

Circular Calibration of Depth Extraction in Stereo Configuration

Zulfiqar Ibrahim, Zulfiqar Ali Bangash
Department of Computing and Technology
Iqra University, Islamabad Campus
Islamabad, Pakistan

Muhammad Zeeshan
School of Electrical Engineering and Computer Science
National Institute of Science and Technology (NUST)
Islamabad, Pakistan

Abstract—Lens distortion is defined as departure from rectilinear projection of an imaging system which affects the accuracy of almost all vision applications. This work addresses the problem of distortion with investigating the effects of camera's view angle and spherical nature of lens on image, and then derives a closed-form solution for the correction of distorted pixel's angle in image according to geometric shape of lens. We first propose technique that explores the linear relation between lens and charge-coupled device in intrinsic environment of camera, through analysis of pixel's angle in field of view. Second technique for depth extraction through linear transformation in rectangular configuration is achieved by considering the camera's background in field of view which provides optimal results in closed environment. As the object moves away from the center of image in image plane, depth accuracy starts to deteriorate due to radial distortion. To rectify this problem, we finally propose circular calibration methodology which addresses this inaccuracy and accommodate radial distortion to achieve optimal results up to 98%, in great depth with very large baseline. Results show the improvement over established stereo imaging techniques in depth extraction where the presented considerations are not observed. This methodology ensures high accuracy of triangulated depth with very large base line.

Keywords—Stereo imaging; depth extraction; triangulation; radial distortion; lenses; rectilinear projection component

I. INTRODUCTION

Camera calibration is the process of finding the internal and external parameters of camera so that correspondence between object point in object space and its respective image point in image plane can be computed accurately. Due to distortion factor, the relationship between image point and its object point can be difficult to establish. These distortions result in inaccurate measurements of depth estimation in images. Generally, a camera comprised of lens and image plane, light ray from world enters the image plane through lens which describes the transformation exist between world space and image space. This transformation cannot be described perfectly by using perspective transformation due to the distortion. Distortion develops between the points of the world space and location of the images of those points. The camera calibration techniques allow us to model those points on image plane. These techniques will always be approximation to the real relationship of image points to world point. There is always a room for improvement in these approximations. Many techniques have been presented to eradicate this distortion in images with different levels of accuracy and robustness.

Earlier some techniques [1] require proper lab equipment to find the parameters of distortion but most recent techniques [2]-[5] utilized nonlinear iterative mathematical model to estimate the coefficient of distortion model. These techniques suffer from high computation, complexity and inaccuracy which lead to suboptimal solutions. In [2], [3] managed external and internal parameters of camera in his calibration process by using pin hole camera model and special complex techniques for one-pixel width calibration such as 1-D FFT [3]. When this model [2] is introduce to high distortion lenses, significant error occurs in the calibration process which results in inaccuracy of data and loss of robustness [6].

Zhang et al. [4] proposes famous versatile calibration algorithm in which planner surface with known measurements is utilized to find the correspondence between object space and image space. Homography is also being applied between the target plan and its image plane for the calibration process. Lens distortion coefficients are filtered by minimizing algebraic distance defined by associated camera model. To minimize the algebraic distance, nonlinear refinement by Levenberg-Maruardt algorithm [7] is used. Heikkila et al. [5] introduced a very distinctive calibration process in which approximation of forward lens distortion model is being applied that associate with distorted image coordinates with equivalent undistorted image points. The distortion is minimized through nonlinear iterative refinement in distorted image space. For lens Distortion model, authors incorporated two tangential and two radial distortion terms.

F. Devernay et al. [8] presented field of view (FOV) distortion correction model for fish eye lens. The pixel of image from wide fish eye lens, has measured distance from principle point. This distance is approx. proportional to angle between 3D point, optical center and optical axis. The field of view is proportional to image resolution along image radius. Field of view (FOV) [9] is the ability of the lens to absorb the angular extent of the world scene and it is both dependent on lens and image sensor. This angular extent of scene generates the radial distortion due to low quality circular nature of lenses. A.W. Fitzgibbon [10] presented a division model for the distortion correction. The tangential distortion is not consideration in this model. The nature of this model is polynomial. The advantage of division model over polynomial model is its ability to address higher distortion in lower order. Polynomial model is linear model which address the radial distortion due to the hemispherical shape of lens. The paper by Hartley and Zisserrman [11] holds very important status in the

subject of epipolar geometry. Our research is based on multi view geometry which is heavily influenced by this research work from R. Hartly and A. Zisserman [11].

In this work, a homography relates to two views of the world scene when transformations between two cameras are strictly rotational about an axis through its principle point. Epipolar geometry is intrinsic projective geometry between two views which is independent of structure of the scene. It solely depends on internal parameters of camera and relative pose. Disparity is used to extract the depth of point in image plane which is comparable to our research. This paper extensively borrows work from D.C. Brown et al. [1] and F. Devernay et al. [8] for radial distortion correction by using straight line method which is also known as “Plumb Line” method. The reason for the selection of this paper to be compared with our method is the feasibility of widespread deployment of the cameras in local market. The depth extraction algorithm presented by R. Hartley and Zisserman [11] doesn't do not require high end hardware and can easily be developed using OpenCv tool. The availability of the low cost cameras in local market allows this algorithm to be deployed extensively where depth is required for the certain operations and also keep the cost of development low.

Distortion correction is the most important procedure in a camera calibration process where true correspondence between pixel and its respective world point is approximately established. Radial distortion is the deviation of lens from rectilinear projection where straight lines in world appear as curved lines in image. This form of distortion occurs in wide angle and fish eye lenses due to substandard use of materials. Radial distortion can be removed in post processing software by using different methodologies. First we purpose linear transformation technique in stereo configuration where the division model is used to find the angle of object's pixel. This angle is then considered in triangulation with the help of second camera with known baseline, for the depth extraction of the observed pixel. Linear transformation technique produces very optimal results in object's depth for close-range and when object lies in the center of image. In second technique, Linear transformation in rectangular configuration is applied where two camera are placed perpendicularly towards the object for depth extraction. This technique is suitable for indoor environment where background of object in world plane is known. In third technique, circular calibration in stereo configuration, we derive closed form mathematical treatment to eradicate radial distortion by considering non-linear nature of relationship between lens and CCD where the pixel's depth on both sides of image starts distorting in first and second technique. This later technique helps to maintain 98% accuracy when great depth is under consideration.

Followings are the contributions of this work:

- Linear transformation in stereo configuration is the first technique that we purpose to extract the depth information of object in close range without considering radial distortion correction.
- In second technique, we present linear transformation in rectangular configuration where two cameras are

perpendicularly placed around object for depth extraction in indoor environment.

- We also propose Circular Calibration in Stereo Configuration technique where radial distortion correction method is applied for the great depth extraction with very large baseline.

