Deep shadows in a shallow box

Xiang Huang, Ankit Mohan and Jack Tumblin

Department of Electrical Engineering and Computer Science
Northwestern University
Evanston, IL, 60201, USA

ABSTRACT

We present a fast, low-cost technique to gather high-contrast 'relightable' photographs of desktop-sized objects. Instead of an elaborate light stage, we follow Mohan et al.; we place the object and a digitally-steered spotlight inside a white cardboard box, aim the spotlight at the box interior, and move the spot to light the object from N repeatable lighting directions. However, strong ambient lighting from box interreflections causes 'shallow' shadows and reduces contrasts in all basis images. We show how to remove this ambient lighting computationally from the N images, by measuring an $N \times N$ matrix of coupling factors between lighting directions using a mirror-sphere light probe. This linear method, suitable for any light stage, creates physically accurate 'deep shadow' basis images, yet imposes only a modest noise penalty, and does not require external light metering or illumination angle measurements. Results from our demonstration system support these claims.

Keywords: image based relighting, rendering

1. INTRODUCTION

Stray unwanted light limits the performance of almost all designs for 4D, 6D, and 8D light-stages. These devices, intended to capture photographs of objects lit from a wide range of controllable illumination directions, include many ingeniously diverse designs, ranging from large, elaborately complex high-performance systems for real-time capture for actors and dancers[1] to a simple hand-held light and a video camcorder.[2,3] These various light stages each capture a set of directionally-lit photographs or 'basis images', raw data variously processed to describe illumination fields, reflectance fields, and derived quantities such as silhouettes, surface normals, and detailed 4D,6D and 8D reflectance estimates.

For re-lightable photographs, good basis images capture high dynamic range (HDR) images of objects lit from a single direction, forming deep, sharp shadows and rich, directionally-lit textures. Weighted sums of these basis images using positive-only weights produce images that exactly match those made by applying the weights to the light sources themselves.[4] While these re-lighting results are physically accurate, they cannot exceed the contrasts of their basis images.

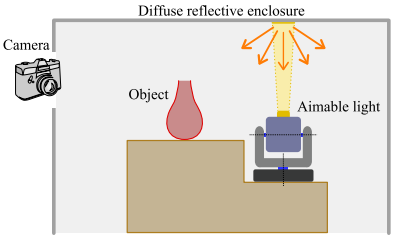
Ideally, all the light that arrives at the photographed object in a light-stage comes directly from a computer-controlled light source, but any practical system will also have unwanted secondary light sources, such as light leakage from lamp ventilation holes, scattering from smudges on light-source lenses, inward-pointing light sources that illuminate each other, and broad light beams that spread light beyond the photographed object. Every illuminated surface of the apparatus reflects some light, and multiple interreflections scatter measurable ambient light everywhere. In each basis image, this ambient light "fills in" surface cracks and deep shadows to form a shallower, lower-contrast photo. Accordingly, many light-stage designs combat stray light by absorbing or or re-directing it; open-frame designs let light escape, flat-black paint reduces interreflections, and baffles, light traps, and black-velvet curtains around light sources can absorb most of the light energy of each interreflection step. However, these stray-light reduction strategies are difficult or impossible for one particularly convenient, low-cost light-stage design introduced by Mohan et al.,[5,6] show in Figure 1(a))

Using this or a similarly simple light stage, we wish to make high quality "relightable" digital photos of desktop-sized objects as easily, quickly, simply, and cheaply as possible, because we believe such photos could be

xhuang, ankit, jet@cs.northwestern.edu

a tremendous help to museums as they devise sensible strategies for digital archiving of their collections. It may help them to share, conserve and compare the vast set of precious objects in their collections that otherwise may never be included in a public display. Simple, low-risk designs encourage widespread use: rather than design a more elaborate light stage to minimize ambient light, we chose to remove the ambient light computationally.

Our simple light stage borrows its basic design from Mohan et al.[6] We place the object inside a diffuse white cardboard box, and aim a steerable spotlight or disco light at its inner surface, and the bright spot projected on the box interior acts as our movable light source. We use a personal computer to place the spot precisely and repeatably at a series of N locations spread across the entire hemisphere of incoming light directions, and trigger the camera to capture each of these N directionally-lit basis images.



