Radiometric Calibration from a Single Image

Stephen Lin[†] Jinwei Gu[‡] Shuntaro Yamazaki[§] Heung-Yeung Shum[†]

[†]Microsoft Research Asia [‡]Tsinghua University* [§]University of Tokyo*

Abstract

Photometric methods in computer vision require calibration of the camera's radiometric response, and previous works have addressed this problem using multiple registered images captured under different camera exposure settings. In many instances, such an image set is not available, so we propose a method that performs radiometric calibration from only a single image, based on measured RGB distributions at color edges. This technique automatically selects appropriate edge information for processing, and employs a Bayesian approach to compute the calibration. Extensive experimentation has shown that accurate calibration results can be obtained using only a single input image.

1 Introduction

Many computer vision algorithms implicitly assume that image intensities are linearly related to scene radiance. This linear relationship is needed for brightness values to have some physical meaning and for images from different cameras to be accurately compared. Most cameras, though, are designed to have a non-linear mapping from scene radiance to image intensity, to emulate the characteristics of film and to account for non-linearities in display systems. To obtain scene radiance information from images, it therefore becomes necessary to transform the non-linear mapping into a linear one by calibrating the radiometric response of the camera system.

Several previous works have addressed this problem, most of which require as input a set of registered images taken with varying camera exposures. For known exposure ratios among the images, methods have been presented to solve for parametric [6] and smooth non-parametric [1] inverse response functions. Starting with a rough estimate of the exposure ratio, Mitsunaga and Nayar [7] iteratively compute a polynomial inverse response function and refine estimates of exposure ratios. Iterative methods have also been proposed by Tsin et al. [9] and Mann [5] for estimating non-parametric inverse response functions. Grossberg and Nayar [4] presented a camera response model derived from PCA of a camera response data set, and in [3] they avoid the need for spatial correspondences among images by employing a histogram analysis to relate intensity values between two images of different exposure. Instead of varying the exposures among different images, Nayar and Mitsunaga [8] obtain various exposures at the expense of spatial resolution by using an image filter with spatially varying transmittance.

Accurate calibration results can be obtained using these prior techniques; however, they are based on certain limiting requirements, such as having multiple images under different exposures or using a special filter. Although this calibration data could be captured if the camera were in hand, often images to be processed by vision algorithms have been captured by unknown cameras or by cameras that are not readily available. Such scenarios include applications for personal photographs, and analysis of images on the web. To recover the scene radiances in such images, a method is needed for performing radiometric calibration from only a single input image taken with unknown camera settings. Though a technique for gamma correction of a single input image has previously been proposed [2], the response function of a camera can differ significantly from a gamma curve.

A mapping from image intensity to scene radiance can be determined by fitting a function to corresponding values of intensity and radiance. For intensity values in measured images, however, their corresponding radiance values are generally unknown. Previous methods all deal with this problem by utilizing different camera exposures which modulate the scene radiance captured at the CCD array. For a pair of images, the ratio of captured radiance at corresponding pixels is equal to the exposure ratio between the two images.

*This work was done while Jinwei Gu and Shuntaro Yamazaki were interns at Microsoft Research Asia.

Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'04)
1063-6919/04 $20.00 © 2004 **IEEE**

With a set of corresponding points in the image pair, a camera response function can be fit to their intensity values and radiance ratios. Since the function is fit to radiance ratios instead of absolute radiance values, it maps image intensities to values linearly related to scene radiance.

In the case of a single input image, correspondences and variable exposures are not available to determine ratios of captured radiance. Furthermore, radiance ratios among pixels in a single image cannot be established reliably because variations in radiance result from a number of scene factors such as geometry, lighting environment and bi-directional reflectance functions.

To address the problem of single-image calibration, we present an approach fundamentally different from previous methods in that no radiance information needs to be recovered. Instead, our method obtains calibration information from how a non-linear radiometric response affects the measured image, in particular its edge color distributions. In a local edge region, the captured radiance colors form a linear distribution in RGB space because of linear color blending at edge pixels. The edge colors in the measured image, however, form a non-linear distribution because of the non-linear mapping from captured radiance to intensity in camera response functions. We show in this paper that the inverse response function that maps image intensity to captured radiance can be estimated as the function that maps the non-linear color distributions of edge regions into linear distributions. With this approach, multiple images, correspondences and variable exposures are not needed. In our extensive experimentation, we have found that this calibration technique produces accurate results using only a single input image.

