UVG UNIVERSIDAD DEL VALLE DE GUATEMALA

ING | FACULTAD DE INGENIERÍA

Geometría Proyectiva y Alineación

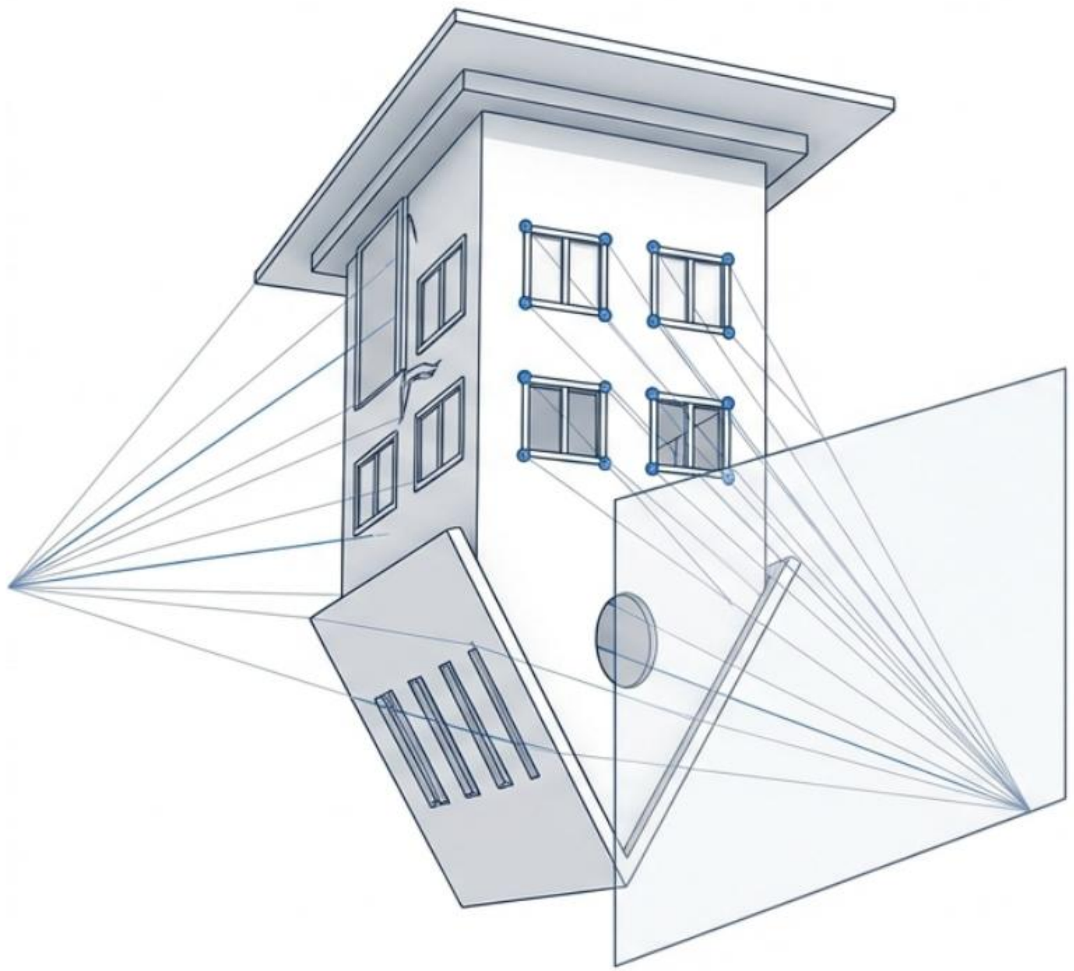
Visión Por Computadora

2026

Introducción

- **Meta:** Mapear el mundo 3D a 2D.
- **Problema:** La geometría euclidiana es insuficiente.
- **Herramienta:** Coordenadas Homogéneas y Matrices.
- **Modelo:** Cámara Pinhole (Intrínsecos/Extrínsecos).
- **Transformación:** Homografías y DLT.
- **Algoritmo:** Estimación robusta con RANSAC.
- **Aplicación:** Panoramas, SLAM, Realidad Aumentada.

El Desafío la Perspectiva



$$x' = Hx$$

El mundo es tridimensional, pero nuestros sensores son planos 2D. Para reconstruir la estructura o unir imágenes, necesitamos mapear puntos de un plano a otro matemáticamente.

El Problema Proyectivo

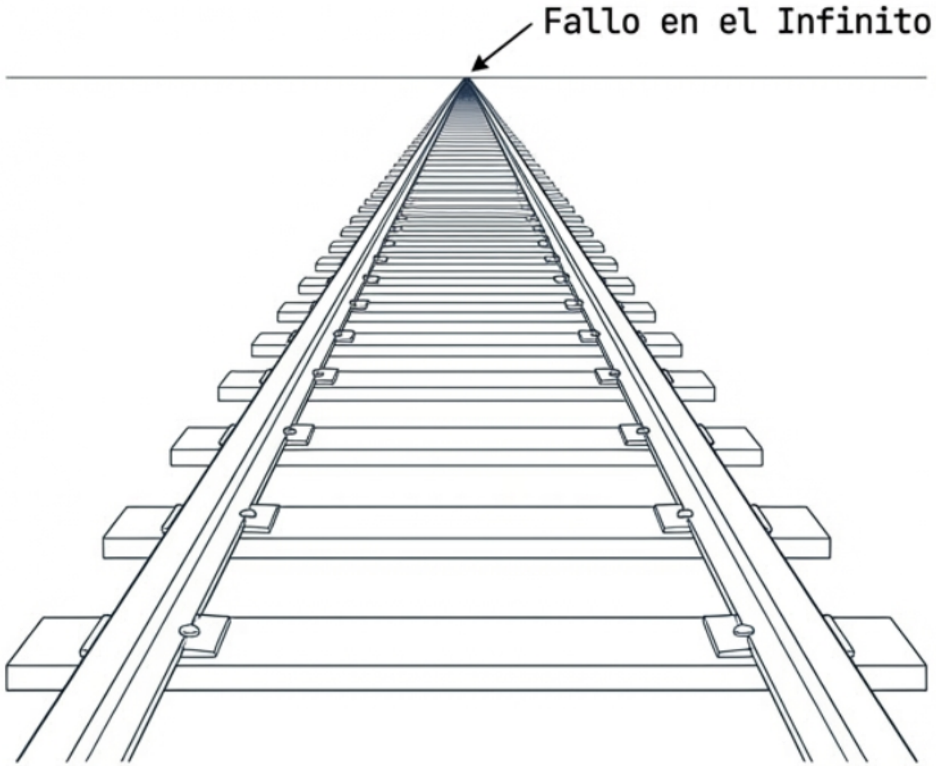
- La geometría euclidiana conserva longitudes y ángulos.
- Las cámaras **no** conservan estas propiedades.
- Las líneas paralelas convergen en puntos de fuga.
- Los objetos se hacen más pequeños a medida que se alejan.
- Necesitamos un nuevo marco matemático.
- La geometría proyectiva modela el proceso de obtención de imágenes.
- Maneja puntos en el infinito de forma natural.

El Espacio Proyectivo P**2

- Punto Euclideano $(x, y) \in \mathbb{R}^2$.
- Punto Homogéneo $\mathbf{x} = [\quad , x, y, w]^T \in \mathbb{P}^2$,.
- Conversión: Dividir por $w \to \left(\quad , \frac{x}{w}, y/w\right)$.
- **Invariancia de Escala:** $\mathbf{x} \sim k\mathbf{x}$ para $k \neq 0$.
- **Puntos al Infinito:** Ocurren cuando $w = 0$.
- Permite representar transformaciones como matrices lineales.
- Simplifica: Proyección se vuelve multiplicación matricial.

El Lenguaje: Coordenadas Homogéneas

Euclidiano



Homogéneo

$$(x, y) \longrightarrow \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

- Añadimos una dimensión extra (factor de escala)
- (x, y) ⇒ (x, y, 1)
- Convierte divisiones en multiplicaciones de matrices
- Invarianza de escala: k(x,y,w) ~ (x,y,w)

Euclideanas vs Homogéneas

Euclidean

$$x' = x + t$$

Non-Linear (Addition)

Homogeneous

$$(x,\ y) \rightarrow (x,\ y,\ 1)$$

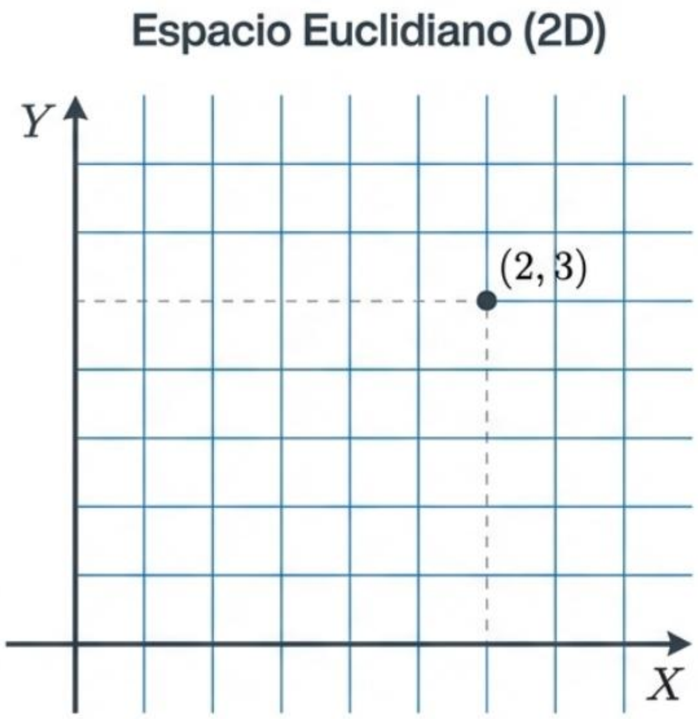
$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & t_x \\ 0 & 1 & t_y \\ 0 & 0 & 1 \end{bmatrix} \times \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

