A Review of Selected Multiple Light Source Detection Methods

Nathan Funk
University of Alberta
Assignment 2 for CMPUT 605

October 20, 2004

1 Hougen and Ahuja

This 1993 paper [2] is the first method to approach the multiple light source problem. It uses a combination of stereo, light source estimation, and shape from shading methods for determining the shape of objects. This summary will focus on the proposed light source estimation technique and it's integration in the system. Four main assumptions are stated:

1. The light sources are distant point light sources.
2. There is no shadowing on visible objects.
3. There are no interreflections between visible objects.
4. The luminance must appear as a multiplicative constant in the reflectance function.

They also mention that the technique works best if the reflectance function is not highly specular.

For a single light source the irradiance from a surface region is defined by

$$E(x,y) = \rho(x,y) \int_{\Theta} R(\mathbf{n}(x,y), \mathbf{l}(\theta)) \mathrm{d}\Omega, \tag{1}$$

where E is the total image irradiance, ρ is the albedo, R is the reflectance map, $\mathbf{n}$ is the normal, and $\mathbf{l}$ is the source distribution. θ represents the angular position parameters, and the integral is taken over the sphere of possible incident directions.

The source distribution function is assumed to be composed of a set of m point light sources. An ambient term λ_0 is introduced and the equation is discretised into a summation formulation. Assuming that k measurements of the irradiance are taken, at a single point j in the image, the irradiance equation for $j = 1, ..., k$ is

$$E_j = \rho_j(\lambda_0 \tilde{R}_0 + \sum_{s=1}^{m} \lambda_s \tilde{R}(\mathbf{n}_j, \tilde{\mathbf{l}}_s)), \tag{2}$$

where $\tilde{l}_s$ is a unit vector for a particular light source, and λ_s is the associated intensity. This set of k equations is linear with respect to the light source intensities $\lambda_0 ... \lambda_m$. Hence, if all other variables are known, it is possible to solve for the intensities when $k \geq m + 1$.

This means that it is necessary to know the directions of the light sources prior to solving for the intensity. The authors simply choose a roughly uniform distribution of light sources and expect the intensity to be

calculated as 0 at the locations where no light sources are present. This is different from other multiple light source detection techniques which attempt to determine the exact location of each light source.

The intensity values can physically not be negative, so it becomes necessary to constrain the solution to nonnegative values. The authors claim that it is not necessary to consider nonnegativity constraints when the only concern is surface shape recovery. A detailed reason is not provided. Ignoring the nonnegativity constraint for only the ambient term is justified by the voltage gain cutoff in the digitization process. They explain that the cutoff results in reducing all intensity values in the image by a fixed amount. This is not absolutely accurate since some intensities might then be reduced below 0. The general idea that an unconstrained ambient term will *help* compensate for this effect is however correct.

The geometric information about the scene is obtained using a stereo method. With this information, the lighting is estimated, and then finally the shape from shading method is applied.

The performance of the lighting estimation algorithm is tested using synthetic images with known illumination conditions. Sets of 5, 9, and 17 light sources (model sources) are employed in recovering the distribution. To obtain a measurement of the error recovery accuracy, the image is rendered using the recovered lighting, and the root mean squared (RMS) difference between the original and reconstructed images is calculated. In most cases, increasing the number of model sources reduced the error in the recovered image. Using unconstrained sources (virtual sources) also reduced the error.

In comparison to other techniques, this method is a relatively simplistic approach to the multiple source problem. It defines the irradiance as a linear system with respect to the light source intensities, then solves for the unknowns using a constrained least-squares approach. No information is provided on the performance, although it is expected to be considerable when a large number of modeling sources are employed. Nevertheless, it appears to be relatively robust and straightforward to implement.

2 Sato, Sato and Ikeuchi

This is the first approach [4] to determine the illumination distribution using the intensity information from an object's shadows rather than using the shading on the object. Similar to Hougen and Ahuja's method, the illumination is recovered as a discretised distribution over a sphere. The main assumption again is that the sources are distant.

