



# An Improvement Stereo Vision Images Processing for Object Distance Measurement

**Tsung-Shiang Hsu and Ta-Chung Wang\***

*Institute of Civil Aviation, National Cheng Kung University, Taiwan*

(Received 6 January 2014; Accepted 17 March 2014; Published on line 1 June 2015)

\*Corresponding author email: [tchung@mail.ncku.edu.tw](mailto:tchung@mail.ncku.edu.tw)

DOI: [10.5875/ausmt.v5i2.460](https://doi.org/10.5875/ausmt.v5i2.460)

**Abstract:** Human has the ability to roughly estimate the distance of objects because of the stereo vision of human's eyes. In this paper we proposed an improved stereo vision system to accurately measure the distance of objects in real world. Object distance is very useful for obstacle avoidance and navigation of autonomous vehicles. Recent researches have used stereo cameras for different applications such as 3D image construction, distance measurement, and occlusion detection. The proposed measurement procedure is a three-phase process: object detection, segmentation, and distance calculation. In distance calculation, we proposed a new algorithm to reduce the error. The result shows our measurement system is capable of providing objects distance with less than 5% of measurement error.

**Keywords:** stereo vision; distance measurement

## Introduction

Object distance measurement is becoming more and more important in mobile autonomous systems. Information of the distance of objects are useful for navigation and avoidance of autonomous mobile systems. The system can easily avoid obstacles if the accurate position of obstacles can be acquired [1].

With the advancement of technology, the stereo vision technology has become more and more mature. Many related applications have hence emerged, such as 3D movie, face recognition [2], obstacle detection and recognition [3]. Previous research works have obtained object distances from images using one vision sensor. For example, in [4], the system utilized a single camera and a laser pointer to obtain distance measurement. Hai-Sung Baek et al. discussed the differences on crossing camera setups and parallel camera setups [5]. They also suggest a method to measure the distance when an object is located outside the optical axes in the overlapping area of two cameras.

Recent researches have used multiple vision sensors for the purpose of object distance and size measurements. Mustafah et al. utilized stereo vision to

measure object distance and size (height and width) [6]. They not only measured object distance but also proposed a method to measure the object size. The other example is the work of A-Lin et al. in which they used stereo vision to measure the safe driving distance of a vehicle [1].

Although there are many existing successful works related to the object distance measurement, the underlying distance calculation formulae have omitted the variation in the image distance to the lenses. In this paper, we proposed a revised algorithm using stereo vision for distance measurement. With the proposed algorithm, the accuracy of distance measurement is better than previous works. Several experiment results showed that we can accurately obtain multi-object distances using the proposed approach.

## Object Distance Measurement

The flow of the object measurement system that we proposed started with stereo vision image capture. Then, on both images, a preprocessing will be applied by object detection and segmentation. We then segment the images to isolate the objects in view. After that, feature points of objects on both images are compared to get the distance measurements. More details will be illustrated in



the following section. The flow of the object measurement system is shown in Figure 1.

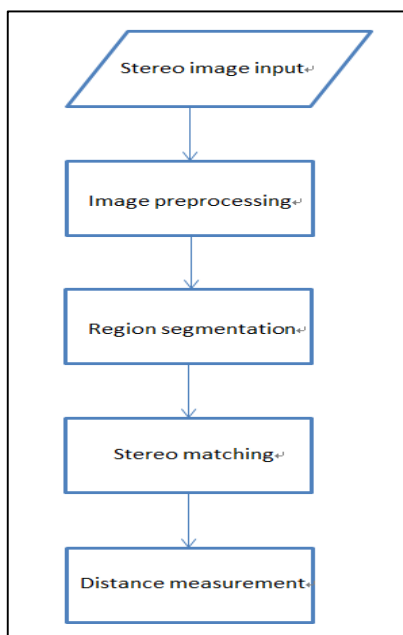


Figure 1. The flow of the object measurement system.

