Calibration of Rotating Line Spherical Camera based on Checkerboard Pattern on Multiple Planes and its Accuracy Assessment

Xiaoyin Guan[1]
http://www.uclan.ac.uk/scitech/adsip
Lik-Kwan Shark[1]
http://www.uclan.ac.uk/scitech/adsip
Geoff Hall[1]
http://www.uclan.ac.uk/scitech/adsip
Wei Deng[2]
http://www.dem-solutions.com

[1] ADSIP Research Centre
University of Central Lancashire
Preston, UK

[2] DEM Solutions
47 Queen Street
Edinburgh, UK

Abstract

This paper focuses on calibrating a rotating line spherical camera with a fish-eye lens to provide a large vertical field of view for 360° panoramic imaging. With simplicity in mind, the proposed method uses a simple calibration object based on a small five-sided open box with a checkerboard pattern on each plane, a simple equidistance camera projection model with an additional camera parameter to characterise the vertical field of view, as well as a simple iterative computation process to estimate the required camera projection parameters based on non-coplanar corner points and a hemisphere calibration image. Also presented in the paper are some experimental results to demonstrate the achievable calibration accuracy, and to show the influence of the number of corners used and the number of planes employed on calibration accuracy.

1 Introduction

Image based immersive virtual reality often requires capturing high resolution images with a large field of view at the image acquisition step. Many different types of digital cameras are available for this purpose, namely, omnidirectional cameras with wide-angle or fish-eye lens, central catadioptric cameras combined with lenses and mirrors, and cameras with multiple sensors [1-3].

Different calibration methods have been proposed for these special cameras. One calibration method is based on the use of a large calibration room/field attached with target points [4-5]. Although this method has the advantage of accuracy with more than 20 imaging parameters generated for the camera model, it has the disadvantages of a large space requirement and time consuming preparation. To enable direct use of real world environment, a calibration method has been proposed based on straight line segments extracted from man-made environment [6] and it relies on a sufficient number of straight lines available in the

scene. A further calibration method requires multiple views, whereby the epipolar geometry between two views is utilized to enable image correspondences being searched for camera calibration [1, 7].

Different projection models have also been used in calibration of these special cameras. For cameras with fish-eye lenses, the proposed projection models include the simple equidistance function [8], a general higher order polynomial form with the result modified by circular control points [2], and a rational function with image points 'lifted' to higher power and the size of the camera projection matrix increased to 3 by 6 [9].

This paper focuses on calibration of a rotating line spherical camera with a fish-eye lens (SpheroCam developed by SpheronVR AG). The proposed method was developed with simplicity in mind. It uses a calibration object based on a small five-sided open box with a checkerboard pattern on each plane, a simple equidistance projection model with an additional camera parameter to characterise the vertical field of view, and a simple iterative computation process to estimate the required camera projection parameters. Also presented in the paper are the experimental results to show the achievable accuracy of the proposed calibration method as well as the influence of the number of corner points and the number of planes on calibration accuracy.

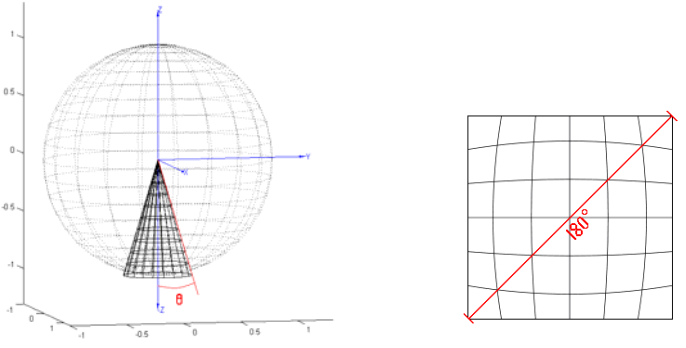
The paper is organised in five sections. Section 2 presents the basics of the SpheroCam, the vertical field of view as the additional camera parameter, and the calibration box. This is followed by the simple calibration methodology based on the corner-point grid patterns on multiple planes in Section 3, and experimental results in Section 4 to show the accuracy as a function of the number of corner points and planes used for calibration. The last section concludes the work.

