A Calibration Method for an Omnidirectional Multi-Camera System

Sei Ikeda, Tomokazu Sato and Naokazu Yokoya

Graduate School of Information Science, Nara Institute of Science and Technology,
8916–5 Takayama, Ikoma, Nara, 630-0101 Japan

ABSTRACT

Telepresence systems using an omnidirectional image sensor enable us to experience remote site. An omnidirectional multi-camera system is more useful to acquire outdoor scenes than a monocular camera system, because the multi-camera system can easily capture high-resolution omnidirectional images. However, exact calibration of the camera system is necessary to virtualize the real world accurately. In this paper, we describe a geometric and photometric camera calibration and a panorama movie generation method for the omnidirectional multi-camera system. In the geometric calibration, intrinsic and extrinsic parameters of each camera are estimated using a calibration board and a laser measurement system called total station. In the photometric calibration, the limb darkening and color balances among the cameras are corrected. The result of the calibration is used in the panorama movie generation. In experiments, we have actually calibrated the multi-camera system and have generated spherical panorama movies by using the estimated camera parameters. A telepresence system was prototyped in order to confirm that the panorama movie can be used for telepresence well. In addition, we have evaluated the discontinuity in generated panoramic images.

Keywords: Omnidirectional Multi-camera system, Camera Calibration, Panorama Movie Generation

1. INTRODUCTION

Telepresence systems using an omnidirectional image sensor enable us to experience remote site. These systems are expected to be used in the fields of education and entertainment. A curved mirror[1,2] and a fish eye lens[3,4] have been widely used for these telepresence systems because they can be attached to a standard camera. However, these methods have problems of shortage and position-dependency of image resolution, because these sensors capture an omnidirectional scene by a single camera.

On the other hand, an omnidirectional multi-camera system can obtain a high-resolution image with uniform resolution, although exact camera calibration is necessary to generate a panoramic image from multiple images accurately. In related work, some calibration methods for omnidirectional multi-camera systems are proposed, which are based on using markers of known 3D positions.[5] In these methods, grid points of iterative patterns are used for markers and are projected on a large-sized immersive projection screen. However, it is difficult to obtain highly accurate calibration using such approaches, because the distribution of the markers is limited on a projection screen. Further more, there is no discussion about the precision of generated panoramic images.

In this paper, we propose a camera calibration and panorama movie generation method for the omnidirectional multi-camera system "Ladybug". Figure 1 shows the multi-camera system Ladybug made by Point Grey Research Inc. The camera unit consists of six cameras (Figure 1 (left)): Five configured in a horizontal ring and one pointing vertically. Figure 1 (right) shows a storage unit, which consists of four hard disks. The camera system can collect movies covering more than 75% of the full spherical view with almost the same apparent point of view.

This paper is constructed as follows. In Section 2, we describe a method of geometric and photometric calibration for the omnidirectional multi-camera system Ladybug. Section 3 describes a method of generating

Further author information: (Send correspondence to S.I.)
S. I.: E-mail: sei-i@is.aist-nara.ac.jp, Telephone: +81(0)743 72 5296
T. S.: E-mail: tomoka-s@is.aist-nara.ac.jp, Telephone: +81(0)743 72 5296
N. Y.: E-mail: yokoya@is.aist-nara.ac.jp, Telephone: +81(0)743 72 5290

Figure 1. Camera head (left) and storage unit (right) of Ladybug.

a spherical panorama movie using the calibration result. The effect of disparity among multiple cameras is also discussed. In Section 4, the Ladybug camera system is actually calibrated geometrically and photometrically by using the proposed method. The spherical panorama movie composed of six input images is also computed. A telepresence system is also prototyped in order to confirm that the panorama movie can be used for telepresence well. In addition, we evaluate the discontinuity in computed panoramic images.

2. CAMERA CALIBRATION FOR OMNIDIRECTIONAL MULTI-CAMERA SYSTEM

This section describes a method of geometric and photometric calibration for an omnidirectional multi-camera system. In the geometric calibration, intrinsic and extrinsic parameters of each camera are estimated using a calibration board and a laser measurement system called total station. In the photometric calibration, limb darkening and color balances among multiple cameras are corrected.

