A. Overview

We organize this supplemental material as follows. In Section B, we provide more detailed experimental results. In Section C, we describe the technical proofs for all the propositions in the main paper. In Section D, we show the scenes we used in this paper.

B. More Experimental Results

B.1. More Visual Comparison Results

Figure 6 shows more visual comparisons between our approach and baseline approaches. Again, our approach produces alignments that are close to the underlying ground-truth. The overall quality of our alignments is superior to that of the baseline approaches.

B.2. Cumulative Density Function

Figure 7 plots the cumulative density functions of errors in rotations and translations with respect to a varying threshold.

B.3. Illustration of Dataset

To understand the difficulty of the datasets used in our experiments, we pick a typical scene from each of the Redwood and ScanNet datasets and render 15 out of 30 ground truth point clouds from the same camera view point. From Figure 9 and Figure 8, we can see that ScanNet is generally harder than Redwood, as there is less information that can be extracted by looking at pairs of scans.

C. Proofs of Propositions

We organize this section as follows. In Section C.1, we provide key lemmas regarding the eigen-decomposition of a connection Laplacian, including stability of eigenvalues/eigenvectors and derivatives of eigenvectors with respect to elements of the connection Laplacian. In Section C.2, we provide key lemmas regarding the projection operator that maps the space of square matrices to the space of rotations. Section C.3 to Section C.6 describe the proofs of all the propositions stated in the main paper. Section C.7 provides an exact recovery condition of a rotation synchronization scheme via reweighted least squares. Finally, Section C.8 provides proofs for new key lemmas introduced in this section.

C.1. Eigen-Stability of Connection Laplacian

We begin with introducing the problem setting and notations in Section C.1.1. We then present the key lemmas in Section C.1.2.

C.1.1 Problem Setting and Notations

Consider a weighted graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ with n vertices, i.e., $|\mathcal{V}| = n$. We assume that $\mathcal{G}$ is connected. With $w_{ij} > 0$ we denote an edge weight associated with edge $(i,j) \in \mathcal{E}$. Let $\overline{L}$ be the weighted adjacency matrix (Note that we drop $\boldsymbol{w}$ from the expression of $\overline{L}$ to make the notations uncluttered). It is clear that the leading eigenvector of $\overline{L}$ is $\frac{1}{\sqrt{n}}\mathbf{1} \in \mathbb{R}^n$, and its corresponding eigenvalue is zero. In the following, we shall denote the eigen-decomposition of $\overline{L}$ as

$$\overline{L} = \overline{U}\overline{\Lambda}\overline{U}^T,$$

where

$$\overline{U} = (\overline{\boldsymbol{u}}_2, \cdots, \overline{\boldsymbol{u}}_n) \text{ and } \overline{\Lambda} = \mathrm{diag}(\overline{\lambda}_2, \cdots, \overline{\lambda}_n)$$

collect the remaining eigenvectors and their corresponding eigenvalues of $L(\boldsymbol{w})$, respectively. Our analysis will also use a notation that is closely related to the pseudo-inverse of $\overline{L}$:

$$\overline{L}_t^+ := \overline{U}(\overline{\Lambda} + tI_{n-1})^{-1}\overline{U}^T, \quad \forall |t| < \overline{\lambda}_2. \tag{12}$$

Our goal is to understand the behavior of the leading eigenvectors of $\overline{L} \otimes I_k + E$[2] for a symmetric perturbation matrix $E \in \mathbb{R}^{nk \times nk}$, which is a $n \times n$ block matrix whose blocks are given by

$$E_{ij} = \left\{ \begin{array}{cc} 0 & i = j \\ -w_{ij}N_{ij} & (i,j) \in \mathcal{E} \end{array} \right.$$

where N_{ij} is the perturbation imposed on R_{ij}.

We are interested in $U \in \mathbb{R}^{nk \times k}$, which collects the leading k eigenvectors of $\overline{L} \otimes I_k + E$ in its columns. With $\lambda_1 \leq \lambda_2 \cdots \leq \lambda_k$ we denote the corresponding eigenvalues. Note that due to the property of connection Laplacian, $\lambda_i \geq 0, 1 \leq i \leq k$. Our goal is to 1) bound the eigenvalues $\lambda_i, 1 \leq i \leq k$, and 2) to provide block-wise bounds between U and $\frac{1}{\sqrt{n}}\mathbf{1} \otimes Q$, for some rotation matrix $Q \in SO(k)$.

Besides the notations introduced above that are related to Laplacian matrices, we shall also use a few matrix norms. With $\|\cdot\|$ and $\|\cdot\|_{\mathcal{F}}$ we denote the spectral norm and Frobenius norm, respectively. Given a vector $\boldsymbol{v} \in \mathbb{R}^n$, we denote $\|\boldsymbol{v}\|_\infty = \max_{1 \leq i \leq n} |v_i|$ as the element-wise infinity norm. We will also introduce a norm $\|\cdot\|_{1,\infty}$ for square matrices, which is defined as

$$\|A\|_{1,\infty} = \max_{1 \leq i \leq n} \sum_{j=1}^{n} |a_{ij}|, \quad \forall A = (a_{ij})_{1 \leq i,j \leq n} \in \mathbb{R}^{n \times n}.$$

We will also use a similar norm defined for $n \times n$ block matrices $E \in \mathbb{R}^{nk \times nk}$ (i.e., each block is a $k \times k$ matrix):

$$\|E\|_{1,\infty} = \max_{1 \leq i \leq n} \sum_{j=1}^{n} \|E_{ij}\|, \quad \forall E = (E_{ij})_{1 \leq i,j \leq n} \in \mathbb{R}^{nk \times nk}.$$

[2]Note that when applying the stability results to the problem studied in this paper, we always use $k = 3$. However, when assume a general k when describing the stability results.

Figure 6: We show the results of ground truth result (column I), Rotation Averaging [12]+Translation Sync. [26] (column II), Geometric Registration [15] (column III), and Our Approach (column IV). These scenes are from Redwood Chair dataset.

C.1.2 Key Lemmas

This section presents a few key lemmas that will be used to establish main stability results regarding matrix eigenvectors and matrix eigenvalues. We begin with the classical result of the Weyl's inequality:

Lemma C.1. ***(Eigenvalue stability)*** *For $1 \leq i \leq k$, we have*

$$\lambda_i \leq \|E\|. \tag{13}$$

We proceed to describe tools for controlling the eigenvector stability. To this end, we shall rewrite U as follows:

$$U = \frac{1}{\sqrt{n}}\mathbf{1} \otimes X + Y.$$

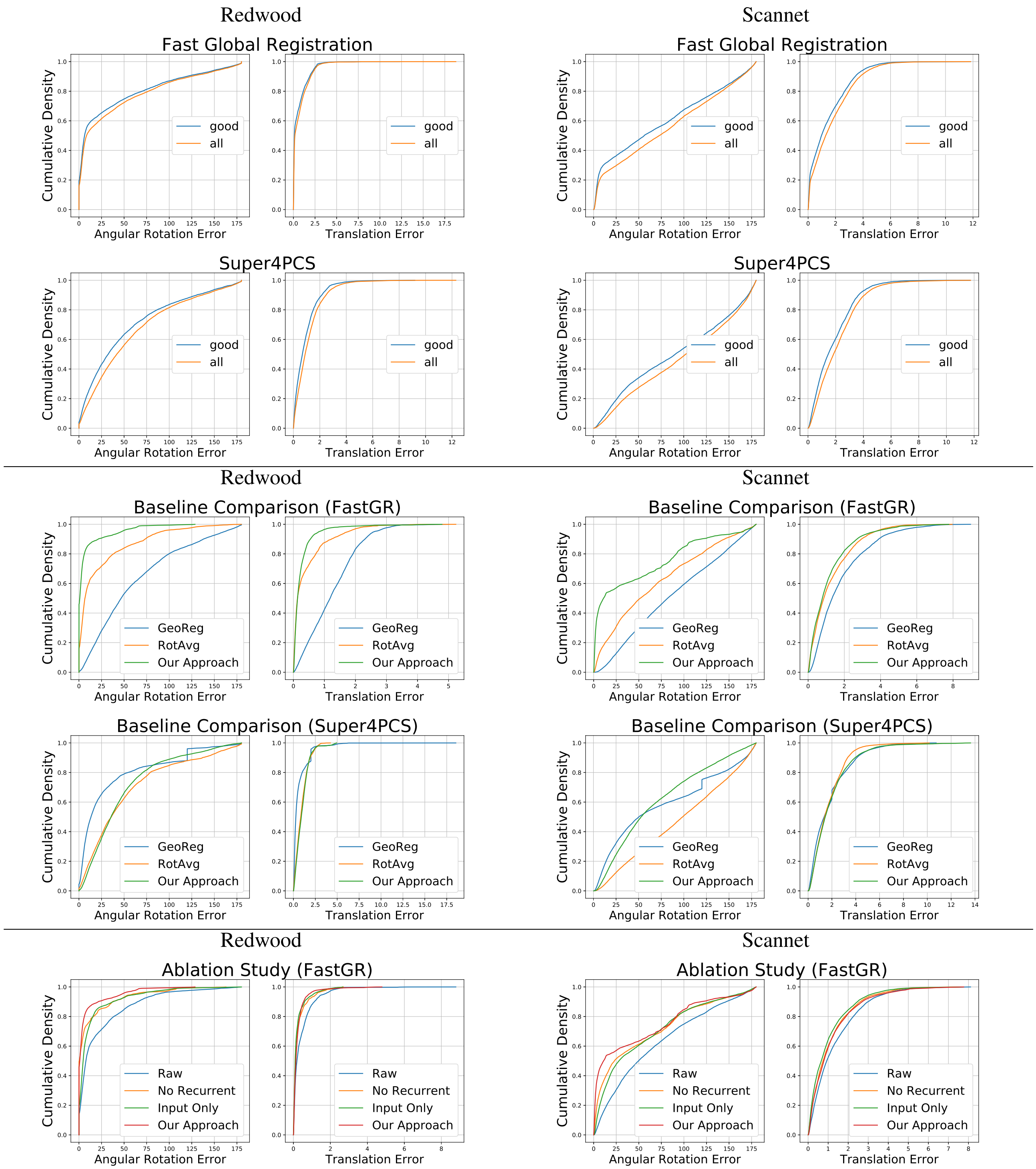


Figure 7: Corresponding cumulative density function (CDF) curves. For the top block, we plot CDF from different input sources. Here ”all” corresponds to errors between all pairs and ”good” corresponds to errors between selected pairs. The pairs were selected by 1) computing ICP refinement, 2) computing overlapping region by finding points in source point clouds that are close to target point clouds (i.e. by setting a threshold), 3) for these points, we compute their median distance to the target point clouds. For the middle block, we report the comparison of baselines and our approach. Results from different input sources are reported separately. For the bottom block, we report the comparison between variants of our approach using Fast Global Registration as the input pairwise alignments.