2 Spherical Camera and Calibration Setting

SpheroCam, together with PanoCam and SpheroCam HDR, are in the family of SpheronVR cameras for capturing 360º panoramas in one scanning cycle. Compared with PanoCam that is used to produce cylindrical images, SpheroCam captures spherical images with a much larger vertical field of view. Figure 1 shows a SpheroCam. Essentially, it is a rotating line spherical camera that consists of a line CCD with an AF Fisheye-Nikkor 16mm f/2.8D lens and an integrated DC motor at its bottom, where it is connected to a tripod. The DC motor rotates itself 360º in the horizontal direction in one scan to enable all surrounding scene to be captured by the line CCD camera except the area underneath the tripod. The camera provides a vertical resolution up to 5300 pixels and horizontal resolution up to 10600 pixels (with fish-eye lens). It takes about 8 minutes to scan an image with maximum resolution, and less than one minute to capture an image with minimum resolution of 640 x 1571 pixels.

Figure 1: SpheroCam

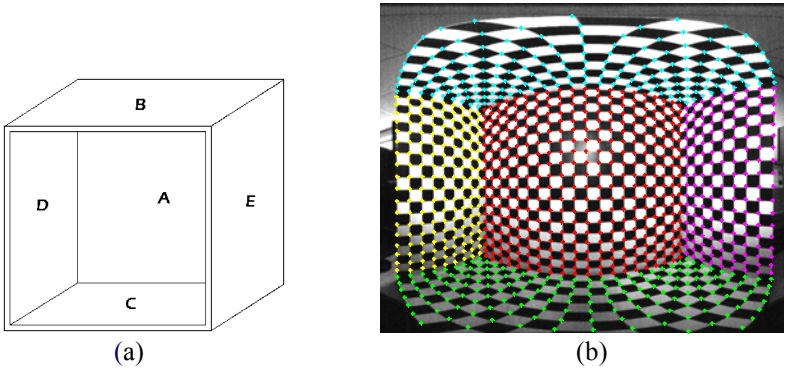
An important camera parameter of the SpheroCam is the vertical field of view that is commonly approximated to 180º. With the Nikon full-frame fish-eye lens [10] being designed to provide 180º field of view in the diagonal direction, the actual vertical field of view can be significantly less than 180º. With the SpheroCam being placed at the centre of the sphere as illustrated in Figure 2 , Figure 2(a) shows the actual field of view of the camera, where the cone area underneath the bottom of the camera is out of view, and Figure 2(b) shows the corresponding rectangle full-frame image taken by this type of fish-eye lens. In the proposed camera calibration method, the vertical field of view is taken into account as an additional camera parameter, and is denoted as $(180^{o} - \theta)$, where θ is the angle between any generatrix line of the cone and the z axis as shown in Figure 2(a).



(a) (b)

Figure 2: (a) Camera field of view, and (b) Typical spherical image

As illustrated in Figure 2(b), the direct output of the SpheroCam is a flat spherical image with object features distorted. By geometrical transformation of each pixel from its image coordinates to the corresponding spherical coordinates, the flat image can be mapped to the inner surface of the sphere to enable correct visualization of the actual scene at the centre of the sphere without image distortion. The first and last column of the image should be correctly aligned if the camera was correctly leveled during image acquisition. If this is not the case, it is caused by a tumbling error due to an incorrect set up of the camera tripod. Although the tumbling effect can be corrected through calibration [4], it was avoided in the camera set up stage by using the camera leveling indicators.



(a) (b)

Figure 3: (a) Calibration box, and (b) Hemisphere calibration image

To perform camera calibration based on the image acquired from a grid pattern on multiple planes, a special calibration object based on a five-sided open box was constructed. As shown in Figure 3(a), the calibration box has an inner physical size of 57 x 57 x 27cm (height x width x depth) and its insider faces are covered by a checkerboard pattern with a square size of 3 x 3cm. In image acquisition set up, the camera is placed around the middle of the calibration box facing directly towards the front panel denoted by A in Figure 3(a), with the panels above and below the camera denoted by B and C, and the panels on the left and right sides of the camera denoted by D and E. Furthermore, the camera head is dipped into the calibration box (as near to panel A as possible) to ensure a hemisphere coverage by the checkerboard pattern. During image acquisition, the horizontal field of view for the SpheroCam is set to 180^{o} (half of one scanning cycle) to capture only half of its surrounding. Figure 3(b) shows an example of the hemisphere calibration image acquired (640 x 785 pixels), where corner points of the checkerboard squares in five different sides of the calibration box are highlighted using different colours. It should be apparent that the proposed calibration method is based on a reasonable assumption of symmetrical hemisphere.

