

# Dense 3D Measurement of the Near Surroundings by Fisheye Stereo

Nobuyuki KITA

AIST (National Institute of Advanced Industrial Science and Technology)  
Tsukuba Central 2, Tsukuba 305-8568 Japan  
e-mail:n.kita@aist.go.jp

## Abstract

*In this paper, dense 3D measurements are tested by using parallel stereo camera system with fisheye lenses when the measurement space is almost spherical at the distance of two or three times of baseline length from the stereo camera system. The 3D measurement method and the experimental results are introduced.*

## 1. Introduction

For the 3D measurement, the stereo vision system is set at fairly distant position from the measurement space. But in some cases, the measurement space becomes very close to the stereo vision system and moreover surrounds it. For example, in the case of surveillance system for the inside of a family car, the measurement space is whole inside of the car and the vision system must be set in the space. When a humanoid robot handles fairly big object like a newspaper by its dual arms, the measurement space is close to and surrounds the cameras in the head of the humanoid robot.

Some of stereo vision systems [1] can get 3D measurement in the almost spherical viewing area. But they assume that the distance between the stereo vision system and the measurable space is enough long comparing to the baseline length. Active stereo vision system enables the 3D measurement in the close and surrounding measurement space [2] but it is not possible to get the measurement for the whole measurable space simultaneously.

In this paper, the 3D measurement of close and surrounding measurement space is tested by the stereo vision system. Catadioptric system can get spherical view and be used for stereo vision system [3] but it lacks the central view, then it is not suitable for some applications like ones mentioned above. Some lenses like fisheye lens can get spherical view without lack of central view, but have some problems such as brightness loss at periphery and weak resolution power at any portion of wide view. But recently high quality fisheye lens with very little brightness loss at periphery and high resolving power at whole image was developed [4], then it is now possible to get clear and bright spherical images by fisheye camera. As for the stereo measurement, since the measurement space is very close to the stereo vision system as shown in the figure 1, the following problems arise.

(1) More accurate calibration for both inner and outer parameters is demanded comparing to the case that the measurement space is far.

(2) Since the viewing direction and distance from the cameras to the visual target are different between right

and left camera, the apparent shape variation of visual target can no more be negligible.

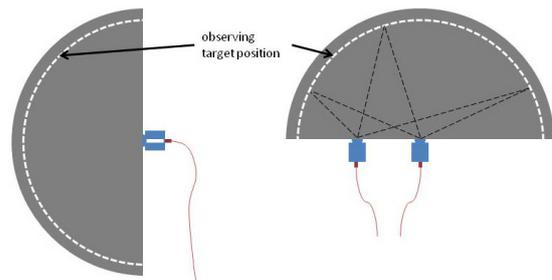


Figure 1. The measurement space and a stereo system.

In the rest of the paper, the calibration method performed to get very high accuracy is briefly explained in the section 2. The stereo measurement method considering the apparent shape difference between right and left images is proposed in the section 3. The evaluation experiments and the results are introduced in the section 4 and the summary is mentioned in the section 5.

## 2. Calibration

The calibration of inner and outer parameters of a stereo camera system is indispensable for the stereo measurement.

### 2.1. Inner parameters calibration

Though lots of automatic or semi-automatic calibration methods have been proposed for inner parameters of a wide view camera[5] [6] [7], we manually do it to get the ultimate accuracy. We made the special tool shown in the figure 2 by which a camera and a visual target can be precisely positioned on the L shape rail independently. Firstly image center (we assume projection center and radial distortion center is coincident.) is calibrated using the fact that a line in the 3D space is projected as a line on the image plane only when the line on the image center. Next the relations between the image height and the incident ray angle of the target center are measured while changing the positions of a camera and a target for four radial directions from the image center. The accuracy of the measurement of incident ray angle is about 0.002

degree and one of image height 0.1 pixel.

