Articulation Estimation Using Depth Sensing

Suren Kumar
Mechanical and Aerospace Engineering
State University of New York at Buffalo
Buffalo, NY, USA
Email: surenkum@buffalo.edu

Vikas Dhiman
Electrical Engineering
University of Michigan
Ann Arbor, MI, USA
Email: dhiman@umich.edu

***Abstract*—**

- **Detect distinctly moving clustered points/voxels/objects in a scene.**
  - **Use Kinect fusion to create a static map.**
  - **Use some kind of noise threshold to detect object movement independent of camera movement.**
  - **Trigger algorithm (may be use RANSAC ?? etc.) that will segment out the object that just moved. The object should be spatially clustered and should be explained by the same rigid 3D motion.**
  - **Maintain a pairwise relative localization graph of the scene.**
- **Semantic reasoning in map update of these objects and their localization. Reason about Physical support and articulated linkage.**
- **Build a 3D reconstruction of these objects.**
- **Find similar unmapped static objects in the scene. May be use Jeff's detection and segmentation code.**
- **Try algorithm for long term mapping (a week) by using auto charging turtlebots in a living room and compare with existing algorithms.**

I. Introduction

Imagine a robot moving in a typical living room environment which encounters indoor objects such as doors, drawers and chairs etc. We posit that in order for the robot to understand, map or interact with such objects, the robot needs to be able to understand the articulation. Pyschophysical experiments on human motion understanding have demonstrated that human first distinguish between competing motion models (translation, rotation and expansion) and then estimate the motion conditioned on motion model [1].

II. Articulation Classification

We consider the problem of motion model identification from point correspondences of motion. In the current work, we consider revolute, prismatic and general motion. Consider motion of two points x_0, x_1 (represented in an inertial frame) on a rigid body at time t_0 and at some subsequent times t_1, t_2. The most general form of rigid body motion of a point can be represented using a rotation matrix R and an associated translation vector T

$$x_0^{t_1} = R_{t_0}^{t_1} x_0^{t_0} + T_{t_0}^{t_1} \tag{1}$$

where the superscript on the point denotes the time. For points lying on a prismatic joint such as a drawer, rotation w.r.t inertial frame remains the same (R=I), resulting in $x_1^{t_1} - x_0^{t_1} = x_1^{t_0} - x_0^{t_0}$. This is essentially saying that the vector joining two points on a prismatic joint remains the same before and after the motion.

For further distinction between revolute and general motion, we need information from more than one time step. For points lying on a body undergoing revolute motion such as a door, the points have same translation vector over time. Hence estimating the translation vector from two time steps $T_{t_0}^{t_1} = T_{t_1}^{t_2}$ is a sufficient condition to classify a joint as revolute joint. For general motion such as a book that can be rotated and translated anywhere in the space both the rotation and translation matrix will be different.

III. Scene Understanding

Analysis by method such as ours is essential in order to decompose the scene into types of motion that a robot can influence on the scene. For example: Understanding the way a drawer can be opened, fridge door can be opened, what can be moved around in the scene

References

[1] Shuang Wu, Hongjing Lu, and Alan L Yuille. Model selection and velocity estimation using novel priors for motion patterns. In D. Koller, D. Schuurmans, Y. Bengio, and L. Bottou, editors, *Advances in Neural Information Processing Systems 21*, pages 1793–1800. Curran Associates, Inc., 2009.