3 Calibration Methodology

The geometry for the proposed camera calibration method is illustrated in Figure 4, where $\vec{P}_w = [x_w, y_w, z_w]^T$ denotes a physical point corresponding to a corner of the checkerboard pattern on a side of the calibration box in the world coordinate system (formed by X_w, Y_w and Z_w with the origin at the upper left corner of the calibration box), $\vec{P}_c = [x_c, y_c, z_c]^T$ denotes the coordinates of the projection of $\vec{p}_w$ on the spherical retina in the camera coordinate system (formed by X_c, Y_c and Z_c with the origin at the camera position) and $P_i = [x_i, y_i]^T$ denotes the coordinates of $\vec{p}_w$ appeared in the hemisphere calibration image acquired.

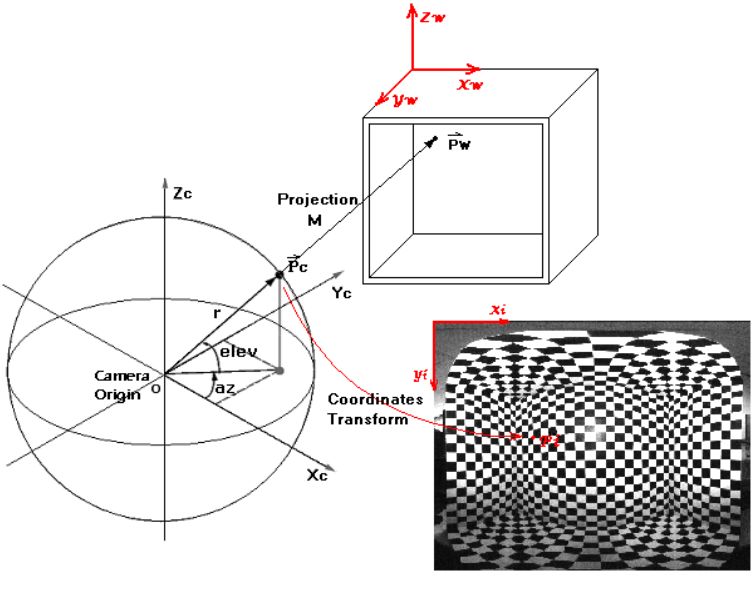


Figure 4: Camera calibration geometry

Since each latitude circle and each longitude line of the sphere correspond to a horizontal and vertical line on the image, a pixel position denoted by $P_i = [x_i , y_i]^T$ in the image coordinate system can be expressed using the corresponding azimuth and elevation angles in the spherical coordinate system, namely, $\bar{P}_c = [az , elev , r]^T$, where az, $elev$ denote azimuth and elevation angles respectively, and $r = 16mm$ is the fish-eye lens focal length. If the image dimension is m by n pixels, then

$$az = \frac{2\pi(x_i - 1)}{(n-1)} - \pi \tag{1}$$

where $az \in [-\pi , \pi]$, and

$$elev = \frac{(m - y_i)(\pi - \theta)}{(m-1)} - \left(\frac{\pi}{2} - \theta\right) \tag{2}$$

where $elev \in \left[-\left(\frac{\pi}{2} - \theta\right), \frac{\pi}{2}\right]$, and $(\pi - \theta)$ is the camera's vertical field of view with θ being a camera parameter to be determined. For a given value θ, expressing the spherical coordinates $\bar{P}_c = [az , elev , r]^T$ in the corresponding Cartesian coordinates $\bar{P}_c = [x_c , y_c , z_c]^T$ gives

$$x_c = r \cdot \cos(elev) \cdot \cos(az) \tag{3}$$

$$y_c = r \cdot \cos(elev) \cdot \sin(az) \tag{4}$$

$$z_c = r \cdot \sin(elev) \tag{5}$$

With $\bar{P}_c = [x_c , y_c , z_c]^T$ being the projection of $\bar{P}_w = [x_w , y_w , z_w]^T$, it gives

$$\begin{bmatrix} x_c \\ y_c \\ z_c \end{bmatrix} = M \cdot \begin{bmatrix} x_w \\ y_w \\ z_w \\ 1 \end{bmatrix} \tag{6}$$

where M denotes the camera projection matrix with a size of 3 x 4 [11] and it can be decomposed into an intrinsic matrix denoted by A, a rotation matrix denoted by R, and a translation vector denoted by T, that is, $M = A \cdot [R \;\; T]$ [12]. Given a set of the image coordinates of the checkerboard square corner points extracted from the hemisphere calibration image P_i, their correspondences on the camera's spherical retina $\bar{P}_c$ can be calculated by using equations (1-5) for a given vertical field of view $(\pi - \theta)$. Using $\bar{P}_c$ obtained from P_i and their corresponding physical positions of the checkerboard square corners on the calibration box $\bar{P}_w$, M can be estimated by applying equation (6) to sufficient pairs of corresponding points (at least 8 pairs) to form a set of simultaneous equations and using singular value decomposition (SVD) to solve the resulting equation system for camera parameters contained in M [13].

