REAL TIME OBSTACLE DETECTION AND AVOIDANCE USING KINECT DEPTH SENSOR AND PIONEER LX MOBILE ROBOT

Thanh-Huong Nguyen, Minh-Hoang Le
International Research Institute MICA, Hanoi University of Science and Technology
E-mail: thanh-huong.nguyen@mica.edu.vn

ABSTRACT This research has been conducted on an Adept Mobile Robots Pioneer LX under the framework of research for assistive systems for blind people. Traditionally blind people used to employ long cane or guided dog to take part in mobility activities. The Mobile Robots Pioneer is designed for carrying up-to-60 kg weight and working in indoor environment, which is appropriate for a wheelchair accessible. The Kinect 3D sensor has the ability to detect and calculated relatively accurate the distance to the obstacle, therefore it can help the robot to avoid the obstacle and implement navigation more easily. In this paper, a Kinect depth based real time stationary obstacle detection algorithm was presented. Besides, a scenario of completing a trajectory along a long corridor using a combination of Mobile Robot Pioneer LX and Kinect Sensor was conducted.

1. INTRODUCTION

According to the World Health Organization in 2014, there are around 286 million visually-impaired people including 69 million blind people (WHO, 2014). Vision loss results in independence loss, especially limited mobility. Lack of mobility is a serious hindrance for the individual. Blind people find difficulties to move independently, as they cannot specify in advance the traveling direction and object location along their path and in the surrounding environment.

The traditional aided mobility devices for most of the blind people are the white cane and the guide dog. These mobility choices support blind people to move through known and unknown environment independently. The white cane involves a simple cane and can detect stationary obstacles through tactile feedback (Armstrong, 1975). The guide dog can distinguish complex situations such as cross walks, staircase and so on. The resulted information is received via tactile feedback. Another option is to use their audition to compensate the vision loss. Recently blind people have strongly used the sound in order to travel and maintain their safety. Besides, the information from the surrounding permits the blind people to recognize through the sources and sounds from the surrounding, such as traffic lights, traffic, noises from coming from the machinery, animals or people, etc.

Wheelchair is one of the novel Electronic Travel Aid System (ETA) which is not only used for blind people but for handicapped people in general. The ETA systems are based on three interfaces: the input interface, the processing interface and the output interface. The input interfaces collect the environmental data and are classified in: ultrasound, laser, and artificial vision and GPS system. The processing interface is composed of the techniques and the software for processing all the acquired information and for transforming it into required data for the output interface. The output system represents the model for transmitting the information from the device to the user. It should be as much concise and clear as possible, in order not to confuse and disturb the user.

Our proposed system utilizes the Kinect sensor as the input interface and processes the environmental information by the computer inside the mobile robot or an integrated laptop. After processing the depth information from the Kinect, the distance and the direction of the obstacles are extracted. Then command will be determined to avoid obstacle and the instruction will be transferred to the robot automatically.

2. RELATED WORK

An autonomous and intelligent mobile robotic assistant system is proposed in this paper to make it possible for the disabled to live independently, safely and comfortably rather than move to a costly healthcare facility. The system works in a coordinated and efficient manner to carry out the tasks. The intelligent mobile robots navigate autonomously through the home environment and transmit the data through wireless network to the remote control center.

A robot manipulator is developed to assist workers with disabilities (Hun et al., 2005). This robot manipulator is designed to carry out the task as PCB circuit testing and inspection of soldering. An electromyographic (EMG) based semiautonomous human – robot interaction system is presented (Rani et al., 2005). It allows the disabilities can send high-level commands to robot for some daily living activities. A guide dog robot and a stereotyped motion following a person are developed. The guide dog robot consists of a mission planner, digital map, interactive navigator, vision system and undercarriage system (Mori and Sano, 1991).

A whole-field target tracking and following mobile robot system is developed based on a pan/tilt/zoom CCD vision system. The vision system scans and locks the pose of the moving target and commands the tracking mobile robot to follow the target while avoiding obstacles (Lee et al.,2013). This paper proposed a novel appearance design and fabrication on an existing mobile robot to assist blind people to achieve independent living and communicate through established social networking. The mobile robot control system is presented in two other separate papers as the intelligent control system and visual servo control system (Lee and Chiu, 2013).

In this paper, the designed system is a guide robot connected to a depth camera Kinect which later can be developed as a wheelchair. This human machine interface allows a convenient operation. The resulting automatic navigation provides the optimal and collision-free path. It significantly reduces the risk for blind users in unfamiliar environment.

