IMAGE DENOISING FOR REAL IMAGE SENSORS

Name: Zhang, Jiachao

APPROVED BY:

Keigo Hirakawa, Ph.D.
Advisory Committee Chairman
Assistant Professor, Electrical
and Computer Engineering

Eric Balster, Ph.D.
Committee Member
Professor, Electrical
and Computer Engineering

Kenneth Barnard, Ph.D.
Committee Member
Dr., Sensor Directorate,
Air Force Research Laboratory

John G. Weber, Ph.D.
Associate Dean
School of Engineering

Eddy M. Rojas, Ph.D., M.A., P.E.
Dean, School of Engineering
ABSTRACT

IMAGE DENOISING FOR REAL IMAGE SENSORS

Name: Zhang, Jiachao
University of Dayton

Advisor: Dr. Keigo Hirakawa

This paper describes a study aimed at comparing the real image sensor noise distribution to the models of noise often assumed in image denoising designs. Quantile analysis in pixel, wavelet, and variance stabilization domains reveal that the tails of Poisson, signal-dependent Gaussian, and Poisson-Gaussian models are too short to capture real sensor noise behavior. A new Poisson mixture noise model is proposed in this work to calibrate the mismatch of tail behavior. Based on the fact that noise model mismatch results in image denoising that undersmoothes real sensor data, we offer a new Poisson mixture image denoising scheme to overcome the problem. Experiments with real sensor data verify that the undersmooth is effectively improved.
To my dear advisor
Enlighten me with your profound thoughts

To my dear labmates in ISSL
Encourage me with your selfless supports

To my dear friends at UD
Accompany me with love and happiness
I would like to express my deepest gratitude to my committee chair, Dr. Hirakawa, for his patient guidance and providing me with an excellent atmosphere for doing research. This paper would not have been finished without his persistent help. I would also like to thank my committee members, Professor Asari and Dr. who offer valuable suggestions to this paper. In addition, I would like to thank all members in Intelligent Signal Systems Laboratory for their continuously help during the whole research processing. Finally, special thanks to my family who always encourage and support me.
# TABLE OF CONTENTS

ABSTRACT ................................................................. iii

DEDICATION ............................................................... iv

ACKNOWLEDGEMENTS ......................................................... v

LIST OF FIGURES ............................................................ viii

I. INTRODUCTION ........................................................... 1

II. ANALYSIS OF SENSOR NOISE DISTRIBUTION ......................... 4
  2.1 Sources of Noise ...................................................... 4
  2.2 Heteroscedastic Pixel Noise Models .................................. 5
  2.3 Noise Model Validation In Pixel Domain ............................. 7
     2.3.1 Experiment Setup ............................................... 8
     2.3.2 Quantile Analysis ............................................. 9
  2.4 Noise Model Validation Under Linear Transformation ............ 10
     2.4.1 Discrete Wavelet Transform ................................ 10
     2.4.2 Discrete Cosine Transform ................................. 12
     2.4.3 Multiscale Multiplicative Innovation ....................... 13
  2.5 Noise Model Validation Under Variance Stabilization .......... 16
  2.6 Discussions ...................................................... 17
III. IMPROVED WAVELET NOISE MODELING . . . . . . . . . . . 20
  3.1 Mixture of Poisson . . . . . . . . . . . . . . . . . . . . . . . 20
  3.2 Denoising of Poisson Mixture . . . . . . . . . . . . . . . . . 23
  3.3 Denoising Results . . . . . . . . . . . . . . . . . . . . . . . . 25

IV. CONCLUSION . . . . . . . . . . . . . . . . . . . . . . . . . . . . 30

BIBLIOGRAPHY . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 31
LIST OF FIGURES

1.1 Real vs. simulated sensor noise. The variance of synthetic Poisson noise in (b) matches that of the real sensor data in (a). The denoised image in (d) proves that the Poisson image denoising technique of [1] successfully denoises Poisson corrupted image in (b). The denoised image in (c) suffers from residual artifacts (resembles salt-and-pepper noise), pointing to the discrepancies between the Poisson and real image sensor noise distributions. .............................. 2

2.1 (a) ANOVA over regions A-E ensure uniformity of each Colorchecker patch. (b) Noise variance scales linearly with signal. Red dot shows electronic noise (see text). .................................................. 8

2.2 QQ plot comparing the distribution of the sensor measurements to the noise models of (2.1-2.3). ............................. 15

