Estimation of illumination distribution based on arbitrary shape objects

Xiaoyan Tong, Hideo Saito
Department of Information and Computer Science
Keio University, Yokohama, Japan
{vickyan, saito}@ozawa.ics.keio.ac.jp

Abstract—This paper describes a new method for estimating the illumination distribution of a real scene from a radiance distribution inside shadows cast by any arbitrary object in the scene. First, a multi-camera system captures several planar images of a calibration board in order to get the projective matrix of every camera. Then, since the illumination distribution is represented by intensities of sampled points, the virtual view can be defined. With the fundamental matrix between camera and virtual view, serials images of virtual views are obtained. According to the character of a point light source and camera, the silhouettes of the object are considered as cast shadow regions and are extracted. Lastly, the illumination distribution of the scene is recovered from the radiance distribution inside cast shadows. For demonstrating the effectiveness of the proposed method, an experiment for overlay an object onto a real scene with correct shadow is also performed.

I. INTRODUCTION

With the development of the computer technology, integration of virtual objects with an image of a real scene is an important step toward the long-term goal of achieving photographic realism of synthesized images. Such techniques for merging virtual objects with a real scene, which are called as "augmented reality", attract a great deal of attention in the fields of both computer graphics and computer vision research recently. The synthesized world via such augmented reality technology allows us to handle phenomena not only in the real world but also in virtual world, while virtual reality technologies immerse a user in a fully computer-generated.

In order to achieve a high quality synthesized image in augmented reality system, three consistencies have to be taken into account: the consistency of geometry, the consistency of time and the consistency of illumination.

Many studies about consistency of geometry and time have been intensively investigated in the field of augmented reality[3], [4]. We can superimpose a virtual object onto the image that we are looking for in real time by using either artificial markers or 3D position sensors of various modalities.

On the other hand, for the complex of the illumination distribution of the real scene, there are few researches on the third consistency. Fournier et al [7] proposed pioneering work in this field, which is short on the requirement of specifying all the 3D shapes of objects in the scene. Recently, Debevec [5] introduced a new method of estimation of radiance distribution by using some special equipment like a spherical mirror. Sato [1], [2] presented a novel idea, which combines the illumination analysis with an estimation of the reflectance properties of a surface in side shadows. From Sato's method we can suppose the illumination distribution with image took by a color CCD camera simply instead of special equipments. But with the limitation of known shape object, it cannot be applied in wide field.

In this paper, an extended method is presented to overcome the limitations discussed above. The illumination distribution is represented by intensities of sampled points on a spherical extended light source. For estimating the intensities in the object scene, shape of cast shadow generated by a point light source at each sampled point must be provided. In our method, we estimate the image of sampled point by the geometric relationship between multi cameras and virtual view; obtain the shape of the cast shadow by transforming the silhouette of the object. Then according to Sato's theory, the illumination distribution can be estimated from the shape of cast shadow by a point light source at the sampled point. In our experiment, a multi-camera system with 8 calibrated cameras [8] is used to capture the object scene. Then from the images of 8 cameras, we can recover the illumination distribution represented by intensities of the camera positions.

The rest of the paper refers to one camera of the multi-camera system as *main camera*, to the other cameras as *illumination cameras*, to the image taken by *main camera* as *shadow image*, to the object, which casts shadow onto the scene as *occluding object*, and to the images, which are taken by *illumination cameras* as *illumination images*.

II. THEORY

- Relating illumination radiance with image irradiance.

To take illumination from all directions into account, discrete sampling over the entire surface of the extended light source as shown in Figure 1. See from the center point *A*, the illumination radiance $L_0(\theta_i, \phi_i)$ per unit solid angle coming from the direction (θ_i, ϕ_i); and then integrate the product of the bi-directional reflectance distribution function (BRDF) $f(\theta_i, \phi_i, \theta_e, \phi_e)$ and the illumination radiance over the entire hemisphere, the pixel value of the *shadow image* $P(\theta_e, \phi_e)$ is compute as

$$P(\theta_e, \phi_e) = \sum_{i=1}^{n} f(\theta_i, \phi_i, \theta_e, \phi_e) L(\theta_i, \phi_i) S(\theta_i, \phi_i) \cos\theta_i \quad (1)$$

where $S(\theta_i, \phi_i)$ are occlusion coefficients; $S(\theta_i, \phi_i) = 0$ if $L_0(\theta_i, \phi_i)$ is occluded by the occluding object; otherwise $S(\theta_i, \phi_i) = 1$.

