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

**INSTRUCTOR:** Prof. Achuta Kadambi
**TA:** Guangyuan Zhao

**NAME:** Qiong Hu
**UID:** 405065032

# QUIZ 3 : ON 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 quiz 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 quiz, you will gain elementary command of multipath interference. In the first part of the quiz, 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.

define $brightness = red + green + blue$ (for any pixel)

for each pixel in the range of image:

$L_{max}$ (at the pixel) = $\arg\max_{r,g,b}\{brightness$ of all images at the pixel$\}$

$L_{min}$ (at the pixel) = $\arg\min_{r,g,b}\{brightness$ of all images at the pixel$\}$

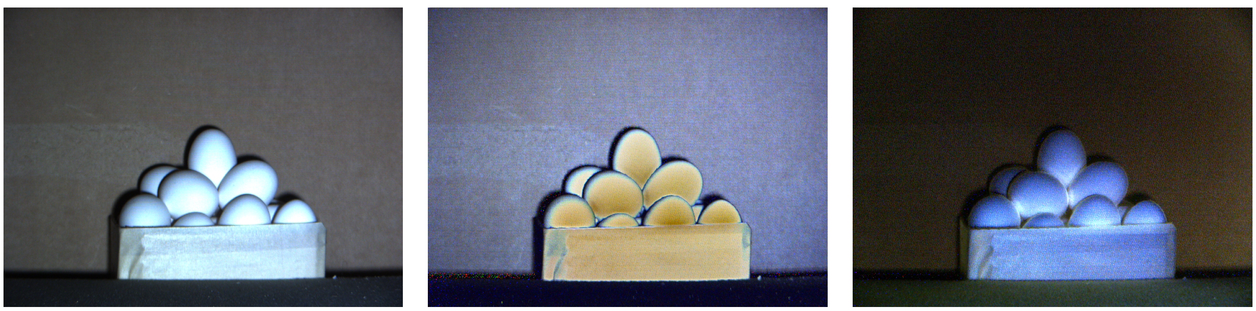
define brightness of a deactivated source element $b \in [0,1]$

$$b = \frac{L_{black}}{L_{white}} \quad (\because L_{white} = L_d + L_g,\ L_{black} = b * (L_d + L_g))$$

$$L_d = \frac{1}{1-b} * (L_{max} - L_{min})$$
$$L_g = \frac{2}{1-b^2} * (-b * L_{max} + L_{min})$$
$$\left(\because L_{max} = L_d + \frac{1+b}{2} * L_g,\ L_{min} = b * L_d + \frac{1+b}{2} * L_g,\ \text{when } \alpha = \frac{1}{2}\right)$$

#### 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.



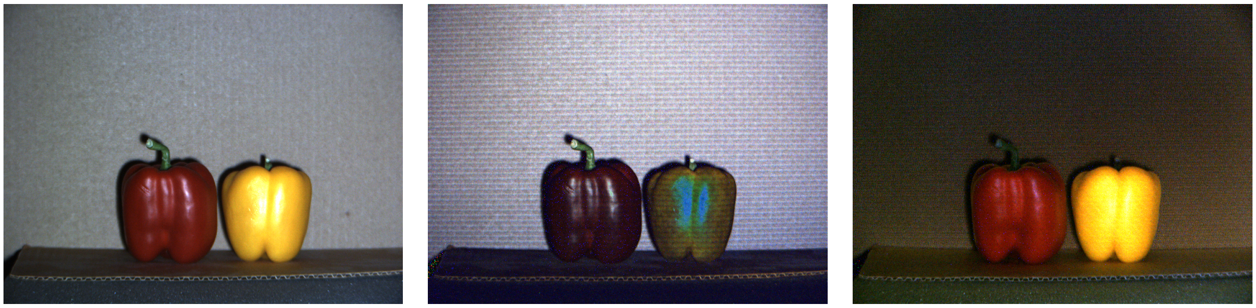
(a) Scene 1: Eggs

The global image includes the strong diffuse interreflections between the eggs[3], and also between the eggs and the paper box.

The reason why the direct image of the eggs and the paper box looks plain yellow in a similar way is that, the surface of the egg shell and the paper are both loose structure, with empty space between cells or fibres. This porous structure would mainly reflect light with a wavelength close to yellow light, and scatter or refract light with other wavelengths.

