

TP2 – Structure-from-motion

Estimation du changement de pose

La matrice essentielle \mathbf{E} estimée à partir de deux images I_1 et I_2 d'une scène 3D est fonction du changement de pose de la caméra, caractérisé par un vecteur de translation $\mathbf{t}_{1/2}$ (coordonnées du centre de projection \mathbf{C}_1) et une matrice de rotation $\mathbf{R}_{1/2}$, tous deux exprimés dans le repère \mathcal{R}_2 :

$$\mathbf{E} = [\mathbf{t}_{1/2}]_{\wedge} \mathbf{R}_{1/2} \quad (1)$$

De ces deux éléments, $\mathbf{t}_{1/2}$ est le plus facile à estimer à partir de \mathbf{E} . En effet, la droite portée par les centres de projection \mathbf{C}_1 et \mathbf{C}_2 passe par les épipôles \mathbf{e}_1 et \mathbf{e}_2 , considérés comme des points 3D. Or, comme l'épipôle \mathbf{e}_2 n'est pas associé à une unique droite épipolaire de I_1 , ce qui signifie, d'après la contrainte épipolaire $\tilde{\mathbf{p}}_2^{\top} \mathbf{F} \tilde{\mathbf{p}}_1 = 0$:

$$\tilde{\mathbf{e}}_2^{\top} \mathbf{F} = 0 \quad \iff \quad \mathbf{F}^{\top} \tilde{\mathbf{e}}_2 = 0 \quad (2)$$

où $\tilde{\mathbf{e}}_2$ désigne le vecteur des coordonnées homogènes de \mathbf{e}_2 , exprimées dans le repère pixels. D'après la relation $\tilde{\mathbf{p}} = \mathbf{K} \mathbf{w}$, et sachant que $\mathbf{E}^{\top} = \mathbf{K}^{\top} \mathbf{F}^{\top} \mathbf{K}$, le point 3D défini par $\mathbf{w}_2 = \mathbf{K}^{-1} \tilde{\mathbf{e}}_2$, qui se trouve lui aussi sur la droite $(\mathbf{C}_1, \mathbf{C}_2)$, vérifie donc :

$$\mathbf{E}^{\top} \mathbf{w}_2 = 0 \quad (3)$$

ce qui signifie que \mathbf{w}_2 , et par conséquent $\mathbf{t}_{1/2}$, sont des vecteurs du noyau de \mathbf{E}^{\top} , dont nous verrons plus loin qu'il est de dimension 1. Dans la mesure où $\mathbf{t}_{1/2}$ peut être seulement estimé à un facteur près, nous imposons arbitrairement $\|\mathbf{t}_{1/2}\| = 1$, mais il existe encore deux solutions en $\mathbf{t}_{1/2}$ égales et opposées. Cette ambiguïté résiduelle sera levée dans l'exercice 2.

Si la matrice $[\mathbf{t}_{1/2}]_{\wedge}$ était inversible, il serait immédiat, connaissant \mathbf{E} et $[\mathbf{t}_{1/2}]_{\wedge}$, de déduire $\mathbf{R}_{1/2}$ de (1). Mais comme cela n'est pas le cas, il nous faut emprunter une autre voie pour déterminer $\mathbf{R}_{1/2}$.

Nous verrons plus loin que les deux valeurs singulières non nulles de \mathbf{E} sont égales. Sa décomposition SVD s'écrit donc sous la forme suivante :

$$\mathbf{E} = \pm \mathbf{U} \begin{bmatrix} \sigma & 0 & 0 \\ 0 & \sigma & 0 \\ 0 & 0 & 0 \end{bmatrix} \mathbf{V}^{\top} \quad (4)$$

où $\sigma > 0$ n'est pas connue a priori, puisque \mathbf{E} est définie à un facteur près.

