CENTER FOR
MACHINE PERCEPTION

CZECH TECHNICAL
UNIVERSITY

VÝZKUMNÁ ZPRÁVA

ISSN 1213-2365

Matematický model kamery v afinním prostoru

(Verze 1.0)

Jan Šochman, Tomáš Pajdla

sochmj1@cmp.felk.cvut.cz, pajdla@cmp.felk.cvut.cz

CTU–CMP–2002–11

27. května 2002

Lze získat na
ftp://cmp.felk.cvut.cz/pub/cmp/articles/sochman/Sochman-TR-2002-11.pdf

Tato práce byla podpořena granty GAČR 102/01/0971 a MSM 212300013

Research Reports of CMP, Czech Technical University in Prague, No. 11, 2002

Published by

Centrum strojového vnímání, Katedra kybernetiky
Fakulta elektrotechnická ČVUT
Technická 2, 166 27 Praha 6
fax: (02) 2435 7385, tel: (02) 2435 7637, www: http://cmp.felk.cvut.cz

1 Motivace

Mějme kameru a snímejme s ní okolní svět. Výstupem takového snímání je rovinný obraz světa. Chtěli bychom nějak využít tohoto obrazu světa k popisu nasnímané scény. Známe však pouze souřadnice bodů, (u, v), v obraze.

Jelikož je však kamera sama umístěna v prostoru, který sleduje, a rovina obrazu odpovídá obrazové rovině kamery, lze se na obrazové body dívat jako na body ve světě (prostoru A). Naším cílem je najít vztah mezi souřadnicemi (u, v) a souřadnicemi v prostoru A, neboli najít takové zobrazení f, které bodu z prostoru A přiřadí souřadnice v obraze.

Aby bylo možno činit nějaká další prohlášení, je nejdříve nutno blíže specifikovat prostor A. V našem případě použijeme afinní prostor[1].

2 Afinní prostor

Afinní prostor je definován jako trojice $(\mathcal{P}, \mathcal{V}, \varphi)$, kde $\mathcal{P}$ je množina bodů, $\mathcal{V}$ vektorový prostor a φ zobrazení $\varphi : \mathcal{P} \times \mathcal{P} \to \mathcal{V}$. Navíc musí být splněno následující:

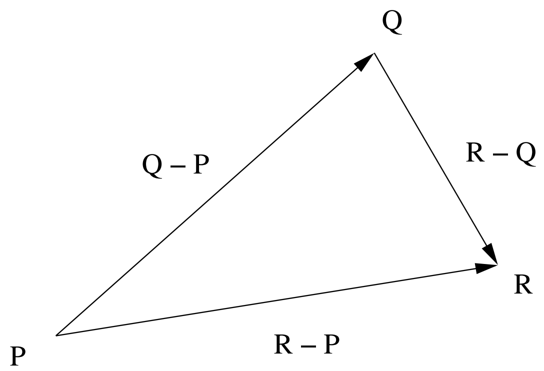
1. $\forall P, Q \in \mathcal{P}\ \exists! \vec{v} \in \mathcal{V} : \varphi(P, Q) = \vec{v}$
2. $\forall P \in \mathcal{P}\ \forall \vec{v} \in \mathcal{V}\ \exists! Q \in \mathcal{P} : \varphi(P, Q) = \vec{v}$
3. $\forall P, Q, R \in \mathcal{P} : \varphi(P, Q) + \varphi(Q, R) - \varphi(P, R) = 0$

Předpokládá se, že $\mathcal{V}$ má dimenzi n. Pak hovoříme o afinním prostoru dimenze n, který značíme A^n. Vektorovému prostoru $\mathcal{V}$ říkáme *zaměření* afinního prostoru A^n.

První axiom dává do souvislosti dva *body* z $\mathcal{P}$ s jedním *vektorem*. Říká, že φ je *funkce* z množiny bodů do vektorového prostoru. Také je možno se na tvrzení dívat tak, že φ přiřazuje dvojicim bodů, na kterých nemusí být definovány žádné operace, prvek z vektorového prostoru, ve kterém už umíme sčítat a násobit.

