ICP-SLAM as LS Optimization problem

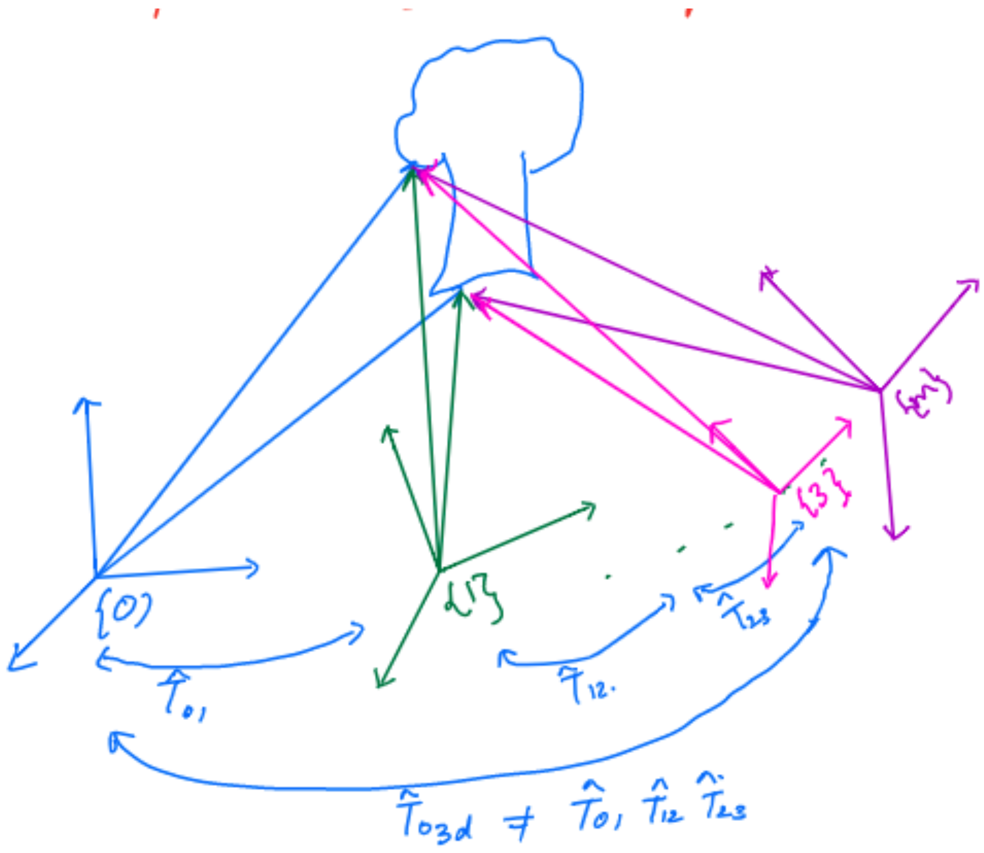
Comments	
Order (according to me)	
STATUS	
Week (according to Prof Madhav's original email)?	

Prof's notes on ICP-SLAM

Prof Madhav's notes on ICP-SLAM

(This Notion page link if you're viewing a PDF)

0. Need for SLAM Backend or Multiview ICP "Optimization"



In the figure:

- Blue frame $\{0\}$ 🖼️ —— $\mathrm{T}_0 = [\mathbf{I}|\mathbf{0}]$ is the origin.
- Blue tree 🌴—— $\mathrm{X}_{0j} = \begin{bmatrix} x_{0j} & y_{0j} & z_{0j} \end{bmatrix}^\mathrm{T}$: where $j = \{1, \ldots, n\}$ are the coordinates of n points in frame $\{0\}$ obtained from depth, azimuth and elevation measurements.
  - Since the depth for every pont j in frame $\{0\}$ is in error, hence X_{0j} is also in error.
  - So is the case for X_{1j} (frame 1) and hence, relative pose estimates between successive frames $\{i, i+1\}$ i.e. $\widehat{\mathrm{T}}_{01}, \widehat{\mathrm{T}}_{12}, \ldots, \widehat{\mathrm{T}}_{(m-1,\,m)}$ are all in error.
- How to alleviate this error?
  - Filtering methods: Last topic of the semester (After Vision)
  - Optimization methods: Now — Pose this as **"multiview optimization".**