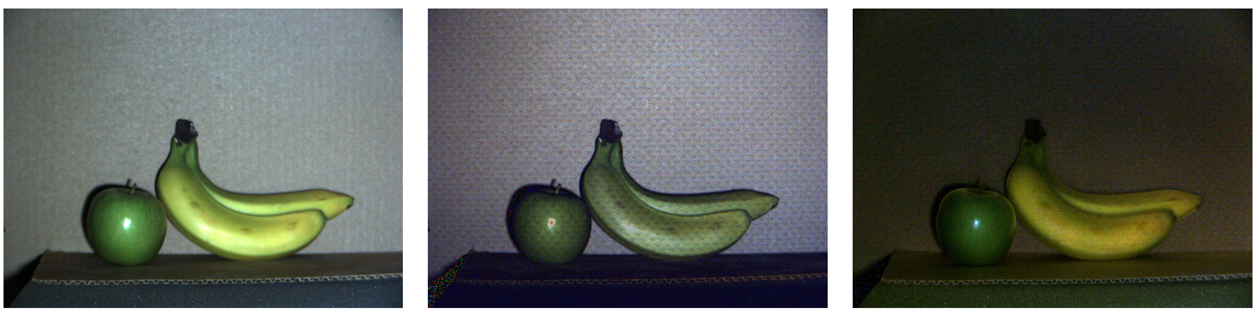
Also in the direct image, the near-edge of the eggs are highlighted and the edges are in shallow; the reason is that the angle between the egg's round surface and the camera is perpendicular at the edge of the eggs, so light is easily scattered at the edge and would hardly enter the camera, so the direct image at the edge is dark.

The wall looks brighter in the direct image and darker in the global image, suggesting that the material of the wall paint is mainly a diffuse reflector.



(b) Scene 2: Peppers

The appearance of the peppers is dominated by subsurface scattering, as seen from the global image. The direct image mainly includes specular reflections, except in the case of the green stalks that are rather more diffuse reflectors[3]



(c) Scene 3: Green apple and Banana

The colors of the green apple and banana in the global image are mostly due to global illustration by subsurface scattering[3], similar to the effects of the peppers.

The upper side of the cardboard and the desktop are only lit up in the global image, meaning that these surface are mostly lit up by the global illustration, such as the diffusion light from the surface of the fruits.

The wall has a coherent behavior in all three scenes, indicating that the analysis of the wall under scene 1 can be generalized to all the scenes in the experiment.

Figure 3: Separation results here. First image in each row describing the scene has been provided. Second image is the direct component. Third image is the global component.

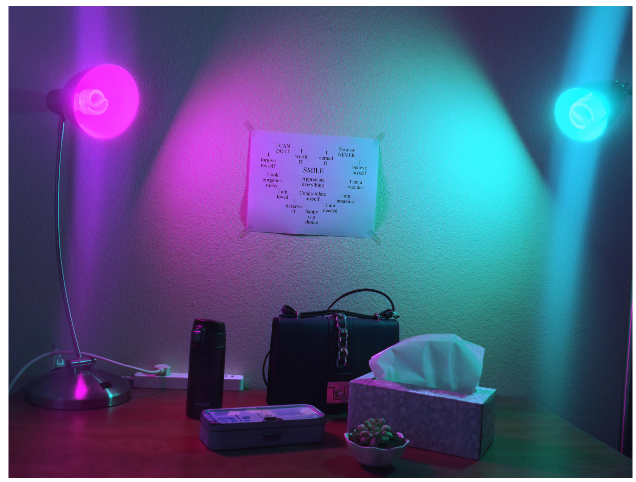
### 2.2 Image-based Relighting

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 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: An ordinary photograph of the scene.