(a) Disco light casts bright spot on the foam-core box. Ambient interreflections act as secondary unwanted light sources.[5]

(b) Ambient light weakens shadows caused by the bright spot.

(c) Computed deep shadow image removes ambient light.

Figure 1.

Ideally, the object will receive light only from the one bright spot on the box interior, but as we can see in Figure 3(b), interreflections within the box distribute small fractions of that light to each of the other $N-1$ light-spot locations as well; these other spots are not black. Instead, each one acts a weak directional light source as well, and the intensity of each source is a fixed fraction, a 'coupling fraction' of the bright-spot's intensity, depending on spot location, box shape, and its surface material.

Suppose we could somehow measure all those coupling fractions for all the basis images; each fraction would tell us just how much of the unwanted light came from each of the unwanted illumination directions in each of the basis images. Recall that each of the N basis images shows us how the object looks when lit *mostly* from just one chosen direction. If we subtract just a tiny fraction of one basis image A from another basis image B, then we will remove some of the unwanted illumination in B from the direction of the bright spot in A. As Section 3 will show, we can measure a matrix of coupling fractions with a light-probe, and use it to find just how much of each basis image we should subtract to remove nearly all the contributions from all ambient light to all basis images.

Figure 1 (b) shows an unmodified basis image with shallow shadows caused by box interreflections, but applying our ambient light removal method produces (c), a new, non-negative, physically-realizeable basis image with deep shadows as if the box itself had absorbed all ambient light. We will describe how to make these results after a brief review of previous work.

2. RELATED WORK

Image-based relighting is a powerful approach to display real objects under given lighting conditions in computer graphics.[3,4,7]

Like most previous image-based lighting methods, we apply the observations formalized by Nimeroff[8] that lights and materials interact linearly. The basic principle is that a photograph of an object lit from each direction provide a basis for relightable photographs. A weighted sum of these photographs will match a photo of the object lit by a weighted sum of light sources.

Given the light intensities and light directions, Debevec et al.[4] showed how to compute 8D directional reflectance as well. Using an early lightstage design, Hawkins et al.[7] made a pioneering application to museum tasks by measuring reflectance with a rotating-arm that hold multiple lights. To measure each lighting direction, they made a second set of photographs where a mirror ball light probe replaced the object, just as we have done. Malzbender et al.[9] avoided mirror ball calibrations by using a small geodesic dome of electronic flash units for capturing directional illumination effects. Sen et al.[10] applied Helmholz reciprocity to interchange light source and camera positions in an attempt to speed up the image capture process. To reduce the complexity and cost of equipment, several other researchers devised novel ways to capture lighting basis images with handheld light sources.[3,5,11,12] Others found lighting directions without mirror ball light probes, such as the dimensionality reduction method of Winnemöller et al.,[2] or with handheld light sources ,[3] or the planar light probe of Alldrin.[13]

To keep our costs low, we apply the simple card-board box light stage introduced by Mohan et al.,[5] but we found that inter-reflections within that box severely reduced basis image contrasts. In the next section, we describe a practical solution to this problem. Our solution can be applied to any of above light-stages to improve the contrast of resulted image.

3. PROBLEM FORMULATION

Under a fixed viewpoint, the amount of light a camera pixel (x, y) receives is a linear combinations of incoming light, because radiance is additive:[12]

$$I(x,y) = \int_{\Omega} R_{xy}(\omega) L(\omega) d\omega. \tag{1}$$

Here $L(\omega)$ denotes the incoming light distribution on incoming angle $\omega = (\theta, \phi)$. $R_{xy}(\omega)$ is called the reflectance field.[4] It represents how much light is reflected to the pixel (x, y) when the object is illuminated by unit incoming light from direction ω. The reflectance field $R_{xy}(\omega)$ subsumes the geometric property as well as the BRDF of the object. (Each R, G, B color channel is processed separately in our paper.)

In an ideal light stage the incoming light only comes from one direction: $L_i(\omega) = \delta(\omega_i)$, where $\delta()$ is the Dirac delta function. According to Equation 1, $I_i(x, y) = R_{xy}(\omega_i)$, so the ith captured images is the direct measurement of R sampled at directions ω_i. A linear combination of captured images of the object produces a synthetic image with the new illumination in any form of sampled illumination.

