

# 3D Reconstruction from Specialized Wide Field of View Camera System Using Unified Spherical Model

Ahmad Zawawi Jamaluddin<sup>(✉)</sup>, Cansen Jiang, Olivier Morel, Ralph Seulin,  
and David Fofi

Université Bourgogne Franche-Comté, Laboratoire Le2i, UMR6306, FRE CNRS,  
12 rue de la Fonderie, 71200 Le Creusot, France

{ahmad.jamaluddin,cansen.jiang,olivier.morel,ralph.seulin,  
david.fofi}@u-bourgogne.fr  
<http://le2i.cnrs.fr/>

**Abstract.** This paper proposed a method of three dimensions (3D) reconstruction from a wide field of view (FoV) camera system. This camera system consists of two fisheye cameras each with  $180^\circ$  FoV. The fisheye cameras placed back to back to obtain a full  $360^\circ$  FoV. A stereo vision camera is placed to estimate the depth information of anterior view of the camera system. A novel calibration method using unified camera model representation has been proposed to calibrate the multiple camera systems. An effective fusion algorithm has been introduced to fuse multi-camera images by exploiting the overlapping area. Moreover, direct and fast 3D reconstruction of sparse feature matches based on the spherical representation are obtained using the proposed system.

**Keywords:** Fisheye camera calibration · Unified spherical model · 3D reconstruction · Interior Point Optimization algorithm

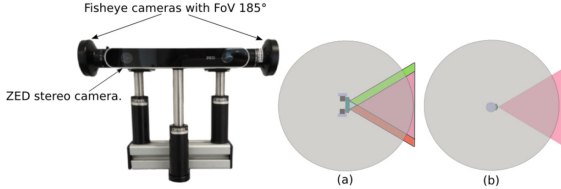
## 1 Introduction

Sensors that provide wider FoV are preferred in the robotic applications, as both navigation and localization can benefit from the wide FoV [1–3]. This paper presents a novel camera system which provides scenes of  $360^\circ$  FoV and the depth information concurrently. The system minimizes the use of equipment, image-data and it has the ability to acquire sufficient information of the scene. The robotic applications of wide FoV camera system are mainly mapping [4, 5], object tracking [6], video surveillance [7–9], virtual reality [10] and structure from motion [11, 12].

### 1.1 Proposed System

The proposed vision system consists of two fisheye cameras (omnidirectional camera) [13, 14] with each has  $185^\circ$  FoV. The cameras are placed back to back

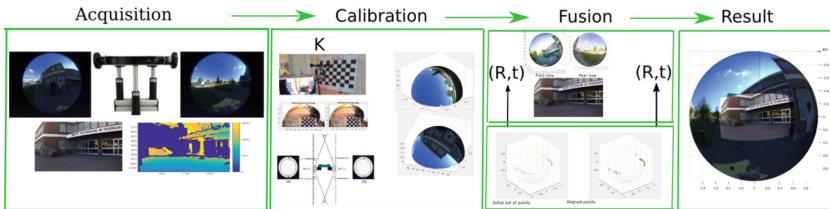
so that they cover the whole 360° of the scene. A high-resolution stereo vision camera, named ZED [16] is placed in front of the rig so that its baseline is in parallel with the baseline of the fisheye cameras. Figure 1 shows the proposed system and the predicted FoV from camera rig.



**Fig. 1.** The front view of the proposed vision system. The illustration (a) and (b) are the predicted FoV viewed from top and right side of the camera rig. The gray color are referred to the FoV of fisheye cameras. The red and green colour are referred to the FoV of stereo vision camera. (Color figure online)

The major contributions of this paper are:

1. We proposed an omni-vision system alongside with a stereo-camera, which offers immense information on 360° FoV of the environment as well as detailed depth information.
2. A new camera calibration method taking the advantages of Unified Camera Model representation has been proposed, which outperforms the state-of-the-art methods.
3. An Interior Point Optimization algorithm (IPO) based on pure rotation matrix estimation approaches has been proposed to fuse the two fisheye and ZED images, which offers seamless images stitching results.
4. A projective distortion has been proposed to be added to the ZED image before projecting onto the unit sphere, which results in enhancing the quality of the overlapping image.



**Fig. 2.** Block acquisition contains the proposed camera system with the images of fisheye and ZED cameras. Block calibration consists of calibration method to estimate intrinsic and extrinsic parameters. Block fusion consists of the method to fuse all images onto unit sphere. Block result consist of final result with image from fisheye and ZED cameras fused together.