2.3 QQ plot comparing the distribution of variance stabilized Canon 550D measurements to the noise models of (2.12-2.14). (a) is variance stabilization in pixel domain; (b) and (c) are VS in DWT domain with zero mean and nonzero mean separately; (d) is VS in DCT domain. .......................... 17

3.1 Log-scaled Histogram and QQ plot comparing the distribution of measurements to the noise models of (2.1-2.3) and (3.1) ................. 21
3.2 Denoised results of Poisson mixture image denoising method. A 5-level wavelet transform is performed for denoising. (d) is the difference between $z_1 \cdot \hat{f}_P(h_M; \alpha z_1, \beta)$ and $z_2 \cdot \hat{f}_P(h_M; \alpha z_2, \beta)$ in wavelet domain.

3.3 (a)-(c) are results of [36, 24, 37] denoising algorithms. (d)-(f) are results of Poisson mixture image denoising.
CHAPTER I

INTRODUCTION

Image sensor noise is present in all commercial, professional, and scientific cameras. Measurement noise is quite complex, and an exact analytical form of noise distribution is unknown. Yet, noise distribution model plays a critical part of the post-capture image denoising, aimed at computationally reversing the effects of the image quality degradation caused by noise[2, 3, 4, 5, 6, 7, 1, 8, 9, 10, 11, 12]. The discrepancies between the model of image sensor noise and the actual distribution of real sensor noise acquired by real sensor hardware have profound effects on image denoising.

To illustrate this point, consider Figure 1.1. Taken with Nikon D90 in raw sensor mode in low light, Figure 1.1(a) is the actual 12-bit readout from the image sensor. We compare this to a synthetically generated noisy image in Figure 1.1(b)—the noise distribution is a Poisson, with noise variance in each square matching that of Figure 1.1(a). Under visual inspection, there are no obvious differences between Figures 1.1(a) and 1.1(b). However, the differences between the Poisson denoised images in Figures 1.1(c) and 1.1(d) are stark. With no discrepancy between the synthetic noise model in Figure 1.1(b) and the image denoising algorithm in [1], the output image shown in Figure 1.1(d) was satisfactory (proving that denoising method performs as designed). However, the same denoising method applied to the real sensor data in Figure 1.1(a) resulted in residual artifacts resembling salt-and-pepper noise. We assure readers that these are not defective pixels—no such dots are evidenced when
an image is taken in ample light. Refer to Chapter 2.3 for the full experiment setup.

We attribute the differences between Figures 1.1(c) and 1.1(d) to the failure in the image denoising method to capture the characteristics of real sensor noise. Specifically, denoising method in [1] was designed to estimate the pixel intensity of the Poisson count variables. The mismatch between Poisson distribution and the actual image sensor noise proved detrimental to image denoising. We conclude that simulated evaluation of image denoising algorithms does not necessarily reflect the actual performance for real imaging applications.
The goal of this paper is two fold. In Chapter II we rigorously investigate discrepancies between the model of image sensor noise commonly used in image denoising algorithms and the distribution of real sensor noise acquired by real sensor hardware.\footnote{We extend on our preliminary findings reported in [13, ?] by incorporating more camera settings and noise models.} There are two important differences between the model validation study we describe in this paper and the many existing empirical studies [14, 15, 16, 17, 18, 19, 20]. First, existing studies overwhelmingly focused on the relationship between the pixel intensity and noise variance, but there has been little emphasis on the actual distribution of the noise. As the tail behavior of the noise distribution greatly influences denoising performance, we provide detailed analysis of this. Second, most modern image denoising techniques incorporate linear and nonlinear transformations that give rise to energy compaction and sparse signal representation. Therefore we focus our investigation on the accuracy of noise model in the transform domain rather than the noise distribution in the pixel intensity domain.

In Section 3.1 we propose modifications to the canonical noise models to better approximate the distribution of real sensor noise leveraging the intuition we gained from analysis in Chapter II. Based on this improved noise model, we make similar modifications to existing image denoising schemes to improve their denoising performance in real sensor data. Though the modifications are simple, they are surprisingly effective and cost effective.
CHAPTER II

ANALYSIS OF SENSOR NOISE DISTRIBUTION

2.1 SOURCES OF NOISE

Noise falls typically into three categories: fixed pattern noise, banding noise and random noise. *Fixed pattern* noise refers to pixel deviations that has almost the same distribution under various imaging conditions, often stemming from manufacturing variabilities of pixel circuitry. Similarly, *banding noise* is a column-shaped variability that emerges as a result of biases in the bank of A/D converters that read the a few columns of sensor data in parallel. While fixed pattern noise and banding noise can be reduced to a degree with stricter manufacturing tolerances, the repeatable and predictable noise pattern can also be calibrated and largely corrected in post-capture processing.

