

# Estimating radiometric response functions from image noise variance

Jun Takamatsu<sup>1\*</sup>, Yasuyuki Matsushita<sup>2</sup>, and Katsushi Ikeuchi<sup>1\*\*</sup>

<sup>1</sup> Microsoft Institute for Japanese Academic Research Collaboration (MS-IJARC)

`j-taka@is.naist.jp`, `ki@cvl.iis.u-tokyo.ac.jp`

<sup>2</sup> Microsoft Research Asia

`yasumat@microsoft.com`

**Abstract.** We propose a method for estimating radiometric response functions from observation of image noise variance, not profile of its distribution. The relationship between radiance intensity and noise variance is affine, but due to the non-linearity of response functions, this affinity is not maintained in the observation domain. We use the non-affinity relationship between the observed intensity and noise variance to estimate radiometric response functions. In addition, we theoretically derive how the response function alters the intensity-variance relationship. Since our method uses noise variance as input, it is fundamentally robust against noise. Unlike prior approaches, our method does not require images taken with different and known exposures. Real-world experiments demonstrate the effectiveness of our method.

## 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 actual observation due to non-linear *camera response functions*. Linearization of observed image intensity is important for many vision algorithms to work, therefore the estimation of the response functions is needed.

Scene radiance intensity (input)  $I$  and observed image intensity (output)  $O$  are related by the response function  $f$  as  $O = f(I)$ . Assuming it is continuous and monotonic, the response function can be inverted to obtain the *inverse response function*  $g (= f^{-1})$ , and measured image intensities can be linearized using  $I = g(O)$ . Since only observed output intensities  $O$  are usually available, most estimation methods attempt to estimate the inverse response functions.

### 1.1 Prior work

Radiometric calibration methods assume known characteristics of radiance at a camera to estimate unknown response functions. One class of these methods uses information about the ratio of input radiance intensities. The Macbeth color checker is used for estimating the response functions by several research groups such as [1]. Nayar and Mitsunaga [2] use an optical filter with spatially

---

\* Nara Institute of Science and Technology; part of this work was done during visiting Microsoft Research Asia.

\*\* Institute of Industrial Science, University of Tokyo

varying transmittance; the variation corresponds to the ratio. Instead of using such special equipment, some methods use a set of input images taken with different exposures from a fixed viewpoint so that the ratio becomes available. These methods are divided into parametric and non-parametric approaches.

In the parametric framework, Mann and Picard [3] 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 [4]. 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 [6] uses the *comparametric* equation that defines the relationship between two images taken with different exposure times.

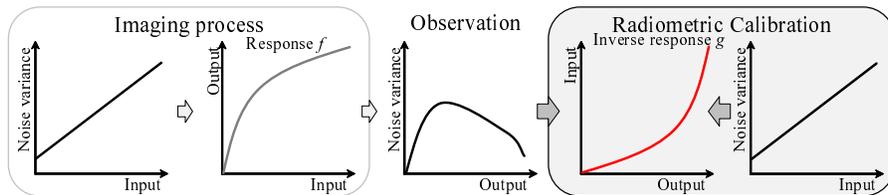
Debevec and Malik estimate response functions with a non-parametric representation using a smoothness constraint [7]. Tsin *et al.* estimate non-parametric forms of response functions using the statistical analysis [8]. Pal *et al.* propose to use a Bayesian network consisting of probabilistic imaging models and prior models of response functions [9]. While non-parametric approaches have greater descriptive power, the large number of unknown variables 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 [10] estimate the function from images with scene motion using brightness histograms, which avoid difficult pixelwise image registration. Kim and Pollefeys [11] propose a method that allows free camera motion by finding the correspondence among images. These methods can handle more general cases, but still require images taken with multiple exposures.

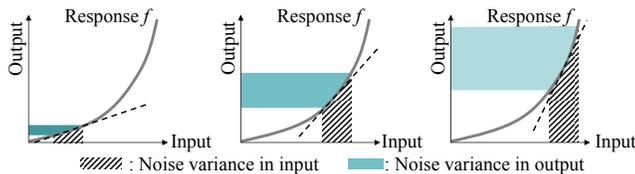
Another class of the estimation methods is based on the physics of the imaging process. Lin *et al.* estimate the function from the color blending around edges in a single image [12]. For grayscale images, a 1D analogue of the 3D color method is presented in [13]. Wilburn *et al.* propose to use the temporal blending of irradiance [14]. Matsushita and Lin propose an estimation method from profiles of image noise distributions [15]; their method relies on the assumption that profiles of noise distributions are symmetric. More recently, Takamatsu *et al.* [16] propose to use a probabilistic intensity similarity measure [17] for the estimation. Unlike these approaches, our method only uses *noise variance*, which is considered a lower level of information.

## 1.2 Proposed method

In this paper, we propose a method for estimating camera response functions from observation of image noise variance. Figure 1 illustrates the overview of the proposed method. As shown in the figure, noise variance in the input domain has an affine relationship with input intensity. This relationship, however, is not maintained in the output domain due to the non-linearity of camera response functions. By estimating a function that projects the non-affine intensity-variance relationship to an affine one, our method derives the response functions.