La décomposition SVD (4) est utile à double titre. D'une part, la troisième colonne \mathbf{u}_3 de la matrice \mathbf{U} est un vecteur de $\text{Ker } \mathbf{E}^{\top}$ de norme unitaire. D'après ce qui précède, nous pouvons donc affirmer que :

$$\mathbf{t}_{1/2} = \pm \mathbf{u}_3 \quad (5)$$

D'autre part, la décomposition (4) permet de déduire $\mathbf{R}_{1/2}$ de \mathbf{E} . Comme \mathbf{U} est une matrice orthogonale, ses trois colonnes $(\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3)$ constituent une base orthonormée. Nous déduisons donc de (5) :

$$[\mathbf{t}_{1/2}]_{\wedge} \mathbf{U} = \pm [\mathbf{u}_3 \wedge \mathbf{u}_1, \mathbf{u}_3 \wedge \mathbf{u}_2, \mathbf{u}_3 \wedge \mathbf{u}_3] = \pm [\mathbf{u}_2, -\mathbf{u}_1, 0] \quad (6)$$

Par ailleurs, il est facile de vérifier l'identité suivante :

$$[\mathbf{u}_2, -\mathbf{u}_1, 0] = \mathbf{U} \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad (7)$$

Sachant que $\mathbf{U}^{-1} = \mathbf{U}^{\top}$, puisque \mathbf{U} est orthogonale, nous tirons de (6) et (7) :

$$[\mathbf{t}_{1/2}]_{\wedge} = \pm \mathbf{U} \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \mathbf{U}^{\top} \quad (8)$$

Une astuce du raisonnement consiste à remarquer que :

$$\begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \text{et} \quad \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (9)$$

En introduisant la matrice $\mathbf{W} = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$, qui est la matrice d'une rotation de $+\pi/2$ autour du 3^{ème} axe, nous tirons de (8) et (9) :

$$[\mathbf{t}_{1/2}]_{\wedge} = \mathbf{U} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \mathbf{W} \mathbf{U}^{\top} \quad \text{ou} \quad [\mathbf{t}_{1/2}]_{\wedge} = \mathbf{U} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \mathbf{W}^{\top} \mathbf{U}^{\top} \quad (10)$$

Comme les matrices \mathbf{U} et \mathbf{W} sont orthogonales, il en va de même des produits matriciels $\mathbf{W} \mathbf{U}^{\top}$ et $\mathbf{W}^{\top} \mathbf{U}^{\top}$. Les égalités (10) constituent donc deux décompositions SVD de la matrice $[\mathbf{t}_{1/2}]_{\wedge}$ (il est rappelé que la décomposition SVD d'une matrice n'est pas unique). En reportant ces décompositions dans (1), il vient :

$$\mathbf{E} = \mathbf{U} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \mathbf{W} \mathbf{U}^{\top} \mathbf{R}_{1/2} \quad \text{ou} \quad \mathbf{E} = \mathbf{U} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \mathbf{W}^{\top} \mathbf{U}^{\top} \mathbf{R}_{1/2} \quad (11)$$

qui constituent deux décompositions SVD de \mathbf{E} . Notons au passage que cela valide a posteriori la forme (4) de la décomposition SVD de \mathbf{E} . Par identification de (11) à (4), nous déduisons $\pm \mathbf{V}^{\top} = \mathbf{W} \mathbf{U}^{\top} \mathbf{R}_{1/2}$ ou $\pm \mathbf{V}^{\top} = \mathbf{W}^{\top} \mathbf{U}^{\top} \mathbf{R}_{1/2}$. Comme $\mathbf{W}^{-1} = \mathbf{W}^{\top}$, la matrice de rotation $\mathbf{R}_{1/2}$ s'écrit donc sous l'une des quatre formes suivantes :

$$\mathbf{R}_{1/2} = \pm \mathbf{U} \mathbf{W}^{\top} \mathbf{V}^{\top} \quad \text{ou} \quad \mathbf{R}_{1/2} = \pm \mathbf{U} \mathbf{W} \mathbf{V}^{\top} \quad (12)$$

Il n'existe en réalité que deux solutions en $\mathbf{R}_{1/2}$, car dans chacune des expressions (12), le signe doit être choisi de manière à garantir la contrainte $\det \mathbf{R}_{1/2} > 0$ (une matrice de rotation est forcément directe).

