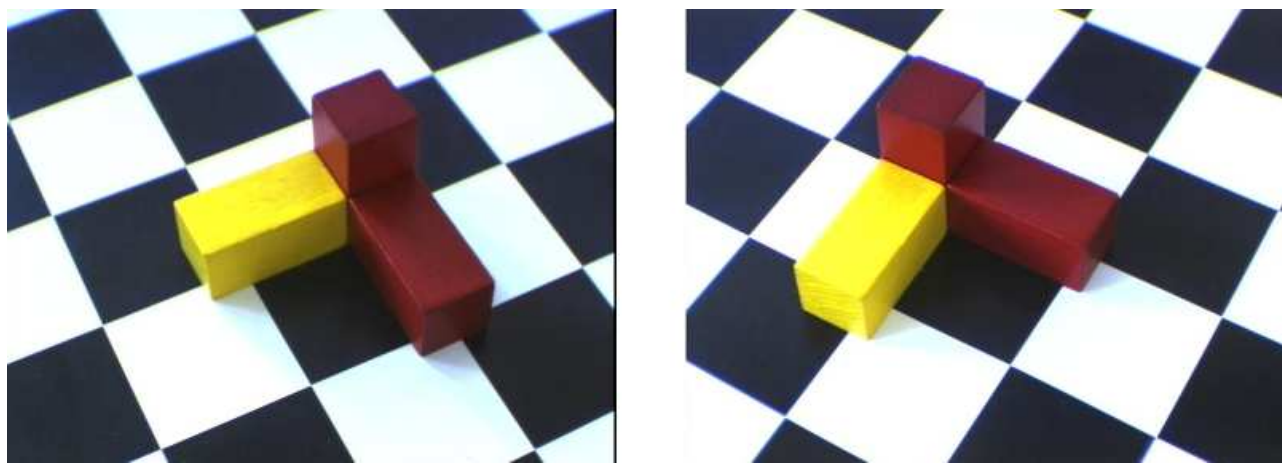


## Úloha 2 - Rekonstrukce 3D scény

### 1 Zadání

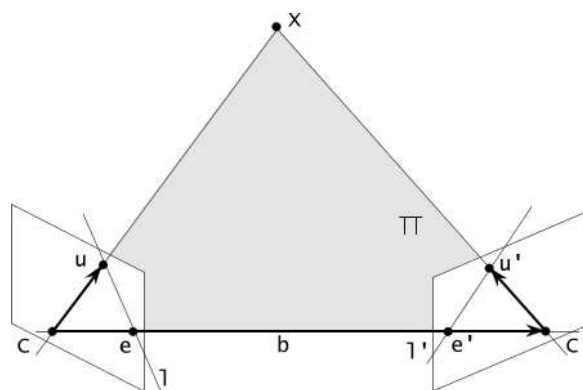
Proveďte rekonstrukci scény ze dvou obrazů. Nalezněte epipóly a epipoláry a transformační matice obou kamer a na jejich základě vypočtete souřadnice bodů rekonstruované scény.



Obr. 1. Snímky pro rekonstrukci

### 2 Úvod

Nejdříve za pomoci nástroje CORRGUI v Matlabu nalezneme množinu dvojic korespondujících bodů z prvního a druhého obrázku. Tyto korespondence pak použijeme k výpočtu fundamentální matice  $\mathbf{F}$ .



Obr. 2. Epipolární geometrie

Co to vlastně fundamentální matice je, si vysvětlíme na obrázku 2. Mějme dvě kamery  $C$  a  $C'$  sledující scénu, ve které se nachází také bod  $X$ . Tento bod se promítne do bodů  $u$  a  $u'$  příslušných

obrazů. Body  $e$  a  $e'$  jsou obrazy druhé kamery v prvním obrazu, resp. první kamery v druhém obrazu; nazývají se *epipóly*. Přímkou spojující obrazy  $u$ ,  $u'$  a epipóly se nazývají *epipoláry*.

