Jurnal
Teknologi

Full Paper

EVALUATION OF SCENE PARAMETERS FOR OPTIMUM PERFORMANCE OF LOCALIZATION USING KINECT

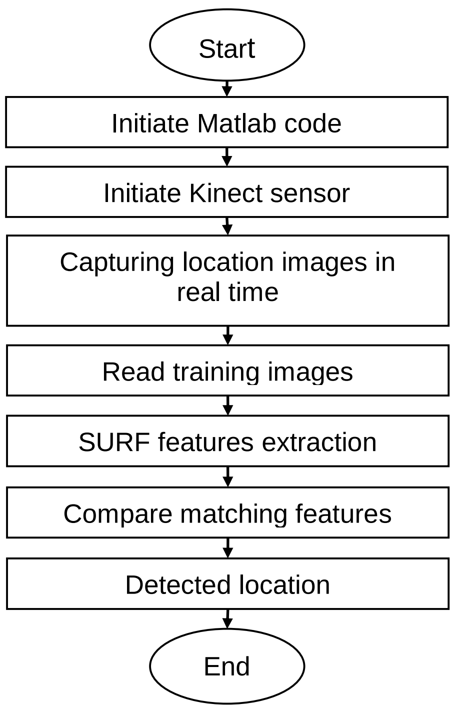
Bukhari Ilias*, Shazmin Aniza Abdul Shukor, Sazali Yaacob, Abdul Hamid Adom, Mohd Firdaus Haja Hussain

School of Mechatronic Engineering, University Malaysia Perlis, 02600 Perlis, Malaysia

Article history
Received
1 June 2015
Received in revised form
13 July 2015
Accepted
20 August 2015

*Corresponding author
bukhari@unimap.edu.my

Graphical abstract



Abstract

The needs of having locations identification are so important in this technological era. Locations identification has been implemented in so many areas nowadays. The challenge in this system is to identify location in real-time. Here, real-time refers to the process of capturing the location and recognize it at the same time. This research proposes a location identification system using Kinect, which is a high speed vision sensor that has been used for Xbox 360 video games. This research is to evaluate the optimum image detection distance, image size and processing time in identifying the locations that have been trained. The algorithm uses the Point Matching Features from the MATLAB software. From here, features are generated from both real time and training images being compared and the matching points between them will be used to identify the locations.

Keyword: Kinect sensor, MATLAB, high speed vision system, point matching features

Abstrak

Keperluan mempunyai pengenalan lokasi sangat penting dalam era teknologi ini. Pengenalan lokasi telah dilaksanakan di dalam berbagai bidang pada masa kini. Cabaran dalam sistem ini adalah untuk mengenal pasti lokasi dalam masa nyata. Di sini, masa nyata merujuk kepada proses merakam lokasi dan mengenalinya pada masa yang sama. Kajian ini mencadangkan satu sistem pengenalan lokasi menggunakan Kinect, yang merupakan sensor penglihatan kelajuan tinggi yang telah digunakan untuk permainan video Xbox 360. Kajian ini adalah untuk menilai jarak pengesanan imej, saiz imej dan masa pemprosesan yang optimum dalam mengenal pasti lokasi yang telah dilatih. Algoritma menggunakan Ciri-ciri Titik Pemadanan daripada perisian MATLAB. Dari sini, ciri-ciri yang dihasilkan dari kedua-dua masa sebenar dan latihan imej dibandingkan dan mata yang sepadan di antara mereka akan digunakan untuk mengenal pasti lokasi.

Kata kunci: Sensor Kinect, MATLAB, sistem penglihatan kelajuan tinggi, ciri-ciri titik pemadanan

1.0 INTRODUCTION

Localization is the most common problem in mobile robot navigation. Robots sometimes would face difficulties to know where is its positions and difficult to make its next moves. Some of the robots use Global Positioning System (GPS). However, GPS may not be available to be used inside buildings, factories and power plants. If the localization system can be used indoor, the information can be used by an autonomous mobile robot for the purpose of localization and navigation [1]. Localization is an important problem in robotics, especially for mobile robots working within domestic and office environment [2]. Current trends show that autonomous mobile robots are now playing an important role with the modern society. A relevant type of autonomous mobile robots is directed to the surveillance and any other indoor safety task [3]. Beside autonomous mobile robots, localization and navigations are also have been implemented in wheelchairs to help disable persons [4].

