PRaDA: Projective Radial Distortion Averaging

Supplementary Material

A. Implementation Details

Following best practices from COLMAP, we estimate initial pairwise models in parallel with image matching. This adds almost no overhead to the feature-matching process. Our method then starts with already estimated pairwise models. We implement the proposed method in C++. Unlike in incremental reconstruction, the global approach allows parallel computation of all components, providing true scalability, so we parallelize every possible step. We use TBB [42] for parallel execution on the CPU. Since all derivatives are computed in closed form, a CUDA implementation is also feasible but beyond the scope of the current work.

For the optimization phase of LO-RANSAC [9], we compute the derivatives of the Sampson error with respect to both cameras and fundamental matrix according to the parameterization described in Sec. 4.2. Our experiments show that, typically, ten iterations of local refinement are sufficient. The weights in Eq. (16) are based on the area covered by the matches. This way, the actual overlap between images is taken into account compared to the naive number of inliers. This handles outlier cameras that may occur after pairwise estimation. The final global refinement (see Sec. 4.5) is done with analytical derivatives using nonlinear solvers from ceres-solver [1]. We run it several times, re-estimating the inliers between successive runs. In the first pass of the global projective refinement, we fix the camera centers and optimize only the distortions and fundamental matrices. The camera centers are then optimized while all other camera parameters are fixed.

B. Performance by sequence length

We evaluate the performance of the proposed algorithm on the ScanNet++ dataset under varying numbers of input frames. For each sequence length N, ranging from 2 to 40, we randomly sample 1000 N-frame subsequences (20 from each of the 50 ScanNet++ test sequences). We then run our method on each subsequence and plot the average focal-adjusted reprojection error in Fig. 7. Frames are 5 frames apart, relative to the original ScanNet++ frame rate.

Comparing with the full-sequence results in Tab. 1, we observe that with 31 input frames, the proposed algorithm achieves a mean error (1.0 px) lower than that of Colmap [48, 49] (2.0 px) and Glomap [41] (1.8 px), even when these methods process the entire frameset.

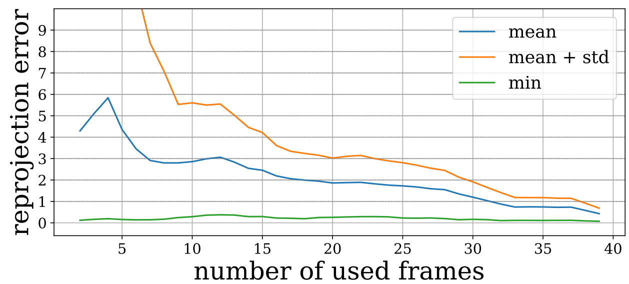

Figure 7. **Mean error for randomly sampled N-frame sequences:** We study the behavior of the proposed algorithm with different numbers of input frames. We sample 20 random sequences of frames for each *number of frames* configuration. We find that 31 frames are sufficient to obtain a competitive 1px mean error.

C. Distortion averaging for the multiplicative model

Our radial distortion averaging formulation can be adapted for any distortion parametrization $d_\theta(r)$. In terms of the model parameters, θ, the associated optimization problem (repeated here for convenience) is formulated as:

$$\bar{\theta} = \underset{\theta}{\operatorname{argmin}} \sum_{i=1}^{n} \omega_i \int_0^R \left\| d_\theta(r) - d_{\theta_i}(r) \right\|^2 r^3 \, \mathrm{d}r. \tag{22}$$

For our method, we use the divisional distortion model: $d_\theta(r) = 1/h_\theta(r)$, where $h_\theta(r) = \sum_{j=0}^{k} \theta_j r^j$ is a degree-k polynomial with coefficients given by the vector θ. Primarily because of the availability of the well-proven F10 solver [30] and computationally efficient derivatives for the Sampson error [15, 44].

For the division model, Eq. (22) cannot be evaluated exactly, so we optimize over a numerical discretization with uniform spacing. Interestingly, it turns out that for the *multiplicative distortion model*, the average distortion as defined by Eq. (22) corresponds to averaging the distortion parameters. That is:

$$\bar{\theta} = \frac{\sum_{i=1}^{k} w_i \theta_i}{\sum_{i=1}^{k} w_i} \tag{23}$$

when $d_\theta(r)$ is parametrized as a degree-k polynomial: $d_\theta(r) = h_\theta(r)$. To see this, let us rewrite Eq. (22) for the multiplicative distortion model:

$$\bar{\theta} = \underset{\theta}{\operatorname{argmin}} \underbrace{\sum_{i=1}^{n} \omega_i \int_0^R \left\| \sum_{j=0}^{k} \theta_j r^j - \sum_{j=0}^{k} \theta_j^i r^j \right\|^2 r^3 \, \mathrm{d}r}_{:=L(\theta)} \tag{24}$$

To find the minimum, we take the derivatives of $L(\theta)$ with respect to the coefficients θ_t:

$$\begin{aligned} \frac{dL}{d\theta_t} &= 2 \sum_{i=1}^{n} \omega_i \int_0^R \left(\sum_{j=0}^{k} \theta_j r^j - \sum_{j=0}^{k} \theta_j^i r^j \right) r^{3+t} \, \mathrm{d}r \\ &= 2 \sum_{i=1}^{n} \sum_{j=0}^{k} w_i A_{tj} (\theta_j - \theta_j^i). \end{aligned} \tag{25}$$

Where

$$A_{ti} = \int_0^R r^{j+t+3} \, \mathrm{d}r = \frac{R^{j+t+4}}{j+t+4}. \tag{26}$$

Rearranging and solving for $\frac{dL}{d\theta_t} = 0$ we get:

$$\sum_{j=0}^{k} A_{tj} \theta_j = \frac{1}{\sum_{i=1}^{n} \omega_i} \sum_{j=0}^{k} A_{tj} \sum_{i=1}^{n} \omega_i \theta_j^i \tag{27}$$

which in matrix form corresponds to

$$A\theta = A\bar{\theta} \tag{28}$$

where $\bar{\theta}$ is the weighted average parameter, as defined in Eq. (23). Since A is invertible, we get that the solution to Eq. (24) is $\theta = \bar{\theta}$. In this case, θ is independent of the radius R. This means that for *multiplicative distortion model*, averaging across an image is equivalent to averaging across the entire space of $\mathbb{R}^2$, including areas outside the image.