**Fig. 1.** The noise variance in the input domain has an affine relationship with input intensity level. Due to the non-linearity of the response function, the affine relationship is lost in the output domain. The proposed method estimates the response functions by recovering the affinity of the measured noise variances.



**Fig. 2.** The relationship between response function and noise variances in input and output domains. The magnitude of output noise variance (height of the filled region) varies with the slope of the response function with a fixed input noise variance.

Figure 2 depicts the relationship between the response function and noise variances in input and output domains. The ratio of the input/output noise variance is closely related to the slope of the response function.

Unlike Matsushita and Lin’s work [15], our method does not rely on profiles of noise distributions which are more difficult to obtain in practice. The proposed method has the following two advantages over their method. First, it uses the more reasonable assumption; the affinity of intensity-variance relationship in the input domain. We derive this relationship from the nature of photon (electron) noise [18, 19]. Second, the proposed method uses only noise variances rather than noise profiles, therefore it only requires less amount of information about noise. Practically, it makes the procedure of noise measurement easier.

This paper has two major contributions. It first theoretically derives how the radiometric response function alters the intensity-variance relationship, and vice versa. While it has been pointed out that these two quantities are related [8, 20], the exact alternation has not been explicitly described. Second, it introduces a method that has a wide range of applicability even with noisy observations. While many existing algorithms break down when the noise level is high, we show that the proposed method is not sensitive to the noise level because it uses the noise as information.

## 2 Noise Variance and Response Functions

This section provides a theoretical relationship between the response function and noise variances in the input and output domains. We show that the levels of noise variance in the input and output domains are related by the square of the first derivative of the camera response function. This result is later used to develop the estimation algorithm in Section 3.

## 2.1 Derivation of the relationship between noise variance and response functions

We treat the relationship between input or output intensity and noise variance as the function similar to Liu *et al.* [20]. The noise variance function  $\sigma_O^2(\tilde{O})$  in the output domain, where the noise-free output intensity is  $\tilde{O}$ , can be described as

$$\sigma_O^2(\tilde{O}) = \int (O - \mu_O)^2 p(O|\tilde{O}) dO = \int (f(I) - \mu_O)^2 p(I|\tilde{I}) dI.$$

The conditional density function (cdf)  $p(O|\tilde{O})$  represents the noise distribution in the output domain, i.e., the probability that the output intensity becomes  $O$  when the noise-free intensity is  $\tilde{O}$ . Likewise, the cdf  $p(I|\tilde{I})$  represents the noise distribution in the input domain when the noise-free input intensity is  $\tilde{I}$ . The function  $f$  is the response function, and  $\tilde{O}$  and  $\tilde{I}$  are related by  $f$  as  $\tilde{O} = f(\tilde{I})$ .  $\mu_O (= \mu_{O(\tilde{O})})$  is the expectation of the output intensity with the cdf  $p(O|\tilde{O})$ .

Using the Taylor series of  $f(I)$  around  $\tilde{I}$  and assuming that the second- and higher-degree terms of the series are negligible (discussed in Section 2.3), we obtain

$$\sigma_O^2(\tilde{O}) \simeq f'^2(\tilde{I}) \sigma_I^2(\tilde{I}), \quad (1)$$

where  $\sigma_I^2(\tilde{I})$  is the function of noise variance in the input domain when the noise-free input intensity is  $\tilde{I}$ . This equation shows the relationship between the noise variance and the response functions. This relationship has been pointed out in [8, 20], however, the correctness of the equation was not thoroughly discussed.

To simplify the notation in derivation of Eq. (1), we define:

$$\begin{aligned} \mu_I &= \mu_I(\tilde{I}) = \int I p(I|\tilde{I}) dI, \quad I_d = \mu_I - \tilde{I}, \\ M_n &= M_n(\tilde{I}) = \int (I - \mu_I)^n p(I|\tilde{I}) dI \quad (n \geq 0). \end{aligned} \quad (2)$$

$\mu_I$  is the expectation, or mean, of input intensity  $I$  with the cdf  $p(I|\tilde{I})$ .  $I_d$  denotes the difference between the expectation  $\mu_I$  and the noise-free intensity  $\tilde{I}$ .  $M_n$  is the  $n$ -th moment about the mean. Note that we do not put any assumptions about the profile and model of the noise distributions.

The expectation of the output intensity  $\mu_O$  where the noise-free output intensity is  $\tilde{O}$  can be written as

$$\mu_O = \mu_{O(\tilde{O})} = \int f(I) p(I|\tilde{I}) dI = \tilde{O} + \sum_{j=1}^{\infty} N_j \frac{f^{(j)}(\tilde{I})}{j!}, \quad (3)$$

where  $N_j$  is the  $j$ -th moment about  $\tilde{I}$  defined as follows.

$$N_j = \int (I - \tilde{I})^j p(I|\tilde{I}) dI = \int ((I - \mu_I) + I_d)^j p(I|\tilde{I}) dI = \sum_{k=0}^j \binom{j}{k} I_d^{j-k} M_k. \quad (4)$$

Note that the response function  $f$  is represented using its Taylor series.