By *random noise* we mean non-repeatable variabilities in sensor measurements. As described by quantum electrodynamics, electron-positron annihilation stemming from light interacting with matter result in refractions and specular reflections. The distribution of this photon emission is referred to as *shot noise* and is classically modeled as Poisson [21]; and the number photons accumulated during sensor integration is therefore stochastic.

Pixel sensors make measurements on light intensity by a process known as *photon transfer* or *photon recapture*. Photodiode is a type of integrating detector that con-
verts photons into current by *pair production* (momentum of one photon energy is absorbed to produce one electron-hole pair) and by *photoelectric effect* (photon energy breaks covalent bond, causing lattice vibration). Though the resultant photocurrent is proportional to the photon energy *on average*, the measurement is stochastic due to the random nature of the photoelectric effects. The canonical measure of stochasticity is Fano factor[22], or a ratio $\kappa$ of noise variance over photon count—$\kappa = 0$ corresponds to the deterministic process of electron-hole pair production; $\kappa = 1$ denotes Poisson distribution, etc. The Fano factor for silicon (and by extension, Fano factor for CMOS sensor) is $\kappa = 0.12$, meaning that photon transfer in a typical image sensor results in heteroskedastic noise.

Besides the noise associated with photon emission and recapture, various sources of electrical noise exist on the detector circuit. The dark current stemming from in-circuit electron excitation and source flower contribute to *thermal noise*, whose power is proportional to the exposure time but is typically independent of the signal strength. *Reset noise* is caused by the improperly discharged capacitors, where measurement biases are introduced by the electrical charges previously held by the capacitors. Rounding and truncation during digitization of analog voltage result in *quantization error*.

### 2.2 HETEROSCEDASTIC PIXEL NOISE MODELS

Modeling each source of noise separately is a daunting task (if not impossible) since phenomenon such as Fano noise are not well understood. It is also challenging to isolate one source of noise from another in hardware, making it difficult to validate each model. Characterizing the overall noise is not only more pragmatic but also has sensible justifications as well. When central limit theorem and law of small numbers
are in force, combined noise approach canonical noise distributions such as normal and Poisson. Unsurprisingly, additive white Gaussian noise (AWGN), Poisson, and Gaussian-Poisson are three of the most well-studied image denoising problems in the literature [14, 15, 16, 17, 18, 19, 20] (as we demonstrated in the example of Figure 1.1). These models provide the basis for image denoising methods that extend the imaging device’s capabilities.

In Figure 2.1(b) signal strength is plotted against the variance of sensor measurements. This figure presents a convincing evidence that noise variance scales linearly with signal strength. The linearity is usually attributed to the Poisson process of the photon emission in the literature [9, 14, 15, 16, 6, 7, 1, 8, 10, 11, 12, 23]. Under this scenario, a sensor observation $h$ is modeled as $h_P$, an affine transformed Poisson count data $g_P$:

$$h_P := \alpha \cdot g_P + \beta, \quad g_P \sim \mathcal{P}(f_P), \quad (2.1)$$

where $f_P$ is the latent intensity, and the subscript $P$ denotes Poisson-based model. The parameters $\alpha$ and $\beta$ are learned from regressing signal strength and variance in Figure 2.1(b).

Another way to capture the signal dependence of noise is to couple the variance of a normal random variable[20]:

$$h_G \sim \mathcal{N}(\alpha f_G + \beta, \alpha^2 f_G). \quad (2.2)$$

where the subscript $G$ denotes signal-dependent Gaussian noise model ((2.2) is the normal approximation of $h_P$).
Poisson-Gaussian hybrid model treats signal-dependent and signal-independent noises separately [24, 18, 2, 3]. The observation $h_H = \alpha g_S + g_C$ is a combination of signal $g_S$ and circuit $g_C$ noise:

$$g_S \sim \mathcal{P}(f_S), \quad g_C \sim \mathcal{N}(\mu_C, \sigma_C^2),$$  \hspace{1cm} (2.3)

where $(\mu_C, \sigma_C^2)$ is the signal-variance pair of the electrical noise (i.e. sensor noise when incoming light is blocked; indicated by red dot in Figure 2.1(b)). Subscript $H$ denotes hybrid model.

