Combination of RGB Camera and Depth Sensor

Daiming Yang, Jingwei Feng, Jing Liu, Zhifan Wang, Zhaohan Lu, Yin Jin, Jianan Kang, Jiahui Ren, Lifu Gao

Beijing Haidian Foreign Language Shi Yan School (14-0349)

Combination of RGB Camera and Depth Sensor

Abstract

A depth sensor is capable of detecting distances between objects and robots. It may be enhanced with an addition of a regular RGB camera. Binding them allows robots to see a clearer world and move more efficiently. However, a problem of pairing coordinates of both visions soon arises, because the two devices are viewing from different angles. After experiments, common equations for pairing visions of simple assemblies are established.

1.0 Introduction

This year, Botball brings a new electronic sensor, depth sensor, to the competition. This kind of sensor is able to detect distances between objects and the sensor. However, it cannot identify which objects it is observing, and its range from 0.8 to 5 meters is limited. Despite disadvantages discussed before, it is still powerful and could be more powerful bound with a regular RGB camera. A camera can capture colorful pictures, and, after analyzing pictures, a controller may recognize which direction its target is. Robots with both a depth sensor and a camera are therefore capable of determining exact positions of their targets.

2.0 Mechanical Combination

The assembly should first be achieved mechanically. A depth sensor, however, does not contain any camera itself; a camera therefore has to be attached to it. The following photo displays how this is done.

The camera, as shown above, is specifically fixed on the middle of the depth sensor. At the beginning, their main angles of visions are particularly set parallel. This is because pixels of both visions are easier to pair in this kind of set (discussed in the next section).

3.0 Pairing Corresponding Pixels of Both Visions

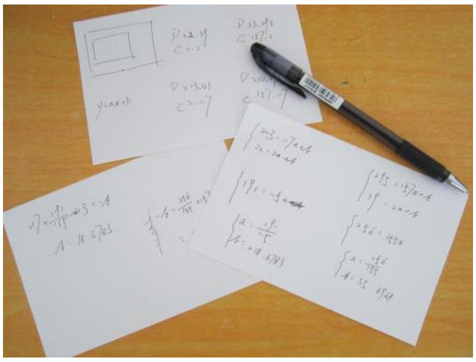
3.1 Abstract

There are still remaining problems with binding. Even though it is not hard to achieve mechanical combination, a problem of pairing data from both kits arises soon after. Since a camera has to be attached to a depth sensor, they are not looking at items from the same angle, despite of parallel main visionary lines; hence, a comparing data table is built to transform separated information into pairs. Thus, equations for a general transformation are established.

3.2 Methods

At first, utilizing functions of “get depth world point” was considered and soon abandoned, because it turns out to be extremely inefficient. Despite accuracy the method assures, links have to analyze each pixel in its updated image, calculate distances and angles, and then transform to corresponding pixels in the camera’s vision. This process can take as long as a minute to finish, and opposite way happens to be a disaster, because images captured by a camera are obscure and can never find the right paired pixel in an image of a depth sensor, except taking a great long time.

Another method is therefore thought of. Instead of using “get depth world point”, simply use “get depth value” turns out to be efficient. Admittedly, view angles of depth sensors are different from those of cameras, which means images from different distances will contain different pairs of pixels. However, differences are small enough to neglect. As a result, users may do an experiment determining a table of corresponding pixels from a certain distance, like the way the picture shows below, and they are able to determine how far their targets are, given where targets are in the camera’s view.



3.3 Experiment Data

Based on experiments, the following table is recorded (simplified), noticing unshared areas of two visions.

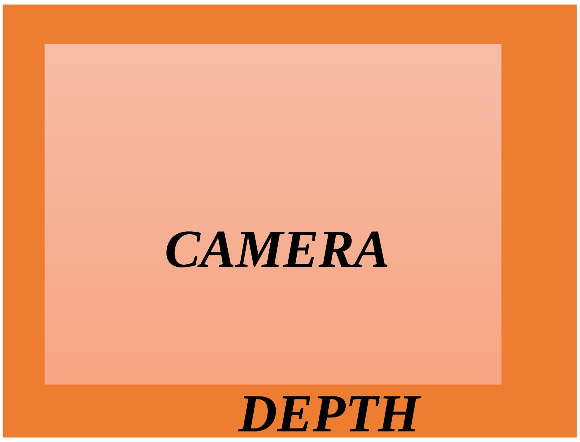
5, 5	5, 39	5, 155	5, 293	5, 314
--	--	--	--	--
22, 5	22, 39	23, 155	23, 293	23, 314
--	2, 2	81, 2	157, 2	--
100, 5	100, 40	101, 156	101, 294	101, 314
--	2, 61	81, 61	157, 61	--
213, 5	213, 41	214, 157	214, 295	214, 314
--	2, 117	81, 117	157, 117	--
234, 5	234, 41	234, 157	234, 295	234, 314
--	--	--	--	--

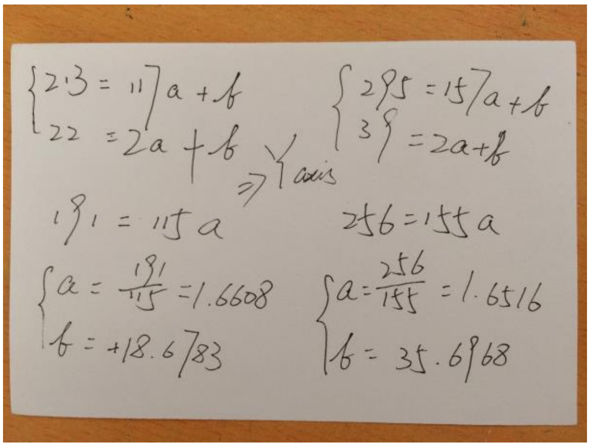
Depth	Camera

For actual experiments, finding at least 25 pairs of pixels is highly suggested, which makes sure successful transforming equations.

3.4 Derived Common Equations

Following is a simplified display showing how two visions overlap each other.





It is easy to discover linear relations between corresponding x-y coordinates. However, to make relationships readable by robots, equations are required, and pairing coordinates are taken into them. Then, common transforming equations from a camera to a depth sensor are generated, and they are specific for the assembly shown before.

$$\begin{cases} Row = 1.6608 * Y + 18.6783 & (2 <= Y <= 117) \\ Col = 1.6516 * X + 35.6968 & (2 <= X <= 157) \end{cases}$$

This is just simply a pair of linear functions. It is easy to interpret for boundaries of a vision of a camera is rectangular, while the outmost boundaries are eliminated because they are unstable and usually return nonsense.

4.0 Application

An assembly of a camera and a depth sensor is capable of providing accurate information on positions of targets. This allows robots operating more effectively. Rather than modifying their positions again and again, robots are able to find exactly not only where their objectives are, but where they are supposed to be as well, so they can move right to the positions and act as we expect them to. It can be more powerful if the minimum distance a depth sensor decreases, while now it is 0.5 meters.

5.0 Conclusion

Combining a regular RGB camera with a depth sensor can dramatically improve a robot's view of objects because it enables the robot to know exactly where all its targets are. However, detecting from different angles of views raises a problem of pairing pixels. Experiments are therefore enforced and successfully lead to a common equation solving the problem. It is therefore more efficient to control robots.

6.0 Reference

Xtion Pro Specifications (2014), ASUS[online]：

http://www.asus.com.cn/Multimedia/Xtion_PRO/specifications/ [Accessed: 06/10/2014]