Epipolar Geometry

Let's add a second camera.

For now, assume:

\[ K_1 = I_{3 \times 3} \]
\[ R_1 = I_{3 \times 3} \]
\[ t_1 = \mathbf{0} \]

For a given (homogeneous) point \( \mathbf{p} \) in image 1, where might it appear in image 2?

- The 3D point that projected to \( \mathbf{p} \) could be anywhere along the ray from \( C_1 \) in the direction of \( \mathbf{p} \).
- Those points all project into image 2 along a line!

What's the equation (homog. coordinates) of that line?

The plane projecting to that line is spanned by \( \mathbf{p} \) and \( t_2 \), so its coordinates are \( t_2 \times \mathbf{p} \).

However, those coordinates are wrt image 2! What does camera 2 see?

Observation: the plane (line) \( t_2 \times \mathbf{p} \) goes through \( C_2 \)'s COP, so all we need to see the line from \( C_2 \)'s view is to rotate it (or equivalently, the camera):

\[ l_{\mathbf{p}} = R_2(t_2 \times \mathbf{p}) = R_2(t_2)^{T} \mathbf{p} \]
We have $l_p^2 = R_2[t] x \hat{p}$.

Any point $\hat{q}$ in image 2 that lies on this line satisfies:

$$\hat{q}^T l_p^2 = 0$$

or:

$$\hat{q}^T R_2[t] x \hat{p} = 0$$

$E$ = the essential matrix

If $K_1 = K_2 = I_{3 \times 3}$, then for any pixel $\hat{p}$ in image 1, the corresponding point $\hat{q}$ in image 2 satisfies

$$\hat{q}^T E \hat{p} = 0$$

What if $K_1$ and $K_2$ are not $I_{3 \times 3}$, but are known?

Pixel $p$ in image 1: $\hat{p} = K_1^{-1} p$

$q$ in image 2: $\hat{q} = K_2^{-1} q$

Plug into the above:

$$\hat{q}^T R_2[t] x \hat{p} = 0$$

$$q^T K_1^{-1} R_2[t] K_1^{-1} p = 0$$

$F$ = the fundamental matrix

For pixel coordinates $p, q$, if they correspond to the same 3D point, then $q^T E p = 0$
Properties of F and epipolar geometry:

- F has rank 2: \( Fp \) maps to a 1D solution space.
- All epipolar lines go through the epipole; the baseline vector \( t \) spans all epipolar planes and passes through both epipoles.
- \( e_1 \) spans the null space of F: \( Fe_1 = 0 \)
- \( e_2 \) spans the null space of \( F^T \): \( F^T e_2 = 0 \)

We derived all this assuming \( k_1, k_2, R_2, t_2 \) are known. Can we find F, then use it to determine camera parameters?
- Sort of. (i.e. yes, but it's hard.)

Finding F: Like finding \( H \), we can set up a homogeneous least squares system using equations from the 8-point algorithm.

\( K, R, t \) from F: this is harder, but there are strategies.

In general, reconstructing camera pose and 3D geometry is done jointly (chicken-and-egg); this is called "Structure from motion" (3D geom.) \( \rightarrow \) (camera pose)