Vybrané kapitoly matematiky a aplikací

Jaroslav Hrdina, Lenka Zalabová

November 19, 2020

Analytická geometrie a binokulární vidění

Smyslem binokulárního vidění je z obrazu dvou kamer (projekcí) zrekonstruovat objekt v prostoru. Jako objekt zvolíme přímku. Jako model kamery budeme používat modifikaci dírkové kamery (pinhole cammera, camera obscura). V dírkové kameře místo objektivu používáme otvor v jedné stěně. Světelný paprsek pak projde otvorem (dírkou) D a zobrazí se na projekční rovinu π. Dírka D pak slouží jako ohnisko a obraz se na projekční rovinu zobrazuje zrcadlově. Tento princip se opravdu u starých fotoaparátů používal a je možné si ho doma i vyrobit [3]. Na obrázku 1 se čtyřstěn $\mathcal{S}$ zobrazí na trojúhelník $\mathcal{T}$. Binokulární vidění pak

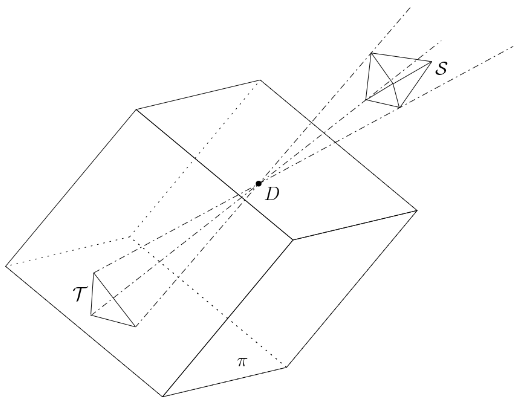

Figure 1: Dírková kamera

ze dvou takových projekcí zrekonstruuje původní objekt. Princip binokulárního vidění je na obrázku 4, jednotlivé projekce spolu s ohniskem kamery tvoří rovinu a hledaná přímka je pak v průniku dvou rovin.

Projekční rovina

Máme k dispozici projekce 3D přímky do projekčních rovin jednotlivých kamer. To jsou pro nás vstupní data ze kterých budeme přímku rekonstruovat. V praxi to znamená, že máme k dispozici dva obrázky v nějakém dostačujícím rozlišení. Nejprve musíme pomocí analýzy obrazu nalézt na každém z obrázků přímku. Přímku najdeme například tak, že nalezneme dva body které na přímce leží, nebo pomocí filtru detekce hran. Nalezené body si označíme jako B a C, jak je vidět na obrázku 2.

Pro popis přímky je ale nepraktické pamatovat si ji formou dvou bodů, které na ní leží. Je pak obtížné ověřit, jestli jiný bod na přímce leží, a také je obtížné vygenerovat další body, které na přímce leží. Přímka v rovině může být obecně určena různými kombinacemi dat:

i) Dvěma body: C a D - *afinní vyjádření*

ii) Bodem C a směrovým vektorem s - *parametrické vyjádření*

iii) Bodem A průniku s osou x a úhlem θ který svírá přímka s osou x (směrnicí) - *směrnicové vyjádření*

iv) Bodem B a normálovým vektorem n - *obecné vyjádření*

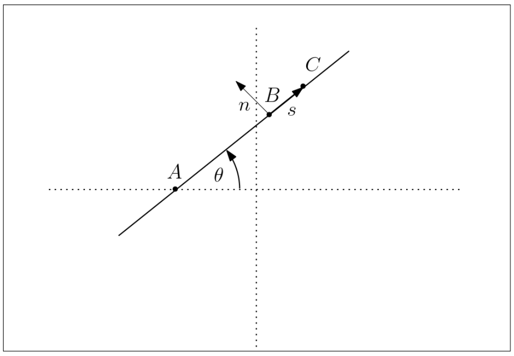

Figure 2: Projekce přímky na projekční rovinu

Pokud máme přímku určenou dvěma body B a C, máme směrový vektor $s = C - B = (s_1, s_2)$ a normálový vektor $n = (-s_2, s_1)$. Všimněme si základního rozdílu mezi parametrickým a obecným vyjádřením přímky. Parametrické vyjádření můžeme přepsat pomocí zápisu

$$[x_1, x_2] = B + ts = [B_1, B_2] + t(s_1, s_2), \;\; t \in \mathbb{R},$$

který říká, že pro libovolné $t \in \mathbb{R}$ dostaneme bod na přímce, a dostaneme tak všechny body. Obecné vyjádření nám popisuje přímku jako následující množinu

$$\{X = [x_1, x_2] : \;\; (B - X) \cdot n = 0\}, \tag{0.1}$$

kde $\cdot$ označuje Eukliedovský skalární součin definovaný v $\mathbb{R}^3$ předpisem

$$(u_1, u_2, u_3) \cdot (v_1, v_2, v_3) = u_1 v_1 + u_2 v_2 + u_3 v_3.$$

Pomocí obecného vyjádření je možné ověřit, jestli bod X leží, nebo neleží na přímce. Podrobněji je možné podmínku v definici množiny (0.1) přepsat jako

$$(B - X) \cdot n = ([B_1, B_2] - [x_1, x_2]) \cdot (n_1, n_2) = B_1 n_1 + B_2 n_2 - x_1 n_1 - x_2 n_1 = 0. \tag{0.2}$$

