Radiometric Calibration Using Temporal Irradiance Mixtures

Bennett Wilburn
Microsoft Research Asia
Beijing, China
bwilburn@microsoft.com

Hui Xu*
TsingHua University
Beijing, China
xu-h06@mails.tsinghua.edu.cn

Yasuyuki Matsushita
Microsoft Research Asia
Beijing, China
yasumat@microsoft.com

Abstract

We propose a new method for sampling camera response functions: temporally mixing two uncalibrated irradiances within a single camera exposure. Calibration methods rely on some known relationship between irradiance at the camera image plane and measured pixel intensities. Prior approaches use a color checker chart with known reflectances, registered images with different exposure ratios, or even the irradiance distribution along edges in images. We show that temporally blending irradiances allows us to densely sample the camera response function with known relative irradiances. Our first method computes the camera response curve using temporal mixtures of two pixel intensities on an uncalibrated computer display. The second approach makes use of temporal irradiance mixtures caused by motion blur. Both methods require only one input image, although more images can be used for improved robustness to noise or to cover more of the response curve. We show that our methods compute accurate response functions for a variety of cameras.

1. Introduction

Many computer vision algorithms assume that image pixel intensities are linearly related to scene irradiance. For most cameras, however, this is not the case. Camera manufacturers often deliberately engineer nonlinear response functions into their cameras in order to match film characteristics or to account for nonlinear characteristics of computer displays and the human visual system. Even if the response is intended to be linear, the analog circuitry in the image sensors themselves introduces small nonlinearities. Thus, the camera's *radiometric response function*, which maps scene irradiance to measured pixel intensities, is generally a nonlinear function. Therefore, we must estimate the response function in order to linearize the image intensities and improve the performance of vision algorithms.

The response function f is represented as

$$M = f(I), \tag{1}$$

where I represents the irradiance at sensor, and M is the measured pixel intensity. Conventionally the axes are normalized to measurement and irradiance ranges in $[0, 1]$. This function is assumed to be constant throughout the image. Imaging and computer vision algorithms are concerned with irradiance, so we generally solve for the inverse radiometric response function, $g = f^{-1}(M)$.

We propose a radiometric calibration approach based on temporally mixing two different irradiances within a single camera exposure. Cameras accumulate incoming light for the duration of their exposure, and scene brightness at any given pixel may vary due to motion blur or illumination changes. We constrain the incident brightness to be one of only two values. For example, we could use motion blur across the boundary between two uniformly colored regions. Although we do not know the accumulated irradiance at each pixel, we do know that if the motion is linear, the brightness in the blurred region will vary linearly along the direction of motion. For a camera with a non-linear response, however, the measured pixel intensities will not vary linearly. Similarly, we can use a simple illumination change, like a rapidly blinking light, to create brightness mixtures. In this case, the irradiance at each pixel is a linear blend of the irradiances with the light on and off. Linearly increasing the portion of the exposure that the light is on creates a linearly increasing series of accumulated irradiances. Again, this linearity does not hold for the measured pixel intensities if the response function is nonlinear. Our method exploits these nonlinear intensity measurements to estimate the response functions.

In the rest of this section, we describe how our approach complements prior work in radiometric calibration. Section 2 explains the general idea of temporal irradiance mixing. In sections 3 and 4, we present two specific methods to use this idea for radiometric calibration. Section 5 shows results with various cameras validating our techniques, and section 6 discusses the implications of our work and av-

---

*This work was done while Hui Xu was an intern at Microsoft Research Asia.

enues for future research.

1.1. Prior Work

Radiometric calibration methods require means to collect samples of the radiometric response function with some known relationship. For example, one approach is to use an image of a uniformly illuminated Macbeth color chart [1], which has color patches with known reflectances. Mann and Picard [9] calibrate using multiple registered images of a static scene with different known exposure times. The ratio of the irradiance at the same pixel in different images is equal to the ratio of the exposure times of the images. They use this relationship to solve for a parametric response function of the form $M = \alpha + \beta I^\gamma$. Debevec and Malik [2] also use multiple images with different exposures, but solve for a smooth non-parametric function. Mitsunaga and Nayar's method [11] requires only rough estimates of the exposure times and iteratively computes a polynomial inverse response function and more accurate estimates of the exposure times. Tsin et al. [13] use a statistical model of the CCD image formation process to iteratively estimate non-parametric inverse response functions.

