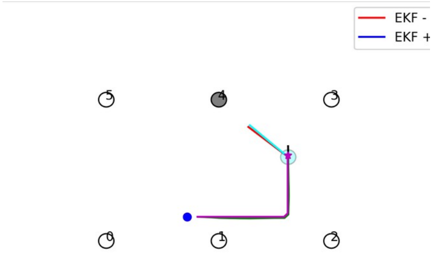
Lecture 9. Visual Data Association

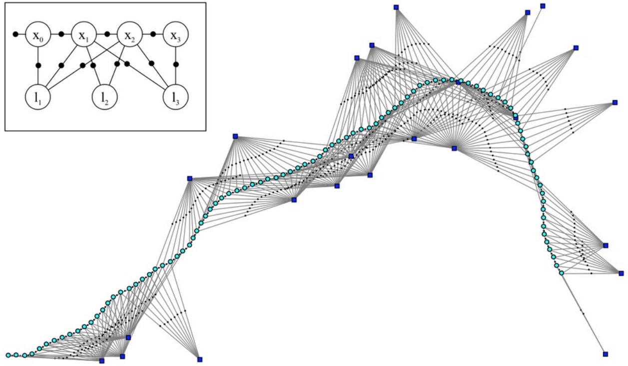
CAIT
Mobile Robotics Lab
Perception in Robotics course

Motivation. Why do we need data association?



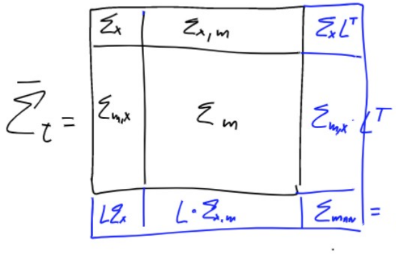
PS2. An observation of **4-th landmark**.

No need for DA



Factor graph (**trajectory** and **observations**) of **Square Root SAM**.

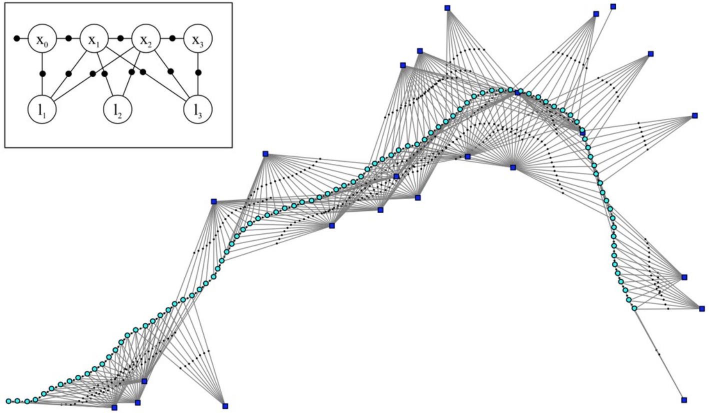
DA is required



EKF-SLAM covariance matrix

What if we don't know **what landmark** are **the observation associated with**?

Front-end and back-end of a SLAM system



Construct the graph (**Front-end**), **depends on the data and the method**

Sensor data

Solve the graph (**Back-end**)

Solved poses

Back-end is the **Mathematical Optimization framework**

- Extended Kalman Filter
- Particle Filter
- Graph SLAM
- Bundle Adjustment (**Visual SLAM**)

Front-end (**Visual SLAM**):

- Feature extraction
- Feature association

Landmarks for Visual SLAM?

state 0

state 1

$$x_0 \xrightarrow{x_1 = R_{01}x_0 + t_{01}} x_1$$

Sparse local features. Keypoints

Green circles are **keypoints** detected on both images

Sparse local features. Descriptors

256d. Real-valued vector.
L2 or **cosine similarity**

0.1	0.5	0.9	...	-0.3	-0.1

256d. Binary vector.
Hamming distance

1	0	0	...	1	1

Sparse local features. Associations

NN matching

Consider two sets of descriptors:

$$d_i^1 \in \{d_0^1, \ldots, d_{n_1-1}^1\}, \quad ||d_i^1||_2 = 1$$

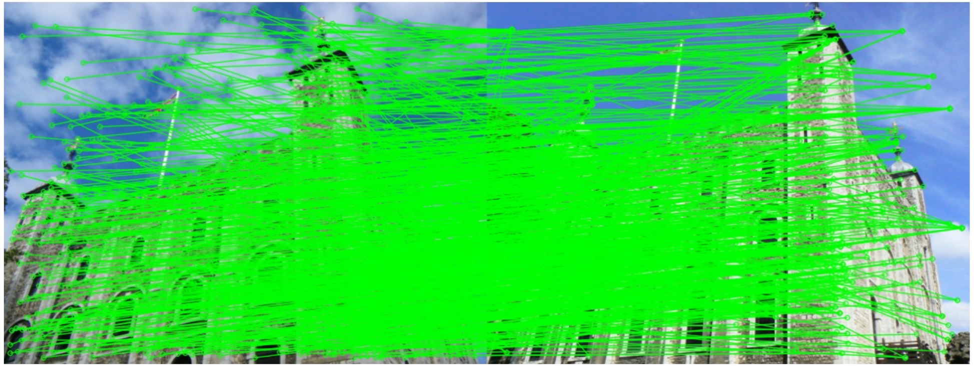
$$d_i^2 \in \{d_0^2, \ldots, d_{n_2-1}^2\}, \quad ||d_i^2||_2 = 1$$

Then the NN for **each index** from the **first set** can be found as:

$$m_i^1 = \underset{j \in \{0, \ldots, n_2-1\}}{argmin} ||d_i^1 - d_j^2||_2$$

single match

matches 1->2



Association filtering. Lowe ratio. Mutual NN

Lowe ratio test

Let $\tilde{m}_i^1 = j$ be retained then it holds:

$$\frac{||d_i^1 - d_j^2||_2}{||d_i^1 - d_k^2||_2} < r, \quad r \in [0,1]$$

where

$$k = \underset{l \in \{0,\ldots,j-1,j+1,\ldots,n_2-1\}}{argmin} ||d_i^1 - d_l^2||_2$$

Mutual NN

Enforce: $\tilde{m}_i^1 \to j \;\cap\; \tilde{m}_j^2 \to i$