z podmínky, že d_1, d_2 a b leží v jedné rovině
- epipolární geometrii je možné zcela zahrnout do jedné 3×3 matice, tzv. fundamentální matice
- obrazový bod $\mathbf{x}$ v prvním pohledu a bod $\mathbf{x}'$ z druhého pohledu jsou projekce společného bodu X pak platí:

$$\mathbf{x}'^T \mathbf{F} \mathbf{x} = 0 \tag{10}$$

- tento vztah se nazývá **epipolární podmínka**
- matice $\mathbf{F}$ se nazývá **fundamentální matice**

▶



- pokud $\mathbf{F}$ popisuje vztah kamer $\mathbf{P_1}$ a $\mathbf{P_2}$ pak transponovaná matice $\mathbf{F^T}$ pak popisuje kamery $\mathbf{P_2}$ a $\mathbf{P_1}$
- epipolární přímka v druhém obraze je vyjádřena jako $\mathbf{l}' = \mathbf{Fx}$ a podobně $\mathbf{l} = \mathbf{F^T x}'$
- pro epipól v druhém obraze platí $\mathbf{e}'^{\mathbf{T}}\mathbf{F} = \mathbf{0}$, tedy je to levý nulový prostor fundamentální matice, a podobně pro epipól v prvním obraze $\mathbf{Fe} = \mathbf{0}$ je pravý nulový prostor.

DEPARTMENT OF CYBERNETICS

Fundamentální matice

- fundamentální matice **F** je čtvercová homogenní matice velikosti 3×3
- hodnost matice je 2, je singulární ($det(\mathbf{A}) = 0$), a má 7 stupňů volnosti
- má pravý a levý nulový prostor, které odpovídají epipólům
- fundamentální matice může být určena numerickým výpočtem pouze z několika sobě korespondujících bodů
- princip výpočtu vychází ze zmíněné epipolární podmínky, která musí platit pro každý pár bodů

DEPARTMENT OF CYBERNETICS

Určení fundamentální matice

- mějme tedy několik sobě korespondujících bodů $x \leftrightarrow x'$
- nechť platí epipolární podmínka $\mathbf{x}'\mathbf{F}\mathbf{x} = \mathbf{0}$
- pokud je $\mathbf{x} = (x, y, 1)$ a $\mathbf{x}' = (x', y', 1)$, pak epipolární podmínku můžeme pro jeden pár rozepsat jako:

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

- tento vztah určuje sloupcový 9×1vektor $\mathbf{f}$
- $\mathbf{f}$ představuje postupně řádky matice $\mathbf{F}$
- parametry předchozí rovnice formují jeden řádek (rovnici) lineární soustavy rovnic

$$(x'x, x'y, x', y'x, y'y, y', x, y, 1) \tag{12}$$

DEPARTMENT OF CYBERNETICS

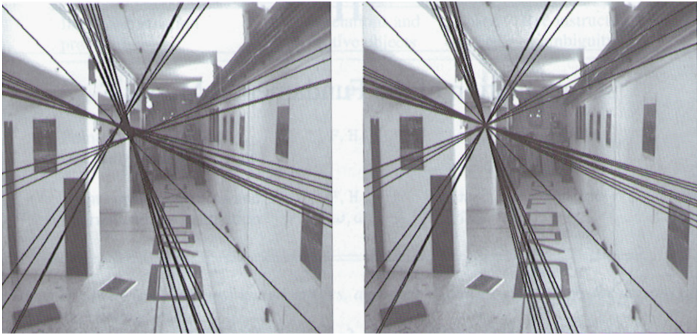
- pro n vzájemné korespondujících bodů v prvním a druhém obraze získáme maticový zápis n rovnic :

$$\mathbf{Af} = \begin{bmatrix} x'_1x_1 & x'_1y_1 & x'_1 & y'_1x_1 & y'_1y_1 & y'_1 & x_1 & y_1 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x'_nx_n & x'_ny_n & x'_n & y'_nx_n & y'_ny_n & y'_n & x_n & y_n & 1 \end{bmatrix} \mathbf{f=0} \tag{13}$$

- pokud je hodnost matice **A** právě 8, pak existuje právě jedno netriviální řešení, což je pravý nulový prostor matice **A**
- v důsledku přítomnosti šumu je tato soustava přeurčená a tedy často nemá řešení
- základní postup nalezení odhadu řešení je výpočet takového **f** s nejmenším kvadrátem chyby ... postup je totožný s postupem použití SVD rozkladu např. při určení homografie
- tedy minimalizujeme normu $||\mathbf{Af}||$ s podmínkou měřítka $||\mathbf{f}|| = 1$
- v tomto kontextu nalezení řešení používáme označení **8-bodový algoritmus**.

