# MoRe: Monocular Geometry Refinement via Graph Optimization for Cross-View Consistency

Supplementary Material

## 1. Performance Analysis

### 1.1. Runtime and Memory

| Type | Resolution | # Match | $k$ of kNN | GPU (L40S) | time |
|---|---|---|---|---|---|
| Default | $512 \times 377$ | 5000 | 4 | 838 MiB | 1.6 s |
| Res. ↑ | **$640 \times 471$** | 5000 | 4 | 1354 MiB | 2.5 s |
| Matches ↑ | $512 \times 377$ | **10000** | 4 | 850 MiB | 1.6 s |
| KNN k ↑ | $512 \times 377$ | 5000 | **16** | 1152 MiB | 2.0 s |

Table 1. Runtime and Memory. The default represents our experimental setting, while the other types correspond to modified settings for analysis.

Table 1 shows runtime and GPU memory usage with respect to resolution, the number of matching points, and the number of nearest neighbors $k$.

### 1.2. Parameter Stability

| Type ↓ | $\lambda_p$ | $\lambda_r$ | $\lambda_s$ | $\lambda_n$ | $k$ of kNN | rel↓ | $\tau$ ↑ |
|---|---|---|---|---|---|---|---|
| Default | 30 | 50 | 0.1 | 10 | 4 | 3.12 | 70.29 |
| $\lambda_p \downarrow$ | **15** | 50 | 0.1 | 10 | 4 | 3.13 | 70.23 |
| $\lambda_p \uparrow$ | **45** | 50 | 0.1 | 10 | 4 | 3.13 | 70.20 |
| $\lambda_r \downarrow$ | 30 | **25** | 0.1 | 10 | 4 | 3.13 | 70.29 |
| $\lambda_r \uparrow$ | 30 | **75** | 0.1 | 10 | 4 | 3.14 | 70.16 |
| $\lambda_s \downarrow$ | 30 | 50 | **0.05** | 10 | 4 | 3.13 | 70.21 |
| $\lambda_s \uparrow$ | 30 | 50 | **0.2** | 10 | 4 | 3.13 | 70.29 |
| $\lambda_n \downarrow$ | 30 | 50 | 0.1 | **5** | 4 | 3.13 | 70.22 |
| $\lambda_n \uparrow$ | 30 | 50 | 0.1 | **20** | 4 | 3.13 | 70.21 |
| k ↑ | 30 | 50 | 0.1 | 10 | **16** | 3.12 | 70.36 |

Table 2. Sensitivity to Parameter Settings. The default represents our experimental setting, while the other types illustrate stability checks.

Table 2 reports the performance in terms of Absolute Relative Error (rel) and Inlier Ratio ($\gamma$) with a threshold of 1.03 under varying parameter settings. The results show overall consistent performance across different settings. A larger number of nearest neighbors $k$ improves performance, but considering computational cost, we fix $k = 4$ in our experiments.

## 2. Additional Experimental Results

### 2.1. 3D Reconstruction

We present additional qualitative results of our point alignment method. Figure 1 shows the results for Case 1 (with given poses), and Figure 2 shows the results for Case 2 (without given poses). We used the Tanks and Temples [5], ETH3D [7], ScanNet [3], Matterport3D [2], KITTI [4], and DTU [1] datasets in this experiment.

### 2.2. Novel View Synthesis

To supplement Table 4 in the main paper, we present per-scene experimental results in terms of PSNR, SSIM, and LPIPS, as shown in Table 3, 4 and 5.

## References


[1] Henrik Aanæs, Rasmus Ramsbøl Jensen, George Vogiatzis, Engin Tola, and Anders Bjorholm Dahl. Large-scale data for multiple-view stereopsis. *International Journal of Computer Vision*, 120:153–168, 2016. 1

[2] Angel Chang, Angela Dai, Thomas Funkhouser, Maciej Halber, Matthias Niessner, Manolis Savva, Shuran Song, Andy Zeng, and Yinda Zhang. Matterport3d: Learning from rgb-d data in indoor environments. *arXiv preprint arXiv:1709.06158*, 2017. 1

[3] Angela Dai, Angel X Chang, Manolis Savva, Maciej Halber, Thomas Funkhouser, and Matthias Nießner. Scannet: Richly-annotated 3d reconstructions of indoor scenes. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 5828–5839, 2017. 1

