Accurate All-round 3D Measurement Using Trinocular Spherical Stereo via Weighted Reprojection Error Minimization

Wanqi Yin, Sarthak Pathak, Alessandro Moro, Atsushi Yamashita and Hajime Asama
Department of Precision Engineering
The University of Tokyo, Tokyo, Japan
Email: {yin, pathak, moro, yamashita, asama}@robot.t.u-tokyo.ac.jp

***Abstract*—Comparing to perspective cameras, the all-round 3D measurement of the environment can be done by spherical cameras in a more efficient way. However, the measurement using binocular spherical stereo has two singularity points at the epipoles of each spherical camera, where the measurement result gets extremely sensitive to the error when getting close to the epipoles and along the epipolar directions. This affects the accuracy of 3D reconstruction along with the epipolar directions. A three-way measurement method using three spherical cameras with trinocular spherical stereo setup is proposed in this paper to achieve accurate all-round 3D measurement. The improved accuracy of 3D measurement by the implementation of weighted reprojection error optimization was verified in experiments.**

***Keywords*-3D Measurement; Spherical Cameras;**

I. Introduction

3D measurement using only cameras has wide applications in the robotics field such as teleoperation and navigation. The information of the environment with low accessibility for humans, such as the disaster environment, can be measured and detected using cameras. The 3D models generated by cameras are helpful in the visualization of the environment and vision aided robot motion planning tasks. Currently, one of the most mature and commonly used methods for 3D measurement with cameras is stereo vision system [1] [2], with the setup of two calibrated perspective cameras. Due to the limited field of view (FoV) of the perspective cameras, extra post-processing such as stitching of the models is needed to perform the 3D measurement with larger FoV [3].

The characteristic that differs the spherical camera from the perspective camera is the different FoV and the shape of the image plane. Unlike the normal perspective cameras, the original image of the spherical camera is on the surface of a sphere, named as spherical image. For decreasing the processing complexity, the spherical images are commonly extended to an equirectangular image using equirectangular projection as shown in Fig. 1.

The spherical camera is similar to the omnidirectional camera [4], which has the $360° \times 180°$ all-round FoV and can capture the whole environment in one shot. In order to capture more area in 3D measurement, binocular stereo with spherical cameras, which is named as binocular spherical stereo was proposed [5]. Though the spherical camera can capture the all-round environment, the 3D all-round measurement via binocular spherical stereo is not able to be achieved due to the existence of singularity points position at the epipoles of each spherical camera. As shown in Fig. 2, with a certain measurement error for each spherical camera, there exists a corresponding confidence area for each point to be measured in the environment. Considering the existence of the measurement error in a certain angle, the measurement result of the point can be any point in the confidence area. The greater the confidence area is, the higher the error rate at that position. The confidence area depends on the position of the point to be measured and the confidence area is greater when getting close to the epipoles and along with epipolar directions, where the 3D measurement accuracy decreases. Therefore, it cannot take advantage of the wide FoV of the spherical camera in 3D measurement.

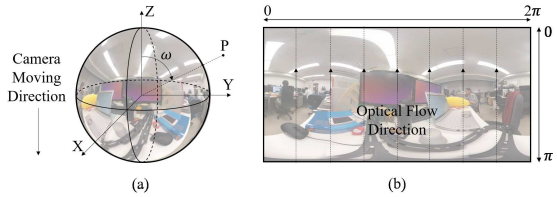


Figure 1. (a) The original spherical image. (b) The extended equirectangular image.

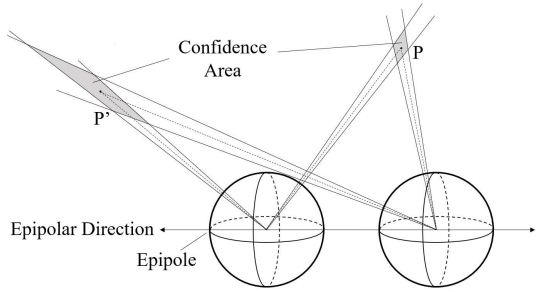


Figure 2. Low measurement confidence along with epipolar direction.

As can be seen from Fig. 2 that the confidence of measurement, which is illustrated by the confidence area, is proportional to the angle of the measurement direction with the baseline, which alingn with the epipolar direction. The closer the point to the epipolar direction, the greater the confidence area and the lower the measurement accuracy is.

If three spherical cameras are used for measurement [6], more than one set of binocular spherical stereo can be formed. Thus, there are more than one baseline, which means there are several values of angle from the point to the baseline. With a proper fusion algorithm that apply the weight according to its confidence when combining the measurement result from different stereo pairs can help in solving the singularity problem. The importance of the result during fusion should depends on the confidence, which can be reflected in the weight.

The three-way measurement method proposed in this paper is a novel joint measurement system aims to improve the all-round measurement accuracy via the trinocular spherical stereo setup with weighted reprojection error optimization implemented. In this paper, the related work is discussed, the setup of the system, the implementation of the algorithm is explained and the experiments result with quantitatively evaluated 3D measurement accuracy is shown to be improved.