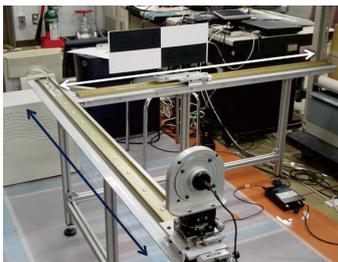


Figure 2. Special tool for calibration.

## 2.2. Outer parameters calibration

In order to get highly accurate outer parameters easily, we put targets within the wide space viewed from the fisheye camera and measured those 3D positions by a total station. From those 3D positions and the projected position of the targets, we estimated 3D pose and position of both cameras by bundle adjustment method.

## 3. Stereo Measurement

It is indispensable to take correspondences between right and left image features. Lots of descriptors have been proposed to measure the similarity between features. In this paper, a feature is simply described by a small image patch and the correspondences are taken by block matching at the possible position by epipolar constraint. Whole procedure of stereo measurement is as follows.

- 1) Feature selection.  
Though there are lots of filtering methods to select salient local features, here features are selected by manual or simple sampling with given interval on the right image.
- 2) Generate depth candidates.  
Under the assumption that the minimum and maximum distance from a camera to target are known, given numbers of depth candidates for each feature are generated between them.
- 3) Projected positions of all depth candidates on the left image are calculated.
- 4) Predict shape.

The variation of a feature's apparent shape on the left image from the right image is affected by not only radial distortion also the differences of viewing direction and distance. While the former variation can be predicted from the camera's inner parameter, the later can also be predicted by the homography if the neighborhood area of a feature can be approximated as a plane and its surface normal and the feature's 3D position are known. In the proposed method, since the 3D position of a depth candidate is known, only surface normal is unknown. In the experiment, targets are put on a sphere and the right camera put at the center of the sphere, then the surface normal of any feature is given as the direction toward the right camera.

- 5) Similarity measurement.  
Four SSD (Sum of squared difference) values between a predicted shape and the four image patch centering at the four neighbor pixels of the projected

position on the left image are calculated for each depth candidate. The similarity of the depth candidate is given by bilinear interpolation of those four SSD values.

- 6) Depth decision.

The depth of each feature is decided by the comparison of the similarities of its all depth candidates. If no similarity satisfies the given threshold, the depth of the feature is indefinite.

## 4. Experiments

### 4.1. Setup

The lenses developed by the cooperative research with NIKON Corporation [4] are used. They are the foveated lenses which have wide field of view and high resolution at the center. By the effort of lens design by NIKON, very little brightness loss at periphery and high resolving power of more than 80 line pairs /mm at whole image was achieved. Since the diameter of the image circle of our lens is 9.6 mm, Prosilica GE-560 camera whose image sensor size is 1.2 inch is used to capture whole image circle. These fisheye cameras are aligned as parallel with 150 mm baseline length.

### 4.2. Calibration

#### 4.2.1 Inner parameter calibration

The right fisheye camera is calibrated by the method described at 2.1. The image height is measured for incident ray angle from 0 degree to 81.5 degree with the 0.5 degree interval. The result is shown in the figure 3. The horizontal axis is incident ray angle (degree) and the vertical axis is the image height (mm). The differences between all the results for four radial directions are very small and negligible. As for the left fisheye camera, the image height measured at one radial direction was quite same as one of the right camera, measurement for other three radial directions are skipped.

#### 4.2.2 Outer parameter calibration

Outer parameters are calibrated as explained in 2.2. The targets are 6 points as shown in the figure 4. The size of input image is 1280 x 1280 pixels and the average of reprojection error is less than 0.5 pixel.

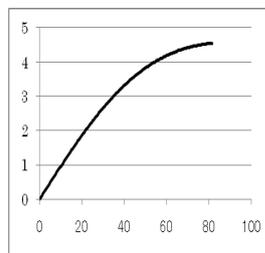


Figure 3. Projection curve.



Figure 4. 3D references.

