Radiometric Calibration with Illumination Change for Outdoor Scene Analysis

Seon Joo Kim[1] Jan-Michael Frahm[1] Marc Pollefeys [1,2]

Department of Computer Science
[1]University of North Carolina at Chapel Hill
sjkim,jmf@cs.unc.edu

Department of Computer Science
[2]ETH Zurich
marc.pollefeys@inf.ethz.ch

Abstract

The images of an outdoor scene collected over time are valuable in studying the scene appearance variation which can lead to novel applications and help enhance existing methods that were constrained to controlled environments. However, the images do not reflect the true appearance of the scene in many cases due to the radiometric properties of the camera : the radiometric response function and the changing exposure. We introduce a new algorithm to compute the radiometric response function and the exposure of images given a sequence of images of a static outdoor scene where the illumination is changing. We use groups of pixels with constant behaviors towards the illumination change for the response estimation and introduce a sinusoidal lighting variation model representing the daily motion of the sun to compute the exposures.

1. Introduction

There are millions of webcams worldwide providing videos of streets, buildings, natural sites such as mountains and beaches, and etc. The images of an outdoor scene collected over time provide a rich source of information and can lead to novel computer vision applications such as computing intrinsic images [21], building webcam synopsis [17], and geolocating webcams [7]. They are also valuable in studying the scene appearance variation which can help develop more novel computer vision applications and enhance existing computer vision methods that were constrained to controlled indoor environments. For this purposes, Narasimhan et al. introduced a database of images of a fixed outdoor scene with various weather conditions captured every hour for over 5 months in [16]. Another database of images were introduced by Jacobs et al. in [6] where they collected more than 17 million images over 6 months from more than 500 webcams across the United States. In their work, it was shown that the image sets from static cameras have consistent correlations over large spatial and temporal extents.

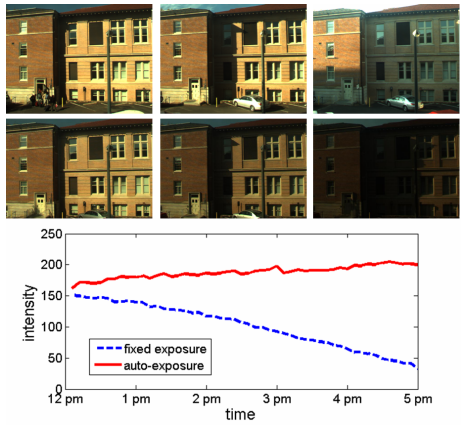

Figure 1. Effect of auto-exposure. (Top) Images taken at different times with auto-exposure (Middle) Images taken with exposure fixed (Bottom) Pixel values of a point over time.

The scene appearance depends on multiple factors including the scene geometry and reflectance, illumination geometry and spectrum, and the viewing geometry. For outdoor scenes, the weather has a large effect on the scene appearance. An important factor for determining the image appearance of a scene that is often not considered is the radiometric properties of the camera. In many computer vision systems, an image of a scene is assumed to directly reflect the appearance of the scene. However, this is not the case for most cameras as the camera response function is nonlinear. In addition, cameras usually operate in the auto-exposure mode where the exposure settings are automatically adjusted according to the dynamic range of the scene which may change the appearance of the scene in the images. Note also that this is often a necessity for outdoor scenes undergoing significant lighting variation during the day. The effect of the auto-exposure on the images is illustrated in Fig. 1 where pixel values of a point over time recorded with auto-exposure are compared with those recorded with a fixed exposure value. The sun is moving away from the scene so the radiance of the points in the

scene are decreasing as shown by the pixel values of the fixed exposure sequence. But the camera compensates for the decrease in the overall brightness of the scene resulting in increase of the pixel values. While this behavior is good for the viewing purposes, it has an ill effect on many computer vision methods that rely on the scene radiance measurement such as photometric stereo, color constancy and also on the methods that use image sequences or time-lapse data of a long period of time such as in [6], [7], and [21] since the pixel values do not reflect the actual scene radiance.

In this paper, we introduce new algorithms to compute the radiometric response function of the camera and the exposure values of images given a sequence of images of a static outdoor scene taken at a regular interval for a period of time. While the underlying assumption for our method is that the surfaces are Lambertian, our method deals with non-Lambertian surfaces such as windows and specular materials by filtering out those points in the system. Radiometric calibration on this type of data is a challenging problem because the illumination for each image is changing causing the exposure of the camera to change. Most of the previous radiometric calibration methods cannot be applied because they are based on using differently exposed images taken with constant illumination. In particular, exposures will only change in response to lighting changes which makes it hard to separate the effect of both. We solve the problem of lighting change by first selecting groups of pixels that have constant behaviors with regard to the illumination change. This means that the pixels in a group are either all in the shadows or in the non-shadow regions at a certain time in addition to having the same surface normal. The effect of the exposure and the lighting is constant for the selected pixels and the intensity difference between these pixels are due to their albedo difference assuming Lambertian surface which should remain constant over time since the albedo is a property of the material. We exploit this property to compute the response function using images with varying illumination. Estimating the exposure value for each image in the sequence after linearizing the images with the computed response function is still a difficult problem because the change in the intensity is due to the change in both the exposure and the illumination. There are countless combinations of the exposure and the illumination change that results in the same intensity change. To solve this problem, we model the illumination variation according to the motion of the sun since we are dealing with outdoor scenes.