### Stereo Image Capture

Stereo image capture is done by using two video cameras which are aligned in parallel with a fixed relative position. The object distance can be measured when it enters the overlapping views of the two cameras. Figure 2 illustrates the stereo vision setup.

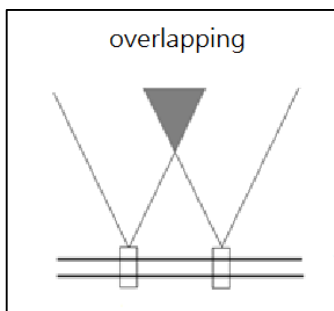


Figure 2. Stereo vision setup.

### Preprocessing

Image preprocessing is an important and common method in a computer vision system. It can enhance the images quality and improve the computational efficiency. In our system, after the stereo images are captured, the resolution of the images is down-scaled to improve the computation speed. For example, the original resolution which is 2784x1856 pixels is downscaled to 320x240 pixels. The result shows that the reduction of resolution does not affect the accuracy of the system. Another way we used to improve the speed is through converting the images from RGB color space to gray scale color space. RGB color space requires three times more computation and memory space compared to gray scale. Although some information may be lost in gray scale color space, it does not have significant effects the distance measurement accuracy as will be shown in the experiment.

### Region Segmentation

**Local Threshold Selection.** For a simple image, we can easily find a threshold to separate objects of interest from background with Otsu’s algorithm [7]. Let the image that isolate the object from the background be called as the binary image. Assume that the considered image can be classified into two clusters  $C_1$  and  $C_2$  as shown in Figure 3. One idea to choose the threshold value is to decide a gray level value,  $T$ , using which the variance between the two clusters is the largest, and the sum of the within-cluster variance in  $C_1$  and  $C_2$  is the smallest.  $T$  would be the optimal threshold to separate the object from the background.

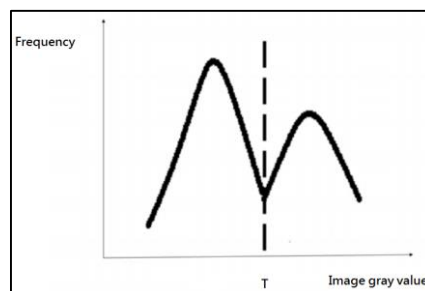


Figure 3. Histogram of gray value distribution with threshold  $T$ .

For complicated images, we cannot acquire a binary image using only one global threshold, because the gray scale of each object may be different. Thus, in our system, the image is divided into smaller blocks. We divide the image into sixteen blocks as shown in Figure 4. For each block, the threshold value for the block is calculated separately to find the corresponding block binary image. Finally, the full binary image is composed from all the block binary images.

**Tsung-Shiang Hsu** received the B. S. degree in mechanical engineering from National Chung Cheng University, Taiwan, and M. S. degree in aeronautics and astronautics from National Cheng Kung University, Taiwan.

**Ta-Chung Wang** received the B. S. and M. S. degrees in mechanical engineering from National Taiwan University, Taipei, Taiwan, in 1995 and 2000, respectively. He received the Ph.D. degree in aeronautics and astronautics from Stanford University, Stanford, CA. in 2007. Currently, he is an assistant Professor of Institute of Civil Aviation, National Cheng-Kung University, Tainan, Taiwan. He teaches course in multivariate analysis, electronics and engineering mathematics. His research interests include operations research in the aviation industry and solving large problems using parallel approaches and nonlinear system analysis.



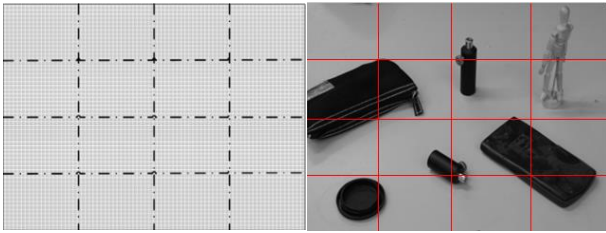


Figure 4. An image is divided into 16 blocks.

