

动机: the camera parameters can be jointly optimised as learnable parameters with NeRF training, through a photometric reconstruction

任务：

1) to estimate camera parameters for each image

2) to learn a volumetric scene representation through a photometric reconstruction loss, achieving high fidelity NVS results comparable to existing two-stage COLMAP-NeRF pipelines while removing the pre-processing step for camera parameters.

we propose to jointly optimise camera poses, intrinsics and a NeRF in an end-to-end manner on forward-facing scenes.

we conduct thorough analyses on our method and COLMAP-NeRF under various camera motions, showing that both systems can tolerate up to ±20◦ of rotation and ±20% translation perturbations, and that the joint optimisation is more favourable than COLMAP in translational perturbations, but less competitive in rotational perturbations.

Camera Intrinsics can be expressed as a focal length f and principle points cx and cy for a pinhole camera model. Without losing generality, we consider the centre of sensor as the camera principle points, i.e., cx ≈ W/2 and cy ≈ H/2, where H and W denote the height and the width of the image, and all input images are taken by the same camera. Thus, estimating the camera intrinsics only refers to finding the focal length f in this case.

总的来说，nerf—的核心是将位姿估计（李代数方法）融入到网络里，但仅适用于前向拍摄的数据集

能否优化环绕场景或者变化剧烈的场景