A Unified Framework For Scene Illuminant Estimation

Wei Zhou , Chandra Kambhamettu
Video/Image Modeling and Synthesis (VIMS) Lab
Department of Computer and Information Sciences
University of Delaware, Newark DE 19716, USA,
wzhou,chandra@cis.udel.edu

Abstract

Most illuminant estimation algorithms work with the assumption of one specific type of the light source (e.g. point light source or directional light source). This assumption brings up two main limitations which significantly restrict the applicability of the algorithms: First, the knowledge about the type of the light source presented in the scene is needed a priori; second, it can not handle complex scenes where multiple different types of light sources co-exist. To overcome these limitations, we remove the assumption about the source type and develop a general light source model for all different types of light sources. Based on this general light source model, we propose a unified framework to estimate multiple illuminants of different types. Within the framework, we use an experiment setup where a calibration sphere with a specular surface is utilized to probe the scene illuminants and a novel ray tracing and matching algorithm is devised to estimate the light source parameters. Experiment results demonstrate the accuracy of our method on a variety of real images.

1 Introduction

Lots of light source models have been used in computer vision and computer graphics, among which the most common ones are point light source, directional light source and area light source. Each of these models provides an appropriate approximation to a specific category of the real light sources. In real applications, the situation could be very complex that multiple different types of light sources co-exist in the scene. Unfortunately, most light source estimation algorithms work with the assumption of one specific type of light source and there presently exists no method in computer vision for inferring the types of light sources from images. This is problematic because choosing which model to apply in a lighting related problem requires the prior knowledge about the types of the light sources presented in the scene. It is desirable to have a general light source model which allows us to estimate the light sources of different types within a single framework.

BMVC 2004 doi:10.5244/C.18.95

1.1 Previous works

Most previous approaches for illuminant estimation make the assumption of directional light sources. For example, many schemes have been proposed in the context of shape from shading [14] to estimate a single directional light source. Recently, many attempts have been made to detect multiple directional light sources by exploring different cues: The shadows cast by an object of known shape onto another object in the scene is used in [9] [12] to estimate the illumination distribution; [16] utilize the specular properties of a shining surface to estimate the light source directions; [10] [13] introduce the idea of occluding boundary to constrain the light source directions; [15] extend the idea of occluding boundary and estimate lighting direction from critical points that have surface normals perpendicular to an illuminant direction. Since the detection of critical points is sensitive to noise, [11] further segment the surface into ”virtual light patches” by the critical boundaries and illuminants are estimated by minimizing a global error. [6] integrate the shading, shadow and specular cues into a single framework which results in more robust estimation.

All the previous methods assume directional light sources, which are reasonable approximations to the real scene illuminants provided that the distances between the light sources and the illuminated objects are large enough compared against the sizes of both the light sources and the illuminated objects. However, this approximation may not be valid in some cases, especially when light sources are within a limited range, e.g. a small room. The proximal point source model is used in this situation such as in [7] [5], by introducing another parameter, i.e. the distance of the light source. Nevertheless, all those algorithms are designed for one specific type of light source and require prior knowledge about the type of the light source presented in the scene.

To estimate the complex scene illuminations, [2] introduces a light-based model. The model is constructed by mapping reflections on a spherical mirror onto a geometric model of the scene. Although the algorithm works well on superimposing virtual objects onto the real image with convincing shadings, the geometries of all objects in the local scene have to be given a priori, while shape modeling of real objects itself is a difficult research topic. [8] addresses the problem with a radiance-map model. They first build a rough geometric model of the scene as a triangular mesh from a pair of omni-directional images taken from different locations, then the radiance-map is constructed by mapping the radiance of the scene onto the reconstructed geometrical model. Although these models have been successfully utilized to render virtual object onto the real images, they were designed specifically for the Augmented Reality related applications, therefore, can not be used as a general light source model for all lighting related applications in both computer vision and computer graphics. Specifically, these models suffer from three limitations: First, the direct and indirect illuminations are not explicitly distinguished, which is problematic when the applications only need the information about the direct illuminations; second, the illuminations are estimated as a radiance map while the parameters of the light sources such as the distance and the size of the light source are not explicitly estimated; finally, the geometric model of the scene should be built a priori or be roughly estimated as a triangular mesh, therefore, the estimations of the light sources are either expensive or inaccurate. These limitations may not be problems for the applications of Augmented Reality, while for most computer vision algorithms, such as shape from shading and photometric stereo, accurate estimation of the parameters of the light sources such as the directions and the distances of the light sources (direct illuminations) is crucial for the success of the algorithms.

