INVESTIGATING THE EFFECTS OF STEREO CAMERA BASELINE ON THE ACCURACY OF 3D PROJECTION FOR INDUSTRIAL ROBOTIC APPLICATIONS

Wutthigrai Boonsuk, Eastern Illinois University

Abstract

Stereo cameras have been used in robotics for decades. With recent advances in sensor technology, they have become one of the most widely used sensors for automation in robotic applications. Fulfilling the need for rapid detection and recognition of complex environments, stereo cameras allow robots to quickly sense an environment by reconstructing 3D scenes from disparity maps. However, typical stereo cameras have a fixed baseline, which lessens the quality of disparity maps when objects get too close to the cameras. The baseline of a stereo camera needs to be adjusted to accommodate the minimum range between the cameras and the object. As the accuracy of 3D scenes can be essential for certain robotic tasks, such as welding, riveting, measuring, and assembly, the focus of this current study was on investigating the effects of different stereo camera baselines on the accuracy of 3D projections generated from disparity maps. The results showed a correlation between stereo camera baseline and valid surface areas of the target object. This finding can be useful for researchers wanting to design and develop an effective stereo camera system and improve the quality of its 3D projection. Although this current study focused on a stereo vision system for industrial robotic applications, the results may also be applied to other robotic applications such as navigational robots and autonomous vehicles.

Introduction

Although stereo cameras have been utilized in the robotics field for decades, only recently have they become one of the most widely used sensors for automated robotics applications. Stereo cameras allow robots to sense their surroundings by reconstructing 3D scenes of the environment from disparity maps. Despite their wide use, however, typical stereo cameras have a fixed baseline—the distance between the two cameras—so disparity maps become distorted when objects get too close to the cameras. Figure 1 illustrates a typical configuration of a stereo camera system with a fixed baseline. In order to reduce this problem, the baseline of the stereo cameras needs to be shortened in order to accommodate the minimum range between the cameras and the object. In this study, the author investigated the effects of different stereo camera baselines on the accuracy of 3D reconstruction of scenes from disparity maps. This information is crucial for finding the optimal baseline of a stereo camera system when a fixed baseline is used, and as a first step towards the development of a flexible baseline stereo camera system.

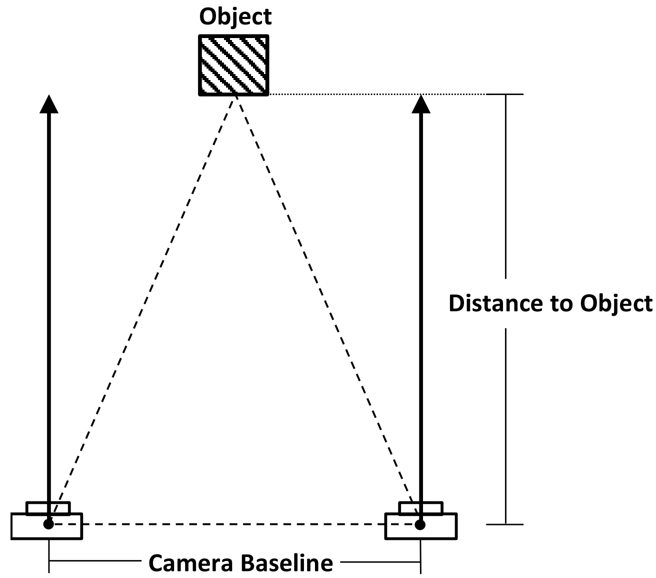

Figure 1. Stereo Camera System with a Fixed Baseline

The fundamental idea behind this study stems from the active stereo vision approach in robotics applications. This approach is a framework of stereo vision that manipulates camera parameters, lighting conditions, and motor controllers [1-3]. Conventional stereo cameras for robotic arms have a fixed baseline length [4], which can cause problems when the object gets too close to the cameras. Multi-baseline stereo cameras have been introduced in earlier research as superior to a set of stereo cameras with different baseline lengths [5]. By using a multi-baseline and multi-resolution approach [6], the depth error can be kept constant because the camera baseline and resolution can be varied in proportion to the depth. Rather than capturing a series of images from stereo cameras, another approach [7] is a stereo system that utilizes a high-speed slider to adjust the baseline length of the stereo camera. A similar method for constructing a stereo camera system was utilized in this present study. In a previous study by Rovira-Mas et al. [8], the authors identified the best combination of baseline length and focal length lenses suitable for agricultural robotic vehicles with a working range of 6-12 meters; however, past studies have not examined the relationship between baseline adjustment and the accuracy of the 3D projection from the cameras, and have tended to pick the best combination from the pool. The purpose of this current study was to extend prior work by investigating the relationship between baseline adjustment in stereo cameras and the accuracy of the 3D projection from the camera. The findings are intended for use in industrial robotic arm applications.