DEPARTMENT OF CYBERNETICS

- pokud není $\mathbf{F}$ singulární ($rank(\mathbf{F} > 2)$), pak se epipolární přímky neprotínají v jednom bodě (chyba)



- pro zajištění hodnosti 2 můžeme provést znovu SVD rozkladu
- opravenou $\mathbf{F}'$ získáme zpětně přenásobením $\mathbf{F} = \mathbf{UDV^T}$, kde upravíme $\mathbf{D} = diag(r, s, t)$ s tím $r \geq s \geq t$
- a fundamentální matice $\mathbf{F}' = \mathbf{U}diag(r, s, 0)\mathbf{V^T}$ minimalizuje Frobeniovu normu $||\mathbf{F} - \mathbf{F}'||$ a jedná se o nejbližší singulární matici k původní matici $\mathbf{F}$

DEPARTMENT OF CYBERNETICS

Další možnosti zpřesnění nalezeného odhadu řešení fundamentální matice:

- **A** je obecně předurčena ke špatné numerické stabilitě výpočtu, protože **A** je špatně podmíněná (velká rozdílnost prvků ... v řádech)
- proto se často matice před vlastním výpočtem normalizuje tzv. umělé snížení rozdílnosti její prvků
- další úprava pro zvýšení robustnosti výpočtu fundamntální matice je použití pouze 7 bodů a následný postprocesing (tzv. 7 bodový algoritmus)

DEPARTMENT OF CYBERNETICS

Automatický výpočet fundamentální matice

- často nemáme jistotu, že všechny nalezené páry jsou korespondenti (jsou nalezeny automaticky)
- pak musíme použít nějakou metodu iterativního zkoušení a hledání řešení, např.:
  - použití *Least median of squared residuals* (LMedS) (r. 1984), metoda předpokládá alespoň 50 % všech párů jsou správní korespondenti, nebo
  - Random Sample Consensus (RANSAC) (r. 1981) až 90 % bodů může být špatně korespondující

DEPARTMENT OF CYBERNETICS

Esenciální matice

- historicky byla esenciální matice popsána dříve než matice fundamentální (r. 1981)
- tato matice má méně stupňů volnosti než fundamantální matice, jen 5 stupňů volnosti
- označme projekční matice obou pohledů jako:
  $\mathbf{P} = \mathbf{K}[\mathbf{I}, -\mathbf{C_1}]$
  $\mathbf{P}' = \mathbf{K}'\mathbf{R}[\mathbf{I}, -(\mathbf{C_1} + \mathbf{b})]$
  a vektor
  $\mathbf{b} = (b_1, b_2, b_3)$
- matice **R** představuje relativní rotaci druhé kamery vůči kameře první a vektor **b** je posun středu druhé kamery od první kamery

$$\mathbf{x}'^{\mathbf{T}}\mathbf{F}\mathbf{x} = \mathbf{x}'^{\mathbf{T}}\mathbf{K}'^{-\mathbf{T}}\mathcal{E}\mathbf{K}^{-\mathbf{1}}\mathbf{x} = \mathbf{0} \tag{14}$$

- matice $\mathcal{E}$ se nazývá esenciální matice a tedy

$$\mathbf{F} = \mathbf{K}'^{-\mathbf{T}}\mathcal{E}\mathbf{K}^{-\mathbf{1}} \tag{15}$$

DEPARTMENT OF CYBERNETICS

- a naopak přenásobením kalibrační maticí zprava a zleva získáme vztah pro výpočet esenciální matice z fundamentální matice jako:

$$\mathcal{E} = \mathbf{K}'^{\mathsf{T}}\mathbf{F}\mathbf{K} \tag{16}$$

- pro esenciální matici dále platí:

$$\mathcal{E} = \mathbf{R}\mathbf{S}(\mathbf{b}) \tag{17}$$

- kde $\mathbf{S}(\mathbf{b})$ je antisymetrická matice:

$$\mathbf{S}(\mathbf{b}) = \begin{bmatrix} 0 & -b_3 & b_2 \\ b_3 & 0 & -b_1 \\ -b_2 & b_1 & 0 \end{bmatrix} \tag{18}$$