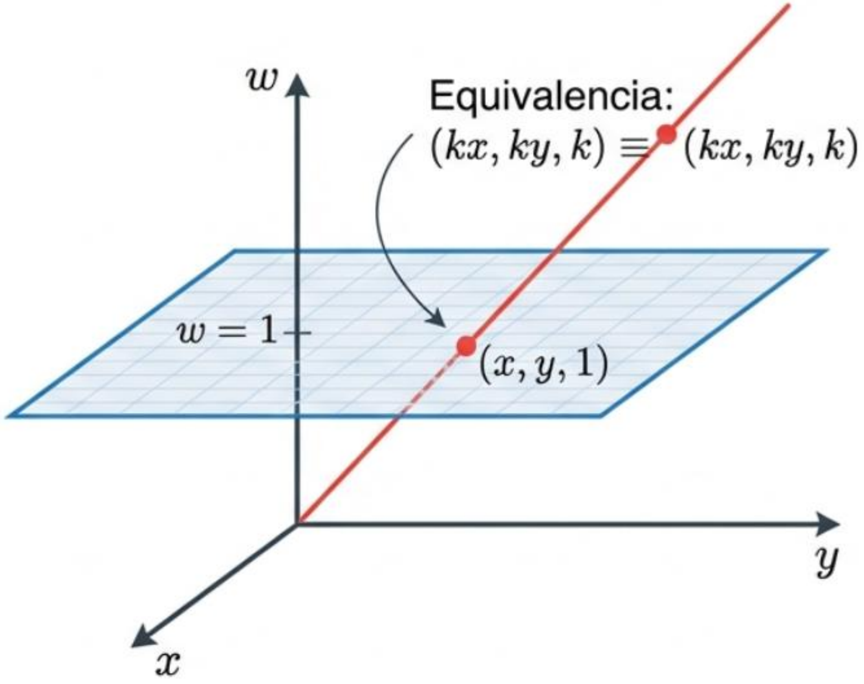
La Solución: Aumentar una dimensión unifica rotación y traslación en una sola operación: **Multiplicación de Matrices.**

Coordenadas Homogeneas

- Las coordenadas estándar 2D son (x, y).
- Las escalas homogenas agregan escala: (x, y, w).
- Se convierte al dividir: $(\frac{x}{w}, \frac{y}{w})$.
- Permite representar matrices de traslados
- Los puntos en el infinito tienen $w = 0$.
- Forma estándar: $\tilde{x} = [\quad , x, y, 1]^T,$.
- Invariante de escala: $k[x, y, w] \equiv [x, y, w]$

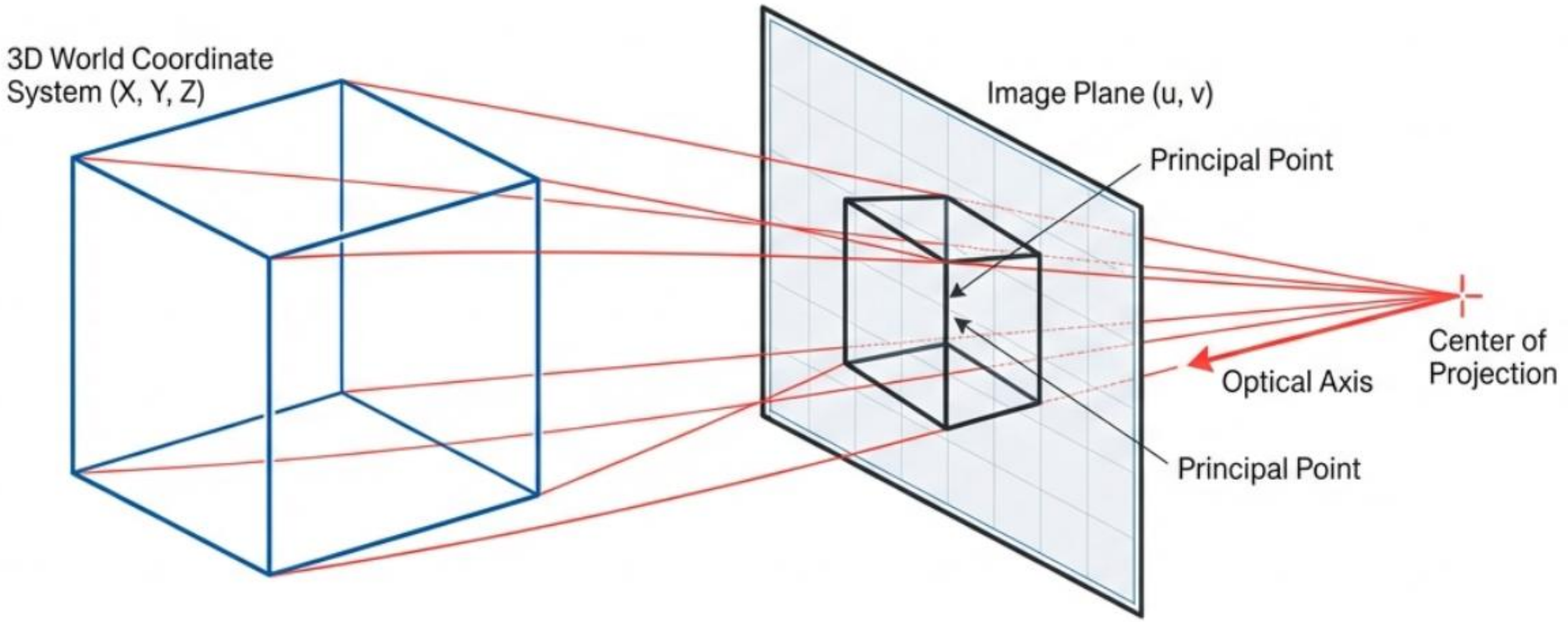
Coordenadas Homogéneas: La clave Proyectiva





Añadimos una dimensión extra. Esto linealiza la proyección perspectiva y nos permite representar puntos en el infinito (donde w=0).

De la Geometría 3D a la Imagen Digital



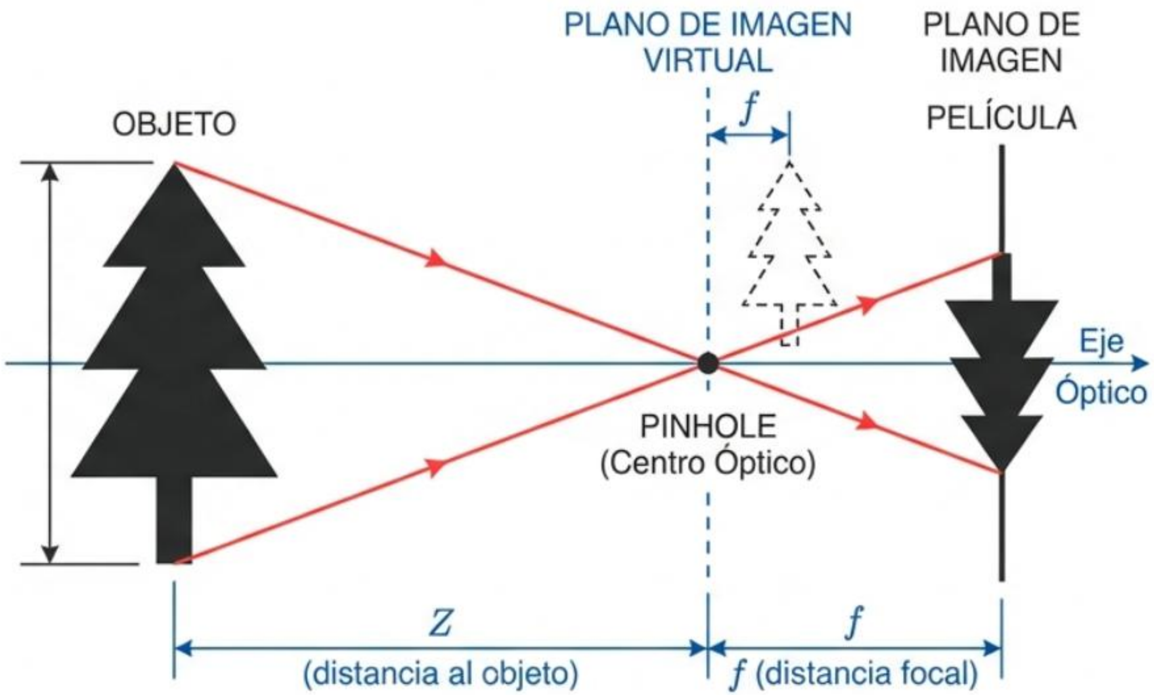
Una exploración técnica del modelo de cámara estenopeica y la solución algebraica para la estimación de parámetros.

