SIMULTANEOUS ESTIMATION OF LIGHT SOURCES POSITIONS AND CAMERA ROTATION

Masahiro Oida, Fumihiko Sakaue and Jun Sato
Nagoya Institute of Technology, Gokiso, Showa, Nagoya 466-8555, Japan

Keywords: Mixed reality, Photometric calibration, Geometric calibration.

Abstract: For mixed reality and other applications, it is very important to achieve photometric and geometric consistency in image synthesis. This paper describes a method for calibrating camera and light source simultaneously from photometric and geometric constraints. In general, feature points in a scene are used for computing camera positions and orientations. On the other hand, if the cameras and objects are sticked and move together, the changes in shading information of the objects in images also include useful information on geometric camera motions. In this paper, we show that if we use both shading information and feature point information, we can calibrate cameras from smaller number of feature points than the existing methods. Furthermore, it is shown that the proposed method can calibrate light sources as well as cameras. The accuracy of the proposed method is evaluated by using real and synthetic images.

1 INTRODUCTION

Recently, mixed reality systems are studied extensively(Milgram and Kishino, 1994). In these systems, real images are taken by cameras and virtual information is added to images. In order to achieve realistic mixed reality, it is important to obtain scene environment such as lighting information and geometric information such as camera parameters.

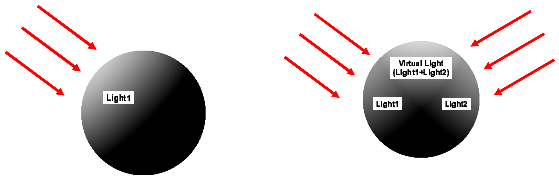
In general, geometric information of cameras is obtained from coordinates of feature points in images by using camera calibration techniques(Hartley and Zisserman, 2000; Faugeras and Luong, 2001). On the other hand, lighting information is obtained from images of specular sphere(Powell et al., 2001), Lambertian sphere(Takai et al., 2003) and so on(Sato et al., 1999b; Sato et al., 1999a). In these existing methods, geometric information and photometric information were obtained separately in different way. However, these informations are actually closely related to each other. For example, if an object is fixed to the camera, the illumination of the object changes according to camera motions, and the change in intensity of the object in the camera image provides us very useful information for estimating camera motions. Also, if a light source is attached on a camera and moves together with the camera, the illumination of a static object changes according to the camera motions, and the change in intensity of the scene provides us useful information for estimating camera motions.

In this paper, we propose a method which enables us to calibrate lighting information and geometric information simultaneously by combining photometric and geometric information. In this method, lighting information is obtained from observation of a reference object, and camera parameters are computed by combining photometric and geometric information, such as shading information and image coordinates of feature points in images. Since there are many light sources in general real scenes and their distributions vary, we consider a geodesic dome around the 3D scene and light sources are distributed on the geodesic sphere. Then, the distribution of light sources is estimated and used for recovering camera motions from photometric and geometric informations. By using both photometric and geometric informations, camera calibration can be achieved from smaller number of feature points.

2 ESTIMATION OF LIGHT SOURCE POSITIONS

We first consider the estimation of light source positions from images. In this paper, we consider a scene, where a known object with Lambertian surface exists with other unknown objects, and they are illuminated

Oida M., Sakaue F. and Sato J..
SIMULTANEOUS ESTIMATION OF LIGHT SOURCES POSITIONS AND CAMERA ROTATION.
DOI: 10.5220/0003374901680174
In *Proceedings of the International Conference on Computer Vision Theory and Applications* (VISAPP-2011), pages 168-174
ISBN: 978-989-8425-47-8



(a) under single light source (b) under multiple light sources

Figure 1: Image intensities under (a) single light source and (b)multiple light sources.

