

# Estimating Camera Response Functions using Probabilistic Intensity Similarity

Jun Takamatsu

Microsoft Institute for Japanese Academic Research Collaboration (MS-IJARC)

Suginami-ku, Tokyo, Japan

j-taka@cvl.iis.u-tokyo.ac.jp

Yasuyuki Matsushita

Microsoft Research Asia

Beijing, P.R. China, 100080

yasumat@microsoft.com

Katsushi Ikeuchi

Microsoft Institute for Japanese Academic Research Collaboration (MS-IJARC)

Suginami-ku, Tokyo, Japan

ki@cvl.iis.u-tokyo.ac.jp

## Abstract

We propose a method for estimating camera response functions using a probabilistic intensity similarity measure. The similarity measure represents the likelihood of two intensity observations corresponding to the same scene radiance in the presence of noise. We show that the response function and the intensity similarity measure are strongly related. Our method requires several input images of a static scene taken from the same viewing position with fixed camera parameters. Noise causes pixel values at the same pixel coordinate to vary in these images, even though they measure the same scene radiance. We use these fluctuations to estimate the response function by maximizing the intensity similarity function for all pixels. Unlike prior noise-based estimation methods, our method requires only a small number of images, so it works with digital cameras as well as video cameras. Moreover, our method does not rely on any special image processing or statistical prior models. Real-world experiments using different cameras demonstrate the effectiveness of the technique.

## 1. Introduction

Many computer vision algorithms rely on the assumption that image intensity is linearly related to scene radiance recorded at the camera sensor. However, this assumption does not hold with most cameras; the linearity is not maintained in the observation domain due to non-linear *camera response functions*. Linearization of measured image intensity is important for many vision algorithms to work. To linearize measured image intensities, we must estimate the camera response function.

Camera response functions describe the relationship between the measured image intensity  $O$  and irradiance at a

camera  $I$  as  $O = f(I)$ . Assuming they are continuous and monotonic, response functions can be inverted to obtain the *inverse response function*  $g (= f^{-1})$ . Measured image intensities can be linearized using  $I = g(O)$ . Usually only observed intensities  $O$  are available, so most estimation methods attempt to estimate the inverse response function  $g$ .

Our method uses a small number of images of a static scene captured from the same viewpoint with fixed camera parameters. Due to noise, these images exhibit slight differences in intensity. Our method uses these intensity fluctuations to estimate camera response functions. More specifically, we use the *probabilistic intensity similarity* [13], which represents the likelihood that two intensity observations originated from the same scene radiance. The similarity varies with the underlying noise distributions and the shape of the response functions. Our method uses this relationship between the intensity similarity and the response function to estimate response functions. As illustrated in Figure 1, our method uses a physics-based noise model [6] to define a model-based intensity similarity. Taking a few images captured from the same viewpoint with fixed camera parameters, our method estimates the response function by maximizing the similarity among the input images. The method outputs the estimate of the response function as well as the data-specific intensity similarity measure. The contributions of this paper are as follows:

- We show a theoretical relationship between intensity similarity and response functions.
- We develop a practical noise-based method that only requires a minimum of two images. We also extend the method to handle a single-image case.

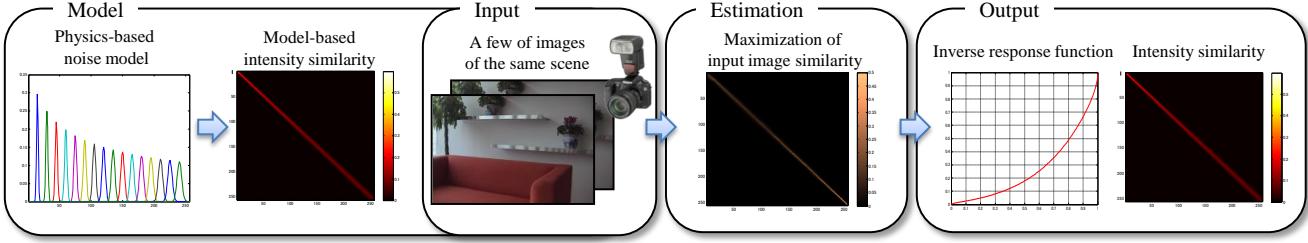


Figure 1. Overview of the proposed method. Our method uses a probabilistic intensity similarity based on a physics-based image noise model. We estimate camera response functions by maximizing the similarity measure of images of the same scene computed from the model-based intensity similarity. We produce estimates of the inverse response functions and data-specific intensity similarity measures.

Throughout this paper, we use the terms *input signal* and *output signal* of a response function. The input signal originates from the scene radiance and is considered the input to the response function. This signal is converted into a digital output signal by a series of circuits including a photodetector, amplifiers, and possibly signal processing blocks. This output signal is the measured pixel intensity in images. We normalize the input and output ranges of the response functions.

The rest of this paper is organized as follows: after reviewing prior work in Section 2, we describe the theoretical background and our estimation method in Section 3. In Section 4, we describe implementation details. In Section 5 we show experimental results to verify the effectiveness of the method and present some of its applications. Section 6 concludes the paper.

