Input-Dependent Uncorrelated Weighting for Monte Carlo Denoising (Supplementary Report)

Jonghee Back
Gwangju Institute of Science and Technology
South Korea
jongheeback@gm.gist.ac.kr

Binh-Son Hua
Trinity College Dublin
Ireland
binhson.hua@gmail.com

Toshiya Hachisuka
University of Waterloo
Canada
thachisu@uwaterloo.ca

Bochang Moon
Gwangju Institute of Science and Technology
South Korea
moonbochang@gmail.com

ABSTRACT

In this supplementary document, we provide a proof for Theorem 2 in the main paper, and then discuss some variants of Theorem 2. We also provide additional comparisons between our kernel and two input-independent kernels for input estimates with high sample counts.

CCS CONCEPTS

• **Computing methodologies → Ray tracing**.

KEYWORDS

Monte Carlo denoising, input-dependent weighting, uncorrelated weighting, unbiased denoising

ACM Reference Format:
Jonghee Back, Binh-Son Hua, Toshiya Hachisuka, and Bochang Moon. 2023. Input-Dependent Uncorrelated Weighting for Monte Carlo Denoising (Supplementary Report). In *SIGGRAPH Asia 2023 Conference Papers (SA Conference Papers '23), December 12–15, 2023, Sydney, NSW, Australia.* ACM, New York, NY, USA, 4 pages. https://doi.org/10.1145/3610548.3618177

1 PROOF OF THEOREM 2

Let us recall the combination function simplified from the previous combination function [Back et al. 2020]:

$$\hat{\mu}_c = y_c + \sum_{i \in \Omega_c} k_i (\Delta z_{ci} - \Delta y_{ci}). \tag{1}$$

In the combination, the input estimates y and z are assumed to be statistically independent of each other and unbiased estimates [Back et al. 2020].

We define the sub-averages of the estimate Δz_{ci} as $\Delta z^1_{ci}, \cdots, \Delta z^B_{ci}$ followed by a distribution $D(\mu_c - \mu_i, \sigma^2/n)$, where B is the number of sub-averages, σ^2 is the variance of the correlated samples, and n is the sample counts for a single sub-average. The sub-averages are independent and identically distributed. We assume that the distribution of $\Delta z^j_{ci}, \forall j \in [1, B]$ is a symmetric distribution.

SA Conference Papers '23, December 12–15, 2023, Sydney, NSW, Australia

Δz_{ci} is the average of all sub-averages, $\Delta z^1_{ci}, \cdots, \Delta z^B_{ci}$, and this functional relationship is represented by using a function f:

$$\Delta z_{ci} = f(\Delta z^1_{ci}, \cdots, \Delta z^B_{ci}) = \frac{1}{B} \sum_{j=1}^{B} \Delta z^j_{ci}. \tag{2}$$

Note that the average function f has the following properties:

$$\begin{aligned} f(\Delta z^1_{ci} + \alpha, \cdots, \Delta z^B_{ci} + \alpha) &= f(\Delta z^1_{ci}, \cdots, \Delta z^B_{ci}) + \alpha, \\ f(-\Delta z^1_{ci}, \cdots, -\Delta z^B_{ci}) &= -f(\Delta z^1_{ci}, \cdots, \Delta z^B_{ci}), \end{aligned} \tag{3}$$

where α is an arbitrary value.

Next, we define that a function k is a function of the estimates $\Delta z^1_{ci}, \cdots, \Delta z^B_{ci}$ (i.e., $k(\Delta z^1_{ci}, \cdots, \Delta z^B_{ci})$) that follows the conditions:

$$\begin{aligned} k(\Delta z^1_{ci} + \alpha, \cdots, \Delta z^B_{ci} + \alpha) &= k(\Delta z^1_{ci}, \cdots, \Delta z^B_{ci}), \\ k(-\Delta z^1_{ci}, \cdots, -\Delta z^B_{ci}) &= k(\Delta z^1_{ci}, \cdots, \Delta z^B_{ci}). \end{aligned} \tag{4}$$

By using two functions, f and k, the combination formula (Eq. 1) is represented as below:

$$\hat{\mu}_c = y_c + \sum_{i \in \Omega_c} k(\Delta z^1_{ci}, \cdots, \Delta z^B_{ci}) \left(f(\Delta z^1_{ci}, \cdots, \Delta z^B_{ci}) - \Delta y_{ci} \right). \tag{5}$$

