Task 1

Integrantes

- Sergio Orellana 221122
- Rodrigo Mansilla 22611
- Ricardo Chuy 221007

1. Una homografía H es una matriz de 3 × 3. Explique matemáticamente por qué, aunque tiene 9 elementos, solo posee 8 grados de libertad (GDL)

Una homografía actúa sobre puntos en coordenadas homogéneas $x \in \mathbb{P}^2$ mediante la relación de equivalencia

$$x' \sim Hx,$$

donde $x' \sim y$ significa que existen escalares no nulos $\lambda \neq 0$ tales que $x' = \lambda y$.

En consecuencia, si multiplicamos H por cualquier escalar no nulo $\alpha \neq 0$, la transformación inducida en el plano proyectivo no cambia:

$$x' \sim Hx \quad \Longrightarrow \quad x' \sim (\alpha H)x.$$

Para verlo de forma explícita en coordenadas euclidianas, sea $x = (u, v, 1)^\top$ y

$$Hx = \begin{bmatrix} p \\ q \\ r \end{bmatrix}, \qquad r \neq 0.$$

La proyección a coordenadas no homogéneas es

$$\pi(Hx) = \left(\frac{p}{r}, \frac{q}{r}\right).$$

Ahora, si usamos αH:

$$(\alpha H)x = \alpha \begin{bmatrix} p \\ q \\ r \end{bmatrix} \quad \Rightarrow \quad \pi((\alpha H)x) = \left(\frac{\alpha p}{\alpha r}, \frac{\alpha q}{\alpha r}\right) = \left(\frac{p}{r}, \frac{q}{r}\right) = \pi(Hx).$$

Por consiguiente, la homografía no es un objeto con 9 parámetros "observables", sino una clase de equivalencia $[H]$ definida hasta escala. Dicho de otro modo, H vive en el espacio proyectivo $\mathbb{P}^8$, así que sus grados de libertad son

$$9 \text{ parámetros } - \text{ } 1 \text{ escala } = \text{ } 8 \text{ GDL}.$$

a. Adicionalmente, respoda. Si tuviéramos una cámara que solo rota sobre su eje

óptico (sin traslación ni cambio de perspectiva), ¿la matriz de transformación sigue teniendo 8 GDL o se reduce? Demuestre la estructura de dicha matriz simplificad

Si la cámara solo rota alrededor de su eje óptico, el movimiento es una rotación 3D $R_z(\theta)$ (sin traslación). En un modelo pinhole con intrínsecos constantes K, la homografía entre dos vistas viene dada por

$$H = KRK^{-1}.$$

En particular, para una rotación alrededor del eje óptico:

$$R_z(\theta) = \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Si, además, tomamos un caso estándar sin skew con f y punto principal (c_x, c_y),

$$K = \begin{bmatrix} f & 0 & c_x \\ 0 & f & c_y \\ 0 & 0 & 1 \end{bmatrix}, \qquad K^{-1} = \begin{bmatrix} \frac{1}{f} & 0 & -\frac{c_x}{f} \\ 0 & \frac{1}{f} & -\frac{c_y}{f} \\ 0 & 0 & 1 \end{bmatrix}.$$

Entonces

$$H = KR_z(\theta)K^{-1} = \begin{bmatrix} \cos\theta & -\sin\theta & c_x(1-\cos\theta) + c_y\sin\theta \\ \sin\theta & \cos\theta & -c_x\sin\theta + c_y(1-\cos\theta) \\ 0 & 0 & 1 \end{bmatrix}.$$

En contraste con una homografía general (8 GDL), aquí toda la transformación queda parametrizada únicamente por θ cuando K es fijo. Por lo tanto, los grados de libertad del movimiento (y de la homografía inducida) se reducen a 1 GDL.

Por ejemplo, si trabajamos en un sistema de coordenadas cuyo origen está en el punto principal (es decir, $c_x = c_y = 0$), la matriz se simplifica aún más:

$$H = \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Aun cuando H siga siendo una matriz 3×3, la restricción geométrica “solo rotación sobre el eje óptico” elimina prácticamente todos los grados de libertad de una homografía arbitraria, dejando solo el ángulo de rotación.

2. En el algoritmo DLT (Direct Linear Transform), convertimos el problema $x' = Hx$ en un sistema de la forma $Ah = 0$. Explique por qué buscamos el vector singular asociado al menor valor singular de A en lugar de simplemente invertir la matriz. ¿Qué representa geométricamente ese "menor valor singular" cuando los datos tienen ruido?

En DLT, partimos de la relación proyectiva $x' \sim Hx$. Una manera estándar de eliminar la escala desconocida es usar el producto cruz:

$$x' \times (Hx) = 0.$$

Esta ecuación genera restricciones lineales en los elementos de H (apilados en $h = \mathrm{vec}(H)$), de modo que obtenemos

$$Ah = 0,$$

donde típicamente $A \in \mathbb{R}^{2N \times 9}$ si usamos N correspondencias (dos ecuaciones independientes por punto).