3. OVERVIEW OF THE SYSTEM

The obstacle detection module takes scene information from a mobile Kinect. In our prototype, the obstacle detection is running on a laptop mounted on a backpack of the visually impaired people and mobile Kinect is the Kinect with battery so that it can be mounted easily on the human body for collecting data and transferring data to the laptop. The scene information, in our case, is the color image, depth image, and accelerometer information provided by Kinect.

Concerning the actuator, the pioneer Mobile Lx was used to perform the mobility based on the information that the Kinect processed. In our work, we consider indoor environment where obstacles are defined as objects in front, obstructing or endangering while blind people are moving. Specifically, we focus on detecting static objects (e.x. trash, plant pots, fire extinguisher). In the following, we will describe in detail the obstacle detection.

3. MICROSOFT KINECT

After introducing the Kinect sensor for Microsoft's Xbox 360 in November 2010 and a version for Windows on February 1, 2012 many computer scientists have used the Kinect as a robotic sensor. It is low cost sensor that includes both an RGB channel and a 3D depth channel, which can provide more information of the scene, work well in a low light environment and they are efficient for real-time processing.

RGB-D camera captures both RGB images and depth maps at a resolution of 640× 480 pixels with 30 frames per second. The effective depth range of the Kinect RGB-D camera is from 0.4 to 3.5 m. The Kinect color stream supports a speed of 30 frames per second (FPS) at a resolution of 640 x 480 pixels. The main purpose of using Microsoft Kinect sensor is to reconstruct 3D scene in front of the user from Color-Depth that represents a crucial data which is necessary information for visually impaired people. The system proposed using a laptop for processing color-depth images to extract accurate information about the obstacles.

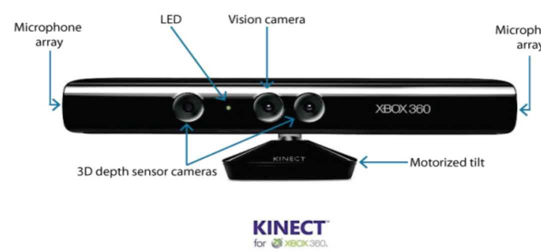


Fig 1. The structure of Microsoft Kinect Sensor

4. OBSTACLE DETECTION

The depth information was processed using the Point Cloud Library (PCL). The PCL support the OpenNI and Kinect SDK interface. In addition to PCL, some libraries can be also used such as OpenCV, Boost, Eigen and so on. To detect obstacles, first we use data from Kinect (color image, depth image and accelerometer data) to build point cloud. Then we detect ground plane and walls plane in the image by using plane segmentation in point cloud proposed by Holz et al., 2012. After that, we detect all obstacles in the scene including static obstacle and human and check for the nearest obstacle to make the instruction for the robot. This whole process can be seen in the Fig. 2.

4.1. Reconstruction

This step contains reconstruction, filtering and rotating point cloud:

