Extended Multiple View Geometry for Lights and Cameras from Photometric and Geometric Constraints

Kazuki KATO Fumihiko SAKAUE Jun SATO
Department of Computer Science and Engineering
Nagoya Institute of Technology
Nagoya 466-8555, JAPAN
{kato@cv.nitech.ac.jp, sakae@nitech.ac.jp, junsato@nitech.ac.jp}

Abstract—In this paper, we derive a novel multilinear relationship for close light sources and cameras. In this multilinear relationship, image intensities and image point coordinates can be handled in a single framework. We first derive a linear representation of image intensity taken under a general close light source. We next analyze multiple view geometry among close light sources and cameras, and derive novel multilinear constraints among image intensity and image coordinates. In particular, we study the detail of the multilinear relationship among 7 lights and a camera. Finally, we show some experimental results, and show that the new multilinear relationship can be used for linearly generating images illuminated by arbitrary close light sources.

Keywords—multiple view geometry; photometry; geometry; close light source

I. INTRODUCTION

Multiple view geometry for multiple 2D cameras has been studied extensively in computer vision. The geometry provides very wide applications such as 3D reconstruction [1], new view synthesis and so on. Recently, extended multiple view geometry is proposed [2], which enables us to handle not only shape, but also motions. Furthermore, extended multiple view geometry for mixed dimensional cameras [3] realized a description of the relationship among cameras and many other sensors such as range sensors.

On the other hand, it is known that shading in images includes much information on 3D shapes and light sources [4], [5]. Many applications were proposed [6], [7], [8] based on the shading information. In particular, images taken under a close light source includes much more information than those taken under a light source at infinity, and are used for recovering 3D shapes in recent years [9], [10]. However, in order to handle images taken under close light sources, non-linear optimization should be solved, since the image intensities are represented by non-linear equations in this case.

In this paper, we introduce a new linear representation method of image intensities observed under close light source illumination. The key idea of our linear model is to represent image intensities in higher dimensional space. Based on the linear representation, we consider image shading information as well as image point coordinates in the multiple view geometry. This enables us to derive a novel multilinear constraints among image intensities and image point coordinates. The combination of geometric information and photometric information enables us to derive more general and efficient image representation. The proposed multilinear constraints can be used for transforming image intensities as well as image point coordinates, and can be used for generating new images illuminated by arbitrary close light sources.

II. LINEAR IMAGE REPRESENTATION UNDER CLOSE LIGHTING

A. Image Representation Under Close Lighting

We first discuss an illumination model of images taken under a close light source. Under the assumption of Lambertian surface, the intensity $I$ can be described by a light position $S = [s_x, s_y, s_z]^T$, 3-d surface point $X = [X, Y, Z]^T$ and the surface normal $n = [n_x, n_y, n_z]^T$ as follows:

$$I = \frac{1}{||S - X||^2} E \rho \max(n^T(S - X), 0)$$  (1)

where $E$ and $\rho$ denote irradiance and albedo respectively. $1/||S - X||^2$ indicates the attenuation of image intensity according to the distance between the surface point and the light source. It is called as attenuation term in this paper. The term $(S - X)/||S - b\rho X||$ indicates that the light orientation is different at each surface point. Thus, images taken under a close light source cannot be represented by a simple linear equation unlike images taken under an infinite light source.

B. Linear Image Representation Using Approximate Intensity Model

We next approximate Eq. (1) which describes image intensities as follows:

$$I = E \rho \max(n^T(S - X), 0) / ||S - X||^2.$$  (2)

We found, from our experiments, the approximation model is valid in most of the case except extremely close light,
and provides much better approximation than ordinary infinite light source model.

For deriving a linear image representation for close light sources, we define a new object point $\mathbf{W}$ in 8 dimensional space as follows:

$$
\mathbf{W} = \begin{bmatrix}
  X \\
  Y \\
  Z \\
  \rho_{n_x} \\
  \rho_{n_y} \\
  \rho_{n_z} \\
  \mathbf{n}^\top \mathbf{X} \\
  \mathbf{X}^\top \mathbf{X}
\end{bmatrix}^\top, \quad (3)
$$

where $X, Y, Z, n_x, n_y$ and $n_z$ are components of $\mathbf{X}$ and $\mathbf{n}$. Furthermore, in order to project the point $\mathbf{W}$ to an image intensity $i$, we newly define a light projection matrix $\mathbf{L}$ as follows:

$$
\mathbf{L} = \begin{bmatrix}
  0 & 0 & 0 & E s_x & E s_y \\
  -2s_x & -2s_y & -2s_z & 0 & 0 \\
  E s_z & -E & 0 & 0 & 0 \\
  0 & 0 & 1 & \mathbf{S}^\top \mathbf{S}
\end{bmatrix}, \quad (4)
$$

