A Review of Selected Multiple Light Source Detection Methods

Nathan Funk
University of Alberta
Assignment 3 for CMPUT 605

November 3, 2004

1 Wang and Samaras (2003)

This approach [9] combines estimation from shading with estimation from shadows. The estimation from shading technique is based on Zhang and Yang's paper [11] which was modified and made more robust in [8]. This summary will first describe the two detection techniques, then outline how they are used in combination.

1.1 Shading-based illuminant detection

This approach does not recover an illumination distribution, but rather a set of distant point light sources. Objects are assumed to have a known geometry and a Lambertian surface. The visible points on the object are mapped to a sphere by matching the normal directions. The rest of the algorithm is then performed on the intensity information mapped to the sphere.

The approach used to detect critical points is changed from the original method to increase the robustness. The search for critical points is still performed along arcs by analysing the intensity values on the sphere. However, not only the parameters (b_i, c_i) are used for comparing segments along the arc. The tangent angles on the intensity curve are also compared for each segment. They claim that at critical points, the angle between the tangent lines to either side of the point is larger than $180°$. The final critical point search is implemented by comparing adjacent (b_i, c_i) parameters within segments, then determining the location of a critical point using the tangent angles.

The critical boundaries are then detected with the Hough transform. They use one-third of the highest vote count as the threshold to consider a (ζ, θ) pair as a critical boundary. This is considerably low. But they note that they aim to detect more critical boundaries than real critical boundaries.

With the critical boundaries as edges, the sphere is segmented into regions. For each region, the combined illumination of all light sources on that region is equivalent to one *virtual light source*. The virtual light source can be determined using an approach similar to Hougen and Ahuja's [2]. Using at least four intensity values inside a region, the values (L_x, L_y, L_z, α) can be calculated using a least squares approach. $L = (L_x, L_y, L_z)$ is the virtual light vector, and α is an ambient light term.

With the direction of the virtual light source for each region it is possible to determine the light sources associated with each critical boundary. It can be shown that for two regions separated by a critical boundary, their virtual light sources L_1 and L_2 can be used to determine the light source for that boundary. The *pre-direction* L_{pre} is calculated as $L_{pre} = L_1 - L_2$. The true direction of the source is then either L_{pre} or $-L_{pre}$. Since for most light sources multiple pre-directions will be detected, all pre-directions within a $5°$ radius are merged.

1.2 Shadow-based illuminant detection

The shadow-based detection technique recovers a distant distribution over a geodesic dome. It is based on the work of Sato et al. [6]. A simple linear equation system is constructed where each individual equation is of the form

$$E = \sum_{i=1}^{n} L_i S_i \cos(\theta_i), \tag{1}$$

where E is the total irradiance, L_i is the illumination radiance per solid angle $\delta = 2\pi/n$, and S_i is the occlusion coefficient. Each of the variables in the sum is indexed with i which indicates a particular direction (θ_i, ϕ_i) to a node on the geodesic dome. The occlusion coefficient is 0 when the incoming light from the direction is occluded, and 1 otherwise. With a least squares approach, the equation system is finally solved to determine the unknown radiance values L_i.

1.3 Integration of the two techniques

The combined method starts with the detection from shading first. Critical points are detected as described above. Then they are grouped into critical boundaries with the Hough transform. After the initial pre-directions are determined, the lighting distribution from shadows is calculated. They note that due to the occlusion of shadows by the object, some regions of the distribution can not be estimated so they are excluded from the following steps. The set of pre-directions is compared to the non-excluded distribution regions. If a pre-direction is close to zero in the distribution, it is rejected. If peaks in the distribution are not matched with existing pre-directions, then new pre-directions are added. From here, the algorithm proceeds according to the shading-based technique.

The accuracy of the combined method is reported to be better than either of the individual methods. The computational time is also reduced compared to running only the shading-based method. The reason is that many of the pre-directions are excluded after comparing them to the shadow results.

2 Li, Lin, Lu and Shum (2003)

Similar to Wang and Samaras, this approach [3] also combines use of shading and shadows, but additionally introduces the use of specularities. Li et al. propose a method of integrating these multiple cues by checking for consistency among them. For example if a shadow edge is found, there should also be an associated critical boundary in the shading.

