Supplementary Material: Pose Synchronization under Multiple Pair-wise Relative Poses

In the supplementary material we give proof details of the propositions in the method section and add extra visualization of RGB-D scene recovery results over ScanNet.

1. Proofs of Propositions and Theorems in Section 3

1.1. Proof of Prop. 1

In this proof, we are going to show that when $\alpha > 0$, both $Q_{\alpha,l_{\max},0}^{\{\sigma_R^l\},R,S_i}$ and $Q_{\alpha,l_{\max},0}^{\{\sigma_R^l\},R,S_i}$ are well defined and converge to $Q_{\alpha,\infty,0}^{\{\sigma_R^l\},R,S_i}$, $Q_{\alpha,\infty,0}^{\{\sigma_{\boldsymbol{t}}^l\},\boldsymbol{t},S_i}$ with the following convergence rate:

$$\|Q_{\alpha,l_{\max},0}^{\{\sigma_R^l\},R,S_i} - Q_{\alpha,\infty,0}^{\{\sigma_R^l\},R,S_i}\|_1 \leq \sqrt{\frac{e^{\alpha\lambda_1}}{l_{\max}!}(\alpha\lambda_1)^{l_{\max}}n}, \quad (1)$$

$$\|Q_{\alpha,l_{\max},0}^{\{\sigma_{\boldsymbol{t}}^l\},\boldsymbol{t},S_i} - Q_{\alpha,\infty,0}^{\{\sigma_{\boldsymbol{t}}^l\},\boldsymbol{t},S_i}\|_1 \leq \sqrt{\frac{e^{\alpha\lambda_1}}{l_{\max}!}(\alpha\lambda_1)^{l_{\max}}n}. \quad (2)$$

We first prove that the rotation part $Q_{\alpha,l_{\max},0}^{\{\sigma_R^l\},R,S_i}$ converges to $Q_{\alpha,\infty,0}^{\{\sigma_R^l\},R,S_i}$. For the translation part we apply the same technique.

We begin with a practical assumption that $\{R_{ij}|(i,j) \in \mathcal{E}\}$ can be discretized, meaning that the group in $SO(3)$ generated by $\{I_3\} \cup \{R_{ij}|(i,j) \in \mathcal{E}\}$ is finite. We define $\mathcal{R}$ to be a finite supergroup in $SO(3)$ of the group generated by $\{I_3\} \cup \{R_{ij}|(i,j) \in \mathcal{E}\}$ for the flexibility of discretization resolution, which is implied by $|\mathcal{R}|$. We index the group elements in $\mathcal{R}$ as $\{g_1, ..., g_{|\mathcal{R}|}\}$ and fix $g_1 = I_3$. The group adjacency matrix over the transformation graph $\mathcal{G}$ and group $\mathcal{R}$ is defined as an $n\times n$ block matrix $A_{\mathcal{R}}(\mathcal{G}) \in \mathbb{R}^{n|\mathcal{R}|\times n|\mathcal{R}|}$. Each $|\mathcal{R}|\times|\mathcal{R}|$ block A_{ij} is a $(0,1)$-matrix with $A_{ij}(k,l) = \delta(R_{ji}g_l, g_k)$, where $\delta(\cdot,\cdot)$ is the Kronecker Delta. $A_{\mathcal{R}}(\mathcal{G})$ can be interpreted as an adjacency matrix of a graph $\mathcal{G}_{\mathcal{R}}$ that splits each vertex S_i in $\mathcal{G}$ into $|\mathcal{R}|$ copies $\{S_{ik}\} := \{S_{i1}, ..., S_{i|\mathcal{R}|}\}$, and all the edges e_{ij} in $\mathcal{G}$ are also split into $|\mathcal{R}|$ copies for a complete bipartite matching between $\{S_{ik}\}$ and $\{S_{jk}\}$.

Given a fixed finite σ_R, $P^{\sigma_R,R}(\cdot)$ becomes a Gaussian distribution with only one parameter R, and the un-normalized distribution of R_p over all the paths p with a given length $|p| = l$ is a mixture of gaussian distributions with different means and weights, but identical standard deviation $\sqrt{l}\sigma_R$. Thus this Gaussian mixture distribution can be encoded with a block vector $X^{(l)} \in \mathbb{R}^{n|\mathcal{R}|\times 1}$. Each block $X_i^{(l)}$ of length $|\mathcal{R}|$ encodes the distribution of the means $R_p \in \mathcal{R}$ for $p \in \mathcal{P}_{1i}^{(l)}$, and each entry $X_i^{(l)}(j)$ in the block indicates the weight of g_j. For $l = 0$, we have one candidate $g_1 = I_3$ for the rotation from S_1 to itself, and thus we setup the initialization and iteration for computing the distribution of R_p of all the path lengths as

$$\begin{aligned} X^{(0)} &= e_1, \\ X^{(l)} &= A_{\mathcal{R}}(\mathcal{G})X^{(l-1)}, \quad l \geq 1, \end{aligned} \quad (3)$$

Where e_1 is the one-hot vector in its first entry. For simplicity we assume that $\sigma_R = \infty$, and $Q_{\alpha,l_{\max},0}^{\{\sigma_R^l\},R,S_i}(\cdot)$ can be computed with

$$\begin{aligned} Q_{\alpha,l_{\max},0}^{\{\sigma_R^l\},R,S_i}(g_j) &= \sum_{l=1}^{l_{\max}} \frac{\alpha^l}{l!} X_i^{(l)}(j) \\ &= \left(\sum_{l=1}^{l_{\max}} \frac{(\alpha A_{\mathcal{R}}(\mathcal{G}))^l}{l!} X^{(0)}\right)_i (j), \end{aligned} \quad (4)$$

and the residual is

$$\begin{aligned} \Delta_{l_{\max}} :&= Q_{\alpha,\infty,0}^{\{\sigma_R^l\},R,S_i} - Q_{\alpha,l_{\max},0}^{\{\sigma_R^l\},R,S_i} \\ &= \sum_{l=l_{\max}}^{\infty} \frac{(\alpha A_{\mathcal{R}}(\mathcal{G}))^l}{l!} X^{(0)}. \end{aligned} \quad (5)$$

Since $A_{\mathcal{R}}(\mathcal{G})$ is real symmetric, let $A_{\mathcal{R}}(\mathcal{G}) = U\Lambda U^T$ denote the diagonalization of $A_{\mathcal{R}}(\mathcal{G})$, with the eigenvalues $\lambda_1, ..., \lambda_{n|\mathcal{R}|}$ distributed in the diagonal of Λ in descending order. (5) is then reformed as

$$\Delta_{l_{\max}} = U\left(\sum_{l=l_{\max}}^{\infty} \frac{(\alpha\Lambda)^l}{l!}\right) U^T X^{(0)}. \quad (6)$$

To show that (6) converges, it suffices to prove the convergence of the infinite series $\left\{\frac{(\alpha\lambda_k)^l}{l!}\right\}, 1 \leq k \leq n|\mathcal{R}|$.