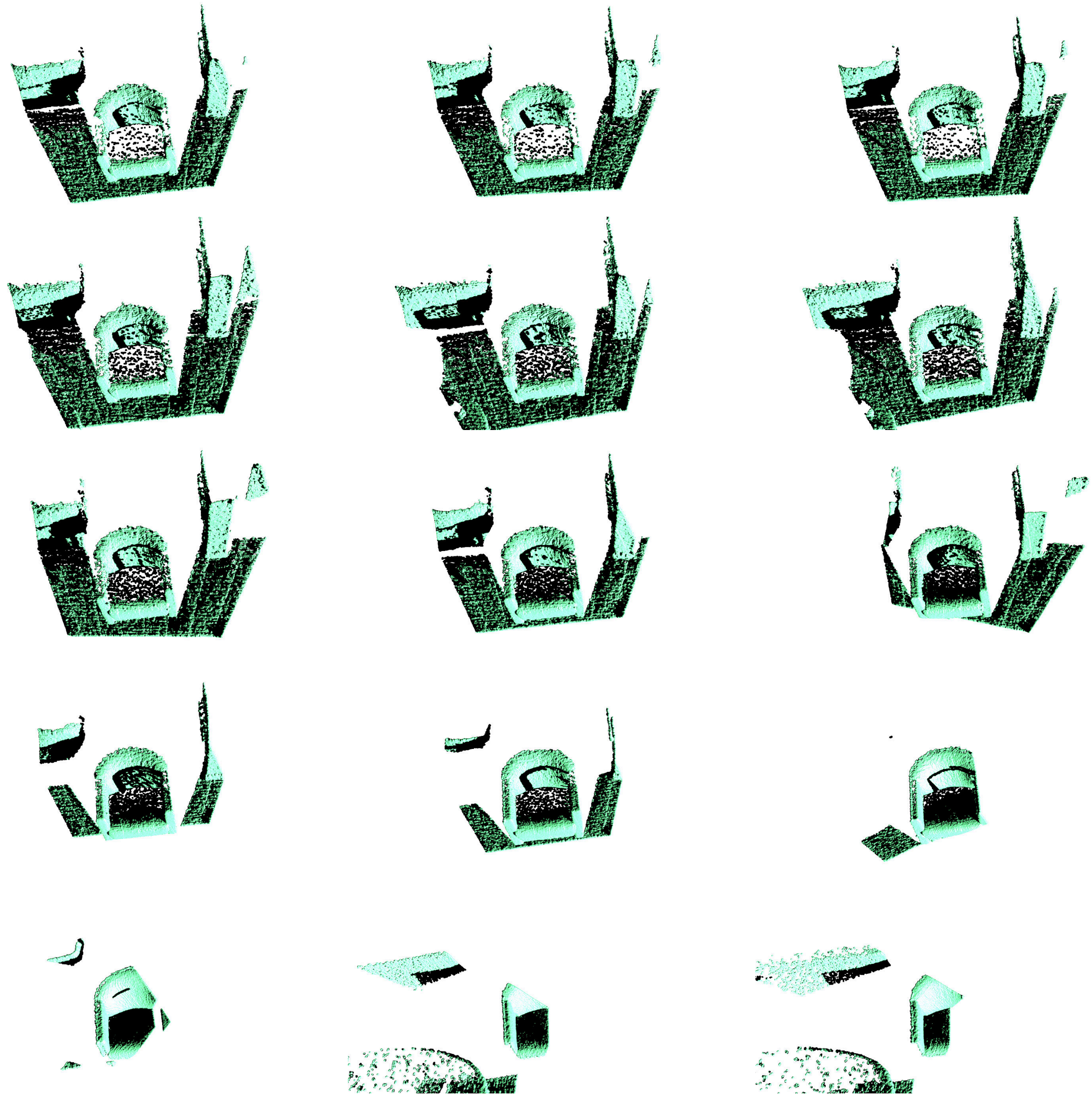


Figure 8: A typical example of the a Redwood Chair scene: the 1st, 3rd, 5th, 7th, . . ., 29th of the selected scans are rendered from the same camera view point. Each scan is about 40 frames away from the next one.

Our goal is to bound the deviation between X and a rotation matrix and blocks of Y.

We begin with controlling X, which we adopt a result described in [8]:

Lemma C.2. ***(Controlling** X **[8])** If*

$$\|E\| < \frac{\overline{\lambda}_2}{2},$$

then there exists $Q \in SO(k)$[3] such that

$$\|X - Q\| \leq 1 - \sqrt{1 - \left(\frac{\|E\|}{\overline{\lambda}_2 - \|E\|}\right)^2}.$$

In particular,

$$\|X - Q\| \leq \left(\frac{\|E\|}{\overline{\lambda}_2 - \|E\|}\right)^2$$

[3] If not, we can always negate the last column of U.

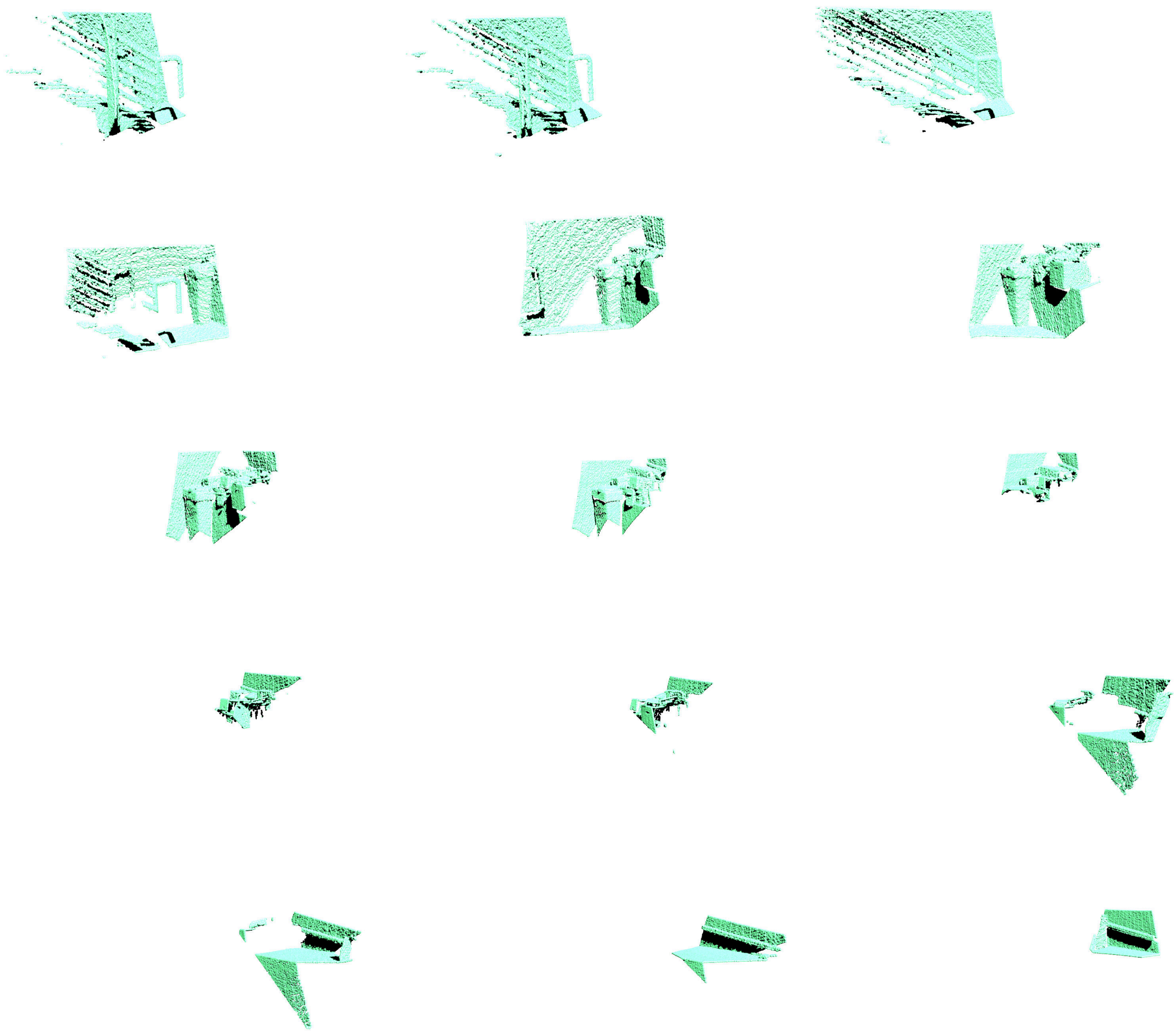


Figure 9: A typical example of the a ScanNet scene: the 1st, 3rd, 5th, 7th, . . ., 29th of the selected scans are rendered from the same camera view point. Each scan is about 40 frames away from the next one.

It remains to control the blocks of Y. We state a formulation that expresses the column of Y using a series:

Lemma C.3. *Suppose* $\|E\| < \frac{\overline{\lambda}_2}{2}$*, then* $\forall 1 \leq j \leq k$*,*

$$Y\boldsymbol{e}_j^{(k)} = -\frac{1}{\sqrt{n}}\sum_{l=1}^{\infty}\Big((\overline{L}_{-\lambda_j}^{+}\otimes I_k)E\Big)^l(\mathbf{1}\otimes X)\boldsymbol{e}_j^{(k)}. \quad (14)$$

We conclude this section by providing an explicit expression for computing the derivative of the leading eigenvectors of a connection Laplacian with its elements:

Lemma C.4. *Let* L *be an* $N\times N$ *non-negative definite matrix and its eigen-decomposition is*

$$L = \sum_{i=1}^{N}\lambda_i \boldsymbol{u}_i\boldsymbol{u}_i^T \quad (15)$$

where $0 \leq \lambda_1 \leq \lambda_2 \leq \ldots \lambda_N$*.*

Suppose $\lambda_k < \lambda_{k+1}$*. Collect the eigenvectors corresponding to the smallest* k *eigenvalues of* L *as the columns of matrix* U_k*. Namely,* $U_k = [\boldsymbol{u}_1,\ldots,\boldsymbol{u}_k]$ *where* $\sigma_1,\ldots,\sigma_k$ *are the smallest* k *eigenvelues of* L*.*

Notice that L can have different decompositions in (15) when there are repetitive eigenvalues. But in our case where $\lambda_k < \lambda_{k+1}$, we claim that $U_k U_k^T$ is unique under different possible decomposition of L so that $\mathrm{d}(U_k U_k^T)$ is well-defined and has an explicit expression:

$$\mathrm{d}(U_k U_k^T) = \sum_{i=1}^{k} \sum_{j=k+1}^{N} \frac{\boldsymbol{u}_j^T \mathrm{d}L \boldsymbol{u}_i}{\sigma_i - \sigma_j} (\boldsymbol{u}_i \boldsymbol{u}_j^T + \boldsymbol{u}_j \boldsymbol{u}_i^T) \quad (16)$$

Moreover, the differentials of eigenvalues are

$$\mathrm{d}\sigma_i = \boldsymbol{u}_i^T \mathrm{d}L \boldsymbol{u}_i. \quad (17)$$

C.2. Key Lemma Regarding the Projection Operator

This section studies the projection operator which maps the space of square matrices to the space of rotation matrices. We begin with formally defining the projection operator as follows:

Definition 1. *Suppose $\det(M) > 0$. Let $M = \sum_{i=1}^n \sigma_i \boldsymbol{u}_i \boldsymbol{v}_i^T$ be the singular value decomposition of square matrix M where $U = [\boldsymbol{u}_1, \ldots, \boldsymbol{u}_1]$ and $V = [\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n]$ are both orthogonal matrices, and all coefficients σ_i are non-negative. Then we define the rotation approximation of M as*

$$R(M) := \sum_{i=1}^{n} \boldsymbol{u}_i \boldsymbol{v}_i^T = UV^T.$$

It is clear that $R(M)$ is a rotation matrix, since 1) both U and V^T are rotations, and 2) $\det(UV^T) > 0$.