Similarly, consistancy within a single cue is also considered. For example, when critical points are found on the surface, other critical points belonging to the same boundary should also be present. Through introducing these requirements the analysis becomes more robust and edges introduced by texture can be handled since they will be ignored after the consistancy checks. However if the texture becomes too dense, they mention that number of false correlations becomes greater and has a negative effect on the results.

The approach is different from most others in the way it detects features. Typically, features such as critical points are searched for over then entire object, then used in determining the location of light sources. This paper instead uses hypothetical light source directions, and calculates where the effects (critical points, shadow edges, and specular highlights) would be expected to be located. Then it checks the image if the points are in fact present.

So for a particular light source direction L, three sets of expected feature locations are determined:

1. D(L) The expected unoccluded shadow edges.
2. C(L) The set of critical points corresponding to D(L).

3. S(L) The expected specularity peaks.

Using a Canny edge detector, the shadow receiving surface is searched for edges. The edge detector is expected to find edges from both texture and shadows. For each point x_D in $D(L)$, $P_D(x_D)$ is set to 1 if an edge is found at that location, or set to 0 otherwise. For the critical points, a similar approach is taken. They are detected by searching perpendicular to the expected critical boundary. $P_C(x_C)$ is set to 1 when expected points are found. Finally $Sp(L)$ is set to 1 if a specular peak is found at the expected location.

With these three measures it is now possible to check whether the cues found actually also agree with each other. This is done by calculating a consistancy measure $\gamma(L)$ using

$$\gamma(L) = \frac{1}{|D(L)|} \sum_{x_D \in D(L)} [P_D(x_D) \wedge P_C(x'_D)] \vee [P_D(x_D) \wedge Sp(x_D)], \tag{2}$$

where x'_D is a critical point corresponding to the shadow edge point x_D. $|D(L)|$ is the total number of expected points on the shadow contour. It serves as a normalization for the measure. The equation shows that at least two cues are required in combination, and the shadow must be one of them. Since texture edges will likely not have corresponding critical points or shadow edges they do not affect the results to the most part. By thresholding $\gamma(L)$, possible source directions are determined and then refined.

The intensity of each detected light source still needs to be determined. For this, the image is segmented into regions of constant lighting conditions. So for each region, the lighting is constant within that region. Then the boundaries dividing the regions are searched in attempt to find sections that are not affected by texture. Then the intensity values on either side of the boundaries are used to determine the light source intensities with a least squares minimization method. The calculation can be performed with multiple colour channels separately, allowing calculation of the light source colour as well. The intensity calculation however only provides a relative measure of the intensities. This is due to the ambiguity between illumination intensity and albedo (which is assumed to be unknown).

3 Okabe, Sato and Sato (2004)

In this paper [4] the authors compare Spherical Harmonics to Haar Wavelets as a basis for recovery from shadows. Their main assumptions are that the geometry is known, the light sources are distant, there is no interreflection between the surfaces, and that the shadows fall on a Lambertian plane with a uniform albedo.

The image irradiance equation is stated as

$$E(x) = \int_0^{2\pi} \int_0^{\pi/2} V(x, \theta, \phi) L(\theta, \phi) \cos\theta \sin\theta d\theta d\phi \tag{3}$$

where x is a point on the shadow receiving plane, $V(x, \theta, \phi)$ is the visibility function, and $L(\theta, \phi)$ is the illumination radiance distribution. At each point x, the visibility function is 0 from all occluded directions and 1 otherwise. Since it is assumed that the geometry is known, V can be computed for all points visible in the image.