The expectation of $\hat{\mu}_c$ can be expressed in the following manner:

$$\begin{aligned} E[\hat{\mu}_c] &= E\left[y_c + \sum_{i \in \Omega_c} k(\Delta z^1_{ci}, \cdots, \Delta z^B_{ci}) \left(f(\Delta z^1_{ci}, \cdots, \Delta z^B_{ci}) - \Delta y_{ci} \right) \right] \\ &= \mu_c + \sum_{i \in \Omega_c} E\left[k(\Delta z^1_{ci}, \cdots, \Delta z^B_{ci}) \left(f(\Delta z^1_{ci}, \cdots, \Delta z^B_{ci}) - \Delta y_{ci} \right) \right] \\ &= \mu_c + \sum_{i \in \Omega_c} \Big\{ E[k(\Delta z^1_{ci}, \cdots, \Delta z^B_{ci}) f(\Delta z^1_{ci}, \cdots, \Delta z^B_{ci})] \\ &\qquad - E[k(\Delta z^1_{ci}, \cdots, \Delta z^B_{ci})](\mu_c - \mu_i) \Big\}. \end{aligned} \tag{6}$$

Before proving its unbiasedness, we first prove the uncorrelatedness between $f(\Delta z^1_{ci}, \cdots, \Delta z^B_{ci})$ and $k(\Delta z^1_{ci}, \cdots, \Delta z^B_{ci})$, referring to the proof of Hogg's Theorem [Hogg 1960]. For a more intuitive proof of the unbiasedness, we temporarily change the notation of the estimates $\Delta z^1_{ci}, \cdots, \Delta z^B_{ci}$ to $x_1, \cdots, x_B$. The covariance of two statistics from functions f and k, which is denoted as C, is represented in the following integral form:

$$C = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} (f(x_1, \cdots, x_B) - \lambda) k(x_1, \cdots, x_B) \\ D(x_1) \cdots D(x_B) dx_1 \cdots dx_B, \tag{7}$$

where λ indicates the expected value of $f(x_1, \cdots, x_B)$, and D is the probability density function of $x_1, \cdots, x_B$.

By using the transformation of variables $x_j = \dot{x}_j + \lambda, \forall j \in [1, B]$, the covariance can be expressed as below:

$$C = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} (f(\dot{x}_1 + \lambda, \cdots, \dot{x}_B + \lambda) - \lambda) k(\dot{x}_1 + \lambda, \cdots, \dot{x}_B + \lambda) \\ D(\dot{x}_1 + \lambda) \cdots D(\dot{x}_B + \lambda) d\dot{x}_1 \cdots d\dot{x}_B. \tag{8}$$

By the properties of the functions f and k (i.e., $f(\dot{x}_1+\lambda, \cdots, \dot{x}_B+\lambda) = f(\dot{x}_1, \cdots, \dot{x}_B) + \lambda$ in Eq. 3 and $k(\dot{x}_1+\lambda, \cdots, \dot{x}_B+\lambda) = k(\dot{x}_1, \cdots, \dot{x}_B)$ in Eq. 4, the covariance in Eq. 8 can be changed in the following manner:

$$C = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f(\dot{x}_1, \cdots, \dot{x}_B) k(\dot{x}_1, \cdots, \dot{x}_B) \\ D(\dot{x}_1 + \lambda) \cdots D(\dot{x}_B + \lambda) d\dot{x}_1 \cdots d\dot{x}_B. \tag{9}$$

When employing a change of variables $\dot{x}_j = -\ddot{x}_j, \forall j \in [1, B]$, the covariance is as follows:

$$C = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f(-\ddot{x}_1, \cdots, -\ddot{x}_B) k(-\ddot{x}_1, \cdots, -\ddot{x}_B) \\ D(-\ddot{x}_1 + \lambda) \cdots D(-\ddot{x}_B + \lambda) d\ddot{x}_1 \cdots d\ddot{x}_B. \tag{10}$$

By the other properties of the functions f and k (i.e., $f(-\ddot{x}_1, \cdots, -\ddot{x}_B) = -f(\ddot{x}_1, \cdots, \ddot{x}_B)$ and $k(-\ddot{x}_1, \cdots, -\ddot{x}_B) = k(\ddot{x}_1, \cdots, \ddot{x}_B)$ in Eqs. 3 and 4) and the symmetric property of the distribution (i.e., $D(-\ddot{x}_j + \lambda) = D(\ddot{x}_j + \lambda)$ for all j), the covariance in Eq. 10 is converted into the covariance in Eq. 9 while inverting the sign:

$$C = -\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f(\ddot{x}_1, \cdots, \ddot{x}_B) k(\ddot{x}_1, \cdots, \ddot{x}_B) \\ D(\ddot{x}_1 + \lambda) \cdots D(\ddot{x}_B + \lambda) d\ddot{x}_1 \cdots d\ddot{x}_B = -C. \tag{11}$$

Therefore, the covariance C is zero.