2 Background

Before describing our algorithm, we review some background on radiometric calibration. The radiometric response function f of a camera system relates captured scene radiance I, also known as image irradiance, to its measured intensity M in the image:

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

Since photometric methods should operate on image irradiance values rather than image intensity measurements, radiometric calibration methods solve directly for the inverse response function $g = f^{-1}$. Response functions f are invertible since sensor output increases monotonically with respect to I.

Computation of the inverse response function in previous works has been based on the relationship

$$g(m_A) = kg(m_B) \tag{2}$$

where m_A denotes measured image intensities in image A, m_B represents intensities of corresponding points in image B, and k denotes the exposure ratio between A and B. Evident in this equation is the need to capture multiple registered images with different exposure settings. In [3], registration is circumvented by forming correspondences through histogram equalization, assuming that the scene radiance distribution has not changed significantly from one image to the next. Prior methods that do not assume known exposure ratios iteratively solve for k and g in their calibration processes.

A major obstacle in computing inverse response functions arises from exponential ambiguity, or u-ambiguity. From Eq. (2), it can be seen that if g and k are solutions for a set of images, then g^u and k^u can be valid solutions as well:

$$g^u(m_A) = k^u g^u(m_B).$$

To deal with this ambiguity, prior methods require a rough initial estimate of k and assumptions on the structure of the radiometric model, as detailed in [3].

It is generally assumed that the sensor response does not change over the image grid, such as from vignetting or fixed pattern noise that results from CCD manufacturing variances. The response functions of the RGB channels, though, can differ from one another.

3 Edge Color Distributions

For single image input, the relationship of Eq. (2) cannot be used for calibration. Our method instead is based on the relationship between the inverse response function and the edge color distributions in a measured image. In this section, we describe the effects of sensor nonlinearities on edge color distributions and how these distributions provide information for radiometric calibration.

3.1 Color Sensing

Images are formed on a CCD sensor array that records radiance from the scene. Because the array is limited in resolution, each array element $\boldsymbol{x}$ images a solid angle of the scene, and we denote the set of image plane points within an array element as $S(\boldsymbol{x})$.

For color imaging, each array element is coupled with a color filter k, typically red, green or blue. The

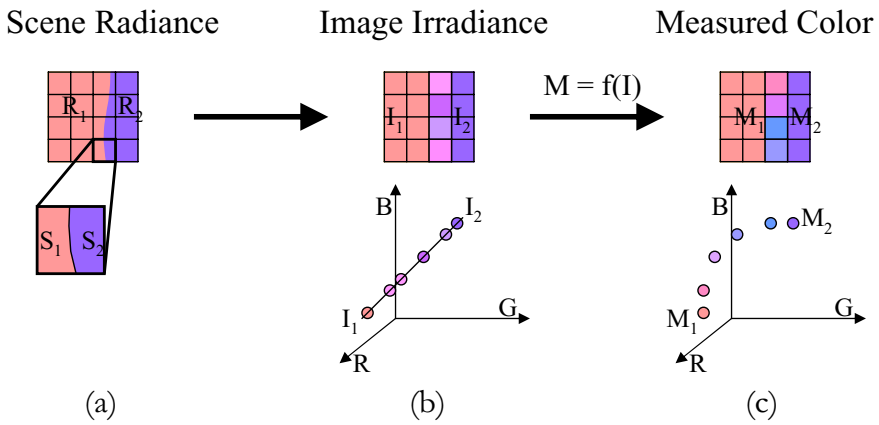

Figure 1. Non-linear distribution of measured colors in edge regions. (a) The first two columns of pixels image only a color region with radiance R_1, and the last column images a color region with radiance R_2. The third column contains pixels that image both regions. (b) The irradiances of pixels in the first two columns map to the same point I_1 in RGB color space, while the irradiances in the last column map to a single color point I_2. The colors of the blended pixels lie on a line defined by I_1 and I_2. (c) A non-linear camera response f warps the image irradiance colors into a non-linear distribution. (Figures best viewed in color)

