

Indoor localization is a challenging problem addressed extensively by both robotics and computer vision communities. Most existing approaches focus on using either cameras or laser scanners as the primary sensor for pose estimation. In laser scan matching based localization, finding scan point correspondences across scans is a challenging problem as individual scan points lack unique attributes. In camera based localization, one has to deal with images with little or no visual features as well as scale factor ambiguities to recover absolute distances. In this paper, we develop a multimodal approach for indoor localization by fusing a camera and a laser scanner in order to alleviate the drawbacks of each individual modality. Specifically, we use structure from motion to estimate the pose of a moving camera-laser rig which is subsequently used to compute scan correspondence estimates which are refined using a window based search method for scan point projections on the images. We have demonstrated our proposed system, consisting of a laser scanner and a camera, to result in a 0.4% loop closure error for a 60m loop around the interior corridor of a building.