Rest of the paper is organized as: Section II discusses the prominent works in the field of epipolar geometry and camera calibration. In Section III, we provide brief discussion about the mathematical modeling of three proposed techniques. In Section IV, we analyze the data acquired through proposed techniques in previous section and compare the results with one of the established techniques and in Section V, we did model validation and Section VI concludes the overall discussion and also discusses future work.

II. RELATED WORK

In recent years, many distortion models have been proposed to address the needs various applications of image analysis. Daniel Herrera et al. [12] propose a unique technique to calibrate the Kinect Sensor with the help of in build depth camera. In this paper, authors presented a new technique to calibrate both color and depth camera simultaneously and estimate the pose between them. An image of planer surface is required from different poses of color camera as depth camera is not used during the calibration process which provides robustness and flexible to noise. A new depth distortion model is also proposed when this technique is applied to Kinect sensor. Radial and tangential distortion is usually used to correct the position of pixel in color cameras; depth cameras require more complex techniques to address the distortion model due to the 3D nature of pixel.

Significant distortion is developed in rolling shutter cameras when they capture images during the course of movement. Intrinsic parameters such as focal length and distortion coefficient, modeling of shutter timing must be accurately calibrated. In this paper [13], authors purpose a software solution which required video of known calibration pattern. A continuous time trajectory model is proposed combined with rolling shutter model which provides solution of batch estimate problem to estimate the camera pose and time delay of shutter. In [14], researchers use pinhole camera model with lens distortion model in order to remove inaccuracy in distortion correction model by computing the Jacobean matrix analytically for the representation of 3D world points. Another self-calibration technique [15] is proposed by authors, using single image combined with vanishing points and ellipse/circle. The orthogonal vanishing points are then utilized to compute principal distance and principal points. Then this principle points are considered as distortion center which then are corrected to acquire the distortion free image.

Antonin Mikis et al. [16] described theoretical analysis of radial distortion in wide angle camera lenses and its effects on change of object location on image plane with respect of optical system. Authors present approximation formulas to model such change from object space to image space. The derivation of approximation formulas to model radial distortion are within the validity of third order aberration theory. The

quality of image is being measured by taking center of image as object point and its measure of radial distortion in image plane around center.

In [17], researchers purpose a combined radial and tangential distortion correction algorithm for lens distortion. They propose straight lines quick remedy algorithm to correct lens bending to an expected normal mistake under 1 pixel. Trial comes about demonstrate that the algorithm can advance the accuracy and reinforce the heartiness, in the meantime, significantly lessen the computation time. In this work [18], a new radial distortion correction method is introduced which require single image with radial distortion and noise. The classical pinhole camera model is applied for the distortion correction model. The first step in this algorithm is to find the position of corners in distorted image by using procedure defined by Geiger et al. [19]. A division model to reduce the radial distortion lower order is used. When there is slight change in position of center of distortion, a triangle is formed between undistorted and distorted point against feature corner point. Different values of radial distortion parameters and different set of noises being used in these experiments to check the validity of work.

Hamza Alzaroket et al. [20] propose the idea of utilizing extrinsic parameters of camera using the calibration process. The 3D extrinsic information is obtained through set of 2D frames then this information is used to find the proper location of the camera mounted on robot using new mathematic formula. Zhang et al. [4] calibration method is used to acquire the external and internal parameters due to its flexibility in distortion correction model. Extrinsic parameters obtained through camera calibration are then utilized to compute the location of camera where the best detection can be achieved. In another work [21], authors discuss an idea to use the camera calibration process to estimate the speed of incoming vehicle in camera frame. An equilateral triangle is drawn on image to extract camera parameters. Optical flow vector is used with camera parameters to determine the speed of vehicle in image frames.

In another technique [22], authors purpose a unique method to address problems related to object-triangulation and 3d object tracking within the limits of unified Bayesian framework for joint multi tracking and camera calibration. When an image is observed from cameras located at known distance from each other, binocular disparity starts to affect the data. Due to this effect, depth perception is achieved. In another application of camera calibration [23], a new approach is proposed by the authors to classify the road lanes using on board mounted camera. First they detect the road lanes using linear parabolic lane model then automatic on board camera calibration is deployed. First step is to detect the near field road lines by linear function and parabolic function to detect far field lanes so the curved lines can easily be tracked. Another application of camera calibration is presented by researchers in [24] to extract the human height by acquiring information of line from background, pedestrians and estimation of vanishing points of extracted lines. Vanishing points are extracted from background and pedestrians are then used to calibrate the cameras to estimate the human height.

3D reconstruction of the scene from multiple images of two or more cameras is being under consideration for very long time. R. Hartley et al. [11] presented techniques to resolve this issue which are derived from Projective geometry and photogrammetry. The authors described extensively the geometric principles with respect to their algebraic representation in terms of camera projection matrices, the fundamental matrix and trifocals tensor. Authors presented the epi-polar geometry of two cameras and projective reconstruction from resulting image map correspondence which helps to identify the depth of image point. This work is extensively used in speed estimation of incoming vehicles on the road due to its ease in deployment, development and maintainability. Our work is derived from work presented by the R. Hartley et al. [11] because of its close bonded relationship of projective geometry.

III. LINEAR TRANSFORMATION MODEL

View angle of camera describes the angular area of the scene imaged by a given camera which can be measured vertically, horizontally and diagonally. It is dependent on focal length of lens and size of the image sensor. Lenses produce rectilinear images of distant object, from which the view angle (VA) can be completely determined by the effective focal length and image sensor dimensions. But due to distortion factor presents in lenses, the calculations required for VA become very complex and produce very inconvenient results for applications. In the first technique of linear transformation in stereo configuration, we use division model to calculate the angle of observed pixel which helps in triangulation for depth extraction.

This methodology provides optimal results for close range environment. In second technique, we present the linear calibration in rectangular configuration where cameras are placed perpendicularly from object. This technique only suitable for closed environment where the object's background in world plane is known. When the distance between object and two cameras increases and object shifts from the center of image, the depth calculated through above techniques starts deteriorating due the radially distorted image. To resolve this problem, we provide another methodology where close form equation is derived to sustain circular nature of lens on CCD in pixel's angle. The angle of pixel on the sides of image will be according to the rounded nature of lens which will produce accurate depth for the observed pixel, in triangulation.

A. Linear Transformation in Stereo Configuration

Horizontal view angle of given camera is computed by using the topology presented by E. McCollough [25]. A division model is proposed that will calculate the angle of each pixel distributed horizontally in image:

$$A.P.P = \frac{\text{View angle}}{\text{Resolution on x - axis}}$$

A.P.P is defined as angle per pixel where view angle of a camera is divided with the image's resolution on x-axis where each pixel has equal portion of the total view angle in image. Fig. 1 is basic triangulation, Object lies in the view angles of Right and Left Cameras where A'B' are internal angles of

object's pixel and A, B are external angles with baseline. a, b are total angles from object's line to the baseline.

