The Kinect Sensor in Robotics Education

Michal Tölgyessy#1, Peter Hubinský*2

#Institute of Control and Industrial Informatics, Faculty of Electrical Engineering and Information Technology, Slovak University of Technology in Bratislava

Ilkovičova 3, Bratislava, Slovakia

1michal.tolgyessy@stuba.sk

2peter.hubinsky@stuba.sk

***Abstract*— This paper deals with education in the field of robot sensing abilities. It briefly introduces the commonly known and used concepts and sensors, but focuses mainly on the recent Kinect sensor. Technical information and background on the Kinect are provided. The last part of the article deals with possible applications of the sensor in various robotic fields with emphasis on the educational process.**

***Keywords*—** sensor, Kinect, depth map, camera, 3D modelling

I. Introduction

One of the basic attributes of a robot is its ability to interact with the world around it. Students have to learn that activities such as sensing and interpreting the surroundings, which are absolutely intuitive for a human, are very complicated tasks in the case of robotic systems. In order to understand how robots accomplish these tasks students have to get acquainted with sensors. The most commonly used sensors in mobile robotics are the active ones, such as infrared, ultrasonic and laser sensors [1]. They help the robot to understand the environment providing it with depth information which can be used to avoid collisions, bypass obstacles or to create more or less sophisticated maps. Students usually have to learn to understand the basic functionality of these sensors, they need to be able to analyse the analogue and digital outputs and interpret them in order to accomplish given tasks. For example a line following robot is usually equipped with three or more infrared sensors that are used to detect a black or white line placed on the floor underneath a simple wheeled robot. After a correct algorithm implementation the robot should be able to properly process the sensor information to follow the line. By implementing algorithms and accomplishing tasks such as the one mentioned students gradually develop their understanding of the interaction between robot and its local environment. Over time they fully grasp the fact that it is a large set of discrete operations that have to be implemented on real-time digital systems. What we consider a higher level in the educational process is using a digital RGB camera as a sensing unit. The use of camera introduces the vast technical field of visual systems, where an M x N matrix of pixels represents the reality around the robot as a 2D image. In order to use a camera as a sensor effectively students have to learn to extract features from images. This includes image pre-processing, segmentation, edge detection, blob detection, object recognition and other operations [2].

The purpose of this article is to present an even more sophisticated sensor that combines the advantages of standard distance sensors and RGB cameras. The Kinect sensor originally developed by PrimeSense and Microsoft for the Xbox 360 gaming console has been hacked and is being used in many hobby and robotic applications. There already is a commercially available robotic platform called Bilibot, which is based on the iRobot Create platform and uses Kinect as its main sensor [3].

Fig. 1 The Bilibot platform [3]

Authors of this article believe that complex depth sensing combined with RGB sensing will become a trend in robotics and Human-computer interaction (HCI) in general. A 3D visual sensor simplifies many common perception tasks and can be a powerful tool in education because working with directly mapped depth data on particular image pixels is much more intuitive than working with either pure RGB data or pure depth data.

II. The Kinect Sensor

The Kinect entered the global market in the beginning of November 2010 and was immediately a huge success selling more than one million products in ten days. To access the Kinect data with non-proprietary software the USB communication had to be reverse-engineered. Using data grabbed by a USB analyser the Spanish hacker, Héctor Martín, was the first one to access and display the Kinect RGB and depth data on a PC using the Linux operating system with OpenGL and OpenCV libraries [4].

The basic parts of the Kinect are (Fig. 2):

- RGB camera
- 3D depth sensing system
- Multi-array microphone
- Motorized tilt

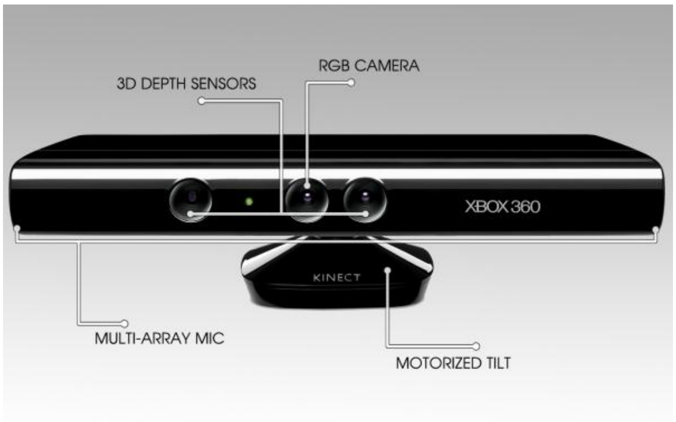

Fig. 2 The Kinect sensor

Kinect is able to capture the surrounding world in 3D by combining the information from depth sensors and a standard RGB camera. The result of this combination is an RGBD image with 640x480 resolution, where each pixel is assigned a color information and a depth information (however some depth map pixels do not contain data, so the depth map is never complete). In ideal conditions the resolution of the depth information can be as high as 3 mm [3], using 11 bit resolution. Kinect works with the frequency 30 Hz for both RGB and depth cameras. On the left side of the Kinect is a laser infrared light source that generates electromagnetic waves with the frequency of 830 nm. Information is encoded in light patterns that are being deformed as the light reflects from objects in front of the Kinect. Based on these deformations captured by the sensor on the right side of RGB camera a depth map is created. According to PrimeSense this is not the time-of-flight method used in other 3D cameras [5].