From the definition, the noise variance function  $\sigma_{\tilde{O}}^2(\tilde{O})$  where the noise-free output intensity is  $\tilde{O}$  can be derived as

$$\sigma_{\tilde{O}}^2(\tilde{O}) = \sum_{j=1}^{\infty} \left( \frac{f^{(j)}(\tilde{I})}{j!} \right)^2 L_{j,j} + 2 \sum_{j=1}^{\infty} \sum_{k>j}^{\infty} \frac{f^{(j)}(\tilde{I})}{j!} \frac{f^{(k)}(\tilde{I})}{k!} L_{j,k}, \quad (5)$$

where  $L_{j,k}$  is defined as

$$L_{j,k} = \int ((I - \mu_I)^j - N_j)((I - \mu_I)^k - N_k)p(I|\tilde{I})dI = N_{j+k} - N_j N_k. \quad (6)$$

As can be seen in the above equation,  $L_{j,k}$  is commutative ( $L_{j,k} = L_{k,j}$ ).

A detailed calculation gives us  $L_{1,1} = M_2 = \sigma_{\tilde{I}}^2(\tilde{I})$ ,  $L_{1,2} = 2I_d \sigma_{\tilde{I}}^2(\tilde{I}) + M_3, \dots$ . Substituting these terms into Eq. (5), we obtain:

$$\begin{aligned} \sigma_{\tilde{O}}^2(\tilde{O}) &= f'^2(\tilde{I})L_{1,1} + f'(\tilde{I})f''(\tilde{I})L_{1,2} + \dots \\ &= f'^2(\tilde{I})\sigma_{\tilde{I}}^2(\tilde{I}) + f'(\tilde{I})f''(\tilde{I})(2I_d \sigma_{\tilde{I}}^2(\tilde{I}) + M_3) + \dots \end{aligned} \quad (7)$$

Eq. (7) is the exact form of the relationship between the response function and noise variances in the input and output domains. By discarding the second- and higher-degree terms of Eq. (7), Eq. (1) is obtained. We discuss the validity of this approximation in Section 2.3.

## 2.2 Noise variance function

Eq. (1) shows the relationship between the response function and noise variance functions in the input and output domains. The input to our method is the measured noise variance in the output domain. This section models the noise variance function  $\sigma_{\tilde{I}}^2(\tilde{I})$  in the input domain so that the estimation algorithm for the inverse response function  $g$  can be developed.

Input intensity  $I$  with camera noise can be written as

$$I = aP + N_{DC} + N_S + N_R, \quad (8)$$

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 [21]<sup>3</sup>. The noise-free input intensity  $\tilde{I}$  equals to  $aP$ .

Now we consider the noise variance function in the input domain. We assume the different noise sources are independent. From Eq. (8), the noise variance function in the input domain,  $\sigma_{\tilde{I}}^2(\tilde{I})$ , can be written as

$$\sigma_{\tilde{I}}^2(\tilde{I}) = \tilde{I}\sigma_S^2 + \sigma_{DC}^2 + \sigma_R^2, \quad (9)$$

where  $\sigma_*^2$  denotes the variances of the different noise sources [8]. Eq. (9) can be written in a simplified form as

$$\sigma_{\tilde{I}}^2(\tilde{I}) = A\tilde{I} + B, \quad (10)$$

where  $A = \sigma_S^2$  and  $B = \sigma_{DC}^2 + \sigma_R^2$ . This equation clearly shows the affine relationship between the noise-free input intensity  $\tilde{I}$  and the noise variance  $\sigma_{\tilde{I}}^2(\tilde{I})$ .

<sup>3</sup> The effect of the fixed-pattern noise is included in the term  $P$ .

### 2.3 Validity of the approximation

It is important to consider the validity of the approximation in Eq. (7). In this section, we show it in the following steps. First, we show that  $L_{i,j}$  becomes exponentially smaller as  $i+j$  increases. Second, the second largest term in Eq. (7),  $L_{1,2}$ , is small enough to be negligible compared with  $L_{1,1}$  through a detailed calculation. Hereafter, we normalize the input and output intensity ranges from 0 to 1.

*Relationship between  $L_{i,j}$  and  $i+j$*  By assuming independence of different noise sources, the second- and higher-order moments can be computed by summing up the moments of noises from different sources. Three types of noise sources must be considered: dark current noise, shot noise, and readout noise [21]. We do not consider low-light conditions [22], so the effect of the dark current noise becomes small.

The probability density of the readout noise can be considered to have a normal distribution with a mean value equal to the noise-free input intensity  $\tilde{I}$ . The moment can be written as

$$N_{R_i} = \begin{cases} 0 & (i \text{ is odd}) \\ \prod_{j=1}^{i/2} (2j-1) \sigma_R^2 & (i \text{ is even}). \end{cases} \quad (11)$$

In Eq. (11),  $(2j-1) \sigma_R^2 \ll 1$ , so the  $i$ -th moment of the readout noise  $N_{R_i}$  about the noise-free input intensity becomes exponentially smaller as  $i$  increases.

