NEAR LIGHT SOURCE LOCATION ESTIMATION USING ILLUMINATION OF A DIFFUSED PLANAR BACKGROUND

Nopporn Chotikakamthorn

Faculty of Information Technology
King Mongkut's Institute of Technology Ladkrabang
Bangkok, Thailand

ABSTRACT

The problem of light source location estimation is considered. It is shown that the location of a near light source can be estimated from an optical-depth image pair using information available from an illuminated Lambertian planar background. The method estimates the projected source location on the planar background from the surface illumination gradient. The distance of the light source from the planar background, which is equivalent to its elevation angle, is estimated by fitting the radiance of the background surface as observed by an optical image, with those synthesized at different light distances. The fitting equation is formulated such that the possible existence of environment light can be taken into account. Experimental results with real images are provided.

Index Terms— Light source estimation, Computer vision, Augmented reality, Shape from X

1. INTRODUCTION

Light source location estimation from 2D images is an important problem for various computer vision and computer graphics applications, such as shape from shading, object recognition, image-based rendering and augmented reality. For example, knowledge of light source positions can be used in augmented reality applications to provide consistent lighting effects for synthetic objects [1]. Most existing light source estimation methods are based on the distant light source assumption [2-6]. These methods differ in what illumination cues are utilized for the estimation (see Fig. 1). Most of them utilize either scene object surface shading [2-3], or shadow information [4-7]. Among the methods based on shading information, they can be further categorized into those relying on specular reflection (e.g., [2]), and those using diffused reflection (e.g. [4]). Some methods such as [7-8], however, make use of both types of reflections.

Under some lighting environment, such as an indoor scene, near light sources are commonplace, where few existing methods are applicable. In [7], by using a spherical object and multiple-view images, specular and diffused reflection are exploited for light source location estimation. By combining information from specular and diffused reflections, a single-view solution that is capable of estimating a single near light source is described in [8]. With two identical spheres, a method in [3] can deal with multiple near light sources, as well as directional light sources. These methods require placement of a calibration spherical object(s). Another approach uses a fish-eye camera [9]. Such method, however, requires an additional specialized camera as well as careful camera placement and calibration.

In this paper, a method is developed for the estimation of a near light source. The method is applied to an optical-depth image pair, or an RGB-D image. Such image becomes increasingly available, with the arrival of some consumer depth and stereoscopic cameras. In addition to the depth data, the method utilizes both shading and shadow information available in an optical image. By analyzing light reflected from a background planar surface, a near light source location can be estimated. It assumes that the flat surface has a Lambertian reflection property. A prior knowledge on the geometry of the scene object(s) is not required. Instead, the method makes use of the available depth data directly without first reconstructing scene object models. As for the technique in [5], by exploiting cast shadow information, the method proposed here allows for the existence of environment (ambient and/or directional) lights, in addition to a near light source.

Previous works on the use of cast shadow for light source estimation include those of [4] and [5]. Both consider the case of directional light sources. In [4], a method that uses both self and cast shadows for directional light source estimation is described. Detection of a self-shadow is performed by processing of shading on a calibration sphere. The method described here uses gradient extracted from shading on a flat surface and is applicable for a near light source. A method that exploits information on a cast shadow on a flat surface has been proposed before in [5]. However, such method is only applicable for directional lights. The method in [10] utilizes information captured by an RGB-D camera for light source estimation. It is based on the method in [5] and thus limited to a directional light source case.

The paper is organized as follow: in Section II, problem formulation and a methodology used are first described. Key finding regarding to relationship between the planar surface illumination and the near light source position is provided in Section III. Experimental result is given in section IV. Section V gives a conclusion and points of possible improvement.
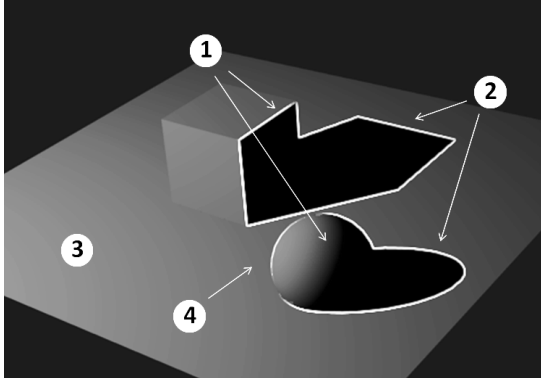

Fig. 1. Illumination cues used for computer vision problems such as shape from shading and light source estimation: 1) attached (self) shadow, 2) cast shadow, 3) shading, 4) occluding contour.