III. Basic Advantages

The first and major advantage of Kinect in robotics and the educational process is its impact on image segmentation tasks. The simplification comes from the depth data. With a single camera it is impossible to distinguish objects of similar colors. For instance, if a white box stands 1 meter in front of a white wall a robot with a camera is not able to find differences between these two objects from the RGB data. However, with the Kinect providing a 3D map the segmentation is very simple using just a single distance threshold. Without its application, there is a lot of noise and unwanted objects in the image (Fig. 3). After its application the desired objects can be easily segmented (Fig. 4).

Fig. 3 Depth data without a threshold

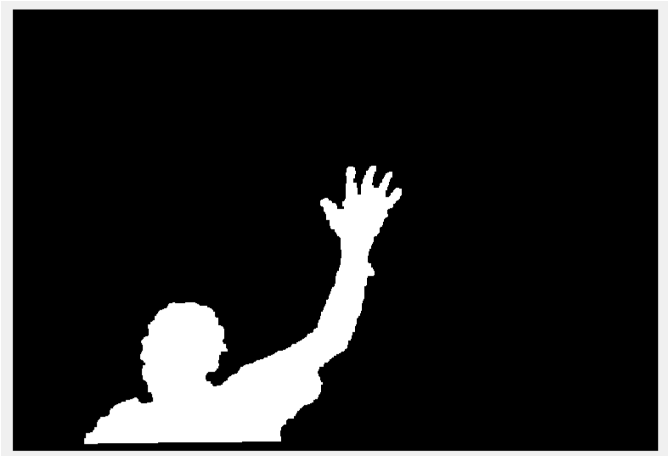

Fig. 4 Depth data with the threshold applied

Using this thresholding method is great for the educational process, because students can immediately see the results of the operations applied, whereas with IR or ultrasonic sensors they see only numerical output provided by the microcontroller via serial interface and have to interpret the data themselves. With 3D data combined with color RGB data a lot of this interpretation is intuitive and automatic.

IV. Applications

The Kinect sensor can be used in a variety of robotics applications being an addition to older methods or a complete substitution. Students can compare the pros and cons of several approaches and find out first-hand what suites their particular project more.

A. Data Fusion

Students can learn to fuse different data. The RGB information can be converted to any commonly used color space, such as Normalized RGB, HIS, HSV, HSL, TSL, YCbCr, CIELAB or CIELUV. All of these color spaces are commonly used for different tasks in visual systems applications. One of their primary utilizations is object detection and segmentation, which is also used in computer and robotic vision. Hence, by using Kinect as a sensor for a robotic system students can learn to fuse depth information with different color space data.

B. Obstacle Avoidance And Collision Detection

Providing a depth map with good resolution the Kinect can be used for collision detection and obstacle avoidance. By applying safety thresholds the robot can be informed any time an unknown object crosses a defined distance. Based on this information it can interrupt its current activity and stop motion or change the direction of its movement. For this the depth map alone is good enough, however for more sophisticated tasks the combination with RGB image is beneficial. Robots can be trained to behave differently when encountering different type of obstacles. When there are several obstacles with same color characteristics, such as tree trunks (brown), concrete objects (gray) or water (blue), the robot will be able to gather more information and classify each object with more detail. Thus, more sophisticated decision trees can be implemented resulting in greater and more reliable autonomy.

C. Object Recognition

When accomplishing autonomous tasks robots might need to recognize certain known objects, such as inscriptions, signs, faces, holes or cracks. All these objects can be primarily recognized by visual systems methods, but the depth information comes in as a very helpful addition. It can help in determining the vertical position of objects, informing the robot wether the recognized object is above, beneath or on the ground level. It can be hard to find cracks and holes from RGB image, but is much easier to do so with depth information. Combining visual algorithms with depth algorithms is a beneficial fusion in object recognition tasks.

D. Gesture Control

A part of the Human-robot interaction (HRI) is controlling a robotic system with hand gestures and body poses. It has been the subject of many research works. Interacting with digital systems without the need of a mouse, keyboard or joystick is the future of modern households. Many algorithms using only a single RGB camera have been proposed. However, the 3D information the Kinect provides is a major help in accomplishing both static and dynamic gesture recognition algorithms. Segmentation and skin detection is an important step in finding hands and is greatly simplified with depth information.

E. Localization And Navigation

One of the basic goals of autonomous robots is their ability to localize themselves and successfully navigate to a defined destination. The depth information provided by the Kinect can be of great help in map creation and localization. If the global map used by the robot contains color information, the interactive online color object recognition can be used to enhance the localization precision.

