Depth Invariant Reflectance Estimation

Henryk Blasinski
Department of Electrical Engineering
Stanford University,
Stanford, CA, USA.
hblasins@stanford.edu

Abstract

Spectral reflectance is often derived from a sequence of images acquired under different, narrowband, infinitely far away illuminants, which means their intensity is constant throughout the scene. When small LEDs, located close to the target are used the light intensity becomes scene depth dependent. We propose to use an RGB+depth camera which lets us model the spatial light intensity falloff and consequently reliably recover scene reflectance irrespectively of the object's distance to the light source. The algorithm including the effects of illuminant non-uniformities allows us to achieve RMS error performance comparable to systems using distant light sources.

1. Introduction

Spectral reflectance describes the ratio between the incident and reflected light intensity at different wavelengths. The reflectance spectrum is an intrinsic property of the material and therefore it is much more informative that just its RGB image. This additional information can be used for image analysis and classification purposes and can be helpful in a variety of computer vision tasks in fields such as biology, medicine or industrial quality control [1, 3].

One popular method to estimate spectral reflectance is to acquire images of a scene with an RGB camera and with a small number of narrowband lights. The reflectance can be derived by solving an inverse estimation problem, provided the spectral characteristics of the camera and the light sources are known. So far all implementations of this method operated under the assumption that the narrowband lights are infinitely far away, and consequently the light intensity is constant and invariant across the imaged scene [4, 5]. This condition is clearly violated when point sources, such as LEDs, located close to the target scene are used instead. At a particular spatial location the light intensity emitted form a point source depends on its distance from this light source. This means that the light intensity distribution depends on the scene geometry and cannot be calibrated for in advance. There are many applications where applications where point light sources are used, and hence light falloff models are indispensable for correct reflectance estimation. One example is the medical field where endoscopes, laparoscopes or microscopes use their own LED source [2].

In this paper we address the problem of non-uniform illuminant energy distribution in the observed scene. Our solution adds a depth sensing camera the original active illumination reflectance estimation setup. This additional component will allow to measure the distance between the imaged object and the light source and model the spatial light intensity falloff thus providing a reliable light intensity estimate for every image pixel. To the best of our knowledge this is the first hardware system and estimation algorithm that allows to reliably recover scene reflectance under illumination from point sources.

2. Reflectance estimation

The linear image formation model [10] is used in all active illumination reflectance estimation systems. This model allows to express the measured pixel intensity m_{ij} captured by a camera as a linear function of the illuminant intensity l_j and the scene reflectance r

$$m_{ij} = g_{ij} \sum_{\lambda=1}^{\Lambda} q_\lambda c_{i,\lambda} r_\lambda l_{j,\lambda}, \tag{1}$$

where q_e is the sensor quantum efficiency and c_i is the transmissivity of the ith color filter. The scalar g_{ij} represents the effects of cameras ISO gain, exposure duration, aperture setting and spectral sampling step. Furthermore, real world reflectances are smooth and slowly varying functions, therefore they are often approximated with a small number

of linear basis functions [6]

$$r = \sum_{k=1}^{K} b_k w_{r,k} = B_r w_r, \tag{2}$$

where b_k is the kth reflectance basis function and $w_{r,k}$ is the corresponding weight. When a particular surface is imaged through I filters and under J illuminants, the surface reflectance may be estimated by solving an inverse problem [4]

$$\begin{aligned} &\text{minimize} \;\; \|M - C^T \mathbf{diag}(B_r w_r) L\|_F^2 + \delta \|R B_r w_r\|_2^2 \\ &\text{subject to} \;\; 0 \leq B_r w_r \leq 1, \end{aligned} \tag{3}$$

where R is a bi-diagonal matrix computing the discrete, first order difference. The first term of the objective function represents how well the model fits the data. It is supplemented by a penalty, adjusted with a scalar δ, on solutions with large variations. Finally, the solution space is restricted to nonnegative vectors with entries no larger than one. This constraint follows from the fact that light is a non-negative quantity and that passive surfaces cannot generate new photons.

The image formation model (5) is accurate only when the light source is infinitely far away, in which case the light intensity is constant and independent of the position in the scene. This assumption is violated when point sources close to the imaged scene are used. The light intensity I_d at some distance d from the point source is governed by the inverse-square law and can be expressed as [7]

$$I_d = \frac{I_0}{d^2}. \tag{4}$$

The effects of this phenomenon are illustrated in Figure 1 which shows an image of a white test target placed at four different distances away from the camera and the light source. The overlay shows green channel intensity profiles measured along a single scan line marked in cyan. There is a clear difference between pixel intensities in the close and far away images, as the far away image is much darker than the close one. Furthermore, there is a large amount of intensity variation even across the same planar surface.

