

PRIMER UPORABE PROJEKTIVNE GEOMETRIJE V RAČUNALNIŠKEM VIDU: SAMODEJNA IZDELAVA 3D MODELOV

dodatek k seminarski nalogi
**PROJEKTIVNA GEOMETRIJA:
UPORABA V RAČUNALNIŠTVU**

seminarska naloga pri predmetu
DIFERENCIALNA IN RAČUNSKA GEOMETRIJA

Avtor: Peter Peer

Mentor: doc. dr. Neža Mramor Kosta

V Ljubljani, 13.9.2000

Kazalo

1	Uvod	4
2	Formacija slike	4
2.1	Vrednost osnovne matrike perspektivne projekcije	4
2.2	Notranji parametri kamere	8
2.3	Zunanji parametri kamere	12
2.4	Povzetek	13
3	Stereo vid	13
3.1	Epipolarna omejitev	15
3.2	Računanje epipolarne geometrije	16
3.3	Delo v normaliziranem koordinatnem sistemu	17
3.4	Določanje odmika kamere na podlagi korespondence točk . .	18
3.4.1	Iskanje korespondenčnih točk	20
3.4.2	Algoritem osmih točk	21
3.5	Rekonstrukcija 3D točk	22
	Dodatki	24
A	3D rotacija	24
A.1	Splošne lastnosti rotacijske matrike	24
A.2	Rotacija okoli koordinatnih osi	25
B	Računski primer vpliva zunanjih parametrov kamere na 3D točko M	25
C	Skica računskega primera rekonstrukcije 3D točke	26
	Literatura	29

Slike

1	Nastanek slike v točkasti kameri.	4
2	Model točkaste kamere.	5
3	Goriščna ravnina \mathcal{F} je vzporedna s slikovno ravnino \mathcal{R} in je od nje oddaljena za goriščno razdaljo (ang. <i>focal length</i>) f . . .	5
4	Vrstični vektorji matrice $\tilde{\mathbf{P}}$ podajajo goriščno ravnino in premico, definirano z optičnim centrom in točko o na slikovni ravnini.	7
5	Spreminjanje koordinatnega sistema slikovne ravnine. Star koordinatni sistem ima izhodišče v točki c , ki jo dobimo kot presečišče optične osi s slikovno ravnino in ima na obeh oseh enake enote. Sedaj gremo iz starega koordinatnega sistema v novi koordinatni sistem, ki ima izhodišče v c_n in lahko ima različne enote na oseh (zaradi lastnosti kamere, s katero zajemamo podatke).	9
6	Relacija med realno (pravo) in normalizirano slikovno ravnino.	11
7	Spreminjanje koordinatnega sistema 3D prostora.	12
8	Problem 3D računalniškega vida.	14
9	Epipolarna geometrija.	15
10	Dve kameri in njuni normalizirani slikovni ravnini.	17
11	Omejitev parametrov gibanja, ki jih dobimo iz korespondenčnih točk.	19
12	Princip korelacijske metode.	20

1 Uvod

V seminarski nalogi je predstavljen postopek 3D rekonstrukcije z vidika projektivne geometrije. Najprej je izpeljan postopek formacije slike v kameri, kjer izpostavimo tudi pomembne relacije za postopek rekonstrukcije. Ker delamo rekonstrukcijo na podlagi dveh slik, je v naslednjem poglavju opisan stereo sistem. Tukaj izpeljemo epipolarno omejitev, preslikavo v normaliziran koordinatni sistem, odmik kamere na podlagi korespondence točk in samo rekonstrukcijo 3D točke.

2 Formacija slike

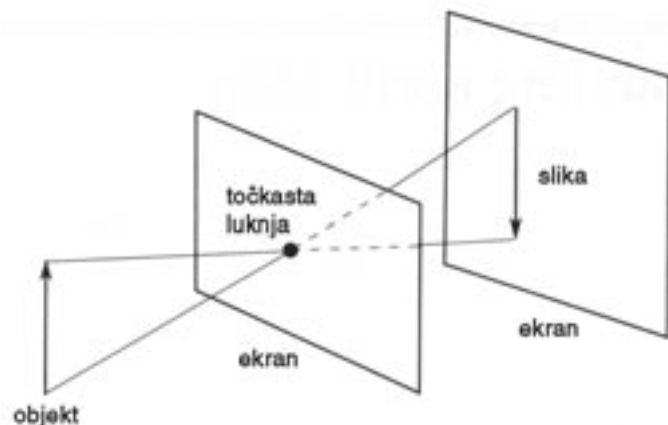
Označimo s $\tilde{\mathbf{P}}$ matriko reda 3×4 , ki predstavlja projektivno transformacijo iz 3D svetovnega koordinatnega sveta v 2D koordinatni sistem slike. Oglejmo si postopek izpeljave matrike $\tilde{\mathbf{P}}$, ki jo imenujemo tudi matrika perspektivne projekcije ter njeno uporabo pri formaciji slike:

Na vrednost $\tilde{\mathbf{P}}$ vplivajo trije dejavniki:

- vrednost osnovne matrike perspektivne projekcije
- notranji parametri kamere in
- zunanji parametri kamere.

Notranje in zunanje parametre kamere dobimo ob kalibraciji sistema.

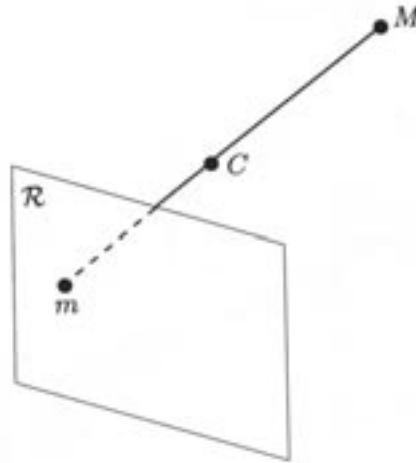
2.1 Vrednost osnovne matrike perspektivne projekcije



Slika 1: Nastanek slike v točkasti kameri.

Oglejmo si sistem s slike 1. Sistem sestavljata dva ekrana, prvi pa vsebuje točkasto luknjo v središču. Skozi to točko prehajajo svetlobni žarki, ki jih

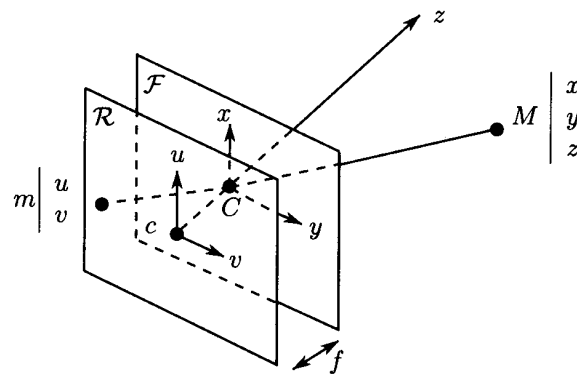
oddaja objekt ali pa so od njega odbiti. Tako na drugem ekranu dobimo invertirano sliko tega objekta.



Slika 2: Model točkaste kamere.

Na podlagi tega sistema lahko zgradimo geometrijski model točkaste kamere (ang. *pinhole camera*) (slika 2). Model je sestavljen iz slikovne ravnine \mathcal{R} in operacije perspektivne projekcije: 3D točka C , ki predstavlja optični center kamere in je od slikovne ravnine \mathcal{R} oddaljena za razdaljo f (goriščna razdalja optičnega sistema), je uporabljena pri preslikavi 3D točke M v 2D točko m tako, da poiščemo presečišče premice, ki je definirana s točkama M in C , s slikovno ravnino \mathcal{R} .

Na tem mestu pa definirajmo še en pojem za lažje razumevanje nadaljnje snovi: Optična os je premica, ki gre skozi optični center in je pravokotna na slikovno ravnino \mathcal{R} .



Slika 3: Goriščna ravnina \mathcal{F} je vzporedna s slikovno ravnino \mathcal{R} in je od nje oddaljena za goriščno razdaljo (ang. *focal length*) f .

Iz slike 3 je razvidno, da relacija med koordinatnim sistemom slike in 3D prostorom (je očitna, ampak) ni linearna¹:

$$-\frac{f}{z} = \frac{u}{x} = \frac{v}{y}. \quad (1)$$

Če uporabimo projektivne oziroma homogene² koordinate, relacija postane linearna:

$$\begin{bmatrix} U \\ V \\ S \end{bmatrix} = \begin{bmatrix} -f & 0 & 0 & 0 \\ 0 & -f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \quad (2)$$

in velja:

$$u = \frac{U}{S}, v = \frac{V}{S}, S \neq 0. \quad (3)$$

Kaj pa se zgodi, če je $z = 0$? Točka M , ki leži na goriščni ravnini \mathcal{F} , nima svoje preslikave na slikovni ravnini \mathcal{R} zato, ker je premica, ki je definirana s točkama M in C , vzporedna s slikovno ravnino \mathcal{R} , torej je ne seka! Če govorimo projektivno, jo seka v neskončnosti; $S=0$ je enačba premice v neskončnosti slikovne ravnine \mathcal{R} .

Enačba (2) je zapisana s homogenimi koordinatami (U, V, S) točke m , torej je definirana do multiplikativnega faktorja natančno. Je torej projektivna, zato jo lahko zapišemo z uporabo homogenih koordinat (X, Y, Z, T) točke M :

$$\begin{bmatrix} U \\ V \\ S \end{bmatrix} = \begin{bmatrix} -f & 0 & 0 & 0 \\ 0 & -f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ T \end{bmatrix},$$

kar lahko v matrični obliki zapišemo kot:

$$\tilde{\mathbf{m}} = \tilde{\mathbf{P}}\tilde{\mathbf{M}}, \quad (4)$$

kjer je $\tilde{\mathbf{m}} = [U, V, S]^T$, $\tilde{\mathbf{M}} = [X, Y, Z, T]^T$, $\tilde{\mathbf{P}}$ pa je osnovna matrika perspektivne projekcije.

