ROB317 - Vision 3D

TP1 Homographie

COELHO RUBACK Arthur

SANTOS SANO Matheus

Palaiseau, France

Septembre 2023

Estimation et correction d'une homographie

Question Q1

La transformation entre deux images est une homographie dans le cas d'une rotation pure autour d'un centre optique, en d'autres termes, les images d'une scène 3D sont capturées à partir du même point de vue. Dans ce cas, toutes les images sont liées par des homographies.

Un autre cas où la transformation est une homographie est celui où la même scène 3D est capturée dans deux poses 3D différentes, c'est-à-dire que les images sont capturées à partir de n'importe quel point de vue. Dans ce cas, l'image capturée d'un point de vue est liée au plan en 3D par homographie, et une autre image capturée d'un autre point de vue est également liée au plan par homographie. Il existe donc une transformation homographie entre ces deux images, même si elles ne sont pas prises du même point de vue.

Par exemple, une caméra ne tourne pas exactement autour de son centre de projection, mais elle est équivalente à un seul point de vue, car la scène est éloignée de la caméra. Si la scène est très éloignée, on peut supposer que la scène elle-même est un plan à l'infini et que, par conséquent, les homographies sont valables même si votre point de vue change quelque peu.

Question Q2

Comme l'homographie a huit degrés de liberté, au moins huit équations indépendantes sont nécessaires pour résoudre les huit éléments inconnus de la matrice d'homographie 3x3. Pour créer un système d'au moins huit équations, il faut au moins quatre paires de points correspondants (deux équations par paire). S'il y a moins de quatre paires de points, le système d'équations est sous-déterminé et ne contient pas suffisamment d'informations pour déterminer la matrice d'homographie. De cette façon, le nombre minimum de points que on doit sélectionner dans l'image est de quatre.

Question Q3

Dans ce contexte, l'intérêt de ce traitement est de corriger la distorsion géométrique de l'image originale. Comme l'angle de vue de l'image de la mosaïque "*Pompei.jpg*" n'étant pas en vue verticale, la transformation homographique est mise en œuvre pour corriger les distorsions géométriques générées par cet angle de vue. Après traitement, l'image résultante présente un angle de vision perpendiculaire.

Pour ce traitement, on considère $N \geq 4$ correspondances de points 2D à 2D (X_{init}, X_{final}). Les coordonnées des points initiaux et finaux ont été normalisées afin d'éviter d'éventuels grands écarts de taille entre les images initiales et finales :

$$\begin{cases} \tilde{X}_{init} = T_{norm} X_{init} \\ \tilde{X}_{final} = T_{norm} X_{final} \end{cases}$$

Soit w le nombre de colonnes et h le nombre de lignes de l'image initiale, la matrice T_{norm} est :

$$T_{norm} = \begin{bmatrix} 1/w & 0 & -1 \\ 0 & 1/h & -1 \\ 0 & 0 & 1 \end{bmatrix}$$

Après le calcul de la matrice d'homographie $\tilde{H}$ avec les coordonnées normalisées, il faut dé-normaliser la matrice $\tilde{H}$:

$$H = T_{norm}^{-1} \tilde{H} T_{norm}$$

Une fois la matrice H obtenue, on l'applique sur l'image initiale, obtenant une image finale avec une vue verticale de la mosaïque. Une démonstration de ce traitement est présentée ci-dessous.

Soit les points sélectionnés par l'image initiale (X_{init}) et leurs points correspondants (X_{final}) présentés dans le Tableau 1.

X_{init}	X_{final}
(144 ; 48)	(50 ; 50)
(111 ; 254)	(50 ; 250)
(369 ; 262)	(250 ; 250)
(355 ; 46)	(50 ; 250)

TABLE 1 – Points sélectionnés par l'image initiale X_{init} et finale X_{final}.

A partir de ces points, la matrice d'homographie H a été calculée et appliquée sur l'image initiale. Le résultat de l'homographie mise en œuvre est présenté dans la Figure 1.

(a) Image avant l'homographie. (b) Image après l'homographie.

FIGURE 1 – Transformation d'image à partir de la matrice d'homographie estimée.

L'indice de confiance qui peut être utilisé pour évaluer la fiabilité du résultat fourni est de comparer la matrice d'homographie implémentée avec la matrice générée par la fonction *cv2.getPerspectiveTransform* de la bibliothèque "*cv2*". Cette fonction génère la matrice d'homographie après avoir reçu en entrée les points initiaux (X_{init}) et les points finaux (X_{final}). Dans notre code, la matrice H implémentée et celle obtenue par la fonction OpenCV sont les mêmes :

$$H = \begin{bmatrix} 1,0885 & 0,2324 & -113,4068 \\ 0,0235 & 1,3522 & -13,7947 \\ 0,0002 & 0,0012 & 1 \end{bmatrix}$$

Création de panorama

Quesiton Q4

Après avoir effectué le calcul de la matrice d'homographie et vérifié qu'il est correct, il devient possible de créer un panorama à partir de deux images. Toutefois, afin de créer le panorama, il est essentiel de respecter une condition : les deux images sont générées à partir d'une rotation pure autour du même centre optique. En d'autres termes, les deux images doivent être prises depuis le même plan de vue, la hauteur et l'orientation de la caméra.

