A New Method for Omni-RGB+D Camera Rig Calibration and Fusion using Unified Camera Model

Ahmad Zawawi Jamaluddin, Cansen Jiang, Osama Mazhar, Olivier Morel, Ralph Seulin, David Fofi

***Abstract*— The proposed vision system is compact and rigid with two fisheye cameras that provide a 360-degree field of view. Alongside, a high-resolution stereo vision camera is mounted to monitor anterior field of view for precise depth perception of the scene. To effectively calibrate the proposed camera system, we offer a novel camera calibration approach taking the advantages of Unified Camera Model representation. The proposed calibration method outperforms the state-of-the-art methods. Moreover, we proposed more affective algorithm in fusing the two fisheye images into a single unified sphere, which offers seamless stitching results. This new omni-vision rig system is designed to obtain sufficient information to be used on a robot for object detection and recognition. A large scale SLAM and dense 3D reconstruction can be achieved taking the advantage of the large field of view.**

I. INTRODUCTION

At all time, the living creatures have been observed by the scientist and the researcher. Their special abilities have been transformed into the usable form by the aid of technology and finally to be used by human or robot. The artificial systems have been developed either to obtain a better and precise result or to replace a human for dangerous missions[1]. A vision system is one of the important components in this research. The research and development on computer vision are extremely increasing in parallel with the development on robotics technology.[2]. The fabrication of hybrid camera systems with the wide field of view, combine with the computer vision technique makes this research more interesting. The objectives of this research were to fully utilise vision sensor as the useful kit for robotics navigation. This paper presents the novel camera system which can provide a 360° field of view and the depth information concurrently. The system minimizes the use of equipment and image-data and has the ability to acquire sufficient information on the scene. The applications of omnidirectional cameras are mainly for robotics such as localization and mapping[3], robot navigation[4], object tracking[5], visual servoing[6], structure from motion/motion from structure[7][8] and virtual reality[9].

A. Proposed System.

The proposed vision system consists of two CCD cameras mounted with a fisheye lens with each has more than 180° field of view. The fisheye cameras placed back to back so that it cover the whole 360° vertically and horizontally. A high-resolution stereo vision camera placed in front of the rig, so that its baseline is in parallel with the baseline of the fisheye cameras. The stereo vision camera named ZED camera, provides a high-resolution RGB image with the depth information. Fig.1 shows the illustration and the views from of camera rig.

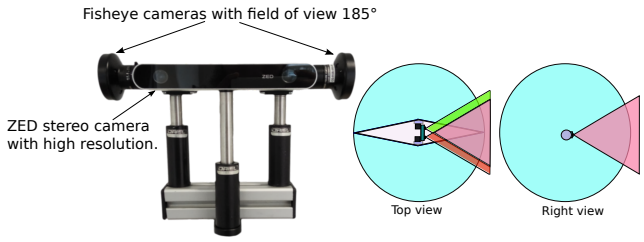

Fig. 1. The front view of the proposed system with the desire field of view.

B. Concept and Motivation

Most of the living creatures have their own capability to look on the object. Some animals have the ability to look from far like an eagle, wide range of view on left and right side like a horse, a motion detector like a fly, thermal view like a snake and stereo and night vision view like a tiger. These abilities of view depend upon the position, form and structure of the eyes. The omni vision cameras rig has been developed by referring to the two categories of animals, prey, and predator. A prey animal has a wide field of view but a small binocular view. They use the eyes for observing the environment from the predator or other threats. While the predator has a large frontal binocular view for targeting and attacking. A creature like jumping spider has the ability to be as prey-like and predator-like. They have four set of eyes for targeting and observing. The same vision system can be realized using multiple camera for the purpose of object observation, recognition and detection.
The major contributions of this paper are three folded:

1) We proposed a very compact Omni-vision system alongside with a Stereo-camera, which offers immense information of 360-degree views of the environment as well as detailed depth information from observation of Stereo-camera.
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) To fuse the two fisheye images, an Interior Point Optimization based pure rotation matrix estimation approach has been proposed, which offers seamless image stitching results.

II. HYBRID CAMERA SYSTEM

A. Omnidirectional Camera

There are three main types of the omnidirectional camera which each has their own advantages.

1) Dioptric: A dioptric system or fisheye[10][11], has a field of view more than 180°. It consists of a single camera mounted with a special lens to increase the field of view.

2) Catadioptric: A catadioptric camera is another type of omnidirectional camera[12]. This camera has more than 180° field of view with a single image. It consists of a camera system with a cone shape reflected mirror. This camera produces also a black spot which is the image of the camera itself.

3) Polydioptric: The Greek words 'poly' means many. This camera consist of several overlapping and non-overlapping identical camera to obtain a panoramic 360° field of view.

Fig.2 illustrates the different omnidirectional vision systems.