A derivation of the image irradiance in the shadow region is provided. Starting with the scene irradiance, the occlusion of incoming light is considered by adding the occlusion coefficients $S(\theta_i, \phi_i)$. For a particular point on the surface the value of S is either 0 or 1 depending on whether the incoming light is occluded from the direction (θ_i, ϕ_i). After developing the initial expression from scene irradiance to scene radiance, then finally to image irradiance, the final equation for the *shadow image* $P(\theta_e, \phi_e)$ is presented as

$$P(\theta_e, \phi_e) = k \int_{-\pi}^{\pi} \int_{0}^{\frac{\pi}{2}} f(\theta_i, \phi_i; \theta_e, \phi_e) L_0(\theta_i, \phi_i) S(\theta_i, \phi_i) \cos(\theta_i) \sin(\theta_i) \mathrm{d}\theta_i \mathrm{d}\phi_i. \tag{3}$$

Here, (θ_e, ϕ_e) is the direction from the surface to the camera and (θ_i, ϕ_i) is the incoming direction of the light. k is the scaling factor between the scene radiance and a pixel value. The BRDF is represented by $f(\theta_i, \phi_i; \theta_e, \phi_e)$, and $L_0(\theta_i, \phi_i)$ is the incoming light intensity. Note that only the upper hemisphere is considered. Although this is not generally suitable, for the experiments in this paper it is appropriate. The object is placed on a horizontal surface and illumination from below the surface does not affect the shadows.

The authors choose to discretise the hemisphere using the nodes of a *geodesic dome*. The area of the hemisphere is divided up into patches located at n nodes, each with an equal solid angle of $\delta\omega = 2\pi/n$. The

equation of the shadow image can then be written as

$$P(\theta_e, \phi_e) = \sum_{i=0}^{n} f(\theta_i, \phi_i; \theta_e, \phi_e) L(\theta_i, \phi_i) S(\theta_i, \phi_i) \cos(\theta_i), \tag{4}$$

where $L(\theta_i, \phi_i)$ is the radiance per solid angle including the scaling factor k.

From here the analysis is broken down into three different cases with respect to the shadow receiving surface. 1) The surface is Lambertian and the reflectance properties (albedo) are known. 2) The surface is non-Lambertian and the reflectance properties are known. 3) The surface is Lambertian but the reflectance properties are not known.

1) When the surface is Lambertian, the BRDF reduces to a constant K_d. It is simple to now set up an linear equation system with n unknown illumination radiance values L_i $(i = 1, 2, ..., n)$. Using enough intensity measurements $P_j (j = 1, 2, ..., m)$ in the shadow region, it is possible to solve for L_i. This of course requires that the occlusion information S is known which in turn requires geometric information about the shadow caster and shadow receiver.

2) If the surface is not Lambertian, it is still possible to determine the illumination distribution. The authors give an example using a simplified Torrance-Sparrow reflection model which includes constants K_d and K_s as measures for the diffuse and specular reflection respectively.

3) With a Lambertian surface, but an unknown albedo K_d one can still estimate the illumination distribution. An additional image called the *surface image* is captured without the shadow casting object in the scene. Then the ratio of the intensities in the two images is used instead of just the intensity in the shadow image. It is also necessary to calibrate the camera with a white surface for this approach to work.

Experiments were performed with a block on a flat surface under typical office illumination including fluorescent lights and windows. To evaluate the recovery results, a new image is synthesized with the recovered lighting and compared with the original image (much like Hougen and Ahuja's evaluation).

Results are obtained with varying numbers of nodes on the geodesic dome. Generally an increasing number of nodes improved the results obtained. With 521 nodes the synthesized image was nearly indistinguishable from the original image. They note that the number of nodes might be lowered depending on the complexity of the scene in order to reduce the computational cost. The results were good for both known and unknown reflectance properties.

From this paper it appears that using shadows for illumination recovery is feasible when the geometry of the shadow caster and receiver are known. The reflectance model of the shadow receiving surface does not need to be Lambertian as long as the resulting equations are linear.

3 Debevec

Debevec's paper [1] focuses on the development of a technique to insert realistic synthetic objects into real scenes. He notes that for realism to be achieved, consistent lighting between the real and synthetic objects is of crucial importance.

To simplify the problem slightly, the scene is divided up into three parts. The first is the *distant scene* which is assumed to not be affected by other parts of the scene. The *local scene* interacts with the *synthetic objects* through shadows, lighting, and interreflection. This division assists in reducing the complexity when inserting the synthesized objects into the original scene.

In order to achieve consistent lighting, an accurate measure of the lighting in the local scene is required. Debevec proposes a method which uses a mirrored sphere placed in the scene. To improve the quality of the results, a *high dynamic range* technique is employed. Multiple images are captured with exposure times ranging from 1/4 to 1/10,000 second. These are then fused together into a single high dynamic range

image. This technique not only improves the visual appearance of the synthesized objects, it is also a more accurate measurement of the true lighting. It could also be applied to other lighting estimation methods as a preprocessing step.