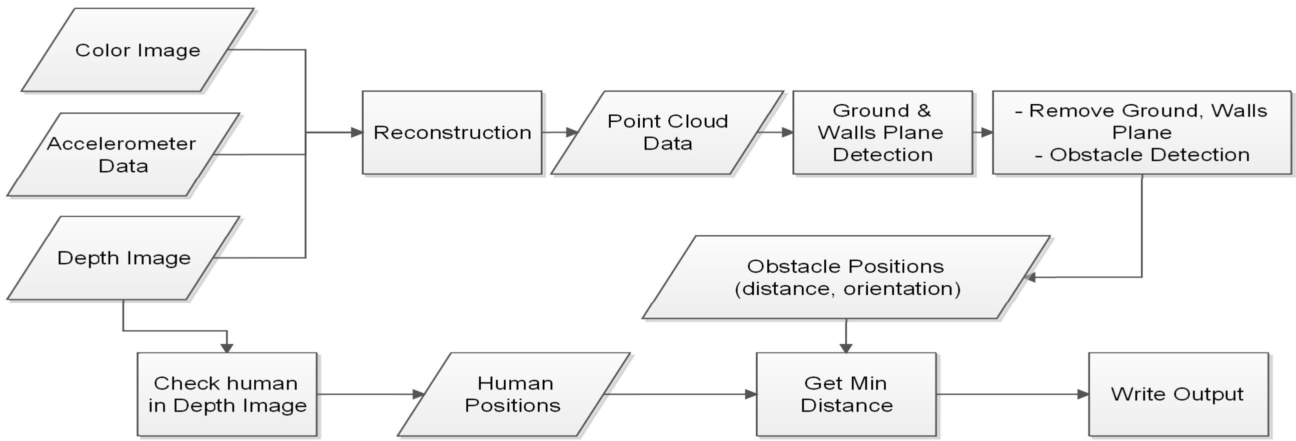


Fig 2. Obstacle Detection Process

– Reconstruction: In this stage, depth and color image will be combined to make a 3D Point Cloud using Point Cloud Library (PCL).Kinect is a low-cost RGB-D camera that can provide various types of data including color image, depth image, accelerometer data, skeleton information, sound from microarrays. However, the color and depth image were captured by two different sensors hence they are not aligned. That means given a pixel in the color image, we cannot get correspondent pixel in depth image directly as well as 3D coordinate. To make a 3D Point Cloud from Kinect data, with each pixel in both color and depth image, we must know exactly the location of this pixel in the 3D coordinate to create an RGB-XYZ point in Point Cloud. To solve that problem, a lot of work has focused on developing a good calibration method in order to transform between color-coordinate, depth coordinate and real world coordinate such as Microsoft Kinect SDK.

In this project, we used Microsoft Kinect SDK to convert depth coordinate to color coordinate, then use parameter from to convert to 3D coordinates. Given a depth and color image. For each pixel in the depth image, we can find its 3D coordinate in meter by this formula:

$$P3D.x = (x_c - cx_c) * depth(x_c, y_c)/fx_c$$

$$P3D.y = (y_c - cy_c) * depth(x_c, y_c)/fy_c$$

$$P3D.z = depth(x_c, y_c)$$

where x_c and y_c is the pixel coordinate in color image, cx_c , cy_c , f x_c , f y_c is taken from color intrinsic matrix, depth(x_c , y_c) is the depth value of pixel. This process is illustrated by Fig. 3.

– Filtering: Because there are a lot of points in point cloud (about 300.000 points with VGA resolution), so the system becomes time-consuming and cannot run in the real-time. To reduce the execution time, point cloud will be down-sampled using 2x2 block. Consequently, the number of points in the cloud will be reduced by 4 times.

– Rotating Point Cloud: As mentioned in section 3, our system using mobile Kinect, which means Kinect mounted on the body. Therefore, while the visually impaired people moving, because Kinect is shocked, shaking so that the point cloud will be rotated due the changing of Kinect direction. In our project, we used accelerometer data provided by Kinect SDK to rotate point cloud in order to align the ground plane with the xz-plane in reference system.

The accelerometer data is actually a 3-D vector point- ing in the direction of gravity with coordinate system is centered on the sensor shown in the Fig. 5. With the default Kinect configuration (horizontal) is represented by the (x, y, z, w) vector whose value is (0, -1.0, 0, 0). We use this vector to build rotation matrix and then apply it into point cloud data in order to rotate point cloud. Fig. 4 shown the output of this stage.

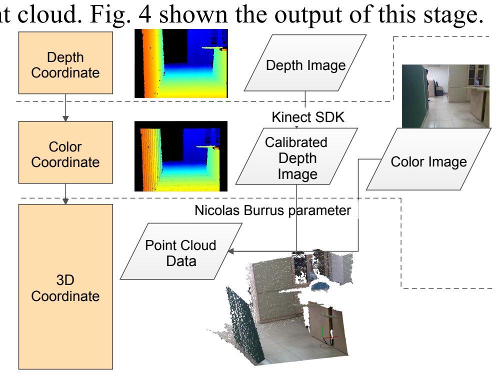


Fig 3. Coordinate Transformation Process

4.2. Ground and walls plane detection

– Plane Segmentation: Finally, point cloud will be segmented into dominant planes. This is very important step because our algorithms based on the ground plane detection. The plane based segmentation using in this project is based on the algorithm which uses the normal vector to segment point cloud data into multiple planes in real time. The main idea and also the advantages of this algorithm is that plane segmentation can be done very fast using both information in image structure and point cloud data. Because after converting color and depth image to point cloud data, each pixel is the point in the 3D space and the relationship between pixels is lost. For example, when we want to find the neighbors of the point in the point cloud, we must calculate the distance between this point with all remaining points in the point cloud or do some sorting algorithms like KD-tree, this process is time consuming. In this algorithm, the authors proposed a new normal vector estimation using an integral image, so it can run in real time. This algorithm can be illustrated in Fig. 5.

