

Robotické vidění

4. ledna 2007

<http://robotika.cz/guide/vision>

Vidění - motivace

- levné kamery (webcam, mobilní telefony, . . .)
- obraz je bohatý na množství informace
- vidění je hlavní senzorický systém u lidí

Vidění - demotivace

- náročné na zpracování
- obraz obsahuje i množství „nezajímavé informace“
- velmi citlivé na světelné podmínky (tma, sluneční osvětlení, . . .)

Robotické vs. počítačové vidění

- nutnost real-time (v nejhorším 1 snímek za sekundu)
- spíše videostream než řada obrázků
- možnost aktivního vidění (směřovat kameru na důležitá místa)

video - monocularVision-autonomousDriving.mov

Základní úlohy (high level)

- vyhybání se překážkám
- lokalizace
- mapování (vytváření modelu světa)
- sledování pohyblivého objektu

Základní úlohy (medium level)

- extrakce 3D informace z 2D snímků
- extrakce pohyblivých objektů
- matching / rozpoznání objektů

Základní úlohy (low level)

- kalibrace kamery
- extrakce zajímavých míst (feature extraction)
- sledování a matchování bodů
- výpočet transformací

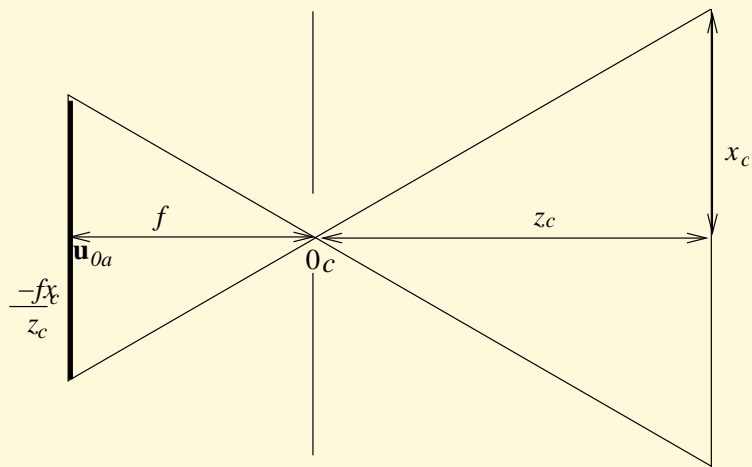
3D model světa

- *Obtaining 3D Models With a Hand-Held Camera* / Marc Pollefeys, SIGGRAPH 2001
- velmi dobrý úvod do problematiky
- je třeba rozumnět vstupním obrázkům videa
- sledování mnoha bodů scény



Model štěrbinové kamery

- idealizovaný jednoduchý model kamery
- paprsek světla vychází z bodu scény, prochází štěrbinou kamery a dopadá do roviny obrázku (*image plain*)



Souřadnicové soustavy

souřadnice vnějšího světa — index w

kamerové souřadnice — počátek ve středu štěrbin, osa Z_c směřuje od kamery, index c

eukleidovské souřadnice v obrázku — osy rovnoběžné s kamerovými souřadnicemi, ale X_i, Y_i leží v rovině obrázku, index i

afinní souřadnice obrázku — podobné předchozím, osy U, V, W, U nemusí svírat pravý úhel s V, W .

Projekce prováděná kamerou

- bod v prostoru \mathbf{x} , světové souřadnice $\mathbf{x}_w = [x_w, y_w, z_w]^T$
- převod ze světových do kamerových souřadnic: posunutí t a otočení R , tedy platí

$$\mathbf{x}_c = \begin{bmatrix} x_c \\ y_c \\ z_c \end{bmatrix} = R(\mathbf{x}_w - t)$$

- označíme-li si f ohniskovou vzdálenost, z podobnosti trojúhelníků pro převod do roviny obrázku platí:

$$\mathbf{u}_c = \left[-\frac{f x_c}{z_c}, -\frac{f y_c}{z_c}, -f \right]^T$$

- zajímají nás především afinní souřadnice tohoto bodu

Afinní souřadnice

principiální bod — průsečík optické osy s rovinou obrázku. V afinních odpovídá bodu $\mathbf{u}_{0a} = [u_0, v_0, 0]^T$.

homogenní souřadnice — bod \mathbf{u} v homogenních souřadnicích můžeme reprezentovat jako $\tilde{\mathbf{u}} = [U, V, W]^T$, kde platí $\mathbf{u} = [u, v]^T = [U/W, V/W]^T$.

