Adapting Stereo Vision From Objects To 3D Lunar Surface Reconstruction with the StereoLunar Dataset (supplementary material)

Anonymous ICCV submission

Paper ID 8

1. Additional Results on Real Lunar Landing Images

We present three additional stereo pairs from the Chang'e 4 landing to test our method on real lunar imagery: a dynamic case with altitude and viewpoint variation (Fig. 1), an oblique-oblique pair near touchdown (Fig. 2), and a vertical nadir descent with altitude difference (Fig. 3). These examples demonstrate our model's robustness to diverse real-world configurations. Each figure compares the output of MASt3R (right column) with our fine-tuned model (left column), including slope maps, hillshaded depths, and 3D reconstructions.

2. Additional Simulated Results on Challenging Cases

We present three synthetic stereo pairs designed to stress-test both our proposed pose estimation and 3D reconstruction under adverse conditions. These edge cases include: (1) a low-light scene over flat terrain, (2) a pair of non-overlapping views, and (3) two nearly identical images with minimal disparity. Despite these challenges, our model recovers accurate poses and coherent 3D structures in all three cases, whereas MASt3R struggles, producing noisy or collapsed reconstructions. These results are illustrated in Fig. 4.

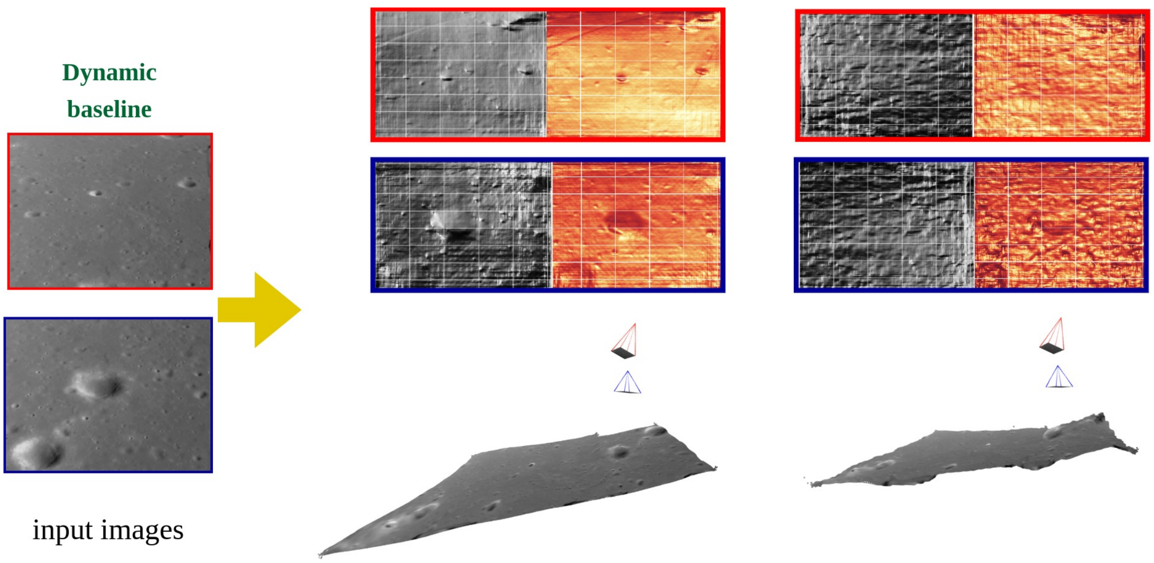

Figure 1. Dynamic trajectory with varying altitudes and oblique viewpoints, similar to our dataset. For each of the two input views, from real Change'E landing (outlined in red and blue, respectively), we show the predicted hillshaded depth maps and slope maps (heat colormap), followed by the reconstructed 3D scene (bottom). **Left:** *Ours*. **Right:** *MASt3R*. Both models recover plausible poses, but only our method reconstructs fine-scale terrain details: crater rims and shadowed slopes appear sharper and more structurally consistent.