2. PROBLEM DESCRIPTION

The problem considered here is to identify the location of a single near light source, from an optical-depth image pair of illuminated scene objects with a planar surface in the background. The following assumptions are made.

A-1 Scene illumination is due to a near point light source with uniform luminosity, and possibly one or more directional as well as ambient light sources.
A-2 Scene objects are located in front of (or lie above) a planar surface (e.g., a floor or a wall). It is modeled as a Lambertian surface with constant albedo.
A-3 At least some shadow of scene objects cast on the planar surface must appear in the captured images.

In this paper, the world coordinate is aligned with the planar surface $\mathcal{P}$ (see Fig. 2). The origin is the nearest point on $\mathcal{P}$ to the camera center. From Fig. 2, at any point P on $\mathcal{P}$, its normal vector is denoted by $\mathbf{n}_{\mathcal{P}}$. Let L_0 denote the near point light source. In addition, L_k $(i = 1, \ldots, K)$ and L_a are the kth directional light source and the ambient light respectively. The point light source location with respect to P is generally defined by $(\theta_{p,0}, \varphi_{p,0}, d_{p,0})$, where $d_{p,0}$ is its distance from P. Alternatively, the source position can be given by $(\mathbf{l}'_0, d_{0,\perp})$, where $\mathbf{l}'_0$ is the location of the projected point L'_0 of L_0 on $\mathcal{P}$, and $d_{0,\perp}$ is the distance from $\mathcal{P}$ of the near light source.

From the pair of optical and depth images (e.g., as observed by an RGB-D camera), the problem is to find the near light source position $\mathbf{l}_0$ through the estimation of $(\mathbf{l}'_0, d_{0,\perp})$. Although a near point light source position is of major interest here, in practice the existence of other light sources is unavoidable and must be taken into account. For example, ambient light always exists due to reflection of light from the main source by surrounding environment. Here, factors due to directional and ambient lights are taken into account by incorporating the method in [5] into the formulation.
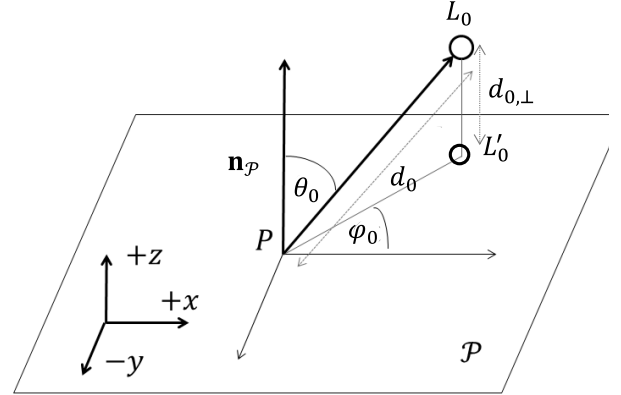

Fig. 2. The coordinate system and the near light source position with respect to a point P on $\mathcal{P}$.

3. NEAR LIGHT SOURCE LOCATION ESTIMATION FROM PLANAR SURFACE ILLUMINATION

3.1. Projected Light Source Position Estimation from Illumination Gradient on a Planar Surface

From Fig. 2, to obtain the light source position, consider first the estimation of $\mathbf{l}'_0$. It can be shown that, an illumination gradient $\mathbf{g}(x, y)$ of the light reflected at any point $[x \quad y \quad 0]^T$ on $\mathcal{P}$ points toward L'_0. In practice, $\mathbf{g}(x, y)$ can be obtained by analyzing scene illumination from an optical image. Let the image-space gradient $\mathbf{f}(m, n)$ be the gradient of image irradiance as observed by an optical image. Under the assumptions that each of the three axes of $\mathcal{P}$ is parallel to that of the camera, and that the image is formed by orthographic projection, $\mathbf{f}(m, n)$ can be used in place of $\mathbf{g}(x, y)$. To deal with the case where any of the two assumptions is invalid, first define $\mathbf{p}(m, n)$ as a world-space point corresponding to its projected point (m, n) on the optical image. Next, define

$$\mathbf{p}_{\mathrm{u}}(m, n) = \mathbf{p}(m, n + 1) - \mathbf{p}(m, n) \qquad (1)$$
$$\mathbf{p}_{\mathrm{v}}(m, n) = \mathbf{p}(m - 1, n) - \mathbf{p}(m, n) \qquad (2)$$