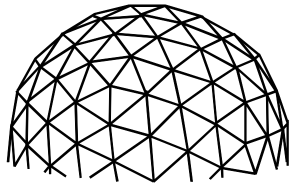


Figure 2: Sampling of light source directions using geodesic dome.

by infinite light sources. Suppose a point $\mathbf{X}$ on the known object is projected to a point $\mathbf{m}$ in the image. Then, the intensity I of the image point $\mathbf{m}$ is determined by a surface normal $\mathbf{n}$ at point $\mathbf{X}$ on the object and a light source direction $\mathbf{s}$ as follows:

$$I = \max(E\rho\mathbf{n}^\top\mathbf{s}, 0), \tag{1}$$

where ρ is the albedo of the surface and E is the power of the light source. If there is a single light source in the scene as shown in Fig.1 (a), we can estimate the light source position from an image by using Eq. (1). However, if there are multiple light sources as shown in Fig.1 (b), it is not easy to estimate light source positions, since some areas are illuminated by multiple light sources, and some areas are not. The number of light sources which illuminate a point on the surface varies depending on the surface normal of the point and the distribution of the light sources. Under such environment, we consider a method which can estimate camera motions and light source distributions simultaneously.

In this paper, we consider two cases separately. We first consider a case where there is only a single light source in the scene. We next consider a case where there are multiple light sources.

2.1 Light Source Estimation under Single Light

We first consider a case where there is only a single light source in the scene. Let us consider an object with Lambertian surface in the scene, whose shape (including surface normal and albedo) is known. If the distance from a light source to the object is sufficiently large, the intensity of a point on this object can be described by Eq. (1). Suppose we have M points $\mathbf{X}_i$ $(i = 1, \cdots, M)$ on the object surface, and their surface normal and albedo are $\mathbf{n}_i$ and ρ_i respectively. Then, the following matrix $\mathbf{W}$, which represents the surface geometry, is known:

$$\mathbf{W} = \begin{bmatrix} \rho_1\mathbf{n}_1 & \cdots & \rho_p\mathbf{n}_p \end{bmatrix}^\top \tag{2}$$

Let $\mathbf{I}$ be a vector which consists of the image intensity I_i $(i = 1, \cdots, M)$ at these points:

$$\mathbf{I} = \begin{bmatrix} I_1 & \cdots & I_p \end{bmatrix}^\top \tag{3}$$

Then the intensity $\mathbf{I}$ can be described by using the light power E, the light source direction $\mathbf{s}$ and the surface geometry $\mathbf{W}$ as follows:

$$\mathbf{I} = E\mathbf{W}^\top\mathbf{s} \tag{4}$$

Note, $\mathbf{s}^\top\mathbf{s} = 1$. By using the least means square method, $E\mathbf{s}$ can be estimated from $\mathbf{I}$ and $\mathbf{W}$ as follows:

$$E\mathbf{s} = \left(\mathbf{W}^+\right)^\top\mathbf{I} \tag{5}$$

where $\mathbf{W}^+ = \mathbf{W}^\top(\mathbf{W}^\top\mathbf{W})^{-1}$ is the pseudo inverse of matrix $\mathbf{W}$. The number of components of a vector $E\mathbf{s}$ is 3, and thus we can estimate light source direction from 3 or more than 3 different intensities.

Note that if the intensity estimated from a surface normal and a light source direction is negative, we cannot use the pixel for the linear light source estimation. Thus, we have to eliminate such pixels as outliers by using the robust estimation methods, such as RANSAC(Hartley and Zisserman, 2000) for correct estimation.

2.2 Light Source Estimation under Multiple Lights

We next consider a case where there are multiple light sources in the scene. In this case, intensity I_i of i-th pixel can be described by using the surface normal $\mathbf{n}_i$ and light source directions $\mathbf{s}_j$ as follows :

$$I_i = \sum_j \max(E_j\rho_i\mathbf{n}_i^\top\mathbf{s}_j, 0), \tag{6}$$

where ρ_i denotes the albedo of i-th point and E_j denotes the light power of j-th light source. In this case, we cannot linearly estimate light sources positions from the equation, since non-linear mapping by $\max(I_j, 0)$ is included in this equation.