Coming back to our notations for the combination, two statistics, $f(\Delta z_{ci}^1, \cdots, \Delta z_{ci}^B)$ and $k(\Delta z_{ci}^1, \cdots, \Delta z_{ci}^B)$, are uncorrelated:

$$\begin{aligned} &E[f(\Delta z_{ci}^1, \cdots, \Delta z_{ci}^B) k(\Delta z_{ci}^1, \cdots, \Delta z_{ci}^B)] \\ &= E[f(\Delta z_{ci}^1, \cdots, \Delta z_{ci}^B)] E[k(\Delta z_{ci}^1, \cdots, \Delta z_{ci}^B)] \\ &= (\mu_c - \mu_i) E[k(\Delta z_{ci}^1, \cdots, \Delta z_{ci}^B)]. \end{aligned} \tag{12}$$

Therefore, by plugging the statistical relationship between two statistics (Eq. 12) into Eq. 6, it can be shown that the expected value of $\hat{\mu}_c$ is equal to μ_c, which means that $\hat{\mu}_c$ is an unbiased estimate of μ_c:

$$\begin{aligned} E[\hat{\mu}_c] = \mu_c + \sum_{i \in \Omega_c} \Big\{ &E[k(\Delta z_{ci}^1, \cdots, \Delta z_{ci}^B) f(\Delta z_{ci}^1, \cdots, \Delta z_{ci}^B)] \\ &- E[k(\Delta z_{ci}^1, \cdots, \Delta z_{ci}^B)](\mu_c - \mu_i) \Big\} \\ = \mu_c. & \end{aligned} \tag{13}$$

2 VARIANTS OF THEOREM 2

Theorem 2 shows the unbiasedness of a denoised estimate $\hat{\mu}_c$ via the combination function using the kernel k_i depending on the sub-averages $\Delta z_{ci}^1, \cdots, \Delta z_{ci}^B$, under an assumption that the sub-averages have a symmetric distribution. Depending on which the sub-averages that have a symmetric distribution we set, we can design different types of input-dependent uncorrelated kernels satisfying the unbiased denoising.

2.1 Input-Dependent Kernel on Independent Estimates

In this section, we demonstrate that the combination (Eq. 1) with an input-dependent kernel on the sub-averages of the estimate Δy_{ci}, which are defined as $\Delta y_{ci}^1, \cdots, \Delta y_{ci}^B$, satisfies its unbiasedness under an assumption that the sub-averages have a symmetric distribution:

Theorem 2.1. *Let the kernel k be a bounded function of estimates $\Delta y_{ci}^j, \forall j \in [1, B]$, i.e., $k(\Delta y_{ci}^1, \cdots, \Delta y_{ci}^B)$ satisfying the following conditions:*

$$\begin{aligned} k(\Delta y_{ci}^1 + \alpha, \cdots, \Delta y_{ci}^B + \alpha) &= k(\Delta y_{ci}^1, \cdots, \Delta y_{ci}^B), \\ k(-\Delta y_{ci}^1, \cdots, -\Delta y_{ci}^B) &= k(\Delta y_{ci}^1, \cdots, \Delta y_{ci}^B), \end{aligned} \tag{14}$$

where α is an arbitrary value. By assuming that Δy_{ci}^j has a symmetric distribution, the denoised output $\hat{\mu}_c$ (Eq. 1) is an unbiased estimate of the ground truth μ_c, i.e., $E[\hat{\mu}_c] = \mu_c$.

Proof. We define an average function as f, and Δy_{ci} is represented by the function f that takes the sub-averages $\Delta y_{ci}^1, \cdots, \Delta y_{ci}^B$:

$$\Delta y_{ci} = f(\Delta y_{ci}^1, \cdots, \Delta y_{ci}^B) = \frac{1}{B} \sum_{j=1}^{B} \Delta y_{ci}^j. \tag{15}$$

The sub-averages $\Delta y_{ci}^1, \cdots, \Delta y_{ci}^B$ are independent and identically distributed, and the function f has the following properties:

$$\begin{aligned} f(\Delta y_{ci}^1 + \alpha, \cdots, \Delta y_{ci}^B + \alpha) &= f(\Delta y_{ci}^1, \cdots, \Delta y_{ci}^B) + \alpha, \\ f(-\Delta y_{ci}^1, \cdots, -\Delta y_{ci}^B) &= -f(\Delta y_{ci}^1, \cdots, \Delta y_{ci}^B), \end{aligned} \tag{16}$$

where α is an arbitrary value.