Vzhledem k tomu, že body  $X$ ,  $C$  a  $C'$  leží v jedné rovině, musí platit (souřadnou soustavu spojenou s prvním obrázkem označíme  $\beta$ , soustavu druhého obrázku  $\beta'$ ):

$$u_\beta \cdot (b_\beta \times u'_\beta) = 0 \quad \Rightarrow \quad u_\beta \cdot (b_\beta \times \mathbf{A}u'_{\beta'}) = 0,$$

kde  $\mathbf{A}$  je matice přechodu z báze  $\beta'$  do báze  $\beta$ , tedy z báze druhého obrázku do prvního. Tuto rovnici můžeme také zapsat tak, že vektorový součin vyjádříme pomocí maticového násobení:

$$u_\beta [b_\beta]_\times \mathbf{A} u'_{\beta'} = 0,$$

nebo také

$$u_\beta \mathbf{F} u'_{\beta'} = 0.$$

Fundamentální matice nám tedy dává do souvislosti dva obrazy jednoho bodu zachycené různými kamerami. Tato matice má hodnotu 2, vzhledem k tomu, že matice pro vektorový součin má hodnotu 2 a matice přechodu  $\mathbf{A}$  má hodnotu 3.

Souřadnice bodů  $u_\beta$ ,  $u'_{\beta'}$  máme již k dispozici - jsou to homogenní souřadnice korespondujících bodů:  $u_\beta = (u, v, 1)^T$ ,  $u'_{\beta'} = (u', v', 1)^T$ . Matici  $\mathbf{F}$  nalezneme řešením následující soustavy rovnic:

$$\begin{pmatrix} u_1 u'_1 & u_1 v'_1 & u_1 & v_1 u'_1 & v_1 v'_1 & v_1 & u'_1 & v'_1 & 1 \\ u_2 u'_2 & u_2 v'_2 & u_2 & v_2 u'_2 & v_2 v'_2 & v_2 & u'_2 & v'_2 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ u_n u'_n & u_n v'_n & u_n & v_n u'_n & v_n v'_n & v_n & u'_n & v'_n & 1 \end{pmatrix} \begin{pmatrix} f_{11} \\ f_{12} \\ f_{13} \\ f_{21} \\ f_{22} \\ f_{23} \\ f_{31} \\ f_{32} \\ f_{33} \end{pmatrix} = 0,$$

$$\mathbf{F} = \begin{pmatrix} f_{11} & f_{12} & f_{13} \\ f_{21} & f_{22} & f_{23} \\ f_{31} & f_{32} & f_{33} \end{pmatrix}.$$

Matice  $\mathbf{F}$  je určena osmi parametry (jeden můžeme zvolit díky singularitě). Proto musíme vyjít minimálně z 8 korespondencí. Soustavu můžeme řešit např. pomocí rozkladu na singulární čísla (SVD). Může se nám stát, že výsledná matice nebude mít hodnotu 2, ale 3, pokud jsme použili více než osm korespondencí a naměřené hodnoty jsou nepřesné (a to jsou téměř vždy). Potom nalezenou matici  $\mathbf{F}$  opět rozložíme pomocí SVD, upravíme rozklad SVD tak, že vynulujeme třetí prvek na diagonále prostřední (diagonální matice), a potom matice získané rozkladem opět mezi sebou vynásobíme.

Nyní, když už máme matici  $\mathbf{F}$ , můžeme vypočítat a vykreslit epipóly a epipoláry. Epipóly jsou body obrazu, do kterých se zobrazí kamery. Platí:

$$e^T \mathbf{F} = 0,$$

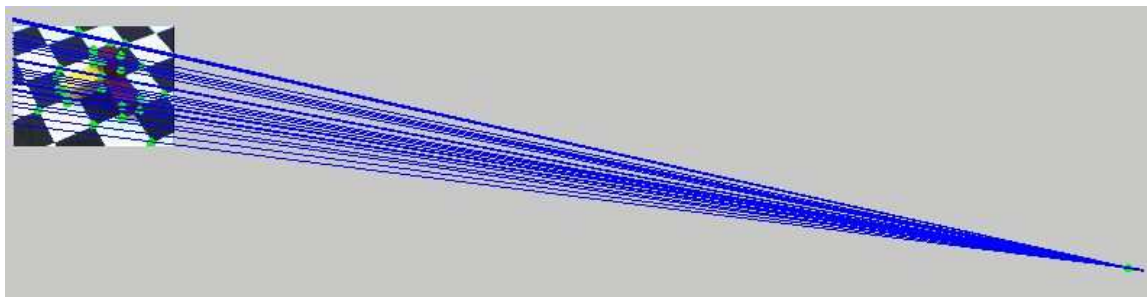
$$\mathbf{F} e' = 0.$$

Jinými slovy vyjádřeno, epipóly představují bázi levého a pravého nulového prostoru matice  $\mathbf{F}$ . Epipoláry jsou přímky, procházející epipóly, do kterých se zobrazují body druhého obrazu:

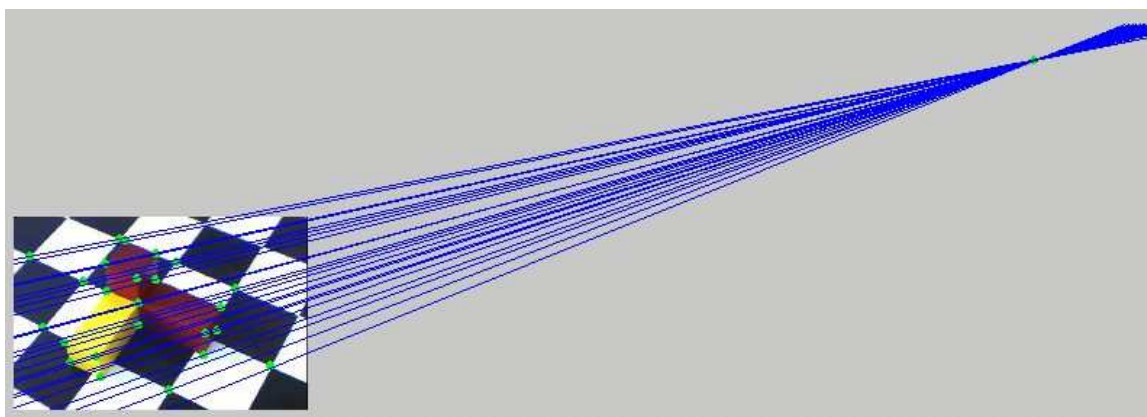
$$l = \mathbf{F} u',$$

$$l' = u^T \mathbf{F}.$$

Složky vektorů  $l$ ,  $l'$  nám udávají rovnici přímky; můžeme proto jednoduše zobrazit do obrázku epipoláry (zvolíme si  $x$ -ové souřadnice krajních bodů úsečky a  $y$ -ové dopočítáme). Výsledek je na obrázcích 3 a 4.



Obr. 3. Epipoláry pro první obrázek



Obr. 4. Epipoláry pro druhý obrázek

### 3 Rekonstrukce scény

V této části si ukážeme postup, jak z obrazů  $u$ ,  $u'$  bodů scény můžeme tuto scénu rekonstruovat. Nejdříve budeme hledat vhodné projekční matice pro obrazy; ukážeme si dva způsoby - faktorizaci a výpočet projekčních matic pomocí fundamentální matice. V prvním případě jsme jako vedlejší produkt získali souřadnice rekonstruovaných bodů, v druhém je musíme dopočítat. Nakonec musíme najít afinní zobrazení (homografii), které nám dává do souvislosti body rekonstruovaného a skutečného prostoru.

#### 3.1 Nalezení projekčních matic faktorizací

Princip faktorizace spočívá v řešení soustavy rovnic, vyjadřující vztahy mezi všemi body scény, obrazy i kamerami. Vztah mezi bodem a jeho obrazem můžeme zapsat jako

$$\alpha u = \mathbf{P}X,$$

kde  $\alpha \in \mathbb{R} \setminus \{0\}$  a projekční matice  $\mathbf{P} \in \mathbb{R}^{3,4}$ . Máme-li  $n$  bodových korespondencí v  $k$  obrazech, bude mít tato soustava tvar

$$\begin{pmatrix} \alpha_{11}u_{11} & \alpha_{12}u_{12} & \cdots & \alpha_{1n}u_{1n} \\ \alpha_{21}u_{21} & \alpha_{22}u_{22} & \cdots & \alpha_{2n}u_{2n} \\ \vdots & \vdots & & \vdots \\ \alpha_{21}u_{21} & \alpha_{22}u_{22} & \cdots & \alpha_{2n}u_{2n} \end{pmatrix} = \begin{pmatrix} \mathbf{P}_1 \\ \mathbf{P}_2 \\ \vdots \\ \mathbf{P}_k \end{pmatrix} \cdot (X_1 X_2 \cdots X_n),$$