In this study, the author analyzed the relationship between three parameters: 1) baseline of the stereo camera system, 2) distance to object, and 3) accuracy of the 3D scenes constructed from the disparity map. The accuracy of the disparity map is essential for building a high-quality 3D surface, particularly for automated robotic tasks that require high precision, such as welding, riveting, measuring, and assembly tasks. Thus, a set of experiments was designed in this study to identify how different stereo camera baselines may influence the disparity map when the distance to the object varies.

Methodology

In this study, a custom-built stereo camera rig consisting of two Raspberry Pi cameras (5 megapixels; focal length = 3.6 mm) was used. Figure 2 shows these cameras connected to two Raspberry Pi boards (model B+), mounted on linear motor screws that allow the cameras to be separated from 30-120 mm (1.1-4.7 inches). An Arduino Pro Mini (microcontroller board) was used to control the rotation of the motors in order to adjust the baseline. Stereo video signals were wirelessly transmitted via Wi-Fi. On a desktop computer, OpenCV and Point Cloud Library (PCL) were used to analyze real-time video signals and develop disparity maps using the Semi-Global Block Matching (SGBM) algorithm. A series of experiments was conducted in the laboratory. Figure 3 shows that the centers of the cameras were separated (i.e., baseline) at 15 mm increments (30, 45, 60, 75, 90, 105, and 120 mm).

This camera system was planned to be implemented on a FANUC LR Mate 220iC, which has a maximum arm reach of approximately 700 mm (27 inches). Thus, Figure 4 illustrates that the object was first placed at 700 mm from the cameras and moved closer in increments of 100 mm, up to 250 mm, which was the minimum arm reach of this robot. Figure 5 shows the target object, a 30x30 mm plastic cube with a 20-mm diameter hole. This target has typically been used for a robot gripping test in the lab and to test the accuracy of the disparity map. The disparity map was generated for each pair of baseline and object distances. The 3D scene was then reconstructed and analyzed for its accuracy. The higher the surface area that the camera picked up, the more effective the baseline would be at that specific distance to the object.