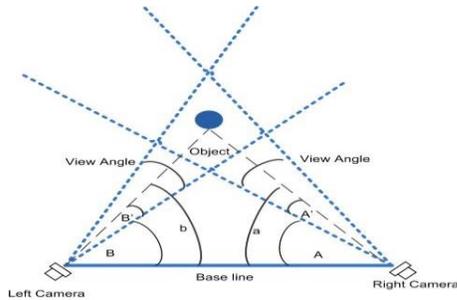


Fig. 1. Linear calibration in stereo configuration, two cameras are looking at object with known distance between them.

Internal angles A'B' of object's pixels within the view angles of left and right cameras are computed as

$$A' = x_axis's \text{ pixel of object} \times A. P. P$$

$$B' = x_axis's \text{ pixel of object} \times A. P. P$$

The total angle (a, b) between object line and baseline is the addition of internal angles A'B' and external angles with baseline A, B which is $b = B + B'$ and $a = A + A'$. The external angles (A, B) are measured manually with the help of protractor. The total angle of triangle is 180° so the angle of the object in world space with respect to two cameras is $c = 180^\circ - a - b$. In Fig. 2, we have the triangle with all the sides $\bar{A}, \bar{B}, \bar{C}$ and angle (a, b, c), the distance between two cameras is known which helps to calculate other parameters. By using the Law of sines, the sides of above triangle are determined as follows:

$$\frac{\bar{A}}{\sin b} = \frac{\bar{B}}{\sin a} = \frac{\bar{C}}{\sin c} \quad (1)$$

By using 1, the side B can be computed as

$$\bar{B} = \sin a \times \frac{\bar{C}}{\sin c}$$

Side B and angle "a" is calculated from above equations. Now to calculate the perpendicular distance, trigonometric functions of right angle triangle is utilized which will extract the depth of object from two cameras. We have

$$\sin \theta = \frac{\text{Perpendicular}}{\text{hypotenuse}} \quad (2)$$

After applying (2) in our scenario, we have

$$\sin b = \frac{\text{Perpendicular}}{\bar{B}}$$

We know the side B and angle "b" so the perpendicular distance of object from two cameras will be

$$\sin b \times \bar{B} = \text{perpendicular}$$

This is called the linear transformation in stereo configuration where view angle is utilized to find the angle of object's pixel which is further used to extract the depth of object. This technique produces optimal results of depth of

object in stereo images where the distance is not very large and object's pixel remains in central area of image where distortion can be neglected.

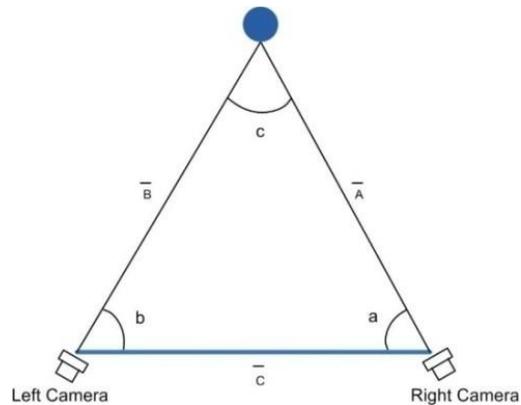


Fig. 2. Stereo configuration of two cameras. Triangle of stereo configuration with all the sides A, B, C and angles (a, b, c).

B. Linear Transformation in Rectangular Configuration

Now we introduce another technique to extract the depth of object in indoor environment where the measurement of background of each camera is known. These cameras are placed perpendicular to the object which makes Rectangle in this basic configuration. This is the rectangular configuration of cameras where object lies in the center of both cameras in Fig. 3, the measurements of both camera's background (x, y) is known. Now we need to find the length of principal axis of each camera to object placed in the middle of rectangular configuration. In this derivation, I.P.P stands for inches per pixel. Initially, we assume that each camera's view angle covers entire background. For first camera, we have:

$$I. P. P = \frac{y}{\text{width of first camera's image}}$$

$$\text{Distance}_{\text{from first camera}} = I. P. P \times X_{\text{Object}}$$

For second camera, we have

$$I. P. P = \frac{x}{\text{width of second camera's image}}$$

$$\text{Distance}_{\text{from second camera}} = I. P. P \times X_{\text{Object}}$$

When the view angle does not cover entire background then we have to make certain adjustment in above equations. In a real scenario, when the view angle of given camera does not cover the entire background i.e. Fig. 4, then I.P.P equations are changed slightly to accommodate this problem. This change is given as:

$$I. P. P = \frac{x}{\text{width of second camera's image}}$$

$$\text{Distance}_{\text{from second camera}} = I. P. P \times X_{\text{Object}} + \alpha$$

" α " is considered when the second camera is on the left side of the wall covering object. We present another situation when the view angle is larger than the background in Fig. 5. In this situation, the change in I.P.P is given as:

$$I.P.P = \frac{y}{\text{width of first camera's image}}$$

$$\text{Distance}_{\text{from first camer}} = I.P.P \times X_{\text{Object}} - \alpha$$

“ α ” is considered when the second camera is on the left side of the wall. This technique is best suited for indoor environment where the backgrounds of cameras can be measured. This process can be used to analyze the speed of coming car where the background of camera can be measured.

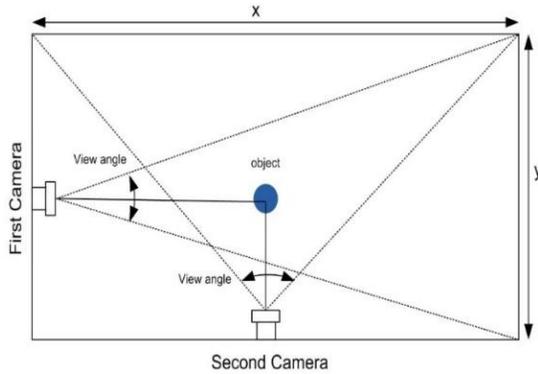


Fig. 3. Linear calibration in rectangular configuration where two cameras are perpendicularly placed from each other and object lies in the center of view angles of both cameras.

