Information and Media Technologies 8(4) : 1095–1099 (2013)
reprinted from: IPSJ Transactions on Computer Vision and Applications 5: 104–108 (2013)


**Express Paper**

# Refining Outdoor Photometric Stereo Based on Sky Model

Kenji Inose[1,a)] Shota Shimizu[2,b)] Rei Kawakami[3,c)]
Yasuhiro Mukaigawa[3,d)] Katsushi Ikeuchi[4,e)]

Received: March 11, 2013, Accepted: April 24, 2013, Released: July 29, 2013

***Abstract:*** We propose an outdoor photometric stereo method, which considers environmental lighting for improving the performance of surface normal estimation. In the previous methods, the sky illumination effect has been either assumed to be constant throughout the scene, or to be removed by pre-processing. This paper exploits a sky model which can derive the entire sky luminance from a sky zenith image; then, sky illumination effect on a surface can be correctly calculated by iteratively refining its normal direction. This paper also extends a two-source photometric stereo method by introducing RANSAC, so that the input images of this method can be taken in a day. Experimental results with real outdoor objects show the effectiveness of the method.

***Keywords:*** outdoor photometric stereo, sky model, 3D reconstruction

## 1. Introduction

Acquiring three-dimensional shape information of small objects, buildings, up to the entire city has attracted much interest as it can be utilized for entertainment, navigation, driving simulation, and heritage preservation. Laser range finders and photogrammetry methods, such as multi-view stereo methods [6], are widely used to obtain accurate shape information. However, the resolution of the three-dimensional points of those methods is not as high as that of two-dimensional photographs, since the lasers need to sweep the target space, or the stereo methods recover structures based on feature points which are not necessarily densely observable in the target scene. Also, their surface normal information is relatively unreliable, due to the sensor noise or lower resolution.

Photometric stereo methods, on the other hand, can be a solution to reconstruct dense and accurate surface normals. The basic idea of the photometric stereo is that surface normals can be estimated by observing how the shading of a surface varies under a point light source located at different positions [17]. The early photometric stereo methods assume that the surface is a Lambertian surface, and the light source is located at more than three different positions that are not coplanar to each other. Many methods have been proposed to relax those assumptions, but they are generally applicable to indoor objects where the illumination is easier to control. Handling large-scale outdoor objects by the photometric stereo methods is still challenging.

To tackle the outdoor photometric stereo, Ackermann et al. [3] propose a photometric stereo method that can handle general lighting, by extending the example-based method [9] combined with the multi-view stereo method [7]. They successfully recover surface normals of outdoor objects. Shen et al. [15] model the outdoor illumination with the spherical harmonics, and combine the multi-view stereo method to have rough estimate of surface normal from collection of images from the Internet, so that the method can estimate the weather of each image. Haber et al. [8] propose a method that uses multi-view stereo method and wavelet basis, to recover illumination environment and surface reflectance from the photo collections on the Internet. Those methods can handle general lighting, but tend to fail when the illumination cannot be modeled by the basis (e.g., a clear sky with strong sunlight).

There are several methods that explicitly model the outdoor illumination with the sun and the sky illumination. Ackermann et al. [2] propose a method which estimates surface normal as well as camera response function, and the bases of materials by separating image intensity into sun and sky contributions. Abrams et al. [1] extend Ackermann et al.'s method to be computationally efficient. However, those methods assume that the sky illumination effect is constant throughout the scene. Furthermore, all of the methods mentioned above assume that the light sources are not coplanar, which means that the video or images need to be collected from different times of the year, so that the sun positions are sufficiently different. Sato et al. [14] successfully realize outdoor photometric stereo method that uses images acquired only

1 Graduate School of Interdisciplinary Information Studies, The University of Tokyo, Bunkyo, Tokyo 113–0033, Japan
2 Graduate School of Information Science and Technology, The University of Tokyo, Bunkyo, Tokyo 113–8656, Japan
3 The Institute of Scientific and Industrial Research, Osaka University, Ibaraki, Osaka 567–0047, Japan
4 Interfaculty Initiative in Information Studies, The University of Tokyo, Bunkyo, Tokyo 113–0033, Japan
a) inoken@cvl.iis.u-tokyo.ac.jp
b) shimizu@cvl.iis.u-tokyo.ac.jp
c) rei@am.sanken.osaka-u.ac.jp
d) mukaigaw@am.sanken.osaka-u.ac.jp
e) ki@cvl.iis.u-tokyo.ac.jp

Information and Media Technologies 8(4) : 1095–1099 (2013)
reprinted from: IPSJ Transactions on Computer Vision and Applications 5: 104–108 (2013)


in a day. However, the method requires the surface to be smooth, and the number of materials of the object to be few.