Let $p_{\mathrm{u}}(m, n)$ and $p_{\mathrm{v}}(m, n)$ be the Euclidean norm of $\mathbf{p}_{\mathrm{u}}(m, n)$ and $\mathbf{p}_{\mathrm{v}}(m, n)$ respectively. In addition, let $\omega_{m,n}$ be

the angle between $\mathbf{p}_u(m,n)$ and $\mathbf{p}_v(m,n)$, and let $f_u(m,n)$ and $f_v(m,n)$ be the two scalar components of $\mathbf{f}(m,n)$. The gradient $\mathbf{g}(m,n)$ (as a function of the image-space point (m,n) on the optical image) is obtained from

$$\mathbf{g}(m,n) \ = \ a_{m,n}\,\mathbf{p}_v(m,n) \ + \ b_{m,n}\,\mathbf{p}_v\!\perp(m,n) \qquad (3)$$

where

$$a_{m,n} \ = \ f_v(m,n)/p_v(m,n) \qquad (4)$$

$$b_{m,n} = \{p_v(m,n)f_u(m,n) - p_u(m,n)f_v(m,n)\cos(\theta_{m,n})\}/c_{m,n} \quad (5)$$

$$c_{m,n} \ = \ p_v(m,n)p_u(m,n)\sin(\theta_{m,n}) \qquad (6)$$

$$\mathbf{p}_v\!\perp(m,n) \ = \ \frac{1}{p_u(m,n)\sin(\theta_{m,n})}\{\mathbf{p}_u(m,n) - \frac{p_u(m,n)\cos(\theta_{m,n})}{p_v(m,n)}\,\mathbf{p}_v(m,n)\} \quad (7)$$

Theoretically, with the minimum of two gradient vectors corresponding to two distinct points on $\mathcal{P}$, $\mathbf{l}_0'$ can be obtained as an intersection point of the two vectors by solving a standard linear equation. In practice, inaccuracy in light reflection measurement (as obtained by an optical image) and depth measurement (as obtained by a depth image) makes the estimate based on this minimum amount of data highly unreliable.

With the following result, a more reliable estimate of $\mathbf{l}_0'$ can be obtained by making use of more data available in the optical and depth images.

Proposition 1: Let $\mathcal{C}$ be a set of points on $\mathcal{P}$ which form a closed circular area. The average vector of $\mathbf{g}(x,y)$, $(x,y) \in \mathcal{C}$, points to the same direction as that of $\mathbf{g}(x_c,y_c)$, where (x_c,y_c) is the center of $\mathcal{C}$.

Proof: Let (x,y) be any point in $\mathcal{C}$. Due to the nature of the shape of $\mathcal{C}$, apart from the point (x_c,y_c), there can always be the point $(x',y')\,) \in \mathcal{C}$ which is symmetrical to (x,y) with respect to the axis formed by the line passing through (x_c,y_c) and the point L_0'. By using a standard result in geometry, and the fact that all gradient vectors point to L_0', $\mathbf{g}(x,y) + \mathbf{g}(x',y') = \alpha\,\mathbf{g}(x_c,y_c)$, where α is a positive scalar factor. Therefore, the average of all gradient vectors corresponding to $\mathcal{C}$ yields $\mathbf{g}(x_c,y_c)$ up to a scaling factor. The proposition is thus proved.