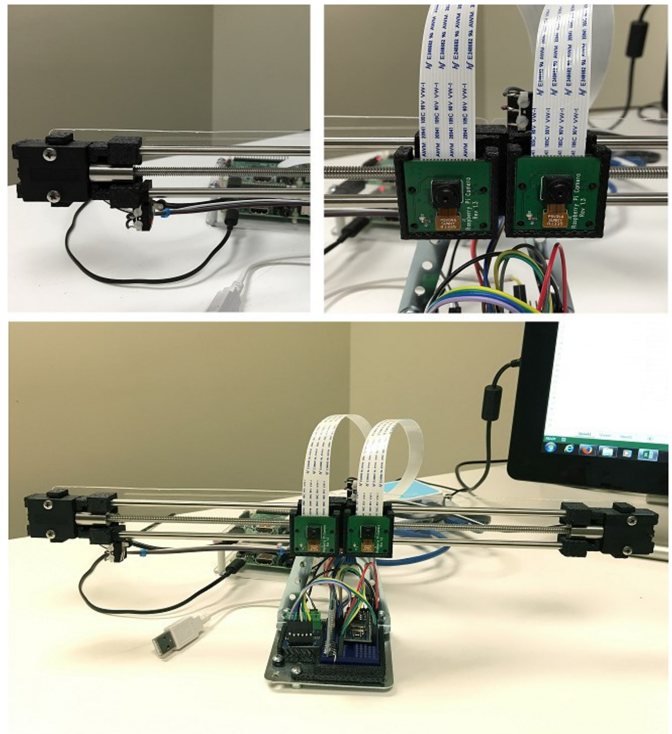

Figure 2. Custom-built Stereo Camera System

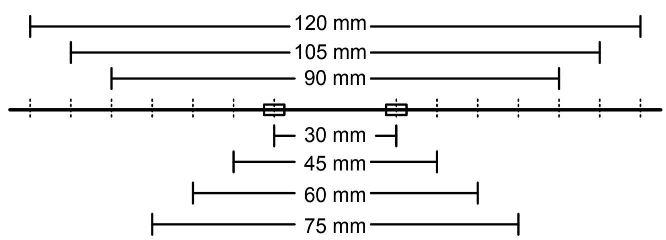

Figure 3. Baselines Used in the Experiments

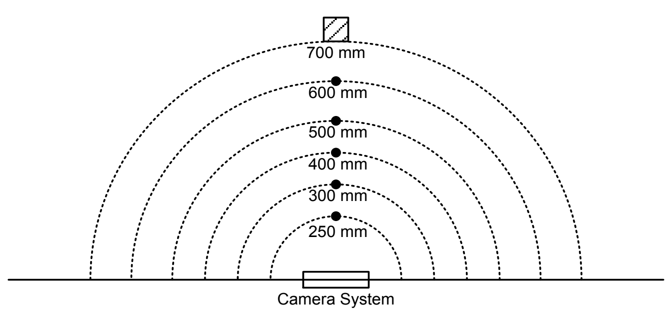

Figure 4. Distance to Target Object from the Camera System

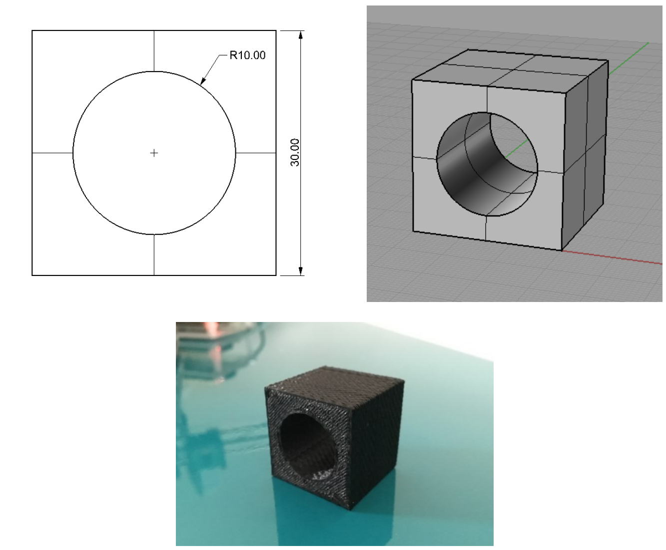

Figure 5. Target Object

Experiments and Results

Figure 6 shows the start of the experiments in which the stereo camera system was calibrated using a checkerboard. After calibration, camera matrices were generated and stored as text files that later were used to compute disparity maps and generate surface point clouds of the target object. Figure 7 shows how, for each baseline setting, the target object was moved from 700 to 250 mm. A recalibration was performed for each baseline adjustment in order to obtain proper disparity maps. A pair of images from the stereo camera system was then recorded for generating the disparity map and point cloud.

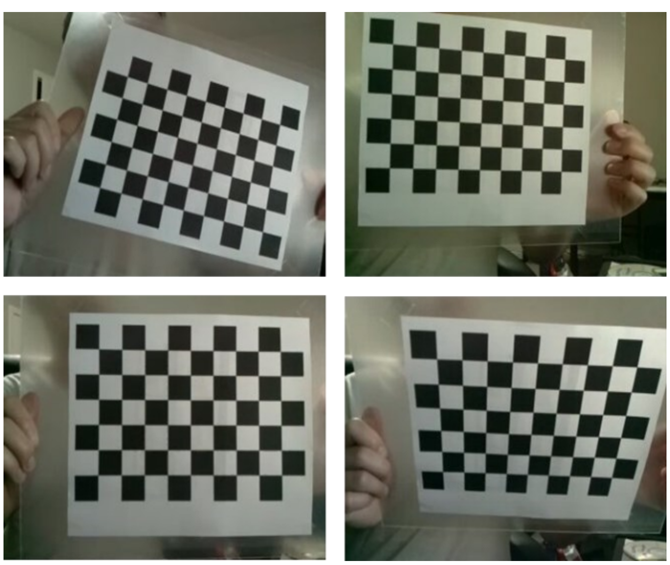

Figure 6. Camera Calibration Process

The position of the target and camera baseline significantly affected the quality of the disparity map. At baseline lengths greater than or equal to 90 mm, disparity maps from the stereo camera generated extreme surface errors at all target distances, implying that baseline lengths of 90 mm and wider were unsuitable for detecting objects in the range under 700 mm. Therefore, the results of this study were discussed based on baselines of 30, 45, 60, and 75 mm in the settings.