#### 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).

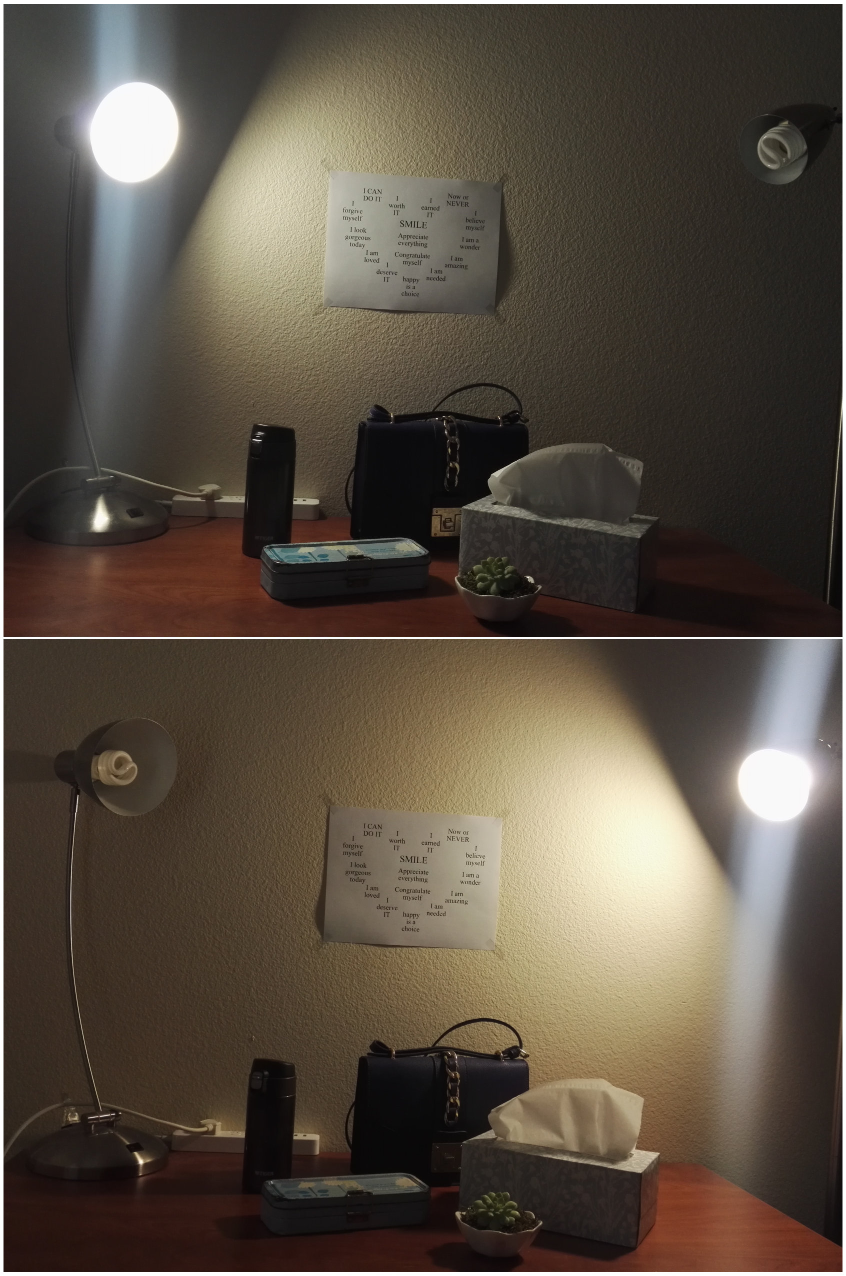
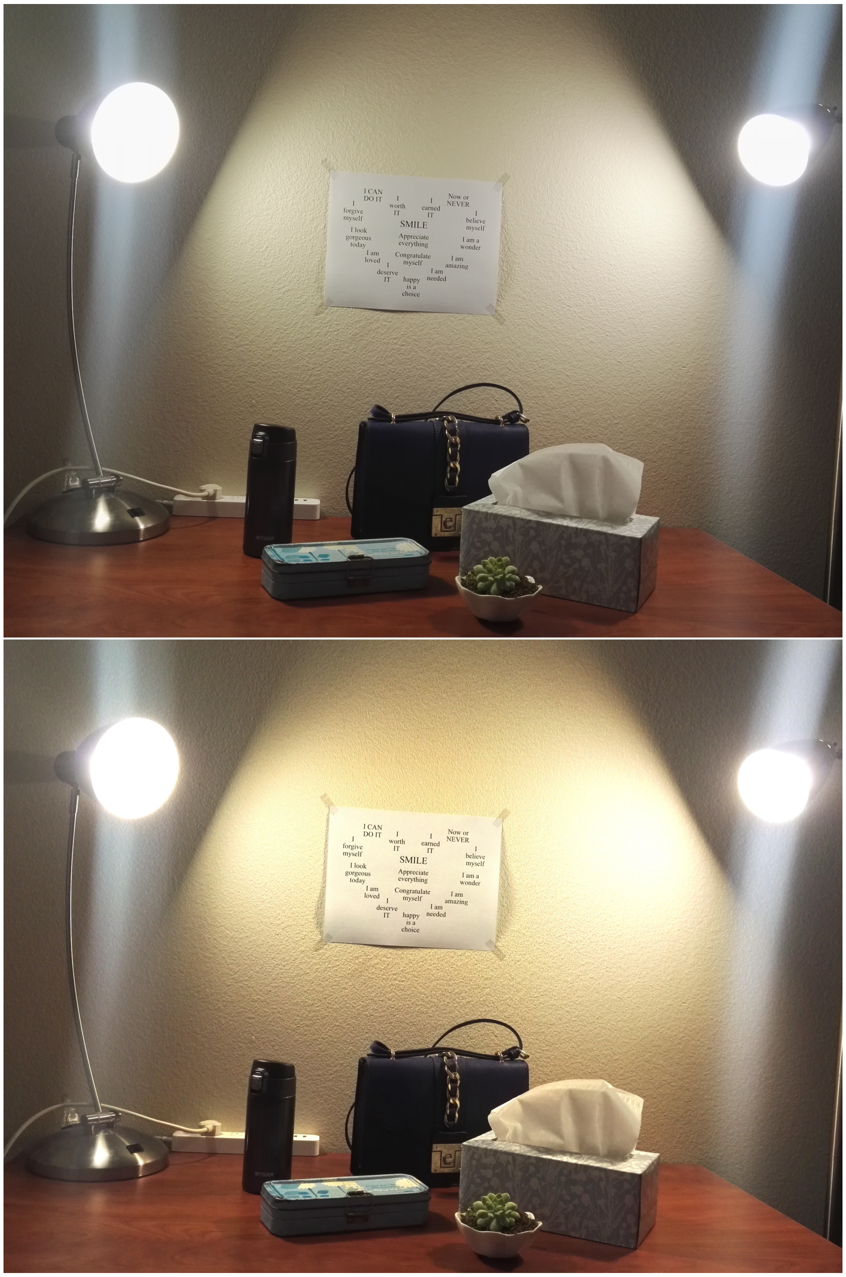