In a real light stage such as ours, some nonzero amount of light will arrive from directions other than the desired direction. To compute R, we divide the the whole hemisphere of incoming light directions into N small solid angle regions $\Omega_1, \Omega_2, \cdots, \Omega_N$. Equation 1 becomes:

$$\begin{aligned} I(x,y) &= \sum_{i=1,\cdots,N} \int_{\omega \in \Omega_i} R_{xy}(\omega) L(\omega) d\omega \\ &= \sum_{i=1,\cdots,N} \left(R_{xy}(\omega_i) \int_{\omega \in \Omega_i} L(\omega) d\omega \right), \end{aligned} \tag{2}$$

The second equality holds according to the first mean value theorems for integration,[14] where ω_i is some angle in region Ω_i. When Ω_i is small enough, we suppose ω_i is approximately the center of Ω_i. In actual computation, we partition the incoming light space by the Voronoi diagram of the centers of bright spots as shown in Figure 2.

The incoming light can be measured by taking pictures of a mirror ball in the same place or near the object (see Section 4.2 for details). We have N unknown variables $R_{xy}(\omega_i)$ in Equation 2. By taking N photos $I_1, I_2, \cdots, I_N$ under different lighting conditions by aiming the spotlight on different positions all over the cardbox. We have N equations, which can expressed as:

$$\begin{bmatrix} I_1(x,y) \\ I_2(x,y) \\ \vdots \\ I_N(x,y) \end{bmatrix} = \begin{bmatrix} L_{11} & L_{12} & \cdots & L_{1N} \\ L_{21} & L_{22} & \cdots & L_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ L_{N1} & L_{N2} & \cdots & L_{NN} \end{bmatrix} \begin{bmatrix} R_{xy}(\omega_1) \\ R_{xy}(\omega_2) \\ \vdots \\ R_{xy}(\omega_N) \end{bmatrix}, \tag{3}$$

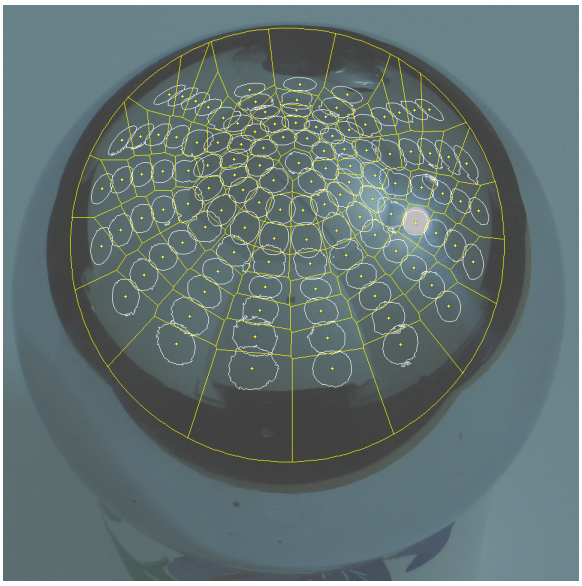


Figure 2. White circles show the positions of the bright spots. Yellow lines show how we divide the hemisphere of incoming light directions by the Voronoi diagrams of the bright spot centers.

or in compact matrix form by

$$\mathbf{I} = \mathbf{LR}, \tag{4}$$

where $L_{ij} = \int_{\omega \in \Omega_j} L_i(\omega) d\omega$ is the amount of light coming from solid angle Ω_j in the ith capturing.

The solution of $\mathbf{R} = (R_{xy}(\omega_1), R_{xy}(\omega_2), \cdots, R_{xy}(\omega_N)^T)$ is then given by

$$\mathbf{R} = \mathbf{L^{-1}I}. \tag{5}$$

We can process each pixel (x, y) using the same L matrix. Because the lighting beams originated from one point on the box to different positions of the object have almost the same angle, when the target object size is much smaller than the box size.

