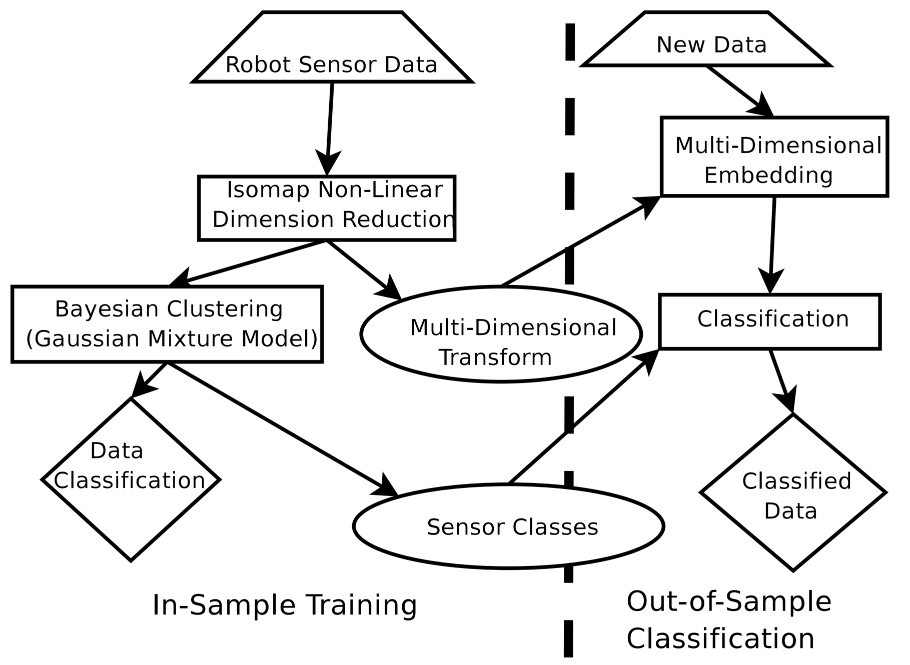
Extensible Data-driven Classification of Robot Sensor Data

Daniel H Grollman and Odest Chadwicke Jenkins and Frank Wood
Brown University
{dang,cjenkins,fwood}@cs.brown.edu

Objective

Physical space classes (Region types) are typically selected by humans and taught to robots. We demonstrate a method where a robot develops its own classes based upon the characteristics of its sensor data. In addition, this approach is extensible, allowing for the fast application of learned classes to new sensor data.



Method

Treating sensor readings as vectors in a high-dimensional space, we

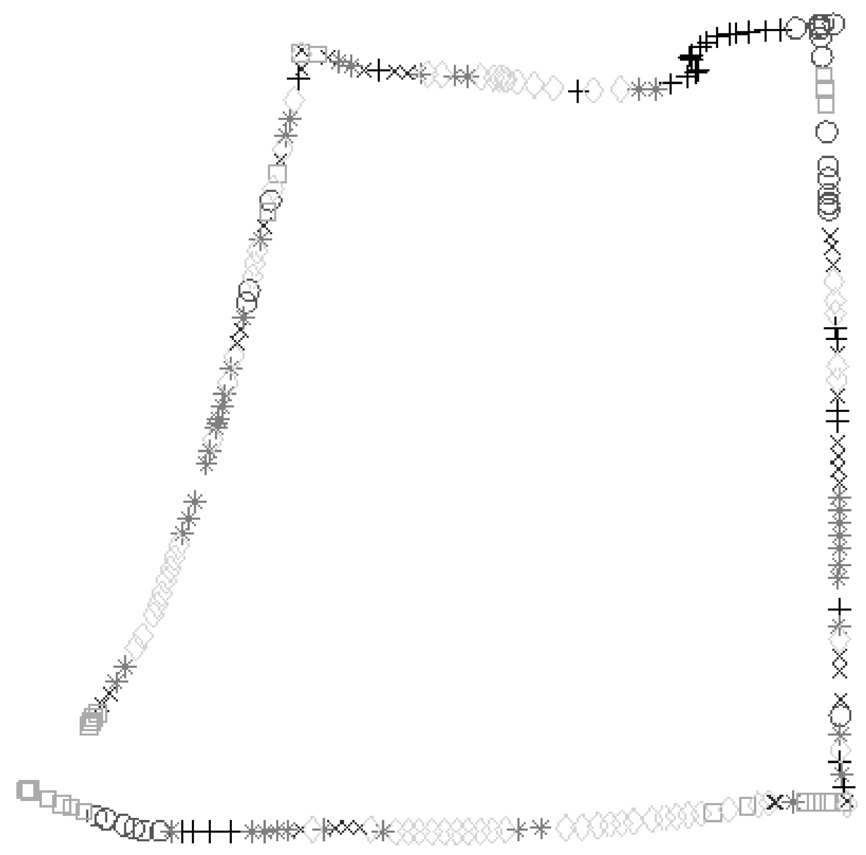
1. Approximate the underlying manifold with Isomap nonlinear dimension reduction.
2. Cluster in the reduced space with a Gaussian Mixture Model.
3. Use model identification techniques (BIC and holdout) to discover the number of classes in the data.