## 2 Methodology

### 2.1 Unified Spherical Camera Model

The unified spherical camera model has been proposed by Geyer [14] and Barreto [17]. The image formation of dioptric camera was effected by the radial distortion. Due to that, the point on the scene is nonlinear with the point in the dioptric image. The 3D points  $\chi$  are projected to the image point  $\mathbf{x}$  using the pin-hole camera model representation. It is also considered the representation of a linear and non linear transformation mapping function which is depending on the type of camera. The model was extended by Mei [18] and an omnidirectional camera calibration toolbox [23] has been developed. This model has been used as a reference to map the image on the unified spherical model. All points  $m$  are projected to the image plane using  $K$ , which is a generalized camera projection matrix. The value of  $f$  and  $\eta$  should be also generalized to the whole system (camera and lens).

$$p = Km = \begin{bmatrix} f_1 \eta & f_1 \eta \alpha & u_0 \\ 0 & f_2 \eta & v_0 \\ 0 & 0 & 1 \end{bmatrix} m, \quad (1)$$

where the  $[f_1, f_2]^T$  are the focal length,  $(u_0, v_0)$  are the principal point and  $\alpha$  is the skew factor. By using the projection model, the point on the normalized camera plane can be lifted to the unit sphere by the following equation:

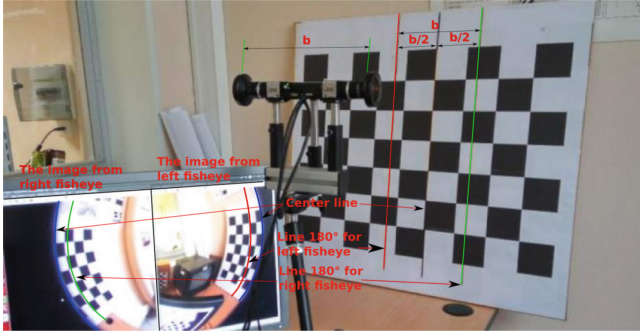
$$\hat{h}^{-1}(m) = \begin{bmatrix} \frac{\xi + \sqrt{1 + (1 - \xi^2)(x^2 + y^2)}}{x^2 + y^2 + 1} x \\ \frac{\xi + \sqrt{1 + (1 - \xi^2)(x^2 + y^2)}}{x^2 + y^2 + 1} y \\ \frac{\xi + \sqrt{1 + (1 - \xi^2)(x^2 + y^2)}}{x^2 + y^2 + 1} - \xi \end{bmatrix}, \quad (2)$$

where the parameter  $\xi$  quantifies the amount of radial distortion of the dioptric camera.

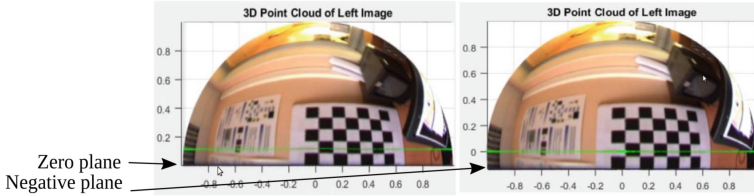
### 2.2 Camera Calibration Using Zero-Degree Overlapping Constraint

A new multi-camera setup is proposed, where two  $185^\circ$  fisheye cameras are rigidly attached in opposite direction to each other. Since the fisheye camera has more than  $180^\circ$  of FoV, the proposed setup contributes an overlapping area along the periphery of the two fisheye cameras. Taking the advantage of overlapping FoV of the two fisheye cameras, we propose a new fisheye camera calibration using the constraint on overlapping zero-degree with Unified Camera Model under the following assumption:

- If  $\xi$  is estimated correctly, the  $180^\circ$  line of the fisheye camera should ideally lay on the zero degree plane of the unit sphere. See Figs. 3 and 4.
- A correct calibration (registration) of multi-fisheye camera setup contributes to a correct overlapping area.



**Fig. 3.** Experimental setup to calibrate the value of  $\xi$ . The baseline of the camera rig, noted as  $b$  (from left fisheye lens to right fisheye lens) is measured and two parallel lines with the same distance to each other as well as a centre line is drawn on a pattern. The rig is faced and aligned in front of the checkerboard pattern such that the centre line touches the edges of both fisheye camera images.