[4] Andreas Geiger, Philip Lenz, Christoph Stiller, and Raquel Urtasun. Vision meets robotics: The kitti dataset. *The international journal of robotics research*, 32(11):1231–1237, 2013. 1

[5] Arno Knapitsch, Jaesik Park, Qian-Yi Zhou, and Vladlen Koltun. Tanks and temples: Benchmarking large-scale scene reconstruction. *ACM Transactions on Graphics (ToG)*, 36(4): 1–13, 2017. 1

[6] Dmytro Kotovenko, Olga Grebenkova, and Björn Ommer. Edgs: Eliminating densification for efficient convergence of 3dgs. *arXiv preprint arXiv:2504.13204*, 2025. 4

[7] Thomas Schops, Johannes L Schonberger, Silvano Galliani, Torsten Sattler, Konrad Schindler, Marc Pollefeys, and Andreas Geiger. A multi-view stereo benchmark with high-resolution images and multi-camera videos. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 3260–3269, 2017. 1

[8] Shuzhe Wang, Vincent Leroy, Yohann Cabon, Boris Chidlovskii, and Jerome Revaud. Dust3r: Geometric 3d vision made easy. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 20697–20709, 2024. 4

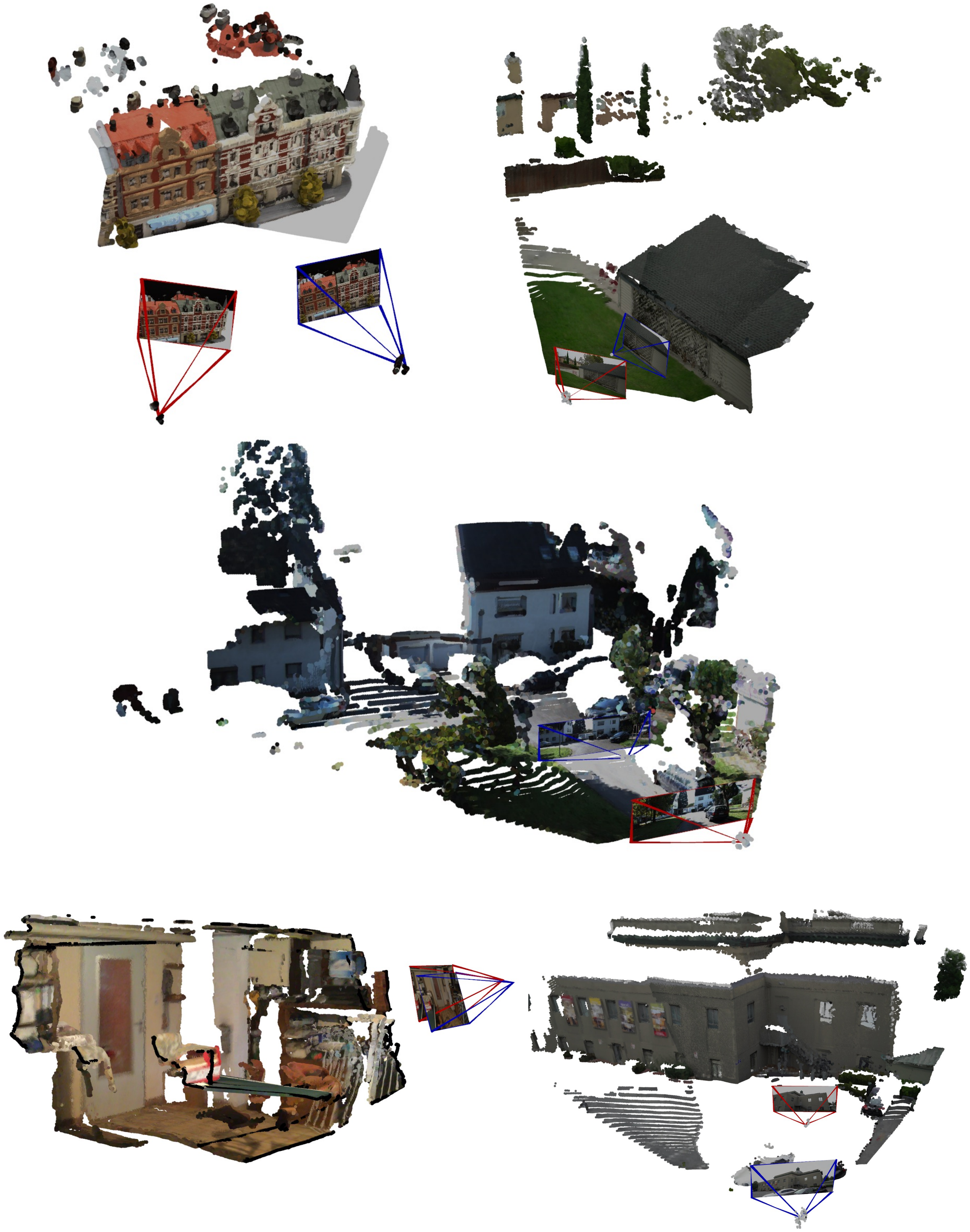

Figure 1. **Additional Qualitative Results of MoRe**

Figure 2. **Additional Qualitative Results of MoRe**

| Method | Steps | Ballroom | | | Barn | | | Family | | | Francis | | | Horse | | | Ignatius | | | Museum | | | Mean | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | PSNR ↑ | SSIM ↑ | LPIPS ↓ | PSNR ↑ | SSIM ↑ | LPIPS ↓ | PSNR ↑ | SSIM ↑ | LPIPS ↓ | PSNR ↑ | SSIM ↑ | LPIPS ↓ | PSNR ↑ | SSIM ↑ | LPIPS ↓ | PSNR ↑ | SSIM ↑ | LPIPS ↓ | PSNR ↑ | SSIM ↑ | LPIPS ↓ | PSNR ↑ | SSIM ↑ | LPIPS ↓ |