Based on the above result, once the estimates of the gradient vectors $\mathbf{g}(x_{c1},y_{c1})$ and $\mathbf{g}(x_{c2},y_{c2})$ at two distinct points on $\mathcal{P}$ by means of gradient averaging, the position $\mathbf{l}_0'$ of L_0' can be calculated by solving the following linear equation

$$\alpha_1\mathbf{g}(x_{c1},y_{c1}) + \mathbf{p}_{c1} \ = \ \alpha_2\mathbf{g}(x_{c1},y_{c1}) + \mathbf{p}_{c2} \quad (8)$$

where $\mathbf{p}_{c1}$ and $\mathbf{p}_{c2}$ are the center positions of the two circle areas. Once α_1 or α_2 is obtained from the above equation, $\mathbf{l}_0'$ is estimated as

$$\mathbf{l}_0' \ = \ \alpha_1\mathbf{g}(x_{c1},y_{c1}) + \mathbf{p}_{c1}, \qquad (9)$$

which is the point where $\mathbf{g}(x_{c1},y_{c1})$ and $\mathbf{g}(x_{c2},y_{c2})$ intersect.

The above result can be applied under the presence of one or more directional light sources, and the ambient light term (except for the surface points near or on a shadow boundary). This is due to the fact that, on a planar surface, an illumination gradient attributed to directional or ambient light is zero.

3.2. Estimating Light Source Distance from Planar Surface Radiance

To complete near light source location estimation, the distance $d_{0,\perp}$ from $\mathcal{P}$, in addition to $\mathbf{l}_0'$, is needed. Based on the scene illumination model composed of a near point light source and directional light sources, the method for estimating $d_{0,\perp}$ is described below.

First, based on the assumptions as described in Section 2.1, the planar surface irradiance at a point P can be given by

$$E_p = \ I_0 \cos(\theta_{p,0})/d_{p,0}^2 \ + \ \sum_{k=1}^{K}\mathcal{L}_k\,\omega_k\cos(\theta_{p,k}) \ + I_a \ (10)$$

where I_0 is the near point light intensity. To model environment lighting, $\mathcal{L}_k$ is the radiance of the kth directional light source, and I_a is the intensity of the ambient light. By dividing the hemisphere above $\mathcal{P}$ equally into K patches (e.g., as a geodesic dome used in [5]), all directional lights have an equal solid angle ω_k, and the term $\mathcal{L}_k\omega_k$ can be replaced by the light source intensities I_k, $k = 1, …, K$.

With the calibrated optical camera, the image irradiance (and intensity) is known to be proportional to the Lambertian planar surface radiance, which is in turn proportional to E_p regardless of the viewing angle. Let $I_c(m,n)$ be the optical image intensity at (m,n), the relationship between $I_c(m,n)$ and the light source parameters is as described by the following equation.

$$(\,\mathbf{A}\odot\mathbf{B}\,)\,\mathbf{q} \ = \ \mathbf{h} \qquad (11)$$

where

$$\mathbf{A} \ = \ \begin{bmatrix} a_{1,0} & a_{1,1} & … & a_{1,K} \\ a_{2,0} & a_{2,1} & … & a_{2,K} \\ \vdots & \vdots & \ddots & \vdots \\ a_{J,0} & a_{J,1} & … & a_{J,K} \end{bmatrix} \qquad (12)$$

$$\mathbf{B} \ = \ \begin{bmatrix} \cos(\theta_{1,0})/d_{1,0}^2 & \cos(\theta_{1,1}) & … & \cos(\theta_{1,K}) \\ \cos(\theta_{2,0})/d_{2,0}^2 & \cos(\theta_{2,1}) & … & \cos(\theta_{2,K}) \\ \vdots & \vdots & \ddots & \vdots \\ \cos(\theta_{J,0})/d_{J,0}^2 & \cos(\theta_{J,1}) & … & \cos(\theta_{J,K}) \end{bmatrix} \ (13)$$

$$\mathbf{q} \ = \ [k_d I_0 \quad k_d I_1 \quad … \quad k_d I_K]^T \qquad (14)$$

$$\mathbf{h} \ = \ [h_1 \quad h_2 \quad … \quad h_J]^T \qquad (15)$$

From the above equations, $\odot$ is a Hadamard product, $\theta_{j,k}$ is the incident angle between the kth source and the plane normal at the jth point, h_j is the image pixel intensity corresponding to the jth surface point, and k_d is the planar surface reflection coefficient. From Eq. (12), $a_{j,k} \in \{0, 1\}$ is the visibility factor regarding to the kth light when viewed from the jth point. For example, for a point j on a planar surface which is occluded from a light source k, $a_{j,k}$ is equal to 0, or is 1 otherwise. This calculation is achieved by using the point cloud data of the foreground object surface points, made available by an RGB-D camera. With L_k being modeled as directional light sources uniformly distributed over the hemisphere above the plane, the ambient light term is treated as part of L_k's.

