Data Driven Scene Understanding

Rahul Mitra

March 30, 2016

Indian Institute Of Technology Bombay

Creating Transformation and Illumination Invariant Descriptor

- Using *ConvNets* to learn a descriptor using pairs of similar and dissimilar image patches Figure 1. It should be invariant to transformations.
- For training purposes large number of scenes each containing more than 250 images are downloaded from the internet.

(a) **(b)**

Figure 1: Here using neighborhood of point correspondences in the two images (marked in Red), we form a similar pair of patches.

Data driven Estimation of 6DOF Pose

- Estimate 6DOF pose of objects from a single image using data driven techniques.

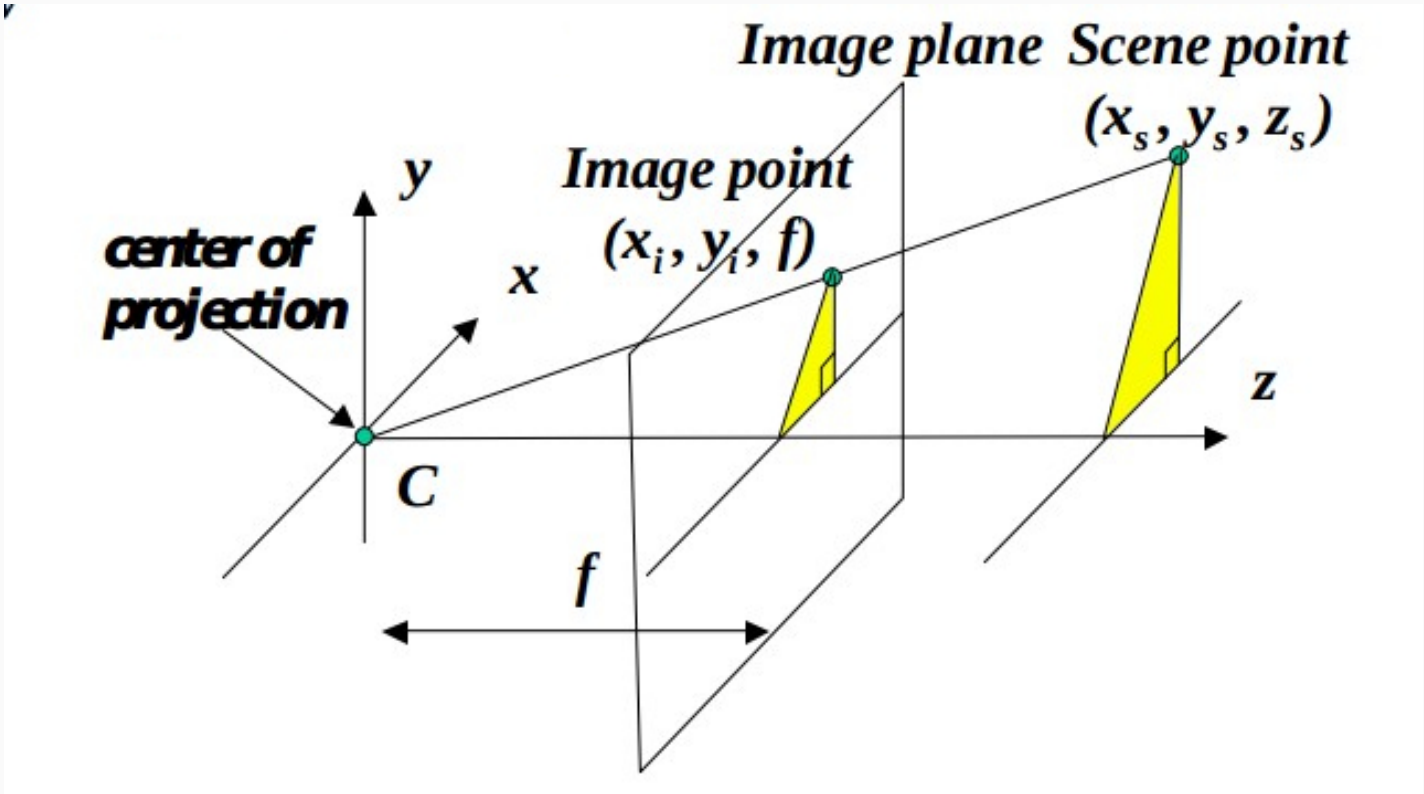

Figure 2: Model of a protective camera

Data Driven SLAM

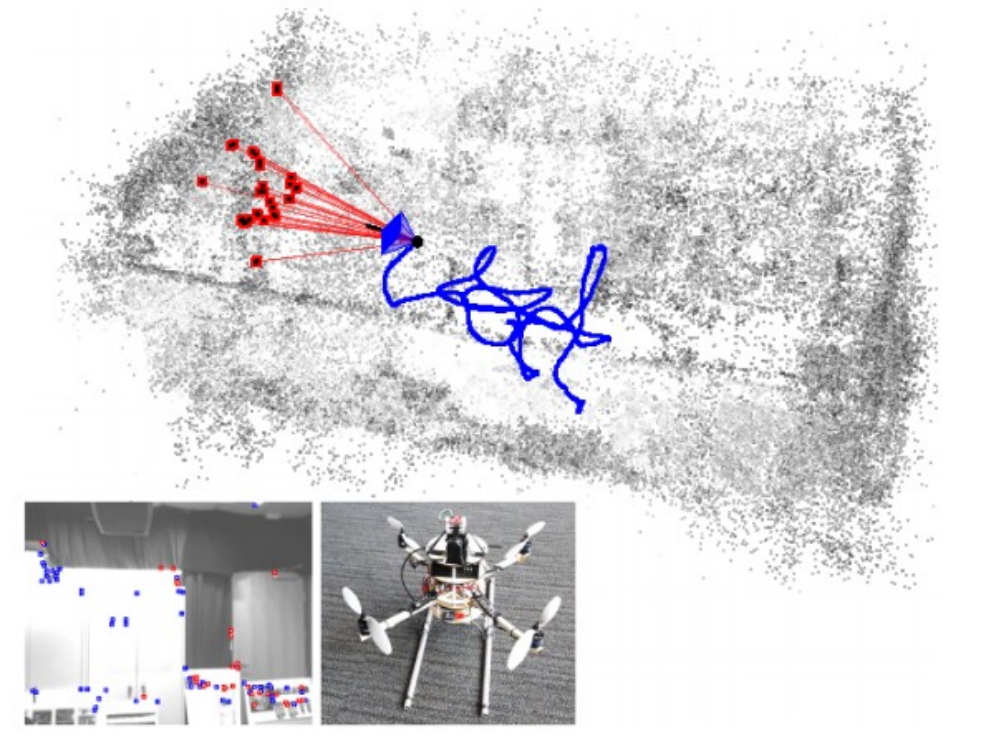

Figure 3: Blue line shows the trajectory predicted. Red points shows correspondences between 2D and 3D points.

References

- Multiple View Geometry in Computer Vision, R. Hartley and A. Zisserman, Cambridge University Press, 2000
- Real-time Image-based 6-DOF Localization in Large-Scale Environments, H. Lim, S.N. Sinha, M.F. Cohen, M. Uyttendaele, CVPR 2012, pg 1043–1050.