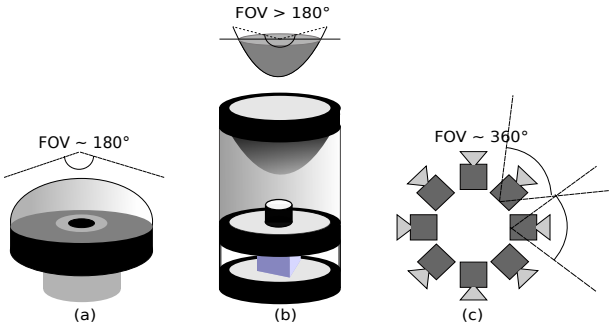

Fig. 2. The illustration shows the omniderectional vision systems (a) Dioptric, (b) Catadioptric and (c) Polydioptric.

The linear perspective geometry has been preserved from any camera system with a single viewpoint. Normally, a single viewpoint exists in all omnidirectional camera. It allows the omnidirectional camera to extract a different view, from perspective to panoramic view[13]. However, for polydioptric which consists of several identical cameras, it considered either they have a unique viewpoint[10][14]. The multi cameras system has a stereovision capability or to enhance the field of view, but it is impossible for the camera rig to have a single effective viewpoint. Fig. 3 shows the illustration of sphere view for the omnidirectional camera.

B. Stereovision Camera

Stereovision is a technique used to build three dimensional description of a scene observed from different viewpoints [15]. If no additional lightening of the scene is required, for example from the laser, this technique is known as passive stereo vision[16]. This technique is used to perceive depth information through generating disparity maps, which then is used to detect obstacles in the environment [17]. This is a classical technique that helps in the field of robotics for

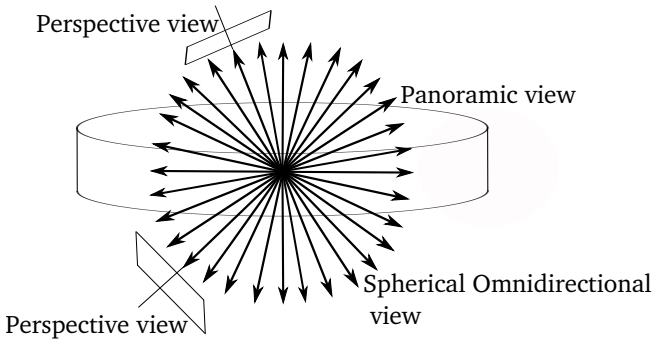

Fig. 3. The illustration shows the sphere view for omnidirectional camera. Image courtesy from Nayar et al: 1997

localization, navigation and obstacle detection since several decades. The development of high-resolution optical sensors and dedicated graphics processing units are helping engineers to design better stereovision cameras for the use in robotics and relevant fields. Lately the release of Bumblebee2 and ZED Stereo Camera[18] enable the researchers to get high-resolution three dimensional depth sensing. Fig. 4 shows the illustration of stereo vison system.

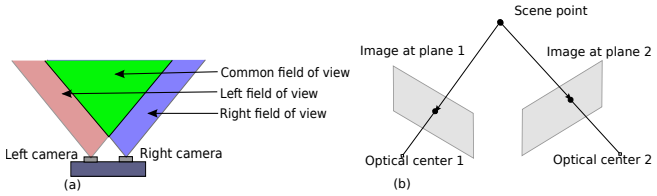

Fig. 4. The Illustration shows the stereo vision system. (a): Left and right cameras have their own view, the green colour is the common view or stereovision. (b): A simple stereovision model.

III. METHODOLOGY

A. Unified Spherical Camera Model

The unified spherical camera model has been proposed by Barreto J.P[14][19]. The image formation on dioptric camera was effected by the radial distortion. Due to that, the point on the scene is not linear with the point on the dioptric image. The image formation in the dioptric camera with radial distortion is in three steps procedure. Following a pinhole camera model, a world point χ originates a projective ray $x = P\chi$, where P is a conventional 3x4 projection matrix. This ray is suppressed into a two dimensional projective point $\mathbf{x} = KP\chi$, where K is the intrinsic parameters matrix. This is the input for the second step, where $\hbar(\cdot)$ (in Equation (10)) is a non-linear transformation in the case of dioptric cameras , where ξ defines the amount of radial distortion.

