

Evaluation of 3D Analysis Error Caused by SVP Approximation of Fisheye Lens

Nobuyuki Kita

Intelligent Systems Research Institute, National Institute of Advanced Industrial Science and Technology, Tsukuba, Japan

Keywords: Fisheye Camera, Viewpoint Shift, Single Viewpoint, Error Evaluation, Self-localization, Stereo Calibration, Epipolar Constraint, Stereo Measurement.

Abstract: We have been doing research about visual SLAM, 3D measurements and robot controls by using images obtained through fisheye lenses. Though fisheye lens is non-single viewpoint (NSVP), we established the 3D analysis methods based on single viewpoint (SVP) model. In this paper, we call such substitution SVP for NSVP as “SVP approximation” and evaluate 3D analysis errors caused by the SVP approximation in the case using two different types of fisheye lenses.

1 INTRODUCTION

Fisheye lens has been often used to obtain views of very wide field in the computer/robot vision research. We also have been doing research about visual SLAM (Kita, 2008), 3D measurements (Kita, 2012) and robot controls (Kita, 2011a, 2013) by using images obtained through fisheye lenses. Though fisheye lens is non-single viewpoint (NSVP), we established the 3D analysis methods based on single viewpoint (SVP) model as same as that most of methods using fisheye lens (Abraham, 2005) (Schwalbe, 2005). Gennery proposed a lens model considering shift of viewpoint and a calibration method based on his lens model. He also evaluated the calibration error when using and not using viewpoint shift (Gennery, 2006). But no literatures evaluating the 3D analysis error when using and not using viewpoint shift can be found. In this paper, we call such substitution SVP for NSVP as “SVP approximation” and evaluate 3D analysis errors caused by the SVP approximation in the case using two different types of fisheye lenses.

Section 2 explains an overview of this evaluation. Section 3 describes about evaluation of self-localization error. Section 4 describes about evaluation of stereo calibration error. Section 5 describes about evaluation of stereo 3D measurements error. Section 6 summarizes.

2 OVERVIEW

Most of cameras have a single viewpoint as shown in the figure 1. But fisheye lens’s viewpoint usually moves on a line. Then it is a kind of axial model (Ramalingam, 2006). But it is a special case that the

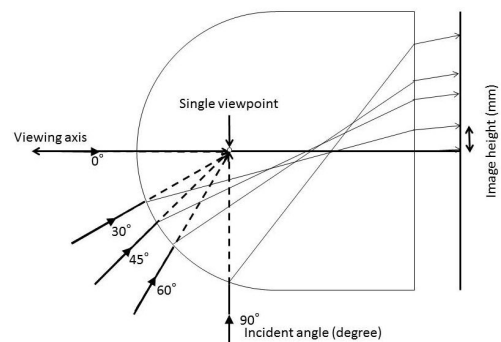


Figure 1: SVP model.

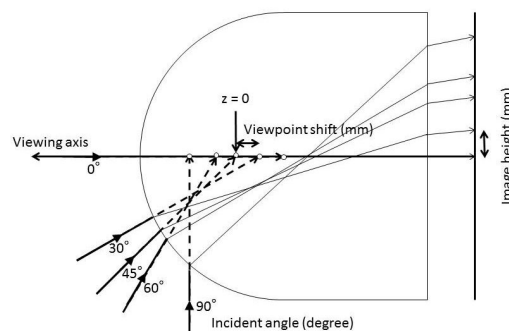


Figure 2: ShiftVP model.

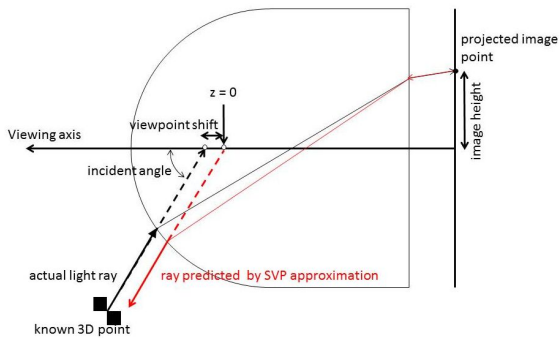


Figure 3: A light ray predicted by SVP approximation.

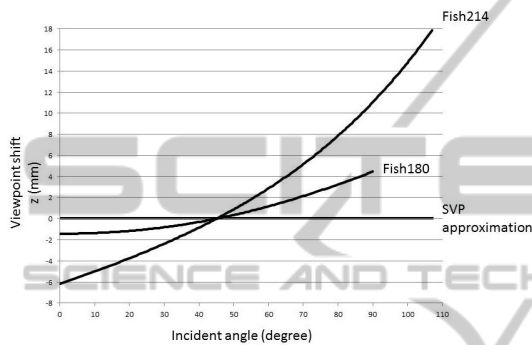


Figure 4: Incident angle and viewpoint shift.

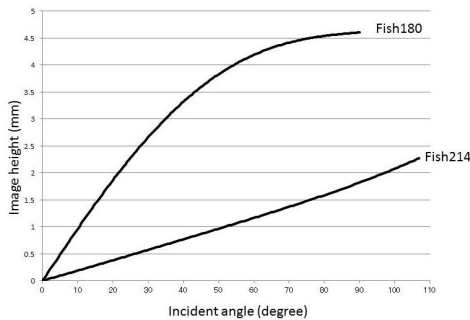


Figure 5: Incident angle and image height.

line is coincident with a viewing axis as shown in the figure 2. Here we call the projection model of fisheye lens as “shifting viewpoint (ShiftVP) model” following Gennery (Gennery, 2006) and represent it by the two relations, one for between incident angle and viewpoint shift and another for between incident angle and image height (figure 2). By using these two relations, a projected image point of a known 3D point to a camera whose pose is known can be obtained as shown in the figure 3. Furthermore a light ray projected backward from an image point is obtained. SVP model which we use for 3D analysis has the same relation for between incident angle and image height but a constant viewpoint position as shown in the figure 1. A red arrow in the figure 3

shows a light ray predicted by SVP approximation which has error from the actual incoming light ray.

In the paper, we use two different types of fisheye cameras Fish180 and Fish214 to evaluate SVP approximation error. The two relations of the cameras are shown in the figure 4 and 5. The lens of Fish180 is a foveated fisheye lens whose viewing angle is 180degree and image height curve is based on $f \cdot \sin \theta$. The circular viewing field is projected on the image whose size is 640×640 pixels. The lens of Fish214 is a spherical fisheye lens whose viewing angle is 214degree and image height curve is based on $f \cdot \theta$. The circular viewing field is projected on the image whose size is 1536×1536 pixels. These particular curves correspond to the fisheye lenses which we are using (Kita, 2011b) (Kita, 2011a).