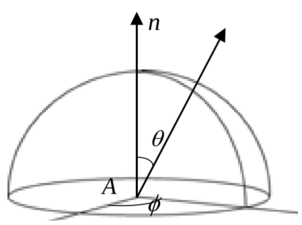


Fig 1 Illumination discrete sampling

Suppose the surface is a Lambertian surface; BRDF $f(\theta_i, \phi_i, \theta_e, \phi_e)$ for a Lambertian surface is known to be a constant, a linear equation is obtained for each image pixel of the shadow image as

$$a_1 L_1 + a_2 L_2 + a_3 L_3 + \cdots + a_{1n} L_n = P \quad (2)$$

where $L_i (i = 1,2,\ldots,n)$ are n unknown illumination radiance specified by n node directions of a geodesic dome. The coefficients $a_i (i = 1,2,\ldots,n)$ represent in equation 1. Therefore by selecting a number of image pixels, a set of linear equations is obtained as

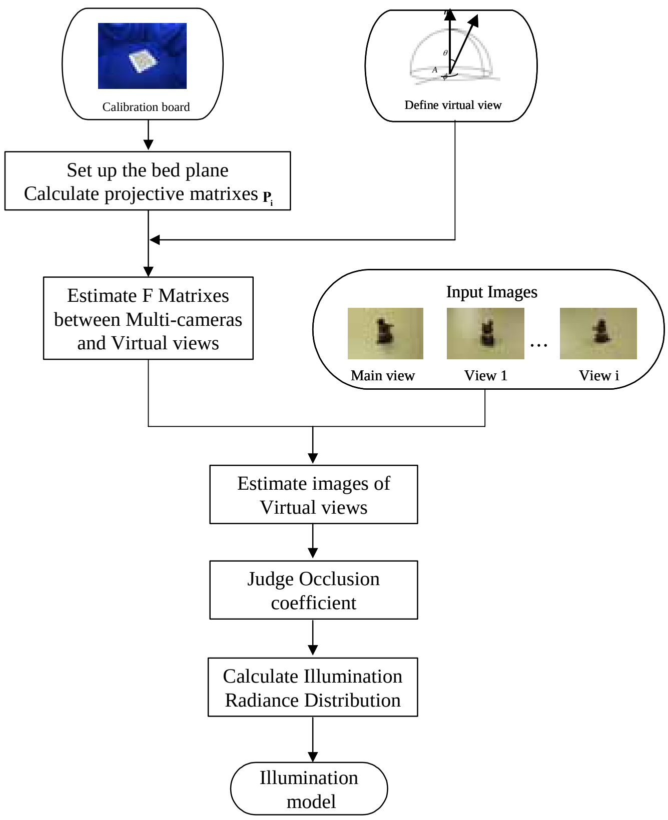
$$\begin{aligned} a_{11} L_1 + a_{12} L_2 + a_{13} L_3 + \cdots + a_{1n} L_n &= P_1 \\ a_{21} L_2 + a_{22} L_2 + a_{23} L_3 + \cdots + a_{2n} L_n &= P_2 \\ &\vdots \\ a_{m1} L_1 + a_{m2} L_2 + a_{m3} L_3 + \cdots + a_{mn} L_n &= P_m \end{aligned} \quad (3)$$

Solution the equations, the set of unknown L_i can be obtained.

III. ESTIMATION SYSTEM

A. Process flowchart

Figure 2 shows the flow of the process of the proposed method.



For the difference with the conventional methods, we use a multi-camera system. As the pre-processing, a calibration board is used to define the world coordinate system. At the same time, multi-camera system with several planar patterns of the calibration board obtains the camera projective matrix of every camera.

Here, we find a relationship between the shadow cast by the point light source and the silhouette of the object by camera. Projecting every silhouette image of virtual view to *main camera*, the shadows casts from sampled points are obtained. Therefore, a serial of occluding coefficients are obtained.

Lastly, by using the formula of Sato described above, the illumination distribution of sampled points could be calculated

B. Bed plane

In order to estimate the position of the shadows and superimpose the object onto the real scene in suitable position, we should set up the world coordinate beforehand.

With the calibration board, we set the plane of board is Z=0.

C. Camera projective matrix

The calibration board is a board with regularly grids, e.g., 10×9 grids in our experiments.

We take a serial of calibration board images with different angles as planar patterns.