Then the model was extended by Christopher Mei[20] and he developed the calibration toolbox to calibrate the omni-directional camera. The model was enhanced by introducing another type of distortion called tangential distortion which is incorporated with the radial distortion. Fig. 5 shows the Mei's projection model from camera to unified spherical model. The eccentricity, ξ parameter which defines the amount of

distortion has been explained. Mei's projection model [20] has been used as a reference and it provides procedures to map the image on an unified spherical model.
Let consider a 3D point $\chi = (X,Y,Z)^T$ in the world, and project it to the unit sphere with C_m as a center of the sphere.

$$\chi_s = \frac{\chi}{\|\chi\|} = (X_s, Y_s, Z_s)^T, \tag{1}$$

Then the point χ_s mapped to the new reference frame with the new center C_p

$$(\chi_s)_{F_m} \rightarrow (\chi_s)_{F_p} = (X_s, Y_s, Z_s)^T, \tag{2}$$

Next, the point projected onto the normalised plane

$$m_u = (\frac{X_s}{Z_s+\xi}, \frac{Y_s}{Z_s+\xi}, 1)^T = \hbar X_s, \tag{3}$$

The model of distortion (tangential and radial) are added to the projection model. It consists of three radial and two tangential distortion parameters.

$$x_c = x1 + k_1 r^2 + k_2 r^4 + k_5 r^6 + 2k_3 xy + k_4(r^2+2x^2), \tag{4}$$

$$y_c = y1 + k_1 r^2 + k_2 r^4 + k_5 r^6 + 2k_4 xy + k_3(r^2+2y^2), \tag{5}$$

where:

$$r = \sqrt{x^2+y^2}, \tag{6}$$

and the sum of distortion is

$$m_d = m_u + D(m_u, V), \tag{7}$$

where V contains the coefficients of distortion.

$$V = (k_1, k_2, k_3, k_4, k_5, k_6), \tag{8}$$

and finally, the point m_d is projected to the image plane using K, which is a generalized camera projection matrix. The value f and η should be also generalized to the whole system (camera+lens).

$$p = Km_d = \begin{bmatrix} f_1\eta & f_1\eta\alpha & u_0 \\ 0 & f_2\eta & v_0 \\ 0 & 0 & 1 \end{bmatrix} m_d, \tag{9}$$

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

$$\hbar^{-1}(m_u) = \begin{bmatrix} \frac{\xi+\sqrt{1+(1-\xi^2)(x^2+y^2)}}{x^2+y^2+1} x \\ \frac{\xi+\sqrt{1+(1-\xi^2)(x^2+y^2)}}{x^2+y^2+1} y \\ \frac{\xi+\sqrt{1+(1-\xi^2)(x^2+y^2)}}{x^2+y^2+1} - \xi \end{bmatrix}, \tag{10}$$

B. Camera Calibration using Zero-degree Overlapping Constrain

For the proposed multi-camera setup, there are two 185^o fisheye cameras rigidly attached opposite to each other. Since the field-of-view (FOV) of the fisheye camera has more than 180^o, the proposed setup has an overlapping area between the two fisheye cameras periphery. We are taking the advantages of overlapping FOV of two fisheye cameras, we propose to use a new fisheye Camera calibration using the constrain

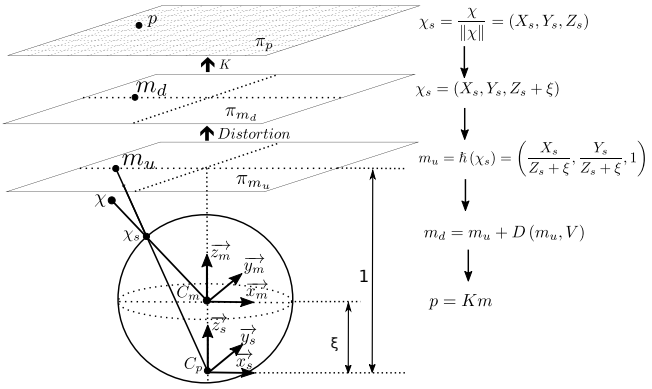

Fig. 5. Mei's projection model from camera to unified spherical model. This figure is the courtesy of Christopher Mei [34].

of overlapping Zero-degree lines of the two fisheye cameras using a Unified Camera Model. Fig. 6 shows experimental setup to re-estimate the value of ξ.

Assumption:

- If ξ is estimated correctly, the 180^o line of the Fisheye camera on the zero degree plane of the Unit Sphere by projecting the fisheye image into a Unified camera model.
- A correct calibration (registration) of multi-fisheye camera setup comes from a correct overlapping area.

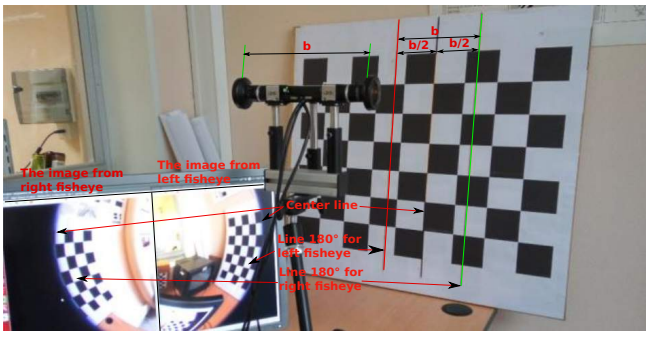

Fig. 6. Experimental setup to calibrate the value of ξ. The baseline of the camera rig (from left fisheye lens to right fisheye lens) is measured and two parallel lines with the same distance to each other as well as a centre line is drawn on a pattern. The rig is faced and aligned in front of the pattern such that the centre line touches the edges of both fisheye camera images