Moreover, robot ability to interact with the world around it is also important task rather than just for navigation or localization. All these tasks require installation of specific sensors within the robot. There are a lot of most commonly used sensors in mobile robotics such as infrared, ultrasonic and laser sensors. For the development of autonomous mobile robots, several aspects must be taken into consideration, such as the type of application and the basic robot platform and sensors adopted which compose the robot's hardware [5].

There are several sensors that could be used to collect information about the robot and the environment. In term of localization, there are three main problems in mobile robotics: the mapping of the environment, localization and navigation of the robot. Indoor localization systems have been conducted through many studies such as indoor messaging system (IMES), the radio frequency identification (RFID) system and ZigBee. These three localization systems have their own disadvantages, such as high cost, plus they do not have a high accuracy, especially in the depth directions [6]. Accuracy is an important element in design consideration for an autonomous mobile robot because it may cause fatal injuries for human beings. Besides the cost and accuracy, they also take a lot of time to process the information.

The use of Kinect sensor may help in term of cost, accuracy, and processing time for an autonomous mobile robot [7]. Besides Kinect, there are others sensors that can provide the same advantages but are costly, like the Hokuyo laser sensor. Kinect sensors have been adapted in many robotics applications and leisure time activities. The well-known robot that has been developed and commercially available which uses Kinect as its main sensors is Bilibot [8-9].

This research is to build a high speed vision system that work with microcontrollers to allow a mobile robot know its location. For vision systems, Kinect sensors is used and being attached to the mobile robot to determine its location. Kinect is chosen because of its advantages as discussed earlier. Software called 'MATLAB' has been used to interpret Kinect sensors through computer. This software is preferred because it can help Kinect to process image and provide the output to the microcontroller.

2.0 EXPERIMENTAL

2.1 Sensing Method

Basically there are a lot of sensing elements or hardware that can be used for localization such as sonar-based sensors, vision-based sensors and laser-based sensors. These three types of sensors can be used to get 2D and 3D data [10]. Although camera can be used as it can produce 2D images, it is better to use 3D processing images because of its relative to the global horizon plane. Conventional RGB camera could not generate the depth value.

Laser based sensors such as laser range finders, uses slope sensors and depth information to produce 3D data. However, it has no colour information. Meanwhile, for a range camera, it uses IR projection and capturing. Low resolution image will be produces from IR projection so the 3D information obtained will not be clear and accurate. A new sensing method called trinocular camera system also can be used. It has an arrangement of three RGB camera aligned together accurately to capture the same scene. As it uses three cameras, this method is more accurate compared to stereo camera which uses only two cameras. There have been researches on using a stereo vision camera-bumble bee, trinocular camera systems and Kinect sensors [11]. In this experiment, Kinect sensor has been chosen because it has the features of advanced vision sensor.

2.2 Feature Based Scan Matching

Scan matching is a process where it detects location by comparing its surroundings with the map environment [12]. It is the commonly used technique for localization. This scans matching process work by finding the transformation and tries to minimize the difference between current images with the reference images. If the best transformation is found, it can be used to correct and estimate the location. Feature based method depends on corners, edges or planes from the scan and then compare it. The feature extraction phase requires complex computations. The extraction features may have a lot of numbers of points which would slow down the matching process in order to recognize the location.

Experiments will be conducted to determine the speed of detection of the image based on this method.

2.3 Colour and Depth Segmentation

The color and depth segmentation technique are simpler process compared to other techniques. This technique straight away uses color and depth information in order to determine location. Even if the object is in the same color it able to differentiate it through the depth data. The probabilistic image technique also is used to determine or identify objects with the same color [13-14]. In term of localization, it seems that colour and depth segmentation are the best technique to implement. Moreover, the algorithm is easy and does not require a large processing unit and memory in order to process it [15-17].

2.4 Experimental Flowchart

Figure 1 shows flowchart of the point matching localization system:

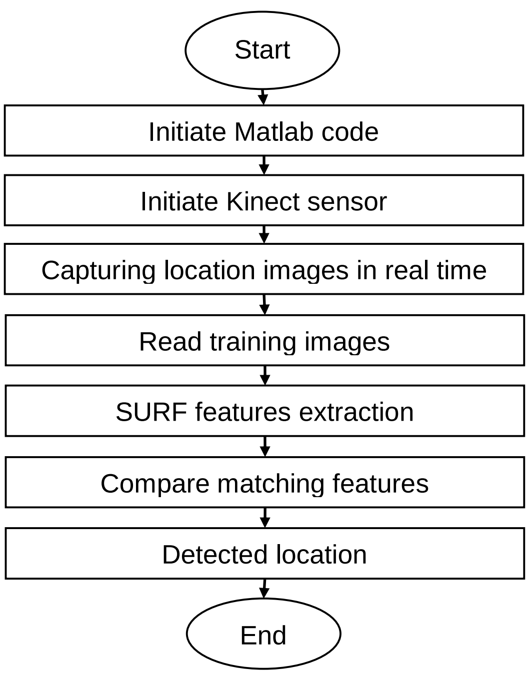


Figure 1 Flowchart of the system

2.5 Image Capturing Stage

During the image capturing stage, Kinect sensor, with the help of MATLAB, will capture the desired images. Kinect is chosen as the sensor used to capture the images due to its capability of capturing in high speed without any loss on the quality of the pictures. There are several other software that may be used with Kinect sensors but in this research, MATLAB has been chosen due to its ability to interact with Kinect sensors without any errors.

2.6 Image Processing Stage

In order to recognize the location, the image needs to be processed. In this stage, extraction and detection of SURF features of the training and real-time images will be extracted in order to locate the place from the images. The point matching between both real-time and training images will aid the process of identifying the location.

2.7 Image Recognition Stage

Image recognition stage is the last stage in identifying the location. In order to identify the location, several final processes needed to be done to obtain the results. The matching points between the real-time and training images will be displayed. A detection box will be drawn if it have sufficient amount of matching points between the real-time and training images.

2.8 Selected Locations

In this research, two distinctive things in a defined location (laboratory) have been selected for location recognition process. These things are selected based on its special features in the locations that could make the system recognize the location in real time. The distinctive images are shown in Figure 2 and 3.

Figure 2 Entrance A

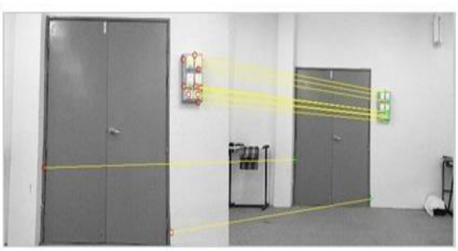
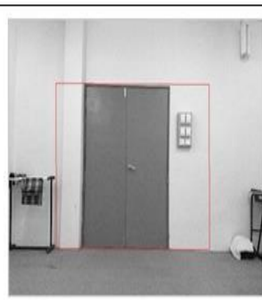
Figure 3 Switch board

3.0 RESULTS AND DISCUSSION

3.1 Image Processing Sequences

In order to obtain image detection stage, several sequences are needed to follow as shown in Table 1.

Table 1 Image process sequences

Sequences	Image
Real-time image capturing	
Read training images	
Generating feature points for real-time images	
Matching points between real-time and training image	
Recognition	

3.2 Image Size and Effect of Distances on Detection Rate

In this analysis, image size and distance have been considered as an important aspect in localizations. The accuracy of detection may effect if these two variables are been manipulated. In order to get the best image size and accuracy, the analysis need to be done. The further the distance of the locations from the Kinect, the smaller the size will be. Therefore, close analysis is done using different distance length from 10cm to 300cm. The best suitable range and image size need to be defined for the system to recognize the locations. In order to get the best detection result, several image sizes are used. The chosen sizes of images for the analysis are 640x480, 480x480, 320x240 and 240x240. These sizes are chosen in order to test the effectiveness of the system. Thirty sample tests were performed for each image size. Table 2 shows the results of image size versus detection percentage.