Then, the expected value of a denoised estimate (Eq. 1) using the defined functions f and k is as follows:

$$\begin{aligned} E[\hat{\mu}_c] &= E\left[y_c + \sum_{i \in \Omega_c} k(\Delta y_{ci}^1, \cdots, \Delta y_{ci}^B) \left(\Delta z_{ci} - f(\Delta y_{ci}^1, \cdots, \Delta y_{ci}^B)\right)\right] \\ &= \mu_c + \sum_{i \in \Omega_c} E\left[k(\Delta y_{ci}^1, \cdots, \Delta y_{ci}^B) \left(\Delta z_{ci} - f(\Delta y_{ci}^1, \cdots, \Delta y_{ci}^B)\right)\right] \\ &= \mu_c + \sum_{i \in \Omega_c} \Big\{ E[k(\Delta y_{ci}^1, \cdots, \Delta y_{ci}^B)](\mu_c - \mu_i) \\ &\qquad - E[k(\Delta y_{ci}^1, \cdots, \Delta y_{ci}^B) f(\Delta y_{ci}^1, \cdots, \Delta y_{ci}^B)] \Big\}. \end{aligned} \tag{17}$$

As shown in the proof of Theorem 2 (Sec. 1), given the properties of the functions f and k (i.e., Eqs. 16 and 14) and the condition that the sub-averages $\Delta y_{ci}^1, \cdots, \Delta y_{ci}^B$ are independent and identically distributed variables following a symmetric distribution, two statistics, $f(\Delta y_{ci}^1, \cdots, \Delta y_{ci}^B)$ and $k(\Delta y_{ci}^1, \cdots, \Delta y_{ci}^B)$, are uncorrelated:

$$\begin{aligned} &E[f(\Delta y_{ci}^1, \cdots, \Delta y_{ci}^B)k(\Delta y_{ci}^1, \cdots, \Delta y_{ci}^B)] \\ &= E[f(\Delta y_{ci}^1, \cdots, \Delta y_{ci}^B)]E[k(\Delta y_{ci}^1, \cdots, \Delta y_{ci}^B)] \\ &= (\mu_c - \mu_i)E[k(\Delta y_{ci}^1, \cdots, \Delta y_{ci}^B)]. \end{aligned} \tag{18}$$

By inserting Eq. 18 into Eq. 17, the expected value of the denoised estimate $\hat{\mu}_c$ becomes μ_c.

$$\begin{aligned} E[\hat{\mu}_c] = \mu_c + \sum_{i \in \Omega_c} \Big\{ &E[k(\Delta y_{ci}^1, \cdots, \Delta y_{ci}^B)](\mu_c - \mu_i) \\ &- E[k(\Delta y_{ci}^1, \cdots, \Delta y_{ci}^B)f(\Delta y_{ci}^1, \cdots, \Delta y_{ci}^B)] \Big\} \\ = \mu_c. \end{aligned} \tag{19}$$

Therefore, the denoised estimate $\hat{\mu}_c$ is an unbiased estimate of ground truth μ_c. □

2.2 Input-Dependent Kernel on Both Independent and Correlated Estimates

This section provides another variant of Theorem 2, a denoised estimate (Eq. 1) using an input-dependent kernel on the sub-averages of the estimates $\Delta z_{ci} - \Delta y_{ci}$ (i.e., $\Delta z_{ci}^1 - \Delta y_{ci}^1, \cdots, \Delta z_{ci}^B - \Delta y_{ci}^B$) is an unbiased estimate of ground truth μ_c under an assumption:

Theorem 2.2. *Let the kernel k be a bounded function of estimates $\Delta v_{ci}^j \equiv \Delta z_{ci}^j - \Delta y_{ci}^j, \forall j \in [1, B]$, i.e., $k(\Delta v_{ci}^1, \cdots, \Delta v_{ci}^B)$ satisfying the following conditions:*

$$\begin{aligned} k(\Delta v_{ci}^1 + \alpha, \cdots, \Delta v_{ci}^B + \alpha) = k(\Delta v_{ci}^1, \cdots, \Delta v_{ci}^B), \\ k(-\Delta v_{ci}^1, \cdots, -\Delta v_{ci}^B) = k(\Delta v_{ci}^1, \cdots, \Delta v_{ci}^B), \end{aligned} \tag{20}$$

where α is an arbitrary value. By assuming that Δv_{ci}^j has a symmetric distribution, the denoised output $\hat{\mu}_c$ (Eq. 1) is an unbiased estimate of the ground truth μ_c, i.e., $E[\hat{\mu}_c] = \mu_c$.