## 2. Prior work

One class of methods for estimating response functions uses knowledge of irradiance ratios for measured intensities. The Macbeth color checker is used for estimating the response functions by several research groups such as [2]. Nayar and Mitsunaga [16] use an optical filter with spatially varying transmittance; the variation corresponds to the irradiance ratio. To avoid using such special equipment, some methods use a set of images of a static scene from a fixed viewpoint, taken with different exposure times, so the irradiance ratio is known. These methods are classified into parametric and non-parametric approaches.

In the parametric framework, Mann and Picard [12] propose a method that assumes the response functions can be approximated by gamma correction functions. Mitsunaga and Nayar use a polynomial function for the representation [15]. Grossberg and Nayar apply principal component analysis (PCA) to a database of real-world response functions (DoRF) and show that the space of response functions can be represented by a small number of basis functions [5]. Mann [11] uses a comparometric equation that relates two input images taken with different exposure times.

Debevec and Malik estimate response functions with a non-parametric representation using a smoothness con-

straint [3]. Tsin *et al.* robustly estimate non-parametric forms of response functions using the statistics of image noise [19]. Pal *et al.* use a Bayesian network consisting of probabilistic imaging models and prior models of response functions [18]. While non-parametric approaches have greater descriptive power, the large number of unknowns often lead to computationally expensive or unstable algorithms.

A few authors have proposed more general estimation methods that allow camera motion or scene motion between input images. Grossberg and Nayar [4] propose a method that allows a small amount of motion in the scene. Their use of the brightness histograms efficiently avoids the problem of pixel-wise image registration. Kim and Pollefeys [7] propose a method that allows free camera motion by finding the correspondences between images. While these methods can estimate the function from less restricted setups, they still require images taken with multiple different exposures.

Another class of the estimation methods is based on the physics of the imaging process. Lin *et al.* [8] estimate the response function by analyzing the linear blending of colors along edges in images. For grayscale images, a 1D analogue of the 3D color method is presented in [9]. While the use of spatial color mixtures has theoretical importance, these authors note that image noise can be a major problem. Matsushita and Lin later proposed a noise-based method using the assumption that profiles of input noise distributions are symmetric [14]. Although the method is robust against noise, it generally requires a large number of images of a static scene to collect noise distributions. If the scene contains many uniformly colored, uniformly illuminated, flat surfaces, then one image can suffice. This limits its applicability only to videos in practice. Our method is a noise-based method as well, but in contrast to the prior work it only requires a small number of images of a natural scene.

## 3. Method

In this section we first provide the theoretical relationship between the response function and probabilistic intensity similarity. We then describe the algorithm for estimating response functions.

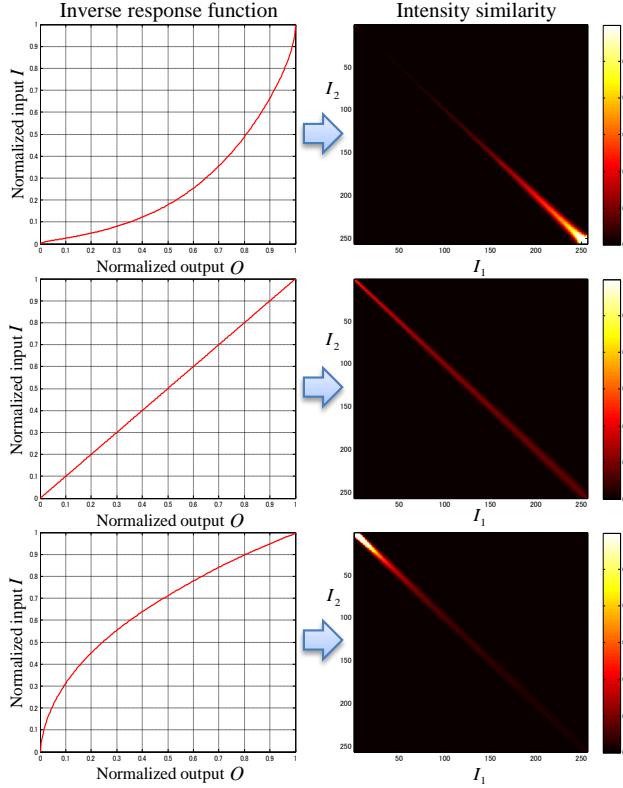


Figure 2. Relationship between inverse response function and probabilistic intensity similarity. The graphs in the left column show inverse response function shapes. The plots in the right column show the intensity similarity measures for any two pixel values. The similarity measure is color-coded as indicated by the bar on the right. These plots show how the intensity similarity varies with the inverse response function.

### 3.1. Response functions and intensity similarity

Our method uses the probabilistic intensity similarity proposed by Matsushita and Lin [13]. The intensity similarity represents the likelihood of two observations deriving from the same scene radiance. As pointed out in their work as well as by Liu *et al.* [10], probability density functions of image noise distributions vary with the scene radiance. Nonlinear camera response functions also alter the intensity similarity function. In fact response functions and the intensity similarity measure are closely related. Figure 2 shows how the intensity similarity varies with the shape of the response function.