Na kamero lahko gledamo kot na sistem, ki izvede linearno projektivno transformacijo iz projektivnega prostora \mathcal{P}^3 na projektivno ravnino \mathcal{P}^2 . To je le eden izmed mnogih primerov, kjer se s pomočjo projektivne geometrije stvari poenostavijo: namesto, da se ukvarjamo z nelinearnimi enačbami (kot

¹Iz slike 3 je razvidno tudi, da je točka $m = (u, v)$ na ravnini slike tudi točka v prostoru, katere koordinate so $m_{3D} = (u, v, -f)$.

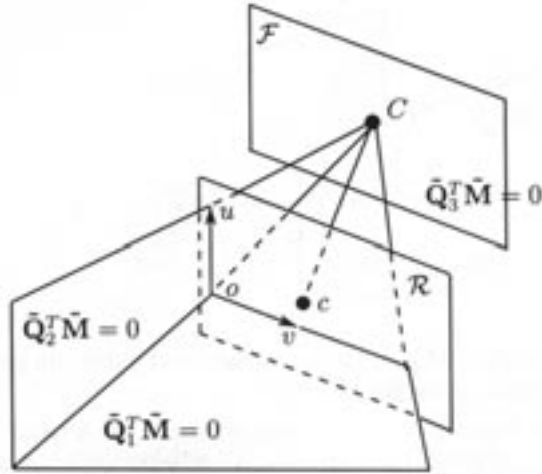
²Za projektivne oziroma homogene koordinate velja, da skaliranje ne vpliva na njih: $(X, Y, Z) = (\lambda X, \lambda Y, \lambda Z)$.

v primeru enačbe (1)), lahko uporabimo linearno relacijo (kot v primeru enačbe (2)) in izkoristimo moč linearne algebre.

Zapišimo geometrijsko interpretacijo vrstičnih vektorjev matrike $\tilde{\mathbf{P}}$:

$$\tilde{\mathbf{P}} = \begin{bmatrix} \tilde{\mathbf{Q}}_1^T \\ \tilde{\mathbf{Q}}_2^T \\ \tilde{\mathbf{Q}}_3^T \end{bmatrix} = \begin{bmatrix} \mathbf{q}_1^T & q_{14} \\ \mathbf{q}_2^T & q_{24} \\ \mathbf{q}_3^T & q_{34} \end{bmatrix}, \quad (5)$$

kjer je $\tilde{\mathbf{Q}}_i$, $i = 1, 2, 3$, vektor velikosti 4×1 in $\tilde{\mathbf{Q}}_i^T = [\mathbf{q}_i^T \quad q_{i4}]$. Vsak vektor predstavlja projekтивно ravnino z enačbo $\tilde{\mathbf{Q}}_i^T \tilde{\mathbf{M}} = 0$. Glede na enačbo (2), ravnina z enačbo $\tilde{\mathbf{Q}}_3^T \tilde{\mathbf{M}} = 0$ opisuje točke na slikovni ravnini \mathcal{R} , za katere velja $S = 0$, torej točke v neskončnosti. Opisuje torej goriščno ravnino \mathcal{F} . Ravnina z enačbo $\tilde{\mathbf{Q}}_1^T \tilde{\mathbf{M}} = 0$ in ravnina z enačbo $\tilde{\mathbf{Q}}_2^T \tilde{\mathbf{M}} = 0$ opisujeta točke na slikovni ravnini \mathcal{R} , za katere velja $U = 0$ in $V = 0$. Presek teh dveh ravnin je premica, ki gre skozi optični center kamere C in točko o na slikovni ravnini \mathcal{R} (slika 4). Pri tem je potrebno poudariti, da v splošnem ta premica ni optična os kamere.



Slika 4: Vrstični vektorji matrike $\tilde{\mathbf{P}}$ podajajo goriščno ravnino in premico, definirano z optičnim centrom in točko o na slikovni ravnini.

Na tem mestu pa podajmo še dve koristni informaciji, ki ju lahko izračunamo s pomočjo matrike perspektivne projekcije $\tilde{\mathbf{P}}$ in sta pomembni za nadaljno razlago:

- Iz slike 4 je razvidno, da je optični center kamere C definiran kot presek treh ravnin z enačbami $\tilde{\mathbf{Q}}_i^T \tilde{\mathbf{M}} = 0$, $i = 1, 2, 3$. C lahko torej izračunamo tako, da rešimo sistem treh linearnih enačb:

$$\tilde{\mathbf{P}} \begin{bmatrix} \mathbf{C} \\ 1 \end{bmatrix} = \mathbf{0}. \quad (6)$$

Če matriko $\tilde{\mathbf{P}}$, ki je reda 3×4 zapišemo kot $[\mathbf{P} \tilde{\mathbf{p}}]$, pri čemer je \mathbf{P} matrika reda 3×3 , $\tilde{\mathbf{p}}$ pa vektor reda 3×1 in predpostavljamo, da je rang matrike \mathbf{P} enak tri, potem lahko optični center kamere izračunamo kot:

$$\mathbf{C} = -\mathbf{P}^{-1}\tilde{\mathbf{p}}, \quad (7)$$

pri čemer je po enačbi (5) vektor $\tilde{\mathbf{p}} = [q_{14}, q_{24}, q_{34}]^T$.

Zaskrbljenost, ker računamo evklidske koordinate točke iz matrike $\tilde{\mathbf{P}}$, ki je definirana zgolj do skalarnega faktorja natančno, je odveč, saj če $\tilde{\mathbf{P}}$ nadomestimo z $\lambda\tilde{\mathbf{P}}$, $\tilde{\mathbf{p}}$ z $\lambda\tilde{\mathbf{p}}$ in \mathbf{P}^{-1} z $1/\lambda\mathbf{P}^{-1}$, vidimo, da se rezultat točke C ne spremeni.

- Formalno v \mathcal{P}^2 (projektivni ravnini) ni razlike med točkami in premicami – točke in premice predstavimo na enak način. Ta lastnost je znana pod imenom *Princip dualnosti*: Na točko, predstavljeno s koordinatnim vektorjem \mathbf{x} , lahko gledamo kot na množico premic, ki gredo skozi njo. Te premice predstavimo s koordinatnim vektorjem \mathbf{u} , za katerega velja $\mathbf{u}^T \mathbf{x} = 0$. Inverzno pa velja, da lahko na premico \mathbf{u} gledamo kot na množico točk \mathbf{x} , za katere velja ista enačba.

Na podlagi tega velja naslednji izrek, ki ga bomo potrebovali v nadaljevanju:

IZREK 1³

Premico, ki gre skozi točki m_1 in m_2 , zapišemo kot vektorski produkt $m_1 \times m_2$.

Drugo uporabno informacijo dobimo iz enačbe 3D premice, ki je določena s slikovnim elementom m (v prostoru) in optičnim centrom C . To premico imenujemo optični žarek, ki ga definira točka m na slikovni ravnini. Po enačbi (7) poznamo že točko C , ki leži na tej premici, druga točka, ki leži na tej premici pa je točka v neskončnosti $\tilde{\mathbf{D}}$, katere projektivne koordinate so $[\mathbf{D}^T, 0]^T$.⁴ Po enačbi (4) za \mathbf{D} velja:

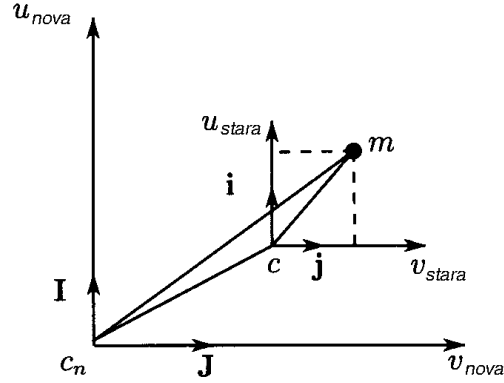
$$\tilde{\mathbf{m}} = \tilde{\mathbf{P}}\tilde{\mathbf{D}} = \mathbf{P}\mathbf{D} \quad \Rightarrow \quad \mathbf{D} = \mathbf{P}^{-1}\tilde{\mathbf{m}}. \quad (8)$$

2.2 Notranji parametri kamere

Če spremenimo izhodišče koordinatnega sistema slike in uvedemo različne enote za slikovne osi k_u in k_v , dobimo tri notranje parametre kamere: koordinati starega izhodišča u_0 in v_0 ter kvocient k_u/k_v .

³Dokaz si lahko ogledate v [2], na strani 15.

⁴ \mathbf{D} je točka preseka omenjenega optičnega žarka in ravnine v neskončnosti [2] (stran 25); po principu dualnosti v \mathcal{P}^3 (projektivnem prostoru) točke in ravnine predstavimo na enak način [2] (stran 23).



Slika 5: Spreminjanje koordinatnega sistema slikovne ravnine. Star koordinatni sistem ima izhodišče v točki c , ki jo dobimo kot presečišče optične osi s slikovno ravnino in ima na obeh oseh enake enote. Sedaj gremo iz starega koordinatnega sistema v novi koordinatni sistem, ki ima izhodišče v c_n in lahko ima različne enote na oseh (zaradi lastnosti kamere, s katero zajemamo podatke).