### 2.3 NOISE MODEL VALIDATION IN PIXEL DOMAIN

Poisson, signal-dependent Gaussian, and Poisson-Gaussian distributions in (2.1)-(2.3) faithfully model the linear coupling between the signal strength and the noise variance in image sensors described in Figure 2.1(b). The failed denoising experiment in Figure 1.1 suggests that there are profound differences between the models and real sensor noise besides the signal-variance coupling, however. To investigate this discrepancy in a systematic manner, we conduct experiments aimed at understanding the impact of noise models on image denoising. In the remainder of Chapter 2.3 we provide a quantile analysis of noise in the pixel domain. In Chapter 2.4 we then modify the experiment to understand the impact of the noise models in wavelet, discrete cosine (DCT), and multiscale multiplicative innovation (MMI) domains—meaningful because contemporary image denoising methods operate on the transform coefficients. In Section 2.5 we also repeat the experiment for variance stabilization transforms, an alternative to heteroskedastic noise models in (2.1)-(2.3) .
2.3.1 Experiment Setup

We obtained samples of noisy sensor data by capturing X-Rite ColorChecker (Figure 2.1(a)) at 1.2 lux using Nikon D90, Canon 550D, and Fuji Pro 1 in raw sensor format with 1/200 second exposure (1/250 for Fuji), ISO 200, and f/4.5 (4 for Fuji). Color filter array sampled data is partitioned into red, blue, and green measurements, which are treated separately (i.e. $3 \times 24 = 72$ ColorChecker patches). Under an ideal scenario, measured pixel component values from the same ColorChecker patch are drawn from the same probability distributions.

Despite our best efforts, however, uneven lighting, vignetting, and camera angle introduce additional variabilities. For this reason, we detect non-uniformity of ColorChecker patches by the analysis of variance (ANOVA) over five $10 \times 10$ regions (labeled A-E in Figure 2.1(a)) cropped from each ColorChecker patch. Any ColorChecker patch that rejects the null hypothesis (i.e. means of A-E are equal) at the 99% confidence level is removed from the experiment. We also removed green pixels from blue-green rows because color crosstalk contaminations affect green pixels in
red-green and the blue-green rows differently[25]. Each accepted ColorChecker patch has over 25,000 samples.

We also acquired another image (with the same camera settings) with a cap placed over the lens. By blocking the incoming light, this “blank” image offers an indication for the circuit noise that is independent of the signal strength.

2.3.2 Quantile Analysis

The noise models above are heuristic approximations at best, and model discrepancies deteriorate the image denoising performance. We employ quantile analysis to robustly compare the distribution of sensor data and the noise model. Consider a parametric curve of the form [26]:

\[
x(t) = F_{\text{data}}^{-1}(t), \quad y(t) = F_{\text{model}}^{-1}(t)
\]

where \( F_{\text{data}}, F_{\text{model}} : \mathbb{R} \to [0, 1] \) are the cumulative distribution functions of data and model distribution, respectively. This so-called quantile-quantile plot (QQ plot) lies on the 45° line if the empirical data and noise model distributions are well-matched. If the data variables are found to be an affine transformation of the model variables, then the QQ plot forms an affine line as well (but not necessarily on 45° line with zero intercept). QQ plot is useful for detecting the deviations of the model from data, particularly in the tails of distributions where the samples are sparse.

The QQ plots shown in Figure 2.2(a) compare the distribution of the measured pixel data within a ColorChecker patch against the models in (2.1-2.3). For \( F_{\text{data}}(h) \), the empirical histogram of the pixels measured within each ColorChecker patch was used.
For $F_{\text{model}}(h)$, we derived the model parameters for each ColorChecker patch by:

$$f_P = f_G = \frac{\mu_H - \beta}{\alpha}, \quad f_S = \frac{\mu_H - \mu_C}{\alpha},$$

(2.5)

where $\mu_H$ is the sample mean of each ColorChecker patch, and $\alpha$, $\beta$ and $\mu_C$ are as described in Section 2.2.