The intensity similarity  $S_O(O_1, O_2)$  between two measured intensities  $O_1$  and  $O_2$  is defined in [13] as

$$S_O(O_1, O_2) = \int p(O_1|\tilde{O})p(O_2|\tilde{O})p(\tilde{O})d\tilde{O}, \quad (1)$$

where  $\tilde{O}$  represents the noise-free output intensity, the probability density function (pdf)  $p(\tilde{O})$  is the prior distribution of having a noise-free output intensity  $\tilde{O}$  in the scene, and the conditional density function (cdf)  $p(O|\tilde{O})$  describes the

probability of observing the intensity  $O$  when the noise-free intensity is  $\tilde{O}$ . The similarity in the input domain  $S_I(I_1, I_2)$  is defined in the same manner. Now we derive the relationship between the intensity similarity in the input and output domains. For the derivation, we use the transformation of pdf's from output to input domain. For all outputs  $O = f(I)$ , the pdf in the output domain  $p(O)$  and its transformation  $p(I)$  in the input domain satisfy the following relationship:

$$p(I)dI = p(O)dO.$$

Using  $dO = f'(I)dI$ , we obtain

$$p(I) = f'(I)p(O). \quad (2)$$

The cdf  $p(I|\tilde{I})$  in the input domain can be transformed in a similar manner. Using Equation (2), Equation (1) can be written as:

$$\begin{aligned} S_O(O_1, O_2) &= \int p(O_1|\tilde{O})p(O_2|\tilde{O})p(\tilde{O})d\tilde{O} \\ &= \int \frac{p(I_1|\tilde{I})}{f'(I_1)} \frac{p(I_2|\tilde{I})}{f'(I_2)} \frac{p(\tilde{I})}{f'(\tilde{I})} f'(\tilde{I})d\tilde{I} \\ &= \frac{1}{f'(I_1)f'(I_2)} \int p(I_1|\tilde{I})p(I_2|\tilde{I})p(\tilde{I})d\tilde{I} \\ &= \frac{1}{f'(I_1)f'(I_2)} S_I(I_1, I_2), \end{aligned} \quad (3)$$

where  $I_1 = g(O_1)$  and  $I_2 = g(O_2)$ . This equation shows that the intensity similarities in the input and output domains are related by the first-order derivatives of the response function.

### 3.2. Image similarity

As described above, our method estimates camera response functions from the similarity among several images. Suppose we have as input a set of images of the same scene taken from the same viewpoint with the same camera parameters. Using all pairs of pixels from different images with the same pixel coordinates, we can compute the cdf's of the observations and the image similarity measures in the output domain.

We begin by defining the *image similarity* in the input domain as follows:

$$E_1 \stackrel{\text{def}}{=} \frac{1}{Z} \iint p(I_1|I_2)S_I(I_1, I_2)dI_1dI_2, \quad (4)$$

where  $p(I_1|I_2)$  is the likelihood of observing input  $I_1$  after observing input  $I_2$  from the same scene radiance. Note that  $p(I_1|I_2)$  and  $S_I(I_1, I_2)$  cannot be directly computed from the observed images. We explain how the observed images and the physics-based noise model are used in this equation in Section 3.3 and Section 3.4, respectively. In Equation (4),

the normalization term  $Z$  is defined as

$$Z = \sqrt{\iint (S_I(I_1, I_2))^2 dI_1 dI_2}.$$

By defining the similarity in the input domain, we can evaluate Equation (4), despite varying camera specifications.

Equation (4) can be viewed as the sum of the intensity similarity  $S_I$  over all inputs weighed by the cdf  $p(I_1|I_2)$ . The weighting is derived by normalizing the joint distribution of  $I_1$  and  $I_2$  by the probability density of observing either  $I_1$  or  $I_2$ . For example, in the case that  $p(I_2)$  is used for normalization, we have

$$p(I_1|I_2) = \frac{p(I_1, I_2)}{p(I_2)}.$$

### 3.3. Estimation method

Now we transform Equation (4) to the output domain.  $p(O_1|O_2)$  is more suitable for estimating response functions because we can compute it from image measurements. Using Equation (2), Equation (4) becomes:

$$\begin{aligned} E_2 &= \frac{1}{Z} \iint \frac{1}{g'(O_1)} p(O_1|O_2) S_I(g(O_1), g(O_2)) \\ &\quad g'(O_1) g'(O_2) dO_1 dO_2 \\ &= \frac{1}{Z} \iint g'(O_2) p(O_1|O_2) S_I(g(O_1), g(O_2)) dO_1 dO_2. \end{aligned} \quad (5)$$

The image similarity measure  $E_2$  evaluates how similar images are. If the correct response function  $g$  is assumed, the image similarity  $E_2$  of our input images is expected to become high because the only fluctuations are image noise.

In Equation (5),  $p(O_1|O_2)$  is treated as a known parameter because it is directly measured from the input images. Our goal is to estimate the two unknown parameters  $g$  and  $S_I$ . We formulate the estimation problem in an energy maximization framework as

$$(\hat{g}, \hat{S}_I) = \underset{g, S_I}{\operatorname{argmax}} E_2(g, S_I),$$

where  $\hat{g}$  and  $\hat{S}_I$ , respectively, are the estimates of the inverse response function and data-specific intensity similarity in the input domain.