image irradiance I for color k at $\boldsymbol{x}$ depends on the sensitivity q_k of the element-filter pair and the incoming scene radiances R incident upon image plane points p in $S(\boldsymbol{x})$:

$$I(\boldsymbol{x},k) = \int_{\lambda_k}\int_{p\in S(x)} R(p,\lambda)q_k(\lambda)\; dp\; d\lambda \qquad (3)$$

where λ_k denotes the range of transmitted light wavelengths by color filter k. Although just a single color filter is paired with each array element, it is typically assumed that all three R, G, B colors are accurately measured with q_R, q_G, q_B filter sensitivities at each pixel, because of effective color value interpolation, or *demosaicing*, in the camera.

3.2 Nonlinearity of Edge Colors

In analyzing an edge color distribution, we consider an image patch P that contains two regions each having distinct but uniform colors, as illustrated in Fig. 1(a). Because of the limited spatial resolution of the image array, the image plane area $S(\boldsymbol{x})$ of an edge pixel will generally image portions of both regions. For an edge pixel $\boldsymbol{x}$ divided into regions $S_1(\boldsymbol{x})$ and $S_2(\boldsymbol{x})$ with respective scene radiances $R_1(\lambda)$ and $R_2(\lambda)$, the overall radiance incident at $\boldsymbol{x}$ can be expressed as

$$\int_{p\in S(x)} R(p,\lambda)dp = \int_{p\in S_1(x)} R_1(\lambda)dp + \int_{p\in S_2(x)} R_2(\lambda)dp$$

$$= \alpha R_1(\lambda) + (1-\alpha)R_2(\lambda) \qquad (4)$$

where $\alpha = \int_{p\in S_1(x)} dp$ and $S(\boldsymbol{x})$ is of unit area.

Substituting (4) and (3) into (1) gives the measured color of $\boldsymbol{x}$ as

$$m(\boldsymbol{x},k) = f\left[\alpha\int_{\lambda_k} R_1(\lambda)q_k(\lambda)d\lambda + (1-\alpha)\int_{\lambda_k} R_2(\lambda)q_k(\lambda)d\lambda\right]$$

$$= f\left[\alpha I_1(\boldsymbol{x},k) + (1-\alpha)I_2(\boldsymbol{x},k)\right] \qquad (5)$$

If there were a linear relationship between image irradiance I and measured color $f(I)$, then the following property would hold:

$$f\left[\alpha I_1 + (1-\alpha)I_2\right] = \alpha f(I_1) + (1-\alpha)f(I_2)$$

meaning that the measured colors of the blended pixels lie on a line in RGB color space between the measured region colors $f(I_1)$ and $f(I_2)$. Since f is typically non-linear, a plot of the measured edge colors forms a curve rather than a line, as illustrated in Fig. 1(c). The non-linearity of a measured color edge distribution provides information for computing the inverse camera response.

3.3 Transformation to Linear Distributions

The inverse response function should transform measured image intensity into values linearly related to scene radiance, and therefore linearly related to image irradiance. Since image irradiance at edge regions form linear distributions as exemplified in Fig. 1(b), the inverse response function should transform non-linear edge color distributions into linear distributions. For an edge patch with measured region colors M_1 and M_2, the inverse response function g should map the measured color M_p of each point p in the patch to a line

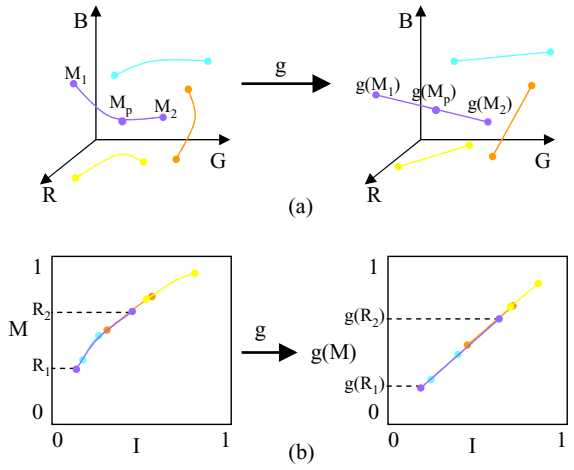

Figure 2. Transformation of non-linear distributions to linear distributions by the inverse response function g. (a) Color space, before and after transformation. (b) Plot of measured red color vs. irradiance, before and after transformation