In [4], Langer *et. al* proposed that the analysis of different types of light sources could be performed within a 4-D light source hypercube. They demonstrated that the set of rays within a free space is a 4-D manifold, and various light sources can be classified into light sources with different dimensions ranging from 0 to 4. In this paper, we extend the idea of the 4-D hypercube and propose a general light source model which is more suitable for the purpose of the light source estimation. With this general light source model, different types of light sources can be treated in the same manner, which makes it possible to design a single unified estimation framework. In section 2, we describe the light source modeling in detail.

1.2 Overview

In this work, we use the similar experimental setup as described in [16], where a calibration sphere with a specular surface is used to probe the scene illuminant. Images are taken by a well calibrated camera. Figure 1 outlines the whole estimation process.

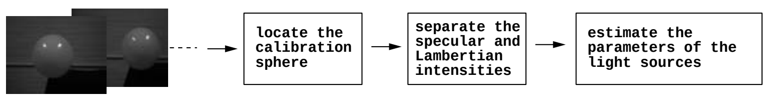


Figure 1: Framework outline

First, a set of images are taken with the calibration sphere placed in different locations of the scene. Then the locations of the sphere are estimated. The surface of the calibration sphere is not assumed to be pure Lambertian or pure specular, instead, we consider that the intensities of the sphere surface may contain both lambertian and specular intensities. We separate the Lambertian intensities from the specular intensities, the specular information is used to estimate the location parameters of the light sources and the Lambertian intensities is used to estimate the light source intensities.

It should be pointed out that, unlike the work of [16] which estimate the directional light sources, the main purpose of this work is to develop a unified framework, which removes the assumption of the source types. Although our work can be viewed as an extension of [16] from some stand points, we argue that the idea of the unified illuminant estimation is the main contribution of our work which could be applied to any other experimental setups.

The rest of this paper is organized as follows: In section 2, we propose a general light source model for all different types of light sources. Section 3 describes the unified estimation framework in detail. Here, we assume that the calibration sphere has already been located and the specular intensities have already been separated from the Lambertian intensities, (For details of these two topics, we refer the readers to the work of [16]). Section 4 shows the experiment results on real images and section 5 gives conclusions and future work.

2 light source modeling

In [4], Langer *et. al* proposed a theoretical framework to compare the different types of the light sources. They demonstrated that the set of rays (including non-source rays and source rays) within a free space is a 4-D manifold, and the set of source rays can be modeled as a 4-D light source hypercube. Figure (2a) illustrates the idea: Place a plane $\mathscr{P}$ between the light source and the scene, a ray r piercing the plane through a point (x_0, y_0) in direction $(p_0, q_0, 1)$ is parameterized as (x_0, y_0, p_0, q_0). So, a set of source rays can be defined as:

$$\mathscr{M}_{src} = \{(x, y, p, q) : x \in [-\frac{h_x}{2}, \frac{h_x}{2}], y \in [-\frac{h_y}{2}, \frac{h_y}{2}], p \in [-\frac{h_p}{2}, \frac{h_p}{2}], q \in [-\frac{h_q}{2}, \frac{h_q}{2}]\}. \quad (1)$$

This set is a 4-D hypercube, and the four parameters h_x, h_y, h_p, h_q define the sizes of the four dimensions of the hypercube. This hypercube structure provides a generic model for the rays emitted from the light sources of different types.