### 4.3. Stereo images input

The spherical styrene form whose inner surface is covered with pieces of newspaper is put just in front of the right camera as shown in the figure 1 and the images are captured. From the 2048 x 2048 pixels images, 1280 x

1280 pixels part of ROI set as including whole image circle are downsized by 2x2 binning and 640 x 640 pixels images are sent to PC. The exposures of two cameras are set same. The obtained images are shown in the figure 5.



Figure 5. Obtained left and right images.

#### 4.4. Stereo measurement

##### Experiment 1

6 features are selected on the right image by manual as shown in the figure 6(b). The depth candidates for each feature are generated with minimum depth 200 mm, maximum depth 500 mm and the candidate number 200. The figure 6(a) shows the projected position of each depth candidates on the left image.

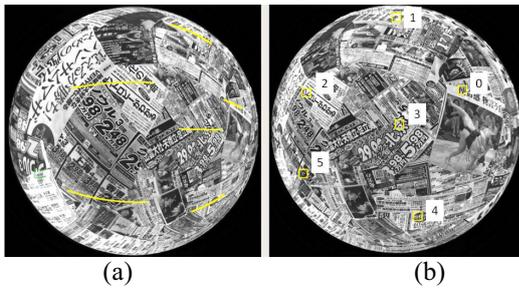


Figure 6. Epipolar constraint.

Similarity is measured for each of 6 features. Figure 7 shows the results for the feature 0. The top row is the plot of SSD values between the image patch with predicted shape and the image patch at the projected position on the left image. The image patch size is 21 x 21 pixels. The vertical axis is the SSD value and horizontal axis is the number of depth candidates. Number 0 is minimum depth. The next two rows show the image patch with predicted shape and at projected position for every 10s depth candidates. For the comparison, the results of similarity

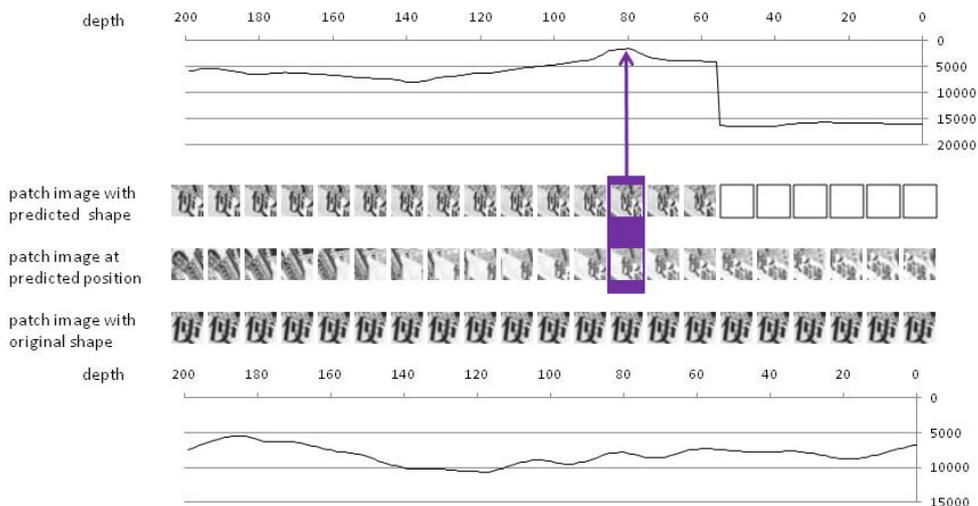


Figure 7. Similarity measurement of depth candidates of feature 0.

measurement without predicting shape variation are shown in the next two rows. In order to generate the predicted shape, the information of the pixels surrounding the original image patch is necessary in some case. In this experiment, the predicted shape is not generated when more than two times size of image patch information is necessary. This is the reason why the image patches of predicted shape for the depth candidates from 0 to 50 are empty.