Several methods have been developed that use multiple exposures, but do not require precise registration. Grossberg and Nayar [3] use the relationship between the intensity histograms of two scenes imaged with different exposures, because intensity histograms are relatively unaffected by small changes in the scene. Kim and Pollefeys [5] compute point correspondences between images. Mann [8] estimates response functions from a rotating and zooming camera.

Lin et al. [6] compute the radiometric response function from RGB distributions along color edges in a single image. They use a prior model of response functions compiled by Grossberg and Nayar [4] to compute the radiometric response function as the one which maps nonlinear measured RGB intensity distributions to linear ones. Lin and Zhang [7] adapt this idea to grayscale images by measuring the nonuniformity of edge intensity histograms. Recently, Matsushita and Lin [10] proposed a method to compute camera response functions from noise distributions. Analyzing noise distributions requires much input data, although a single image can be used under very carefully controlled conditions.

Our work differs from the prior art in that we use temporal color mixtures to directly sample the response function. We do not need special equipment or multiple registered images with precise exposure control. Our methods require only a single image, do not rely on statistical priors over the response function, and work with color or grayscale images. In some ways our approach complements the work of Lin et al. [6] in that we use temporal mixtures, while their methods use spatial mixtures. Although they mention defocus and motion blur as sources of linear irradiance blends, they do not test this idea. Moreover, calibration from temporal color mixtures frees us from reliance on indirect, statistical methods and priors over camera response functions.

2. Calibration Using Temporal Irradiance Mixtures

Following the notation of Mitsunaga and Nayar [11], we can describe the relationship between image irradiance I and scene radiance L as

$$E(t) = L(t)\frac{\pi}{4}\left(\frac{d}{h}\right)^2 \cos^4(\phi), \qquad (2)$$

where h is the lens focal length, d is the aperture diameter, and ϕ is the angle subtended by the principal ray from the imaging axis. We will simplify this to $E = kL(t)$, where $k = \frac{\pi}{4}(\frac{d}{h})^2 \cos^4(\phi)$. We assume, for now, that k is constant. Image sensors and film are integrating devices, so the total integrated irradiance for an exposure that lasts from time t_0 to time t_1 is

$$I = \int_{t_0}^{t_1} kL(t)dt. \qquad (3)$$

Prior methods assume that $L(t)$ is constant. Let us assume, instead, L(t) can be one of only two values, L_0 or L_1. Then the integrated irradiance at the sensor becomes

$$I = \alpha I_0 + (1 - \alpha)I_1, \qquad (4)$$

where $0 \leq \alpha \leq 1$ is the portion of time that the scene radiance is L_0. I_0 and I_1 are the accumulated irradiances corresponding to a constant scene radiance of L_0 or L_1, respectively, for the entire exposure, i.e., $I_0 = \int_{t_0}^{t_1} kL_0(t)dt$ and $I_1 = \int_{t_0}^{t_1} kL_1(t)dt$. We can generate any convex combination of I_0 and I_1 by varying α. We propose to use these blends to sample the response function for radiometric calibration. Our method can be applied to grayscale or color images; for color images, we can use the same mixtures to sample and independently calibrate the three color channels.

