Research Paper

International Journal of Research and Review
[Indexed, Peer-Reviewed Journal]
Volume 12; Issue: 5; May 2025
Website: www.ijrrjournal.com
E-ISSN: 2349-9788; P-ISSN: 2454-2237

Optimising Baseline Distance in Stereo Cameras: An Experimental Approach to Enhance Object Distance Accuracy

Fariz Imam[1], Arief Setyanto[2], Kusnawi[3]

[1,2,3]Master In Informatics Technology Business Intelligence Concentration, Universitas Amikom Yogyakarta. Indonesia

Corresponding Author: Fariz Imam

DOI: https://doi.org/10.52403/ijrr.20250512

ABSTRACT

This research explores how the variation of baseline distance influences the accuracy of object distance estimation in stereo camera systems based on a self-collected stereo image dataset. The dataset contains stereo images of people recorded during a laboratory session where participants were placed at distances between 2 and 50 meters, while the baseline distances could be set between 25 cm to 275 cm. While this dataset offers optimal and stringent evaluation conditions, it suffers from the notable drawback of being restricted to a single object category (human subjects) and relatively low levels of dynamism for the surrounding environment. A quantitative experimental methodology was implemented that included object detection by Mask R-CNN and depth estimation through disparity maps using Stereo Semi-Global Block Matching (SGBM). Results indicated a clear correlation between baseline distance and measurement accuracy, with the 75 cm baseline achieving optimal performance (MAE: 0.23 m; RMSE: 0.34 m). Conversely, a shorter baseline (25 cm) resulted in significant errors, especially for distant objects (MAE: Measurements taken using baselines longer than 175 cm showed increased errors when measuring short distances. The RMSE value at a 75 cm baseline was measured at 0.34 m, which corresponds to only 0.68% of the maximum tested distance (50 m). This result highlights the effectiveness of the experimental design and approach, as it is significantly lower than the commonly recognized RMSE threshold for stereo-based distance estimation. Nevertheless, additional validation using more different data records and more complex environmental conditions is extremely important to improve the reliability and applicability of these results for actual use exceeding the limits of current data records.

Keywords: *stereo camera, baseline, object distance, mask R-CNN, disparity map*

INTRODUCTION

Stereo camera systems employ two cameras aligned in parallel to collect images from distinct perspectives. This method uses disparity (the positional difference of an object in two photographs) to ascertain the depth of an object in three-dimensional space (Kardovskyi and Moon 2021). Stereo cameras determine object distances by analyzing disparity, or the difference in object position across images captured from two perspectives. Traditional stereo vision employs two cameras aligned in parallel to record images from distinct perspectives. The stereo vision system can produce depth maps and precisely assess object distances by utilizing the variances in viewpoint between the two cameras. This work proposes a flexible setup wherein the two cameras are not fixed and may be positioned flexibly, enabling baseline variations for diverse purposes such as mapping and remote monitoring (Sumetheeprasit et al. 2023). In some cases, stereo cameras are positioned with their optical axes tilted inward (toe-in), where these axes converge (Lin, Tsai, and Tran 2021). Stereo vision relies on the principle of stereopsis, wherein multiple cameras capture images from distinct perspectives to provide visual parallax, facilitating the computation of an object's depth. The sensor rotation-based configuration generates a stereo effect without the necessity of two distinct cameras, although the fundamental method persists, specifically acquiring images from varying perspectives. This arrangement offers greater error tolerance in distance estimation than conventional in-line stereo setups. The most effective distance measurement devices often necessitate supplementary 3D data from stereo cameras, active depth sensors, or laser scanners, which inevitably raises the overall system cost (Zhang et al. 2021).

In recent decades, stereo cameras have become a significant technology in robotics, autonomous vehicles, and augmented reality applications. Stereo cameras are among the most prevalent range sensors, capable of delivering precise distance measurements and adaptable to various platforms (Seo, Park, and Choi 2022). Stereo cameras obtain depth information by employing two images captured from distinct perspectives, enabling the system to ascertain the distance of objects in three-dimensional space. This technology is crucial for numerous applications, such as robotic navigation and accurate distance measurement in autonomous cars. Cameras are the most accessible and often utilised sensor devices in autonomous vehicle applications. Precisely quantifying the distance between cars constitutes a significant and intricate difficulty in image processing. This technology is crucial for various systems, including Driving Safety Support Systems (DSSS), autonomous vehicles, and traffic mobility management solutions (Zaarane et al. 2020). Stereo cameras are especially beneficial for visual Simultaneous Localisation and Mapping (SLAM) because they can acquire extensive depth information from visual landmarks, resulting in more accurate robot posture estimates (Raoui and Amraoui 2024). Nonetheless, the capacity of vision systems to operate well under poor settings continues to pose a considerable problem that remains inadequately addressed (Gehrig et al. 2021). The ability to accurately assess distances is crucial for real-time navigation and decision-making in these contexts. Current research predominantly focuses on algorithm development, neglecting the influence of baseline distance on the precision of object distance estimate. This study aims to investigate the influence of variations in stereo camera baselines on the precision of object distance measurements. This study seeks to provide practical guidelines for creating appropriate baselines in various visual technology applications requiring accurate distance measurements.