II. Related Work

There are many methods proposed to achieve wide FoV 3D structure with spherical cameras. For the setup using multiple sets of binocular spherical stereo systems, various fusing methods are also proposed.

A 3D reconstruction method using the commercialized Richo Theta spherical cameras is proposed [7]. The combination of manually labeled ground truth and feature matching is used to achieve image rectification therefore to perform the dense 3D reconstruction. Since only two spherical images are used as the input of the method, the low accuracy along with epipolar directions and singularity at epipoles is not able to be avoided.

Kim [8] proposed a method using multiple calibrated spherical stereo pairs at different positions for 3D reconstruction. Stitching of 3D mesh and selection of the reliable surface is needed in this method to avoid the low accuracy area in the reconstructed 3D mesh. The combined final 3D mesh is not the fusion of multiple measurements but the selection of a most reliable surface at the overlapped area based on the statistic relationship between accuracy and the relationship of surface normal vector and camera viewing direction. Images taken at different positions is needed in this method. Thus it is difficult to reconstruct a moving environment unless a complex camera setup at different position can be designed.

Li [6] proposed a trinocular spherical stereo setup with two baselines perpendicular to each other. The algorithm to achieve all-round measurement with this setup is the fusion of two sets of binocular stereo measurement result by weighted average, where the weight is based on the geometry information. Though it can obtain the all-round measurement result, the weak epipolar constraint in this method lowers the overall measurement accuracy. The trinocular spherical stereo setup proposed in [6] is helpful in solving the problem exists in binocular spherical stereo with some depth result fusion algorithm that can gain the optimized depth.

Yin [9] tried to improve the performance of the 3D measurement based on the trinocular spherical setup proposed by Li [6]. Reprojection error minimization was implemented in their method to integrate the disparity maps during reconstruction. However, they ignored that the confidence of the measurement from two disparity maps generated along two different baselines are not identical, which means the importance of the reporjection error in the optimization should not be treated equally. The optimized result might have the trend to move to the less accurate direction while considering large the reprojection error near singulariy of the measurement.

When dealing with the fusion of several depth maps that do not agree with each other, Merrell [10] proposed a depth map fusion algorithm, which fuses several depth maps based on stability and confidence. The concept of stability of a 3D point is defined and evaluated base on the occlusion situation and the number of free space violation. In this fusion algorithm, the point in all the other views is re-rendered into the reference view. A hypothesis 3D point based on the confidence is generated and then the stability is evaluated. The fused depth map is the depth information with the most stable hypothesis 3D points at each pixel. This method is suitable when fusing more than two depth maps.

A multi-baseline spherical stereo model for 3D reconstruction was proposed in [11], in which a series of points within a certain distance range on a backprojection ray are backprojected to all the other images. The photometric cost was applied to each point and the point with the lowest cost is the optimal point for reconstruction. However, the aligned baseline results in the singularity along with epipolar directions.

III. Proposed Approach

A. Three-way Joint Measurement Model

The concept of the three-way measurement model proposed in [9] is based on the physical constraint that one point can only have one physical position in the 3D environment which is missing in [6]. In this paper, the fact that measurement confidence along two baselines is different and may seriously affect the reconstruction accuracy, which is not considered in [9] is realized and emphasized. In the implementation of optimization algorithm, the weight, which is related to the position of the point to be measured is introduced. With the weight applied, the optimization process can

consider the more accurate measurement result as a more important input while considering the measurement result with low accuracy less.

The concept is further systematized and completed for a more stable and robust performance in this work, which can be shown by the quantitative evaluations of the experiment results. The proposed three-way measurement model is proved to work in various setup configuration at random measurement positions in this work.

The model uses the trinocular spherical stereo setup as shown in Fig. 4. The three cameras are setup in equilateral right triangle shape. Different from the trinocular setup with perspective cameras where the third camera is mainly to provide more information to reduce the error caused by weak epipolar constraint, each spherical camera in the trinocular spherical stereo setup performs an important role in the all-round 3D measurement. The not aligned baselines ensure that one stereo pair can provide the information that is missing or not accurate in another pair. In the other hand, considering the existence of error, the depth map generated by the binocular stereo pairs might not agree with each other in the overlapped area, and neither of it can be used directly for reconstruction since both results may include error due to the weak epipolar constraint. The confidence can be implied by the position of the point to be reconstructed. With a fusion algorithm applying the weighted optimization according to the confidence, the proposed three-way joint measurement model is able to find out the position which is the closest to the accurate position.

The three-way joint measurement model is integrated and implemented as shown in Fig. 3. The model mainly includes four parts, image configuration and rectification, generation of the calculated points, generation of the reprojected points and weighted reprojection error optimization. The details of the algorithm of each part is explained in the following sections.