  - esenciální matice zachycuje relativní pozici druhé kamery oproti první kameře
  - hodnost matice $rank(\mathcal{E}) \leq 2$ (5 stupňů volnosti $=$ 3 rotace $+$ 3 translace -1 nejednoznačnost rozkladu - první dvě vlastní čísla jsou stejné a třetí je nulové

DEPARTMENT OF CYBERNETICS

Odhad pohybu kamery

- odhad pohybu kamery spočívá v určení vektoru báze $\mathbf{b}$ a matice rotace $\mathbf{R}$ ze známé množiny korespondujících bodů $k \geq 8$
- jde tedy o nalezení a rozklad esenciální matice
- v základu může být esenciální matice určena z fundamentální matice použitím kalibračních matic obou pohledů
- pokud neznámě tyto dvě projekční matice ... odhad esenciální matice stejným způsobem jako odhad fundamentální matice
  - jak? Pokud první pohled platí $\mathbf{P} = \mathbf{K}[\mathbf{R}|\mathbf{t}]$ a platí, že $\mathbf{x} = \mathbf{PX}$
  - pak použijeme vztahu reprojekce s kalibrační maticí a získáme bod $\hat{\mathbf{x}} = \mathbf{K}^{-\mathbf{1}}\mathbf{x}$, který nazýváme bod v *normalizovaných souřadnicích*
  - jinými slovy bod v těchto souřadnicích získáme projekcí s jedničkovou kalibrační maticí $\mathbf{I}$ (ohnisková vzdálenost je 1) a bod měřený v obraze vydávat za tento bod

- potom pro projekci mluvíme o *matici normalizované kamery*, ztrácíme měřítko.
- dostaneme první pohled jako $\mathbf{P} = [\mathbf{I}|\mathbf{0}]$ a druhý pohled se stejnou projekcí, ale s pootočením a posunem je $\mathbf{P}' = [\mathbf{R}|\mathbf{b}]$
- epipolární podmínka pro formulování rovnic získá tvar:

$$\hat{\mathbf{x}}'^{\mathbf{T}} \mathcal{E} \hat{\mathbf{x}} = \mathbf{0} \tag{19}$$

- esenciální matici pak můžeme určit stejným postupem jako matici fundamentální, např. 8-bodovým algoritmem

DEPARTMENT OF CYBERNETICS

Rozklad esenciální matice

- rozklad esenciální matice na $\mathcal{E} = \mathbf{SR}$ použijeme známý SVD rozklad
- nejprve si definuje pomocné matice

$$\mathbf{W} = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{20}$$

matice $\mathbf{W}$ je ortogonální

$$\mathbf{Z} = \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \tag{21}$$

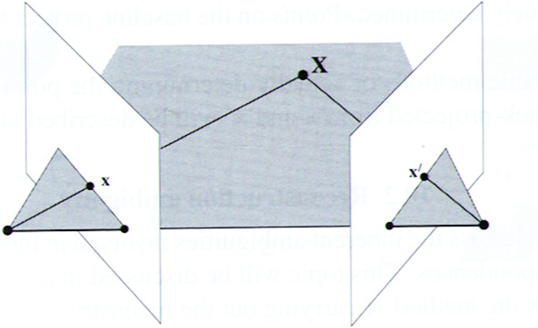
a matice $\mathbf{Z}$ je antisymentrická

DEPARTMENT OF CYBERNETICS

- pokud SVD rozklad $\mathcal{E} = \mathbf{UDV^T}$
- matice rotace je možné určit dvěma způsoby, buď jako $\mathbf{R} = \mathbf{UWV^T}$ nebo jako $\mathbf{R} = \mathbf{UW^TV^T}$ (neznámé znamínko)
- dále matice $\mathbf{S}$ je dána jako $\mathbf{S} = \mathbf{UZU^T}$ resp. $\mathbf{b} = \mathbf{V}(\mathbf{0}, \mathbf{0}, |\beta|)^\mathbf{T}$ (opět neznámé znamínko)
- jinými slovy máme dvě možnosti jak určit matici $\mathbf{R}$ a dvě možnosti (znaménko) jak určit posun
- v kombinaci získáme celkem čtyři možná (správná) řešení pro matici $\mathbf{P}'$
- avšak pouze jedno z těchto řešení představuje pozorování scény před oběma pohledy (kamerami).

