How to cite this paper:

Yasir Dawood Salman, Ku Ruhana Ku-Mahamud, & Eiji Kamioka. (2017). Distance measurement for self-driving cars using stereo camera in Zulikha, J. & N. H. Zakaria (Eds.), Proceedings of the 6th International Conference of Computing & Informatics (pp 235-242). Sintok: School of Computing.

DISTANCE MEASUREMENT FOR SELF-DRIVING CARS USING STEREO CAMERA

Yasir Dawood Salman[1], Ku Ruhana Ku-Mahamud[2], and Eiji Kamioka[3]

[1]*Universiti Utara Malaysia, Malaysia, yasir.dawod@gmail.com*
[2]*Universiti Utara Malaysia, Malaysia, ruhana@uum.edu.my*
[3] Shibaura Institute of Technology, *Japan, kamioka@shibaura-it.ac.jp*

ABSTRACT. Self-driving cars reduce human error and can accomplish various missions to help people in different fields. They have become one of the main interests in automotive research and development, both in the industry and academia. However, many challenges are encountered in dealing with distance measurement and cost, both in equipment and technique. The use of stereo camera to measure the distance of an object is convenient and popular for obstacle avoidance and navigation of autonomous vehicles. The calculation of distance considers angular distance, distance between cameras, and the pixel of the image. This study proposes a method that measures object distance based on trigonometry, that is, facing the self-driving car using image processing and stereo vision with high accuracy, low cost, and computational speed. The method achieves a high distance measuring accuracy of up to 20 m. It can be implemented in real time computing systems and can determine the safe driving distance between obstacles.

Keywords: distance measurement, self-driving car, stereo camera, image processing

INTRODUCTION

In the past two decades, self-driving cars have drawn considerable attention from both the academia and the industry with their potential applications in military, transportation, and industrial production. Self-driving cars are expected to perform various missions to replace humans in different fields (Li, Wang, Wang, & Zhang, 2015).

Self-driving cars technologies mostly involve computer systems that support human drivers by automating vehicle control parts. These technological parts possess a range of competences, from forward collision warning and antilock brakes to lane keeping and adaptive cruise control, to fully automate driving (Administration, 2013).

Vehicle automation has become one of the main trends in automotive research and development. Expectations are high in terms of increasing road safety and driving comfort by partially or completely relieving drivers of driving tasks (Hohm, Lotz, Fochler, Lueke, & Winner, 2014). Automating the responsibilities of drivers may reduce crashes and improve roadway efficiency significantly (Goodall, 2014).

Self-driving cars have several specific challenges compared to other automated robot types. These challenges result from the combination of dynamic environment, complex, and high-speed movement. The vehicle needs to detect the positions of other vehicles in any type of traffic, which involves object detection, relative speed estimation, and distance. These factors have to be determined for the vehicle to adjust its route and/or speed (Ilas, 2013).

Object distance measurement is becoming increasingly important in mobile autonomous systems (Hsu & Wang, 2015). Recognizing surroundings accurately and quickly is one of the most essential and challenging tasks for autonomous systems, such as self-driving cars. Two methods for measuring distance exist, namely, active and passive. The active method sends signals to the object to measure distance (Mrovlje & Vrancic, 2008; Walcher, 2014). Systems that use sensors, such as structured light, laser scanners, ultrasound, or time-of-flight, to search for objects (Häne, Sattler, & Pollefeys, 2015) are developed to increase awareness about the environment more than humans are capable of (Appiah & Bandaru, 2011). The downside of these systems is the cost involved in deploying them. Thus, the use of cameras and computer vision techniques in place of these systems aids in cost reduction.

In contrast, passive methods receive information on the object's position from passive measurements of the scene, such as camera images. Object distance measurement using image processing is one of the important research areas in computer vision (Alizadeh & Zeinali, 2013). The implication of extracting useful information from images and videos finds application in a variety of fields, such as robotics, remote sensing, virtual reality, and industrial automation (Appiah & Bandaru, 2011). The passive method has the advantages of working over a wide range of weather and lighting conditions, high resolution, and using low-cost cameras (Häne et al., 2015). The most popular method relies on stereoscopic measurement. This method uses two cameras. The object's distance is calculated from the relative difference of the object's position on both cameras (Carnicelli, 2005; Mrovlje & Vrancic, 2008).