El Modelo Pinhole (Cámara Estenopeica)

- La forma más simple de capturar la realidad es la "Cámara Oscura". La luz pasa por un agujero minúsculo, proyectando una imagen invertida en el plano opuesto. Utilizamos la **regla de los triángulos semejantes** para modelar esto.

$$-\frac{x}{f} = \frac{X}{Z} \rightarrow x = -f\left(\frac{X}{Z}\right)$$

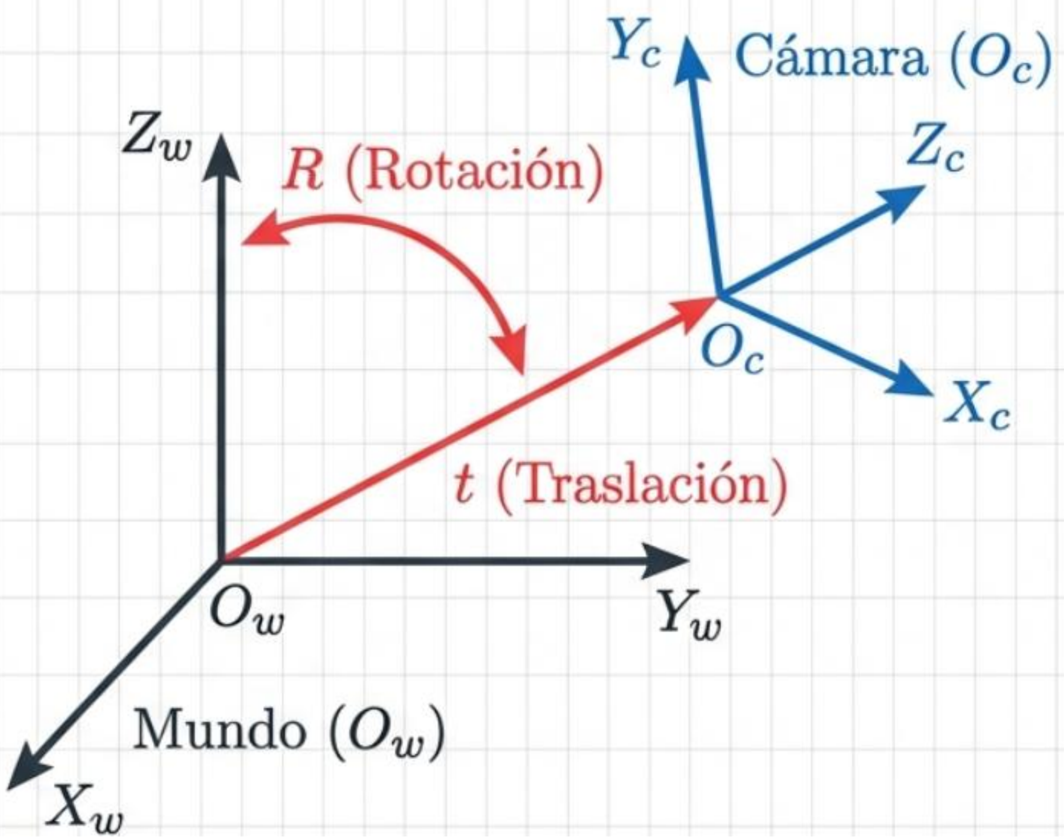
- El signo negativo indica la inversión de la imagen



Modelo de Cámara Pinhole

- La luz pasa a través de una pequeña abertura.
- Proyecta una imagen invertida sobre el plano del sensor.
- Mapea coordenadas de cámara 3D a pixeles.
- Ecuación:
  - $\mathbf{x} = K[R \mid t]\mathbf{X}$
  - $\mathbf{x} = K[I \mid \mathbf{0}]\mathbf{X}_{cam}$.
- **Intrínsecos (K):** Distancia focal y centro..
  - K describe las propiedades internas de la cámara.
- **Extrínsecos (R, t):** Rotación y traslación.
  - [R|t] describe la posición de la cámara en el mundo.
- K es triangular superior de 3×3.
- f_x, f_y :Distancia focal en unidades de pixel.
- (c_x, c_y) :Punto principal (centro óptico).
- s: Factor de "skew" (usualmente 0).
- Define la geometría interna del sensor.

Parámetros Extrínsecos (Posición y Orientación)



Definen **dónde** está la cámara.
6 Grados de Libertad.

- **Rotación** (R): Matriz 3×3 que alinea la orientación.
- **Traslación** (t): Vector 3×1 de posición.

Matriz Combinada: $[R \mid t]$

Parámetros Intrínsecos (La Matriz K)

Distancia Focal (zoom/escala en píxeles)

$$K = \begin{bmatrix} f_x & s & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix}$$

Punto Principal (centro óptico)

Skew (sesgo de ejes)

Definen **cómo** la cámara forma la imagen.
Son parámetros internos de la lente y el sensor.

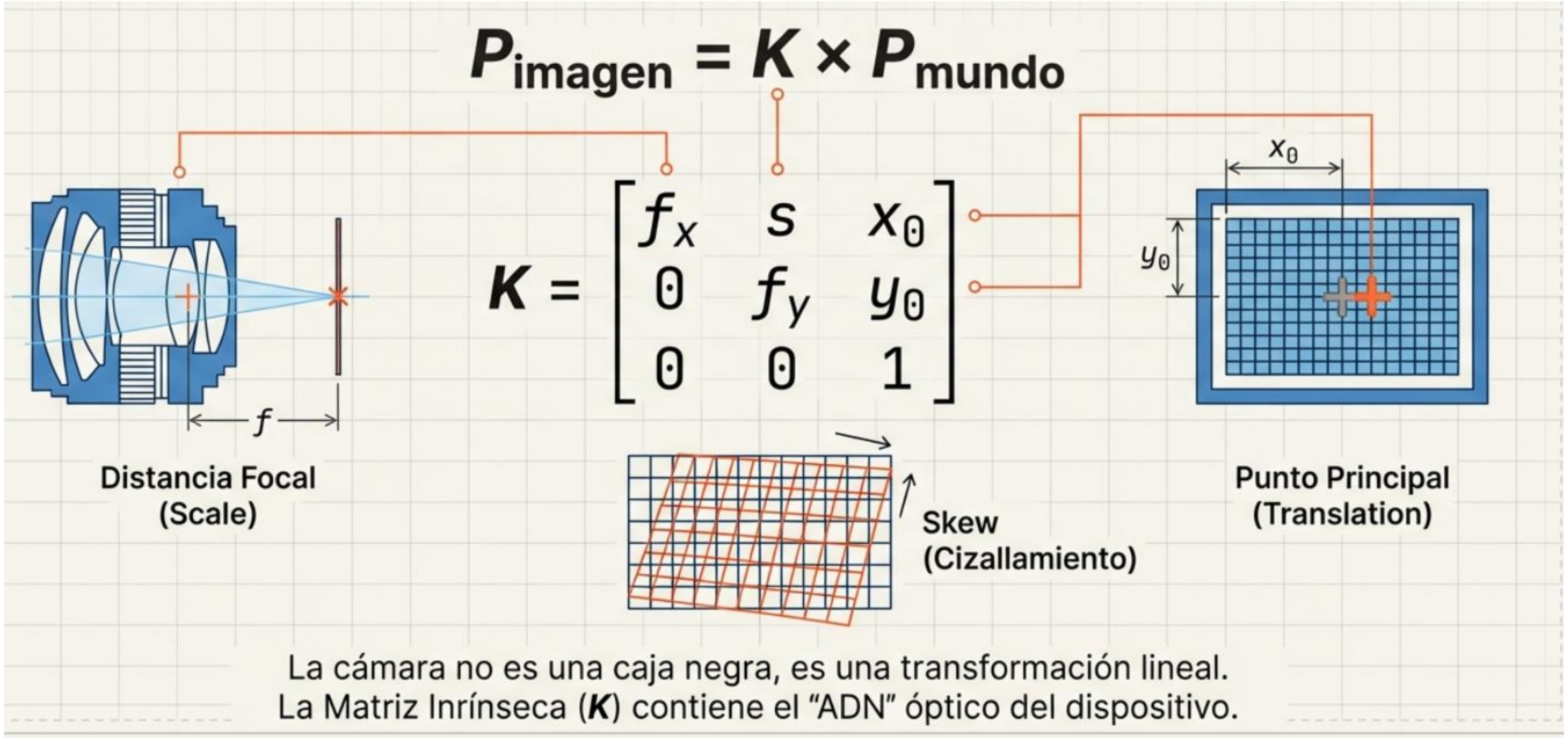
La Matriz de Proyección P

$$\boldsymbol{x} = \boldsymbol{P}\boldsymbol{X}$$

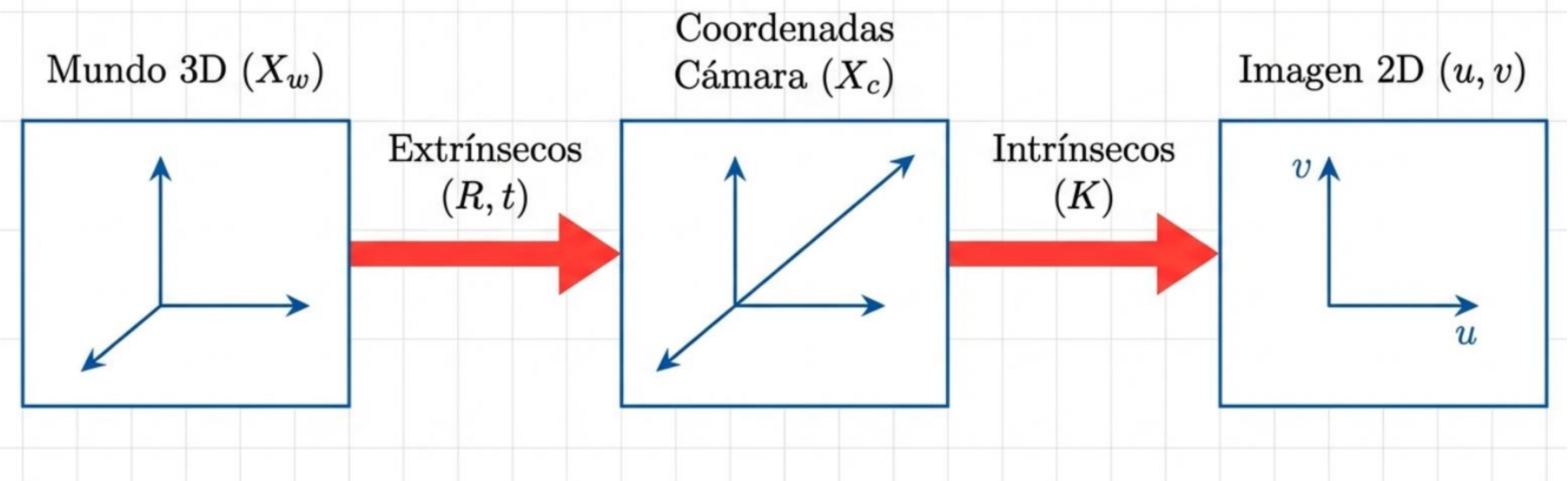
$$\boldsymbol{P} = \underbrace{\boldsymbol{K}}_{\text{Intrínsecos}} \times \underbrace{[\boldsymbol{R} \mid \boldsymbol{t}]}_{\text{Extrínsecos}}$$

Una matriz de 3×4 que mapea directamente coordenadas homogéneas del mundo ($\boldsymbol{X}$) a coordenadas homogéneas de imagen ($\boldsymbol{x}$).