To overcome the problems, this paper proposes a method which models the sun and the sky illumination by using a sky model [13]. In order to calculate sky illumination effect on a surface correctly, surface normal itself has to be known beforehand. To tackle this, we iteratively estimate the surface normal and refine its corresponding sky effect. Furthermore, the proposed method extends the two-source photometric stereo method [18] by introducing RANSAC [4], so that the estimation becomes more robust. The method enables us to obtain surface normal of an object with higher accuracy, from an image set taken in a day.

The contribution of this paper is twofold: First, by utilizing a sky model, we refine the surface normal estimation of outdoor photometric stereo. Second, we extend Zhang et al.'s two light photometric stereo [18] by combining RANSAC, and propose an overall framework so that the method can be applied to an image set that is taken in a day.

The rest of this paper is structured as follows. Section 2 describes the sky model used in this paper, and a method to calculate sky illumination effect. Section 3 provides a photometric stereo method introducing the refinement of surface normal with a sky model and RANSAC. Section 4 shows experiments with real data. Finally we present conclusions and suggestions for future work.

## 2. Sky Illumination Effect

**Sky Model** We will use the sky model proposed by Preetham et al. [13] in the field of computer graphics. The model is an extension of Perez et al.'s [12] which is proposed in atmospheric science to simulate sky luminance distribution. The fraction of luminance $Y$ of the sky in CIE $xyY$ space is calculated from the sun's zenith and azimuth angles and a parameter turbidity $T$ which expresses the haziness of the sky. Namely,

$$\frac{Y}{Y_{ref}} = \frac{F(\theta, \gamma, T, \theta_s, \phi_s)}{F(\theta_{ref}, \gamma_{ref}, T, \theta_s, \phi_s)}, \tag{1}$$

where $Y_{ref}$ is the luminance of a reference point, and F(.,.) is a function proposed by Perez et al. with the extension of Preetham et al. [13]. $\theta$ is the zenith angle of the view direction, and $\gamma$ is the angle between the view direction and the sun direction. $\theta_s$ and $\phi_s$ are the sun's zenith and azimuth angles. **Figure 1** shows the notations. The reference point can be anywhere, but we usually use the zenith of the sky. Using the model, turbidity can be estimated by fitting the model to the brightness of the sky in the image. Namely, we can estimate turbidity $T$ as follows:

$$T_{est} = \arg\min_{T} \sum_{i=1}^{n} \left| \frac{Y_i(T)}{Y_{ref}(T)} - \frac{J_i}{J_{ref}} \right|, \tag{2}$$

where $n$ is the number of the sampled pixels, and $J$ is the image brightness calculated by taking the sum of each channel in an RGB image.

To estimate the sun direction, we assume that the date and time of the input image is available, as well as the geo-location of the target object and the camera pose. Note that, in our method, we estimate turbidity from an auxiliary fish-eye image of the sky taken along with the input images for photometric stereo, and assume that the turbidity is constant in a day. However, it is possible to directly use the partial sky captured in the images for photometric stereo, and readers can refer to Ref. [19] for further details.

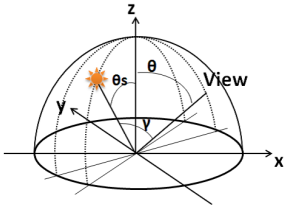


**Fig. 1** The coordinates for specifying the sun position and the viewing direction in the sky hemisphere.

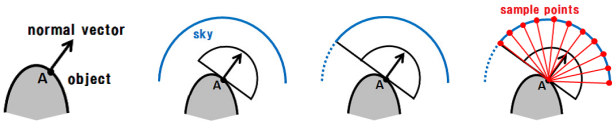


**Fig. 2** Calculating influence of the skylight.

**Sky Effect on a Surface** In order to incorporate sky illumination effect to photometric stereo, surface normal of the point in interest needs to be known beforehand. This is illustrated in **Fig. 2**. Let us assume that the surface normal of a point A is known as in Fig. 2 (a). Then, all the rays coming from the hemisphere above the point A needs to be considered to calculate the effect of sky illumination (irradiance from the sky) as in Fig. 2 (b). Specifically, as in Fig. 2 (c), we calculate the tangent plane of the surface normal, and define the range of the sky need to be considered. The angles that is lower than the horizontal plane are also excluded from the range. Then, we integrate the radiance of the sky in the defined range by sampling the points as in Fig. 2 (d) to calculate sky illumination effect on the surface. Figure 2 is illustrated in two dimension, but the process is actually done in three dimension. Geodesic dome is used to partition the hemisphere equally. How to obtain the initial surface normal is explained in the next section.