| DUSt3R [8] | 200 | 10.65 | 0.202 | 0.592 | 15.60 | 0.497 | 0.490 | 11.54 | 0.351 | 0.615 | 16.85 | 0.520 | 0.508 | 12.92 | 0.512 | 0.523 | 15.16 | 0.306 | 0.587 | 13.65 | 0.312 | 0.592 | 13.77 | 0.386 | 0.558 |
| EDGS [6] | 200 | 9.98 | 0.338 | 0.533 | 5.99 | 0.214 | 0.610 | 7.93 | 0.326 | 0.558 | 10.93 | 0.224 | 0.505 | 3.71 | 0.144 | 0.624 | 10.59 | 0.368 | 0.521 | 12.37 | 0.440 | 0.486 | 8.79 | 0.293 | 0.548 |
| Ours-align | 200 | 17.79 | 0.561 | 0.340 | 17.78 | 0.606 | 0.364 | 16.46 | 0.644 | 0.340 | 18.94 | 0.637 | 0.395 | 15.08 | 0.596 | 0.425 | 19.62 | 0.523 | 0.402 | 20.29 | 0.689 | 0.254 | 17.99 | 0.608 | 0.360 |
| Ours-full | 200 | 17.31 | 0.338 | 0.547 | 17.03 | 0.365 | 0.365 | 18.70 | 0.697 | 0.291 | 18.73 | 0.652 | 0.350 | 17.63 | 0.670 | 0.370 | 19.69 | 0.396 | 0.396 | 20.28 | 0.689 | 0.242 | 18.48 | 0.544 | 0.366 |
| DUSt3R [8] | 1000 | 10.45 | 0.170 | 0.533 | 17.36 | 0.509 | 0.415 | 11.40 | 0.307 | 0.591 | 17.55 | 0.501 | 0.468 | 12.63 | 0.468 | 0.500 | 15.03 | 0.270 | 0.510 | 14.26 | 0.305 | 0.538 | 14.10 | 0.362 | 0.508 |
| EDGS [6] | 1000 | 17.03 | 0.546 | 0.383 | 18.90 | 0.628 | 0.383 | 19.31 | 0.670 | 0.404 | 20.86 | 0.652 | 0.389 | 16.79 | 0.657 | 0.402 | 18.38 | 0.545 | 0.418 | 18.46 | 0.625 | 0.333 | 18.53 | 0.617 | 0.387 |
| Ours-align | 1000 | 18.32 | 0.245 | 0.568 | 18.09 | 0.579 | 0.292 | 18.51 | 0.639 | 0.248 | 21.02 | 0.641 | 0.323 | 17.17 | 0.594 | 0.309 | 19.32 | 0.283 | 0.283 | 20.75 | 0.682 | 0.184 | 19.03 | 0.523 | 0.315 |
| Ours-full | 1000 | 17.62 | 0.251 | 0.549 | 18.32 | 0.282 | 0.591 | 21.10 | 0.729 | 0.210 | 21.18 | 0.650 | 0.287 | 20.98 | 0.721 | 0.243 | 19.47 | 0.273 | 0.273 | 20.82 | 0.682 | 0.171 | 19.93 | 0.513 | 0.332 |

