Enhanced Computer Vision with Microsoft Kinect Sensor: A Review

Yufeng Jiang

May 21, 2018

1 Kinect Software Tools

Kinect software refers to the Kinect developmment library and the algorithmic components included in the library. Currently, there are some aviliable tools including OpenNI, Microsoft Kinect SDK and OpenKinect. The Microsoft SDK is only available for Windows, but OpenNI can be used at multiplatform and its source is open. Table 1 gives a comparison between OpenNI and the Microsoft SDK about their algorithmic components. It is worth highlighting that the new version of OpenNI allows users th install Microsoft Kinect SDK on the same machine. And it also allows to run both packages using the the Microsoft Kinect driver. It means that the OpenNI is now compatible with the Kinect driver.

Table 1: Comparisons of the OpenNI Library and the Microsoft SDK

	OpenNI	Microsoft SDK
Camera calibration	√	√
Automatic body calibration	×	√
Standing skeleton	√	√
Seated skeleton	×	√
Body gesture recognition	√	√
Hand gesture analysis	√	√
Facial tracking	√	√
Scene analyzer	√	√
3-D scanning	√	√
Motor control	√	√

2 Kinect Performance Evaluation

There are a few papers that evalute the performance of Kinect from either the hardware or the software perspective. In [1] , the authors investigate the depth measurement of Kinect about resolution and precision from experiments. The result of this investigation reveals that Kinect is superior in accuracy to the TOF camera and close to a medium-resolrtion stereo camera.

3 Object Tracking and Recognition

In this section, authors begin our exploration of the technical principles, overview what Kinect can do and how people change and enhance related Kinect techniques in order to solve vision problems. Object detection and tracking are hot topics in RGB-based image and video analysis applications. With the availability of the low-cost Kinect depth camera, researchers establish a stable background model which is resistant to changes in illumination or lack of contrast. Figure 1 illustrates this.

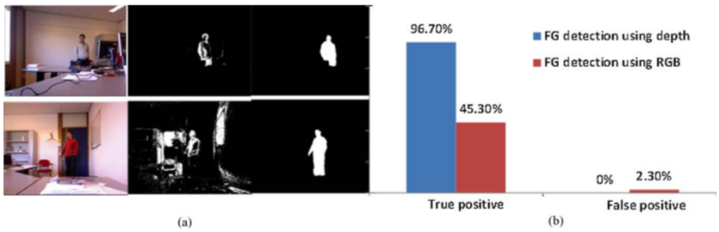

Figure 1: (a) From the left to the right, RGB images, FG mask using RGB data and FG mask using depth data, respectively. The example on the top shows that it is hard to distinguish clothing from the background, and the one at the bottom reports the results when the person suddenly turns on the lighting. (b) True positive and false positive of FG detection by using depth data and RGB data respectively. Here, the lighting is stable but the foreground and background are similar.

References

[1] M. J. J. Smisek and T. Pajdla, "Accuracy and resolution of kinect depth data for indoor mapping applications," *Proc. IEEE ICCV Workshops*, pp. 1154–1160, 2011.