The remainder of the paper is organized as follows. We begin by reviewing previous works on radiometric calibration in the next section. In section 3, we introduce a method for computing the camera response function using images with illumination change. Then we develop methods for computing the exposure value for each image in a sequence in section 4. We evaluate our methods with experiments in section 5 and conclude with discussion about our algorithm and future works.

2. Related Works

Majority of existing radiometric calibration methods uses multiple images taken with different exposure values to compute the camera response function. Assuming constant irradiance value which implies constant illumination, the change in intensity is explained just by the change in exposure. Using this property to solve for the response function, different models for the response function such as gamma curve [13], polynomial [15], non-parametric [1], and PCA model [5] were proposed. These methods required both the scene and the camera to be static. Some methods relieve these restrictions using the intensity mapping from one image to other computed by relating the histograms of differently exposed image [4] and by dynamic programming on the joint histogram built from correspondences [8].

A different framework for radiometric calibration was introduced by Lin et al. in [10]. In this work, a single image was used instead of multiple images by looking at the color distributions of local edge regions. They compute the response function which maps the nonlinear distributions of edge colors into linear distributions. Lin and Zhang further extended the method to deal with a single grayscale image by using the histograms of edge regions [11]. In [14], Matsushita and Lin use the asymmetric profiles of measured noise distributions to compute the response function which complements the edge based methods which may be sensitive to image noise.

Closely related to this paper are the works that use differently illuminated images to compute the radiometric response function. In [12], Manders et al propose a radiometric calibration method by using superposition constraints imposed by different combinations of two (or more) lights. Shafique and Shah also introduced a method that uses differently illuminated images in [18]. They estimate the response function by exploiting the fact that the material properties of the scene should remain constant and use cross-ratios of image values of different color channels to compute the response function. The response function is modeled as a gamma curve and a constrained non-linear minimization approach is used for the computation. Compared to these works, the algorithm proposed in this paper is more general in that the we use natural lighting condition and allow exposure changes compared to the method in [12]. In addition, we allow for more general model of the response function, do not require information across different color channels, and the response function is computed linearly as compared to the method in [18].

3. Computing the Radiometric Response Function with Illumination Change

In this section, we first introduce a method for computing the response function of a camera given multiple images of a static scene (dominantly Lambertian surface) with illumination change. The following equation explains the image formation process.

$$I_{it} = f(k_t a_i M_{it}) \tag{1}$$

The response function f transforms the product of the exposure value k, the illumination M, and the albedo a to the image intensity I. The indexes i and t denote pixel location and time respectively. The illumination M is the inner product between the surface normal $\mathbf{N}$ and the directional light $\mathbf{L}$ which in our case is the sunlight. The illumination also includes ambient lighting $L_{ambient}$.

$$M_{it} = \mathbf{N}_i \cdot \mathbf{L}_t + L_{ambient} \tag{2}$$

Eq. (1) can also be written as follows.

$$f^{-1}(I_{it}) = k_t a_i M_{it} \tag{3}$$

$$g(I_{it}) = K_t + \alpha_i + log(M_{it}) \tag{4}$$

where $g = log f^{-1}$, $K = log(k)$, and $\alpha = log(a)$.

If two points in the image have the same surface normals and both points are either both in a shadow or a non-shadow region, the amount of lighting is the same for the two points ($M_{it} = M_{jt}$)[1]. Then the relationship between the two points can be stated as follows.

$$g(I_{jt}) - g(I_{it}) = \alpha_j - \alpha_i \tag{5}$$

By using the points with same lighting conditions, the relationships between the image intensities of the points are explained only with the albedos of the points. Since the albedo of a point is constant over time, we can use Eq. (5) to compute the response function g from multiple images with different illumination.

3.1. Finding Pixels with Same Lighting Conditions

The first step necessary to compute the camera response function is to find pixels that have the same lighting conditions in all images that are used for the radiometric calibration. For different pixels to have the same lighting conditions, the surface normals of the points have to be the same and if one point is in the shadows, the other points also have to be in the shadows at that time. We modify the method proposed in [9] by Koppal and Narasimhan in which they cluster the appearance of the scene according to the surface normals. The key observation is that appearance profiles for iso-normal points exhibit similar behaviors over time (Fig. 2). An appearance profile is a vector of measured intensities at a pixel over time and they use the extrema location in the profiles to cluster the appearance. In this paper, we compute the similarity of the lighting between two pixels by simply computing the normalized correlation between the appearance profiles of the two points. With this similarity measure, we use the k-means algorithm to cluster pixels with same lighting conditions over time (Fig. 3). The clusters at this point may contain errors due to non-Lambertian regions, motions in the scene, and reflections. To deal with these errors, we then divide the image into blocks of the same size and filter out regions where all the pixels do not fall into the same cluster as illustrated in Fig. 3. Blocks with uniform intensity such as in sky are also filtered out since they don't provide valuable information for the radiometric calibration.

