Homework 4: Dense SLAM with Point-based Fusion

16-833 Robot Localization and Mapping

Austin Windham

1 Overview

No sections required for the report.

2 Iterative Closest Point (ICP)

2.1 Project data association

Question 1:

For a point p to be projected to a vertex map, the point must lie within the projected vertex map to be a valid correspondence. So, if a vertex map's height and width is given by H and W, then u, v, and d must satisfy the following conditions:

d >0 ; 0 <= u < W ; 0 <= v < H.

Question 2:

It was necessary to filter the correspondences by thresholds to remove outliers that would cause inaccuracies in the transformation. These thresholds help ensure geometric consistency by discarding point pairs where the distance between points is too large or the angle between their normals exceeds a certain threshold. Without this filtering, erroneous correspondences could negatively impact the ICP optimization, leading to unstable or inaccurate transformations.

2.2 Linearization

Completed in handwriting below.

2.2 Linearization

Rewrite $r_i(\delta R, \delta t)$ in form of

$$r_i(\alpha, \beta, \gamma, t_x, t_y, t_z) = A_i \begin{bmatrix} \alpha \\ \beta \\ \gamma \\ t_x \\ t_y \\ t_z \end{bmatrix} + b_i$$

$$\delta R = \begin{bmatrix} 1 & -\gamma & \beta \\ \gamma & 1 & -\alpha \\ -\beta & \alpha & 1 \end{bmatrix}$$

linear: $n_{q_i}^T (R p_i' + t - q_i)$

nonlinear: $n_{q_i}^T (\delta R p_i' + \delta t - q_i)$

$$\Rightarrow r(\alpha, \beta, \gamma, t_x, t_y, t_z) = [n_1 \; n_2 \; n_3] \left(\begin{bmatrix} 1 & -\gamma & \beta \\ \gamma & 1 & -\alpha \\ -\beta & \alpha & 1 \end{bmatrix} \begin{bmatrix} p_x' \\ p_y' \\ p_z' \end{bmatrix} + \begin{bmatrix} \delta t_x \\ \delta t_y \\ \delta t_z \end{bmatrix} - \begin{bmatrix} q_x \\ q_y \\ q_z \end{bmatrix} \right)$$

$$= n_1(p_x' - \gamma p_y' + \beta p_z' + \delta t_x - q_x) + n_2(\gamma p_x' - p_y' - \alpha p_x' + \delta t_y - q_y)$$

$$+ n_3(\beta q_x' + \alpha q_y' + p_z' + \delta t_y - q_z)$$

2.2 cont.

$\Rightarrow$

$$r_i = \underbrace{\begin{bmatrix} -n_2 p_z' + n_3 p_y' \\ n_1 p_z' - n_3 p_x' \\ -n_1 p_y' + n_2 p_x' \\ n_1 \\ n_2 \\ n_3 \end{bmatrix}^T}_{A_i} \underbrace{\begin{bmatrix} \alpha \\ \beta \\ \gamma \\ t_x \\ t_y \\ t_z \end{bmatrix}}_{x} + \underbrace{\left[n_1(p_x' - q_x) + n_2(p_y' - q_y) + n_3(p_z' - q_z) \right]}_{b_i}$$

2.3 Optimization

Question 1: Completed in handwriting below and qr method was used in code.

2.3 Optimization

$$\sum r_i^2(\alpha, \beta, \gamma, t_x, t_y, t_z) = \sum \| A_i \begin{bmatrix} \alpha \\ \beta \\ \gamma \\ t_x \\ t_y \\ t_z \end{bmatrix} + b_i \|^2$$

X

$$\Rightarrow \underset{X}{\operatorname{argmin}} \| A_i X + b_i \|^2$$

$$\Rightarrow \frac{\partial}{\partial X} \| A_i x + b_i \|^2 = 0$$

$$A_i^T A_i x + A^T b_i = 0$$

$$\Rightarrow \underline{A_i^T A_i x = - A^T b}$$

& solved with qr method in code.

Question 2:

The results for before and after ICP with the source as frame 10 and the target as frame 50 can be seen below.