2.1. Geometric Calibration

In the geometric calibration, intrinsic parameters (focal length f, lens distortion parameters (κ_1, κ_2), center of distortion (C_x, C_y), aspect ratio s_x) and extrinsic parameters (position $\boldsymbol{T_c}$, orientation $\boldsymbol{R_c}$) of each camera c should be estimated. The extrinsic parameters are estimated in a unified coordinate system to obtain the relationship among six cameras.

In camera calibration, spatial arrangement of many markers of known 3D position is required to estimate camera parameters accurately. First, arrangement of markers are described. In our method, grid points of a lattice pattern on a calibration board are used as markers. 3D coordinates of four corners of the calibration board are measured by the total station and all 3D positions of the markers on the board are calculated by linear interpolation among its corners. The calibration board is arranged and measured at several different positions for acquiring a large number of makers' 3D positions and 2D positions on images as shown in Figure 2. In our method, the total station must be fixed on the same position while the whole geometric calibration procedure is performed. Consequently, 3D positions of all markers captured by all the cameras can be represented in a single coordinate system made by the total station. Grid points on images are detected at sub-pixel resolution by calculating intersection of two lines as shown Figure 3.

Next, the camera parameters of each camera are estimated. The intrinsic parameters are estimated by Tsai's method.[6] In Tsai's method, the undistorted image coordinates (X_u, Y_u) of markers are transformed to the distorted coordinates (X_d, Y_d) by using following equations:

$$X_u = s_x(X_d - C_x)(1 + \kappa_1 r^2 + \kappa_2 r^4), \tag{1}$$

$$Y_u = (Y_d - C_y)(1 + \kappa_1 r^2 + \kappa_2 r^4), \tag{2}$$

$$r = \sqrt{X_d^2 + Y_d^2}. \tag{3}$$

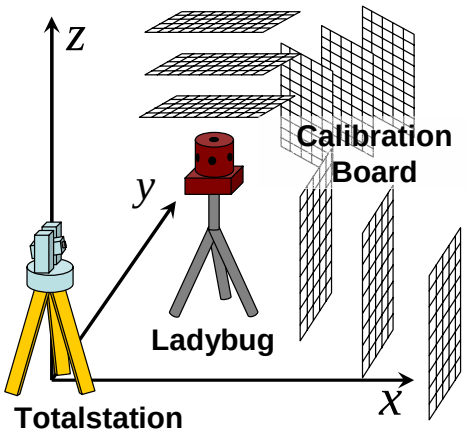


Figure 2. Arrangement of calibration board in space.

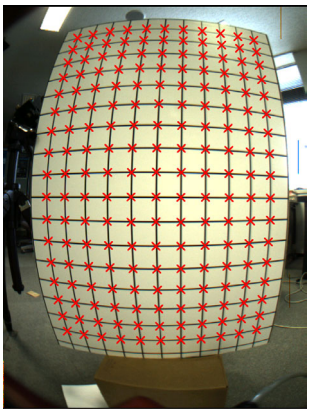


Figure 3. Detected grid points.

Note that the distortion parameter κ_2 is also considered because input images from the cameras are highly distorted. The extrinsic parameters of each camera are estimated by minimizing the sum of re-projection errors E_c of makers.[7] The re-projection error of a marker m is defined as a squared distance on the image between the projected position $\boldsymbol{v_m}$ of measured 3D position and the detected 2D position $\boldsymbol{u_m}$ of the marker m. The measure E_c is computed by:

$$E_c = \sum_m (\boldsymbol{u_m} - \boldsymbol{v_m})^2. \tag{4}$$

In this paper, extrinsic parameters of each camera are represented as a world-to-camera transformation matrix $\boldsymbol{M_c}$ using position $\boldsymbol{T_c}(t_1, t_2, t_3)$ and orientation $\boldsymbol{R_c}(r_1, r_2, r_3)$ of each camera c ($c = 1, 2, ..., 6$) on the world coordinate system:

$$\boldsymbol{M_c} = \begin{bmatrix} \boldsymbol{R_c} & \boldsymbol{T_c} \\ 0 & 1 \end{bmatrix}, \tag{5}$$

$$= \begin{bmatrix} c_1c_3 + s_1s_2s_3 & s_1c_2 & -c_1s_3 + s_1s_2c_3 & t_1 \\ -s_1c_3 + c_1s_2s_3 & c_1c_2 & s_1s_3 + c_1s_2c_3 & t_2 \\ c_2s_3 & -s_2 & c_2c_3 & t_3 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \tag{6}$$

where,

$$\begin{array}{lll} s_1 = \sin r_1, & s_2 = \sin r_2, & s_3 = \sin r_3, \\ c_1 = \cos r_1, & c_2 = \cos r_2, & c_3 = \cos r_3. \end{array} \tag{7}$$