Table 3. **Breakdown results** on Tanks & Temples dataset for novel-view synthesis with **3 training views**. Red, orange, and yellow indicate the first, second, and third best performing algorithms for each metric.

| Method | Steps | Ballroom | | | Barn | | | Family | | | Francis | | | Horse | | | Ignatius | | | Museum | | | Mean | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | PSNR ↑ | SSIM ↑ | LPIPS ↓ | PSNR ↑ | SSIM ↑ | LPIPS ↓ | PSNR ↑ | SSIM ↑ | LPIPS ↓ | PSNR ↑ | SSIM ↑ | LPIPS ↓ | PSNR ↑ | SSIM ↑ | LPIPS ↓ | PSNR ↑ | SSIM ↑ | LPIPS ↓ | PSNR ↑ | SSIM ↑ | LPIPS ↓ | PSNR ↑ | SSIM ↑ | LPIPS ↓ |
| DUSt3R [8] | 200 | 11.31 | 0.220 | 0.608 | 16.08 | 0.519 | 0.485 | 12.09 | 0.372 | 0.615 | 17.17 | 0.527 | 0.508 | 13.68 | 0.522 | 0.515 | 15.72 | 0.327 | 0.595 | 14.41 | 0.358 | 0.596 | 14.35 | 0.406 | 0.560 |
| EDGS [6] | 200 | 10.98 | 0.466 | 0.451 | 6.40 | 0.333 | 0.529 | 8.98 | 0.470 | 0.446 | 11.11 | 0.326 | 0.428 | 4.33 | 0.280 | 0.512 | 12.68 | 0.448 | 0.466 | 15.42 | 0.560 | 0.427 | 9.98 | 0.412 | 0.466 |
| Ours-align | 200 | 19.02 | 0.611 | 0.319 | 19.77 | 0.681 | 0.310 | 18.14 | 0.680 | 0.301 | 19.67 | 0.671 | 0.361 | 16.35 | 0.640 | 0.399 | 20.32 | 0.548 | 0.398 | 21.60 | 0.716 | 0.254 | 19.27 | 0.650 | 0.334 |
| Ours-full | 200 | 19.09 | 0.644 | 0.303 | 19.64 | 0.670 | 0.307 | 18.51 | 0.680 | 0.297 | 23.09 | 0.705 | 0.275 | 17.58 | 0.669 | 0.350 | 20.28 | 0.547 | 0.395 | 21.72 | 0.722 | 0.245 | 19.99 | 0.662 | 0.310 |
| DUSt3R [8] | 1000 | 11.56 | 0.190 | 0.551 | 19.28 | 0.557 | 0.395 | 12.20 | 0.324 | 0.590 | 19.29 | 0.526 | 0.455 | 13.76 | 0.475 | 0.478 | 16.28 | 0.304 | 0.507 | 15.70 | 0.357 | 0.538 | 15.44 | 0.390 | 0.502 |
| EDGS [6] | 1000 | 18.84 | 0.655 | 0.297 | 20.95 | 0.682 | 0.314 | 22.16 | 0.749 | 0.304 | 23.73 | 0.737 | 0.303 | 20.46 | 0.747 | 0.301 | 20.82 | 0.604 | 0.354 | 20.69 | 0.708 | 0.257 | 21.09 | 0.697 | 0.304 |
| Ours-align | 1000 | 20.17 | 0.640 | 0.218 | 21.17 | 0.672 | 0.229 | 20.19 | 0.690 | 0.212 | 20.29 | 0.697 | 0.318 | 19.81 | 0.682 | 0.269 | 20.54 | 0.545 | 0.263 | 22.63 | 0.734 | 0.170 | 20.69 | 0.666 | 0.240 |
| Ours-full | 1000 | 20.69 | 0.689 | 0.202 | 21.60 | 0.668 | 0.220 | 20.03 | 0.692 | 0.213 | 24.27 | 0.731 | 0.234 | 20.31 | 0.692 | 0.238 | 20.59 | 0.551 | 0.258 | 23.15 | 0.750 | 0.157 | 21.52 | 0.682 | 0.217 |