Poznámka: esenciální (fundamentální) matici je možné použít přímo pro 3D rekonstrukci z páru nezkalibrovaných kamer (až r. 1990) Existuje několik postupů a doplňujících podmínek jak získat správnou metrickou rekonstrukci.

3D rekonstrukce - Linear triangulation



- pokud známe dvě projekční matice $\mathbf{P}$ a $\mathbf{P}'$
- a současně máme korespondující pár, tedy obrazový bod v prvním pohledu $\mathbf{x}$ a obrazový bod v druhém pohledu $\mathbf{x}'$
- pak můžeme určit 3D souřadnici neznámého bodu

Poznámka: jediné body,které nelze rekonstruovat jsou body ležící na spojnici mezi středy kamer

DEPARTMENT OF CYBERNETICS

- pro tento pár musí platit epipolární podmínka, tedy $\mathbf{x}'^{\mathsf{T}}\mathbf{F}\mathbf{x} = \mathbf{0}$
- pro bod $\mathbf{x}$ můžeme očekávat bod $\mathbf{x}'$ na epipolární přímce
- rovina spojující prostorový bod $\mathbf{X}$ a středy kamer určuje epipolární přímky v obou pohledech
- platí $\mathbf{x} = \mathbf{P}\mathbf{X}$ a $\mathbf{x}' = \mathbf{P}'\mathbf{X}$ a můžeme převést problém na formu $\mathbf{A}\mathbf{X} = \mathbf{0}$
- měřítko dané homogenní souřadnicí je eliminováno vektorovým součinem např. pro první pohled jako $\mathbf{x} \times \mathbf{P}\mathbf{X} = 0$

$$
\begin{aligned}
x(\mathbf{p}^{3\mathsf{T}}\mathbf{X}) - (\mathbf{p}^{1\mathsf{T}}\mathbf{X}) &= 0 \\
y(\mathbf{p}^{3\mathsf{T}}\mathbf{X}) - (\mathbf{p}^{2\mathsf{T}}\mathbf{X}) &= 0 \\
x(\mathbf{p}^{2\mathsf{T}}\mathbf{X}) - y(\mathbf{p}^{1\mathsf{T}}\mathbf{X}) &= 0
\end{aligned}
$$

DEPARTMENT OF CYBERNETICS

- jen 2 rovnice jsou vždy lineárně nezávislé
- po vypuštění třetí rovnice dostáváme vztah pro určení matice **A** složené ze čtyř rovnic (odpovídající dvou bodům)
- každý bod po dvou souřadnicích korespondenčního páru
- řešením soustavy získáváme neznámý vektor 4×1 **X**, tedy 3D rekonstrukci

$$\mathbf{A} = \begin{bmatrix} x\mathbf{p}^{3\mathsf{T}} - \mathbf{p}^{1\mathsf{T}} \\ y\mathbf{p}^{3\mathsf{T}} - \mathbf{p}^{2\mathsf{T}} \\ x'\mathbf{p}'^{3\mathsf{T}} - \mathbf{p}'^{1\mathsf{T}} \\ y'\mathbf{p}'^{3\mathsf{T}} - \mathbf{p}'^{2\mathsf{T}} \end{bmatrix} \tag{22}$$

- často měření obsahuje šum a soustava nemá netriviální řešení
- nalezení nejlepšího odhadu nezámého bodu je podobné předchozím případům
- tedy SVD rozklad $\mathbf{A} = \mathbf{UDV^T}$ a řešení je poslední sloupec matice $\mathbf{V}$ odpovídající nejmenšímu vlastnímu číslu
- nalezený odhad řešení dehomogenizujeme tak, $\mathbf{X} = (x_1, x_2, x_3, x_4)^T$ podělíme poslední (homogenní) souřadnicí, tedy:
  $(x, y, z)^T = (x_1/x_4, x_2/x_4, x_3/x_4)^T$

Algoritmus tzv. řídké 3D stereo rekonstrukce:

- použiji nějaký detektor významných bodů, detekce tzv. stabilních bodů ... invariantních změně a posunu pohledu
- spočítat lokální deskriptor každého dobu v obou pohledech
- provést předběžné párování nalezených bodů porovnáním vektorů z deskriptoru
- nalézt pouze správné korespondenty např. s pomocí metody **RANSAC** a splněním epipolární podmínky
- provést 3D rekonstrukci těchto správných korespondentů