Topic – Depth map generation from RGB image sequences

Advisor – Brian Scassellati, along with grad students Alessandro Roncone and Olivier Mangin

Background –

Depth map generation is a well-researched field, as depth information is mission-critical for robots interacting with their environments. However, many methods of depth map generation make use of cameras designed specifically to extract depth information (often referred to as RGB+Depth cameras, or RGB+D), such as the Microsoft Kinect or the Leap Motion controller. Due to the low cost of medium resolution RGB cameras and the abundance of processing power, using RGB-only video streams for depth map generation would pose a cost-effective and flexible method of gaining valuable data about a robot’s environment.

The problem of stereo-vision depth estimation is a well-researched one, using correspondence between pixels and a known distance between the cameras to infer distance data between a chosen point in three-dimensional space and a selected image ROI. Optical flow techniques, based on assumptions regarding relative motion between the environment and the camera, have also been developed on RGB+D cameras.

Goals –

This project will attempt to create a module that feeds on a live stream of RGB-only images and camera position information and outputs a point cloud, refreshed at a high frequency. The module will be tested on the Yale Social Robotics Lab’s Baxter robot from Rethink Robotics. Baxter’s jointed arms can provide both position data of the robot’s extremities and RGB video streams via the cameras at the end of the arms. The high refresh rate of the image is meant to mitigate the complications caused by moving objects in the images.

Code will likely be written in Python, making use of the OpenCV library.

Deliverables –

A depth map generation module, deployable on Baxter
A set of test run data featuring object configurations of varying size
A point cloud visualization (?)
A live demo (?)