2.1. Scale and Offset Ambiguities

The irradiances in our mixtures are unknown, so our measurements have two ambiguities. Figure 1 shows this graphically. We are blending a low irradiance I_{lo} and a high irradiance I_{hi}, corresponding to measured intensities M_{lo} and M_{hi}, respectively. Because I_{lo} is unknown, there may be a global offset to the curve. We alleviate this by making I_{lo} as close to zero as possible. I_{hi} is also unknown, leading to a scale ambiguity for the curve in the measured region. Often, we are able to set I_{lo} below the noise floor and I_{hi} above the saturation level, enabling us to capture the entire response curve. Regardless, even with both ambiguities, it is still possible to linearize the measured intensities.
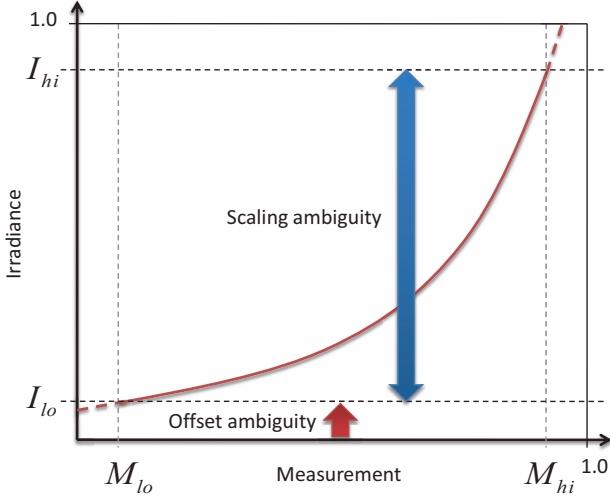
Figure 1. Scale and offset ambiguities for radiometric calibration. The response curve computed from our method suffers from two ambiguities. I_{lo} is unknown, so there may be a global offset to the curve. We alleviate this by making I_{lo} as close to zero as possible. I_{hi} is also unknown, leading to a scale ambiguity for the curve in the measured region.
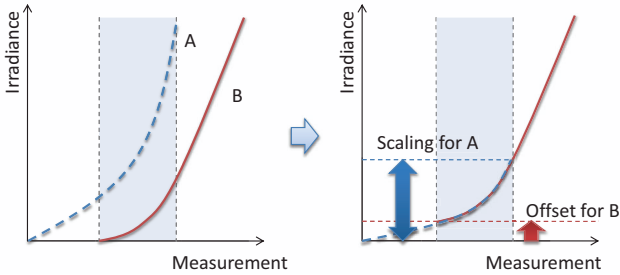

Figure 2. Curve merging. On the left, we see an example with two measured irradiance blends. Each curve has an offset and scale ambiguity. To merge the data from both curves, we must compute the scale and offset that causes the two curves to coincide. To do this, we simultaneously solve for a camera inverse response function $g(M)$ and scale/offset values for each mixture that best align the data to $g(M)$. The desired result is shown on the right.

2.2. Combining Multiple Irradiance Blends

We may want to use data from many different irradiances mixtures to improve our estimate of the response function. For example, we could increase the range of our measurements by using multiple mixtures that span a broader range of irradiances. Across a motion-blurred edge, we can use many lines of pixel data to improve robustness to noise. In either case, we estimate the low and high irradiances from each mixture in order to combine the data. Figure 2 shows this process graphically.

Mathematically, the problem is as follows. We observe n irradiance mixtures. The number of observations in the j-th mixture, for $j \in (0, ..., n-1)$, is l_j. Each mixture is represented by a set of measurement-irradiance pairs, (M_{ij}, I_{ij}), for $i \in (0, ..., l_j - 1)$. M_{ij} is the pixel value for observation i in mixture j. I_{ij} is the relative irradiance for that measurement, up to the scale and offset ambiguity. Thus, it could correspond to α from equation 4. Or, for evenly distributed samples between two irradiances (a linear ramp), the I_{ij} could simply be 0,1,2,...,$(l_j - 1)$, i.e., $I_{ij} = i$.

We want to compute an inverse response function and a set of irradiance scales and offsets that aligns these measurements. We denote the offsets by r_j and the scales by s_j. We assume the inverse response function $g(M)$ can be modeled by a polynomial of degree d with coefficients a_k, i.e., $g(M) = \sum_{k=0}^{d} a_k M^k$. We solve for a, r and s that minimize the following error function:

$$E = \sum_{j=0}^{n-1} \sum_{i=0}^{l_j} \left(s_i I_{ij} + r_j - \sum_{k=0}^{d} a_k M_{ij}^k \right)^2. \quad (5)$$

We add a monotonicity constraint that forces the first derivative of the inverse response function to be non-negative. The scale and offset of the final curve is still ambiguous, so we normalize it by setting a_0 to zero and adding a constraint that $\sum_k a_k = 1$.

