DEPARTMENT OF ELECTRICAL AND COMPUTER ENGINEERING, UCLA
ECE 239AS: COMPUTATIONAL IMAGING

**INSTRUCTOR:** Prof. Achuta Kadambi
**TA:** Pradyumna Chari

**NAME:** Swapnil Sayan Saha
**UID:** 605353215

PSET 4 : LIGHT TRANSPORT

| PROBLEM | TOPIC | MAX. POINTS | GRADED POINTS | REMARKS |
|---|---|---|---|---|
| 2.1.1 | Reproducing the Results from Nayar et al. | 1.0 | | |
| 2.1.2 | Separation of Scene Images | 1.0 | | |
| 2.2.1 | Set up a Static Scene | 0.5 | | |
| 2.2.2 | Capturing Images | 0.5 | | |
| 2.2.3 | Synthesizing Novel Images- I | 0.5 | | |
| 2.2.4 | Synthesizing Novel Images- II | 0.5 | | |
| 3.1 | Difference in Images | 2.0 | | |
| 3.2 | Measuring the Light Transport Matrix | 1.0 | | |
| 3.3 | Understanding the Topology of the Light Transport Matrix | 1.0 | | |
| 3.4 | Sampling part I | 0.5 | | |
| 3.5 | Sampling Part II | 0.5 | | |
| 3.6 | Sampling Part III | 1.0 | | |
| | **Total** | 10 | | |

# 1 Motivation

*At the top-level, this PSet helps you understand the complex interactions occurring between light and a scene point.*

The study of *light transport* probes how light travels from a light source to a camera. This class, geared toward macroscopic scenes, deals with a ray-based model of light (i.e. no wave effects). Powerful effects of ray-based light transport can be studied using elementary mathematical methods.

In this PSet, you will gain elementary command of multipath interference. In the first part of the PSet, we will simplify multipath interactions into two categories — **direct illumination** and **global illumination** [1].

Direct Illumination represents a scenario where scene points are directly illuminated by the light source. Global illumination represents the light a scene point receives that is reflected, refracted or scattered off other scene elements. Consequently, a captured image of the scene can be decomposed into two components: direct component and global component. Figure 1, courtesy of *Nayar et al.* [3, 4], shows this separation of a scene image into its direct and global component. The direct and global components convey different information. The direct component gives us a measurement of first bounce reflections in the scene. This is often the most primal form of interaction between a light, material, and the camera. The global component, on the other hand, conveys complex optical interactions between different objects in the scene[3].

(a) Scene Image (b)Direct Component (c) Global Component

Figure 1: **Separation of Direct and Global Components.** A scene image is decomposed into two components. Objects and media in the scene (a) with different optical properties. Direct component of the scene image (b) and global component (c) show different physical phenomenon such as subsurface scattering, volumetric scattering, reflection, etc. Figure from [3].

In the second part of the experiment, we will leverage our understanding of the light transport matrix to synthesize photo-realistic images in novel lighting conditions [5].

## 2 Experimental Component

### 2.1 Separation of global and direct component

We have captured experimental data for you. Download the data from CCLE. The data is organized as follows:

1. Scenes are organized into 3 different folders. Each scene has 28 images.
2. 'white.png' was captured by projecting the white pattern on the scene and 'black.png' was captured by projecting a black pattern.
3. Checkerboard patterns of $8 \times 8$ pixels in size were shifted 5 times (by 3 pixels each time) in each of the two dimensions, to capture a total of 26 images.

The patterns are as shown in Figure 2. We used a Vankyo projector to project the patterns and a FLIR Blackfly camera to capture the images.

(a) White pattern (b) Black Pattern (c) Checkerboard pattern

Figure 2: Projection Patterns used to capture images.

#### 2.1.1 Reproducing the Results from Nayar et al. (1 points)

Please read the Nayar paper [3] and write a program to reproduce the results as in the paper above. In the box below, place the pseudo-code of your program.