3 SELF-LOCALIZATION

Evaluation of self-localization errors caused by the SVP approximation in the case using fisheye lens is explained. Self-localization is based on known landmarks. Several environments of landmarks are prepared. To avoid image processing error, the inputs to the localization method is the set of image coordinates which are numerically calculated from known landmarks positions and a known camera pose (true pose) by ShiftVP models. The algorithm of the self-localization method is a bundle adjustment. A camera is initially set to arbitrarily pose which is a little different from true pose. The camera pose is updated as decreasing the difference between image coordinates calculated by SVP model and inputs. We exploited the utility function of the Levenberg-Marquardt method from OpenCV. We prepared 9 environments of landmarks. The first one (Whole) has 33 landmarks which are set 2.0m far from the origin in the space $z < 0$. The second one (Inner) has 15 landmarks which are set as same as the first one but just around z axis. The third one (Outer) has 18 landmarks which are set as same as the first one but just not around z axis. For the following threes, the distance between the origin and landmarks are changed from 2.0m to 0.5m. For the following threes, the distance between the origin and landmarks are changed to 0.2m. The camera is set at the origin and gazing minus z direction. The 3D views on the columns 1 and 2 of the figure 6 depict the true pose of a camera by a white square and a line and the 3D positions of landmarks by \blacksquare . The order of rows follows the order of environments explanation above. The third and fourth column of

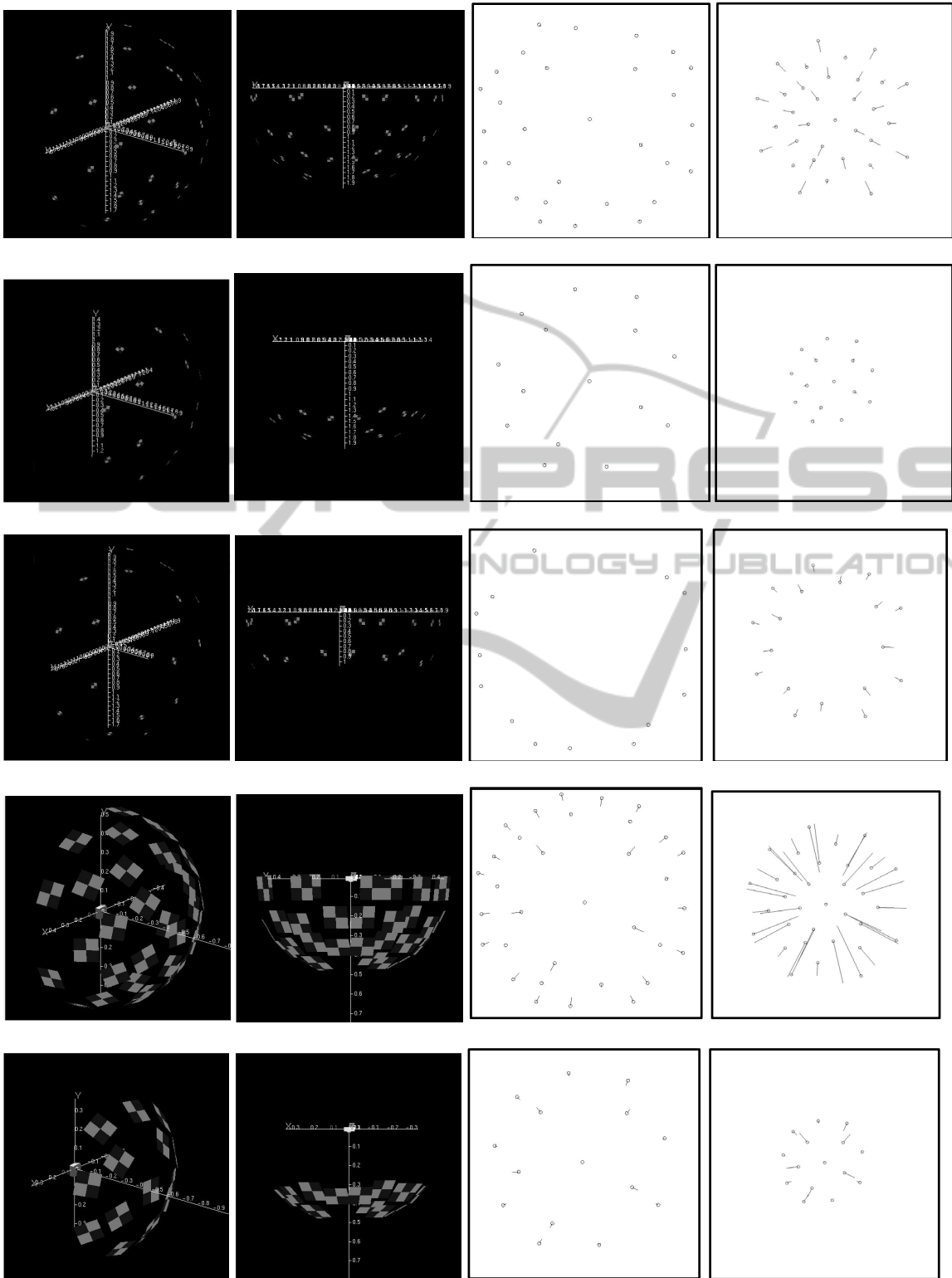


Figure 6: Self-localization error evaluation results.