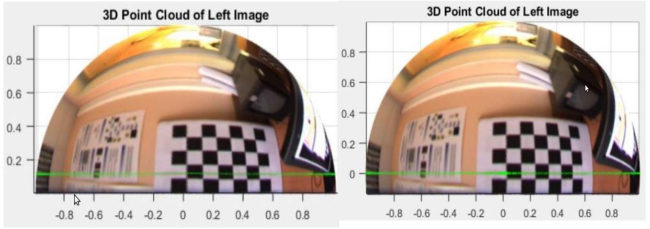

Fig. 7. The left image was projected with initial estimate of ξ; the 180^o lines should ideally lie on the zero plane. After the iterative estimation of ξ, the 180^o line now lie on the zero plane.

1) Zero-crossing Plane Distance Minimization: Let $\{\mathbf{x}_i\}_{i=1}^n$ and $\{\chi_i\}_{i=1}^n$ be a set of points located on the 180^o

line of the fisheye image and their projections onto a unit sphere, receptively. Fig. 7 shows the unit sphere enfolded using initial and estimated ξ and the calibration problem can be simplified as a minimization problem such that the distance between the 180^o-line and the zero-crossing plane of unified sphere is minimized. In other words, the distance between χ and the Zero-crossing plane of the unit sphere is required to be minimized. It denoted as:

$$\min_{\xi} \sum_{i=1}^{n} \psi(\|f(\mathbf{x}_i, \xi)\|_2), \tag{11}$$

where $f(\mathbf{x}, \xi)$ equation (10) is the mapping function from fisheye image ($\mathbf{x}_i$) to the unified camera model (χ_i), operator $\|\cdot\|$ stands for the *l2*-norm, $\psi(\cdot)$ function is the adopted Huber-Loss function for robust estimation purpose. Since the mapping function $f(\mathbf{x}, \xi)$ in equation (11) is not linear, we suggest the Interior Point Optimization Algorithm scheme which is a global non-linear optimization method.

2) Pure Rotation Registration: One of the major objective of our setup is to produce a high quality 360^o FOV unit sphere for handy visualization. To do so, a common way 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 camera (projected onto unit sphere) be denoted as χ^L and χ^R, respectively. The transformation between the two fisheye camera is noted as $T \in R^{4\times4}$, such that:

$$\chi^L = T\chi^R, \tag{12}$$

The transformation matrix, T was estimated, as discussed in [21], the Extrinsic Calibration method using [21] or Singular Value Decomposition (SVD)[24] is not able to correct recover the transformation matrix T due to its pure rotation property. We use a pure rotation matrix to solve this problem by enforcing the transformation matrix contained zero translation[22], which represented as:

$$\min_{R} \sum_{i=1}^{n} \Psi(\|\chi_i^L - R\chi_i^R\|), \quad \text{s.t.} \quad \mathrm{RR}^\mathsf{T} = 1, \det(\mathrm{R}) = 1, \tag{13}$$

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

3) Fusion of Multi-camera Images: In our setup (or other 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 $\mathbf{x}^{Fl}$ be image feature correspondences between ZED camera and fisheye camera, respectively. Let χ^Z and χ^{Fl} be the feature correspondences (mapped from $\mathbf{x}^z$ and $\mathbf{x}^{Fl}$) on a Unified Sphere. The fusion of the ZED camera and the fisheye camera can be framed as a minimization problem from the feature correspondences on a unified sphere, which can be defined as:

$$\underset{\theta_{x,y,z}}{\operatorname{argmin}} \sum_{i=1}^{n} \Psi\left(\left\|\chi_S^{Lf} - \chi_S^Z(\theta_{x,y,z})\right\|_2\right), \tag{14}$$

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

$$\chi(\theta_{x,y,z}) = R(\theta_{x,y,z}) \begin{bmatrix} x_s & . & . & . & . & x_s^n \\ y_s & . & . & . & . & y_s^n \\ z_s & . & . & . & . & z_s^n \end{bmatrix}, \tag{15}$$

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 $\theta_{x,y,z}$. To solve this problem, similar to solving Equation (13), an Interior Point Optimization algorithm is applied.