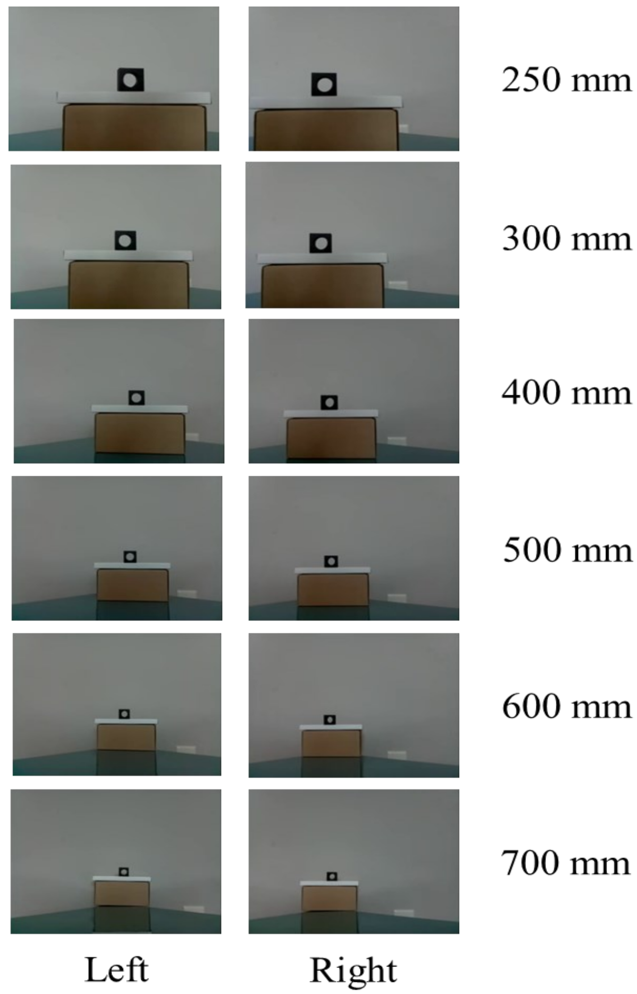

Figure 7. Stereo Images for a Baseline of 45 mm

The three-dimensional surfaces of the target were subsequently generated from the point clouds. Post-processing procedures were applied to the point clouds in order to remove noise and irrelevant points from the scene. Figure 8 illustrates an example of a point cloud after post-processing procedures. Surface areas (in square millimeters) were calculated to assess the ability of the stereo camera to pick up the surface of a target object when different baselines were used.

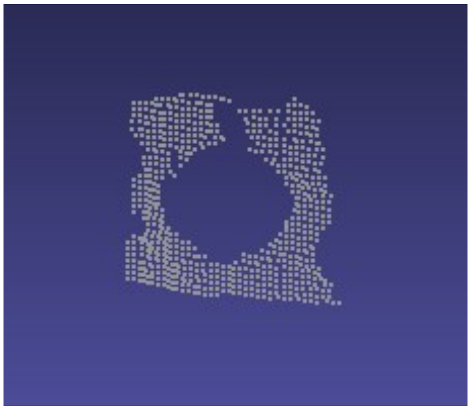

Figure 8. Example of the Point Cloud of a Target Object

Figure 9 and Table 1 present the results. Figure 9 displays a stereo camera with a 30-mm baseline that was able to pick up the most surface areas of the target object at distances of 250 and 300 mm. At 400 mm distance, the 45-mm baseline captured more surfaces than other lengths. At 500 and 600 mm distances, cameras with baselines of 45, 60, and 75 mm had very similar results, but were all better than the 30-mm baseline. At 700 mm, all baselines yielded similar results.