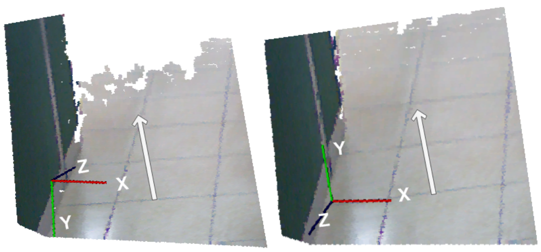


Fig 4. Point Cloud rotation using normal vector of ground plane (while arrow): left: before rotating, right: after
rotating

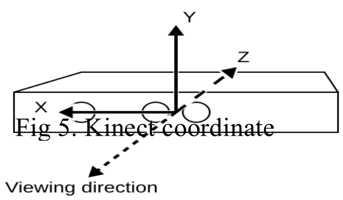


Fig 5. Kinect coordinate

A normal vector of a single point can be calculated be a cross product of two vectors of four neighbor points: bottom-top and left-right. So, in this algorithm, the authors first calculate two maps of tangential vectors, one for x- and the other for y-dimension. After normal estimation, planes can be detected by segmentation in normal space. The result of this step can be shown in Fig. 6.

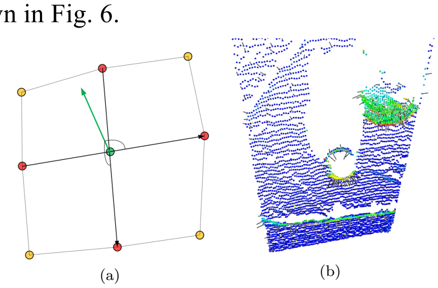


Fig 6. Normal vector estimation algorithms
(a) Normal vector of the center point can be calculated by a cross product of two vectors of four neighbor points (red) and (b) Normal vector estimation in a scene.

– Ground and Wall plane detection: After planes have been segmented, ground and wall planes can be detected easily using some constraints. Because our point cloud has been rotated to align with ground plane in the previous step using gravity vector. So, the ground plane must be satisfied some condition:
– The angle between gravity vector and ground plane's normal vector is almost 0 degree
– Ground plane must be large enough. In our case, we checked number of point inside a ground plane, if the number of points is larger than 10000 points, then we consider it's a ground plane candidate
– Because Kinect's mounted at the human body, so distance between ground plane and Kinect (y- axis coordinates) must be in a range of 0.8-1.2m

Wall is considered as perpendicular plane to the ground plane. So, to detect wall planes, we use similar constraints with ground plane except the angle between gravity vector and wall's normal vector is almost 90 degree and don't need to check distance between wall plane and the Kinect because wall plane can be appear anywhere in our scene. Then, all the point belonging to ground and walls plane will be removed. Fig. 7 shows the ground and wall plane detection results.

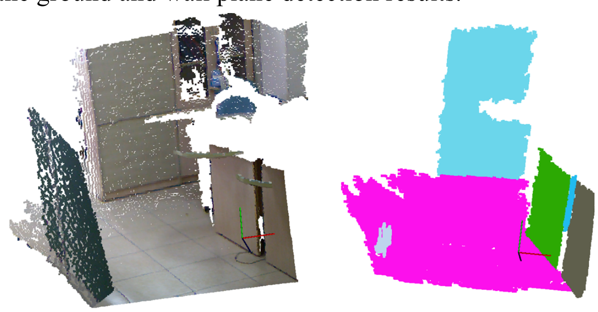


Fig 7. Plane segmentation

4.3. Obstacle detection

In this step, we will detect obstacles from the remaining point cloud. There are two kind of obstacle: human and static object. With human detection, Microsoft Kinect SDK also provided human segmentation data. Kinect can track up to 6 people in a camera field-of-view. This data is encoded as 3 lowest bit for each pixel in depth image and represented index of the person that Kinect has been tracked.

After checking human data in the frame, we remove all point that belong to the human in the point cloud and do clustering to find remaining obstacle in the scene. To do this, firstly, all points in the cloud will be rearranged by coordinate using the KD-tree algorithm. Then, based on the Euclidean clustering algorithm provided by PCL library, each obstacle will be segmented from the point cloud. For obstacles lying on the ground, we calculate the distance to the user to give a warning message.