## 3. Refinement of Outdoor Photometric Stereo with Sky Illuminaition

### 3.1 Two-source Photometric Stereo with RANSAC

We extend Zhang et al.'s two-source photometric stereo by introducing RANSAC. Zhang et al.'s method takes two images and initial normals as inputs while general photometric stereo requires three images. Initial normals are obtained by multi-view stereo also in this paper.

In our method, we exploit many pairs of images that are taken under coplanar light sources. Let us assume that $k$ images are available. First, we generate ${}_kC_2$ candidates of the surface normal and $N$ random vectors. Then, we find the vector from the randomly generated vectors that has the most inliers. Here, inliers are defined as those whose dot product is bigger than a certain threshold. Having found the vector that has most inliers, we estimate the correct surface normal as the mean of the inliers. The algorithm is also shown in **Fig. 3**. To be efficient, we exclude

Information and Media Technologies 8(4) : 1095–1099 (2013)
reprinted from: IPSJ Transactions on Computer Vision and Applications 5: 104–108 (2013)


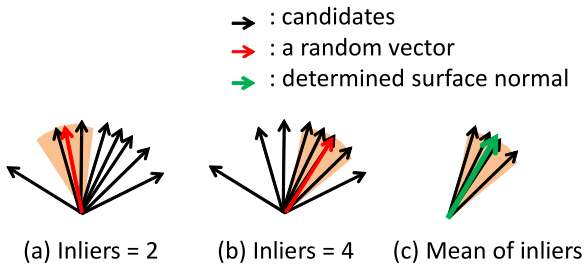


**Fig. 3** Computing surface normal by a RANSAC method [10]. As shown in (a) and (b), we generate random vectors and then find the vector of them which has most inliers. Surface normal is then determined as the mean vector of the inliers.

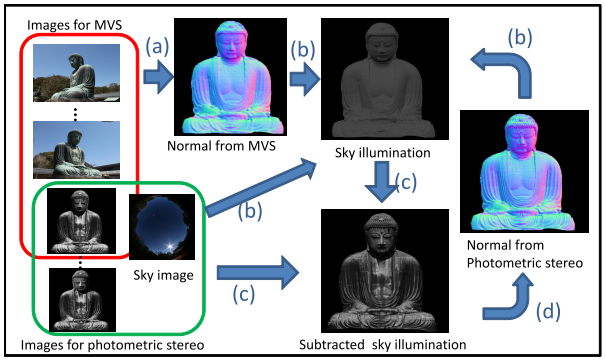


**Fig. 4** The outline of our method. (a) An initial normal map is obtained by multi-view stereo. (b) Sky illumination is rendered by a normal map and a sky model after turbidity of the scene is estimated. (c) The sky illumination effect is subtracted from the images for photometric stereo. (d) A normal map is estimated by photometric stereo.

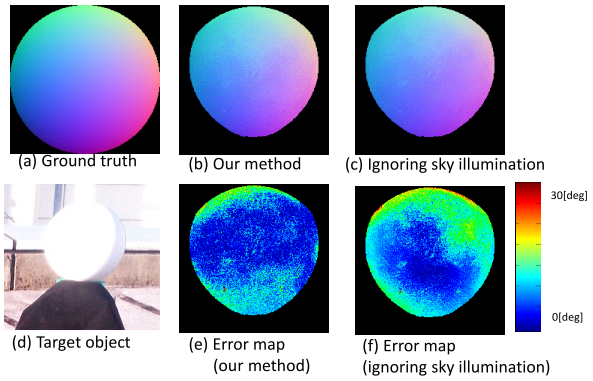


**Fig. 5** Result of a white matte sphere. The shadowed pixels were manually excluded for photometric stereo.

**Fig. 6** An input image of the Kamakura Buddha.

**Table 1** Mean angular error of the estimated surface normal of the white sphere.

| | Mean Error (in degree) |
|---|---|
| our method | 7.27 |
| ignoring sky illumination | 9.79 |

the pairs whose light directions are closer than a threshold, since otherwise the computed solution space will be sensitive to errors.