The illumination distribution and the visibility function grouped together with the geometric factors can both be represented in terms of spherical harmonic coefficients. For example, the light distribution $L(\theta, \phi)$ can be represented as a linear combination of the coefficients L_{lm} and the spherical harmonic basis functions $Y_{lm}(\theta, \phi)$ as shown here:

$$L(\theta, \phi) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} L_{lm} Y_{lm}(\theta, \phi). \tag{4}$$

Due to the mathematical features of the spherical harmonics (orthonormality primarily) the image irradiance can be expressed as

$$E(x) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} L_{lm} T_{lm}(x), \tag{5}$$

where $T_{lm}(x)$ are the coefficients of the transfer function. This expression can be further modified to include only low order spherical harmonics under a limit of l_{max} rather than an infinite sum. By doing this, the number of unknowns is limited to $(l_{max}+1)^2$. This makes it possible to construct an equation system from the intensity measurements in the shadow region. When sufficient samples can be taken within the shadow, the linear system can be solved to determine the coefficients of the distribution L_{lm}.

As studied in depth by Ramamoorthi [5] the coefficients of the transfer function are of great importance. For the clamped cosine function, the coefficients rapidly decay with increasing frequency. This means that the high frequency components of the distribution affect the shading very little.

When considering shadowing, the visibility function is added to the clamped cosine function. This increases the value of the high-frequency components. With the higher coefficients, the high-order coefficients of the illumination have a greater effect on the appearance. So despite noise in the image formation process, it is possible to recover high-frequency components from shadowing.

Okabe et al. also develop a method to determine the maximum frequency of illumination which can be recovered from a specific scene. Depending on the scene geometry, they propose a method to calculate the band limit for the illumination estimation. They note that the limitation of the spherical harmonic approach are that the object typically occludes most of the shadow. This reduces the number of samples that can be taken. They also mention that natural lighting usually contains high-frequency components, which are in many cases difficult to recover using spherical harmonics.

Their Haar wavelet approach is similar, since Haar wavelets also form an orthonormal basis. The advantages they point out are the compact supports with various sizes which allow different resolutions in different regions. They claim sparsity is also important since very few coefficient can be used to represent a natural lighting setting.

The illumination distribution is mapped to a cube and two-dimensional Haar basis functions are used on each face. Similar to the spherical harmonics, the distribution can be expressed as a linear combination of the basis functions $\Phi_i(\theta,\phi)$ and $\Psi_{ijkl}(\theta,\phi)$ as

$$L(\theta,\phi) = \sum_i \left(c_i \Phi_i(\theta,\phi) + \sum_{j,k,l} d_{ijkl} \Psi_{ijkl}(\theta,\phi) \right), \tag{6}$$

where c_i and d_{ijkl} are coefficients of the corresponding basis functions. By substituting equation 6 into equation 3 a linear equation system is obtained. Then the coefficients are computed using a constrained least squares estimation which constrains the resulting distribution to be positive.

Experiments are performed on synthetic and real images. To evaluate the results a root mean square (RMS) error is computed for a re-synthesized scene under the estimated lighting. In all illumination conditions shown, the error is smaller for the Haar wavelet based method. They conclude that even though Haar wavelets might not be the optimal choice, they are still more suited for the distributions tested than spherical harmonics.

4 Bouganis and Brooks (2004)

The general method proposed in this paper [1] is similar to the work of Wei [10], and Wang and Samaras [7] [8] [9]. The method is based on Zhang and Yang's approach [11] and attempts to increase accuracy as well

as combat some of the limitations of the original approach.

The paper first formulates the problem with a special focus on the ambiguities pointed out in the original paper. Zhang and Yang noticed problems when relying on critical points for determining light source directions. Opposite lights share the same critical line so it is not possible to distinguish one from the other. In addition, when opposite lights are present in a scene, more than critical points are necessary to determine that both lights exist. Bouganis and Brooks show that together with undetectable sources (sources for which no critical points are visible), there can exist multiple light source set that all produce the same image. They formulate this in a general manner which has not yet before been stated. In their work they resolve the ambiguity by selecting the configuration of sources that contains the fewest visible lights.

They name their proposed algorithm *Virtual-to-Real* (V2R). Like Wang and Samaras' method, it uses virtual light sources of adjacent regions with constant illumination conditions. By subtracting the two virtual light sources $\mathbf{v}_1$ and $\mathbf{v}_2$ the pre-direction of the source associated with the dividing critical line is determined.