Shot noise is modeled as Poisson distribution [21]. From the theory of generalized Poisson distribution [23], the moment  $M_{S_i}$  about the mean of the distribution is defined as

$$M_{S_i} \simeq a^{i-2} (\sigma_{S_i}^2 + O(\sigma_{S_i}^4)) \quad (i \geq 2), \quad (12)$$

since the minimum unit of the distribution equals to  $a$  (See Eq. (8)).  $\sigma_{S_i}^2$  is the variance of shot noise where the noise-free input intensity is  $\tilde{I}$ . By substituting this into Eq. (4) yields

$$N_{S_i} \simeq I_d^i + \sum_{j=2}^i \binom{i}{j} I_d^{i-j} a^{j-2} \sigma_{S_i}^2. \quad (13)$$

Even in the worst case where  $\binom{i}{j}$  is overestimated as  $2^i$ ,  $N_{S_i}$  becomes exponentially smaller, since Eq. (13) is rewritten as

$$N_{S_i} \leq I_d^i + \sum_{j=2}^i (2I_d)^{i-j} (2a)^{j-2} (2\sigma_{S_i})^2, \quad (14)$$

and we know that  $2I_d \ll 1$ ,  $2a \ll 1$ , and  $(2\sigma_{S_i})^2 \ll 1$ . This equation shows that  $N_{S_i}$  exponentially decreases as  $i$  increases.

The term  $L_{i,j}$  is defined as  $L_{i,j} = N_{i+j} - N_i N_j$  in Eq. (6). Because the  $i$ -th moment of image noise  $N_i$  can be computed as the sum of the readout and shot noise as  $N_i = N_{R_i} + N_{S_i}$ , it also becomes exponentially smaller as  $i$  increases. From these results, we see that the term  $L_{i,j}$  becomes exponentially smaller as  $i + j$  increases.

*Ratio of  $L_{1,1}$  to  $L_{1,2}$*  Now we show that  $L_{1,2}$  is small enough to be negligible compared with  $L_{1,1}$ . A detailed calculation gives us  $L_{1,2} = 2I_d M_2 + M_3$ . The third moment of shot noise  $M_{S_3}$  can be computed from Eq. (12). Also, the third moment of readout noise can be obtained using Eq. (4) as

$$M_{R_3} = N_{R_3} - 3I_d N_{R_2} - I_d^3 = -3I_d \sigma_R^2 - I_d^3. \quad (15)$$

From these results, the following equation is obtained:

$$L_{1,2} = 2I_d M_2 - 3I_d \sigma_R^2 - I_d^3 + a \sigma_{S_I}^2. \quad (16)$$

Since  $M_2 \simeq \sigma_R^2 + \sigma_{S_I}^2$ ,  $a \ll I_d$ , if  $M_2 \ll I_d^2$ , the order of  $L_{1,2}$  is roughly the same as the order of  $I_d M_2$ . Since  $I_d$  is the difference between the noise-free input intensity and the mean which can be naturally considered very small, it is implausible to have cases where  $M_2 \ll I_d^2$ .

From these results, the order of  $L_{1,2}$  is roughly equivalent to the order of  $I_d L_{1,1}$ , and  $I_d$  is small because it is computed in the normalized input domain, e.g., in the order of  $10^{-2}$  ( $\simeq 1/2^8$ ) in 8-bit image case. Therefore,  $L_{1,2}$  is about  $10^{-2}$  times smaller than  $L_{1,1}$ .

To summarize,  $L_{1,2}$  is sufficiently small compared with  $L_{1,1}$ , and  $L_{i,j}$  decreases exponentially as  $i + j$  increases. Also, because response functions are smooth,  $f'(\tilde{I}) \ll f''(\tilde{I})$ . Therefore, Eq. (7) can be well approximated by Eq. (1).

### 3 Estimation Algorithm

This section designs an evaluation function for estimating inverse response functions  $g$ , using the result of the previous section.

#### 3.1 Evaluation function

From Eqs. (1) and (10), the noise variance  $\sigma_O^2(\tilde{O})$  of the output intensity  $O$  is

$$\sigma_O^2(\tilde{O}) \simeq f'^2(\tilde{I}) \sigma_I^2(\tilde{I}) + \sigma_Q^2 = f'^2(\tilde{I})(A\tilde{I} + B) + \sigma_Q^2. \quad (17)$$

$\sigma_Q^2$  is the variance of the quantization noise, which affects after applying the response function. Using the inverse response function  $g$ , Eq. (17) can be rewritten as

$$\sigma_O^2(\tilde{O}) = \frac{1}{g'(\tilde{O})^2} (Ag(\tilde{O}) + B) + \sigma_Q^2. \quad (18)$$

The variance of the quantization noise  $\sigma_Q^2$  becomes  $\sigma_Q^2 = l^2/12$ , where  $l$  is the quantization interval. Since its distribution is uniform, the following equation holds:

$$\sigma_Q^2 = \int_{-\frac{l}{2}}^{\frac{l}{2}} x^2 p(x) dx = \frac{1}{l} \int_{-\frac{l}{2}}^{\frac{l}{2}} x^2 dx = \frac{l^2}{12}. \quad (19)$$

In the following, we use  $\sigma_{O_m}^2(\tilde{O})$  to represent the measured noise variance to discriminate from the analytic form of the noise variance  $\sigma_O^2(\tilde{O})$ . Using Eq. (18) and the measured noise variances  $\sigma_{O_m}^2(\tilde{O})$ , our method estimates the inverse response function  $g$  that minimizes the following evaluation function:

$$E_1(g; \sigma_{O_m}^2(\tilde{O})) = \min_{A,B} \sum_{\tilde{O}} (\sigma_O^2(\tilde{O}) - \sigma_{O_m}^2(\tilde{O}))^2. \quad (20)$$

Eq. (20) involves the estimation of  $A$  and  $B$ , which can be simply solved by linear least square fitting, given  $g$ .

