Feuille d'exercices #2 — Vision par ordinateur

IFT 3345

Question 1. *Backprojection*

Un robot mobile est équipé d'une caméra RGB-D. Un pixel est détecté dans l'image et sa profondeur est connue. Votre tâche est de calculer la position 3D de ce point dans le **repère mondial**.

Supposons que le repère mondial soit "Est-Nord-Haut" - c'est-à-dire x vers la droite, y vers le haut et z sortant du sol. Supposons que le repère du robot soit également "Est-Nord-Haut" - avec x sortant de l'avant du robot, y vers la gauche et l'axe z pointant vers le ciel. Le repère de la caméra a x vers la droite, y vers le bas et z sortant de la caméra.

Les paramètres intrinsèques de la caméra sont donnés par :

$$K = \begin{bmatrix} 500 & 0 & 250 \\ 0 & 500 & 250 \\ 0 & 0 & 1 \end{bmatrix}$$

La caméra est montée à une position de $(0.2, 0.0, 0.2)$ *dans le repère du robot* avec une orientation alignée avec la direction du robot et parallèle au plan du sol.

La position du robot dans le repère mondial est $\mathbf{x} = (1.5, 1.5)$ (sur le plan du sol) avec un cap (yaw) de $\psi = 90^o$.

Une caractéristique est observée au pixel $(u, v) = (400, 200)$ (depuis le coin supérieur gauche). La profondeur associée est $d = 2.5m$. Pour trouver la position de la caractéristique dans le repère mondial, nous devons procéder comme suit :

(a) Convertir le pixel en un point 3D dans le **repère de la caméra**.

(b) Transformer le point dans le **repère du robot**.

(c) Transformer le point dans le **repère mondial**.

Question 2. *Points idéaux (points à l'infini)*

(a) Considérons le point homogène

$$\tilde{\mathbf{x}} = \begin{bmatrix} 3 \\ -2 \\ 0 \end{bmatrix}.$$

Cela représente un point à l'infini (un point idéal) puisque la valeur du troisième élément est 0. Quel est l'ensemble des lignes qui se rencontrent toutes en ce point particulier à l'infini ?

(b) Deux lignes dans l'image sont

$$\ell_1 = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}, \qquad \ell_2 = \begin{bmatrix} 2 \\ -2 \\ 1 \end{bmatrix}.$$

Calculez leur intersection en utilisant le produit vectoriel. L'intersection est-elle finie ou à l'infini ?

Question 3. *Erreur de reprojection*

Supposons la même configuration (intrinsèques de la caméra, extrinsèques de la caméra, pose du robot) que dans la Question 1. Il y a une caractéristique connu sur le plan du sol ($z = 0$) à la position $x = 2$, $y = 4$ (dans le repère mondial). Il est détecté à l'emplacement du pixel :

$$x_{obs} = \begin{bmatrix} 370 \\ 300 \end{bmatrix}$$

Nous souhaitons calculer l'erreur de reprojection. Cela peut se faire selon les étapes suivantes :

(a) Transformer le point 3D du repère mondial vers le repère de la caméra.

(b) Projeter le point dans l'image.

(c) Calculer l'erreur de reprojection en pixels.

Question 4. *Intersection rayon-plan*

Une caméra calibrée montée sur un robot observe un point situé sur le plan du sol. Calculons les coordonnées 3D du point dans le repère mondial de deux manières différentes :

1. En intersectant le rayon avec le plan du sol
2. En calculant et en appliquant l'homographie inverse entre le plan du sol et le plan du capteur

Vous devriez obtenir la **même réponse** avec les deux méthodes.
Comme précédemment, supposons la matrice intrinsèque de la caméra :

$$K = \begin{bmatrix} 500 & 0 & 250 \\ 0 & 500 & 250 \\ 0 & 0 & 1 \end{bmatrix}$$

Le pixel qui détecte le point est :

$$p = \begin{bmatrix} 400 \\ 400 \end{bmatrix}$$

Cette fois, la caméra est située à 1 m au-dessus du plan du sol (et toujours orientée parallèlement au plan du sol). Supposons que le repère mondial soit placé sur le plan du sol directement sous la caméra. Nous allons calculer le point sur le plan du sol en utilisant les deux méthodes ci-dessus. Voici la répartition des tâches :

(a) Backprojecter le pixel pour obtenir le rayon visuel dans le repère de la caméra.

(b) Intersecter le rayon avec le plan du sol dans le repère de la caméra.

(c) Calculer la position du point du sol dans le repère mondial.

(d) Calculer l'homographie induite par le plan du sol vers l'image.

(e) Utiliser l'homographie inverse pour calculer la position du point du sol dans le repère mondial.

Question 5. *Gradients d'image et détection de coins Harris*

Considérons le patch d'image en niveaux de gris 5×5 suivant, normalisé entre 0 et 1 :

$$I = \begin{bmatrix} 0.1 & 0.1 & 0.1 & 0.1 & 0.1 \\ 0.1 & 0.5 & 0.5 & 0.1 & 0.1 \\ 0.1 & 0.5 & 0.9 & 0.1 & 0.1 \\ 0.1 & 0.1 & 0.1 & 0.1 & 0.1 \\ 0.1 & 0.1 & 0.1 & 0.1 & 0.1 \end{bmatrix}$$

(a) En utilisant un « filtre à différences centrales », calculez I_x et I_y pour les pixels intérieurs 3×3. Le filtre à différences centrales approxime le gradient en x comme : $\frac{\partial f}{\partial x} = \frac{f(x+1,y)-f(x-1,y)}{2}$, implémenté par le filtre $1/2 \begin{bmatrix} -1 & 0 & 1 \end{bmatrix}$, et $\frac{\partial f}{\partial y} = \frac{f(x,y+1)-f(x,y-1)}{2}$, implémenté par le filtre $1/2 \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$.

(b) Calculez la magnitude et l'orientation du gradient du pixel central.

(c) En utilisant une fenêtre uniforme 3×3 centrée sur le pixel central, calculez le tenseur de structure (matrice des moments du 2 ordre) pour ce pixel.

(d) Calculez la réponse Harris, R, du pixel central en supposant $\alpha = 0.04$. Classifiez le pixel central comme région plate, bord ou coin.