Parámetros Intrínsecos (La Matriz K)



Flujo de Transformación de Coordenadas



$$P_{\text{pixel}} = [\text{Matriz Intrínseca}] \times [\text{Matriz Extrínseca}] \times P_{\text{mundo}}$$

Transformaciones Planares

- **Traslación:** Mueve puntos y conserva la orientación.
- **Euclidiana:** Rotación + Traslación (Rígida).
- **Similitud:** Añade escala uniforme (conserva los ángulos).
- **Afín:** Las líneas paralelas permanecen paralelas.
- **Proyectiva (Homografía):** Las líneas rectas permanecen rectas.
- Proyectiva es la función lineal más general.
- Representada por matrices de 3×3.

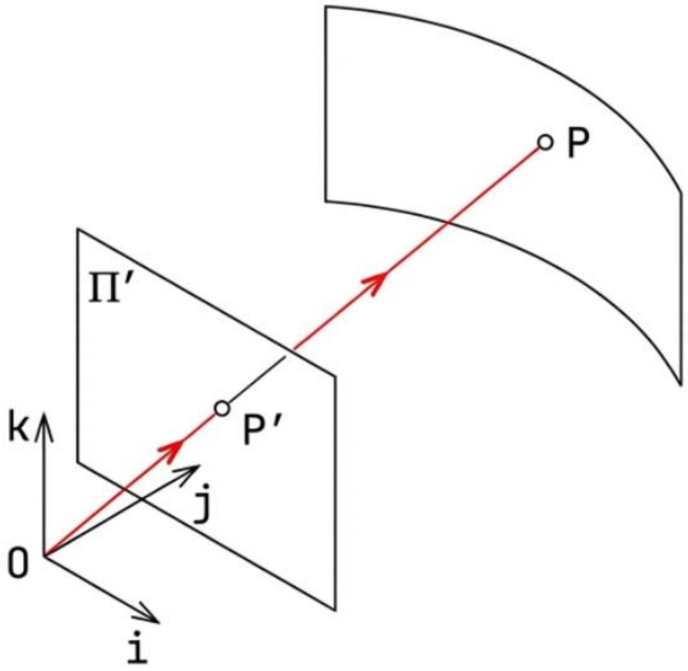
Transformaciones Proyectivas 2D

- **Isometría:** Rígida (Rotación + Traslación). 3 GDL.
- **Similitud:** Añade escala isotrópica. 4 GDL.
- **Afín:** Paralelismo conservado. 6 GDL.
- **Proyectiva (Homografía):** 8 GDL.
- Líneas rectas se mantienen rectas.
- Matriz H de 3×3 no singular.
- Fundamental para "Image Stitching" (Panoramas).

La Transformación Proyectiva (Homografía)

La Transformación Proyectiva (Homografía)

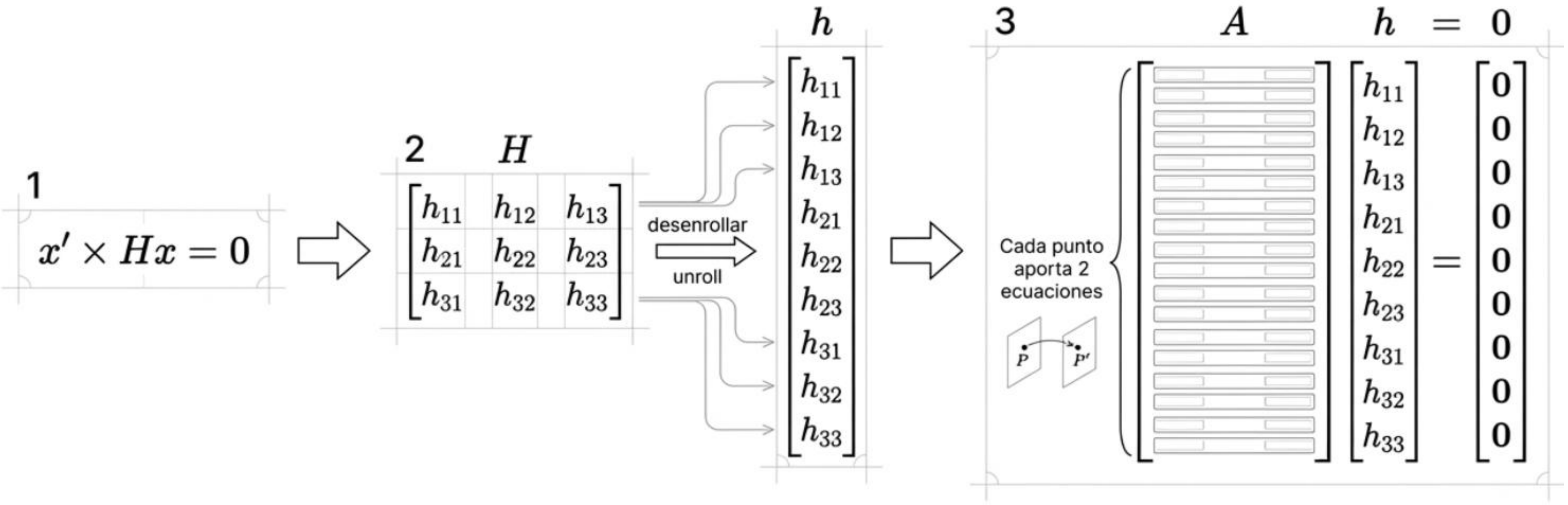
$$H = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix}$$



La Homografía encapsula la relación geométrica entre dos planos proyectivos.

- 8 grados de libertad (escala ambigua)
- Preserva la colinealidad (líneas rectas)
- Permite rectificación y panoramas

DLT (Direct Linear Transformation)

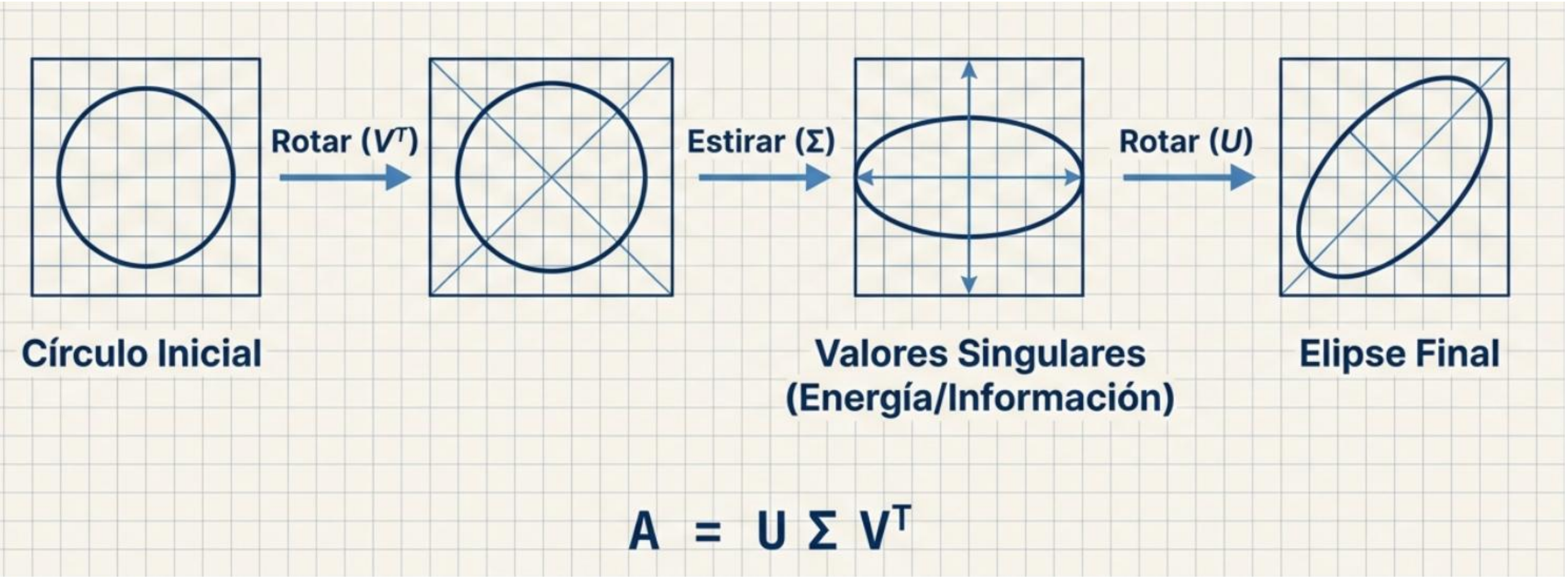


Restricción: Mínimo 4 puntos correspondientes para resolver 8 incógnitas.

