2004－CVIM－143　(9)
2004／3／5

A CALIBRATED PINHOLE CAMERA MODEL FOR SINGLE VIEWPOINT IMAGING SYSTEMS

Daniel Moldovan, Yuichi Yoshioka and Toshikazu Wada

Department of Computer and Communication Science, Wakayama University,
930 Sakaedani, Wakayama, 640-8510, Japan
E-mail: daniel@vrl.sys.wakayama-u.ac.jp

Abstract This paper presents a perspective imaging model that can be used to represent any single viewpoint imaging system. Our imaging model consists of an optical center, whose position is well determined in the 3D space, and a virtual screen on which undistorted images taken with the uncalibrated camera are displayed. This model uses two screens as virtual image planes, and the optical center is estimated as the converging point of those rays that are passing through the corresponding points on these planes. This method has the advantages that 1) it can be applied to any single viewpoint camera, and 2) it can remove any type of distortions. In the experiments, results for normal and omni-directional cameras demonstrate the effectiveness of our method.
Keywords single viewpoint imaging system, calibrated pinhole camera.

1. INTRODUCTION

Several techniques have appeared in the last decade capable of recovering the complete three-dimensional information from multiple images. They employ either calibrated or uncalibrated cameras. An important feature of algorithms using uncalibrated cameras is that no knowledge of the camera's positions and internal parameters is required. However, in the presence of nonlinear lens distortion the pinhole camera model is no longer valid and the technique breaks down. On the other hand, mathematical models of lens distortion developed for calibrated cameras compensated the errors but could not completely eliminate them.

This paper is presenting a perspective imaging model that can be used to represent any single viewpoint imaging systems, imposing no restriction to the properties of the imaging system: it can be perspective or non-perspective.

Our imaging model is based on a perspective projection and consists of an optical center, whose position is well determined in the 3D space, and a virtual screen on which we'll display undistorted images taken with the uncalibrated camera. Virtually, this new imaging model can be regarded as a calibrated pinhole camera.

Our technique for detecting the optical center uses calibration objects in order to find 3D corresponding points that are projected on the same pixel on the image plane. The approach is based on a virtual modelling in which calibration objects are replaced by virtual screens (Fig. 1) that have the same dimension and structure with the calibration objects. By finding the 3D points on these virtual screens that are projected on the same pixel on the actual image plane, we can track the rays of light that converge to the optical center of the imaging model. Thus, the 3D position of the optical center can be estimated as the intersection point of a large number of rays. Moreover, our virtual model allows a visualization of the optical center too.

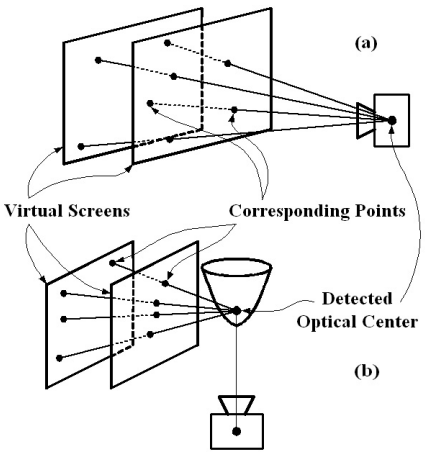

Fig. 1. Imaging systems that can be emulated by our imaging model: (a) a perspective camera, (b) a single viewpoint catadioptric system.

Images taken with uncalibrated camera are then back-projected on these virtual screens based on our interpolation techniques that are adjusted for each type of imaging system. Because of the regular shape of the

virtual screens, the interpolation process allows the removal of the geometric distortions obtaining in the end a distortion free virtual screen.

1.1. Related Works

Although carried out independently, our work contains a few aspects that are similar with the ones that are present in Grossberg's work [1]. Thus, in order to represent any arbitrary imaging system, they introduced a new imaging model that was based on determining the locus of viewpoints (i.e. the caustic) of the imaging system by using the Jacobian method. Their model proved to be similar with ours for the particular case of perspective imaging systems. Both techniques converged toward the determination of the optical center of the perspective camera.

However, the implementation methods differ. Thus, in order to determine the ray-to-image mapping, they binary coded each pixel of an active LCD display. In our method we used an bilinear interpolation process that was applied to a chess board type calibration pattern.