To make the algorithm robust against the measuring errors, namely the erroneous component in the measured noise, we use weighting factors. Eq. (20) is changed to

$$E_2(g; \sigma_{O_m}^2(\tilde{O})) = \min_{A,B} \frac{1}{\sum w(\tilde{O})} \sum_{\tilde{O}} w(\tilde{O}) (\sigma_O^2(\tilde{O}) - \sigma_{O_m}^2(\tilde{O}))^2, \quad (21)$$

where the weight function  $w(\tilde{O})$  controls the reliability on the measured noise variance  $\sigma_{O_m}^2(\tilde{O})$  at the intensity level  $\tilde{O}$ . We use a Cauchy distribution (Lorentzian function) for computing the weight function  $w(\tilde{O})$ :

$$w(\tilde{O}) = \frac{1}{e^2 + \rho}, \quad (22)$$

where  $e$  is defined as  $e = \sigma_O^2(\tilde{O}) - \sigma_{O_m}^2(\tilde{O})$ . A damping factor  $\rho$  controls the relationship between the difference  $e$  and weight  $w(\tilde{O})$ . As  $\rho$  becomes smaller, the weight  $w(\tilde{O})$  decreases more rapidly as the difference  $e$  increases.

We also add a smoothness constraint to the evaluation function, and the evaluation function becomes

$$E_3(g; \sigma_{O_m}^2(\tilde{O})) = \frac{1}{\sum_{\tilde{O}} \sigma_{O_m}^2(\tilde{O})} E_2 + \lambda_s \frac{1}{n_{\tilde{O}}} \sum_{\tilde{O}} g''(\tilde{O})^2, \quad (23)$$

where  $n_{\tilde{O}}$  is the number of possible noise-free output intensity levels, e.g., 256 in 8-bit case.  $\lambda_s$  is a regularization factor that controls the effect of the smoothness constraint.  $1/\sum_{\tilde{O}} \sigma_{O_m}^2(\tilde{O})$  is a normalization factor that makes  $E_2$  independent of the degree of noise level.

Our method estimates the inverse response function  $\hat{g}$  by minimizing Eq. (23) given the measured noise variance  $\sigma_{O_m}^2(\tilde{O})$ :

$$\hat{g} = \underset{g}{\operatorname{argmin}} E_3(g; \sigma_{O_m}^2(\tilde{O})). \quad (24)$$

### 3.2 Representation of inverse response functions

To reduce the computational cost, we represent the inverse response functions using a parametric model proposed by Grossberg and Nayar [5]. In their method,

principal component analysis (PCA) is performed on the database of real-world response functions (DoRF) to obtain a small number of eigenvectors that can represent the space of the response functions. As done by Lin *et al.* [12, 13], we compute the principal components of the inverse response functions using the DoRF. Using the principal components, we represent the inverse response function  $g$  as  $\mathbf{g} = \mathbf{g}_0 + \mathbf{H}\mathbf{c}$ , where  $\mathbf{g}_0$  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 Lin *et al.* [12, 13], we use the first five eigenvectors. Using this representation, the number of unknown variables is significantly decreased, e.g., from 256 to 5 in the case of 8-bit images.

### 3.3 Implementation

In our implementation, we set the damping factor  $\rho$  to the variance of the difference  $e$  in Eq. (22). The regularization factor  $\lambda_s$  is set to  $5 \times 10^{-7}$  from our empirical observation. Minimization is performed in an alternating manner. We perform the following steps until convergence:

1. minimize the evaluation function in Eq. (23) with fixing the weight function  $w(\tilde{O})$
2. recompute the values of the weight function  $w(\tilde{O})$  using the current estimation result

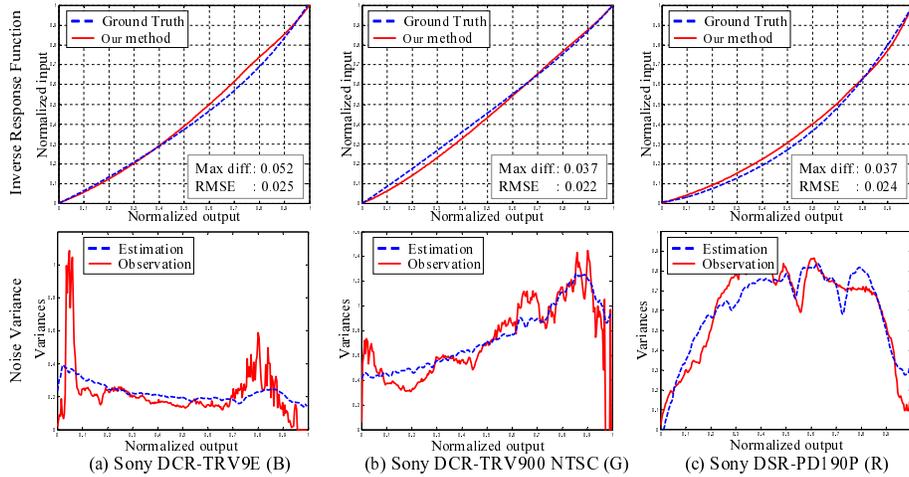
We use the Nelder-Mead Simplex method [24] as the minimization algorithm implemented in Matlab as a function `fminsearch`. The values of the weight function  $w(\tilde{O})$  are set to one for every  $\tilde{O}$  at the beginning. During the experiments, we used five initial guesses for the inverse response function  $g$  as the input to the algorithm. The converged result that minimizes the energy score is finally taken as the global solution.