Determining the illumination conditions is very simple in comparison to most other techniques. A mirrored sphere with known size is placed at a known position. The camera parameters such as location and focal length must also be known. With this information, it is trivial to trace the path of the light from the image through the focal point of the camera. Finally, the ray is reflected from the sphere into the environment.

Debevec identified a few problems with this technique. One of them is that the camera and possibly also the photographer are visible in the sphere. The mirror finish also has a much greater interaction with the scene than other surfaces, since it does not attenuate the reflected light as much. The sphere also casts shadows which are visible in the reflection. Furthermore, there is no reflection of the scene directly behind the sphere. In addition, most of the environment behind the sphere is poorly sampled due to the foreshortening near the edge of the sphere. It is possible to reduce the effects of some of these issues through careful placement of the sphere and camera. One can also take multiple images from different directions and stitch the resulting environment maps together. According to the paper, using two images taken from 90 degrees apart works well to remove the camera's appearance and avoid poor sampling.

The rendering of the synthesized objects is performed with a global illumination package. A great part of the paper focuses on how their shadows and reflections are worked into the real image.

This technique is undoubtedly very useful for scenes that can be manipulated. The results are accurate and do not suffer from the limitations that most other lighting estimation methods show. The mirror BRDF greatly simplifies the analysis since only light from a single direction contributes to the appearance of a point on the surface. Hence for situations where light source estimation is required and inserting an object into the scene is not a problem, this technique is well suited.

4 Singh Ahuja

Just like the previous three papers, this method [5] recovers a distant illumination distribution. It also assumes that the observed surface is Lambertian with a constant albedo. A specific shape is not assumed, although the geometry of the object must be known. This technique is different from most others in that it is an iterative approach. It repeatedly estimates the lighting distribution by adjusting the previous estimate to fit within a set of constraints.

The algorithm behind this technique is not new. The POCS framework (Projection Onto Convex Sets) is used for restoring signals that were distorted by noise. It is also applied to image restoration. The similarity between image recovery and illumination estimation is that both can be formulated as a deconvolution problem.

The lighting distribution is described with the function $L(\theta_s, \phi_s)$. The image is represented by $I(i, j)$. A simple model of the image formation process is constructed by discretising the domain of L. The two spherical coordinate parameters are replaced by integers i and j, and the distribution over this discrete domain is referred to as $\tilde{L}$. Note that this form of sampling on the sphere is not uniform. Towards the two poles, the sampling rate becomes higher. Singh and Ahuja do not mention whether a more uniform sampling approach would have any benefits. A convolution kernel $h(i, j, k, l)$ which includes the BRDF and geometric factors is formulated to further simplify the equation describing the image formation. Finally, the image can be written as

$$I(i,j) = \sum_{(k,l)} h(i,j,k,l)\tilde{L}(k,l). \tag{5}$$

The basic concept of POCS is to define a set of constraints C that apply to the function to be recovered. In the case of lighting recovery, two constraints are that the recovered lighting is non-negative, and that the recovered lighting produces the same image as the original lighting. More specifically, in this paper, a constraint $C_{i,j}$ is defined for each pixel (i, j) of the image. A projection operator P is derived for each constraint. This operator, when applied to an estimate, will alter the estimate enough for the according constraint to be satisfied. If all projection operators could be applied at once the recovered estimate would meet all the constraints. But since only one projection is applied at once, each projection might cause previous constraints to no longer be satisfied. For this reason it is necessary to iterate.

For each iteration step, all of the projection operators are applied. This can be written as

$$\tilde{L}_{k+1}(i,j) = P_A P_{N_1,N_2} P_{N_1,N_2-1} \cdots P_{1,2} P_{1,1}(\tilde{L}_k(i,j)), \tag{6}$$

where P_A is the amplitude projection and $P_{i,j}$ is the residual projection for each pixel (i, j).

The amplitude constraint simply constrains the illumination within a range from 0 to A:

$$C_A = \{L(i,j) : 0 \leq L(i,j) \leq A \quad \forall (i,j)\}. \tag{7}$$

The associated projection operator P_A is

$$P_A[x(k,l)] = \begin{cases} 0: & x(k,l) < 0 \\ x(k,l): & 0 \leq x(k,l) \leq A \\ A: & A < x(k,l) \end{cases}. \tag{8}$$

For this constraint it is easy to see how the projection operators function. Values that are below the acceptable range are moved to the lower limit, and values above the range are reduced to the upper limit A.