B. Setup, Image Configuration, and Rectification

As shown in Fig. 4, two pairs of binocular spherical stereo with baseline of length t perpendicular to each other are used to form the trinocular spherical stereo. Perpendicular baseline setup gives the optimal measurement since the areas with lowest confidence along the epipolar direction in one binocular spherical stereo set can be compensate by the area with highest confidence along with the epipolar direction in another binocular spherical stereo set. In the world coordinate system, camera C at the origin is the common camera in two pairs of binocular spherical stereo. It is identified as the reference camera in the measurement model. Camera R is the camera placed along x-axis and camera L is the camera placed along z-axis. Image $\mathrm{S_c}$, image $\mathrm{S_r}$ and image $\mathrm{S_l}$ is the spherical images taken by the three cameras respectively.

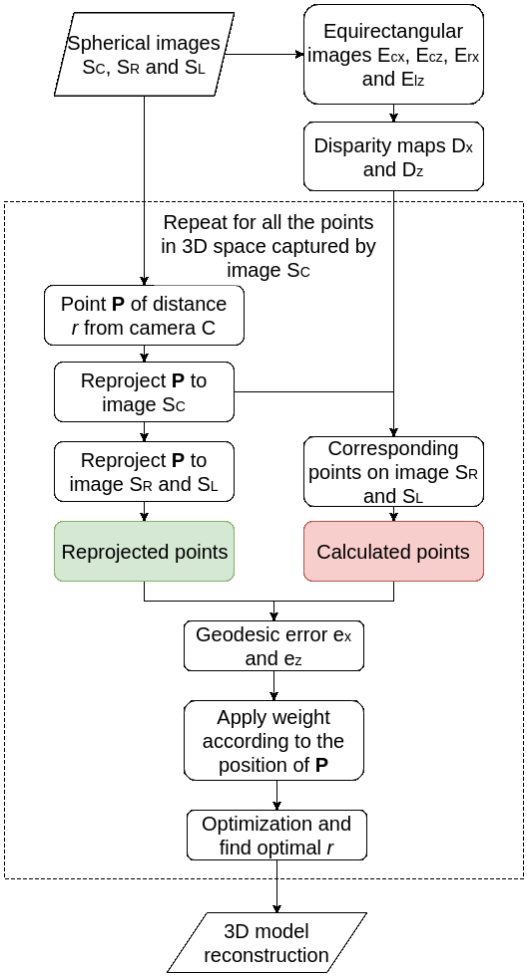


Figure 3. The flowchart of the three-way joint measurement model.

The stereo systems using spherical camera usually do not process the spherical images directly but project the spherical images to the equirectangular images as shown in Fig. 1, which have the planer image plane. Processing the equirectangular images significantly reduces the complexity of computation. The spherical image $\mathrm{S_c}$ is projected equirectangular image $\mathrm{E_c}$ and spherical image $\mathrm{S_r}$ and $\mathrm{S_l}$ is reprojected to equirectangular image $\mathrm{E_r}$ and $\mathrm{E_l}$ respectively.

Considering the characteristic of spherical images, when a set of stereo spherical images are projected to equirectangular images, the epipolar lines in the equirectangular images may not be straight lines, and the curves may vary according to the direction of the stereo baseline. As shown in Fig. 5(a), the epipolar lines are curved in this way when the baseline of the spherical stereo pair is along x-axis. It is difficult to search along the curved epipolar lines to find disparity values of each pixel. The curved epipolar lines in the equirectangular images can be rectified to straight lines and become parallel to the epipolar direction as shown in Fig. 5(b) [12]. This step significantly reduces

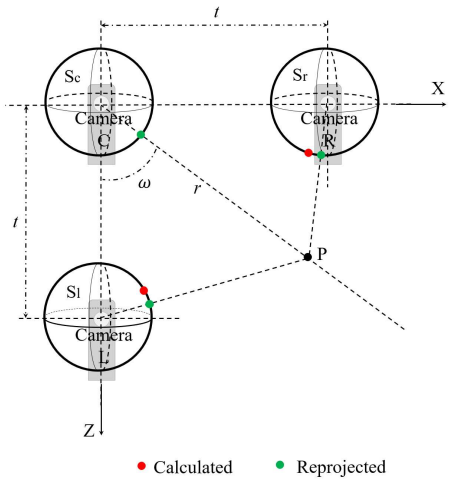


Figure 4. Trinocular spherical stereo and setup of the method.

the computation complexity of calculation of the disparity map.

In the stereo pair along x-axis in this setup, image $\mathrm{E_c}$ and image $\mathrm{E_r}$, which has the epipolar lines as in Fig. 5(b) needs to be rectified. The images after rectification are identified as $\mathrm{E_{cx}}$ and $\mathrm{E_{rx}}$, which has epipolar lines as in Fig. 5(b). Since the images of the stereo pair along z-axis which is $\mathrm{E_c}$ and image $\mathrm{E_l}$ originally has straight epipolar lines, no rectification is needed when generating $\mathrm{E_{cz}}$ and $\mathrm{E_{lz}}$ in this case.