Proof. We first express the combination function (Eq. 1) by replacing $\Delta z_{ci} - \Delta y_{ci}$ into Δv_{ci} for a more intuitive representation:

$$\hat{\mu}_c = y_c + \sum_{i \in \Omega_c} k_i \Delta v_{ci}. \tag{21}$$

The estimate Δv_{ci} is the average of the independent and identically distributed sub-averages $\Delta v_{ci}^1, \cdots, \Delta v_{ci}^B$, which can be expressed by using an average function f as below:

$$\Delta v_{ci} = f(\Delta v_{ci}^1, \cdots, \Delta v_{ci}^B) = \frac{1}{B} \sum_{j=1}^{B} \Delta v_{ci}^j. \tag{22}$$

The average function f exhibits the characteristics outlined below:

$$\begin{aligned} f(\Delta v_{ci}^1 + \alpha, \cdots, \Delta v_{ci}^B + \alpha) = f(\Delta v_{ci}^1, \cdots, \Delta v_{ci}^B) + \alpha, \\ f(-\Delta v_{ci}^1, \cdots, -\Delta v_{ci}^B) = -f(\Delta v_{ci}^1, \cdots, \Delta v_{ci}^B), \end{aligned} \tag{23}$$

where α is an arbitrary value.

Subsequently, the expectation of the denoised estimate $\hat{\mu}_c$ introducing two functions f and k is calculated in the following manner:

$$\begin{aligned} E[\hat{\mu}_c] &= E\left[y_c + \sum_{i \in \Omega_c} k(\Delta v_{ci}^1, \cdots, \Delta v_{ci}^B) f(\Delta v_{ci}^1, \cdots, \Delta v_{ci}^B) \right] \\ &= \mu_c + \sum_{i \in \Omega_c} E[k(\Delta v_{ci}^1, \cdots, \Delta v_{ci}^B) f(\Delta v_{ci}^1, \cdots, \Delta v_{ci}^B)]. \end{aligned} \tag{24}$$

In the proof of Theorem 2, it is shown that two statistics from the functions f and k of the sub-averages $\Delta v_{ci}^1, \cdots, \Delta v_{ci}^B$ are statistically uncorrelated, under the conditions where two functions have the aforementioned properties (Eqs. 23 and 20) and the sub-averages are independent and identically distributed variables having a symmetric distribution:

$$\begin{aligned} &E[f(\Delta v_{ci}^1, \cdots, \Delta v_{ci}^B)k(\Delta v_{ci}^1, \cdots, \Delta v_{ci}^B)] \\ &= E[f(\Delta v_{ci}^1, \cdots, \Delta v_{ci}^B)]E[k(\Delta v_{ci}^1, \cdots, \Delta v_{ci}^B)]. \end{aligned} \tag{25}$$

Since the expected value of Δv_{ci} (i.e., $E[f(\Delta v_{ci}^1, \cdots, \Delta v_{ci}^B)]$) is zero, $E[f(\Delta v_{ci}^1, \cdots, \Delta v_{ci}^B)k(\Delta v_{ci}^1, \cdots, \Delta v_{ci}^B)]$ in Eq. 25 becomes zero as well. Thus, the summation in Eq. 24 is zero, so that the expectation of $\hat{\mu}_c$ is equivalent to the ground truth μ_c, indicating an unbiased estimation. □

3 ADDITIONAL COMPARSIONS AT HIGH SAMPLE COUNTS

Fig. 1 illustrates comparisons between our kernel and two input-independent kernels (a uniform kernel and a cross-weighting kernel) when applied to input estimates with high sample counts. It is shown that the input-independent kernels do not effectively reduce some remaining errors inherent in the input estimates (e.g., random noise in path-traced images with independent sampling and structured error in path-traced images sampled with a common random number (CRN)), even when the number of samples for the inputs is high. On the other hand, our kernel exhibits much more effective error reduction with the introduction of a slight bias. It is noteworthy that the assumption made in Theorem 2 (i.e., input estimate having a symmetric distribution) becomes increasingly valid as the sample count rises, resulting in consistent error reduction in the inputs.