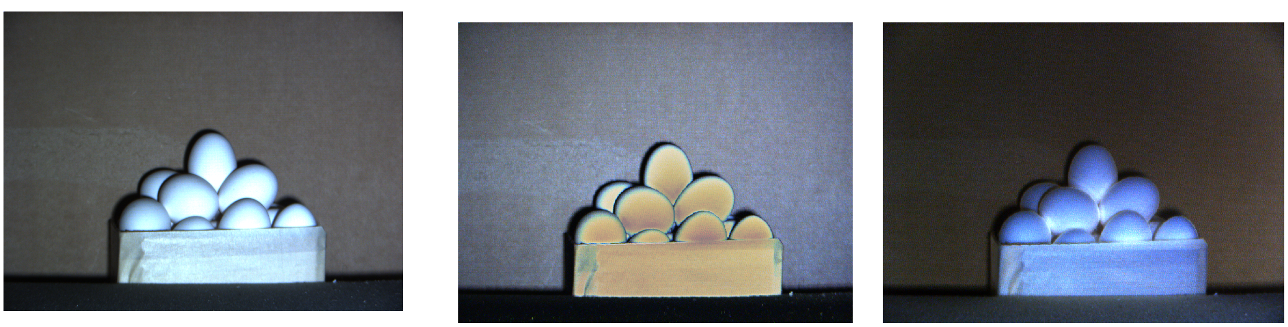
$n$ = number of scene images except white and black images.
$i_k$ = current image in scene images, $L_g$ = global component, $L_d$ = direct component
Assumption: $b = 1$ and $\alpha = 0.5$

$L_{max,bw}$ = uint8(zeros(964,1288))
$L_{min,bw}$ = uint8(zeros(964,1288)*255)
$L_{max} = i_{white}$
$L_{min} = i_{black}$
*Intensity computation, finding max. and min. luminescence picture index and masking*
**for** $k = 1$ to $n$ **do**
$L_{max,bw}$ = max(rgb2gray($i_k$),$L_{max,bw}$)

mask = uint8(zeros(size($L_{max,bw}$)))
mask($L_{max,bw}$ - rgb2gray($i_k$)) $\neq$ 0) = 1
**for** $j$ = 1 to 3 **do**
  $L_{max}(:,:,j) = L_{max}(:,:,j)\times$ mask + $i_k(:,:,j)\times$(1 - mask)
**end for**
$L_{min,bw}$ = max(rgb2gray($i_k$),$L_{min,bw}$)
mask = uint8(zeros(size($L_{min,bw}$)))
mask(rgb2gray($i_k$)-$L_{min,bw}$) $\neq$ 0) = 1
**for** $j$ = 1 to 3 **do**
  $L_{min}(:,:,j) = L_{min}(:,:,j)\times$ mask + $i_k(:,:,j)\times$(1 - mask)
**end for**
**end for**
*Calculate direct and global components:*
$L_g = L_{min} \times 2$
$L_d = L_{max}$ - ($L_g$/2)
**return** $L_g$, $L_d$

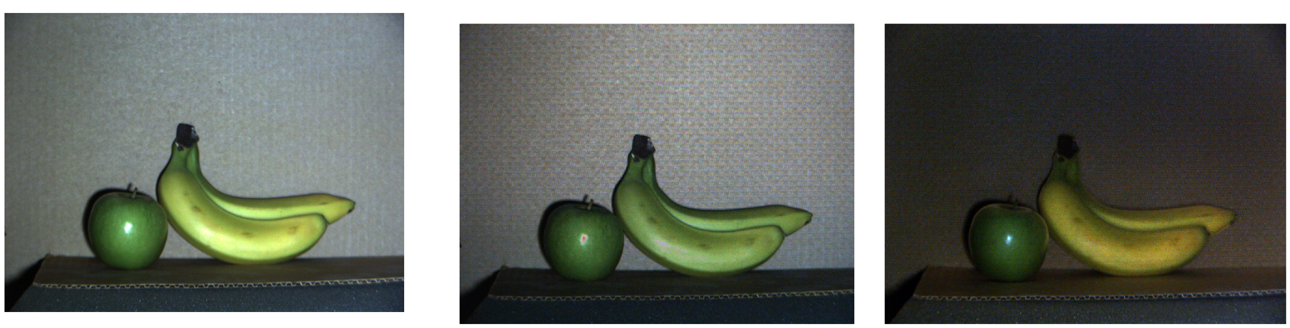
#### 2.1.2 Separation of Scene Images (1 points)

Since scenes must be captured with a projector and camera, we have done this for you. Perform the separation on the data for three scenes. For credit, please place your results for all scenes in the format as shown below. Replace the placeholder images with your results. For each scene, try to explain why the separated images look the way they do.