**Fig. 4.** The left image was projected with initial estimate of  $\xi$ . The  $180^\circ$  lines should ideally lay on the zero plane. After the iterative estimation of  $\xi$ , the  $180^\circ$  linear now lay on the zero plane.

**Pure Rotation Registration.** One of the major objective of our setup is to produce a high quality  $360^\circ$  FoV unit sphere which ables to handle the visualization. A common way to do this is to calibrate the camera setup such that the relative poses between the cameras are known. Let the features from the left and right fisheye cameras (projected onto the unit sphere) be denoted as  $\mathbf{x}^{Lf}$  and  $\mathbf{x}^{Rf}$ , respectively. The transformation between the two fisheye cameras is noted as  $T \in R^{4 \times 4}$ , such that:

$$\mathbf{x}^{Lf} = R\mathbf{x}^{Rf}, \tag{3}$$

The estimation of the transformation matrix,  $T$  as discussed in [19,21,22] contains the rotation  $R$  and also the translation  $t$ .

In our method, we used a pure rotation matrix to solve this problem by enforcing the transformation matrix contains zero translation [20], which represented as:

$$\min_R \sum_{i=1}^n \Psi(\|\mathbf{x}^{Lf} - R\mathbf{x}^{Rf}\|), \quad \text{s.t.} \quad RR^T = 1, \det(R) = 1, \tag{4}$$

where  $R$  is the desired pure rotation matrix,  $\Psi(\cdot)$  is the Huber-Loss function for robust estimation. By solving the above equation, a pure rotation matrix that minimizes the registration errors between fusion of the two fisheye cameras. Here, we adopt the Interior Point Optimization algorithm (IPO) to solve the system.

**Fusion of Perspective Camera onto Unit Sphere.** A new method has been proposed to fuse perspective image onto the unit sphere. A major difference between perspective and spherical images is the existence of distortion which deforms the object on the scene. The direct matching point from the perspective is deficient to match the features on the unit sphere. This due to the characteristic of unit spheres which has several levels of distortion. The projective distortion parameters have been proposed to be added to the perspective image plane before projecting onto the unit sphere.

$$P = K^{-1} \cdot H \cdot I, \quad \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = K^{-1} \cdot H \cdot I, \quad (5)$$

where,  $H$  is a projective distortion parameter.  $K$  is a camera matrix.  $I$  is an image frame.

$$\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} f_x & \delta & v_o \\ 0 & f_y & \nu_o \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix} \cdot \begin{bmatrix} v \\ \nu \\ 1 \end{bmatrix}, \quad (6)$$

the value  $x$ ,  $y$  and  $\xi$  ( $\xi = 0$  for the perspective camera) are replaced into mapping function (in Eq. (2)) for projecting onto the unit sphere.

**Fusion of Multi-camera Images.** In our multi-camera setup, the fusion of fisheye cameras alongside with the ZED camera based on a unified model representation can be achieved in a similar manner.

Let  $\mathbf{x}^z$  and  $\mathbf{x}^{Lf}$  be the feature correspondences (mapped from  $\chi^z$  and  $\chi^{Lf}$ ) on a Unified Sphere. The fusion of the ZED camera and the fisheye camera can be framed as a minimization problem of the feature correspondences on a unified sphere, which is defined as:

$$\operatorname{argmin}_R \sum_{i=1}^n \Psi (\|\mathbf{x}^{Lf} - \mathbf{x}^Z(R)\|_2), \quad (7)$$

where  $\Psi(\cdot)$  is the Loss function for the purpose of robust estimation, while

$$\chi(\theta_{x,y,z}) = R(\theta_{x,y,z}) \begin{bmatrix} x_s \dots x_s^n \\ y_s \dots y_s^n \\ z_s \dots z_s^n \end{bmatrix}, \quad (8)$$

stands for the registration of ZED camera sphere points to the left fisheye camera (the reference), where  $R(\theta_{x,y,z})$  is the desired pure rotation matrix with estimated rotation angles  $\theta_{x,y,z}$ . This can be solved on a similar manner solving Eq. (4), by applying an IPO algorithm.