In addition, the perceived intensity of a point on a surface will also depend on the surface orientation and will depend on the the angle between the surface normal and a vector directed towards the light source. The augmented image formation model, including the light intensity falloff and surface orientation effects becomes

$$m_{ij} = g_{ij} \cos\theta \sum_{\lambda=1}^{\Lambda} q_\lambda c_{i,\lambda} r_\lambda \frac{l_{0j,\lambda}}{d^2}, \tag{5}$$

where θ is the angle between the surface normal and the unit vector pointing from the surface towards the light source.

3. Intensity estimation

It is important to note that the depth d in (4) is not the distance between the scene and the camera, but rather then distance between the light source and the camera. For this reason, in order to correctly estimate the distance light travels, it is necessary to know the position of the light source, relative to the camera. Let d_{cam} represent a distance between a point in the scene and the camera, and t be the position of the point source with respect to the camera, then

$$d^2 = d_{\text{cam}}^2 + t^2 - 2 d_{\text{cam}} t \cos\phi, \tag{6}$$

where ϕ is the angle between the lines joining the point in the scene with the camera and the light source and the camera.

Given a point source an iso-intensity surface forms a sphere centered at the light position in 3D space. Therefore, given a set of equal light intensity points 3D coordinates, the position of the sphere (x_s, y_s, z_s) and its radius d can be found by minimizing [9]

$$\text{minimize} \;\; \sum_i \| \sqrt{(x_c - x_i)^2 + (y_c - y_i)^2 + (z_c - z_i)^2} - d \|_2^2. \tag{7}$$

The localization process can be made more robust by observing that all iso-intensity spheres will differ by their radii only and that the radii are related through point intensities

$$\frac{I_1}{I_2} = \left(\frac{d_2}{d_1} \right)^2. \tag{8}$$

This means that the center and the radius can be fitted to multiple iso-intensity clouds of 3D points at the same time, as long as the corresponding radii are scaled accordingly

$$\text{minimize} \;\; \sum_i \| \sqrt{(x_c - x_i)^2 + (y_c - y_i)^2 + (z_c - z_i)^2} - \sqrt{\frac{I_1}{I_i}} d_1 \|_2^2. \tag{9}$$

We arbitrarily chose the distance to the first point in the set as d_1 the reference value.

Rather than propagating light intensities from the point source intensity I_0, it is simpler to acquire a reference image of a white target I_{ref} and the corresponding depth map d_{ref}. To compute the intensity at a particular distance d_q it is sufficient to scale that reference image by the ratio of the distances

$$I_q = I_{\text{ref}} \left(\frac{d_{\text{ref}}}{d_q} \right)^2. \tag{10}$$

Even this simple model can produce very good light intensity estimates as demonstrated in Figure 2. The reference intensity is sampled using a white target was placed

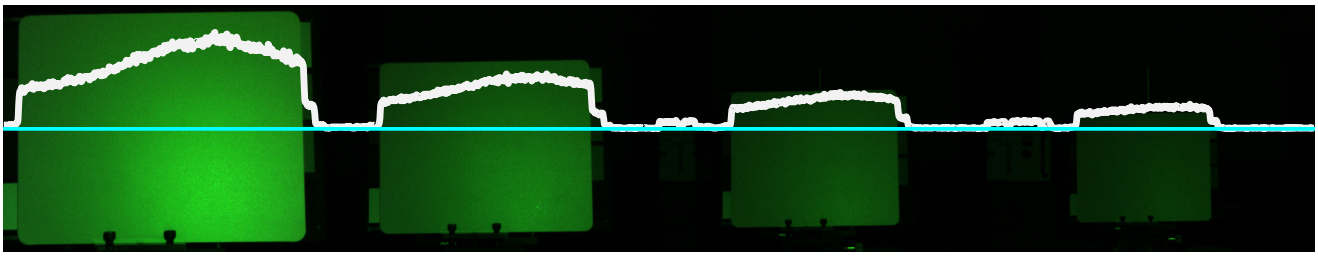


Figure 1: Light intensity decreases with scene depth. A sequence of white target images taken at depths of 50cm (left), 70cm, 90cm and 110cm (right).