To estimate θ, its value is increased from minimum to maximum expected angles in equation (2) to yield different $elev$ values for each given y_i (the increase can be made in a small angular interval in order to give a sufficient angular resolution). Since different $elev$ values will result in different sets of $\bar{P}_c$ being generated from equations (3-5), a number of different camera projection matrices will be generated for M. By applying each camera projection matrix estimated and the corresponding θ value used in estimation, the physical positions of the checkerboard square corners are reprojected onto the image plane. If the reprojected image points are denoted by $P_i' = [x_i' , y_i']^T$, then the root-mean-square (rms) error

between the reprojected image points and the corner points appeared in the hemisphere calibration image P_i is given by

$$RMS\ Error = \frac{1}{n}\sum_{i=1}^{n}\sqrt{(x_i - x_i')^2 + (y_i - y_i')^2} \qquad (7)$$

where n denotes the total number of the checkerboard square corner points. With the rms error varying for different θ values, it forms the cost function of the iterative search for the minimum rms error with the corresponding θ value and the camera projection matrix identified to represent the actual camera parameters.

4 Accuracy Assessment with Experimental Results

Presented in this section are some experimental results to demonstrate the achievable accuracy as well as the effect on the accuracy due to the numbers of control points and planes used.

As an example, Figure 5 shows the characteristics of the rms error as a function of θ for determination of the vertical field of view. It is generated based on the all the corner points available in all 5 planes using the image shown in Figure 3(b) with θ increasing from 5° to 25º in a step of 0.1º. The minimum rms error of 1.2155 pixels is seen to occur at θ = 16.5°, resulting in the vertical field of view given by 180º - 16.5º = 163.5º.

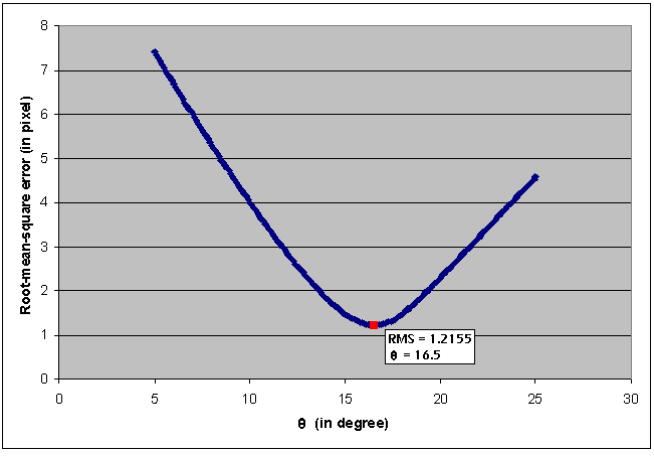


Figure 5: rms error curve for determination of vertical field of view

Based on the image shown in Figure 3(b), the first investigation on calibration accuracy was based on all checkerboard square corner points available in different plane combinations. The use of corner points in any one single plane for calibration was found to yield high rms error (between 200 and 400 pixels). Table 1 lists the results obtained for non-coplanar corners points, and the lowest rms error of 1.2155 pixels with θ = 16.5° is seen to be generated by all the available corner points in all the planes. In terms of the number of planes used for calibration, an increase in the number of planes, resulting in a higher number of corner points involved in the calibration process, is seen to reduce the variation of rms errors in each group. In terms of the number of control points used for calibration within each

group, an increase in the number of control points by using the large front plane does not lead to a decrease in the rms error. For the group using all available corner points in 2 planes, the lowest and the highest rms errors (1.2837 pixels and 2.0404 pixels) are seen to be generated by using similar numbers of corner points, with the combination of two side planes (D and E planes) involving the lowest number of corner points to produce the lowest rms error. The same outcome is also seen from the group using all available corner points in 3 planes. The lowest and the highest rms errors (1.2159 and 1.4034 pixels) are again generated by using similar number of corner points, with the combination of two side plane plus the bottom plane (D, E and C planes) involving the lowest number of corner points to produce the lowest rms error. Although the same could not be said for the group using all available corner points in 4 planes, the difference between the rms errors in this group is much smaller (less than 0.0907 pixels).