### 2.3 Epipolar Geometry of Omnidirectional Camera

The epipolar geometry for an omnidirectional camera has been studied [17] and it was originally used as a model for catadioptric camera. The study was then extended to the dioptric or fisheye camera system. Figure 5 shows the epipolar geometry of fisheye camera. Lets consider the two positions of a fisheye camera observed from point  $P$  in the space. The points  $P_1$  and  $P_2$  are the projection point  $P$  onto the unit sphere image in two different fisheye's positions. The points  $P, P_1, P_2, O_1$  and  $O_2$  are coplanar, such that:

$$\overline{O_1O_2} \times \overline{O_2P_1} \cdot \overline{O_2P_2} = 0, \quad O_1^2 \times P_1^2 \cdot P_2 = 0, \tag{9}$$

where,  $O_1^2$  and  $P_1^2$  are the coordinates of  $O_1$  and  $P_1$  in coordinate system  $O_2$ . The transformation between system  $X_1, Y_1, Z_1$  and  $X_2, Y_2, Z_2$  can be described by rotation  $R$  and translation  $t$ . The transformation equations are:

$$O_1^2 = R \cdot O_1 + t = t, \quad P_1^2 = R \cdot O_1 + t, \tag{10}$$

$O_1^2$  is the pure translation. By substituting (10) in (9) we get:

$$P_2^T E P_1 = 0, \tag{11}$$

where  $E = [t]_{\times} R$  is the essential matrix which consists of rotation and translation. In order to estimate the essential matrix, the points correspondence pairs on the fisheye images are stacked into the linear system, thus the overall epipolar constraint becomes:

$$Uf = 0, \text{ where } U = [u_1, u_2, \dots, u_n]^T, \tag{12}$$

and  $u_i$  and  $f$  are vectors constructed by stacking column of matrices  $P_i$  and  $E$  respectively.

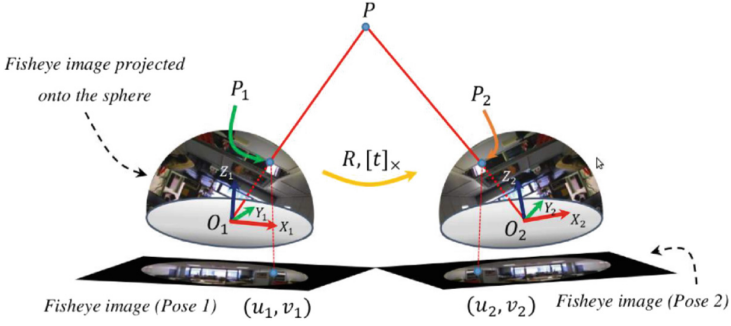
$$P_i = P_i P_i'^T, \tag{13}$$

$$E = \begin{bmatrix} f_1 & f_4 & f_7 \\ f_2 & f_5 & f_8 \\ f_3 & f_6 & f_9 \end{bmatrix}. \tag{14}$$

The essential matrix can be estimated with linear least square by solving Eqs. (12) and (13), where  $P_i^i$  is the projected point which corresponds to  $P_2$  of the Fig. 5,  $U$  is  $n \times 9$  matrix and  $f$  is  $9 \times 1$  vector containing the 9 elements of  $E$ . The initial estimated essential matrix is then utilized for the robust estimation of essential matrix. An iterative reweighted least square method [15] is proposed to re-estimate the essential matrix of omnivision camera. This assigns minimal weight to the outliers and noisy correspondences. The weight assignment is performed by the residual  $r_i$  for each point.

$$r_i = f_1 x'_i x_i + f_4 x'_i y_i + f_7 x'_i z_i + f_2 x_i y'_i + f_5 y_i y'_i + f_8 y'_i z_i + f_3 x_i z'_i + f_6 y_i z'_i + f_9 z_i z'_i, \tag{15}$$

$$err \rightarrow \min_f \sum_{i=1}^n \left( w_{S_i} f^T u_i \right)^2, \tag{16}$$



**Fig. 5.** The diagram of epipolar geometry of fisheye camera for 3D reconstruction.

$$w_{S_i} = \frac{1}{\nabla r_i}, \text{ where } r_i = (r_{x_i}^2 + r_{y_i}^2 + r_{z_i}^2 + r_{x_i'}^2 + r_{y_i'}^2 + r_{z_i'}^2)^{\frac{1}{2}}, \quad (17)$$

where  $w_{S_i}$  is the weight (known as Sampson's weighting) that will be assigned to each set of corresponding point and  $\nabla r_i$  is the gradient;  $r_{x_i}$  and so on are the partial derivatives found from Eq. (15), as  $r_{x_i} = f_1 x_i' + f_2 y_i' + f_3 z_i'$ .