Iz slike 5 je razvidno, da za slikovni element m velja:

$$\overline{c_n m} = \overline{c_n c} + \overline{c m} . \quad (9)$$

$\overline{c m}$ lahko sedaj v starem koordinatnem sistemu zapišemo kot $\overline{c m} = u_{stara} \mathbf{i} + v_{stara} \mathbf{j}$, zaradi spremembe velikosti enot na novem koordinatnem sistemu pa velja $\mathbf{i} = s \mathbf{I}$ in $\mathbf{j} = s \mathbf{J}$, kjer je

$$s = \begin{bmatrix} k_u & 0 \\ 0 & k_v \end{bmatrix} .$$

Če $\overline{c_n c}$ označimo kot vektor translacije \mathbf{t} , lahko enačbo (9) zapišemo kot:

$$\begin{aligned} \overline{c_n m} &= \overline{c_n c} + \overline{c m} = \overline{c_n c} + u_{stara} \mathbf{i} + v_{stara} \mathbf{j} = \\ &= \overline{c_n c} + u_{stara} s \mathbf{I} + v_{stara} s \mathbf{J} = \mathbf{t} + u_{stara} s \mathbf{I} + v_{stara} s \mathbf{J}, \end{aligned}$$

kar je v projektivnih koordinatah enako:

$$\tilde{\mathbf{m}}_{nova} = \tilde{\mathbf{H}} \tilde{\mathbf{m}}_{stara} , \quad (10)$$

kjer je $\tilde{\mathbf{H}}$ matrika reda 3×3 :

$$\tilde{\mathbf{H}} = \begin{bmatrix} \mathbf{s} & \mathbf{t} \\ \mathbf{0}_2^T & 1 \end{bmatrix} .$$

Ker po enačbi (4) velja:

$$\tilde{\mathbf{m}}_{stara} = \tilde{\mathbf{P}}_{stara} \tilde{\mathbf{M}},$$

sledi, da je

$$\tilde{\mathbf{P}}_{nova} = \tilde{\mathbf{H}} \tilde{\mathbf{P}}_{stara}. \quad (11)$$

Če označimo koordinate \mathbf{t} z (u_0, v_0) , potem velja:

$$\tilde{\mathbf{P}}_{nova} = \begin{bmatrix} -fk_u & 0 & u_0 & 0 \\ 0 & -fk_v & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}. \quad (12)$$

Za konec pa, zaradi nadaljnjih izpeljav, definirajmo še normaliziran koordinatni sistem kamere, ki nam omogoča normalizacijo koordinat slikovne ravnine. Ta koordinatni sistem nam pride prav pri stereo aplikacijah (poglavje 3), saj nam omogoča, da lahko ignoriramo specifične lastnosti kamer in gledamo na njo, kot da je idealna. Glede na enačbo (12), spremenimo slikovni koordinatni sistem tako, da lahko $\tilde{\mathbf{P}}$ zapišemo kot:

$$\tilde{\mathbf{P}}_{nova} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}. \quad (13)$$

Glede na enačbi (11) in (12) ima $\tilde{\mathbf{H}}$ sedaj naslednjo obliko:

$$\tilde{\mathbf{H}} = \begin{bmatrix} \frac{1}{-fk_u} & 0 & \frac{u_0}{fk_u} \\ 0 & \frac{1}{-fk_v} & \frac{v_0}{fk_v} \\ 0 & 0 & 1 \end{bmatrix}.$$

Po enačbi (10) tako dobimo novi slikovni koordinati:

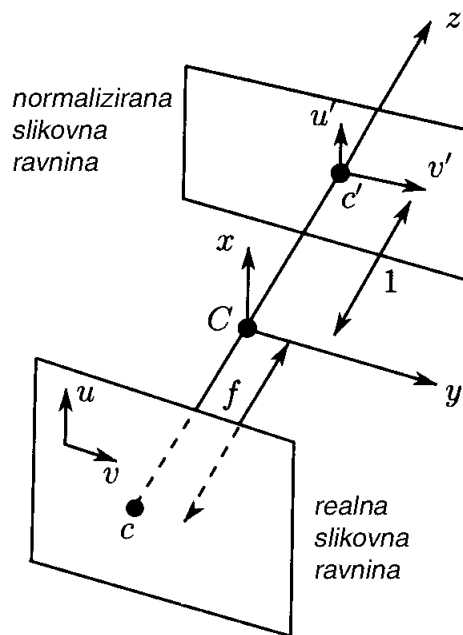
$$\begin{aligned} u' &= \frac{u-u_0}{-fk_u} \\ v' &= \frac{v-v_0}{-fk_v}. \end{aligned} \quad (14)$$

Iz slike 6 je razvidno, da ravnina, ki je vzporedna s slikovno ravnino in je za enoto oddaljena od optičnega centra kamere, skupaj z optičnim centrom definira t.i. normalizirano kamero.

Vidimo lahko, da je normalizirana slikovna ravnina na drugi strani optičnega centra C , torej je tukaj slika, v primerjavi z originalno, invertirana. Po enačbi (4) oziroma enačbi (2), v kateri uporabimo enačbo (12), velja:

$$\begin{bmatrix} U \\ V \\ S \end{bmatrix} = \begin{bmatrix} -fk_u & 0 & u_0 & 0 \\ 0 & -fk_v & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$$

in po enačbi (3) velja:



Slika 6: Relacija med realno (pravo) in normalizirano slikovno ravnino.

$$\begin{aligned} u &= \frac{U}{S} = -fk_u \frac{x}{z} + u_0 \\ v &= \frac{V}{S} = -fk_v \frac{y}{z} + v_0 \end{aligned} .$$

Za normalizirano situacijo pa po enačbi (4) oziroma enačbi (2), v kateri uporabimo enačbo (13), velja:

$$\begin{bmatrix} U \\ V \\ S \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$$

in po enačbi (3) velja:

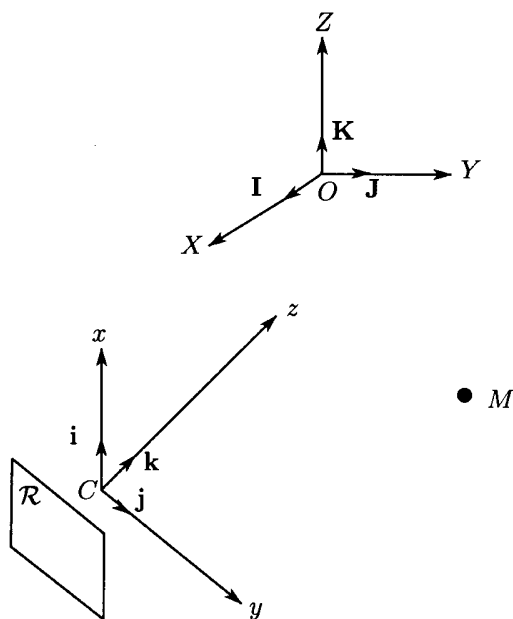
$$\begin{aligned} u &= \frac{U}{S} = \frac{x}{z} \\ v &= \frac{V}{S} = \frac{y}{z} \end{aligned} .$$

Za k_u in k_v po enačbi (14) in zgornjih enačbah tako velja:

$$\begin{aligned} k_u &= \frac{u-u_0}{-f\frac{x}{z}} \\ k_v &= \frac{v-v_0}{-f\frac{y}{z}} \end{aligned} .$$

2.3 Zunanji parametri kamere

S spremembo položaja in orientacije kamere v prostoru dobimo šest zunanjih parametrov kamere, tri za rotacijo \mathbf{R} (matrika velikosti 3×3) (dodatek A) in tri za translacijo \mathbf{T} (vektor velikosti 3×1).



Slika 7: Spreminjanje koordinatnega sistema 3D prostora.

Kot je razvidno iz slike 7, gremo iz starega koordinatnega sistema z izhodiščem v optičnem centru C v nov koordinatni sistem z izhodiščem v točki O z rotacijo \mathbf{R} in translacijo $\mathbf{T} = \overrightarrow{CO}$. Glede na notacijo slike in po podobnem principu kot v primeru notranjih parametrov, velja:

$$\overrightarrow{CM} = \overrightarrow{CO} + \overrightarrow{OM}.$$

Če izrazimo \overrightarrow{OM} v novem koordinatnem sistemu, dobimo:

$$\overrightarrow{OM} = x_{nova}\mathbf{I} + y_{nova}\mathbf{J} + z_{nova}\mathbf{K},$$

ob upoštevanju rotacije \mathbf{R} iz starega koordinatnega sistema v novi koordinatni sistem, pa velja:

$$\mathbf{I} = \mathbf{R}\mathbf{i}, \quad \mathbf{J} = \mathbf{R}\mathbf{j}, \quad \mathbf{K} = \mathbf{R}\mathbf{k}.$$

Če sedaj zapišemo vektor \overrightarrow{CO} v starem koordinatnem sistemu kot $\mathbf{T} = [t_x, t_y, t_z]^T$, velja:

$$\begin{aligned}
\overline{\mathbf{CM}} &= \overline{\mathbf{CO}} + \overline{\mathbf{OM}} = \mathbf{T} + \overline{\mathbf{OM}} = \mathbf{T} + x_{nova}\mathbf{I} + y_{nova}\mathbf{J} + z_{nova}\mathbf{K} = \\
&= \mathbf{T} + x_{nova}\mathbf{Ri} + y_{nova}\mathbf{Rj} + z_{nova}\mathbf{Rk} \\
&= \mathbf{T} + \mathbf{R} \begin{bmatrix} x_{nova} \\ y_{nova} \\ z_{nova} \end{bmatrix}.
\end{aligned}$$

Na podlagi zgornje enačbe v homogenih koordinatah dobimo:

$$\tilde{\mathbf{M}}_{stara} = \tilde{\mathbf{K}}\tilde{\mathbf{M}}_{nova}, \quad (15)$$

kjer je $\tilde{\mathbf{K}}^5$ matrika reda 4×4 :

$$\tilde{\mathbf{K}} = \begin{bmatrix} \mathbf{R} & \mathbf{T} \\ \mathbf{0}_3^T & 1 \end{bmatrix}.$$

Računski primer vpliva zunanjih parametrov kamere na 3D točko je podan v dodatku B.