## 4 Experiments

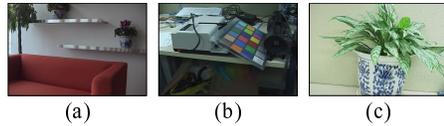
We used two different setups to evaluate the performance of the proposed algorithm; one is with multiple images taken by a fixed video camera, the other is using a single image. The two setups differ in the means for collecting noise variance information.

### 4.1 Multiple-images case

In this experiment, the measurements of noise variances are obtained by capturing multiple shots of a static scene from a fixed viewpoint with fixed camera parameters. From multiple images, a histogram of output intensities is created for each pixel. From the histogram, the noise-free output intensity  $\tilde{O}$  is determined by taking the mode of the distribution, assuming that the noise-free intensity should correspond to the most frequently observed signal. The pixelwise histograms are then merged together to form the histogram  $h(O, \tilde{O})$  for each



**Fig. 3.** Results of our estimation method. Top row: comparison of inverse response functions. Bottom row: measured noise variance and fitting result.



**Fig. 4.** Recorded scenes corresponding to the results in Figure 3 (a-c).

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

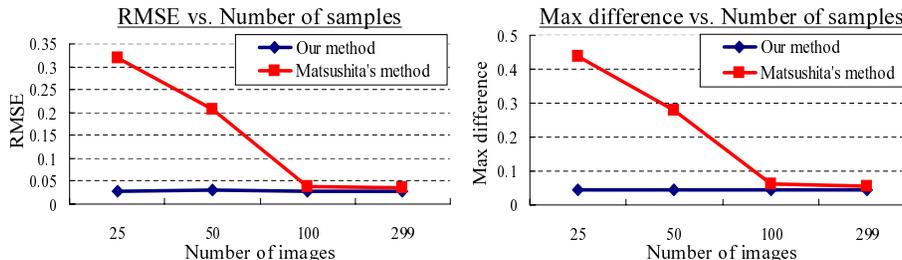
Camera	Mean RMSE	Disparity
A. DCR-TRV9E	0.026	0.053
B. DCR-TRV900	0.024	0.040
C. DSR-PD190P	0.033	0.055

output intensity level  $\tilde{O}$ . Finally, the noise distribution  $p(O|\tilde{O})$  is computed by normalizing the histogram  $h(O, \tilde{O})$  as

$$p(O|\tilde{O}) = \frac{h(O, \tilde{O})}{\sum_o h(O, \tilde{O})}. \quad (25)$$

*Results* We used three different video cameras for this experiment: *Sony DCR-TRV9E* (Camera A), *Sony DCR-TRV900 NTSC* (Camera B), and *Sony DSR-PD190P* (Camera C). To obtain the ground truth of Camera C, we used Mitsunaga and Nayar’s method [4], and the Macbeth color checker-based method [1], and combined these results by taking the mean. For Camera A and B, we used only the Macbeth color checker-based method [1] to obtain the ground truth because the exposure setting was not available in these cameras. The results obtained by the proposed method are compared with the ground truth curves.

Figure 3 shows the results of our algorithm. The top row shows the plot of the estimated inverse response functions with the corresponding ground truth curves. The bottom row shows the estimated and measured distributions of noise variances; the horizontal axis is the normalized output, and the vertical



**Fig. 5.** Comparison between our method and Matsushita and Lin’s method [15]. Our method uses noise variance, but not profiles of noise distributions. Our method works well even when the sampling number is relatively small.

axis corresponds to the noise variance. Figure 4 shows the scenes used to obtain these results.

Figure 3 (a) shows an estimation result using the blue channel of Camera A. The maximum difference is 0.052 and the RMSE is 0.025 in terms of normalized input. As shown in the bottom of (a), the noise variances in lower output levels contain severe measured errors. Our algorithm is robust against such errors because of the use of adaptive weighting factors. Figure 3 (b) shows the result of Camera B (green channel). The maximum difference is 0.037 and the RMSE is 0.022. Figure 3 (c) shows the estimation result of Camera C (red channel). The input frames are obtained by setting the camera gain to 12 db which causes high noise level. The maximum difference is 0.037 and the RMSE is 0.024.

Table 1 summarizes all the experimental results. For each camera, three different scenes are used. The algorithm is applied to RGB-channels independently, therefore 9 datasets for each camera are used. Disparity represents the mean of maximum differences in normalized input. From these results, the proposed method performs well even though the algorithm only uses the noise variance as input.

## 4.2 Comparison with another noise-based estimation method

Figure 5 shows the comparison between our method and Matsushita and Lin’s method [15]. Unlike other estimation methods, these two methods take noise as input. We use Camera B for the comparison.