To every camera, there is a relationship between world coordinate and image coordinate as below:

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \underbrace{\mathbf{A}_i [\mathbf{R}_i \mid \mathbf{t}_i]}_{\mathbf{P}_i} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \qquad (4)$$

where $\mathbf{A}_i$ is the intrinsic parameter and $[\mathbf{R}_i \mid \mathbf{t}_i]$ is the extrinsic parameter of camera_i.

To every planar pattern of the calibration board, there is also a relationship between plane coordinate and world coordinate as below:

$$\begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} = \underbrace{\begin{bmatrix} \mathbf{p}^j & \mathbf{q}^j & \mathbf{d}^j \\ 0 & 0 & 1 \end{bmatrix}}_{\mathbf{Q}^j} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \qquad (5)$$

where, $\mathbf{p}^j$, $\mathbf{q}^j$ are the vector of vertical axis and horizontal axis, $\mathbf{d}^j$ is the original position from the world coordinate.

Form equation (4) and (5),

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \underbrace{\mathbf{P}_i \mathbf{Q}^j}_{\mathbf{H}_i^j} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \qquad (6)$$

For $\mathbf{H}_i^j$ is a 3×3 matrix, with knowing at least four group corresponding points of image_i and plane_j, we can obtain the matrix $\mathbf{H}_i^j$. By using the Singular Value Decomposition, the project camera matrix $\mathbf{P}_i$ can be obtained.

D. Silhouette Image in Virtual View

With knowing the projective matrixes of every camera and virtual view, we can obtain the fundamental matrixes between camera and virtual view with defined equation (7).

$$\begin{aligned} \mathbf{F}_{vi} &= [\mathbf{P}_i \tilde{\mathbf{c}}]_\times \mathbf{P}_i \mathbf{P}_v^- \\ \mathbf{P}_v^- &= \mathbf{P}_v^T (\mathbf{P}_v \mathbf{P}_v^T)^{-1} \end{aligned} \qquad (7)$$

Where $\mathbf{F}_{vi}$ is the fundamental matrixes between camera and virtual view, $\mathbf{P}_i, \mathbf{P}_v$ are projective matrixes of camera and virtual view accordingly.

Then as show in Figure 3, from every camera, epipolar line is drawn to the virtual view. The crossing part can be regarded as silhouette image in virtual view.

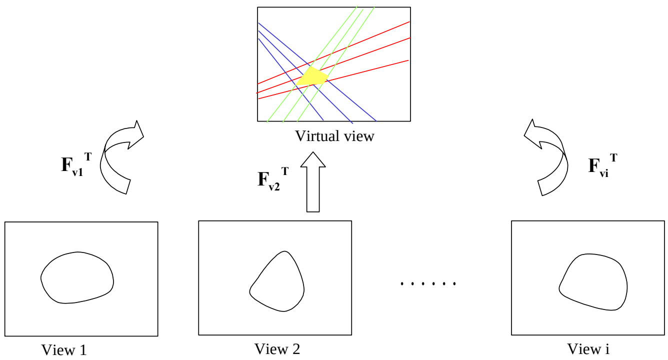


Fig 3 Estimate Virtual view image

E. Images Transfer

We extent our method from Sato's theory with the advantage of abandon the limitation of the knowledge of object's shape, so that it can be used in wider field. From equation 1 we can find the point to solute it is to decide the occlusion coefficients. That is to say how to judge the occluding object occludes the illumination or not. The conventional method based on the condition of known shape object, specified the 3D coordinates of a point light source and the occluding objects, it is easy to calculate the shadow's position related to the sampled point by geometry relationship.

In fact, however, to get all the 3D coordinates of all objects in the scene is impossible. In this method we are try to find if there is any relationship between the cast shadow of point light source and the image of camera.

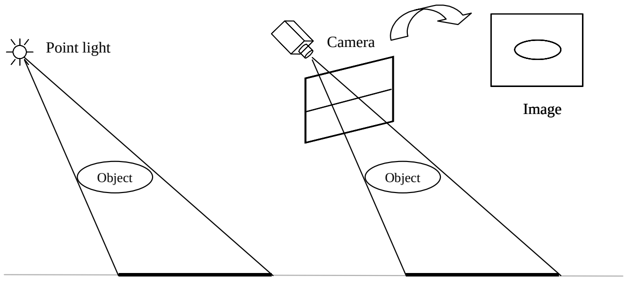


Fig 4 Relationships between light and camera

From the definition as show in Figure 4. Shadow is produced because an occluding object occludes the light. In fact, the shape of the shadow is similar to the shape of the occluding object faced to the point light source. It just makes some transformation. Consider that with the same position and orientation of a point light source, the silhouette image of object can be regarded as the projection of the shadow.