Estimating $\boldsymbol{M_c}$ for six parameters $(t_1, t_2, t_3, r_1, r_2, r_3)$ is a non-liner minimization problem. Since there exist problems concerning calculation cost and local minima, we first estimate $\boldsymbol{M'_c}$ for twelve parameters by minimizing E_c linearly. Then, the estimated camera parameter $\boldsymbol{M'_c}$ is linearly adjusted to reduce the degree of freedom to six by assuming that the direction of optical axis is correctly estimated. Finally, the sum of re-projection error E_c is minimized by a gradient method for optimizing $\boldsymbol{M_c}$.

2.2. Photometric Calibration

In the photometric calibration, first the limb darkening and color balances are corrected. First, the correction of the limb darkening is described. The limb darkening is a gradually decreasing effect of brightness in peripheral regions in images. Generally, images from an omnidirectional multi-camera system suffer from the limb darkening effect problem because wide-angle lenses are equipped. The $\cos^4\theta$ phenomenon[8] and the vignetting[9] are well known as limb darkening effects. It is known that the $\cos^4\theta$ phenomenon is proper to a wide-angle lens and the

vignetting is proper to a zoom lens. Thus we only treat the $\cos^4\theta$ phenomenon and correct image intensities using the following equation.

$$I' = \frac{\pi l^2 \cos^4\theta}{f^2} I, \tag{8}$$

where l is a lens diameter, f is a focal length, I is a radiance of surface of an object in the direction toward the lens, I' is an irradiance of the image and θ is an angle between the incident light and the optical axis.

Next, we describe a method for correcting color balances among cameras. Generally it is known that the irradiance of the image I' and the radiance of the surface I have a linear relation shown as $I' = aI + b$. The color balances are different due to the difference in color transform parameters a and b among cameras. In this paper, the color balance between two cameras c and c' is adjusted by estimating parameters a_c and b_c in the linear transformation $I_{c'} = a_c I_c + b_c$. The parameters a_c and b_c are determined by equalizing the normalized RGB histograms $h_c(i)$ and $h_{c'}(i)$ where i is intensity of an image pixel. The following error function $e(a_c, b_c)$ is minimized for equalizing the histograms $h_c(i)$ and $h_{c'}(i)$:

$$e(a_c, b_c) = \sum_i \left\{ h_{c'}(i) - \frac{1}{a_c} h_c\left(\frac{i - b_c}{a_c}\right) \right\}^2. \tag{9}$$

3. GENERATING OMNIDIRECTIONAL PANORAMA MOVIE

This section describes a method of generating omnidirectional panoramic images. The method is based on re-projecting calibrated input images on a spherical virtual image surface. We also discuss the effect of disparity between multiple cameras which do not satisfy the single viewpoint constraint.

3.1. Generation of a Panorama Movie

In generation of a panoramic image, limb darkening and the color balance of input images are corrected as a pretreatment, as shown in Figure 4. Then, corrected images are projected on a spherical projection surface by using the intrinsic and the extrinsic parameters of the geometric calibration result. In this section we describe a projection method of the corrected images.