Algorithm 1: Linear Transformation in Stereo Configuration

- Result:** Find depth from stereo camera configuration
- 1 Get Horizontal view angle of cameras;
 - 2 READ images;
 - 3 Get distance between cameras;
 - For Left Camera;
 - 4 Divide view angle with x-resolution of image(A.P.P);
 - 5 Get internal angle within view angle of left camera(B’);
 - 6 Calculate angle between Right hand of view angle and base line (B);
 - 7 Add the internal angles(B’) with external(B) to get total angle (b);
 - 8 Get side of left camera by using the Law of sines;
 - For Right camera;
 - 9 Divide view angle with x-resolution of image(A.P.P);
 - 10 Get internal angle within view angle of right camera(A’);
 - 11 Calculate angle between left hand of view angle and base line (A);
 - 12 Add the internal angles(A’) with external(A) to get total angle (a);
 - 13 $c = 180^\circ - a - b$;
 - 14 Get side of Right camera by using the Law of sines;
 - 15 Use right angled trigonometric function for perpendicular distance from left camera according to the position of object’s image point(x,y);

C. Circular Calibration in Stereo Configuration

In previous two proposed techniques, view angle of the given camera is divided with the width of image to acquire angle of certain pixel which is then utilized in the triangulation technique for the extraction of the object’s distance from two cameras. But as the distance of object increases and object shifts away from center of image, these linear techniques start to give less accurate results of depth. The lens is circular body and the sensor is linear so there is no one-to-one relationship between object point on lens and its respective image point on sensor plane. This is the reason when object is in the center of image and closer to camera then depth extraction from linear techniques improves but as the object moves to the sides of image and distance between camera and object increases, the results of depth deteriorate. To improve this situation, we have to come up with closed form equation for the rectification of radially distorted images.

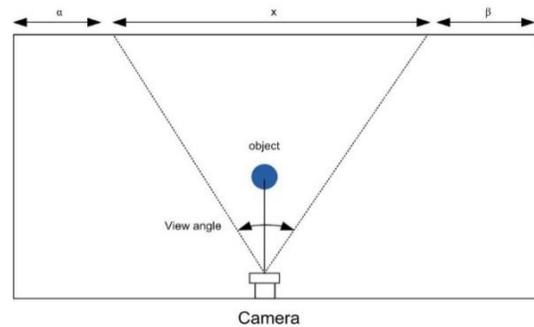


Fig. 4. Rectangular Configuration with small view angle, when the view angle is small compared to its background then we update the equations of linear calibration in rectangular configuration.

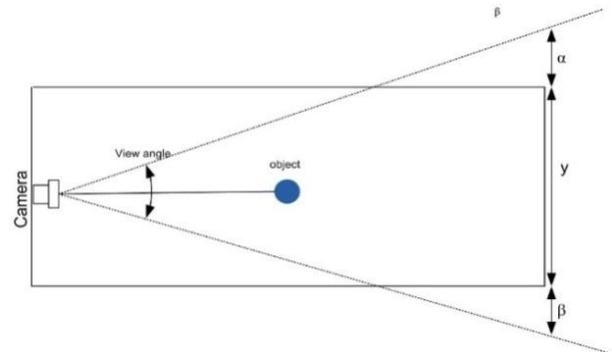


Fig. 5. Rectangular Configuration with large view angle. When the view angle is large compared to its background then we update the equations of linear calibration in rectangular configuration.

In Fig. 6 image and its field of view angle is divided among a/b and α/β , respectively, according to object location (x, y) on image plane. View angle of the given camera is the area cover by a sensor. Anything outside the image is also outside from view angle. A point in 3D world object (x, y, z) is represented by point $p(x, y)$ in an image plane by projective transformation. The view angle and image plane is related to each other mathematically. The view angle of the given camera is already known .which is divided in α and β horizontally according to the position of object in world plane. Its respective image is divided in "a" and "b" similarly according to the position of object in image plane. By using law of sines,

relation between α/β and a/b is established in 3, 4. After multiplying $\frac{1}{a}$ on both sides (5), $\frac{1}{b}$ on both sides of (6) and inverting these equations we acquire (7), (8).

$$\beta = \text{view angle} - \alpha$$

$$b = \text{image}_{\text{width}} - a$$

$$a: R = \sin \alpha : \sin \gamma \quad (3)$$

$$b: R = \sin \beta : \sin \gamma \quad (4)$$

$$\frac{a}{R} = \frac{\sin \alpha}{\sin \gamma} \quad (5)$$

$$\frac{b}{R} = \frac{\sin \beta}{\sin \gamma} \quad (6)$$

$$\frac{1}{R} = \frac{\sin \alpha}{a \sin \gamma}, \quad \frac{1}{R} = \frac{\sin \alpha}{b \sin \gamma} \quad (7)$$

$$R = \frac{a \sin \gamma}{\sin \alpha} \quad (8)$$

$$R = \frac{b \sin \gamma}{\sin \beta} \quad (9)$$

Equations (7) and (8) are equal, therefore we get:

$$\frac{a \sin \gamma}{\sin \alpha} = \frac{b \sin \gamma}{\sin \beta}$$

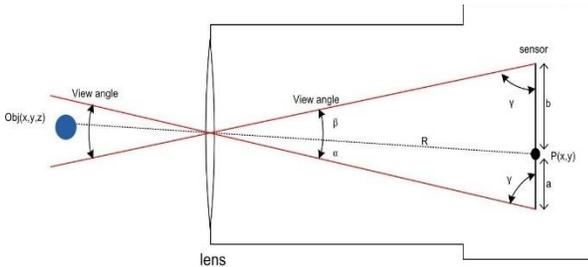


Fig. 6. Circular Calibration between lens and CCD. Ratio between the view angle and its image plane is utilized for the derivation of circular equation.

Multiply both sides with $\frac{1}{\sin \gamma}$ and shift "b" to left and $\sin \alpha$ to right, we have:

$$\frac{a}{\sin \alpha} = \frac{b}{\sin \beta}, \quad \frac{a}{b} = \frac{\sin \alpha}{\sin \beta}$$

For the ease of derivation of the final formula, we denote $\text{Cir} = \frac{a}{b}$ and use the difference of angles in sin formula. Let's also assume that view angle of our camera is equal to 40° then equation 9 becomes as follows, we have already discussed that $\beta = \text{view angle} - \alpha$ so above equation becomes:

$$\frac{a}{b} = \frac{\sin \alpha}{\text{view angle} - \alpha} \quad (10)$$

$$\text{Cir} = \frac{\sin \alpha}{(\sin 40^\circ \cos \alpha - \cos 40^\circ \sin \alpha)}$$

Now multiply the above equation with $\frac{1}{\sin \alpha}$ on both sides for further simplification, shift "Circular" to Right hand side and Invert the both sides of above equation, we get:

$$\frac{1}{\sin \alpha} \times \text{Cir} = \frac{\sin \alpha}{(\sin 40^\circ \cos \alpha - \cos 40^\circ \sin \alpha)} \times \frac{1}{\sin \alpha}$$

$$\frac{1}{\sin \alpha} = \frac{1}{(\sin 40^\circ \cos \alpha - \cos 40^\circ \sin \alpha) \times \text{Cir}}$$

$$\sin \alpha = (\sin 40^\circ \cos \alpha - \cos 40^\circ \sin \alpha) \quad (11)$$