As in [5], with the sufficient number of points J, the over-determined set of linear equations as described above can be solved for $\mathbf{q}$. However, unlike [5], the main purpose for the use of Eq. (11) is to get the estimate of $d_{0,\perp}$. It can be obtained by solving the following minimization problem.

$$\hat{d}_{0,\perp} = arg \min_{d_{0,\perp} \in \mathcal{D}} \left\| \left(\mathbf{A}(d_{0,\perp}) \odot \mathbf{B}(d_{0,\perp}) \right) \hat{\mathbf{q}}_{\text{LS}}(d_{0,\perp}) - \mathbf{h} \right\| \tag{16}$$

where $\hat{\mathbf{q}}_{\text{LS}}(d_{0,\perp})$ is the non-negative least squares solution of $\mathbf{q}$ in Eq. (11) for the light source distance $d_{0,\perp}$. In addition, $\mathbf{B}(d_{0,\perp})$ is the matrix $\mathbf{B}$ with the term $\cos(\theta_{j,0}) / d_{j,0}^2$ being calculated using a hypothetical light source position $\mathbf{l}_0' + d_{0,\perp}\mathbf{n}_{\mathcal{P}}$. Similarly, $\mathbf{A}(d_{0,\perp})$ is the matrix $\mathbf{A}$ whose first-column elements being recalculated for each $d_{0,\perp} \in \mathcal{D}$. The solution to Eq. (16) is obtained by searching over a set $\mathcal{D}$ of possible light source distance values. In this method, the search dimension for the light source position is reduced by separately computing $\mathbf{l}_0'$ using the method described in Section 3.1. To further reduce computational time in solving Eq. (16), a range of possible values of $d_{0,\perp}$ can be first obtained by analyzing the height above $\mathcal{P}$ of the main light occluding object from information available in $I_d(m,n)$.

The estimation method above relies on $I_c(m,n)$, which represents the optical image, and the corresponding depth image $I_d(m,n)$. Here, the subscript c may represent any one of the three color channels (R, G, and B), the converted grayscale version of the original color image, or a pre-processed image by other means, e.g. a bright channel [11]. From $I_d(m,n)$, along with the available camera intrinsic parameters, the corresponding 3-dimensional position $\mathbf{p}(m,n)$ of each projected point (m,n) is calculated. In summary, the steps for performing near light source location estimation is explained below.

Step 1: Perform planar background segmentation on $I_c(m,n)$ using depth information from $I_d(m,n)$. As a result of the segmentation, let Φ be the set of illuminated points on $\mathcal{P}$ which are observable by $I_c(m,n)$. In addition, let $\mathbf{g}_i(x_i, y)$, $(x_i, y_i, 0) \in \Phi$,

be the ith illumination gradient at each of the points specified by Φ. Such gradient vector can be obtained from the corresponding image gradient through 2D-3D mapping using $\mathbf{p}(m,n)$ as described.

Step 2: Construct the two largest circle areas $\mathcal{C}_1$ and $\mathcal{C}_2$, both of which are subsets of Φ. By using the illumination gradients within each of these areas, $\mathbf{l}_0'$ is estimated from Eqs. (8-9)

Step 3: Choose J points from $I_c(m,n)$ to construct $\mathbf{h}$. Then, contruct the matrices $\mathbf{A}(d_{0,\perp})$ and $\mathbf{B}(d_{0,\perp})$, and solve for $\hat{\mathbf{q}}_{\text{LS}}(d_{0,\perp})$ using Eq. (11). From all $\hat{\mathbf{q}}_{\text{LS}}(d_{0,\perp})$, $d_{0,\perp} \in \mathcal{D}$, find the estimate $\hat{d}_{0,\perp}$ using Eq. (16).