C. Epipolar Geometry of Omnidirectional Camera.

The epipolar geometry for an omnidirectional camera has been studied and it originally used for a catadioptric camera as a model[17]. The study was extended to the dioptric or fisheye camera system. Fig. 8 shows the epipolar geometry of fisheye camera. Lets consider the two positions of a fisheye camera which observed a point P in the space. Points P_1 and P_2 are the projection of point P onto unit spheres with a coordinate (u_1, v_1) and (u_2, v_2) on the fisheye images[19]. The points P, P_1, P_2, O_1 and O_2 are coplanar,and it can be written as:

$$\begin{aligned} \overline{O_1O_2} \times \overline{O_2P_1} \cdot \overline{O_2P_2} = 0, \\ O_1^2 \times P_1^2 \cdot P_2 = 0, \end{aligned} \tag{16}$$

where, O_1^2 and P_1^2 are the coordinates of O_1 and P_1 in coordinate system X_2, Y_2, Z_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:

$$\begin{aligned} O_1^2 = R \cdot O_1 + t = t, \\ P_1^2 = R \cdot O_1 + t, \end{aligned} \tag{17}$$

By substituting (17) in (16) we get,

$$P_2^T E P_1 = 0, \tag{18}$$

where, $E = [t]R$ the essential matrix which consists of rotation and translation. the computation is always to minimizing the epipolar errors. In order to estimate the essential matrix, the points correspondence pairs on the fisheye imsges are stacked into the linear system, thus the overall epipolar constraint becomes.

$$Uf = 0, \tag{19}$$

where,

$$U = [u_1, u_2, ..., u_n]^T \ ,$$

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

$$P_i = P_i P_i'^T, \tag{20}$$

$$E = \begin{bmatrix} f_1 & f_4 & f_7 \\ f_2 & f_5 & f_8 \\ f_3 & f_6 & f_9 \end{bmatrix}, \tag{21}$$

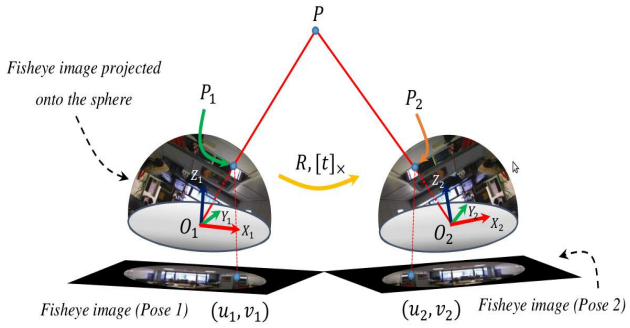

Fig. 8. The epipolar geometry of fisheye camera.

The essential matrix can be estimated with linear least square by solving equation (18) and (19), where P_i^i is the projected point which corresponds to P_2 of the Fig. 6, U is $n \times 9$ matrix and f is 9×1 vector containing the 9 elements of E. The initial estimate of essential matrix is then utilized for the robust estimation of essential matrix and a modified iteratively re-weighted least square method for omnivision cameras is proposed which is originally explained in [24]. This assigns minimal weights to the outliers and noisy correspondences. The weight assignment is performed by finding the residual r_i for each point.

$$r_i = f_1 x_i' x_i + f_4 x_i' y_i + f_7 x_i' z_i + f_2 x_i y_i' \tag{22}$$

$$+ f_5 y_i y' i + f_8 y_i' z_i + f_3 x_i z_i' + f_6 y_i z_i' + f_9 z_i z_i' \ ,$$

$$err \rightarrow \min_f \sum_{i=1}^{n} \left(w_{Si} f^T u_i \right)^2 , \tag{23}$$

$$w_{Si} = \frac{1}{\nabla r_i}, \tag{24}$$

$$\nabla r_i = (r_{xi}^2 + (r_{yi}^2 + (r_{zi}^2 + (r_{xi'}^2 + (r_{yi'}^2 + (r_{zi'}^2)^{\frac{1}{2}}, \tag{25}$$

where, w_{Si} s the weight (known as Sampson's weighting) that will be assigned to each set of corresponding point and ∇r_i is the gradient; r_{xi} and so on are the partial derivatives found from equation (17), as $r_{xi} = f_1 x_i' + f_2 y_i' + f_3 y_i'$.
Once all the weights are computed, U matrix is updated as follow,

$$U = W \times U, \tag{26}$$

where, W is a diagonal matrix of the weights computed using equation (18). 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.

IV. EXPERIMENTAL RESULTS

A. The estimation of intrinsic parameters

- The unknown parameters f_1, f_2, u_0, v_0 and ξ of fisheye cameras are estimated using the omnidirectional camera toolbox provided by Christopher Mei's.
- The fisheye images are projected onto the unit sphere using the Inverse Mapping Function defined in Christopher Mei's camera projection model.
- Fig. 9 shows the images from left and right fisheye cameras projected onto the unit sphere.

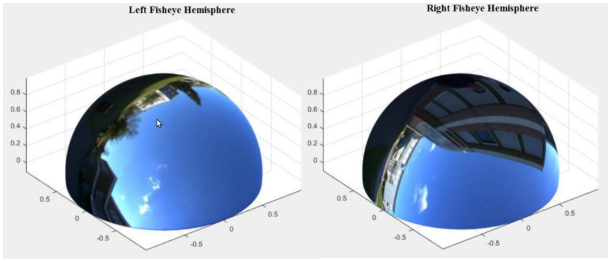

Fig. 9. The images from left and right fisheye cameras projected onto the unit sphere.

