Summaries of Selected Papers on Relighting

Nathan Funk
University of Alberta
Assignment 5 for CMPUT 605

December 3, 2004

1 Marschner & Greenberg

This relighting approach [2] includes an illumination estimation technique (referred to as ”inverse rendering”) coupled with a simple method for relighting the original photograph. It requires a geometric model of the scene which is obtained with a laser range finder in the experiments. The technique is successfully applied by relighting an image of a watering can as well as a picture of a human head.

The authors note in the introduction that it is often neccessary to make modifications to images. Typically this is performed with 2D image editing software. Although great results can be achieved by a skilled artist, tasks such as changing the lighting of a scene are still very difficult and tedious. With information about the 3D geometry of the scene, many new possibilities are opened.

The goal of the first part of the method is to recover the lighting distribution given an image, the camera position, and model of the scene. This procedure uses a least squares solution similar to Hougen & Ahuja's [1]. A renderer $\mathcal{R}$ maps the lighting configuration to an image. This mapping is modelled as a linear operator. The input for $\mathcal{R}$ is approximated by a finite-dimensional linear system. In other words, the lighting distribution is modelled as the sum of a set of basis functions on the sphere. A lighting distribution L can be expressed as $L = \sum_{i=1}^{n} \alpha_i L_i$, where L_i is one of the n basis functions and α_i is the associated coefficient. A image under an arbitrary lighting condition can then be expressed as the sum of the basis images. This can be written in matrix form as

$$\begin{bmatrix} \vdots \\ v_j \\ \vdots \end{bmatrix} = \begin{bmatrix} & \vdots & \\ \cdots & K_j(\mathcal{R}(L_i)) & \cdots \\ & \vdots & \end{bmatrix} \begin{bmatrix} \vdots \\ \alpha_i \\ \vdots \end{bmatrix}, \tag{1}$$

where K_j is a linear relationship of the image intensity to the pixel value. Using a least squares approach it is possible to determine the α_i values which determine the lighting distribution. The authors note the limitations such as required number of pixels and ill-conditioned nature of the system for most cases. They use a linear least squares algorithm with first-order linear regularization and a generalized singular value decomposition for interactive adjustment of the regularization parameter. They do not mention using a non-negativity constraint although it is likely necessary for an accurate lighting distribution to be obtained.

With the information about the existing lighting, it becomes possible to relight the image under novel lighting conditions. A simple approach would be to render an image of the 3D model under the new lighting conditions. But it is noted that the original picture contains details that could not be reproduced with this method since the 3D model is not accurate enough. To account for the inaccuracies of the 3D model, the ratio between the original image and the rendered image under the same lighting conditions is calculated at each pixel. This ratio image contains more detailed information including subtle surface textures for example. Then a new image is rendered from the 3D model under the new lighting conditions. Finally, this image is multiplied pixel-by-pixel with the ratio image. So the ratio between real and synthetic images is the same for both lighting conditions. Why this is appropriate is not thoroughly justified.

The method is tested on a watering can which was painted with flat paint to achieve a nearly completely diffuse reflecting surface. This object was scanned with a laser range finder to determine the 3D shape. The results of the relighting are good, but since the flat paint did not completely illiminate specular reflection there are some minor noticable artifacts. Although the authors did not comment on this, the inaccuracy in the light recovery also appears to have an effect on the final image.

A second experiment is discussed where a human head is composited into a new scene with the attempt to relight the head in the new environment. The same problems are visible in the results as in the previous experiment. The specular reflections from the face are not captured in the rendering so the ratio image contains the specular highlights and brings them into the final image. So the final composited image contains quite accurate new diffuse reflection, but together with the old specular reflection. The visual effect is not too disturbing in the example shown but is expected to be worse in other cases.

2 Ng et al.

Relighting a scene can be performed by rendering an image of an accurate 3D model as noted in the previous section. If however the scene needs to be relit under many different lighting conditions, perhaps even in real-time, a simple rendering approach will typically not be fast enough. For this reason, methods which precompute lighting information and allow quick generation of an image under arbitrary lighting are desired. This paper [4] is one of them.