Druhý axiom říká, že ke každému bodu z množiny $\mathcal{P}$ a vektoru z vektorového prostoru $\mathcal{V}$ existuje právě jeden bod z $\mathcal{P}$. Je tak tedy definována funkce $\psi : \mathcal{P} \times \mathcal{V} \to \mathcal{P}$.

[1] Afinní prostor dobře modeluje geometrii běžného světa kolem nás. Závěry učiněné na základě axiomů afinního prostoru souhlasí se závěry o světě, ke kterým dojdeme běžným usuzováním na základě zdravého rozumu. Oproti užití zdravého rozumu má formulace a řešení problémů v řeči afinního prostoru tu výhodu, že zkoumání je systematické a lze se při něm opřít o výsledky známé z teorie lineárních prostorů.



Obrázek 1: Trojúhelníková rovnost.

Třetímu axiomu se říká trojúhelníková rovnost. Ta je známější ve formě (viz obrázek 1)

$$(Q - P) + (R - Q) = (R - P), \tag{1}$$

kde P, Q, R jsou opět body. Je však třeba si uvědomit, co znamenají operace $+$ a $-$. Za předpokladu, že body P, Q, R jsou body z A^n, tedy n-dimenzionálního afinního prostoru, představuje operace $-$ zobrazení φ, které přiřadí uspořádané dvojici bodů vektor. Toto zobrazení je někdy označováno jako odčítání bodů. Operace $+$ je sčítání vektorů ve vektorovém prostoru $\mathcal{V}$.

Na funkci ψ bychom narazili, pokud bychom rovnici (1) přepsali do tvaru

$$(Q + (R - Q)) - (P + (R - Q) = (R - P). \tag{2}$$

Zde funkce ψ odpovídá operaci $+$, jejímž parametrem je bod a vektor a výsledek je opět bod.

3 Zobrazení φ

Získat množinu bodů $\mathcal{P}$ a vektorový prostor $\mathcal{V}$ lze snadno. Můžeme vzít například vektorový prostor R^3. Tím však ještě není určeno zobrazení φ, natož jednoznačně. Díky vlastnostem afinního prostoru však nemusíme zobrazení φ definovat pro všechny body, ale postačí nám k tomu čtyři speciálně vybrané. Ukažme postup jeho konstrukce.

Mějme A^3, tedy $\mathcal{P}$, $\mathcal{V}$ dimenze 3 a φ splňující axiomy 1–3, které neznáme, ale o kterém víme, že existuje. Vezměme bod $O \in \mathcal{P}$ a tři lineárně nezávislé vektory $\vec{b}_1$, $\vec{b}_2$, $\vec{b}_3$ z $\mathcal{V}$. Podle druhého axiomu platí $\varphi(O, B_i) = \vec{b}_i$ pro nějaké tři body B_i, $i = 1 \ldots 3$. Získali jsme tak další tři body z $\mathcal{P}$.

Ukažme, že je φ takto definované na čtyřech dvojicích bodů jednoznačně dáno i na zbylých bodech. Nejdříve ukážeme jednoznačnost pro dvojice (O, X).

Z prvního axiomu víme, že pro takovouto dvojici existuje právě jeden vektor $\vec{v} \in \mathcal{V}$. Jelikož jsou $\vec{b}_1, \vec{b}_2, \vec{b}_3$ lineárně nezávislé, tvoří bázi a vektor $\vec{v}$ lze zapsat jako jejich lineární kombinaci. Máme tedy

$$\varphi(O, X) = \vec{v} = x_1\vec{b}_1 + x_2\vec{b}_2 + x_3\vec{b}_3 = x_1\varphi(0, B_1) + x_2\varphi(0, B_1) + x_3\varphi(0, B_1),$$

což jednoznačně přiřazuje této dvojici vektor z $\mathcal{V}$. Naopak z druhého axiomu získáme pro každý vektor $\vec{v} \in \mathcal{V}$ bod X takový, že $\varphi(O, X) = \vec{v}$. Zobrazení φ je tedy pro dvojici (O, X), kde X je libovolné, určeno jednoznačně. Vektorům $\vec{v}$ splňujícím rovnost $\vec{v} = \varphi(O, X)$ pro nějaký bod X říkáme *zaměření bodu* X.

