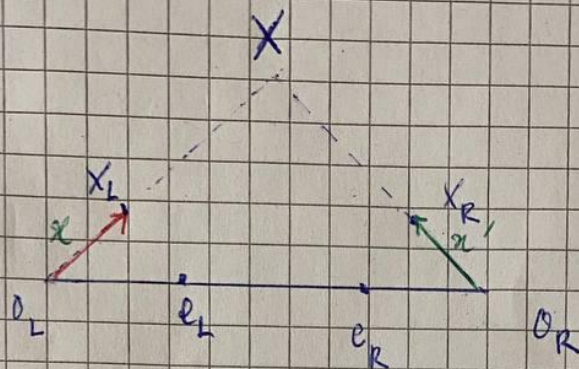


Name: Cuong Ha; Matrik. Nr: 03749410

Part 2: Epipolar constraint

Derive the computation of the essential matrix from the transformation between two cameras



We assume $t = \vec{O_R O_L}$ i.e. translation from right camera to left camera; and R is the rotation from the left camera to the right.

$$\Rightarrow x' = R(x - t)$$

Since x, t, x' are coplanar, then:

$$x^T (t \times x) = 0$$

$$\text{and } (x - t)^T (t \times x) = 0$$

$$\Rightarrow (x'^T R) (t \times x) = 0$$

$$\Leftrightarrow x'^T R \cdot [t_x] \cdot x = 0 \text{ where } [t_x] \text{ is the cross product matrix of } t$$

$$\Rightarrow x'^T E x = 0 \text{ where } E = R \cdot [t_x]$$

Therefore

$$x_R^T E x_L = 0 \text{ where } E = R \cdot [t_x]$$

Part 4: Bag-of-Words for Place Recognition

Compare the number of candidate pairs:

For n frame pairs i.e. $2n$ images.

- Match all: $2n * (2n - 1) / 2 = 2n(n-1)$ candidate pairs
- Match bow: between $2n * (\text{num_bow_candidates} - 1)$ and $2n * \text{num_bow_candidates}$

If $\text{num_bow_candidates}$ is set to 25.

- Case 2 x 82 images:
 - Match all: 13284 candidate pairs
 - Match bow: between 3936 and 4100 candidate pairs.
- Case 2 x 1000 images:
 - Match all: 1998000 candidate pairs
 - Match bow: between 48000 and 50000 candidate pairs.

When the number of images grows, “Match bow” significantly reduces the number of candidate pairs compared to “Match all”.