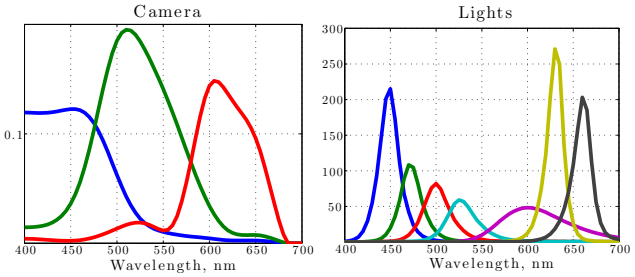
about 50cm away from the light source. We then compute the light intensity distribution for a flat white surface that is about 90cm away from the source. Since image uses a depth map obtained from an actual depth camera some of the pixels are missing and pre-processing block artifacts are also visible. Note however, that the intensity distribution is very similar to a ground truth image acquired for that depth.

4. Experiments

The experimental setup consisted of an Intel RealSense R200 RGB + depth camera and seven narrowband LEDs in the 400 to 700nm range. The RealSense camera contained a full HD (1920×1080) Bayer sensor and two VGA resolution near-infrared (NIR) sensors. The depth was estimated through stereo matching and triangulation using the NIR images and then projected onto the RGB stream using cameras intrinsic parameters and relative positions and orientations. Since a low resolution image was projected onto a high resolution one, missing depth values were estimated by image closing.

Spectral properties of the RGB sensor were derived experimentally by imaging a white reference surface illuminated with narrowband lights generated with a monochromator (Oriel Cornerstone 130). Lights spectra were measured with a PR715 monochromator. Responsivity curves were derived by solving a ridge regression problem as described in [8]. Figure 3 shows the spectral responsivity curves of the camera and the spectral distributions of the LED lights.

We conducted a simple experiment which consisted of two phases. In the first phase images of a white target, placed at three different distances away from the camera, and under all seven illuminants were acquired. This set of images was used for calibration purposes: to determine LEDs positions relative to the camera using the described sphere fitting algorithm and to capture, for every pixel, a reference intensity I_{ref} at reference depth d_{ref}. In the second part of the experiment targets placed at 50, 70, 90, and 110 cm away from the camera were acquired. The targets were a Macbeth color test chart and uniform white surface



(a) Camera (b) Lights

Figure 3: Spectral characteristics of the experiment setup.

to measure the ground truth illuminant intensity.

4.1. Results

Macbeth spectral reflectance was estimated using an unconstrained version of the estimation algorithm (3) and 12 basis functions derived from the Macbeth reflectances set [5]. Estimation RMS error values, averaged over 24 Macbeth patches, are presented in Figure 4. This figure plots the errors for three different cases. The ideal case, plotted in blue, represents the reflectance estimates using the ground truth light intensity data, these errors are invariant of the scene depth. The second curve plots the errors when the light intensity falloff is ignored altogether and it is assumed that it remains constant as the depth varies. The errors rapidly accumulate and are substantially higher than for the ideal case. Finally, the green curve shows the results obtained with the proposed algorithm. The performance is just slightly inferior to the ideal case and is significantly better than the naive estimate.

Figure 5 plots the reflectance curves for two different Macbeth chart patches at the four tested depths. Again, the results obtained using the ideal data and the proposed method are very close to ground truth reflectances. As the depth increases the naive method consistently underestimates the reflectance. This is not surprising since the algorithm overestimates the light intensity reflected of a white patch.

Figure 2: Spatial Light intensity estimation. Reference intensity image I_{ref} at about 50cm (left), estimated intensity image at about 90cm (middle) and the true intensity image at 90cm (right).