Resolviendo la Homografía (DLT)

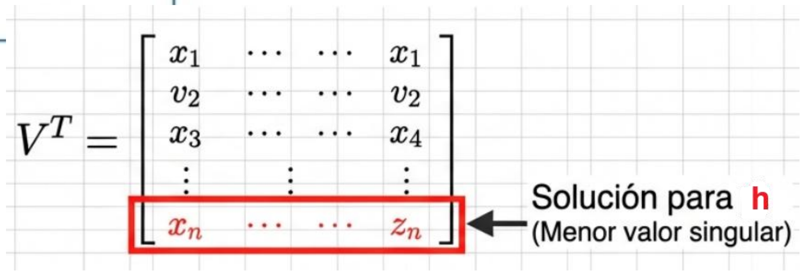
- Objetivo: Hallar H tal que $\mathbf{x}_i' \times H\mathbf{x}_i = \mathbf{0}$.
  - H es una matriz de 3x3 no-singular, y tiene 8 grados de libertad
- Genera sistema lineal homogéneo $Ah = \mathbf{0}$.
- A es matriz $2N \times 9$ ($N \geq 4$ puntos).
- Solución trivial $h = 0$ no es útil.
- Restricción: Minimizar $\| Ah \|$ sujeto a $\| h \| = 1$.
- **Solución:** Vector singular de A asociado al menor valor singular.
- Se obtiene vía SVD (Singular Value Decomposition).

Singular Value Decomposition



El Motor del Cálculo: SVD (Singular Value Decomposition)

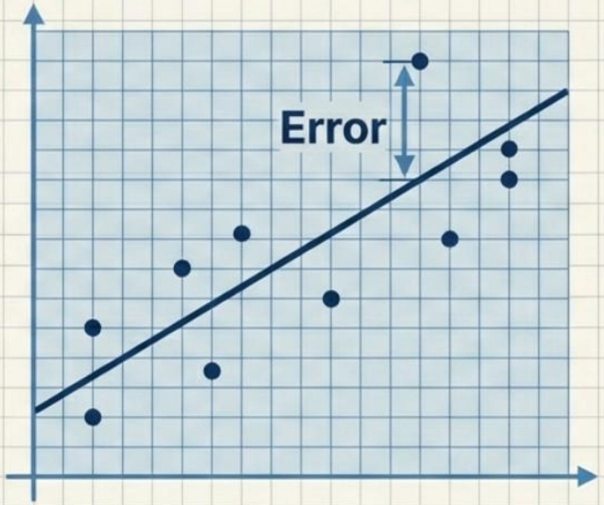
El problema Ah=0 tiene solución trivial h=0. Para evitarla, minimizamos ||Ah|| sujeto a ||h||=1.