Differences between these two works are best revealed for the case of single viewpoint imaging systems that exhibit a non-perspective projection. While their method is determining the envelope of the reflected rays (i.e. the catacaustic), we are estimating the point of intersection of the incident rays. Also, in their calibration technique of the rotationally symetric imaging systems or of the imaging systems with an arbitrary field of view they neglected the deviation between the rotation axis and the optical center. This deviation will obstruct the correct convergence of the reflected rays toward the optical center of the camera. Our paper presents also a method to compensate this deviation.

Recently, a geometric method for estimating the intrinsic parameters of a catadioptric system consisting of a parabolic mirror and an orthographic lens have been investigated by Geyer and Daniilidis [2]. They introduced the geometry of catadioptric line projection and showed that the vanishing points lie on a conic section which encodes the entire calibration information.

Aliaga [3] used an omnidirectional camera based on the design of Nayar [4] to develop a calibration model that relaxes the assumption of an ideal projection system and compensates for mild perspective projections in addition to radial distortion and mirror misalignment. The projection of undistorted images on virtual screens was used by Yamazawa [5]. They used a grid pattern and by employing a bilinear interpolation process they projected the captured image on a perfectly rectilinear pattern. In this way they succeeded to eliminate the errors induced by lens artifacts. In order to perform the camera calibration they subsequently applied Tsai's method [6].

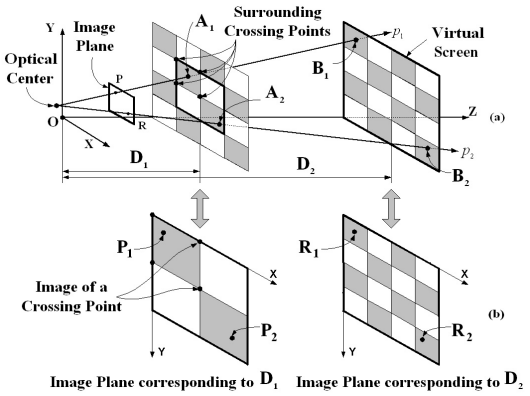

Fig. 2. Setup for Perspective Camera

2. PERSPECTIVE CAMERA

A schematic of our virtual model for perspective camera is depicted in Fig. 2(a). Our virtual screen resembles the shape of the chess board type calibration object that was placed consecutively at distances $\mathbf{D}_1$ and $\mathbf{D}_2$ in front of the camera. We depicted representations of the images on the corresponding image planes in Fig. 2(b).

A ray of light proceeding from the camera's optical center and passing through 3D point $\mathbf{A}_1$ will intersect the corresponding image plane in point $\mathbf{P}_1$. Similarly, a ray of light passing through $\mathbf{B}_1$ and the optical center will have its corresponding point $\mathbf{R}_1$. If $\mathbf{P}_1$ and $\mathbf{R}_1$ are identical (have the same coordinates), the corresponding rays of light will have in common two points (the optical center and the point on the image plane), a sufficient condition for the two rays to be identical. This result implies that points $\mathbf{A}_1$, $\mathbf{B}_1$, point $\mathbf{P}$ on the image plane and the optical center are collinear.

By using this reasoning, we'll be able to find the location of the optical center as the crossing point of a multitude of rays of light generated by pairs of 3D points that have the same projection on the image plane. In order to detect these pairs we established first the correspondences between points on the virtual screens and their projections on the image plane. Based on a bilinear interpolation transformation, each point on the screen can be expressed as a function of the surrounding crossing points.

Knowing the 3D world coordinates of the crossing points and the coordinates of their images on the image plane, and by using the same interpolation function, a correspondence between each pixel in the image plane and 3D points on the virtual screen could be found. Subsequently, for each intermediate distance we built a lookup table that matched the pixels in the image plane with the corresponding 3D points located on the calibration pattern. By iterating a search process on the resulting two tables, 3D points that correspond to the same pixels could be easily detected.

Using the same interpolation technique, pixels from

the image plane can be back-projected on the virtual screens. Because of the regular shape of the virtual screens, the obtained images will be distortion free.

3. SINGLE VIEWPOINT OMNIDIRECTIONAL CAMERA

For the case of single viewpoint omnidirectional cameras, we used a catadioptric system consisting of a perspective camera looking into a hyperbolic mirror (Fig. 3).