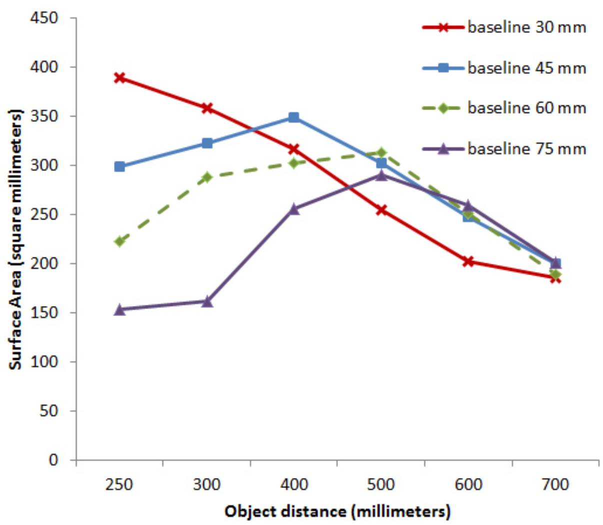

Figure 9. Chart of Surface Areas of the Target

Table 1. Surface Areas of the Target at Different Baselines

Object Distance	Surface Areas (Square Millimeters)			
	30 mm	45 mm	60 mm	75 mm
250 mm	389.32	298.61	222.57	153.36
300 mm	357.21	322.54	287.54	161.24
400 mm	316.47	347.93	301.26	254.87
500 mm	254.32	301.59	312.25	289.65
600 mm	201.97	247.63	251.27	259.31
700 mm	185.68	199.87	188.63	201.23

Discussion

Results from the experiments showed that stereo camera baselines could affect the 3D projection of a target object at different distances. In general, shorter baselines performed better when the object was at shorter distances, whereas longer baselines tended to perform better when the object was at greater distances. In addition, each baseline seemed to have an optimal point where it could perform best at a particular distance. Several factors might need to be considered before implementing this system, including calibration parameters, object size, and lighting conditions. Obtaining optimal calibration parameters for a camera system is challenging. These parameters appeared to be unique for each baseline. Future experiments can extend this study by identifying the relationship between baseline and those calibration parameters, as well as creating an automatic system that can adjust without the need for recalibration each time. In this study, the size of the target object remained constant. Using various sizes of objects may help identify the effectiveness of the 3D projection for larger baselines. Finally, lighting conditions can influence the performance of the camera system, especially when the target object is placed further from the camera. In this study, the lighting conditions were held constant across all experiments. Future research can manipulate the lighting conditions to confirm the results of the study.

Conclusions

In summary, the author of this current study implemented a preliminary stereo vision system for industrial robotics. The results suggest that adjusting the stereo camera baseline can influence the accuracy of 3D projection from disparity maps. An optimal stereo system should have a dynamic baseline that varies depending on the distance between the cameras and the target object.

References

[1] Hogue, A. & Jenkin, M. R. (2014). Active Stereo Vision. *Computer Vision: A Reference Guide*, Springer US, 8-12.

[2] Klarquist, W. N., & Bovik, A. C. (1998). Fovea: A foveated vergent active stereo vision system for dynamic three-dimensional scene recovery. Robotics and Automation. *IEEE Transactions, 14*(5), 755-770.

[3] Murray, D. W., McLauchlan, P. F., Reid, I. D., & Sharkey, P. M. (1993). Reactions to peripheral image motion using a head/eye platform. *Proceedings of International Conference on Computer Vision*, (pp. 403-411). Los Alamitos, CA.

[4] Saxena, A., Driemeyer, J., & Ng, A. (2008). Robotic grasping of novel objects using vision. *The International Journal of Robotics Research, 27*(2), 157-173.

[5] Okutomi, M., & Kanade, T. (1993). A Multiple-Baseline Stereo. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 15*(4), 353-363.

[6] Gallup, D., Frahm, J. M., Mordohai, P., & Pollefeys, M. (2008). Variable baseline/resolution stereo. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, (pp. 1-8). Anchorage, Alaska.

[7] Nakabo, Y., Mukai, T., Hattori, Y., Takeuchi, Y., & Ohnishi, N. (2005). Variable baseline stereo tracking vision system using high-speed linear slider. *Proceedings of IEEE International Conference on Robotics and Automation*, (pp. 1567-1572). Barcelona, Spain.

[8] Rovira-Mas, F., Wang, Q., & Zhang, Q. (2010). Design parameters for adjusting the visual field of binocular stereo cameras. *Biosystems Engineering, 105* (1), 59-70.

Biography

WUTTHIGRAI BOONSUK is an assistant professor of Applied Engineering and Technology at Eastern Illinois University. He earned his MS and PhD degrees in Industrial and Manufacturing System Engineering from Iowa State University. Dr. Boonsuk also received his second MS degree in Human Computer Interaction from Iowa State University. His research interests include 3D stereoscopic applications, manufacturing systems, rapid prototyping, robotic and controller systems, virtual reality, and geographic information systems (GIS). Dr. Boonsuk may be reached at wboonsuk@eiu.edu