Algorithm 2: Linear Transformation in Rectangular Configuration

Result: Find depth from Rectangular camera configuration

- 1 Get Horizontal view angle of cameras;
- 2 READ images;
- 3 Background=100;
- 4 view-angle;
- if view-angle-camera==Background then
- For First camera;
- 5 Divide "y" background of first camera of with width of image(I.P.P);
- 6 $\text{Depth} = I.P.P \times X_{\text{Object}}$;
- For Second camera;
- 7 Divide "x" background of second camera with width of image(I.P.P);
- 8 $\text{Depth} = I.P.P \times X_{\text{Object}}$;
- else if view-angle-camera < Background then
- For First camera;
- 9 Divide "y" background of first camera with width of image(I.P.P);
- 10 $\text{Depth} = I.P.P \times X_{\text{Object}} + \alpha$;
- For Second camera;
- 11 Divide "x" background of first camera with width of image(I.P.P);
- 12 $\text{Depth} = I.P.P \times X_{\text{Object}} + \alpha$;
- else if view-angle-camera > Background then
- For First camera;
- 13 Divide "y" background of first camera with width of image(I.P.P);
- 14 $\text{Depth} = I.P.P \times X_{\text{Object}} - \alpha$;
- For Second camera;
- 15 Divide "x" background of second with width of image(I.P.P);
- 16 $\text{Depth} = I.P.P \times X_{\text{Object}} - \alpha$;

Now for ease of representation, we have:

$$A = \text{Cir} \sin 40^\circ$$

$$B = \text{Cir} \cos 40^\circ$$

Now we will use above equations in equation 10 for further simplification:

$$\sin \alpha = A \cos \alpha - B \sin \alpha$$

Multiply both sides of above equation with $\frac{1}{\sin \alpha}$ we get

$$1 = \frac{A \cos \alpha - B \sin \alpha}{\sin \alpha}$$

$$1 = \frac{A \cos \alpha}{\sin \alpha} - B$$

According to the Trigonometric identities, $\cot^{-1} a = \frac{\sin \alpha}{\cos \alpha}$,
so after applying that we have

$$1 = A \cot^{-1} \alpha - B$$

After re arrangements of the above equation, we get:

$$\alpha = \cot^{-1} \frac{(B + 1)}{A}$$

Algorithm 3. Circular transformation in Stereo Configuration

Result: Find depth from stereo configuration using circular transformation

- 1 Get Horizontal view angle of cameras(view angle=40°);
- 2 READ images;
- 3 Get x,y coordinates of object in image from both cameras;
- 4 According to object position(x,y), calculate "a" in both cameras;
- 5 Calculate the distance between two cameras;
For Left camera;
- 6 $\beta = \text{view angle} - \alpha$;
- 7 $b = \text{image width} - a$;
- 8 $\text{Circular} = a/b$;
- 9 $A = \text{Circular} \sin 40^\circ$;
- 10 $B = \text{Circular} \cos 40^\circ$;
- 11 $\alpha = \cot^{-1}(\frac{B+1}{A})$;
- 12 Calculate angle between right hand of view angle and base line (B);
- 13 $b = \alpha + B$;
- 14 $C = 180^\circ - a - b$;
- 15 Get side of left camera by using the Law of sines;
For Right camera;
REPEAT step 6,7,8,9,10,11,12,13,15;
- 16 Use right angled trigonometric function for perpendicular distance of object from baseline;

This is the circular angle "α", calculated according to the geometric shape of the lens. Now depth extraction from the method of triangulation, discussed earlier, is improved significantly when the object lies on the sides of image or lens. This angle is added with external angles (A, B), discussed in Fig. 1. For the triangulation to extract the depth of object, object present at great distance will also be accurately triangulated according to this proposed technique.

In the initial two proposed techniques, due to the linear transformation, the produced data has some significant radial distortion which affects the accuracy of the overall system. To address this problem, we derive the third technique which reduces the radial distortion in an angle to exhibits accuracy of the depth of the object. If we have to acquire the depth of the object which is for example 400 Km away from the cameras then we need very high resolution cameras to capture the object. Limitation in our third technique, the acquired depth exhibits radial distortion when object moves away from the center of the image.

IV. RESULTS AND DISCUSSIONS

We used industrial grad cameras for the experiments to validate three proposed techniques. These cameras have 640×480 resolution with 1/3" optical format and produce 259 frames per second. To measure the success of the results produce by these techniques, we use measuring instruments to acquire the distance of the object from cameras. This data is then compared to the data produced by the three techniques to visualize the efficiency of the results. Calculation of intrinsic parameter of the camera required deployment of the precision engineering techniques to produce accurate results. Each technique required more than 100 experiments to acquire stable results. In initial phase of the testing we have used 5000 frames to check the authenticity of the depth data. The results can be produced from continuous video feed of the cameras which benefits many applications related to the surveillance or inspection. For first and third technique, we performed in indoor and outdoor environments to check the validity of the results. Whereas second technique requires the measurement of the object's background which allows us to place the camera setup only in indoor environment. We label the right image which is produced by the right camera and left image which produced by the left camera to differentiate between the objects images.

For analysis of the proposed technique, we compared depth extraction of our work with S. Sengupta's [26] method. Sengupta's [26] defined:

$$Z = \lambda + \frac{\lambda \Delta x}{x_L - (x_R + \Delta x)}$$

Where λ is focal length of camera and (x_L, x_R) are x-coordinates of detected object of left and right cameras with known distance (Δx) between them. By using this equation, in our experiment, we get around 740 mm while the actual distance of object from cameras is 440 mm.

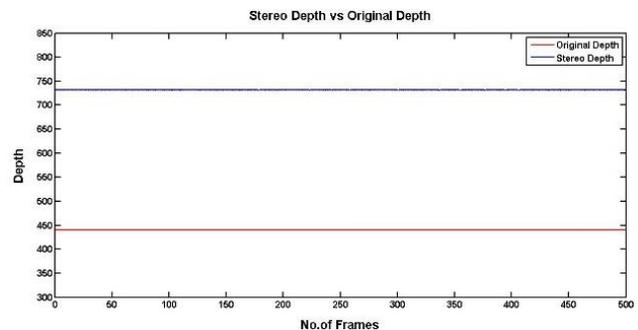


Fig. 7. Comparison between Original depth versus stereo calibrated depth. The data between original depth and stereo calibrated one is compared.

Blue line represents the stereo vision's depth through above mentioned technique, comparable to original depth. Red line represents original depth (440 mm). As we can see in Fig. 7, the computed depth from the above mention technique has 25% error margin compared to the original depth (440mm). Total number of frames is 500 to examine the accuracy of depth, acquired by the given technique.