RANSAC type of estimation is useful for photometric stereo [10], since the target object may not be necessarily a Lambertian surface, and therefore might include specular refleciton. To be self-contained, Zhang et al.'s method is briefly explained in Appendix.

### 3.2 Iterative Estimation

We take an iterative approach to estimate sky illumination and surface normal. **Figure 4** shows the overview of our method. As a pre-process, we reconstruct an initial normal by multi-view stereo method. The turbidity is estimated from the auxiliary sky image and the whole sky is recovered from the sun direction and the estimated turbidity.

In the first iteration, the sky illumination effect on each surface is calculated from the initial surface normal, and then subtracted from the input images. The photometric stereo with RANSAC is then performed to the images. This will be repeated until the estimates are converged. Empirically, the estimates converge in a few iterations.

## 4. Results

We validated our method using real world data: a white matte sphere shown in **Fig. 5** (d) and the Buddha in Kamakura shown in **Fig. 6**. To estimate the turbidity and render the sky illumination, auxiliary sky images were taken with a fish-eye lens mounted on a Canon 5D. Other images for photometric stereo and multi-view stereo were taken with a Canon 5D Mark2. Gamma correction was turned off to obtain images that are proportional to incoming light intensity.

### 4.1 White Sphere

To quantify the error of normal estimation caused by sky illumination, we used a basic photometric stereo (Eq. (A.2)) instead of the method described in Section 3 and chose a white sphere as a target whose shape is known. We captured three pairs of the target and sky zenith images under different sun directions.

Figure 5 (b) and (c) show the estimated normal maps, with and without considering the sky illumination effect, respectively. Figure 5 (e) and (f) show the angular error maps w.r.t. the ground truth shown in Fig. 5 (a). The error maps show that the error of the surface normal can be decreased by considering the sky illumination effect. Particularly, the accuracy can be improved in the upper region where the sky illumination effect is significant. In such area, the improvement of normal estimation is around 20 degrees. **Table 1** shows the mean angular error in degrees.

### 4.2 The Buddha in Kamakura

We conducted another experiment to evaluate the proposed method, using images of the great Buddha in Kamakura. We took around 60 images for photometric stereo from 11:00 am to 12:30 pm and 20 images for multi-view stereo at a particular time with a Canon 5D Mark2. For Structure from Motion (SfM) followed by the multi-view stereo, the captured images at the full resolution were used. We then crop the region of interest containing the object, yielding images of around 2,800 by 2,800 pixels for photometric stereo.

Information and Media Technologies 8(4) : 1095–1099 (2013)
reprinted from: IPSJ Transactions on Computer Vision and Applications 5: 104–108 (2013)


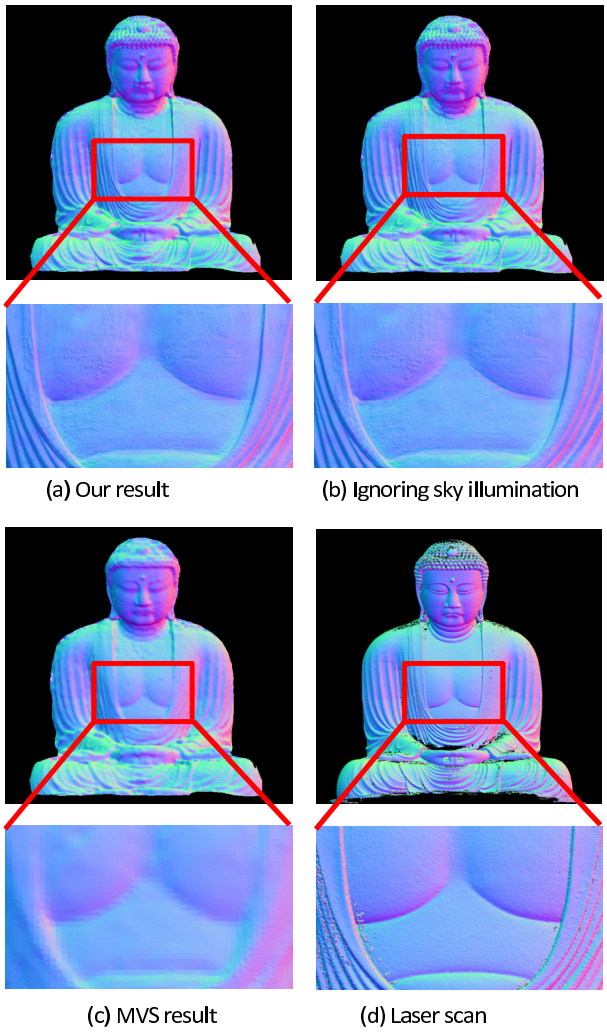


**Fig. 7** Reconstructed surface normal of the Buddha. Our result (a) is more detailed than the MVS result (c).

A sky zenith image was captured with a fish-eye lens to estimate turbidity. The sun direction was computed from known location and timestamps for each image. The turbidity for the sky model was estimated to be 1.88. The radiance of the sun and sky was then computed from the sky model.