where $s_x, s_y$ and $s_z$ are components of $\mathbf{S}$. By using the point $\mathbf{W}$ and the matrix $\mathbf{L}$, an image intensity $\bar{I}$ in (2) can be represented linearly as follows:

$$
\lambda \begin{bmatrix}
  I \\
  1
\end{bmatrix} = \mathbf{L} \bar{\mathbf{W}} \quad (5)
$$

where $\lambda$ is a real scalar and $\bar{\mathbf{W}}$ is the homogeneous representation of $\mathbf{W}$ in (3). The equation (5) indicates that illumination model under a close light source can be represented linearly in 8 dimensional space. Furthermore, the equation indicates that image intensities are affected by not only surface normal $\mathbf{n}$, but also point position $\mathbf{X}$, unlike the case of infinite light source. In the following sections, we will show a relationship among lights and cameras by using the linear equation (5).

III. MULTIPLE VIEW GEOMETRY AMONG LIGHTS AND CAMERAS

The traditional multiple view geometry is defined for multiple 2D cameras. On the other hand, Kozuka et al.[3] showed that there exists a multiple view geometry for mixed dimensional cameras, i.e. cameras with different dimension. For example, a geometric relationship among 1D projective cameras, 2D projective cameras and 3D projective cameras in 4D space can be formulated. In this paper, we consider a relationship among cameras and lights by using the framework, and derive the multilinear relationship among image intensities and image point coordinates.

A. Extended camera projection matrix

In general, the projection from 3D space to 2D image can be described by a $3 \times 4$ camera projection matrix. However, in order to handle intensity projection and camera projection in the same framework, we in this paper extend the ordinary image projection to a projection from 8-dimensional space to 2-dimensional image as follows:

$$
\kappa \bar{x} = \begin{bmatrix}
  \mathbf{p}_1 \\
  \mathbf{p}_2 \\
  \mathbf{p}_3 \\
  0 & 0 & 0 & 0 & 0 & \mathbf{p}_4
\end{bmatrix} \bar{\mathbf{W}}, \quad (6)
$$

where $\kappa$ is a real scalar, $\mathbf{p}_i$ is the $i$-th column of the ordinary $3 \times 4$ camera projection matrix and $\bar{x} = [x^1, x^2, x^3]^\top$ is the homogeneous representation of an image point $\mathbf{x} = [x, y]^\top$. By using the new camera projection matrix, extended multiview geometry among close light source and camera can be formulated as shown in the next section.

B. Multiple View Geometry among Lights and Cameras

In this section, multilinear constraints among lights and cameras are introduced. Let $\mathbf{L}_i$ be a light projection matrix of the $i$-th light described in section II-B, and $\mathbf{P}_j$ be an extended camera projection matrix of the $j$-th camera defined in section III-A. The 8 dimensional point $\mathbf{W}$ is projected to an intensity $I$ and a point $\mathbf{x}$ as follows:

$$
\lambda_i \bar{I}_i = \mathbf{L}_i \bar{\mathbf{W}} \quad (7)
$$

$$
\kappa_j \bar{x}_j = \mathbf{P}_j \bar{\mathbf{W}} \quad (8)
$$

where, $\bar{I}_i = [I^1_i, I^2_i]^\top$ is the homogeneous representation of intensity $I$ illuminated by the $i$-th light. These equations can be rewritten as follows:

$$
\begin{bmatrix}
  \mathbf{L}_1 \\
  \vdots \\
  \mathbf{L}_{n_L} \\
  \mathbf{P}_1 \\
  \vdots \\
  \mathbf{P}_{n_p}
\end{bmatrix}
\begin{bmatrix}
  \bar{I}_1 \\
  \vdots \\
  \bar{I}_{n_L} \\
  \bar{x}_1 \\
  \vdots \\
  \bar{x}_{n_p}
\end{bmatrix} = \mathbf{M} \bar{\mathbf{W}} = \mathbf{0} \quad (9)
$$

where $n_p$ and $n_L$ are the number of cameras and lights. The submatrix (which size is $(9 + n_L + n_p) \times (9 + n_L + n_p)$) of $\mathbf{M}$ has null space, and its determinant vanishes, which represents multilinear constraints among lights and cameras.

The number of rows of the matrix $\mathbf{M}$ is $2n_L + 3n_p$ and the number of columns is $9 + n_L + n_p$. For deriving the multilinear constraints, the number of rows should be larger than the number of columns, and thus, the following inequality must hold:

$$
2n_L + 3n_p \geq 9 + n_L + n_p \quad (10)
$$

Furthermore, we should select more than 2 rows from each camera in order to derive valid constraints [1], and thus, the following inequality must also hold:

$$
9 + n_L + n_p \geq 2n_L + 2n_p \quad (11)
$$

Also, the number of lights must be larger than 6, since camera projection matrices has lots of 0 components and we cannot obtain sufficient constraints if the number of lights is less than 6. Therefore, the multiple view geometry among lights and cameras exists in the case of nine lights and no camera, eight lights and one camera, seven lights and one camera, six lights and two cameras, and six lights and three
cameras. In the following section, the multilinear constraints among seven lights and one camera will be discussed in detail.

