Projective Geometry: Homogeneous Points

Homogeneous coordinates: math hack

Allows us to represent translations using linear transformations (matrix multiplication).

homgenize

$$\begin{bmatrix} x \\ y \end{bmatrix} \rightarrow \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} ; \begin{bmatrix} x \\ y \\ w \end{bmatrix} \rightarrow \begin{bmatrix} x/w \\ y/w \end{bmatrix}$$

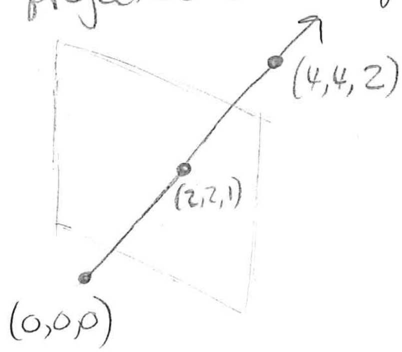
dehamogenize

$$\begin{bmatrix} x \\ y \\ w \end{bmatrix} \xrightarrow{\text{normalize}} \begin{bmatrix} x/w \\ y/w \\ 1 \end{bmatrix}$$

Mathematically speaking, 2D homogeneous coordinates live in

2D Projective space $\mathbb{P}^2$.

A nice geometric interpretation: objects in $\mathbb{P}^2$ are objects from $\mathbb{R}^3$, projected onto a plane using the origin $(0,0,0)$ as the COP.

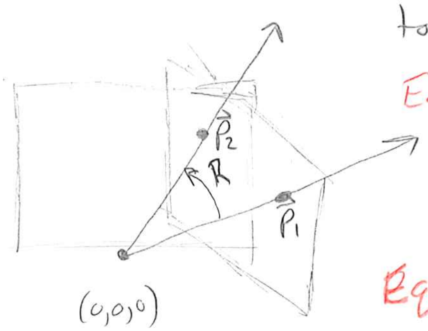


The projection means all points on the ray from $(0,0,0)$ in the direction of $\begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$ are equivalent: they project to the same point on the plane.

Interpreting Homographies

Projecting rays onto a different plane (with the same COP) is like applying a rotation in 3D to the homogeneous coordinates (eqn 1). If pixel coordinates are different from camera coordinates, i.e., $K \neq I$ we need to map from pixels to camera, rotate, then from camera to pixels (eqn 2)

rays:



Eqn 1:

$$\vec{p}_2 = R\vec{p}_1$$

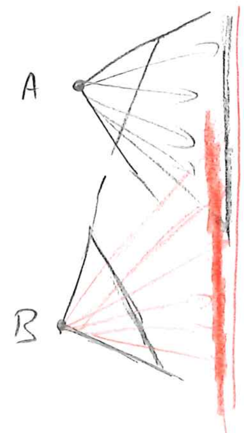
Eqn 2: $$\vec{p}_2 = \underbrace{KRK^{-1}}_{\text{3x3 matrix: homography!}}\vec{p}_1$$

Stereo Rectification

What we want:

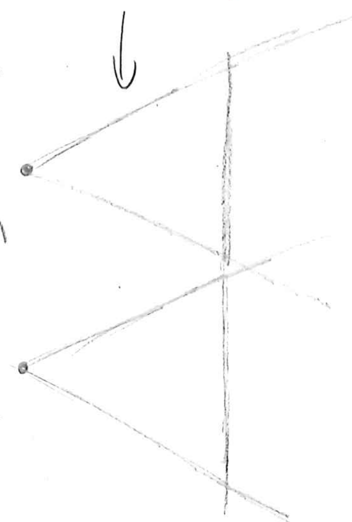
Same orientation
Same F
X translation only.

What we get:

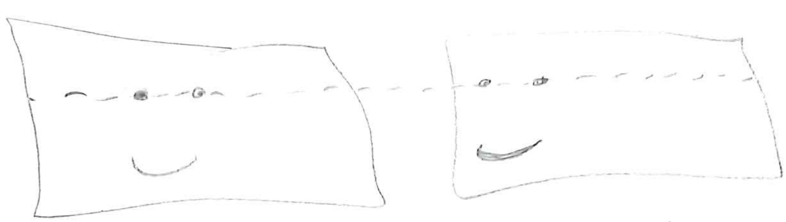


1. Let A be at origin (WLOG).
2. Project images onto a common plane (just a homography!) Simply rotating each camera.

Stereo pairs can be __rectified__ by mapping their images onto a common plane using a homography for each image. Geometrically, this is simply applying a rotation to each camera (and possibly adjusting for differing intrinsics.



Once rectified, our stereo pairs look friendly:



We can search along rows for matching windows to find disparity and compute depth, because the only translation is in X.