4. EXPERIMENT

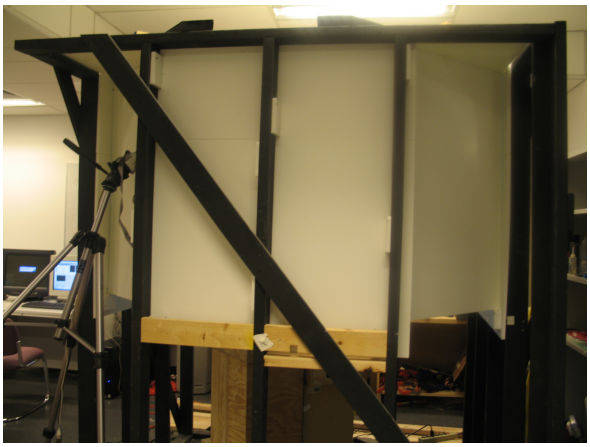
4.1 Equipment Setup

In our experiment, we need to take a large set of photos viewed at a fixed position under a large variety of different lighting directions. Although it is possible for us to manually move the light for every image like as was done by Masselus2002 et al.,[3] it can be extremely time consuming and tedious. In addition, the results are not repeatable, as we may accidentally move the light source closer or farther away from the object.

The center of the experiment is a special piece of equipment we designed to help us collect data for analysis. It is a simple box-like enclosure made of wood, lined with white foam-core boards that serves as a reflector. The white foam-core boards are specially treated with a matte finish paint to better approximate a Lambertian surface. Coupled with a computer-steered moving head spotlight, the box allows us to illuminate our object of interest from almost any angle. As we move the light around using the computer control software, we use our camera to record photos. Our setup is intended for high-quality photographic lighting for small objects.

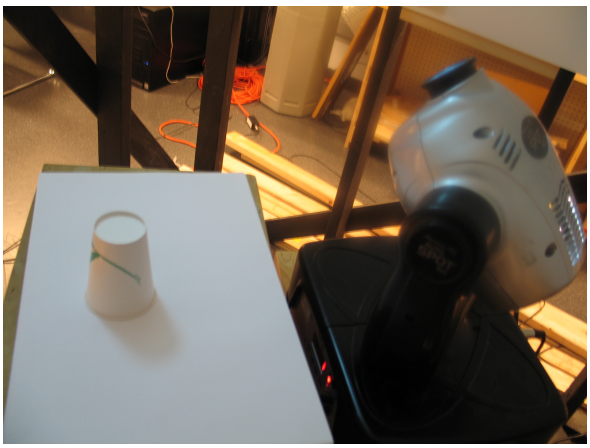
We place the object in the center of the equipment for photographing. To record the amount of light from each part of the reflector, we replace the object with a mirror ball and photograph the mirror ball.

Figure 3 shows our real light stage.



(a) overall view

(b) disco light casts bright spot on the foam-core box



(c) the object is placed inside the box

(d) a mirror ball is placed inside the box

Figure 3. Our light stage setup.