Lemma C.5. *Let $A \in \mathbb{R}^{nk \times k}$ be a block matrix of form*

$$A = \begin{bmatrix} A_1 \\ \vdots \\ A_n \end{bmatrix}$$

where $A_i \in \mathbb{R}^{k \times k}$. Use a_{ij} to denote the element on position i, j in A. Then we have

$$\sum_{i=1}^{n} \|A_i\|^2 \leq k \|A\|^2$$

We then present the following key lemma regarding the stability of the projection operator:

Lemma C.6. *Let M be a square matrix and $\epsilon = \|M - I\|$. Suppose $\epsilon < \frac{1}{3}$, then*

$$\|R(M) - I\| \leq \epsilon + \epsilon^2.$$

Lemma C.7. *Regarding $R(M)$ as a function about M, then the differential of $R(M)$ would be*

$$\mathrm{d}R(M) = \sum_{i \neq j} \frac{\boldsymbol{u}_i^T \mathrm{d}M \boldsymbol{v}_j - \boldsymbol{u}_j^T \mathrm{d}M \boldsymbol{v}_i}{\sigma_i + \sigma_j} \boldsymbol{u}_i \boldsymbol{v}_j^T$$

where all notations follow Definition (1).

C.3. Robust Recovery of Rotations

We state the following result regarding robust recovery of rotations using the connection:

Proposition 3. *Suppose the underlying rotations are given by $R_i^\star, 1 \leq i \leq n$. Modify the definition of E such that*

$$E_{ij} = \begin{cases} -w_{ij} R_j^\star (R_j^{\star T} R_{ij} R_i^\star - I_k) R_i^{\star T} & (i,j) \in \mathcal{E} \\ 0 & \text{otherwise} \end{cases}$$

Define

$$\epsilon_1 := \frac{2\|E\|_{1,\infty}}{\overline{\lambda}_2}, \quad \epsilon_2 := \|\overline{L}^+\|_{1,\infty} \|E\|_{1,\infty}. \quad (18)$$

Suppose $\epsilon_1 < 1$, $\epsilon_2 < 1$, and

$$\delta := \Big(\frac{\epsilon_1}{2 - \epsilon_1}\Big)^2 + \sqrt{k} \cdot \Big(1 + \big(\frac{\epsilon_1}{2 - \epsilon_1}\big)^2\Big) \cdot \frac{\epsilon_2 (1 + \epsilon_2)}{1 - \epsilon_2 (1 + \epsilon_2)} < \frac{1}{3}.$$

Then the optimal solution $R_i, 1 \leq i \leq n$ to the rotation synchronization step satisfies that there exists $Q \in SO(k)$,

$$\max_{1 \leq i \leq n} \|R_i - R_i^\star Q\| \leq \delta + \delta^2. \quad (19)$$

Proof of Prop. 3: Without losing generality, we assume $R_i^\star = I_k, 1 \leq i \leq k$ when proving Prop. 3. In fact, we can always apply an unitary transform to obtain $\mathrm{diag}(R_1^\star, \cdots, R_n^\star)^T L \mathrm{diag}(R_1^\star, \cdots, R_n^\star)$, which does not impact the structure of the eigen-decomposition, and which satisfies the assumption.

Before proving Prop.3, we shall utilize two Lemmas, whose proofs are deferred to Section C.8.

Lemma C.8. *Under the assumptions described above, we have*

$$\|\overline{L}_{-\lambda_j}^+\|_{1,\infty} \leq \|\overline{L}^+\|_{1,\infty} (1 + \|\overline{L}^+\|_{1,\infty} \|E\|_{1,\infty}). \quad (20)$$

Lemma C.9. *Given a $k \times k$ matrix A, we have*

$$\|A\| \leq \sqrt{k} \max_{1 \leq j \leq k} \|A \boldsymbol{e}_j^{(k)}\|. \quad (21)$$

Now we proceed to complete the proof of Prop.3. First of all, applying Lemma C.2, we obtain that there exists $Q \in SO(k)$ such that

$$\|X - Q\| \leq \big(\frac{\epsilon_1}{2-\epsilon_1}\big)^2. \tag{22}$$

Applying Lemma C.3, we have $\forall 1 \leq j \leq k$,

$$\begin{aligned}
&\sqrt{n}\|(\boldsymbol{e}_i^{(n)} \otimes I_k) Y \boldsymbol{e}_j^{(k)}\| &(23)\\
&\leq \sum_{l=1}^{\infty} \|(\overline{L}_{-\lambda_j}^{+}) E\|_{1,\infty}^{l} \|X\| \\
&\leq \sum_{l=1}^{\infty} \big(\|\overline{L}_{-\lambda_j}^{+}\|_{1,\infty}\|E\|_{1,\infty}\big)^{l}\|X\| \\
&= \frac{\|\overline{L}_{-\lambda_j}^{+}\|_{1,\infty}\|E\|_{1,\infty}}{1-\|\overline{L}_{-\lambda_j}^{+}\|_{1,\infty}\|E\|_{1,\infty}}\|X\| \\
&\overset{(20)}{\leq} \frac{\|\overline{L}^{+}\|_{1,\infty}\|E\|_{1,\infty}(1+\|\overline{L}^{+}\|_{1,\infty}\|E\|_{1,\infty})}{1-\|\overline{L}^{+}\|_{1,\infty}\|E\|_{1,\infty}(1+\|\overline{L}^{+}\|_{1,\infty}\|E\|_{1,\infty})}\cdot\|X\| \\
&\leq \frac{\epsilon_2(1+\epsilon_2)}{1-\epsilon_2(1+\epsilon_2)}\cdot\Big(1+\big(\frac{\epsilon_1}{2-\epsilon_1}\big)^2\Big). &(24)
\end{aligned}$$

We can now conclude the proof by combining (22), (24), Lemma C.9, and Lemma C.6. □

C.4. Robust Recovery of Translations

In the same spirit as the preceding section, we assume the underlying ground-truth satisfies

$$R_i^{\star} = I_k, \boldsymbol{t}_i^{\star} = \mathbf{0}, \quad 1 \leq i \leq n. \tag{25}$$

In other words, a correct measurement along edge $(i,j) \in \mathcal{E}$ should satisfy $R_{ij} = I_k, \boldsymbol{t}_{ij} = \mathbf{0}$. As we will see later, this assumption makes the error bound easier to parse. It is easy to see that the more general setting can always be converted into this simple setup through factoring out the rigid transformations among the coordinate systems associated with the input objects.

We present a formal statement of Prop. 4.2 of the main paper as follows:

Proposition 4. *Consider the assumption of (25). Define*

$$\epsilon_3 := \max_{1\leq i\leq n} \sum_{j\in\mathcal{N}(i)} w_{ij}\|t_{ij}\| \tag{26}$$

Intuitively, ϵ_3 measures the cumulative transformation error associated with each object. Under the same assumption as Prop. 3 for the connection Laplacian $L = \overline{L}\otimes I_k + E$, we can bound the error of the translation synchronization step as

$$\max_{1\leq i\leq n}\|\boldsymbol{t}_i\| \leq \frac{\|\overline{L}^{+}\|_{1,\infty}\epsilon_3}{1-4\epsilon_2}. \tag{27}$$

Proof of Lemma 4: First of all, note that $(\mathbf{1}\otimes I_k)^T\boldsymbol{b} = 0$. Thus we can factor out the component in E that corresponds the the subspace spanned by $\mathbf{1}\otimes I_k$. Specifically, define

$$E' = \big((I - \frac{1}{n}\mathbf{1}\mathbf{1}^T)\otimes I_k\big)E\big((I-\frac{1}{n}\mathbf{1}\mathbf{1}^T)\otimes I_k\big).$$

It is easy to check that

$$\|E'\|_{1,\infty} \leq 4\|E\|_{1,\infty}.$$

Moreover,

$$\boldsymbol{t} = (\overline{L}\otimes I_k + E')^{+}\boldsymbol{b}.$$

This means

$$\max_{1\leq i\leq n}\|\boldsymbol{t}_i\| \leq \|(\overline{L}\otimes I_k + E')^{+}\|_{1,\infty}\epsilon_3.$$

Note that

$$(\overline{L}\otimes I_k + E')^{+} = \sum_{l=0}^{\infty}\big((\overline{L}^{+}\otimes I_k)E'\big)^{l}(\overline{L}^{+}\otimes I_k).$$

It follows that

$$\begin{aligned}
&\|(\overline{L}\otimes I_k + E')^{+}\|_{1,\infty} \\
&\leq \sum_{l=0}^{\infty}\|\big((\overline{L}^{+}\otimes I_k)E'\big)\|_{1,\infty}^{l}\|(\overline{L}^{+}\otimes I_k)\|_{1,\infty} \\
&= \sum_{l=0}^{\infty}(\|\overline{L}^{+}\|_{1,\infty}\|E'\|_{1,\infty})^{l}\|\overline{L}^{+}\|_{1,\infty} \\
&\leq \frac{\|\overline{L}^{+}\|_{1,\infty}}{1-4\|\overline{L}^{+}\|_{1,\infty}\|E\|_{1,\infty}}.
\end{aligned}$$

□

C.5. Proof of Proposition 1 in the Main Paper

Applying Lemma C.7, we have

$$\mathrm{d}R_i = \sum_{1\leq s,t\leq k}\frac{\boldsymbol{v}_s^{(i)T}\mathrm{d}U_i\boldsymbol{w}_t^{(i)} - \boldsymbol{v}_t^{(i)T}\mathrm{d}U_i\boldsymbol{w}_s^{(i)}}{\sigma_s^{(i)}+\sigma_t^{(i)}}\boldsymbol{v}_s^{(i)}\boldsymbol{w}_t^{(i)T}. \tag{28}$$

We further divide the computation of $\mathrm{d}U_i$ into two parts.