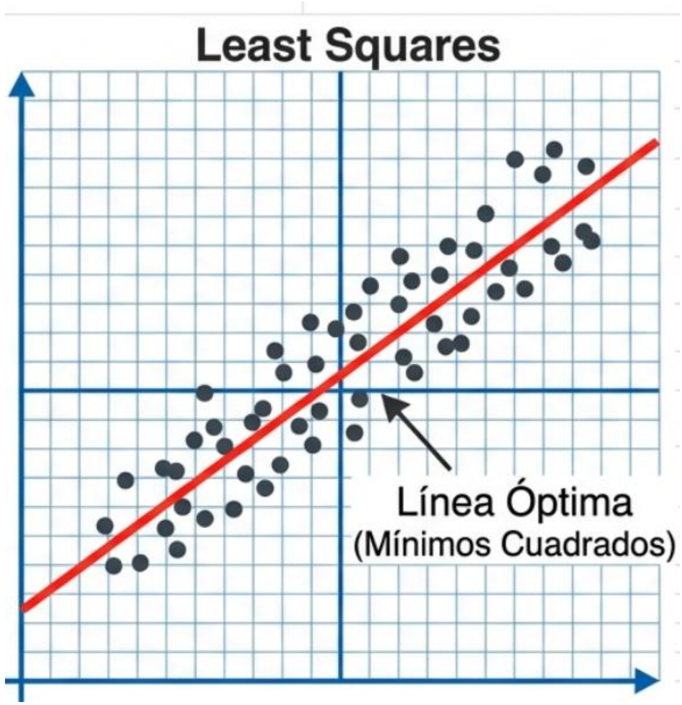
Pseudo Inversa y Mínimos Cuadrados

$$\tilde{x} = A^{\dagger}B$$

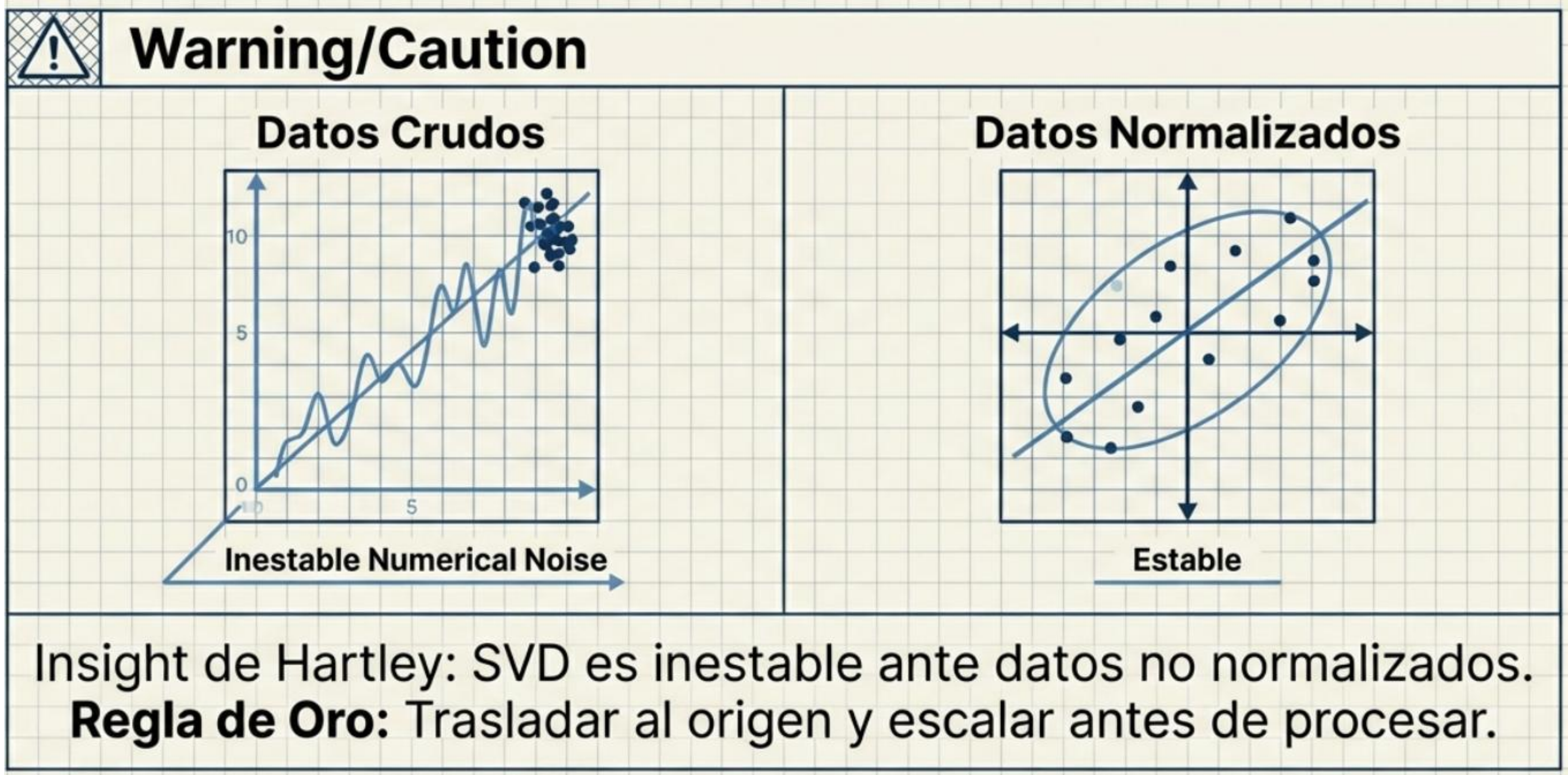
$$A^{\dagger} = V\Sigma^{-1}U^{T}$$



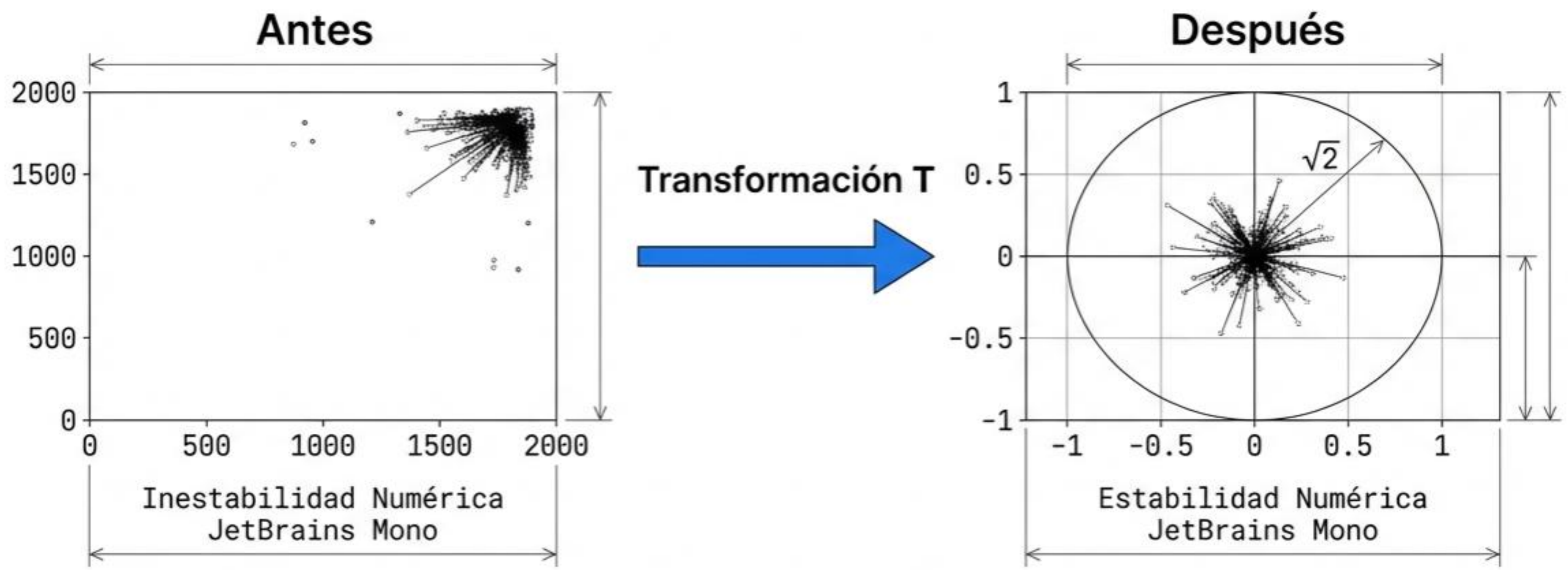
Dado que ***A*** no es cuadrada, no existe inversa. La Pseudoinversa encuentra la solución de **Mínimos Cuadrados** (Least Squares).



Normalizar los Datos

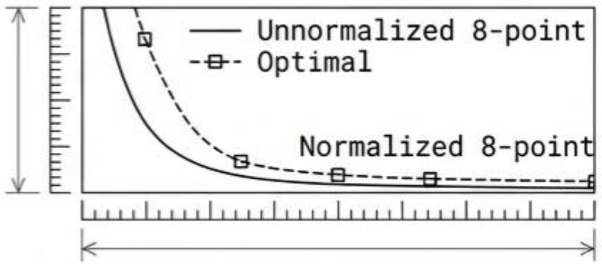


Normalización de Datos



Pasos de la Normalización:

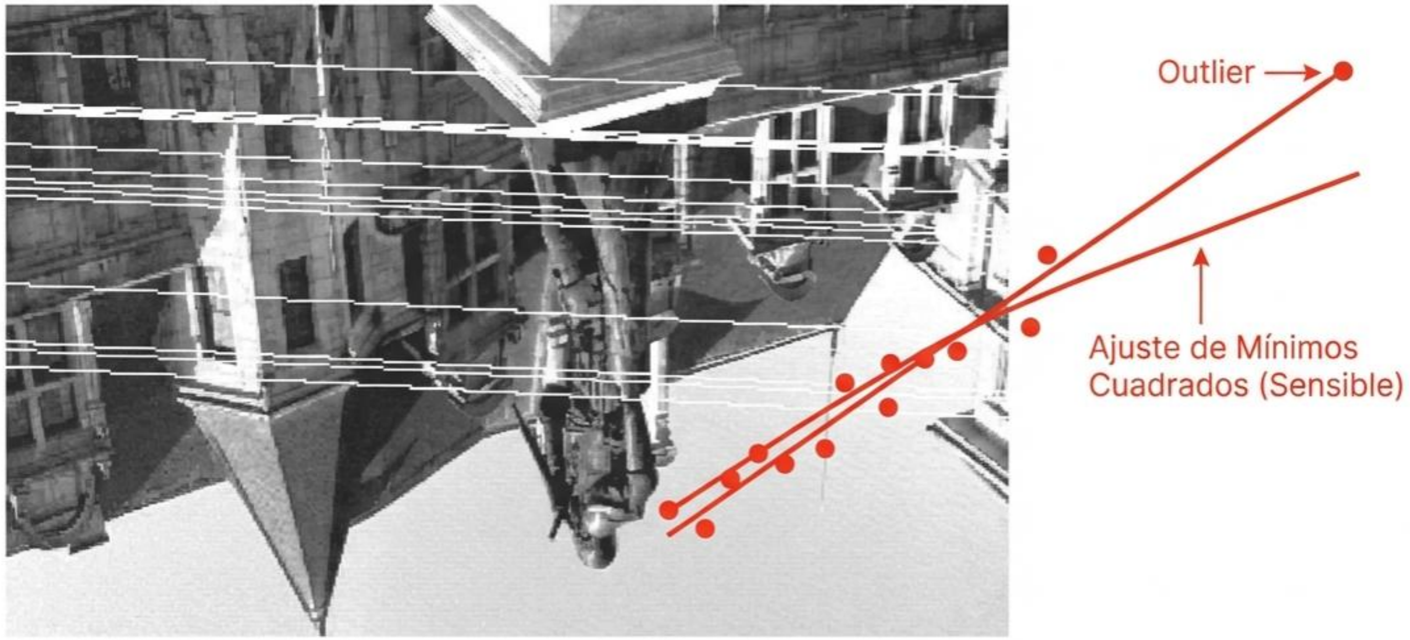
1. Trasladar centroide al origen (0,0)
2. Escalar distancia promedio a $\sqrt{2}$
3. Calcular SVD
4. Des-normalizar ($H_{final} = T^{-1} H T$)



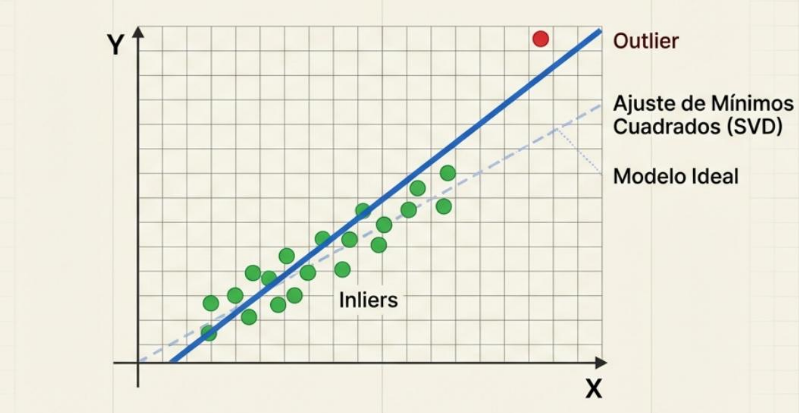
El Desafío de los "Outliers"

- Detectamos características (SIFT, ORB, Harris).
  - Emparejamiento de características (SIFT) genera errores
- Coincidimos descriptores entre dos imágenes.
- Las coincidencias contienen outliers (coincidencias erróneas).
  - **Outliers:** Emparejamientos falsos que no siguen el modelo.
- Mínimos Cuadrados (DLT) asume ruido Gaussiano en todo.
  - Trata de encajar todo
  - Un solo outlier desvía la solución enormemente.
- El promedio no es robusto ante valores extremos.
- Necesitamos un enfoque de "consenso/votación", no de promedio.
- Solución estándar en industria: **RANSAC.**