Dostaneme pak lineární rovnici $x_1 n_1 + x_2 n_1 + D = 0$, kde $D = -(B_1 n_1 + B_2 n_2)$. Obecné vyjádření přímky v rovině je tedy jedna lineární rovnice, kterou body splňují právě tehdy když na přímce leží. Dá se tedy říct, že parametrické vyjádření body na přímce generuje a obecné vyjádření body na přímce ověřuje. Podíváme se na obecné vyjádření jako na rovnici jejíž řešení hledáme. Máme dvě proměnné v jedné rovnici, která má řešení. Musíme tedy zavést jeden parametr. Parametrické řešení rovnice je pak parametrický zápis stejné přímky. V tom jsou tyto dvě reprezentace význačné a je rozumné mezi nimi přecházet.

Umístění kamery v prostoru

Podívejme se jak popsat vnitřní parametry dírkové kamery a její umístění v prostoru. Poznamenejme, že ve skutečnosti pracujeme s kamerou vybavenou optikou a dírková kamera je jen vhodným modelem. Ve skutečnosti se obraz na projekční rovině neotáčí, to dosáhneme v modelu umístěním ohniska F za projekční rovinu π.

Kamera je tedy v prostoru určená pozicí projekční roviny π, ohniska F a svou orientací, která je určena jednotkovým vektorem s, který hraje roli osy x. Vnitřní parametry kamery jsou ohnisko F a ohnisková vzdálenost $|F - D|$, kde D je projekce bodu na rovinu π, jak je znázorněno na obrázku 3. Stejně jako v rovině máme k dispozici parametrické a obecné

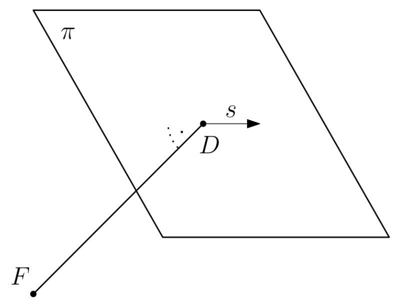

Figure 3: Umístění kamery v prostoru

vyjádření roviny. V našem případě je rovina π přirozeně zadaná bodem $D = [d_1, d_2, d_3]$ a normálovým vektorem $n = F - D = (n_1, n_2, n_3)$ a její obecná rovnice je tedy

$$0 = (X - D) \cdot n = x_1 n_1 + x_2 n_2 + x_3 n_3 + E, \text{ kde } E = -(d_1 n_1 + d_2 n_2 + d_3 n_3) \in \mathbb{R}$$

Umístění kamery v prostoru je tedy určeno jednou lineární rovnicí $x_1 n_1 + x_2 n_2 + x_3 n_3 + E = 0$ určující projekční rovinu, její orientací s a bodem F určujícím ohnisko.

Dalším krokem je umístit přímku na projekční rovinu v prostoru. Přímka v prostoru je určena dvěma lineárními rovnicemi, které určují dvě roviny v jejichž průniku přímka leží. V našem případě je jednou z rovin projekční rovina π a pro druhou rovinu musíme využít projekci přímky.

Všimněme si, že je výhodné pokud máme přímku popsanou ve směrnicovém tvaru, který je určený bodem $A = [a_1, 0]$ ležícím v průniku přímky a osy x a úhlem θ, který svírá průvodič přímky s osou x. Nalezneme tak lehce dva body na projekční rovině π, které určují 3D pozici detekované přímky. Protože vektor s hraje roli osy x, dostaneme bod

$$P_1 := D + a_1 s.$$

Roli osy y hraje jednotkový vektor v, který doplňuje vektory n a s do pravotočivé soustavy souřadnic (s, v, n). Takový vektor nalezneme pomocí vektorového součinu, který si zavedeme v textu o kousek dál. V našem případě označíme $v = s \times n$ a dostaneme

$$P_2 := D + \sin(\theta)(s \times n).$$

Body P_1, P_2 spolu s ohniskem F tvoří druhou rovinu určující přímku π_2.

Binokulární vidění

V obrázku 4, vidíme jak binokulární vidění funguje. Máme dvě kamery s projekčními rovinami π_1 a π_2. Na obou projekčních rovinách máme přímku určenou dvěma body a to konkrétně body P_1, P_2 na rovině π_1 a body Q_1, Q_2 na rovině π_2. Tyto přímky umíme vyjádřit v obecném vyjádření pomocí dvojice rovin. Jedna z nich je vždy příslušná projekční rovina π_i. Druhá rovina je určena dvěma body přímky a ohniskem příslušné kamery, tyto roviny označíme σ_i. Výsledek binokulárního vidění je pak hledaná přímka která je průniku rovin σ_1 a σ_2. Její obecné vyjádření je přímo dvojice rovnic určujících obě roviny σ_i.