The residual constraint is a little more involved. Here, for every pixel, the intensity generated from the estimated lighting is compared to the original intensity value. An maximum residual magnitude δ_o is allowed, although in experiment it is set to 0. The constraint is expressed as

$$C_{i,j} = \{L(k,l) : |I(i,j) - \sum_{(k,l)} h(i,j,k,l) L(k,l)| < \delta_o)\}, \tag{9}$$

and the associated projection operator is defined as

$$P_{i,j}[x(k,l)] = \begin{cases} x(k,l) + \frac{d^x(i,j) - \delta_o}{\sum_m \sum_n h^2(i,j,m,n)} h(i,j,m,n): & d^x(i,j) > \delta_o \\ x(i,j): & |d^x(i,j)| \leq \delta_o \\ x(k,l) + \frac{d^x(i,j) - \delta_o}{\sum_m \sum_n h^2(i,j,m,n)} h(i,j,m,n): & d^x(i,j) < -\delta_o \end{cases}, \tag{10}$$

where

$$d^x(i,j) = I(i,j) - \sum_{(k,l)} h(i,j,k,l) x(k,l). \tag{11}$$

Singh and Ahuja implemented the POCS approach and ran a few simulations using the framework. The results they present show multiple circular and rectangular light sources, with up to two in each case. The recovered distribution is a slightly blurred version of the original lighting (for 50 iterations). They note that the recovered lighting features artifacts when the residual projections are applied in order. By applying these projections in a random order, the artifacts were almost completely removed and the rate of convergence was greatly improved.

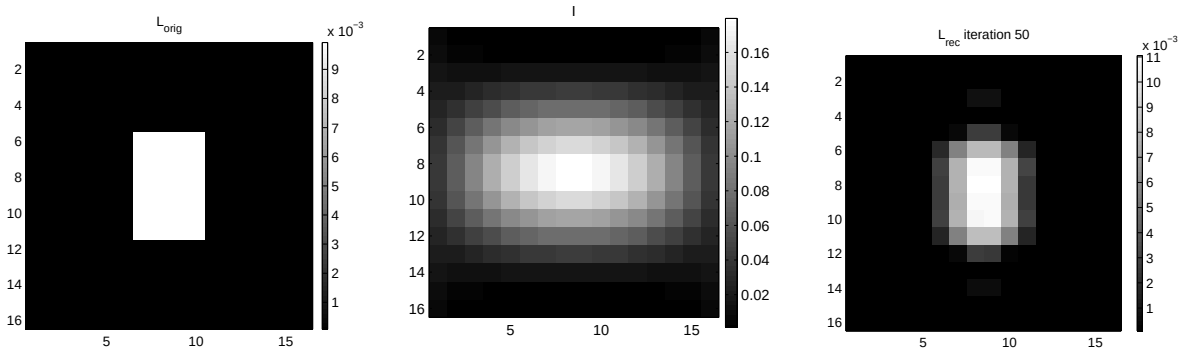

Figure 1: Single area light source. The *original light distribution*, *image of the reflection*, and *recovered light distribution* respectively. All images are shown as 2D maps on the discretised spherical coordinates.

5 Implementation of Singh and Ahuja's method

Singh and Ahuja's method was implemented during the summer period. No major changes were made to the implementation due to time limitations. The focus here is on reproducing the results of Singh and Ahuja. Three tests were conducted: The first two simulate the some of the results published the original paper. The third looks at the convergence behaviour of the algorithm.

All experiments were performed with a 16x16 discretisation of the spherical coordinates. The results in the paper are at a 64x64 resolution. A courser resolution was chosen since the computational time would have been significant for anything finer. This choice does not appear to affect the general appearance of the results. Other parameter values are $\delta_o = 0$ and $A = 1$. Since $A = 255$ in the paper, it is likely that calculations were performed using 8 bit integers. Here, double precision floating point values are used. A significant performance increase is expected by using lower precision arithmetic. The number of iterations for the first two test cases is 50, according to the number used in the published results.

5.1 Single area light source

This first test attempts to reproduce the results from the third row of results in Figure 2 of the original paper. The recovered lighting shown in Figure 1 appears to be as accurate as the published results. The lower resolution does not allow precise comparison, but it is expected that the results would be identical if the higher resolution were used.

Just like in the original implementation, the order in which the residual projection operators are performed is randomised. This reduced the appearance of artifacts. The fact that the recovered distribution is not as sharp as the original distribution might be related to the limitations of recovery from a Lambertian surface which were pointed out by Ramamoorthi in 2001 [3]. Singh and Ahuja do not discuss why complete recovery is not achieved.