Table 2 Results of image size and detection percentage

Image No	Distance (cm)	Image Size			
		640 x 480	480 x 480	320 x 240	240 x 240
1	10	Not Detected	Not Detected	Not Detected	Not Detected
2	20	Not Detected	Not Detected	Not Detected	Not Detected
3	30	Not Detected	Not Detected	Not Detected	Not Detected
4	40	Detected	Detected	Detected	Detected
5	50	Detected	Detected	Detected	Detected
6	60	Detected	Detected	Detected	Detected
7	70	Detected	Detected	Detected	Detected
8	80	Detected	Detected	Detected	Detected
9	90	Detected	Detected	Detected	Detected
10	100	Detected	Detected	Detected	Detected
11	110	Detected	Detected	Detected	Detected
12	120	Detected	Detected	Detected	Detected
13	130	Detected	Detected	Detected	Not Detected
14	140	Detected	Detected	Detected	Not Detected
15	150	Detected	Detected	Detected	Not Detected
16	160	Detected	Detected	Detected	Not Detected
17	170	Detected	Detected	Detected	Not Detected
18	180	Detected	Detected	Not Detected	Not Detected
19	190	Detected	Detected	Not Detected	Not Detected
20	200	Detected	Detected	Not Detected	Not Detected
21	210	Detected	Detected	Not Detected	Not Detected
22	220	Detected	Detected	Not Detected	Not Detected
23	230	Detected	Detected	Not Detected	Not Detected
24	240	Detected	Detected	Not Detected	Not Detected
25	250	Detected	Not Detected	Not Detected	Not Detected
26	260	Detected	Not Detected	Not Detected	Not Detected
27	270	Detected	Not Detected	Not Detected	Not Detected
28	280	Detected	Not Detected	Not Detected	Not Detected
29	290	Detected	Not Detected	Not Detected	Not Detected
30	300	Detected	Not Detected	Not Detected	Not Detected
Detection Rate		**90%**	**70%**	**46.7%**	**30%**

Based on the testing result, the images size and distance plays an important role on the rate of detections. The detection rate will decrease if the image sizes are decrease. Besides that the detection rate also decreases if the distance of the location from the Kinect is increase. The most effective image size is 640x480 because it has the highest detection rate and longest detection range compared to others.

3.3 Analysis of Image Size and Detection Time

Thirty data detection time are collected for each image in order to find the fastest or slowest detection time. The average time also has been calculated for each image sizes. Table 3 shows the results. Based on Table 3, the fastest detection time is for 240x240 image sizes and the slowest detection time is for images with 640x480 image sizes. This is because 640x480 images has a larger pixel and require a longer processing time.

Table 3 Results of image size with detection time

Image No	Detection Time (sec)			
	640 x 480	480 x 480	320 x 240	240 x 240
1	2.73	2.63	2.41	2.41
2	2.75	2.72	2.42	2.42
3	2.83	2.74	2.45	2.34
4	2.92	2.82	2.45	2.34
5	3.12	2.81	2.54	2.34
6	3.13	2.9	2.56	2.42
7	3.25	3	2.48	2.48
8	3.11	2.9	2.52	2.38
9	3.12	2.72	2.43	2.33
10	3.55	2.72	2.48	2.36
11	3.32	2.81	2.53	2.35
12	3.72	2.32	2.58	2.36
13	3.33	2.45	2.54	2.34
14	3.48	2.53	2.62	2.45
15	3.45	2.51	2.53	2.46
16	3.16	2.32	2.54	2.45
17	3.23	2.31	2.54	2.45
18	3.42	2.4	2.52	2.53
19	3.43	2.34	2.53	2.35
20	3.46	2.41	2.54	2.35
21	3.56	2.32	2.56	2.45
22	3.2	2.35	2.54	2.46
23	2.78	2.35	2.42	2.41
24	2.83	2.4	2.32	2.31
25	2.9	2.41	2.28	2.28
26	3.1	2.46	2.32	2.28
27	3.33	2.62	2.32	2.29
28	3.23	2.72	2.36	2.28
29	3.25	2.83	2.46	2.27
30	3.1	2.85	2.38	2.28
Average Detection Time	3.19	2.59	2.47	2.37

4.0 CONCLUSION

This research has been developed to identify location using high speed Kinect sensor. Kinect sensor was selected because of its capability of producing high speed vision. In order to test the effectiveness of the location identifications systems, several parameters are being considered such as distances, image sizes, and speed. Although 240x240 image has faster detection time, 640x480 image size was selected due to higher detection rate and longer detection range. The detection time with an average of 3.1 second are considered to be fast enough at this moment. The effective distance for the Kinect sensors is from 40cm to 300cm and the Kinect sensors should be in these distances in order to detect locations correctly.

Acknowledgement

This work is supported by the RAGS research grant from Ministry of Higher Education (no. 9018-00050).

References

[1] Tamura, Y., Takabatake, Y., Kashima, N. and Umetani, T. 2012. Localization System using Microsoft Kinect for Indoor Structures. The Japan Society of Plasma Science and Nuclear Fusion Research, Konan University, Japan, 7: 2406036-1- 2406036-4. DOI: 10.1585/pfr.7.2406036.