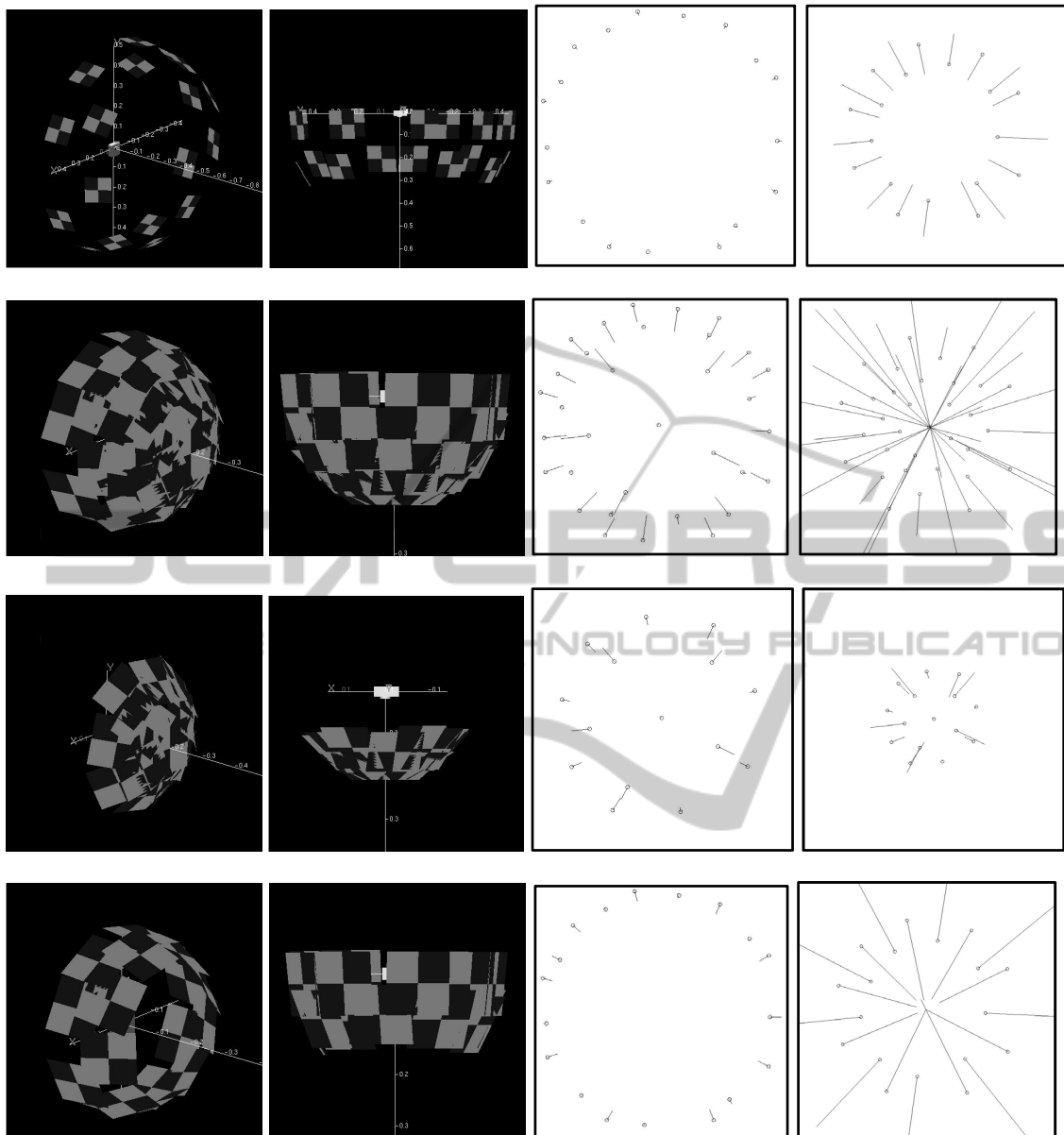


Figure 6: Self-localization error evaluation results (cont.).

the figure 6 depict the image positions of landmarks projected by ShiftVP model by small circles. Short lines depict the image residuals after the Levenberg-Marquardt iterations for the Fish180 and Fish214 respectively. The length of the short lines are magnified times 100.

Table 1 and 2 show the localization errors and average of residuals after iterations for Fish180 and Fish214 respectively. The axis error means the angle difference between the true and the estimated viewing axes. For Fish180, position errors of x and y directions which are the directions perpendicular to

the viewing axis are less than 0.1mm. Position errors of z direction which is the direction of the viewing axis for the Whole environments are from -0.11mm to -0.16mm. Position errors of z direction for the Inner environments are 0.24mm. Position errors of z direction for the Outer environments are -1.57mm, -1.67mm and -1.69mm. Axis errors are very small and maximum is 0.04degree. For Fish214, position errors of x and y directions which are the directions perpendicular to the viewing axis are less than 0.3mm. Position errors of z direction which is the direction of the viewing axis for the Whole

Table 1: Self-localization error for Fish180.

environment		estimation error				residual (pixel)
radius (m)	range	x (mm)	y (mm)	z (mm)	axis (degree)	
2.0	Whole	-0.0221	-0.0144	-0.1148	0.0011	0.0399
	Inner	-0.0660	0.0669	0.2372	0.0035	0.0281
	Outer	-0.0343	-0.0209	-1.5733	0.0037	0.0176
0.5	Whole	-0.0137	0.0136	-0.1614	0.0044	0.1642
	Inner	-0.0675	0.0683	0.2363	0.0140	0.1131
	Outer	-0.0246	0.0087	-1.6713	0.0122	0.0759
0.2	Whole	-0.0138	0.0136	-0.1631	0.0112	0.4136
	Inner	-0.0686	0.0695	0.2369	0.0355	0.2847
	Outer	-0.0245	0.0080	-1.6894	0.0308	0.1914

Table 2: Self-localization error for Fish214.

environment		estimation error				residual (pixel)
radius (m)	range	x (mm)	y (mm)	z (mm)	axis (degree)	
2.0	Whole	-0.0387	0.0276	-5.4709	0.0040	0.5764
	Inner	-0.2616	0.2729	0.2148	0.0137	0.1052
	Outer	0.0098	0.0112	-7.1621	0.0029	0.5007
0.5	Whole	-0.0397	0.0250	-5.6004	0.0159	2.3892
	Inner	-0.2654	0.2857	0.2195	0.0566	0.4237
	Outer	0.0108	0.0090	-7.3370	0.0112	2.0930
0.2	Whole	-0.0411	0.0153	-5.8982	0.0387	6.4501
	Inner	-0.2727	0.2928	0.2195	0.1452	1.0735
	Outer	0.0125	0.0012	-7.7370	0.0248	5.7515

environments are from -5.4mm to -5.9mm. Position errors of z direction for the Inner environments are 0.22mm. Position errors of z direction for the Outer environments are from -7.1mm to -7.7mm. Axis errors are very small and maximum is 0.15degree. Position errors in z direction for Fish214 are fairly large, while ones for Fish180 are much small. It is because the amount of viewpoint shift of Fish214 is much larger than one of Fish180. As for the landmarks distribution, the distance between the camera and landmarks looks not affect the localization errors, but the range of landmarks looks affect them. The fact suggests us to select landmarks so that their images are projected in the whole field of view to get the stable localization results.