Consider the j-th column of $\mathrm{d}U_i$:

$$\begin{aligned}\mathrm{d}U_i\boldsymbol{e}_j^{(k)} &= ({\boldsymbol{e}_i^{(n)}}^T \otimes I_k)\mathrm{d}U\boldsymbol{e}_j^{(k)} \\ &= ({\boldsymbol{e}_i^{(n)}}^T \otimes I_k)\mathrm{d}\boldsymbol{u}_j \\ &= ({\boldsymbol{e}_i^{(n)}}^T \otimes I_k)\sum_{l\neq j}\boldsymbol{u}_l\frac{\boldsymbol{u}_j^T\mathrm{d}L\boldsymbol{u}_l}{\lambda_j-\lambda_l} \\ &= ({\boldsymbol{e}_i^{(n)}}^T \otimes I_k)\Big(\sum_{\substack{l=1\\ l\neq j}}^{k}\frac{\boldsymbol{u}_j^T\mathrm{d}L\boldsymbol{u}_l}{\lambda_j-\lambda_l}\boldsymbol{u}_l+ \\ &\quad \sum_{l=k+1}^{kn}\frac{\boldsymbol{u}_j^T\mathrm{d}L\boldsymbol{u}_l}{\lambda_j-\lambda_l}\boldsymbol{u}_l\Big) \\ &= \mathrm{d}U_i^{(inner)}\boldsymbol{e}_j^{(k)} + \mathrm{d}U_i^{(outer)}\boldsymbol{e}_j^{(k)} \qquad (29)\end{aligned}$$

where

$$\mathrm{d}U_i^{(inner)}\boldsymbol{e}_j^{(k)} = ({\boldsymbol{e}_i^{(n)}}^T \otimes I_k)\sum_{\substack{l=1\\ l\neq j}}^{k}\frac{\boldsymbol{u}_l\boldsymbol{u}_l^T}{\lambda_j-\lambda_l}\mathrm{d}L\boldsymbol{u}_j \qquad (30)$$

$$= \sum_{\substack{l=1\\ l\neq j}}^{k}(U_i\boldsymbol{e}_l^{(k)})\frac{\boldsymbol{u}_l\boldsymbol{u}_l^T}{\lambda_j-\lambda_l}\mathrm{d}L\boldsymbol{u}_j \qquad (31)$$

$$\mathrm{d}U_i^{(outer)}\boldsymbol{e}_j^{(k)} = ({\boldsymbol{e}_i^{(n)}}^T \otimes I_k)\sum_{l=k+1}^{kn}\frac{\boldsymbol{u}_l\boldsymbol{u}_l^T}{\lambda_j-\lambda_l}\mathrm{d}L\boldsymbol{u}_j \qquad (32)$$

In (30), we used the fact that $({\boldsymbol{e}_i^{(n)}}^T \otimes I_k)\boldsymbol{u}_l$ is just $U_i\boldsymbol{e}_l^{(k)}$ by definition of U_i.

Since $\mathrm{d}R_i$ is linear with respect to $\mathrm{d}U_i$, we can divide $\mathrm{d}R_i$ similarly:

$$\mathrm{d}R_i = \mathrm{d}R_i^{(inner)} + \mathrm{d}R_i^{(outer)}$$

$$\mathrm{d}R_i^{(inner)} := \sum_{1\leq s,t\leq k}\frac{{\boldsymbol{v}_s^{(i)}}^T\mathrm{d}U_i^{(inner)}\boldsymbol{w}_t^{(i)} - {\boldsymbol{v}_t^{(i)}}^T\mathrm{d}U_i^{(inner)}\boldsymbol{w}_s^{(i)}}{\sigma_s^{(i)}+\sigma_t^{(i)}}\boldsymbol{v}_s^{(i)}{\boldsymbol{w}_t^{(i)}}^T \qquad (33)$$

$$\mathrm{d}R_i^{(outer)} := \sum_{1\leq s,t\leq k}\frac{{\boldsymbol{v}_s^{(i)}}^T\mathrm{d}U_i^{(outer)}\boldsymbol{w}_t^{(i)} - {\boldsymbol{v}_t^{(i)}}^T\mathrm{d}U_i^{(outer)}\boldsymbol{w}_s^{(i)}}{\sigma_s^{(i)}+\sigma_t^{(i)}}\boldsymbol{v}_s^{(i)}{\boldsymbol{w}_t^{(i)}}^T \qquad (34)$$

Then the derivative we would like to compute can be written as

$$\begin{aligned}\mathrm{d}(R_iR_j^T) &= \mathrm{d}R_iR_j^T + R_i\mathrm{d}R_j{}^T \\ &= \mathrm{d}R_i^{(inner)}R_j^T + R_i\mathrm{d}{R_j^{(inner)}}^T \\ &\quad + \mathrm{d}R_i^{(outer)}R_j^T + R_i\mathrm{d}{R_j^{(outer)}}^T. \qquad (35)\end{aligned}$$

From (34) and (32) it can be easily checked that the formula in Proposition 1 of the main aper that we want to prove is just (35) except the extra terms $\mathrm{d}R_i^{(inner)}R_j^T + R_i\mathrm{d}{R_j^{(inner)}}^T$. Hence in the remaining proof it suffices to show that

$$\mathrm{d}R_i^{(inner)}R_j^T + R_i\mathrm{d}{R_j^{(inner)}}^T = 0.$$

To this end, we define a k-by-k auxiliary matrix C as

$$C_{lj} = \frac{\boldsymbol{u}_j^T\mathrm{d}L\boldsymbol{u}_l}{\lambda_j-\lambda_l}$$

for all $l \neq j$ and $C_{jj} = 0$. Since L is symmetric, C would be skew-symmetric that means $C + C^T = 0$. Fist of all, notice that

$$\mathrm{d}U_i^{(inner)} = \sum_{\substack{l=1\\ l\neq k}}^{k}(U_i\boldsymbol{e}_l^{(k)})\frac{\boldsymbol{u}_j^T\mathrm{d}L\boldsymbol{u}_l}{\lambda_j-\lambda_l} = U_iC.$$

Also it is clear that

$${\boldsymbol{v}_s^{(i)}}^TU_i = \sigma_s\boldsymbol{w}_s^T$$

by using simple properties of SVD. It follows that

$$\mathrm{d}R_i^{(inner)} = \sum_{1\leq s,t\leq k}\frac{\sigma_s{\boldsymbol{w}_s^{(i)}}^TC\boldsymbol{w}_t^{(i)} - \sigma_t{\boldsymbol{w}_t^{(i)}}^TC\boldsymbol{w}_s^{(i)}}{\sigma_s^{(i)}+\sigma_t^{(i)}}\boldsymbol{v}_s^{(i)}{\boldsymbol{w}_t^{(i)}}^T \qquad (36)$$

$$\begin{aligned}&= \sum_{1\leq s,t\leq k}{\boldsymbol{w}_s^{(i)}}^TC\boldsymbol{w}_t^{(i)}\boldsymbol{v}_s^{(i)}{\boldsymbol{w}_t^{(i)}}^T \qquad (37)\\ &= \sum_{1\leq s,t\leq k}\boldsymbol{v}_s^{(i)}{\boldsymbol{w}_s^{(i)}}^TC\boldsymbol{w}_t^{(i)}{\boldsymbol{w}_t^{(i)}}^T \qquad (38)\\ &= V^{(i)}{W^{(i)}}^TC \qquad (39)\\ &= R_iC. \qquad (40)\end{aligned}$$

In the derivations above, we used the fact that C is a skew-symmetric matrix for deriving the first equality (36). In addition, we used the fact that ${\boldsymbol{v}_s^{(i)}}^TC\boldsymbol{v}_t^{(i)}$ is a scalar for deriving the second equality (37). When deriving (38), we used the expansion of $U^{(i)}{V^{(i)}}^T$ and the fact that $\{\boldsymbol{v}_1^{(i)}, \ldots, \boldsymbol{v}_k^{(i)}\}$ form an orthonormal basis:

$$U^{(i)}{V^{(i)}}^T = \sum_{s=1}^{k}\boldsymbol{u}_s^{(i)}{\boldsymbol{v}_s^{(i)}}^T,\ I = \sum_{t=1}^{k}\boldsymbol{v}_s^{(i)}{\boldsymbol{v}_s^{(i)}}^T.$$

(39) uses the definition of R_i. Finally, plugging (40) into (35) gives

$$\mathrm{d}R_i^{(inner)}R_j^T + R_i\mathrm{d}{R_j^{(inner)}}^T = R_iCR_j^T + R_iC^TR_j^T = 0,$$

which completes our proof. □

C.6. Proof of Proposition 2 in the Main Paper

The proof is straightforward, since

$$0 = d(L \cdot L^{-1}) = dL \cdot L^{-1} + L \cdot d(L^{-1}),$$

meaning

$$dL = -L^{-1} dL L^{-1}.$$

In the degenerate case, we replace L^{-1} by L^{+}. This is proper since the only null space of L is $\mathbf{1} \otimes I_k$, which does not affect the solution $\boldsymbol{t}$. □

C.7. Exact Recovery Condition of Rotation Synchronization

Similar to [26], we can derive a truncated rotation synchronization scheme (the generalization to transformation synchronization is straight-forward). Specifically, consider an observation graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$. Let $\mathcal{E}_{bad} \subset \mathcal{E}$ be the edge set associated with incorrect rotation measurements. Starting from $\mathcal{G}$, at each iteration, we use the solution $R_i^{(k)}, 1 \leq i \leq n$ at the kth iteration to prune input rotations whenever $\|R_{ij} - R_j^{(k)} R_i^{(k)T}$
$2\gamma^k$, where $\gamma < 1$ is a constant. Using Prop. 3, we can easily derive the following exact recovery condition:

Proposition 5. *The truncated rotation synchronization scheme recovers the underlying ground-truth if*

$$\|L_{\mathcal{G}}^{+}\|_{1,\infty} d_{\max}(\mathcal{E}_{bad}) \leq \frac{1}{16}, \quad \gamma > 0.95, \tag{41}$$

where $L_{\mathcal{G}}$ is the graph Laplacian of $\mathcal{G}$, and $d_{\max}(\mathcal{E}_{bad})$ is the maximum number of bad edges per vertex. Note that the constants in (41) are not optimized.