C. Disparity Generation and Reprojection

The processing of images of this three-way joint measurement model are divided into the processing of equirectangular images and processing of spherical images.

The input images of the first step are the rectified equirectangular images $\mathrm{E_{cx}}$, $\mathrm{E_{cz}}$, $\mathrm{E_{rx}}$ and $\mathrm{E_{lx}}$. For the binocular spherical stereo pairs on x-axis and z-axis, disparity maps are generated using DeepFlow [13]. Disparity map $\mathrm{D_x}$ is generated using image $\mathrm{E_{cx}}$ and $\mathrm{E_{rx}}$ while disparity map $\mathrm{D_z}$ is generated using $\mathrm{E_{cz}}$ and $\mathrm{E_{lz}}$.

With the know disparity of a point in stereo system, the distance of the point from the reference camera center, r, can be calculated. From the disparity maps $\mathrm{D_x}$ and $\mathrm{D_z}$, the distance maps $\mathrm{R_x}$ and $\mathrm{R_z}$ can be calculated accordingly. For each pixel in each disparity map, the distance is calculated by solving the triangulation problem using Eq. (1) [12].

$$r = t \times \frac{\sin(\omega + d)}{\sin(d)}, \tag{1}$$

where t is the baseline length, d is the disparity value of the pixel in the disparity map in rad and ω is the angle between the epipolar direction and the direction of point P as shown in Fig. 4.

The next step is to process the spherical images $\mathrm{S_c}$, $\mathrm{S_r}$ and $\mathrm{S_l}$. The corresponding points of $\mathrm{S_c}$ in $\mathrm{S_r}$ and $\mathrm{S_l}$ are calculated using $\mathrm{D_x}$ and $\mathrm{D_y}$. The corresponding points are identified as calculated point, which is illustrated as the red point in Fig. 4.

For each point on $\mathrm{S_c}$, assume it can be reconstructed to a point P in the environment, with a distance r_{assume} to the camera center, an assumed distance map can be constructed as $\mathrm{R_{assume}}$. Reproject each assumed point P back to the $\mathrm{S_c}$, $\mathrm{S_r}$, and $\mathrm{S_l}$ at the same time and identified as reprojected point, which is illustrated as the green point in Fig. 4.

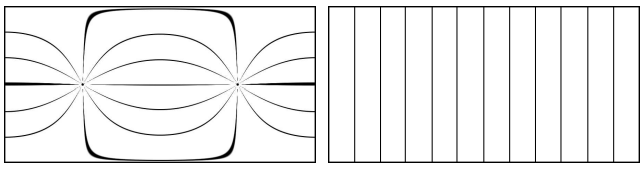


Figure 5. The epipolar lines in equirectangular image (a) before and (b) after the rectification for the spherical stereo pairs with baseline along x-axis.

D. Weighted Optimization and 3D Reconstruction

Weighted optimization is applied for the fusion of distance maps. Since the accuracy confidence of the distance maps is unknown, any fusion algorithm based on the statistical method cannot ensure that the fused distance map is the optimal result, but this can be achieved by optimization based on the physical constraint proposed in [9]. Since the confidence of each disparity map is not identical neither the importance of the reprojection error in each direction is identical. Considering the geodesic reprojection error in x-axis and z-axis equally during optimization may include the error in the final result. To solve this problem, the weight is applied to the reprojection error during optimization to take care of the confidence of disparity maps and maintain the balance between the errors in two different directions.

In the ideal case, if the assumed point P is at its true position, where r_{assume} is equal to the true value r_{true}, the calculated point should overlap with the reprojected point in both $\mathrm{S_r}$ and $\mathrm{S_l}$. The length of the curve between calculated point and reprojected point on $\mathrm{S_r}$ and on $\mathrm{S_l}$ can be represented by the geodesic reprojection error as illustrated in Fig. 6 and identified by e_x and e_z. Ideally both e_x and e_z should be zero. Thus, the goal of optimization is to optimize the $\mathrm{R_{assume}}$ so that the reprojected point gets as close as possible to the calculate point for each point in $\mathrm{S_c}$. If the geodesic reprojection error reaches its minimum, the corresponding $\mathrm{R_{assume}}$ should reach optimal value according to Eq. (2),

$$r = \mathrm{argmin}(w_1 e_x(r) + w_2 e_z(r)). \tag{2}$$

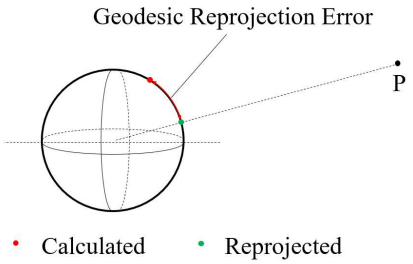


Figure 6. The definition of geodesic error in this method which is the length of curve between the calculated point and reprojected point.