Each QQ plot describes the variation of real and model noise within one ColorChecker patch only.\footnote{We cannot infer signal-noise dependence from the QQ plots as noise samples in each plot are drawn from the same distribution with same mean/variance. Only representative examples shown due to page limit.} As evidenced by the $45^\circ$ line formed by a portion of the QQ plot, the noise model is accurate near the median. However, sensor measurements are clearly more heavy tailed than the model (Nikon D90’s short negative tail is likely due to saturation).

### 2.4 NOISE MODEL VALIDATION UNDER LINEAR TRANSFORMATION

#### 2.4.1 Discrete Wavelet Transform

Though we are interested in the impact of linear transforms on the noise, quantile analysis in the transform domain is challenging. With the precise form of real noise distribution unknown, an analytical form of noise distribution in discrete wavelet (DWT) and cosine (DCT) transform domains cannot be derived. Yet, unlike the pixel domains, it is difficult to obtain a large number of real noisy DWT/DCT coefficients drawn from the same distribution. We rejected the idea to take the raw sensor data from a video sequence of a scene—although applying DWT/DCT to each frame yields a large number of noisy coefficients, noise is unnaturally coupled with the temporal
hysteresis of reset noise.

We developed a new strategy to obtain a large number of noisy coefficients from the ColorChecker image. Consider Haar wavelet transform (HWT)—finest level noisy wavelet \( w(n) \) and scaling \( s(n) \) coefficients at location \( n \) are

\[
\begin{align*}
    w(n) &= h(2n) - h(2n + 1) \\
    s(n) &= h(2n) + h(2n + 1)
\end{align*}
\]

where \( h(2n) \) and \( h(2n + 1) \) are neighboring pixels. Thanks to sparsity, the majority of DWT coefficients have zero mean:

\[
E(w(n)) = 0 \iff E(h(2n)) = E(h(2n + 1)).
\]  \hfill (2.7)

Hence \( h(2n) \) and \( h(2n + 1) \) are assumed to be drawn from the same distribution. We obtain a large number of \( w(n) \) samples corresponding to a mean zero DWT coefficient by taking a difference between two observed samples drawn at random from the same ColorChecker patch. By contrast, coefficients with nontrivial mean have the property:

\[
E(w(n)) \neq 0 \iff E(h(2n)) \neq E(h(2n + 1)).
\]  \hfill (2.8)

Hence \( h(2n) \) and \( h(2n + 1) \) are drawn from a different distributions. We may obtain a large number of \( w(n) \) samples by taking a difference between two noisy samples drawn at random from two predesignated ColorChecker patches.

For analysis, \( F_{data}(w) \) was computed from the noisy DWT coefficients obtained by the above scheme. For \( F_{model}(w) \), the DWT noise model derived from (2.1-2.3) have
the form:

\[
\begin{align*}
\alpha^{-1} w_P &\sim \text{Skellam}(f_P, f'_P) \\
w_G &\sim \mathcal{N}(\alpha(f_G - f'_G), \alpha^2(f_G + f'_G)) \\
w_H &= \alpha w_S + w_C, \quad \begin{cases} \\
w_S &\sim \text{Skellam}(f_S, f'_S) \\
w_C &\sim \mathcal{N}(0, 2\sigma_C^2). \\
\end{cases}
\end{align*}
\]

(2.9)

where \(\{f_P, f_G, f_S\}\) and \(\{f'_P, f'_G, f'_S\}\) are parameters derived from ColorChecker patches corresponding to \(h(2n)\) and \(h(2n + 1)\), respectively.\(^2\) The QQ plots shown in Figure 2.2(b-c) compare the distribution of empirical DWT coefficients against their models. Though noise models are accurate near the median, the models clearly shorten the tails.

2.4.2 Discrete Cosine Transform

DCT is defined for \(k \in \{0, 1, ... N - 1\}\) as follows [27]:

\[
d(k) = \sum_{n=0}^{N-1} \frac{h(n)}{\sqrt{N}} \cos \left( \frac{\pi}{2N} (2n - 1)k \right). \quad (2.10)
\]

2D DCT applies (2.10) to horizontal and vertical directions. For quantile analysis in DCT domain, \textit{randomly drawn} sensor measurement samples from each of \(N\) predesignated ColorChecker patches \((N^2\text{ patches for 2D DCT})\) are respectively assigned to \(\{h(0), \ldots, h(N - 1)\}\) to yield a single DCT coefficient \(d(k)\) via (2.10). Repeating this experiment yields a large number of DCT coefficients drawn from the same distribution and \(F_{\text{data}}(d)\). For \(F_{\text{model}}(d)\), \textit{randomly drawn} samples of \(h_P\) in (2.1) from the \textit{models of \(N\) predesignated ColorChecker patches} are assigned to \(\{h(0), \ldots, h(N - 1)\}\)

\(^2\text{Skellam}(f_P, f_P)\) is known as Irwin distribution.
to compute DCT coefficient $d_P(k)$. We followed the same procedure to yield $d_G(k)$ and $d_H(k)$ from (2.2-2.3).