5.2 Two area light sources

This test attempts to reproduce the results from the forth row of results in Figure 2 of the original paper. Similar to the first test, the results shown in Figure 2 appear to be the same as those achieved in the original paper. The interesting aspect here is that even though the image intensity values only show a single peak region, the POCS algorithm still correctly restores the two separate area light sources.

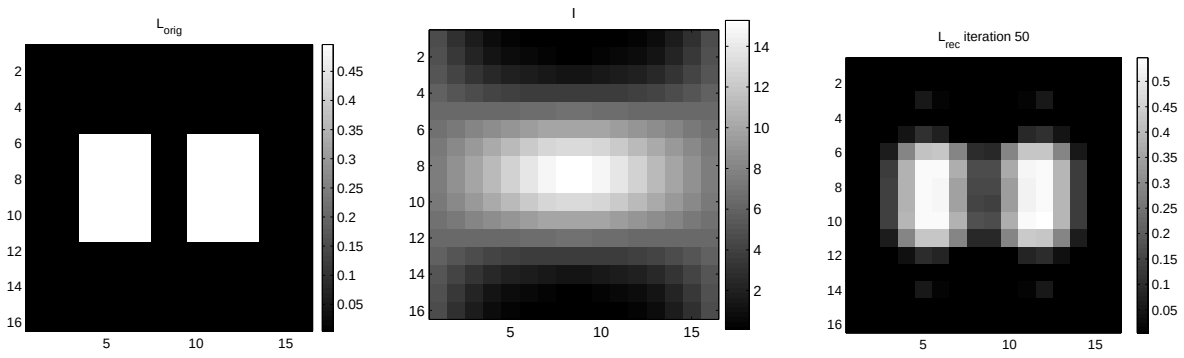

Figure 2: Two area light sources. The *original light distribution*, *image of the reflection*, and *recovered light distribution* respectively. All images are shown as 2D maps on the discretised spherical coordinates.

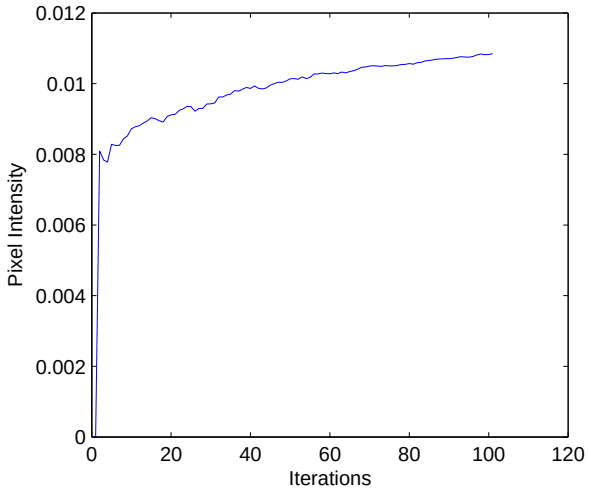

Figure 3: This graph shows the progression of the recovered light intensity at a single point over 100 iteration steps. For other test cases the resulting trend was very similar.

5.3 Convergence

The convergence properties of the algorithm are of interest since it is useful to know how many iterations are necessary. Figure 3 shows a typical trend for the intensity of a single recovered point. The first iteration changes the values the most, after that the change with each iteration is relatively small. Note that even at 100 iterations the trend has not completely flattened out. Attempts were made to accelerate the convergence by introducing an additional factor in equation 10. They did however not have a significant effect and in other cases only caused the results to diverge.

References

[1] P. Debevec. Rendering synthetic objects into real scenes: Bridging traditional and image-based graphics with global illumination and high dynamic range photography. In *Proc. SIGGRAPH 98*, pages 189–198, Orlando, Florida, July 1998.

[2] D. R. Hougen and N. Ahuja. Estimation of the light source distribution and its use in integrated shape recovery from stereo and shading. In *IEEE 4th Int. Conf. On Computer Vision*, pages 148–155, 1993.

[3] R. Ramamoorthi and P. Hanrahan. On the relationship between radiance and irradiance: determining the illumination from images of a convex lambertian object. *Journal of the Optical Society of America A*, 18(10):2448–2459, October 2001.

[4] I. Sato, Y. Sato, and K. Ikeuchi. Estimation of illumination distribution by using soft shadows. *Center for Spatial Information Science Discussion Papers*, (3):1–14, November 1998.

[5] M. Singh and N. Ahuja. Estimating light sources. In *Indian Conference on Computer Vision, Graphics and Image Processing*, pages 76–81, New Delhi, India, December 1998.