4.2 Methodology

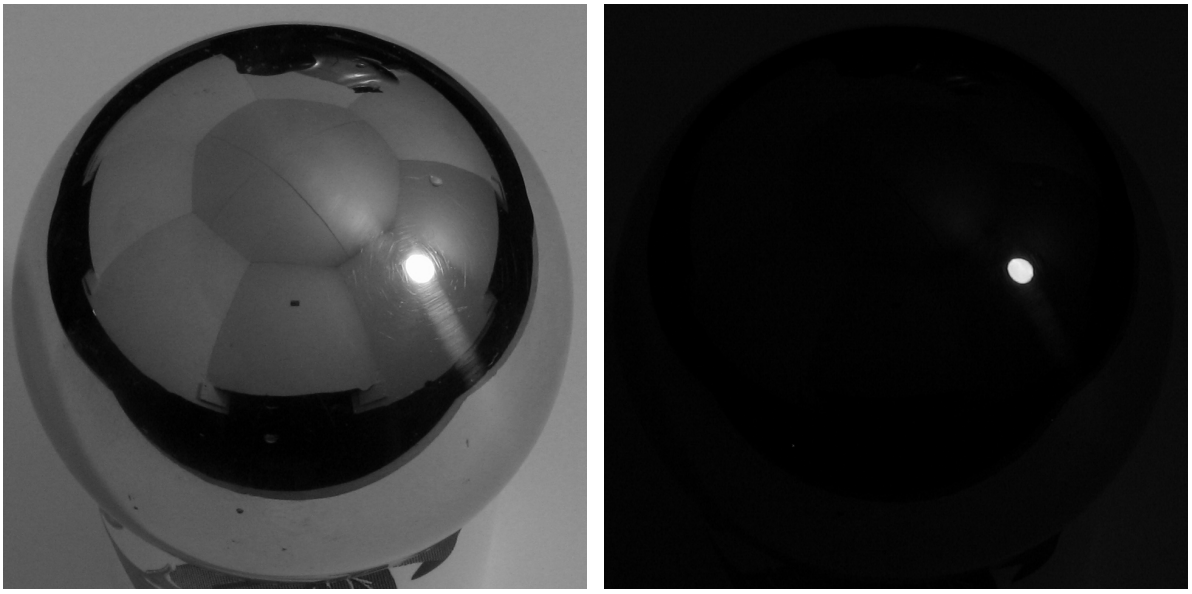
1. Collect images of the objects from a wide range of lighting directions. Place the object of interest in the center of the rig. Use the automated computer software to take object images from $N = 119$ different disco light aiming angles. The photos are taken by a fixed camera positioned outside of the rig, near the top of the rig.

2. Replace the object with a highly reflective mirror ball. The mirror ball images capture the illumination directions and intensities in the rig for each disco light aiming angle. To collect high dynamic range (HDR) photo, we took images at two different exposures for each aiming angle, as shown in Figure 4. We use the technique of Debevec and Malik[15] to merge two images with different exposure times.

   We assume the card-board has a perfectly diffuse surface. When we illuminate one point on the card-board, it reflects light evenly in all directions, so that it evenly illuminates every point on the photographed object (or mirror sphere); it will not make the front of the object brighter or darker than the left, right, top, or bottom.

3. Compute the L matrix. Each pixel in the photo of a mirror sphere describes the radiance leaving one point P on the hemisphere in the direction of the mirror sphere. It is the radiance received everywhere on the object from that point P, because we know that point P emits light evenly in all directions.

   $L_{ij} = \int_{\omega \in \Omega_j} L_i(\omega) d\omega$ is the amount of light coming from solid angle Ω_j in the ith capturing.



(a) Mirror ball long exposure 1 second (b) Mirror ball short exposure 0.05 second

Figure 4. To avoid under-exposure or over-exposure problem, we take pictures with two exposure time.

Once the data is collected, we want to know the radiance of light from all N directions for each of the N light aiming positions, to fill our L matrix. Since we have one image for each light direction, each image of the mirror ball would have one single bright spot. We can measure the radiance by measuring the pixel value of the center of the bright spot.

The radiance $L(x \leftarrow \Psi_i)$ can be computed by a mirror ball image that records the incident illumination conditions at a particular point in space.[16] A mirror ball has constant reflectivity and is made by highly specular steel. We place a mirror ball in the same place as the object and capture the image with the camera in the same position. The irradiance $L(x, \leftarrow \Psi_i)$ is proportional to the the radiance $E(x, y)$, where the point x makes light incoming angle Ψ_i equal to the light outgoing angle Θ. Section 4.2 shows how the to find the bisection point x for each directions.

We assume the distance from the camera to mirror sphere is large enough to reasonably approximate the camera projection as orthographic. In other words, all rays from this sphere to the camera are very nearly parallel. This is shown in Figure 5(a).