Proof: Denote $\epsilon_4 := \|L_{\mathcal{G}}^{+}\|_{1,\infty} d_{\max}(\mathcal{E}_{bad})$. Consider an arbitrary set $\mathcal{E}_r \subseteq \mathcal{E}_{bad}$. Introduce the graph that collects the corresponding remaining observations $\mathcal{G}_{cur} = (\mathcal{V}, E_{cur})$, where $E_{cur} = \mathcal{E} \setminus (\mathcal{E}_{bad} \setminus \mathcal{E}_r)$. Suppose we apply rotation synchronization step to $\mathcal{G}_{cur}$ and the associated observations, it is easy to show that (c.f.[26])

$$\epsilon_2 \leq 2\frac{\epsilon_4}{1-\epsilon_4} \cdot \max_{(i,j)\in\mathcal{E}_{cur}} \|N_{ij}\|, \quad \epsilon_1 \leq 2\epsilon_2.$$

Using Prop. 3 and after simple calculations, we can derive that the truncated scheme described above will never remove good measurements, which end the proof. □

Remark 1. This exact recovery condition suggests that if we simply let the weighting function to be small when the residual is big, then if the ratio of the incorrect measurements is small. It is guaranteed to remove all the incorrect measurement. Yet to maximize the effectiveness of the weighting scheme, it is suggested to learn the optimal weighting scheme from data. The approach presented in the main paper is one attempt in this direction.

C.8. Proofs of Key Lemmas

C.8.1 Proof of Lemma C.3

We first introduce the following notations, which essentially express E in the coordinate system spanned by $\frac{1}{\sqrt{n}}(\mathbf{1} \otimes I_k)$ and $\overline{U} \otimes I_k$:

$$\begin{aligned}
E_{11} &:= \frac{1}{n}(\mathbf{1} \otimes I_k)^T E (\mathbf{1} \otimes I_k), \\
E_{12} &:= \frac{1}{\sqrt{n}}(\mathbf{1} \otimes I_k)^T E (\overline{U} \otimes I_k), \\
E_{21} &:= \frac{1}{\sqrt{n}}(\overline{U} \otimes I_k)^T E (\mathbf{1} \otimes I_k), \\
E_{22} &:= (\overline{U} \otimes I_k)^T E (\overline{U} \otimes I_k).
\end{aligned}$$

Let $Y := (U \otimes I_k)\overline{Y}$. Substituting $U = \frac{1}{\sqrt{n}}(\mathbf{1} \otimes I_k) + (\overline{U} \otimes I_k)\overline{Y}$ into

$$(\overline{L} \otimes I_k + E)U = U\Lambda,$$

we obtain

$$\begin{aligned}
&(\overline{L} \otimes I_k + E)(\frac{1}{\sqrt{n}}(\mathbf{1} \otimes I_k) + (\overline{U} \otimes I_k)\overline{Y}) \\
=&(\frac{1}{\sqrt{n}}(\mathbf{1} \otimes I_k) + (\overline{U} \otimes I_k)\overline{Y})\Lambda.
\end{aligned}$$

Multiply both sides by $(U \otimes I_k)^T$, it follows that

$$E_{21}X + (\overline{\Lambda} \otimes I_k)\overline{Y} + E_{22}\overline{Y} = \overline{Y}\Lambda.$$

Since $\|E\| < \frac{\overline{\lambda}_2}{2}$, we have

$$\begin{aligned}
Y\boldsymbol{e}_j^{(k)} &:= (\overline{U} \otimes I_k)\overline{Y}\boldsymbol{e}_j^{(k)} \\
&= -(\overline{U} \otimes I_k)\big((\overline{\Lambda} - \lambda_j) \otimes I_k - E_{22}\big)^{-1} E_{21} X \boldsymbol{e}_j^{(k)} \\
&= -\sum_{l=0}^{\infty} (\overline{U} \otimes I_k)\big(((\overline{\Lambda} - \lambda_j)^{-1} \otimes I_k)E_{22}\big)^l \\
&\quad \cdot \big((\overline{\Lambda} - \lambda_j)^{-1} \otimes I_k\big) E_{21} X \boldsymbol{e}_j^{(k)} \\
&= -\sum_{l=1}^{\infty} \big((\overline{U}(\overline{\Lambda} - \lambda_j)^{-1}\overline{U}^T) \otimes I_k E\big)^l \frac{1}{\sqrt{n}} \mathbf{1} \otimes X \boldsymbol{e}_j^{(k)} \\
&= -\frac{1}{\sqrt{n}} \sum_{l=1}^{\infty} \big(L_{-\lambda_j}^{+} E\big)^l (\mathbf{1} \otimes X \boldsymbol{e}_j^{(k)}).
\end{aligned}$$

□

C.8.2 Proof of Lemma C.4

Let

$$L = \sum_{i=1}^{N} \sigma_i \boldsymbol{u}_i \boldsymbol{u}_i^T = \sum_{i=1}^{N} \sigma_i \boldsymbol{u}_i' \boldsymbol{u}_i'^T$$

be two different decompositions of L. It can be written in matrix form

$$L = U\Lambda U^T = U'\Lambda U'^T$$

where $U = [\boldsymbol{u}_1, \dots, \boldsymbol{u}_N]$, $\Lambda = \mathrm{diag}(\sigma_1, \dots, \sigma_N)$, $U' = [\boldsymbol{u}'_1, \dots, \boldsymbol{u}_N]$. Then we have

$$(U'^T U)\Lambda = \Lambda(U'^T U)$$

Let $A = U'^T U$ and the element of position (i, j) on A be a_{ij}, then we have

$$a_{ij}\sigma_j = \sigma_i a_{ij},$$

which means $a_{ij} = 0$ for all $\sigma_i \neq \sigma_j$.

Since we have assumed $\sigma_1 \leq \dots \leq \sigma_k < \sigma_{k+1} \leq \dots \leq \sigma_N$, the matrix A would have form

$$\begin{bmatrix} A_{k,k} & O_{k,N-k} \\ O_{N-k,k} & A_{N-k,N-k} \end{bmatrix}.$$

But we have known that A is an orthogonal matrix, thus $A_{k,k}$ is also an orthogonal matrix. In this way $A_{k,k} = U_k'^T U_k$ can be rewritten as

$$U_k = U'_k A_{k,k}$$

and furthermore we have

$$U_k U_k^T = U'_k(A_{k,k}A_{k,k}^T)U_k'^T = U'_k U_k'^T.$$

Since eigen-decomposition is a special case of SVD when dealing with symmetric matrix, (53) gives

$$\begin{aligned} \mathrm{d}\boldsymbol{u}_i &= \sum_{j\neq i} \frac{\sigma_i \boldsymbol{u}_j^T \mathrm{d}L\boldsymbol{u}_i + \sigma_j \boldsymbol{u}_i^T \mathrm{d}L\boldsymbol{u}_j}{\sigma_i^2 - \sigma_j^2}\boldsymbol{u}_j \\ &= \sum_{j\neq i} \frac{\boldsymbol{u}_i^T \mathrm{d}M\boldsymbol{u}_j}{\sigma_i - \sigma_j}\boldsymbol{u}_j \end{aligned}$$

in which we used the fact that $\mathrm{d}L$ is also symmetric in the last step.

Finally the differential of $U_k U_k^T$ can be written as

$$\begin{aligned} &\mathrm{d}(U_k U_k^T) \\ =&\mathrm{d}\left(\sum_{i=1}^k \boldsymbol{u}_i\boldsymbol{u}_i^T\right) \\ =&\sum_{i=1}^k (\mathrm{d}\boldsymbol{u}_i\boldsymbol{u}_i^T + \boldsymbol{u}_i\mathrm{d}\boldsymbol{u}_i^T) \\ =&\sum_{i=1}^k \sum_{\substack{j=1 \\ j\neq i}}^N \frac{\boldsymbol{u}_i^T\mathrm{d}M\boldsymbol{u}_j}{\sigma_i - \sigma_j}(\boldsymbol{u}_j\boldsymbol{u}_i^T + \boldsymbol{u}_i\boldsymbol{u}_j^T) \\ =&\sum_{i=1}^k \sum_{j=k+1}^N \frac{\boldsymbol{u}_i^T\mathrm{d}M\boldsymbol{u}_j}{\sigma_i - \sigma_j}(\boldsymbol{u}_j\boldsymbol{u}_i^T + \boldsymbol{u}_i\boldsymbol{u}_j^T) + \\ &\sum_{i=1}^k \sum_{j=i+1}^k \left(\frac{\boldsymbol{u}_i^T\mathrm{d}M\boldsymbol{u}_j}{\sigma_i - \sigma_j} + \frac{\boldsymbol{u}_i^T\mathrm{d}M\boldsymbol{u}_j}{\sigma_j - \sigma_i}\right)(\boldsymbol{u}_j\boldsymbol{u}_i^T + \boldsymbol{u}_i\boldsymbol{u}_j^T) \\ =&\sum_{i=1}^k \sum_{j=k+1}^N \frac{\boldsymbol{u}_i^T\mathrm{d}M\boldsymbol{u}_j}{\sigma_i - \sigma_j}(\boldsymbol{u}_j\boldsymbol{u}_i^T + \boldsymbol{u}_i\boldsymbol{u}_j^T) \end{aligned}$$

As for formula (17), taking differential of equation $L\boldsymbol{u}_i = \sigma_i\boldsymbol{u}_i$, we obtain

$$\mathrm{d}L\boldsymbol{u}_i + L\mathrm{d}\boldsymbol{u}_i = \mathrm{d}\sigma_i\boldsymbol{u}_i + \sigma_i\mathrm{d}\boldsymbol{u}_i$$

Let us multiply both sides by $\boldsymbol{u}_i^T$ and notice that $\boldsymbol{u}_i^T\boldsymbol{u}_i = 1$, $L\boldsymbol{u}_i = \sigma_i\boldsymbol{u}_i$, and $\boldsymbol{u}_i^T\mathrm{d}\boldsymbol{u}_i = 0$, we conclude that the equation above can be simplified to

$$\mathrm{d}\sigma_i = \boldsymbol{u}_i^T\mathrm{d}L\boldsymbol{u}_i.$$

□

C.8.3 Proof of Lemma C.5

It is well-known that $\|X\| \leq \|X\|_F$ for any matrix X where $\|\cdot\|_F$ represents the Frobenius norm. Thus

$$\begin{aligned} k\|A\|^2 = \sum_{j=1}^k \|A\|^2 &\geq \sum_{j=1}^k \sum_{i=1}^{kn} a_{ij}^2 \\ &= \sum_{i=1}^{kn}\sum_{j=1}^k a_{ij}^2 = \sum_{i=1}^n \|A_i\|_F^2 \\ &\geq \sum_{i=1}^n \|A_i\| \end{aligned}$$

completes our proof. □

C.8.4 Proof of Lemma C.6

Suppose $M = \sum_{i=1}^n \sigma_i\boldsymbol{u}_i\boldsymbol{v}_i^T$ is the SVD decomposition of M. By definition of $R(\cdot)$,

$$R(M) = \sum_{i=1}^n \boldsymbol{u}_i\boldsymbol{v}_i^T.$$

First we have a simple lower bound on ϵ:

$$\epsilon = \|M - I\| \geq |(M - I)\boldsymbol{v_i}| = |\sigma_i - 1|. \tag{42}$$