Key Insight

- If there is a frame $\{q\}$ in which the depth measurements and the point cloud X_{qj} are particularly noiseless: Can this be used to alleviate other views in terms of poses and 3D points estimated in those views
- If a set of n points are viewed in m frames or observations, what is the best estimate for these n points and m poses.
    - **Multiview Aggregation**
    - **Multiview Consistency**
- Let the points $j = \{1, ..., n\}$ be represented in frames $i = \{1, ..., m\}$ as $\mathrm{X}_{ij} = \begin{bmatrix} x_{ij} & y_{ij} & z_{ij} \end{bmatrix}^{\mathrm{T}}$ as mn observations. $\widehat{\mathrm{X}}_{ij}$ **are estimated from ICP while** X_{ij} **are obtained directly through a sensor say LiDAR.**
- Also based on the observation of j in $\{0\}$ which is X_{0j} and $\widehat{\mathrm{T}}_{0i}$ **estimated from ICP** as $\widehat{\mathrm{T}}_{0i} = \widehat{\mathrm{T}}_{01}\widehat{\mathrm{T}}_{12} \ldots \widehat{\mathrm{T}}_{(i-1,\, i)}$ we predict what is X_{0j} in frame i as

$$\widehat{\mathrm{X}}_{ij} = \widehat{\mathrm{T}}_{i0}\mathrm{X}_{0j} \tag{0.1}$$

Why is this $\widehat{\mathrm{X}}_{ij}$ different from the observations X_{ij}?

- Two reasons! Participate by commenting! (for students)
    1. `TODO` through participation
    2. `TODO` through participation

Various ways to approach this "multiview ICP"

We go with the first procedure below.

1. When I aggregate the same n points from multiple views $\widehat{\mathrm{X}}^{i}_{0j} = \widehat{\mathrm{T}}_{0i}\mathrm{X}_{ij}$, I get m sets of n points in frame $\{0\}$ that I average as

$$\widehat{\mathrm{X}}_{0j} = \sum_{i=1}^{m} \frac{\widehat{\mathrm{X}}^{i}_{0j}}{m} \tag{0.2}$$

   or such aggregation over the m frames.

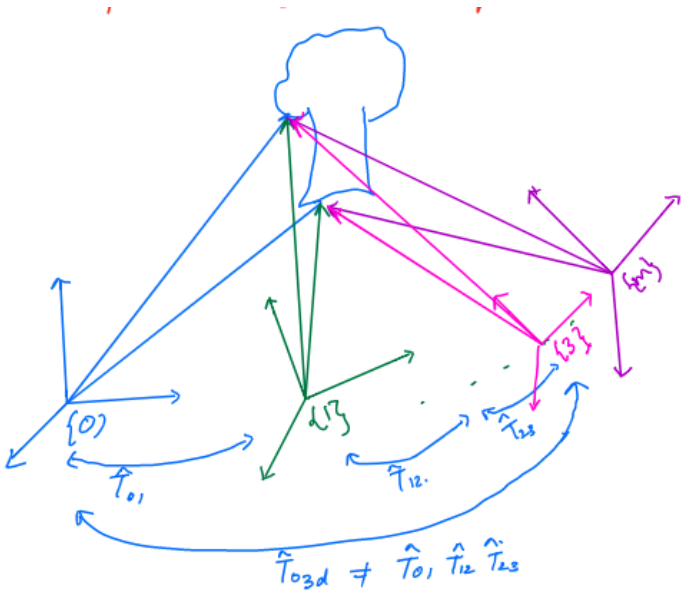
2. I can also use an ICP estimate directly between the points in frames $\{0\}$ and $\{i\}$ and get a $\widehat{\mathrm{T}}_{0i\mathbf{d}}$ and say that the 2 sets of points should be the

same which is

$$\widehat{\mathrm{X}}_{ij\mathbf{d}} = \widehat{\mathrm{T}}_{i0\mathbf{d}}\mathrm{X}_{0j} \tag{0.3}$$

should be the same as $\widehat{\mathrm{X}}_{ij}$ predicted/estimated by (0.1). The difference is that here, we are estimating ICP directly while above, we are estimating ICP between immediate pair of point clouds.

1. ICP-SLAM as Optimization



$\vec{X}$ (represented by the arrow above) are direct measurements (say from LiDAR) and $\widehat{T}$ are obtained from ICP. ($\vec{X}_{0p}$ below could also refer to 1. described above but for the sake of this derivation, we will keep things simple and refer to it as direct LiDAR measurement).

$$\sum_{i=1}^{m} \left\| \vec{X}_{ip} - \widehat{T}_{i0}\vec{X}_{0p} \right\|_2^2$$

Over n points from $j = 1 \to n$ (Remember m is frames)

$$\sum_{i=1}^{m}\sum_{j=1}^{n} \left\| \vec{X}_{ij} - \widehat{T}_{i0}\vec{X}_{0j} \right\|_2^2 \tag{1.1}$$

$$\hat{T}_{i0} = \left[\hat{T}_{0i}\right]^{-1} = \left[\hat{T}_{01}.\hat{T}_{12}\ldots\hat{T}_{i-1\,i}\right]^{-1}$$

What are the optimum locations of the 3D points $\vec{X}_{0j}, j = 1 \rightarrow n$ and the location of the Mobile Robot $\vec{T}_{0i}, i = 1 \rightarrow m$ so that $\sum_{i=1}^{m} \sum_{j=1}^{n} \left\| \vec{X}_{ij} - \widehat{T}_{i0} \vec{X}_{0j} \right\|_2^2$ or

$$\sum_{i=1}^{m} \sum_{j=1}^{n} \| r_{ij} \|_2^2$$

is minimized?