Po enačbi (4) in enačbi (15) velja:

$$\tilde{\mathbf{m}} = \tilde{\mathbf{P}}_{stara}\tilde{\mathbf{M}}_{stara} = \tilde{\mathbf{P}}_{stara}\tilde{\mathbf{K}}\tilde{\mathbf{M}}_{nova},$$

torej velja, da je

$$\tilde{\mathbf{P}}_{nova} = \tilde{\mathbf{P}}_{stara}\tilde{\mathbf{K}}.$$

2.4 Povzetek

Zapišimo enačbo, ki pove, kako se spreminja matrika perspektivne projekcije $\tilde{\mathbf{P}}$, ko prestavimo koordinatni sistem slikovne ravnine in/ali 3D prostora:

$$\tilde{\mathbf{P}}_{nova} = \tilde{\mathbf{H}}\tilde{\mathbf{P}}_{stara}\tilde{\mathbf{K}}.$$

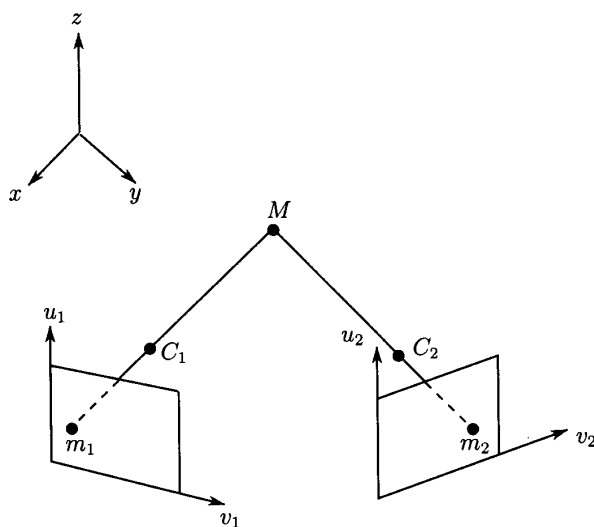
3 Stereo vid

Stereo vid predstavlja zelo pomemben praktičen problem računalniškega vida: rekonstrukcija 3D koordinat točk na sceni iz dveh (ali več) slik, zajetih s kamerama, za kateri poznamo relativno lokacijo in orientacijo v prostoru⁶ (kalibrirani kameri [2] (stran 51)).

Osnovni problem stereo vida je prikazan na sliki 8. Točkasti kameri tvorita sliki \mathcal{R}_1 in \mathcal{R}_2 (slikovni ravnini), v katerih se fizična točka M preslika v točki m_1 in m_2 . Opravitni imamo s tremi koordinatnimi sistemi: dvema

⁵ $\tilde{\mathbf{K}}$ je matrika, katere determinanta je različna od nič in definira linearno transformacijo (kolineacijo) iz \mathcal{P}^3 v \mathcal{P}^3 . [2] (strani 10 in 26)

⁶Poznamo torej matriki $\tilde{\mathbf{P}}$, torej tudi notranje in zunanje parametre obeh kamer.



Slika 8: Problem 3D računalniškega vida.

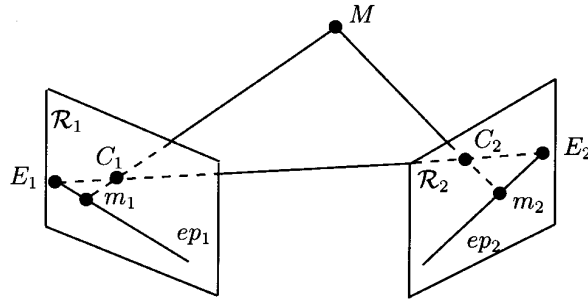
slikovnima koordinatnima sistemoma $((u_1, v_1)$ in $(u_2, v_2))$ in 3D koordinatnim sistemom $((x, y, z))$. Glede na formirani sliki \mathcal{R}_1 in \mathcal{R}_2 , želimo rešiti naslednja dva problema:

Problem korespondence - Za točko m_1 v slikovni ravnini \mathcal{R}_1 najdi korespondenčno točko m_2 v slikovni ravnini \mathcal{R}_2 . Korespondenca pomeni, da točki m_1 in m_2 predstavljata sliko iste fizične točke M .

Problem rekonstrukcije - Na podlagi dane točke m_1 in točke m_2 izračunaj 3D koordinato fizične točke M .

Problem korespondence je dvoumen: točka m_1 iz slikovne ravnine \mathcal{R}_1 je lahko v korespondenci s katerokoli točko m_2 iz slikovne ravnine \mathcal{R}_2 . Zato uporabimo določene omejitve, s katerimi zmanjšamo število potencialnih ujemanj za dano točko m_1 . Za nadaljne razumevanje snovi seminarske naloge si bomo v nadaljevanju ogledali Epipolarno omejitev.

Rešitev problema rekonstrukcije pa lahko po sliki 8 zapišemo kot presečišče premic, ki sta definirani s točkama C_1 in m_1 ter C_2 in m_2 .⁷ Seveda je rezultat močno odvisen od natančnosti kalibracijskega postopka sistema kamer, torej kako natančno poznamo lokacijo 3D točk C_1 in C_2 (optičnih centrov) in slikovnih ravnin \mathcal{R}_1 in \mathcal{R}_2 v 3D koordinatnem prostoru. O rekonstrukciji 3D točk bomo več spregovorili v poglavju 3.5.



Slika 9: Epipolarna geometrija.

3.1 Epipolarna omejitev

Iz slike 9 je razvidno, da vse možne 3D fizične točke M , ki ustvarijo dano slikovno točko m_1 , ležijo na neskončnem poltraku, ki je definiran s točkama m_1 in C_1 . Posledica tega je, da vse možne korespondenčne točke m_2 iz slikovne ravnine \mathcal{R}_2 točki m_1 ležijo na sliki tam, kamor se ta poltrak preslika. Na sliki \mathcal{R}_2 se ta preslikan poltrak imenuje epipolarna premica ep_2 točke m_1 , ki gre skozi točko E_2 . E_2 se imenuje epipol in predstavlja presečišče premice, ki je definirana s točkama C_1 in C_2 ter slikovne ravnine \mathcal{R}_2 . Na točko E_2 slikovne ravnine \mathcal{R}_2 lahko gledamo tudi kot na preslikavo optičnega centra C_1 prve kamere na slikovno ravnino \mathcal{R}_2 .

Ustrezna omejitev tako pravi, da ob dani točki m_1 iz slikovne ravnine \mathcal{R}_1 , možne korespondenčne točke na slikovni ravnini \mathcal{R}_2 ležijo na premici. Tako smo zmanjšali preiskovalni prostor iz dveh dimenzij na eno. Epipolarna omejitev je seveda simetrična. Epipolarni premici ep_1 in ep_2 dobimo kot rezultat presečišča slikovnih ravnin \mathcal{R}_1 in \mathcal{R}_2 z ravnino, ki je definirana s točkama C_1 , C_2 in M . To ravnino imenujemo epipolarna ravnina točke M .

Velikokrat zaradi enostavnosti obravnave sistema predpostavljamo, da sta slikovni ravnini \mathcal{R}_1 in \mathcal{R}_2 vzporedni s premico, ki je definirana s točkama C_1 in C_2 . Na ta način greta oba epipola v neskončnost, epipolarni premici ep_1 in ep_2 pa sta vzporedni. Splošen primer izračuna epipolarne premice je podan v naslednjem poglavju, tukaj pa si oglejmo, kako je le-ta določena v tem posebnem primeru: Kot rečeno sta presečišči epipolarne ravnine točke M in slikovnih ravnin \mathcal{R}_1 in \mathcal{R}_2 epipolarni premici ep_1 in ep_2 , ki sta vzporedni. Torej je epipolarna premica v tem primeru določena že zgolj z eno točko (enačba (16)) in omenjenim dejstvom vzporednosti s premico, ki je definirana s točkama C_1 in C_2 . Pri delu v normaliziranem koordinatnem sistemu (poglavje 3.3) pa vemo, da je translacija dovoljena le po osi x in osi y , ne pa tudi po osi z (sicer ne velja več pravilo vzporednosti s premico, ki

⁷Kako vemo, da ti dve premici nista mimobežni? Dokaz se skriva v epipolarni omejitvi, predstavljeni v naslednjem poglavju; bodite pozorni na pojem epipolarna ravnina točke M . [2] (izrek na strani 24)

je definirana s točkama C_1 in C_2), rotacija pa je (iz istega vzroka) omejena le na os z .

3.2 Računanje epipolarne geometrije

Po enačbi (4) velja:

$$\tilde{\mathbf{m}}_1 = \tilde{\mathbf{P}}_1 \tilde{\mathbf{M}} \text{ in } \tilde{\mathbf{m}}_2 = \tilde{\mathbf{P}}_2 \tilde{\mathbf{M}} .$$

Kot sledi iz enačbe (6), dobimo koordinati obeh optičnih centrov C_i ($i = 1, 2$) v 3D prostoru tako, da rešimo naslednja sistema treh linearnih enačb:

$$\tilde{\mathbf{P}}_i \tilde{\mathbf{M}} = \mathbf{0} , \quad i = 1, 2 .$$

Epipole E_i izračunamo po enačbi:

$$\tilde{\mathbf{e}}_i = \tilde{\mathbf{P}}_i \begin{bmatrix} \mathbf{C}_j \\ 1 \end{bmatrix} , \quad i, j = 1, 2 , \quad i \neq j ,$$

saj je epipol E_i projekтивna preslikava optičnega centra kamere j na slikovno ravnino i . Z $\tilde{\mathbf{e}}_i$ v tem primeru označimo epipol E_i , zapisan v homogenih koordinatah projekтивne ravnine \mathcal{P}^2 .