Pro libovolné dva body $A, B \in \mathcal{P}$ je pak zobrazení $\varphi(A, B)$ definováno jednoznačně z trojúhelníkové rovnosti jako

$$\varphi(A, B) = \varphi(O, B) - \varphi(O, A).$$

Takovéto zobrazení existuje pro všechny dvojice, neboť výrazy na pravé straně jsou definovány pro všechny body a je i jednoznačné díky jednoznačnosti výrazů na pravé straně.

4 Souřadná soustava kamery

Vraťme se k naší úloze nalezení vztahu mezi souřadnou soustavou obrazu a souřadnou soustavou okolního světa. Zatím jsme specifikovali okolní svět jako afinní prostor A^3. Dalším krokem je definice souřadné soustavy obrazu vzhledem ke kameře.

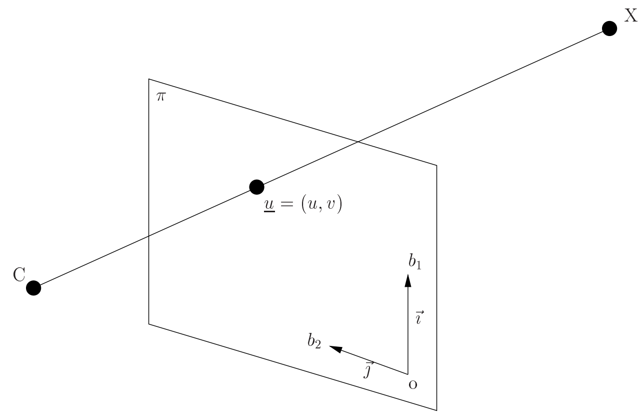
Pokud se nám toto podaří, převedeme náš problém na nalezení zobrazení jedné báze zaměření prostoru A^3 na druhou. Kamera je totiž umístěna v prostoru A^3 a tak je tedy báze spojená se souřadnou soustavou kamery i bází zaměření prostoru A^3.

Souřadnou soustavu kamery lze definovat mnoha způsoby, takže si ukážeme jednu z možných a ukážeme si, proč je právě tato výhodná.

Na obrázku 2 je naznačena situace v kameře[2]. C je optický střed kamery a π obrazová rovina. V rovině π již máme souřadnou soustavu (o, b_1, b_2), které odpovídají bázové vektory $\vec{\imath}$, $\vec{\jmath}$. Tato souřadná soustava odpovídá souřadné soustavě obrázku.

Zde je třeba pozastavit se nad pojmem „rovina“. Takovýto pojem v našem afinním prostoru zatím nemáme. Je třeba vyvarovat se toho, že si afinní prostor A^3, ve kterém se pohybujeme, představujeme jako R^3. V R^3 již víme,

[2] Předpokládáme dírkový model kamery.



Obrázek 2: Promítnutí bodu do obrazu.

co je rovina. Je to podprostor dimenze 2. Intuitivně tak ztotožňujeme rovinu v R^3 s rovinou v A^3, což není správně.

Rovina v afinním prostoru $A^3 = (\mathcal{P}, \mathcal{V}, \varphi)$ je opět podprostor dimenze 2, ovšem *afinní* podprostor. Tedy $A^2 = (\mathcal{P}', \mathcal{V}', \varphi')$, kde $\mathcal{V}'$ je podprostor $\mathcal{V}'$ dimenze 2, φ' je „zúžení“ φ na $\mathcal{V}'$ a $\mathcal{P}'$ jsou body z $\mathcal{P}$, na kterých je definováno φ'.

Je tedy nutné vzít v potaz zmenšení dimenze nejenom u vektorového prostoru $\mathcal{V}$, ale také pro zobrazení φ a množinu bodů $\mathcal{P}$. Jak přesně je definováno ono „zúžení“ φ a výběr bodů z $\mathcal{P}$ je vidět na následující formální definici roviny v afinním prostoru.