5. PIONEER LX MOBILE ROBOT

The Pioneer LX is an advanced mobile robotics research platform based on the Adept Lynx industrial AIV (Autonomous Intelligent Vehicle). This ground robot is programmable, and easy to add, switch and customize different sensors, effectors and other equipment for new projects

Fig 8. Pioneer LX Robot

The Pioneer LX has been designed for continuous non-stop industrial service and can operate up to 13 hours before recharging. This autonomous mobile robot can carry payloads of up to 60 kg over indoor surfaces in wheelchair-accessible facilities that resemble a wheelchair.

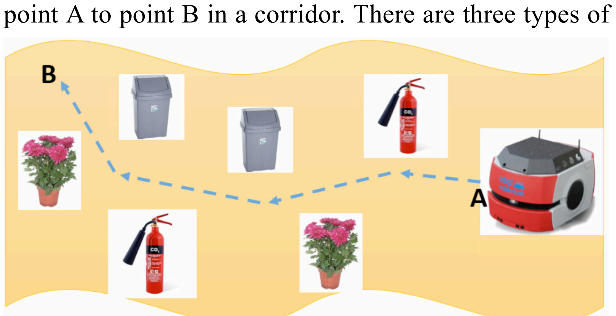
The Pioneer LX can travel at speeds up to 2 m/s with full payload. As with other MobileRobot platforms, the Pioneer LX includes extensive Pioneer SDK, a set of software applications and libraries to accelerate the pace of development. All of our robotics platforms can also be used in a "semi-autonomous" fashion in which the robots will navigate autonomously but respond to commands from a remote control computer.

6. EXPERIMENT

In order to test the obstacle detection program, we connect and send the instruction to the robot based on the obstacle detection result. The following Fig.9 shows the scenario to test.

Fig 9. The test scenario

The robot was required to perform the route from point A to point B in a corridor. There are three types of



objects on the path, the fire extinguisher, the flower pot and the dust bin. To finish the route, the instruction for the robot has to be made based on the constraints. In our case, we used the angle constraint because the robot can rotate around itself and then go straight. The following Fig.10 shows the angle condition.

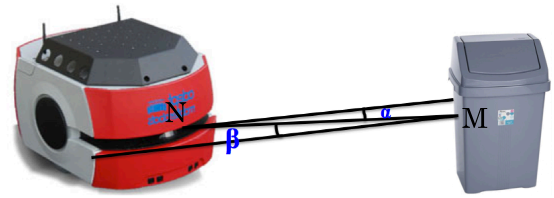


Fig 10. Angle regulation

M and N are the mid point of the side of object (ex: dust-bin) and the front of robot. The angle **α** is the angle between MN line and the line connecting the edge of object and midpoint N of the robot while the angle **β** is the angle between MN line and the line connecting the edge of robot and the midpoint M of the object. Kinect can also calculate the angle through vector calculating. The robot when detecting the object will stop and calculate these angles. After that it will rotate an angle = (**α** + **β**) in order not to collide with the object. Then it will continue to travel until detecting the next objects.

7. RESULTS

For obstacle detection evaluation, we tested our program with 248 images collected from MICA hallway. Our system was evaluated on a notebook with an In1tel Core i3 2328M processor and 6GB memory inside. The system operates at an average speed of 2 Hz (493 ms/frame) with downsample block is 2x2 (about 75000 points in point cloud) , which is fast enough to be used in practice.

Fig. 11 shows average detection time of each step and the whole process.

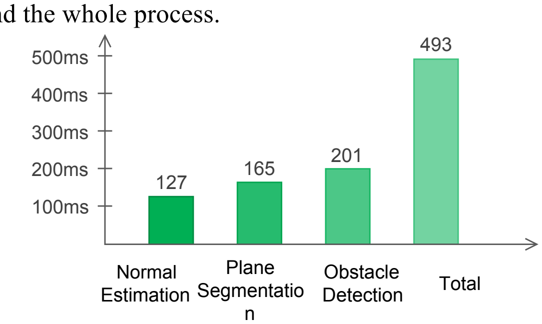


Fig 11. Detection time

To make evaluation, we used precision, recall and F-measure measurement:

$$Precison = \frac{TP}{TP + FP}$$

$$Recall = \frac{TP}{TP + FN}$$

$$F = 2\frac{Precision * Recall}{Precision + Recall}$$

where: TP: True Positive (true detection) FP: False Positive (object detected not in ground truth) FN: False Negative (miss detection)

Table 1. Object detection result

TP	FP	FN	Precision	Recall	F-Measure
344	71	154	82.9%	69%	75.3%

We made evaluations on this dataset on object detection. We used Watershed algorithm on depth image to segment object from the background and making the ground-truth. To evaluate, we back project the point cloud into 2D images to make binary mask of original image where white pixels is the detected obstacle as shown in Fig. 12.