We used the Bundler software by Snavely et al. [16], Patch based multi-view stereo (PMVS) [5] and Poisson Surface Reconstruction (PSR) [11] to generate the initial surface normal. In such a way, a mesh which has 70,000 vertices was obtained. From the initial normal map which was computed by the mesh, surface normals were estimated by the proposed method. **Figure 7** shows the results of the proposed method.

Figure 7 shows that the proposed method is able to estimate detailed surface normals as accurate as the one obtained from a laser scanner. However, there are inaccurate normals in shadowed regions since the method currently does not take the shadow effect into account.

**Figure 8** shows the normal vectors on one scanline of the nose. In Fig. 8, normal vectors of our result are closer to those of laser scanner, compared to those when sky illumination is ignored.

## 5. Conclusion and Future Work

In this paper, we have presented a method to improve the accuracy of outdoor photometric stereo by subtracting sky illumination effect from images. The method utilizes a sky model to estimate sky illumination and surface normals effectively. We also extend a two-source photometric stereo by introducing RANSAC, so that the input images of the method can be taken in a day. The method would be practical to be applied to objects in outdoors.

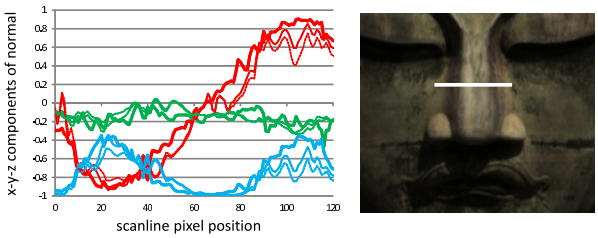


**Fig. 8** Left: The normal vectors on one scanline at the nose. The x, y and z components are indicated in red, green and blue respectively. The bold solid lines, the dotted lines and the thin solid lines, respectively, show results of laser scan, photometric stereo ignoring sky illumination, and our method. Right: The scanline marked in white.

Currently, we ignore specular reflection, inter-reflection, ambient occlusion and shadow effect. To estimate more accurate surface normals, we will try to account for those effects and relax the constraint that the reflectance model has to be Lambertian. Another issue is how to model other illumination effect such as reflections from clouds and ground.

**Acknowledgments** This research is, in part, granted by the Japan Society for the Promotion of Science (JSPS) through the "Funding Program for Next Generation World-Leading Researchers (NEXT Program)," initiated by the Council for Science and Technology Policy (CSTP), and by SCOPE (Strategic Information and Communications R&D Promotion Program) of Ministry of Internal Affairs and Communications.

### References

[1] Abrams, A., Hawley, C. and Pless, R.: Heliometric Stereo: Shape From Sun Position, *European Conference on Computer Vision* (*ECCV*), pp.357–370 (2012).
[2] Ackermann, J., Langguth, F., Fuhrmann, S. and Goesele, M.: Photometric stereo for outdoor webcams, *CVPR*, pp.262–269 (2012).
[3] Ackermann, J., Ritz, M. and Goesele, M.: Removing the Example from Example-based Photometric Stereo, *European Conference on Computer Vision* (*ECCV*), pp.197–210 (2010).
[4] Fischler, M.A. and Bolles, R.C.: Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography, *Commun. ACM*, Vol.24, No.6, pp.381–395 (1981).
[5] Furukawa, Y. and Ponce, J.: Accurate, Dense, and Robust Multi-View Stereopsis, *IEEE Trans. Pattern Anal. Mach. Intell.*, Vol.32, No.8, pp.1362–1376 (2010).
[6] Goesele, M., Curless, B. and Seitz, S.M.: Multi-View Stereo Revisited, *CVPR*, CVPR '06, pp.2402–2409, Washington, DC, USA, IEEE Computer Society (2006).
[7] Goesele, M., Snavely, N., Curless, B., Hoppe, H. and Seitz, S.M.: Multi-View Stereo for Community Photo Collections, *ICCV*, pp.1–8 (2007).
[8] Haber, T., Fuchs, C., Bekaert, P., Seidel, H.-P., Goesele, M. and Lensch, H.P.A.: Relighting Objects from Image Collections, *CVPR*, pp.627–634 (2009).
[9] Hertzmann, A. and Seitz, S.M.: Example-Based Photometric Stereo: Shape Reconstruction with General, Varying BRDFs, *IEEE Trans. Pattern Anal. Mach. Intell.*, Vol.27, No.8, pp.1254–1264 (2005).
[10] Higo, T., Miyazaki, D. and Ikeuchi, K.: Multi-View Photometric Stereo using Rough Shape Data, *MIRU*, pp.1280–1287 (2008).
[11] Kazhdan, M., Bolitho, M. and Hoppe, H.: Poisson surface reconstruction, *Proc. 4th Eurographics Symposium on Geometry Processing*, SGP '06, pp.61–70, Aire-la-Ville, Switzerland, Eurographics Association (2006).