We use bar graph representation of comparison between original depth of object and calibrated depth using

conventional stereo vision technique in Fig. 8. Original depth measured using measuring scale is 440 millimeter illustrated as red color while Blue color is calibrated through above mentioned technique. Bar graphs are also used to better visualize the depth acquired by the given technique.

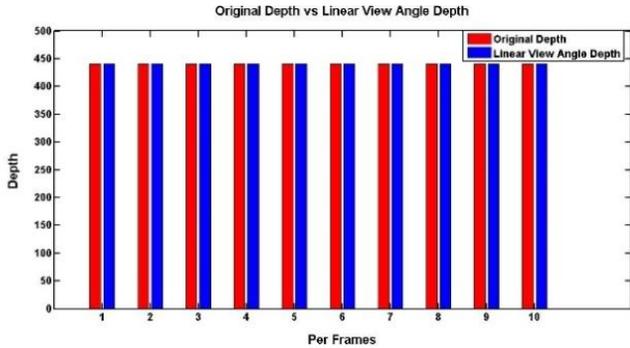


Fig. 8. This is the bar graph representation of comparison between Original depth and Stereo calibrated depth.

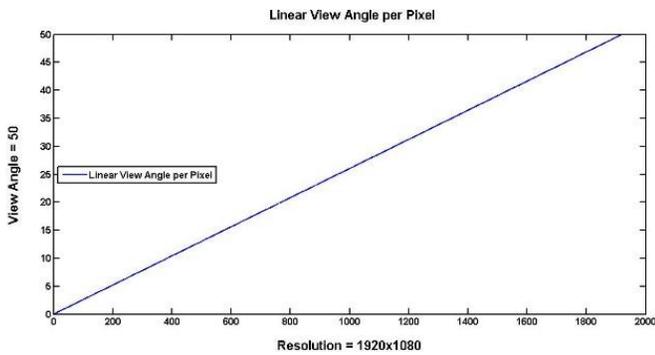


Fig. 9. Linear distribution of the angle per pixel. This is distribution of the angle per pixel through linear calibration technique among the pixels of image.

Fig. 9 shows the linear division of a view angle among the pixels of image. Initial and final pixels shows the sides of an image and it can be seen that this where the graph is straight line which results in less accurate depth calculation when the object lies on the sides of image. We place the object in center of image, in Fig. 10 where there is perfect one-to-one mapping between lens and image resolution. This is where the results are very accurate but due to linear division of the view angle on image resolution, the depth starts to deviate from its result on the sides of image.

In Fig. 11, Depth calculated using linear transformation technique in stereo configuration has staggering accuracy against original depth. It can be seen in the bar graphs about the accuracy of depth. Fig. 12 illustrates the data acquired by linear transformation in rectangular calibration in which left Camera is represented by light blue color and yellow color represents right camera. Red and light blue color shows Original depth for Left and Right cameras respectively. This technique is suitable in indoor environment where the background of each camera is known. The above bar graphs show very prominent results.

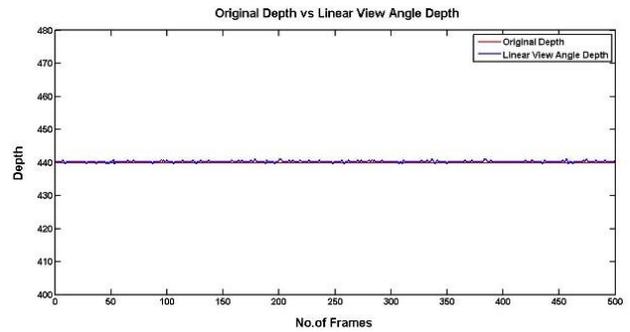


Fig. 10. Comparison of depths acquired through linear transformation (Linear view angle) in stereo configuration and original depth.

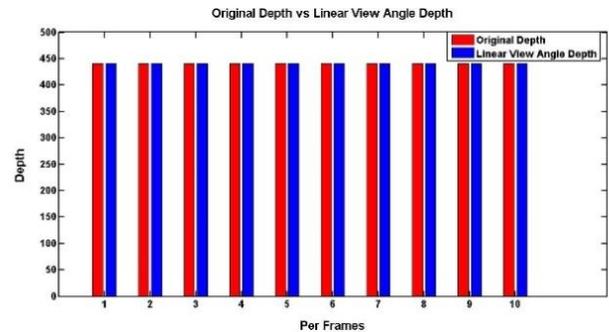


Fig. 11. This is the bar graph representation of linear transformation.

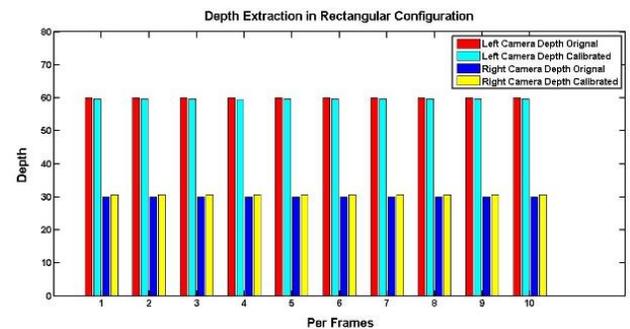


Fig. 12. Depth extraction in Rectangular Configuration. This figure illustrates comparison of rectangular calibration technique between two cameras versus original depths from cameras.

In Fig. 13 Circular Calibration technique is used to extract the depth of object from frames in stereo configuration. We can observe the accuracy of depth acquired using circular calibration against original depth. Red bar shows the Original depth and Blue bar shows calibrated depth. An image is divided among A and B where "A" is right side of detected object's pixel and B is the left side of image resolution. We move image center of the detected object from right to left as A increases and B decreases. Then we calculate the angle of each pixel calculated by the Circular angle per pixel formula which is already being derived.

Hartley and Zizzerman et al. [11] method is closest to our method and also very famous for their results. Comparatively, our method shows better stability in Depth extraction in

Fig. 14. The actual distance is 440 mm from cameras and our method produced less noisy results as compare to Zizzerman et al. [11] method. This method relays on the disparity map of the stereo image of pixel where motion or difference of color is considered for depth extraction. This consideration allows the depth extraction with noisy results. This technique is ideal when there are low cost cameras available and required low maintenance. Our technique is heavily influenced by this technique but produce less noisy results in depth extraction which further improve the functionality of imaging system with low cost development, deployment and maintainability. The purpose of choosing this state of the art technique is its ease of development and deployment from the local market perspective. Keeping these parameters in mind, we have produced better results as compared to Zizzerman's [11] technique.