En outre, pour superposer des images, il est nécessaire que les deux images aient des points communs afin qu'elles puissent être utilisées pour calculer l'homographie. Une fois la matrice d'homographie calculée, la transformation d'homographie est appliquée à l'une des images. On superpose ensuite l'image transformée et l'autre image qui reste inchangée. De cette façon, le panorama est réalisé en fusionnant les deux images.

Par ailleurs, il est possible d'obtenir les paramètres de transformation à partir de la matrice d'homographie, en particulier la rotation, la translation et la mise à l'échelle. La translation est obtenue à partir des éléments de la matrice d'homographie. La rotation R, à son tour, peut être extraite à partir de la décomposition de la matrice d'homographie en valeurs singulières :

$$[U, S, V] = svd(H)$$
$$R = VU^T$$

Un exemple de panorama est présenté dans la Figure 2, où l'image de droite est transformée et superposée à l'image de gauche, qui reste inchangée. Il convient de noter que 4 points ont été sélectionnés manuellement dans chaque image.

Figure 2 – Panorama à partir de deux images avec sélection manuelle des points.

Comme le montre la Figure 2, le panorama a été bien créé en sélectionnant manuellement 4 points. Plus la sélection des points est précise, mieux l'image transformée est ajustée par rapport à l'image de référence. Comme le montre la Figure 2, l'image transformée (à droite) est légèrement déformée par rapport à l'image de référence (à gauche). Cela s'explique par le fait que la sélection des points n'était pas très précise, puisqu'elle a été effectuée manuellement.

Une solution consiste à utiliser l'algorithme SIFT (Scale-invariant feature transform) pour la sélection précise de points. Cet algorithme consiste à localiser les points clés de l'image où se produisent des changements significatifs dans l'intensité des pixels. Une autre approche qui améliore l'image est l'algorithme RANSAC (RANdom SAmple Consensus) qui calcule la matrice d'homographie. Cet algorithme consiste à estimer des modèles mathématiques à partir d'un ensemble de données et, surtout, supprimer les points qui ne correspondent pas au modèle. Ainsi, en appliquant les algorithmes SIFT et RANSAC, il est possible d'obtenir un meilleur panorama que celui présenté dans la Figure 2, où la matrice d'homographie est implémentée et où les points sont sélectionnés manuellement.

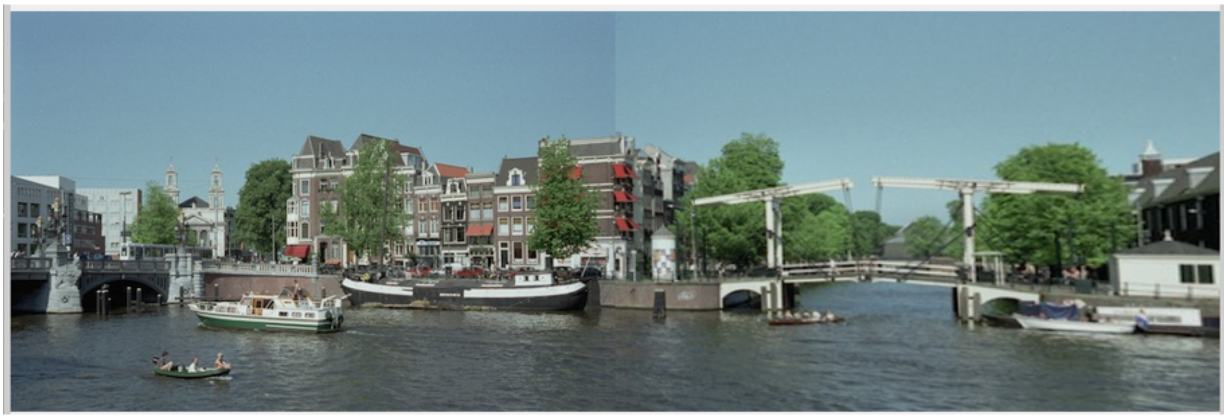

FIGURE 3 – Panorama à partir de deux images avec les algorithmes SIFT et RANSAC.

Dans la Figure 3, on peut voir que l'image transformée (à droite) ne présente plus les déformations de la Figure 2, ce qui montre l'amélioration apportée par les deux algorithmes. Cependant, on peut constater aussi que le panorama de la Figure 3 est très similaire à celui de la Figure 2, ce qui montre également que le calcul de l'homographie H et la construction du panorama sont bien implémentés.