Humans have the ability to roughly estimate the distance of objects because of the stereo vision of human eyes (Hsu & Wang, 2015). An improved stereo vision system is proposed to accurately measure the distance between the vehicle and the objects around it using two cameras. The use of two horizontal cameras allows the calculation of depth information in the viewing direction. The main idea for object detection is to classify points in the 3D space based on height, width, and traversable slope relative to the neighbouring points. The detected object points can be mapped into convenient projection planes for motion planning (Appiah & Bandaru, 2011).

The paper is structured as follows. The next section discusses related works. The third section explains the measurement methods. The conclusion and the future work are presented in the last section.

RELATED WORK

Autonomous vehicles should be able to perceive and discriminate objects in the environment accurately and reliably. Hence, numerous approaches have been presented in different application areas and scenarios in the past years using stereo vision or 2D/3D sensor technologies (Appiah & Bandaru, 2011).

Using a single camera for the sensor system is widely used method in the existing image-based 3D reconstruction (Fathi & Brilakis, 2014). This imposes a constraint on the generated results. The scene can only be reconstructed to an unknown scale factor if a single camera is used for image acquisition (Pollefeys et al., 2008). In infrastructure applications that require spatial data collection in the Euclidean space, this limitation is of great concern. The use of a calibrated stereo camera set addresses this problem (Fathi & Brilakis, 2014).

Stereo vision applications are depending on configuring two-camera in calibration setup, where each camera delivers two 2D representations for the same scene at the same time, as shown in Figure 1. Stereo vision is achieved through the extraction of 3D information by processing two or more 2D images for the same scene. The process of extracting 3D information creates a map that describes what point in the 2D images corresponds to the same point in the 3D scene (Calin & Roda, 2007).

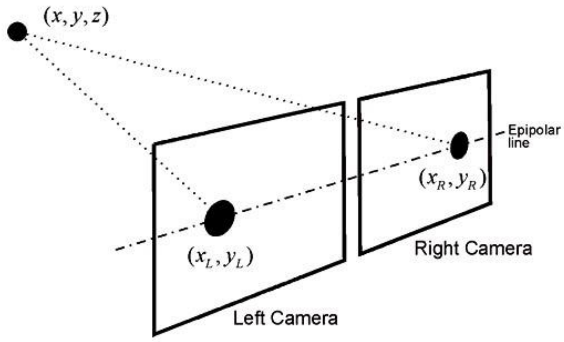

Figure 1. Typical Stereo implementations.

Recent research used multiple vision sensors for object distance and size measurements (Hsu & Wang, 2015). Kuhnert and Stommel (2006) proposed a new method for the reconstruction of the environment of mobile robots based on the fusion of the sensor outputs of a stereo camera and a PMD-camera. This shows that each technique balances the disparities of other techniques while preserving its characteristic benefits. As a result, dense depth maps with reliable values on both homogeneous regions and precise and robust values on the edges are achieved.

Santoro, AlRegib, and Altunbasak (2012) presented a correction method for reducing depth errors resulting from camera shift. Their method uses a real-time motion estimation approach to correct the alignment errors between the stereo cameras. The natural disparity between stereo views is incorporated into a constrained motion estimation framework. Their method accurately estimates synthetic misalignments due to translation, rotation, scaling, and perspective transformation.

A study on the visual ego-motion estimation algorithm for a self-driving car equipped with multi-camera system was conducted by Lee, Faundorfer, and Pollefeys (2013). They showed that a 2-point minimal solution of the generalized essential matrix could obtain full relative motion, including the metric scale, by modeling the multi-camera system as a generalized camera. They verified the validity of their assumptions of the motion model by comparing their results with a real-world data set collected with a multi-camera system mounted on a car.

Mahammed, Melhum, and Kochery (2013) focused on finding the distance of an object using two webcams as sensing elements, which resulted in low accuracy. Image processing is used in template matching, and Matlab code is written to execute the distance measurements. However, their approach contains a major limitation because it can only be used under a certain condition and for short distances (less than 1.2 m).

Although numerous successful works related to object distance measurement exists, the underlying distance calculation formulae omits the variation in the image distance to the lenses. A revised method using stereo vision for distance measurement is proposed.

STEREO IMAGE MEASUREMENT METHOD

Stereo vision is a technique used for recording and representing stereoscopic (3D) images. A pair of cameras is required to implement stereo vision. An illusion of depth is created by using two pictures taken at slightly different positions. Stereo image is captured using a Nikon 5300 DLSR camera with EF 50-mm f/1.8L lens and a horizontal angle of view in DX-format equal to 31.5°. The size of the image sensor is 23.5 mm 15.6 mm. The lens is primed with a fixed zoom to 50 mm. The cameras are aligned in a parallel stand with a fixed distance. The object's distance is measured when it enters the overlapping views of the two cameras. Figure 2 illustrates the stereo vision setup.