Note that

$$\sum_{l=0}^{\infty} \frac{(\alpha\lambda_k)^l}{l!} = e^{\alpha\lambda_k} \quad (7)$$

is finite for every k, which gives $\lim_{l_{\max}\to\infty} \Delta_{l_{\max}} = 0$.

Next we approximate the convergence rate of $\Delta_{l_{\max}}$. Since $A_{\mathcal{R}}(\mathcal{G})$ is the adjacency matrix of $\mathcal{G}_{\mathcal{R}}$, Λ is bounded by

$$d(\mathcal{G}) = d(\mathcal{G}_{\mathcal{R}}) \geq \lambda_1 \geq -\lambda_{n|\mathcal{R}|}, \quad (8)$$

Where $d(\cdot)$ is the graph degree. Moreover, $\mathcal{G}_R$ can be viewed as $|\mathcal{R}|$ identical copies of $\mathcal{G}$, which means λ_1 also equals the spectral norm of the adjacency matrix of $\mathcal{G}$. Apply the Taylor remainder theorem, we have

$$\begin{aligned} \left| \sum_{l=l_{\max}}^{\infty} \frac{(\alpha\lambda_k)^l}{l!} \right| &\leq \left| \frac{\max(1, e^{\alpha\lambda_k})}{l_{\max}!} (\alpha\lambda_k)^{l_{\max}} \right| \\ &\leq \frac{e^{\alpha\lambda_1}}{l_{\max}!} (\alpha\lambda_1)^{l_{\max}}. \end{aligned} \quad (9)$$

Since $X^{(0)}$ is a unit vector, we have

$$\|\Delta_{l_{\max}}\|_2 \leq \frac{e^{\alpha\lambda_1}}{l_{\max}!} (\alpha\lambda_1)^{l_{\max}}. \quad (10)$$

(10) holds for arbitrary discretization resolution $|\mathcal{R}|$ and in the remaining part of this proof we consider the case where the group $\mathcal{R}$ generated by $\{I_3\} \cup \{R_{ij} | (i,j) \in \mathcal{E}\}$ is infinite, with $\hat{\mathcal{R}}$ being a discrete approximation. Let

$$\hat{R_{ij}} := \underset{\hat{R}\in\hat{\mathcal{R}}}{\operatorname{argmin}} \|\hat{R} - R_{ij}\|_{\mathcal{F}} \quad (11)$$

be the projection of R_{ij} on $\hat{\mathcal{R}}$ and $\hat{Q}$ be the distribution generated by $\{\hat{R_{ij}}\}, (i,j) \in \mathcal{E}$. Apply (10) to $\hat{Q}$ and we know that $\hat{Q}_{\alpha,l_{\max},0}^{\{\sigma_R^l\},R,S_i}$ converges to $\hat{Q}_{\alpha,\infty,0}^{\{\sigma_R^l\},R,S_i}$.

We define the difference between $\hat{Q}_{\alpha,l_{\max},0}^{\{\sigma_R^l\},R,S_i}$ and $\hat{Q}_{\alpha,\infty,0}^{\{\sigma_R^l\},R,S_i}$ as $\left(\hat{\Delta}_{l_{\max}}\right)_i$. In the above we view $\left(\hat{\Delta}_{l_{\max}}\right)_i$ as the i-th block of the vector $\hat{\Delta}_{l_{\max}}$ indicating the strength at $\forall \hat{g}_j \in \hat{\mathcal{R}}$. Now instead, we consider it as a function $\left(\hat{\Delta}_{l_{\max}}\right)_i(\cdot) : SO^3 \to R$, composed of weighted Dirac delta functions at $\forall g_j \in \hat{\mathcal{R}}$ with $\sigma^R = \infty$. Suppose that $(\Delta_{l_{\max}})_i(\cdot)$ is also well defined with respect to Q. We have

$$\int_{R\in SO(3)} (\Delta_{l_{\max}})_i (R) dR = \int_{R\in SO(3)} \left(\hat{\Delta}_{l_{\max}}\right)_i (R) dR, \quad (12)$$

since on both sides we are counting over all the possible paths of length greater than $l_{\max}$. Given that $|\hat{\mathcal{R}}|$ is a finite constant, combining (10) and (12) shows that

$$\sum_{i=1}^{n} \| (\Delta_{l_{\max}})_i \|_1 \leq \sqrt{\frac{e^{\alpha\lambda_1}}{l_{\max}!} (\alpha\lambda_1)^{l_{\max}} n |\mathcal{R}|}. \quad (13)$$

In fact, (12) holds for arbitrary σ^R since the 1-norm of any probability distribution function equals 1. For the extreme case where we set $|\mathcal{R}| = 1$, we finish the proof with

$$\sum_{i=1}^{n} \| (\Delta_{l_{\max}})_i \|_1 \leq \sqrt{\frac{e^{\alpha\lambda_1}}{l_{\max}!} (\alpha\lambda_1)^{l_{\max}} n}. \quad (14)$$

□

1.2. Proof of Prop. 2

We hereby recall Prop. 2 first:

Suppose all T_{ij} are independent and identically follow a normal distribution $R \sim \mathcal{N}(\mu_R, \sigma_R), \boldsymbol{t} \sim \mathcal{N}(\mu_{\boldsymbol{t}}, \sigma_{\boldsymbol{t}})$ with probability w_1 and a uniform distribution with probability w_2. Let P_1, P_2 be the PDFs of the two distributions. If $(1 - w_1^{\alpha\lambda_1})P_2(t) \ll 1$, then almost surely the global maximums of $Q_{\alpha,\infty,0}^{\{\sigma_R^l\},R,S_i}$ and $Q_{\alpha,\infty,0}^{\{\sigma_{\boldsymbol{t}}^l\},\boldsymbol{t},S_i}$ are with in the Δ-neighborhood of the ground-truth, where

$$\Delta \sim O(\lambda_1^{-\frac{1}{6}} (\ln \alpha\lambda_1)^{\frac{2}{3}}).$$

Moreover, this neighborhood is almost convex when we choose $\sigma_R, \sigma_{\boldsymbol{t}} \sim O(\Delta)$.