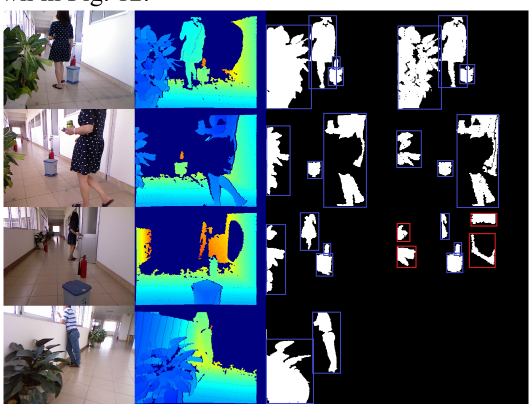


Fig 12. Obstacle Detection Result

For the scenario test, the Pioneer LX can finish without collision 36 trials in total 48 trials (75%).

CONCLUSION

The Mobile Robots Pioneer is designed for carrying up-to-60 kg weight and working in indoor environment, which is appropriate for a wheelchair accessible. The Mobile Robot itself has the laser rangefinder sensor, ultrasonic sonar sensor, an Ubuntu-based embedded computer and a complete robot control system to navigate in the flat topology. The obstacle detection algorithm was presented for stationary obstacles. The Kinect 3D sensor has shown that it can calculate relatively accurate the distance to the obstacle and can help the robot to avoid the obstacle. This work focused on indoor environments and detection of static objects. A scenario of completing a route along corridor using a combination of Mobile Robot Pioneer LX and Kinect Sensor was conducted. Three classes of obstacles were considered and placed along the path: fire extinguisher, flower pot, dustbin. The resulted system gives relatively accurate detection and can instruct the robot to complete the scenario without bumping into given obstacles. The obstacle detection of 82.9% of precision and 75% of route completion without collision.

ACKNOWLEDGEMENT

This work was supported and funded by Hanoi University of Science and Technology (Hanoi, Vietnam) under grant number T2015-055 on the research project "Obstacle detection and classification from mobile Kinect for robot in indoor environment".

REFERENCES

Fact Sheet on Visual impairment and blindness,WHO, www.who.int/mediacentre/factsheets/fs282/en, 2014

Armstrong, J.D., Evaluation of man-machine systems in the mobility of the visual handicapped. *Human Factors in Health Care R. M. Pickett and T. J. Triggs Eds. Lexington Book*, Massachusetts, 1975

Hun C.P., Rae P.S., J. Je Hyung, Hyun P.S., Development of a robot arm assisting people with disabilities at working place using task-oriented design. *9th International Conference on Rehabilitation Robotics*, (ICORR 2005), pp.482 – 487, 28 June-1 July 2005.

Rani P., Sarkar M.S., EMG-based high level human-robot interaction system for people with disability, *IEEE International Workshop on Robot and Human Interactive Communication*, (ROMAN 2005), pp.280 – 285, 13-15 Aug. 2005

Mori H., and Sano M., Guide dog robot Harunobu-5 following a person, *Proceedings of the IEEE/RSJ International Workshop on Intelligent Robots and Systems* - IROS '91, Osaka, Japan, pp. 397-402, November 3-5, 1991.

Lee M.F.R. and Lee K.H.E., Autonomous target tracking and following mobile robot, *Journal of the Chinese Institute of Engineers, Transactions of the Chinese Institute of Engineers,Series A*, vol. 36, pp. 502-529, 2013.

Lee M.F.R. and Chiu F.H.S., A networked intelligent control system for the mobile robot navigation, *Proc. 2013 IEEE/SICE International Symposium on System Integration*, Kobe, Japan ,December 15-17, 2013.

Lee M.F.R. and Chiu F.H.S., A hybrid visual servo control system for the autonomous mobile robot*, Proc. 2013 IEEE/SICE International Symposium on System Integration*, Kobe, Japan, December 15-17, 2013.

Dirk Holz, Stefan Holzer, Radu Bogdan Rusu, and Sven Behnke, Real-time plane segmentation using rgb-d cameras, *In RoboCup 2011: Robot Soccer World Cup XV*, pp. 306–317,Springer, 2012.

Thanh Huong Nguyen received the B.E. (2007) at Hanoi University of Science and Technology (HUST), M.S. (2010) at University of Twente, and PhD (2014) degrees at University of Grenoble. She is now a lecturer and researcher at HUST.