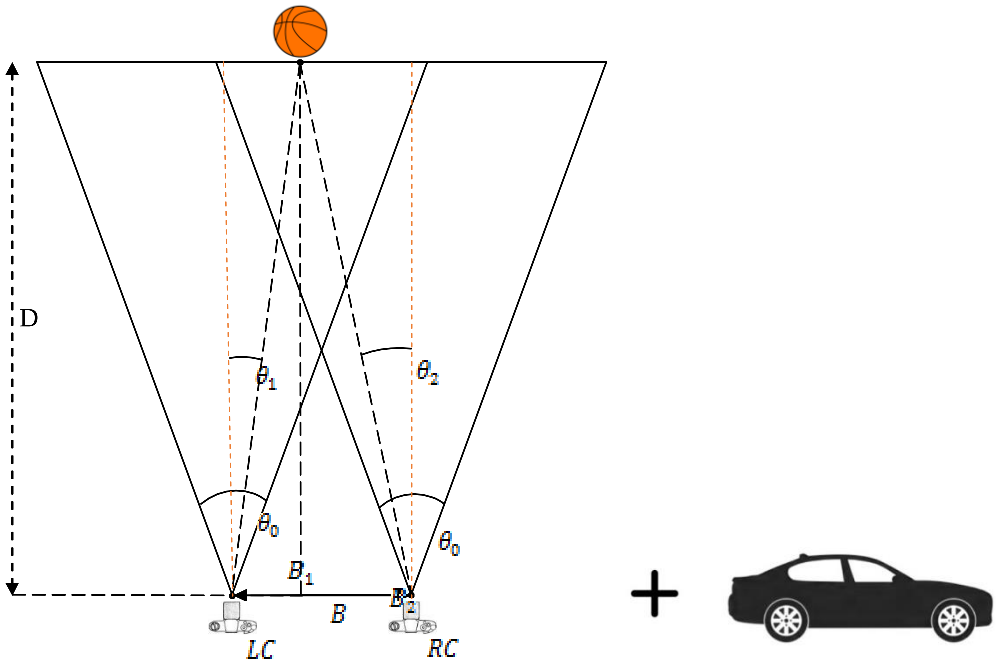

Figure 2. The stereo camera for self-driving car setup.

Image pre-processing is an important and common method in computer vision systems. Image pre-processing enhances image quality and improves computational efficiency (Hsu & Wang, 2015). In this work, after capturing the stereo images, the image resolution is downscaled to improve the computational speed. For example, the original resolution of 6000 × 4000 pixels is downscaled to 600 × 400 pixels. According to Hsu and Wang (2015), the reduction of resolution to a certain level does not affect the accuracy of the system. The conversion of images from RGB color to grayscale is another way to improve computational speed. The information that may be lost in the grayscale color space has insignificant effect on distance measurement accuracy (Appiah & Bandaru, 2011; Hsu & Wang, 2015). Therefore, the images were converted to grayscale in this work.

The two cameras were positioned on horizontal base with a distance of 15 cm from each other. The experiment started by allowing the right camera (RC) to take the first picture and left camera (LC) to take the second picture, as depicted in Figure 2. B represents the distance between the two cameras and θ_0 is the camera's horizontal angle of view. The object's position (distance D) can be calculated using the following geometrical derivations (Mrovlje & Vrancic, 2008) as follow:

The distance B is expressed as the sum of distances B_1 and B_2 , as shown in Figure 2. Since $B_1 = D\tan\theta_1$ and $B_2 = D\tan\theta_2$ as shown in Figure 3, then

$$D = \frac{B}{\tan\theta_1 + \tan\theta_2} \tag{1}$$

using Figure 3 and trigonometry

$$D = \frac{\tan\theta_1}{x_1} = \frac{\tan\left(\frac{\theta_0}{2}\right)}{\frac{x_0}{2}} \quad \text{in LC view and} \quad D = \frac{\tan\theta_2}{-x_2} = \frac{\tan\left(\frac{\theta_0}{2}\right)}{\frac{x_0}{2}} \quad \text{in RC view}$$