Since the centers of projection of multiple cameras of the omnidirectional multi-camera system are different from each other, the single viewpoint perspective projection model is not applicable for this system. However, when the distance of a target from the system is sufficiently large, the centers of projection can be considered as the same. Therefore, we assume that the target scene is far enough from the system and set the spherical projection surface S far enough from the camera system. The center of the spherical projection surface is set at a center of gravity of the centers of projection of all the cameras. A panoramic image is generated by projecting all the pixels of all the images onto the spherical surface S. Note that a blending method is used for generating a smooth panoramic image, when a point s on the spherical surface S is projected from image surfaces of different cameras. The intensity $I_S(s)$ of the point s is determined by the following equation.

$$I_S(s) = \frac{\sum_{c \in \boldsymbol{C}(s)} \gamma_c I_c(u_c)}{\sum_{c \in \boldsymbol{C}(s)} \gamma_c}, \tag{10}$$

where u_c is the point in the image of the camera c which is projected on the point s in the spherical surface S, γ is the distance between u_c and border of the image c, and $\boldsymbol{C}(s)$ is a set of cameras to which the point s in the S can be projected.

3.2. Effect of Disparity

We discuss the effect of disparity in camera position. As shown in Figure 5, when the target x is captured at u_c and $u_{c'}$ on the images of the two cameras c and c', they are projected to different position s_c and $s_{c'}$ on the spherical surface S in the proposed method. Since s_c and s'_c are not consistent, a panoramic image doubly shifted is generated. If the circumference of the spherical surface S is composed of N pixels, the effect of disparity is limited within one pixel if the condition $\angle s_c G s_{c'} < \frac{2\pi}{N}$ is satisfied. When the radius of the spherical surface S is

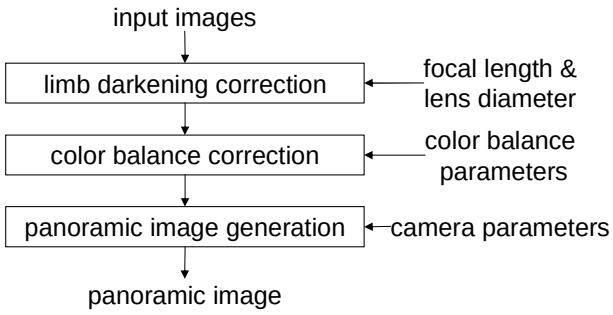


Figure 4. Flow diagram of generating a panoramic image.

assumed to be infinite, $\angle s_c G s_{c'}$ is approximated by $\angle s_c x s_{c'} < \frac{2\pi}{N}$. Consequently, the condition $\angle s_c G s_{c'} < \frac{2\pi}{N}$ is approximated by $\angle T_c x T_{c'} < \frac{2\pi}{N}$.

When we assume that the distance between the two cameras is d and the distance from the each camera to the target x is the same, the condition $\angle T_c x T_{c'} < \frac{2\pi}{N}$ is simplified that the distance λ of the target x from the center point of the two camera is represented as follows:

$$\lambda > \frac{d}{2 \tan \frac{\pi}{N}}. \tag{11}$$

4. EXPERIMENT

We have calibrated Ladybug system and have generated a panorama movie. This section describes some experimental results concerning geometric and photometric calibration, panorama movie generation, and telepresence application.

4.1. Calibration of Ladybug

In the geometric calibration, the calibration board was captured by each camera of Ladybug fixed on a tripod. The calibration board was arranged and captured at different depth from fixed cameras three times. For each of five horizontal cameras, 187 grid points on the calibration board were used, then totally 561 set of 3D position and 2D position on captured images were obtained. For a vertical camera, 170 grid points on the calibration board were used, then totally 510 set of 3D position and 2D position on captured images were obtained. The 3D position was measured by a total station LEICA TCR1105 XR.

Figure 6 shows the result of reducing distortion of an input image using the estimated intrinsic parameters. We can confirm that the intrinsic parameters are fairly good. Figure 7 shows the result of correcting the limb darkening. The left is an input image. The right is a corrected image for the shortage of brightness in peripheral regions of the input image. Figure 8 shows a part of panoramic images before and after the color balance adjusting. In the right image, the boundary lines do not stand out compared with the non-corrected image in the left.

4.2. Generation of a Panorama Movie

A panorama movie was generated by using the result of calibration described in the previous section. The input movie was obtained by setting Ladybug put on a moving car as shown in Figure 9. Figure 10 shows an example set of input images (resolution: 768×1024). A panoramic image (resolution: 3840×1920) generated from these input images is also shown in Figure 11. This panoramic image is unfolded by polar coordinates from the spherical surface. A black part at the bottom of the panoramic image is the part of no input images. Since any boundary lines are hardly found in Figure 11, we can confirm that geometrical and photometrical calibration is successfully achieved.

