Nonlinear Noise2Noise for Efficient Monte Carlo Denoiser Training

Andrew Tinits and Stephen Mann, University of Waterloo

SIGGRAPH Asia 2025 Conference Papers

Author Email:
amtinits@uwaterloo.ca

Project Page:
https://tinits.com/nn2n/

Introduction

Learning-based denoisers are trained on pairs of noisy and clean images, but clean images can be hard to obtain. Noise2Noise solves this problem by training on pairs of noisy images [Lehtinen et al. 2018]. For zero-mean noise, the L_2 loss is used to recover the expectation of the noisy targets $\mathbb{E}[\hat{y}]$, which is equal to the clean target y [Lehtinen et al. 2018].

Noise2Noise has a major limitation: **nonlinear functions applied to the noisy targets will skew the results** [Lehtinen et al. 2018]. This bias occurs because the nonlinearity φ makes the expected value of the noisy targets $\mathbb{E}[\varphi(\hat{y})]$ different from the clean target $\varphi(\mathbb{E}[\hat{y}])$ [Lehtinen et al. 2018]. Nonlinear functions are common in image processing, so avoiding them limits the preprocessing that can be performed on the noisy targets.

$$\mathbb{E}[\varphi(\hat{y})] \neq \varphi(\mathbb{E}[\hat{y}])$$

Bounding the Jensen Gap

Our main idea is that we can measure the difference between these means. This difference is known as the *Jensen gap* [Jensen 1906]. By ensuring that the Jensen gap is small, we can **apply nonlinear functions to the noisy targets without adding significant bias to the results**.

We use the method of Liao and Berg [2019] to bound the Jensen gap. This bound is based on the curvature of φ at the clean target y, and the variance of the noisy targets $\hat{y}$:

$$J(y)\mathrm{Var}(\hat{y}) \leq \mathbb{E}[\varphi(\hat{y})] - \varphi(\mathbb{E}[\hat{y}]) \leq J(y)\mathrm{Var}(\hat{y})$$

References

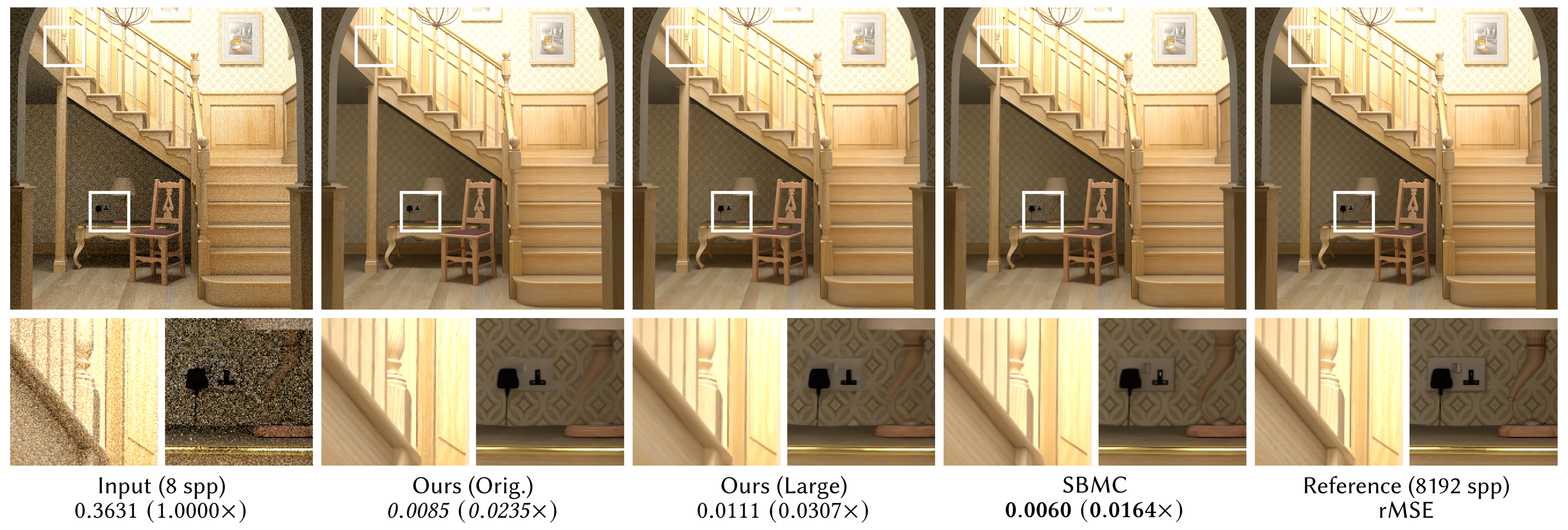
Rajendra Bhatia and Chandler Davis. 2000. A Better Bound on the Variance. The American Mathematical Monthly 107, 4 (April 2000), 353–357. doi:10.1080/00029890. 2000.12005203