When presented with new sensor readings, the physical space class from which they were taken is quickly and easily determined:

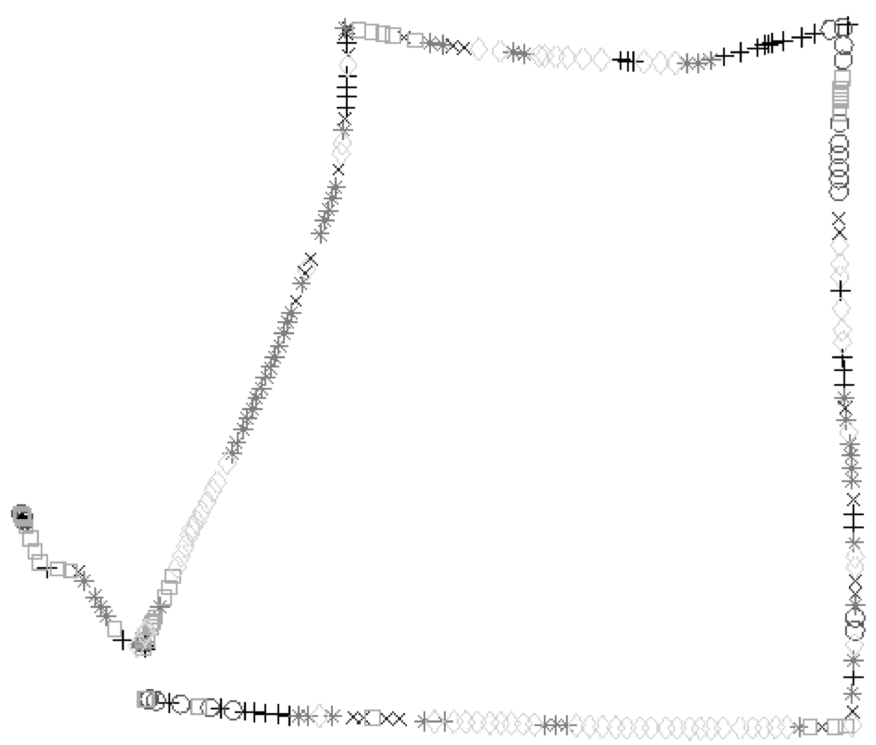
1. Place the new point on the previously discovered manifold.
2. Classify it according to the GMM.

Crunch

Data from a small, inverted pendulum robot was used in a streaming-data scenario to test the speed and reliability of the system. Sensor readings from an exploratory first trip were used to discover physical space classes, and readings from a second trip were classified in real time.