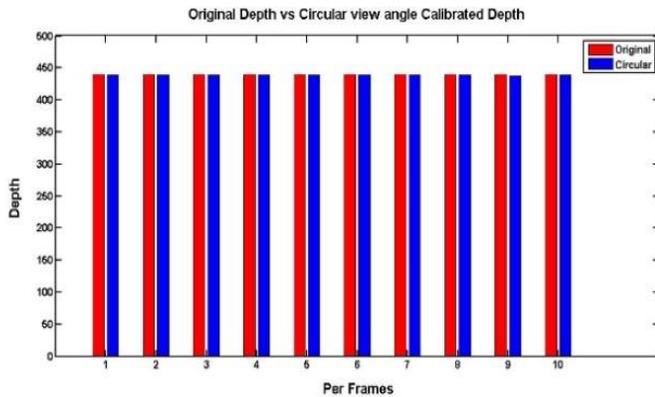


Fig. 13. Bar representation of comparison between depths acquired through Circular calibration technique and original depth.

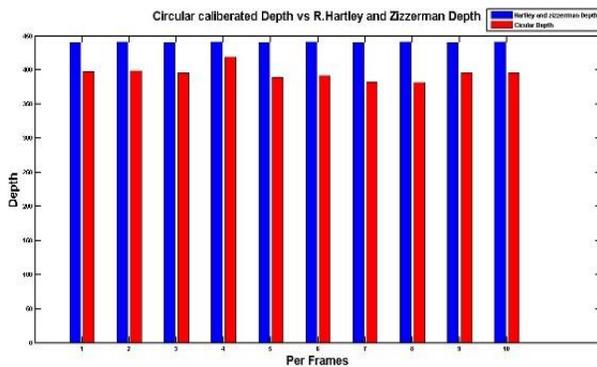


Fig. 14. Bar graph representation of comparison between Circularly Calibrated depth and R. Hartley, Zizzerman's [11] method.

There is a curve in Fig. 15 which is undetectable through naked eye but if we combined both graphs of linear and circular transformation of stereo configuration then we can see the circular angle per pixel depicts the $\cot^{-1} \alpha$ graph around 25° angle as shown in figure below. When the distance is great and object lies away from center of image then Circular calibration tends to hold the accuracy of depth close to 98% due to radial distortion correction of observed pixel.

In Fig. 16, the change in circular angle per pixel increases as it reaches towards the center of image and decreases when it leaves the center. This change in linear angle per pixel remains

the same which contributes to the inaccuracy in depth extraction when the distance between object and camera increases and object lies to the sides of image. The comparison shows difference between linear and circular angle per pixel.

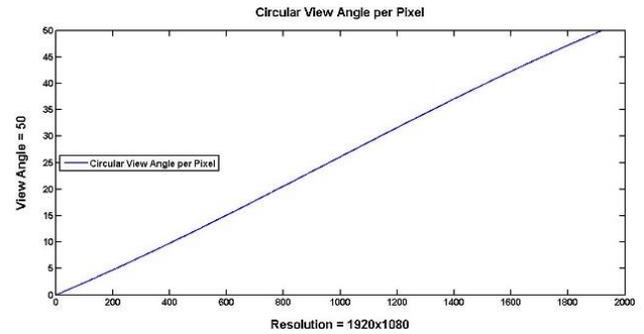


Fig. 15. Plot for circular angle A plot of circular angle per pixel for each pixel of 1920 resolution image.

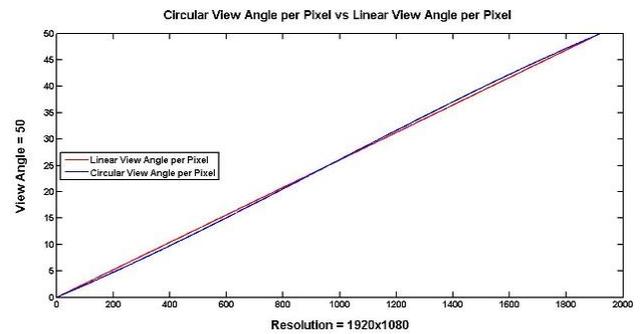


Fig. 16. Linear view angle per pixel versus Circular view angle per pixel Comparison between two graphs: linear view angle per pixel and Circular angle per pixel.

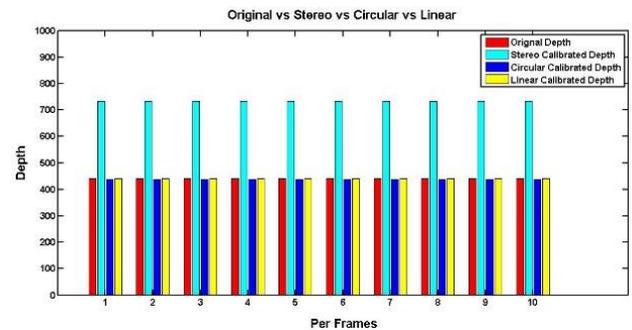


Fig. 17. Original vs. Circular View Angle vs. Linear View angle

In Fig. 17 linear calibration shows the one-to-one mapping between camera lens and image plane which gives us wrong approximation of an angle. This unrealistic result does not produce significant marginal error when the object is placed very near to cameras and lies in the center of image but when we extend the distance between object and cameras, also object shifts away from center of image then radial distortion starts playing its role in accuracy of depth which cannot be ignored.

Now question arises how much distance should exist between two cameras? There is no direct answer for this question as depth calculation or the perpendicular distance relies on angle made by two incident rays from object. The bigger the angle (distance between cameras increases) the more

accurate result we extract but if we increases the distance between cameras, we must increase their resolution so that incident rays coming from object must be properly absorbed by CCD through lens therefore inter-pixel accuracy will not be affected by less resolution camera.

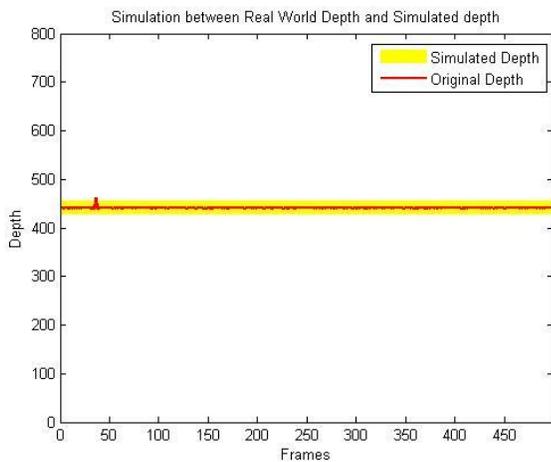


Fig. 18. Original depth acquired through cameras is projected as red color line and the depth acquired through simulation is presented as thick yellow line.

V. MODEL VALIDATION

We have derived the final mathematical formula for the radial distortion in stereo depth extraction. In this section, we have performed a simulation of our technique with comparison of the real world data. The view angle in the given equations is set to 40° which is same as in cameras which are being tested. In the real world scenario, we have placed the cameras apart with known distance and then applied circular calibration in triangulation for depth extraction. The same depth is also acquired through simulation of the equations by giving the same values of distance and view angle. This is done to see the amount of difference these equations produced when tested in real world scenario and in simulation.