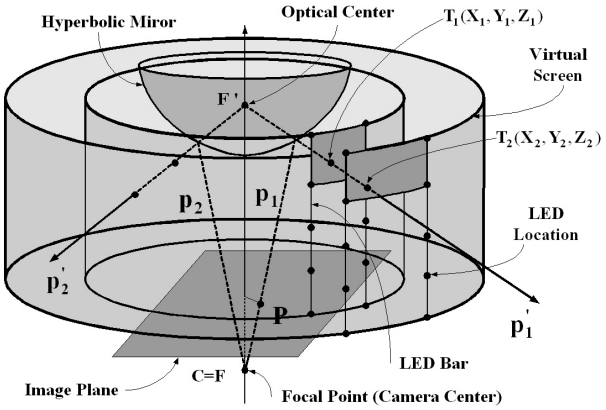

Fig. 3. Catadioptric System with a Hyperbolic Mirror

The ray $\mathbf{p}_1$ proceeding from the camera is reflected by the mirror as a ray $\mathbf{p}_1^{'}$. In the case of catadioptric systems with hyperbolic mirrors, all the reflected rays $\mathbf{p}^{'}$ intersect at the focal point of the mirror $\mathbf{F}^{'}$ and the camera center of projection **C** coincides with the second focal point of the mirror, **F**. Because of the uniqueness of the projection center **C**, if two 3D points from the surrounding environment have the same representation on the image plane we can say that they lie on the same reflected ray that will include also the mirror's focal point. Our purpose is to find the optical center which is represented by the mirror's focal point.

Like in the previous case, we built a virtual model that consists of two concentric cylinder-shaped virtual screens, with the omnidirectional camera at their center. The corresponding points will be found from these two cylinders. Each of them resembles the surface generated by an imaginary rotation of an LED bar around the camera. In order to detect the pairs of corresponding points we used two aproaches: the first one approximated the virtual screen with a right prism that had a variable number of faces while the second one considered it as a cylinder. First approach employed a bilinear interpolation method while the second employed a combination of linear and circular interpolation.

In order to construct these cylinders we appealed to a practical method. We built an LED bar that was placed at a certain distance in front of the omnidirectional camera. Subsequently, we performed a complete rotation of the camera around its revolution axis in consecutive steps from 15 to 15 degrees. For each intermediate angle, we lit up one by one all LEDs and recorded their image plane coordinates. This process is equivalent with finding the image plane coordinates of lighted LEDs obtained by rotating the LED bar around the camera.

Knowing the 3D coordinates of the LEDs on the virtual screen and the coordinates of their projections on the image plane, a correspondence between 3D points that have the same projection on the image plane could be found. By decreasing the size of the rotation angle, the size of the rectangles will decrease also and the final shape of the virtual screen will resemble the shape of a cylinder with a higher fidelity. By using the same interpolation technique, undistorted images can be projected on the virtual screen.

4. ESTIMATING THE OPTICAL CENTER

Two lines in 3D generally don't intersect at a point. They may be parallel (no intersection) or they may be coincident (infinite intersections) but most often only their projection onto a plane intersect. When they don't exactly intersect at a point they can be connected by a line segment, the shortest line segment is unique and is often considered to be their intersection in 3D.

In order to find the point that is located at the shortest distance to all of the generated rays of light, we resorted to the use of a nonlinear least squares method that computes an approximate solution by minimizing the sum of all the distances from an arbitrary point to all the rays.

Considering Fig. 4, the vector **OR** from an arbitrary point **O** to a line **d** in the direction of a vector **v** that includes point **P** can be expressed as follows:

$$\mathbf{OR} = \mathbf{PR} - \mathbf{PO}$$

$$\text{with } \mathbf{PR} = (\mathbf{PO} \cdot \mathbf{v})\,\mathbf{v} = \mathbf{v}^T\mathbf{PO}\,\mathbf{v}$$

As a result:

$$\mathbf{OR} = \mathbf{v}^T\mathbf{PO}\,\mathbf{v} - \mathbf{PO} = \mathbf{v}^T(\mathbf{P}-\mathbf{O})\,\mathbf{v} - (\mathbf{P}-\mathbf{O}) =$$

$$= \mathbf{v}^T\mathbf{P}\,\mathbf{v} - \mathbf{v}^T\mathbf{O}\,\mathbf{v} - \mathbf{P} + \mathbf{O}$$