The resultant QQ plot in Figure 2.2(d) suggests that the distribution of DCT coefficients stemming from the real sensor data is well approximated by the models. The improved match is likely due to central limit theorem, which is in force as a result of DCT in (2.10) taking weighted average of measurements $\{h(0), \ldots, h(N-1)\}$. When $N$ is small, the deviation of tails is still observable (not shown).

2.4.3 Multiscale Multiplicative Innovation

Another commonly used transformation Multiplicative Innovation (MMI)[28] works grounded on Haar wavelet transform. The model is established as:

$$
\frac{h_P - \beta}{\alpha} | s_P \sim Binomial\left( \frac{s_P - 2 \beta}{\alpha}, \frac{f_P - f'_P}{2(f_P + f'_P)} \right)
$$

(2.11)

where $f_P$ and $f'_P$ are derived as in Section 2.4.1. For Quantile analysis in MMI domain, to start with sensor measurement samples are randomly drawn from two selected Colorchecker patches and specified as $h(2n)$ and $h(2n+1)$. One pair of $w(n)$ and $s(n)$ are attained via (2.6). If the scaling coefficient $s(n)$ is an approximation to $c$ (ground truth of scaling coefficient), then the paired wavelet coefficient $w(n)$ is assigned as one MMI coefficient $\theta(k)$. Repeating this experiment yields substantive MMI coefficients and $F_{data}(\theta)$. For $F_{model}(\theta)$, process Haar wavelet transform for $h_P$ in (2.1) from the models of two preselected patches and operate approximate comparison of scaling coefficient with $c$ to yield MMI coefficient $\theta_P(k)$. The $\theta_G(k)$ and $\theta_H(k)$ from (2.2-2.3) are generated along with the same procedure.

The QQ plots shown in figure 2.2(e) indicate that the distribution of MMI coefficients derived from model is well matched to measurement data. The multiscale Bayesian
estimator, which resembles filtration to HWT, evidently attenuate the mismatch of tails.
Figure 2.2: QQ plot comparing the distribution of the sensor measurements to the noise models of (2.1-2.3).
2.5 NOISE MODEL VALIDATION UNDER VARIANCE STABILIZATION

Variance stabilization (VS) is an invertible function that recovers homoscedasticity given heteroscedastic noise. VS is often combined with additive white Gaussian noise (AWGN) image denoising to address camera sensor noise. Bartlett/Anscombe VS transforms Poisson counts into a normal variable [29, 30]:

\[ \eta\{h_P\} = 2\sqrt{(h_P - \beta)/\alpha + k} \sim \mathcal{N}(2\sqrt{f_P + k}, 1), \]  

(2.12)

where the constant value is \( k = 1/2 \) for Bartlett and \( k = 3/8 \) for Anscombe. Poisson-Gaussian hybrid \( h_H \) in (2.3) is also stabilized by the generalized Anscombe transform [2, 3]:

\[ \eta'\{h_H\} = 2\sqrt{(h_H - \mu_C)/\alpha + 3/8 + \sigma_C^2/\alpha^2}. \]  

(2.13)

However, one can prove \( \eta\{h\} \equiv \eta'\{h\} \) by the fact that \( \sigma_C^2 = \alpha\mu_C - \alpha\beta \) must hold in Figure 2.1(b). Haar-Fisz (HF) transform is a more contemporary VS treatment that modify Haar wavelet (\( w \)) and scaling (\( s \)) coefficients [31]:

\[ \gamma\{h_P\} = \frac{w_P}{\sqrt{\alpha(s_P - 2\beta)}} \sim \mathcal{N}\left(\frac{f_P - f'_P}{\sqrt{f_P + f'_P}}, 1\right). \]  

(2.14)

This procedure is repeated for coarser level wavelet representations.