In Fig. 18, we have acquired the data from 500 frames for the purpose of comparison to analyze with simulated data. We set the original depth line to thick yellow line because the difference between simulated and original data is very small. This simulation shows the accuracy of the data produced by circular calibration in depth extraction and insures the application of this technique in real world scenarios.

VI. CONCLUSION AND FUTURE WORK

In this work, we propose three techniques to extract the depth of object from image. In first technique, a linear transformation in stereo configuration is applied where view angle of the given camera is divided on image resolution to extract the distribution of view angle in pixels. Due the spherical shape of the lens, triangulation results are accurate in the central area of the image in close range but less accurate on the sides in long range. In the quest for improvement in depth extraction, we derive another technique, linear transformation in Rectangular Configuration where two cameras are placed perpendicularly from object. The backgrounds of both cameras are measured for the depth extraction. This procedure is most

suitable for indoor environment where the length of camera's background is known. In both techniques, when the objects shift away from center of image and distance between object and cameras increases, the distortion starts to hamper the accuracy of depth. To improve this situation, we derive circular equation to extract the angle of observed pixel according to geometric shape of the lens, which is used in triangulation for the depth extraction. Results show the accuracy of proposed solution where the depth of object is approximately 98% accurate, compared to the earlier techniques. We have also insured the cost of deployment and maintainability of the system is less than previous techniques.

We are currently working to further improve the accuracy because with the increase in resolution of the image, division of pixel's angle starts behaving linearly. In future work, we plan to acquire the height of the object from the surface using single camera by deploying circular calibration technique which will have significance scope in the surveillance and inspection applications.

REFERENCES

- [1] Brown, D.C.,1966.Decentering distortion of lenses,photogrammetric Engineering and Remote Sensing, , vol. 32, no. 3, pp.440-462.
- [2] Tsai, R.,1987.A versatile camera calibration technique for high-accuracy 3D machine vision metrology using off-the-shelf TV cameras and lenses. IEEE Journal on Robotics and Automation, Vol. 3,no. 4, pp. 323-344.
- [3] Lenz, R.K. and R.Y. Tsai, 1988. Techniques for calibration of the scale factor and image center for high accuracy 3-D machine vision metrology. IEEE Transactions on pattern analysis and machine intelligence, Vol. 10,no. 5, pp. 713-720.
- [4] Zhang, Z., 2000. A flexible new technique for camera calibration. IEEE Transactions on pattern analysis and machine intelligence, Vol. 22,no. 11, pp. 1330-1334.
- [5] Heikkila, J.,2000.Geometric camera calibration using circular control points. IEEE Transactions on pattern analysis and machine intelligence, Vol. 22,no. 10,pp. 1066-1077.
- [6] Nomura, Y., et al.,1992Simple calibration algorithm for high-distortion lens camera. IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 14,no. 11,pp. 1095-1099.
- [7] Levenberg, K.,1944.A method for the solution of certain non-linear problems in least squares. Quarterly of applied mathematics, Vol.2,no. 2,pp. 164-168.
- [8] Devernay, F. and O. Faugeras,2001.Straight lines have to be straight. Machine vision and applications, 13(1): p. 14-24.
- [9] Dobbert, T.,2006.Matchmoving: the invisible art of camera tracking.: John Wiley & Sons.
- [10] Fitzgibbon, A.W.2001.Simultaneous linear estimation of multiple view geometry and lens distortion. in Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on. IEEE.
- [11] Hartley, R. and A. Zisserman,2005.Multiple view geometry in computer vision. Robotica, Vol. 23,no. 2,pp. 271-271.
- [12] Herrera, D., J. Kannala, and J. Heikkilä,2012.Joint depth and color camera calibration with distortion correction. IEEE Transactions on Pattern Analysis and Machine Intelligence,Vol. 34, no. 10,pp. 2058-2064.
- [13] Oth, L., et al.2013.Rolling shutter camera calibration. in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition.
- [14] Rahman, T. and N. Krouglicof,2012.An efficient camera calibration technique offering robustness and accuracy over a wide range of lens distortion. IEEE Transactions on Image Processing,Vol. 21,no. 2,pp. 626-637.
- [15] Sun, Q., et al.,2016.Camera self-calibration with lens distortion. Optik-International Journal for Light and Electron Optics, Vol. 127,no.10,pp. 4506-4513.

- [16] Miks, A. and J. Novak, 2012. Dependence of camera lens induced radial distortion and circle of confusion on object position. *Optics & Laser Technology*, Vol. 44, no. 4, pp. 1043-1049.
- [17] Zhou, L., G.-h. Gu, and N. Shao, 2016. Straight lines fast correction algorithm for lens array distortion. in *SPIE Optical Engineering+ Applications*. International Society for Optics and Photonics.
- [18] Andrade, J. and L.J. Karam, 2016. Robust radial distortion correction based on alternate optimization. in *Image Processing (ICIP)*, IEEE International Conference on. 2016. IEEE.
- [19] Geiger, A., et al. 2012. Automatic camera and range sensor calibration using a single shot. in *Robotics and Automation (ICRA)*, IEEE International Conference on. 2012. IEEE.
- [20] Alzarok, H., S. Fletcher, and A.P. Longstaff, 2016. A new strategy for improving vision based tracking accuracy based on utilization of camera calibration information. in *Automation and Computing (ICAC)*, 2016 22nd International Conference on. IEEE.
- [21] Do, V.-H., et al. 2015. A simple camera calibration method for vehicle velocity estimation. in *Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON)*, 12th International Conference on. 2015. IEEE.
- [22] Houssineau, J., et al., 2016. A unified approach for multi-object triangulation, tracking and camera calibration. *IEEE Transactions on Signal Processing*, Vol. 64, no. 11, pp. 2934-2948.
- [23] de Paula, M.B. and C.R. Jung, 2015. Automatic detection and classification of road lane markings using onboard vehicular cameras. *IEEE Transactions on Intelligent Transportation Systems*, Vol. 16, no. 6, pp. 3160-3169.
- [24] Jung, J., et al. 2016. Human height analysis using multiple uncalibrated cameras. in *Consumer Electronics (ICCE)*, IEEE International Conference on. 2016. IEEE.
- [25] McCollough, E., 1893. *Photographic topography*. Industry: A Monthly Magazine Devoted to Science, Engineering and Mechanic Arts, Vol. 54, pp. 65.
- [26] Sengupta, S., 1997. Effects of unequal focal lengths in stereo imaging. *Pattern recognition letters*, Vol. 18, no. 4, pp. 395-400.