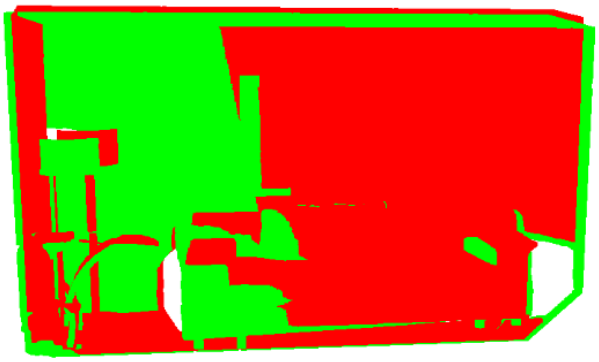

Figure 1: Before ICP (Frame 10)

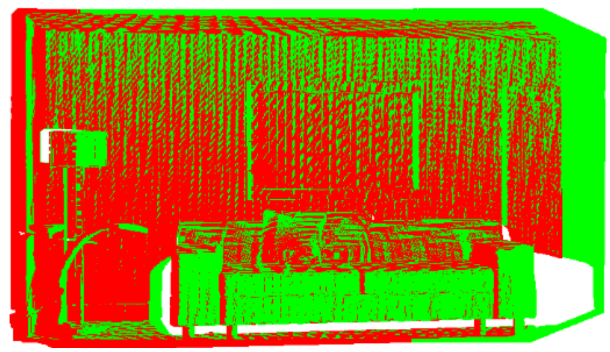

Figure 2: After ICP (Frame 50)

The results for before and after ICP with the source as frame 10 and the target as frame 100 can be seen below.

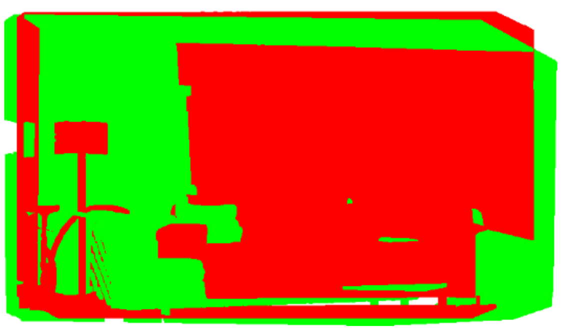

Figure 3: Before ICP (Frame 10)

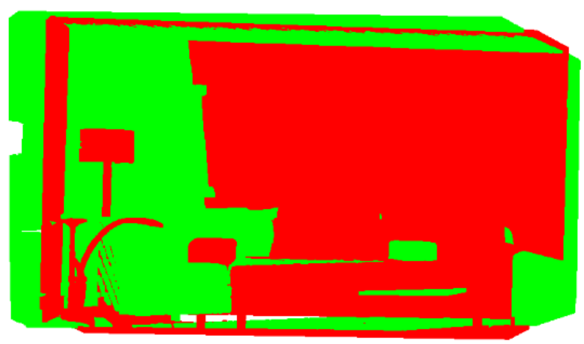

Figure 4: After ICP (Frame 100)

The point cloud registration between frame 10 and frame 50 is successful because the frames are temporally closer, meaning their geometric structures and visual features are more similar. This leads to a higher number of valid correspondences and fewer noisy matches, allowing the ICP algorithm to reduce the loss effectively and produce an optimal transformation.

On the other hand, for frame 10 and frame 100, you can see the registration fails and this is because the frames are farther apart in the sequence, which causes a significant difference in their geometric and visual features. This results in fewer valid correspondences and more noisy matches, which prevents the ICP algorithm from minimizing the loss effectively. Without sufficient valid correspondences, the transformation is suboptimal, and the alignment is poor.

3 Point-Based Fusion

3.1 Filter

The code is included in the submission.

3.2 Merge

The code is included in the submission, and the question is answered on the handwritten next page.

3.2 Merge Question

weight average of positions in p, q, R_c^w, t_c^w, w

$$\bar{p} = \frac{wp + R_c^w q + t_c^w}{w+1}$$

weight average of normals in $n_p, n_q, R_c^w, t_c^w, w$

$$\bar{n}_p = \frac{w n_p + R_c^w n_q}{w+1}$$

3.3 Addition

The code is included in the submission.

3.4 Results

The visualization for the ground truth and the normal map can be seen below. The final number of points obtained in the map was 1,104,379. The frame resolution was 640 x 480, and we used a downsample factor of 2, and we had 200 frames, so the total number of points of all the input was 320 x 240 x 200, which is 15,360,000. The estimated compression ratio is then 1104379/15360000, which is approximately 7.19 %.