C. Multiple View Geometry among 7 Lights and one Camera

In this section, we describe a multiple view geometry among seven lights and one camera. The multilinear constraints in this case are described by light projection matrices $L_i$ ($i = 1, \cdots, 7$) and a camera projection matrix $P_1$ as follows:

$$
\begin{bmatrix}
 L_1 & \tilde{I}_1 \\
 \vdots & \ddots \\
 L_7 & \tilde{I}_7 \\
 P_1 & \tilde{x}_1
\end{bmatrix}
\begin{bmatrix}
 \tilde{W} \\
 -\lambda_1 \\
 \vdots \\
 -\lambda_7 \\
 -\kappa_1
\end{bmatrix} = 0 \quad (12)
$$

This equation shows that a matrix in the left term has right null space and the determinant of the square submatrix of it is equal to 0. By using the tensor notation, the expansion of the determinant can be described as follows:

$$
\begin{aligned}
x^a I^b_1 I^c_2 I^d_3 I^e_4 I^f_5 I^g_6 I^h_7 C_{abcdefgh} &= 0 \quad (13)
\end{aligned}
$$

where $x^j$ is the $j$th component of $\tilde{x}$, and $I^j_i$ is the $j$th component of $\tilde{I}_i$ respectively. This equation represents a multilinear constraint for 7 lights and one camera. In this case, a multilinear tensor $C$ has 105 non-zero components, and the equation (13) provides us a single linear constraint for $C$. Therefore, the tensor $C$ can be estimated linearly from 104 corresponding points.

The constraint also indicates that we can estimate an image intensity from an image point, 6 intensities and the tensor $C$, as the image point transfer in three view geometry. This property enables us to estimate images taken under different light environment from multilinear tensor and 6 images.

Note, in this multiple view geometry, image points and intensities correspond to each other automatically, since any image points correspond to their own intensities in the image.

IV. EXPERIMENTAL RESULTS

In this section, experimental results from our proposed method are shown. In this experiment, we used three objects shown in Fig. 1. The size of these objects is about 10cm. The objects and a camera were placed as shown in Fig. 2. The point light source was moved manually around the objects. The distance from the objects to the light source is about 80cm. Seven images were taken under different light conditions, and two examples of them are shown in Fig. 2. In this experiment, image points and intensities on the cone and the sphere were used for estimating the tensor. The image intensities of the sculpture in one image were estimated from the tensor and its image intensities in the other six images.

For comparison, we also generated an intensity image by using the ordinary subspace method. In this method, a subspace was constructed from 6 images by singular value decomposition (SVD). The coefficients of three principal bases were estimated from the images of cone and sphere, and the image of sculpture was estimated from the coefficients.

Figure 3 shows the experimental results. Figure 3 (a) shows the ground truth, and (b) and (c) show images recovered from the proposed method and the subspace method respectively. The red pixels in the recovered images indicate negative intensities. Figure 3 (d) and (e) show errors in images recovered from the proposed method and the subspace method, where the color indicates the amount of errors as shown in the color bar. The average intensity errors, $d$, are also shown.

From these figures, we find that the proposed method provides us accurate estimates in most part of the image except the neck area. The error in the neck area is large, since the area has cast shadow, while the proposed method can not represent the cast shadow.

Also, from Fig. 3 (b) and (c), we find the proposed method can generate much better images than the subspace method.

Figure 1. Target objects

Figure 2. Experimental setup and example images

Figure 3. Experimental results
The image generated by the subspace method has errors in most part, since the subspace cannot represent images taken under a close light source, while the proposed method can represent it.

We next compare the intensity error of images generated from the proposed method and the subspace method by changing the relative distance between the light source and the object. The synthetic images are used in this experiment for estimating intensity errors quantitatively. Figure 4 shows the relationship between the intensity error and the relative distance of light. The red line shows the results from the proposed method, and the blue line shows those from the subspace method. Although the accuracy of both methods improves as the distance of light becomes large, the proposed method provides us much better accuracy when the distance of light is small.

These results indicate that the multiple view geometry among close light sources and cameras proposed in this paper is quite useful in image generation under arbitrary light sources.

V. CONCLUSIONS

In this paper, we proposed a novel extended multiple view geometry among close light sources and cameras. We showed that the relationship among image intensities and image points can be described by multilinear constraints in 8D space. In order to derive the multilinear constraints, we introduced a linear representation of image intensities taken under a close light source. Some experimental results indicate that the proposed multilinear constraints can describe the relationship among lights and cameras. The newly defined relationship on lights and cameras can be applied to many computer vision problems, such as shape from shading under close light sources.

REFERENCES