Figure 2.3 compares the normal distribution models of (2.12-2.14) against the distributions of variance stabilized sensor measurements \( \eta\{h\} \) and \( \gamma\{h\} \) (and their DWT/DCT). Although Anscombe transform stabilized noise variance, the tails of empirical VS coefficients \( \eta\{h\} \) are clearly longer than normality. The tails of the
Empirical variance stabilized DWT coefficients also deviate from normal probability. By contrast, HF stabilized coefficients $\gamma\{h\}$ appear to be normally distributed, as evidenced by straight QQ plot line. However, the gentle slope of QQ plot suggests that data variance is smaller than 1. By comparison, QQ plot followed the 45° line very closely when scene was well lit (not shown). We conclude that HF VS succeeds in “Gaussianizing” the observed data, but it did not achieve homoscedasticity (i.e. variance depends on signal intensity).

### 2.6 DISCUSSIONS

Poisson, signal-dependent Gaussian, and Poisson-Gaussian distributions are commonly used to model the linear coupling between the signal strength and the noise variance in image sensors. Our quantile analysis definitively proved that the tails of the noise models are too short to describe the actual distribution of the measurement.
noise. The trends we described in this paper are common among a variety of camera manufacturers, red/green/blue color components, and camera settings.

Our study is not without limitations. In Sections 2.4.1 and 2.4.2, random sampling of the pixels within Colorchecker patch were linearly combined to yield a large number of empirical DWT and DCT coefficients. This procedure is only valid if the random phenomena occurring in the spatially neighboring pixels are independent. Thermal noise, for example, is not always spatially white since electron leakage affects neighboring pixels. Hence the computed DWT/DCT coefficient noise in our study is noisier than the actual coefficients computed from an image sensor with significant leak. However, the main conclusions of this work—that the tail behavior of sensor noise is heavier than the models—remains valid. Our lab’s capabilities today do not allow for measurements with integrating spheres, which guarantees uniformity of the scene beyond our current ANOVA testing. Most commercial cameras have safety features that prevent pictures from being taken while the camera optics are removed, increasing the risks of vignetting.

Recall that models in in (2.1)-(2.3) provide the basis for image denoising methods that extend the imaging device’s capabilities. What practical impact does the model mismatch play in image denoising? How should image denoising methods be improved for handling real image sensor data? Most modern image denoising methods operate in (linear) transform domain. However, we showed that the noise models also fail in Haar wavelet domain, where the model tails insufficiently account for large noise coefficients (i.e. heavy tail). The practical impact of the model mismatch is the undersmoothing of noise—when denoising algorithms designed with (2.1-2.3) in mind are applied to real sensor measurements, a large DWT coefficient is incorrectly attributed to the signal since noise model does not account for it. Indeed, the artifacts
evidenced in Figure 1.1(c) stem from the DWT coefficients inappropriately preserved by the wavelet-based denoising method in [32].

An alternatives approach is to combine VS transforms with AWGN image denoising. We concluded earlier that while Anscombe VS achieves homoscedasticity (i.e. noise variance decoupled from signal strength), the overall distribution profile is far from normality. In working with Anscombe VS, one would need to extend the noise model tails (e.g. scale mixture of AWGN) in order to improve denoising performance. On the other hand, Haar-Fisz VS successfully transformed sensor data into normal random variables, but the claim of the homoscedasticity could not be substantiated. Indeed, one is at risk of oversmoothing noise with conventional AWGN denoising method used in conjunction with Haar-Fisz VS, since noise variance in low light regions is lower than 1. In order that denoising work more effectively in conjunction with Haar-Fisz, one must first determine the true relationship between signal strength and noise variance in the VS transformed domain.

Model mismatch is less significant in DCT domain and MMI domain. It suggests that DCT-based and MMI-based denoising in simulation and with real image sensor data are likely to agree. Nevertheless, the analysis of noise heteroscedasticity in the DCT domain is far more complicated than that of wavelet noise, and DCT-based denoising without VS is impractical. The combination of Anscombe VS and DCT (as was investigated in [2, 3], for example) is the only scheme in our investigation that satisfactorily yielded a homoscedastic model matching the observation distribution.
CHAPTER III

IMPROVED WAVELET NOISE MODELING

3.1 MIXTURE OF POISSON

One of the main conclusions reached in the previous section is that the noise models fail in Haar wavelet domain, where the model tails insufficiently account for large noise coefficients. This discrepancy results in undersmoothing of noise, as already evidenced also in Figure 1.1(c). Although recent work in DWT-based denoising (such as [33, 34]) have reportedly outperformed DCT-based solutions (such as [2, 3]) in simulation, the mismatch between real image sensor noise in and their models have to be resolved in order for these methods to be useful in real-world low light imaging.