Figure 6: Scene images with (a) LAMP1 on and (b) with LAMP2 on.

#### 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).



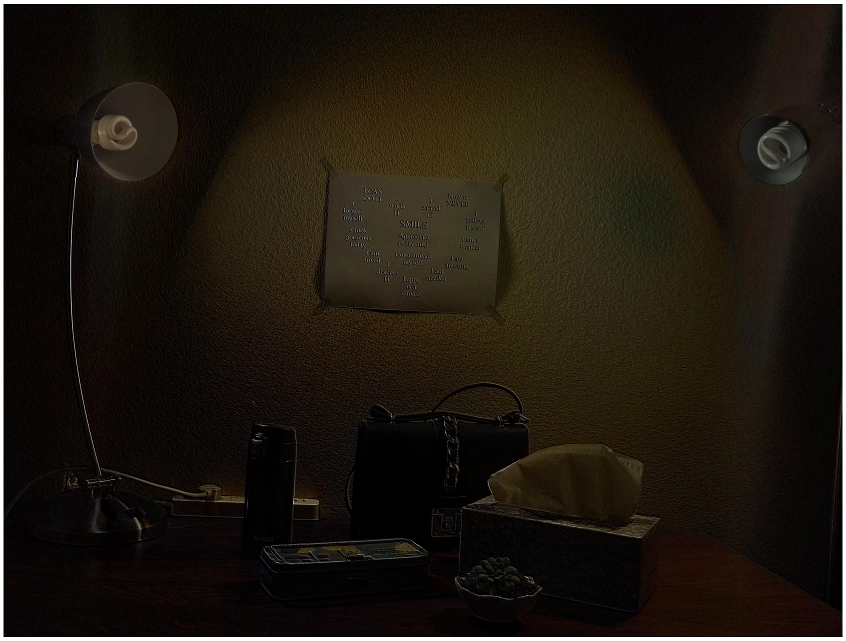



Figure 7: (a) Scene with both lamps on for comparison with synthetically generated image. (b) Synthesized image obtained by addition of images from Figure 6. (c) Difference between captured image and synthesized image.