1) Re-estimating parameter ξ: The baseline of the camera rig (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 center line is drawn on a pattern. The rig is faced and aligned in front of the pattern such that the center line touches the edges of both circular fisheye camera images (see Fig. 6 and 7). The parallel line corresponding to the edge of fisheye lens of each camera is then forced to the zero plane when the fisheye image is projected onto the unit sphere. This is done by developing a cost function to estimate ξ that minimizes the z-component of pixels on the selected line using interior point optimization algorithm. This has been explained in the Section III.

B. The estimation of extrinsic parameters

1) Rigid transformation between two fisheyes image: Our system has an overlapping features about 5^o between the left and right hemispheres. The selected points are projected onto the unit sphere. The rigid 3D transformation matrix are estimated using the overlapping features. The Interior Point Optimization Algorithm is used to estimate the rotation between a set of projected points. Fig. 10 shows that the set of projected points are aligned together. The rotation matrix

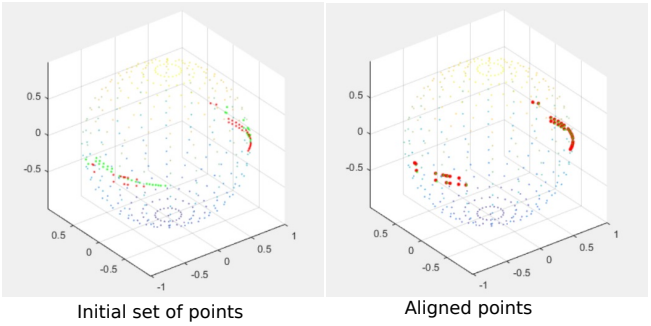

Fig. 10. The selected points are aligned together.

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

The transformation matrix of two fisheye cameras using our method.,

$$T_{Our} = \begin{bmatrix} -1.0000 & -0.0048 & 0.0085 & 0 \\ -0.0045 & 0.9994 & -0.0335 & 0 \\ -0.0087 & -0.0334 & -0.9994 & 0 \\ 0 & 0 & 0 & 1.0000 \end{bmatrix}, \quad (27)$$

The transformation matrix of two fisheye cameras using Singular Value Decomposition(SVD).

$$T_{SVD} = \begin{bmatrix} -1.0000 & -0.0017 & 0.0073 & -0.0039 \\ -0.0019 & 0.9997 & -0.0250 & 0.0085 \\ -0.0073 & -0.0250 & -0.9997 & 0.1143 \\ 0 & 0 & 0 & 1.0000 \end{bmatrix}, \quad (28)$$

The transformation results using SVD though are very close to pure rotation. It assumed that translation is also as a parameter to align the set of points.

Fig. 11 shows that the two hemispheres are fused together using the estimated transformation matrix. Once the rotation

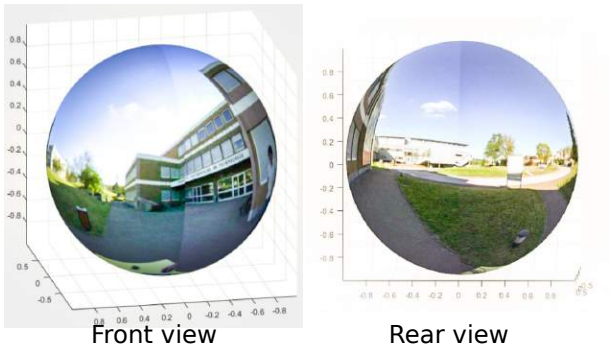

Fig. 11. The unit sphere; images front and rear view from the camera rig.

matrix is estimated, fusion of the two hemispheres is only two step procedure. The points on the hemispheres that are beyond the zero plane are first eliminated. Then the transformation (rotation only) is applied on the hemisphere of the right fisheye camera and the point matrices are concatenated to get a full unit sphere.

2) 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. The transformation matrix image ZED camera refers to left fisheye.

$$T_{ztolf} = \begin{bmatrix} -0.0143 & -0.0290 & -0.9995 & 0 \\ -0.0062 & 0.9996 & -0.0289 & 0 \\ -0.9999 & 0.0058 & -0.0145 & 0 \\ 0 & 0 & 0 & 1.0000 \end{bmatrix}, \quad (29)$$

As shown in fig. 12, the RGB and depth images from ZED camera are overlapped onto the unit sphere. It is also recovered the scale between the ZED and fisheye cameras. The high-resolution RGB image and the depth image are fused onto the unit sphere by the help of transformation matrix estimated. It should be noted that, the fusion of zed image onto the unit sphere is an approximation. The result obtained is reasonable after handling the strong distortion on the fisheye images.