Distance estimation in stereo cameras is accomplished through triangulation techniques, which rely on the disparity between two images captured from distinct perspectives by two cameras. In stereo imaging, triangulation is employed to calculate the depth of points in a 3D scene (Haider and Hel-Or 2022). Data from both cameras is utilized for depth triangulation, which also serves to mitigate uncertainty in determining the position of visual landmark (Sabato, Valente, and Niezrecki 2020). Several factors, including the baseline distance between the cameras, significantly affect distance estimation accuracy in stereo cameras. The baseline, a critical parameter in stereo systems, directly influences distance measurement capabilities, with a more extended baseline enabling the system to measure greater distances (Seo, Park, and Choi 2022). A hierarchical approach to stereo matching is employed, beginning with coarse matching using macro pixels and progressing to fine matching at the pixel level, thereby reducing matching ambiguities (Liu and Hou 2024). For flexible baseline adjustments, a more comprehensive baseline is used for distant objects, while a narrower baseline is applied for closer objects to minimize estimation errors caused by occlusion (Sumetheeprasit et al. 2023).

Selecting the proper baseline distance between two stereo cameras can significantly improve the accuracy of object distance estimation. Measurement distance positively correlates with measurement errors, while baseline length is negatively correlated with measurement errors (W. Wang et al. 2024). Experiments using real-time images of stereo vision systems, taken from various obstacles at different distances, showed that calculating the distance to the obstacle was entirely accurate (Adil, Mikou, and Mouhsen 2022). For stereo vision on more distant targets, a more considerable baseline distance can be used, especially in vertical stereo configurations, when a suitable horizontal configuration is unavailable and causes a loss of horizontal overlap (Sumetheeprasit et al. 2023). The use of deep learning models can increase the accuracy of distance measurements in stereo-vision systems (Z. Wang et al. 2023) (Varma et al. 2018) (Bourja et al. 2021) (Mirbod et al. 2023).

There is no explicit provision regarding the optimal baseline distance for various environmental conditions. A semi-autonomous UAV type drone has been developed to enhance the capabilities of stereo vision systems. In long-term 3D reconstruction tasks, a more comprehensive baseline, up to 10 meters, is applied (Sumetheeprasit et al. 2024). Stereo cameras have limited measurement range, particularly with short baselines (Lin, Tsai, and Tran 2021). The stereo camera system often produces poor results or is inaccurate when measuring the estimated distance of a near or distant object. Flexible stereo configurations with variable baselines enable medium to large scale mapping, resulting in several advantages including reduced travel distance, reduced processing time per area, and improved real-time monitoring capabilities (Sumetheeprasit et al. 2024). Measurement accuracy can decrease at long distances, mainly due to the limitations of disparity maps produced by stereo cameras (Lin, Tsai, and Tran 2021). Distance measurement accuracy is highly dependent on good stereo camera calibration; poor calibration can cause significant errors in distance measurement (Bourja et al. 2021).

Stereo vision facilitates the calculation of object distances by examining the discrepancy in viewing angles between two parallel cameras. This method, known as stereo vision, utilises the idea of disparity, which is the positional difference of objects viewed in the left and right images. Technologies Stereo SGBM (Semi-Global Block Matching) utilised by OpenCV offers enhanced accuracy in distance measurement using stereo image analysis (Erwin Syahrudin, Ema Utami, and Anggit Dwi Hartanto 2024). The baseline, or the distance between the two cameras, is a crucial factor that must be calibrated to significantly improve the precision of distance measurement.

Most research emphasizes algorithm development, but not much compares system performance with various baselines to detect human distance at multiple distances accurately. Development of a distance measurement algorithm using a stereo camera, but does not explicitly discuss baseline variations for human distance detection (Adil, Mikou, and Mouhsen 2022) (Zhou et al. 2021). Unsupervised deep learning with minimal baseline distances fails to produce accurate disparity maps and does not reach the right solution due to disappearing gradients (Imran et al. 2020). Pedestrian trajectory prediction uses stereo cameras but emphasizes algorithmic aspects for pose estimation and does not compare system performance with baseline variations in detecting human distance accurately (Bourja et al. 2021).

Numerous studies concentrate on objects at relatively short distances; however, what occurs at extended distances where objects become undetectable? The detection range of stereo cameras is restricted, particularly at extended distances, and precision frequently diminishes when objects are outside the ideal detection range (Seo, Park, and Choi 2022). At extended distances, stereo cameras may encounter heightened measurement inaccuracies, chiefly because to the restricted resolution of disparity maps (Sumetheeprasit et al. 2024). The stereo setup permits modification of the baseline, orientation, and tilt angle to enhance the quality of stereo matching and depth estimation. Previous research has mostly focused on the creation of algorithms for stereo image processing; however, few studies have rigorously investigated the effect of baseline variation on the accuracy of distance estimate. Improving baseline distances under varying ambient conditions and object distances necessitates further research. In applications such as autonomous vehicles, where the accuracy of distance measurement affects safety, the choice of an appropriate baseline is essential.

Triangulation employed in stereo cameras is often impractical, particularly in low-light circumstances or with low-contrast subjects. In low light settings, stereo camera measurements may be compromised, mostly due to challenges in catching low-contrast objects (Zhou et al. 2021). Stereo camera systems encounter difficulties in low-light conditions and when objects possess attributes that are challenging to quantify, particularly if the object's dimensions are diminutive or variable (Z. Wang et al. 2023). Employing colour pictures for depth perception diminishes the efficiency of object detection in low-light conditions (Varma et al. 2018). This indicates that the performance of a stereo vision system, including triangulation, may be compromised by suboptimal lighting conditions.