Prior to this paper, Sloan et al. [6] proposed a pre-computed radiance transfer method which focused on low-frequency lighting environments. It allows fast rendering of an image but is limited to distant lighting that changes smoothly in space. With this method, the illumination can also contain high frequency components. The main reason is the transition from a spherical harmonic basis to a wavelet basis.

So called high-resolution transport matrices are pre-computed. They combine the BRDF, visibility and geometric factors into a 6x64x64 cube map. This transport matrix is projected onto a wavelet basis, making it sparse due to the coherence of the data. Then the rendering can be performed quickly using a sparse matrix multiplication algorithm.

Two different cases are considered in the analysis. For *geometry relighting* the viewpoint is allowed to change but the surfaces are constrained to be diffuse reflecting. The second case is *image relighting* where the viewpoint is fixed, but arbitrary reflection models are allowed for the surfaces. When illuminating the scene with a distant environment map, the pixel intensity $B(\mathbf{x}, \omega_o)$ at a point in the scene $\mathbf{x}$ can be calculated with

$$B(\mathbf{x}, \omega_o) = \int\int_{\Omega} L(\omega) S(\mathbf{x}, \omega) f_r(\mathbf{x}, \omega \to \omega_o)(\omega \cdot n(\mathbf{x})) d\omega, \tag{2}$$

where $\mathbf{x}$ is the location of the point, ω_o is the viewing direction, ω is the incident direction, L is the environment map, S is the visibility function, and f_r is the reflection function at location $\mathbf{x}$. Depending on the case, a different transport function can be constructed based on the assumptions. For *geometry relighting* the f_r depends only on the location, so the transport function can be expressed as

$$T(\mathbf{x}, \omega) = L(\omega) S(\mathbf{x}, \omega) f_r(\mathbf{x})(\omega \cdot n(\mathbf{x})). \tag{3}$$

For *image relighting*, the viewing direction ω_o depends only on the point location $\mathbf{x}$, so the transport function can be written as

$$T(\mathbf{x}, \omega) = S(\mathbf{x}, \omega) f_r(\mathbf{x}, \omega \to \omega_o(\mathbf{x}))(\omega \cdot n(\mathbf{x})). \tag{4}$$

Since for both cases ω_o is now either not present or is a function of $\mathbf{x}$, we can now write $B(\mathbf{x}, \omega_o)$ as a function of only $\mathbf{x}$. Together with numerical cubature of the integral (the normalization weights are not shown), it may be written as

$$B(\mathbf{x}_i) = \sum_j T(\mathbf{x}_i, \omega_j) L(\omega_j), \tag{5}$$

or in matrix notation as $B = TL$.

In this matrix formulation it is possible to convert the representation of L from an explicit tabular form to any orthonormal basis such as spherical harmonics or wavelets. This requires each row of T to be projected onto the same basis. In this paper, the 2D Haar wavelet basis is applied to each cubemap face. If the full transform was applied to the light distribution and transport matrix, the results would be just as accurate as before but also require just as much computation. By quantizing and discarding all zero coefficients, the representations become sparse and make sparse matrix multiplication applicable.

Three different schemes for selecting wavelet basis lights for a specific distribution are analysed. From these three, the *area-weighted selection*, which scales the priority of each wavelet light by its area, is determined to give the least image error. When using 100 coefficients to represent the lighting compared to the same number of spherical harmonics, the wavelet based representation consistently results in a lower image error.

The main conclusion drawn in this paper is that wavelets are much more appropriate for representating certain lighting distributions than spherical harmonics. They do note that using wavelets for mostly smoothly varying distributions only provides a very small or no improvement over spherical harmonics.

3 Séquin & Smyrl - Parameterized Ray Tracing

Similar to the previous technique, this method [5] also assumes that the scene is modelled in the computer. Through use of precomputed information, not only the lighting, but also other parameters of the scene can be changed with little computational time to produce the final resulting image. Since this paper was published in 1989, a significant amount of effort is put into reducing the memory and computational requirements so the approach could be implemented on the current computers. Although the goal of this paper is not stated as relighting, it is still essentially the same. The authors describe the process of artists rendering a scene and needing to quickly adjust certain parameters such as surface properties or lighting. This is the main motivation.