then

$$\frac{x_1}{\frac{x_0}{2}} = \frac{\tan\theta_1}{\tan\left(\frac{\theta_0}{2}\right)}, \frac{-x_2}{\frac{x_0}{2}} = \frac{\tan\theta_2}{\tan\left(\frac{\theta_0}{2}\right)} \quad \rightarrow \tan\theta_1 = \frac{2x_1\tan\left(\frac{\theta_0}{2}\right)}{x_0}, \tan\theta_2 = \frac{-2x_2\tan\left(\frac{\theta_0}{2}\right)}{x_0}$$

$$\tan\theta_1 + \tan\theta_2 = \frac{2\tan\left(\frac{\theta_0}{2}\right)(x_1 - x_2)}{x_0}$$

by implementing in Eq. (1), distance D can be measured as follows:

$$D = \frac{Bx_0}{2\tan\left(\frac{\theta_0}{2}\right)(x_1 - x_2)} \tag{2}$$

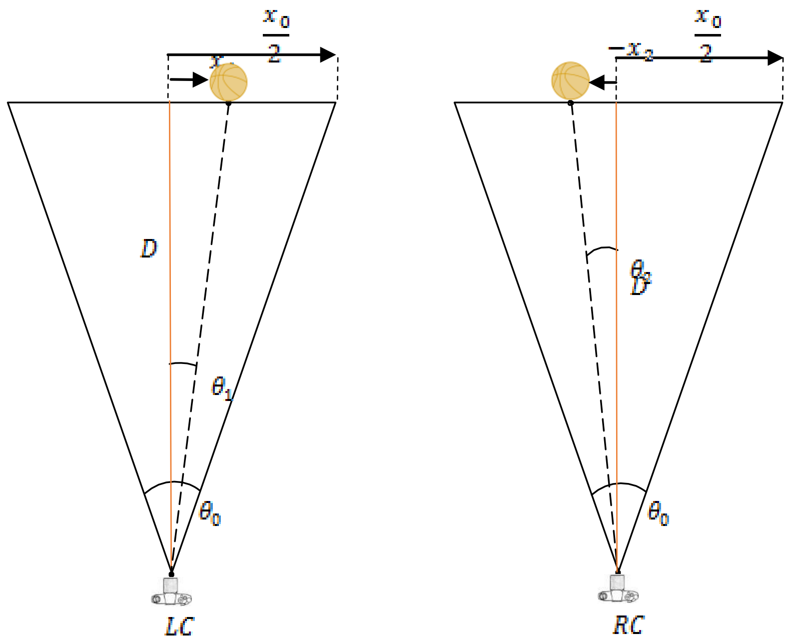

Figure 3. Position of the objects and the cameras.

To find the distance D of the targets' objects, both cameras obtained two images that contain the targets for the same scene. The resolutions of the two images were decreased to 600 × 400 pixels and converted to grayscale from RGB. The disparity $(x_1 - x_2)$ was calculated from the difference between the target object's coordinates in the x-axis for the right and left camera images.

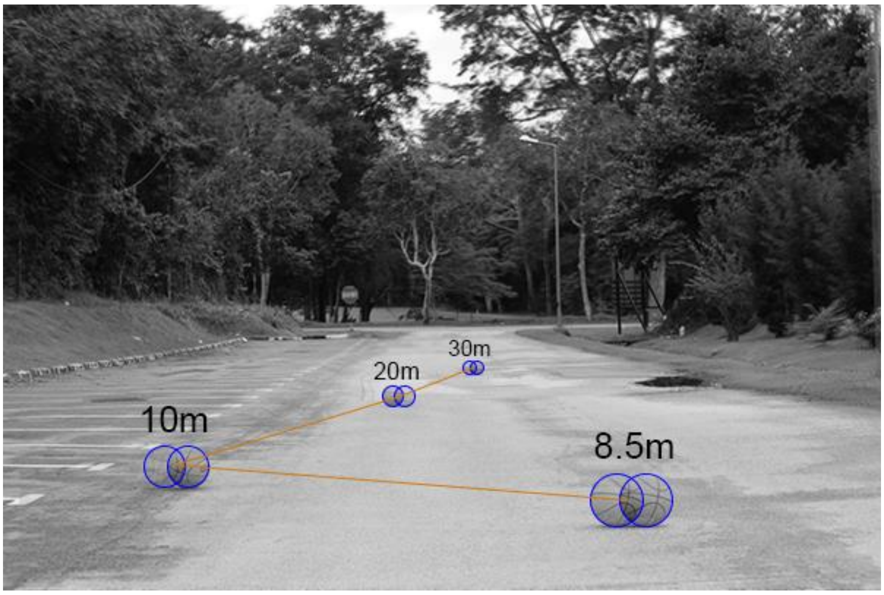

Figure 4. The baseball distance measurement.

