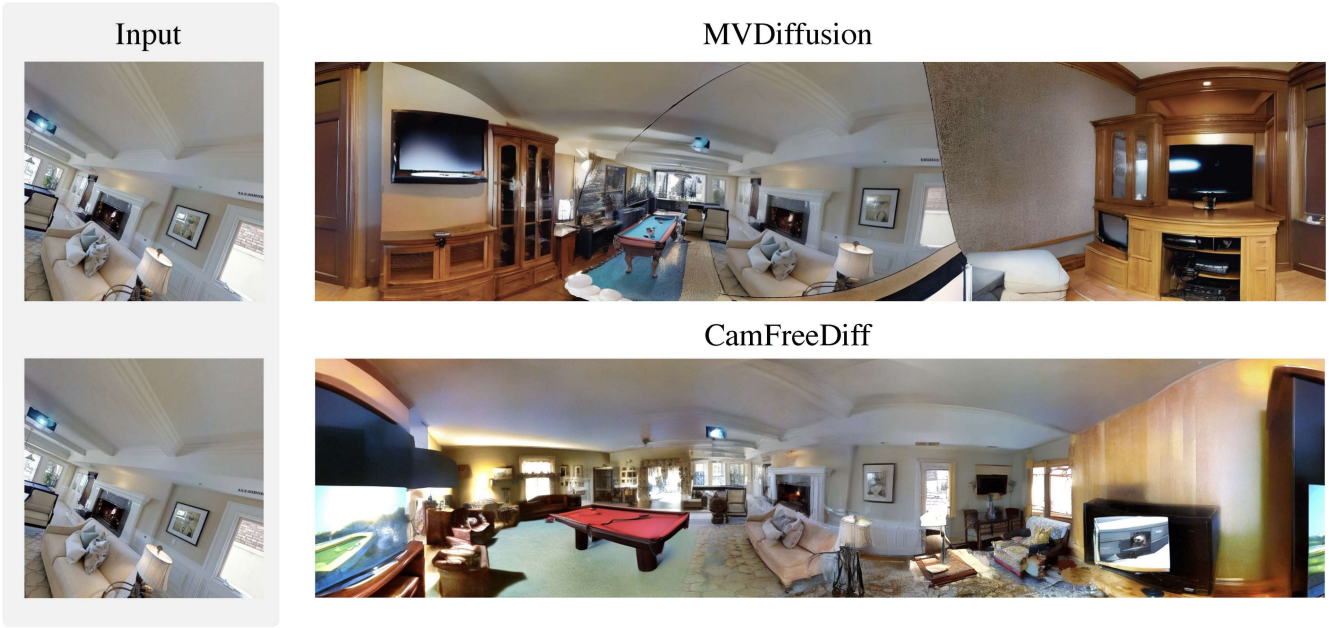
CamFreeDiff: Camera-free Image to Panorama Generation with Diffusion Model

Supplementary Material



7. More Qualitative Results

In Figure above we provide a qualitative comparison between baseline MVDiffusion and our CamFreeDiff on the same hard example. Both MVDiffusion and CamFreeDiff are trained on the Matterport3D dataset, but CamFreeDiff is designed and trained to handle arbitrary camera parameters, while MVDiffusion is not. Note that we provide camera parameter estimation results to MVDiffusion for a fair comparison since it has not been trained to handle arbitrary camera views. For this hard example, the camera estimator doesn't provide perfect camera parameters to rectify the input image. Compared to the MVDiffusion, we still observe that CamFreeDiff generates more consistent and reasonable panoramic output.

8. Ablation on the Learning Objectives of Homography Estimators

We compared classification and regression as different types of homography estimators. Cross-entropy loss is used as the objective for the classifier, and mean squared error (MSE) loss is for the regression model. The classifier gives the best input view estimation results instead of the regression model, as shown in table Table 5.

Objectives	fov↓	phi(ϕ)↓	psi(ψ)↓
MSE (reg.)	10.6	2.5	2.4
CE (cls.)	**7.9**	**1.8**	**1.5**

Table 5. Ablation study of different types of homography estimation objectives. Experiments are based on HomographyNet and evaluated on randomly warped images from Matterport3D. Predictions are evaluated in terms of Mean Absolute Error.