Stereo vision is relatively simpler compared to other sensors such as Light Detection and Ranging (LiDAR). However, the accuracy of 3D reconstruction depends on several critical parameters, with one of the most significant being the camera separation (baseline distance) (Sumetheeprasit et al. 2023). A larger baseline allows for more accurate detection of distant objects, yet it increases the likelihood of occlusion and requires more complex handling of perspective differences. Conversely, an overly short baseline may restrict the measurement range because the parallax difference becomes too small. To address these challenges, the concept of variable baseline and flexible configuration has emerged in stereo camera systems (Sumetheeprasit et al. 2024). In this configuration, the two cameras are no longer rigidly fixed together but can be set apart (for instance, on separate drones) and rearranged according to the desired measurement range. A shorter baseline is more suitable for nearby objects, whereas a longer baseline is needed for those at greater distances. Furthermore, such flexibility enables orientation adjustments (horizontal, vertical, or angled) to minimize occlusions depending on the shape and position of the objects.

A very long baseline (30 meters or more) may be required to improve measurement accuracy over distances of several kilometers (Yang, Lu, and Li 2024). However, this introduces higher parallax and substantial differences in viewing angles between the left and right cameras. When observing dynamic targets such as the ocean surface, synchronizing both cameras becomes more challenging, potentially causing rectification errors. Traditional stereo matching methods, whether block-based or feature-based, often fail to handle these difficulties—particularly when changes in illumination or contrast occur rapidly, such as over ocean waves.

Environmental factors also influence stereo camera performance, including lighting conditions, lens quality, and medium changes (air and water) (Hu et al. 2023). In underwater applications, baseline selection becomes more intricate due to refraction effects caused by variations in the densities of water, the camera housing glass, and air.

These phenomena not only alter the camera's intrinsic parameters but also affect feature matching and the triangulation process. Recent studies confirm that selecting and optimizing the baseline for stereo cameras significantly improves distance measurement accuracy and 3D mapping precision.

Overall, the quality of 3D reconstruction and object distance measurements is strongly affected by various parameters, including baseline distance (Seo, Park, and Choi 2022). A larger baseline can extend the effective measurement range, particularly for more distant objects. Nevertheless, it also introduces limitations, such as increased system size, difficulties in maintaining camera stability, and higher distortion potential in certain areas of observation. On the other hand, short-baseline approaches facilitate integration into smaller platforms (compact robots or drones), but they often restrict the sensing range and render the system more sensitive to lighting disruptions. Therefore, selecting the appropriate baseline becomes key to balancing accuracy requirements, placement flexibility, and data processing complexity.

One of the primary issues in utilizing stereo cameras is ascertaining the ideal baseline distance between the two devices. The baseline refers to the horizontal distance separating the two cameras in a stereo system and is essential for the precision of object distance readings. An excessively short baseline may result in considerable measurement inaccuracies over extended distances, whereas an overly long baseline can exacerbate mistakes at shorter distances. Consequently, it is essential to establish the appropriate baseline in accordance with the specific conditions and requirements of a certain application. This research is vital because determining the optimal baseline distance will improve the accuracy of distance estimation in stereo camera systems and expand the practical applications of this technology in various industries, including autonomous vehicles and robotics. This study aims to explore how the effect of baseline distance on stereo cameras can improve accuracy in measuring the estimated distance of objects.

MATERIALS & METHODS

This study uses a quantitative experimental approach. It evaluated the effect of baseline distance variation on stereo camera systems on the accuracy of object distance measurement. In the research, to detect objects in the image will use the Mask R-CNN algorithm. Furthermore, to calculate the disparity using the Stereo Semi-Global Block Matching (SGBM) algorithm. Disparity is used to create an accurate depth map. The main focus of this study is the analysis of distance measurement accuracy based on baseline variations. The flowchart of this research can be seen in Figure 1.

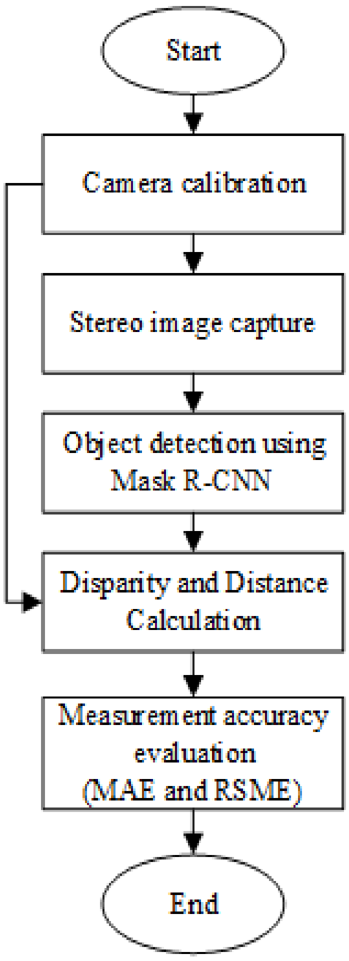

Figure 1. Flow of The Research Process

1 Camera Calibration

Camera calibration is an essential procedure for guaranteeing precise measurement outcomes. Calibration procedure flaws can lead to substantial inaccuracies in the final measurement outcomes (Gao et al. 2021). The calibration process uses a chessboard calibration pattern, where the size or distance between the points is known. Some images of the calibration pattern were taken from various angles and distances, which were then used to extract the corner points. Based on these images, the OpenCV calibration algorithm calculates the camera's intrinsic parameters, including focal length and lens distortion. These parameters are saved for use at a later stage, especially in disparity and depth calculations. Camera calibration documentation can be seen in Figure 2.

