Supplemental Material: Adaptive Lighting for Data-Driven Non-Line-of-Sight 3D Localization and Object Identification

Sreenithy Chandran
schand56@asu.edu
Arizona State University
Tempe, USA
Suren Jayasuriya
sjayasur@asu.edu

1 Overview

In this supplemental material, we present additional material concerning the derivation of our adaptive lighting algorithm based on radiosity, implementation details for our experiments, and a small analysis of visual saliency in our network performance.

2 Adaptive Lighting

In the main paper, we presented our adaptive lighting algorithm that optimized over B_{NLOS}. In this section, we present the complete analytic derivations for radiosity for three bounce light used in that algorithm. We mainly follow the approach of Klein et al. [3] in calculating our radiosities.

Let S_1,S_2....S_N be the N patches on the reflective LOS wall, light source denoted as p, camera denoted as C and the NLOS patch denoted as $NLOS$. To calculate the radiosity along a ray for three bounce light, we must first calculate its first and second bounces.

First Bounce (LOS):

$$p \Rightarrow S_i \Rightarrow C \quad \forall i \in \{1,N\}.$$

When light travels from the source to a diffuse wall and bounces back to the camera, the associated radiosity is given as the product of the reflectance of the surface ρ_i, the radiosity of the incident light B_p, and the form factor F_{ip} between the p and the ith patch, and the visibility term V_i [1, 2]:

$$B_{first} = B_i = \rho_i B_p F_{ip} V_i, \quad \forall i \in \{1,N\}. \tag{1}$$

The form factor calculates how much light is transferred from one patch to another. Since the wall is divided into N patches, the first bounce radiosity associated with all the N patches is calculated. It takes into account the distance between the surfaces, computed as the distance between the center of each of the surfaces, and their orientation in space relative to each

other, computed as the angle between each surface's normal vector and a vector drawn from the center of one surface to the center of the other surface:

$$F_{ij} = \frac{\cos\theta_j \cos\theta_i}{\pi(x_j - x_i)^2} \cdot A_i \tag{2}$$

A visual depiction is shown in Figure 1. However the above equation does not account for occlusion between the two patches. This is accounted by the visibilty term $V(i,j)$:

$$V_i(\vec{x_a}, \vec{x_b}, \vec{N_a}) = \begin{cases} 0, & \text{if } k >= \frac{\pi}{2} \text{ and } k <= \frac{3\pi}{2} \\ 1, & \text{otherwise} \end{cases}$$

where $k = (\vec{x_a} - \vec{x_b}) \cdot \vec{N_a}$.

Substituting form factor and visibility terms into (1), we get the following expression for first bounce radiosity:

$$B_{first} = B_i = \rho_i \cdot B_p \cdot \left(\frac{\cos\theta_p \cos\theta_i}{\pi(x_p - x_i)^2}\right) \cdot A_i \cdot V_i, \quad \forall i \in \{1,N\}. \tag{3}$$

Second Bounce (LOS):

$$p \Rightarrow S_i \Rightarrow S_j \Rightarrow C$$

for $\forall i \in \{1,N\}, \forall j \in \{1,N\}, j \neq i$. In this case, light from the illuminating source hits the diffuse wall and the gets reflected to another patch on the diffuse wall. This can be viewed as the light taking two bounces and containing only LOS scene information when it reaches the camera. Using the radiosity calculated from Equation(3) as the radiosity illuminating a second bounce patch, we get the expression for B_i:

$$B_j = \rho_j B_i F_{ji} V_j \tag{4}$$

$$\forall i \in \{1,N\}, \forall j \in \{1,N\}, j \neq i.$$

Third Bounce: For third bounces, we now have two subcases: when the ray only interacts with LOS patches, and when the ray interacts with the NLOS patch. We treat each case separately in our derivations.

LOS Condition

$$p \Rightarrow S_i \Rightarrow S_j \Rightarrow S_k \Rightarrow C,$$

$$\forall i \in \{1,N\}, \forall j \in \{1,N\}, j \neq i, \forall k \in \{1,N\}, k \neq j.$$

Consider the scenario where the light after bouncing off two diffuse wall patches strikes another diffuse wall patch. This can be viewed as the light taking three bounces and containing only LOS scene information when it reaches the camera. Using the radiosity calculated from Equation (4) as the radiosity illuminating a third bounce patch, we get the expression for B_k:

$$B_k = \rho_k B_j F_{kj} V_k \tag{5}$$

NLOS Condition

$$p \Rightarrow S_i \Rightarrow NLOS \Rightarrow S_k \Rightarrow C$$

In the scenario where, after light after undergoing first bounce LOS reaches the NLOS scene and then bounces to a diffuse wall patch beore reaching the camera, we can view that as three bounce NLOS light. After the first bounce, the incident radiosity is given by the condition Equation (3).

$$B_n = \rho_n B_i F_{ni} V_n \tag{6}$$
$$\forall i \in \{1,N\}, \forall n \in \{1,N\}$$

Using Equation (6) as the radiosity of the light reaching the diffuse wall patch, we obtain the third bounce as below:

$$B_k = \rho_k B_n F_{kn} V_k, \tag{7}$$
$$\forall k \in \{1,N\}$$

Using Equation (3), Equation (4), Equation (5), and Equation (7), the total radiosity can be calculated as the contribution due to the NLOS radiosity and LOS radiosity as below,

$$B_{total} = B_{NLOS} + B_{LOS}$$

Using these radiosity contributions, we can then solve the optimization problems formulated in the main paper, Section 4. The full steps are summarized in Algorithm 1.