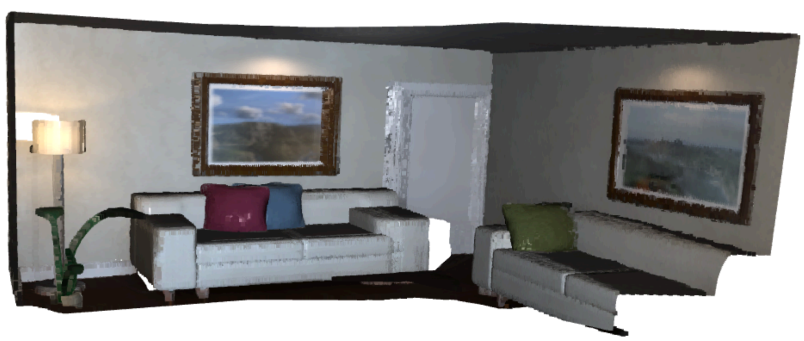

Figure 5: Ground Truth Visualization

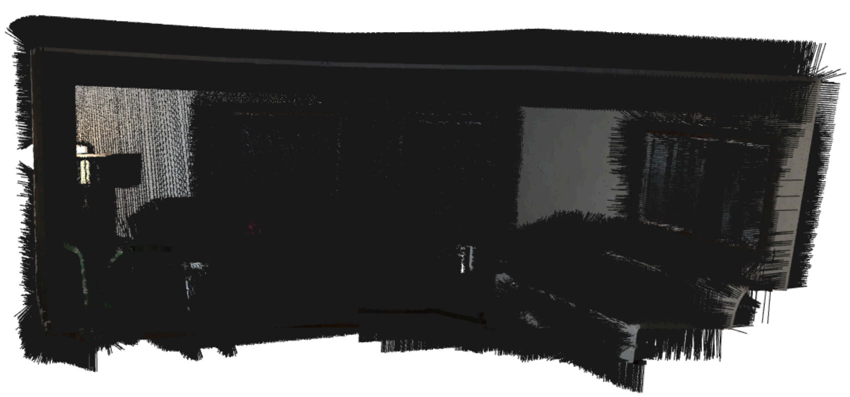

Figure 6: Normal Map Visualization

4 The Dense SLAM System

Question 1:

The source is the current RGB-D frame being aligned, and the target is the map that has been built up to the current time step. The roles cannot be swapped because the map points must be projected onto the vertex map of the current RGB-D frame to establish correspondences. This projection step requires the vertex map of the RGB-D frame, which is readily available. If the roles were reversed, we would not have a vertex map representation of the map, making it impossible to establish correspondences or compute the transformation. Thus, the source must remain the RGB-D frame, and the target must remain the map to ensure proper alignment and mapping consistency.

Question 2:

The fusion visualization can be seen below as well as the project trajectory compared to the ground truth. From the figures below, we can see that there is some drift. We also see that the trajectory follows a similar shape as the ground truth, but over time the errors accumulate and the estimated position is much different from the ground truth position.

Figure 7: Fusion Visualization

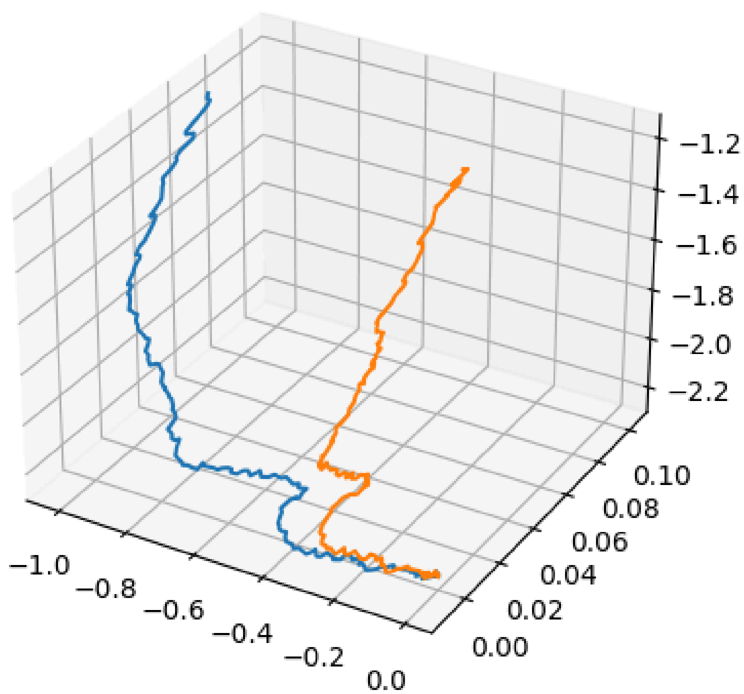

Figure 8: Ground Truth (blue) and Estimated (orange) Trajectories