Levenberg-Marquardt method is used in optimization and the initial value of $\mathrm{R_{assume}}$ for optimization is the average value of $\mathrm{R_x}$ and $\mathrm{R_z}$.

The w_1 and w_2 in Eq. (2) is the weight applied during optimization. It is correlated to the point position from epipolar direction, ω shown in Fig. 4, where ω is within the range of $(0,\frac{\pi}{2})$. The weight is calculated by Eqs. (3) and (4),

$$w_1 = \frac{2\omega}{\pi}, \tag{3}$$

$$w_2 = 1 - w_1. \tag{4}$$

With the optimized $\mathrm{R_{assume}}$ known, the 3D reconstruction can be achieved.

This three-way joint measurement method is able to get the all-round 3D measurement without information loss at the areas near epipoles and along with epipolar directions. By considering the result from three cameras at the same time, not only the epipole areas, the measurement accuracy of the whole environment can be improved by optimization.

IV. Experiment

The main experiment environment was an indoor classroom model[1] in Blender, three spherical cameras were arranged at corresponding positions according to the setup plan. The four epipolar directions in this experiment environment were along x-axis and z-axis, which cross the window, the wall, the ceiling, and the floor.

Blender was used to render the three equirectangular images using the 360°×180° all-round FoV spherical cameras and rotation was applied to the images rendered using the reference camera C. The input images of the three-way measurement model are shown in Fig. 7, which can be seen that the epipolar directions align for the camera pair in each direction. The epipoles of the x-axis stereo pair are illustrated in the images of the stereo pair along the z-axis and vice versa. The disparity maps generated from the input images in Fig. 7 are shown in Fig. 8.

[1]Available under CC0 license in http://www.blender.org.

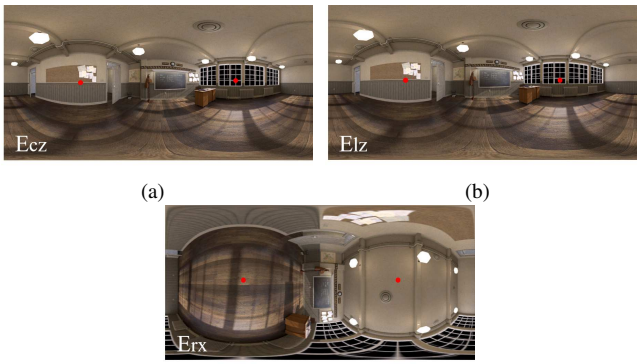


Figure 7. Input images of the experiment. The red points marked in (a) $\mathrm{E_{cz}}$ and (b) $\mathrm{E_{lz}}$ are the epipoles of the stereo setup along x-axis and the red points marked in (c) $\mathrm{E_{rx}}$ are the epipoles of the stereo setup along z-axis. In (c), $\mathrm{E_{rx}}$ is the equirectangular image after rectification.

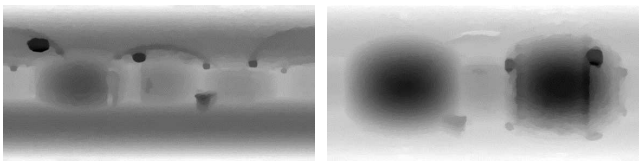
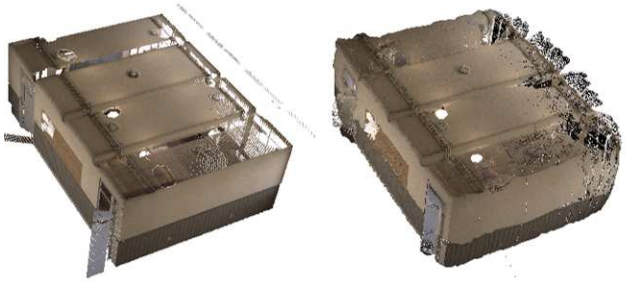


Figure 8. The disparity map generated by Deepflow from the input equirectangular images.

The advantage of using Blender to generate the input image is that it can generate the depth ground truth of the 3D model at the same time. From the depth map, the ground truth in the point cloud format can be calculated in the same way of 3D measurement, while getting the ground truth in real environment is almost impossible. The ground truth value makes the quantitative evaluation of the measurement result possible and the performance of the algorithm can be shown intuitively by the quantitative evaluation.

The objective of the experiments are to test the performance of the proposed three-way measurement model in improving the all-round measurement accuracy with various configurations, and under various conditions. Several sets of the experiments were performed. First of all, qualitative evaluation was done to show the improvement of the measurement accuracy intuitively.