The depth for each feature is decided as the depth whose candidate has the minimum SSD value when it is below of the given threshold 5000. Figure 8 shows the corresponded pair for 6 features. The top row is the number of features, next original patch image, next the patch image with predicted shape and bottom the corresponded patch image on the left image. The corresponded pairs are also depicted in the figure 7.

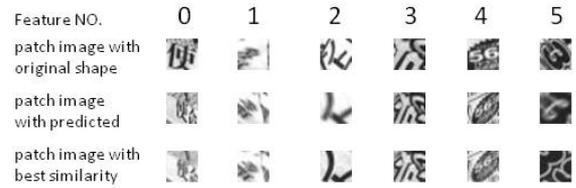


Figure 8. Matching results.

By the prediction of shape variation, the better similarity are gotten and correct correspondences are obtained for the 5 features within the 6 features.

##### Experiment 2

On the whole right image, features are selected every 24 pixels for both of horizontal and vertical directions. The rest of 3D measurement conditions are the same as experiment 1. Figures from 9 to 12 show the 3D measurement results. At each figure, the upper 2 show the decided 3D position of features and position and viewing axis of cameras from top and side and the lower two show the right and left image position of features whose depth are decided. Figures 9 and 10 are the results using and without using shape prediction for the SSD threshold value is 5000. Figures 11 and 12 are for the SSD threshold value is 2000.

For the SSD threshold 5000, within the 140 degrees viewing field, 3D measurements are obtained. By predicting the shape variation, the number of mismatching decreased and the accuracy of 3D depth improved. Since

the similarity is very high when predicting the shape variation, many features still obtained 3D measurements even for the very severe SSD threshold 2000 while the mismatching drastically decreases. Without shape prediction, very few features got 3D measurements.

## 5. Summary

The possibility of the dense 3D measurement of close and surrounding measurement space by the fisheye stereo system was shown.

Since the inner parameters of the fisheye lenses especially the radial distortion parameters which was obtained by the manual measurement in this paper have very high accuracy and can be regarded as ground truth, the calibration accuracy of some automatic or semi-automatic calibration methods will be evaluated by compared with them.

The stereo measurement method will be extended so that it can be applied for the case the surface normal information of the target is ambiguous. Specific application for the handling clothes by humanoid robot is progressing.

### Acknowledgment

This work was supported by a Grant-in-Aid for Scientific Research, KAKENHI (22240019). We are thankful to NIKON Corp. for letting us their original lenses.

## References

- [1] Abraham, S., Förstner, W.: Fish-eye-stereo calibration and epipolar rectification. *Photogrammetry and Remote Sensing* 59, 278–288 (2005).
- [2] S. Rougeaux, Y. Kuniyoshi, Robust real-time tracking on an active vision head, in: *IEEE International Conference on Intelligent Robots and Systems*, 1997, pp. 873-879.
- [3] T. Svoboda, T.Pajdla. Epipolar Geometry for Central Catadioptric Cameras. *IJCV*, 49(1):23-37, Kluwer August 2002.
- [4] K. Wakamiya, T. Senga, K. Isagi, N. Yamamura, Y. Ushio and N. Kita, "A new foveated wide angle lens with high resolving power and without brightness loss in the periphery", *Proc. SPIE* 6051, 605107 (2005); doi:10.1117/12.650967.
- [5] S.B. Kang, "Semi-Automatic Methods for Recovering Radial Distortion Parameters from a Single Image," *Technical Reports Series CRL 97/3*, DEC, Cambridge Research Labs, May 1997.
- [6] Swaminathan, R. and Nayar, S. K.: "Nonmetric calibration of wide-angle lenses and polycameras", *TPAMI*, 22(10):1172-1178, 2000.
- [7] R. Hartley and S. B. Kang: Parameter-free Radial Distortion Correction with Centre of Distortion Estimation, *Proceedings of the Tenth IEEE International Conference on Computer Vision*, pp.1834-1841, 2005.