Information and Media Technologies 8(4) : 1095–1099 (2013)
reprinted from: IPSJ Transactions on Computer Vision and Applications 5: 104–108 (2013)


[12] Perez, R., Seals, R. and Michalsky, J.: All-weather model for sky luminance distribution – Preliminary configuration and validation, *Solar Energy*, Vol.50, No.3, pp.235–245 (1993).
[13] Preetham, A.J., Shirley, P. and Smits, B.: A practical analytic model for daylight, *Proc. 26th Annual Conference on Computer Graphics and Interactive Techniques*, SIGGRAPH '99, pp.91–100, New York, NY, USA, ACM Press/Addison-Wesley Publishing Co. (1999).
[14] Sato, Y. and Ikeuchi, K.: Reflectance analysis under solar illumination, *Physics-Based Modeling in Computer Vision*, pp.180–187 (1995).
[15] Shen, L. and Tan, P.: Photometric stereo and weather estimation using internet images, *CVPR*, pp.1850–1857 (2009).
[16] Snavely, N., Seitz, S.M. and Szeliski, R.: Photo tourism: Exploring photo collections in 3D, *SIGGRAPH Conference Proceedings*, pp.835–846, New York, NY, USA, ACM Press (2006).
[17] Woodham, R.J.: Photometric method for determining surface orientation from multiple images, *Optical Engineering*, Vol.19, No.1, pp.139–144 (1980).
[18] Zhang, Q., Ye, M., Yang, R., Matsushita, Y., Wilburn, B. and Yu, H.: Edge-preserving photometric stereo via depth fusion, *CVPR*, pp.2472–2479 (2012).
[19] Zhao, H.: Estimation of Atmospheric Turbidity from a Sky Image and Its Applications, PhD Thesis, The Graduate School of the University of Tokyo (2012).

## Appendix

Assuming Lambertian surface, observed image intensity $i$ is written as follows:

$$\rho \boldsymbol{s} \cdot \boldsymbol{n} = i, \tag{A.1}$$

where $\boldsymbol{n}$ is surface normal, $\boldsymbol{s}$ is light source direction, and $\rho$ is albedo. When three or more image intensities taken under different illuminant directions are observed, Eq. (A.1) can be rewritten in a matrix form

$$\boldsymbol{S}\boldsymbol{n} = I, \tag{A.2}$$

where $I$ is the image data matrix and $\boldsymbol{S}$ is the light vector matrix. While surface normal can be computed by solving this linear system at each point, this equation is however unsolvable when the rank of $\boldsymbol{S}$ is lower than 3.

In two light conditions, solution space of surface normal is expressed as a plane through origin and an interline of $\Pi_1$ and $\Pi_2$ (see **Fig. A·1**). Zhang et al. [18] compute surface normal by projecting an initial normal, which is measured by a depth sensor, onto the solution plane.

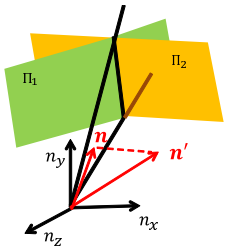


**Fig. A·1** The illustration of two-source photometric stereo by Zhang et al. [18]. $\Pi_1$ and $\Pi_2$ are plane functions of surface normal expressed as Eq. (A.1). Note that however the albedo changes, the interline of $\Pi_1$ and $\Pi_2$ is located on the unique solution plane.

(Communicated by *Hajime Nagahara*)