In 3D measurement the baseline length and measurement position are two main factors that usually affect the measurement result. The purpose of the experiments is to show that with all kinds of factor settings, the proposed measurement model can improve the measurement accuracy. The first quantitatively evaluated experiment was to show the improvement of the overall measurement accuracy using input images sets is independent of the baseline setting. Another experiment was to show that the improvement of measurement accuracy is independent of camera position, and environment. Furthermore, the worst case evaluation was



(a) (b)

Figure 9. (a) the 3D model generated from ground truth value. (b) The 3D reconstruction result via the proposed three-way measurement method.

done to show the improvement at two epipolar directions.

The evaluation of the experiment result was based on the mean and standard deviation of the error at every reconstructed point, where the error was defined as the difference in distance between the measurement result and the ground truth value. The experiment evaluated the effectiveness of the proposed method.

A. Visualization of Measurement Results

In this session, figures of the reconstructed 3D model are shown for intuitive visualization of the improvement by the implementation of optimization. In Fig. 9, the module reconstructed using ground truth value and by the proposed reconstruction model is shown. It is well reconstructed in the overall shape, however, the corners and the connections of different planes are not as sharp as the ground truth model. Some distortion due to the occlusion and aperture problem can be seen from the figure.

The area where accuracy improves the most is along with epipolar directions. As illustrated in Fig. 7, there are four areas in total that should encounter the singularity during measurement. In Fig. 10, a view that can visualize all the four areas at the same time and the error are colored according to the heat-map color scale provided. The rectangular visualized in Fig. 10 is the sliced side view of the classroom model, where the top and bottom surface is the ceiling and floor, the surfaces at sides are walls. It can be clearly seen that after performing optimization, the accuracy improved significantly.

The distribution of the measurement error among all the points before and after optimization of the experiment result shown in Fig. 10 is shown in Table I. The difference in the amount of points within a certain error range before and after optimization is shown in percentage. Before optimization, there exists measurement error up to 3.3m and only 88.21% of the points are within the error range of 0.5m, the proposed method successfully eliminated the error within 0.7m and 94.18% of the points are within the error range of 0.5m.

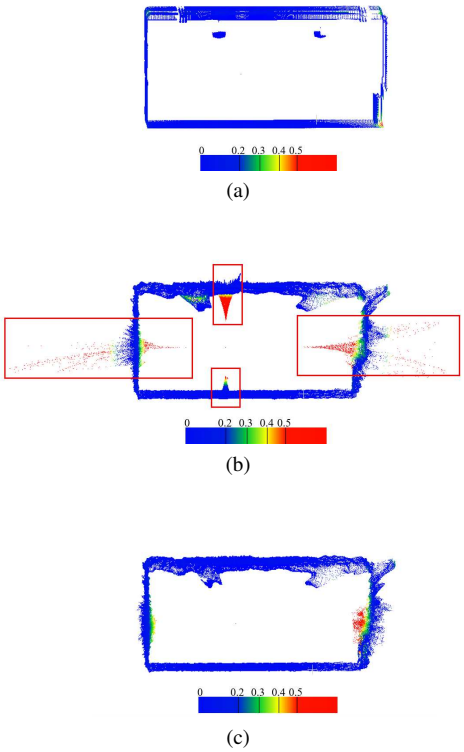


Figure 10. The sliced cross section view of the 3D model along y-axis, which can include the measurement result along with all epipolar directions. (a) The ground truth; (b) Before optimization; (c) After optimization.

B. Quantitative Evaluation of Measurement Results

The baseline of the stereo setup is usually adjusted according to the size of the environment to be measured. In order to prove that the performance of our proposed measurement model is independent to the change of baseline length, the experiment was done with same input image resolution at same camera position with various baseline length. In this experiment, the resolution was controlled at 1500 $\times$ 750 pixel for all cases while the baseline increased from 0.3m to 0.5m with a step of 0.1m.

In Fig. 11, the mean values of the measurement error before and after optimization are shown by the color bars and the standard deviation values are shown by the error

Table I
DISTRIBUTION OF MEASUREMENT ERROR IN PERCENTAGE

Error (m)	Unoptimized	Optimized
<0.1	14.16	16.13
<0.2	73.32	78.63
<0.3	88.21	94.18
<0.4	93.43	98.24
<0.5	96.08	99.69
<0.6	97.11	99.97
<0.7	97.73	100.00
<0.8	98.95	100.00

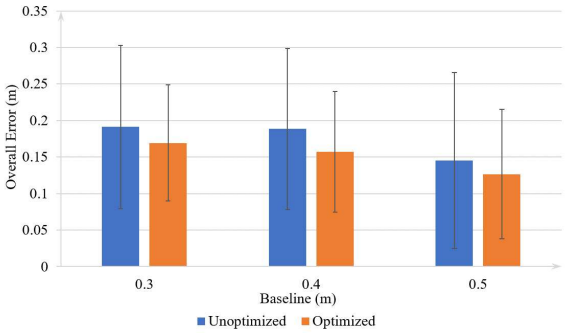


Figure 11. The mean and standard deviation of overall error with various baseline length. Resolution: 750 × 1500 pixel.