**Object Segmentation.** The binary image obtained from the threshold selection stage is then processed for segmentation. We applied a connected component labeling method on the image to separate the objects in the image. The connecting component labeling is used to connect or aggregate the interesting regions. This approach starts with a point (called a seed) and grows the region following the specified growing directions as shown in Figure 5. The idea of region growing is to check if the neighboring pixels are within the same binary region of the binary image.

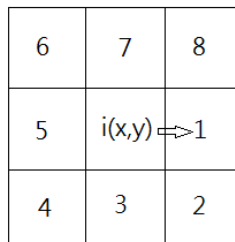


Figure 5. The seed grow directions.

In order to improve the performance of component labeling, we would like to rule out the regions generated due to noise in the image. A simple example is shown in Figure 6. Using a simple labeling region size restriction approach, we can remove the small region being considered as noise shown in Figure 6(b).

*Stereo Matching*

Before making distance measurements, we need to have the information of the disparity value of the objects in the two images. [6] took the pixel differences between the two horizontal centroids of the objects of interest as the disparity as shown in Figure 7.

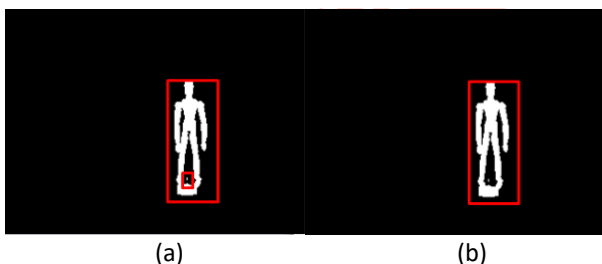


Figure 6. (a) Non restrictions (b) With labeling region size restriction.

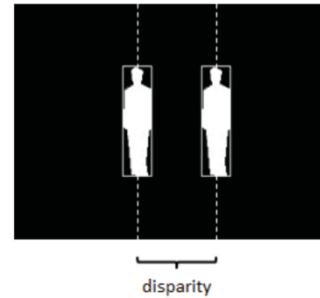


Figure 7. Object disparity.

Unlike Mustafah’s method, we use Speed Up Robust Feature (SURF) to find the disparity between two images. SIFT algorithm generates feature with Gaussian images [8, 9], while SURF algorithm generates them with integral images [10]. SURF is a novel scale- and rotation-invariant interest point detector and descriptor; it approximates or even outperforms previously proposed schemes with respect to repeatability, distinctiveness, and robustness, yet can be computed and compared much faster. This is achieved by relying on integral images for image convolutions

In our proposed system, we detect features in both images and extract feature descriptors after preprocessing. We then match features in both images by using their descriptors. Some matching error will be deleted if the matching accuracy is too low. Finally, we take the average of every matched feature distance as the object disparity. The process is shown in Figure 8.

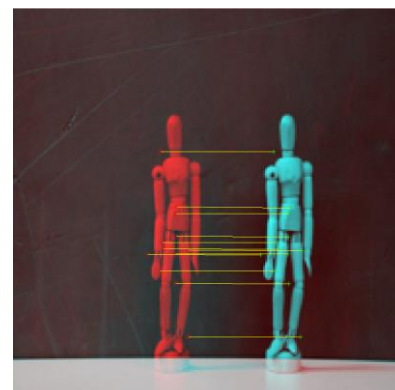


Figure 8. Calculating the matched feature distance as disparity.

*Distance Measurement*

In the previous work [5, 6], the distance,  $D$ , of the object to the equivalent lens can be calculated using the following equation:

$$D = \alpha d^{-1}, \tag{1}$$

where  $\alpha$  is a fixed parameter given by:

$$\alpha = wf,$$

where  $w$  is the distance between the two cameras, and  $f$  is the focal length of the cameras as shown in Figure 9.

