The Perception Project

Monday, April 8, 2019 7:54 PM

Project Summary
In this project we must locate and identify various objects from RGBD camera data. These objects are in a cluttered environment and have various geometries.

CAMERA SET UP

- We'll learn about and use RGBD to create a 3D point-cloud i.e. a cloud of points.

FILTERING/BASIC SEGMENTING

- We'll do some camera calibrating, filtering out the noise, and segment i.e. breaking up the point-cloud data in objects using a RANSAC algorithm.

ADVANCED SEGMENTING

- We'll make the segmentation more advanced using clustering

OBJECT RECOGNITION

- And finally perform object recognition using prior-knowledge/feature intuition, and a trained classifier algorithm (support vector machine)

PRODUCT

- The end result will be the creating of messages with the locations of specific objects that will be sent to a .yaml file for the pick and place task.

Sensor Basics and Overview

Sensors can be put into three categories: passive, active, and hybrid.
Passive sensors only measure energy in the environment
Active sensors send out some energy and measure the way it interacts with the environment
Hybrid sensors mix active and passive technology

Active Sensors

- Lidar
  - super good,
  - super expensive and bulky
- T.O.F. - pulse runtime & phase shift continuous wave
  - Fast/real-time,
  - small,
  - has issues with interferences and detecting glass
- Ultrasonic
  - Accurate and
  - can detect glass but is
  - low-res, short range, and
  - has issues with lots of physical things

Passive Sensors

- Monocular
  - Cheap and
  - rich in data but
  - lacking in range and is
  - computationally heavy and
  - light dependent
- Stereo
  - Same pros and cons as monocular but
  - has color

Hybrid

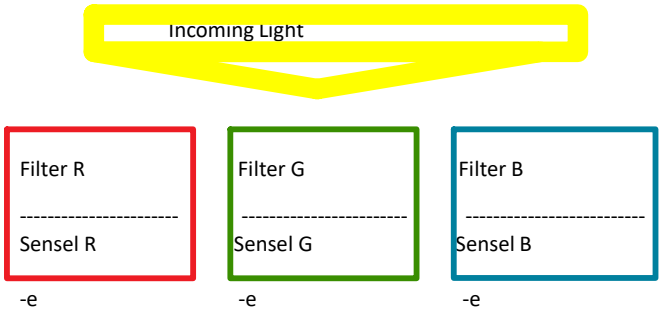
- RGBD
  - IR sensor + RGB sensor + projector
  - Cheap,
  - easy to use,
  - low-res,
  - poor range

RGBD

Incoming light passes through filters which isolate colors, then

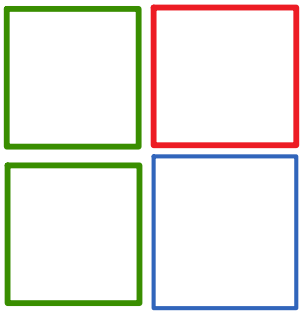
RGBD

RGB



Incoming light passes through filters which isolate colors, then Sensels tuned to particular wavelengths are stimulated and electrons build up on the measuring end which is read as a voltage Proportional to the incoming signal intensity.

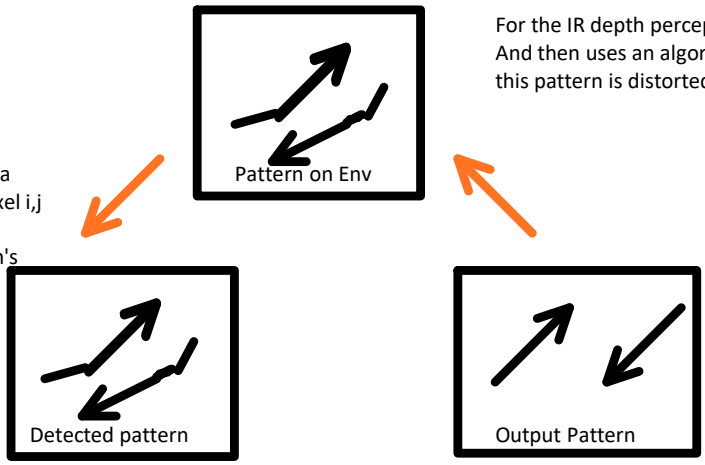
Bayer imaging sensor. This is a standard sensor type.



It's gridded and has 2x more green pixels than red or blue, this supposedly mimics human visual perception. The intensity data from the pixels is averaged In a process called **demosaicing**. I'm not sure why this is efficient but perhaps it's because of simplicity or ease of manufacturing.

The data we get back is point cloud data
That looks like [i,j][R,G,B,D] for each pixel i,j

We can easily process this using python's
PCL - point cloud library



For the IR depth perception part the camera sends out a pattern And then uses an algorithm to determine distances based of how this pattern is distorted with comparison to the original image