Visual localization is actually affected by image quantization etc. Whether the errors caused by SVP approximation can be neglected or not depends on processing methods and application demands.

4 STEREO CALIBRATION

Evaluation of stereo calibration errors caused by the

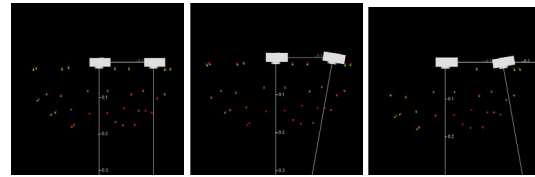


Figure 7: Stereo configurations.

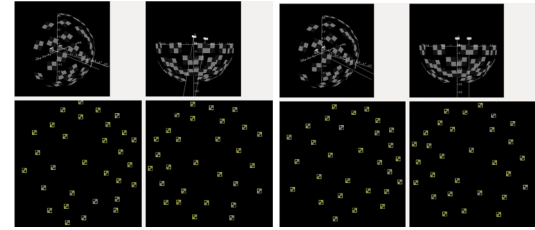
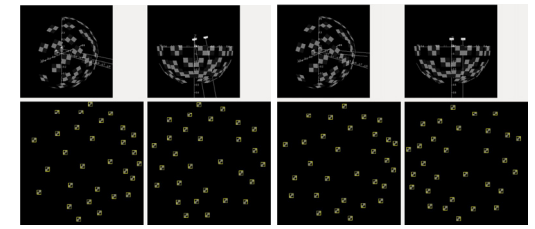
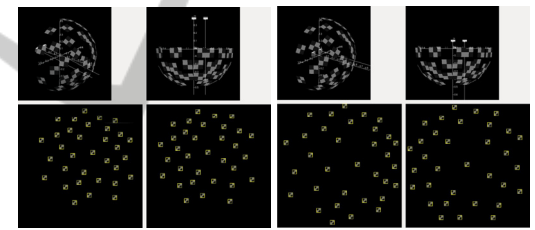
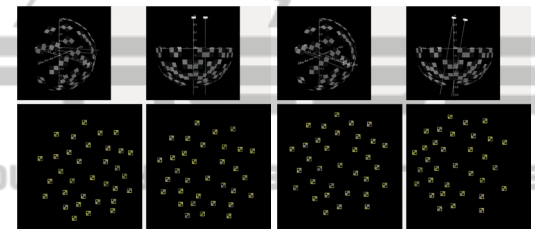
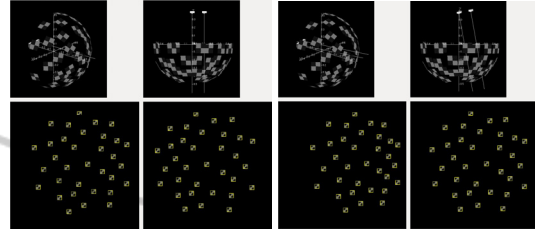


Figure 8: An example of input gathering for stereo calibration.

SVP approximation in the case using fisheye lens is realized. Inputs for the calibration procedure are the

set of image coordinates which are numerically calculated from known landmarks positions and known stereo camera poses by ShiftVP models. Three different types of stereo configurations are tested. As shown in the figure 7, the viewing axis of the left camera is set as parallel to one of the right camera for the parallel type, the viewing axis of the left camera is set 10degree inside for the vergent type, and the viewing axis of the left camera is set 10degree outside for the divergent type. For all types the baseline length is 0.15m. For the calibration targets, Whole environments of radius 2.0m, 0.5m and 0.2m are used. To gather the stereo calibration inputs, each stereo rig set at the 10 different poses. An example to gather the stereo calibration inputs for vergent stereo using 0.5m radius Whole environment is shown in the figure 8.

The algorithm of the stereo calibration follows the function “cvStereoCalibration” in OpenCV (Bradski, 2008). Table 3 and 4 show the stereo calibration errors and average of residuals after iterations for Fish180 and Fish214 respectively. The position errors mean the differences between correct and estimated left camera positions in the right

camera coordinate system. The axis errors mean the angle differences between correct and estimated left camera’s viewing axes. For Fish180, both of position and axis errors for three stereo configurations are very small. For Fish214, some values in the table could not be gotten by any bug of program, but soon it will be debugged and the table will be fill up. As long as evaluating the gotten results, both of position and axis errors for three stereo configurations are rather larger than those of Fish180.

5 STEREO MEASUREMENTS

Evaluations of stereo epipolar constraint and depth measurement errors caused by the SVP approximation in the case using fisheye lens are realized. As same as section 4, three different types of stereo configurations and Whole environments of radius 2.0m, 0.5m and 0.2m are used. The stereo camera is set as the right camera gazes minus z direction from the origin for Fish180 and (0.0,0.0,-0.1) for Fish214. Inputs for the evaluation are the set of image coordinates which are numerically calculated from known landmarks positions and a known stereo camera pose by ShiftVP models.

Figure 9 and 10 show the evaluation results for Fish180 and Fish214 respectively. In the figure of the first column, curved lines depict epipolar lines on the left images which are predicted from the image coordinates in the right image and the stereo parameters by using SVP model. Here the correct stereo parameters are used to clarify the errors caused by SVP approximation. Vertical lines in the figure of the first column depict the vertical errors between the actual image coordinates on the left image and the predicted epipolar lines. The length of the short lines are magnified times 100 in the figure 9 and magnified times 10 in the figure 10. Error2 is an average of the vertical errors in pixel. In the figures of the second and third columns, red and green dots depict real and measured 3D positions of the targets respectively. Error3 is an average of the 3D distances between real and measured 3D positions of the targets.