(a) Scene 1: Eggs

(b) Scene 2: Peppers



(c) Scene 3: Green apple and Banana

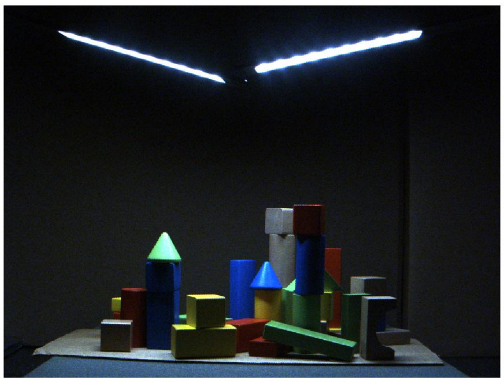
Figure 3: Original image, direct component and global components for 3 scenes.

### 2.2 Image-based Relighting

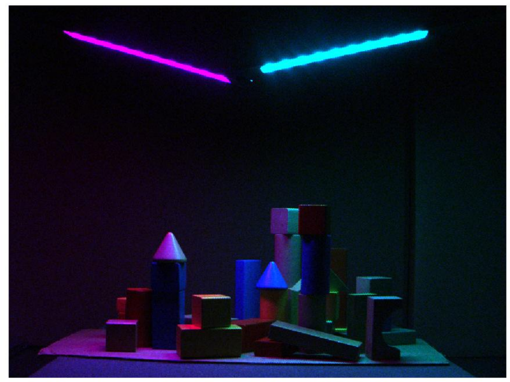
The interaction of light with a scene can be described mathematically by a linear relation known as the light transport equation:

$$\mathbf{p} = \mathbf{T}\mathbf{l}, \tag{1}$$

where the vector $\mathbf{l} \in \mathbb{R}^N$ represents the energy emitted by $N$ controllable light sources, the vector $\mathbf{p} \in \mathbb{R}^M$ represents the radiant energy incident on each of $M$ pixels during the same exposure period, and $\mathbf{T}$ is the scene's $M \times N$ light transport matrix [5]. Thus with the knowledge of the light transport matrix, we can estimate the appearance of a scene that is subject to any illumination condition. We can also photo-realistically synthesize novel images by modifying the lighting condition vector $\mathbf{l}$.



(a) Scene image



(b) Generated Image with novel lighting

Figure 4: **Image based relighting**: For the scene as shown in (a), we modify the illumination vector $\mathbf{l}$ to generate an image (b) with with novel lighting condition.

#### 2.2.1 Set up a static scene (0.5 points)

To do this experiment, you will need two lamps and a couple of scene objects. Set up a static scene similar to the one shown in Figure 4a. For credit, place your image in the box below (replace our block toys scene with your own scene).


Figure 5: Scene image with both lamps on.

#### 2.2.2 Capturing Images (0.5 points)

Let us label the two lamps as LAMP1 and LAMP2. Make sure the position of the camera is stationary throughout the experiment. Capture the image of the scene by turning on LAMP1 only. Repeat the process by turning on LAMP2 only. For credit, place your images in the box below (replace our example).


Figure 6: Scene images with (a) LAMP1 on and (b) with LAMP2 on.

#### 2.2.3 Synthesizing Novel Images - I (0.5 points)

Now, our goal is to obtain a photo with both of the lamps on. We can do this in two ways. We could, of course, turn on both lamps and capture a photograph. However, we can also synthetically generate an image to create the effect of having both lamps on. We do this by summing together the images with individual lamps on. For credit, please insert three images: (a) the scene with both lamps on; (b) synthesized image of having both lamps; and (c) the difference image of (a) and (b). When dealing with difference images, pay careful attention to re-scaling (c) to the full available dynamic range. You can do this by performing the subtraction and then dividing each element of the image matrix with the maximum value in the image. (replace our example).