As shown in the result, the estimation results are equivalent when the number of images is relatively large. However, Matsushita and Lin’s method breaks down when the number of samples becomes small, and our method shows significant superiority. In statistics, it is known that variance of measured from samples’ variance is inversely proportional to the number of the samples. Therefore, the measured variance becomes more stable than the profile of noise distribution does, as the number of samples increases. In addition, Matsushita and Lin’s symmetry criterion naturally requires large number of samples to make the noise profiles smooth, while it does not hold in the lower number of samples in Figure 5. These are why our method works well when the number of samples is relatively small.

### 4.3 Single-image case

We describe a single-image case where only one shot of the scene is available. In this setup, the distribution of noise variances are collected from uniformly colored image regions. However, the measured noise distribution is often insufficient to determine the inverse response functions because the limited measurements do not span the entire range of output levels. To better constrain the problem, we use a prior model  $p(g)$  of the inverse response functions obtained from the DoRF as done in [12] and [15].

Using the prior model  $p(g)$ , the MAP (maximum a posteriori) estimation is performed by maximizing the cdf  $p(g|\sigma_{O_m}^2(\tilde{O}))$  which represents the probability of the inverse response function being  $g$  when the measured noise variances are  $\sigma_{O_m}^2(\tilde{O})$  as

$$\hat{g} = \operatorname{argmax}_g p(g|\sigma_{O_m}^2(\tilde{O})) = \operatorname{argmax}_g (\log p(\sigma_{O_m}^2(\tilde{O})|g) + \log p(g)). \quad (26)$$

The likelihood  $p(\sigma_{O_m}^2(\tilde{O})|g)$  is defined as

$$p(\sigma_{O_m}^2(\tilde{O})|g) = \frac{1}{Z} \exp(-\lambda_p E_3(g; \sigma_{O_m}^2(\tilde{O}))), \quad (27)$$

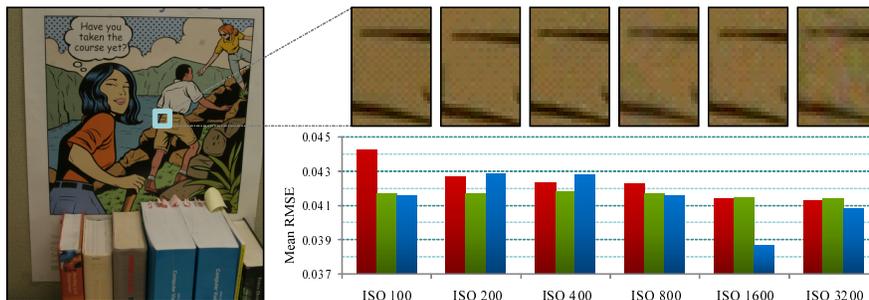
where  $Z$  is the normalization factor, and  $\lambda_p$  is a regularization coefficient that determines the weight on the evaluation function  $E_3$ . We empirically set  $\lambda_p$  to  $2 \times 10^4$  in the experiments. The prior model  $p(g)$  is formed using a multivariate Gaussian mixture model as

$$p(g) = \sum_{i=1}^K \alpha_i \mathcal{N}(g; \mu_i, \Sigma_i), \quad (28)$$

where  $\mathcal{N}$  represents a normal distribution with mean  $\mu$  and covariance matrix  $\Sigma$ , and  $\alpha_i$  is a weight factor. The prior model is obtained using the PCA coefficients of the inverse response functions in the DoRF by applying the cross-entropy method [25]. The number of normal distributions  $K$  is set to 5 in our experiments.

*Results* We used a *Canon EOS-20D* camera for the experiment. To obtain the ground truth, we used Mitsunaga and Nayar’s method [4] using images taken with different exposures. Since our focus is on estimating the inverse response functions from the measured noise variances, we photographed a scene composed of relatively flat and uniformly colored surfaces, so that the noise variances can be easily obtained. The left image in Figure 6 shows one of two scenes used for the experiment. We photographed them five times each at six different camera gains (ISO 100  $\sim$  3200). We manually selected 21 homogeneous image regions to obtain the noise variances as input. In total, we ran our estimation algorithm 60 times ( $= 2$  scenes  $\times 5$  shots  $\times 6$  ISO levels) for each RGB color channel.

Figure 6 summarizes the results of estimation at different ISO levels. The noise level increases with the ISO gain level, as shown by the cropped images on the top. The results indicate that the estimation is unaffected by the greater noise level. The mean RMSE is almost constant across the different ISO levels, which verifies that our method is not sensitive to the noise level.



**Fig. 6.** Relationship between the noise level and mean RMSE of the estimates. Left image shows one of the photographed scenes. Top row shows magnification of a part of the image at different ISO levels. Bottom row shows the mean RMSE of RGB channels at each ISO gain level, and demonstrates that our estimation method is independent of noise levels.

## 5 Conclusions

In this paper, we have proposed the method for estimating a radiometric response function using noise variance, not noise distribution, as input. The relationship between the radiometric response function and noise variances in input and output domains is explicitly derived, and this result is used to develop the estimation algorithm. The experiments are performed for two different scenarios; one is with multiple shots of the same scene, and the other is only from a single image. These experiments quantitatively demonstrate the effectiveness of the proposed algorithm, especially its robustness against noise. With our method, either special equipment or images taken with multiple exposures are not necessary.