The general idea is to store an expression for each pixel then simplify the expression based on what can be assumed to be constant. For example if only the lighting will be changed, then all other parts of the expression are evaluated and stored. After this, generating an image only requires substituting in the variable values and evaluating the expression.

The shading equation used allows for reflective and transparent surfaces in addition to the standard ambient, diffuse and specular terms. The equation used for the intensity calculation is

$$\mathbf{I} = \mathbf{i}_a k_d \mathbf{c} + \sum_{lights} \mathbf{i}_l k_d \mathbf{c}(\hat{n} \cdot \hat{l}) + \sum_{lights} \mathbf{i}_l k_s (\hat{n} \cdot \hat{h})^p (m\mathbf{c} + (1-m)) + \mathbf{I}_s k_s (m\mathbf{c} + (1-m)) + \mathbf{I}_t k_t \mathbf{c}, \tag{6}$$

where the individual terms in order are ambient, diffuse, specular, reflected, and transmitted contributions. The geometry coefficients are $\hat{n}$, $\hat{l}$, and $\hat{h}$. None of these can be changed after the precomputation step. All the remaining variables such as the light source intensities $\mathbf{i}_l$ can be changed in the final implementation.

Since transparent surfaces and specular reflection between surfaces is considered, the intensity of a single point can also depend on the intensity of other surfaces. A tree is constructed where each node corresponds to one surface interaction. For each reflection or transparent component a child node is created. So during evaluation of the expression tree, the child nodes are evaluated first, then their results are passed up the tree to determine the final pixel intensity. Since the surface properties are assumed to be constant for an entire object, the information stored in each node can be shared among all trees containing a node for one object. This greatly reduces the amount of memory necessary.

Before optimizing the program, a preliminary stage was implemented for which only the surface color $\mathbf{c}$ can be changed. They noted that the memory consumption and computational time of the initial implementation were too high to be practical. This prompted a number of refinements to optimize the renderer. One optimization is to share the expression trees if the structure and coefficients are similar enough. Simply comparing all pixel intensities to find similar pixels was not considered acceptable since it can produce coincidental matches. Only grouping of pixels generated with similar expressions is desired. Another optimization is the selective expression updating. This is accomplished by setting flags in nodes to indicate whether the surface properties of that node have been changed. For example if the color of a single object is being changed, only certain areas of the image will be affected. The other regions do not need to be re-evaluated.

Further optimizations include runlength encoding and 2-D runlength encoding, which take advantage of horizontal and vertical coherence. A special hashing scheme is also implemented which uses the similarity of subtrees to quickly determine which subtrees can be grouped together. All these optimizations reduce the memory usage for an example scene from over 20Mb to 1Mb. With the optimizations implemented, the algorithm could be generalized to allow modification of not only the surface color but all other surface properties and the lighting. The final implementation allows updates to the scene to be calcluated within 8 seconds which is less than 1/10th of the initial ray tracing time.

When examining the intensity equation (Equation 6), it can be noted that if the ambient term was removed, the entire equation could be expressed as a single sum over the light sources (similar to the Ng et al. approach). This difference between Ng and Séquin's work allows Ng's approach and similar techniques to benefit from the formulation as matrix multiplication. It is likely that this formulation could also be exploited in Séquin's more general framework to improve the overall performance of the approach.

4 Masselus et al. - Relighting with 4D Incident Light Fields

This method is an image-based relighting technique [3] which allows relighting of a scene from a single viewpoint with a 4D light field. The main idea is to take a set of basis images under controlled lighting conditions then combine these basis images to relight the scene under an arbitrary 4D light field. This concept has been implemented earlier but not for spacially varying illumination.

With this method the authors wanted to avoid using a 3D model of the scene. For many objects such as the hairy monkeys used as example in this paper, the creation of a 3D model is difficult, whether it is modelled by hand or using a computer vision technique. So among the other relighting techniques that use 3D information to estimate the results, there is still need for techniques that operate on this set of difficult-to-model objects and surfaces.