afinní transformace — celou transformaci lze vyjádřit jako násobení matice 3×3 . Neznámé a , b a c — sklonění a změna měřítka:

$$\tilde{\mathbf{u}} = \begin{bmatrix} U \\ V \\ W \end{bmatrix} = \begin{bmatrix} a & b & u_0 \\ 0 & c & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} -\frac{fx_c}{z_c} \\ -\frac{fy_c}{z_c} \\ 1 \end{bmatrix} = \begin{bmatrix} -fa & -fb & u_0 \\ 0 & -fc & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \frac{x_c}{z_c} \\ \frac{y_c}{z_c} \\ 1 \end{bmatrix}$$

Vnitřní a vnější parametry kamery

- po přenásobení z_c dostaneme:

$$\begin{aligned} z_c \tilde{\mathbf{u}} &= \begin{bmatrix} -fa & -fb & u_0 \\ 0 & -fc & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_c \\ y_c \\ z_c \end{bmatrix} = \begin{bmatrix} -fa & -fb & u_0 \\ 0 & -fc & v_0 \\ 0 & 0 & 1 \end{bmatrix} R(\mathbf{x}_w - t) = \\ &= KR(\mathbf{x}_w - t) \end{aligned}$$

- K - kalibrační matice kamery, koeficienty označovány jako vnitřní (*intrinsic*) parametry kamery
- R a t - poloha kamery vzhledem ke světovým souřadnicím, vnější (*extrinsic*) parametry kamery

Barva

- barva odpovídá určité vlnové délce
- různé tvorové mají různě „nastavené senzory“ (lidé, včely, hadi, . . .)
- většina algoritmů pracuje pouze s 2D funkcí
- převod barevného vstupu na šedotonový obraz
- jasová složka nese nejvíce informace ($Y=R*0.299+G*0.587+B*0.114$)

Feature Extraction (1/2)

- znak musí být v malém okolí jedinečný



- podobnost vzorků obrazu (např. 7×7)
 - sum-of-square-differences (SSD)
 - normalised cross-correlation (NCC)

Feature Extraction (2/2)

- derivace podle x a y :

$$I_x(x, y) = \frac{I(x + 1, y) - I(x - 1, y)}{2}$$

$$I_y(x, y) = \frac{I(x, y + 1) - I(x, y - 1)}{2}$$

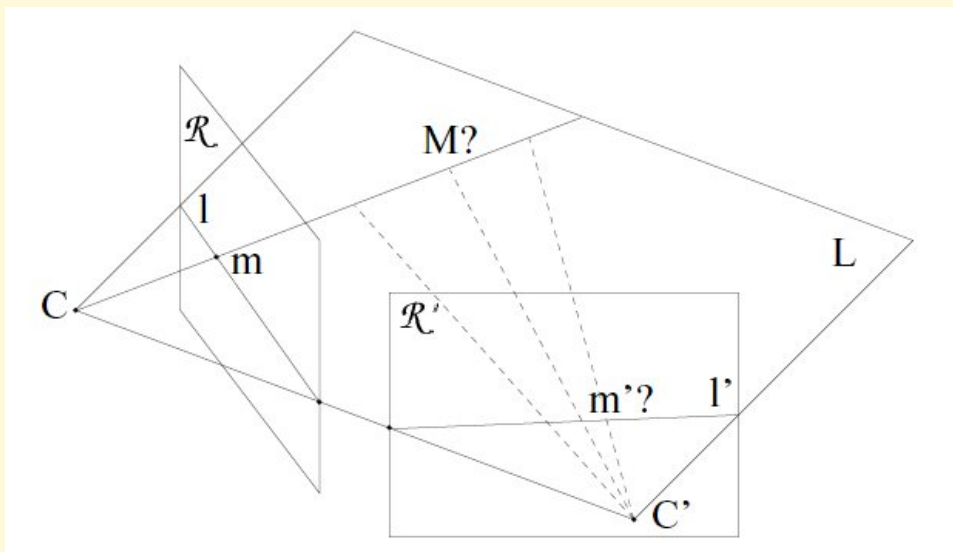
- gradient G (pro optický tok):

$$G = \sum_{(x,y) \in W} \begin{bmatrix} I_x^2(x, y) & I_x(x, y)I_y(x, y) \\ I_x(x, y)I_y(x, y) & I_y^2(x, y) \end{bmatrix}$$

demo Lucas Kanade Feature Tracker

Geometrie dvou pohledů

- projekce potenciálních podobností (epipolar line)