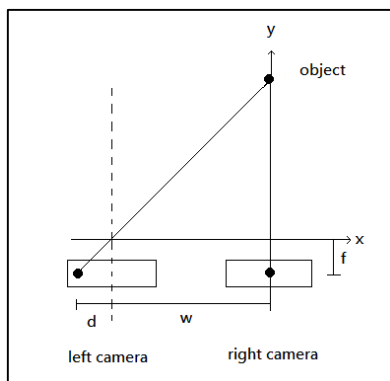


Figure 9. Object beyond the optical axis of left and right camera.

During the experiment, we found that the distance between sensor and the equivalent lens is not precisely equal to the focal length,  $f$ . Therefore we try to revise the formula with thin lens formula:

$$\frac{1}{a} + \frac{1}{b} = \frac{1}{f}, \tag{2}$$

where  $a$  is the object distance,  $b$  is the image distance and  $f$  is the focal length of the lens. According to this formula, the image distance will be equal to the focal length only when the object is located at infinity.

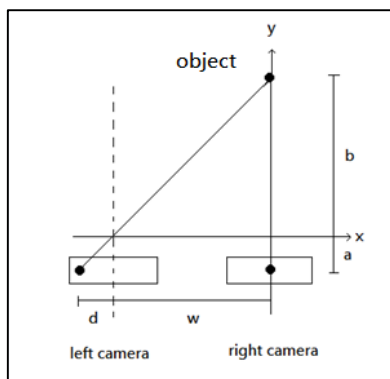


Figure 10. Our proposed system relationship.

According to geometry shown in Figure 10, the distance can be calculated as the following equation:

$$b = \frac{wa}{d}, \tag{3}$$

From (2) and (3), the new distance of object can be calculated as the following equation:

$$b = \frac{fd + fw}{d} \tag{4}$$

where  $d$  is the disparity between two images. Although the derivation assumes the object lies directly in front of one camera, the result also holds if this alignment condition does not hold.

During the experiment, we cannot measure the distance between the equivalent lens and the sensor, marked as  $a$  in Figure 10, and the distance between the equivalent lens and the object, marked as  $b$  in Figure 10, since the location of the equivalent lens is not visible. However, we can have the information about the total distance between the sensor and the object,  $a+b$ . We calculate  $b$  from (4), and get  $a$  from (3), and add them together to get the distance from object to the image sensor.

### Result and Discussion

The following results were developed using MATLAB run on a computer with one Intel Core i7 2.67 GHz processor. Tests were conducted to test the accuracy of the object distance, and the impact of changing the focal length.

The camera we used is a Canon 5D Mark II DLSR camera with EF 24-70mm f/2.8L lens. The size of the image sensor is 36x24mm. The lens can zoom from 24mm to 70mm. The camera and zoom lens are shown in Figure 11.



Figure 11. The camera and zoom lens.

Due to the lack of equipment, the stereo image is done by using only one camera moving on the track. The object can be measured for its distance when it enters the overlapping view of the camera at two different locations. Figure 12 below illustrates the stereo vision setup. If the object enters the overlapping zone, we first take the image at position  $a$ , then take another picture by moving the camera to position  $b$ . We then calculate the perpendicular distance from the object to the image plane parallel to the track.

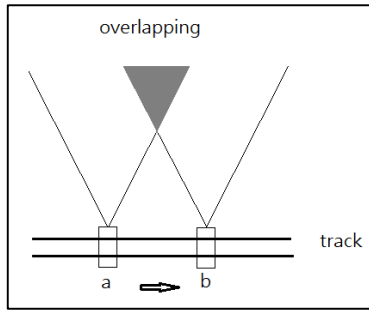


Figure 12. Stereo vision setup.

Table 1. Object distance measurement result on fixed focal length 70mm.