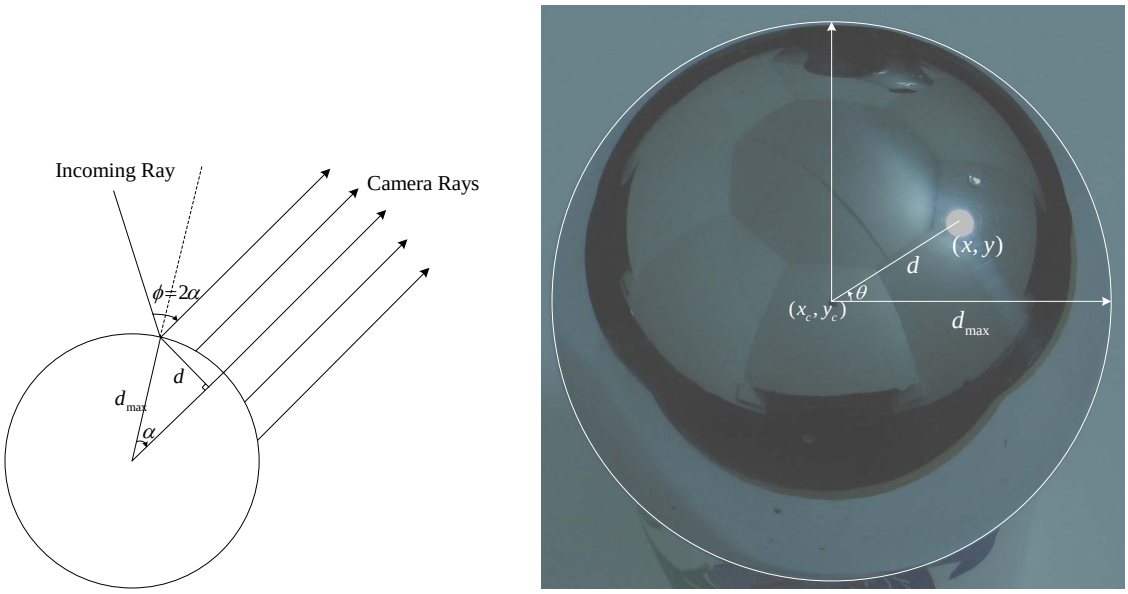
Each pixel in this mirror ball image measures a different fraction of solid angle, and this fraction depends on the pixel's distance from the center of the sphere's image as shown in Figure 5(b). Let us examine the mapping from pixel position (x, y) to illumination direction $\omega = (\theta, \phi)$. First, we find the pixel (x_c, y_c) at the center of the mirror ball image and measure the distance d_{max} from that center to the sphere's silhouette. Next, we find the distance d from camera pixel (x, y) to to (x_c, y_c) and divide this distance by d_{max}. The incoming lighting angle $\omega = (\theta, \phi)$ can be computed according to $\phi = 2\alpha = 2\arcsin(d/d_{max})$, and $\cos(\theta) = (x - x_c)/d, \sin(\theta) = (y - yc)/d$.

4) Use the inverse of L matrix to compute the deep-shadow image. This step is straightforward by using equation (5). We process each color channel and each pixel of image individually. Making 119 ambient-free 3072×2304 resolution images takes around one hour in Matlab 7.0 running on an computer with Intel Core 2 Duo $2.13GHz$ CPU and $2GB$ memory. As all the operations are linearly combinations and matrix multiplications, it should be straightforward to accelerate by GPU programming.

5. RESULTS AND ANALYSIS

5.1 Results

Our deep shadows algorithm is able to selectively enhance the contrast of shadows in the image without affecting other parts of the image, which Histogram equalization is not able to do.



(a) Light reflected to the camera by the mirror ball

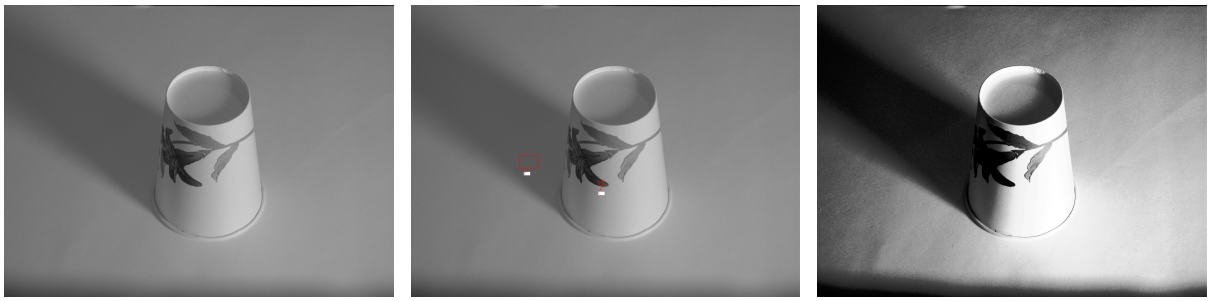
(b) Pixel position on the sphere's image indicates the angel of corresponding incoming light

Figure 5. Compute the L matrix by mirror ball images.