Rekonstrukce 3D

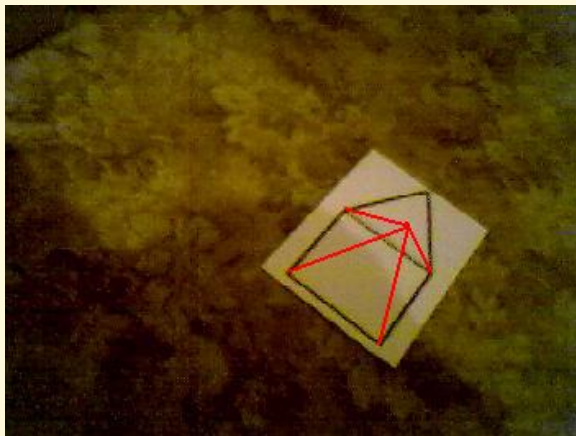
- máme-li 7 (8) odpovídajících dvojic, je možné dopočítat 3D rotaci a posun
- singularity:
 - pohled na rovinu
 - rotace kamery kolem osy
- co dělat, pokud více dvojic, ale některé jsou špatné?

RANSAC — RANdOm SAmpling Consensus

- Fischler a Bolles (1981)
- náhodně vybereme k vzorků (pro 3D rekonstrukci vzorek = 7 bodů)
- každý vzorek ohodnotíme (např. jak moc sedí projekce zbylých bodů)
- po dostatečném opakování dostaneme s vysokou pravděpodobností nejlepší volbu

Loňské cvičení — hledání domečku

- úkol: nalézt v obrázku ikonu domu
- domeček = čtverec + rovnostranný trojúhelník
- vstup: barevný obrázek 320×240



Letos — hledání cesty

- motivace: RoboTour / Grand Challenge



OpenCV

- podpora počítačového vidění
- původně byla vytvořena firmou Intel
- open source
- <http://www.intel.com/research/mrl/research/opencv/>

OpenCV — IplImage

- automaticky alokovaný obraz

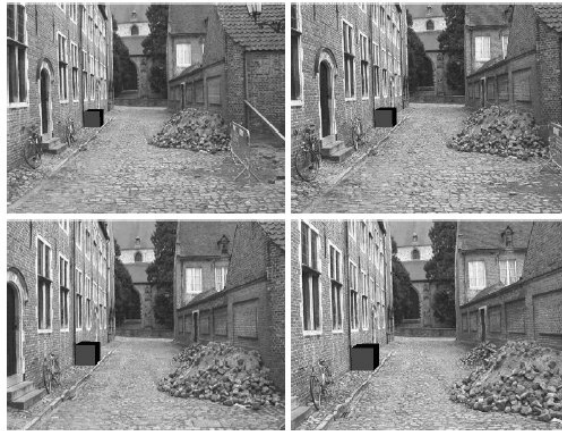
```
IplImage* image = cvLoadImage("park.jpg");  
IplImage* gray  = cvCreateImage(cvSize(image->width,  
                                     image->height), image->depth, 1);  
cvCvtColor(image, gray, CV_BGR2GRAY);  
cvSaveImage("gray.jpg", image);  
cvReleaseImage(&gray);  
cvReleaseImage(&image);
```

OpenCV — použití kamery

```
IplImage* image = 0;
CvCapture* capture = cvCaptureFromCAM(0);
cvNamedWindow("TestCam", CV_WINDOW_AUTOSIZE);
while(1)
{
    image = cvQueryFrame( capture );
    if(image == 0)
        break;
    myProcessing(image);
    cvShowImage("TestCam", image );
    if( cvWaitKey(10) == 27 )
        break;
}
cvReleaseCapture(&capture);
cvDestroyWindow("TestCam");
```


Augmented reality

- augmented = „rozšířená“
- kombinace virtuální reality se skutečností (video)
- aplikace: architektura, opravářství, medicína. . .



ProMIS

- Professional Minimum Invasion Surgery trainer
- kombinace reálných laparoskopických nástrojů s virtuální realitou
- reálný nebo simulovaný endoskop
- pouze vizuální zpětná vazba
- info: www.haptica.com



Na co příště určitě zapomeneme...

- Seminář z mobilní robotiky
- Eurobot 2007
- RoboTour 2007
- Robotický den (dobrovolníci?)
- anketa