To evaluate the performance of the proposed method, four tests were conceded to find the distance of a target object. The selected target in these tests was a baseball, as shown in Figure 4. In all four tests, the blue circle indicates the position of the ball in each image. Three more experiments were conducted to determine the farthest distance that can be detected by the proposed method, and the result showed that it could detect up to 160 m. In the experiment, $B = 0.15$ is the 15-cm lens distance setup, $x_0 = 600$ is the pixel width, and $\theta_0 = 31.5°$ is the lens view angle. Given these, the experiment was conducted to calculate the distance of the object, as shown in Figure 4. Figure 5 shows the accuracy of the calculated distance results compared to the measured distance. Correspondingly, Table 1 shows the real distance, calculated distance, and the error percentage of the results. It shows that this method provides a high detection accuracy of up to 20 m. Furthermore, the experimental results showed that this method achieves a detection accuracy while running in real-time that is sufficient for practical applications.

Table 1: Real and calculated object distances in left and right images.

Test	Real distance	Calculated distance	Error %
1	8.5	8.5	0
2	10	10	0
3	20	20	0
4	30	32	6.25
5	50	53.3	6.19
6	130	133.3	2.47
7	160	160	0
		Total	2.13

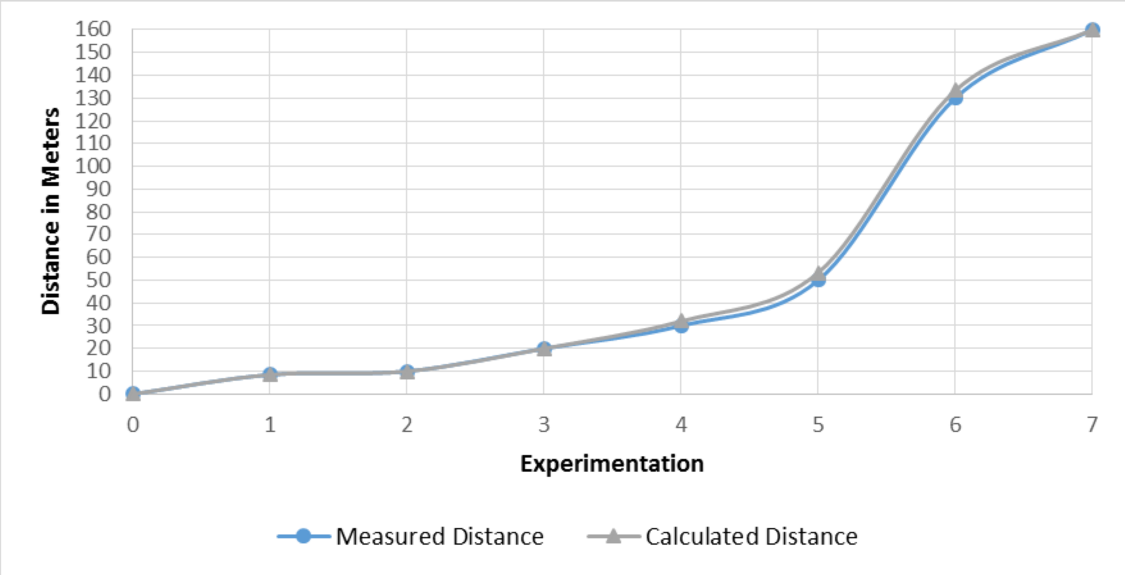

Figure 5. Objects measured distance and calculated distance.

According to Murray and Little (2000), stereo vision mapping is very sensitive to errors because the process of collapsing the data from 3D to 2D introduces errors in the form of points to be propagated into the map. Furthermore, the distance of the object varies inversely with the disparity and the accuracy of distance measurements depends on the resolution of the camera pictures, the optical properties of the lens, and the separation between the optical axes of both cameras (Mahammed et al., 2013). Therefore, the use of compressed images resulted in accrued error, and the measurement accuracy decreased to only one pixel because x_1 and x_2 were measured using pixels. Therefore, the results are true only for objects within 160 m from the location of the car/camera.