defined by $g(M_1)$ and $g(M_2)$, as illustrated in Fig. 2(a). A function g satisfies this property if the distance from $g(M_p)$ to line $\overline{g(M_1)g(M_2)}$ is zero, formulated as

$$\frac{|[g(M_1) - g(M_2)] \times [g(M_p) - g(M_2)]|}{|g(M_1) - g(M_2)|} = 0 \qquad (6)$$

where $\times$ is the cross product operation between two vectors.

An inverse function g computed with respect to a single edge patch with region colors $M_1 = (R_1, G_1, B_1)$ and $M_2 = (R_2, G_2, B_2)$ will transform measured red colors in the range delimited by R_1 and R_2, measured green colors in the range delimited by G_1 and G_2, and measured blue colors in the range delimited by B_1 and B_2 to values linearly related to irradiance. To more fully cover the ranges of measured R, G and B colors, g is computed with respect to all obtained edge patches in the image.

For a given image, we collect all obtained edge color triples into an observation set $\Omega = \{< M_1, M_2, M_p >\}$, and define the total distance as

$$D(g;\Omega) = \sum_{\Omega} \frac{|[g(M_1) - g(M_2)] \times [g(M_1) - g(M_p)]|}{|g(M_1) - g(M_2)|}. \qquad (7)$$

The desired inverse response function g is the one that gives the smallest total distance.

The linear relationship between estimated irradiance given by g and absolute irradiance will have unknown scale factors $\gamma = [\gamma_R, \gamma_G, \gamma_B]^T$ for each color channel, such that $g(M) = \gamma I$ by element-wise multiplication.

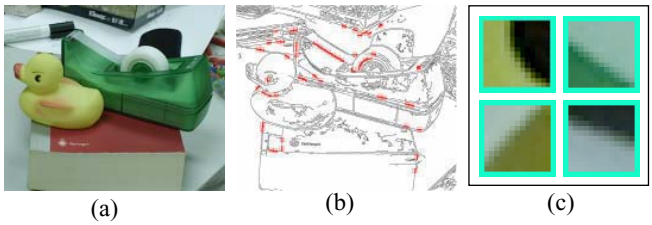

Figure 3. Edge selection for observation set. (a) Input image; (b) Automatically identified patches on the edge image; (c) Closeups of selected patches

To deal with difference in scalings, we normalize both the domain and co-domain of the function g so that it satisfies $g(0) = 0$ and $g(1) = 1$, as done in [4].

4 Formation of Observation Set

The set of edge color triples Ω is collected from fixed-size (15×15) image windows containing monotonic color edges between uniform region colors. Our method selects non-overlapping candidate windows centered on a Canny-detected edge whose path divides the window into exactly two regions.

To determine whether a window contains valid edge colors for inclusion into the observation set, some color analysis is performed. The edge path is dilated by three pixels, and with the two partitioned non-edge regions S_1 and S_2, the mean color and the color variance with respect to Euclidean RGB distance is computed. If both regions have color variances below a specified threshold, then the regions are considered uniform. The mean colors must lie at least a specified distance from each other, since image noise could otherwise dominate the distance computation of Eq. (6). Additionally, because of the monotonicity of response functions, the R, G and B edge colors must lie within the range delimited by the two region colors. This requirement excludes edges that exhibit ringing. For each valid edge window, the color triple for each edge pixel is added to the observation set.

Fig. 3 exemplifies the selection of valid edge windows for a given image. The color variance threshold is fixed at a low value in our implementation to ensure the quality of the edge triples.