### 3.4. Modeling intensity similarity

In practice, it is difficult to treat  $S_I(I_1, I_2)$  in Equation (5) in a nonparametric form. To parameterize  $S_I(I_1, I_2)$ , we use the physics-based noise model described by Healey and Kondepudy [6]. As described in [6], the input  $I$  with sensor noise can be written as

$$I = aP + N_{DC} + N_S + N_R,$$

where  $a$  is a factor of photon-to-electron conversion efficiency with amplification, and  $P$  is the number of photons.  $N_{DC}$ ,  $N_S$ , and  $N_R$  indicate dark current noise, shot noise,

and readout noise, respectively. Without noise, the noise-free input  $\tilde{I}$  equals  $aP$ . Assuming the noise sources are independent, the cdf  $p(I|\tilde{I})$  is the convolution of the noise distributions.

Shot noise is known to obey a Poisson distribution. We assume natural lighting conditions where a large number of photons are available. The profile of the Poisson distribution approaches a Gaussian distribution as the number of events increases, so we can approximate the Poisson distribution by a Gaussian distribution with zero-mean and variance  $A\tilde{I}$ , where  $A$  is some positive real number. The distributions of the dark current noise and the readout noise also follow Gaussian distributions. Assuming these noise sources are independent, the combined noise model can be represented as a single normal distribution whose average is zero and variance is some real number  $B$ . As a result,  $p(I|\tilde{I})$  can be written as

$$p(I|\tilde{I}) = \frac{1}{\sqrt{2\pi(A\tilde{I} + B)}} \exp\left(-\frac{(I - \tilde{I})^2}{2(A\tilde{I} + B)}\right). \quad (6)$$

We compute the similarity  $S_I(I_1, I_2)$  is calculated using a numerical approximation of the integral, such as the midpoint method, from this equation:

$$\begin{aligned} S_I(g(O_1), g(O_2)) &= S_I(I_1, I_2) \\ &= \int p(I_1|\tilde{I}) p(I_2|\tilde{I}) p(\tilde{I}) d\tilde{I}. \end{aligned} \quad (7)$$

In the above equation,  $p(\tilde{I})$  is considered a uniform distribution [13]. In this way,  $S_I(I_1, I_2)$  can be parameterized by only two variables  $A$  and  $B$ .

## 4. Implementation

This section describes two key points of our implementation: the parameterization of the response functions and the solution algorithm.

### 4.1. Representation of inverse response functions

To reduce the computational cost, we represent the inverse response functions using the parametric model proposed by Grossberg and Nayar [5]. We chose this parameterization because of its compactness and ease of use. Although we use this parameterization, we do not use the prior distributions in our energy function for estimating the response function. Grossberg and Nayar use principal component analysis (PCA) of the database of real-world response functions (DoRF) to obtain a small number of eigenvectors that represent the space of the response functions. Using the principal components, we represent the inverse response function  $g$  as  $g = g_0 + \mathbf{H}\mathbf{c}$ , where  $g_0$  is the mean vector of all the inverse response functions,  $\mathbf{H}$  is a matrix in which a column vector represents an eigenvector, and  $\mathbf{c}$  is a vector

of PCA coefficients. Following the original paper [5], we use the first five eigenvectors. With this representation, the number of unknown variables decreases significantly, e.g., from 256 to 5 in the case of 8-bit images. Our method is not restricted to this parameterization. Other parameterization methods such as polynomial representations [15] could also be used.

## 4.2. Maximization of image similarity

We maximize the image similarity measure defined in Equation (5) using an iterative alternating algorithm. We continue the following steps until convergence:

**Algorithm:** Maximization of  $E_2$

**Input:** Images taken from the same viewpoint with fixed camera parameters.

**Output:** Inverse response function  $\hat{g}$  and data-specific similarity  $\hat{S}_I$ .

**Initialize** the similarity measure in the input domain  $S_I(I_1, I_2)$  using Equations (6)(7) with initial parameters  $A$  and  $B$ .

**Repeat until**  $E_2$  converges **do**

1. Maximize  $E_2$  with the fixed input similarity function  $S_I(I_1, I_2)$  to estimate the inverse response function  $g$  with  $\hat{g} = \text{argmax}_g E_2$ .
2. Update  $g$ :  $g \leftarrow \hat{g}$ .
3. Fix the inverse response function  $g$  and maximize  $E_2$  according to  $\hat{S}_I = \text{argmax}_{S_I} E_2$ . Note  $S_I$  is parameterized by  $A$  and  $B$ .
4. Update  $S_I$ :  $S_I \leftarrow \hat{S}_I$  by updating  $A$  and  $B$ .

**done**

We use the Matlab `fminsearch` algorithm, an implementation of the Nelder-Mead Simplex method [17] to maximize the energy function. We empirically set  $A = 0.06$  and  $B = 1$  as the initial values for 8-bit images. The maximization usually converges within ten iterations. Like all multi-dimensional non-linear optimization problems, our optimization can converge to a local maxima. We try to avoid this by restarting the optimization from neighborhoods of the current best estimate.