In this prove we focus on the property of translations $\boldsymbol{t}$ and for rotations R a similar technique can be applied.

Without losing generality, we consider the case where $\boldsymbol{t}_{ij}^{\star} = \mathbf{0}, \forall (i,j) \in \mathcal{E}$, meaning that the input $\boldsymbol{t}_{ij}$ is a distribution

$$\boldsymbol{t}_{ij} = \mathcal{D}_{\boldsymbol{t}} := w_1 \mathcal{N}(\mathbf{0}, \sigma_1^2 I) + w_2 \mathcal{U}(d), \quad (15)$$

where $\mathcal{U}(d)$ denotes a uniform distribution over the ball in $\mathbb{R}^3$ of radius d. However, directly taking the distribution as an input can cause ambiguity in defining the accumulated translation along a path p. Therefore we still start from a discrete setting. Moreover, since $\boldsymbol{t}$ is independent along each dimension, we consider its component along one dimension and call it t.

Let $\mathcal{R}, A_{\mathcal{R}}(G), X^{(0)}$ be defined the same as in the proof of Prop **??**. For each block A_{ij}, if we denote the translation implied by the (u,v)-th entry as $t(u,v)$, then $A_{ij}(u,v)$ is assigned the value $\frac{1}{Z} P_{t\sim\mathcal{D}_t}(t(u,v))$, where Z is a global regularization coefficient. It is obvious that all the blocks A_{ij} for $(i,j) \in \mathcal{E}$ are identical, which we denote as $D_{t,\mathcal{R}}$ and as the resolution goes to infinity, A converges to $\mathcal{D}_t$. In fact, if $A(G)$ denotes the adjacency matrix of G, then

$$A_{\mathcal{R}}(G) = A(G) \otimes D_{t,\mathcal{R}}. \quad (16)$$

We can view the construction of $A_{\mathcal{R}}(G)$ as fixing a set of evenly sampled $\{t(u,v)\}$ in distribution $\mathcal{D}_t$. For a path $p = (e_1, e_2, ..., e_{|p|})$ with finite length $|p|$, each $t_(e_i), 1 \leq i \leq$

$|p|$ takes value from $\{t(u,v)\}$ with probability $D_{t,\mathcal{R}}(u,v)$. By enforcing a high resolution $|\mathcal{R}|$, we can approximate the distribution of the translation $\boldsymbol{t}(p)$ along any path p with a bounded length $|p| \leq l_{\max}$ as

$$P(t(p)) = \prod_{i=1}^{|p|} P_{t\sim\mathcal{D}_t}(t(e_i)). \quad (17)$$

Thus, for any $l \leq l_{\max}$, $X^{(l)} = A_{\mathcal{R}}(G)^l X^{(0)}$ encodes the distribution of all paths with length l. For the k-th block $X_k^{(l)}$, $|\mathcal{P}_{1k}^{(l)} = \|X_k^{(l)}\|_1$ is the number of paths p from S_1 to S_k with length l and $X_k^{(l)}/\|X_k^{(l)}\|_1$ is the distribution from which all such paths p are i.i.d. sampled. To be specific,

$$X_k^{(l)}/\|X_k^{(l)}\|_1 = (D_{t,\mathcal{R}})^l e_1. \quad (18)$$

Next we bound the difference between the distribution of sampled p and the real distribution $(D_{t,\mathcal{R}})^l e_1$. In the context of discretized distribution, we set a label $r(p)$ to a path p if $t(p)$ falls into the $r(p)$-th bin among all the $|\mathcal{R}|$ bins, and $P_r := (D_{t,\mathcal{R}})^l e_1(r)$ be the probability that p falls into bin r. We index all the paths in $\mathcal{P}_{1k}^{(l)}$ as $p_1, p_2, ..., p_m$. Let $Y_{i,r}$ be the indicator variable of $r(p_i) = r$. Apply Chernoff bound over $Y_{i,r}, 1 \leq i \leq m$ and any $\delta > 0$:

$$P(|\sum_{i=1}^m Y_{i,r} - mP_r| > \delta m P_r) < 2\exp\frac{-mP_r\delta^2}{3}. \quad (19)$$

Then apply union bound over all r, we have

$$P(|\sum_{i=1}^m Y_{i,r} - mP_r| < \delta m P_r, \forall r) > 1 - 2\sum_{r=1}^{|\mathcal{R}|}\exp\frac{-mP_r\delta^2}{3}. \quad (20)$$

To analyze the concentration of $t(p_i), 1 \leq i \leq m$, suppose that bin 1 has a width of 2Δ, i.e., for any p such that $r(p) = 1$, we have $-\Delta < t(p) < \Delta$. Compare the mass of the central bin with bin r in the distribution of $t(p)$, we have

$$\frac{P_{r=1}}{P_r} = \frac{\int_{-\Delta}^{\Delta}\left(w_1^l \frac{1}{\sqrt{2\pi}\sigma_1} e^{\frac{-t^2}{2l\sigma_1^2}} + C(1-w_1^l)\right)dt}{\int_{(2r-1)\Delta}^{(2r+1)\Delta}\left(w_1^l \frac{1}{\sqrt{2\pi}\sigma_1} e^{\frac{-t^2}{2l\sigma_1^2}} + C(1-w_1^l)\right)dt}, \quad (21)$$

Where C is a small constant to include the uniform part of the distribution of $t(p)$. When $\Delta \ll d$, (21) is bounded by

$$\frac{P_{r=0}}{P_r} \geq e^{\frac{(2r-1)^2\Delta^2}{2l\sigma_1^2}}. \quad (22)$$

In order to cluster $t(p)$ for $p \in \mathcal{P}_{1k}^{(l)}$, we apply a gaussian kernel with standard deviation σ_2 at each sampled $t(p)$ and get a Gaussian mixture model

$$g^{(l)}(t) = \sum_{p\in\mathcal{P}_{1k}^{(l)}} e^{-\frac{(t-t(p))^2}{2\sigma_2^2}}. \quad (23)$$

For any $-\Delta \leq t \leq \Delta$,

$$g^{(l)}(t) \geq \sum_{\substack{p\in\mathcal{P}_{1k}^{(l)},\\ r(p)=1}} e^{-\frac{4\Delta^2}{2\sigma_2^2}}. \quad (24)$$