En primer lugar, **no “invertimos”** A porque:

1. A no es cuadrada en general ($2N \times 9$), así que no existe una inversa clásica.
2. Incluso si fuera cuadrada en algún caso particular, el sistema es homogéneo: h solo se define hasta escala ($h \sim \alpha h$), por lo que hablar de “solución por inversión” no es el enfoque correcto.
3. Con ruido, el sistema **no tiene solución exacta** $Ah = 0$.

En lugar de eso, buscamos el vector h que minimiza el error algebraico bajo una normalización (para evitar la solución trivial $h = 0$). Es decir:

$$\min_{||h||=1} ||Ah||.$$

Si hacemos la descomposición en valores singulares (SVD)

$$A = U\Sigma V^\top,$$

y escribimos $h = Vz$ con $||h|| = ||z||$, entonces

$$||Ah|| = ||U\Sigma V^\top Vz|| = ||\Sigma z||.$$

La norma $||\Sigma z||$ se minimiza eligiendo z como el vector canónico asociado al **menor** valor singular $\sigma_{\min}$, es decir, h es la **última columna de** V.

Geométricamente, cuando hay ruido, $\sigma_{\min}$ mide “qué tan cerca” está A de ser rango-deficiente (el caso ideal donde existe un nullspace exacto de dimensión 1). En otras palabras, $\sigma_{\min}$ cuantifica el residual mínimo posible:

$$\sigma_{\min} = \min_{||h||=1} ||Ah||.$$

Por consiguiente:

- si los datos fueran perfectos, esperaríamos $\sigma_{\min} \approx 0$ (nullspace casi exacto);
- sin embargo, con ruido, $\sigma_{\min} > 0$ y su magnitud refleja la inconsistencia introducida por el ruido y/o una mala configuración geométrica (por ejemplo, puntos casi colineales o mala normalización).

En síntesis, el "menor valor singular" es el indicador de la mejor compatibilidad algebraica de una homografía con los datos observados bajo ruido.

3. Si usted selecciona 4 puntos para calcular H, pero 3 de ellos son colineales (están en la misma línea recta), el algoritmo fallará. Explique algebraicamente qué le sucede a la matriz A del sistema DLT en este caso y por qué no tiene solución única.

En el caso ideal (sin degeneración), con 4 correspondencias obtenemos 8 ecuaciones independientes y, típicamente, $\operatorname{rank}(A) = 8$, de modo que el nullspace de A tiene dimensión 1 y define h de forma única (hasta escala).

Ahora supongamos que tres puntos de entrada x_1, x_2, x_3 son colineales. Entonces existe una recta l tal que

$$l^\top x_i = 0 \quad \text{para } i \in \{1, 2, 3\}.$$

De igual forma consideremos cualquier vector $u \in \mathbb{R}^3$ y defina una familia de matrices

$$H(\tau) = H + \tau\, u\, l^\top, \qquad \tau \in \mathbb{R}.$$

Para los puntos colineales:

$$H(\tau)x_i = Hx_i + \tau\, u\, (l^\top x_i) = Hx_i + \tau\, u \cdot 0 = Hx_i.$$

Es decir, **todas** las matrices $H(\tau)$ inducen exactamente la misma proyección sobre esos tres puntos (porque el término añadido se anula en la recta l). Por lo tanto, las ecuaciones que esos puntos generan en $Ah = 0$ no "ven" ciertas combinaciones de parámetros: hay infinitas soluciones compatibles con esas tres correspondencias.

Aunque el cuarto punto x_4 agrega restricciones adicionales, en general no alcanza para eliminar toda la ambigüedad creada por la colinealidad. En consecuencia, a nivel matricial ocurre que:

- varias filas de A se vuelven linealmente dependientes (o casi dependientes si hay ruido),
- $\operatorname{rank}(A)$ cae por debajo de 8,
- el nullspace de A pasa a tener dimensión mayor que 1 (no hay unicidad).

Una señal clara en SVD es que, en lugar de un único valor singular cercano a cero, aparecen **varios** valores singulares pequeños, reflejando múltiples direcciones h con residuales muy parecidos:

$$Ah \approx 0 \quad \text{para varios } h \text{ distintos.}$$

En síntesis, la colinealidad hace que el sistema DLT quede subdeterminado: la matriz A pierde rango y, por lo tanto, no existe una solución única para la homografía.

Referencias

- 41 Homographies – Foundations of Computer Vision. (n.d.). https://visionbook.mit.edu/homography.html
- Hartley, R., & Zisserman, A. (2004). Multiple view geometry in computer vision. In Cambridge University Press eBooks. https://doi.org/10.1017/cbo9780511811685
- OpenCV: Basic concepts of the homography explained with code. (n.d.). https://docs.opencv.org/4.x/d9/dab/tutorial_homography.html