Vertical errors of epipolar lines increase when the targets become close to a stereo camera. They also seem to have relation with stereo configurations though it is not clear. In most of cases, the average sizes of errors increase in the order of parallel, vergent and divergent. For the Fish180, the averages of the vertical errors are between 0.004pixel to

Table 3: Stereo calibration error for Fish180.

enviro nment	stereo style	estimation error				residual (pixel)
		x (mm)	y (mm)	z (mm)	axis (degree)	
2.0	Parallel	-0.0070	-0.0013	-0.0051	0.0005	0.0430
	Vergent	0.0200	-0.0028	0.0208	0.0020	0.0429
	Divergent	-0.0510	-0.0025	-0.0106	0.0033	0.0430
0.5	Parallel	-0.0290	-0.0008	-0.0491	0.0061	0.1428
	Vergent	-0.0130	-0.0014	-0.0313	0.0030	0.1418
	Divergent	-0.0470	-0.0007	-0.0479	0.0076	0.1463
0.2	Parallel	-0.0160	-0.0157	-0.0749	0.0086	0.2696
	Vergent	-0.0320	-0.0041	-0.1835	0.0273	0.2518
	Divergent	-0.0020	-0.0025	-0.0106	0.0093	0.0430

Table 4: Stereo calibration error for Fish214.

enviro nment	stereo style	estimation error				residual (pixel)
		x (mm)	y (mm)	z (mm)	axis (degree)	
2.0	Parallel					
	Vergent	0.0720	0.0410	0.2882	0.0508	0.5903
	Divergent	-0.0390	-0.0355	0.0550	0.0492	0.5814
0.5	Parallel	0.3630	0.0781	0.1083	0.0474	2.3258
	Vergent	-0.1350	-0.0436	0.0002	0.1624	2.2687
	Divergent	0.6590	0.0198	0.0742	0.1062	2.3784
0.2	Parallel					
	Vergent					
	Divergent					

0.61pixel. For the Fish214, the averages of the vertical errors are between 0.03pixel to 9.3pixel. The errors of Fish214 are much larger than ones of Fish180 provably because of the differences of the shift amounts of viewpoint and the pixel sizes of camera images. For the every case, the error size greatly varies depending on the target position in the left image and looks some pattern which is interesting and will need more evaluations. If a stereo matching algorithm relies on an epipolar constraint, more than one pixel error will drastically affect the reliability of the matching results.

For the Fish180, the averages of the 3D position measurement errors are between 1.6mm and 8.6mm. For the Fish214, the averages of the 3D position measurement errors are between 6.4mm and 43.4mm. No clear relation between the error sizes and the distances to targets/stereo configurations is found.

6 CONCLUSIONS

The errors caused by SVP approximation are evaluated for self-localization, stereo calibration, epipolar constraints and 3D depth measurements.

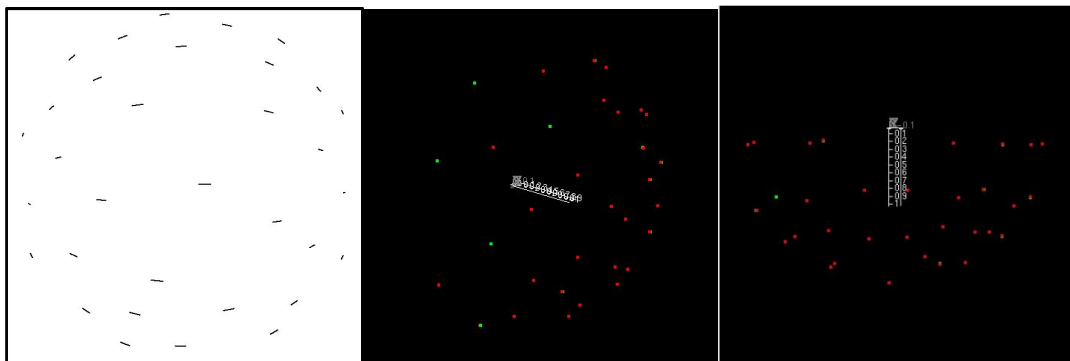
Visual analysis is actually affected by image quantization etc. Whether the errors caused by SVP approximation can be neglected or not depends on processing methods and application demands.

If a stereo camera of Fish214 is used for the stereo analysis of a target which is close to the stereo camera, shift amount of viewpoint must be considered.

REFERENCES

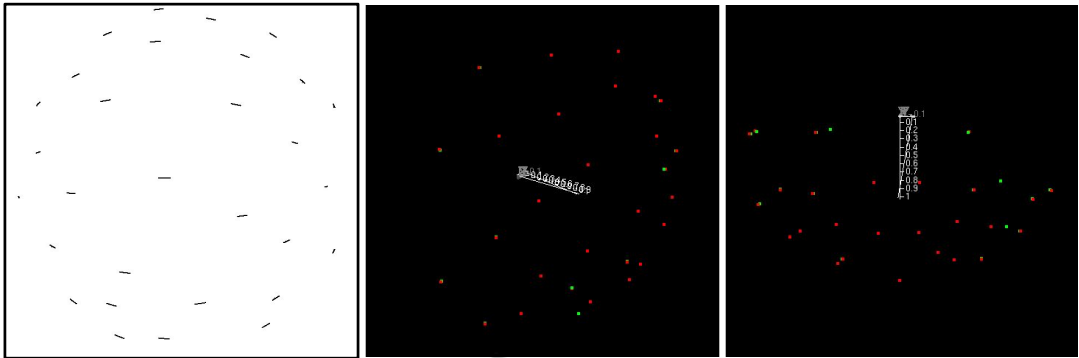
- Kita, Nobuyuki, 2008. Improvements of MonoSLAM performance by Wide Angle of View. DIA2008, O2-2 (in Japanese).
- Kita, Nobuyuki, 2012. 3D Shape Measurement of a Large Cloth close to a Fisheye Stereo, IEEE/SICE International Symposium on System Integration, 895–900.
- Kita, Nobuyuki, 2011a. Direct floor height measurement for biped walking robot by fisheye stereo, 11th IEEE-RAS International Conference on Humanoid Robots, 187–192.
- Kita, Nobuyuki, 2013. Obstacle Detection for a Bipedal Walking Robot by a Fisheye Stereo, IEEE/SICE International Symposium on System Integration, 119–125.
- Abraham, Steffen, and Förstner, Wolfgang, 2005. Fish-eye-stereo calibration and epipolar rectification, ISPRS Journal of Photogrammetry and Remote Sensing, vol. 59, pp. 278-288.
- Schwalbe, E., 2005. Geometric modelling and calibration of fisheye lens camera systems, 2nd Panoramic Photogrammetry Workshop.
- Gennery, B., Donald, 2006. Generalized Camera Calibration Including Fish-Eye Lenses. International Journal of Computer Vision 68(3), 239–266.
- Ramalingam, Srikanth, Sturm, Peter, and Lodha, K., Suresh, 2006. Theory and Calibration for Axial Cameras. Asian Conference on Computer Vision, 704–713.
- Kita, Nobuyuki, 2011b. Dense 3D Measurement of the Near Surroundings by Fisheye Stereo, IAPR Conference on Machine Vision Applications, 148–151.
- Bradski, Gary, and Kaehler, Adrian, 2008. Learning OpenCV.