Sedaj pa si oglejmo, kako lahko izračunamo epipolarno premico ep_2 , če poznamo slikovno točko m_1 . Premica je definirana z dvema točkama:

- Epipolom E_2 , za katerega po enačbi (4) in enačbi (7) velja:

$$\tilde{\mathbf{e}}_2 = \tilde{\mathbf{P}}_2 \tilde{\mathbf{C}}_1 = \tilde{\mathbf{P}}_2 \begin{bmatrix} -\mathbf{P}_1^{-1} \tilde{\mathbf{p}}_1 \\ 1 \end{bmatrix} .$$

- Drugo točko ($\tilde{\mathbf{m}}_2$) dobimo s pomočjo točke v neskončnosti $\tilde{\mathbf{D}}$ optičnega žarka (ta je, kot rečeno, določen s slikovnim elementom m_1 (v prostoru) in optičnim centrom C_1), torej točko, ki jo dobimo s presekom optičnega žarka z ravnino v neskončnosti (stran 8). Glede na enačbi (4) in (8), dobimo točko m_2 , ki je preslikava točke D_1 na drugo slikovno ravnino, po naslednji enačbi:

$$\tilde{\mathbf{m}}_2 = \tilde{\mathbf{P}}_2 \tilde{\mathbf{M}} = \tilde{\mathbf{P}}_2 \tilde{\mathbf{D}}_1 = \mathbf{P}_2 \mathbf{D}_1 = \mathbf{P}_2 \mathbf{P}_1^{-1} \tilde{\mathbf{m}}_1 . \quad (16)$$

Projektivna predstavitev ep_2 je sedaj, po izreku 1 (stran 8), enaka vektorskemu produktu $\tilde{\mathbf{e}}_2 \times \tilde{\mathbf{m}}_2$.

Če vektorju $\tilde{\mathbf{e}}_2$ priredimo matriko $\tilde{\mathbf{E}}_2$, za katero velja:

$$\tilde{\mathbf{E}}_2 \mathbf{y} = \tilde{\mathbf{e}}_2 \times \mathbf{y} , \quad \text{za vsak vektor } \mathbf{y} ,$$

potem lahko $\tilde{\mathbf{e}}_2 \times \tilde{\mathbf{m}}_2$ po enačbi (16) zapišemo kot:

$$\tilde{\mathbf{e}}_2 \times \tilde{\mathbf{m}}_2 = \tilde{\mathbf{E}}_2 \tilde{\mathbf{m}}_2 = \tilde{\mathbf{E}}_2 \mathbf{P}_2 \mathbf{P}_1^{-1} \tilde{\mathbf{m}}_1 = \mathbf{F} \tilde{\mathbf{m}}_1 . \quad (17)$$

Elementi matrike $\tilde{\mathbf{E}}_2$ so pri tem razvidni iz naslednje enačbe:

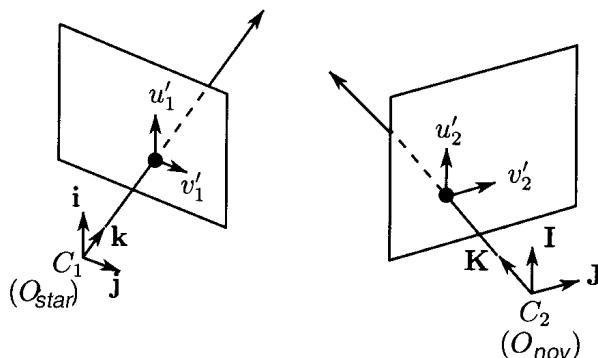
$$\tilde{\mathbf{e}}_2 = \begin{bmatrix} a \\ b \\ c \end{bmatrix} \Rightarrow \tilde{\mathbf{E}}_2 = \begin{bmatrix} 0 & -c & b \\ c & 0 & -a \\ -b & a & 0 \end{bmatrix} . \quad (18)$$

Vsak slikovni element m_2 , ki leži na epipolarni premici ep_2 , katero določa slikovni element m_1 , tako zadovoljuje enačbo⁸ (epipolarna omejitev):

$$\tilde{\mathbf{m}}_2^T \mathbf{F} \tilde{\mathbf{m}}_1 = 0 , \quad (19)$$

ki se imenuje Longuet-Higginsova enačba in jo bomo s pridom uporabili v nadaljevanju.

3.3 Delo v normaliziranem koordinatnem sistemu



Slika 10: Dve kameri in njuni normalizirani slikovni ravnini.

Če uporabimo normaliziran model kamere (poglavje 2.2) in predpostavimo relacije iz slike 10, se opis epipolarne geometrije poenostavi. Uporabimo notacijo, ki smo jo uporabili v poglavju 2.3: $\mathbf{T} = \mathbf{C}_1 \mathbf{C}_2$, $\mathbf{R}[\mathbf{i}, \mathbf{j}, \mathbf{k}] = [\mathbf{I}, \mathbf{J}, \mathbf{K}]$, matrika $\tilde{\mathbf{K}}$ pa je enaka:

$$\tilde{\mathbf{K}} = \begin{bmatrix} \mathbf{R} & \mathbf{T} \\ \mathbf{0}_3^T & 1 \end{bmatrix} .$$

$\tilde{\mathbf{P}}_1$ zapišemo kot:

⁸Natančnejša razlaga je podana v poglavju 3.4.

$$\tilde{\mathbf{P}}_1 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} .$$

Matrika $\tilde{\mathbf{P}}_2$ v novem koordinatnem sistemu C_2XYZ je enaka $\tilde{\mathbf{P}}_1$ ⁹. Matrika $\tilde{\mathbf{P}}_2$ izražena v koordinatnem sistemu C_1xyz pa je enaka:

$$\tilde{\mathbf{P}}_1 = \tilde{\mathbf{P}}_2 \tilde{\mathbf{K}} \Rightarrow \tilde{\mathbf{P}}_2 = \tilde{\mathbf{P}}_1 \tilde{\mathbf{K}}^{-1} = \begin{bmatrix} \mathbf{R}^T & -\mathbf{R}^T \mathbf{T} \end{bmatrix} .$$

Sedaj lahko izračunamo koordinati epipolov E_1 in E_2 :

$$\tilde{\mathbf{e}}_1 = \tilde{\mathbf{P}}_1 \tilde{\mathbf{C}}_2 = \tilde{\mathbf{P}}_1 \begin{bmatrix} \overrightarrow{\mathbf{C}_1 \mathbf{C}_2} \\ 1 \end{bmatrix} = \tilde{\mathbf{P}}_1 \begin{bmatrix} \mathbf{T} \\ 1 \end{bmatrix} = \mathbf{T} ,$$

$$\tilde{\mathbf{e}}_2 = \tilde{\mathbf{P}}_2 \tilde{\mathbf{C}}_1 = \tilde{\mathbf{P}}_2 \begin{bmatrix} \mathbf{0}_3 \\ 1 \end{bmatrix} = -\mathbf{R}^T \mathbf{T} .$$

Če slikovni element predstavimo z $\tilde{\mathbf{m}}_1 = [u'_1, v'_1, 1]^T$, je epipolarna premica ep_2 , glede na zgornje enačbe in enačbo (17), sedaj predstavljena z enačbo:

$$\tilde{\mathbf{e}}_2 \times \tilde{\mathbf{m}}_2 = \tilde{\mathbf{e}}_2 \times \mathbf{P}_2 \mathbf{P}_1^{-1} \tilde{\mathbf{m}}_1 = -\mathbf{R}^T \mathbf{T} \times \mathbf{R}^T \tilde{\mathbf{m}}_1 ,$$

kar je projektivno¹⁰ enako:

$$\mathbf{R}^T (\mathbf{T} \times \tilde{\mathbf{m}}_1) .$$

Podobno velja tudi za $\tilde{\mathbf{m}}_2 = [u'_2, v'_2, 1]^T$, kjer je epipolarna premica ep_1 predstavljena z enačbo:

$$\tilde{\mathbf{e}}_1 \times \tilde{\mathbf{m}}_1 = \tilde{\mathbf{e}}_1 \times \mathbf{P}_1 \mathbf{P}_2^{-1} \tilde{\mathbf{m}}_2 = \mathbf{T} \times \mathbf{R} \tilde{\mathbf{m}}_2 . \quad (20)$$

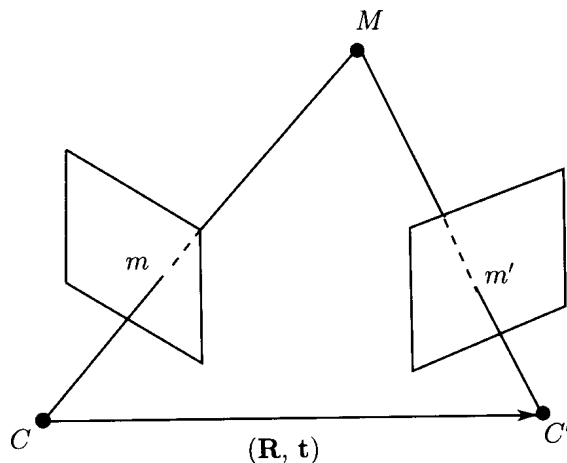
3.4 Določanje odmika kamere na podlagi korespondence točk

Predpostavimo, da imamo opravka z normaliziranim koordinatnim sistemom, kar torej pomeni, da moramo kalibrirati kamero in izračunati njene notranje parametre. Ti definirajo transformacijo slikovnih koordinat v normalizirane slikovne koordinate. S transformacijo dobimo normalno kamero z enotsko goriščno razdaljo f in koordinatni sistem $Cxyz$ z izhodiščem v optičnem centru kamere (slika 6). Na premik kamere lahko tako gledamo kot na premik tega koordinatnega sistema. Odmik D_i iz koordinatnega sistema ene kamere v koordinatni sistem druge kamere, kar je ekvivalentno premikanju kamere iz lokacije in orientacije ob času t_i v lokacijo in orientacijo

⁹Smo v normaliziranem koordinatnem sistemu; notranji parametri kamere ostajajo enaki, spremembe zunanjih parametrov pa tukaj ne upoštevamo.