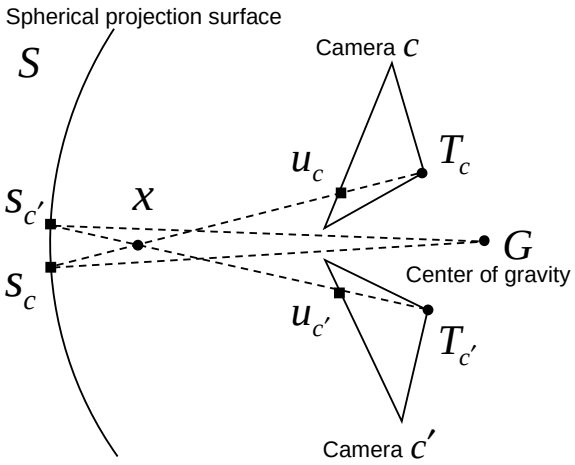
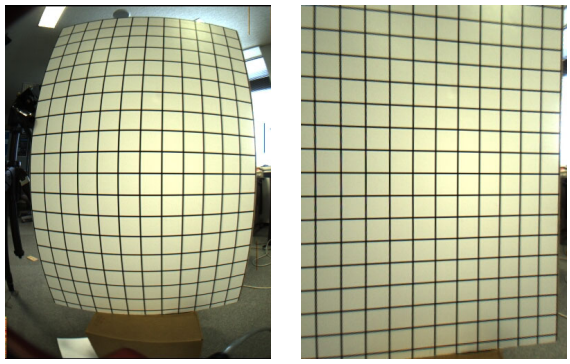


Figure 5. Effect of disparity.



(a) input image (b) corrected image

Figure 6. Lens distortion correction.

(a) input image (b) corrected image

Figure 7. Limb darkening correction.

(a) non-corrected image (b) corrected image

Figure 8. Color adjustment.

Figure 12 gives a panoramic image generated without the blend method blending for comparing it with the blended one. We can find some boundaries in the non-blended panoramic image; close objects and sky region containing saturated pixels due to direct rays of sun. On the other hand, we can not perceive any boundary lines in the blended panoramic image except very close objects.

4.3. Prototype Telepresence System

A telepresence system was prototyped in order to confirm that the panorama movie can be used for telepresence application well. The system is composed of a spherical immersive display, a controller and a personal computer as shown in Table 1. Our system enables users to interactively view an omnidirectional view field scene projecting a specified part of a panorama movie in real time. The system performance of generating images is 20fps, which is faster than the performance 15fps of image capturing by Ladybug. Due to the resolution limitation 1024×768 of the immersive display, the resolution of a panorama movie was resized to 2048×1024 in the experiment. Figure 13 shows user's appearance in telepresence experiment. We have confirmed that our telepresence system exhibits high-presence in respect of resolution, field of view, and interaction.

4.4. Quantitative Evaluation

We have evaluated errors in panoramic image generation. The error is defined as an angle between two lines connecting the spherical projection center G and the two corresponding projected points of the same point in space on the spherical projection surface.

Figure 9. Ladybug mounted on a car.

Figure 10. Input image 5: upper camera (right lower) and horizontal camera (others).

There are common regions in input images of adjacent cameras of Ladybug as shown in Figure 14. A circle marker is captured in these regions to specify the corresponding points of adjacent two camera images. The center of gravity of this marker is calculated with sub-pixel accuracy. In this experiment, the angle $\angle s_c G s_{c'}$ shown in Figure 5 was evaluated as error ϵ, where u_c and $u_{c'}$ are centers of gravity of these markers. Because the distance decreasing the effect of disparity within one pixel on the panoramic image is about 24m under assumptions of $d = 4$cm and $N = 3840$, the marker was located about 30m away from Ladybug. More than one hundred corresponding points were obtained for each camera.

Table 2 shows maximum and average values of error ϵ in generation a panoramic image. The number 6

Figure 11. One frame of panorama movie (blending).

Figure 12. One frame of panorama movie (no blending).