[1] We assume that the ambient lighting is the same for all points within a patch at a specific time

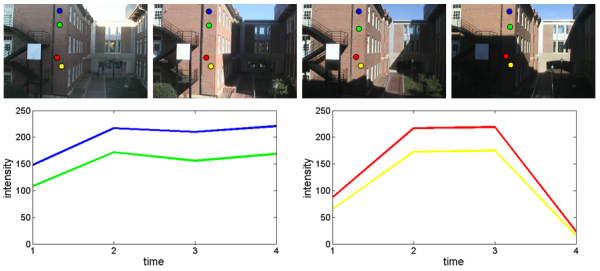

Figure 2. Using appearance profile to cluster pixels with same lighting conditions : (Top) Images used to compute the camera response function (Bottom) Appearance profiles of points with the same lighting conditions. Note that even though all the points have the same normal in the example, they have different profiles due to shadows.

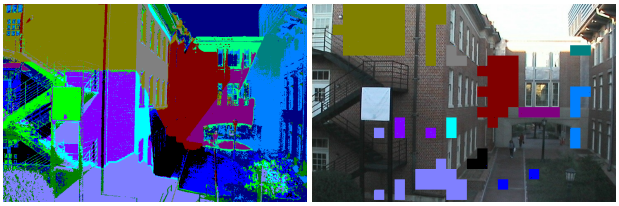

Figure 3. (Left) Pixels clustered by the appearance profile with k-means algorithm using 4 images shown in Fig. 2. (Right) Regions with non-uniform clusters are filtered out. Most of the non-Lambertian regions are filtered out at this stage.

3.2. Pixel Selection

After clustering the pixels, we then select pixels from each cluster for the response function estimation. First, we randomly pick a number of points (300 points in our experiments) from each cluster. Due to image noise and non-uniform surface, the appearance profiles for the selected pixels will be significantly disturbed by noise as shown in

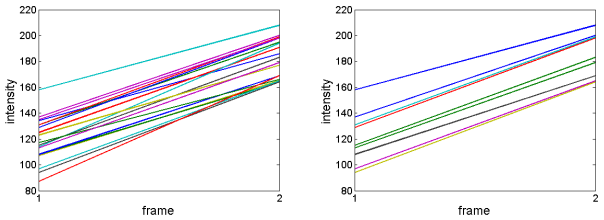

Figure 4. Pixel profiles for two frames (Left) Originally selected pixels and their profiles (Right) Profiles after postprocessing.

Fig. 4. Profiles of two pixels under the same lighting conditions crossing each other means that the albedo difference between the two points changed even though it should stay constant throughout. It is essential to filter out these outliers which can otherwise have a serious effect on the estimation results.

To remove outliers from the selected pixels for each cluster, we use the order of the pixels as the cue. The idea is that if a pixel has the lowest intensity in one frame, the intensity of that pixel should also be the lowest in the following frames. Assuming that there are n points selected for a cluster, we build a vector d_{it} of size n for each pixel i at time t where each element is :

$$\mathbf{d}_{it}(j) = \begin{cases} +1 & \text{if } I_{it} > I_{jt} \\ -1 & \text{if } I_{it} < I_{jt} \\ 0 & \text{if } I_{it} = I_{jt} \end{cases} \tag{6}$$

The dot product between d_{it} and d_{it+1} gives us how much support the pixel i has in terms of orders from other pixels in the cluster. We iteratively remove pixels with the worst support until all the pixels are in order between frames. An example of this process is shown in Fig. 4.

3.3. Radiometric Response Function Estimation

To model the response function g, we use the Empirical Model of Response (EMoR) introduced by Grossberg and Nayar in [5]. The model is a PCA based mth order approximation :

$$g(I) = \mathbf{g}_0(I) + \sum_{s=1}^{m} c_s \mathbf{h}_s(I) \tag{7}$$

where the g_0 is the mean function and the c_k's are the coefficients for the basis functions $\mathbf{h}_k$'s. Combining Eq. (5) and Eq. (7), we have

$$\sum_{s=1}^{m} c_s(\mathbf{h}_s(I_{jt}) - \mathbf{h}_s(I_{it})) - \alpha_{ji} = g_0(I_{it}) - g_0(I_{jt}) \tag{8}$$

where $\alpha_{ji} = \alpha_j - \alpha_i$.