Num of variables

$\vec{X}_{0j}, j = 1 \rightarrow n; \ \vec{T}_{0i}, i = 1 \rightarrow m \implies$ $12\mathrm{M} + 3\mathrm{N}$ $\boxed{12m+3n}$

$$\begin{aligned} X_{ij} &= f_{ij}\left(R_i, t_i, X_j\right) \qquad \text{dropping suffix 0} \\ &= f\left(R_i, t_i, X_i\right)_{|R_{iO},\, t_{iO},\, X_{jO}} + J_{ij}\delta_{ij} \end{aligned} \tag{1.2}$$

$$J_{ij\ (3\times15)} = \begin{bmatrix} \dfrac{\partial \mathbf{f_{ij}}}{\partial \mathbf{T}_i}_{\ (3\times12)} & \dfrac{\partial \mathbf{f_{ij}}}{\partial \mathbf{X}_j}_{\ (3\times3)} \end{bmatrix} \tag{1.3}$$

$$X_{ij} = f_{ij}\left(T_i(\xi_i), X_j\right) \tag{1.4.1}$$

Rotation about arbitrary axis links:

- Moodle file
- Notion link

$$J_{ij\ (3\times9)} = \begin{bmatrix} \left(\dfrac{\partial \mathbf{f_{ij}}}{\partial \mathbf{T}_i} \dfrac{\partial \mathbf{T}_i}{\partial \xi_\mathbf{i}}\right)_{(3\times6)} & \dfrac{\partial \mathbf{f_{ij}}}{\partial \mathbf{X}_j}_{\ (3\times3)} \end{bmatrix} \tag{1.4.2}$$

$$J_{ij\ (3\times9)} = \begin{bmatrix} \left(I_{3\ (3\times3)} \quad -\left[\mathrm{T} \oplus \mathrm{X_j}\right]_{X\ (3\times3)}\right)_{(3\times6)} & \dfrac{\partial \mathbf{f_{ij}}}{\partial \mathbf{X}_j}_{\ (3\times3)} \end{bmatrix} \tag{1.5}$$

$$\delta_{ij\ (9\times1)}$$

- What is $\oplus$? It is a pose composition whose can simply be written as a $matrix \times vector$ as follows:

$$\mathrm{T} \oplus \mathrm{X_j} = \mathrm{TX_j} \quad = [R_i X_j + t_i]$$

$$= [R_{i0} \vec{X}_{0j} + t_{i0}]$$

The product $\mathrm{TX_j}$ is now a vector. What is $[\mathrm{TX_j}]_X$ then?

$$[\mathrm{TX_j}]_X \text{ is a skew symmetric matrix version of } [\mathrm{TX_j}]$$

If $\mathrm{P} = [x, y, z]^\mathrm{T}$ is a vector, it's skew symmetric form is given by:

$$[\mathrm{P}]_X = \begin{bmatrix} 0 & -z & y \\ z & 0 & -x \\ -y & x & 0 \end{bmatrix}$$

- *Cross product as skew-symmetric matrix:*

$$\begin{aligned} \mathbf{a} \times \mathbf{b} &= \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix} \\ &= (a_2 b_3 - a_3 b_2)\,\mathrm{i} + (a_3 b_1 - a_1 b_3)\,\mathrm{j} + (a_1 b_2 - a_2 b_1)\,\mathrm{k} \end{aligned}$$

$$\mathbf{a} \times \mathbf{b} = [\mathbf{a}]_\times \mathbf{b} = \begin{bmatrix} 0 & -a_3 & a_2 \\ a_3 & 0 & -a_1 \\ -a_2 & a_1 & 0 \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$$

$$\mathbf{a} \times \mathbf{b} = [\mathbf{b}]_\times^\mathrm{T} \mathbf{a} = \begin{bmatrix} 0 & b_3 & -b_2 \\ -b_3 & 0 & b_1 \\ b_2 & -b_1 & 0 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix}$$

J_{ij} can be split as:

$$J_{ij} = \left[\underbrace{\mathrm{J_{ij\,L}}_{(3\times6)}}_{Localization\ Jacobian} \;\middle|\; \underbrace{\mathrm{J_{ij\,M}}_{(3\times3)}}_{Mapping\ Jacobian} \right]$$

- **Localization Jacobian**: Associated with the pose derivatives
- **Mapping Jacobian**: Associate with the map (point cloud) derivatives