nebo ve zkráceném zápisu  $\alpha_{ij}u_{ij} = \mathbf{P}_i X_j$ . V případě, že chceme jen najít nějaké řešení (a nezáleží nám na tom, jestli zachovává topologii scény), můžeme v matici na levé straně rovnice vynechat projektivní hloubky (koeficienty  $\alpha_{ij}$ ). Označme tedy  $\mathbf{R} = (u_{ij})$ . Hledáme rozklad této matice na součin dvou matic  $\mathbf{P}$  a  $\mathbf{X}$ . Tento rozklad nalezneme pomocí již několikrát zde zmíněného SVD. Musíme se ale postarat (omezením matic  $\mathbf{U}$ ,  $\mathbf{D}$ ,  $\mathbf{V}$  získaných z SVD), aby výsledné matice měly požadované rozměry ( $\mathbf{R} \in \mathbb{R}^{3k,n}$ ,  $\mathbf{P} \in \mathbb{R}^{3k,4}$ ,  $\mathbf{X} \in \mathbb{R}^{4,n}$ ):

$$\begin{aligned} \mathbf{R} &= \mathbf{U}\mathbf{D}\mathbf{V}^T \\ \mathbf{U} &\rightarrow \mathbf{U}' \in \mathbb{R}^{3k,4} \\ \mathbf{D} &\rightarrow \mathbf{D}' \in \mathbb{R}^{4,4} \\ \mathbf{V}^T &\rightarrow \mathbf{V}'^T \in \mathbb{R}^{4,n} \\ \mathbf{P} &= \mathbf{U}'\mathbf{D}', \quad \mathbf{X} = \mathbf{V}'^T. \end{aligned}$$

Matice  $\mathbf{P}$  potom bude obsahovat  $k$  pod sebou umístěných projekčních matic pro jednotlivé kamery; ty jsou určeny jednoznačně až na afinní zobrazení  $\mathbf{H}$  (viz odstavec 3.4). Rekonstruované souřadnice bodů udávají sloupce matice  $\mathbf{X}$ .

### 3.2 Nalezení projekčních matic volbou jedné a výpočtem druhé

Jiný přístup k tomuto problému spočívá v odvození projekčních matic z fundamentální matice  $\mathbf{F}$ . Je třeba si uvědomit, že i když nám dvě projekční matice  $\mathbf{P}$ ,  $\mathbf{P}'$  určují fundamentální matici jednoznačně, opačné tvrzení neplatí a my můžeme nalézt k jedné fundamentální matici celou třídu projekčních matic  $\mathbf{P}$ ,  $\mathbf{P}'$ ; tyto třídy se budou lišit o afinní transformaci  $\mathbf{H}$ . Mějme dva obrazy  $u$ ,  $u'$  bodu  $X$  pořízené kamerami  $\mathbf{P}$ ,  $\mathbf{P}'$ . Tyto body jsou však také obrazy bodu  $\mathbf{H}^{-1}X$  vzhledem ke kamerám  $\mathbf{P}\mathbf{H}$ ,  $\mathbf{P}'\mathbf{H}$  pro libovolnou afinní transformaci  $\mathbf{H}$ . Při hledání vyhovujících projekčních matic si můžeme tedy jednu z nich zvolit a druhou dopočítat. Často se používá kanonický tvar první projekční matice  $\mathbf{P} = [\mathbf{E} | \mathbf{0}]$  (rekonstrukce bude v souřadné soustavě první kamery). Označíme-li si  $X = [x^T | 1]^T$  a  $\mathbf{P}' = [\mathbf{A} | b]$ , potom má platit:

$$\begin{aligned} u^T \mathbf{F} u' &= 0, \\ (\mathbf{P}X)^T \mathbf{F} \mathbf{P}' X &= 0, \\ x^T \mathbf{F} \mathbf{A} x + x^T \mathbf{F} b &= 0. \end{aligned}$$

Pokud zvolíme  $b = e'$ , zbavíme se tak druhého členu na levé straně poslední rovnice. Vzhledem k tomu, že  $x^T \mathbf{F} \mathbf{A} x = 0$ , musí být matice  $\mathbf{F} \mathbf{A}$  antisymetrická. Matici  $A$  získáme řešením následující soustavy rovnic:

$$\begin{pmatrix} f_{11} & f_{12} & f_{13} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & f_{21} & f_{22} & f_{23} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & f_{31} & f_{32} & f_{33} \\ f_{21} & f_{22} & f_{23} & f_{11} & f_{12} & f_{13} & 0 & 0 & 0 \\ 0 & 0 & 0 & f_{31} & f_{32} & f_{33} & f_{21} & f_{22} & f_{23} \\ f_{31} & f_{32} & f_{33} & 0 & 0 & 0 & f_{11} & f_{12} & f_{13} \end{pmatrix} \begin{pmatrix} a_{11} \\ a_{12} \\ a_{13} \\ a_{21} \\ a_{22} \\ a_{23} \\ a_{31} \\ a_{32} \\ a_{33} \end{pmatrix} = 0,$$