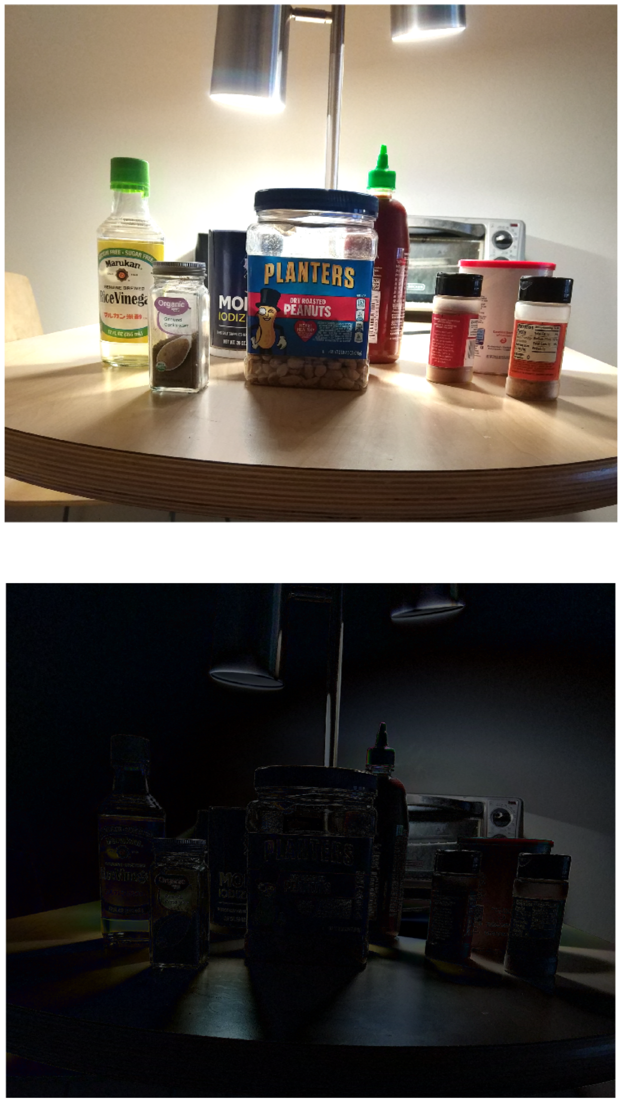



Figure 7: (a) Scene with both lamps on for comparison with synthetically generated image. (b) Synthesized image obtained by addition of images from Figure 6. (c) Difference between captured image and synthesized image.

#### 2.2.4 Synthesizing Novel Images - II (0.5 points)

In the previous box, we validated the equation $\mathbf{p} = \mathbf{T}\mathbf{l}$ by showing the linearity of light transport. Now, we can exploit this linearity in more exotic ways. Create a new image that is in the RGB color-space, where the red channel is a photograph with LAMP1 turned on and the blue channel is a photo taken with LAMP2 turned on. For credit, place your images in the box below (replace our example).


Figure 8: Image with novel lighting generated by weighted addition of red and blue channel of photographs captured with single lamps turned on in Figure 6.

# 3 Assessment Component

## 3.1 Difference in Images (2 points)

You will observe from Section 2.2.3, that the synthesized image is not same as the captured image. In other words, the difference image is not exactly zero. Please precisely explain the origin of this phenomena.

The residual error is due to signal dependent photon shot noise, resulting from the discrete nature of photons. The captured image is a random variable whose variance depends on the brightness, resulting in the error being predominant in bright regions of the scene.

## 3.2 Measuring the Light Transport Matrix (1 point)

For a given scene, the light transport matrix **T** is not known *a priori*, and could take any form (e.g. while it is necessarily of size $m \times n$, it need not be diagonal, etc). Hence, one needs to measure or, as we say in vision/graphics "optically compute" the diagonal matrix. Here, the complexity of optical compute would be measured by the number of images captured. Using pseudo-code, please write an algorithm to optically compute the light transport matrix, assuming $m$ camera pixels and $n$ light sources. In addition, please describe the complexity of optical compute in your algorithm.