APPENDIX

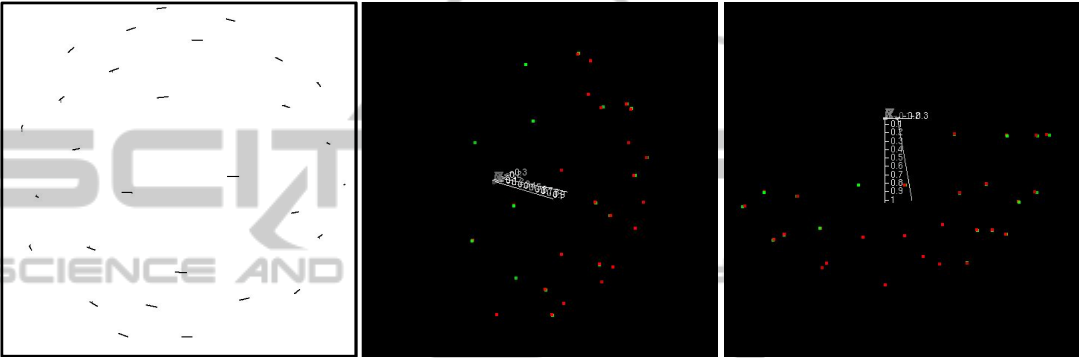


Error2 = 0.004pixel Error3 = 4.43mm

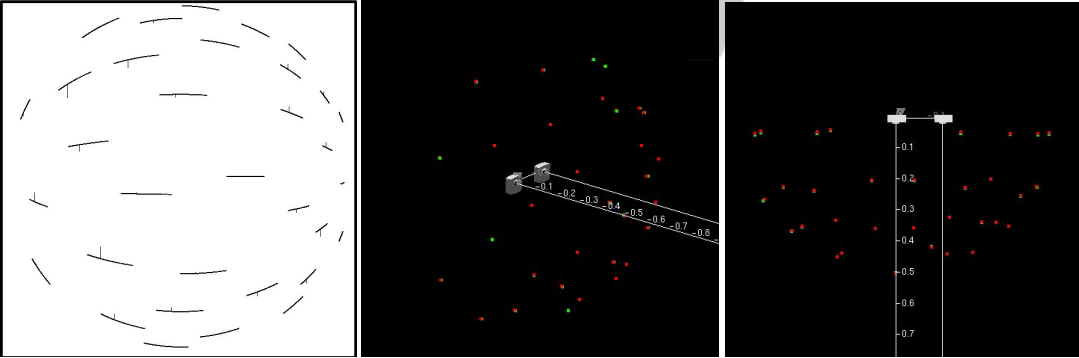
Figure 9: Stereo measurements error evaluation results for Fish180.



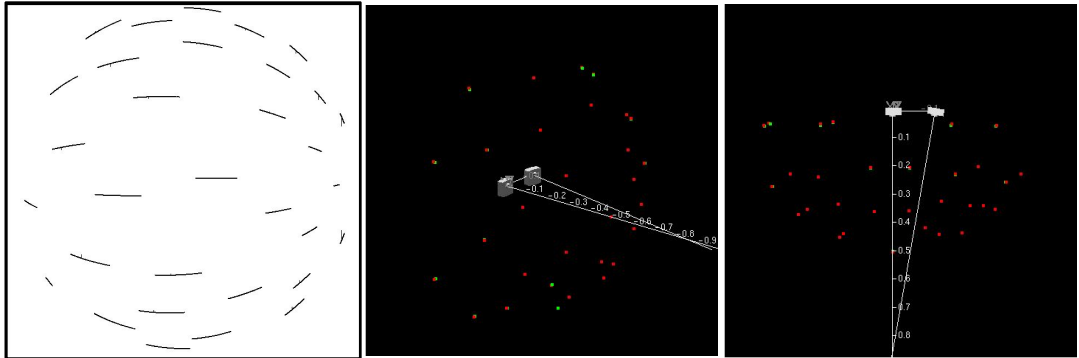
Error2 = 0.017pixel Error3 = 5.90mm



Error2 = 0.070pixel Error3 = 8.56mm



Error2 = 0.070pixel Error3 = 2.13mm



Error2 = 0.039pixel Error3 = 1.64mm

Figure 9: Stereo measurements error evaluation results for Fish180 (cont.).