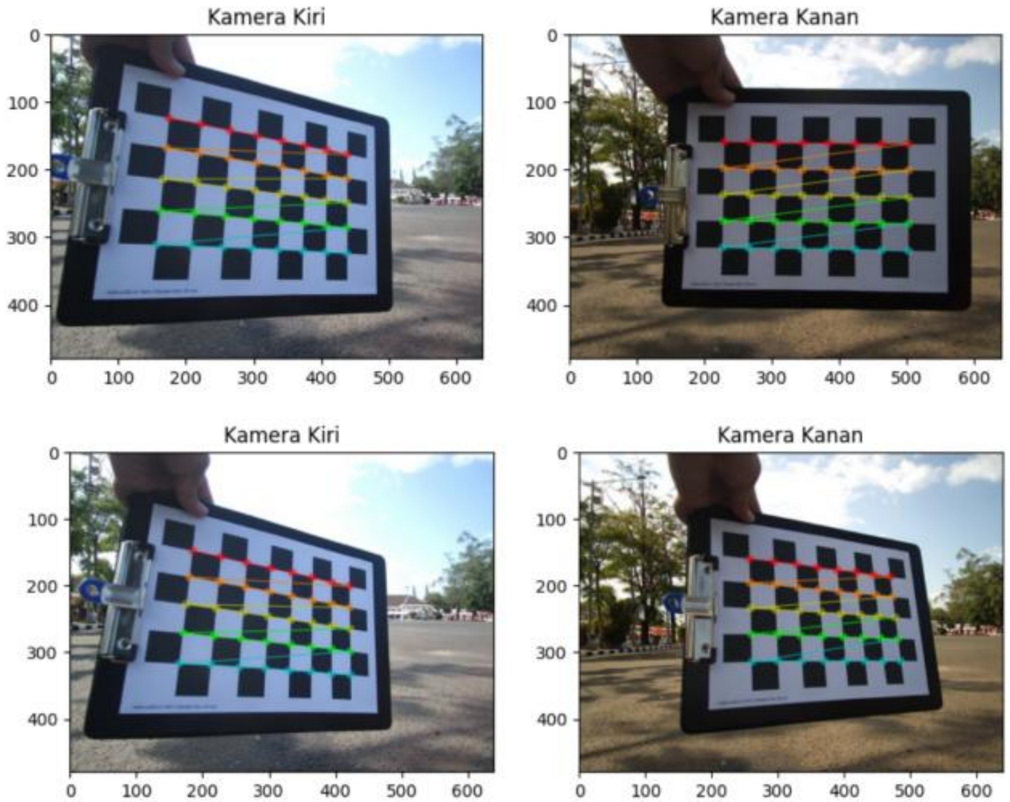

Figure 2. Camera Stereo Calibration Process

2. Stereo Image Capture

In this study, the test subjects were human objects observed using a stereo system consisting of two identical cameras mounted horizontally with an adjustable base distance. Before the shooting began, the stereo system was calibrated to determine the intrinsic and extrinsic parameters of the camera, namely focal length, lens distortion, and relative position between the cameras. The baseline used was 25 cm to 275 cm with each multiple of 50 cm. Each baseline configuration was tested at various object distances, ranging from 2 m to 50 m. The camera used was a webcam type with a resolution of 1920 x 1080p. Lighting conditions remained stable during the recording procedure to minimize external variables that could affect the results. In addition, the location and orientation of the camera were maintained consistently to ensure that any observed variance was only due to variations in the base distance rather than external influences such as camera movement. Documentation of stereo image capture can be seen in Figure 3.

Figure 3. Stereo Image Capture Process

3. Object Detection Using R-CNN Masks

Upon acquiring the stereo image, the Mask R-CNN technique is employed to identify the items inside the image. The Mask R-CNN is a preeminent object detection method, distinguished by its capability to recognize items with bounding boxes and to generate a mask for each identified object (Seo, Park, and Choi 2022). The Mask R-CNN approach was chosen for its superior capacity to identify things with increased precision, especially when the objects have complex or irregular shapes. Mask R-CNN provides a bounding box and a mask for each detected object, enabling the stereo system to focus on a specific area during disparity calculation. This ensures that the resulting distance measurement relates exclusively to the object of interest, omitting the backdrop or other objects in the image. Figure 4 illustrates the outcomes of object detection with Mask R-CNN.

Figure 4. Object Detection using Mask R-CNN

4. SGBM Algorithm Parameter Optimization

After the Mask R-CNN identifies the item, the subsequent stage involves computing the disparity map via the SGBM algorithm. This approach calculates the disparity map, representing the positional difference of the identical object in the left and right images. This discrepancy indicates the extent of the object's positional change in the stereo picture, which will subsequently be utilized to compute the object's depth (Gao et al. 2021).

The SGBM technique contains several options that significantly influence the quality of the resulting disparity map, including minDisparity, numDisparities, window size, P1 and P2. The parameter optimization approach used is Optuna, an open-source library specifically designed for hyperparameter optimization, to identify the most optimal Stereo SGBM configuration. Unlike traditional grid search, Optuna employs adaptive sampling strategies such as the Tree-structured Parzen Estimator (TPE) to concentrate on the most promising parameter combinations. An objective function was designed to calculate the error metric (MAE) based on the outputs of Stereo SGBM, which employed various parameters. The optimization process was conducted 500 trials for each measured ground distance, thereby yielding a parameter configuration that minimizes errors while enhancing efficiency in terms of both time and computational resources. The optimization results can be seen in Figure 5.