For n pixels in the same cluster l at time t, we have $n-1$ linear equations $\mathbf{A}_{tl}\mathbf{x_t} = \mathbf{b}_{tl}$ as follows.

$$\mathbf{A}_{tl} = [\mathbf{A'}_{tl} \quad \mathbf{I}_{n-1}] \tag{9}$$

$$\mathbf{A'}_{tl}(y,x) = \mathbf{h}_x(I_{y+1,t}) - \mathbf{h}_x(I_{1t}), \quad 1 \le y \le n-1,\ 1 \le x \le m, \tag{10}$$

$$\mathbf{b}_{tl}(y) = g_0(I_{1t}) - g_0(I_{y+1,t}), \quad 1 \le y \le n-1 \tag{11}$$

$$\mathbf{x_t} = [\mathbf{c}, \mathbf{a}_l]^T \tag{12}$$

where $\mathbf{I}_{n-1}$ is an identity matrix of size $n-1$ by $n-1$, $\mathbf{c} = [c_1, c_2, \ldots, c_m]$ and $\mathbf{a}_l = [\alpha_{21}, \alpha_{31}, \ldots, \alpha_{n1}]$.

Since we have m+n-1 unknowns with n-1 equations, the system above is underconstrained. We can add more equations to the system by incorporating the temporal information of multiple frames. The number of points n is typically bigger than the number of basis functions (m = 5 in this paper), so as few as two frames are enough to solve for the response function. Since one cluster typically does not provide enough range of intensities for accurate estimation, we combine equations from different clusters. Adding multiple clusters at multiple frames, we can compute the response function by solving the following least squares problem $\mathbf{Ax} = \mathbf{b}$ with (assuming we are using 3 clusters from 2 frames for simplicity)

$$\mathbf{A} = \begin{pmatrix} \mathbf{A'}_{11} & \mathbf{I}_{n-1} & \mathbf{0} & \mathbf{0} \\ \mathbf{A'}_{21} & \mathbf{I}_{n-1} & \mathbf{0} & \mathbf{0} \\ \mathbf{A'}_{12} & \mathbf{0} & \mathbf{I}_{n-1} & \mathbf{0} \\ \mathbf{A'}_{22} & \mathbf{0} & \mathbf{I}_{n-1} & \mathbf{0} \\ \mathbf{A'}_{13} & \mathbf{0} & \mathbf{0} & \mathbf{I}_{n-1} \\ \mathbf{A'}_{23} & \mathbf{0} & \mathbf{0} & \mathbf{I}_{n-1} \end{pmatrix} \tag{13}$$

$$\mathbf{b} = [\mathbf{b}_{11}, \mathbf{b}_{21}, \mathbf{b}_{12}, \mathbf{b}_{22}, \mathbf{b}_{13}, \mathbf{b}_{23}]^T \tag{14}$$

$$\mathbf{x} = [\mathbf{c}, \mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3]^T \tag{15}$$

In practice, rows of $\mathbf{A}$ and $\mathbf{b}$ are weighted according to the intensity of the pixel used for the row. The weights are included because response functions typically have a steep slope near the extreme intensities causing the data to fit poorly in those regions. We used a Gaussian function centered at the intensity of 128 with the standard deviation ranging from 0.25 to 2.0 as the weighting function.

The solution to the problem $\mathbf{Ax} = \mathbf{b}$ above suffers from the exponential (or scale in the log space) ambiguity [4] which means that the whole family of $\gamma\mathbf{x}$ are the solutions to the problem. To fix the scale of the response function, we set the value of the response function at the image value 128 to a value τ. We will discuss this ambiguity later in Section 4.

4. Exposure Estimation from Images with Different Illumination

By using the computed response function, we can linearize the images as in Eq. (3). While the images taken at different times are now linearly related, the images may not reflect the true appearance of the scene due to the exposure

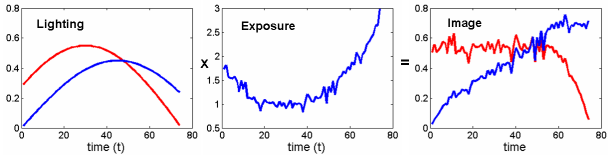

Figure 5. Relationship between lighting, exposures, and image appearance. We need at least two pixel profiles to compute exposures since many combination of lighting and exposure can result in a same profile.

change in the camera. However, there is an inherent ambiguity in computing the exposures from images with different illumination similar to the exponential ambiguity mentioned in the previous section. As can be seen from Eq. (3), there is an infinite number of combinations of the exposure and the lighting that result in the same image intensity. To compute the exposure, assumptions on the lighting have to be made.

In this section, we introduce a method to estimate the exposure values given a sequence of images of an outdoor scene taken over a period of time. For the outdoor scenes, the dominant source of lighting is the sun. We model the lighting change according to the motion of the sun and use the model to compute the exposures. We make the assumption that the sunlight was not blocked by clouds when the images were taken.