In order to estimate light source distributions under multiple lights, Sato et al(Sato et al., 1999b) represented the light source distributions by using a geodesic dome as shown in Fig.2. In this model, a light source distribution is represented by light source power E_k $(k = 1, \cdots, N)$ in N light source directions

$\mathbf{s}_k$ $(k = 1, \cdots, N)$. Then, the image intensity I_i of ith point can be described as follows:

$$I_i \sim \sum_{k=1}^{N} E_k \rho_i v(i,k) \mathbf{n}_i^\top \mathbf{s}_k, \qquad (7)$$

where $v(i,k)$ denotes a function which takes 1 if the i-th pixel is illuminated from the k-th light source direction, and takes 0 in the other case. Since $\mathbf{s}_k$ $(k = 1, \cdots, N)$ is predefined on the geodesic dome, the estimation of a light source distribution is same as the estimation of E_k $(k = 1, \cdots, N)$.

In our case, we know $v(i,k)$ since we know the object shape, and thus E_k is the only unknown variable in Eq.(7). Therefore, we can estimate E_k from $\mathbf{I}$ by solving the following linear equations:

$$\mathbf{I} = \begin{bmatrix} \rho_1 v(1,1)\mathbf{n}_1^\top \mathbf{s}_1 & \cdots & \rho_1 v(1,N)\mathbf{n}_1^\top \mathbf{s}_N \\ & \vdots & \\ \rho_M v(M,1)\mathbf{n}_M^\top \mathbf{s}_1 & \cdots & \rho_M v(M,N)\mathbf{n}_M^\top \mathbf{s}_N \end{bmatrix} \begin{bmatrix} E_1 \\ \vdots \\ E_N \end{bmatrix} \qquad (8)$$

Thus, we can linearly estimate light source distributions on the geodesic dome from the image. Althought estimated E_k may not be identical with the real light source environment, the estimated light source distributions can be used for generating images which are identical with the input images.

3 SIMULTANEOUS ESTIMATION OF LIGHT SOURCES AND CAMERA MOTIONS

In this section, we consider a method which enables us to estimate camera motions and lighting environments simultaneously. Let us consider a case, where a camera moves in the scene and images are taken by the camera under light sources which are attached on the camera. This is equivalent to the case where objects move under a fixed camera and lights as shown in Fig.3. We consider the estimation of camera motions in such cases.

3.1 Estimation of Camera Position using Geometric Information

In general, camera motions are estimated from geometric information such as markers in the scene. Let a 3D point $\mathbf{X}$ in a 3D scene be projected into the camera image as $\mathbf{m}$ as follows:

$$\lambda \tilde{\mathbf{m}} = \mathbf{A} \begin{bmatrix} \mathbf{R} & \mathbf{T} \end{bmatrix} \tilde{\mathbf{X}}, \qquad (9)$$

where $\tilde{}$ denotes the homogeneous coordinates, and $\mathbf{A}$, $\mathbf{R}$ and $\mathbf{T}$ denote the intrinsic parameters, rotation and the translation of the camera.

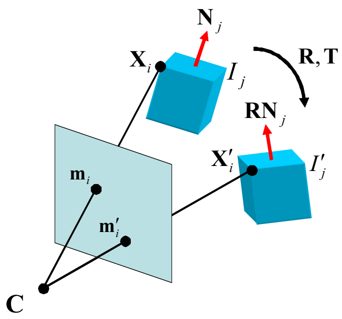


Figure 3: Camera motion and scene motion.

If the intrinsic parameters $\mathbf{A}$ and object shape $\mathbf{X}$ are known, the number of unknown variables in Eq.(9) is 6. We can obtain 2 constraints from Eq.(9) for each image point $\mathbf{m}$. Thus, we can calibrate extrinsic parameters, $\mathbf{R}$ and $\mathbf{T}$, from the projections of 3 known points.

3.2 Estimation of Camera Rotation under Single Light Source

As shown in the previous section, we need 3 projected points in order to calibrate the extrinsic parameters in general. However, we can calibrate camera parameters from fewer feature points. if we use lighting information for calibration.