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

In the previous box, we validated the equation $\mathbf{p} = \mathbf{Tl}$ 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).

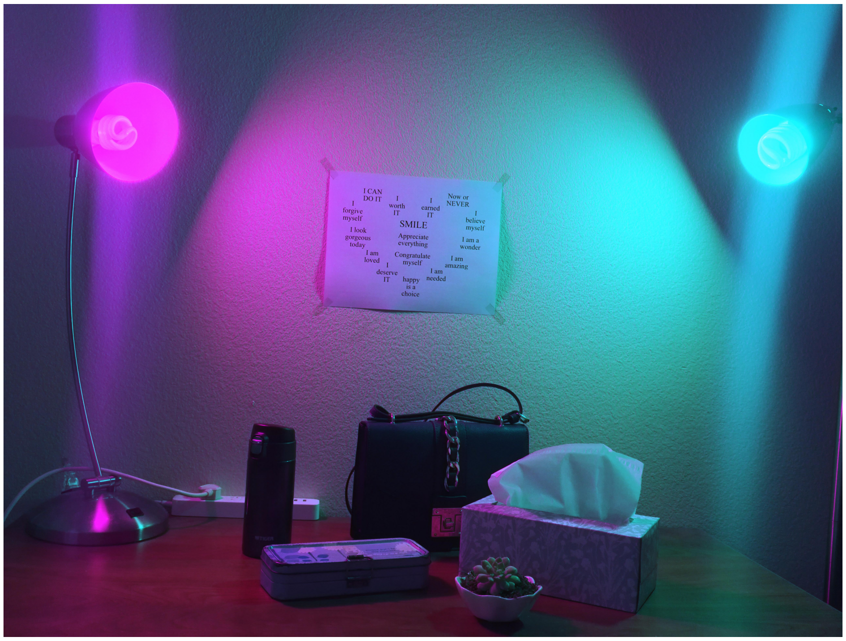



Figure 8: Image with novel lighting generated by weighted addition of red and blue channel of photographs captured with single lamps turned on in Figure 6.

# 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.

From the non-zero area of the difference image, we can notice that the two bulbs and many reflectors are lit up, especially the specular reflection surfaces (such as the metal lamp holders, metal chain of the bag, the surface of the metal pensil case, and the surface of the power strip), and the strong diffuse reflection areas (such as the strongly-lit areas of the wall, and the edge of the letters on the printed paper.

A possible explanation to the difference in reflection surface is that, when both lamps are on, the lighting resources in the scene are not only the lamps, but also the reflection surfaces, so there would be more lighting near those areas, which would not be included in the synthesis image of simple addition.

A possible explanation to the difference in bulbs is that, when the photos are taken, the bulb of the lamp that is turned on would be overexposed, but the RGB value at the pixel are always limited to 255, so there is some lost information of the scene near the bulbs. One thing to notice is that, whether the bulbs show up in the difference image depends on the way to do the subtraction:

- If $img_{diff} = uint8(img_{lamp1} + img_{lamp2} - img_{lamps})$, the color value at the bulb pixel cancels out in own-lamp-on image and both-lamp-on image, and left out in the-other-lamp-on image, so the bulbs would appear in the difference image;

- If $img_{diff} = uint8(img_{lamp1} + img_{lamp2}) - img_{lamps}$, the RGB value at the areas with the bulbs would be limited to 255 in the synthetic addition image, and be cancelled out when subtracting the both-lamp-on image, so the bulbs would not appear in the difference image.

In my case, I used the first method, so there are clear bulbs in the difference image.