## 5. Experiments

We used two different conditions to evaluate the performance of the proposed algorithm. The first uses multiple images taken by a fixed camera with the same camera parameter settings, and the second uses only a single image. These two setups differ only in the way of computing the cdf's  $p(O_1|O_2)$  of the input dataset.

### 5.1. Multiple-image case

In this experiment, we run our algorithm using multiple shots of a static scene captured from a fixed view-

point with the same camera parameters. We compute a histogram  $h(O_1, O_2)$ , which represents the co-occurrence of two output intensity values  $O_1$  and  $O_2$ , by gathering all pairs of pixels with the same pixel coordinates. The cdf's of the image intensities  $p(O_1|O_2)$  are obtained from the histogram  $h$  by:

$$p(O_1|O_2) = \frac{h(O_1, O_2)}{\sum_{O_2} h(O_1, O_2)}.$$

**Results** We used the following four cameras for this experiment: *Point Grey Dragonfly2* (Camera A), *Sony DCR-TRV9E Video camera* (Camera B), *Sony DCR-TRV900NTSC Video camera* (Camera C), and *Canon EOS Kiss Digital*, the original model (Camera D). For estimating the response function of camera D, we used six images ( $3072 \times 2048$  pixels with ISO 400 gain level). For the other cameras, we used ten video frames (Camera A  $640 \times 480$  pixels, Camera B  $720 \times 576$  pixels, and Camera C  $720 \times 480$  pixels). These are considerably fewer images than were used by the prior noise-based estimation method [14].

To obtain ground truth inverse response functions for Cameras A and D, we used Mitsunaga and Nayar's method [15]. Cameras B and C do not have exposure controls, so we computed their response functions using the Macbeth color checker-based method [2].

Figure 3 shows typical results of our method. The images in the first column are the scenes used for the experiment. The second column shows the cdf's  $p(O_1|O_2)$  obtained from the input data. The third column shows the estimated intensity similarity in the input domain, and the fourth column shows the plot of the estimated inverse response functions and the corresponding ground truth curves. The horizontal and vertical axes represent normalized output and input, respectively. These results verify the accuracy of our method; the estimated curves are very close to the ground truth curves.

Table 1 summarizes the results of the experiment. For each camera, three different scenes are used. Because the algorithm is applied to the RGB channels independently, we use 9 datasets for each camera. The disparity is the mean of the maximum differences in the normalized input domain. These results verify the stability as well as the accuracy of our method.

**Estimation accuracy and the number of images** We performed another experiment to observe the relationship between the number of images and the estimation accuracy. One would expect higher accuracy as more images are used because of the increased statistical stability. We ran our algorithm using two, five, ten, and twenty frames from Camera C. For each run, we used 9 different datasets. Figure 4 shows the results. We see a slight improvement in accuracy with increasing numbers of images, but the improvement is not significant. Thus, our method works well with a small number of images.