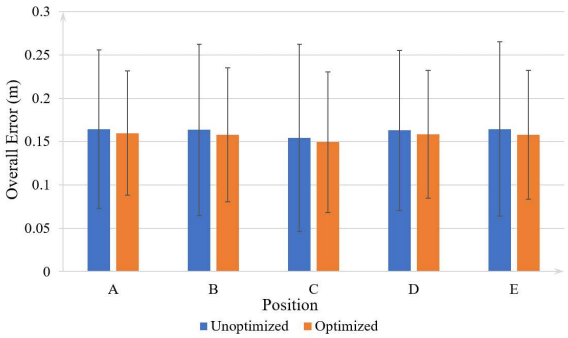


Figure 12. The mean and standard deviation of overall error from the images taken at various position. Resolution: 750 × 1500; Baseline: 0.3m.

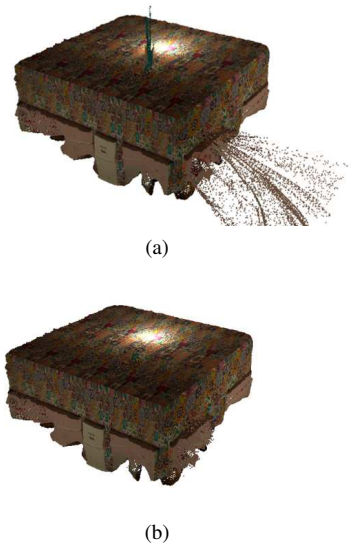
bars. As shown in Fig. 11, the accuracy and precision improved after performing optimization for all the cases with various baseline length. The effectiveness of the algorithm can be proved by the experiments. Within the reasonable range of baseline which depends on the environment to be measured, the proposed measurement model can improve the the measurement accuracy in spite of the change of the baseline length.

During the measurement process, the position to perform the measurement may change according to the environment structure and restriction on the setup. In order to prove the performance of the proposed measurement model at random positions, the experiment was performed at various positions in the classroom for the all-round 3D reconstruction task.

Since if the camera is too close to the ground, ceiling or walls, the reconstruction accuracy might be affected. The position of the cameras were restricted within the area that is 0.5m far from the ground, ceiling, and walls. Five random positions were generated in side the restricted area, with the image resolution 750 × 1500 pixel and basline value 0.3m. The experiment result is shown in Fig. 12. The color bars indicate the mean error in meters before applying weighted optimization algorithm proposed by us and the error after applying the proposed weighted optimization algorithm. The error bars indicate the standard deviation of the error in meters.

From the quantitative evaluation result shown in Fig. 12, the average and standard deviation of the error decreased in all the test cases at different positions in the classroom, especially the standard deviation, which indicates the precision of the measurement and 3D reconstruction. It can be proved by this experiment that the performance of the three-way measurement model is independent of the camera pose within the reasonable measuring area.

With the result of the above experiments, we can infer that the proposed three-way joint measurement model can successfully perform the measurement task with superior accuracy and precision in any random environment with



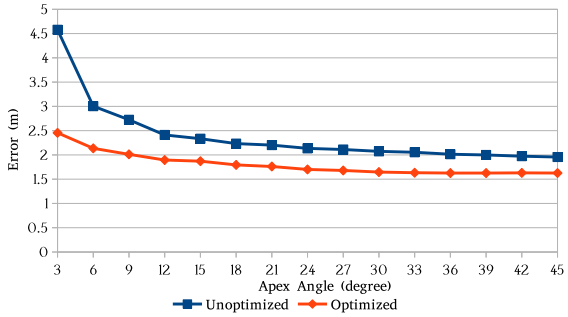
(a)

(b)

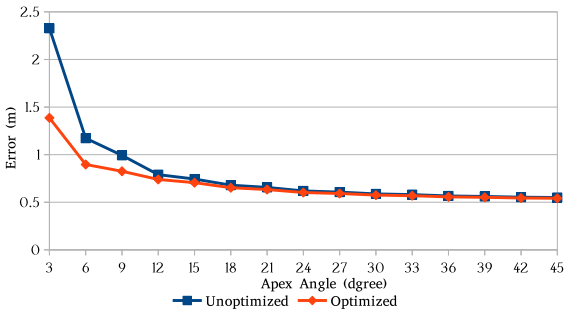
Figure 13. The reconstruction result of a cafe environment (a) before and (b) after applying the proposed measurement model.Resolution: 750 × 1500; Baseline: 0.6m

suitable baseline length at any camera pose. To test and show this, a successful reconstruction of a different environment is shown in Fig. 13. It proved the guess and showed that the improvement by proposed measurement model is independent of the environment.