3. Display-Based Calibration Method

If the radiance of a computer monitor were linearly related to the displayed pixel value, we could easily measure a camera's response curve by displaying different known radiances and recording the intensities measured by the camera for a fixed exposure time. Most displays, however, have a non-linear relationship between the display pixel value and emitted intensities. Using temporal color mixtures, we can use an uncalibrated display to generate a linear ramp of irradiances. We take advantage of the high frame rate of computer displays to show several different images during a single camera exposure. For some fraction α of the exposure, we display pixel value P_{lo}, and for the remainder we display P_{hi}. As in equation 4, this allows us to make known linear combinations of the irradiances I_{lo} and I_{hi} corresponding to the two selected pixel values.

For calibration, we choose P_{lo} to be completely dark, and P_{hi} high enough to saturate the image sensor. If we can display N images in one camera exposure, then we can use a sequence of exposures to generate $N+1$ samples of the response function that are evenly spaced in irradiance. Fitting a curve through the measured values produces a continuous representation of the camera response function.

As described, this method requires one image for each irradiance measurement. We can accelerate the data capture by dividing the displayed image into a grid of patches, and using each patch to display a different linear combination of radiances. In this way, we can calibrate the camera

response curve using just one image. Figure 3 shows an example of this process. For cameras with very nonlinear responses, one irradiance blend might poorly sample some part of the response curve. In these cases, we extend the method again by simultaneously displaying multiple grids to produce multiple blends. We then combine the data from the blends as described in section 2.2. In this way, we can add more measurements to poorly sampled regions.

Implementation. We chose to implement our display-based calibration using a regular LCD computer monitor. Provided the camera exposure duration can be set to a multiple of the display time for the entire image sequence, we do not need to synchronize the camera and display. We simply continuously repeat the sequence of images for the grid pattern (see figure 3).

LCD monitors might have other characteristics that could interfere with our measurements: unequal on and off transition times for the LCD material, backlight flicker, or inversion (switching the polarity of the applied LCD voltage at each pixel every frame to prevent damage to the LCD material). To minimize these effects, we use a relatively long exposure time of one second, display each image in the sequence for an even number of frames (four), and ensure that each pixel only switches its display value once during the sequence.

One phenomena we observed when implementing our grid-based method was that the brightness of each patch was heavily influenced by bright surrounding regions. We found that separating the patches by a black border eliminated this effect. We use 30×30 pixel patches with 30 pixel black borders between all patches. To prevent artifacts due to non-uniform illumination of the display or vignetting in the camera, we made sure that the overall size of the displayed pattern was small relative to the display and camera image dimensions.

4. Calibration from Motion Blur

Our display-based method requires a display with known frame rate and a camera with accurate control of the exposure duration. For situations which preclude that method, we offer a second approach: calibration from motion blur. Figure 4 shows a simple way to use motion blur for radiometric calibration. We create a target with a step edge between two uniformly colored regions, place the target fronto-parallel to the camera, then take a picture of the target while moving the camera so the direction of blur is across the edge. Of course, one could also move the target. Pixels in the blurred region along the edge will measure a blend of the two uniform irradiances. If the target is uniformly illuminated and the motion is constant and parallel to the plane of the target, the irradiance at the camera will increase (or decrease) linearly from pixel to pixel across the edge. This produces a series of irradiances at the sensors that are linearly spaced between the irradiances for the two uniform regions.

By choosing one region to be black and the other to be bright enough to saturate the camera, we can capture the entire range of the camera response. Moving the camera quickly enough (or exposing for long enough) to blur over many pixels will produce many samples of the response function. We can combine the blurred values from many lines to estimate the camera response function as described in section 2.2.

Implementation. Some care must be taken to calibrate the camera radiometric response using motion blur. The relative motion between the camera and the target must be constant, roughly parallel to the plane of the target, and with some component across the edge. The target must also be uniformly illuminated. Because the linear irradiance blend is distributed across many pixels in the image, radiometric falloff must be negligible over the motion blurred region. The direction of camera motion need not be exactly perpendicular to the edge between the two colored regions, as long as the edge is straight and the uniform regions are large enough that the blurred region sees only the two intended irradiances. Also, we carefully focus the camera to prevent adding additional blur to the irradiance profile.