¹⁰Uprabljamu homogene koordinate.

ob času t_{i+1} , opišemo kot rotacijo (\mathbf{R}) okoli osi, ki gre skozi C_i , kateri sledi transformacija $\mathbf{t} = \overrightarrow{C_i C_{i+1}}$.



Slika 11: Omejitev parametrov gibanja, ki jih dobimo iz korespondenčnih točk.

Kot je razvidno iz slike 11, se 3D točka M preslika v korespondenčni slikovni točki m in m' . Sedaj nas zanima, kakšna omejitev velja za \mathbf{R} in \mathbf{t} ? Iz slike je razvidno, da sta točki m in m' v korespondenci le v primeru, če so prostorski vektorji $\mathbf{m} = \overrightarrow{Cm}$, $\mathbf{m}' = \overrightarrow{C'm'}$ in $\mathbf{t} = \overrightarrow{CC'}$ koplanarni. Če želimo to omejitev zapisati v obliki enačbe, moramo te vektorje najprej prestaviti v skupni koordinatni sistem. Za vektor \mathbf{m}' , ki je izražen v koordinatnem sistemu druge kamere, velja, da je enak vektorju \mathbf{Rm}' v koordinatnem sistemu prve kamere (poglavje 2.3). V koordinatnem sistemu prve kamere lahko torej omejitev zapišemo kot¹¹:

$$\mathbf{m} \cdot (\mathbf{t} \times \mathbf{Rm}') = 0 .$$

Na podlagi zgornje enačbe in enačbe (20) lahko epipolarno omejitev v koordinatnem sistemu prve kamere zapišemo kot:

$$\tilde{\mathbf{m}}_1 \cdot (\mathbf{t} \times \mathbf{R}\tilde{\mathbf{m}}_2) = 0 .$$

Če vektorju \mathbf{t} priredimo matriko \mathbf{T} , za katero velja isto kot za matriko v enačbi (18), potem lahko omejitev zapišemo kot (enačba (19)):

$$\tilde{\mathbf{m}}_1^T \mathbf{T} \mathbf{R} \tilde{\mathbf{m}}_2 = \tilde{\mathbf{m}}_1^T \mathbf{E} \tilde{\mathbf{m}}_2 = 0 . \quad (21)$$

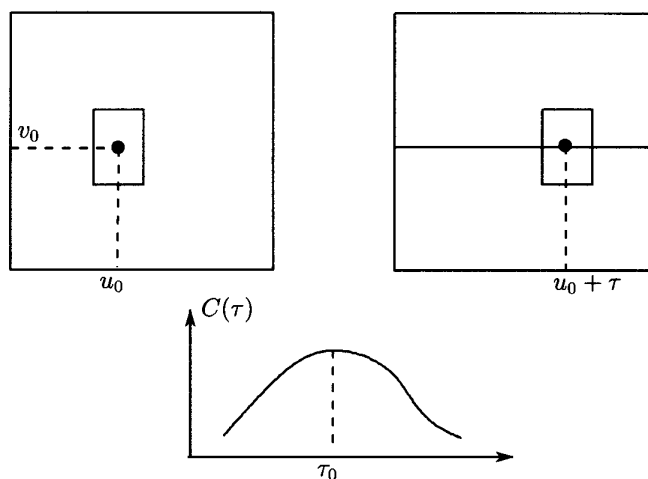
¹¹Enačbo dobimo na podlagi dveh pravil: rezultat vektorskega produkta je vektor, ki je pravokoten na ravnino, katero tvorita ta dva vektorja, skalarni produkt dveh pravokotnih vektorjev pa je enak nič.

Osnovna matrika \mathbf{E} (ang. *essential matrix*) predstavlja osnovo linearne metode za izračun odmika D_i , ki se imenuje Algoritem osmih točk (ang. *The eight-point algorithm*). Ta ob znanih korespondenčnih točkah na izhodu poda \mathbf{R} in \mathbf{t} ¹².

Preden pa si ogledamo to metodo v detajlih, pa na kratko naprej opišimo, kako poiščemo ustrezne korespondenčne točke.

3.4.1 Iskanje korespondenčnih točk

Veliko je raziskav korelacijske tehnike intenzitetnih področij slik za komercialne aplikacije stereofotogrametrije, hkrati pa le-te predstavljajo ene najstarejših metod računalniškega vida.



Slika 12: Princip korelacijske metode.

Princip korelacijske metode je razviden iz slike 12. Da bi našli koordinati slikovnega elementa na drugi sliki, ki ustrežata koordinati (u_0, v_0) slikovnega elementa na prvi sliki, računamo korelacijo $C_{12}(\tau)$ kvadratnega okna velikosti $(2P + 1) \times (2N + 1)$, katerega središče je (u_0, v_0) , z drugo intenzitetno sliko vzdolž vrstice ($v_2 = v_0$). Pri tem seveda upoštevamo izračunano epipolarno premico. Krivulja $C_{12}(\tau)$ ima navadno en maksimum, ki ga dosežemo v točki τ_0 . Dispariteto¹³ slikovnega elementa (u_0, v_0) tako izenačimo s τ_0 .

Ena izmed pomankljivosti te metode je predpostavka, da so epipolarne premice vrstice slike. Če vhodne slike ne ustrezajo tej zahtevi, moramo slike predhodno procesirati s postopkom izravnave (ang. *rectification*).

¹²V nadaljevanju predpostavljamo, da je $\mathbf{t} \neq \mathbf{0}$ ([2], stran 249).

¹³Definicija disparitete v primeru vzporednih kamer: ob danem slikovnem elementu m_1 s koordinatama (u_1, v_1) v prvi slikovni ravnini in korespondenčnim slikovnim elementom m_2 s koordinatama (u_2, v_2) v drugi slikovni ravnini, je dispariteta definirana kot razlika $v_2 - v_1$.

Ker korelacija in izravnava presegata namen te seminarske naloge, si lahko bralec več o teh metodah prebere v [2] (stran 188).

3.4.2 Algoritem osmih točk

Na podlagi dejstva, da je enačba (21) linearna in homogena glede na koeficiente matrike \mathbf{E} , sta Longuet in Higgins razvila metodo, s pomočjo katere lahko najdemo te koeficiente, če imamo na voljo osem takšnih enačb.

Metoda je sestavljena iz dveh korakov:

- ocena matrike \mathbf{E}
- izračun \mathbf{t} in \mathbf{R} iz \mathbf{E} .

Ocena matrike \mathbf{E}

Zapišimo \mathbf{X} kot vektor velikosti 9×1 :

$$\mathbf{X} = [\mathbf{e}_1^T, \mathbf{e}_2^T, \mathbf{e}_3^T]^T,$$

kjer je \mathbf{e}_i i -ta vrstica matrike \mathbf{E} .

Sedaj lahko enačbo (21) zapišemo kot:

$$\mathbf{a}^T \mathbf{X} = 0,$$

kjer je \mathbf{a} vektor velikosti 9×1 ¹⁴:

$$\mathbf{a} = [x_1 \tilde{\mathbf{m}}_2^T \quad y_1 \tilde{\mathbf{m}}_2^T \quad \tilde{\mathbf{m}}_2^T]^T.$$

Če imamo n korespondenčnih parov točk, nam vsak par da enačbo tipa $\mathbf{a}^T \mathbf{X} = 0$.

Ker je \mathbf{E} , podobno kot \mathbf{t} , definirana le do multiplikativnega faktorja natančno, imamo opravka le z osmimi neznankami. Torej je najmanjša vrednost za n , s katero lahko ocenimo \mathbf{X} , $n = 8$ (in od tod ime algoritma).

Enačbe tipa $\mathbf{a}^T \mathbf{X} = 0$ lahko združimo:

$$\mathbf{A}_n \mathbf{X} = \mathbf{0}, \tag{22}$$

kjer je \mathbf{A}_n matrika velikosti $n \times 9$:

$$\mathbf{A}_n = \begin{bmatrix} \mathbf{a}_1^T \\ \vdots \\ \mathbf{a}_n^T \end{bmatrix}.$$

¹⁴Ker imamo opravka z normaliziranim koordinatnim sistemom, je $\tilde{\mathbf{m}}_1 = [x_1, y_1, 1]^T$ in $\tilde{\mathbf{m}}_2 = [x_2, y_2, 1]^T$.

Če predpostavljamo, da obstaja v 3D prostoru množica točk M_1, \dots, M_8 , za katere velja, da je \mathbf{A}_8 reda 8, potem lahko iz enačbe (22) dobimo vrednosti koeficientov \mathbf{X} . Tako dobimo oceno matrike \mathbf{E} .