The model for the incident light field is four-dimensional, meaning that an incoming intensity measurement is taken with respect to four different parameters. The first two parameters are the azimuth angle ϕ_p and the tilt angle θ_p. They identify a point on a hemisphere at which the incident light map is measured using the two other parameters ϕ_a and θ_a. The 4D space of all possible incident directions Ω is then $\Omega = [0, 2\pi] \times [0, \pi/2] \times [0, 2\pi] \times [0, \pi/2]$. Within this space, the four parameters can be grouped together as $\Theta = (\phi_p, \theta_p, \phi_a, \theta_a)$, and the incident light field can be written as $L_{incident}(\Theta)$. Then, the exitant light can be expressed as

$$L_{exitant}(x, y) = \int_{\Omega} R(\Theta, x, y) L_{incident}(\Theta) \mathrm{d}\mu(\Theta), \tag{7}$$

where $R(\Theta, x, y)$ is the reflectance field and $\mu(\Theta)$ is a measure in space Ω.

A set of basis functions B_i are defined by partitioning the space Ω into partitions Ω_i and letting each basis function have the value 1 within Ω_i and 0 outside that partition. The authors note that this is a very simple approach to defining the basis functions and suggest that other basis functions will have other benefits. The incident light can be approximated as a linear combination with

$$L_{incident}(\Theta) \approx \sum_{i=1}^{N} l_i B_i(\Theta), \tag{8}$$

where N is the number of basis functions and l_i are the intensity coefficients for each basis function. Substituting this into equation 7 allows the exitant radiance to be expressed as

$$L_{exitant}(x, y) \approx \sum_{i=1}^{N} l_i R_i(x, y), \tag{9}$$

where $R_i(x, y)$ is calculated by integrating the reflectance function over the partition Ω_i. This $R_i(x, y)$ can also be understood as the image resulting from illumination B_i. Each $R_i(x, y)$ can be captured by controlling the lighting, and taking a picture of the scene.

In their setup, the authors use an LCD projector which is mounted on a movable gantry controling the azimuth angle ϕ_p. The tilt angle is controlled by rotating the object on a turntable. The light projection samples the directions ϕ_a and θ_a within a limited range. The entire range is not neccessary since the region of interest is contained within the sampled region.

In the experiments, 32x7 projector positions were used and a total 16x16 basis images were projected at each position. The images captured resulted in approximately 10Gb of data and took 41 hours to capture. Since the positioning of the projector is not automated, this is a very tedious task. A speedup method is proposed in which entire rows and colums of basis illumination patches are projected at once. Then, by comparing the images from the row and column illumination, the areas that are illuminated in both images of a specific row-column pair define the reflectance images required. This speedup method reduces the complexity from O(M^2) to O(M) where M^2 is the number of patches being projected.

To relight the scene with a specific illumination field, the coefficients for each of the basis images need to be determined. Then the basis images are linearly combined resulting in the final image. The results show a variety of objects under different lighting conditions. One common problem is the aliasing effect present in the lighting. This is due to the low number of basis functions used. Using more functions would however result in even longer aquisition time and larger storage requirements. The authors mention that other basis functions such as Gaussian patterns instead of box patterns would remove the aliasing effects.

References

[1] 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.

[2] S. Marschner and D. Greenberg. Inverse lighting for photography. In *Proc. of the Fifth Color Imaging Conference*, pages 262–265, November 1997.

[3] Vincent Masselus, Pieter Peers, Philip Dutré, and Yves D. Willems. Relighting with 4d incident light fields. *ACM Trans. Graph.*, 22(3):613–620, 2003.

[4] Ren Ng, Ravi Ramamoorthi, and Pat Hanrahan. All-frequency shadows using non-linear wavelet lighting approximation. *ACM Trans. Graph.*, 22(3):376–381, 2003.

[5] C. H. Séquin and E. K. Smyrl. Parameterized ray-tracing. In *Proc. 16th annual conference on Computer graphics and interactive techniques*, pages 307–314. ACM Press, 1989.

[6] Peter-Pike Sloan, Jan Kautz, and John Snyder. Precomputed radiance transfer for real-time rendering in dynamic, low-frequency lighting environments. In *Proc. 29th Conf. Comp. Graphics and interactive techniques*, pages 527–536. ACM Press, 2002.