



ECOLE NATIONALE SUPÉRIEURE DE TECHNIQUES AVANCÉES

Vision 3D

TP Homographies

Davy Araujo Sa Teles

Palaiseau, France 2025

Table des matières

1. Question 1	3
2. Question 2	3
3. Question 3	3
4. Question 4	4

1. Question 1



Figure 1: Homographie

Le premier cas survient lorsque la scène observée est parfaitement plane (comme un tableau, un mur, un terrain de sport ou le sol vu de très haut) ; dans cette configuration, tous les points du monde réel appartiennent au même plan, ce qui permet une correspondance mathématique directe avec le plan de l'image, quel que soit le mouvement de la caméra. Le second cas se présente lorsque la caméra effectue une rotation pure autour de son centre optique (mouvement panoramique sur un trépied) ou un simple changement de focale (zoom), sans aucune translation ; ici, la distance ou le relief des objets n'a aucune importance car, le point de vue restant fixe dans l'espace, les rayons lumineux entrent sous les mêmes angles relatifs, permettant d'aligner les images parfaitement via une matrice 3×3 .

2. Question 2

Il faut donc au minimum 4 paires de points correspondants pour résoudre le système. Ces 4 points ne doivent pas être colinéaires (alignés), sinon le système est dégénéré et ne peut pas être résolu.

3. Question 3

Ce traitement s'appelle une Rectification de Perspective (ou correction de perspective). L'intérêt principal est de passer d'une géométrie projective (celle de l'œil ou de la caméra, où les lignes parallèles se rejoignent vers un point de fuite) à une géométrie euclidienne (celle d'une carte ou d'un plan, où les parallèles restent parallèles).

Applications: La Métrologie, La Reconnaissance de Caractères, La Réalité Augmentée et Mosaïque

Pour avoir une confiance maximale, il faut que l'erreur de reprojection soit faible (pixels) et que vos 4 points soient bien répartis aux 4 coins de la zone d'intérêt (polygone convexe le plus grand possible).

L'objectif central de l'homographie dans ce contexte est le passage de la **géométrie projective** (vue caméra avec fuites) à la **géométrie euclidienne** (vue métrique/plane).

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N \|\mathbf{x}'_i - \text{norm}(H\mathbf{x}_i)\|^2} \quad (1)$$

Où :

- ❖ \mathbf{x}_i : Coordonnées des points dans l'image originale (source).
- ❖ \mathbf{x}'_i : Coordonnées des points dans le plan rectifié (destination).
- ❖ H : Matrice d'Homographie estimée.
- ❖ $\text{norm}(\cdot)$: Normalisation par la coordonnée homogène (division par w).

Critère de Validité : Un $RMSE$ faible (proche de 0 pixel) indique une haute cohérence mathématique. De plus, la confiance géométrique dépend du *conditionnement spatial* : les points sélectionnés ne doivent pas être colinéaires et doivent couvrir la plus grande surface possible de l'objet (polygone convexe large).

4. Question 4

Pour que l'homographie fonctionne parfaitement dans un panorama, il faut idéalement que le centre optique de la caméra ne bouge pas (on tourne la caméra sur elle-même, on ne se déplace pas latéralement). Sinon, les objets proches bougent plus vite que les objets lointains qui rend l'assemblage impossible sans artefacts.

Q5 : Retrouver les paramètres de la transformation à partir de l'homographie

Une matrice d'homographie H (3×3) encode à la fois les paramètres intrinsèques de la caméra et le mouvement géométrique. Nous pouvons l'analyser selon deux niveaux de lecture :



Figure 2: Panorama

1. Interprétation Intuitive (Matricielle)

Si l'on décompose la matrice H :

$$H = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix} \quad (2)$$

- ❖ **Translation** (T_x, T_y) : Les termes h_{13} et h_{23} représentent le déplacement horizontal et vertical (exprimé en pixels si $h_{33} = 1$).
- ❖ **Rotation et Échelle** : Le bloc supérieur gauche 2×2 $\begin{pmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{pmatrix}$ encode la rotation dans le plan, le facteur de zoom et le cisaillement (*shear*).
- ❖ **Perspective** : Les termes h_{31} et h_{32} sont responsables des effets de fuite (trapèze). S'ils sont proches de 0, la transformation est affine.

2. Interprétation Rigoureuse (Décomposition)

Pour isoler la rotation physique R et la translation T , il est nécessaire de connaître la matrice de calibration intrinsèque K de la caméra. Dans le cas d'une rotation pure (panorama), la relation est :

$$H = K \cdot R \cdot K^{-1} \implies R = K^{-1} \cdot H \cdot K \quad (3)$$

Q6 : Stratégie d'uniformisation des couleurs

Les différences d'exposition ou de vignettage entre deux images créent une "cou-ture" visible (*seam*) lors de l'assemblage. Pour y remédier, nous proposons deux approches complémentaires :

1. Compensation de Gain (Gain Compensation)

Avant l'assemblage, on ajuste l'intensité globale pour que les zones communes aient la même luminosité moyenne. On calcule un facteur g tel que :

$$g = \frac{\mu_{I_1}}{\mu_{I_2}} \quad \text{où } \mu \text{ est la moyenne des pixels dans la zone de chevauchement.} \quad (4)$$

2. Fusion Linéaire (Alpha Blending)

Au lieu d'une coupure nette, on définit une zone de transition progressive. On utilise un masque de pondération $\alpha(x)$ variant de 1 à 0 à travers la jonction :

$$I_{final}(p) = \alpha(p) \cdot I_{gauche}(p) + (1 - \alpha(p)) \cdot I_{droite}(p) \quad (5)$$

Cette méthode, aussi appelée *Feathering*, lisse les discontinuités et rend la jonction invisible à l'œil nu.