4. EXPERIMENTS

The experiments were performed under the indoor environment. An object was photographed against a planar background. A planar background was a concrete wall with a matt finish. A Microsoft® Kinect was applied to obtain a depth image, while a consumer-graded SLR camera (Canon® EOS 1100D) was used to obtain an optical image. The SLR camera was mounted on top of the Kinect camera. All cameras were initially calibrated to obtain depth-optical image alignment as well as the optical camera intrinsic parameters. The scene was illuminated by a typical spherical-shaped light bulb. Light source location was measured with respect to the center of the planar background.

To capture the planar background surface illuminated by a single point light source, optical-depth image pairs were taken at the image size of 1728×2592 (with simple up sampling on a depth image). The process was repeated for different camera and light positions, to obtain a total of 20 data sets. The near point light source was placed at various distances between 2.3m – 3.2m from the center of the planar background. The distance between the source and the nearest point on the plane ($d_{0,\perp}$) was between 1 m – 1.7 m. Ten of those images were taken in the presence of an additional distant light source at about 8 meters from the plane. The accuracy on the estimation of $d_{0,\perp}$ is considered first. Given that the actual projected light source location $\mathbf{l}_0'$ is known, the method described in Section 3.2 was applied to estimate $d_{0,\perp}$. With the absence of a directional light, the averaged estimation error is 4.6%. With the presence of the second distant light source, the averaged error is slightly increased to 5.2%. The results were obtained based on Eqs.11-15 with $K = 30$. Solving Eq. (16) without the environment lighting terms ($K = 0$) yielded the averaged estimation error of 25.8% for the case with the additional distant light source.

Next, the accuracy of estimating $\mathbf{l}_0'$ was measured using 1) the difference between the actual distance $\|\mathbf{l}_0'\|$ and that of its estimate, 2) the angle between $\mathbf{l}_0'$ and its estimate. The position $\mathbf{l}_0'$ was estimated using two largest background-only

circular areas in the optical images. The averaged total area used for the estimation is about 30% of the whole image. The averaged angular error is 3°. The averaged distance error, normalized by the actual distance $\|\mathbf{l}'_0\|$, is 20.7%.

When the light source is very close to the object and planar surface (e.g., $d_{0,\perp} < 1.5$ m in our study), inaccuracy in terms of $\|\mathbf{l}'_0\|$ results in noticeable mismatch between the shadow synthesized using the estimated position, and that of the real one. To improve the estimation accuracy, the estimate of $\mathbf{l}'_0$ was further refined by minimizing Eq. (16) with respect to both $d_{0,\perp}$ and $\mathbf{l}'_0$. After such refinement process, averaged estimation error regarding to $\|\mathbf{l}'_0\|$ was substantially reduced. As for $\mathbf{l}'_0$, while the averaged angular estimation error for $\mathbf{l}_0$ is small at 3.7°, the averaged distance estimation error, as normalized by $\|\mathbf{l}_0\|$, is quite high at 42%. Fig. 3 compares some of the (cropped) real images with those synthesized using the light location estimates.

5. CONCLUSIONS

In this paper, the method for a near light source location estimation has been described. The method is divided into two stages. The projected light source position is first estimated, and followed by estimating the source distance from the plane. While angular inaccuracy is comparatively low, error in the distance $\|\mathbf{l}'_0\|$ can make the synthesized shadow noticeably different from the actual one. Search over both $d_{0,\perp}$ and $\mathbf{l}'_0$ is needed to improve the result.

REFERENCES

[1] A. Fournier, A. S. Gunawan, and C. Romanzin, "Common Illumination between Real and Computer Generated Scenes," *Proceedings Graphics Interface*, 1993, pp. 254–262.

[2] K. Hara, K. Nishino, and K. Ikeuchi, "Multiple light sources and reflectance property estimation based on a mixture of spherical distributions," *Tenth IEEE International Conference on Computer Vision*, vol. 2, 2005, pp. 1627 – 1634.

[3] T. Takai, A. Maki, K. Niinuma, and T. Matsuyama, "Difference sphere: An approach to near light source estimation," *Computer Vision and Image Understanding*, vol. 113, no. 9, 2009, pp. 966–978.

[4] Y. Wang and D. Samaras, "Estimation of multiple directional light sources for synthesis of augmented reality images," *Graphical Models*, vol. 65, no. 4, 2003, pp. 185–205.