5 Bayesian Estimation

While the inverse response function could be estimated by minimizing Eq. (7) under the physical constraints that the function be smooth and monotonic, the function could be more accurately determined by

utilizing prior information on real-world camera responses, presented in [4]. This prior data helps to interpolate and extrapolate the inverse response curve over intervals of incomplete color data, and its PCA representation facilitates computation by providing a concise descriptor for g. Using the PCA model of camera responses presented in [4], we represent the inverse response function g by

$$g = g_0 + c\,H \tag{8}$$

where $g_0 = [g_{R0}, g_{G0}, g_{B0}]^T$ is the mean inverse response, and H is a matrix whose columns are composed of the first $N = 5$ eigenvectors. $c = [c_R, c_G, c_B]^T$ is a coefficient vector in $R^{3\times N}$ that represents an inverse response function $g = [g_R, g_G, g_B]^T$. With this prior information, our technique computes a MAP solution of the inverse response function.

5.1 Prior Model

We model the prior $p(g)$ of the inverse response function from the DoRF database compiled by Grossberg and Nayar [4], which contains 201 inverse response functions from various digital cameras and films. From this set of inverse responses, we form the prior as a Gaussian mixture model:

$$p(g) = \sum_{i=1}^{K} \alpha_i \mathcal{N}(g; \mu_i, \Sigma_i). \tag{9}$$

In our implementation, we empirically use five kernels ($K = 5$) obtained using the EM algorithm.

5.2 Likelihood Function

The inverse response function g should yield a low total distance as expressed in Eq. (7), so we model the likelihood $p(\Omega|g)$ by incorporating this distance measure into an exponential distribution:

$$p(\Omega|g) = \frac{1}{Z} exp(-\lambda D(g;\Omega)) \tag{10}$$

where λ is set empirically to 10^4 and Z is a normalization constant.

5.3 Solution Method

After modelling the prior $p(g)$ and the likelihood $p(\Omega|g)$ by Eq. (9) and Eq. (10) respectively, we can formulate the whole problem in a Bayesian framework. Given an input image M, we determine the observation set Ω. The optimal response function g^* is then defined as

$$g^* = \arg\max p(g|\Omega) = \arg\max p(\Omega|g)p(g),$$

which is the MAP solution of the problem. Taking the log of the above equation, g^* also can be written as

$$g^* = \arg\min E(g) = \arg\min \lambda D(g;\Omega) - \log p(g), \tag{11}$$

where g^* can be viewed as the optimal solution of the objective function $E(g)$.

The optimization is computed by the Levenberg-Marquardt method, with the coefficients of g initialized to zero. Since g is represented by principal components in Eq. (8), the first and second derivatives of $g(c)$ are approximated by the first and second differences with a small δc. After the optimization algorithm converges, the result is refined sequentially in each dimension using a greedy local search.

6 Results

We applied our calibration technique on single-image input captured from three kinds of cameras: a Kodak DCS 330, a Canon PowerShot G5 and a Nikon D100. These cameras were not included in the response database used to form the prior model. Examples of calibration results for the respective cameras are shown in Fig. 5. Due to space limitations, only two images and one color response function is shown for each camera. For comparison, we also recovered the functions using the methods of Debevec and Malik [1], Mitsunaga and Nayar [7], and by imaging a Macbeth ColorChecker which consists of measured reflectance patches. The recovered functions of our method are reasonably close to the others, even though our response is determined from a single image while the methods of [1] and [7] utilized a set of 12 registered images captured with different exposure settings.

Table 1 summarizes the performance of our method on a set of 50 images per camera in terms of RMS error and disparity with respect to the average inverse response curve among [1], [7] and the ColorChecker. In [4], RMS error and disparity were computed for their calibration system, and our method yields roughly similar performance.