El desafío de los Outliers

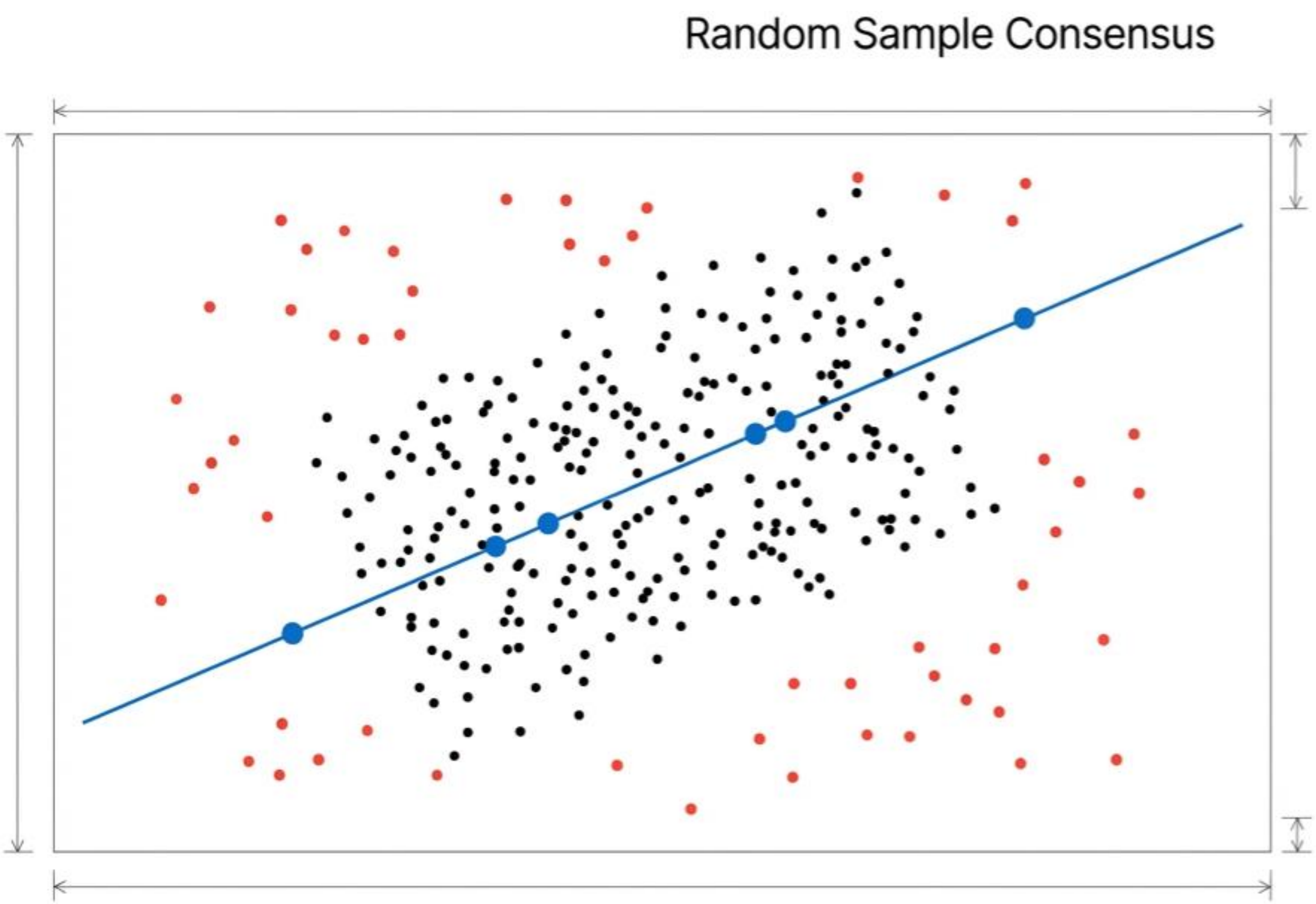


- SVD asume error Gaussiano (ruido suave).
- Un solo emparejamiento incorrecto destruye la estimación.
- Mínimos Cuadrados intenta 'complacer' a todos los puntos, incluso los erróneos.



SVD asume ruido gaussiano. Un solo dato incorrecto ("mentira") arruina la estimación.

Defensa de RANSAC



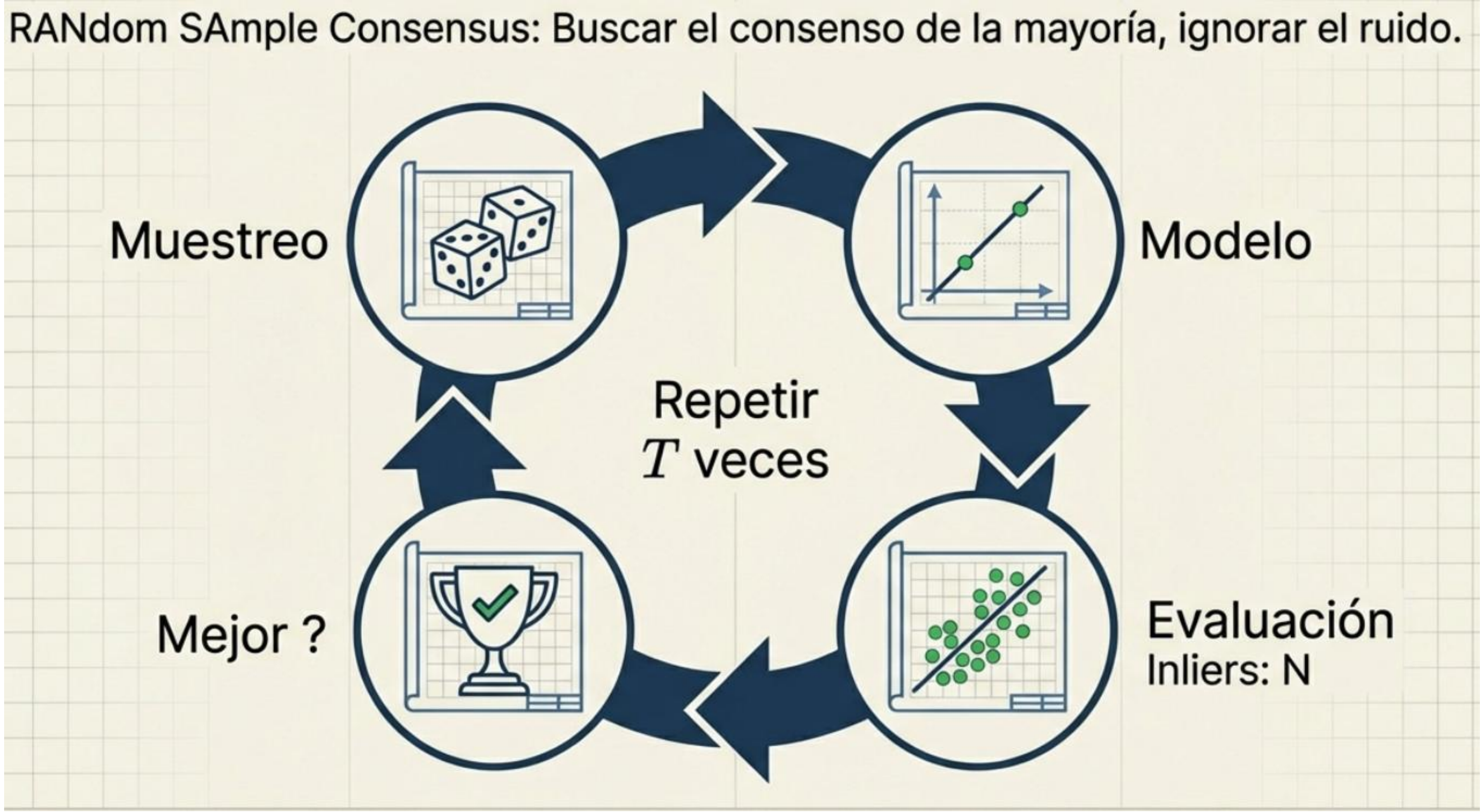
Un enfoque democrático:

1. Usar el mínimo de datos para una hipótesis (4 puntos).

2. Preguntar al resto: "¿Están de acuerdo?"

3. El modelo con más votos (inliers) gana.

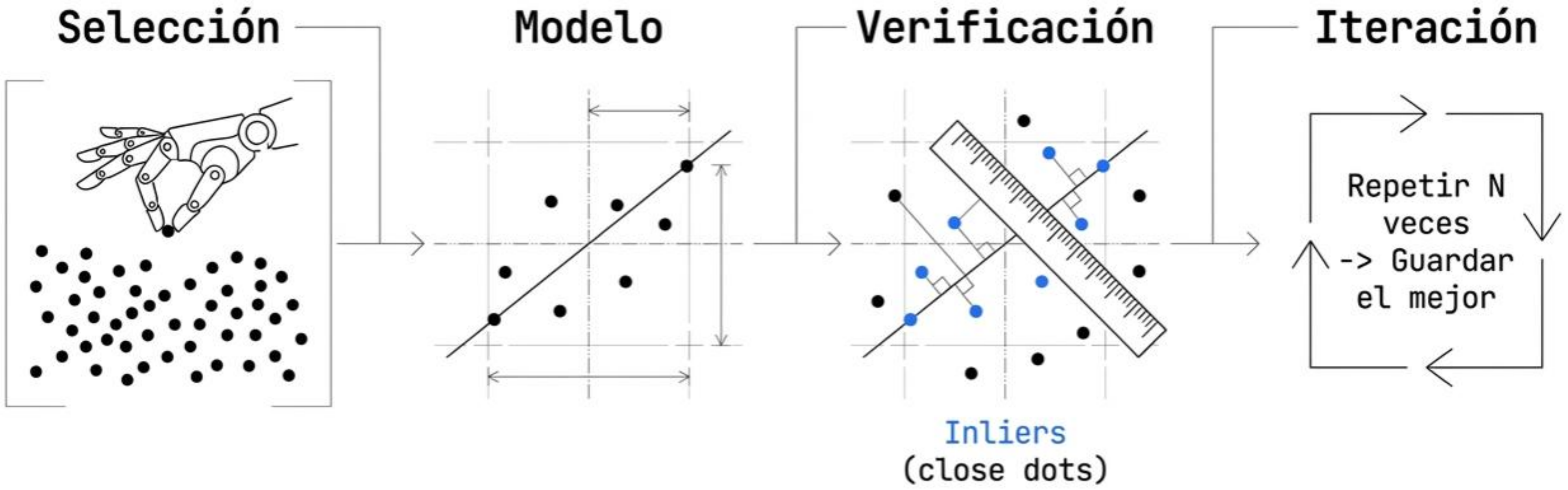
Defensa de RANSAC



Algoritmo RANSAC

- **Paso 1:** Seleccionar s puntos aleatorios (mínimo 4).
- **Paso 2:** Calcular modelo hipotético H_{test}.
- **Paso 3:** Verificar cuántos puntos ajustan (Inliers).
- **Criterio:** Distancia $\| \mathbf{x}' - H_{test}\mathbf{x} \| < umbral$.
- **Paso 4:** Repetir N veces.
- **Paso 5:** Quedarse con el modelo con más Inliers o votos.
- **Paso 6:** Refinar con todos los inliers (SVD final).

Algoritmo de RANSAC en 4 Pasos Generales



Estrategia: Fuerza bruta inteligente

Probabilidad en RANSAC

- No iteramos infinitamente; calculamos el límite
- Fórmula: $N = \frac{\log(1-p)}{\log\left(1-(1-e)^{s}\right)}$.
- s :Puntos necesarios para el modelo (4 para Homografía).
- e: Proporción estimada de outliers (ej. 50%).
- p: Probabilidad de éxito deseada (ej. 0.99).
- Si $e = 50\%$ y $s = 4$, $N \approx 72$ iteraciones.
- Computacionalmente muy eficiente.