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}.$$

Maticе této soustavy má hodnost 5, hodnost množiny řešení je tedy 4. Matici  $\mathbf{A}$  je pak vytvořena libovolnou lineární kombinací čtyř bázových vektorů. Můžeme si vybrat řešení, které nám bude nejvíce vyhovovat (například matice s největším možným poměrem prvních a posledních hodnot na diagonále). Projekční matice  $\mathbf{P}'$  bude mít potom tvar  $\mathbf{P}' = [\mathbf{A} | e']$ . Toto řešení můžeme také vyjádřit pomocí epipólu  $e'$ :

$$\mathbf{P} = [\mathbf{E} | 0], \quad \mathbf{P}' = [[e']_{\times} \mathbf{F} + e'v^T | \lambda e'],$$

kde  $v \in \mathbb{R}^3$  je libovolný vektor a  $\lambda$  nenulové reálné číslo.

### 3.3 Rekonstrukce scény

Ve chvíli, kdy už máme k dispozici matice  $\mathbf{P}$  a  $\mathbf{P}'$  reprezentující kamery, nám nic nebrání v rekonstrukci scény z korespondujících bodů  $u, u'$ . Mezi rekonstruovanými body a jejich obrazy platí vztahy

$$\begin{aligned} \mathbf{P}X &= \alpha u, \\ \mathbf{P}'X &= \alpha' u', \end{aligned}$$

kde  $\alpha, \alpha'$  jsou opět nenulová reálná čísla a  $X = (x, y, z, w)^T$  je bod scény vyjádřen homogenními souřadnicemi. Označíme-li  $\mathbf{P} = [\mathbf{A} | b]$ ,  $\mathbf{P}' = [\mathbf{A}' | b']$  a  $X = [x^T | 1]^T$  ( $x, b \in \mathbb{R}^3$ ,  $\mathbf{A} \in \mathbb{R}^{3,3}$ ), můžeme předchozí rovnice napsat jako

$$\begin{aligned} \mathbf{A}x + b &= \alpha u \quad \Rightarrow \quad x = \alpha \mathbf{A}^{-1}u - \mathbf{A}^{-1}b, \\ \mathbf{A}'x + b' &= \alpha' u' \quad \Rightarrow \quad x = \alpha' \mathbf{A}'^{-1}u' - \mathbf{A}'^{-1}b'. \end{aligned}$$

Tuto soustavu rovnic můžeme řešit pro neznámé  $\alpha, \alpha'$  a zpětným dosazením získat  $x$ . Ve skutečnosti je situace složitější, protože naše data (souřadnice bodů  $u, u'$  v obraze) nejsou přesnými obrazy bodů scény a při zpětné projekci nemusí odpovídat jedinému bodu  $x$ , ale dvěma bodům  $x_1$  a  $x_2$ . Budeme tedy hledat takové řešení, aby tyto body byly k sobě co nejblíže, z následným zprůměrováním (nalezením středu úsečky s krajními body  $x_1, x_2$ ) najdeme bod  $x$ . Chceme tedy minimalizovat výraz  $\|x_1 - x_2\|^2 = (x_1 - x_2)^T(x_1 - x_2)$ , víme-li, že

$$\begin{aligned} x_1 &= \alpha \mathbf{A}^{-1}u - \mathbf{A}^{-1}b, \\ x_2 &= \alpha' \mathbf{A}'^{-1}u' - \mathbf{A}'^{-1}b'. \end{aligned}$$

Zderivováním podmínky podle  $\alpha$ , resp.  $\alpha'$  získáme rovnice

$$\begin{aligned} (\mathbf{A}^{-1}u)^T(x_1 - x_2) &= 0, \\ (\mathbf{A}'^{-1}u')^T(x_1 - x_2) &= 0, \end{aligned}$$

a po dosazení  $x_1, x_2$

$$\begin{aligned} \alpha(\mathbf{A}^{-1}u)^T \mathbf{A}^{-1}u - \alpha'(\mathbf{A}^{-1}u)^T \mathbf{A}'^{-1}u' &= (\mathbf{A}^{-1}u)^T \mathbf{A}^{-1}b - (\mathbf{A}^{-1}u)^T \mathbf{A}'^{-1}b', \\ \alpha(\mathbf{A}'^{-1}u')^T \mathbf{A}^{-1}u - \alpha'(\mathbf{A}'^{-1}u')^T \mathbf{A}'^{-1}u' &= (\mathbf{A}'^{-1}u')^T \mathbf{A}^{-1}b - (\mathbf{A}'^{-1}u')^T \mathbf{A}'^{-1}b', \end{aligned}$$