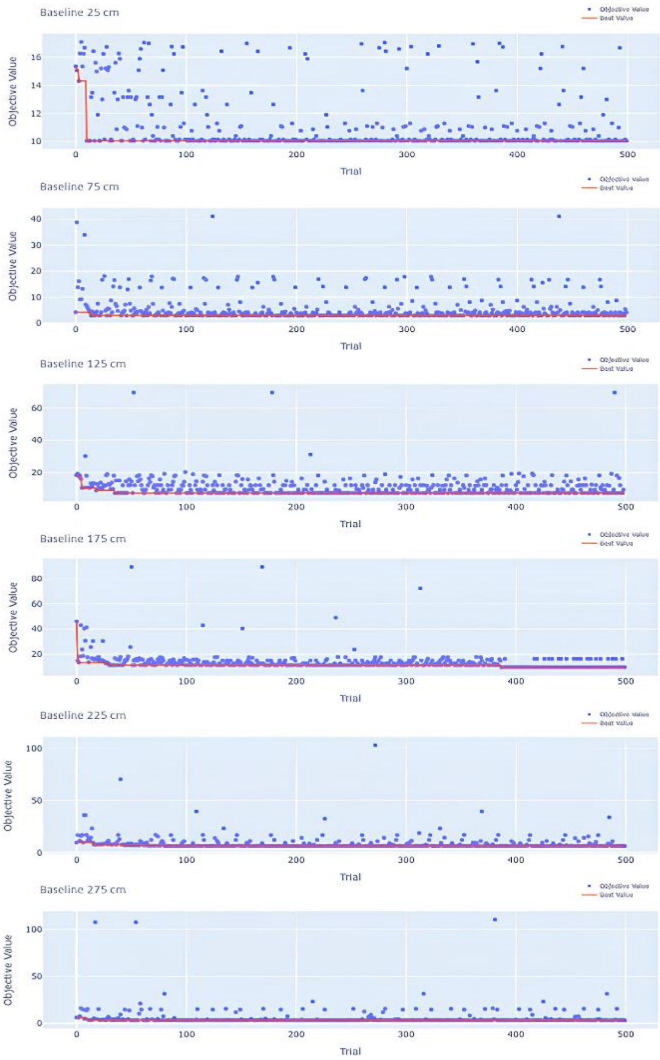

Figure 5. Convergence Plots of the Objective Function in Optuna for Different Baseline Distances

5. Disparity and Distance Calculation

Selecting appropriate parameters is essential for generating precise depth maps, particularly for objects at differing distances. Consequently, we employ the Optima-based optimization technique to identify the ideal parameter combination that minimizes measurement errors in object distances. Figure 6 illustrates an example of the disparity map results.

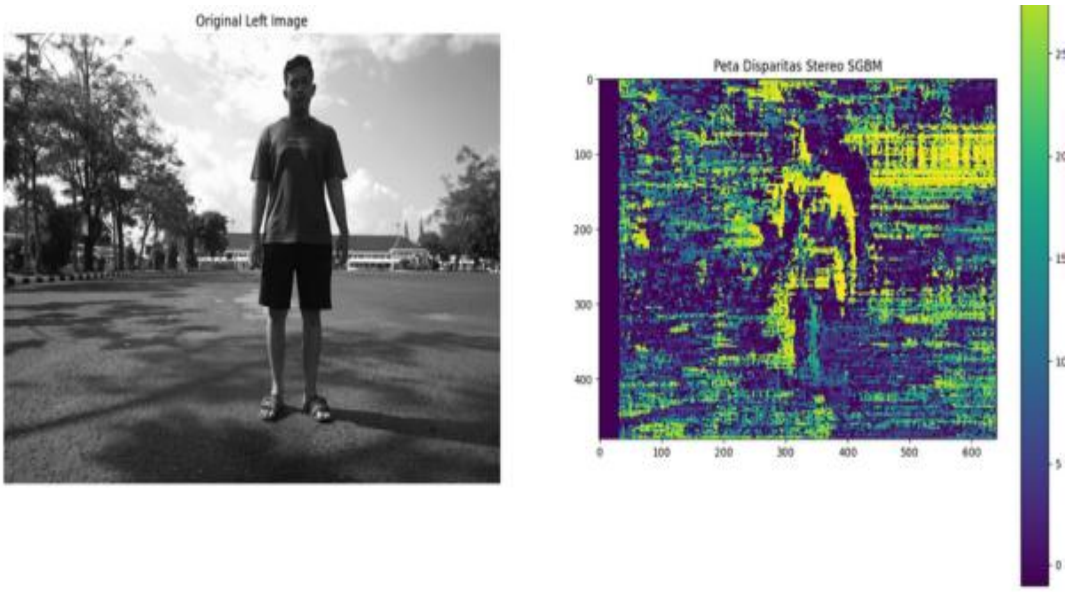

Figure 6. Disparity Maps Results

Measurements of object distance derived from the disparity map are computed utilizing the subsequent stereo vision equation:

$D = \frac{df \ x \ B}{d}$ (1)

D represents the item's distance from the camera (depth), f denotes the focal length derived from calibration findings, B signifies the baseline distance between the two cameras, and d indicates the disparity value computed from the positional difference of the object in the left and right photos. Figure 7 illustrates instances of distance measurement outcomes.