Izračun \mathbf{t} in \mathbf{R} iz \mathbf{E}

Naj bodo \mathbf{c}_i in \mathbf{r}_i , $i = 1, 2, 3$, vektorji kolon matrik \mathbf{E} in \mathbf{R} . Iz enačbe $\mathbf{E} = \mathbf{TR}$ (po enačbi (21)) sledi, da je $\mathbf{c}_i = t \times \mathbf{r}_i$ za vsak i . Iz tega sledi, da so vsi trije kolonski vektorji \mathbf{E} pravokotni na \mathbf{t} , torej je \mathbf{t} vzporeden z vektorskim produktom dveh \mathbf{c}_i . To dejstvo določi \mathbf{t} do predznaka natančno:

$$\mathbf{t} = \pm \mathbf{c}_1 \times \mathbf{c}_2 .$$

Če poznamo predznak ene od komponent \mathbf{t} (na primer t_3), poznamo tudi predznak \mathbf{t} .

Iz relacije $\mathbf{c}_i = t \times \mathbf{r}_i$ lahko sedaj izrazimo:

$$\mathbf{c}_i \times \mathbf{c}_{i+1} = (\mathbf{r}_{i+2} \cdot \mathbf{t}) \mathbf{t} , \quad i = 1, 2, 3 ,$$

kjer je $\mathbf{r}_{i+2} = \mathbf{r}_{((i+1) \bmod 3)+1}$. Matrika kofaktorjev \mathbf{E}^* matrike \mathbf{E} je dana kot:

$$\begin{aligned} \mathbf{E}^* &= [\mathbf{c}_2 \times \mathbf{c}_3, \mathbf{c}_3 \times \mathbf{c}_1, \mathbf{c}_1 \times \mathbf{c}_2]^T = \\ &= [(\mathbf{r}_1 \cdot \mathbf{t}) \mathbf{t}, (\mathbf{r}_2 \cdot \mathbf{t}) \mathbf{t}, (\mathbf{r}_3 \cdot \mathbf{t}) \mathbf{t}]^T = \\ &= (\mathbf{t} (\mathbf{R}^T \mathbf{t})^T)^T = (\mathbf{t} \mathbf{t}^T \mathbf{R})^T . \end{aligned}$$

Na podlagi naslednjega izreka

IZREK 2¹⁵

Za matriko $\mathbf{E} = \mathbf{TR}$ velja, da je matrika $\mathbf{E} \mathbf{E}^T$ odvisna le od translacijskega vektorja \mathbf{t} .

velja:

$$\mathbf{TE} = \mathbf{T}^2 \mathbf{R} = \mathbf{t} \mathbf{t}^T \mathbf{R} - (\mathbf{t} \cdot \mathbf{t}) \mathbf{R} ,$$

od tod pa sledi:

$$(\mathbf{t} \cdot \mathbf{t}) \mathbf{R} = \mathbf{E}^{*T} - \mathbf{TE} ,$$

kar nam da \mathbf{R} v odvisnosti od \mathbf{E} in \mathbf{t} .

3.5 Rekonstrukcija 3D točk

Oglejmo si linearno metodo rekonstrukcije:

Predvidevamo, da imamo dve sliki iz dveh kamer, na katerih smo našli korespondenčni točki m_1 in m_2 , ki imata koordinati (u_1, v_1) in (u_2, v_2) .

¹⁵Dokaz si lahko ogledate v [2], na strani 250.

Če s p_{ij} označimo elemente (znane) matrike $\tilde{\mathbf{P}}$, dobimo iz enačbe (4) naslednje enačbe:

$$\begin{aligned} U &= p_{11}X + p_{12}Y + p_{13}Z + p_{14}T \\ V &= p_{21}X + p_{22}Y + p_{23}Z + p_{24}T \\ S &= p_{31}X + p_{32}Y + p_{33}Z + p_{34}T . \end{aligned}$$

Za prehod iz projektivnih koordinat v evklidske koordinate upoštevamo enačbo (3) in izberemo $X = x$, $Y = y$, $Z = z$ in $T = 1$. Dobimo:

$$\begin{aligned} u(p_{31}x + p_{32}y + p_{33}z + p_{34}) &= p_{11}x + p_{12}y + p_{13}z + p_{14} \\ v(p_{31}x + p_{32}y + p_{33}z + p_{34}) &= p_{21}x + p_{22}y + p_{23}z + p_{24} , \end{aligned}$$

kar lahko zapišemo kot:

$$\begin{aligned} (p_{11} - up_{31})x + (p_{12} - up_{32})y + (p_{13} - up_{33})z + p_{14} - up_{34} &= 0 \\ (p_{21} - vp_{31})x + (p_{22} - vp_{32})y + (p_{23} - vp_{33})z + p_{24} - vp_{34} &= 0 . \end{aligned}$$

V dobljenih enačbah poznamo koeficiente p_{ij} ter slikovne koordinate (u, v) . Ko v enačbi vstavimo vrednosti p_{ij_1} , p_{ij_2} ¹⁶, (u_1, v_1) in (u_2, v_2) , dobimo sistem štirih linearnih enačb s tremi neznankami. Rešitev tega sistema nam da iskane 3D koordinate (x, y, z) točke, katere rekonstrukcijo smo iskali.

Posplošitev postopka na n slik je očitna.

V dodatku C je podana skica računskega primera celotnega postopka rekonstrukcije 3D točke.

¹⁶Matriki $\tilde{\mathbf{P}}_1$ in $\tilde{\mathbf{P}}_2$ morata biti izraženi v istem koordinatnem sistemu (stran 18).

Dodatki

A 3D rotacija

A.1 Splošne lastnosti rotacijske matrike

Rotacijo 3D vektorja \mathbf{v} v 3D vektor \mathbf{v}' lahko predstavimo z linearno transformacijo, ki je definirana z matriko \mathbf{R} , velikosti 3×3 :

$$\mathbf{v}' = \mathbf{R}\mathbf{v} ,$$

za katero velja:

$$\mathbf{R}\mathbf{R}^T = \mathbf{R}^T\mathbf{R} = \mathbf{I} \text{ in } \det(\mathbf{R}) = 1 ,$$

kjer je \mathbf{I} matrika identitete. Prva enačba nam pove, da je inverz matrike \mathbf{R} enak transponirani matriki \mathbf{R} , druga enačba pa nam pove, da transformacija ohranja relativno orientacijo referenčnega okvirja (osi koordinatnega sistema). [2] (strani 20, 21, 29 in 143)

Matrika \mathbf{R} , za katero velja prva enačba oziroma lastnost, se imenuje ortogonalna. Lastnost ortogonalnosti si lahko lažje predstavljamo, če \mathbf{v} in \mathbf{v}' izrazimo v dveh medsebojno ortogonalnih koordinatnih sistemih, ki sta definirana z enotskimi vektorji $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ in $\mathbf{e}'_1, \mathbf{e}'_2, \mathbf{e}'_3$. Iz zamišljenega sledi, da je element r_{ij} matrike \mathbf{R} enak kosinusu kota, ki ga definirata enotski vektorj \mathbf{e}_i in \mathbf{e}'_j :

$$r_{ij} = \mathbf{e}_i^T \mathbf{e}'_j .$$

Na podlagi enačbe $\mathbf{R}\mathbf{R}^T = \mathbf{R}^T\mathbf{R} = \mathbf{I}$ in zgornje enačbe, velja:

$$\sum_{j=1}^3 r_{ij}r_{kj} = \sum_{j=1}^3 r_{ji}r_{jk} = \begin{cases} 0, & i \neq k \\ 1, & i = k \end{cases} ,$$

torej vrstice (in kolone) matrike \mathbf{R} so medsebojno ortogonalni vektorji enotske dolžine [2] (stran 143):

$$\mathbf{r}_i^T \mathbf{r}_j = \begin{cases} 0, & i \neq j \\ 1, & i = j \end{cases} , \quad i, j = 1, 2, 3 , \quad i \leq j ,$$

kjer smo z \mathbf{r}_i ($i = 1, 2, 3$) označili vrstične (kolonske) vektorje matrike \mathbf{R} .

Devet elementov matrike \mathbf{R} torej ustreza zgornji relaciji, ki tvori šest neodvisnih omejitev. To nam zmanjša število prostostnih stopenj (parametrov) 3D rotacije na $9-6=3$ [6]. Obstaja več parametrizacij matrike \mathbf{R} : rotacija okoli koordinatnih osi, rotacija na osnovi osi in kota, rotacija na osnovi Eulerjevih kotov, rotacija na osnovi kvaternionov itd. V nadaljevanju pa si oglejmo prvo.

