Deep Permutation Equivariant Structure from Motion

Dror Moran* Hodaya Koslowsky* Yoni Kasten
Haggai Maron Meirav Galun Ronen Basri

Abstract: Existing deep methods produce highly accurate 3D reconstructions in stereo and multiview stereo settings, i.e., when cameras are both internally and externally calibrated. Nevertheless, the challenge of simultaneous recovery of camera poses and 3D scene structure in multiview settings with deep networks is still outstanding.

Inspired by projective factorization for Structure from Motion (SFM) and by deep matrix completion techniques, we propose a neural network architecture that, given a set of point tracks in multiple images of a static scene, recovers both the camera parameters and a (sparse) scene structure by minimizing an unsupervised reprojection loss. Our network architecture is designed to respect the structure of the problem: the sought output is equivariant to permutations of both cameras and scene points. (Fig. 1)

Notably, our method does not require initialization of camera parameters or 3D point locations. We test our architecture in two setups: (1) single scene reconstruction (Fig. 2) and (2) learning from multiple scenes (Fig. 3). Our experiments, conducted on a variety of datasets in both internally calibrated and uncalibrated settings, indicate that our method accurately recovers pose and structure, on par with classical state of the art methods. Additionally, we show that a pre-trained network can be used to reconstruct novel scenes using inexpensive fine-tuning with no loss of accuracy.

Fig 1. Task Symmetries: Predicting a set of camera positions and 3D points from an input measurement tensor M is equivariant to reordering of the points and the cameras, represented by the pair of permutations $(\tau_{cams}, \tau_{pts})$. This is illustrated in the commutative diagram above.

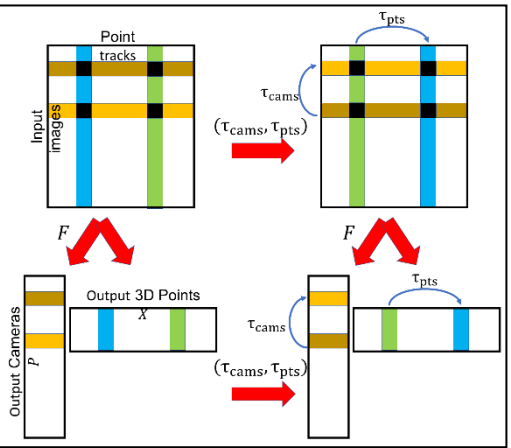
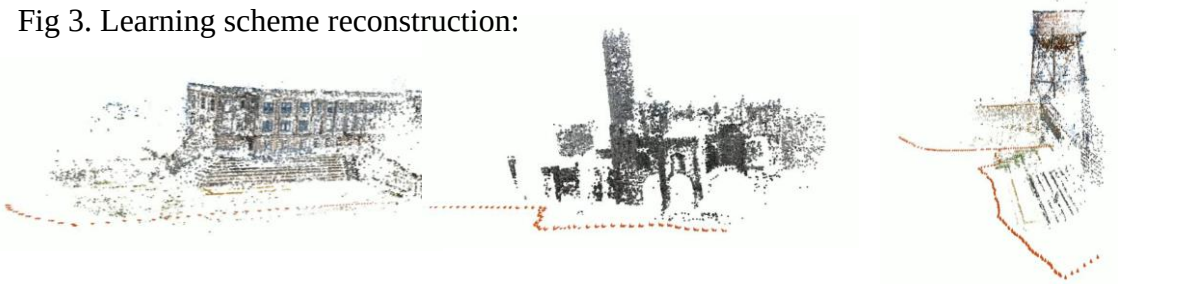


Fig 2. Single scene recovery:

85 images, 84792 points 196 images, 69341 241 images, 67107 points

Fig 3. Learning scheme reconstruction:



133 images, 23674 points 177 images, 30367 points 172 images, 14828 points