According to Matthew O'Toole[5] the residual error is dominated by photon noise, which is the largest in bright regions of the scenes.

## 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.

The pseudo-code to optically compute the light transport matrix $\mathbf{T}$ is as follows:

**Input:** $m$ camera pixels and $n$ light sources

**Output:** light transport matrix $\mathbf{T}$

Initialize a zero $\mathbf{T}$

for $i = 1$ to $n$:

  Illuminate the scene with the $i$-th light source only

  Capture the photo and store in $I^{(i)}$

  Vectorize $I^{(i)}$ into $I_{vec}^{(i)}$

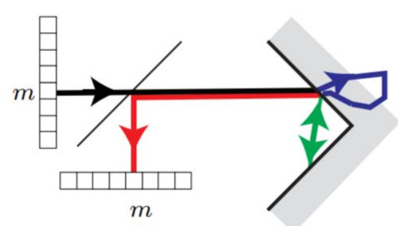
  $\mathbf{T}^{(i)} = I_{vec}^{(i)}$, where $\mathbf{T}^{(i)}$ is the $i$-th column of $\mathbf{T}$

return $\mathbf{T}$

The complexity of optical compute is $O(n)$, because we need to capture $n$ images under $n$ different light sources, in general.

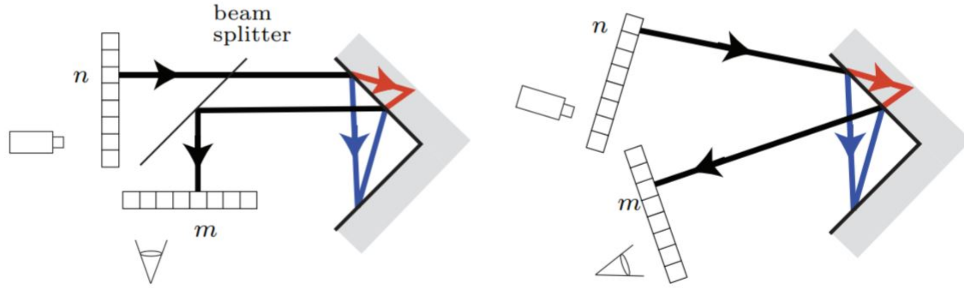
### 3.3 Understanding the Topology of the Light Transport Matrix (1 point)

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?




(c) direct + back-scattering + retro-reflection paths

The figure (c) is cited from Matthrew's paper [5], demonstrating the basic light paths of the diagonal element $\mathbf{T}[m,m]$. $\mathbf{T}[m,m]$ takes into account all light paths that begin at projector pixel $m$ and terminate at sensor pixel $m$. For coaxial arrangements, the fundamental resource of this element is the direct transport path (red). It may also include contributions from back-scattering (blue) and retro-reflection paths (green) begin and end at the same pixel. Therefore, the diagonal entries of a light transport matrix basically describe the direct component of light transport plus retro-reflection and back-scattering.




(a) coaxial arrangement (b) non-coaxial arrangement

The figure (a)-(b) are also cited from Matthre's paper [5], indicating the basic light paths of the off-diagonal element $\mathbf{T}[m,n]$. $\mathbf{T}[m,n]$ takes into account all light paths that begin at projector pixel $m$ and terminate at sensor pixel $n$, where $m \neq n$. For both coaxial arrangement as in figure (a) and non-coaxial arrangement as in figure (b), two such paths are shown: inter-reflections (blue) and sub-surface scattering (red). Therefore, the off-diagonal entries of a light transport matrix basically describe all other indirect components of light transport.

### 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 $\mathbf{T}$?

If the light transport matrix $\mathbf{T}$ is diagonal, it ensures that $m = n$, and the maximum number of non-zero values in $\mathbf{T}$ is $m$. $\mathbf{T}$ can be represented in the form of:

$$\mathbf{T} = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_m \end{bmatrix},$$

where $\lambda_1$, $\lambda_2$, $\cdots$, $\lambda_m$ are eigenvalues of the matrix $\mathbf{T}$. Theoretically, taking only photo with certain known illumination is sufficient to optically compute $\mathbf{T}$, so the cost is one photo, the complexity is O(1).