A.2 Rotacija okoli koordinatnih osi

3D rotacijo lahko izrazimo kot zaporedje treh rotacij okoli koordinatnih osi \mathbf{e}_1 , \mathbf{e}_2 in \mathbf{e}_3 : okoli koordinatne osi \mathbf{e}_1 (x os) za kot α ¹⁷, okoli koordinatne osi \mathbf{e}_2 (y os) za kot β in okoli koordinatne osi \mathbf{e}_3 (z os) za kot γ . Ti koti tako predstavljajo tri parametre matrike \mathbf{R} , vsaka rotacija pa je predstavljena z rotacijsko matriko \mathbf{R}_j , ki rotira okoli osi \mathbf{e}_j :

$$\begin{aligned} \mathbf{R}_1(\alpha) &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \alpha & -\sin \alpha \\ 0 & \sin \alpha & \cos \alpha \end{bmatrix} \\ \mathbf{R}_2(\beta) &= \begin{bmatrix} \cos \beta & 0 & \sin \beta \\ 0 & 1 & 0 \\ -\sin \beta & 0 & \cos \beta \end{bmatrix} . \\ \mathbf{R}_3(\gamma) &= \begin{bmatrix} \cos \gamma & -\sin \gamma & 0 \\ \sin \gamma & \cos \gamma & 0 \\ 0 & 0 & 1 \end{bmatrix} \end{aligned}$$

Matriko \mathbf{R} , ki opisuje celotno rotacijo, dobimo z zmnožkom matrik \mathbf{R}_j ¹⁸:

$$\mathbf{R} = \mathbf{R}_1 \mathbf{R}_2 \mathbf{R}_3 = \begin{bmatrix} \cos \beta \cos \gamma & -\cos \beta \sin \gamma & \sin \beta \\ \sin \alpha \sin \beta \cos \gamma + \cos \alpha \sin \gamma & -\sin \alpha \sin \beta \sin \gamma + \cos \alpha \cos \gamma & -\sin \alpha \cos \beta \\ -\cos \alpha \sin \beta \cos \gamma + \sin \alpha \sin \gamma & \cos \alpha \sin \beta \sin \gamma + \sin \alpha \cos \gamma & \cos \alpha \cos \gamma \end{bmatrix} .$$

Izračun kotov α , β in γ je na podlagi prejšnjega poglavja enostaven, vendar pa se moramo v primeru, da smo matriko \mathbf{R} dobili kot izhod neke numerične metode, vedno prepričati, da je dobljena matrika \mathbf{R} res ortogonalna.

B Računski primer vpliva zunanjih parametrov kamere na 3D točko M

S spremembo položaja in orientacije kamere v prostoru dobimo torej šest zunanjih parametrov kamere, tri za rotacijo \mathbf{R} in tri za translacijo \mathbf{T} (poglavje 2.3). Vzemimo ilustrativen primer, v katerem za tri parametre rotacije (dodatek A) velja: $\alpha=180^\circ$, $\beta = \gamma = 0^\circ$, vektor translacije pa naj bo določen s $\mathbf{T}=[1,1,1]^T$. Zanima nas, kakšne so nove koordinate 3D točke M , ki je v starem koordinatnem sistemu imela koordinate $M_{stara}=(10,10,10)$.

¹⁷V katero smer vrtimo? Če "primemo" koordinatno os za puščico, potem moramo zavrteti koordinatni sistem okoli te osi za dani kot v smeri urinega kazalca.

¹⁸Pri tem je seveda vrstni red množenja pomemben, saj nam drugačno zaporedje da drugačen rezultat v odvisnosti od kotov α , β in γ . Torej tukaj obstaja šest različnih načinov sestavljanja matrike \mathbf{R} .

Rešitev:

Po enačbi (15) velja:

$$\tilde{\mathbf{M}}_{nova} = \tilde{\mathbf{K}}^{-1} \tilde{\mathbf{M}}_{stara} ,$$

kjer je

$$\tilde{\mathbf{K}} = \begin{bmatrix} \mathbf{R} & \mathbf{T} \\ \mathbf{0}_3^T & 1 \end{bmatrix} ,$$

$$\mathbf{T} = [1, 1, 1]^T \text{ in}$$

$$\mathbf{R} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix} .$$

Ko izračunamo

$$\tilde{\mathbf{K}}^{-1} = \begin{bmatrix} 1 & 0 & 0 & -1 \\ 0 & -1 & 0 & 1 \\ 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix} ,$$

dobimo $M_{nova}=(9,-9,-9)$!

C Skica računskega primera rekonstrukcije 3D točke

Podane imamo naslednje podatke¹⁹:

- Delamo z normaliziranimi kamerama (poglavje 2.2).
- Izhodišče 3D prostora je v optičnem centru prve kamere.
- Točka m_2 na normalizirani slikovni ravnini \mathcal{R}_2 (poglavje 3.3) druge kamere:

$$m_2 = \left(\frac{1}{8}, 0\right) .$$

- Translacijski vektor med dvema koordinatnima sistemoma kamer:

$$\mathbf{t} = \left[5, 1, \frac{13}{2}\right]^T .$$

¹⁹Celoten postopek je zapisan tudi v datoteki programa Mathematica: *3DRec.nb*.

- Rotacijska matrika med dvema koordinatnima sistemoma kamer, kjer so parametri $\alpha = \gamma = 0^\circ$, $\beta = 270^\circ$ (dodatek A):

$$\mathbf{R} = \begin{bmatrix} 0 & 0 & -1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}.$$

Sedaj nas zanimajo 3D koordinate točke M , ki se preslika v koordinatnem sistemu druge kamere v točko m_2 .

Rešitev:

Po enačbi (21), kjer je \mathbf{T} po enačbi (18) enak:

$$\mathbf{T} = \begin{bmatrix} 0 & -\frac{13}{2} & 1 \\ \frac{13}{2} & 0 & -5 \\ -1 & 5 & 0 \end{bmatrix},$$

velja:

$$\begin{aligned} \tilde{\mathbf{m}}_1^T \mathbf{T} \mathbf{R} \tilde{\mathbf{m}}_2 &= \\ \tilde{\mathbf{m}}_1^T \begin{bmatrix} 0 & -\frac{13}{2} & 1 \\ \frac{13}{2} & 0 & -5 \\ -1 & 5 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 & -1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} \frac{1}{8} \\ 0 \\ 1 \end{bmatrix} &= \\ \tilde{\mathbf{m}}_1^T \begin{bmatrix} \frac{1}{8} \\ -\frac{57}{8} \\ 1 \end{bmatrix} &= 0. \end{aligned}$$

Točko $m_1 = (u, v)$ na slikovni ravnini prve kamere zapišemo v projektiivnih koordinatah kot $(u, v, 1)$, pri čemer za u in v po zgornji enačbi velja:

$$v = \frac{1}{57}u + \frac{8}{57}.$$

Možni kandidati za točko m_1 na slikovni ravnini prve kamere tako izpolnjujejo zgornji pogoj (epipolarna omejitve). S pomočjo korelacije dobimo korespondenčno točko (poglavje 3.4.1) m_1 , ker pa korelacijska metoda presega namen te seminarske naloge, zapišimo $m_1 = (\frac{1}{7}, \frac{1}{7})$ ²⁰.

Ker se nahajamo v normaliziranem koordinatnem sistemu, vemo, da je:

$$\tilde{\mathbf{P}}_1 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}.$$

²⁰Ta točka je dobljena iz realne situacije: preslikava 3D točke $M = (1, 1, 7)$.

Vemo tudi, da je matrika $\tilde{\mathbf{P}}_2$ v novem koordinatnem sistemu C_2XYZ enaka matriki $\tilde{\mathbf{P}}_1$. Matrika $\tilde{\mathbf{P}}_2$ izražena v koordinatnem sistemu C_1xyz pa je tako enaka:

$$\tilde{\mathbf{P}}_1 = \tilde{\mathbf{P}}_2 \tilde{\mathbf{K}} \Rightarrow \tilde{\mathbf{P}}_2 = \tilde{\mathbf{P}}_1 \tilde{\mathbf{K}}^{-1} = \begin{bmatrix} \mathbf{R}^T & -\mathbf{R}^T \mathbf{t} \end{bmatrix},$$

kjer je $\tilde{\mathbf{K}}$ matrika kolineacije (stran 13):

$$\tilde{\mathbf{K}} = \begin{bmatrix} \mathbf{R} & \mathbf{t} \\ \mathbf{0}_3^T & 1 \end{bmatrix}.$$

$\tilde{\mathbf{P}}_2$ je torej enaka:

$$\tilde{\mathbf{P}}_2 = \begin{bmatrix} 0 & 0 & 1 & -\frac{13}{2} \\ 0 & 1 & 0 & -1 \\ -1 & 0 & 0 & 5 \end{bmatrix}.$$

Sedaj poznamo slikovne koordinate točk m_1 in m_2 ter vrednosti elementov matrik $\tilde{\mathbf{P}}_1$ in $\tilde{\mathbf{P}}_2$, zato lahko izračunamo 3D koordinate točke M (poglavje 3.5) tako, da te podatke vstavimo v enačbi:

$$\begin{aligned} (p_{11} - up_{31})x + (p_{12} - up_{32})y + (p_{13} - up_{33})z + p_{14} - up_{34} &= 0 \\ (p_{21} - vp_{31})x + (p_{22} - vp_{32})y + (p_{23} - vp_{33})z + p_{24} - vp_{34} &= 0. \end{aligned}$$

Rešiti moramo torej sistem štirih linearnih enačb s tremi neznankami:

$$\begin{aligned} -\frac{57}{8} + \frac{1}{8}x + z &= 0 \\ -1 + y &= 0 \\ x - \frac{1}{7}z &= 0 \\ y - \frac{1}{7}z &= 0, \end{aligned}$$

ki nam dajo rešitev:

$$\underline{\underline{M = (1, 1, 7)}}.$$

Literatura

- [1] S. Birchfield, An Introduction to Projective Geometry (for computer vision), <http://robotics.stanford.edu/~birch/projective/>, 1998.
- [2] O. Faugeras, *Three-Dimensional Computer Vision: A Geometric Viewpoint*, MIT Press, 1993.
- [3] E. Kramar, *Linearna algebra*, zapiski iz predavanj, FRI, 1995.
- [4] J. N. Bronštejn, K. A. Semendjajev, *Matematični priročnik: za inženirje in slušatelje tehniških visokih šol*, Tehniška založba Slovenije, v Ljubljani, 1988.
- [5] D. Skočaj, *Avtomatsko modeliranje 3-dimenzionalnih večbarvnih predmetov z uporabo globinskega sensorja*, magistrska naloga, FRI, 1999.
- [6] E. Trucco, A. Verri, *Introductory Techniques for 3-D Computer Vision*, Prentice-Hall, 1998.