Řešením této soustavy pak získáme  $\alpha$  a  $\alpha'$ , pomocí nich vypočteme  $x_1$  a  $x_2$ . Bod  $x$  potom bude ležet ve středu úsečky s krajními body  $x_1, x_2$ .

### 3.4 Nalezení afinní transformace $\mathbf{H}$

Transformační matice  $\mathbf{H}$  v našem případě vyjadřuje vztah mezi body  $\hat{X}$  skutečné a  $X$  rekonstruované scény:

$$\mu X = \mathbf{H}\hat{X},$$

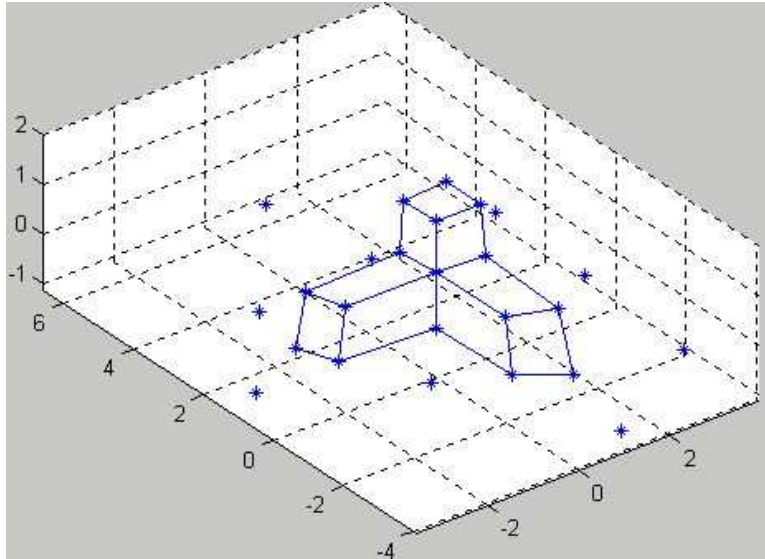
kde  $\mu$  je nenulové reálné číslo specifické pro každý bod. Předpokládejme, že souřadnice bodů  $X$ ,  $\hat{X}$  známe. Z těchto hodnot zjistíme koeficienty matice  $\mathbf{H}$  a hodnoty  $\mu_i$  pro každý bod  $X_i$ . Pro jeden bod vypadá soustava rovnic takto ( $X_i^T = (x_i, y_i, z_i, w_i)$ ,  $\hat{X}_i^T = (\hat{x}_i, \hat{y}_i, \hat{z}_i, \hat{w}_i)$ ):

$$\begin{pmatrix} -x_i & \hat{x}_i & \hat{y}_i & \hat{z}_i & \hat{w}_i & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -y_i & 0 & 0 & 0 & 0 & \hat{x}_i & \hat{y}_i & \hat{z}_i & \hat{w}_i & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -z_i & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \hat{x}_i & \hat{y}_i & \hat{z}_i & \hat{w}_i & 0 & 0 & 0 \\ -w_i & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \hat{x}_i & \hat{y}_i & \hat{z}_i & \hat{w}_i \end{pmatrix} \begin{pmatrix} h_{11} \\ h_{12} \\ h_{13} \\ \vdots \\ h_{44} \end{pmatrix} = 0.$$

Soustava má 4 rovnice a 17 (16 koeficientů matice  $\mathbf{H}$  a  $\mu$ ) neznámých pro jeden bod. Pro  $n$  bodů bude mít  $4n$  rovnic a  $16+n$  neznámých; nejmenší možný počet bodů k potřebný k nalezení řešení je 5 (20 rovnic o 21 neznámých pro jednodimenzionální prostor řešení). Soustavu pro větší počet bodů můžeme řešit např. pomocí SVD, kdy upravíme matici soustavy tak, aby dimenze jejího nulového prostoru byla rovna jedné.