No. of planes	Plane combin-ation	No. of corner points	Error in pixel	θ in degree
2	AB	580	1.5847	15.3
	AC	570	1.3102	17
	AD	562	1.3163	16.7
	AE	562	1.412	16.4
	BC	350	1.5055	16.1
	BD	342	1.7176	16.5
	BE	342	2.0404	15.8
	CD	332	1.8567	16.5
	CE	332	1.8958	16.2
	DE	324	1.2837	16.4
3	ABC	750	1.2272	16.4
	ABD	742	1.3113	16.2
	ABE	742	1.359	16.3
	ACD	732	1.2833	16.8
	ACE	732	1.286	16.6
	ADE	724	1.2666	16.5
	BCD	512	1.4034	16.1
	BCE	512	1.3852	16.2
	BDE	504	1.3771	16.1
	CDE	494	1.2159	16.6
4	ABCD	912	1.222	16.5
	ABCE	912	1.2192	16.5
	ABDE	904	1.2854	16.4
	ACDE	894	1.2222	16.6
	BCDE	674	1.3099	16.4
5	ABCDE	1074	1.2155	16.5

Table 1: Calibration accuracy produced by all available corner points in different plane combinations

The second investigation on calibration accuracy was based on half of the corner points available in different plane combinations, and Table 2 lists the results obtained based on 10

sets of corner points randomly selected from each plane. Comparing Table 2 with Table 1, there are various similarities between them which include the combination of planes to yield the lowest and highest average errors for the groups using corner points in 2 and 3 planes. The lowest average errors in these two groups are again seen to be produced with the least number of corner points without using the front plane (A plane). The trend in error variation is similar between the two investigations. The maximum error is seen to be just over 2 pixels based on corner points selected from 2 planes and is reduced to well below 2 pixels by using corner points in 3 or more planes.

No. of planes	Plane combin-ation	No. of corner points	Averaged error in pixel	θ in degree	Max error in pixel	Standard deviation
2	AB	290	1.57343	15.42	1.7072	0.055
	AC	285	1.34333	16.97	1.3891	0.030
	AD	281	1.37834	16.68	1.4663	0.050
	AE	281	1.42485	16.35	1.5939	0.095
	BC	175	1.53478	16.13	1.5789	0.032
	BD	171	1.7024	16.48	1.8651	0.091
	BE	171	1.97444	15.77	2.1044	0.100
	CD	166	1.89414	16.48	2.1264	0.156
	CE	166	1.96923	16.47	2.1062	0.106
	DE	162	1.33643	16.5	1.389	0.043
3	ABC	375	1.25144	16.49	1.2875	0.022
	ABD	371	1.33164	16.24	1.3865	0.029
	ABE	371	1.37083	16.31	1.417	0.026
	ACD	366	1.29771	16.84	1.3346	0.017
	ACE	366	1.31731	16.57	1.3516	0.022
	ADE	362	1.26916	16.51	1.308	0.017
	BCD	256	1.39906	16.19	1.4559	0.032
	BCE	256	1.39379	16.18	1.4424	0.024
	BDE	252	1.39087	16.14	1.435	0.034
	CDE	247	1.22742	16.57	1.2588	0.012
4	ABCD	456	1.22658	16.53	1.2556	0.013
	ABCE	456	1.23557	16.58	1.2539	0.012
	ABDE	452	1.29772	16.46	1.3234	0.017
	ACDE	447	1.24246	16.63	1.2666	0.015
	BCDE	337	1.31627	16.43	1.3412	0.018
5	ABCDE	537	1.22322	16.54	1.2499	0.011

Table 2: Calibration accuracy produced by half of corner points in different plane combinations

With the error reduced by using corner points in increasing number of planes, the third investigation on calibration accuracy was based on the number of corner points from all five

planes, and Figure 6 shows the results obtained based on 10 sets of randomly selected corner points. Although the average, maximum and minimum errors as well as the standard deviation are seen to increase as the number of the corner points reduces, the proposed method is seen to be able to provide a calibration error well within 2 pixels by using 18 corner points or more.