Naive method for computing **T**:

**Input:** iterations $n$
**Output:** Light transport matrix **T**
**for** $k = 1$ to $n$ **do**
  illuminate scene with source $k$
  capture photo and store the unrolled photo in $\mathbf{i}_k$
  $\mathbf{T}(1{:}m, k) = \mathbf{i}_k$
**end for**
**return** **T**

The optical complexity is linearly related to the number of light sources $n$, i.e. $O(n)$.

## 3.3 Understanding the Topology of the Light Transport Matrix (1 point)

The structure of the light transport matrix describes the manipulation of light as it leaves a light source and reaches a camera. It is a powerful tool that can be used - as we have seen - to relight images. The structure of the light transport matrix also enables us to characterize the scene. Concretely, describe an experimental scenario where-in the light transport matrix is a diagonal matrix. What does this tell us about the captured photograph? Using the terminology of Nayar's paper, how would you describe diagonal entries vs off-diagonal entries of the light transport matrix?

A scenario where one might obtain a diagonal transport matrix (high rank and sparse) is when imaging a convex, diffuse object using a projector-camera setup. This tells us that the photograph only contains direct illumination light paths from the source, with scattering happening only once and the projector high-frequency pattern being transferred to the photograph as a distorted texture, resulting in one-to-one mapping between a scene point and a camera pixel. The diagonal elements of light transport matrix contain direct component of light paths (first bounce) and the off-diagonal entries contain the global components of light paths (e.g. interreflections, subsurface scattering, volumetric scattering and diffusion).

### 3.4 Sampling Part I (0.5 points)

In Section 3.2 we asked you to optically compute $\mathbf{T}$ without assuming anything about the structure of $\mathbf{T}$. Now, assume that you are told that the matrix $\mathbf{T}$ is diagonal. What is the physical meaning of having a diagonal transport matrix? What is the cost of optically computing $\mathbf{T}$?

A diagonal light transport matrix indicates that an image has only direct component of light (light paths that scatter only once or a single specular path connecting an image point to the camera), resulting in $\mathbf{T}$ being high rank and sparse with one-to-one mapping between a scene point and a camera pixel. The optical complexity in this case is linearly related to the number of camera pixels $m$, i.e. $O(m)$.

### 3.5 Sampling Part II (0.5 points)

Although a diagonal assumption does not hold for general scenes, it is still possible to simplify $\mathbf{T}$. Read about "Helmholtz Reciprocity". You can refer to the 1993 paper by Hapke [2], or simply use Wikipedia. What does Helmholtz reciprocity tell you about the structure of matrix $\mathbf{T}$?

Helmholtz reciprocity states that the intensity of light path does not depend on the direction light travels on the path. In other words, swapping the camera and projector position yields a light transport matrix which is equal to $\mathbf{T}^T$. It also tells us that many rows of $\mathbf{T}$ are linearly dependent.

### 3.6 Sampling Part III (1 point)

We spoke in class about how the light transport matrix can very often be assumed to be low rank. Machine Learning research often studies how a low rank and symmetric matrix can be constructed by sparsely computing rows and columns. Read about the generalized Nystrom method for low rank matrices by Williams and Seeger [6]. Similar to Section 3.2, write how one could optically compute the light transport matrix using the Nystrom Method.