Kinect can also be used to implement visual odometry. With bare 2D color information visual odometry is practically impossible. Combining the color image with depth map opens new possibilities to odometry applications. Students can compare for instance incremental sensor outputs with the visual-depth odometry outputs. The incremental sensor data corruption caused by wheel sliding can be corrected by reliably designed odometry algorithms based on the Kinect outputs.

F. 3D Modelling

The depth data acquired from the Kinect can be used to create a 3D map of the environment, however adding the color data allows the creation of a complex color 3D model. If a reliable algorithm is implemented the robot can add static local 3D images together as it moves. The result of this is a 3D color model of a corridor (Fig. 5, Fig. 6). If such corridors are correctly attached to each other a complex 3D map of a local environment can be gradually created.

Fig. 5 3D model of a corridor (A)

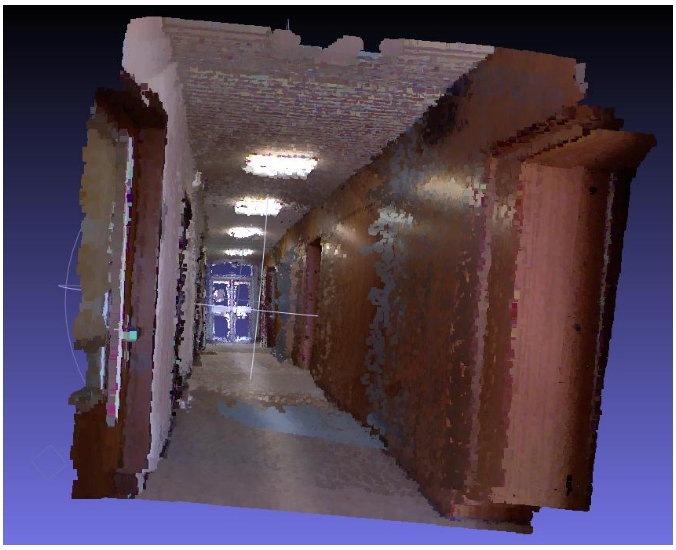

Fig. 6 3D model of a corridor (B)

V. PRACTICAL PROPOSALS

The technical parameters and capabilities of the Kinect are one thing, but its practical application in classes for students is another. Our ideas are forced to be merely theoretical because the Kinect is a very recent sensor. The official Microsoft SDK was released only in June 2011, therefore its inclusion to the educational process is limited by lack of time, despite the existence of many hobby applications. A hobby application requires a functional or semi-functional result with no documentation. On the other hand a University course requires a fully functional platform with an environment and software interface suitable for a particular group of students with certain experience level.

We decided to divide the Kinect's educational application into two areas. The first one focuses on standard practical classes lasting several weeks that usually support lectures, the second one consists of large projects, such as bachelor's and diploma projects.

Of course, the role of the Kinect sensor as a part of a class depends on the subject taught. Generally, we think that it's best to start teaching the least sophisticated algorithms and hardware and only then finish the course with a complex sensor, such as the Kinect. This way the students will learn to use the standard ways, but will also realize what benefits can the Kinect bring comparing to what they already know. We think that the "comparison moment" is very important. Students won't just read and learn what 3D sensing brings to the subject they study, but they will experience it. Through this experience they will creatively come up with their own ideas and solutions.

The second area mentioned includes large projects where Kinect is the central sensor. We advise mounting it on mobile platforms, such as hexapods, wheeled mobile robots or quadrocopters. Student can then pick a task, design and implement a method or algorithm under the supervision of his or her teacher.

VI. CONCLUSIONS

The presented sensor is a great and complex tool that can be used to teach students many common robotic tasks, ranging from the easiest (collision detection) to the most complex (3D mapping). As 3D sensing is a current research trend, studying robotics with Kinect as a part of the sensor equipment is both challenging and motivating. Students can compare the results obtained from depth and color algorithms with the more common methods that use infrared, ultrasonic or laser sensors. We also believe that working with a hacked proprietary sensor sold by a major company can increase their motivation.

ACKNOWLEDGMENT

This work was supported by MS SR under contract VMSP-P-0004-09 and VEGA 1/0177/11.

REFERENCES

[1] R. Siegwart and I. R. Nourbakhsh, *Introduction to Autonomous Mobile Robots*, Massachusetts Institute of Technology, The MIT Press, 2004.

[2] M. Šonka, V. Hlaváč and |R. Boyle, *Image Processing, Analysis and Machine Vision*, Pacific Grove : Brooks/Cole Publishing company, 1999.

[3] The Bilibot website. [Online]. Available: http://ww.bilibot.com

[4] We Have a Winner. [Online]. Available: http://www.adafruit.com/blog/2010/11/10/we-have-a-winner-open-kinect-drivers-released-winner-will-use-3k-for-more-hacking-plus-an-additional-2k-goes-to-the-eff/

[5] Kinect: The company behind the tech explains how it works. [Online]. Available: http://www.joystiq.com/2010/06/19/kinect-how-it-works-from-the-company-behind-the-tech/