We manually specify the motion-blurred region using a very simple user interface shown in figure 4 (c). We automatically locate the monotonically decreasing or increasing ramps of grayscale values in each scanline, and combine data from the specified rows to produce the response function. If we use a symmetric target like that in the figure, we can verify the accuracy of the calibration; if the motion is not linear or the illumination varies too greatly, the response curves calculated from the left and right edges of the black region will not match. On the other hand, if these conditions are satisfied, we expect the two curves to agree. We visually inspect the curves; for a more quantitative method, one could also measure the RMSE between them.

5. Experimental Results

In this section, we evaluate our results against for several cameras. We use the online implementation of Mitsunaga and Nayar's method[11] (which we will call "Mitsunaga-Nayar") as a reliable reference, and also compute our own ground truth data.

Ground Truth Data. We compute ground truth data using multiple exposures of a Macbeth color chart and a technique very similar to Mitsunaga and Nayar's method. For a given color patch, the relative exposure time of each image is the relative irradiance of the patch in that image.
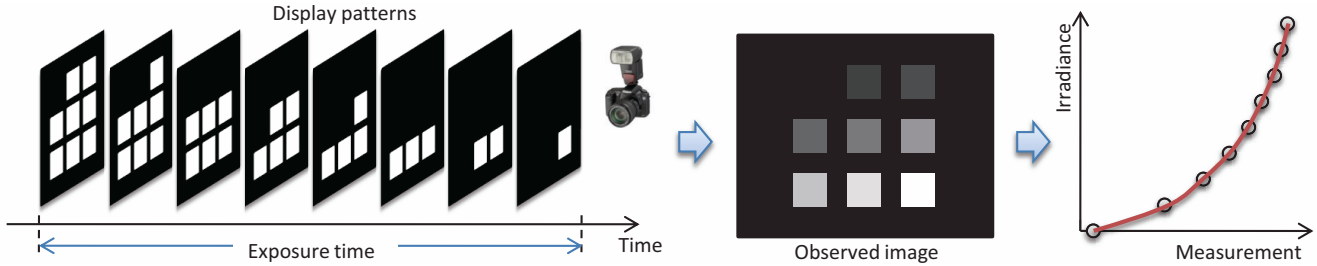
Figure 3. Radiometric calibration by temporally blending two intensities on a computer monitor. In this example, we display a sequence of nine images during a single camera exposure. For each patch, the bright and dark values are shown for different portions of the exposure. This creates a grid of patches in the camera image that have linearly spaced irradiances. Plotting the irradiances versus their measured values in the image produces a sampled version of the inverse response curve.
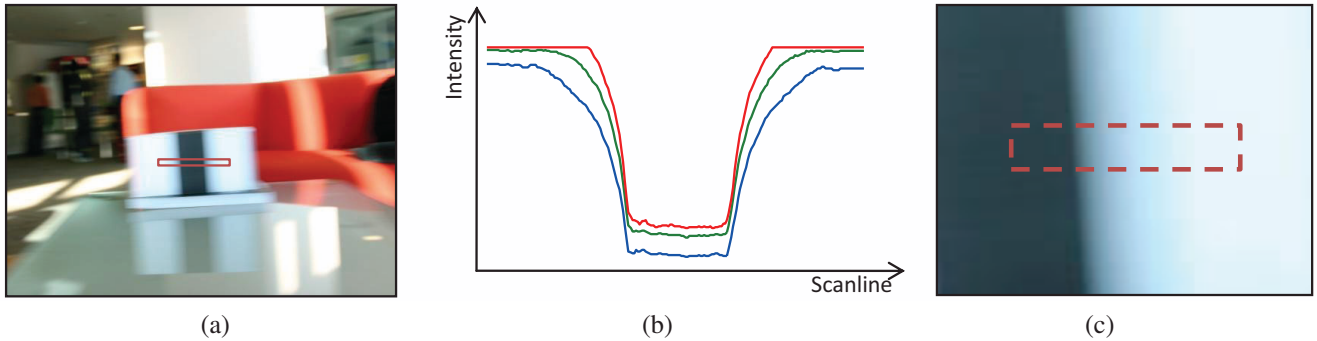

Figure 4. Calibrating from motion blur. (a) An example of motion blur with constant velocity. (b) The pixel intensity for one row of pixels crossing the blurred target. This is a sampled version of the camera response curve. (c) An example of a user interface for specifying the motion blurred region.