Figure 7. Distance Calculation Result

The Stereo Semi-Global Block Matching (SGBM) algorithm was selected for its capacity to generate highly accurate depth maps. The resultant disparity map will be employed to ascertain the object's depth utilizing standard stereo vision equations. This approach also mitigates ambiguity in disparity computations, frequently encountered in photos with intricate details or distant objects.

6. Measurement Accuracy Evaluation

For each baseline configuration, the depth of the item as measured by the stereo system is compared to the distance obtained using a physical measuring instrument, such as a meter. This comparison seeks to assess the precision of the stereo system. The system's accuracy is evaluated using Mean Absolute Error (MAE) and Root Mean Square Error (RMSE) to determine the proximity of the stereo system's measurement findings to the real distance value. Mean Absolute Error (MAE) and Root Mean Square Error (RMSE) are two measures commonly employed to evaluate model performance. MAE quantifies the mean absolute error between expected and actual values, attributing equal significance to all errors. The equation is as follows:

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \qquad (2)$$

Meanwhile, the RMSE calculates the square root of the mean of the squared error. The formula is as follows:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \qquad (3)$$

The value of n is the amount of data used in the calculation. The value of yi is the actual value of the measurement or observation data. The value of $\hat{y}_i$ indicates the predicted or estimated value produced by the model or measurement system.

MAE and RMSE values are computed to assess the proximity of the measurement findings to the true distance. This elucidates the precision of the stereo system in object detection across diverse baseline configurations and offers an additional understanding of how baseline alterations influence measurement accuracy.

RESULTS AND DISCUSSIONS

Based on a series of parameter tests for Stereo SGBM across various baselines, several key findings clarify the impact of configuration settings on both distance measurement accuracy and disparity map quality. The results indicate that a moderately large blockSize (e.g., 7–15) consistently yields smoother disparity maps with fewer mismatches, whereas very small blockSize values (≤5) tend to introduce noise, particularly at medium to long distances. For medium to long baselines (≥75 cm), employing higher numDisparities (≥128) significantly improves accuracy; however, for short baselines (25 cm), increasing num Disparities offers limited benefit due to restricted parallax. Setting P1 and P2 proportionally (with P2 equal to 4×P1) effectively reduces mismatches without sacrificing important details, whereas inappropriate P1/P2 configurations can result in gaps in the disparity map or elevated errors in textured regions. Regarding the Uniqueness Ratio, a moderate range (5–10) balances mismatch rejection with disparity map density, though excessively high values may reduce the overall disparity area, and overly low values risk increasing mismatches. Additionally, Speckle Window Size and Speckle Range values that are not excessively large (5 and 1–2, respectively) generally suppress speckle noise without compromising key object features, while overly aggressive filtering may eliminate essential edge details.

The measurement results show a clear correlation between the baseline distance and the accuracy of stereo camera systems in recognizing object distances. As shown in Table 1, shorter baselines (e.g., 25 cm) result in significant measurement errors, especially when detecting distant objects. These errors occur due to minimal disparities, leading to unreliable depth estimation. Conversely, longer baselines (>175 cm) improve accuracy for distant objects but introduce larger errors in near-range measurements due to exaggerated disparity distortions and increased calibration sensitivity.

Table 1. Distance Measurement Results Based on Baseline

Ground Truth Distance (m)	Calculated distance (m) for baseline (cm)					
	25	75	125	175	225	275
2	2.0	2.0	2.5	3.5	5.9	5.9
4	4.3	4.0	3.9	4.0	5.0	5.5
6	6.2	6.0	6.0	6.0	6.0	6.2
8	7.0	7.7	8.1	7.8	8.0	8.0
10	9.5	10.3	9.9	10.0	10.9	10.0
15	12.0	14.1	15.0	15.0	15.0	15.3
20	13.3	20.0	19.4	20.1	20.0	19.8
25	15.4	25.1	25.1	25.0	24.1	25.4
30	12.8	30.1	29.7	29.7	31.9	30.0
35	21.8	35.4	36.8	35.8	38.7	35.6
40	29.7	39.8	39.3			40.8
45		45.5	46.4	44.0	44.4	45.0
50			43.6	49.5	55.6	

In some baselines, measurement results cannot be obtained because objects are not detected in stereo images. This happens because object detection algorithms, such as Mask R-CNN, fail to recognize objects in the images. Based on the measurement results presented in Table 1, the accuracy of measuring the distance of objects using MAE and RMSE, where the results can be seen in Table 2.

Table 2. Model Performance Evaluation Based on Baseline Distance

No	Baseline (cm)	MAE	RMSE
1	25	5.64	8.12
2	75	0.23	0.34
3	125	0.94	1.91
4	175	0.38	0.61
5	225	1.54	2.34
6	275	0.66	1.25

The results indicate that the most optimal performance is achieved at a 75 cm baseline, yielding the lowest RMSE of 0.34 m and MAE of 0.23 m. This confirms that a 75 cm baseline provides a highly reliable distance estimation, with an error margin of only 0.68% of the maximum measured distance (50 m). Compared to existing literature, which generally reports an acceptable RMSE threshold of around 2%, the RMSE obtained in this study is significantly lower, further demonstrating the effectiveness of the experimental setup.

The reliability of the obtained RMSE value (0.34 m at 75 cm baseline) is considered high for practical applications, particularly in fields such as robotic navigation, autonomous vehicles, and industrial monitoring. Given that acceptable RMSE values for stereo-based distance estimation are typically within 2% of the total measured distance, the 0.68% RMSE achieved in this study suggests that the proposed baseline configuration is highly effective for real-world applications.