Even though histogram equalization can also increase the local contrast of an image, its result is not equivalent to ambient light removal. Through the adjustment, the pixel intensities is better distributed on the histogram: the most frequent intensity values are spread out. We show the original image, deep shadow image, and histogram equalized image in Figure 6 and their histograms in Figure 7.

Our result cannot be obtained by using histogram equalization. First, in histogram equalization, two image pixel of different positions with the same pixel values will be mapped to the same pixel value, this is not the case in our method. Second, in histogram equalization, the output image gray level is a monotonically non-decreasing function of the input, while our method is not. I marked two interesting regions in Figure 6(b). The printed flower petals of small region is darker than the shadow of large region both in original image (Figure 6(a)) and histogram equalized image (Figure 6(c)), but brighter in our deep-shadow image (Figure 6(b)).

We show another original image and deep shadow image in higher resolution in Figure 8.



(a) Original Image

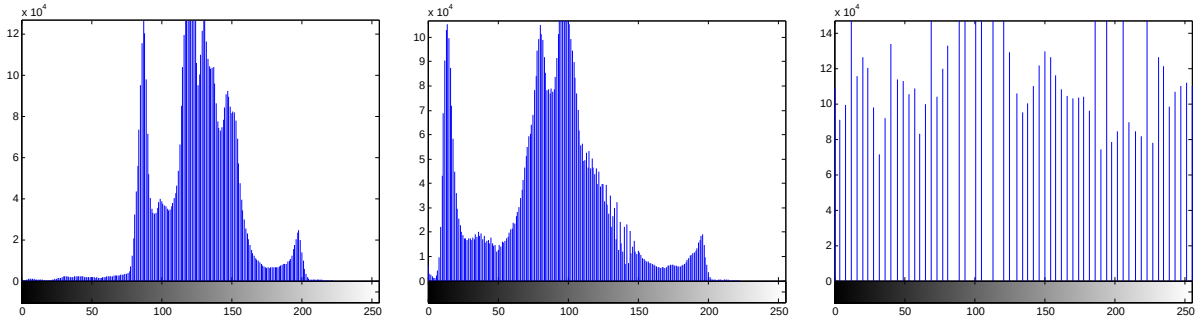
(b) Deep shadow image

(c) Histogram equalized image

Figure 6. Comparing original image, deep shadow image and histogram equalized image: although we have deep shadows and high contrast after histogram equalization, we lost the details of the flowers on the cup!

6. CONCLUSIONS AND FUTURE RESEARCH

There also seems to be some quantization noise (contouring) which is sensible for 8-bit images. One possible solution is to combine our images in the gradient domain instead of the intensity domain. It is shown that



(a) Histogram of the original image (b) Histogram of the deep shadow image (c) Histogram of the histogram equalized image

Figure 7. Comparing Histograms of original, deep shadow, histogram equalized images in Figure 6.

(a) One Input Image Directly Captured by Camera (b) One Deep Shadow Image Synthesized by Linear Combinations of N Captured Photos

Figure 8. Experimental Results, Green Channel Only

Poisson solvers hide zero-mean quantization of gradients very well.[17] This suggests that after finding the weights for each photo, we take the gradients of each photo, make a weighted sum of gradients, then use a Poisson solver to reconstruct the final result. The quantization would be in the gradients and not the intensities of the image, and over-constrained: Poisson solver gives smoother result.

We illuminate the scene by hundreds of spot samples at different positions. An interesting and basic question is what are the best positions and what are the best shapes for the light patterns.

7. ACKNOWLEDGEMENT

We would like to thank Jiang Duan for discussing the HDR concepts. We would like to thank Jiang Duan and Chi-Yin Cheung for helping and feedback with the experiments. This material is based upon work supported by the National Science Foundation.