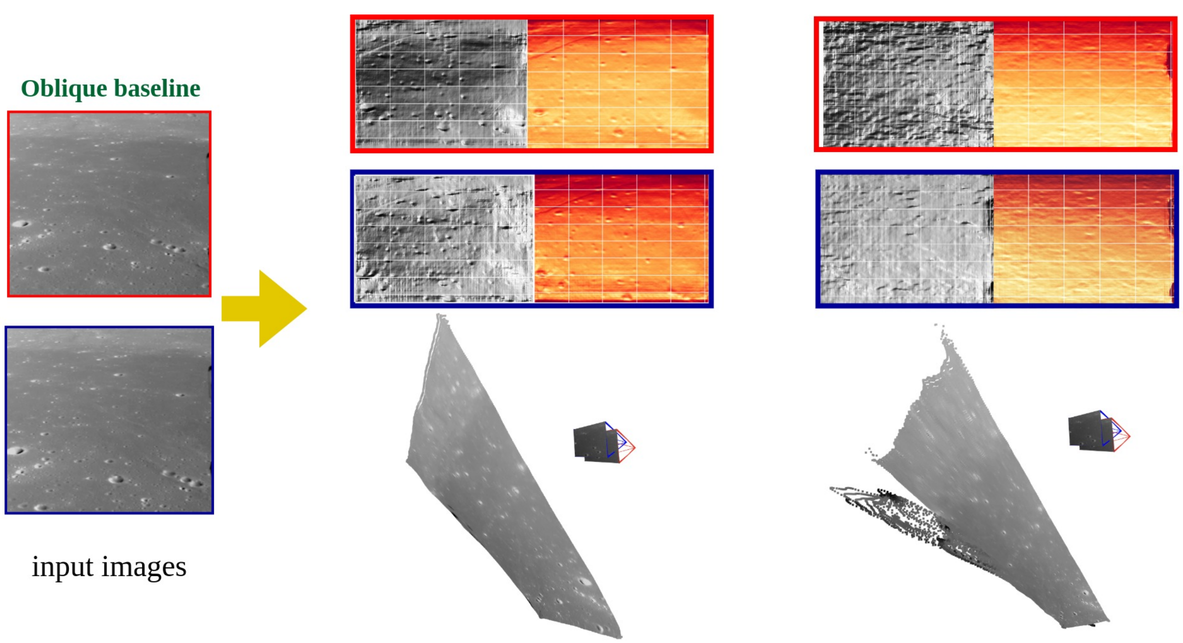

Figure 2. Oblique trajectory near touchdown. As in the previous example, we show for each of the two input views, from real Change'E landing (outlined in red and blue, respectively), we show the predicted hillshaded depth maps and slope maps (heat colormap), followed by the reconstructed 3D scene (bottom). **Left:** *Ours*. **Right:** *MASt3R*. While both methods estimate reasonable poses, our model provides sharper gradient transitions and more coherent topographic discontinuities, particularly along crater rims and slope breaks.

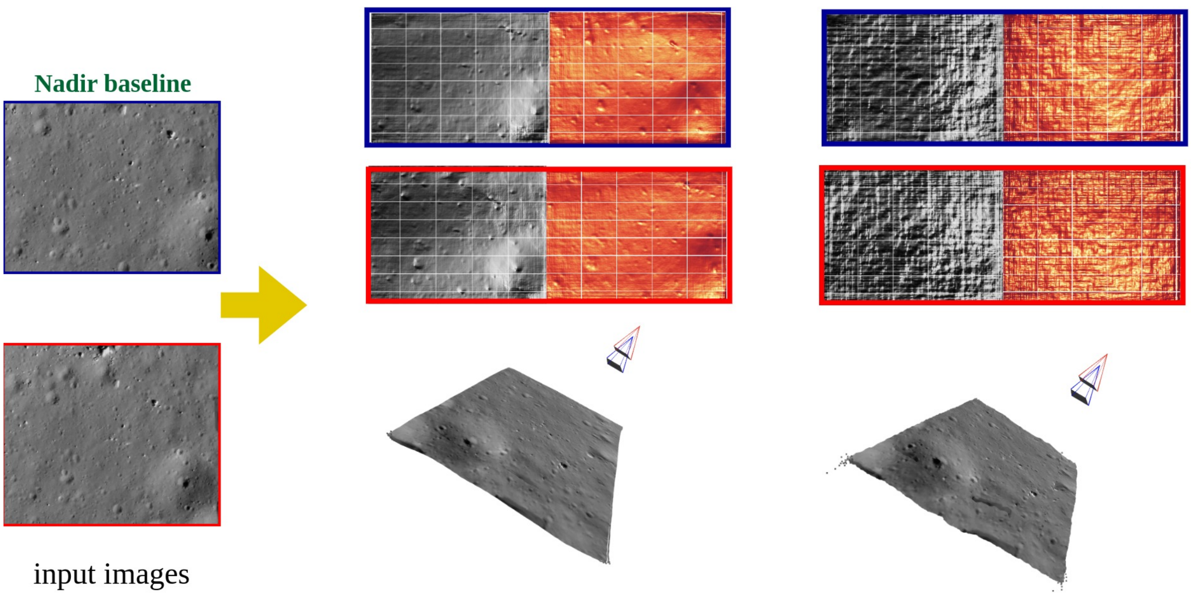

Figure 3. Vertical descent with nadir orientation. This configuration is uncommon in our dataset, where nadir views usually correspond to horizontal motion. For each of the two input views, from real Change'E landing (outlined in red and blue, respectively), we display the predicted hillshaded depth maps and slope maps, followed by the full 3D reconstruction. **Left:** *Ours*. **Right:** *MASt3R*. Both methods estimate consistent poses, but MASt3R produces noisy geometry with limited structural detail, while our method captures more distinct terrain relief, particularly around crater rims.