Exercice 1 : estimation du changement de pose

Importez le fichier `E_estimee.mat` du TP1, puis écrivez la fonction `estimation_4_poses`, appelée par le script `exercice_1`, censée calculer les quatre solutions $(\mathbf{t}_{1/2}, \mathbf{R}_{1/2})$ obtenues en combinant les deux solutions (5) en $\mathbf{t}_{1/2}$ et les deux solutions (12) en $\mathbf{R}_{1/2}$. Les paramètres de sortie `[t_4, R_4]` de cette fonction doivent être, respectivement, de dimensions 3×4 et $3 \times 3 \times 4$. Le script `exercice_1` teste si ces quatre solutions vérifient bien l'égalité (1), avec une certaine tolérance pour tenir compte de possibles écarts numériques.

Exercice 2 : détermination de la « bonne » pose

Importez la fonction `reconstruction_3D` et le fichier `donnees_appariees.mat` du TP1. Le script `exercice_2` affiche les quatre nuages de points 3D correspondant aux quatre couples $(\mathbf{t}_{1/2}, \mathbf{R}_{1/2})$. Parmi ces quatre nuages, un seul est acceptable. En effet, tout point 3D reconstruit doit se situer à l'avant de chacune des deux caméras.

Écrivez la fonction `estimation_pose`, appelée par le script `exercice_2`, qui détermine la « bonne » pose en comptabilisant le nombre de points 3D situés à l'arrière de l'une ou l'autre des caméras, pour chacun des quatre couples $(\mathbf{t}_{1/2}, \mathbf{R}_{1/2})$. Une fois cette solution affichée, vérifiez que la position des caméras soit bien cohérente avec le nuage de points 3D reconstruits.

Importez les fonctions `estimation_E` et `estimation_E_robuste` du TP1, puis lancez le script `sfm`. Ce script estime les $n - 1$ matrices essentielles correspondant aux $n - 1$ paires d'images $\{I_1, I_2\}, \dots, \{I_{n-1}, I_n\}$, pour un jeu de $n > 2$ images d'une même scène, puis calcule les $n - 1$ nuages de points 3D correspondant aux $n - 1$ paires d'images par la méthode « du point milieu » (cf. TP1).

Grâce à l'estimation des changements de pose de la caméra (cf. exercice 2), ces $n - 1$ nuages de points 3D peuvent être affichés dans un repère commun, en l'occurrence le repère caméra de la pose numéro n , sous différentes couleurs, grâce à la fonction `affichage_3D_melange`. Vous constatez, comme cela a été annoncé en cours, que les échelles de ces différents nuages de points 3D ne sont pas cohérentes entre elles.

Estimation de l'échelle a posteriori

La pose $(\mathbf{t}_{1/2}, \mathbf{R}_{1/2})$ a pu être déduite de la matrice essentielle \mathbf{E} par la relation $\mathbf{E} = [\mathbf{t}_{1/2}]_{\wedge} \mathbf{R}_{1/2}$, à ceci près que la norme de la translation $\mathbf{t}_{1/2}$ n'a pu être estimée, car \mathbf{E} est seulement connue à un facteur près. En choisissant comme solution $\mathbf{t}_{1/2} = \pm \mathbf{u}_3$, nous avons supposé implicitement que le centre optique de la caméra s'était déplacé d'une distance $\|\mathbf{t}_{1/2}\| = 1$, puisque \mathbf{u}_3 est la troisième colonne d'une matrice orthogonale.

Pour rendre cohérentes entre elles les échelles des nuages de points 3D reconstruits à partir des paires $\{I_1, I_2\}$ et $\{I_2, I_3\}$, il existe différentes méthodes. La plus simple consiste à estimer a posteriori la distance entre \mathbf{C}_2 et \mathbf{C}_3 , en projetant dans l'image I_3 les m points 3D du nuage reconstruit à partir de la paire $\{I_1, I_2\}$. Si les vecteurs \mathbf{Q}_2^k , $k \in \{1, \dots, m\}$, contiennent les coordonnées de ces points 3D, exprimées dans le repère \mathcal{R}_2 , et que $\alpha \mathbf{t}_{2/3}$ et $\mathbf{R}_{2/3}$ caractérisent la translation et la rotation de la caméra entre les poses 2 et 3, estimées dans le repère \mathcal{R}_3 , alors les coordonnées de ces points 3D s'écrivent, dans le repère \mathcal{R}_3 :

$$\mathbf{Q}_3^k = \mathbf{R}_{2/3} \mathbf{Q}_2^k + \alpha \mathbf{t}_{2/3}, \quad k \in \{1, \dots, m\} \quad (13)$$