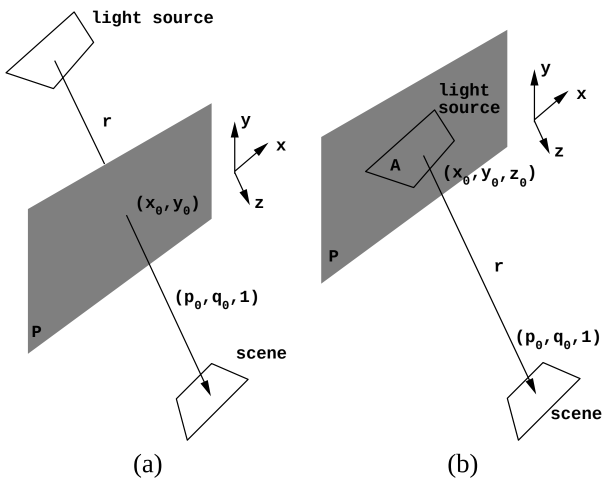


Figure 2: light source modeling

It should be noted that this 4-D hypercube model only captures the general characteristics of a set of source rays in the free space, less concern is given to the information of the light source itself, such as the location, the size, or the radiance of the light source. Here, we extend the idea of the 4-D hypercube and propose a general model of light sources which is more suitable for the purpose of the light source estimation. As shown in figure (2b), instead of putting a plane arbitrarily between the light source and the scene, we choose the plane $\mathscr{P}$ to be one of the tangent planes of the light source oriented in such a way that it maximizes the alignment of its normal with all the rays emitted from the light source to the scene. Similarly, we define a ray r piercing the plane through point (x_0, y_0, z_0) in direction $(p_0, q_0, 1)$ with intensity l_0 as $(x_0, y_0, z_0, p_0, q_0, l_0)$. Since the plane is extremely close to the light source, we make an approximation that the light source lies on the plane. So, a ray $(x_0, y_0, z_0, p_0, q_0, l_0)$ can be described as originating from point (x_0, y_0, z_0), and emitting in direction $(p_0, q_0, 1)$ with intensity l_0. Hereby, a light source can be defined in terms of the set of rays emitting into the scene:

$$\mathscr{M}'_{src} = \{(x, y, z, p, q, l) : (x, y, z) \in A, (p, q) \in Q, l \in \Re^+\}, \quad (2)$$

where A represents the source region on the plane $\mathscr{P}$, and $Q \in \Re^2$ is the set of the directions of all the rays emitting from the light source. Equation (2) provides us a general light source model (GLM). According to this model, the light source estimation process includes estimating the source region A, the directional range Q and the light intensities l.

Now, we have built up a general light source model (GLM) (equation 2) for all types of the light sources. In the next section, we will develop a unified framework for scene illuminant estimation based on this model.

3 A unified framework for scene illuminant estimation

According to the general light source model (GLM) proposed in last section, the light source estimation process includes three components: estimating the source region A, the directional range Q and the light intensity l. Estimating the directional range Q could be a very complicated problem for non-isotropic light sources, for instance, laser lights or flashlights. Since source rays only reach some parts of the scene, all possible directions should be sampled and probed in order to detect the light source. In this work, we assume that the rays emitting from light sources cover the whole range of the scene. Therefore, we simply set the directional range Q to span the whole directional space: $Q = \Re^2$ and put our effort mainly on estimating the source region A and the light intensity l.

3.1 Estimating the source region A

As mentioned previously, we use the similar experimental setup as described in [16], where a calibration sphere with a specular surface is used to probe the scene illuminant. Assuming that the calibration sphere has already been located and the specular patches have already been separated from the Lambertian intensities, the locations of light sources can be estimated by a ray tracing and matching process described in the following.

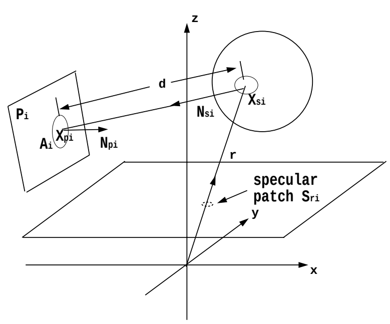


Figure 3: Ray tracing geometry

In GLM, the source region A is defined as a region on a plane $\mathscr{P}$. Once the plane $\mathscr{P}$ is determined, the source region A can be estimated by intersecting the retraced rays with the plane $\mathscr{P}$ (see figure 3). Based on this observation, the estimation of the light source locations is equivalent to determining the plane $\mathscr{P}_i$ for each light source i. We define the plane $\mathscr{P}_i$ in 3-D space as:

$$(\mathbf{X} - \mathbf{X_{pi}}) \cdot \mathbf{N_{pi}} = 0 \tag{3}$$

where point $\mathbf{X_{pi}}$ is a point on the plane $\mathscr{P}_i$, and $\mathbf{N_{pi}}$ is the plane normal. In order to determine the plane $\mathscr{P}_i$, we should find a point $\mathbf{X_{pi}}$ on the plane $\mathscr{P}_i$ and estimate its normal $\mathbf{N_{pi}}$. To find a point $\mathbf{X_{pi}}$ on the plane $\mathscr{P}_i$, we notice that all the retraced source rays will eventually intersect with the plane $\mathscr{P}_i$, and one of the intersection points can be chosen as the point $\mathbf{X_{pi}}$. Suppose N images are taken and one of which is chosen as the reference image. We know that each light source will result in a specular patch in each image. Denote the specular patch in the reference image corresponding to the $i^{'}th$ light source as S_{ri}, and the corresponding specular patch in the $j^{'}th$ image as S_{ji}. Retrace a ray through the centroid point of the specular patch S_{ri} until it hits the sphere surface at a point, call it $\mathbf{X}_{si}$. By the specular reflection property, the ray can be further reflected back towards the light source. Denote the direction vector of the reflected ray as $\mathbf{N}_{si}$. With $\mathbf{X}_{si}$ and $\mathbf{N}_{si}$, we define $\mathbf{X}_{pi}$ and $\mathbf{N}_{pi}$ as follows:

$$\begin{aligned} \mathbf{X}_{pi} &= \mathbf{X}_{si} + d * \mathbf{N}_{si} && (4) \\ \mathbf{N}_{pi} &= \mathbf{R}(\theta,\phi)\mathbf{N}_{si}. && (5) \end{aligned}$$

where $\mathbf{R}(\theta,\phi)$ is a rotation matrix with two variables: θ, ϕ (latitude and longitude angles with $\mathbf{N}_{si}$ as the pole axis respectively), d is the distance from the reflection point $\mathbf{X}_{si}$ to the intersection point $\mathbf{X}_{pi}$. Since we have already localized the sphere and extracted the specular patches, the reflection point $\mathbf{X}_{si}$ and the reflection vector $\mathbf{N}_{si}$ can be calculated easily. Therefore, the plane $\mathscr{P}_i$ corresponding to the $i^{'}th$ light source only depend on three unknowns: d,θ and ϕ. We re-write the plane function as follows:

$$\mathscr{P}_i(d,\theta,\phi) = 0, \quad (d > 0, \theta \in [0,\pi/2], \phi \in [0,2\pi]) \tag{6}$$

Given an initial guess of the three unknowns (d,θ,ϕ), we can determine the plane $\mathscr{P}_i$. Then, the source region A_i can be estimated by the ray tracing process: For each pixel inside the specular patch S_{ji}, a ray originating from the camera projection center is retraced through it; once it hits the surface of the sphere, the ray is further reflected back towards the light source and intersected with the plane $\mathscr{P}_i$. Repeating the retracing process for all the pixels inside the specular patch S_{ji}, we get a set of intersection points, call it P_{ji}. Obviously, the intersection points should all be inside the source region A_{ji}, and can be viewed as the discrete samplings of the source region A_{ji}. Therefore, the source region A_{ji} can be calculated as the convex hull of the intersection point set , i.e. $A_{ji} = H(P_{ji})$, where $H(\bullet)$ is a function which calculates the convex hull of a point set.