This gives us one series of measurements with known irradiances for each patch. We then simultaneously solve for the relative irradiances of the patches (i.e. the irradiance of one patch relative to another patch) and a single polynomial that fits all of the scaled measurements. The relative irradiance is not simply the patch reflectance; other factors such as veiling glare in the camera, brightness offsets, and non-uniform illumination also affect this ratio.

Display-Based Calibration. For the display-based calibration method, we used a DELL 1907FPt monitor and displayed two grid patterns in each image. The low displayed pixel value was always 0, and the high one was 220 (out of 255) for one grid and 100 for the other. The two grids were each three rows of five patches, and displayed one above the other on the screen. We used 60Hz displays and a one second exposure for all cameras, so each pattern in the grid sequence was displayed for four frames. We optimized using the Nelder-Mead simplex method [12]. Figure 5 shows an example image of our pattern, taken with the Canon EOS D20.

Calibration from Motion Blur. To verify the correctness of our motion blur approach, we used a controlled setup to create a very high-contrast edge and guarantee linear mo-
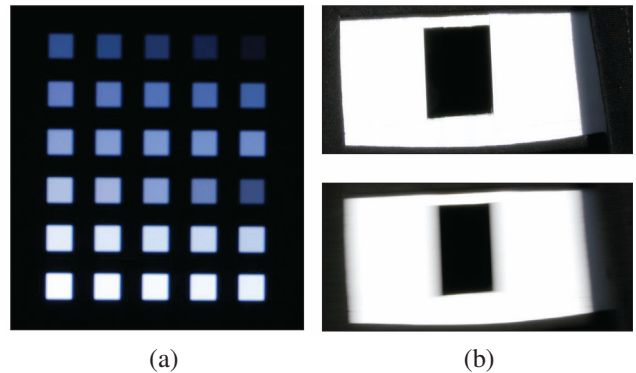

Figure 5. Example inputs for our calibration method. (a) Two 5x3 grids of irradiance blends. (b) A static image of the light trap used for our motion blur experiments (top), and an image taken while the camera was moving (bottom).

tion. We used a light trap as our target. A light trap is a large cavity with black interior walls and a small opening. Light entering the opening is "trapped" inside because it is mostly absorbed by the walls, so the hole emits very little light. For our experiments, the hole is rectangular, and the outside of the trap is white, creating a very high-contrast white-black-white edge, shown in figure 5. We made the light trap ourselves from an ordinary cardboard box. Care

Table 1. Mean RMSE, variance, and disparity of the estimated inverse response functions in the normalized input domain. Ten datasets were used for each method. For Mitsunaga-Nayar, we removed two out of ten estimates that were obvious outliers.

	Display	Motion blur	Mitsunaga & Nayar
RMSE	0.012	0.012	0.006
Variance	0.050	0.052	0.061
Disparity	0.017	0.029	0.017

must be taken to make the edge of the hole straight and thin (so light does not reflect off the sides). We used sheets of white paper to make the border. To ensure linear motion, we mounted each camera on a linear translation stage. The exposure time for the cameras was roughly one second. Still and blurred images of the light trap are shown in figure 5.

Results. Figure 6 shows ground truth and our results for the green channel response of three cameras: a Canon EOS 20D, a Nikon D70, and a PointGrey Research Dragonfly. We chose the Canon and Nikon for their popularity, and the Dragonfly because its response (nearly linear) differs greatly from the other two. The plots were normalized to the domain and range of our measured pixel intensities. A typical range was 15 to 240 (out of 255). Our results agree well with those obtained from the other methods.

Some of these calibration methods might be expected to produce varying results for different input images. To investigate this, we used each method ten times with different inputs. Table 5 shows a comparison of the mean RMSE, variance, and disparity (maximum error) for ten trials of each methods. The RMSE of our methods are roughly twice that of Mitsunaga-Nayar, while our variance is slightly less. For the Mitsunaga-Nayar method, however, we had to remove two curves that were obvious outliers.