It is enough to show that for any unit vector $\boldsymbol{p} \in \mathbb{R}^n$ we have

$$\|R(M)\boldsymbol{p} - \boldsymbol{p}\| \leq (1 + \epsilon)\|M\boldsymbol{p} - \boldsymbol{p}\|. \tag{43}$$

In fact, if (43) is true, then

$$\begin{aligned}
\|R(M) - I\| &= \max_{|\boldsymbol{p}|=1} \|R(M)\boldsymbol{p} - \boldsymbol{p}\| \\
&\leq \max_{|\boldsymbol{p}=1}(1 + \epsilon)\|M\boldsymbol{p} - \boldsymbol{p}\| \\
&\leq (1 + \epsilon)\epsilon \quad \text{(by definition of } \epsilon).
\end{aligned}$$

By noting $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n\}$ are a set of basis on $\mathbb{R}^n$, we can decompose $R(M)$ and M into

$$\begin{aligned}
R(M)\boldsymbol{p} - \boldsymbol{p} &= \sum_{i=1}^{n} \boldsymbol{u}_i \boldsymbol{v}_i^T \boldsymbol{p} - \sum_{i=1}^{n} \boldsymbol{v}_i(\boldsymbol{v}_i^T \boldsymbol{p}) \\
&= \sum_{i=1}^{n} (\boldsymbol{u}_i - \boldsymbol{v}_i)(\boldsymbol{v}_i^T \boldsymbol{p}) \\
M\boldsymbol{p} - \boldsymbol{p} &= \sum_{i=1}^{n} \sigma_i \boldsymbol{u}_i \boldsymbol{v}_i \boldsymbol{p} - \sum_{i=1}^{n} \boldsymbol{v}_i(\boldsymbol{v}_i^T \boldsymbol{p}) \\
&= \sum_{i=1}^{n} (\sigma_i \boldsymbol{u}_i - \boldsymbol{v}_i)(\boldsymbol{v}_i^T \boldsymbol{p})
\end{aligned}$$

To prove (43), it suffices to show that

$$|\boldsymbol{u}_i - \boldsymbol{v}_i| \leq (1 + \epsilon)|\sigma_i \boldsymbol{u}_i - \boldsymbol{v}_i|.$$

Let $\delta = \boldsymbol{u}_i^T \boldsymbol{v}_i$. The case that $\boldsymbol{u}_i = \boldsymbol{v}_i$ is trivial. Also, if $\sigma_i = 0$, then $\epsilon \geq 1$ and the resulting inequality

$$|\boldsymbol{u}_i - \boldsymbol{v}_i| \leq 2|\boldsymbol{v}_i|$$

is trivial. Thus we can always assume $|\sigma \boldsymbol{u}_i - \boldsymbol{v}_i| \neq 0$ and $\sigma_i > 0$. Then by the laws of cosines we have

$$\begin{aligned}
\frac{|\boldsymbol{u}_i - \boldsymbol{v}_i|}{|\sigma_i \boldsymbol{u}_i - \boldsymbol{v}_i|} &= \sqrt{\frac{2 - 2\delta}{1 + \sigma_i^2 - 2\sigma_i \delta}} \\
&= \sqrt{\sigma_i^{-1} + \frac{2 - (\sigma_i^{-1} + \sigma_i)}{1 + \sigma_i^2 - 2\sigma_i \delta}}
\end{aligned} \tag{44}$$

In (44) it is clear that $2 - (\sigma_i^{-1} + \sigma_i) \leq 0$ and $1 + \sigma_i^2 \geq 2\sigma_i \delta$. Hence by monotonicity (44) reaches its maximum when $\delta = -1$ and then

$$\frac{|\boldsymbol{u}_i - \boldsymbol{v}_i|}{|\sigma_i \boldsymbol{u}_i - \boldsymbol{v}_i|} \leq \frac{2}{1 + \sigma_i} = 1 + \frac{1 - \sigma_i}{1 + \sigma_i} \leq 1 + |\sigma_i - 1| \leq 1 + \epsilon$$

C.8.5 Proof of Lemma C.7

For the sake of brevity we simply write R instead of $R(M)$ in the following proof. It is easy to see that

$$M\boldsymbol{v}_i = \sigma_i \boldsymbol{u}_i \quad \boldsymbol{u}_i^T M = \sigma_i \boldsymbol{v}_i^T$$

for $i = 1, \ldots, n$. Taking the differential on both sides we obtain

$$\mathrm{d}M\boldsymbol{v}_i + M\mathrm{d}\boldsymbol{v}_i = \mathrm{d}\sigma_i \boldsymbol{u}_i + \sigma_i \mathrm{d}\boldsymbol{u}_i \tag{45}$$

$$\mathrm{d}\boldsymbol{u}_i^T M + \boldsymbol{u}_i^T \mathrm{d}M = \mathrm{d}\sigma_i \boldsymbol{v}_i^T + \sigma_i \mathrm{d}\boldsymbol{v}_i^T \tag{46}$$

Left multiplying both sides of (45) by $\boldsymbol{u}_j$ with $j \neq i$ and observing that $\boldsymbol{u}_j^T \boldsymbol{u}_i = 0$, we obtain

$$\boldsymbol{u}_j^T \mathrm{d}M\boldsymbol{v}_i + \boldsymbol{u}_j^T M\mathrm{d}\boldsymbol{v}_i = \sigma_i \boldsymbol{u}_j^T \mathrm{d}\boldsymbol{u}_i \tag{47}$$

Similarly right multiplying both sides of (46) by $\boldsymbol{v}_j^T$ with $j \neq i$ gives

$$\mathrm{d}\boldsymbol{u}_i^T M\boldsymbol{v}_j + \boldsymbol{u}_i^T \mathrm{d}M\boldsymbol{v}_j = \sigma_i \mathrm{d}\boldsymbol{v}_i^T \boldsymbol{v}_j \tag{48}$$

Since $\boldsymbol{u}_j^T M = \sigma_j v_j^T,\ Mv_j = \sigma_j \boldsymbol{u}_j^T$, we have

$$\boldsymbol{u}_j^T \mathrm{d}M\boldsymbol{v}_i + \sigma_j \boldsymbol{v}_j^T \mathrm{d}\boldsymbol{v}_i = \sigma_i \boldsymbol{u}_j^T \mathrm{d}\boldsymbol{u}_i \tag{49}$$

$$\mathrm{d}\boldsymbol{u}_i^T \sigma_j \boldsymbol{u}_j + \boldsymbol{u}_i^T \mathrm{d}M\boldsymbol{v}_j = \sigma_i \mathrm{d}\boldsymbol{v}_i^T \boldsymbol{v}_j \tag{50}$$

for all $i \neq j$.

Observe that $\boldsymbol{u}_j^T \mathrm{d}\boldsymbol{u}_i = \mathrm{d}\boldsymbol{u}_i^T \boldsymbol{u}_j,\ \boldsymbol{v}_j^T \mathrm{d}\boldsymbol{v}_i = \mathrm{d}\boldsymbol{v}_i^T \boldsymbol{v}_j$. Combining (49) and (50) and regarding them as a linear equation group about $\boldsymbol{u}_j^T \mathrm{d}\boldsymbol{u}_i$ and $\boldsymbol{v}_j^T \mathrm{d}\boldsymbol{v}_i$ they can be solved out as

$$\boldsymbol{u}_j^T \mathrm{d}\boldsymbol{u}_i = \frac{\sigma_i \boldsymbol{u}_j^T \mathrm{d}M\boldsymbol{v}_i + \sigma_j \boldsymbol{u}_i^T \mathrm{d}M\boldsymbol{v}_j}{\sigma_i^2 - \sigma_j^2} \tag{51}$$

$$\boldsymbol{v}_j^T \mathrm{d}\boldsymbol{v}_i = \frac{\sigma_i \boldsymbol{u}_i^T \mathrm{d}M\boldsymbol{v}_j + \sigma_j \boldsymbol{u}_j^T \mathrm{d}M\boldsymbol{v}_i}{\sigma_i^2 - \sigma_j^2} \tag{52}$$

Since $\boldsymbol{u}_i^T \boldsymbol{u}_i = \|\boldsymbol{u}_i\| = 1$, we have $\boldsymbol{u}_i^T \mathrm{d}\boldsymbol{u}_i = 0$. As $\{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n\}$ form a set of orthogonal basis of $\mathbb{R}^n$, we can write $\boldsymbol{u}_i$ as

$$\mathrm{d}\boldsymbol{u}_i = \sum_{j \neq i} \frac{\sigma_i \boldsymbol{u}_j^T \mathrm{d}M\boldsymbol{v}_i + \sigma_j \boldsymbol{u}_i^T \mathrm{d}M\boldsymbol{v}_j}{\sigma_i^2 - \sigma_j^2} \boldsymbol{u}_j \tag{53}$$

Similarly for $\mathrm{d}\boldsymbol{v}_i$ we have

$$\mathrm{d}\boldsymbol{v}_i = \sum_{j \neq i} \frac{\sigma_i \boldsymbol{u}_i^T \mathrm{d}M\boldsymbol{v}_j + \sigma_j \boldsymbol{u}_j^T \mathrm{d}M\boldsymbol{v}_i}{\sigma_i^2 - \sigma_j^2} \boldsymbol{v}_j \tag{54}$$

Finally we can write $\mathrm{d}R$ as

$$\begin{aligned}\mathrm{d}R &= \sum \boldsymbol{u}_i \mathrm{d}\boldsymbol{v}_i^T + \sum \mathrm{d}\boldsymbol{u}_i \boldsymbol{v}_i^T \\ &= \sum_{i \neq j} \frac{\sigma_i \boldsymbol{u}_i^T \mathrm{d}M \boldsymbol{v}_j + \sigma_j \boldsymbol{u}_j^T \mathrm{d}M \boldsymbol{v}_i}{\sigma_i^2 - \sigma_j^2} \boldsymbol{u}_i \boldsymbol{v}_j^T \\ &\quad + \sum_{i \neq j} \frac{\sigma_i \boldsymbol{u}_j^T \mathrm{d}M \boldsymbol{v}_i + \sigma_j \boldsymbol{u}_i^T \mathrm{d}M \boldsymbol{v}_j}{\sigma_i^2 - \sigma_j^2} \boldsymbol{u}_j \boldsymbol{v}_i^T \\ &= \sum_{i \neq j} \frac{(\sigma_i - \sigma_j) \boldsymbol{u}_i^T \mathrm{d}M \boldsymbol{v}_j - (\sigma_i - \sigma_j) \boldsymbol{u}_j^T \mathrm{d}M \boldsymbol{v}_i}{\sigma_i^2 - \sigma_j^2} \boldsymbol{u}_i \boldsymbol{v}_j^T \\ &= \sum_{i \neq j} \frac{\boldsymbol{u}_i^T \mathrm{d}M \boldsymbol{v}_j - \boldsymbol{u}_j^T \mathrm{d}M \boldsymbol{v}_i}{\sigma_i + \sigma_j} \boldsymbol{u}_i \boldsymbol{v}_j^T\end{aligned}$$