However, some limitations exist in this study. The dataset used was collected under controlled experimental conditions, primarily focusing on human subjects in relatively static environments. As a result, findings may not fully generalize to more dynamic scenarios or different object types. Additional tests with varied textures, colors, lighting conditions, and more complex backgrounds are required to enhance the robustness and generalizability of these results.

This study successfully demonstrated that measuring object distances from a distance is still a challenge for stereo camera systems, especially at suboptimal baseline lengths. As the object distance increases, the results prove that baselines that are too long, such as 225 cm and 275 cm, result in significant error rates. In contrast, at shorter to medium baseline distances, a baseline of 75 cm is shown to be the most appropriate choice, as it can produce high accuracy and minimal error rates.

CONCLUSION

In this study, a 75 cm baseline was identified as optimal based on the experimental results. However, for both shorter (25 cm) and longer (275 cm) baselines, the tests showed significant errors in distance measurements. For the very short baseline (25 cm), the parallax produced is very small, leading to inaccuracies when measuring objects at longer distances. In the context of triangulation, as the baseline decreases, the disparity between the left and right camera views becomes smaller, making depth calculation more sensitive to small errors. This is particularly evident for distant objects, where the small difference in angles leads to larger errors in distance estimation. On the other hand, with a very long baseline (275 cm), while accuracy improves for distant objects, errors increase for close objects. For objects that are very close, the angular difference between the two cameras becomes very large, leading to significant disparity distortion, which can result in rectification errors. This explains why long baselines tend to perform poorly for close-range objects, as the viewing angle becomes too extreme, affecting disparity matching.

To overcome these issues, the use of a dynamic baseline could be a potential solution. By adjusting the distance between the cameras dynamically, the stereo system could change the baseline based on the distance of the detected object. For close objects, the system could use a shorter baseline to reduce disparity distortion and improve accuracy. Conversely, for distant objects, a longer baseline could be employed to enhance distance measurement accuracy. This configuration would allow the system to adapt to varying object distances, resulting in better overall performance and improving the robustness of the stereo camera in dynamic environments.

Furthermore, the environmental conditions during testing were relatively controlled, and the study did not include variations in lighting, weather conditions, or highly complex backgrounds. In real-world scenarios, stereo cameras often encounter challenges such as low-light conditions, backlighting, and cluttered backgrounds, which can significantly reduce the reliability of disparity maps and increase the likelihood of mismatches. These factors were not accounted for in the current testing.

Therefore, future studies are needed to evaluate the system's performance on a wider range of objects (metallic surfaces, textured objects, and those with contrasting colors) and under varying lighting conditions (such as indoor, outdoor, and low-light scenarios). This would provide a more comprehensive understanding of the stereo system's robustness, and ensure its applicability in real-world settings, such as industrial applications, agriculture, and environmental monitoring.

Declaration by Authors

Acknowledgements: We express our sincere gratitude to Universitas Amikom Yogyakarta for the support and resources that made this research possible. We also acknowledge our colleagues' invaluable guidance and encouragement throughout this project. The authors declare no conflict of interest.

Source of Funding: None

Conflict of Interest: No conflicts of interest declared.

REFERENCES

1. Adil, Elmehdi, Mohammed Mikou, and Ahmed Mouhsen. 2022. “A Novel Algorithm for Distance Measurement Using Stereo Camera.” *CAAI Transactions on Intelligence Technology* 7(2): 177–86. doi:10.1049/cit2.12098.
2. Bourja, Omar, Hatim Derrouz, Hamd Ait Abdelali, Abdelilah Maach, Rachid Oulad Haj Thami, and Francois Bourzeix. 2021. “Real Time Vehicle Detection, Tracking, and Inter-Vehicle Distance Estimation Based on Stereovision and Deep Learning Using Yolov3.” *International Journal of Advanced*

Computer Science and Applications 12(8): 915–23. doi:10.14569/IJACSA.2021.01208101.