First Trip

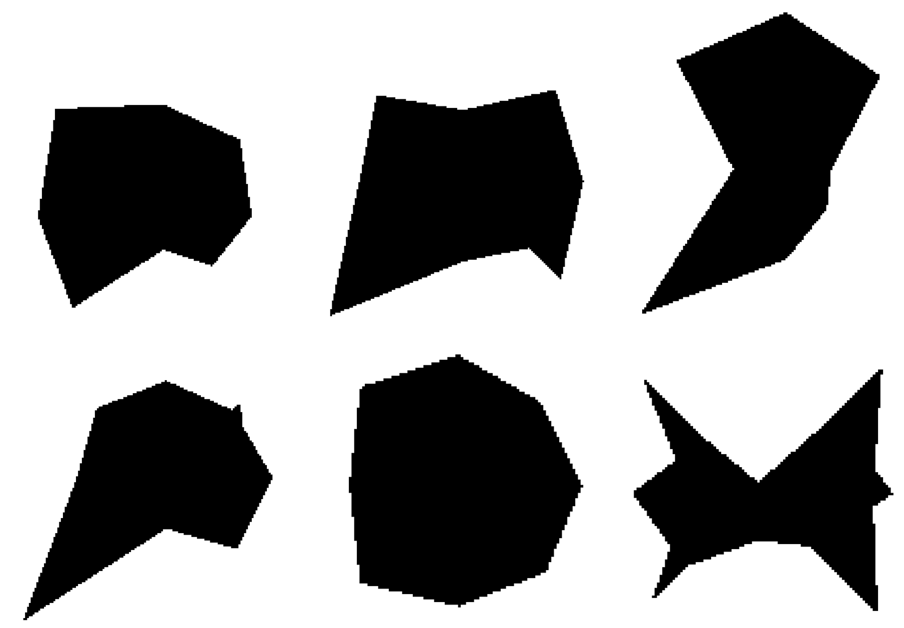


Second Trip

Each sensor reading is coded according to the physical space class it belongs to, and overlaid on the robot's odometry. The learned classes from the first trip were used to classify the data from the second trip in real time. The classifications are quite similar across the two trips, showing that the discovered physical space classes are reusable.

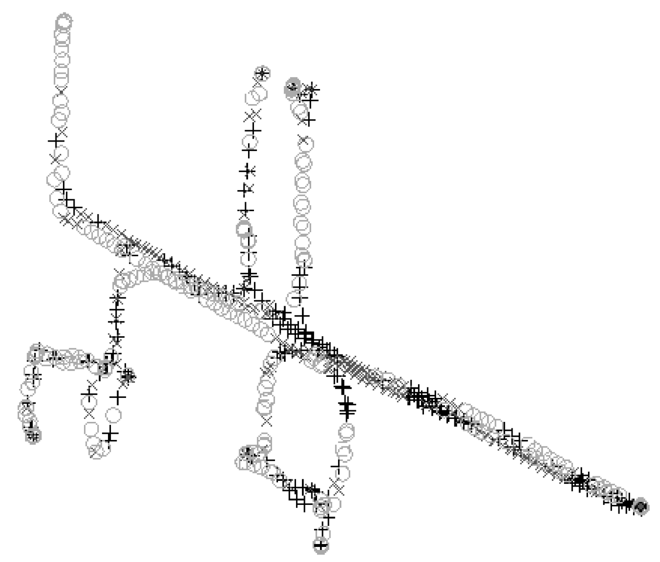
Derived classes

Our system discovered six physical space classes in the area visited by Crunch. Under the standard ray model of Crunch's sensors (sonar and IR), several of them are quite similar, but they are actually distinct enough to be detected.