Once all the weights are computed,  $U$  matrix is updated as follow:  $U = WU$ .

where  $W$  is a diagonal matrix of the weights computed using Eq. (16). The essential matrix is estimated at each step and forced to be of rank 2 in each iteration. The procrustean approach is adopted here and singular value decomposition is used for this purpose.

## 3 Experimental Results

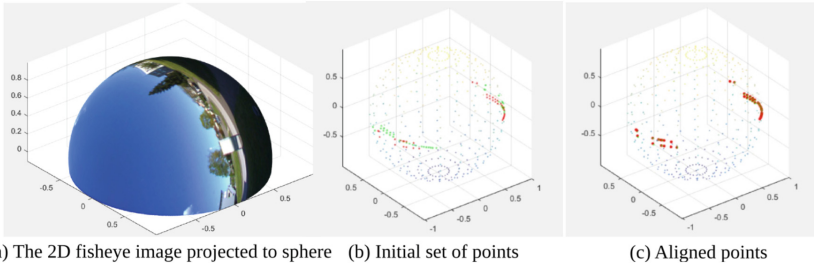
### 3.1 Estimation of Intrinsic Parameters

The unknown parameters  $f_1, f_2, u_0, v_0$  and  $\xi$  of fisheye cameras are estimated using the camera calibration toolbox provided by Mei [23]. The fisheye images are projected onto the unit sphere using the Inverse Mapping Function defined in Mei's projection model. Figure 6(a) shows the two dimension(2D) images from a fisheye camera is projected onto the unit sphere.

### 3.2 Estimation of Extrinsic Parameters

**Rigid Transformation Between Two Fisheyes Image.** The overlapping features are taken along the periphery on the left and right fisheye images. The selected points are projected onto the unit sphere. The rigid 3D transformation matrix are estimated using the selected overlapping features. The IPO algorithm is used to estimate the rotation between the set of projected points. Figure 6(b) shows the set of projected points (green-left fisheye and red-right fisheye) aren't aligned. After using IPO algorithm, the selected points are aligned together as shown in Fig. 6(c).

The rotation matrix is parameterized in terms of Euler angles and cost function is developed that minimize the Euclidean distance between the reference (point projections of left camera image) and the three dimensional points from the right camera



**Fig. 6.** (a) The 2D fisheye image is projected onto unit sphere. (b) The selected points (green and red point) aren't aligned together. (c) After use IPO algorithm, the green and red points are aligned with their respective point. (Color figure online)

image. The transformation results using Singular Value Decomposition (SVD) though are very close to pure rotation. It assumed that translation is also as a parameter to align the set of points. Figure 7 shows the fusion result. The points on the hemispheres that are beyond the zero plane are first eliminated. Then the transformation is applied on the hemisphere of the right fisheye camera and the point matrices are concatenated to get a full unit sphere.



**Fig. 7.** (a) The image from fisheye and ZED cameras are fused together. (b) The fusion result after applying the projective distortion on the ZED image. Focusing to the border between ZED and unit sphere, the ZED image is perfectly over-lapped on the unit sphere.

**Rigid Transformation Between a Fisheyes and ZED Camera.** The same procedures are used to estimate the transformation matrix between the image from ZED camera and two hemispheres.

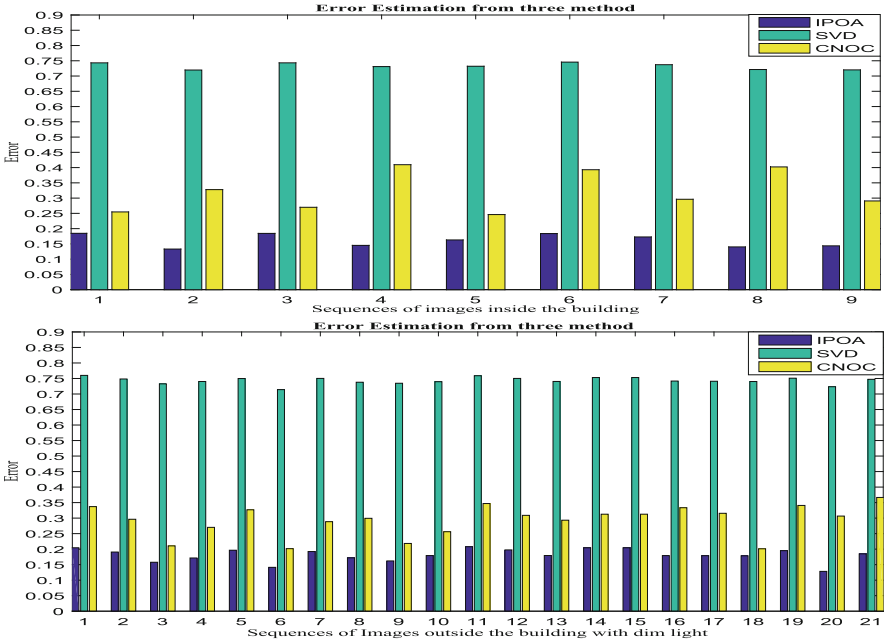
As shown in Fig. 7(a), the RGB images from ZED camera are overlapped onto the unit sphere. It is also recovered the scale between the ZED and fisheye cameras.