Actual distance a+b(mm)	Measured distance				
	a(mm)	b(mm)	a+b(mm)	absolute error (mm)	absolute error(%)
500	83.253	439.705	552.959	22.959	4.59
600	80.370	542.478	622.849	22.849	3.81
700	78.656	636.105	714.761	12.761	1.82
800	77.237	747.074	824.311	24.311	3.04
900	76.538	819.473	896.011	3.989	0.44

**Experiment I.** Table I and Figure 13 show the distance measurement experiment results using lenses with various focal lengths. The results indicate that the measurement of distances have errors of  $\pm 24$  mm. Our result are more accurate compared with the method used in [6], whose error range is  $\pm 250$  mm.

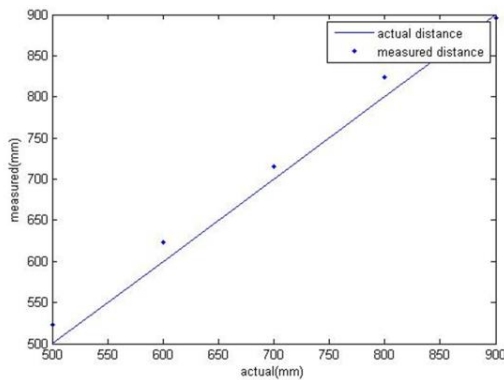


Figure13. Object distance measurement result.

Figure 14 and Table II show the measured distance vs. the actual object distance using lenses with different focal lengths. We can see that the impact on the same distance using different focal lengths. It can be seen that the measurement error decreases if we use a lens with a larger focal length.

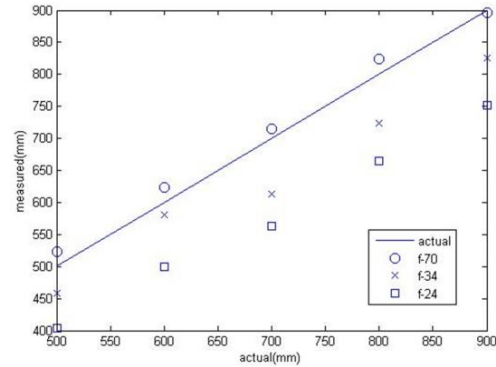


Figure 14. The measurement result in different focal length.

Table 2. The distance measurement error(%) on different focal length.

Actual distance	Measured distance		
	24(mm)(%)	34(mm)(%)	70(mm)(%)
500	96.1(19.22%)	42.0(8.4%)	22.9(4.59%)
600	100.9(16.82%)	18.9(3.14%)	22.8(3.81%)
700	137.5(19.64%)	87.3(12.47%)	14.8(2.11%)
800	135.0(16.88%)	75.7(9.47%)	24.3(3.04%)
900	147.5(16.39%)	74.9(8.32%)	4.0(0.44%)

**Experiment II.** In this experiment, we consider two objects located at different distances as shown in Figure 15. The distance of object 1 is 880mm, and object 2 is located at 1050mm. According to the experiment result in Experiment I, the accuracy of distance measurement using  $f=70$ mm will be better than using other focal lengths. Therefore the focal length we used in this experiment is 70mm.

The calculated results are shown in Table III. We were successful to isolate these two objects in the image and measure their distances using the proposed method.

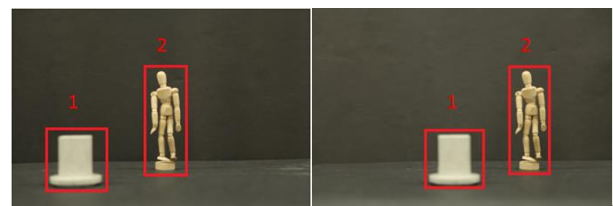


Figure15. Two objects with different distance.

Table 3. The distance measurement error(%) on different focal length.

