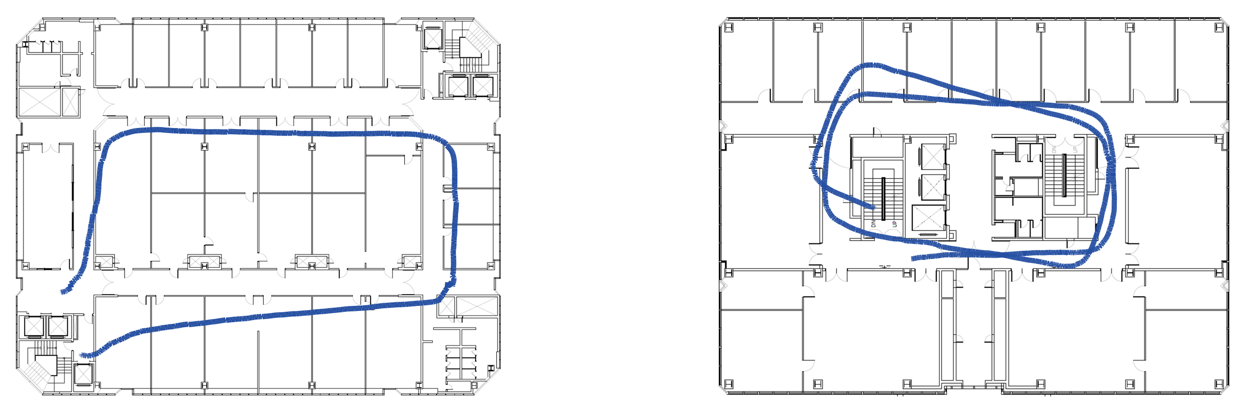
Visual Odometry with Drift-Free Rotation Estimation Using Indoor Scene Regularities

Pyojin Kim[1], Brian Coltin[2], H. Jin Kim[1]

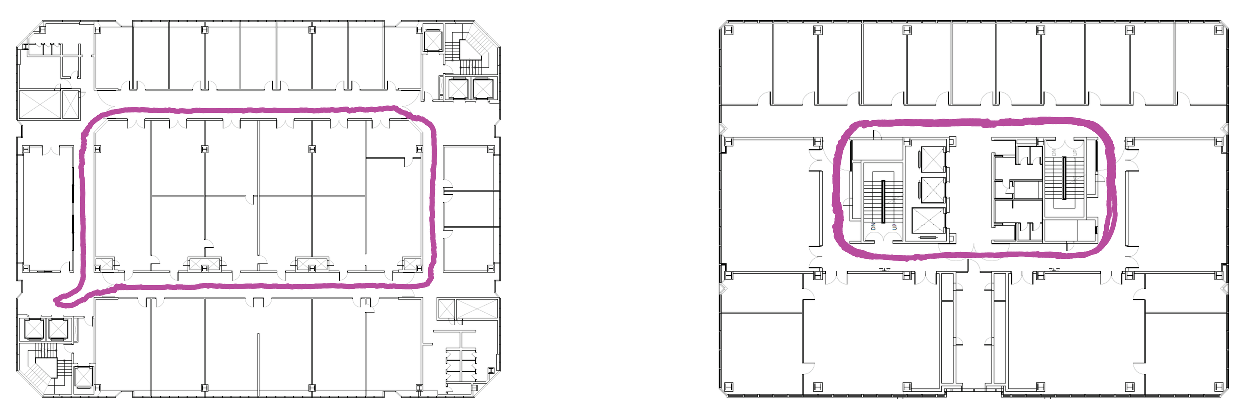
[1] Seoul National University, [2] SGT Inc., NASA Ames Research Center

Motivation

Drift in the rotation estimate is **a main source of positioning inaccuracy** in visual odometry (VO).