For a couple images, the measured R, G and B ranges for each edge patch are shown in Fig. 4. When the measured color data is incomplete, there is a question of whether the correct inverse response can be solved, and whether all the color values in the image will map to a consistent linear relationship with irradiance. Even when the observed edge colors are relatively limited as in Fig. 4(e-h), accurate calibration

Table 1. Average RMSE and Disparity of Inverse Response Curves

	Red		Green		Blue	
	RMSE	Disparity	RMSE	Disparity	RMSE	Disparity
KODAK DCS330	1.95E-02	4.41E-02	1.02E-02	3.28E-02	2.51E-02	4.73E-02
CANON G5	1.56E-02	2.83E-02	5.39E-03	9.60E-03	1.33E-02	2.46E-02
NIKON D100	2.17E-02	3.84E-02	1.04E-02	3.20E-02	2.91E-02	4.15E-02

Figure 6. Images of the same scene taken by the same camera with different exposure, of ratio 0.5. Image matching with radiometric calibration leads to smaller match error than without calibration.

can be determined primarily due to constraints provided by the prior model of camera responses compiled by Grossberg and Nayar [4].

Radiometric calibration is important not only for photometric methods, but also for many techniques that involve more than one camera. For example, two different uncalibrated cameras capturing the same scene will have images with color differences, because of their different non-linear sensor responses. The match quality between the two images would be lower than if the cameras were calibrated. To illustrate this problem, we performed a similar experiment, where we captured two images of the same scene using the same camera but at different exposures.

Fig. 6 displays two images taken with a Kodak DCS 330, where the exposure ratio is 0.5. We use only the first image to compute the inverse response function. For intensity normalization, we simply compute the average $R+G+B$ intensity in each image, and use their ratio to scale the second image. All color values are then normalized to the range $[0, 1]$. Without calibration, the mean-squared RGB matching error between the two images is 0.0830. With calibration, the error drops to 0.0121. It can be seen from this example that when one camera is used for modeling and different one is used for testing in applications such as recognition, radiometric calibration can lead to greater accuracy.

7 Discussion

Our calibration method achieves higher accuracy when the measured edge colors cover a broader range of brightness values in each color channel. Although the dependence on edges may limit the amount of data obtainable in an image, we have found that blended colors within a given edge patch often encompass broad intervals of R, G and B, as exhibited in Fig. 4. By examination of an image's edge color intervals, the need for additional data can be inferred.

For a single image, at least a couple methods could potentially be used to obtain additional color information. One is to more aggressively search for edge data. Our current method for edge patch selection is quite conservative. If augmentation of the RGB ranges is needed, variation in window sizes and incrementally less restrictive thresholds could be employed to progressively acquire more edge windows. Another source of information arises from the Bayer pattern, which is the arrangement of color filters on the CCD array. Since an array element can be paired with only a single color filter, the two other colors that it does not measure must be demosaicked from the neighboring pixels. Recovery of the demosaicing method would uncover additional color relationships among pixels throughout the image which could be used for radiometric calibration. We plan to examine both of these directions in future work.

The amount of edge color data can also be increased by using additional input images captured by the same camera, even if the images are not registered or otherwise corresponded, because edge windows from the images can in principle be used together for calibration. This property is especially useful for applications on personal photographs, which generally consist of many unregistered images taken from the same camera.

For an image that exhibits a rather limited range of colors, the estimated inverse response function may not be very accurate, but it will nevertheless transform the measured colors to be linearly related to irradiance. Although colors outside this range may not transform to this linear relationship, the inaccurate inverse response function can still be useful, because a linear relationship for the observed RGB colors is all that vision algorithms generally need.

Besides coverage of two region colors by a pixel, there exist other possible causes of color blending at