Tímto postupem jsme vypočetli matici transformace bodů skutečné scény na body rekonstruované scény. Opačnou transformaci vyjádříme takto:

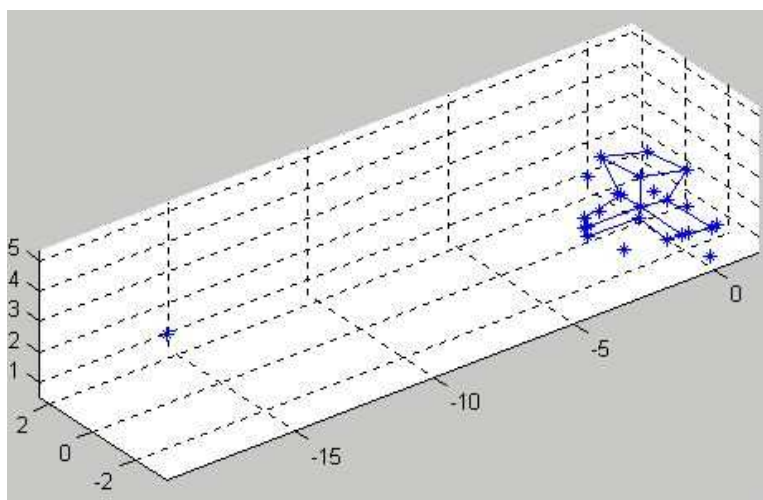
$$\hat{X} = \mu \mathbf{H}^{-1} X.$$



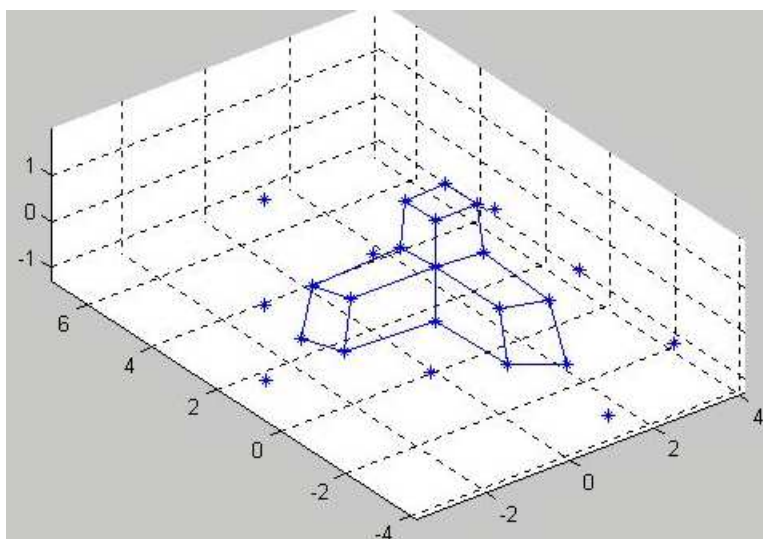
Obr. 5. Výsledek rekonstrukce (faktorizace)

## 4 Závěr

Vyzkoušeli jsme dva způsoby výpočtu projekčních matic z daných bodových korespondencí (přibližně 20 bodů). Faktorizaci jsme prováděli pro tři kamery. Při druhém způsobu výpočtu (volbě