4.1. Modeling the Illumination with the Motion of the Sun

The direction of the sunlight ($\mathbf{L}_t$) at time t and the surface normal of a point i ($\mathbf{N}_i$) can be expressed in Cartesian coordinates as in the following equation where θ's are the azimuth angles and ϕ's are the elevation angles.

$$\begin{aligned}\mathbf{L}_t &= [\cos\phi_t\cos\theta_t, \cos\phi_t\sin\theta_t, \sin\phi_t]^T \\ \mathbf{N}_i &= [\cos\phi_i\cos\theta_i, \cos\phi_i\sin\theta_i, \sin\phi_i]^T\end{aligned} \tag{16}$$

The lighting due to the sun at point i is then

$$\mathbf{N}_i \cdot \mathbf{L}_t = \cos\phi_t\cos\phi_i\cos(\theta_t - \theta_i) + \sin\phi_t\sin\phi_i \tag{17}$$

Without loss of generality we rotate $\mathbf{L}_t$ and $\mathbf{N}_i$ so that $\phi_t = 0$, Eq. (17) becomes

$$\begin{aligned}\mathbf{N}_i \cdot \mathbf{L}_t &= \cos\phi_i'\cos(\theta_t - \theta_i') \\ &= \cos\phi_i'(\cos\theta_i'\cos\theta_t + \sin\theta_i'\sin\theta_t) \\ &= p_i\cos\theta_t + q_i\sin\theta_t\end{aligned} \tag{18}$$

where $p_i = \cos\phi_i'\cos\theta_i'$ and $q_i = \cos\phi_i'\sin\theta_i'$. According to Eq. (18), the lighting variation at a point due to the sun over time is a sinusoidal function with the scale and the phase being the parameter.

4.2. Exposure Estimation

Combining Eq. (2), Eq. (3), and Eq. (18) we have

$$\frac{1}{k_t}f^{-1}(I_{it}) - p_i'\cos\theta_t - q_i' sin\theta_t = 0. \tag{19}$$

In the above equation, p_i' and q_i' are considered to include the albedo term a from Eq. (3). Additionally, we assumed that the effect of ambient lighting is constant over time. Note that at least two appearance profiles of different surface normals are necessary to compute the exposures k_t's using Eq. (19) as shown in Fig. 5.

For simplicity, it is assumed that we have a sequence of η images ($1 \le t \le \eta$) with pixel values (I_{it} and I_{jt}) of two points with different surface normals. From Eq. (19), the exposure k_t for each time t is computed by solving a linear least squares problem $\mathbf{U}\mathbf{y} = \mathbf{0}$ with

$$\mathbf{U} = \begin{pmatrix} \mathbf{S} & \mathbf{0}_{\eta\times 2} & \mathbf{F}_i \\ \mathbf{0}_{\eta\times 2} & \mathbf{S} & \mathbf{F}_j \end{pmatrix}, \tag{20}$$

$$\mathbf{S} = \begin{pmatrix} \cos\theta_1 & \sin\theta_1 \\ \cos\theta_2 & \sin\theta_2 \\ \vdots & \vdots \\ \cos\theta_\eta & \sin\theta_\eta \end{pmatrix}, \tag{21}$$

$$\mathbf{F}_i = \begin{pmatrix} f^{-1}(I_{i1}) & 0 & \cdots & 0 \\ 0 & f^{-1}(I_{i2}) & \cdots & 0 \\ & & \ddots & \\ 0 & 0 & \cdots & f^{-1}(I_{i\eta}) \end{pmatrix}, \tag{22}$$

$$\mathbf{y} = [p^1, q^1, p^2, q^2, k_1', k_2', \cdots, k_\eta'] \tag{23}$$

where $k_t' = 1/k_t$, $\theta_t = 2\pi t(\frac{1}{24}\frac{s_{min}}{60})$, and s_{min} is the sampling interval in minutes.

A set of pixels used to solve the equation above are randomly selected from the clusters used for the response function estimation (Section 3). It is important not to use pixel values at time t in the above equation when the pixels fall into shadows since the lighting model does not apply to shadow regions. From the appearance profile of a pixel, we detect whether the pixel is in shadow by a simple thresholding as in [20]. We also remove the pixels from the equation if the average intensity is too low meaning that the pixels were probably always in the shadow.

4.3. Exponential Ambiguity

In Section 3 we discussed the inherent ambiguity in computing the response function where the elements in Eq. (3) are related exponentially as follows.

$$(f^{-1}(I_{it}))^\gamma = k_t^\gamma a_i^\gamma M_{it}^\gamma \tag{24}$$

We resolved this exponential ambiguity by arbitrarily fixing the scale of the response function which is not a problem