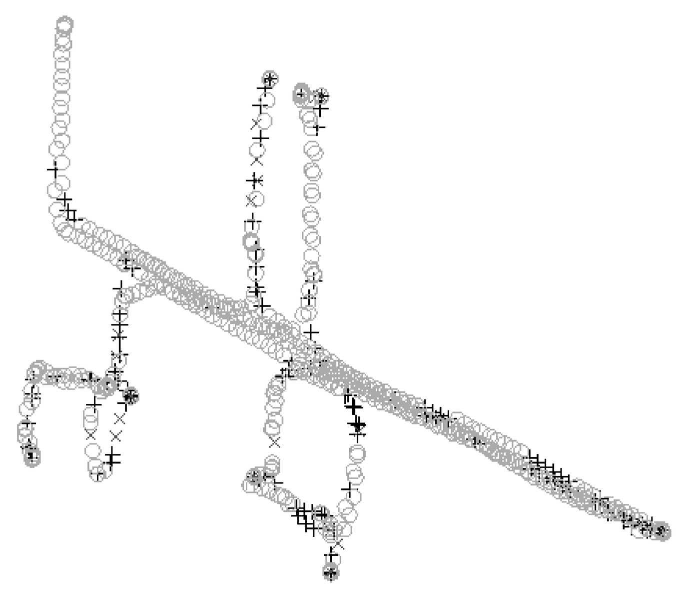


Fr079

Data from the Robotics Data Set Repository[a] was used in a large-dataset scenario to test the system's extensibility. In addition to a full analysis of the data A second analysis was done, using the first 500 datapoints as 'landmarks' and classifying the remaining 2500 datapoints based upon what was learned.



Full Processing



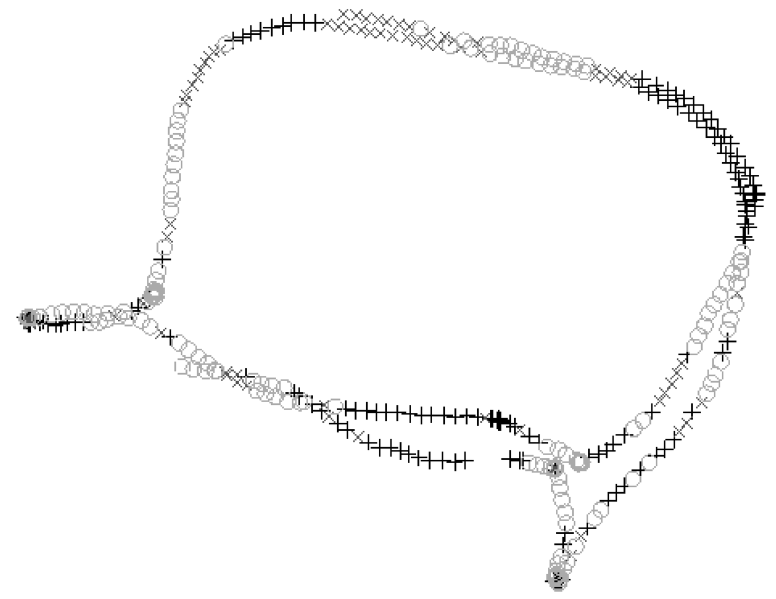
Landmarked

[a] http://radish.sourceforge.net

The results are similar, but there are some striking differences, most likely due to an imbalance in the representations of the various classes in the landmark data. Experiments with random landmark points yielded similar results.

Loops

Part of the Fr079 data set contained loops. As the robot revisited a location, it was detected to belong to the same class each time, showing that a location is robustly classifiable.



Loop Closeup

Derived classes

Laser range finders are well modeled by the standard ray model, and thus the derived classes from the Fr079 dataset are easily visualized and seen to be distinct.

Conclusions and Application

By deriving physical space classes directly from raw sensor data, biases due to particular models of sensor operation and noise are eliminated. The resulting learned classes can be reused to classify new sensor readings as they are aquired.
Robust space classification is an important first step in many mobile robot applications, such as localization and mapping. We believe that our technique can serve as a solid grounding for a topological mapping system.