Definice: Rovinou v afinním prostoru $A^3 = (\mathcal{P}, \mathcal{V}, \varphi)$ rozumíme takový afinní prostor $A^2 = (\mathcal{P}', \mathcal{V}', \varphi')$, kde $\mathcal{V}'$ je podprostor $\mathcal{V}$ dimenze 2 a pro množinu bodů $\mathcal{P}'$ platí

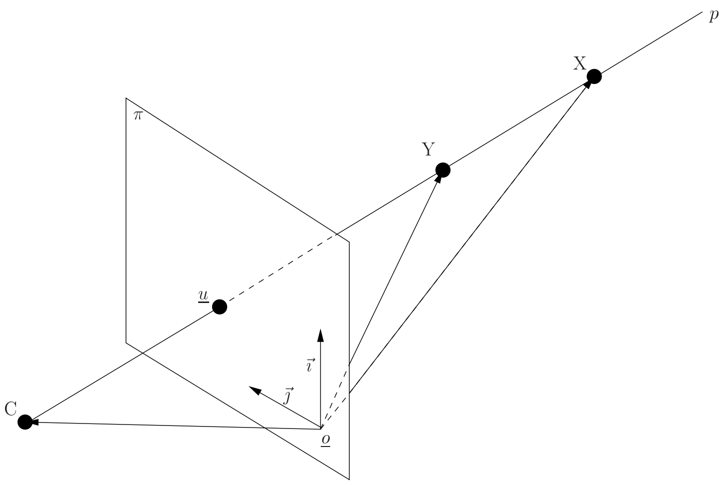
1. pro každý vektor $\vec{v} \in \mathcal{V}'$ a všechny body $P, Q \in \mathcal{P}$, pro které platí $\varphi(P, Q) = \vec{v}$, platí $P, Q \in \mathcal{P}'$,

2. pro každou dvojici bodů $P, Q \in \mathcal{P}'$ platí $\varphi(P, Q) \in \mathcal{V}'$.

Zobrazení φ' je definováno stejně jako φ ovšem pouze na $\mathcal{P}'$ a $\mathcal{V}'$.

Pokud tedy máme bod C, který neleží v rovině π, znamená to, že pro libovolný bod $X \in \pi$ platí $\varphi(C, X) \notin \mathcal{V}'$. Dostáváme tak vektor, který je lineárně nezávislý na vektorech z roviny π.

K nadefinování souřadné soustavy kamery potřebujeme stejně jako v případě hledání souřadné soustavy A^3 tři lineárně nezávislé vektory a jeden bod. Navíc chceme, aby tyto byly svázány s kamerou.



Obrázek 3: Lineární závislost vektorů při nevhodně zvolené souřadné soustavě kamery.

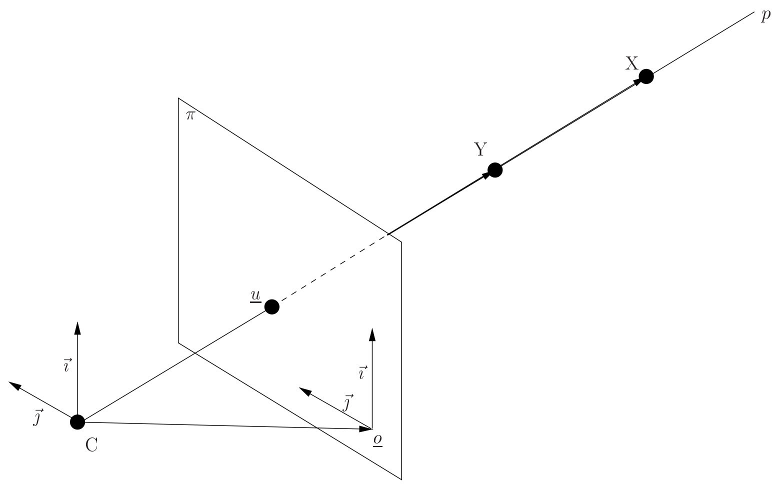
Asi nejjednodušší volbou je použít jako bázi vektory $\vec{\imath}$, $\vec{\jmath}$, $\varphi(\underline{o}, C)$ (víme, že jsou lineárně nezávislé) a jako počátek bod $\underline{o}$. Takováto volba je korektní a má tu výhodu, že body z obrazové roviny v něm budou mít souřadnice $(u, v, 0)$. Dostali jsme tedy snadný převod ze souřadné soustavy obrazu do souřadné soustavy kamery.

Takto zvolená souřadná soustava však má i jeden nedostatek (viz obrázek 3). Později budeme pro zformulování rovnice (3) požadovat, aby vektory zaměřující všechny body na přímce p (například X, Y) byly lineárně závislé na zaměření bodu $\underline{u}$, které reprezentuje projekci bodů na přímce p do roviny π.

