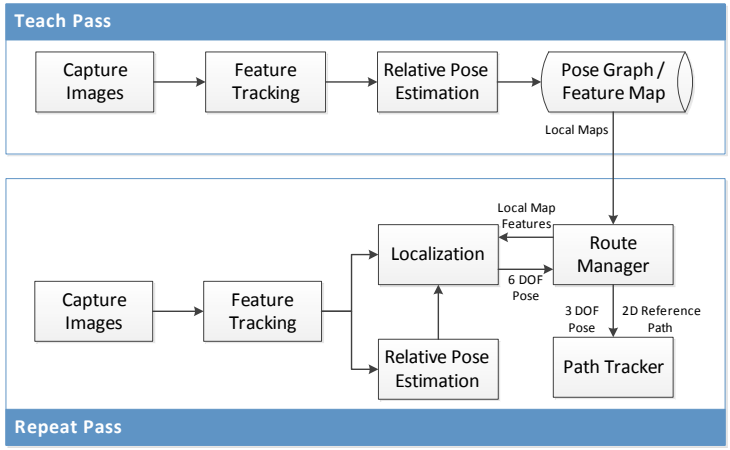
Monocular Vision for Long-range Visual Teach and Repeat in Unstructured Environments

Lee Clement, Jonathan Kelly, and Timothy D. Barfoot

Institute for Aerospace Studies
UNIVERSITY OF TORONTO

Visual Teach and Repeat allows mobile robots to autonomously retrace long routes in unstructured environments using only a camera.



Major Processing Blocks of the Visual Teach and Repeat System

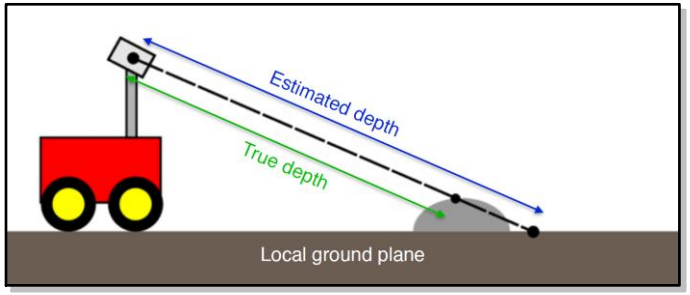
Teach Pass: The system tracks 3D image features to build a database of robot poses and corresponding keyframes while a human drives the robot along a path.

Repeat Pass: The system matches features to identify nearby keyframes and builds locally consistent maps using sliding-window bundle adjustment, then interleaves visual odometry and map-based localization to autonomously repeat the taught route.

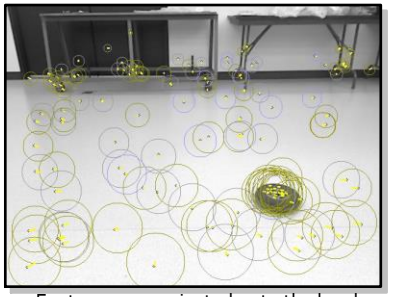
Visual Teach and Repeat relies on 3D data, but many robots are equipped with monocular cameras.

How well do we need to estimate 3D feature positions for Visual Teach and Repeat to operate?

Assuming the world is locally planar provides a simple method of rapidly generating estimates of 3D feature positions from known camera height and orientation relative to the robot.

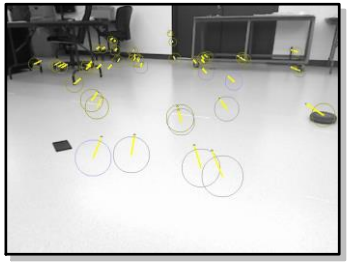


Local ground plane back-projection for feature depth estimation

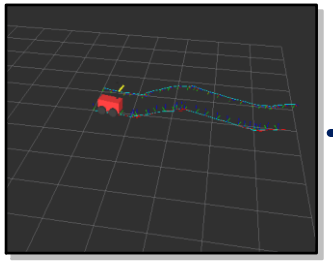


Features are projected onto the local ground plane regardless of their true height

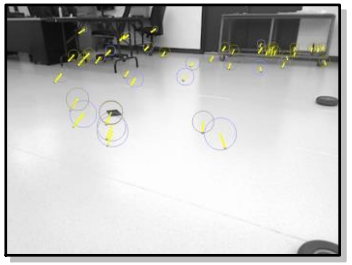
This assumption allows Visual Teach and Repeat to map and accurately retrace reasonably smooth routes, even when many features are far from the local ground plane.



Feature tracks on an indoor obstacle course during the teach pass



Visualization of the taught route



Matched feature tracks during the repeat pass

Summary

Treating the world as locally planar allows Visual Teach and Repeat to operate successfully with monocular vision in non-planar environments, provided the robot motion is smooth.

The question remains to be answered how smooth the robot motion must be for this method to function reliably.

References

Furgale, P., & Barfoot, T. (2010). Visual Teach and Repeat for Long-Range Rover Autonomy. Journal of Field Robotics, Special Issue on Visual Mapping and Nagivation Outdoors, 27(5), 534–560.

McManus, C., Furgale, P., Stenning, B., & Barfoot, T.D. (2012). Lighting-invariant Visual Teach and Repeat Using Appearance-based Lidar. Journal of Field Robotics, 30(2), 254-287.