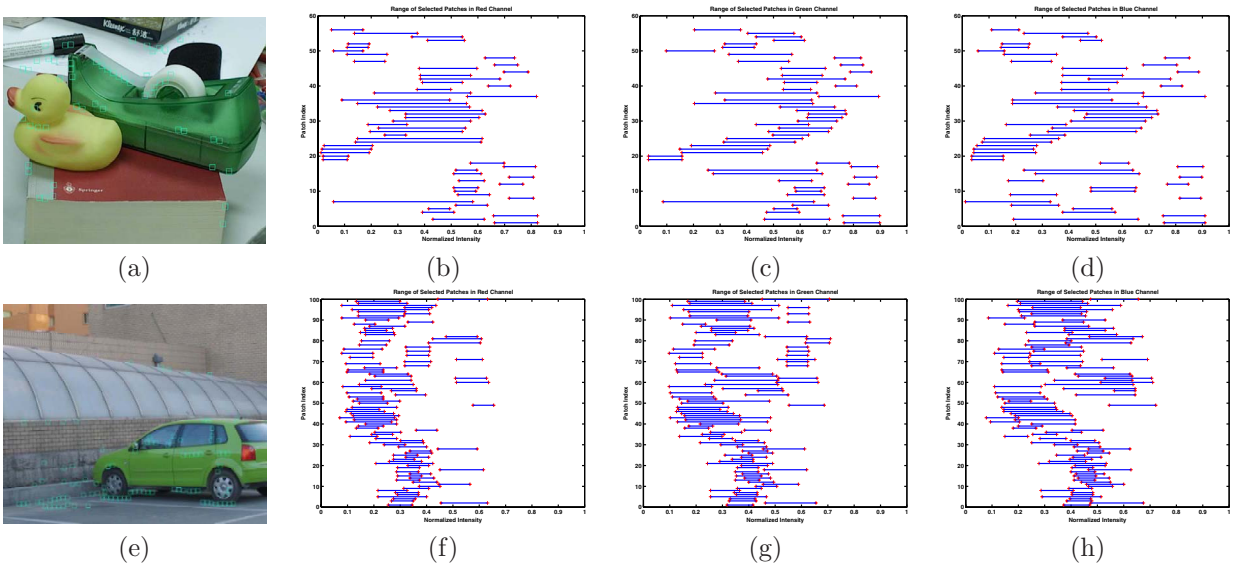

Figure 4. Color range of edge patches in images. (a,e) Input images. (b-d) R, G and B ranges of edge patches in (a). (f-h) R, G and B ranges of edge patches in (e). The vertical axis indexes individual patches in an image.

edges. Defocus and other convolution-type operators convolve each of the R, G and B channels in the same way, resulting in a linear distribution of image irradiance colors that can be exploited by our technique. In contrast, non-linearities of image irradiance colors at edges can result from image artifacts such as chromatic aberration, which can produce fringes of purple at some edges. Local windows with such artifacts, however, are generally rejected for processing by our algorithm, based on the edge patch selection criteria described in Section 4.

Previous calibration methods are faced with the problem of exponential ambiguity in the inverse response function. For a related ambiguity to arise in our approach, multiple inverse response functions would have to give a same minimum energy $E(g)$ in Eq. (11). Given the form of the prior and likelihood function, this occurs only in limited instances such as when both the measured color distribution lies along the $R = G = B$ line and the RGB sensor responses are identical. Since this ambiguity is uncommon and is very unlikely to exist among all the edge windows of an image, it is not a major consideration in this work.

Our proposed method demonstrates the effectiveness of color edge analysis in radiometric calibration, allowing calibration using only a single image or a set of unregistered images. This technique exhibits good performance on a wide range of single-image input, and promising directions exist for increasing the amount of collected color data. Without need for scene radiance or camera information, this single-image approach widens the applicability of radiometric calibration.

References

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

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

[3] M. D. Grossberg and S. K. Nayar. What can be known about the radiometric response from images? In *Proc. Euro. Conf. on Comp. Vision*, pages IV:189–205, 2002.

[4] M. D. Grossberg and S. K. Nayar. What is the space of camera response functions? In *Proc. IEEE Comp. Vision and Pattern Recog.*, pages II:602–609, 2003.

[5] S. Mann. Comparametric imaging: Estimating both the unknown response and the unknown set of exposures in a plurality of differently exposed images. In *Proc. IEEE Comp. Vision and Pattern Recog.*, pages I:842–849, 2001.

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

[7] T. Mitsunaga and S. K. Nayar. Radiometric self calibration. In *Proc. IEEE Comp. Vision and Pattern Recog.*, pages II:374–380, 1999.

[8] S. K. Nayar and T. Mitsunaga. High dynamic range imaging: Spatially varying pixel exposures. In *Proc. IEEE Comp. Vision and Pattern Recog.*, pages I:472–479, 2000.

[9] Y. Tsin, V. Ramesh, and T. Kanade. Statistical calibration of ccd imaging process. In *Proc. Int. Conf. on Comp. Vision*, pages I:480–487, 2001.