Object	Measured distance			
	disparity(mm)	a+b(mm)	actual distance(mm)	Error(mm)(%)
1	9.3	896.1	880	16.1(1.8%)
2	7.5	1073.9	1050	23.9(2.3%)

## Conclusion

An object distance measurement approach using a stereo vision system is proposed. The results show our measurement system algorithm is more accurate than previous research works. One disadvantage of the proposed method is that the preprocessing process may be affected by the environment lighting. Our future work is to combine the system with self-propelled vehicles for obstacle avoidance. We would also like to investigate the application of using multiple distance measurement on an object to depict the surface of the object.

## References

- [1] A.-L. Hou, X. Cui, Y. Geng, W.-J. Yuan, and J. Hou, "Measurement of safe driving distance based on stereo vision," in proceeding of *Sixth International Conference on Image and Graphics (ICIG)*, Hefei, Anhui, China, Aug. 12-15, 2011, pp. 902-907. doi: [10.1109/ICIG.2011.27](https://doi.org/10.1109/ICIG.2011.27)
- [2] M.-H. Yang, D. J. Kriegman, and N. Ahuja, "Detecting faces in images: a survey," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 1, pp. 34-58, 2002. doi: [10.1109/34.982883](https://doi.org/10.1109/34.982883)
- [3] Q. Yu, H. Araujo, and H. Wang, "Stereo-vision based real time obstacle detection for urban environments," in proceeding of *The 11th International Conference on Advanced Robotics*, Coimbra, Portugal, June 30-July 3, 2003.
- [4] K. Muljowidodo, A. R. Mochammad, N. SaptoAdi, and B. Agus, "Vision based distance measurement system using single laser pointer design for underwater vehicle," *Indian Journal of Marine Sciences*, vol. 38, no. 3, pp. 324-331, 2009.
- [5] H.-S. Baek, J.-M. Choi, and B.-S. Lee, "Improvement of distance measurement algorithm on stereo vision system(svs)," in proceeding of *the 5th International Conference on Ubiquitous Information Technologies and Applications (CUTE)*, Sanya, China, Dec. 16-18, 2010, pp. 1-3. doi: [10.1109/ICUT.2010.5678176](https://doi.org/10.1109/ICUT.2010.5678176)
- [6] Y. M. Mustafah, R. Noor, H. Hasbi, and A. W. Azma, "Stereo vision images processing for real-time object distance and size measurements," in proceeding of *International Conference on Computer and Communication Engineering (ICCCCE)*, Kuala Lumpur, Malaysia, July 3-5, 2012, pp. 659-663. doi: [10.1109/ICCCCE.2012.6271270](https://doi.org/10.1109/ICCCCE.2012.6271270)
- [7] N. Otsu, "A threshold selection method from gray-level histograms," *IEEE Transactions on Systems, Man and Cybernetics*, 1999, vol. 9, no. 1, pp. 62-66, 1979. doi: [10.1109/TSMC.1979.4310076](https://doi.org/10.1109/TSMC.1979.4310076)
- [8] D. G. Lowe, "Object recognition from local scale-invariant features," in proceedings of *the Seventh IEEE International Conference on Computer Vision*, Kerkyra, Greece, Sept. 20-27, 1999, vol. 2, pp. 1150-1157 vol.1152. doi: [10.1109/ICCV.1999.790410](https://doi.org/10.1109/ICCV.1999.790410)
- [9] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *International Journal of Computer Vision*, vol. 60, no. 2, pp. 91-110, 2004. doi: [10.1023/b:visi.0000029664.99615.94](https://doi.org/10.1023/b:visi.0000029664.99615.94)
- [10] H. Bay, T. Tuytelaars, and L. Gool, "Surf: Speeded up robust features," *Computer Vision and Image Understanding*, vol. 110, no. 3, pp. 346-359, 2008. doi: [10.1016/j.cviu.2007.09.014](https://doi.org/10.1016/j.cviu.2007.09.014)