Table 1. Components of telepresence system.

Spherical Display	Elumens VisionStaion
Controller	Microsoft SideWinder Game Pad Pro
PC	CPU:Intel Pentium4 1.7GHz, RAM:1GB
Graphics Card	Nvidia Geforce4

represents the vertical camera and the others represent horizontal cameras. The average value of the error ϵ is 0.0063rad which is approximately equivalent to 3 pixels on an input image. We have also found that points having large errors tend to distribute at the borders of input images.

5. CONCLUSION

In this paper, we have described a method of generating a high-resolution spherical panorama movie based on the result of geometric and photometric calibration for an omnidirectional multi-camera system Ladybug. In the geometric calibration, a large number of markers were virtually arranged around the system by using a calibration board and a total station for improving the accuracy of calibration. We have also prototyped a

Table 2. Error ϵ of a generated panoramic image [rad].

Camera Number	1-2	2-3	3-4	4-5	5-1
Number of the data	227	258	229	216	305
Maximum value	0.0214	0.0378	0.0258	0.0227	0.0198
Average value	0.0072	0.0057	0.0054	0.0048	0.0041

Camera Number	6-1	6-2	6-3	6-4	6-5
Number of the data	176	131	127	123	154
Maximum value	0.0178	0.0498	0.0208	0.0107	0.0112
Average value	0.0058	0.0144	0.0095	0.0063	0.0052

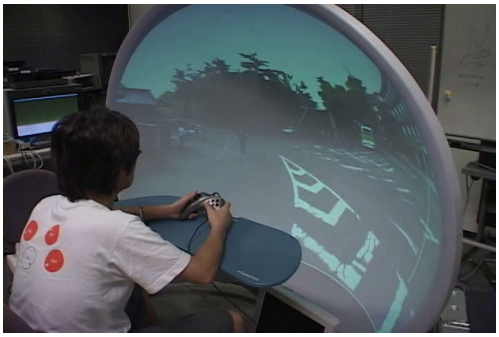


Figure 13. Appearance of telepresence system.

Figure 14. Common regions in input images.

telepresence system having omnidirectional view of field. Experiments have exhibited that panorama movies can be used for telepresence application. Finally, projected errors generating a panoramic image was evaluated. We have found that the average of the error is about 3 pixels on an input image. In future work, a better lens distortion model for wide-angle lens will be investigated for obtaining higher accuracy in generation of panorama movie.

REFERENCES

1. Y. Onoe, K. Yamazawa, H. Takemura, and N. Yokoya, "Telepresence by Real-time View-dependent Image Generation from Omnidirectional Video Streams," *Computer Vision and Image Understanding* **71**(2), pp. 154–165, 1998.
2. M. Fiala and A. Basu, "Line Segment Extraction in Panoramic Images," *Journal of Winter School of Computer Graphics* **10**(1), pp. 179–186, 2002.
3. H. Bakstein and T. Pajdla, "Panoramic Mosaicing with a 180 Field of View Lens," *Proc. IEEE Workshop on Omnidirectional Vision* , pp. 60–67, 2002.
4. K. Miyamoto, "Fish Eye Lens," *Journal of Optical Society of America* **54**(8), pp. 1060–1061, 1964.
5. J. Shimamura, H. Takemura, N. Yokoya, and K. Yamazawa, "Construction of an Immersive Mixed Environment Using an Omnidirectional Stereo Image Sensor," *Proc. IEEE Workshop on Omnidirectional Vision* , pp. 62–69, 2000.
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 Journal of Robotics and Automation* **RA-3**(4), pp. 323–344, 1987.
7. T. Sato, M. Kanbara, N. Yokoya, and H. Takemura, "Dense 3D Reconstruction of an Outdoor Scene by Hundreds-Baseline Stereo Using a Hand-Held Video Camera," *Int. Journal of Computer Vision* **47**(1-3), pp. 110–129, 2002.
8. B. K. P. Horn, *Robot Vision*, ch. 10, pp. 206–209. Mit Press, 1986.
9. N. Asada, A. Amano, and M. Baba, "Photometric Calibration of Zoom Lens Systems," *Proc. Int. Conf. Pattern Recognition* **A**, pp. 186–190, 1996.