$$x_1 n_1 + x_2 n_2 + x_3 n_3 + E = 0,$$
$$x_1 m_1 + x_2 m_2 + x_3 m_3 + F = 0.$$

Pokud nás zajímá směr výsledné přímky, uvědomme si, že pokud přímka leží v průniku dvou

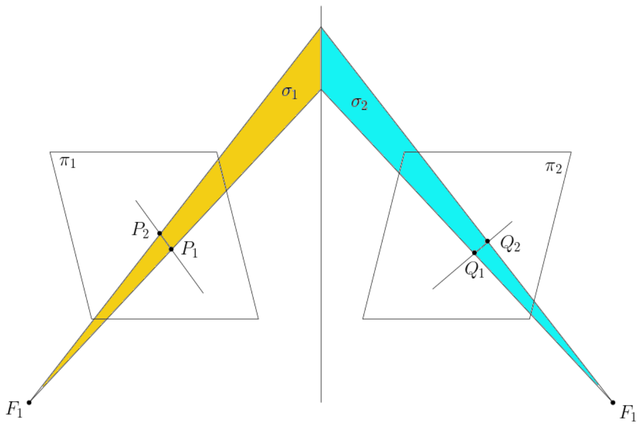

Figure 4: Binokulární vidění

rovin, je její směrový vektor kolmý k normálovým vektorům obou přímek. Hledáme tedy vektor současně kolmý k vektorům $n = (n_1, n_2, n_3)$ a $m = (m_1, m_2, m_3)$.

Vektorový součin

Využijeme vlastnosti determinantu. Sestavíme si determinant, kde do prvního řádku dáme formálně vektor $s = (s_1, s_2, s_3)$ a do dalších dvou řádků vektory m a n. Pak použijeme Lapalaceův rozvoj podle prvního řádku

$$\begin{vmatrix} s_1 & s_2 & s_3 \\ n_1 & n_2 & n_3 \\ m_1 & m_2 & m_3 \end{vmatrix} = s_1 \begin{vmatrix} n_2 & n_3 \\ m_2 & m_3 \end{vmatrix} - s_2 \begin{vmatrix} n_1 & n_3 \\ m_1 & m_3 \end{vmatrix} + s_3 \begin{vmatrix} n_1 & n_2 \\ m_1 & m_2 \end{vmatrix}$$
$$= s \cdot \left(\begin{vmatrix} n_2 & n_3 \\ m_2 & m_3 \end{vmatrix}, - \begin{vmatrix} n_1 & n_3 \\ m_1 & m_3 \end{vmatrix}, \begin{vmatrix} n_1 & n_2 \\ m_1 & m_2 \end{vmatrix} \right)$$

Pokud za s dosadíme vektor m nebo n, determinant matice vlevo je rovný nule. Vektor

$$m \times n = \left(\begin{vmatrix} n_2 & n_3 \\ m_2 & m_3 \end{vmatrix}, - \begin{vmatrix} n_1 & n_3 \\ m_1 & m_3 \end{vmatrix}, \begin{vmatrix} n_1 & n_2 \\ m_1 & m_2 \end{vmatrix} \right)$$

je tedy kolmý v vektorům m a n.

Determinant můžeme chápat jako objem. Pokud vektor $m \times n$ dosadíme místo vektoru s do zmíněné matice, dostaneme jeho velikost na druhou, tj $|m \times n|^2$. Protože objem rovnoběžnostěnu je obsah podstavy krát výška, dostaneme $|m \times n|^2 = |m \times n| V(m, n)$, kde

$V(m,n)$ je obsah lichoběžníku určeného vektory m a n. Pokud jsou vektory m a n lineárně nezávislé tak velikost $|m \times n| \neq 0$ je nenulová a můžeme pomocí ní dělit. Dostaneme výraz

$$|m \times n| = V(m,n),$$

a velikost vektoru $|m \times n|$ je tedy rovna obsahu lichoběžníku určeného vektory m a n. Determinant je ve skutečnosti orientovaný objem a je kladný pokud vektory v řádcích tvoří pravotočivou bázi. Pokud za s dosadíme vektor $m \times n$ je determinant matice vždy kladné číslo $|m \times n|^2$. Vektor $m \times n$ doplňuje tedy dvojici m a n na pravotočivou bázi. Tyto tři vlastnosti určují $m \times n$ jednoznačně a definují na $\mathbb{R}^3$ takzvaný vektorový součin

$$\times : \mathbb{R}^2 \times \mathbb{R}^3 \to \mathbb{R}^3$$

předpisem

$$(u_1, u_2, u_3) \times (v_1, v_2, v_3) = (u_2v_3 - u_3v_2, -u_1v_3 + u_3v_1, u_1v_2 - u_1v_2).$$

References

[1] Horák, P., Janyška, J.: Analytická geometrie, Masarykova univerzita 1997

[2] Zlatoš, P.: Lineárna algebra a geometria (Cesta z troch rozmerov s presahmi do príbuzných odborov) 2011

[3] Helena Macenauerová, Fotografování dírkovou kamerou, článek na internetové stránce, url: https://www.fotoaparat.cz/clanek/45/fotografovani-dirkovou-kamerou-2150/