3. Erwin Syahrudin, Ema Utami, and Anggit Dwi Hartanto. 2024. "Enhanced Yolov8 with OpenCV for Blind-Friendly Object Detection and Distance Estimation." *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)* 8(2): 199–207. doi:10.29207/resti.v8i2.5529.
4. Gao, Zeren, Yue Gao, Yong Su, Yang Liu, Zheng Fang, Yaru Wang, and Qingchuan Zhang. 2021. "Stereo Camera Calibration for Large Field of View Digital Image Correlation Using Zoom Lens." *Measurement: Journal of the International Measurement Confederation* 185. doi:10.1016/j.measurement.2021.109999.
5. Gehrig, Mathias, Willem Aarents, Daniel Gehrig, and Davide Scaramuzza. 2021. "DSEC: A Stereo Event Camera Dataset for Driving Scenarios." *IEEE Robotics and Automation Letters* 6(3): 4947–54. doi:10.1109/LRA.2021.3068942.
6. Haider, Azmi, and Hagit Hel-Or. 2022. "What Can We Learn from Depth Camera Sensor Noise?" *Sensors* 22(14). doi:10.3390/s22145448.
7. Hu, Kai, Tianyan Wang, Chaowen Shen, Chenghang Weng, Fenghua Zhou, Min Xia, and Liguo Weng. 2023. "Overview of Underwater 3D Reconstruction Technology Based on Optical Images." *Journal of Marine Science and Engineering* 11(5). doi:10.3390/jmse11050949.
8. Imran, Saad, Sikander Bin Mukarram, Muhammad Umar Karim Khan, and Chong-Min Kyung. 2020. "Unsupervised Deep Learning for Depth Estimation with Offset Pixels." *Optics Express* 28(6): 8619. doi:10.1364/oe.385328.
9. Kardovskyi, Yurii, and Sungwoo Moon. 2021. "Artificial Intelligence Quality Inspection of Steel Bars Installation by Integrating Mask R-CNN and Stereo Vision." *Automation in Construction* 130. doi:10.1016/j.autcon.2021.103850.
10. Lin, Huei Yung, Chun Lung Tsai, and Van Luan Tran. 2021. "Depth Measurement Based on Stereo Vision with Integrated Camera Rotation." *IEEE Transactions on Instrumentation and Measurement* 70. doi:10.1109/TIM.2021.3073687.
11. Liu, Fei, and Guangqi Hou. 2024. "Depth Estimation from a Hierarchical Baseline Stereo with a Developed Light Field Camera." *Applied Sciences (Switzerland)* 14(2). doi:10.3390/app14020550.
12. Mirbod, Omeed, Daeun Choi, Paul H. Heinemann, Richard P. Marini, and Long He. 2023. "On-Tree Apple Fruit Size Estimation Using Stereo Vision with Deep Learning-Based Occlusion Handling." *Biosystems Engineering* 226: 27–42. doi:10.1016/j.biosystemseng.2022.12.008.
13. Raoui, Younès, and Mohammed Amraoui. 2024. "Simultaneous Localization and Mapping of a Mobile Robot with Stereo Camera Using ORB Features." *Journal of Automation, Mobile Robotics and Intelligent Systems* 18(2): 62–71. doi:10.14313/jamris/2-2024/14.
14. Sabato, Alessandro, Nicholas A. Valente, and Christopher Niezrecki. 2020. "Development of a Camera Localization System for Three-Dimensional Digital Image Correlation Camera Triangulation." *IEEE Sensors Journal* 20(19): 11518–26. doi:10.1109/JSEN.2020.2997774.
15. Seo, Beom Su, Byungjae Park, and Hoon Choi. 2022. "Sensing Range Extension for Short-Baseline Stereo Camera Using Monocular Depth Estimation." *Sensors* 22(12). doi:10.3390/s22124605.
16. Sumetheeprasit, Borwonpob, Ricardo Rosales Martinez, Hannibal Paul, Robert Ladig, and Kazuhiro Shimonomura. 2023. "Variable Baseline and Flexible Configuration Stereo Vision Using Two Aerial Robots." *Sensors* 23(3). doi:10.3390/s23031134.
17. Sumetheeprasit, Borwonpob, Ricardo Rosales Martinez, Hannibal Paul, and Kazuhiro Shimonomura. 2024. "Long-Range 3D Reconstruction Based on Flexible Configuration Stereo Vision Using Multiple Aerial Robots." *Remote Sensing* 16(2). doi:10.3390/rs16020234.
18. Varma, V. S.K.P., S. Adarsh, K. I. Ramachandran, and Binoy B. Nair. 2018. "Real Time Detection of Speed Hump/Bump and Distance Estimation with Deep Learning Using GPU and ZED Stereo Camera." In *Procedia Computer Science*, Elsevier B.V., 988–97. doi:10.1016/j.procs.2018.10.335.
19. Wang, Weidong, Jun Peng, Wenbo Hu, Jin Wang, Xinyue Xu, Qasim Zaheer, and Shi Qiu. 2024. "A Multi-Degree-of-Freedom Monitoring Method for Slope Displacement Based on Stereo Vision." *Computer-Aided*

Civil and Infrastructure Engineering 39(13): 2010–27. doi:10.1111/mice.13173.
20. Wang, Zilong, Yifei Ding, Tianhang Zhang, and Xinyan Huang. 2023. "Automatic Real-Time Fire Distance, Size and Power Measurement Driven by Stereo Camera and Deep Learning." *Fire Safety Journal* 140. doi:10.1016/j.firesaf.2023.103891.
21. Yang, Ying, Cunwei Lu, and Zhenhua Li. 2024. "Deep-Learning-Based Stereo Matching of Long-Distance Sea Surface Images for Sea Level Monitoring Systems." *Journal of Marine Science and Engineering* 12(6). doi:10.3390/jmse12060961.
22. Zaarane, Abdelmoghit, Ibtissam Slimani, Wahban Al Okaishi, Issam Atouf, and Abdellatif Hamdoun. 2020. "Distance Measurement System for Autonomous Vehicles Using Stereo Camera." *Array* 5(January): 100016. doi:10.1016/j.array.2020.100016.
23. Zhang, Yufeng, Lianghui Ding, Yuxi Li, Weiyao Lin, Mingbi Zhao, Xiaoyuan Yu, and Yunlong Zhan. 2021. "A Regional Distance Regression Network for Monocular Object Distance Estimation." *Journal of Visual Communication and Image Representation* 79. doi:10.1016/j.jvcir.2021.103224.
24. Zhou, Changxin, Yazhou Liu, Quansen Sun, and Pongsak Lasang. 2021. "Vehicle Detection and Disparity Estimation Using Blended Stereo Images." *IEEE Transactions on Intelligent Vehicles* 6(4): 690–98. doi:10.1109/TIV.2020.3049008.

How to cite this article: Fariz Imam, Arief Setyanto, Kusnawi. Optimising baseline distance in stereo cameras: an experimental approach to enhance object distance accuracy. *International Journal of Research and Review*. 2025; 12(5): 94-107. DOI: *10.52403/ijrr.20250512*

******