[2] Yii, W., Damayanthi, N., Drummond, T. and Li, W. H. 2014. Visual Localization of a Robot with an external RGBD Sensors. Monash University, Australia retrieved October 12, 2014 from https://ssl.linklings.net/conferences/acra/program/attendee_program_acra2011/includes/files/pap157.pdf.

[3] Correa, D. S. O., Sciotti, D. F., Prado, M. G., Sales, D. O., Wolf, D. F. and Osorio, F. S. 2012. Mobile Robots Navigation in Indoor Environments Using Kinect Sensor. *Second Brazilian Conference on Critical Embedded Systems*. 36-41. DOI 10.1109/CBSEC.2012.18.

[4] Rivera, L. A., DeSouza, G. N. and Franklin L. D. R. 2013. Control of a Wheelchair using an Adaptive Kmeans Clustering. *IEEE Symposium on Computational Intelligence in Rehabilitation and Assistive Technologies (CIRAT), University of Missouri, Columbia*. 24-31.

[5] Keat, H. W. and Ming, L. S. 2012. An Investigation of the Use of Kinect Sensor for Indoor Navigation. *IEEE Region 10 Conference (TENCON 2012), Monash University Sunway campus, Malaysia*. 1-5. DOI 10.1109/TENCON. 2012. 6412246.

[6] Khoshelham, K. 2011. Accuracy Analysis of Kinect Depth Data. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences (ISPRS Calgary 2011 Workshop), University of Twente, Canada*. 133-138.

[7] Tolgyessy, M. and Hubinsky, P. 2014. The Kinect Sensor in Robotics Education. Slovak University of Technology in Bratislava, Slovakia retrieved October, 12, 2014 from http://www.innoc.at/fileadmin/user_upload/_temp_/RiE/Proceedings/69.pdf.

[8] Ajay, A. and Venkataraman, D. 2013. A Survey on Sensing Methods and Features Extraction Algorithms for SLAM Problem. *International Journal of Computer Science, Engineering and Applications (IJCSEA)*. 3(1): 59-63. DOI: 10.5121/ijcsea.2013.3105.

[9] Ilias, B., Abdul Shukor, S. A., Yaacob, S., Adom, A. H. and Mohd Razali, M. H. 2014. A Nurse Following Robot with High Speed Kinect Sensor. *ARPN Journal of Engineering and Applied Sciences.* 9(12): 2454-2459. ISSN 1819-6608.

[10] Emter, T. and Stein, A. 2012. Simultaneous Localization and Mapping with the Kinect Sensor. *Proceedings of ROBOTIK, 7th German Conference on Robotics*. 1-6.

[11] Zhang, Z. 2012. Microsoft Kinect Sensor and Its Effect. Multimedia. *IEEE*. 19(2): 4-10. DOI:10.1109/MMUL.2012.24.

[12] Pradalier, C., and Sekhavat, S. 2002. Current Matching, Localization and Map Building using Invariant Features. Proceedings of the 2002 IEEE/RSJ. *Intl. Conference on Intelligent Robots and Systems, Switzerland*. 514-520.

[13] Andreopoulos, A. and Tsotsos, J. K. 2013. 50 Years of Object Recognition Directions Forward. *Computer Vision and Image Understanding.* 117(8): 827-891. DOI: 10.1016/j.cviu.2013.04.005.

[14] Hernandez-Lopez, J. J., Quintanilla-Olvera, A. L., Lopez-Ramırez, J. L., Rangel-Butanda, F. J., Ibarra-Manzano, M. A and Almanza-Ojeda, D. L. 2012. Detecting Objects Using Color and Depth Segmentation with Kinect Sensor. *The 2012 Iberoamerican Conference on Electronics Engineering and Computer Science*. 196-204. DOI:10.1016/j.protcy.2012.03.021.

[15] Biswas, J. and Veloso, M. 2012. Depth Camera Based Indoor Mobile Robot Localization and Navigation. *Proceedings of IEEE International Conference on Robotics and Automation.* 1697-1702.

[16] Cunha, J., Pedrosa, E., Cruz, C., Neves, A. J. R. and Lau, N. 2011. Using a Depth Camera for Indoor Localization and Navigation. Retrieved October, 12, 2014 from http://mobilerobotics.cs.washington.edu/rgbd-workshop-2011/camera_ready/cunha-rgbd11-localization.pdf.

[17] Hadjakos, A. and Darmstadt, T. 2012. Pianist Motion Capture with the Kinect Depth Data. *Proceedings of the 9th Sound and Music Computing Conference, Copenhagen, Denmark.* 303-310.