Having given 3-D coordinates of

$$\mathbf{P} = \begin{pmatrix} x_0 \\ y_0 \\ z_0 \end{pmatrix}, \mathbf{O} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} \text{ and } \mathbf{v} = \begin{pmatrix} v_x \\ v_y \\ v_z \end{pmatrix},$$

distance

$$\mathbf{L} = \| \mathbf{OR} \|$$

can than be expressed as:

$$\mathbf{L} = \mathbf{AX} - \mathbf{B}$$

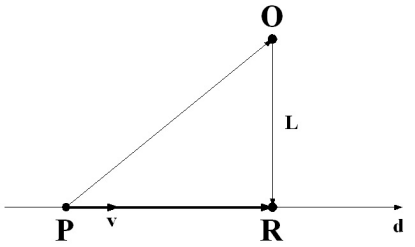

Fig. 4. Distance from a Point to a Given Line

with

$$\mathbf{A} = \begin{pmatrix} v_x v_x - 1 & v_x v_y & v_x v_y \\ v_x v_y & v_y v_y - 1 & v_y v_z \\ v_x v_z & v_y v_z & v_z v_z - 1 \end{pmatrix}$$

$$\mathbf{B} = \begin{pmatrix} (v_x v_x - 1)x_0 + v_x v_y y_0 + v_x v_z z_0 \\ v_x v_y x_0 + (v_y v_y - 1)y_0 + v_y v_z z_0 \\ v_x v_z x_0 + v_y v_z y_0 + (v_z v_z - 1)z_0 \end{pmatrix}$$

$$\text{and } \mathbf{X} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}.$$

The condition that point **O** will have to satisfy in order to coincide with the crossing point of a number of n rays of light will be:

$$\mathbf{L}_i = 0 \text{ for } i = \overline{1,n}$$

were $\mathbf{L}_i$ is the distance from point **O** to ray i.

This is equivalent with a linear system of equations

$$\mathbf{A}_i\mathbf{X} = \mathbf{B}_i \text{ for } i = \overline{1,n}$$

Rather than expecting no error in some equations and large errors in the others it is much better to choose **X** that minimizes the average error in the n equations. The most convenient way to define such an average is the sum of squares:

$$\mathbf{E}^2 = \sum_{i=1}^{n} (\mathbf{A}_i\mathbf{X} - \mathbf{B}_i)^2 \text{ for } i = \overline{1,n}$$

$\mathbf{E}^2$ will have a minimum at the point where

$$\frac{d\mathbf{E}^2}{d\mathbf{X}} = 0.$$

The solution given by the calculus is

$$\mathbf{X} = \frac{\mathbf{A}^T\mathbf{B}}{\mathbf{A}^T\mathbf{A}} \text{ with } \mathbf{A} = \begin{pmatrix} \mathbf{A}_1 \\ \mathbf{A}_2 \\ \vdots \\ \mathbf{A}_n \end{pmatrix}, \mathbf{B} = \begin{pmatrix} \mathbf{B}_1 \\ \mathbf{B}_2 \\ \vdots \\ \mathbf{B}_n \end{pmatrix}$$

and $\mathbf{A}_i, \mathbf{B}_i$ being the corresponding **A** and **B** for line $i = \overline{1,n}$.

5. EXPERIMENTAL RESULTS

For the case of perspective camera, the calibration setup consisted of: (1) a pattern obtained by printing out a chess board like poster using a laser printer, (2) a SONY digital color camera with a resolution of 640 x 480 pixels and (3) a sliding rail. For the case of omnidirectional camera, the calibration setup consisted of: (1) a calibration pattern (an LED bar whose lighting was controlled from the parallel port), (2) an analog color camera with a resolution of 640 x 480 pixels, (3) a catadioptric attachment (paraboloid mirror), (4) a sliding rail, (5) a Pan-Tilt Unit (controlled from the serial port). For both experiments we used a 2GHz PC with 2GB memory, a video capture board and a flat panel display.

5.1. Perspective Camera

In order to check the correctness of the calibration result, the experiments were performed for two distinct combinations of the camera and calibration pattern. First combination required the placement of the calibration pattern at a distance of 25cm and 35cm respectively from the camera (Fig. 5).