Moreover, for any $|t| \geq 5\Delta$,

$$g^{(l)}(t) \leq \sum_{\substack{p\in\mathcal{P}_{1k}^{(l)},\\ r(p)=1}} e^{-\frac{16\Delta^2}{2\sigma_2^2}} + \sum_{\substack{p\in\mathcal{P}_{1k}^{(l)},\\ r(p)\neq 1}} 1. \quad (25)$$

Combining (22) with (25), it is almost sure that

$$g^{(l)}(t) \leq \sum_{\substack{p\in\mathcal{P}_{1k}^{(l)},\\ r(p)=1}} \left(e^{-\frac{16\Delta^2}{2\sigma_2^2}} + \frac{1+\delta}{1-\delta}\sum_{r=3}^{|\mathcal{R}|} e^{-\frac{(2r-1)^2\Delta^2}{2l\sigma_1^2}}\right). \quad (26)$$

Therefore when $\sigma_2^2 \geq l\sigma_1^2$ and $\Delta > \sqrt{\frac{\ln(1+2\delta)}{25}l\sigma_1}$, the global maximum of $g^{(l)}(t)$ is in $(-5\Delta, 5\Delta)$.

Since we do not know the ground truth of $t(p)$, it remains to find out where bin 1 lies in. With a given sample set $t(P_1), ..., t(p_m)$. let W denote the event that we successfully assign all the bins such that $|\sum_{i=1}^m Y_{i,r} - mP_r| < \delta m P_r$ holds for every r. As long as the event in the left-hand-side of 20 happens, $P(W) = 1$ and bin 1 is assigned with at most Δ error. If we discard every $|t(p_i)| \geq 6\Delta, 1 \leq i \leq m$ and set $\sigma_2 \geq 6\Delta$, then the global maximum of $g^{(l)}(t)$ can be reached by gradient descend from any remaining initial point $t(p)$.

Next we generalize the above result to all the paths p from S_1 to S_k with length $|p| \leq l_{\max}$. The number of such paths is bounded by

$$c\lambda_1^l - b \leq \sum_{l=1}^{l_{\max}} |\mathcal{P}_{1k}^l| \leq c\lambda_1^l + b \quad (27)$$

for some constant b, c, where λ_1 is the largest eigenvalue of $A(G)$. For (20) to happen with high probability over all l, we require

$$2\sum_{l<l_{\max}}\sum_{r=1}^{|\mathcal{R}|}\exp\frac{-c\lambda_1^l P_r \delta(l)^2}{3} \leq e^{-c_1} \quad (28)$$

for some $c_1 > 0$, leading to

$$\delta(l) \sim O(\lambda_1^{-\frac{l}{2}} e^{\frac{|\mathcal{R}|}{2l\sigma_1^2}} \ln l_{\max}). \quad (29)$$

Therefore we require

$$\Delta \sim O(\lambda_1^{-\frac{l}{4}} e^{\frac{|\mathcal{R}|}{4l\sigma_1^2}} \ln l_{\max}\sqrt{l}\sigma_1) = O(\lambda_1^{-\frac{l}{4}} \ln l_{\max}\sqrt{|\mathcal{R}|}). \quad (30)$$

Recall that in the proof of Prop. 1, we require $l_{\max} \sim O(\alpha\lambda_1)$ to make all the paths of length greater than $l_{\max}$ negligible. Since $|\mathcal{R}| \sim O(1/\Delta)$, we have $\Delta \sim O(\lambda_1^{-\frac{1}{6}}(\ln(\alpha\lambda))^{\frac{2}{3}})$.

1.3. Proof of Prop. 3

The derivative of Q^T with respect to $\boldsymbol{t}'$ is given by

$$\frac{\partial Q^T}{\partial \boldsymbol{t}'}(T') = -\sum_{k=1}^{K} \frac{w_k}{\sigma_{k,\boldsymbol{t}}^2} e^{-d^2_{\sigma_{k,R},\sigma_{k,\boldsymbol{t}}}(T',T_k)}(\boldsymbol{t}' - \boldsymbol{t}_k)$$

There

$$\frac{\partial Q^T}{\partial \boldsymbol{t}'}(T') = 0 \rightarrow \boldsymbol{t}' = \frac{\sum\limits_{k=1}^{K} \frac{w_k}{\sigma_{k,\boldsymbol{t}}^2} e^{-d^2_{\sigma_{k,R},\sigma_{k,\boldsymbol{t}}}(T',T_k)} \boldsymbol{t}_k}{\sum\limits_{k=1}^{K} \frac{w_k}{\sigma_{k,\boldsymbol{t}}^2} e^{-d^2_{\sigma_{k,R},\sigma_{k,\boldsymbol{t}}}(T',T_k)}}.$$

To compute the derivative of Q^T with respec to R', we parameterize rotations near a current rotation R' as

$$R = \exp(\boldsymbol{c}\times)R', \qquad \forall \boldsymbol{c} \in \mathbb{R}^3.$$

Under this parameterization, we have $\forall 1 \le i \le 3$

$$\begin{aligned}
&\boldsymbol{e}_i^T \cdot \frac{\partial Q^T}{\partial \boldsymbol{c}}(T') \\
=& -\sum_{k=1}^{K} \frac{w_k}{\sigma_{k,R}^2} e^{-d^2_{\sigma_{k,R},\sigma_{k,\boldsymbol{t}}}(T',T_k)} \langle R', \boldsymbol{e}_i \times R' \rangle \\
&+ \sum_{k=1}^{K} \frac{w_k}{\sigma_{k,R}^2} e^{-d^2_{\sigma_{k,R},\sigma_{k,\boldsymbol{t}}}(T',T_k)} \langle R_k, \boldsymbol{e}_i \times R' \rangle \\
=& -\sum_{k=1}^{K} \frac{w_k}{\sigma_{k,R}^2} e^{-d^2_{\sigma_{k,R},\sigma_{k,\boldsymbol{t}}}(T',T_k)} \langle R'R'^T, \boldsymbol{e}_i \times \rangle \\
&+ \sum_{k=1}^{K} \frac{w_k}{\sigma_{k,R}^2} e^{-d^2_{\sigma_{k,R},\sigma_{k,\boldsymbol{t}}}(T',T_k)} \langle R_k, \boldsymbol{e}_i \times R' \rangle \\
=& \sum_{k=1}^{K} \frac{w_k}{\sigma_{k,R}^2} e^{-d^2_{\sigma_{k,R},\sigma_{k,\boldsymbol{t}}}(T',T_k)} \langle R_k, \boldsymbol{e}_i \times R' \rangle \\
=& \langle U\Sigma V^T, \boldsymbol{e}_i \times R' \rangle \\
=& \langle \Sigma, U^T \boldsymbol{e}_i \times UU^T R' V \rangle \\
=& \langle \Sigma, U^T \boldsymbol{e}_i \times U\overline{R} \rangle, \quad \overline{R} = U^T R' V
\end{aligned} \tag{31}$$