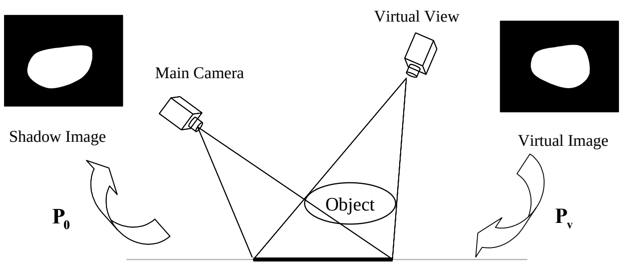


Fig 5 Transfer from illumination image to shadow image.

The virtual image can be transferred to the *shadow image* as show in Figure 5.

Having got the conception of illumination and image, the transfer matrix calculates as below.

The relationship of the image and object is showed as equation 8.

$$\lambda_v \begin{bmatrix} u_v \\ v_v \\ 1 \end{bmatrix} = \mathbf{P_v} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} = \begin{bmatrix} P_{11} & P_{12} & P_{13} & P_{14} \\ P_{21} & P_{22} & P_{23} & P_{24} \\ P_{31} & P_{32} & P_{33} & P_{34} \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \quad (8)$$

Because the shadow is in the bed that Z=0, so

$$\lambda_v \begin{bmatrix} u_v \\ v_v \\ 1 \end{bmatrix} = \mathbf{P_v} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} = \begin{bmatrix} P_{11} & P_{12} & P_{13} & P_{14} \\ P_{21} & P_{22} & P_{23} & P_{24} \\ P_{31} & P_{32} & P_{33} & P_{34} \end{bmatrix} \begin{bmatrix} X \\ Y \\ 0 \\ 1 \end{bmatrix}$$

$$= \begin{bmatrix} P_{11} & P_{12} & P_{14} \\ P_{21} & P_{22} & P_{24} \\ P_{31} & P_{32} & P_{34} \end{bmatrix} \begin{bmatrix} X \\ Y \\ 0 \\ 1 \end{bmatrix} = \mathbf{H_v} \begin{bmatrix} X \\ Y \\ 1 \end{bmatrix} \quad (9)$$

And then the shadow in the bed re-projects to the main view.

$$\lambda_0 \begin{bmatrix} u_0 \\ v_0 \\ 1 \end{bmatrix} = \mathbf{H_0} \begin{bmatrix} X \\ Y \\ 1 \end{bmatrix} = \mathbf{H_0} \mathbf{H_v^{-1}} \begin{bmatrix} u_v \\ v_v \\ 1 \end{bmatrix} \quad (10)$$

From equation 10 we transfer every virtual image to the image of *main camera*, which could be regarded as the *shadow image* according to every sampled point and the occlusion coefficients in these fields must be 0 accordingly.

F. Illumination distribution

Formula for relating illumination radiance with image irradiance has been described in Section II. After knowing the occlusion coefficients, selecting all the pixels in *shadow image*, from equation 3 we are able to estimate unknown L_i of *Virtual View*. Note that, since each pixel consists of 3 color bands (R, G, B), each band of illumination L_i is also estimated from the corresponding color band of the image.

IV. EXPERIMENT RESULTS

We have tested the proposed method by using real images taken in blue back environment.

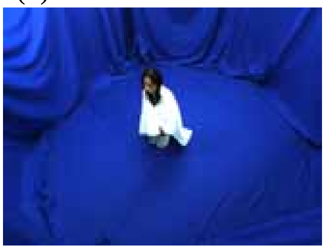
For the beginning of our experiment, we only consider the illumination condition on the camera position and instead the camera images with virtual views. We set beforehand and only suppose that there is once reflection in the scene. We use a multi-camera system with 8 cameras, one is assigned as the *main camera* and the others are assigned as estimated *illumination cameras*. Then an image with a people was taken under the settled illumination environment. The input image is shown in Figure 6(a).

First, the multi-camera system captures several planar images of a calibration board in order to get the projective matrixes as described in Section III C. Regularly spaced grids on the calibration board showed in Figure 6(b). Also, with the calibration board we define the world coordinate system.