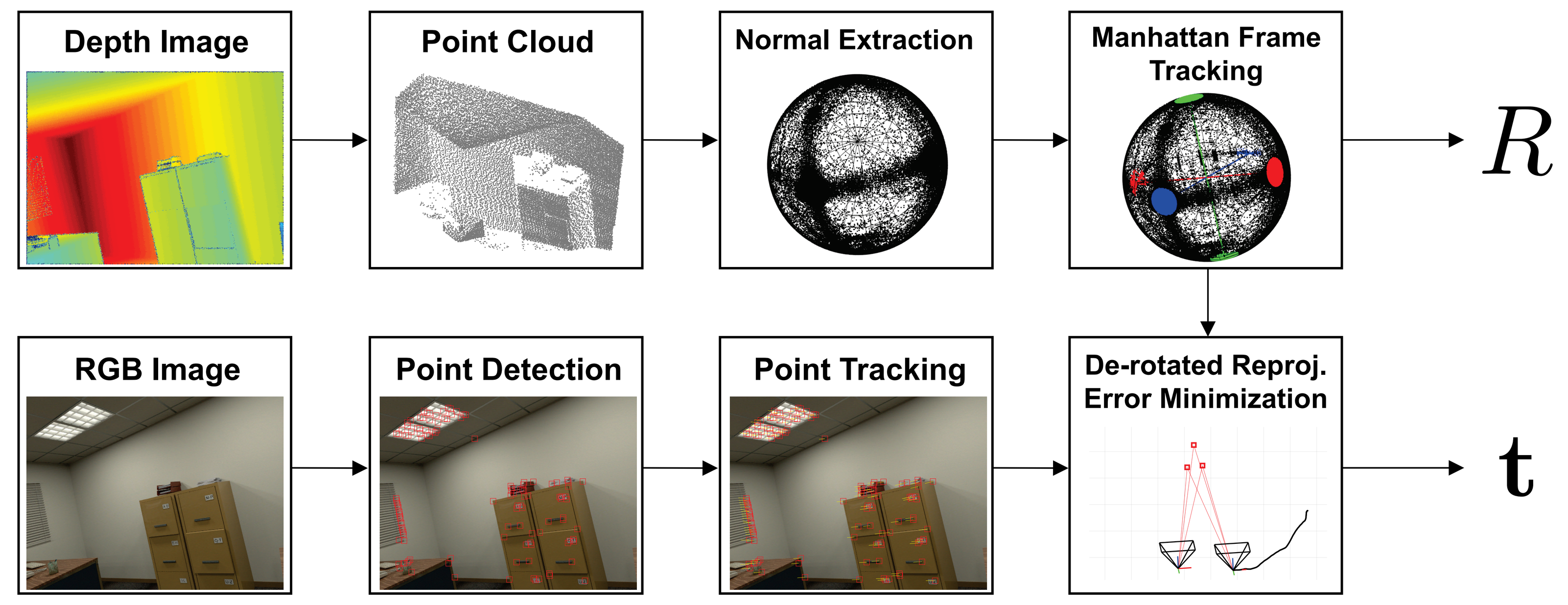


Solution: **Drift-Free Rotation from Structural Regularity**



Goal

The proposed VO separately estimates **the drift-free rotation** and translation motion by exploiting **orthogonal structures**.