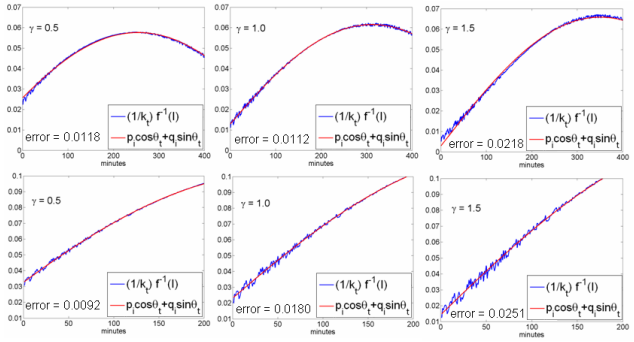

Figure 6. Simulation of the effect of the exponential ambiguity. The exposures and the lighting changes were estimated on the two synthetically generated image profiles similar to Fig. 5 using 400 minutes of data (top) and 200 minutes of data (bottom). The correct γ is 1.0.

for applications that require image intensity alignment since different γ's still result in the same intensity value. However the arbitrary scale causes problems in methods that require accurate or linearly proportional photometric measurements such as photometric stereo or motion blur simulation as in [1]. It also affects our exposure estimation process since our method is based on having the right scale for the response function f. If the scale of the response function is incorrect, then the system is trying to fit a sine function to a measurement that is the exponent of a sine. As shown in [4], prior knowledge about the response function f or k (also a or M in our case) is necessary to find the right γ value. Ideally, the error $\|\mathbf{U}\mathbf{y}\|$ in Section 4.2 gives us the information about the γ. It should be the minimum when the correct scale of the response function is used. However, the error is not distinctive due to image noise and lack of time interval when surfaces of different normals are both in the sunlight as shown in Fig. 6. Alternatively, we need information about the camera or the scene to find the right scale. In this paper, we first estimate the exposures (k_t) and the lighting functions ($p'_i \cos\theta_t + q'_i sin\theta_t$ in Eq. 19) using multiple γ values. The recovered lighting functions will have different phases with different γ's as shown in Fig. 6. We select the γ value that yields the lighting functions to have the peaks at the right time of the day which can be inferred from the orientations of shadows in the image sequence.

5. Experiments

We first evaluate our response function estimation method introduced in Section 3. Two cameras used for the experiments are Sony SNC-RZ30N PTZ camera and Point Grey Dragonfly camera. For the Sony camera, we first computed the response function by using the method introduced in [8] with multiple images of a static scene with constant illumination to test our method. We then computed the response function with our method using four images shown in Fig. 2 and the comparison of the computed response functions is shown in Fig. 7. While only the green channel was used for this example, we can easily combine all channels if necessary. For the Point Grey Dragonfly camera, we compare our result computed with two images with the known linear response function which is shown in Fig. 8. The number of images for accurate estimation depends on the intensity range of each image. While the method does not need a large number of points, it is important to have a well distributed pixel intensities for an accurate estimation.

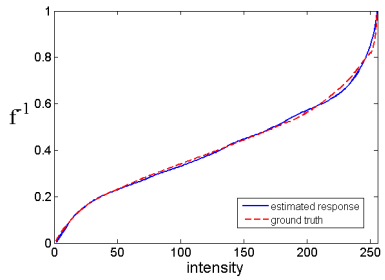

Figure 7. Response function estimation result for Sony SNC-RZ30N PTZ camera

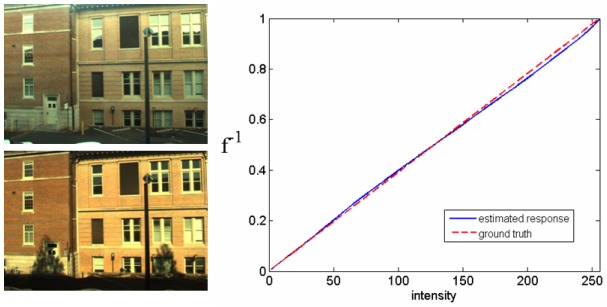

Figure 8. Response function estimation result for Point Grey Dragonfly camera with two images used for our estimation.

To evaluate our exposure estimation method, we recorded images of a scene every minute for a little more than 4 hours with the Point Grey camera when we could observe surfaces with different normals being illuminated by the sun. Some sample images as well as some of the pixel profiles used for the estimation are shown in Fig. 9. Our exposure estimates are compared to the ground truth exposures reported by the camera in Fig. 10. Notice that the exposure estimates start to deviate from the ground truth starting around 1400. The cause for this is the change in ambient lighting as a building in front of the scene started to block the sunlight at that time. Since our method is based on constant ambient lighting, the change in the ambient lighting caused errors in the exposure estimation. However, for a long periods time when the ambient lighting was close to constant, our estimation was accurate as shown in the figure. We can observe the function of the auto-exposure from