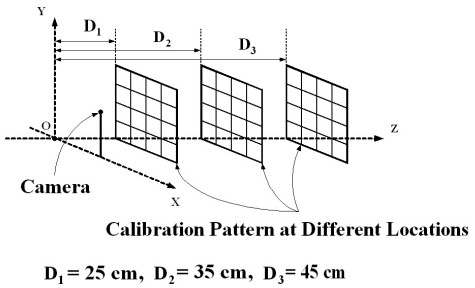

Fig. 5. Experimental Setup for Perspective Camera

We placed the camera along the X axes in such a way that the front part of the lens was located at 5cm from the origin of the world coordinate system **O**. The second combination required the calibration pattern to be placed at 25cm and 45cm respectively from the camera. The results of the calibration process are presented in Tabel 1.

Combination	X[cm]	Y[cm]	Z[cm]
$First$	5.99	5.205	-5.37
$Second$	5.97	5.145	-5.54

Table 1. Calibration Results for Perspective Camera

Negative value of Z coordinate is placing the optical center as was expected, inside of the camera. As can be noticed, the location of the optical center has slight variations (less than 2 mm) on Y and Z directions, mainly caused by the calculations accuracy. The converging rays of light are visualized in Fig. 6.

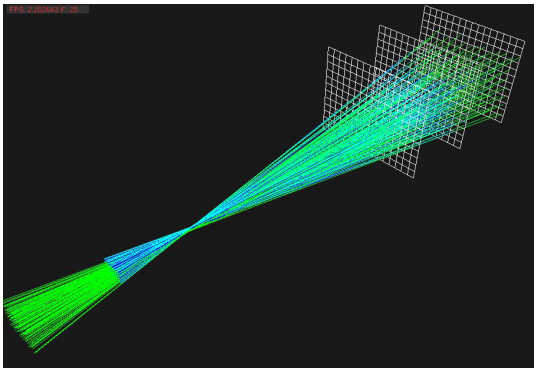

Fig. 6. Converging Rays for Perspective Camera

5.2. Single Viewpoint Omnidirectional Camera

For the case of omnidirectional camera, the generated rays of light will include the optical center if the vertical axes of the LED bar for a certain rotation angle will be coplanar with the rotation axis of the camera (Fig. 7). If this condition is not accomplished and if a deviation angle is inserted, the location of the optical center will be missed (Fig. 8).

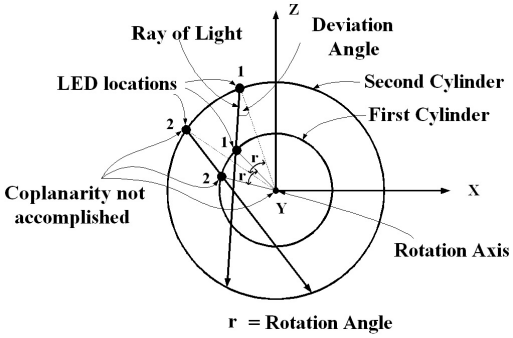

Fig. 7. Coplanarity is required in order to have convergent rays of light.

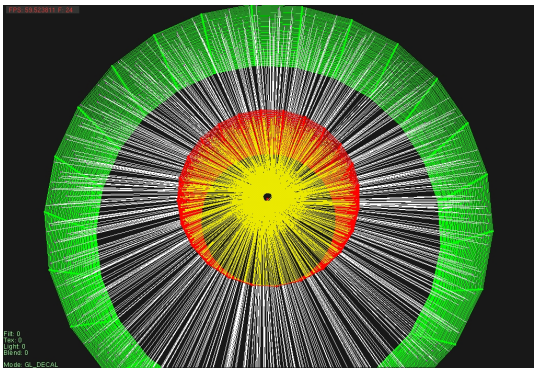

Fig. 8. Erroneous placement of the Calibration Pattern generates a locus of the Optical Center on the circumference of a circle.