CONCLUSION

The proposed method for object distance measurement in self-driving cars designed to rely purely on stereo cameras provide good results. This method is able to aid in measuring the distance between the cars and the objects to determine the safe driving distance.

The method developed in this work uses only two cameras to capture the front obstacles because the purpose of using stereo vision is to replace the existing active method and perception systems. For future work, the same method can be implemented using more cameras or 360° cameras to detect obstacles in more directions.

ACKNOWLEDGMENTS

The authors wish to thank the Ministry of Higher Education Malaysia for funding this study under the Long Term Research Grant Scheme (LRGS/b-u/2012/UUM/Teknologi Komunikasi dan Informasi). The study was also supported by JASSO grant under the collaboration between SIT and UUM.

REFERENCES

Administration, N. H. T. S. (2013). Preliminary statement of policy concerning automated vehicles. *Washington, DC*, 1-14.

Alizadeh, P., & Zeinali, M. (2013). A Real-Time Object Distance Measurement Using A Monocular Camera.

Appiah, N., & Bandaru, N. (2011). Obstacle detection using stereo vision for self-driving cars.

Calin, G., & Roda, V. (2007). Real-time disparity map extraction in a dual head stereo vision system. *Latin American applied research, 37*(1), 21-24.

Carnicelli, J. (2005). Stereo vision: measuring object distance using pixel offset.

Fathi, H., & Brilakis, I. (2014). Multistep explicit stereo camera calibration approach to improve euclidean accuracy of large-scale 3D reconstruction. *Journal of Computing in Civil Engineering, 30*(1), 04014120.

Goodall, N. (2014). Ethical decision making during automated vehicle crashes. *Transportation Research Record: Journal of the Transportation Research Board*(2424), 58-65.

Häne, C., Sattler, T., & Pollefeys, M. (2015). *Obstacle detection for self-driving cars using only monocular cameras and wheel odometry.* Paper presented at the Intelligent Robots and Systems (IROS), 2015 IEEE/RSJ International Conference on.

Hohm, A., Lotz, F., Fochler, O., Lueke, S., & Winner, H. (2014). Automated Driving in Real Traffic: from Current Technical Approaches towards Architectural Perspectives: SAE Technical Paper.

Hsu, T.-S., & Wang, T.-C. (2015). An Improvement Stereo Vision Images Processing for Object Distance Measurement. *International Journal of Automation and Smart Technology, 5*(2), 85-90.

Ilas, C. (2013). *Electronic sensing technologies for autonomous ground vehicles: A review.* Paper presented at the 2013 8th International Symposium on Advanced Topics in Electrical Engineering (ATEE).

Kuhnert, K.-D., & Stommel, M. (2006). *Fusion of stereo-camera and pmd-camera data for real-time suited precise 3d environment reconstruction.* Paper presented at the 2006 IEEE/RSJ International Conference on Intelligent Robots and Systems.

Lee, G. H., Faundorfer, F., & Pollefeys, M. (2013). *Motion estimation for self-driving cars with a generalized camera.* Paper presented at the Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition.

Li, C., Wang, J., Wang, X., & Zhang, Y. (2015). *A model based path planning algorithm for self-driving cars in dynamic environment.* Paper presented at the Chinese Automation Congress (CAC), 2015.

Mahammed, M. A., Melhum, A. I., & Kochery, F. A. (2013). Object distance measurement by stereo vision. *International Journal of Science and Applied Information Technology (IJSAIT), 2*(2), 05-08.

Mrovlje, J., & Vrancic, D. (2008). *Distance measuring based on stereoscopic pictures.* Paper presented at the 9th International PhD Workshop on Systems and Control: Young Generation Viewpoint.

Murray, D., & Little, J. J. (2000). Using real-time stereo vision for mobile robot navigation. *Autonomous Robots, 8*(2), 161-171.

Pollefeys, M., Nistér, D., Frahm, J.-M., Akbarzadeh, A., Mordohai, P., Clipp, B., . . . Merrell, P. (2008). Detailed real-time urban 3d reconstruction from video. *International Journal of Computer Vision, 78*(2-3), 143-167.

Santoro, M., AlRegib, G., & Altunbasak, Y. (2012). *Misalignment correction for depth estimation using stereoscopic 3-d cameras.* Paper presented at the Multimedia Signal Processing (MMSP), 2012 IEEE 14th International Workshop on.

Walcher, H. (2014). *Position sensing: angle and distance measurement for engineers*: Elsevier.