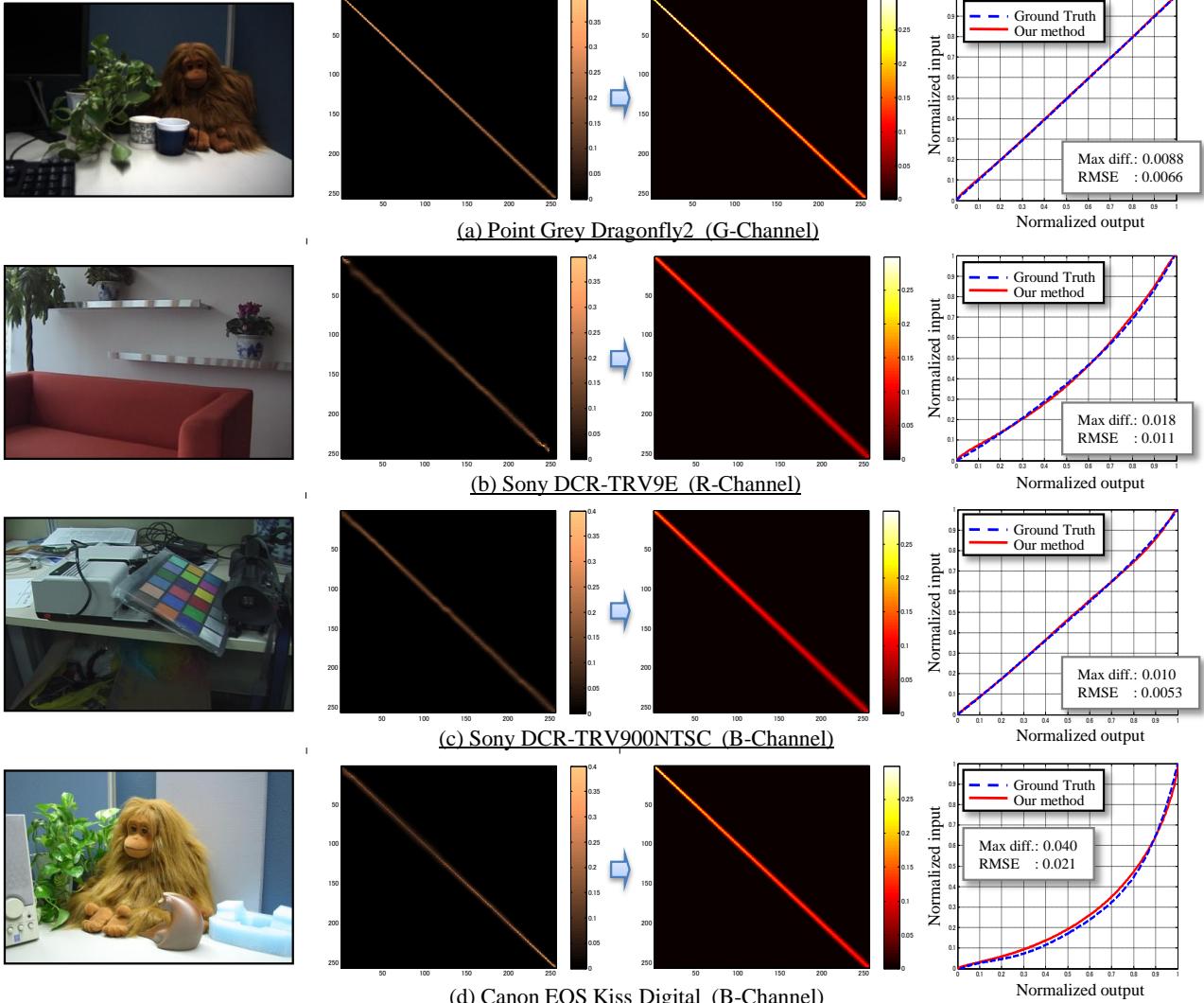


Figure 3. Results of estimating camera response functions from multiple images. From left to right, the first column shows an image from the dataset. The second column shows the computed cdf  $p(O_1|O_2)$ . The estimated intensity similarity and inverse response functions are shown in the third and fourth column, respectively. The estimated inverse response function is plotted with a red solid line, while the ground truth is plotted with a blue dotted line.

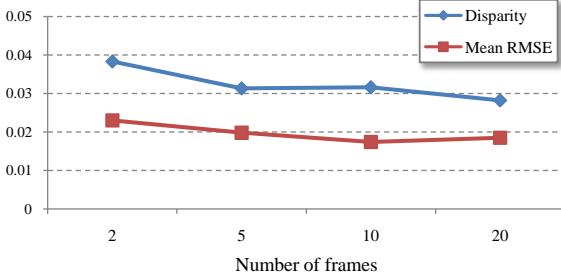


Figure 4. Plot of the number of images versus estimation accuracy. It shows our method performs well even with a small number of images.

## 5.2. Single-image case

We further extend our method to handle a single-image case. The estimation algorithm is the same, but we use a different means to compute the cdf  $p(O_1|O_2)$  because we

do not have registered pixel pairs.

As in [14], we assume that the scene contains a sufficiently large portion of uniformly colored regions. In this experiment, histograms  $h$  are obtained from spatially pairing neighboring pixels. We used four neighbors; up, down, left, and right. To efficiently gather pixel pairs, we evaluate the pixel similarity by a normal distribution with standard deviation  $\sigma$  in intensity differences for each color channel. The values of  $\sigma$  in the images used for this experiment were between 7 and 9, which tended to increase with increasing the ISO gain level. We use only pixel pairs whose similarity is within  $3\sigma$  in the distribution. While it may contain outliers in the histogram  $h$ , the frequency is expected to be dominated by pixel pairs with the same scene radiance.

**Results** For this experiment, we used Camera D with three different gain levels: ISO 400, 800, and 1600. A larger

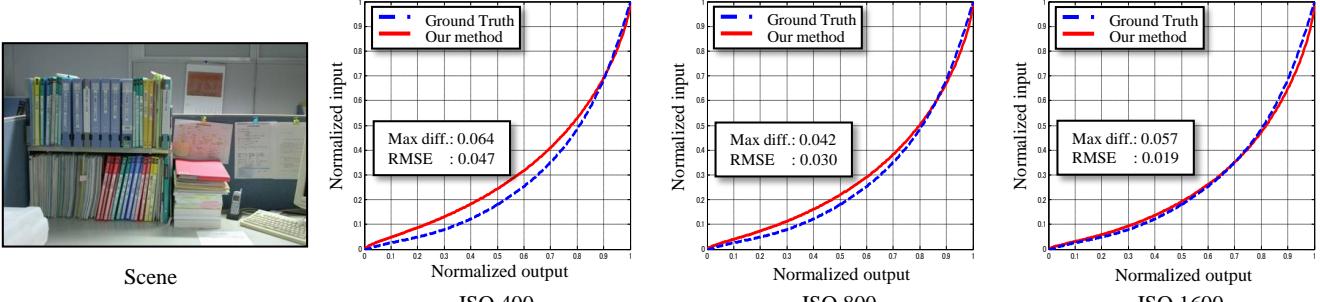


Figure 5. Result of estimating red channel inverse response functions from a single image. From left to right, we show the input image and the results for ISO 400, ISO 800, and ISO 1600.

Table 1. Mean RMSE and disparity of the estimated inverse response functions in the the normalized input domain. Three different scenes were used for each camera.