The whole ray tracing process described above can be formulated as a mapping function: $\Xi(\mathscr{P}_i(d,\theta,\phi),S_{ji}) = A_{ji}$, i.e, given plane $\mathscr{P}_i(d,\theta,\phi)$, the specular patch S_{ji} corresponding to the $i^{'}th$ light source in the $j^{'}th$ image is mapped to the region A_{ji} on the plane $\mathscr{P}_i$. Mapping all the specular patches corresponding to the $i^{'}th$ light source from all the captured images, we get N regions $A_{ji}(j \in [1,N])$. If the estimation of the plane $\mathscr{P}_i$ is correct, all the specular patches resulting from the same light source should map to the same source region, i.e. all mapped regions $A_{ji},(j \in [1,N])$ should be overlapped. Define a function $\chi(A_1,A_2)$ to measure the matching error between two regions:

$$\chi(A_1,A_2) = 1 - \frac{\mathscr{R}(A_1 \cap A_2)}{\mathscr{R}(A_1 \cup A_2)} \tag{7}$$

where $\mathscr{R}(\bullet)$ is a function which calculate the area of a region. The second term of the equation (7) is the ratio of the intersection area and the union area of the two regions.

Ideally, when two regions overlap, the intersection and the union of the two regions should be the same, and the function χ gets its minimum 0.

Considering all the specular patches corresponding to the $i^{'}th$ light source, we define a matching error function with respect to the plane $\mathscr{P}_i(d,\theta,\phi)$:

$$E(\mathscr{P}_i(d,\theta,\phi)) = \frac{1}{N-1}\sum_{j=1,j\neq r}^{N}\chi(\Xi(\mathscr{P}_i(d,\theta,\phi),S_{ji}),\Xi(\mathscr{P}_i(d,\theta,\phi),S_{ri})) \quad (8)$$

The optimal estimation of the plane $\mathscr{P}_i$ can be achieved by minimizing the matching error function. Once the plane $\mathscr{P}_i$ is determined, the source region A_i can be easily estimated by the ray retracing process described above.

3.2 Estimating the light source intensity

In the last subsection, the source region A is estimated as the convex hull of a set of points P_i. So, the light source can be approximated as a set of clustered point light sources. Based on this approximation, the Lambertian intensity of a pixel x in the sphere image can be written as:

$$I_{lam}(x) \quad = \quad I_a + \rho\sum_{i=1}^{n} I_d^{(i)} \sum_{k\in P_i} \frac{N(x)\cdot L_k}{d_k^2} \quad (9)$$

where the $N(x)$ is the normal of the sphere surface point corresponding to the pixel x. ρ is the diffuse-reflection coefficient of the uniform sphere surface. L_k and d_k are the direction and the distance of the light source. I_a is the ambient intensity, and $I_d^{(i)}$ is the illumination intensity contributed by the $i'th$ light source. Let $I_d^{'(i)} = I_d^{(i)} * \rho$, the previous equation becomes:

$$I_{lam}(x) \quad = \quad I_a + \sum_{i=1}^{n} I_d^{'(i)} \sum_{k\in P_i} \frac{N(x)\cdot L_k}{d_k^2} \quad (10)$$

With this, we can form a system of equations for light intensities: I_a, $I_d^{'(i)}$, which can be solved by the following least-squares minimization:

$$arg \min_{I_a, I_d^{'(i)}} \sum_x [I_{lam}(x) - I_a - \sum_{i=1}^{n} I_d^{'(i)} \sum_{k\in P_i} \frac{N(x)\cdot L_k}{d_k^2}]^2 \quad (11)$$

Since the light intensities should be non-negative, non-negative least mean square (NNLMS) is used to solve the equation. Due to the inherent ambiguity between illumination intensity and albedo, only relative intensity values can be recovered up to a scale factor. We set $I_d^{'(1)}$ to 1 for the convenience of calculation.

4 Experiments

In order to test and evaluate our approach in practice, we have performed experiments on real images. The images were acquired by a Canon S30 camera. The intrinsic parameters of the camera was calibrated using Bouguet's calibration toolbox [1]. Figure (4) shows an example which illustrate the estimation process. The first row shows two images captured with the sphere placed in two different locations of the scene illuminated by a desk lamp. The second row shows the specular patches extracted from the images. The last row demonstrates the estimation process for the plane $\mathscr{P}$, where the left graph shows the two mapped source regions on the plane $\mathscr{P}$ with the initial guess of the plane parameters. After minimizing the error function, we get the optimal estimation of the plane $\mathscr{P}$, and the two mapped source regions on the plane $\mathscr{P}$ are almost overlapped as shown in the right graph.

