(*L-3 / Space Grant Stipend*)
Object Recognition in Point Cloud Data using AI Aspects

Brian Fehrman

October 3, 2012

Executive Summary

Computer vision has long been performed in a flat, two-dimensional world. The real world, however, is not two-dimensional. Projecting the world onto a plane has inherent consequences such as ambiguities when trying to recognize objects within a scene. For instance, a picture of an object could be mistaken for the real life item [1]. To overcome the inaccuracies of flattening the world there has been a large effort towards algorithms using three and four dimensional point cloud data. Some of the fuel for the interest has been the recent availability of low cost natural interface devices such as the Microsoft Kinect sensor which can scan the environment and provide a cloud point representation of what it sees. These devices can easily be fitted to robots and ultimately used for tasks such as object recognition within a scene. Here the goal is to improve upon existing object recognition algorithms by investigating artificial intelligence techniques such as evolving a recurrent neural network that can be used to identify items that are encountered.

Current State

One of the most prevalent initiatives in point cloud data algorithms is the Point Cloud Library project. The Point Cloud Library, or *PCL*, is a large scale, stand alone, and open source project for both 2D and 3D point cloud processing. The PCL contains many advanced image processing functions such as 3D point cloud stitching and object recognition. Devices using the OpenNI 3D (e.g., Microsoft Kinect) interface can send their data to a program for processing via the PCL [2].

Aldoma et al researched fast 3D feature based object recognition and pose estimation. Their aim was to take a set of CAD models that represented how we perceive an item in the world and then accurately identify each one of those objects in a real world scene using a depth sensor. The algorithm was an extension of the Viewpoint Feature Histogram (*VFH*) and is more geared towards clustered environments and, as such, is dubbed the Clustered Viewpoint Feature Histogram (*CVFH*). They found that with using a Kinect sensor on a set of 44 objects that their algorithm was better able to recognize objects in the presence of partial occlusion and noise than the original VFH routine [3]

More recently, in collaboration with the PCL project, Aldoma and others have even further extended the CVFH algorithm such that they can repeatably place a reference frame on 3D objects in a scene. This reference frame is matched to a model reference frame in order to more easily determine an item's pose and to identify the item with greater certainty. They found that a substantial improvement was made over the CVFH method alone in terms of accuracy and computational performance. They plan to add CUDA support for future iterations of their algorithms [4].

Specific Goals

The OUR-CVFH algorithm previously described is essentially constrained to a static set of data on which it was trained. This would limit a robot in terms of its ability to adapt and learn in environments where information is encountered that has not previously been seen. The goal here is to investigate the use of Recurrent Neural Networks to identify objects within a scene and to also update their knowledge by providing some sort of a reward or fitness function to sway the decisions to be more accurate. The outcome would be improved robustness over the existing state of the art in terms of object detection within a cluttered, noisy, unpredictable environment.

Research Activities

The first and foremost activity will be an in depth literature review of the past and present research on point cloud data. This will include not only object recognition algorithms, but also more low level routines such as edge detection, segmentation, and other topics that could be beneficial. Research will be conducted on recurrent neural networks and methods for training them such as evolutionary techniques. Since intelligence and biological concepts underlie all of this it will also be worth while to take a closer look at how humans recognize objects and perceive their environment as this sort of view can often give great insight. The algorithms developed will be tested using simulations and a natural interface sensor and performance will be compared to existing approaches.

Metrics

The following lists the specific metrics to define the success of the research:

- Develop recurrent neural network for object recognition in point cloud data
- Show that the neural network will can learn based upon new information
- See that the newly developed solution performs better in terms of accuracy and speed than the current algorithms

Timeline

- Literature review

  Present - February 2013
- Combine information from research into an algorithm to robustly detect objects in a scene and learn new information

  February 2013 - June 2013
- Code, debug, and test the algorithm

  June 2013 - November 2013
- Revise algorithm and test against other solutions

  November 2013 - March 2014
- Finalize research

  March 2014 - May 2014

SDSGC and this Research

Awarding a Space Grant stipend for this research will directly address, among other goals, the SDSGC Objective B.2.2 and Strategy B.2.2.3 by supporting multi-disciplinary graduate research that is aligned with NASA's mission since this technology could be very valuable to autonomous robots in space. This stipend will assist in the research by allowing me to better focus on the work at hand without external financial concerns. Good results from this could jump start my career into the STEM disciplines which is something that I have always sought after. My love of what I do makes me a natural fit for the field which could lead to great accomplishments such as owning a successful business that can continue to provide funds for exciting and useful research.

References

[1] Radu Rusu and Michaul Dixon, editors. *Point Cloud Library on CUDA*. GPU Technology Conference, 2012.

[2] PCL. Pcl introduction. www.pointclouds.org. Online: Accessed 10-1-2012.

[3] A. Aldoma, M. Vincze, N. Blodow, D. Gossow, S. Gedikli, R.B. Rusu, and G. Bradski, editors. *CAD-model recognition and 6DOF pose estimation using 3D cues*. IEEE Computer Vision Workshop, 2011.

[4] Aitor Aldoma, Federico Tombari, Radu Bogdan Rusu, and Markus Vincze. Our-cvfh - orientied, unique, repeatable clustered viewpoint feature histogram for object recognition and 6dof pose estimation. *Lecture Notes in Computer Science*, 7476:113–122, 2012.