In order to detect the deviation angle we applied the following method. Firstly, we determined the line equations that are corresponding to the projections of LEDs on the image plane for two different distances camera-LED bar (Fig. 9). We used least-squares line method to approximate the best fitted lines. Further, we detected their crossing point which represents also the projection of the rotation axis on the image plane. The results obtained confirmed that the two crossing

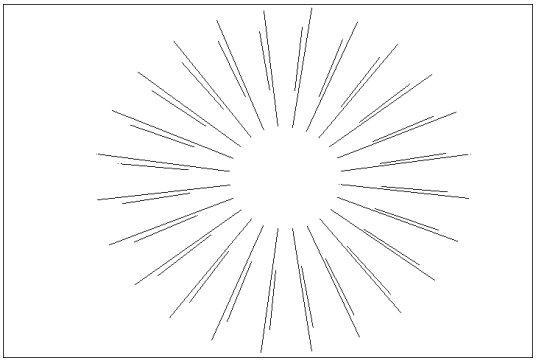

Fig. 9. Lines through the sets of pixels representing the projections of LEDs for two consecutive distances

points were identical. This outcome gave us the common element that was needed in order to compare the two sets of lines. The deviation between the lines corresponding to different distances represents actually the deviation inserted by erroneous placement of the LED bar. Once the deviation angle was detected, we built the cylinders accordingly, the final result being presented in Fig. 10.

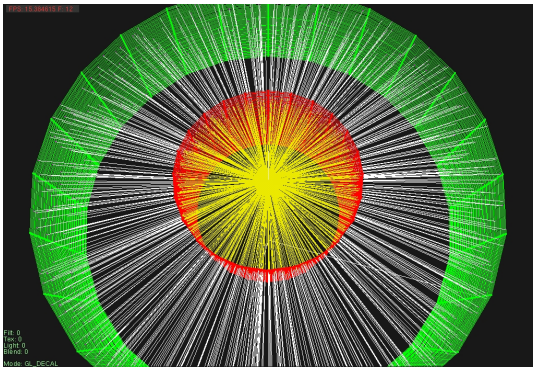

Fig. 10. After proper correction, the locus of the Optical Center will be a point.

In the experiments the sliding rail was oriented perpendicular towards the rotation axis of the omnidirectional camera that was marked by the producer of the catadioptric attachment. The origin of the world coordinate system was placed inside of the calibration patterns and it had the same orientation of the axes as in the previous experiment. The measurements were performed for two distinct combinations of the camera and calibration pattern. First combination required the calibration pattern to be placed at 13cm and 33cm

respectively from the camera. For the second combination, the calibration pattern was placed at 13cm and 23cm respectively. The results of the calibration process are presented in Tabel 2.

Combination	X[cm]	Y[cm]	Z[cm]
$First$	0.0	393.12	0.0
$Second$	0.0	393.16	0.0

Table 2. Calibration Results for Omnidirectional Camera

The images back projected on the virtual screens for the first approach are presented in Fig. 11. As can be noticed, the objects in the image are keeping their proportions and the only present deformation occurs when two rectangles from consecutive angles are joined. However, this distortion can be reduced by decreasing the rotation angle or it can be eliminated by using cylindrical virtual screens Fig. 12.

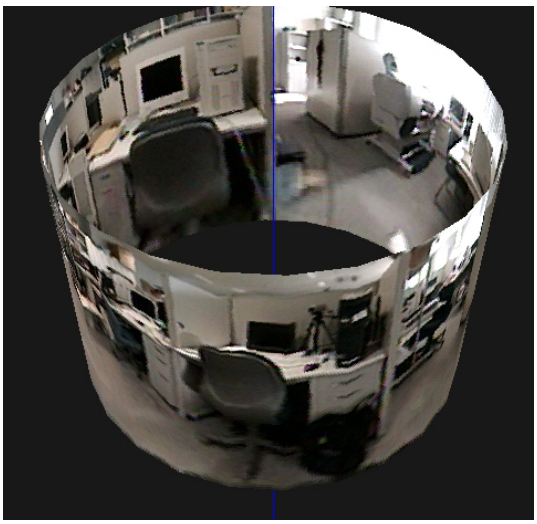

Fig. 11. Image backprojected on the virtual screen - right prism approach

By combining backprojected images with the converging rays we get the final representations of the virtual screens and the optical center estimated as the converging point of those rays (Fig. 13 for right prism approach and Fig. 14 for cylindrical approach).