□

C.8.6 Proof of Lemma C.8

Since $\|E\|_{1,\infty} \leq \frac{\overline{\lambda}_2}{2}$, we have

$$\begin{aligned}\overline{L}^+_{-\lambda_j} &= \overline{U}(\overline{\Lambda} - \lambda_j)^{-1}\overline{U}^T \\ &= \overline{U}\sum_{l=0}^{\infty} \overline{L}^{-(l+1)} \lambda_j^l \overline{U}^T \\ &= \sum_{l=0}^{\infty} (\overline{L}^+)^{l+1} \lambda_j^l, \quad 1 \leq j \leq k.\end{aligned}$$

As $\|E\|_{1,\infty}\|\overline{L}^+\|_{1,\infty} < 1$, it follows that

$$\begin{aligned}\|\overline{L}^+_{-\lambda_j}\|_{1,\infty} &\leq \sum_{l=0}^{\infty} \|\overline{L}^+\|_{1,\infty}^{l+1} \lambda_j^l \\ &\leq \sum_{l=0}^{\infty} \|\overline{L}^+\|_{1,\infty}^{l+1} \|E\|_{1,\infty}^l \\ &= \|\overline{L}^+\|_{1,\infty}\big(1 + \|\overline{L}^+\|_{1,\infty}\|E\|_{1,\infty}\big).\end{aligned}$$

□

C.8.7 Proof of Lemma C.9

In fact, $\forall \boldsymbol{x} \in \mathbb{R}^k$, where $\|\boldsymbol{x}\| = 1$, we have

$$\begin{aligned}\|A\boldsymbol{x}\| &\leq \sum_{j=1}^{k} \|A\boldsymbol{e}_j^{(k)}\||x_j| \\ &\leq \max_{1 \leq j \leq k} \|A\boldsymbol{e}_j^{(k)}\| \sum_{j=1}^{k} |x_j| \\ &\leq \max_{1 \leq j \leq k} \|A\boldsymbol{e}_j^{(k)}\| \sqrt{k}(\sum_{j=1}^{k} x_j^2)^{\frac{1}{2}}.\end{aligned}$$

□

D. Scenes used in this paper

For completeness, we show the scenes we used in this paper. Including 100 scenes from ScanNet [17] dataset and 60 scenes from Redwood Chair dataset. Fig. 10-Fig. 10 and Fig. 11-Fig. 11 show the scenes we used in the paper from ScanNet and Redwood chair dataset, respectively.

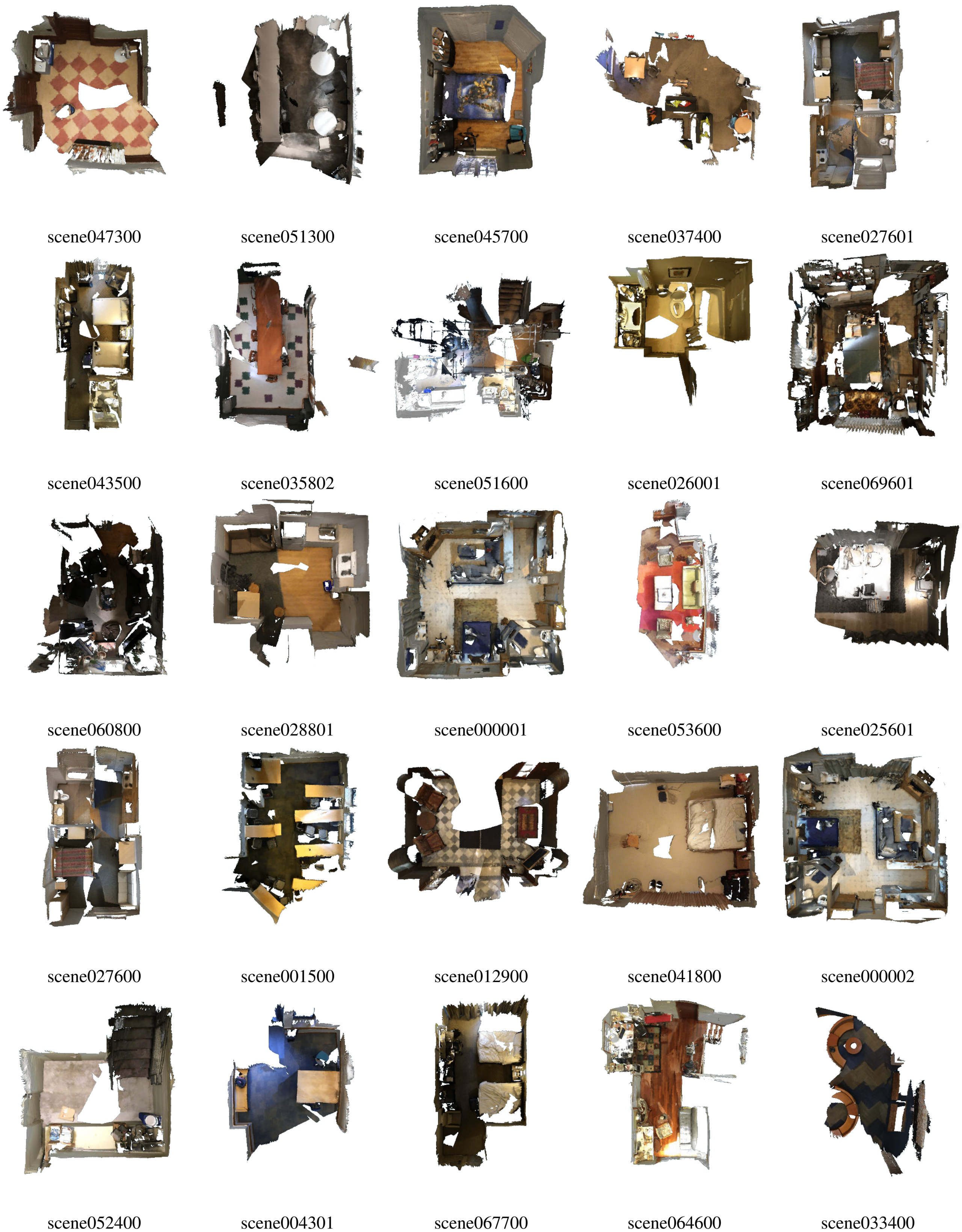


Figure 10: ScanNet Train Dataset (1st to 25th)

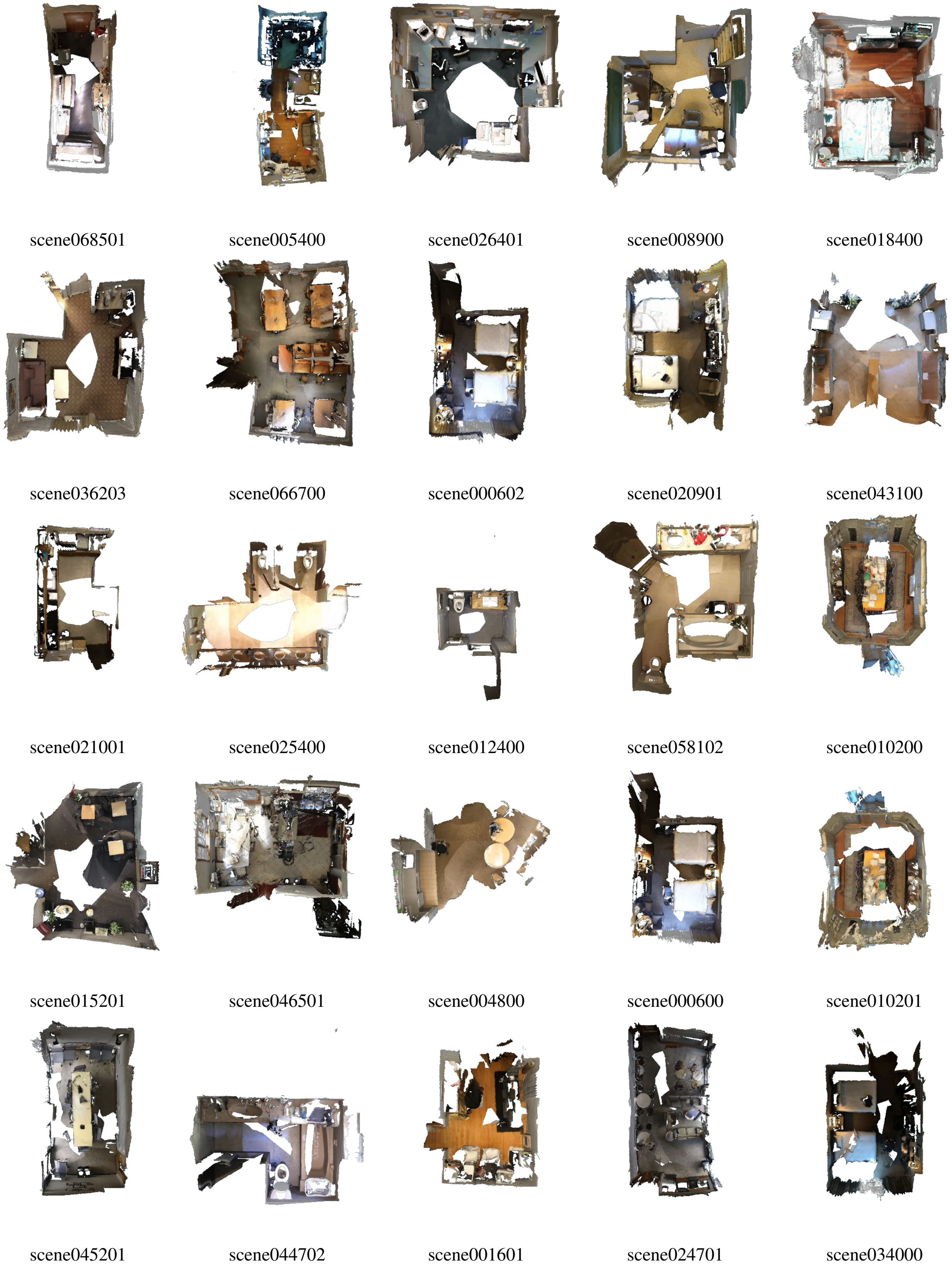


Figure 10: ScanNet Train Dataset (26th to 50th)