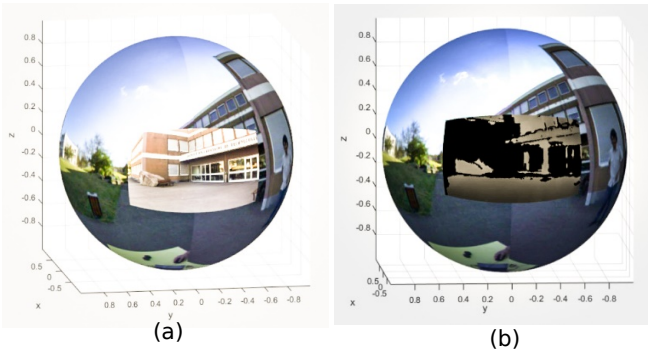

Fig. 12. The unit sphere overlaps with image from ZED camera (a: RGB image, b: Depth image).

C. Estimation the three dimensional registration error.

The computation of registration error during mapping on the unit sphere is done to proof the registration method. The Root Means Square Error (RMSE) is frequently used to calculate the error. The rigid 3D transformation matrix and parameter ξ which was obtained from calibration are used to determine the residuals error which is the difference between the actual values and the predicted values. Three methods have been compared:

1) IPOA : The pure rotation estimated using feature matches and Interior Point Optimization Algorithm.
2) SVD : The transformation (rotation and translation) estimated using features matches with Singular Vector Decomposition[24].
3) CNOC : The method used in Calibration Non Overlapping Cameras[21].

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. Fig. 13, Fig. 14 and Fig. 15 show that the proposed method has the lowest registration errors.

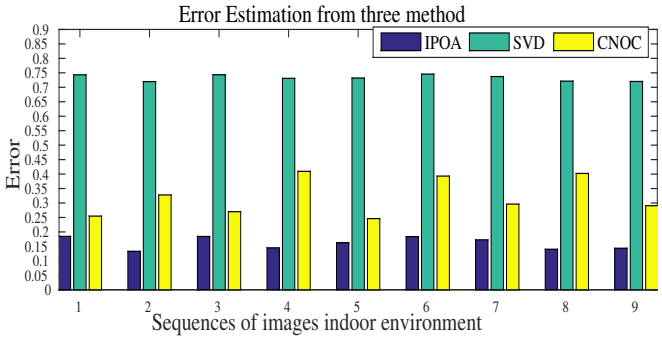

Fig. 13. The registration error estimation using three different methods. For the first experiment, the images sequences have been taken inside the building. The average registration error using proposed method is 0.1612.

V. CONCLUSIONS

The intrinsic parameter of fisheye and ZED cameras are estimated from the camera calibration toolbox. The distortion of fisheye camera is determined by parameter ξ. We were

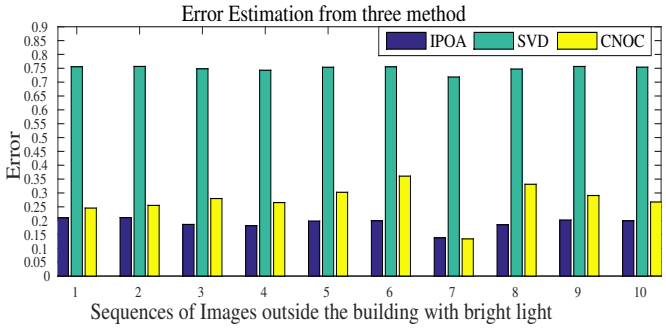

Fig. 14. For the second experiment, the images sequences have been taken outside the building with bright light. The average registration error for the proposed method is 0.1912.

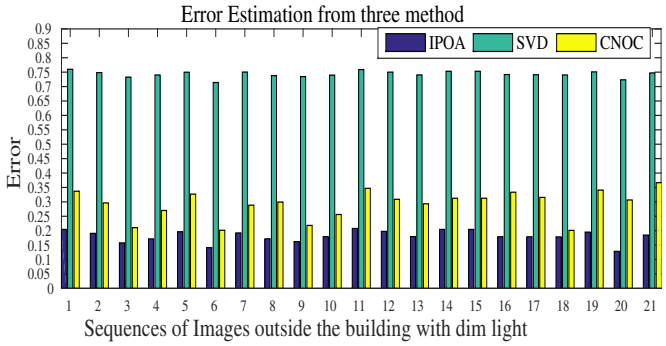

Fig. 15. For the third experiment, the images sequences have been taken outside the building with dim environment (cloudy). The average registration error using proposed method is 0.1812

re-estimating the value of parameter ξ by using a cost function to minimize the z-component of pixels on the selected line using interior point algorithm. The rigid 3D transformation matrix are estimated using the overlapping features assuming that it is a pure rotation. The Interior Point Optimization Algorithm is used to estimate the rotation between a set of projected points. The same procedure is used to fuse ZED camera onto the unit sphere. The overlap area between fisheye and ZED cameras is estimated for the purpose of object tracking and detection. The camera rig is applied to reconstruct 3D scene using features matching and an algorithm based on a spherical model of camera. The registration error is calculated to show the performance of our proposed method.