Denote $U = (\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3)^T$. It is easy to check that

$$U^T(\boldsymbol{e}_i \times)U = \boldsymbol{u}_i \times$$

Therefore, applying (31), we have that if $\frac{\partial Q^T}{\partial R'}(T') = 0$, then

$$0 = \langle \Sigma, \boldsymbol{u}_i \times \overline{R} \rangle, \quad 1 \le i \le 3. \tag{32}$$

Denote $\boldsymbol{u}_i = (u_{i1}, u_{i2}, u_{i3})^T$, $\overline{R} = (\overline{r}_{ij})_{1 \le i,j \le 3}$, and $\Sigma = \mathrm{diag}(\sigma_1, \sigma_2, \sigma_3)$. Expanding (32), we arrive at $\forall 1 \le i \le 3$

$$u_{i1}(\overline{r}_{23}\sigma_3 - \overline{r}_{32}\sigma_2) + u_{i2}(\overline{r}_{31}\sigma_1 - \overline{r}_{13}\sigma_3) + u_{i3}(\overline{r}_{12}\sigma_2 - \overline{r}_{21}\sigma_1) = 0.$$

It follows that

$$\overline{r}_{ij}\sigma_j = \overline{r}_{ji}\sigma_i, \forall 1 \le i < j \le 3. \tag{33}$$

Parameterize $\overline{R}$ using the axis-angle parameterization

$$\overline{R} = \cos(\overline{\theta})I_3 + (1 - \cos(\overline{\theta}))\overline{\boldsymbol{n}}\overline{\boldsymbol{n}}^T + \sin(\overline{\theta})\overline{\boldsymbol{n}}\times \tag{34}$$

where $\overline{\boldsymbol{n}} = (\overline{n}_1, \overline{n}_2, \overline{n}_3)^T$ and $\overline{\theta}$ are the rotation axis and rotation angle respectively.

Substituting (34) into (33), we obtain

$$(\sigma_2 - \sigma_1)\overline{n}_1\overline{n}_2(1 - \cos(\overline{\theta})) = (\sigma_1 + \sigma_2)\overline{n}_3 \sin(\overline{\theta}) \tag{35}$$
$$(\sigma_1 - \sigma_3)\overline{n}_1\overline{n}_3(1 - \cos(\overline{\theta})) = (\sigma_1 + \sigma_3)\overline{n}_2 \sin(\overline{\theta}) \tag{36}$$
$$(\sigma_3 - \sigma_2)\overline{n}_2\overline{n}_3(1 - \cos(\overline{\theta})) = (\sigma_2 + \sigma_3)\overline{n}_1 \sin(\overline{\theta})$$

We show that $\overline{\theta} = 0$. This means $\overline{R} = I_3$ or $R' = UV^T$, which ends the proof. Suppose $\overline{\theta} \neq 0$. It follows that $\overline{n}_i \neq 0, 1 \le i \le 3$. Otherwise $\sin(\overline{\theta}) = 0$ as $\max(|\overline{n}_i|) > 0$. When $\overline{n}_i \neq 0, 1 \le i \le 3$, we have $\sigma_i \neq \sigma_j, \forall 1 \le i < j \le 3$. Without losing generality, we assume $\sigma_1 > \sigma_2 > \sigma_3$. In this case, factoring out $\overline{\theta}$ in (35) to (36), we have

$$\frac{\sigma_2 - \sigma_1}{\sigma_1 - \sigma_3} = \frac{|\overline{n}_3|^2(\sigma_2 + \sigma_1)}{|\overline{n}_2|^2(\sigma_3 + \sigma_1)} \tag{37}$$

which results in a contradiction as the left side (37) is negative while its right side is positive.

□

1.4. Proof of Prop. 4

We begin with simplifying the objective function of (15):

$$\begin{aligned}
&\int_{T' \in \mathbb{R}^{3\times 4}} \Big(w e^{-d^2_{\sigma_R,\sigma_{\boldsymbol{t}}}(T',T_j^\star)} - \sum_{k \in \mathcal{C}_j} w_k e^{-d^2_{\sigma_{k,R},\sigma_{k,\boldsymbol{t}}}(T',T_k)} \Big)^2 \\
=& w^2 \cdot \int_{R' \in \mathbb{R}^{3\times 3}} e^{-\frac{\|R' - R_j^\star\|_{\mathcal{F}}^2}{\sigma_R^2}} \int_{\boldsymbol{t}' \in \mathbb{R}^3} e^{-\frac{\|\boldsymbol{t}' - \boldsymbol{t}_j^\star\|^2}{\sigma_{\boldsymbol{t}}^2}} + \sum_{k,k' \in \mathcal{C}_j} w_k w_{k'} \\
&\int_{T' \in \mathbb{R}^{3\times 4}} e^{-d^2_{\sigma_{k,R},\sigma_{k,\boldsymbol{t}}}(T',T_k) - d^2_{\sigma_{k',R},\sigma_{k',\boldsymbol{t}}}(T',T_{k'})} \\
&- 2w \sum_{k \in \mathcal{C}_j} w_k \int_{R' \in \mathbb{R}^{3\times 3}} e^{-\frac{\|R' - R_j^\star\|_{\mathcal{F}}^2}{2\sigma_R^2} - \frac{\|R' - R_k\|_{\mathcal{F}}^2}{2\sigma_{k,R}^2}} \\
&\int_{\boldsymbol{t}' \in \mathbb{R}^3} e^{-\frac{\|\boldsymbol{t}' - \boldsymbol{t}_j^\star\|_{\mathcal{F}}^2}{2\sigma_{\boldsymbol{t}}^2} - \frac{\|\boldsymbol{t}' - \boldsymbol{t}_k\|^2}{2\sigma_{k,\boldsymbol{t}}^2}}
\end{aligned}$$