V námi navržené souřadné soustavě jsou však vektory $\varphi(\underline{o}, X)$ a $\varphi(\underline{o}, Y)$ lineárně nezávislé na zaměření bodu $\underline{u}$.

Pokusíme se souřadnou soustavu zvolit vhodněji. Vezmeme jako počátek bod C a tři nezávislé vektory $\vec{\imath}$, $\vec{\jmath}$, $\varphi(C, \underline{o})$ (obrázek 4). Ty nám opět definují souřadnou soustavu kamery, ve které máme jednoduchý převod z obrazových souřadnic do souřadné soustavy kamery. Bod (u, v) z obrazu bude mít souřadnice $(u, v, 1)$. Navíc jsme získali lineární závislost vektorů $\varphi(C, X)$ a $\varphi(C, Y)$ na $\varphi(C, \underline{u})$.

Nalezli jsme tedy vhodnou souřadnou soustavu kamery, do které máme snadný převod ze souřadné soustavy obrazu a navíc jsou v ní zaměření všech bodů na jedné přímce (paprsku) procházející středem promítání lineárně závislé na zaměření bodu reprezentujícím projekci těchto bodů. Můžeme tedy pokročit k definici matematického modelu kamery.



Obrázek 4: Vhodně zvolená souřadná soustava kamery.

5 Matematický model kamery

Nyní již máme vše nachystáno k tomu, abychom dokázali vyjádřit náš problém formálně. Označme souřadnice vzhledem k obecné souřadné soustavě afinního prostoru A^3 indexem $_\beta$ a souřadnice vzhledem k souřadné soustavě kamery indexem $_{\beta'}$.

Jak již bylo řečeno, body na jednom paprsku vycházejícím z optického středu kamery by se měly promítnout do stejného bodu v obraze. Pro každý bod X z prostoru A^3 tedy musí platit

$$\alpha \begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = \varphi_{\beta'}(C, X). \tag{3}$$

Neboli vektor $\varphi_{\beta'}(C, X)$ je α-násobkem vektoru $(u, v, 1)$, který v souřadné soustavě kamery odpovídá bodu $\underline{u}$ – promítnutý bod X. Tedy $\varphi_{\beta'}(C, \underline{u}) = (u, v, 1)$.

Obvykle však $\varphi_{\beta'}(C, X)$ neznáme. Pokud máme známý objekt, můžeme však nalézt $\varphi_\beta(C, X)$ tak, že si v A^3 zvolíme libovolnou souřadnou soustavu. Potřebujeme tedy navíc najít zobrazení f takové, že

$$\varphi_{\beta'}(C, X) = f(\varphi_\beta(C, X)).$$

Jelikož se nejedná o nic jiného, než o přechod z jedné báze do druhé, je možno zobrazení f vyjádřit maticí, takže lze ekvivalentně napsat

$$\varphi_{\beta'}(C, X) = A\varphi_\beta(C, X) \tag{4}$$

kde A je odpovídající matice. S využitím (4) dostáváme přepis vztahu (3)

$$\alpha \begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = A\varphi_\beta(C, X) = A(X - C). \tag{5}$$

Tato rovnice je základním vztahem mezi kamerou (reprezentovanou maticí A), bodem v prostoru X a bodem v obraze (u, v). Častěji se s touto rovnicí setkáme ve tvaru

$$\alpha \underline{u} = P \begin{pmatrix} X \\ 1 \end{pmatrix},$$

kde $P \in R^{3\times 4}$ odpovídá matici A a nazývá se *projekční matice*.

Takto vybudovaný model kamery je sice správný, ale ještě ne zcela úplný. Zatím nevíme, co dělat z body v rovině rovnoběžné s π a procházející bodem C. Kam se promítnou? Další zvláštní případ nastane, když budeme zobrazovat přímku kolmou na rovinu π neprocházející středem promítání. Pokud bychom po ní šli pořád dál a dál, až do nekonečna, promítne se nám toto nekonečno do bodu v obraze. Jak se máme chovat k takovýmto bodům?

Je tedy nutno učinit další krok a přejít z afinního prostoru do prostoru projektivního.