Nous supposons que les points 3D reconstruits à partir de leurs projections dans les images I_1 et I_2 sont également visibles dans l'image I_3 . Leurs projections \mathbf{p}_3^k dans I_3 doivent donc vérifier les égalités suivantes :

$$\tilde{\mathbf{p}}_3^k = \frac{1}{Z_3(\mathbf{Q}^k)} \mathbf{K} (\mathbf{R}_{2/3} \mathbf{Q}_2^k + \alpha \mathbf{t}_{2/3}), \quad k \in \{1, \dots, m\} \quad (14)$$

où $\tilde{\mathbf{p}}_3^k = [u_3^k, v_3^k, 1]^T$ désigne le vecteur des coordonnées homogènes de la projection de \mathbf{Q}^k dans I_3 , exprimées dans le repère pixels, et $Z_3(\mathbf{Q}^k)$ la profondeur du point \mathbf{Q}^k relativement au repère \mathcal{R}_3 , c'est-à-dire la troisième composante du vecteur \mathbf{Q}_3^k . En développant ces équations vectorielles, il vient, $\forall k \in \{1, \dots, m\}$:

$$\begin{cases} u_3^k = \frac{1}{Z_3(\mathbf{Q}^k)} [\mathbf{K} (\mathbf{R}_{2/3} \mathbf{Q}_2^k + \alpha \mathbf{t}_{2/3})]_1 \\ v_3^k = \frac{1}{Z_3(\mathbf{Q}^k)} [\mathbf{K} (\mathbf{R}_{2/3} \mathbf{Q}_2^k + \alpha \mathbf{t}_{2/3})]_2 \\ 1 = \frac{1}{Z_3(\mathbf{Q}^k)} [\mathbf{K} (\mathbf{R}_{2/3} \mathbf{Q}_2^k + \alpha \mathbf{t}_{2/3})]_3 \end{cases} \quad (15)$$

où les indices 1, 2 et 3 désignent les trois composantes du vecteur entre crochets. En éliminant $Z_3(\mathbf{Q}^k)$ de (15), nous obtenons un système linéaire de $2m$ équations scalaires en α du type suivant, $k \in \{1, \dots, m\}$:

$$\begin{cases} u_3^k [\mathbf{K} (\mathbf{R}_{2/3} \mathbf{Q}_2^k + \alpha \mathbf{t}_{2/3})]_3 = [\mathbf{K} (\mathbf{R}_{2/3} \mathbf{Q}_2^k + \alpha \mathbf{t}_{2/3})]_1 \\ v_3^k [\mathbf{K} (\mathbf{R}_{2/3} \mathbf{Q}_2^k + \alpha \mathbf{t}_{2/3})]_3 = [\mathbf{K} (\mathbf{R}_{2/3} \mathbf{Q}_2^k + \alpha \mathbf{t}_{2/3})]_2 \end{cases} \quad (16)$$

système linéaire à une seule inconnue $\alpha \in \mathbb{R}$ de la forme $\mathbf{a} \alpha = \mathbf{b}$, où \mathbf{a} et \mathbf{b} sont des vecteurs de \mathbb{R}^{2m} .

Exercice 3 : recalage des $n - 1$ nuages de points 3D

Écrivez la fonction `estimation_echelle`, appelée par le script `exercice_3`, permettant de *recaler* les $n - 1$ nuages de points 3D rendus cohérents par ajustement des échelles a posteriori.

Remarque – À partir d'un triplet d'images $\{I_1, I_2, I_3\}$, certains points d'intérêt sont mis en correspondance seulement entre I_1 et I_2 , ou seulement entre I_2 et I_3 . Ces deux catégories de points d'intérêt permettent d'enrichir le nuage de points 3D, mais pas d'ajuster les échelles. C'est uniquement grâce aux points d'intérêt visibles dans les trois images qu'il est possible de rendre cohérentes entre elles les échelles des deux nuages de points 3D.