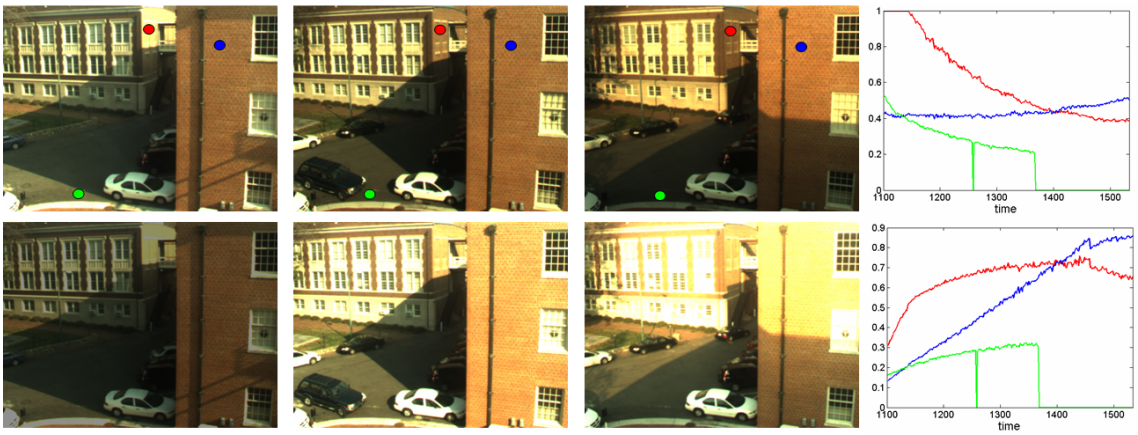

Figure 9. Exposure Estimation. (Top) Sample images from the input sequence and the pixel profiles of the dotted points (Bottom) Images and profiles normalized to a fixed exposure. The 0 values in the profiles represent shadow.

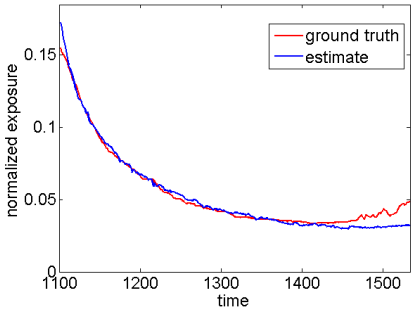

Figure 10. Comparison of the estimation with the ground truth exposure

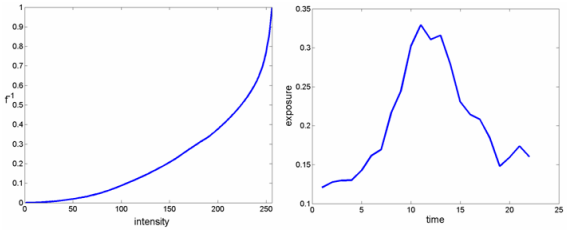

Figure 11. Estimated response function (left) and the exposures (right) using dataset introduced in [6] (Fig. 12)

Fig. 9. The camera adjusts to the brightness change by trying to fix the intensity of dominant pixels constant. This function prohibits images from being under-exposed or saturated as can be seen from the exposure compensated images in the figure. While this is good for viewing, this could affect vision algorithms that rely on photometric measurements since the image intensities do not reflect true radiance of the points. By computing the response function and exposures using our method, we can convert the image intensities to their actual radiance enabling further analysis of the scene.

As our last experiment, we used one of the webcam datasets introduced in [6] as shown in Fig. 12. The images we used were captured every 30 minutes for 11 hours. The estimated response function and the exposures are shown in Fig. 11. Note that we do not have the ground truth for this data since the camera is unknown. We can roughly evaluate the results by comparing the input images and the pixel profiles with the images and the profiles normalized with the estimated exposures as in Fig. 12. Input profiles tend to stay constant unless affected by shadows. However, after normalizing the images with the estimated exposures, the pixel values vary gradually as expected.

6. Conclusion

We have introduced a novel method for computing the camera response function for outdoor image sequences in which the illumination is changing. This is a challenging problem because the image appearance varies due to the changes in both the exposure of the camera and the lighting. Most previous methods cannot deal with the illumination change and the methods that deal with the change are restrained to some special cases [12, 18]. Our approach also computes the exposure values of the camera with the illumination in the scene changing and we believe this work can serve as a basis for more exciting outdoor scene analysis applications.

For the future, we would like to extend our method to find the right scale of the response function automatically. One possibility would be using high-order correlations in the frequency domain such as in [2]. We also plan to enhance our algorithm to take into account the change in ambient lighting as well as the change in lighting due to weather. Additionally, we would also like to expand our method to use images from commodity photo collections such as in [19] and [3] which can be used for texture alignment and also help improve image matching.