Suppose a light source, such as a projector, is fixed to the camera, and they move together in the 3D scene. Then, the intensity of the scene object in the camera image changes according to the camera motion. Then, the information of the camera motion is included in the change in intensity of the object as shown in Fig.4. Therefore, by using the intensity information, extrinsic parameters of the camera can be estimated from fewer feature points than usual.

Now, let us consider a method for estimating camera motions from intensity information. We first consider a scene where a single light source exists. Let us consider 2 images taken by a camera at two different positions under the same lighting condition. In this case, intensities I_j and I'_j of j-th pixel in two camera images can be described as follows:

$$\begin{aligned} I_j &= E\rho_j \mathbf{n}_j^\top \mathbf{s} && (10) \\ I'_j &= E\rho_j \mathbf{n}_j^\top \mathbf{s}' & \\ &= E\rho_j \mathbf{n}_j^\top \mathbf{R}^\top \mathbf{s} && (11) \end{aligned}$$

In this equation, there is no effect of translation $\mathbf{T}$, since we assume infinite light source. Thus, we can obtain constraints for rotation $\mathbf{R}$. Since we can estimate lighting direction $\mathbf{s}$ and $\mathbf{s}'$ for two images using

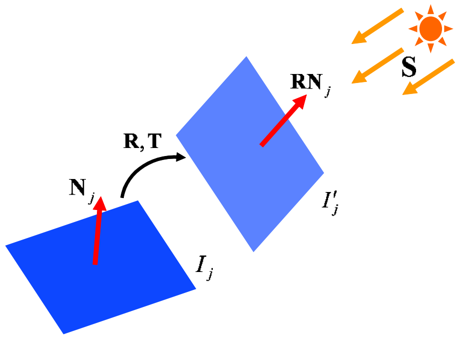


Figure 4: Change in intensity caused by object/camera motion.

Eq.(5), constraints for rotation $\mathbf{R}$ can be obtained as follows:

$$\mathbf{s}' = \mathbf{R}^{\top}\mathbf{s} \tag{12}$$

From this equation, we cannot estimate a rotation around $\mathbf{s}$, since the intensity does not change when a camera rotates around $\mathbf{s}$. Thus, we have to combine geometric information for complete estimation of extrinsic parameters. From the intensity constraints, we can obtain 2 constraints for extrinsic parameters. Thus, for estimating complete extrinsic parameters, we require 2 points in the scene, which provides us 4 additional constraints for extrinsic parameters.

Furthermore, we can estimate optimized light source directions as well as extrinsic parameters by minimizing the following cost function.

$$e = \sum_{l} ||\mathcal{N}(\mathbf{A}[\mathbf{R}\ \ \mathbf{T}]\tilde{\mathbf{X}}) - \mathbf{m}||^2 + \sum_{l}\sum_{i}(I_{i,l} - E\rho_i \mathbf{n}_i^{\top}\mathbf{R}\mathbf{s})^2, \tag{13}$$

where $\mathcal{N}(\cdot)$ indicates the normalization of homogeneous coordinates, and $I_{i,l}$ denotes the intensity of ith pixel in lth image. The equation includes non-linear components, and we use Newton-Raphson method inorder to minimize the equation.

3.3 Estimation of Camera Rotation under Multiple Light Sources

We next consider scenes where multiple light sources exist. In our method, we observe a Lambert sphere for estimating camera and light source information. In general the appearance of a sphere in the image does not change except size and illumination, even if the camera is moved.

When we have multiple images taken under corresponding light sources, an intensity $I_{i,l}$ of ith pixel in lth image can approximately be represented as follows:

$$I_{i,l} \sim \sum_{k=1}^{N} E_k \rho_i v(i,k)\mathbf{n}_i^{\top}\mathbf{R}_l^{-1}\mathbf{s}_k, \tag{14}$$

where, $\mathbf{R}_l$ is the camera rotation of lth camera. By estimating E_k and $\mathbf{R}_l$ from the equation, we can estimate light source environment and camera motion simultaneously.