Camera	Mean RMSE	Disparity
A. Dragonfly	0.0068	0.011
B. DCR-TRV9E	0.026	0.049
C. DCR-TRV900NTSC	0.017	0.032
D. EOS Kiss Digital	0.025	0.046

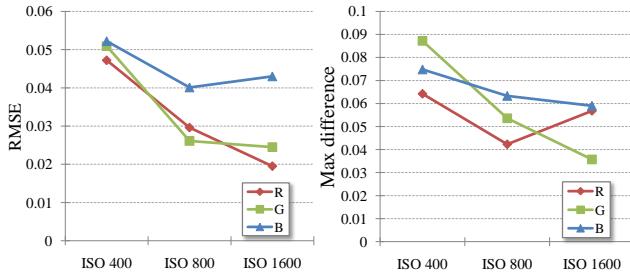


Figure 6. The relationship between ISO level and RMSE (left) and between ISO level and maximum difference (right).

ISO setting corresponds to a higher noise level.

Figure 5 shows the red channel estimation results with different ISO gain levels. We used the same scene to compare the different gains. At the ISO 1600 gain level, the estimation result is entirely satisfactory, even though only a single image is used. In contrast to the other single-image methods [8, 9, 14], our method does not use any special image processing such as edge detection and image segmentation, nor does it require a prior model of the inverse response functions.

### 5.3. Applications

In this sub-section, we introduce applications of our method. We can estimate not only camera response functions but also the function of probabilistic intensity similarity. As a use of the estimated response function, we first show the result of linearization of image intensity. Next, we show the result of edge detection using the estimated probabilistic intensity similarity.

**Linearization of image intensity** Figure 7 shows the result of intensity linearization for image taken by Camera D.



Figure 7. The result of intensity linearization using the estimated camera response functions for Camera D.

For linearization, we estimated the inverse response functions of the RGB channels independently. We use  $I = g(O)$  to produce the linearized image shown on the right.

**Edge detection** We show in this application the effectiveness of our estimated intensity similarity measure for edge detection. We used a  $3 \times 3$  four-neighbor Laplacian filter, i.e.,  $[0, -1, 0; -1, 4, -1; 0, -1, 0]$ . The Laplacian operation at some pixel  $\mathbf{p}$  is written as

$$\sum_{\mathbf{q} \in adj(\mathbf{p})} d(O(\mathbf{p}), O(\mathbf{q})),$$

where  $adj(\mathbf{p})$  is a set of neighboring pixels of  $\mathbf{p}$ ,  $O(\mathbf{p})$  is the intensity value at pixel  $\mathbf{p}$ , and  $d(O_1, O_2)$  is a signed distance metric between  $O_1$  and  $O_2$ . The standard definition of the metric is their difference, i.e.,  $d(O_1, O_2) = O_1 - O_2$ . Using the probabilistic intensity similarity measure, it becomes

$$d(O_1, O_2) \stackrel{\text{def}}{=} -\log(S_O(O_1, O_2)),$$

as described in [13]. We transform the similarity measure from the input to the output domain according to Equation (3). Thresholding after convolving with the Laplacian filter produces a binary edge image.

Figure 8 shows the comparison between the two metrics for edge detection using a noisy input image. The threshold for the binarization was carefully tuned to produce the visually best result; we used 22 for the standard metric and 40 for our metric. The intensity similarity measure produces a better result. Because our proposed method computes data-specific intensity similarity measures as well as response functions, it enables edge detection that is robust to noise. As in [10], we can perform this type of noise-adaptive image processing on any images that are suitable for our single-image method.

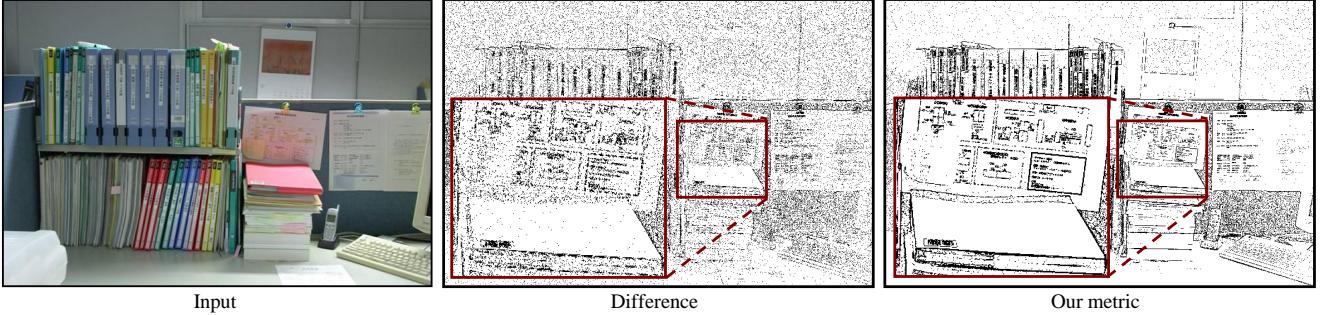


Figure 8. Result of edge detection from a noisy image (ISO-1600) taken by Camera D. The Laplacian filter is used with two different metrics: the standard intensity difference and probabilistic intensity similarity.