Detallando las Iteraciones (T)

$$T = \frac{\log(1-p)}{\log(1-(1-e)^s)}$$

Probabilidad de Éxito (ej. 99%)

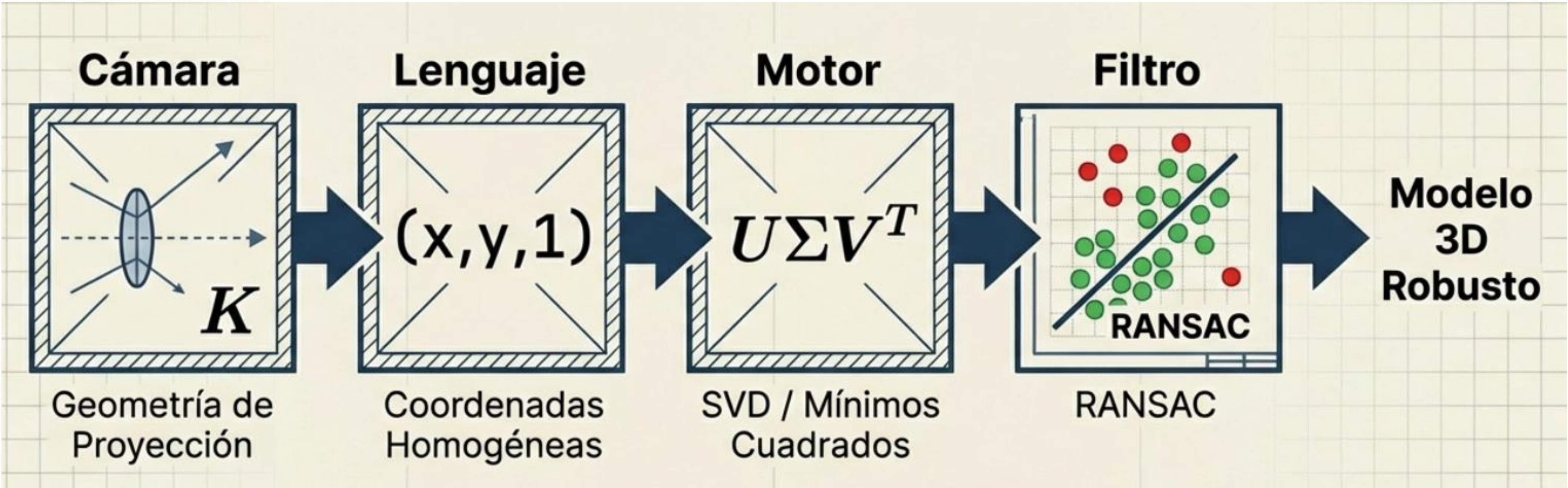
Proporción de Outliers

Puntos mínimos para el modelo

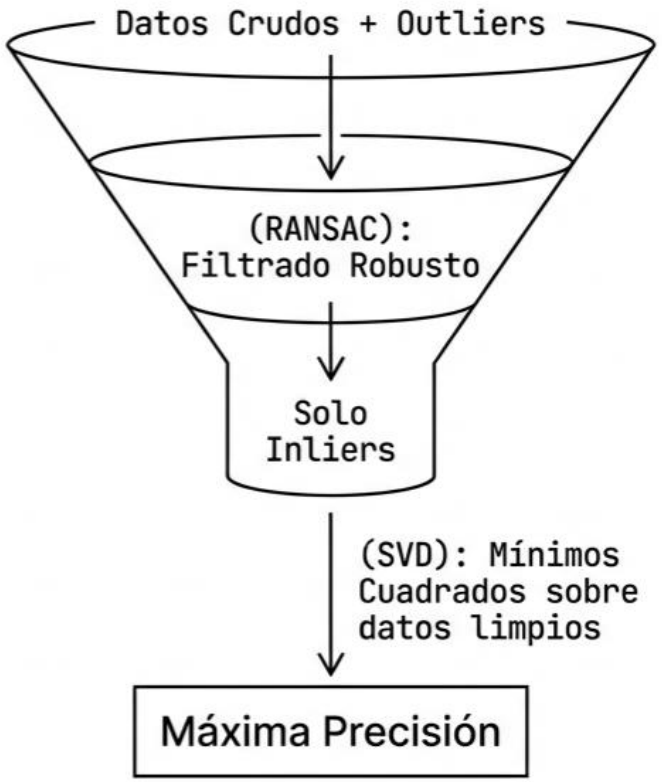
RANSAC garantiza matemáticamente el éxito incluso con datos altamente contaminados.

El Pipeline de Visión Computacional

- Visión por Computadora = Geometría + Álgebra Lineal + Estadística

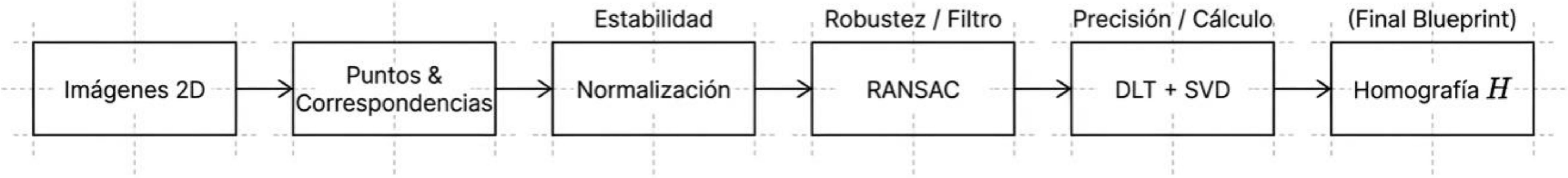


Refinamiento Final

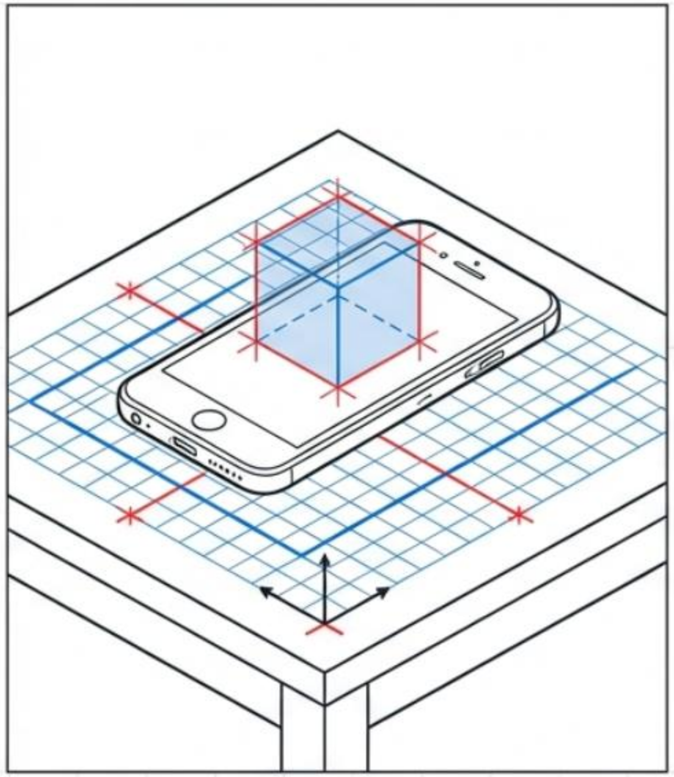


Del Caos a la Estructura

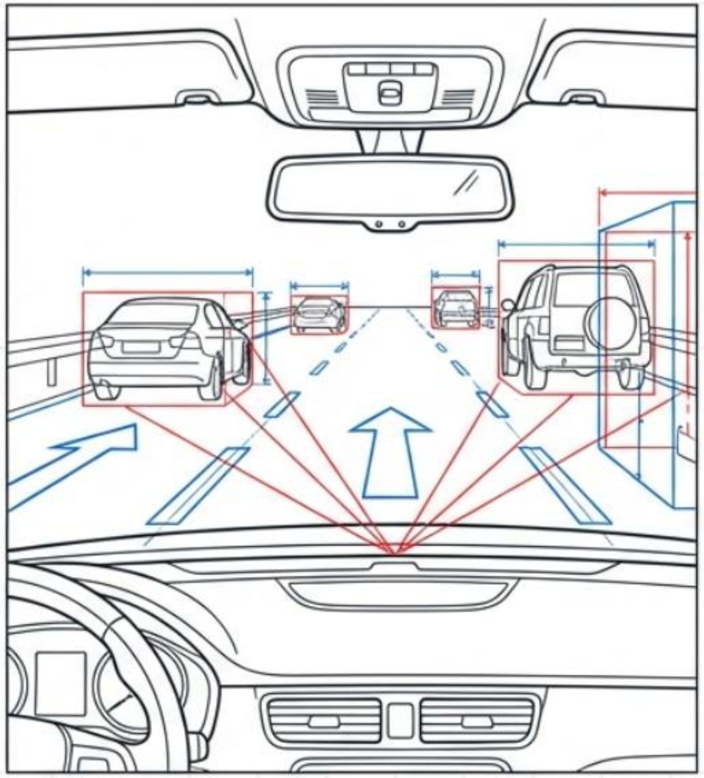
- La visión geométrica robusta no es solo calcular; es saber qué datos ignorar y cómo procesar óptimamente los que quedan



Aplicaciones Reales



Realidad Aumentada



Conducción Autónoma

- **AR/VR:** Inserción de objetos virtuales estables.
- **Robótica:** Estimación de distancias a obstáculos.
- **Reconstrucción 3D:** Modelado digital desde fotografías.

Resumen

- Pinhole → Homogéneas → Homografía → RANSAC.
- Las coordenadas homogéneas linealizan las matemáticas de proyección.
- Modelo estenopeico: $x=PX$
- (Intrínsecos/Extrínsecos).
- Las homografías mapean planos utilizando matrices 3×3.
- Los datos del mundo real son ruidosos.
- RANSAC encuentra geometría a pesar de los valores atípicos significativos.
- La robustez (RANSAC) es más importante que la precisión pura.

¡Gracias!