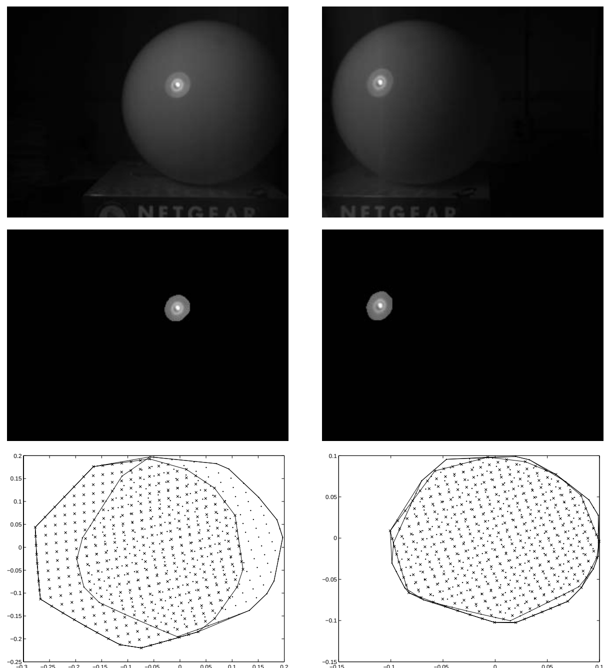


Figure 4: source region estimation

Sixty light sources located in different locations in total were used to create different kinds of images. The locations and the sizes of the light sources were manually measured as ground truth. Here, we only evaluate the center location of the source region in order to simplify the measuring process. As shown in Figure (5), the average error of the estimated locations is about 4% and the average error of the estimated sizes is about 6%. To evaluate the intensity estimation, we use the high dynamic range images of the light sources as the ground truth [3]. Since only the relative intensity value can be estimated, we calculated the intensity ratio of different light sources and the average error is about 1%.

As a unified illuminant estimation framework, our algorithm can handle complex scenes where different types of light sources co-exist. Figure (6) shows a scene where two light sources are present. One is close to the scene which should be better modeled as an area light source, while the other is far away from the scene and can be modeled as a directional light source. We place the calibration sphere on the table to probe the lighting information. With the estimated lighting information, we render a synthetic teapot

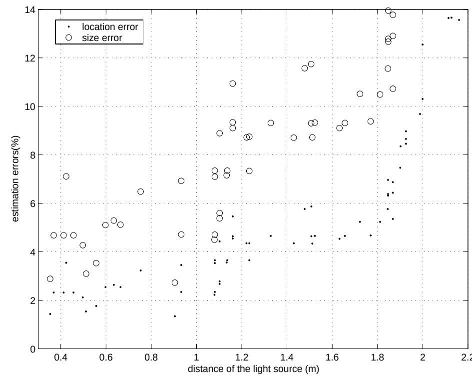
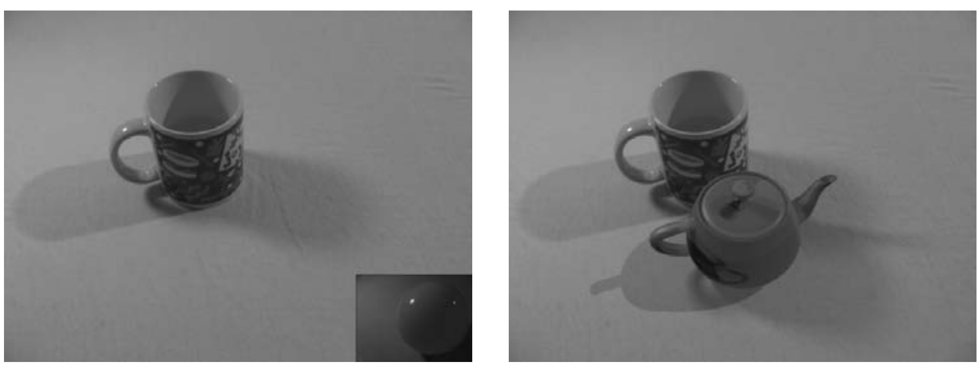


Figure 5: Error estimation for real image

into the real scene. Figure (6a,b) show the original scene and the rendered mixture scene respectively. We can see that the shading and the shadow of the synthetic teapot are correctly rendered. It should be mentioned that since there are two different types of light sources co-exist in the scene, most illuminant estimation algorithms working with one specific source type will definitely fail to capture the correct lighting information. In our estimation framework, the problem is avoided by employing the general light source model. From the rendering result, we can see that the rendered shadow patterns casted by the two light sources are different, one has the sharp edge while the other has the blurred edge which match perfectly with the real shadow patterns of the cup on the table.