REFERENCES

Jonghee Back, Binh-Son Hua, Toshiya Hachisuka, and Bochang Moon. 2020. Deep Combiner for Independent and Correlated Pixel Estimates. *ACM Trans. Graph.* 39, 6, Article 242 (Nov. 2020), 12 pages.

Robert V. Hogg. 1960. Certain Uncorrelated Statistics. *J. Amer. Statist. Assoc.* 55, 290 (1960), 265–267.

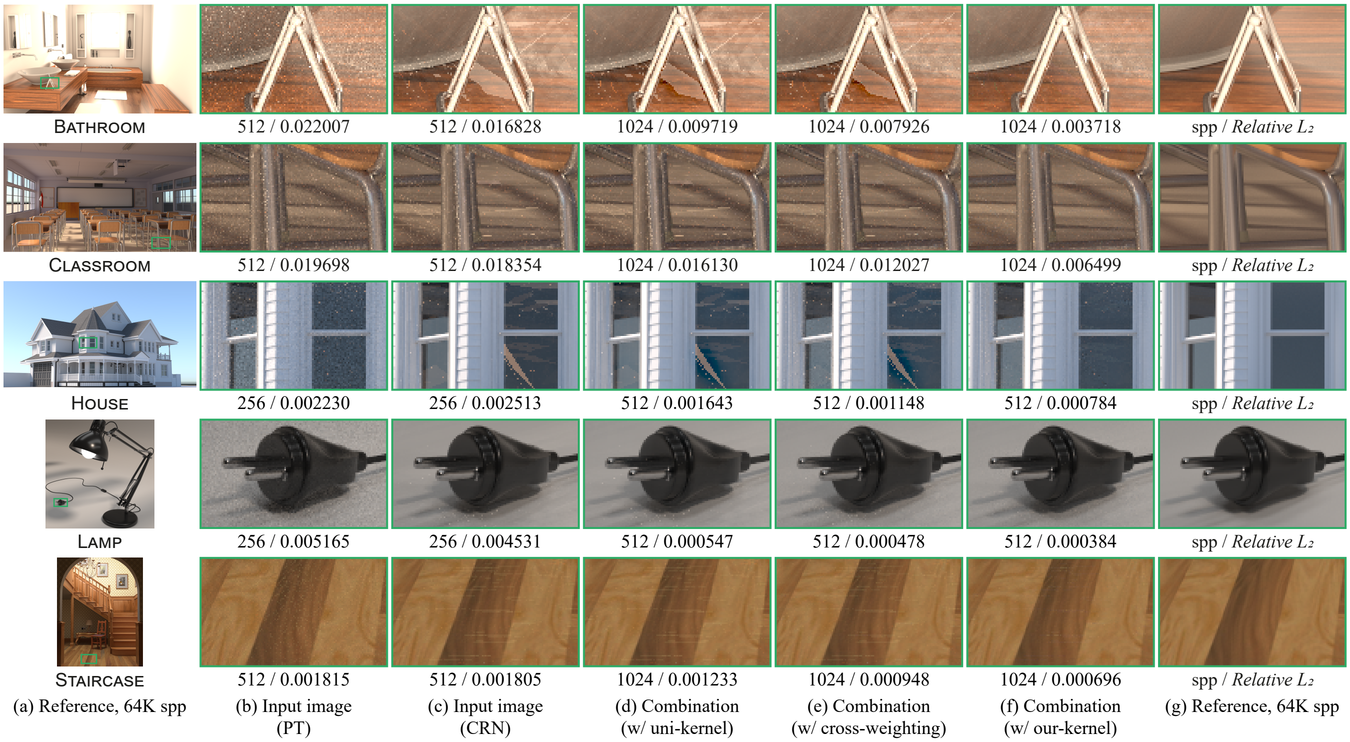

Figure 1: Comparisons with input-independent kernels (a uniform kernel and a cross-weighting kernel) at high sample counts. It can be seen that the input-independent kernels fail to reduce certain errors in the input estimates even with a high number of samples per pixel (e.g., structured artifacts in BATHROOM, HOUSE and STAIRCASE, and spike noise in CLASSROOM and LAMP). Our method effectively mitigates these residual errors.