Homogeneous coordinates: math hack

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

Mathematically speaking, 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 is like applying a rotation in 3D to the homogeneous coordinates. If pixel coordinates are different from camera coordinates, we need to map from pixels to camera, rotate, then from camera to pixels (equation 2).

\[
\text{Eqn 1: } \tilde{\mathbf{p}}_2 = \mathbf{R} \tilde{\mathbf{p}}_1
\]

\[
\text{Eqn 2: } \tilde{\mathbf{p}}_2 = \mathbf{K} \mathbf{R} \mathbf{K}^{-1} \tilde{\mathbf{p}}_1
\]
Stereo Rectification

What we want:

- Same orientation
- Same F
- X translation only.

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.