Object Level Depth Reconstruction for Category Level 6D Object Pose Estimation From Monocular RGB Image

Zhaoxin Fan [1], Zhenbo Song [2], Wenping Zhang[1], Hongyan Liu[3], Jun He[1], Xiaoyong Du[1]

[1] Renmin University of China, [2] Nanjing University of Science and Technology
[3] Tsinghua University

Arxiv Link: https://arxiv.org/pdf/2108.12790.pdf

Motivation

Catastrophic collapse caused by rotation problems needs to be solved.

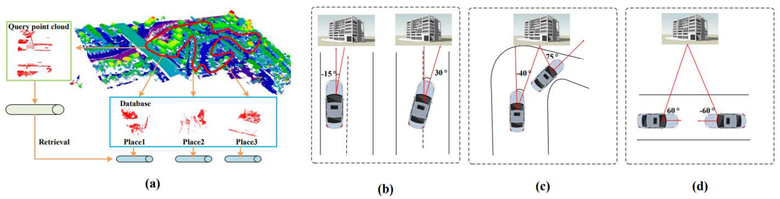

Figure 1. Task definition and motivation. (a) an illustration of point cloud-based large scale place recognition. (b) change lanes causes rotation problems. (c) turning a corner causes rotation problems. (d) driving from different direction causes rotation problems.

Contribution

1. We propose a novel model named RPR-Net, which, to the best of our knowledge, is the first strictly rotation-invariance dense network designed for point cloud-based large scale place recognition.
2. We propose an Attentive Rotation-Invariant Convolution operation, which learns rotation-invariant by mapping low-level rotation-invariant features
3. We achieve the state-of-the-art performance when dealing with rotation problems and achieve comparable results with most of existing methods on several original non-rotated benchmarks.

Ours: RIFs and Network Structure

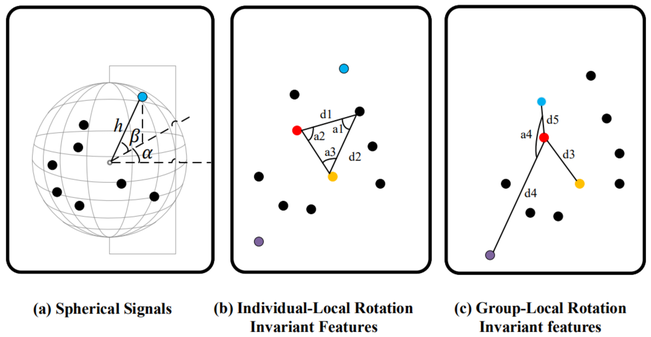

Figure 2. Illustration of three kinds of RIFs. For each points, we extract three kinds of RIFs: Spherical Signals, Individual-level Local Rotation-Invariant Features and Group-level Local Rotation-Invariant Features

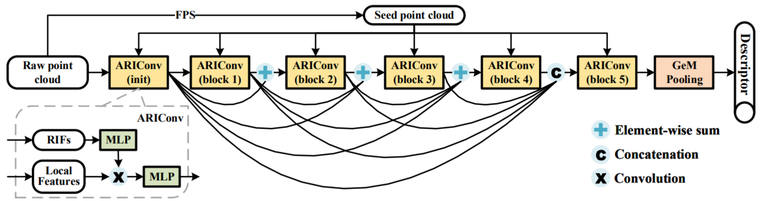

Figure 3. . Illustration of network architecture. 6 ARIConvs and one GeM Pooling layer are used to constitute the skeleton of the network. We adopt the densely connected idea to design a dense network architecture.

Results:

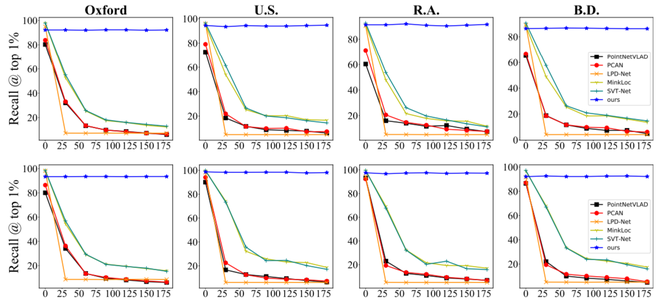

Figure 4. Comparison of our method with state-of-the-art place recognition methods at different rotation levels. The top row is the result at the baseline training setting; the bottom row is the result at the refined training setting. The values along the horizontal axis represents different maximal rotation degrees.

Conclusion

In this paper, we propose a novel rotation-aware large scale place recognition model named RPR-Net, which focus on learning rotation-invariant global point cloud descriptors to solve the catastrophic collapse caused by rotation. We propose a novel ARIConv, which is equipped with three kinds of RIFs and an attentive module to conduct rotation-invariant convolution operation. Taking ARIConv as a basic unit, a dense network architecture is constructed. Experimental comparison with state-of-the-art place recognition models and rotation invariant models has demonstrated the superiority of our method. Our research may be potentially applied to SLAM, robot navigation, autonomous driving, etc, to increase the robustness and the reliability of place recognition.