Motion Blur by Hand. Our motion blur calibration method is not limited to lab settings. In figure 7, we show two examples of motion-blur based calibration where we manually created the motion. For this experiment, we taped a rectangular piece of black cardboard to the inside of a window. When photographed from the inside, this creates a very high-contrast edge. We then photographed the cardboard while waving the camera sideways by hand. If the exposure time for the camera is short enough, we can expect the motion to be roughly linear. We used a $1/30$ second exposure. Because the target has two symmetric edges, we can verify the results by comparing the response curves computed from the blurred regions on each side of the cardboard.

Figure 7 (a) shows one attempt. Figure 7 (b) shows that the inverse response curves obtained from the left and right sides of the target. These coincide very well, indicating
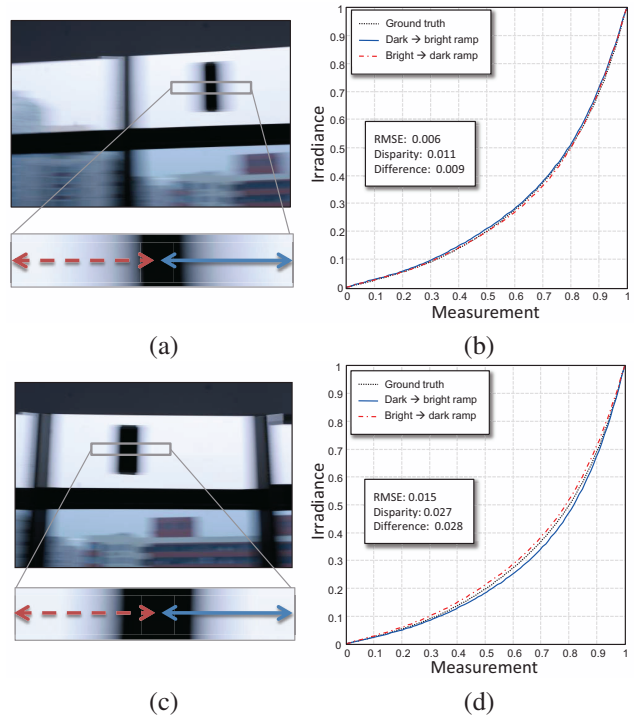

(a)

(b)

(c)

(d)

Figure 7. Calibration from manual motion blur. (a) An image of a piece of cardboard against a window, taken with the Canon EOS 20D while the camera was waved sideways by hand. (b) The inverse response curves computed from the left and right sides of the blur region match each other, showing that this calibration result is good. The difference between the two curves, measured as the RMSE error, is under 1%. Both curves fit the ground truth well. (c) A second motion blur example. (d) This time, the curves computed from both sides do not match each other (an RMSE error of 2.8%), indicating the motion was not linear. These are the best and worst results from six trials. We used a $1/30$ second exposure time for all images.

that the data are good. As we see, the computed curve also matches the ground truth. Figure 7 (c) and (d) show an image and plot from a less successful attempt. The curves do not match each other or the ground truth. These are the best and worst results from six trials.

6. Discussion and Conclusion

We have presented the general idea and two specific methods for using temporal color mixtures to compute camera radiometric response functions. Our display-based method is well-suited for lab environments. For outdoor settings or with cameras whose exposure time cannot be set, calibration from motion blur is recommended. The strength of these techniques lies in their ability to directly sample the response function. We can easily verify that we have covered the full measurement range, and plotting the measured data provides immediate feedback about the shape of the re-
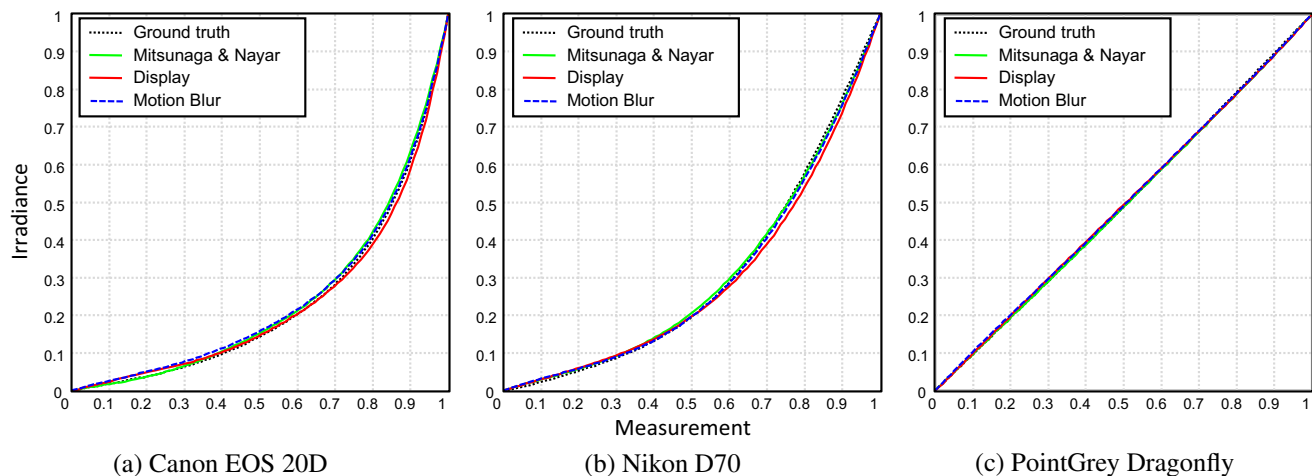
Figure 6. Our calibration results for the green channel of three different cameras, compared to ground truth and Mitsunaga-Nayar. All plots were normalized to the domain and range of our measured pixel intensities. Our results agree well with those obtained from the other methods.