The search for critical points is similar to Wei [10] in that it considers a $w \times w$ window centered around a pixel p. For windows within a constant lighting region, it is possible to estimate the virtual light for that patch. With the normals for each pixel in the window as rows in the $\mathbf{N}_{W_p}$ matrix and the virtual light direction as $\mathbf{v}$, the intensities at each pixel in the window can be expressed as $\mathbf{e}_{W_p} = \mathbf{N}_{W_p}\mathbf{v}$. So the virtual light source can be determined by solving this linear equation system. If the window overlaps a critical line, the equation system will not be correct. The system can be solved using a least squares approach, then applying thresholding to the resultant mean square error f_p which is expressed as

$$f_p = \frac{1}{w^2}(\mathbf{Nv} - \mathbf{e})^T(\mathbf{Nv} - \mathbf{e}). \tag{7}$$

If f_p is above a selected threshold T, then the point p is selected as a critical point candidate. The regions with $f_p > T$ typically need to be thinned, and isolated pixels can be removed, so erosion and dilation operators are applied to the detected regions.

The Hough Transform is applied in a similar way as the original method. Virtual lights are estimated based on the segmentation defined by the critical lines. This estimate however also contains a bias term to account for ambient lighting and intensity offsets caused by digitization. A global bias estimate is estimated from the individual estimates and subtracted from the image intensities for further analysis.

The direction coefficients d_m which are introduced in their model which considers opposite light sources, are estimated using an optimization algorithm which minimizes the difference between the original pixel values and the synthesized results. Finally, in a refinement step, the directions are adjusted with a steepest-decent method to further reduce the errors of estimation.

Bouganis and Brooks analyse the selection of parameters such as window size, critical point threshold and critical line limit in detail, and report their results for their selected image size and noise conditions. They compare the results of their approach to the original method. In nearly all cases, the accuracy is higher, and the robustness to noise is increased.

5 Implementation

The focus of the implementation is the effects of a skeletonization method on the accuracy of light source detection. The basis code which was modified for this assignment uses a hybrid of Zhang and Yang's original method together with Bouganis and Brooks' proposed critical point detection method.

The original implementation uses the results from the Hough Transform as pre-directions for the light sources. The critical points are detected with sub-pixel refinement for increased accuracy. Bouganis and Brooks' method is supposedly more robust to noise, yet does not attempt a very high accuracy for detecting

critical points. The reason is likely because the resulting directions only serve for segmenting the image. The light source directions are determined from comparing adjacent virtual light sources.

The use of the mean square error with thresholding, results in many potential critical points. Keeping the threshold low ensures that all critical lines are detected, but it also causes many critical points to be falsely detected near true critical lines.

Bouganis and Brooks mention using dilation and erosion operators to reduce the thickness of the detected lines. For this assignment, the window-based critical point search method was adjusted according to Bouganis and Brooks' specifications and a skeletonization of the detected critical lines was added.

The skeletonization algorithm used is from the MATLAB imaging toolbox. The bwmorph(image,'skel',Inf) function performs skeletonization of binary images. The results were found to be good enough to study the benefits of skeletonization. Figure 1 shows how the skeletonization routine effectively reduces the thickness of the lines to approximately 1 pixel.