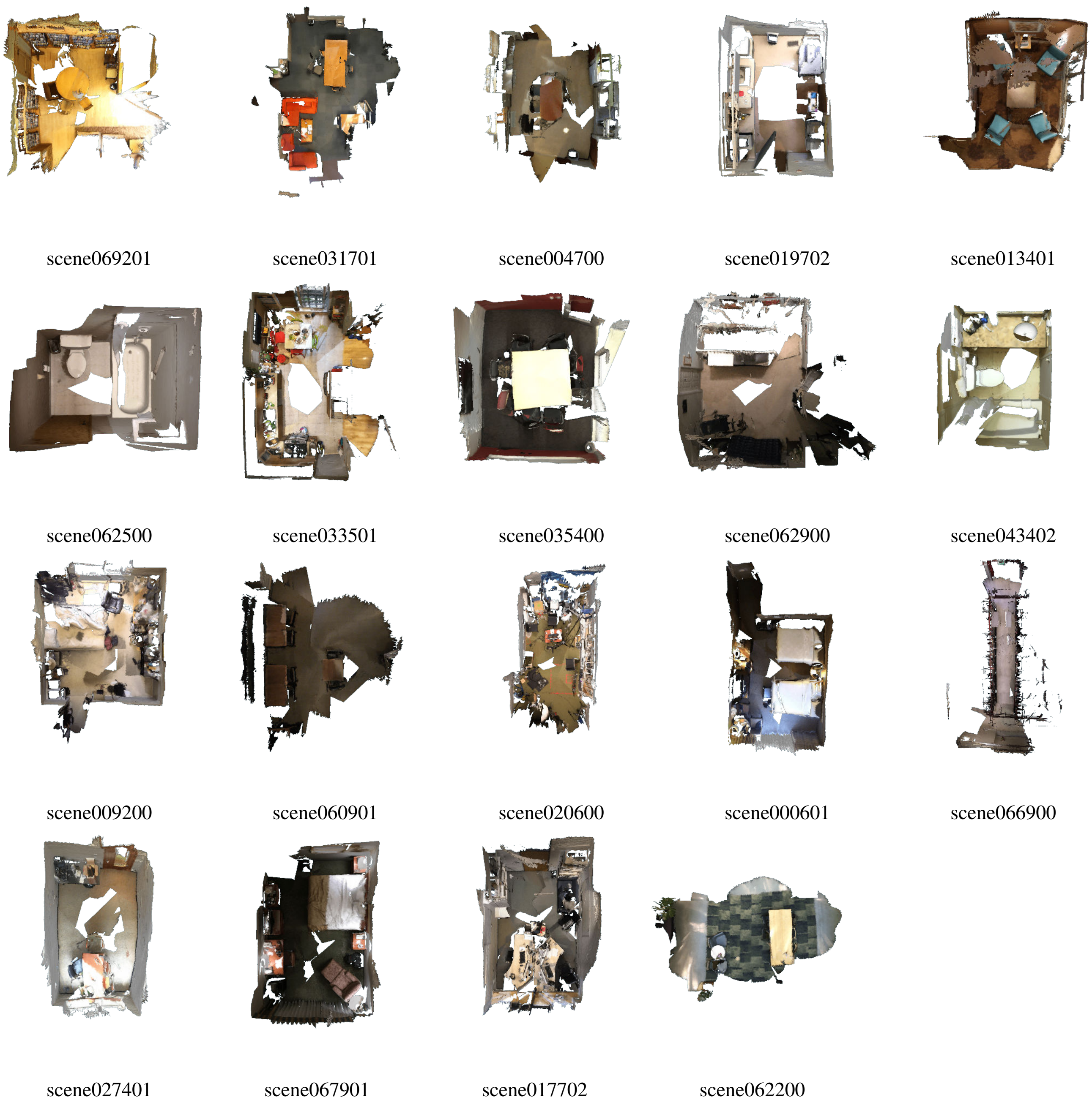
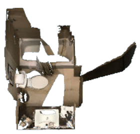


Figure 10: ScanNet Train Dataset (51st to 69th)

Figure 10: ScanNet Test Dataset (1st to 25th)



scene040602

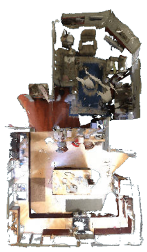
scene035300

scene022300

scene014602

scene020800

scene057502



scene023101

Figure 10: ScanNet Test Dataset (26th to 32nd)

Figure 11: Redwood Chair Train Dataset (1st to 25th)

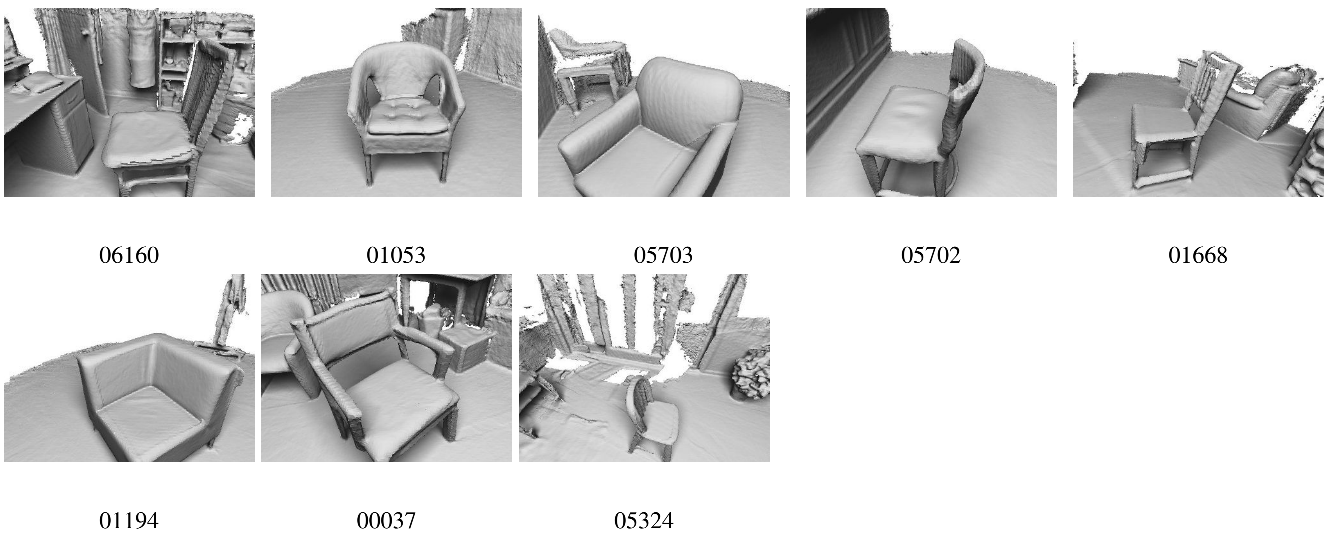


Figure 11: Redwood Chair Train Dataset (26th to 33th)

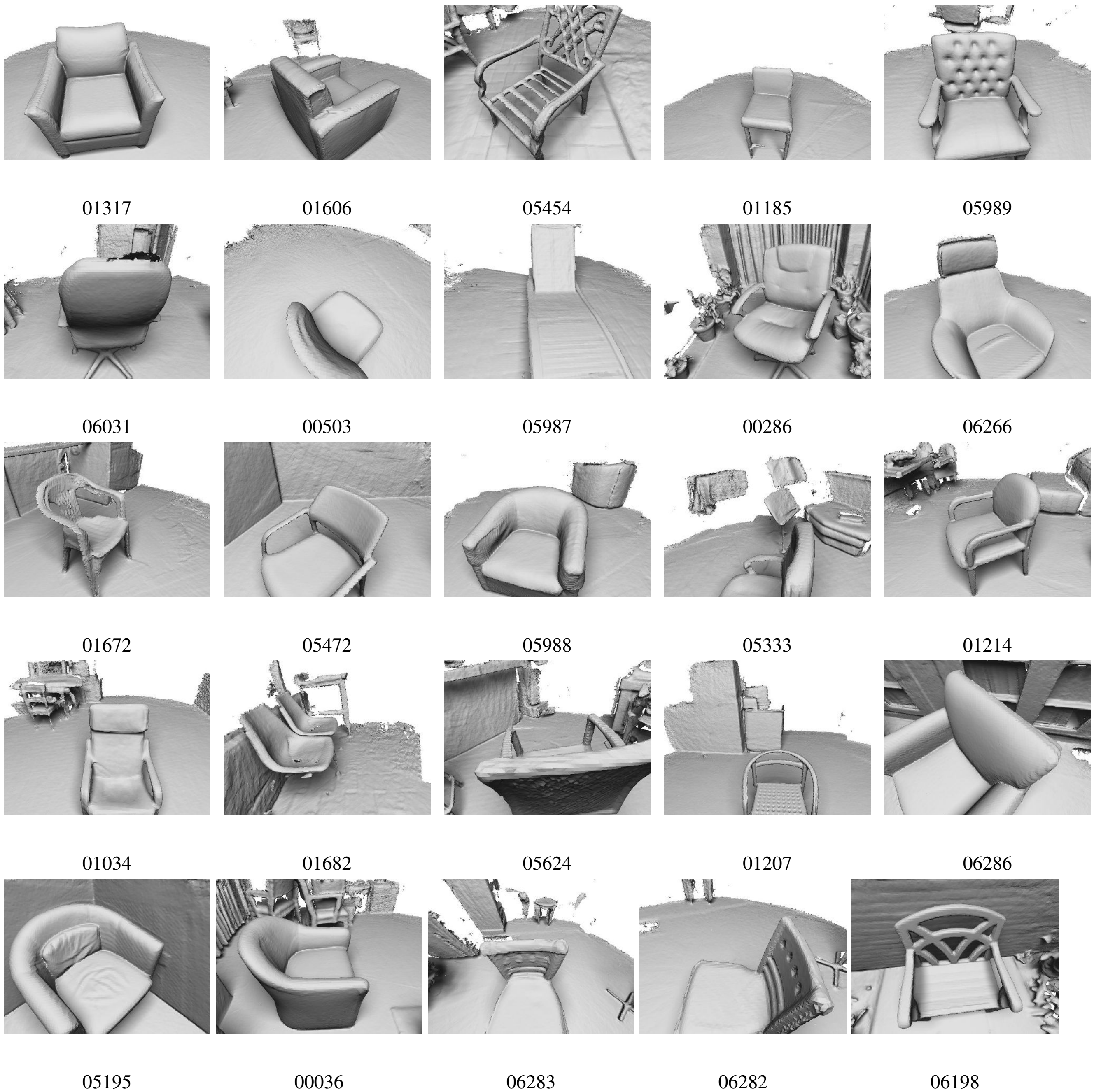


Figure 11: Redwood Chair Test Dataset