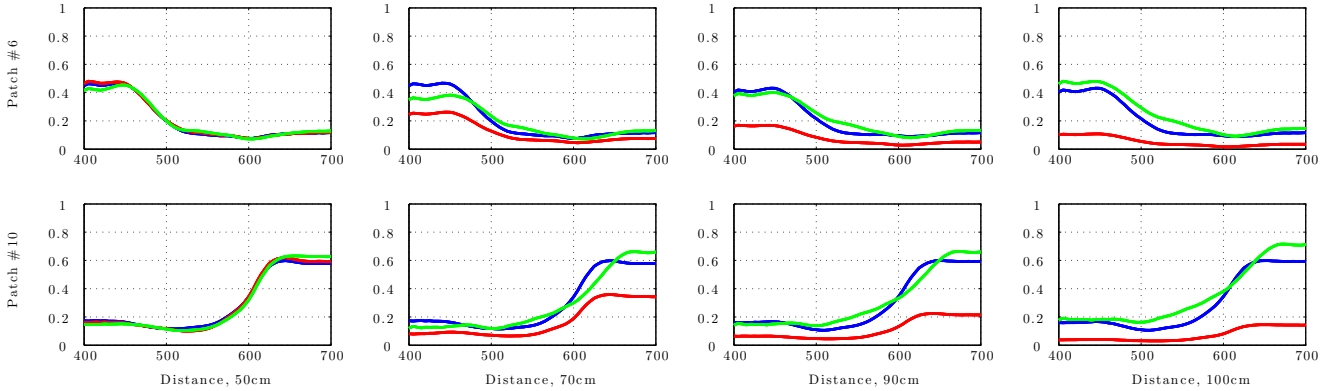


Figure 5: Macbeth reflectance estimation examples.

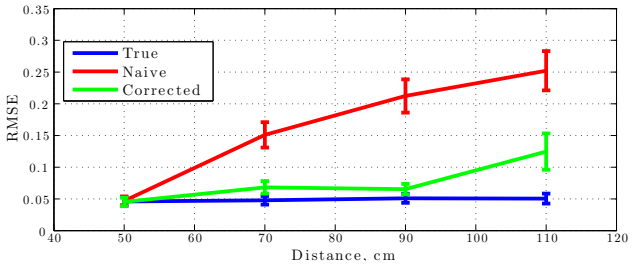


Figure 4: Macbeth reflectance estimation RMS errors.

4.2. Conclusions

In this paper we presented system capable of reliably recovering surface reflectance when illuminated with point light sources located near the scene. Our method uses an RGB+depth camera to estimate relative lights positions and then to model light intensity falloff as a function of the point source to object distance. We demonstrated that our method is almost as accurate as when the ground truth intensity is known and is vastly superior to a naive case, when light intensity is assumed to be constant.

Future work will focus on more accurate spatial light models, as many discrete light sources such as LEDs are not emitting light uniformly in all directions. Furthermore, the algorithm also needs to be evaluated when the target depth and surface orientation vary.

References

[1] A. Gowen, C. O'Donnell, P. Cullen, G. Downey, and J. Frias. Hyperspectral imaging–an emerging process analytical tool for food quality and safety control. *Trends in Food Science & Technology*, 18(12):590–598, 2007.

[2] K. M. Irion. Endoscope with LED illumination, May 4 2004. US Patent 6,730,019.

[3] R. M. Levenson and J. R. Mansfield. Multispectral imaging in biology and medicine: slices of life. *Cytometry part A*, 69(8):748–758, 2006.

[4] J.-I. Park, M.-H. Lee, M. D. Grossberg, and S. K. Nayar. Multispectral imaging using multiplexed illumination. In *Computer Vision, 2007. ICCV 2007. IEEE 11th International Conference on*, pages 1–8. IEEE, 2007.

[5] M. Parmar, S. Lansel, and J. Farrell. An led-based lighting system for acquiring multispectral scenes. In *IS&T/SPIE Electronic Imaging*, pages 82990P–82990P. International Society for Optics and Photonics, 2012.

[6] M. Parmar, S. Lansel, and B. A. Wandell. Spatio-spectral reconstruction of the multispectral datacube using sparse recovery. In *Image Processing, 2008. ICIP 2008. 15th IEEE International Conference on*, pages 473–476. IEEE, 2008.

[7] A. Ryer and V. Light. *Light measurement handbook*. Citeseer, 1997.

[8] Q. Tian, H. Blasinski, S. P. Lansel, H. Jiang, M. Fukunishi, J. E. Farrell, and B. A. Wandell. Automatically de-

signing an image processing pipeline for a five-band camera prototype using the local, linear, learned (l3) method. In *IS&T/SPIE Electronic Imaging*, pages 940403–940403. International Society for Optics and Photonics, 2015.

[9] D. Umbach and K. N. Jones. A few methods for fitting circles to data. *Instrumentation and Measurement, IEEE Transactions on*, 52(6):1881–1885, 2003.

[10] B. A. Wandell. *Foundations of vision*, volume 8. Sinauer Associates Sunderland, MA, 1995.