Michaël Gharbi, Tzu-Mao Li, Miika Aittala, Jaakko Lehtinen, and Frédo Durand. 2019. Sample-based Monte Carlo Denoising Using a Kernel-splatting Network. ACM Trans. Graph. 38, 4 (July 2019), 125:1–125:12. doi:10.1145/3306346.3322954

J. L. W. V. Jensen. 1906. Sur les fonctions convexes et les inégalités entre les valeurs moyennes. Acta Mathematica 30 (Jan. 1906), 175–193. doi:10.1007/BF02418571

Jaakko Lehtinen, Jacob Munkberg, Jon Hasselgren, Samuli Laine, Tero Karras, Miika Aittala, and Timo Aila. 2018. Noise2Noise: Learning Image Restoration without Clean Data. In International Conference on Machine Learning. PMLR, 2965–2974. http://proceedings.mlr.press/v80/lehtinen18a.html

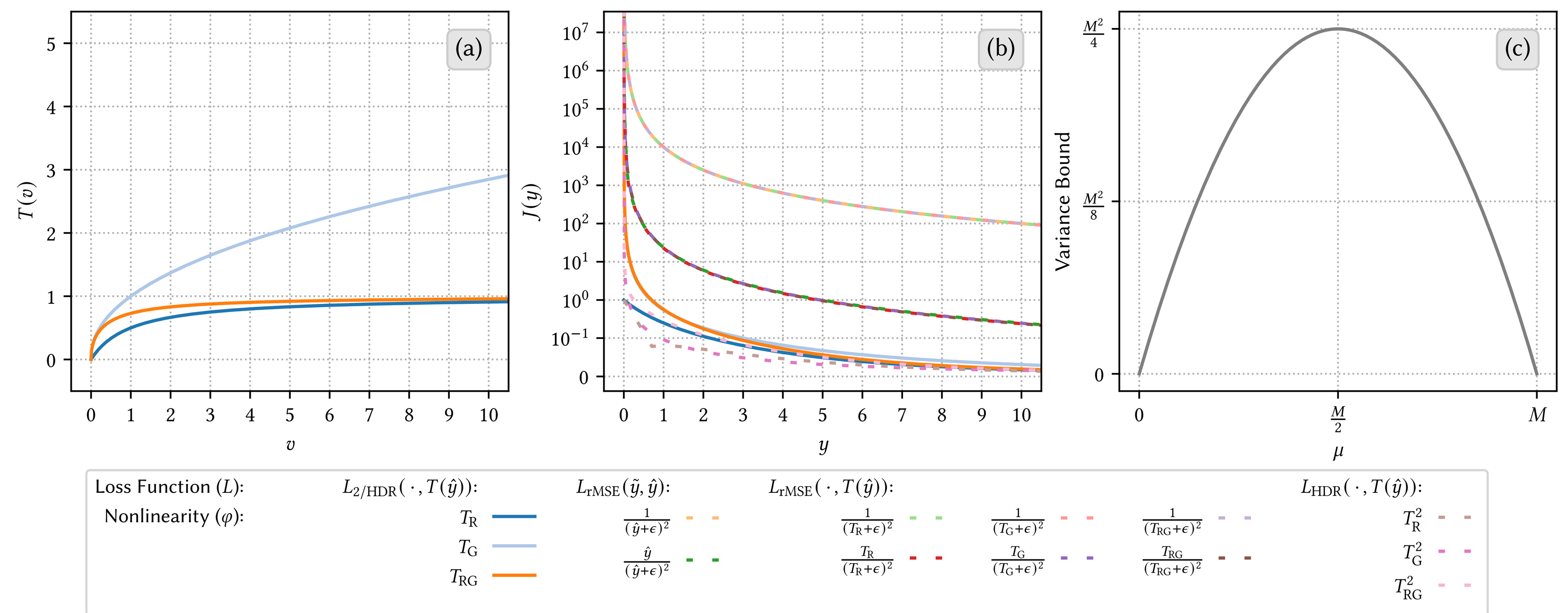
J. G. Liao and Arthur Berg. 2019. Sharpening Jensen's Inequality. The American Statistician 73, 3 (July 2019), 278–281. doi:10.1080/00031305.2017.1419145



A Monte Carlo rendering at 8 samples per pixel (spp), denoised with our models, denoised with the SBMC model, and the reference rendered at 8192 spp. **Our models are trained with only noisy data**, while SBMC is trained with clean references.