$$=w^2\sigma_R^9\sigma_{\boldsymbol{t}}^3\pi^6 + \text{const}$$
$$- 2w\sum_{k\in\mathcal{C}_j} w_k \int_{R'} e^{-\frac{\sigma_R^2+\sigma_{k,R}^2}{2\sigma_R^2\sigma_{k,R}^2}\|R' - \frac{R_j^\star\sigma_{k,R}^2 + R_k\sigma_R^2}{\sigma_R^2+\sigma_{k,R}^2}\|_{\mathcal{F}}^2 - \frac{\|R_j^\star - R_k\|_{\mathcal{F}}^2}{2(\sigma_R^2+\sigma_{k,R}^2)}}$$
$$\int_{\boldsymbol{t}'} e^{-\frac{\sigma_{\boldsymbol{t}}^2+\sigma_{k,\boldsymbol{t}}^2}{2\sigma^2\sigma_k^2}\|\boldsymbol{t}' - \frac{\boldsymbol{t}_j^\star\sigma_{k,\boldsymbol{t}}^2 + \boldsymbol{t}_k\sigma_{\boldsymbol{t}}^2}{\sigma_{\boldsymbol{t}}^2+\sigma_{k,\boldsymbol{t}}^2}\|^2 - \frac{\|\boldsymbol{t}_j^\star - \boldsymbol{t}_k\|^2}{2(\sigma_{\boldsymbol{t}}^2+\sigma_{k,\boldsymbol{t}}^2)}}$$

$$=w^2\sigma_R^9\sigma_{\boldsymbol{t}}^3\pi^6 + \text{const} - 2w\sum_{k\in\mathcal{C}_j}\Big(w_k(2\pi)^6\frac{\sigma_R^9\sigma_{k,R}^9}{(\sigma_R^2+\sigma_{k,R}^2)^{\frac{9}{2}}}$$
$$\frac{\sigma_{\boldsymbol{t}}^3\sigma_{k,\boldsymbol{t}}^3}{(\sigma_{\boldsymbol{t}}^2+\sigma_{k,\boldsymbol{t}}^2)^{\frac{3}{2}}}\cdot e^{-\frac{\|R_j^\star - R_k\|_{\mathcal{F}}^2}{2(\sigma_R^2+\sigma_{k,R}^2)} - \frac{\|\boldsymbol{t}_j^\star - \boldsymbol{t}_k\|^2}{2(\sigma_{\boldsymbol{t}}^2+\sigma_{k,\boldsymbol{t}}^2)}}\Big) \tag{38}$$

For fixed σ_R and $\sigma_{\boldsymbol{t}}$, it is clear that the optimal w is given by

$$w^\star = \sum_{k\in\mathcal{C}_j} w_k\big(\frac{2\sigma_{k,R}^2}{(\sigma_R^2+\sigma_{k,R}^2)}\big)^{\frac{9}{2}}\big(\frac{2\sigma_{k,\boldsymbol{t}}^2}{(\sigma_{\boldsymbol{t}}^2+\sigma_{k,\boldsymbol{t}}^2)}\big)^{\frac{3}{2}}$$
$$\cdot e^{-\frac{\|\boldsymbol{t}_j^\star - \boldsymbol{t}_k\|^2}{2(\sigma_{\boldsymbol{t}}^2+\sigma_{k,\boldsymbol{t}}^2)}} e^{-\frac{\|\boldsymbol{t}_j^\star - \boldsymbol{t}_k\|^2}{2(\sigma_{\boldsymbol{t}}^2+\sigma_{k,\boldsymbol{t}}^2)}} \tag{39}$$

Substituting (39) into (38), we have that the optimal $\sigma_R^\star$ and $\sigma_{\boldsymbol{t}}$ are given by

$$\sigma_R^\star, \sigma_{\boldsymbol{t}}^\star = \arg\max_{\sigma_R,\sigma_{\boldsymbol{t}}} \sum_{k\in\mathcal{C}_j} w_k\big(\frac{2\sigma_{\boldsymbol{t}}\sigma_{k,\boldsymbol{t}}^2}{(\sigma_{\boldsymbol{t}}^2+\sigma_{k,\boldsymbol{t}}^2)}\big)^{\frac{3}{2}} e^{-\frac{\|\boldsymbol{t}_j^\star - \boldsymbol{t}_k\|^2}{2(\sigma_{\boldsymbol{t}}^2+\sigma_{k,\boldsymbol{t}}^2)}}$$
$$\big(\frac{2\sigma_R\sigma_{k,R}^2}{(\sigma_R^2+\sigma_{k,R}^2)}\big)^{\frac{9}{2}} e^{-\frac{\|R_j^\star - R_k\|_{\mathcal{F}}^2}{2(\sigma_R^2+\sigma_{k,R}^2)}} \tag{40}$$

□

1.5. Details on Variance Optimization of (40)

We solve (40) via alternating maximization, starting from $\sigma_R = \underset{k\in\mathcal{C}_j}{\text{median}}(\sigma_{k,R})$ and $\sigma_{\boldsymbol{t}} = \underset{k\in\mathcal{C}_j}{\text{median}}(\sigma_{k,\boldsymbol{t}})$. When $\sigma_{\boldsymbol{t}}$ is fixed, the optimization problem reduces to

$$\max_{\sigma_R} h_{\sigma_{\boldsymbol{t}}}(\sigma_R) \tag{41}$$

where

$$h_{\sigma_{\boldsymbol{t}}}(\sigma_R) := \sum_{k\in\mathcal{C}_j} w_{k,\sigma_{\boldsymbol{t}}}\cdot\big(\frac{2\sigma_R\sigma_{k,R}^2}{\sigma_R^2+\sigma_{k,R}^2}\big)^{\frac{9}{2}} e^{-\frac{\|R_j^\star - R_k\|_{\mathcal{F}}^2}{2(\sigma_R^2+\sigma_{k,R}^2)}}$$
$$w_{k,\boldsymbol{t}} := w_k\cdot\big(\frac{2\sigma_{\boldsymbol{t}}\sigma_{k,\boldsymbol{t}}^2}{\sigma_{\boldsymbol{t}}^2+\sigma_{k,\boldsymbol{t}}^2}\big)^{\frac{3}{2}} e^{-\frac{\|\boldsymbol{t}_j^\star - \boldsymbol{t}_k\|^2}{2(\sigma_{\boldsymbol{t}}^2+\sigma_{k,\boldsymbol{t}}^2)}}.$$