The physical meaning of such a diagonal light transport matrix $\mathbf{T}$ is that: all the light paths interact with only one diffuse point within the scene. Thus, a single specular path connects the diffuse scene point to the projector at the only projector pixel that illuminates this point, and a single specular light path connects the same diffuse scene point to the camera at the only camera pixel that observes this point.

So all the light rays are a one-to-one correspondence between projector pixels and camera pixels. Since only direct light paths are involved, the photo would only have direct component of light transport, for example, the object is a perfect mirror. As a result, displaying a projector pattern directly onto the object will transfer it to the photo in the form of a distorted texture, as it transfers a projector's high-frequency content to the photo [5].

### 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 transfer of radiant energy along a light path does not depend on the direction of propagation, so the value of the scattering throughput function $f(x)$ for a light path $x$ also does not depend on the direction of light's propagation[3]. Therefore, swapping the position of a projector pixel $n$ and camera pixel $m$ produces the same value $\mathbf{T}[m,n]$ in the light transport matrix $\mathbf{T}$:

$$\mathbf{T}[m,n] = \mathbf{T}[n,m]$$

satisfies for any $m$ and $n$. Therefore,

$$\mathbf{T}^{\mathrm{T}} = \mathbf{T},$$

the matrix $\mathbf{T}$ is a symmetric matrix (if the resolution of the camera and the projector matches).

If the resolution of the camera and the projector does not match, then $\mathbf{T}$ is asymmetric, so we can enforce the symmetry by replacing $\mathbf{T}$ with $\mathbf{T}^*$ in the algorithm:

$$\mathbf{T}^* = \mathbf{T}^{\mathrm{T}}\,\mathbf{T}.$$

### 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 [7]. Similar to Section 3.2, write how one could optically compute the light transport matrix using the Nyström Method.

The Kernel Nyström Method reconstructs a low-rank symmetric matrix from sparsely sampled columns, and the Extended Kernel Nyström Method can be applied to asymmetric matrices. The light transport matrix $\mathbf{T}$ can be written as:

$$\mathbf{T}_{(m\times n)} = \begin{bmatrix} \mathbf{A}_{(r\times c)} & \mathbf{R}_{(r\times(n-c))} \\ \mathbf{C}_{((m-r)\times c)} & \mathbf{B}_{((m-r)\times(n-c))} = ? \end{bmatrix},$$

where $\mathbf{A}$, $\mathbf{R}$, $\mathbf{C}$ are known matrices used for estimating $\mathbf{B}$ and thus to compute $\mathbf{T}$. The dimensions of each matrix has been denoted along with the matrix.

The Kernel Nyström Method aims to find a non-linear light transport kernel $f$, such that:

$$\mathbf{K} = f(\mathbf{T}) = \begin{bmatrix} f(\mathbf{A}) & f(\mathbf{R}) \\ f(\mathbf{C}) & f(\mathbf{C})(f(\mathbf{A}))^{+}f(\mathbf{R}) \end{bmatrix},$$

where $\mathbf{A}^{+}$ denotes the Moore-Penrose pseudoinverse of $\mathbf{A}$, which has the property $\mathbf{A}\mathbf{A}^{+}\mathbf{A} = \mathbf{A}$.

The optimized kernel function $f$ would be the one that minimize the rank factor $e_r$ and the inverse function slope $e_s$, the definition and equation of $e_r$ and $e_s$ can be found in Wang et al.'s paper [6].

Therefore, the pseudo-code to measure $\mathbf{T}$ is as follows:

Randomly capture $r$ rows and $c$ columns as initialization

while $e_r > \varepsilon min(r,c)$:

  Estimate $f$ from $\mathbf{A}$

  Estimate rank factor $e_r$ of $f(\mathbf{A})$

  Eandomly capture $r'$ rows and $c'$ columns

  Add new samples to the sample set

  Update $\mathbf{A}$

  $r = r + r'; c = c + c'$

return $f$

$\mathbf{T} = f^{-1}(\mathbf{K})$

## 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] Jiaping Wang, Yue Dong, Xin Tong, Zhouchen Lin, and Baining Guo. Kernel nyström method for light transport. *ACM Transactions on Graphics (TOG)*, 28(3):29, 2009.

[7] 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.