ENGN2502: 3D Photography
Assignment 2b: Robust Pixel Classification

Instructor: Gabriel Taubin*
Assignment written by: Fatih Calakli†

24 March 2012

Introduction

This is a short and natural extension to assignment 2. The goal of this extension is to improve both the quantity of decoded pixels and the quality of the final reconstruction. Specifically, you'll implement the pixel classification method introduced by Xu and Aliaga [2]. This method is easy to implement, and does not require newly designed patterns. Therefore, you will use the same data provided in assignment 2. The algorithm rely on separating the direct and global components of a pixel intensity introduced by Nayar et. al. [1].

1 Direct and Indirect Components of a Scene

The intensity of a pixel can be decomposed into the direct component and indirect (or global) component. The direct component is due to light bouncing off the surface in a single reflection. The indirect component is due to multiple reflections (e.g. inter-reflections, subsurface, scattering etc.).

For a directly illuminated pixel p, its intensity $L(p)$ can be decomposed into two components: direct component $L_d(p)$ and $L_g(p)$. For an indirectly illuminated pixel, its intensity $L(p)$ only contains the indirect component $L_g(p)$. In summary,

$$L(p) = \begin{cases} L_d(p) + \alpha L_g(p) & \text{if p is on} \\ (1-\alpha) L_g(p) & \text{if p is off,} \end{cases} \tag{1}$$

where $0 \leq \alpha \leq 1$ is a fraction of activated source pixels. For example, for a random binary pattern $\alpha = \frac{1}{2}$, and for a white pattern $\alpha = 1$.

2 Separation of Direct and Global Components

You will compute direct component $L_d(p)$ and global component $L_g(p)$ for each pixel using the method introduced by Nayar et. al [1]. You will use two captured images of the scene, where, in the first image the scene is lit with high-frequency illumination and in the second image it is lit with the complementary illumination. Suppose that $L^+(p)$ is the intensity at a pixel p when the scene is lit with the first illumination, and $L^-(p)$ when it is with the second illumination. If $\alpha \geq \frac{1}{2}$, we get

$$L_d(p) = L_{max}(p) - \frac{\alpha}{1-\alpha} L_{min}(p) \tag{2}$$

$$L_g(p) = \frac{1}{1-\alpha} L_{min}(p), \tag{3}$$

*taubin@brown.edu
†fatih_calakli@brown.edu

where $L_{max} = max\{L^+(p), L^-(p)\}$, and $L_{min} = min\{L^+(p), L^-(p)\}$. The value of α depends on the choice of illumination pattern, *i.e.* fraction of white pixels. In this assignment, consider a code pattern and its inverse as the high frequency illumination patterns. Make sure that the bit pattern satisfies $\alpha \geq \frac{1}{2}$. In fact, if α is either close to 1 or 0, the scene will be lit (sampled) very sparsely in one of the two images. A good choice is $\alpha = \frac{1}{2}$ as also recommended in [1].

3 Robust Pixel Classification

Projecting the code pattern i and its inverse yields two values for each pixel. Using both pixel values, $L_i^+(p)$ and $L_i^-(p)$, you will classify p according to the rules presented in [2]:

$$p \text{ is } \begin{cases} \text{uncertain} & \text{if } L_d(p) < m \\ \text{on} & \text{if } L_d(p) > L_g(p) \text{ and } L_i^+(p) > L_i^-(p) \\ \text{off} & \text{if } L_d(p) > L_g(p) \text{ and } L_i^+(p) < L_i^-(p) \\ \text{off} & \text{if } L_i^+(p) < L_d(p) \text{ and } L_i^-(p) > L_g(p) \\ \text{on} & \text{if } L_i^+(p) > L_g(p) \text{ and } L_i^-(p) < L_d(p) \\ \text{uncertain} & \text{otherwise,} \end{cases} \tag{4}$$

where m is a predefined small threshold, say 5 out of 255.

Thus far, we have assumed that when a source pixel is not activated it does not generate any light. In the case of a projector, this is seldom completely true. If we assume the brightness of a deactivated source pixel is a fraction $0 \leq \beta \leq 1$ of an activated pixel, then the expression (1) can be modified as:

$$L(p) = \begin{cases} L_d(p) + (\alpha + \beta - \alpha\beta)L_g(p) & \text{if p is on} \\ \beta L_d(p) + (1 - \alpha + \alpha\beta)L_g(p) & \text{if p is off.} \end{cases} \tag{5}$$

We can compute the direct and global components at each camera pixel the same way:

$$L_d(p) = \frac{(1 - \alpha + \alpha\beta)L_{max}(p) - (\alpha + \beta - \alpha\beta)L_{min}(p)}{(1 - \beta^2)(1 - \alpha)} \tag{6}$$

$$L_g(p) = \frac{-\beta L_{max}(p) + L_{min}(p)}{(1 - \beta^2)(1 - \alpha)}. \tag{7}$$

Note that β should be a small number. To find a good value, experiment with various values of β such as 0.1, 0.2, 0.3, 0.4, and 0.5.

What to turn in: Compute per-pixel direct and global components for a Gray code pattern. Report your choice of the code pattern, and β. Explain your reasoning. Implement a function to decode the Gray code structured light sequences using robust pixel classification rules. Turn in a plot of the decoded correspondences for the focused and defocused cases, for the `man` sequences. Compare these plots and the reconstructions against the ones you have obtained using the standard rules.

References

[1] 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 Trans. Graph.*, 25(3):935–944, July 2006.

[2] Yi Xu and Daniel G. Aliaga. Robust pixel classification for 3d modeling with structured light. In *Proceedings of Graphics Interface 2007*, GI '07, pages 233–240, New York, NY, USA, 2007. ACM.