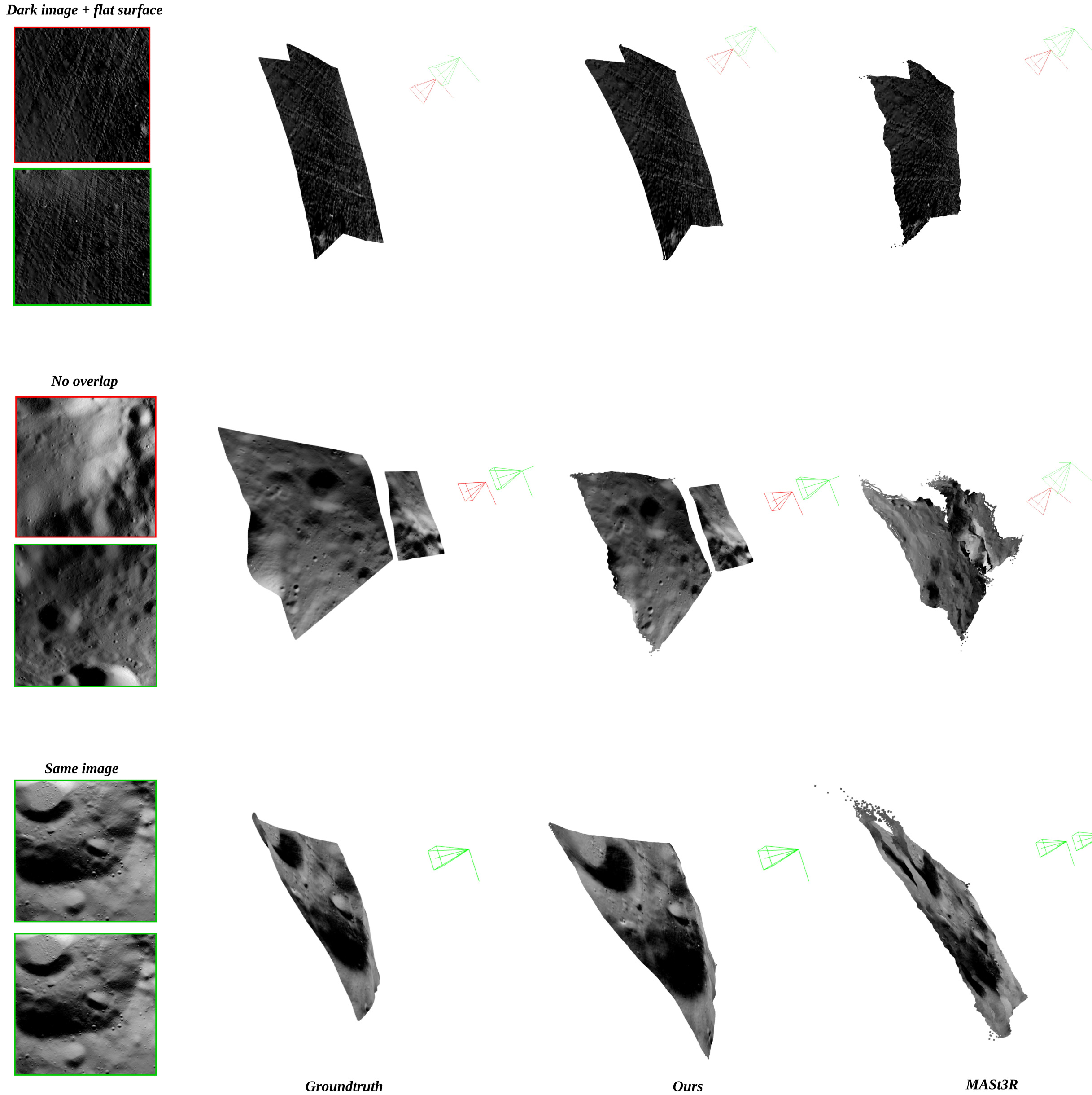

Figure 4. Examples of 3D reconstruction in three challenging simulation cases. Each row shows a stereo image pair followed by reconstructed 3D point clouds from ground truth, MASt3R (middle), and our method (right). **Top row:** low-light images over flat terrain — both models estimate plausible poses, but only ours recovers usable geometry. **Middle row:** two non-overlapping images — MASt3R aligns the views incorrectly and produces collapsed geometry, while our model handles the mismatch robustly. **Bottom row:** a test with a pair including the same image, while MASt3R wrongly estimates a larger displacement between the views and a poor 3D reconstruction, our method correctly estimates the same pose for both views and a more reliable 3D reconstruction.