Table 4. **Breakdown results** on Tanks & Temples dataset for novel-view synthesis with **6 training views**. Red, orange, and yellow indicate the first, second, and third best performing algorithms for each metric.

| Method | Steps | Ballroom | | | Barn | | | Family | | | Francis | | | Horse | | | Ignatius | | | Museum | | | Mean | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | PSNR ↑ | SSIM ↑ | LPIPS ↓ | PSNR ↑ | SSIM ↑ | LPIPS ↓ | PSNR ↑ | SSIM ↑ | LPIPS ↓ | PSNR ↑ | SSIM ↑ | LPIPS ↓ | PSNR ↑ | SSIM ↑ | LPIPS ↓ | PSNR ↑ | SSIM ↑ | LPIPS ↓ | PSNR ↑ | SSIM ↑ | LPIPS ↓ | PSNR ↑ | SSIM ↑ | LPIPS ↓ |
| DUSt3R [8] | 200 | 11.68 | 0.234 | 0.620 | 15.64 | 0.525 | 0.502 | 12.33 | 0.394 | 0.621 | 16.87 | 0.560 | 0.493 | 14.04 | 0.542 | 0.513 | 15.59 | 0.336 | 0.612 | 14.67 | 0.387 | 0.608 | 14.40 | 0.426 | 0.567 |
| EDGS [6] | 200 | 11.58 | 0.519 | 0.412 | 7.00 | 0.423 | 0.479 | 9.60 | 0.556 | 0.385 | 11.60 | 0.369 | 0.423 | 4.74 | 0.349 | 0.468 | 12.93 | 0.482 | 0.442 | 17.11 | 0.626 | 0.368 | 10.65 | 0.475 | 0.425 |
| Ours-align | 200 | 19.35 | 0.663 | 0.286 | 20.06 | 0.702 | 0.295 | 18.12 | 0.702 | 0.299 | 19.39 | 0.670 | 0.365 | 15.98 | 0.650 | 0.394 | 20.31 | 0.564 | 0.398 | 22.10 | 0.739 | 0.240 | 19.33 | 0.670 | 0.325 |
| Ours-full | 200 | 19.07 | 0.655 | 0.293 | 20.11 | 0.698 | 0.292 | 21.14 | 0.736 | 0.212 | 20.60 | 0.700 | 0.330 | 17.64 | 0.670 | 0.370 | 20.20 | 0.554 | 0.396 | 22.08 | 0.730 | 0.234 | 20.12 | 0.678 | 0.304 |
| DUSt3R [8] | 1000 | 11.99 | 0.213 | 0.578 | 19.80 | 0.595 | 0.395 | 13.13 | 0.355 | 0.592 | 21.25 | 0.622 | 0.392 | 14.79 | 0.501 | 0.469 | 16.69 | 0.331 | 0.526 | 16.36 | 0.402 | 0.550 | 16.29 | 0.431 | 0.500 |
| EDGS [6] | 1000 | 19.25 | 0.687 | 0.264 | 22.98 | 0.734 | 0.260 | 23.60 | 0.794 | 0.253 | 24.61 | 0.762 | 0.277 | 21.51 | 0.783 | 0.249 | 21.85 | 0.632 | 0.335 | 22.47 | 0.703 | 0.218 | 22.32 | 0.728 | 0.265 |
| Ours-align | 1000 | 21.92 | 0.728 | 0.186 | 22.68 | 0.698 | 0.292 | 18.70 | 0.697 | 0.291 | 23.01 | 0.720 | 0.280 | 20.36 | 0.705 | 0.271 | 21.24 | 0.593 | 0.273 | 23.93 | 0.778 | 0.159 | 21.69 | 0.703 | 0.250 |
| Ours-full | 1000 | 21.42 | 0.713 | 0.197 | 23.28 | 0.734 | 0.190 | 21.11 | 0.729 | 0.210 | 25.00 | 0.753 | 0.232 | 20.98 | 0.721 | 0.243 | 21.20 | 0.590 | 0.270 | 23.75 | 0.767 | 0.155 | 22.39 | 0.715 | 0.214 |

Table 5. **Breakdown results** on Tanks & Temples dataset for novel-view synthesis with **12 training views**. Red, orange, and yellow indicate the first, second, and third best performing algorithms for each metric.