The fusion has been enhanced by adding the projective distortion to the ZED image. Figure 7(b) shows that the result is much better after handling the distortion on the fisheye images.

### 3.3 Estimation the Three Dimensional Registration Error

The computation of registration error during mapping on the unit sphere is done to prove the registration method. The Root Means Square Error (RMSE) is used to





**Fig. 8.** The registration error estimation used three different methods. The images sequences have been taken inside and outside of the building. The average registration error using proposed method is 0.1612 (inside building) and 0.1812 (outside building).

calculate the error. The rigid 3D transformation matrix and parameter  $\xi$  which was obtained from calibration are used to determine the residual error of the point pairs registration on the unit sphere. Three methods have been compared:

1. IPO: Our method - The pure rotation estimated using feature matches and IPO algorithm.
2. SVD: The transformation matrix is estimated using features matches with SVD [21].
3. CNOC: Calibration Non Overlapping Cameras, Lébraly [19].

The image sequences were taken in several different environments. The feature points were selected on the overlapping area. The same data set are used in all three methods. Figure 8 shows that the proposed method has the lowest registration errors.

### 3.4 3D Reconstruction Using the Camera Rig

The goal of triangulation is to minimize the distance between the two lines toward point  $P$  in 3D the space. This problem can be expressed as a least square problem.

$$\min_{a,b} \|aP_1 - bRP_2 - t\|, \tag{18}$$

$$\begin{bmatrix} a^* \\ b^* \end{bmatrix} = (A^T A)^{-1} A^T t, A = [P_1 - RP_2], \tag{19}$$

By referring to Fig. 5, looking from the first pose, point  $P$ , the line passing through  $O_1$  and  $P_1$  can be written as  $aP$  and the line passing through  $O_2$  and  $P_2$  can be written as  $bRP_2 + t$ , where  $a, b \in \mathbb{R}$ ,  $P$  is a world coordinate point.  $O_1$  and  $O_2$  are the camera center for pose 1 and 2.  $P_1$  and  $P_2$  are the point  $P$  on the unit sphere at pose 1 and 2.  $R$  and  $t$  are the rotation and translation between the two poses.

The 3D point  $P$  is reconstructed by finding the middle point of the minimal distance between the two lines. It can be computed by;

$$P_k = \frac{a^* P_1 + b^* RP_2 + t}{2}, \text{ where, } k = 1, 2, 3, 4 \tag{20}$$

Figure 9 shows the features matching points. All the points are selected manually. For the future works, an automated features matching points algorithm will be developed using the existing features descriptor. Figure 10 shows the results of feature matching and scene reconstruction algorithm developed following the spherical model of the camera.

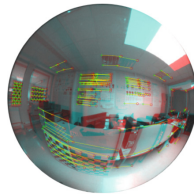


Fig. 9. The features matching points between two different poses of fisheye cameras

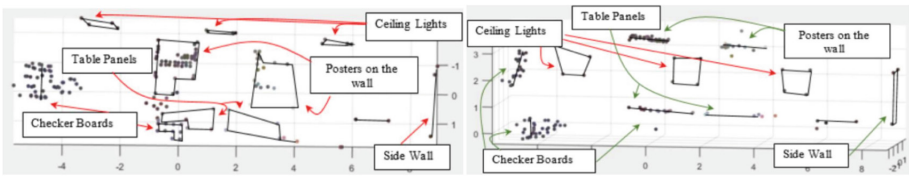


Fig. 10. The front view (left) and top view (right) of the three dimensional reconstruction scenes

## 4 Conclusions

This paper proposed a new camera system which has 360° FoV and detail depth information at anterior. The two fisheye cameras each 180° FoV are placed back to back to obtain 360° FoV. A stereo vision camera is placed perpendicular to obtain depth information at anterior. A novel camera calibration method taking advantages the Unified

Spherical Model has been introduced to calibrate multi camera system. A pure rotation matrix based-on IPO algorithm has been used to fuse images from multi camera setup by exploiting the overlapping area. The result are reduced the registration error and enhance the quality of image fusion. The 3D reconstruction based on the spherical representation has been estimated using the proposed system.