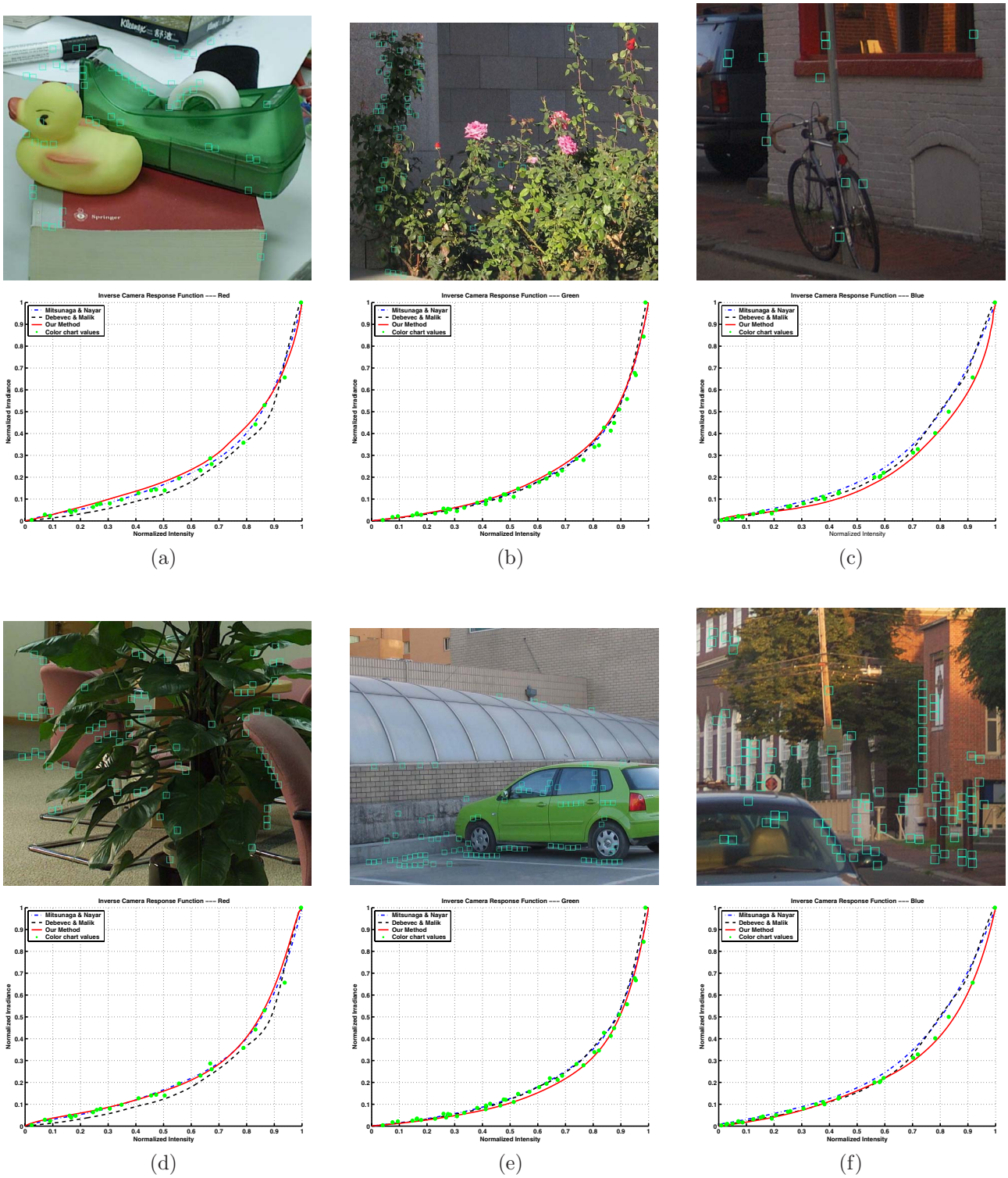

Figure 5. Inverse camera response functions for (a,d) Kodak DCS 330 in the red channel, (b,e) Canon G5 in the green channel, (c,f) Nikon D100 in the blue channel.