$\mathbf{A} \in \mathbb{R}^{r\times c}, \mathbf{C} \in \mathbb{R}^{m\times c}, \mathbf{R} \in \mathbb{R}^{r\times n}, \mathbf{B} \in \mathbb{R}^{m\times n}, \mathbf{T} \in \mathbb{R}^{(r+m)\times(c+n)}$
**start**
done = 0
*column sampling:*
**for** $(k = 1 : c')$
Sample $k$ randomly.
illuminate scene with source $k$ from projector through diffuser
capture photo with prime camera and store the unrolled photo in $\mathbf{i}_k$
$\mathbf{A}(1{:}r, k) = \mathbf{i}_k$
**end for**
*row sampling:*
**for** $(q = 1 : r')$
Sample $q$ based on $k$ using k-means clustering.
illuminate scene with source $q$ from laser pointers (close to prime camera)
capture photo with dual camera (close to projector) and store the unrolled photo in $\mathbf{i}_q$
$\mathbf{A}(q, 1{:}c) = \mathbf{i}_q$
**end for**
$r_{sample} = r'$; $c_{sample} = c'$
**while** (!done) **do**
*calculate light transport kernel and rank factor:*
estimate pixel distribution density $p(x)$ of $\mathbf{A}$ from histogram entries
estimate $f$, $f(x) = x^\gamma$ from $\mathbf{A}$ such that the objective function $g(\gamma)$ is minimized:
Using golden section search:
$g(\gamma) = \frac{||f\mathbf{A}||_*}{||f\mathbf{A}||_2} \int_0^1 \frac{p_x}{f'(x)} \mathrm{dx}, \qquad ||\mathbf{X}||_* = \sum_i \sigma_i$ and $||\mathbf{X}||_2 = \max_i\{\sigma_i\}$
$e_r = \frac{||f\mathbf{A}||_*}{||f\mathbf{A}||_2}$
*check if rank of $f(\mathbf{A})$ is much less than min$(r,c)$*
**if** $(\frac{e_r}{min(r,c)} < \varepsilon)$ **do**
done = 1
**else do**
*column sampling:*
**for** $(k = 1 : c')$
Sample $k$ randomly.
illuminate scene with source $k$ from projector through diffuser
capture photo with prime camera and store the unrolled photo in $\mathbf{i}_k$
$\mathbf{A}(1{:}r, k) = \mathbf{i}_k$
**end for**
*row sampling:*
**for** $(q = 1 : r')$
Sample $q$ based on $k$ using k-means clustering.
illuminate scene with source $q$ from laser pointers (close to prime camera)
capture photo with dual camera (close to projector) and store the unrolled photo in $\mathbf{i}_q$

$\mathbf{A}(q,1{:}c) = \mathbf{i}_q$
**end for**
$r_{sample} = r_{sample} + r'$; $c_{sample} = c_{sample} + c'$
**end if**
**end while**
*Construct $f(\boldsymbol{T})$ using generalized Nyström method and calculate $\boldsymbol{T}$:*
$f(\mathbf{T}) \approx \mathbf{K} = \begin{pmatrix} f(\mathbf{A}) & f(\mathbf{R}) \\ f(\mathbf{C}) & f(\mathbf{C})(f(\mathbf{A}))^{+} f(\mathbf{R}) \end{pmatrix}$, $\mathbf{C} = \mathbf{PA}$, $\mathbf{R} = \mathbf{AQ}$
$\mathbf{T} \approx f^{-1}(\mathbf{K})$
**return T**
**end start**

Reference: Wang, Jiaping, et al. "Kernel Nyström method for light transport." *ACM SIGGRAPH 2009 papers*. 2009. 1-10.

## References


[1] Mohit Gupta. Scene recovery and rendering techniques under global light transport. 2010.

[2] Bruce Hapke. *Theory of reflectance and emittance spectroscopy*. Cambridge university press, 2012.

[3] Shree K Nayar, Gurunandan Krishnan, Michael D Grossberg, and Ramesh Raskar. Fast separation of direct and global components of a scene using high frequency illumination. *ACM Transactions on Graphics (TOG)*, 25(3):935–944, 2006.

[4] Shree K Nayar, Gurunandan Krishnan, Michael D Grossberg, and Ramesh Raskar. Visual chatter in the real world. In *Robotics Research*, pages 13–24. Springer, 2010.

[5] Matthew O'Toole. *Optical Linear Algebra for Computational Light Transport*. PhD thesis, University of Toronto (Canada), 2016.

[6] Christopher KI Williams and Matthias Seeger. Using the nyström method to speed up kernel machines. In *Advances in neural information processing systems*, pages 682–688, 2001.