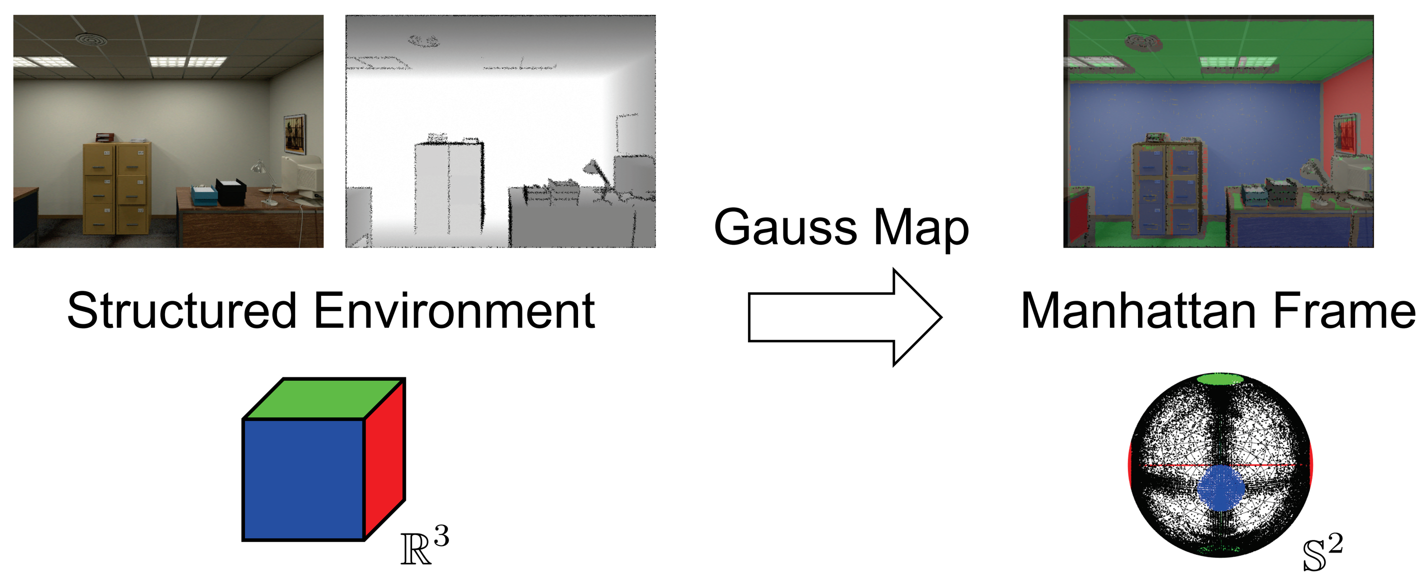


Overview of the proposed algorithm

Contributions

1. Integration of **drift-free rotation estimation** in VO
2. **De-rotated reprojection error** for pure translation
3. Evaluation on **large-scale man-made environments**

1. Drift-Free Rotation Estimation



- Extract **surface normal vectors** for each pixel by cross-product from a smoothed depth image.
- Find the dominant Manhattan frame with **a SO(3) manifold constrained mean-shift algorithm**.
- Keep **tracking the detected Manhattan frame** for drift-free rotation estimation of the RGB-D camera.

2. Translation Estimation

- Track **the Good Features to Track** feature points by using **KLT method**.

known rotational motion

$$\mathbf{X}_i^k = Z_i^k \bar{\mathbf{X}}_i^k = R\mathbf{X}_i^{k-1} + \mathbf{t}$$

$$r_{i_1}(\mathbf{t}) = R_1 - \bar{X}_i^k R_3 \mathbf{X}_i^{k-1} + \mathbf{t}_1 - \bar{X}_i^k \mathbf{t}_3 = 0$$

$$r_{i_2}(\mathbf{t}) = R_2 - \bar{Y}_i^k R_3 \mathbf{X}_i^{k-1} + \mathbf{t}_2 - \bar{Y}_i^k \mathbf{t}_3 = 0$$