Computing the derivative of $h_{\sigma_{\boldsymbol{t}}}(\sigma_R)$ with respect to σ_R, we can see that σ_R is a critical point of $h_{\sigma_{\boldsymbol{t}}}(\sigma_R)$ if

$$0 = \sum_{k\in\mathcal{C}_j} w_{k,\sigma_{\boldsymbol{t}}}\cdot\big(\frac{2\sigma_R\sigma_{k,R}^2}{\sigma_R^2+\sigma_{k,R}^2}\big)^{\frac{7}{2}} e^{-\frac{\|R_j^\star - R_k\|_{\mathcal{F}}^2}{2(\sigma_R^2+\sigma_{k,R}^2)}}$$
$$\Big(\frac{9\sigma_{k,R}^2(\sigma_{k,R}^2-\sigma_R^2)}{(\sigma_R^2+\sigma_{k,R}^2)^2} + \frac{2\sigma_R^2\sigma_{k,R}^2\|R_j^\star - R_k\|_{\mathcal{F}}^2}{(\sigma_R^2+\sigma_{k,R}^2)^3}\Big) \tag{42}$$

Denote

$$c_{k,\sigma_R,1} = \big(\frac{2\sigma_R\sigma_{k,R}^2}{\sigma_R^2+\sigma_{k,R}^2}\big)^{\frac{7}{2}} e^{-\frac{\|R_j^\star - R_k\|_{\mathcal{F}}^2}{2(\sigma_R^2+\sigma_{k,R}^2)}}\frac{9\sigma_{k,R}^2}{(\sigma_R^2+\sigma_{k,R}^2)^2},$$
$$c_{k,\sigma_R,2} = \big(\frac{2\sigma_R\sigma_{k,R}^2}{\sigma_R^2+\sigma_{k,R}^2}\big)^{\frac{7}{2}} e^{-\frac{\|R_j^\star - R_k\|_{\mathcal{F}}^2}{2(\sigma_R^2+\sigma_{k,R}^2)}}$$
$$\cdot\Big(\frac{9\sigma_{k,R}^4}{(\sigma_R^2+\sigma_{k,R}^2)^2} + \frac{2\sigma_{k,R}^2\sigma_R^2\|R_j^\star - R_k\|_{\mathcal{F}}^2}{(\sigma_R^2+\sigma_{k,R}^2)^3}\Big).$$

It is easy to check that (42) is equivalent to

$$\sigma_k^2\cdot\sum_{k\in\mathcal{C}_j} w_{k,\sigma_{\boldsymbol{t}}}c_{k,\sigma_R,1} = \sum_{k\in\mathcal{C}_j} w_{k,\sigma_{\boldsymbol{t}}}c_{k,\sigma_R,2}. \tag{43}$$

(43) leads the following formula for updating σ_R:

$$\sigma_R \leftarrow \sqrt{\sum_{k\in\mathcal{C}_j} w_{k,\sigma_{\boldsymbol{t}}}c_{k,\sigma_R,2} / \sum_{k\in\mathcal{C}_j} w_{k,\sigma_{\boldsymbol{t}}}c_{k,\sigma_R,1}} \tag{44}$$

Similarly, when σ_R is fixed, the optimization reduces to

$$\max_{\sigma_{\boldsymbol{t}}} h_{\sigma_R}(\sigma_{\boldsymbol{t}}) \tag{45}$$

where

$$h_{\sigma_R}(\sigma_{\boldsymbol{t}}) := \sum_{k\in\mathcal{C}_j} w_{k,R}\cdot\big(\frac{2\sigma_{\boldsymbol{t}}\sigma_{k,\boldsymbol{t}}^2}{\sigma_{\boldsymbol{t}}^2+\sigma_{k,\boldsymbol{t}}^2}\big)^{\frac{3}{2}} e^{-\frac{\|\boldsymbol{t}_j^\star - \boldsymbol{t}_k\|^2}{2(\sigma_{\boldsymbol{t}}^2+\sigma_{k,\boldsymbol{t}}^2)}}$$
$$w_{k,R} := w_k\cdot\big(\frac{2\sigma_R\sigma_{k,R}^2}{\sigma_R^2+\sigma_{k,R}^2}\big)^{\frac{9}{2}} e^{-\frac{\|R_j^\star - R_k\|^2}{2(\sigma_R^2+\sigma_{k,R}^2)}}.$$

Computing the derivative of $h_{\sigma_R}(\sigma_{\boldsymbol{t}})$ with respect to $\sigma_{\boldsymbol{t}}$, we can see that $\sigma_{\boldsymbol{t}}$ is a critical point of $h_{\sigma_R}(\sigma_{\boldsymbol{t}})$ if

$$0 = \sum_{k\in\mathcal{C}_j} w_{k,\sigma_R}\cdot\big(\frac{2\sigma_{\boldsymbol{t}}\sigma_{k,\boldsymbol{t}}^2}{\sigma_{\boldsymbol{t}}^2+\sigma_{k,\boldsymbol{t}}^2}\big)^{\frac{1}{2}} e^{-\frac{\|\boldsymbol{t}_j^\star - \boldsymbol{t}_k\|^2}{2(\sigma_{\boldsymbol{t}}^2+\sigma_{k,\boldsymbol{t}}^2)}}$$
$$\Big(\frac{3\sigma_{k,\boldsymbol{t}}^2(\sigma_{k,\boldsymbol{t}}^2-\sigma_{\boldsymbol{t}}^2)}{(\sigma_{\boldsymbol{t}}^2+\sigma_{k,\boldsymbol{t}}^2)^2} + \frac{2\sigma_{\boldsymbol{t}}^2\sigma_{k,\boldsymbol{t}}^2\|\boldsymbol{t}_j^\star - \boldsymbol{t}_k\|^2}{(\sigma_{\boldsymbol{t}}^2+\sigma_{k,\boldsymbol{t}}^2)^3}\Big) \tag{46}$$

Denote

$$c_{k,\sigma_{\boldsymbol{t}},1} = \Big(\frac{2\sigma_{\boldsymbol{t}}\sigma_{k,\boldsymbol{t}}^2}{\sigma_{\boldsymbol{t}}^2+\sigma_{k,\boldsymbol{t}}^2}\Big)^{\frac{1}{2}} e^{-\frac{\|\boldsymbol{t}_j^\star-\boldsymbol{t}_k\|^2}{2(\sigma_{\boldsymbol{t}}^2+\sigma_{k,\boldsymbol{t}}^2)}} \frac{3\sigma_{k,\boldsymbol{t}}^2}{(\sigma_{\boldsymbol{t}}^2+\sigma_{k,\boldsymbol{t}}^2)^2},$$

$$c_{k,\sigma_{\boldsymbol{t}},2} = \Big(\frac{2\sigma_{\boldsymbol{t}}\sigma_{k,\boldsymbol{t}}^2}{\sigma_{\boldsymbol{t}}^2+\sigma_{k,\boldsymbol{t}}^2}\Big)^{\frac{1}{2}} e^{-\frac{\|\boldsymbol{t}_j^\star-\boldsymbol{t}_k\|^2}{2(\sigma_{\boldsymbol{t}}^2+\sigma_{k,\boldsymbol{t}}^2)}} \cdot\Big(\frac{3\sigma_{k,\boldsymbol{t}}^4}{(\sigma_{\boldsymbol{t}}^2+\sigma_{k,\boldsymbol{t}}^2)^2} + \frac{2\sigma_{k,\boldsymbol{t}}^2\sigma_{\boldsymbol{t}}^2\|\boldsymbol{t}_j^\star-\boldsymbol{t}_k\|^2}{(\sigma_{\boldsymbol{t}}^2+\sigma_{k,\boldsymbol{t}}^2)^3}\Big).$$