(a) (b)

Figure 6: Application to Augmented Reality. (a): original scene; (b): a synthetic teapot rendered into the scene

5 Conclusion and future work

This paper proposes a unified framework for scene illuminant estimation. Primary contributions of our work towards illuminant estimation include: First, a general light source model (GLM) is proposed to model different types of light sources; second, a novel illuminant estimation scheme based on GLM is proposed. Our framework removes the assumption of the source types, and can therefore deal with complex scenes where different types of light sources co-exist. Experimental results on real data show both of efficiency and accuracy of our algorithm. In future, our framework will be extended by applying to other experimental setups, where arbitrary scene object could be used by integrating the lighting estimation framework with scene reconstruction algorithms complementarily.

References

[1] J.-Y. Bouguet. *Camera calibration toolbox for matlab*. http://www.vision.caltech.edu/bouguetj/calib_doc/index.html, 2001.

[2] P. Debevec. *Rendering Synthetic Objects into Real Scenes: Bridging traditional and image-based graphics with global illumination and high dynamic range photography*. ACM SIGGraph, Computer Graphics pp.189-198, July 1998.

[3] P. Debevec and J. Malik. *Recovering high dynamic range radiance maps from photographs*. ACM SIGGraph, Computer Graphics pp.369-378, August, 1997.

[4] M. S. Langer and S. W Zucker. *What is a light source?* IEEE Proc. Computer Vision and Pattern Recognition, pp.172-178, 1997.

[5] H. P. A. Lensch, J. Kautz, M. Goesele, W. Heigrich, and H. Seidel. *Image-based reconstruction of spatial appearance and geometric detail*. ACM Transactions on Graphics, Vol.V, No.N, May, 2003.

[6] Y. Li, S. Lin, H. Lu, and H. Shum. *Multiple-cue illumination estimation in textured scenes*. IEEE Proc. 9th International Conference on Computer Vision, vol.2, 2003.

[7] M. W. Powell and D. Goldgof. *Calibration of Light Sources*. IEEE Proc. Computer Vision and Pattern Recognition, vol.2, pp.263–269, 2000.

[8] I. Sato, Y. Sato, and K. Ikeuchi. *Acquiring a radiance distribution to superimpose virtual objects onto a real scene*. IEEE transactions on visualization and computer graphics, vol.5, No.1, March, 1999.

[9] 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*. Proceedings of IEEE International Conference on Computer Vision, pp.875-882, September, 1999.

[10] E. V. Vega and Y. H. Yang. *Default shape theory: With the application to the computation of the direction of the light source*. JOSA-A, 60:285-299, 1994.

[11] Y. Wang and D. Samaras. *Estimation of multiple illuminants from a single image of arbitrary known geometry*. 7th European Conference on Computer Vision part.3, pp.272-288, 2002.

[12] Y. Wang and D. Samaras. *Estimation of multiple directional light sources for synthesis of Augmented Reality images*. Graphical Models, 65, pp.185-205, 2003.

[13] Y. Yang and A. Yuille. *Source from shading*. IEEE Proc. CVPR pp.534–539, 1991.

[14] R. Zhang, Ping-Sing Tsai, J. Cryer, and M. Shah. *Shape from Shading: A Survey*. IEEE Transactions on PAMI, Volume 21, Number 08, pp 690-706, August, 1999.

[15] Y. Zhang and Y. Yang. *Illumination Direction Determination for Multiple Light Sources*. IEEE Proc. CVPR, vol.1, pp.269-276, 2000.

[16] W. Zhou and C. Kambhamettu. *Estimation of Illumination Direction and Intensity of Multiple Light sources*. 7th ECCV part.4, pp.206-220, 2002.