Algorithm 1: Adaptive Lighting to calculate B_{NLOS}

Step 1: Divide the LOS scene into N patches, calculate the surface normal and area per patch.
Step 2: Calculate light source to LOS patch light transfer.
for *LOS patch i=1:N* **do**
  Calculate the first bounce radiosity: $B_i = \rho_i B_a F_{ia}$ where B_a is the radiosity of the illumination source.
end
Step 3: LOS patch to the NLOS object light transfer.
for *LOS patch i=1:N* **do**
  Calculate the second NLOS bounce radiosity using Equation (4) and using Step 2 as the radiosity emitted by each LOS patch
end
Step 4: Third bounce light from LOS to camera.
for *LOS patch i=1:N* **do**
  Calculate the final radiosity using Equation (7)
end
Step 5: Solve corresponding optimization problem (Equation 2 or Equation 3 in the main paper) using radiosities from Step 4.

3 Energy-efficiency of Adaptive Lighting

When a spatially-varying light pattern is used instead of a spotlight source, the illumination power is spread over the entire scene. This is counter to our stated goal of optimizing the energy-efficiency for the lighting, particularly the distributed optimization algorithm in the paper that operates under a power budget.

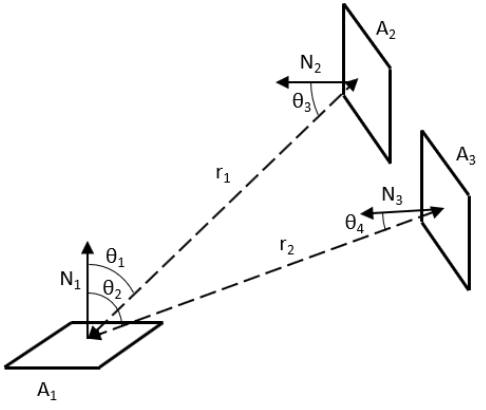

Figure 1: Radiosity measures the radiative transfer of light between diffuse surfaces and emitters based on their reflectance, viewing/occlusion, and geometric form-factors [1]. We formulate an optimization to identify the patches in the LOS which maximize the NLOS radiosity captured by the camera.

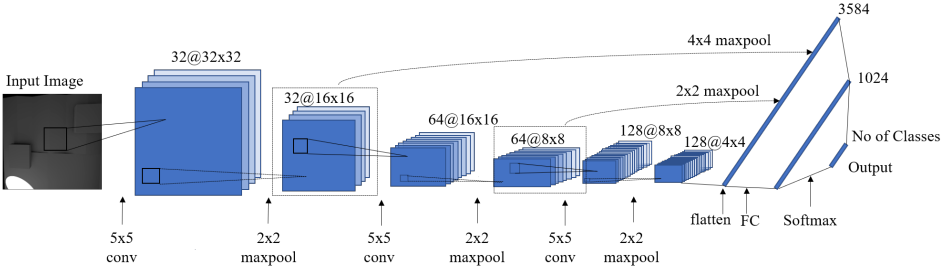

Figure 2: Network Architecture for Identification

To illustrate the effects of loss of power that occurs when you spread the light in a spatial pattern, we conducted the following experiment. Consider a room with the reflective wall divided into 100 patches. We illuminate the scene for a finite number of patches, and compare our adaptive lighting algorithm versus floodlighting the scene. For floodlighting, the incident illumination power is divided by the number of patches considered, and we measure the radiosity returned from NLOS. For our adaptive algorithm, we focus the incident illumination power onto a particular set of patches given by the optimization. In Figure 3, we see that our adaptive algorithm (green) returns higher NLOS radiosity than spreading the light out in a floodlit pattern (blue). We believe this experiment shows the value of not using spatially-varying lighting patterns for the same energy budget.

However, there is an interesting avenue for future research. One advantage is that spatially-varying lighting could improve detection coverage over the NLOS, as opposed to our adaptive lighting method which requires N adaptive lighting patterns at test time to determine where the object is located. We can imagine NLOS imaging schemes which utilize spatially-varying lighting for coarse localization and detection, and then adaptive lighting for finer localization.

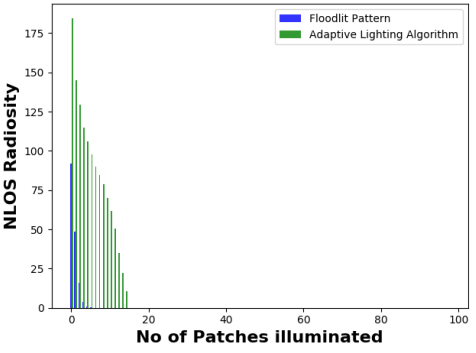

Figure 3: Plotting the NLOS radiosity returned by our adaptive lighting algorithm (green) versus floodlighting the scene (blue). Choosing optimal patches to illuminate returns more NLOS radiosity than spreading the illumination power evenly (as in a lighting pattern or floodlit) for the same illumination power.