sponse curve and the accuracy of the computed fit. This is not true for methods based on analyzing edges or noise distributions. Moreover, we do not rely on statistical priors [4] over the camera response function. This might allow us to calibrate cameras not represented well by existing priors.

One limitation of our method is the offset ambiguity in computed response curve; in practice, we minimize this by observing irradiances as near to zero as possible. Regardless, the scale and offset ambiguities do not prevent us from linearizing the intensities. We assume control over the exposure time for our display-based calibration method, but one could imagine ways to calibrate without this constraint. As one example, we could take a series of images of a single high-speed LED flashing with varying duty cycles. One interesting feature of our display-based method using multiple irradiance blends is that it computes relative irradiances for different displayed pixel values. This could be extended to simultaneously calibrate the monitor and the camera.

We verified the correctness of the motion blur approach using a controlled setup and further demonstrated the applicability of our method to a hand-held camera in a relaxed environment. Conceivably, our method could even be used to calibrate videos or images for which the camera is no longer available.

References

[1] Y.-C. Chang and J. F. Reid. Rgb calibration for color image analysis in machine vision. *IEEE Transactions on Image Processing*, 5(10):1414–1422, 1996.

[2] P. E. Debevec and J. Malik. Recovering high dynamic range radiance maps from photographs. In *ACM SIGGRAPH*, pages 369–378, 1997.

[3] M. D. Grossberg and S. K. Nayar. Determining the camera response from images: What is knowable? *IEEE Trans.*

on Pattern Analysis and Machine Intelligence, 25(11):1455–1467, 2003.

[4] M. D. Grossberg and S. K. Nayer. What is the space of camera response functions? In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, volume 2, pages 602–609, 2003.

[5] S. J. Kim and M. Pollefeys. Radiometric alignment of image sequences. In *IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, volume 1, pages 645–651, 2004.

[6] S. Lin, J. Gu, S. Yamazaki, and H. Y. Shum. Radiometric calibration from a single image. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, volume 2, pages 938–945, 2004.

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

[8] S. Mann. Comparametric imaging: Estimating both the unknown response and the unknown set of exposures in a plurality of differently exposed images. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, volume 1, pages 842–849, 2001.

[9] S. Mann and R. Picard. Being 'undigital' with digital cameras: extending dynamic range by combining differently exposed pictures. In *IS&T, 48th annual conference*, pages 422–428, 1995.

[10] Y. Matsushita and S. Lin. Radiometric calibration from noise distributions. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2007.

[11] T. Mitsunaga and S. K. Nayar. Radiometric self-calibration. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, volume 2, pages 374–380, 1999.

[12] J. Nelder and R. Mead. A simplex method for function minimization. *Computer Journal*, 7:308–313, 1965.

[13] Y. Tsin, V. Ramesh, and T. Kanade. Statistical calibration of ccd imaging process. In *International Conference of Computer Vision (ICCV)*, volume 1, pages 480–487, 2001.