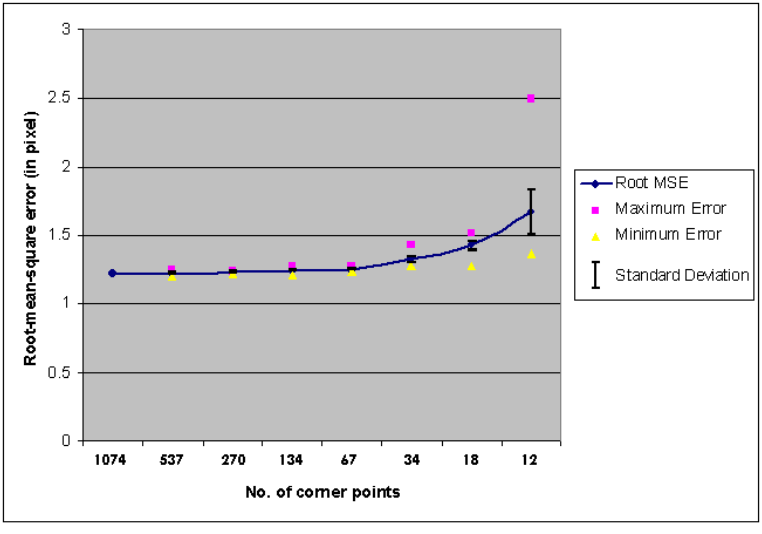


Figure 6: Calibration accuracy versus the number of corner points in five planes

5 Conclusions

In this paper, a simple method for calibration of a rotating line spherical camera has been presented. It is based on a single hemisphere image acquired from a five-sided open box with the insider faces covered by a checkerboard pattern. By include the vertical field of view as an additional camera parameter, the proposed method is seen to provide good calibration accuracy and robustness without requiring a large number of corner points as shown by various experiments. In particular, the influence of the corner points selected from different combination of planes on calibration accuracy has been shown, the achievable accuracy is seen to be very near to 1 pixel, and stable results are seen to be obtained consistently with an error bound of well within 2 pixels by using corner points selected from at least 3 different planes.

References

[1] T. Svoboda, T. Pajdla, Epipolar Geometry for Central Catadioptric Cameras, International Journal of Computer Vision 49(1), pp. 23-37, 2002

[2] J. Kannala, S. S. Brandt, A Generic Camera Model and Calibration Method for Conventional, Wide-Angel, and Fish-Eye Lenses, IEEE Transaction on Pattern Analysis and Machine Intelligence, Vol. 28, No. 8, August 2006

[3] D. Bradly, A. Brunton, M. Fiala, G. Roth, Image-based Navigation in Real Environment Using Panoramas, HAVE 2005 - IEEE International Workshop on Haptic Audio Visual Environment and their Applications, Ottawa, Ontario, Canada, 1-2 October 2005

[4] J. A. Parian, A. Gruen, Panoramic Camera Calibration Using 3D Straight Lines, International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. XXXVI, part 5/W8, ISPRS "Panoramic Photogrammetry Workshop", Berlin, Germany, 24-25 February 2005

[5] D. Schneider, H.-G. Maas, Geometric Modelling and Calibration of a High Resolution Panoramic Camera, Optical 3-D Measurement Techniques VI, Vol. II, pp. 122-129, Institute of Geodesy and Photogrammetry, ETH Zurich, 2003

[6] T. Hirota, H. Nagahara, M. Yachida, Calibration of Rotating Line Camera for Spherical Imaging, ACCV 2006, LNCS 3851, 7th Asian Conference on Computer Vision, pp. 389-398, 2006

[7] J. Fujiki, A. Torii, S. Akaho, Epipolar Geometry Via Rectification of Spherical Images, MIRAGE 2007, LNCS 4418, pp. 461-471, 2007

[8] H Bakstein, T Pajdla, Panoramic Mosaicing with a 180^{o} Field of View Lens, *Proc.* IEEE Workshop Omnidirectional Vision, pp. 60-67, 2002

[9] D. Claus, A. W. Fitzgibbon, A Rational Function Lens Distortion Model for General Cameras, Proc. of the IEEE Conference on Computer Vision and Pattern Recognition, Vol. 1, pp. 213-219, June 2005

[10] AF Fisheye-Nikkor 16mm f/2.8D Instruction Manual, Nikon Corporation, Japan

[11] R. Hartley, A. Zisserman, Multiple View Geometry in Computer Vision, 2nd Ed. Cambridge: CUP, 2003

[12] A Fusiello, E. Trucco, A. Verri, A Compact Algorithm for Rectification of Stereo Pairs, Machine Vision and Applications, Vol. 12, pp. 16-22, 2000

[13] E Trucco, A. Verri, Introductory Techniques for 3-D Computer Vision, Upper Saddle River, New Jersey: Prentice-Hall, 1998