Object	Localization	
	Adaptive	Non-adaptive
BUNNY 1	2.41 cm	3.79 cm
BUNNY 2	1.32 cm	1.65 cm
BALL 1	2.89 cm	4.67 cm
BALL 2	1.61 cm	2.76 cm

Table 1: Additional real data results for four objects trained on the complex LOS wall 2.

4 Implementation Details

For the simulated data, we implement our CNNs using PyTorch version 0.4.1 on a single NVIDIA GeForce GTX 1080 Ti GPU. Our datasets are of size 100,000 images for each specific wall and 64×64 resolution. We train using stochastic gradient descent with momentum 0.9 and learning rate $\lambda = 0.0001$ for 20 epochs until convergence for classification and 16 epochs until convergence for localization, with a 70 : 30 training/testing split.

For our hardware prototype, we built a room setup, constructed using wood, of dimension 35.6 cm $\times$ 35.6 cm $\times$ 35.6 cm. We 3D printed the walls and then spray-painted them to be diffuse white. The real scene we use for the LOS is a variation of Wall 2. The Stanford bunny, sphere and man silhouette of varying sizes were 3D printed and spray painted diffuse white to help improve signal return back to the LOS. The wall was illuminated with an InFocus IN3138HD projector. We used an aperture after the projector of black construction paper with a small hole to focus the spot and emulate a spot light source. A Logitech C615 HD WebCam captured images of the diffuse wall. We capture roughly 10,000 real images to use for our datasets.

5 Additional Real Data Localization Results

For our real data, we also performed an ablative study for localization with other sizes of spheres and bunnies. A bunny (BUNNY 1) with 5.5 cm width and 3.7 cm height was localized with MSE 2.41/3.79 cm respectively for adaptive/non-adaptive method, while a larger bunny (BUNNY 2) with 9 cm width and 7 cm height localized to 1.32/1.65 cm respectively for adaptive/non-adaptive method. A sphere (BALL 1) of diameter 3 cm localized to

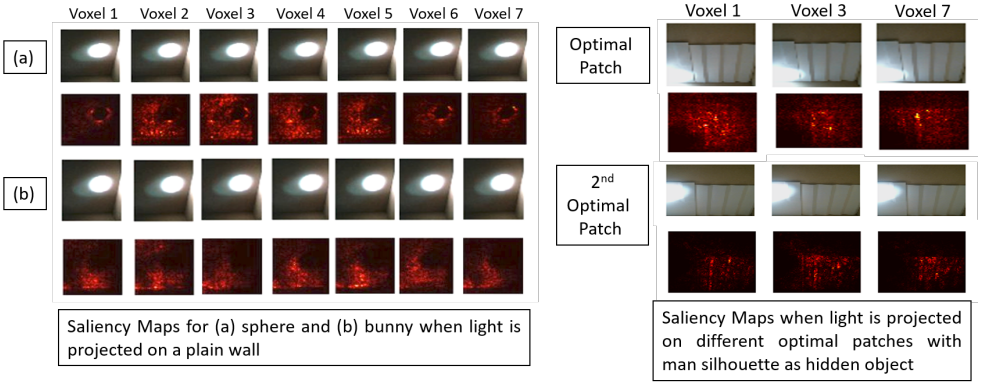

Figure 4: Salient image regions used by Inception network is calculated using the method from [4].

2.89/4.67 cm respectively for adaptive/non-adaptive, and a larger sphere (BALL 2) of diameter 8 cm localized to 1.61/2.76 cm respectively for adaptive/non-adaptive. Note how as the size of objects gets bigger, the localization becomes more accurate in general due to more signal being reflected back from the NLOS.

6 Saliency

To investigate what parts of the image our network is finding the most salient information, we utilize class-saliency maps from Simonyan et al. [4]. In Figure 4(a), we show the input images and saliency maps for the sphere and bunny projected on a planar wall. Note how the saliency of the sphere and bunny look qualitatively different, which probably explains why the network has poor generalisation performance across objects it has never seen in training before. In Figure 4(b), we show how the optimal patch returned by our adaptive lighting algorithm has more saliency for the network compared to the second best patch. This correlates with the improvement benefits we see with adaptive lighting.

References

[1] Michael F Cohen and John R Wallace. *Radiosity and realistic image synthesis*. Elsevier, 2012.

[2] Alexander Keller. Instant radiosity. In *Proceedings of the 24th Annual Conference on Computer Graphics and Interactive Techniques*, pages 49–56. ACM Press/Addison-Wesley Publishing Co., 1997.

[3] Jonathan Klein, Christoph Peters, Jaime Martín, Martin Laurenzis, and Matthias B Hullin. Tracking objects outside the line of sight using 2d intensity images. *Scientific Reports*, 6:32491, 2016.

[4] Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*, 2014.