*Limitations* It is better for our method that the measured noise variances cover a wide range of intensity levels. Wider coverage provides more information to the algorithm, so the problem becomes more constrained. This becomes an issue, particularly in the single-image case. In the single-image case, we used a simple method to collect the noise variances, but more sophisticated methods such like [20] can be used to obtain more accurate measurements that could potentially cover a wider range of intensity levels.

## Acknowledgement

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

## References

1. Chang, Y.C., Reid, J.F.: Rgb calibration for color image analysis in machine vision. *IEEE Trans. on Image Processing* **5** (1996) 1414–1422
2. Nayar, S.K., Mitsunaga, T.: High dynamic range imaging: Spatially varying pixel exposures. In: *Proc. of Comp. Vis. and Patt. Recog. (CVPR)*. (2000) 472–479

3. Mann, S., Picard, R.: Being ‘undigital’ with digital cameras: Extending dynamic range by combining differently exposed pictures. In: Proc. of IS & T 48th Annual Conf. (1995) 422 – 428
4. Mitsunaga, T., Nayar, S.K.: Radiometric self-calibration. In: Proc. of Comp. Vis. and Patt. Recog. (CVPR). (1999) 374–380
5. Grossberg, M.D., Nayar, S.K.: What is the space of camera response functions? In: Proc. of Comp. Vis. and Patt. Recog. (CVPR). (2003) 602–609
6. Mann, S.: Comparametric equations with practical applications in quantigraphic image processing. IEEE Trans. on Image Processing **9** (2000) 1389–1406
7. Debevec, P.E., Malik, J.: Recovering high dynamic range radiance maps from photographs. In: Proc. of ACM SIGGRAPH. (1997) 369–378
8. Tsin, Y., Ramesh, V., Kanade, T.: Statistical calibration of ccd imaging process. In: Proc. of Int’l Conf. on Comp. Vis. (ICCV). (2001) 480–487
9. Pal, C., Szeliski, R., Uyttendale, M., Jojic, N.: Probability models for high dynamic range imaging. In: Proc. of Comp. Vis. and Patt. Recog. (CVPR). (2004) 173–180
10. Grossberg, M.D., Nayar, S.K.: What can be known about the radiometric response function from images? In: Proc. of European Conf. on Comp. Vis. (ECCV). (2002) 189–205
11. Kim, S.J., Pollefeys, M.: Radiometric alignment of image sequences. In: Proc. of Comp. Vis. and Patt. Recog. (CVPR). (2004) 645–651
12. Lin, S., Gu, J., Yamazaki, S., Shum, H.Y.: Radiometric calibration from a single image. In: Proc. of Comp. Vis. and Patt. Recog. (CVPR). (2004) 938–945
13. Lin, S., Zhang, L.: Determining the radiometric response function from a single grayscale image. In: Proc. of Comp. Vis. and Patt. Recog. (CVPR). (2005) 66–73
14. Wilburn, B., Xu, H., Matsushita, Y.: Radiometric calibration using temporal irradiance mixtures. In: Proc. of Comp. Vis. and Patt. Recog. (CVPR). (2008)
15. Matsushita, Y., Lin, S.: Radiometric calibration from noise distributions. In: Proc. of Comp. Vis. and Patt. Recog. (CVPR). (2007)
16. Takamatsu, J., Matsushita, Y., Ikeuchi, K.: Estimating camera response functions using probabilistic intensity similarity. In: Proc. of Comp. Vis. and Patt. Recog. (CVPR). (2008)
17. Matsushita, Y., Lin, S.: A probabilistic intensity similarity measure based on noise distributions. In: Proc. of Comp. Vis. and Patt. Recog. (CVPR). (2007)
18. Janesick, J.R.: Photon Transfer. SPIE Press (2007)
19. Schechner, Y.Y., Nayar, S.K., Belhumeur, P.N.: Multiplexing for optimal lighting. IEEE Trans. on Patt. Anal. and Mach. Intell. **29** (2007) 1339–1354
20. Liu, C., Freeman, W.T., Szeliski, R., Kang, S.B.: Noise estimation from a single image. In: Proc. of Comp. Vis. and Patt. Recog. (CVPR). (2006) 901–908
21. Healey, G.E., Kondepudy, R.: Radiometric ccd camera calibration and noise estimation. IEEE Trans. on Patt. Anal. and Mach. Intell. **16** (1994) 267 – 276
22. Alter, F., Matsushita, Y., Tang, X.: An intensity similarity measure in low-light conditions. In: Proc. of European Conf. on Comp. Vis. (ECCV). (2006) 267–280
23. Consul, P.C.: Generalized Poisson Distributions: Properties and Applications. Marcel Dekker Inc, New York (1989)
24. Nelder, J.A., Mead, R.: A simplex method for function minimization. Computer Journal **7** (1965) 308–312
25. Botev, Z., Kroese, D.: Global likelihood optimization via the cross-entropy method with an application to mixture models. In: Proc. of the 36th Conf. on Winter simul. (2004) 529–535