REFERENCES

1. A. Wenger, A. Gardner, C. Tchou, J. Unger, T. Hawkins, and P. Debevec, "Performance relighting and reflectance transformation with time-multiplexed illumination," *Proc. SIGGRAPH '05, Transactions on Graphics* **24**(3), pp. 756–764, 2005.
2. H. Winnemöller, A. Mohan, J. Tumblin, and B. Gooch, "Light waving: Estimating light positions from photographs alone," *Computer Graphics Forum* **24**, pp. 433–438, sep 2005.

3. V. Masselus, P. Dutré, and F. Anrys, "The free-form light stage," in *In Proceedings of the 13th Eurographics workshop on Rendering*, pp. 247–256, 2002.
4. P. Debevec, T. Hawkins, C. Tchou, H.-P. Duiker, W. Sarokin, and M. Sagar, "Acquiring the reflectance field of a human face," in *Proc. SIGGRAPH 2000, Computer Graphics Proceedings, Annual Conference Series*, K. Akeley, ed., pp. 145–156, Addison-Wesley, (Reading, MA), 2000.
5. A. Mohan, J. Tumblin, B. Bodenheimer, C. Grimm, and R. J. Bailey, "Table-top computed lighting for practical digital photography," in *Rendering Techniques*, pp. 165–172, Eurographics Association, (Aire-la-Ville, Switzerland), 2005.
6. A. Mohan, R. Bailey, J. Waite, J. Tumblin, C. Grimm, and B. Bodenheimer, "Tabletop computed lighting for practical digital photography," *IEEE Transactions on Visualization and Computer Graphics* **13**(4), pp. 652–662, 2007.
7. T. Hawkins, J. Cohen, and P. Debevec, "A photometric approach to digitizing cultural artifacts," in *Proceedings of Conference on Virtual Reality, Archaeology, and Cultural Heritage*, pp. 333–342, ACM Press, (New York), 2001.
8. J. S. Nimeroff, E. Simoncelli, and J. Dorsey, "Efficient re-rendering of naturally illuminated environments," in *Proceedings of the Fifth European Workshop on Rendering*, pp. 359–373, Eurographics Association, (Aire-la-Ville, Switzerland), 1994.
9. T. Malzbender, D. Gelb, and H. Wolters, "Polynomial texture maps," in *Proceedings of SIGGRAPH 2001, Computer Graphics Proceedings, Annual Conference Series*, E. Fiume, ed., pp. 519–528, Addison-Wesley, (Reading, MA), 2001.
10. P. Sen, B. Chen, G. Garg, S. R. Marschner, M. Horowitz, M. Levoy, and H. P. A. Lensch, "Dual photography," *ACM Trans. Graph.* **24**(3), pp. 745–755, 2005.
11. F. Anrys and P. Dutré, "Image-based lighting design," in *The 4th IASTED International Conference on Visualization, Imaging, and Image Processing*, ACTA Press, (Calgary, Canada), 2004.
12. M. Fuchs, V. Blanz, and H.-P. Seidel, "Bayesian relighting," in *Rendering Techniques*, pp. 157–164, Eurographics Association, (Aire-la-Ville, Switzerland), 2005.
13. N. G. Alldrin and D. J. Kriegman, "A planar light probe," in *CVPR '06: Proceedings of the 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 2324–2330, IEEE Computer Society, (Washington, DC, USA), 2006.
14. J. J. O'Connor and E. F. Robertson, *Paramesvara, MacTutor History of Mathematics archive*, 2000.
15. P. Debevec and J. Malik, "Recovering high dynamic range radiance maps from photographs," in *Proceedings of SIGGRAPH 97, Computer Graphics Proceedings, Annual Conference Series*, T. Whitted, ed., pp. 369–378, Addison Wesley, (Reading, MA), 1997.
16. P. Debevec, "Rendering synthetic objects into real scenes: bridging traditional and image-based graphics with global illumination and high dynamic range photography," in *SIGGRAPH '98: Proceedings of the 25th annual conference on Computer graphics and interactive techniques*, 1998.
17. J. Tumblin, A. Agrawal, and R. Raskar, "Why I want a gradient camera," in *IEEE CVPR*, pp. 103–110, IEEE Computer Society, (Washington, DC), 2005.