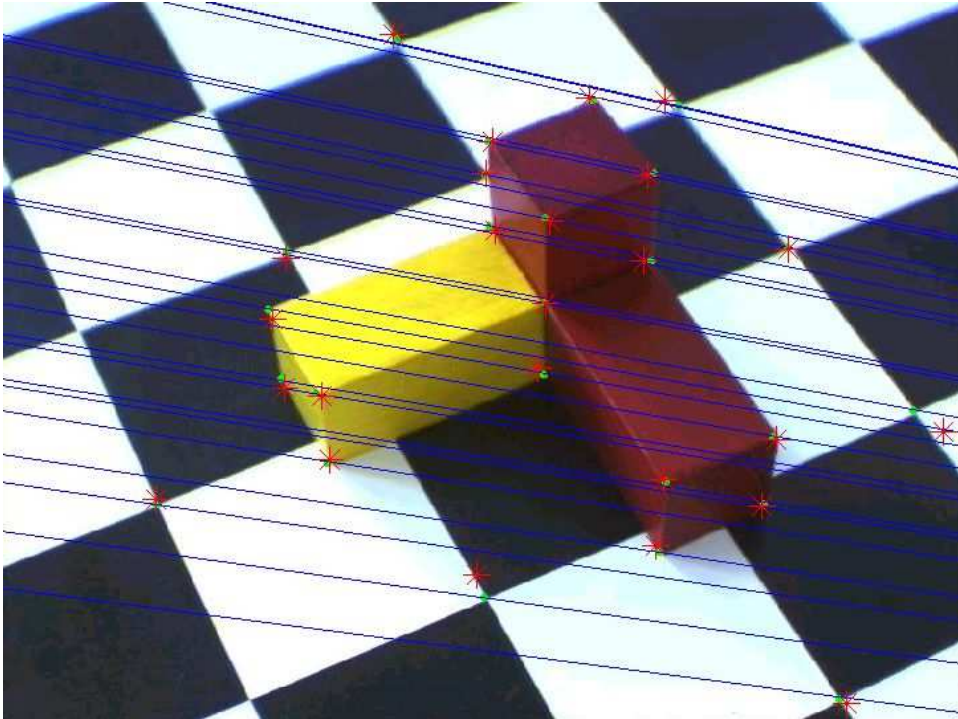


Obr. 6. Výsledek rekonstrukce (výpočet  $\mathbf{P}'$ ) bez normalizace

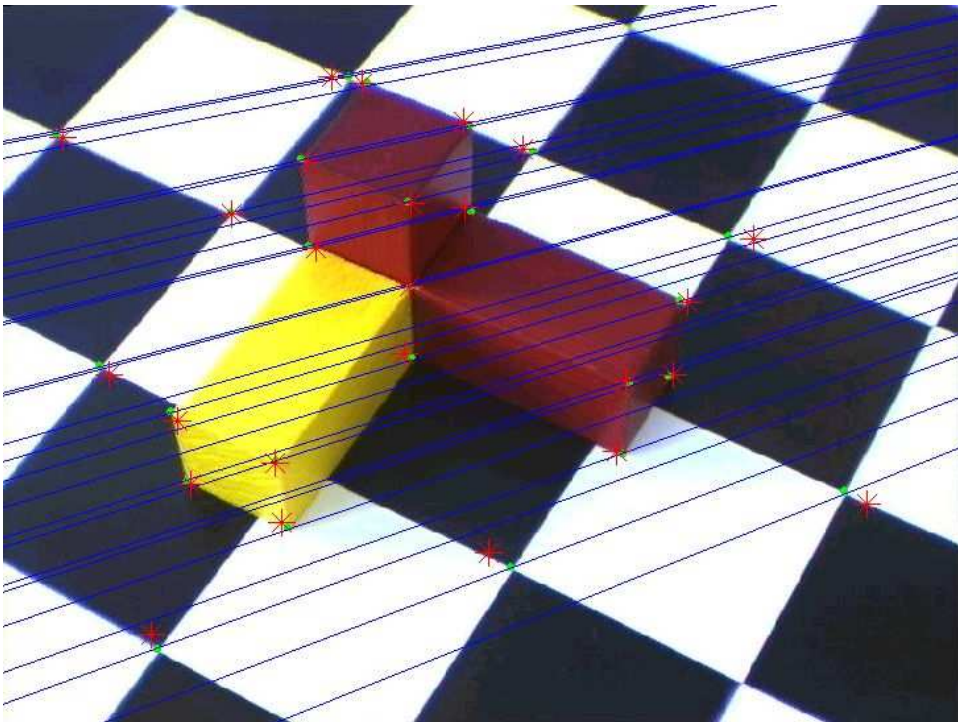


Obr. 7. Výsledek rekonstrukce (výpočet  $\mathbf{P}'$ ) s normalizací

jedné a dopočtení druhé projekční matice) jsme hledali takové řešení, které by bylo co nejvíce “regulární” - s největším podílem prvního a posledního prvku na diagonále matice  $\mathbf{A}$ . V praxi jsme to provedli tak, že jsme náhodně generovali určitý počet matic a z tohoto souboru jsme vybrali nejlepší řešení. Dále jsme zkoušeli vliv normalizace souřadnic vstupních bodů (obrazů) na kvalitu výsledné rekonstrukce. U faktorizace jsme nezaznamenali podstatný rozdíl mezi rekonstrukcemi z normalizovaných a nenormalizovaných bodů. Jiná byla situace u druhé použité metody, kde jsme zaznamenali po zavedení normalizace značné zlepšení.



Obr. 8. Reprojekce rekonstruovaných bodů (faktorizace, první obrázek)



Obr. 9. Reprojekce rekonstruovaných bodů (faktorizace, druhý obrázek)