## 6. Conclusions

We have proposed a method for estimating camera response functions by maximizing the image similarity measure defined as the integral of the probabilistic intensity similarity. Using the image similarity, our method can estimate inverse response functions from only a few images. We also presented an extension that handles the single-image case.

Our experimental results for both the multiple and single image cases quantitatively demonstrate the effectiveness of the proposed algorithm. Our method does not require any prior knowledge about the inverse response functions, unlike [8, 9, 14]. In the future, we plan to explore cases where the physic-based noise model no longer holds, such as low-light conditions.

**Limitations** Our method uses a physics-based image noise model [6]. Therefore, if the noise distribution deviates from this model as described in [1], our approach may produce inaccurate estimates. The single-image method described in this paper assumes a scene containing many local planar patches of uniform colors. If the scene does not fit this assumption, e.g., gradient-colored surfaces, the method may produce an unsatisfactory result. Note that this assumption is not required for the multiple-image case.

Since our method uses noise observations, it is unable to perform the estimation from noise-free images. In practice, noise is always present. Our experimental result suggests that the method works well with noisier images, however, this tendency is expected to be capped at a certain noise level. For future work, we plan to explore the upper bound of the noise level where the method still works.

## Acknowledgement

The authors would like to thank Bennett Wilburn for his useful feedback on this research.

## References

- [1] F. Alter, Y. Matsushita, and X. Tang. An intensity similarity measure in low-light conditions. In *Proc. of European Conf. on Comp. Vis. (ECCV)*, pages 267–280, 2006. 8
- [2] Y.-C. Chang and J. F. Reid. Rgb calibration for color image analysis in machine vision. *IEEE Trans. on Image Processing*, 5(10):1414–1422, 1996. 2, 5
- [3] P. E. Debevec and J. Malik. Recovering high dynamic range radiance maps from photographs. In *Proc. of ACM SIGGRAPH*, pages 369–378, 1997. 2
- [4] M. D. Grossberg and S. K. Nayar. What can be known about the radiometric response function from images? In *Proc. of European Conf. on Comp. Vis. (ECCV)*, volume 2, pages 189–205, 2002. 2
- [5] M. D. Grossberg and S. K. Nayar. What is the space of camera response functions? In *Proc. of Comp. Vis. and Patt. Recog. (CVPR)*, volume 2, pages 602–609, 2003. 2, 4, 5
- [6] G. E. Healey and R. Kondepudy. Radiometric ccd camera calibration and noise estimation. *IEEE Trans. on Patt. Anal. and Mach. Intell.*, 16(3):267 – 276, 1994. 1, 4, 8
- [7] S. J. Kim and M. Pollefeys. Radiometric alignment of image sequences. In *Proc. of Comp. Vis. and Patt. Recog. (CVPR)*, pages 645–651, 2004. 2
- [8] S. Lin, J. Gu, S. Yamazaki, and H. Y. Shum. Radiometric calibration from a single image. In *Proc. of Comp. Vis. and Patt. Recog. (CVPR)*, volume 2, pages 938–945, 2004. 2, 7, 8
- [9] S. Lin and L. Zhang. Determining the radiometric response function from a single grayscale image. In *Proc. of Comp. Vis. and Patt. Recog. (CVPR)*, volume 2, pages 66–73, 2005. 2, 7, 8
- [10] C. Liu, W. T. Freeman, R. Szeliski, and S. B. Kang. Noise estimation from a single image. In *Proc. of Comp. Vis. and Patt. Recog. (CVPR)*, pages 901–908, 2006. 3, 7
- [11] S. Mann. Comparametric equations with practical applications in quantigraphic image processing. *IEEE Trans. on Image Processing*, 9(8):1389–1406, 2000. 2
- [12] 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 Conf.*, pages 422 – 428, 1995. 2
- [13] Y. Matsushita and S. Lin. A probabilistic intensity similarity measure based on noise distributions. *Proc. of Comp. Vis. and Patt. Recog. (CVPR)*, 2007. 1, 3, 4, 7
- [14] Y. Matsushita and S. Lin. Radiometric calibration from noise distributions. *Proc. of Comp. Vis. and Patt. Recog. (CVPR)*, 2007. 2, 5, 6, 7, 8
- [15] T. Mitsunaga and S. K. Nayar. Radiometric self-calibration. In *Proc. of Comp. Vis. and Patt. Recog. (CVPR)*, volume 2, pages 374–380, 1999. 2, 5
- [16] S. K. Nayar and T. Mitsunaga. High dynamic range imaging: Spatially varying pixel exposures. In *Proc. of Comp. Vis. and Patt. Recog. (CVPR)*, volume I, pages 472–479, 2000. 2
- [17] J. A. Nelder and R. Mead. A simplex method for function minimization. *Computer Journal*, 7:308–312, 1965. 5
- [18] C. Pal, R. Szeliski, M. Uyttendaele, and N. Jojic. Probability models for high dynamic range imaging. In *Proc. of Comp. Vis. and Patt. Recog. (CVPR)*, volume 2, pages 173–180, 2004. 2
- [19] Y. Tsai, V. Ramesh, and T. Kanade. Statistical calibration of ccd imaging process. In *Proc. of Int'l Conf. on Comp. Vis. (ICCV)*, volume 1, pages 480–487, 2001. 2