[5] I. Sato, Y. Sato, and K. Ikeuchi, "Illumination distribution from brightness in shadows: adaptive estimation of illumination distribution with unknown reflectance properties in shadow regions," *7th IEEE International Conference on Computer Vision*, 1999, pp. 875 – 882.

[6] A. Panagopoulos, C. Wang, D. Samaras, and N. Paragios, "Simultaneous cast shadows, illumination and geometry inference using hypergraphs.," *IEEE transactions on pattern analysis and machine intelligence*, vol. 35, no. 2, Feb. 2003, pp. 437–49.

[7] W. Zhou and C. Kambhamettu, "A Unified Framework for Scene Illuminant Estimation," *Procedings of the British Machine Vision Conference 2004*, 2004, pp. 95.1–95.10.
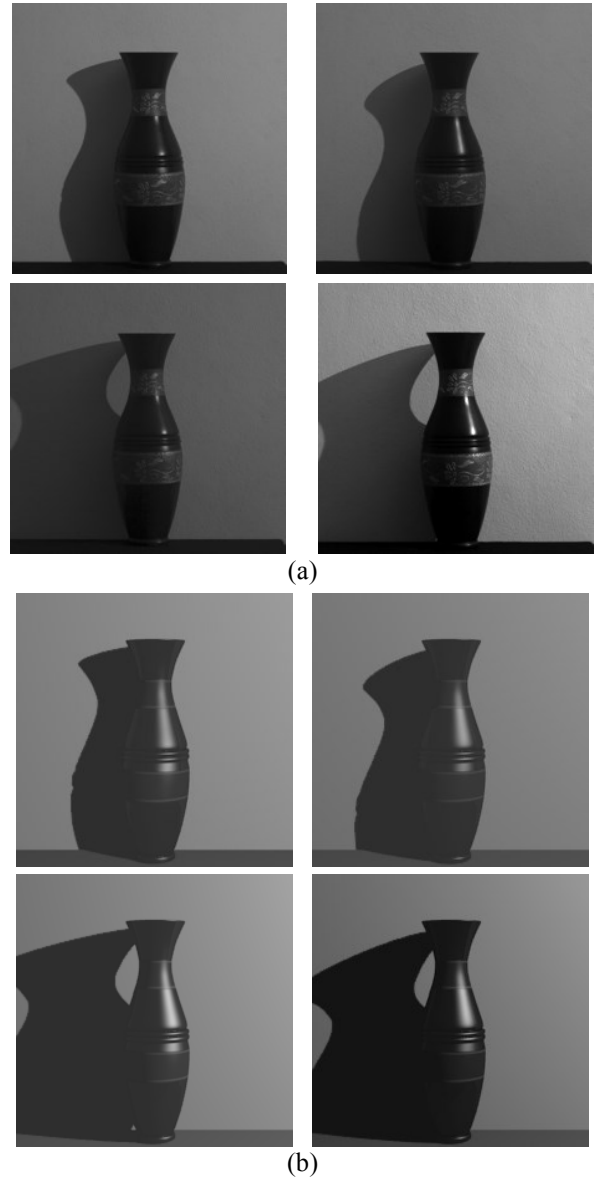
(a)

(b)

Fig. 3 Comparison of a) the original images, b) the synthesized images based on the results obtained using the estimated light source positions.

[8] K. Hara, K. Nishino, and K. Ikeuchi, "Light source position and reflectance estimation from a single view without the distant illumination assumption.," *IEEE transactions on pattern analysis and machine intelligence*, vol. 27, no. 4, Apr. 2005, pp. 493–505.

[9] J. D. Yoo and K. H. Lee, "Real Time Light Source Estimation Using a Fish-Eye Lens with ND Filters," *2008 International Symposium on Ubiquitous Virtual Reality*, Jul. 2008, pp. 41–42.

[10] T. Ikeda, Y. Oyamada, M. Sugimoto, and H. Saito, "Illumination estimation from shadow and incomplete object shape captured by an RGB-D camera," *21st International Conference on Pattern Recognition*, 2012, pp. 165–169.

[11] A. Panagopoulos1, C. Wang, D. Samaras, and N. Paragios, "Estimating Shadows with the Bright Channel Cue," *Proceedings of the 11th European conference on Trends and Topics in Computer Vision*, 2010, pp. 1-12.