Furthermore, the worst cases while performing the measurement with the proposed three-way joint measurement model, which is the measurement along epipolar directions and areas near epipoles were analyzed. The analysis was done for the pixels within the area inside a symmetric double cone on the reference spherical image, where the apex was at the position of reference camera. The apex angle varies from 3 dgree to 45 degree with the step of 3 degrees, which can cover the whole epipole area. The mean error of the points in the reconstructed model corresponding to the pixels within the double cone area was calculated and shown in Fig. 14 for both x-axis and z-axis epipolar directions. It can be clearly



(a)



(b)

Figure 14. (a)Worst case analysis along epipolar direction in x-axis and (b)worst case analysis along epipolar direction in z-axis.Resolution: 750 × 1500; Baseline: 0.6m

seen from Fig. 14 that along epipolar directions at areas around epipoles, the improvement of accuracy is significant.

The experiments successfully tested the performance and effectiveness of the three-way joint measurement model. The model is able to perform 3D measurement and reconstruction of the environment with improved overall accuracy and precision in spite of the baseline value of the setup, the position to perform the measurement, and the environment. Not only the overall measurement accuracy can be improved, the improvement of accuracy in the worst case is more significant and impressive.

V. Conclusions

This paper proposed a novel three-way joint measurement model with weighted optimization algorithm implemented to perform all-round 3D measurement with superior accuracy and precision using trinocular spherical stereo setup. The weighted reprojection error minimization performs an important role in measurement result optimization. With the disparity maps generated from the input images, which include the random errors, the measurement output is of high accuracy and eliminated error. Since the measurement result by two pairs of binocular spherical stereo might be different especially along with the epipolar directions and it is hard to figure out which result is more close to the true value, the proposed three-way joint measurement model is able to take care of the confidence and perform the weighted optimized fusion for a more accurate final measurement result compare to simply taking the weight average of two results.

Since the optimization is performed pixel by pixel, the next step is to ensure the smoothness of the surfaces while improving the accuracy. The smoothness and continuity of the points will be considered in the optimization cost function. The occlusion problem is not yet considered in this paper, while it affects the accuracy of the measurement. The identification of the occlusion area and processing of the information in the occlusion area will be considered in the future work. The experiment of applying the proposed method on a physical camera setup will also be done in the future.

References

[1] Murray, D., Jennings, C., "Stereo vision based mapping and navigation for mobile robots," in Proceedings of the 1997 IEEE International Conference on Robotics and Automation, pp. 1694–1699. IEEE, 1997.

[2] Kriegman, D.J., Triendl, E., Binford, T.O., "Stereo vision and navigation in buildings for mobile robots," IEEE Transactions on Robotics and Automation **5**(6), 792-803, 1989.

[3] Peng, Q., Tu, L., Zhang, K., Zhong, S., "Automated 3D scenes reconstruction using multiple stereo pairs from portable four-camera photographic measurement system," International Journal of Optics **2015**, 2015.

[4] Nayar, S.K., "Catadioptric omnidirectional camera," in Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp. 482–488. IEEE, 1997.

[5] Li, S., "Binocular spherical stereo," IEEE Transactions on Intelligent Transportation Systems **9**(4), 589–600, 2008.

[6] Li, S., "Trinocular spherical stereo," in Proceedings of the 2006 IEEE/RSJ International Conference on Intelligent Robots and Systems, pp. 4786–4791. IEEE, 2006.

[7] Ma, C., Shi, L., Huang, H., Yan, M., "3D Reconstruction from Full-view Fisheye Camera," arXiv e-prints, arXiv:1506.06273, 2015.

[8] Kim, H., Hilton, A., "3D scene reconstruction from multiple spherical stereo pairs," International journal of computer vision **104**(1), 94–116, 2013.

[9] Yin, W., Pathak, S., Moro, A., Yamashita, A., Asama, H., "Improving 3D Measurement Accuracy in Epipolar Directions via Trinocular Spherical Stereo," in Proceedings of the 2019 IEEE Global Conference on Consumer Electronics, pp. 1005–1006. IEEE, 2019.

[10] Merrell, P., Akbarzadeh, A., Wang, L., Mordohai, P., Frahm, J., Yang, R., Nistér, D., Pollefeys, M., "Real-time visibility-based fusion of depth maps," in Proceedings of the 2007 IEEE International Conference on Computer Vision, pp. 1–8. IEEE, 2010.

[11] Gava, C.C., Stricker, D., Yokota, S., "Dense Scene Reconstruction from Spherical Light Fields," in Proceedings of the 2018 IEEE International Conference on Image Processing, pp. 4178–4182. IEEE, 2010.

[12] Pathak, S., Moro, A., Yamashita, A., Asama, H., "Dense 3D reconstruction from two spherical images via optical flow-based equirectangular epipolar rectification," in Proceedings of the 2016 IEEE International Conference on Imaging Systems and Techniques, pp. 140–145. IEEE, 2016.

[13] Weinzaepfel, P., Revaud, J., Harchaoui, Z., Schmid, C., "DeepFlow: Large displacement optical flow with deep matching," in Proceedings of the 2013 IEEE International Conference on Computer Vision, pp. 1385–1392. IEEE, 2013.