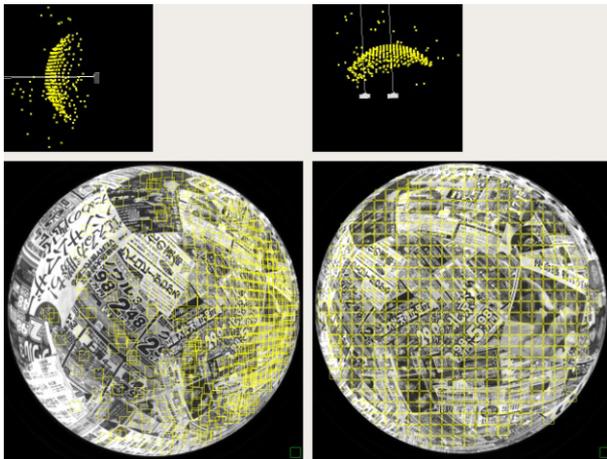


Figure 9. Result with prediction for SSD threshold 5000.

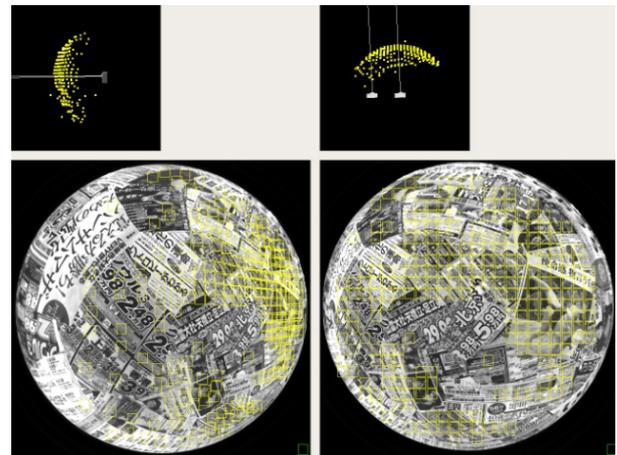


Figure 11. Result with prediction for SSD threshold 2000.

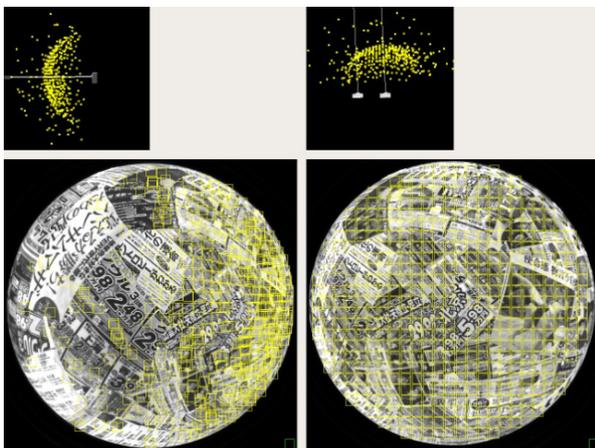


Figure 10. Result without prediction for SSD threshold 5000.

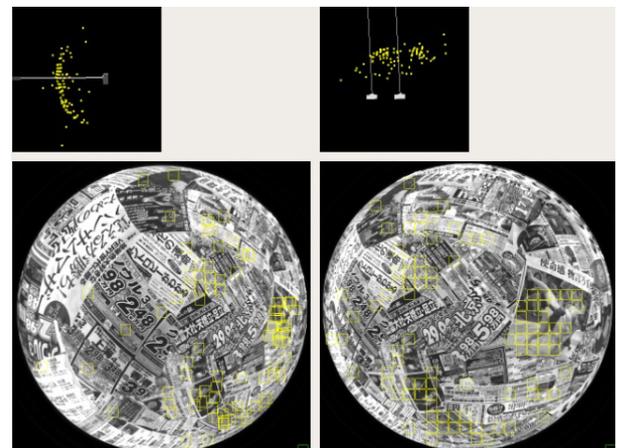


Figure 12. Result without prediction for SSD threshold 2000.