(a) main camera image


(b) calibration board


camera 1


camera 3


camera 5


camera7

(c)

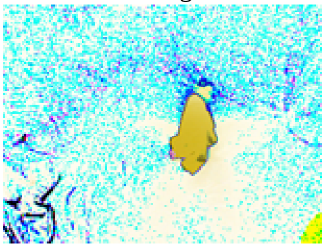
Fig 6 (a) main camera image (b) calibration board (c) Illumination camera images

Figure 6(c) shows the *illumination images* of *illumination camera* 1, 3, 5 and 7.

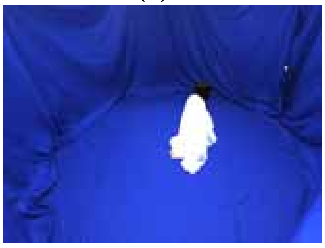
Then the illumination distribution of the 7 *illumination cameras* is estimated using image transfer matrixes as explained in Section III *E* and the formula for relating illumination radiance with image irradiance as explained in Section II.

The result of the illumination distribution is showed in table1

For testing the result of illumination distribution estimated, we tried to synthesize the people in the scene with estimated cast shadow, which from 7 sampled point light source. The superimposed result is shown as Figure 7.


(a)


(b)


(c)

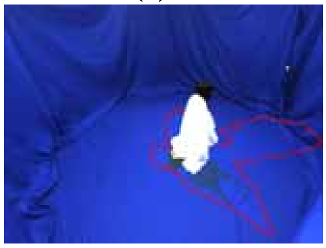

(d)

Fig 7 (a) Differential image (b) Differential image of synthesized image (c) Synthesized image without shadow (d) Synthesized image with estimated shadow

Comparing the synthesized image with the real image, we can find though the shadow cast from the estimated illumination is not as same as the real image, it is clear that (d) of Figure 7 is closer and more natural than (c) because of the shadow.

From the result of our experiment, we can convince that the method we presented is effective. Of course, estimating more complex illumination with the virtual silhouette images and synthesizing more natural image near to the real image are the future work we will effort in.

Table 1

L No.	B	G	R
1	15.023361	12.006111	8.955446
2	23.081326	16.936392	15.059221
3	69.342567	73.598373	55.343613
4	43.724865	44.949982	24.311668
5	-23.95113	-36.15731	-45.49013
6	72.010368	73.460594	64.788857
7	49.547203	53.528347	41.040432

V. CONCLUSION

In this paper, we have proposed an extended method for estimating an illumination distribution of a real scene. By establishing the relationship between the cast shadow of point light source and the silhouette image took by CCD camera, we could estimate an illumination distribution of a real scene with any arbitrary shape objects.

There is the beginning experiment on the proposed method. We only supposed the illumination condition is set with the same position as cameras. To simply real scene at present.

For the future work, we will use the multi-camera system to obtain the image of arbitrary position, and recover the complex illumination distribution environment of the real scene. Also overlay virtual objects with natural shadow with the estimated illumination condition in real time.

REFENCES

[1] Sato, I., Sato, Y., and Ikeuchi, K.: "Acquiring a Radiance Distribution to Superimpose Virtual Objects onto a Real Scene", *IEEE Transactions on Visualization and Computer Graphics*, Vol.5, NO.1, pp.1-12, March 1999.

[2] Sato, I., Sato, Y., and Ikeuchi, K.: "Illumination Distribution from Brightness in Shadows"

[3] R.T.Azuma, "A Survey of Augmented Reality", *Presence: Teleoperators and Virtual Environments*, Vol.6, No.4, pp.355-385, August 1997.

[4] R.T.Azuma and G.Bishop, "Improving static and dynamic registration in an optical see-through HMD", *Proc. SIGGRAPH 94*, pp.197-204, July, 1994.

[5] P.E.Debevec, "Rendering Synthetic Objects into Real Scenes: Bridging Traditional and Image-based Graphics with Global Illumination and High Dynamic Range Photography", *Proc. SIGGRAPH* 98, pp.189-198, July, 1998.

[6] G. Drettakis, L. Robert, S. Bougnoux, "Interactive Common Illumination for Computer Augmented Reality", *Proc. 8th Eurographics Workshop on Rendering*, pp.45-47, 1997.

[7] A. Fournier, A, Gunawan and C. Romanzin, "Common Illumination between Real and Computer Generated Scenes", *Pro. Graphics Interface 93*, pp.254-262, 1993.

[8] Zhang, Z.: F., "Flexible Camera Calibration by Viewing a Plane from Unknown Orientations", Proc. 7th International Conference on Computer Vision, Vol.1, pp.666-673 (1999)