Minimizing the Jensen Gap

The bound will be small when the variance is small, but this is hard to achieve in practice. **The bound will also be small when φ is nearly linear around the clean target** y, so we choose φ to meet this requirement. For bounded distributions, the variance will be small near the minimum and maximum [Bhatia and Davis 2000]. φ can therefore have some curvature near these minimum and maximum values.

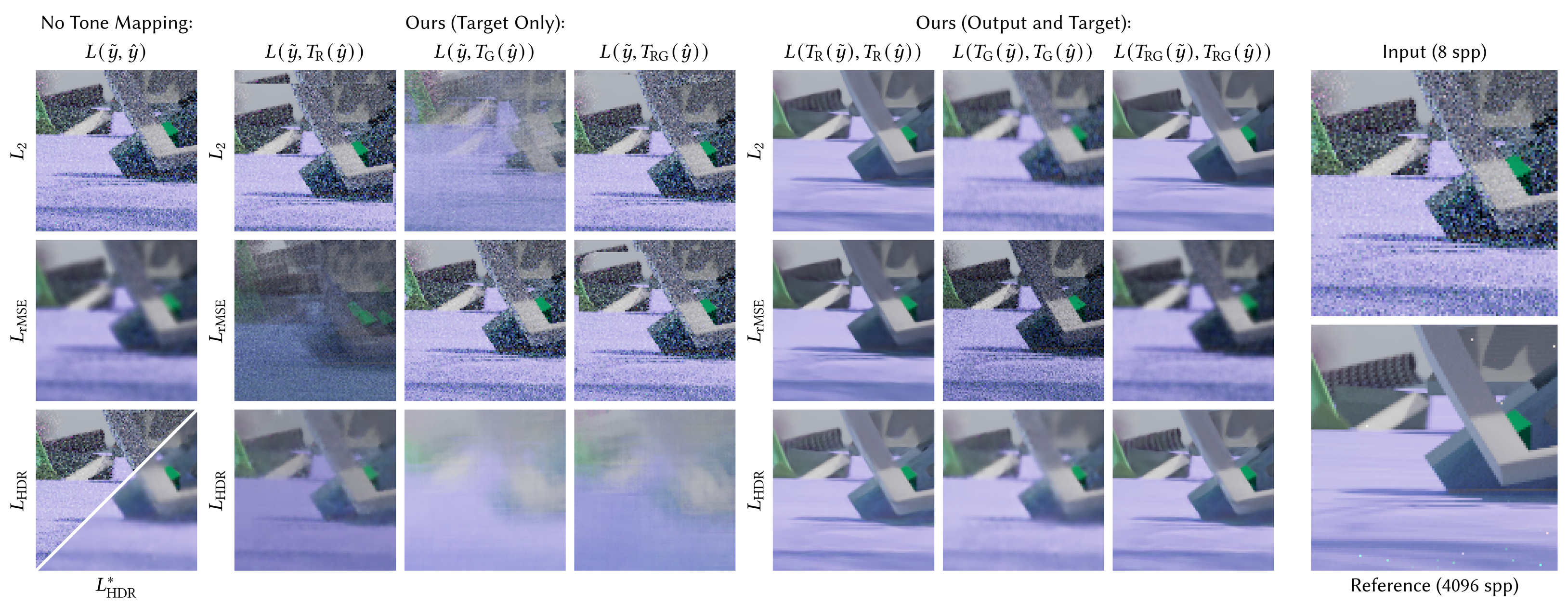


Above, we plot (a) the nonlinear tone mapping functions used for our experimental results, (b) the $J(y)$ functions resulting from combining these tone mapping functions with some L_2-based loss functions, and (c) a graphical representation of the Bhatia-Davis inequality for bounded probability distributions with minimum $m = 0$ and maximum M.

Experimental Results

Model Selection

We evaluated our method on a Monte Carlo denoising task. Noise2Noise can have trouble with high dynamic range (HDR) images, where the L_2 loss is overwhelmed by outliers. We train denoising models using 22 different combinations of loss and tone mapping functions. Our models with tone mapping of both the model output and the target images performed best, **matching our theoretical predictions**.



Comparing to Clean References

We compared our method to training with clean reference images, as represented by the original SBMC implementation [Gharbi et al. 2019]. Our method is represented by the best performing configuration above, and we included a version trained on twice as much noisy data. These models were evaluated on the SBMC test set, which consists of 55 publicly available scenes. **Our models are close to SBMC for all spp in terms of DSSIM**, and are close for relative MSE at high spp, but about an order of magnitude larger at low spp.