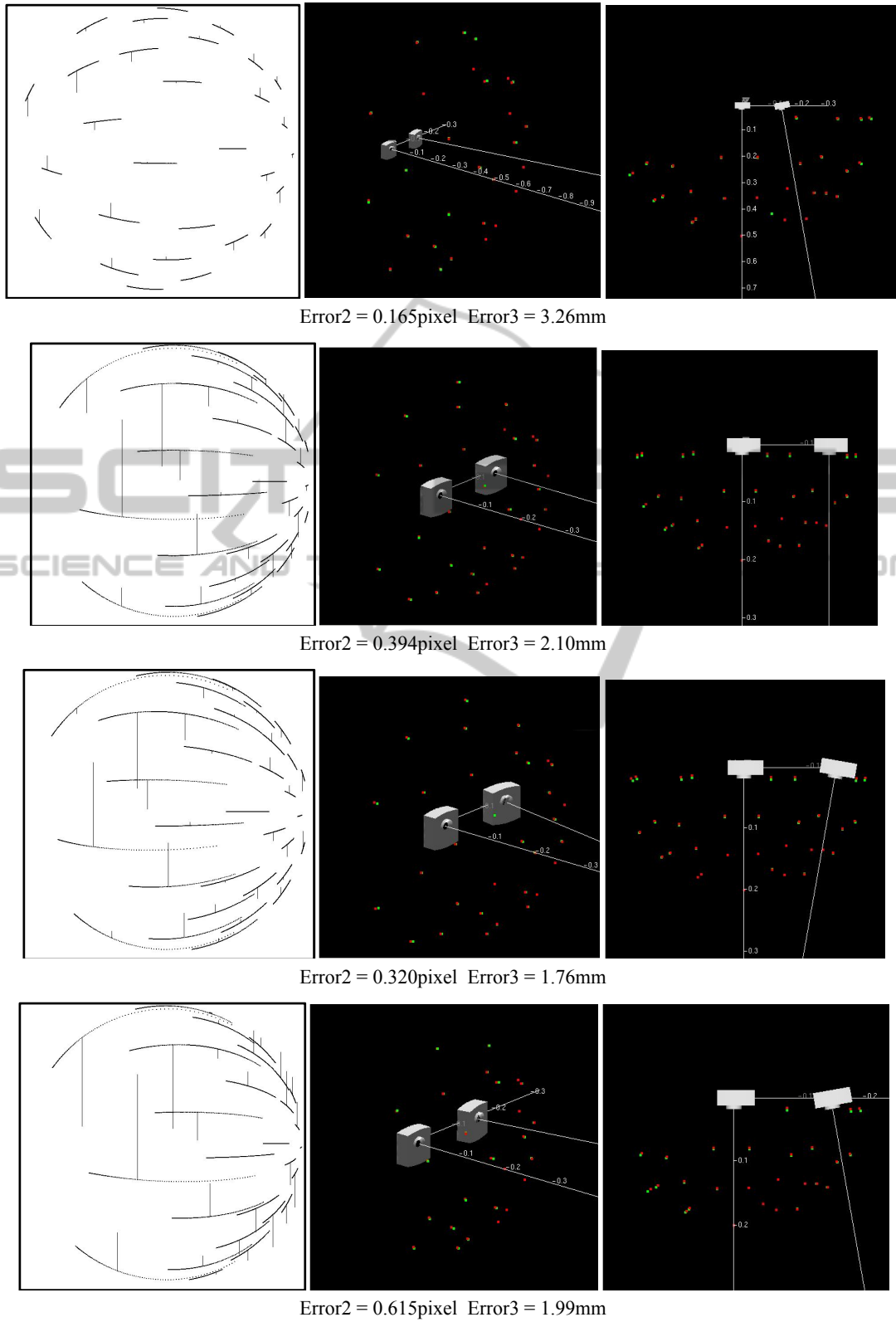
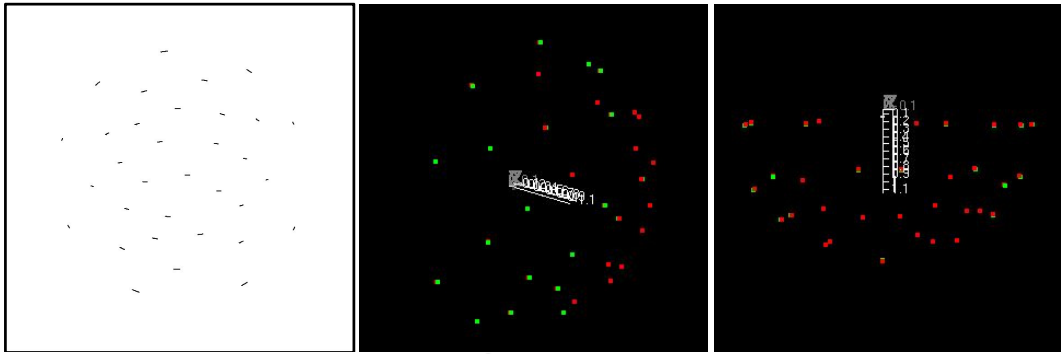
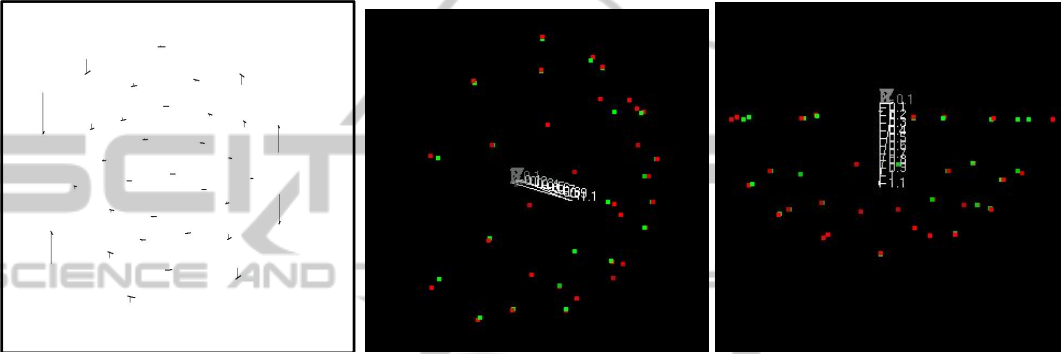


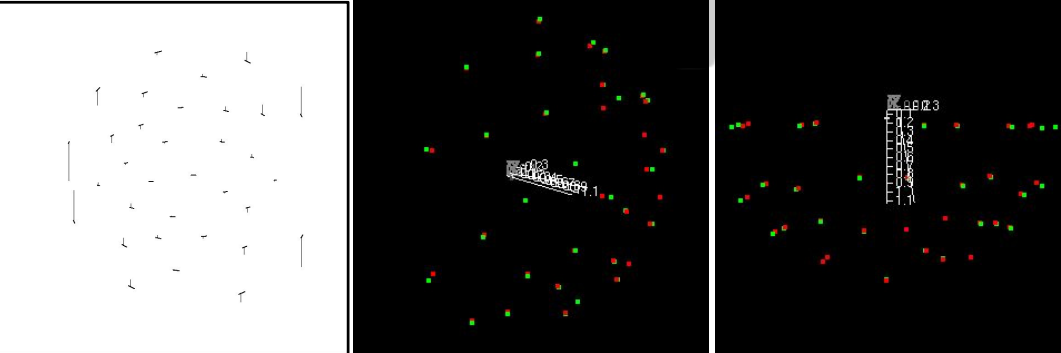
Figure 9: Stereo measurements error evaluation results for Fish180 (cont.).



