A Joint Optimization Approach of LiDAR-Camera Fusion for Accurate Dense 3D Reconstructions

introduce

LiDAR and camera has complementary properties

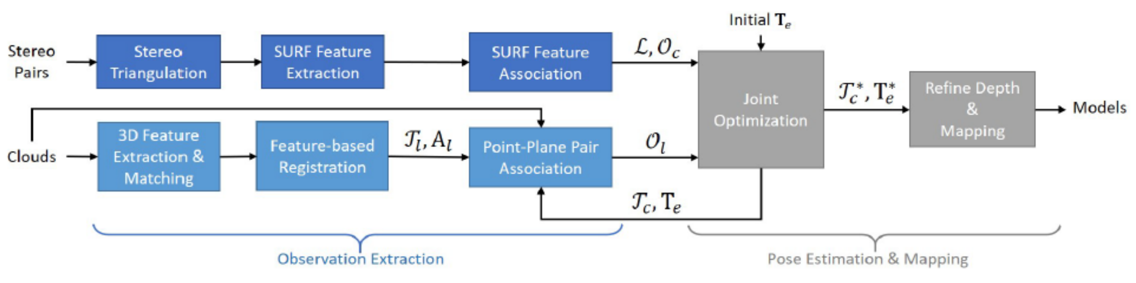
- camera: higher resolution and have colors
- LiDAR: provide more accurate range measurement, robust to low-texture conditions and have a wider Field Of View

But the fusion problem remains challenging

we choose to combine a rotating LiDAR with a wide-baseline, high-resolution stereo
system to increase the density of raw data

This is an offline method

Framework

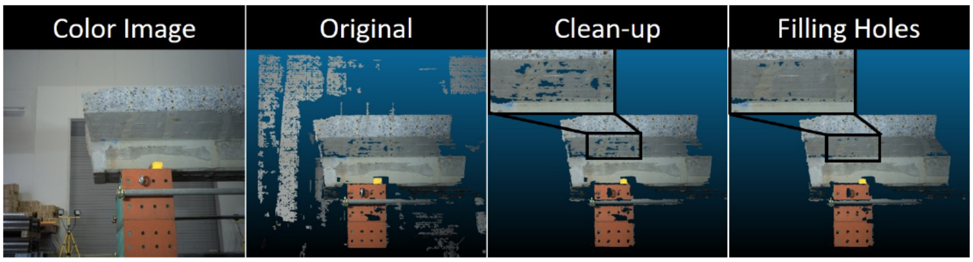


- Camera Observation Extraction
  - stereo triangulation to obtain a disparity image
  - SURF features are extracted from the left image
  - associated with depth value

    - an adjacency matrix Ac encoding the correlation of the images can be obtained
- LiDAR Observation Extraction
    - the Binary Shape Context (BSC) descriptor is used to match
    - another adjacency matrix Al matched cloud pairs is obtained and define the final pose graph
    - a set of points are sampled randomly from each point cloud as the key points
- Joint Optimization
    - the residual Ef and Ed encode landmark reprojection and depth error, while El denotes the point-to-plane distance error

$$f(\mathcal{T}, \mathcal{L}, \mathbf{T}_e) = \frac{1}{2}\sum_{\mathbf{o}_c} w_{\mathbf{o}_c}\left(E_f^2 + E_d^2\right) + \frac{1}{2}\sum_{\mathbf{o}_l} w_{\mathbf{o}_l} E_l^2$$

    - filter out incorrect observations
- Mapping
    - the stereo depth maps typically contain outliers and holes due to triangulation failure
        - the stereo depth to LiDAR depth with a significant difference will be removed
        - LiDAR depth is selectively used to fill holes, only the regions that are local smoothness



result