## References

1. Li, S.: Full-view spherical image camera. In: ICPR 2006, pp. 386–390. IEEE (2006)
2. Zhang, C., Xu, J.: Development of an omni-directional 3D camera for robot navigation. In: AIM, pp. 262–267. IEEE (2012)
3. Kim, S.-H., Park, J.-H., Jung, I.-K.: Global localization of mobile robot using an omni-directional camera. In: IPCV, p. 1 (2014)
4. Liu, M., Siegwart, R.: Topological mapping and scene recognition with lightweight color descriptors for an omnidirectional camera. *IEEE Trans. Robot.* **30**(2), 310–324 (2014)
5. Lukierski, R., Leutenegger, S., Davison, A.J.: Rapid free-space mapping from a single omnidirectional camera. In: ECMR, pp. 1–8. IEEE (2015)
6. Marković, I.: Moving object detection, tracking and following using an omnidirectional camera on a mobile robot. In: ICRA, pp. 5630–5635. IEEE (2014)
7. Cogal, O., Akin, A., Seyid, K., et al.: A new omni-directional multi-camera system for high resolution surveillance. In: SPIE, pp. 91200N–91200N-9 (2014)
8. Depraz, F., Popovic, V.: Real-time object detection and tracking in omnidirectional surveillance using GPU. In: SPIE, pp. 96530N–96530N-13 (2015)
9. Sablak, S.: Omni-directional intelligent autotour and situational aware dome surveillance camera system and method. U.S. Patent No 9,215,358, 15 December 2015
10. Li, D., et al.: Motion interactive system with omni-directional display. In: 2013 International Conference on Virtual Reality and Visualization (ICVRV). IEEE (2013)
11. Chang, P., Hebert, M.: Omni-directional structure from motion. In: Workshop on Omnidirectional Vision, pp. 127–133. IEEE (2000)
12. Micusik, B., Pajdla, T.: Structure from motion with wide circular field of view cameras. In: PAMI, vol. 28, no 7, pp. 1135–1149 (2006)
13. Nayar, S.K.: Catadioptric omnidirectional camera. In: CVPR, pp. 482–488. IEEE (1997)
14. Geyer, C., Daniilidis, K.: A unifying theory for central panoramic systems and practical implications. In: ECCV, pp. 445–461 (2000)
15. Hartley, R.I., Sturm, P.: Triangulation. *Comput. Vis. Image Underst.* **68**(2), 146–157 (1997)
16. ZED Documentation. <https://www.stereolabs.com/documentation/overview/getting-started/introduction.html>. Accessed 15 June 2017
17. Barreto, J.P.: A unifying geometric representation for central projection systems. *Comput. Vis. Image Underst.* **103**(3), 208–217 (2006)
18. Mei, C., Rives, P.: Single view point omnidirectional camera calibration from planar grids. In: ICRA, pp. 3945–3950. IEEE (2007)
19. Lébraly, P., Ait-Aider, O., Royer, E., et al.: Calibration of non-overlapping cameras-application to vision-based robotics. In: BMVC, pp. 10.1-10.12 (2010)
20. Othmani, A.A., Jiang, C.: A novel computer-aided tree species identification method based on burst wind segmentation of 3D bark textures. *MVA* **27**(5), 751–766 (2016)

21. Lourakis, M.I.A., Deriche, R.: Camera self-calibration using the singular value decomposition of the fundamental matrix: from point correspondences to 3D measurements. Ph.D. thesis. INRIA (1999)
22. Jamaluddin, A.Z., Mazhar, O., Morel, O., et al.: Design and calibration of an omni-RGB + D camera. In: URAI, pp. 386–387. IEEE (2016)
23. Mei, C.: Active Vision Group. <http://www.robots.ox.ac.uk/cmei/Toolbox.html>. Accessed 15 June 2017