Then,

$$
J = \left[\begin{array}{cccc|cccc}
\multicolumn{4}{c|}{\overbrace{\hphantom{J_{11_L}\quad 0 \quad \cdots \quad 0}}^{\text{'m' times}}} & \multicolumn{4}{c}{\overbrace{\hphantom{J_{11_M}\quad 0 \quad \cdots \quad 0}}^{\text{'n' times}}} \\
J_{11_{L_{3\times 6}}} & 0_{3\times 6} & \cdots & 0 & J_{11_{M_{3\times 3}}} & 0_{3\times 3} & \cdots & 0_{3\times 3} \\
J_{12_L} & 0 & \cdots & 0 & 0 & J_{12_M} & \cdots & 0 \\
\vdots & & & & 0 & & \ddots & \\
J_{1n_L} & 0 & \cdots & 0 & 0 & & & J_{1n_M} \\
0 & J_{21_L} & \cdots & 0 & J_{21_M} & 0 & \cdots & 0 \\
\vdots & \vdots & & \vdots & \vdots & J_{22_M} & \cdots & 0 \\
0 & J_{2n_L} & \cdots & 0 & 0 & & \ddots & J_{2n_M} \\
\vdots & & & \vdots & & & & \\
0 & & \cdots & J_{m1_L} & J_{m1_M} & & \cdots & 0 \\
\vdots & & & & \vdots & & & \\
0 & & \cdots & J_{mn_L} & 0 & & \cdots & J_{mn_M} \\
\multicolumn{4}{c|}{\underbrace{\hphantom{J_{11_L}\quad 0 \quad \cdots \quad 0}}_{\text{LOCALIZATION JACOBIAN}}} & \multicolumn{4}{c}{\underbrace{\hphantom{J_{11_M}\quad 0 \quad \cdots \quad 0}}_{\text{MAPPING JACOBIAN}}}
\end{array}\right]
\begin{bmatrix} \delta\xi_1 \\ \delta\xi_2 \\ \vdots \\ \delta\xi_m \\ \delta X_{01} \\ \delta X_{02} \\ \vdots \\ \delta X_{0n} \end{bmatrix}
$$

$$
= \mathrm{J}_{(3mn,6m+3n)} \;\; \delta_{(6m+3n)}
$$

- For every one of the m robot locations, there are n points giving rise to $3n$ ICP equations. Note each point gives 3 ICP equations.
- For m such robot locations, we have 3nm equations .
- Each of the m poses has 6 parameters in the tangent vector ξ_i and each point has 3 components: 6m + 3n variables

$$
\hat{X}_{ij} = T_{i0} X_{0j} = \begin{bmatrix} \hat{x}_{ij} \\ \hat{y}_{ij} \\ \hat{z}_{ij} \end{bmatrix}
$$

$\hat{X}_{ij_X}$ is a skew symmetric matrix given by

$$
\left[\hat{X}_{ij}\right]_X = \begin{bmatrix} 0 & -\hat{z}_{ij} & \hat{y}_{ij} \\ \hat{z_{ij}} & 0 & -\hat{x}_{ij} \\ -\hat{y}_{ij} & \hat{x}_{ij} & 0 \end{bmatrix}
$$

$$\xi = \begin{bmatrix} \omega_1 \\ \omega_2 \\ \omega_3 \\ t_x \\ t_y \\ t_x t_z \end{bmatrix}$$

where $[\omega_1, \omega_2, \omega_3]^T$ represent the axis about which the rotation occurred

θ is the magnitude of the rotation $\omega = [\omega_1, \omega_2, \omega_3]^T$ and

$$\theta^2 = \omega^T \omega$$

$$R = \exp\left([\omega]_X\right) = I_{3\times 3} + \left(\frac{\sin\theta}{\theta}\right)\omega_X + \left(\frac{1-\cos\theta}{\theta^2}\right)\omega_X^2$$

Here, $[\omega]_X$ is the skew symmetric matrix of ω.

After LM or Gauss Newton, we get

$$X_{oj}(n+1) = X_{oj}(n) + \delta X_{oj}$$

$$\xi_i(n+1) \leftarrow \xi_i(n) - \delta\xi \tag{2.1}$$

$$\xi_i(n+1) = \begin{bmatrix} \omega_i(n+1)\ _{(3\times1)} \\ n\ _{(3\times1)} \end{bmatrix} \tag{2.2}$$

$$R_i(n+1) \text{ or } R_{i0}(n+1) \leftarrow \exp\left([\omega_i(n+1)]_X\right) \tag{2.3}$$

$$\begin{bmatrix} \delta\xi_i \\ \delta X_j \end{bmatrix} = [J^T J]^{-1} J^T [\vec{X}_{ij} - \hat{R}_{io}\vec{X}_{oj} + \hat{\vec{t}}_{io}]$$