In order to estimate these components, we use iterative estimation method. At first, light source environment for 1st image is estimated from Eq.(8). We next estimate camera motion of lth camera ($l = 2, \cdots, M$) by minimizing the following cost function.

$$e_l = \sum_{i}(I_{i,l} - \sum_{k=1}^{N} E_k \rho_i v(i,k)\mathbf{n}_i^{\top}\mathbf{R}_l^{-1}\mathbf{s}_k,) \tag{15}$$

The minimization can be achieved by using a non-linear minimization method. In this paper, we used Gauss-Newton method. By using estimated $\mathbf{R}_l$, E_k is updated from the following linear estimation.

$$\mathbf{I} = \begin{bmatrix} \mathbf{M}_1 \\ \vdots \\ \mathbf{M}_L \end{bmatrix} \begin{bmatrix} E_1 \\ \vdots \\ E_M \end{bmatrix} \tag{16}$$

where

$$\mathbf{M}_l = \begin{bmatrix} \rho_1 v(1,1)\mathbf{n}_1^{\top}\mathbf{R}_l^{-1}\mathbf{s}_1 & \cdots & \rho_1 v(1,N)\mathbf{n}_1^{\top}\mathbf{R}_l^{-1}\mathbf{s}_N \\ & \vdots & \\ \rho_M v(M,1)\mathbf{n}_M^{\top}\mathbf{R}_l^{-1}\mathbf{s}_1 & \cdots & \rho_M v(M,N)\mathbf{n}_M^{\top}\mathbf{R}_l^{-1}\mathbf{s}_N \end{bmatrix} \tag{17}$$

and, $\mathbf{I}$ is a vector which consists of the intensity vector $\mathbf{I}_l$ of lth image as follow:

$$\mathbf{I} = \begin{bmatrix} \mathbf{I}_1 & \cdots & \mathbf{I}_L \end{bmatrix}^{\top} \tag{18}$$

The estimation of E_k and $\mathbf{R}_l$ is iterated until convergence. Then, we can obtain the camera motion $\mathbf{R}_l$ and light source environment E_k simultaneously. Note that, in this method we can estimate light source environment more accurately than the existing single image method, since we can obtain more information about light sources from multiple images in our method.

4 EXPERIMENTS

In this section, we show some experimental results from the proposed method.

4.1 Experiments under Single Light Source

We first show experimental results under single light source. Experimental environments are shown in Fig.5. In this experiment, a dodecahedron which has Lambertian surface as shown in Fig.6 was used. The

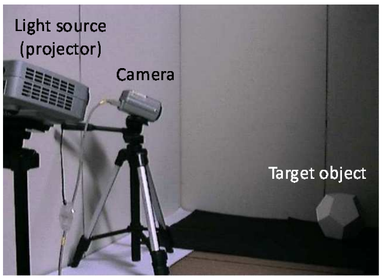


Figure 5: Experimental environment.

(a) Image 1 (b) Image 2

Figure 6: Input images.

distance from the target to the light source is sufficiently large. In this experiment, we moved target object instead of camera. The camera and the light source were fixed. Thus, the relative relationship between camera and target object was changed and the relative relationship between the object and the light source was also changed. Therefore, we estimate the motion of the target object from images instead of camera motions by using our method.

The images were taken before and after the object motion. These images are shown in Fig.6. The vertices of the dodecahedron in the input images were used as feature points, and the position of these points in images were extracted manually. We estimated the object motions and light source directions simultaneously by using the coordinates of feature points and the image intensity of each plane.

The estimated light source directions and object rotations were evaluated by reproducing the second image from the first image by using the estimated light source directions and rotations, and comparing intensities of the reproduced image with those of the real image. The RMS error of intensity was 8.1, and we find the proposed method can estimate light source directions and rotations accurately.

The estimated object motions were also evaluated by reprojecting the feature points by using the estimated object motions, and computing the errors in the reprojected positions. The average error of the reprojected point was 6.2 pixels, and it indicates that our method can estimate relative motions between cameras and objects as well as light source directions under single light source environments.