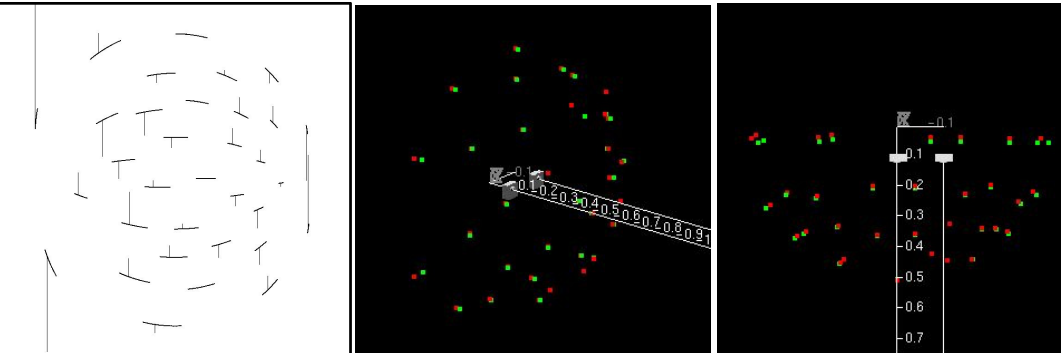
Error2 = 0.026pixel Error3 = 6.43mm



Error2 = 0.295pixel Error3 = 41.97mm



Error2 = 0.339pixel Error3 = 43.38mm



Error2 = 1.506pixel Error3 = 14.20mm

Figure 10: Stereo measurements error evaluation results for Fish214.

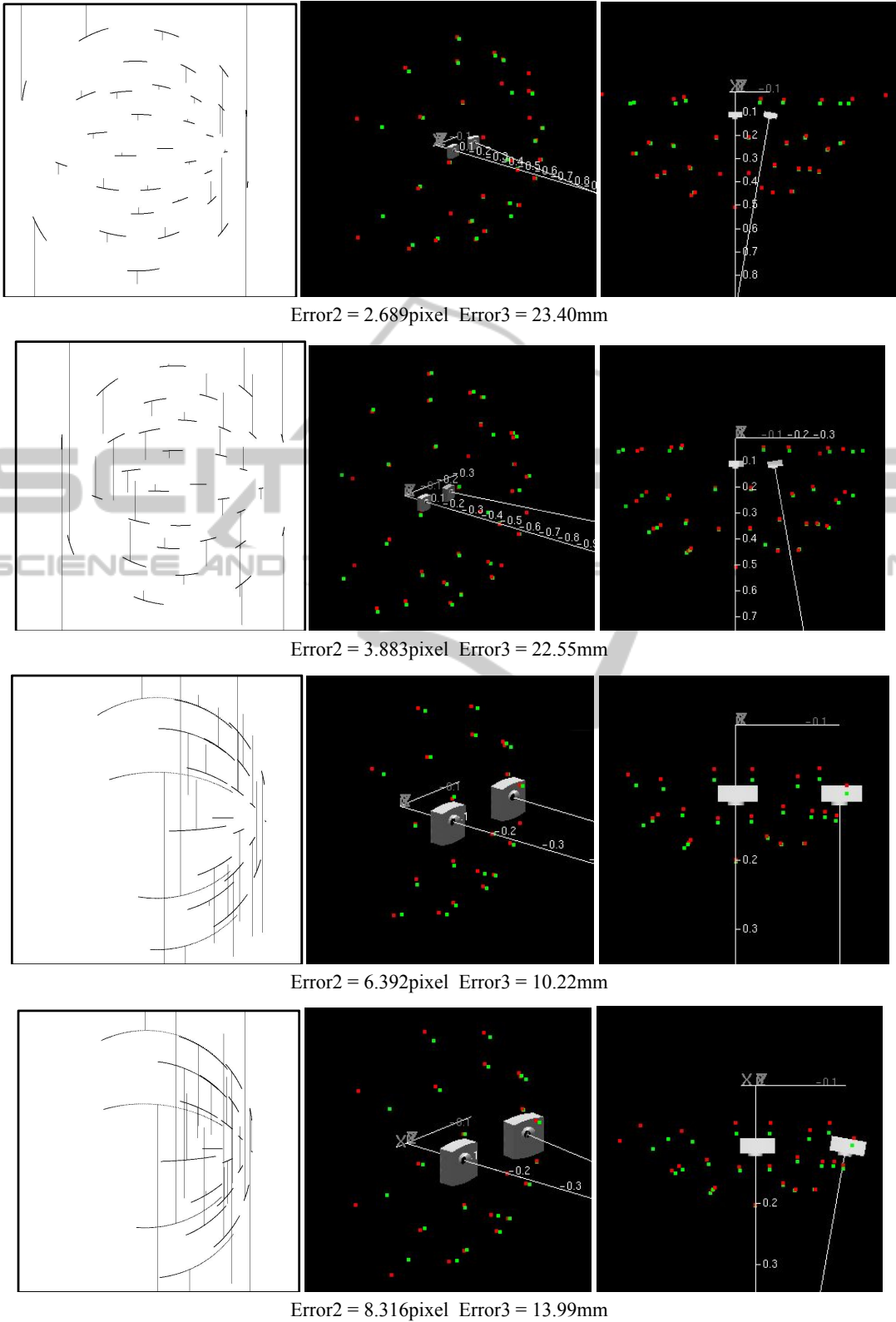
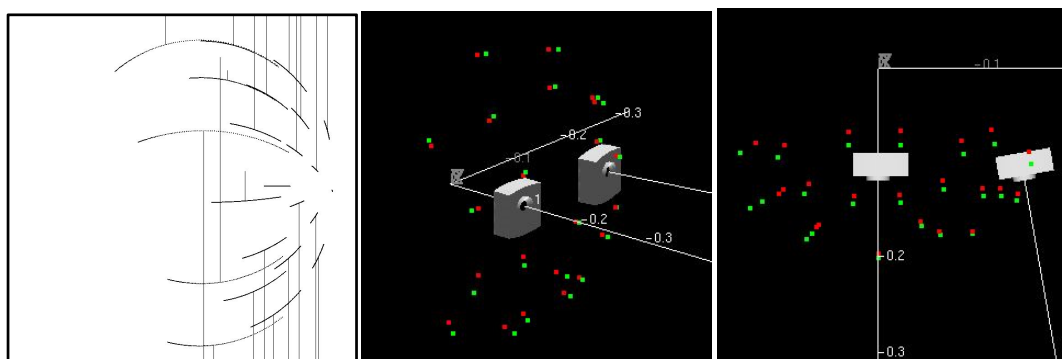


Figure 10: Stereo measurements error evaluation results for Fish214 (cont.).



Error2 = 9.329pixel Error3 = 13.07mm

Figure 10: Stereo measurements error evaluation results for Fish214 (cont.).

SCITEPRESS
SCIENCE AND TECHNOLOGY PUBLICATIONS