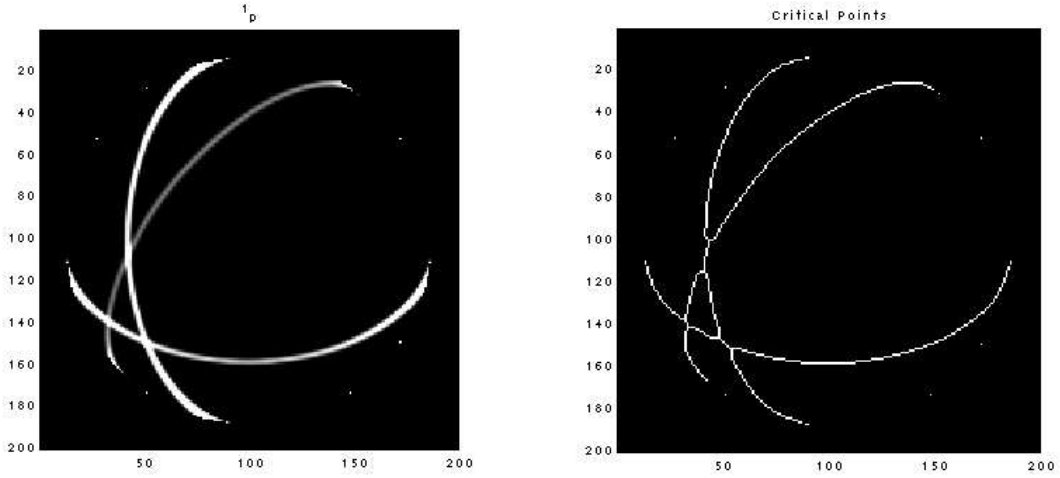

Figure 1: Get slim quick. Results after one function call, or get your money back. These images show the result of skeletonization on an example lighting arrangement with 3 light sources.

Without skeletonization, the Hough Transform space features rough plateus rather than distinct peaks. This was initially combatted by finding the average location of nearby peaks (all within the same plateau). When the plateaus are too large (caused by wide critical lines), the peaks are not grouped and false sources are reported. The skeletonization effectively changes the appearance of the Hough Transform space, by making the peaks more distinct. In the tests performed, the number of critical points is reduced from over 2400 to around 500.

An experiment was performed with the lighting configuration as shown in Figure 1. Results were obtained for three cases:

1. **Original implementation:** The original implementation with refinement was known to perform very well on this configuration. The sum of squared differences (SSD) error in the re-synthesized image is 10^{-4}, and the direction estimates are accurate within 0.2 degrees.

2. **Window-based critical point detection:** The window-based detection method resulted in a false detection of a light source. The SSD error was 4.2, and the maximum direction error was 4.5 degrees.

3. **Skeletonized critical points:** Using skeletonization permitted correct detection of all light sources, however still with a considerable error in their directions. The SSD error was 4.4, and the maximum direction error was 0.5 degrees.

It would be beneficial to conduct more tests, but the results are expected to be representative of the general advantages and disadvantages of each approach. The skeletonization improved the results by eliminating the falsely detected light source. The results are therefor more accurate than the estimates without skeletonization. But the accuracy still does not outperform the original approach. Adding noise to the image caused many false critical point detections when using the window-based method. Skeletonization does not help much in this case. It is expected that by adding the segmentation and virtual light source based direction estimation, the results would however be even more accurate than the original method as Bouganis and Brooks report.

References

[1] C.-S. Bouganis and M. Brookes. Multiple light source detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26(4):509–514, April 2004.

[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] Y. Li, S. Lin, H. Lu, and H. Shum. Multiple-cue illumination estimation in textured scenes. In *Proc. Ninth International Conference on Computer Vision*, pages 1366–1373, 2003.

[4] T. Okabe, I. Sato, and Y. Sato. Spherical harmonics vs. haar wavelets: Basis for recovering illumination from cast shadows. In *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, pages I–50–57, Washington DC., June 2004.

[5] 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.

[6] I. Sato, Y. Sato, and K. Ikeuchi. Illumination distribution from brightness in shadows. In *Proc. IEEE ICCV*, volume 2, pages 875–882, 1999.

[7] Y. Wang and D. Samaras. Estimation of multiple directional light sources for synthesis of mixed reality images. In *Pacific Graphics*, pages 38–47, 2002.

[8] Y. Wang and D. Samaras. Estimation of multiple illuminants from a single image of arbitrary known geometry. In *ECCV 2002*, pages III: 272–288, 2002.

[9] Y. Wang and D. Samaras. Estimation of multiple directional light sources for synthesis of augmented reality images. *Graph. Models*, 65(4):185–205, 2003.

[10] J. Wei. Robust recovery of multiple light source based on local light source constant constraint. *Pattern Recognition Letters 24(1-3)*, pages 159–172, 2003.

[11] Y. Zhang and Y.-H. Yang. Multiple illuminant direction detection with application to image synthesis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23(8):915–920, 2001.