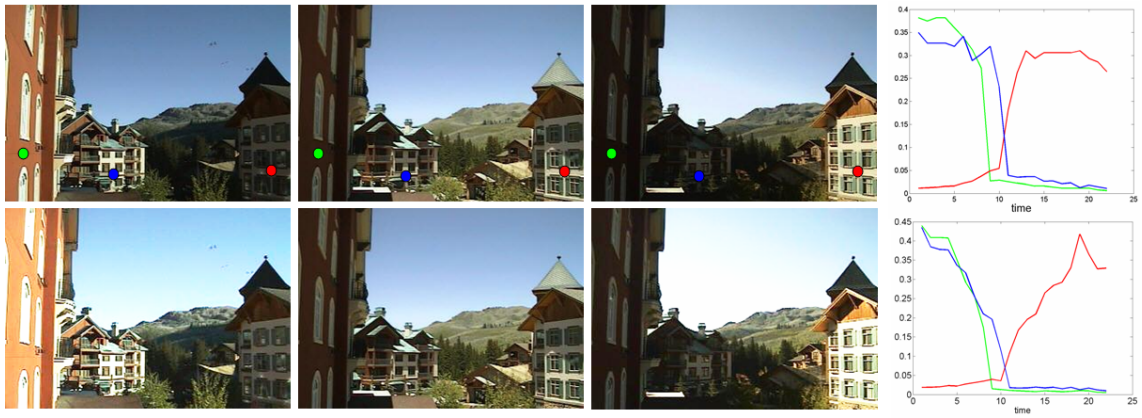

Figure 12. (Top) Sample images from one of the dataset introduced in [6] and the pixel profiles of the dotted points. (Bottom) Images and profiles normalized to a fixed exposure. The right side of the figure is to the east.

Acknowledgments

We gratefully acknowledge the support of the NSF Career award IIS 0237533 as well as a Packard Fellowship for Science and Technology.

References

[1] P. Debevec and J. Malik. Recovering high dynamic range radiance maps from photographs. *Proc. SIGGRAPH'97*, pages 369–378, Aug. 1997. 2, 6

[2] H. Farid. Blind inverse gamma correction. *IEEE Transaction on Image Processing*, 10(10):1428–1433, 2001. 7

[3] M. Goesele, N. Snavely, B. Curless, and S. M. S. H. Hoppe. Multi-view stereo for community photo collections. *Proc. IEEE Int. Conf. on Computer Vision*, 2007. 7

[4] M. Grossberg and S. Nayar. Determining the camera response from images: What is knowable? *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 25(11):1455–1467, 2003. 2, 4, 6

[5] M. Grossberg and S. Nayar. Modeling the space of camera response functions. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 26(10):1272–1282, 2004. 2, 4

[6] N. Jacobs, N. Roman, and R. Pless. Consistent temporal variations in many outdoor scenes. *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, pages 1–6, 2004. 1, 2, 7, 8

[7] N. Jacobs, S. Satkin, N. Roman, R. Speyer, and R. Pless. Geolocating static cameras. *Proc. IEEE Int. Conf. on Computer Vision*, 2007. 1, 2

[8] S. J. Kim and M. Pollefeys. Robust radiometric calibration and vignetting correction. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 30(4):562–576, 2008. 2, 6

[9] S. J. Koppal and S. G. Narasimhan. Clustering appearance for scene analysis. *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, pages 1323 – 1330. 3

[10] S. Lin, J. Gu, S. Yamazaki, and H. Shum. Radiometric calibration from a single image. *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, pages 938–945, 2004. 2

[11] S. Lin and L. Zhang. Determining the radiometric response function from a single grayscale image. *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, pages 66–73, 2005. 2

[12] C. Manders, C. Aimone, and S. Mann. Camera response function recovery from different illuminations of identical subject matter. *Proc. IEEE Int. Conf. on Image Processing*. 2, 7

[13] S. Mann and R. Picard. On being 'undigital' with digital cameras: Extending dynamic range by combining differently exposed pictures. *Proc. IS&T 46th annual conference*, pages 422–428, May 1995. 2

[14] Y. Matsushita and S. Lin. Radiometric calibration from noise distributon. *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, pages 1 – 8. 2

[15] T. Mitsunaga and S. Nayar. Radiometric self-calibration. *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, pages 374–380, June 1999. 2

[16] S. G. Narasimhan, C. Wang, and S. K. Nayar. All the images of an outdoor scene. *Proc. of the European Conference on Computer Vision*, pages 148–162, 2002. 1

[17] Y. Pritch, A. Rav-Acha, A. Gutman, and S. Peleg. Webcam synopsis: Peeking around the world. *Proc. IEEE Int. Conf. on Computer Vision*, 2007. 1

[18] K. Shafique and M. Shah. Estimation of the radiometric response functions of a color camera from differently illuminated images. *Proc. IEEE Int. Conf. on Image Processing*, pages 2339–2342, 2004. 2, 7

[19] N. Snavely, S. M. Seitz, and R. Szeliski. Photo tourism: Exploring photo collections in 3d. *ACM Transactions on Graphics*, 25(3):835–846, 2006. 7

[20] K. Sunkavalli, W. Matusik, H. Pfister, and S. Rusinkiewicz. Factored time-lapse video. *ACM Transactions on Graphics*, 26(3), 2006. 5

[21] Y. Weiss. Deriving intrinsic images from image sequences. *Proc. IEEE Int. Conf. on Computer Vision*, pages 68–75, 2001. 1, 2