It is easy to check that (46) is equivalent to

$$\sigma_k^2\cdot\sum_{k\in\mathcal{C}_j} w_{k,\sigma_R} c_{k,\sigma_{\boldsymbol{t}},1} = \sum_{k\in\mathcal{C}_j} w_{k,\sigma_R} c_{k,\sigma_{\boldsymbol{t}},2}. \tag{47}$$

(47) leads the following formula for updating $\sigma_{\boldsymbol{t}}$:

$$\sigma_{\boldsymbol{t}} \leftarrow \sqrt{\sum_{k\in\mathcal{C}_j} w_{k,\sigma_R} c_{k,\sigma_{\boldsymbol{t}},2}/\sum_{k\in\mathcal{C}_j} w_{k,\sigma_R} c_{k,\sigma_{\boldsymbol{t}},1}} \tag{48}$$

2. More qualitative results over RGB-D datasets

We provide more results over the ScanNet dataset in Figure 1. We observe that Step I and Step II can provide good pose estimations in some of the cases while Step III is capable of refining results with high noise after Step II.

3. Running time comparison over RGB image datasets

We provide the running time of all the baseline methods and ours in Table 1 over the Cornell-Artsquad and Sanfrancisco dataset. SESync takes over 3 hours on Cornell-Artsquad and Shonan Rotation Averaging fails to converge over Cornell-Artsquad.

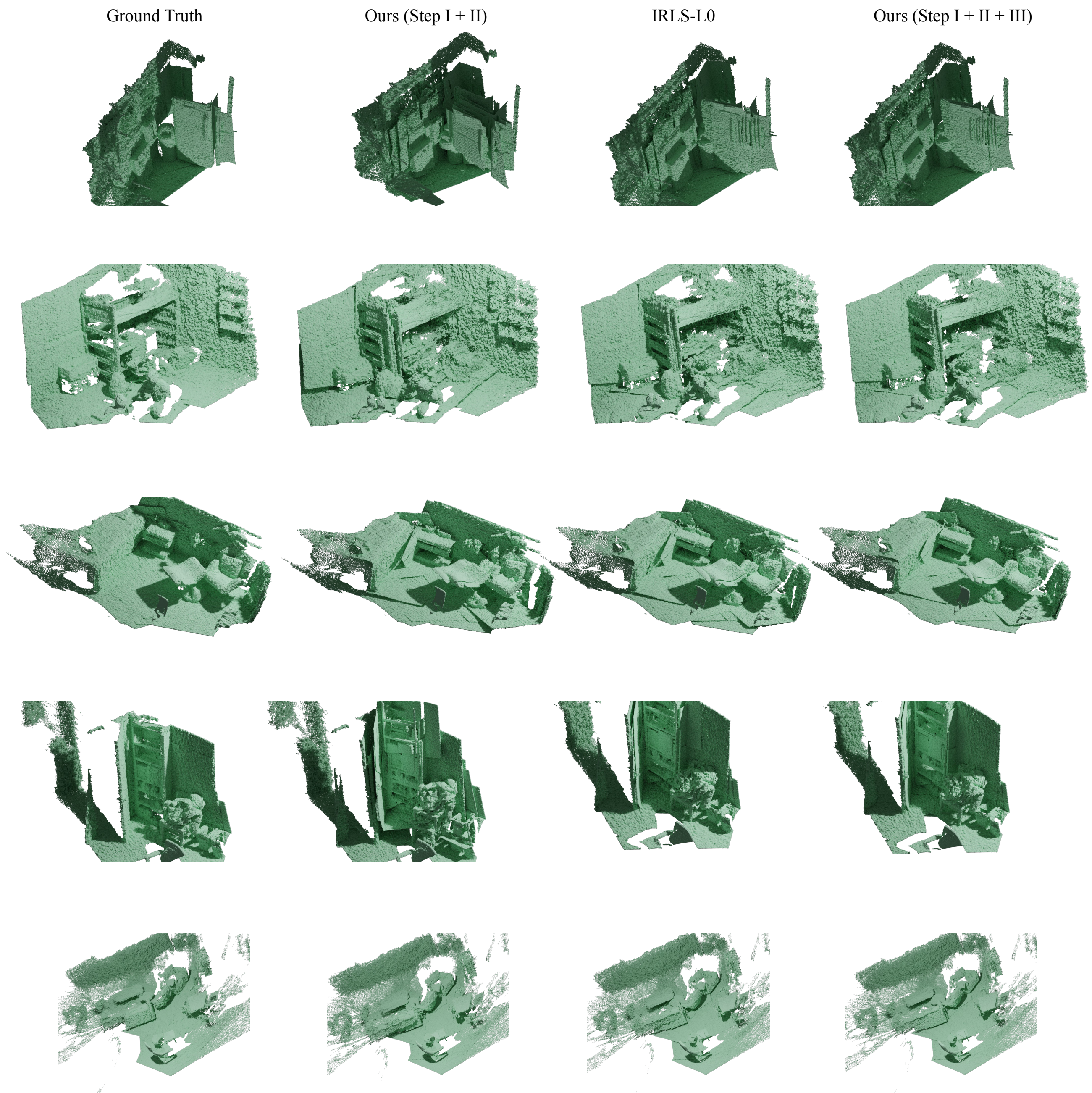

Figure 1. Comparison of qualitative results over 5 scenes from ScanNet. Step I and Step II provides accurate pose estimation in Row 2 and 5 while Step III refines the others considerably.

	IRLSL0	RobustR	SDP	SESync	SFMMRF	SHONAN	TransSync	K-Best	Ours
Cornell-Quad	93	1574	50	-	570	-	5323	302	595
San-francisco	161	763	15	335	433	1327	1819	133	243

Table 1. Running time (sec.) comparison of all baselne methods and ours over RGB image datasets.