REFERENCES

[1] Marefat, F., A. Partovi, and A. Mousavinia. "A hemispherical omni-directional bio inspired optical sensor." 20th Iranian Conference on Electrical Engineering (ICEE2012). IEEE, 2012.

[2] Li, Shigang. "Full-view spherical image camera." 18th International Conference on Pattern Recognition (ICPR'06). Vol. 4. IEEE, 2006.

[3] Liu, Ming, and Roland Siegwart. "Topological mapping and scene recognition with lightweight color descriptors for an omnidirectional camera." IEEE Transactions on Robotics 30.2 (2014): 310-324.

[4] Zhang, Chi, et al. "Development of an omni-directional 3D camera for robot navigation." 2012 IEEE/ASME International Conference on Advanced Intelligent Mechatronics (AIM). IEEE, 2012.

[5] Marković, Ivan, Franois Chaumette, and Ivan Petrovi. "Moving object detection, tracking and following using an omnidirectional camera on a mobile robot." 2014 IEEE International Conference on Robotics and Automation (ICRA). IEEE, 2014.

[6] Depraz, Florian, et al. "Real-time object detection and tracking in omni-directional surveillance using GPU." SPIE Security+ Defence. International Society for Optics and Photonics, 2015.

[7] Chang, Peng, and Martial Hebert. "Omni-directional structure from motion." Omnidirectional Vision, 2000. Proceedings. IEEE Workshop on. IEEE, 2000.

[8] Micusik, Branislav, and Tomas Pajdla. "Structure from motion with wide circular field of view cameras." IEEE Transactions on Pattern Analysis and Machine Intelligence 28.7 (2006): 1135-1149.

[9] Li, Dong, et al. "Motion Interactive System with Omni-Directional Display." Virtual Reality and Visualization (ICVRV), 2013 International Conference on. IEEE, 2013.

[10] Nayar, Shree K. "Catadioptric omnidirectional camera." Computer Vision and Pattern Recognition, 1997. Proceedings., 1997 IEEE Computer Society Conference on. IEEE, 1997.

[11] Neumann, Jan, Cornelia Fermuller, and Yiannis Aloimonos. "Polydioptric camera design and 3d motion estimation." Computer Vision and Pattern Recognition, 2003. Proceedings. 2003 IEEE Computer Society Conference on. Vol. 2. IEEE, 2003.

[12] Gluckman, Joshua, and Shree K. Nayar. "Ego-motion and omnidirectional cameras." Computer Vision, 1998. Sixth International Conference on. IEEE, 1998.

[13] Knill, Oliver, and Jose Ramirez-Herran. "Space and camera path reconstruction for omni-directional vision." arXiv preprint arXiv:0708.2442 (2007).

[14] Geyer, Christopher, and Kostas Daniilidis. "A unifying theory for central panoramic systems and practical implications." European conference on computer vision. Springer Berlin Heidelberg, 2000.

[15] Hartley, Richard I., and Peter Sturm. "Triangulation." Computer vision and image understanding 68.2 (1997): 146-157.

[16] Ayache, Nicholas, and Francis Lustman. "Trinocular stereovision for robotics." IEEE Transactions on Pattern Analysis and Machine Intelligence 13.1 (1991).

[17] ZED Stereo Camera from https://www.stereolabs.com/zed/specs/.

[18] Kumar, Saurav, Daya Gupta, and Sakshi Yadav. "Sensor fusion of laser and stereo vision camera for depth estimation and obstacle avoidance." International Journal of Computer Applications 1.25 (2010): 20-25

[19] Barreto, João P. "A unifying geometric representation for central projection systems." Computer Vision and Image Understanding 103.3 (2006): 208-217.

[20] Mei, Christopher, and Patrick Rives. "Single view point omnidirectional camera calibration from planar grids." Proceedings 2007 IEEE International Conference on Robotics and Automation. IEEE, 2007.

[21] Lébraly, Pierre, et al. "Calibration of non-overlapping cameras-application to vision-based robotics." (2010).

[22] Othmani, Alice Ahlem, et al. "A novel Computer-Aided Tree Species Identification method based on Burst Wind Segmentation of 3D bark textures." Machine Vision and Applications (2015): 1-16.

[23] Harris, Chris, and Mike Stephens. "A combined corner and edge detector." Alvey vision conference. Vol. 15. 1988.

[24] ourakis, Manolis IA, and Rachid Deriche. Camera self-calibration using the singular value decomposition of the fundamental matrix: From point correspondences to 3D measurements. Diss. INRIA, 1999.