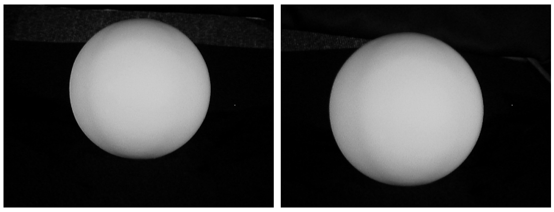


Figure 7: Input images taken from different positions.

4.2 Experiments under Multiple Light Sources

We next show the experimental results under multiple light source environments. In this experiment, a Lambertian sphere is used for estimating the distribution of light sources and the rotational motions of the camera simultaneously.

The sphere was illuminated by 2 lights, and images were taken by a camera placed at two different positions with different orientations. These images are shown in Fig.7. In this experiment, the target object was fixed and the camera was moved.

By using the left and right images in Fig.7, the light source distribution and rotation **R** were estimated simultaneously from the proposed method. Note that, the position and the size of the sphere are different in these two images. Therefore, we computed the position and the size of the sphere by using the Hough transformation and normalized them in two images, so that the corresponding points on the sphere are at the same position in these two images.

The accuracy of the recovered light source distribution and rotation components were evaluated by reproducing the intensity images by using the recovered light source distribution and rotation. For representing the light source distribution, we used 18, 66, 102 and 146 light source directions. The reproduced image in each case is shown in Fig.8.

As shown in this figure, the reproduced images are almost identical to the real input images shown in Fig.7, and the intensity errors become small when the number of sampled lights become large. The results indicate that the proposed method works very well for image reconstruction, and it can estimate light source information properly. Table 1 shows estimated rotations and their ground truth. The results indicate that the proposed method needs some improvements for obtaining better accuracy.

We next show the results under natural lighting environment as shown in Fig.9. The images observed before and after camera motion are shown in Fig.10. The markers on the sphere are used for estimating the ground truth motions. The light source distribution

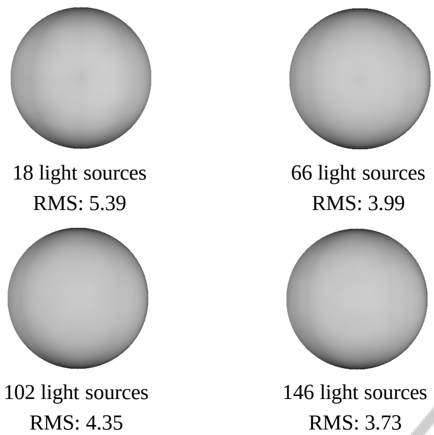


Figure 8: The number of light sources in the light source representation and reconstructed images with RMS errors.

Table 1: Estimated rotations and ground truth(GT) [degree].

	GT	18	66	102	146
θ	2.52	0.38	0.45	0.43	0.49
ϕ	16.6	8.12	7.16	8.51	7.23
δ	2.70	2.35	2.27	2.38	2.21
error	-	5.09	5.63	4.86	5.58

and the camera rotation were estimated from these images. The number of sampled lights is fixed to 102. From the estimated light source distribution and rotation, we reproduced images before and after the camera motion. The reproduced images are shown in Fig. 11. The RMS intensity errors of the reproduced images were 4.77 and 5.37 respectively, and the average error of estimated rotation was 6.38 degrees. The results indicate that the proposed method works well even under natural lighting environments.

4.3 Accuracy Evaluation under Multiple Lights

In this section, we evaluate the accuracy of the proposed method under multiple light sources by using synthetic images. We first evaluate the relationship between the number of actual light sources in the scene and the accuracy of estimation. In this evaluation, the number of actual light sources was changed, and camera rotation was estimated by our method. Figure.12 shows the relationship between the number of actual lights and the errors in estimated rotation.

As shown in this figure, although the accuracy degrades slightly as the number of actual lights increases, the proposed method still provides us sufficient accuracy even under hundreds of lights.