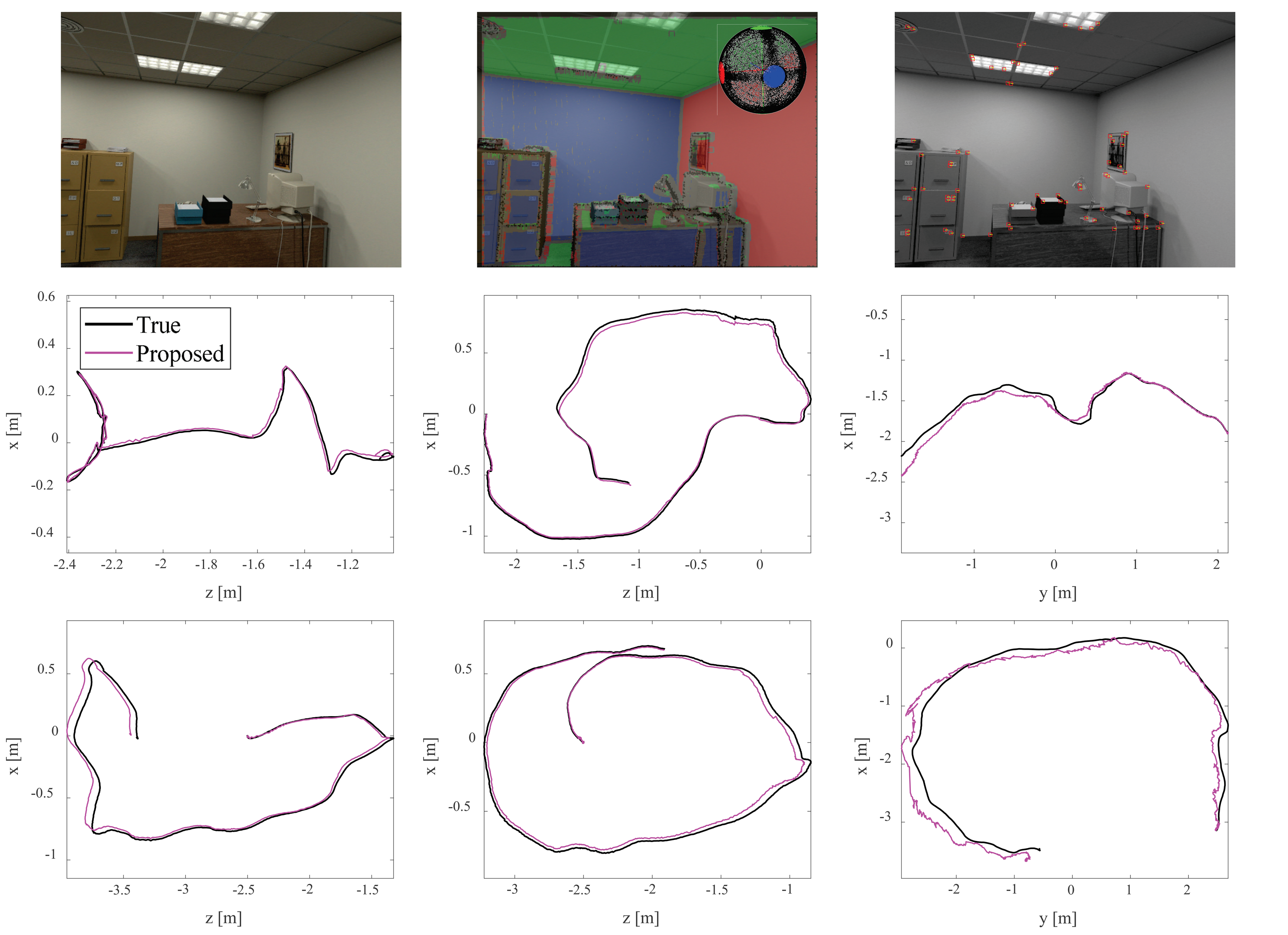
- Given the drift-free rotation, recover the translational motion **by minimizing de-rotated reprojection error**.

$$\mathbf{t}^* = \arg\min_{\mathbf{t}} \sum_{i=1}^{M} (r_{i_1}(\mathbf{t}))^2 + (r_{i_2}(\mathbf{t}))^2$$

- Employ the Levenberg-Marquardt (LM) algorithm to obtain **the optimal 3-DoF translational motion**.

3. Evaluation

ICL-NUIM & *TUM* RGB-D datasets (~10 m)



- Estimated and ground-truth trajectory overlap significantly.

Author-collected RGB-D datasets (~100 m)

- Starting and end points meet at the same place.