6. CONCLUSION

The main contribution of this paper was the presentation of a pinhole camera model that can be used to represent any single viewpoint imaging system. This model uses two screens as virtual image planes, and the optical center is estimated as the converging point of those rays that are passing through the corresponding points on these planes. The advantages of this method are that 1) it can be applied to any single viewpoint camera, and 2) it can remove any type of distortions. In the experiments, results for normal and omnidirectional cameras demonstrate the effectiveness of our method.

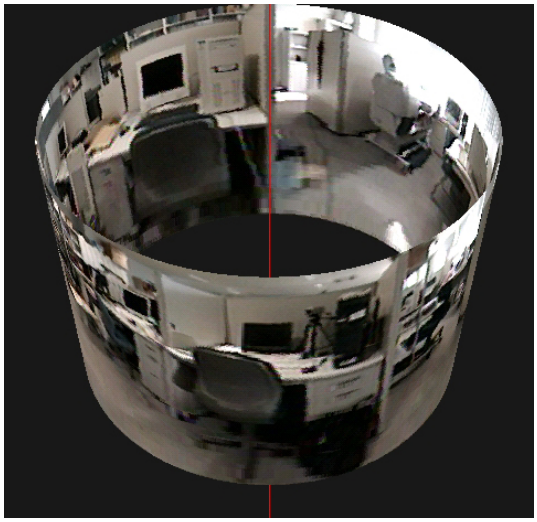

Fig. 12. Image backprojected on the virtual screen - cylinder approach

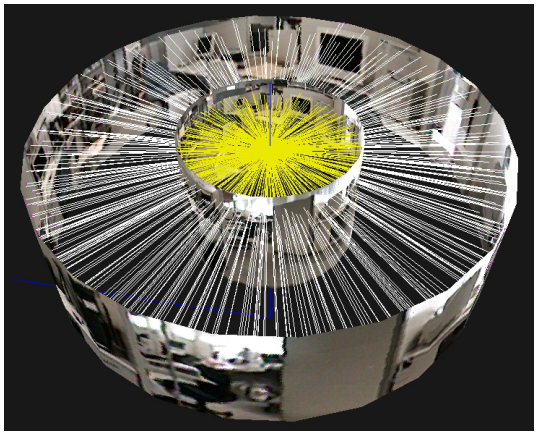

Fig. 13. Right Prism Approach

7. REFERENCES

[1] Shree K. Nayar M.D. Grossberg, "A general imaging model and a method for finding its parameters," in *Proceedings of ICCV*, 2001.

[2] C. Geyer and K. Daniilidis, "Catadioptric camera calibration," in *Proceedings of the Seventh International Conference on Computer Vision*, September 1999, pp. 398–404.

[3] Daniel G. Aliaga, "Accurate catadioptric calibration for real-time pose estimation in room-size environments," in *Proceedings of ICCV*, 2001, pp. 127–134.

[4] S.K. Nayar, "Catadioptric omnidirectional camera," in *CVPR*, 1997, pp. 482–488.

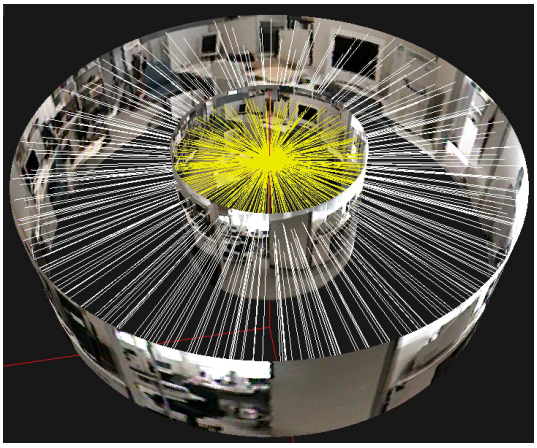

Fig. 14. Cylindrical Approach

[5] Y. Yagi K. Yamazawa and M. Yachida, "Omnidirectional imaging with hyperboloidal projection," in *Proc. of the Int'l Conf. on Robots and Systems*, 1993.

[6] R.Y. Tsai, "A versatile camera calibration technique for high-accuracy 3d machine vision metrology using off-the-shelf tv cameras and lenses," *IEEE J. of Robotics and Automation*, vol. RA-3(4), pp. 323–344, Aug. 1987.