We next evaluate the relationship between the number of sampled lights and the accuracy of rotation estimation. The number of sampled lights was changed from 18 to 400. Figure.13 shows the results. Although the estimation errors are relatively large when the number of sampled lights is small, they are still acceptable, and they become small when we use sufficient number of sampled lights. Computational cost for the estimation does not become extremely large even under hundreds of lights, since our method consists of linear estimations. Thus, we should use a large number of sampled lights for accurate estimation.

Figure 9: Experimental environment in a nutural room.

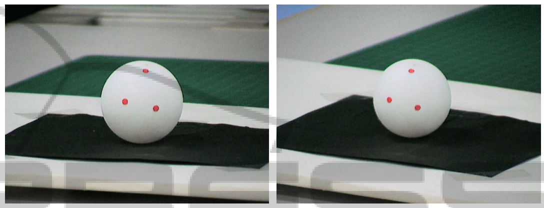


Figure 10: Input images taken under natural room.

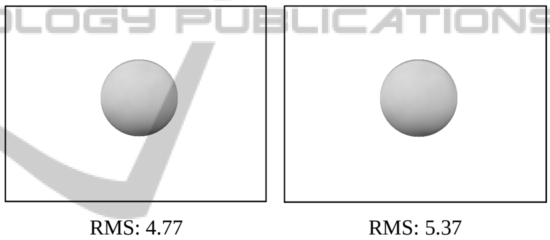


Figure 11: Reconstructed image from estimated results and RMS errors.

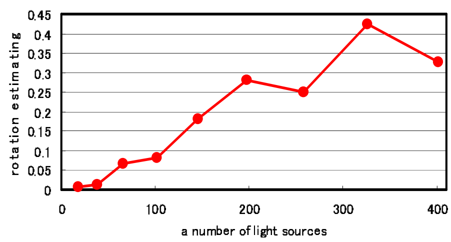


Figure 12: Relationship between a number of actual light sources and errors in rotation estimation.

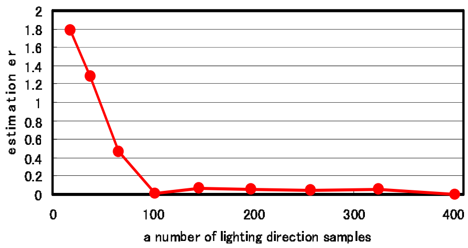


Figure 13: Relationship between a number of sampled light sources and errors in rotation estimation.

5 CONCLUSIONS

In this paper we proposed a method for estimating light source distribution and camera motions simultaneously from photometric and geometric informations. For this objective, we proposed two types of method. One is for the case where a single light source exists in the scene, and the other is for multiple light sources. Under a single light source, we estimated light source direction, camera rotation and translation from image intensities and image coordinates of feature points. Under multiple light sources, we estimated the distribution of light sources and camera rotations simultaneously just from image intensity.

The experimental results indicate that the proposed methods can estimate lighting information and camera rotation simultaneously, even if there are a lot of light sources such as natural rooms. In the future work, we improve the accuracy of camera motion estimation.

REFERENCES

Faugeras, O. and Luong, O. (2001). *The Geometry of Multiple Images*. MIT Press.

Hartley, R. and Zisserman, A. (2000). *Multiple View Geometry in Computer Vision*. Cambridge University Press.

Milgram, P. and Kishino, A. F. (1994). Taxonomy of mixed reality visual displays. *IEICE Transactions on Information Systems*, E77-D(12):1321–1329.

Powell, M. W., Sarkar, S., and D.Goldgof (2001). A simple strategy for calibrating the geometry of light sources. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 23:1022–1027.

Sato, I., Sato, Y., and Ikeuchi, K. (1999a). Acquiring a radiance distribution to superimpose virtual objects onto a real scene. *IEEE Trans. Visualization and Computer Graphics*, 5(1):1–12.

Sato, I., Sato, Y., and Ikeuchi, K. (1999b). Illumination distribution from shadows. In *Proc.IEEE Conference on Computer Vision and Pattern Recognition 99*, pages 306–312.

Takai, T., Niinuma, K., and Matsuyama, T. (2003). 3d lighting environment sensing with reference spheres. *Computer Vision and Image Media*, 137(15):117–124.