In this section, we develop a simple technique to “correct” the short-tailed DWT noise model. We draw inspirations from prior work in [35] where mixture of canonical distributions (e.g. normal, Laplace, binomial) have successfully modeled the heavy-tailed DWT coefficients corresponding to the image signal (i.e. not noise distribution). Leveraging this idea, we propose a Poisson mixture model to make heavier-tailed DWT coefficient distribution of noise (i.e. instead of image signal). Consider a Poisson mixture of the form:

\[ h_M := \alpha \cdot z \cdot g_M + \beta \]  
\[ g_M \mid z \sim \mathcal{P} \left( \frac{f_M}{z} \right). \]
Figure 3.1: Log-scaled Histogram and QQ plot comparing the distribution of measurements to the noise models of (2.1-2.3) and (3.1).
where \( f_M = f_P = f_G \) is as defined in (2.5); and \( z \) is a hidden Bernoulli variable with known camera-specific constants \( z_1 \) and \( z_2 \):

\[
P[z = z_1] = 1 - P[z = z_2] = \pi. \tag{3.3}
\]

The probability \( \pi \) is determined entirely by \( z_1 \) and \( z_2 \), since the linear coupling between the signal strength and the noise variance in image sensors must agree with Figure 2.1(b). Simple algebraic manipulation yields the relation:

\[
\pi = \frac{1 - z_2}{z_1 - z_2}. \tag{3.4}
\]

The DWT noise model corresponding to the Poisson mixture model of (3.1) is a Skellam mixture:

\[
\alpha^{-1} z^{-1} w_M \bigg| z \sim Skellam \left( \frac{f_M}{z}, \frac{f'_M}{z} \right). \tag{3.5}
\]

The significance of the DWT noise model in (3.5) is that Skellam mixture is a heavy-tailed distribution. For each camera model, we trained \( z_1 \) and \( z_2 \) to best match the empirical DWT noise distribution of the sensor measurements in the mean square error sense. One can verify from Figure 3.1 that the Skellam mixture indeed approximates real image sensor noise very closely, including the tail behavior. The proposed DWT noise model\(^1\) in (3.1) is clearly in better agreement with the empirical DWT noise than the existing alternatives in (2.1)-(2.3).

\(^1\)Note also the slight impreciseness of (3.1). Owing to the fact that Poisson variables are integer valued, \( h_M \) technically lives in the union of the cosets \( \Lambda = \{ \alpha z_1 \mathbb{Z} + \beta \} \cup \{ \alpha z_2 \mathbb{Z} + \beta \} \). However, there is no physical significance to \( \Lambda \)—it is rather a mathematical consequence of the approximation to real image sensor we made in (3.1) that behaves more like quantization error.
3.2 DE NOISING OF POISSON MIXTURE

Motivated by the Poisson mixture noise model in (3.1), we develop an image denoising technique designed for real image sensors. The objective of the denoising function is to estimate the rate $f_M$ based on the observations $h_M$, which may be accomplished indirectly by wavelet transform. In the Bayesian paradigm, the following relation holds:

$$\tilde{f}_M(h_M|\alpha, \beta) := \mathbb{E}[f_M|h_M]$$

$$= P[z = z_1|h_M] \cdot \mathbb{E}[f_M|h_M, z = z_1] + P[z = z_2|h_M] \cdot \mathbb{E}[f_M|h_M, z = z_2]$$

(similar relationship holds in the wavelet domain as well). That is, the minimum mean squared error estimate $\tilde{f}_M(h_M; \alpha, \beta)$ is a convex combination of the conditionally minimum mean squared error estimates $\mathbb{E}[f_M|h_M, z = z_1]$ and $\mathbb{E}[f_M|h_M, z = z_2]$. The convex combination weights are determined by the posterior Bernoulli probability

$$\hat{\pi}(h_M) := P[z = z_1|h_M] = 1 - P[z = z_2|h_M]$$

$$= \frac{\pi P[h_M|z = z_1]}{\pi P[h_M|z = z_1] + (1 - \pi) P[h_M|z = z_2]}.$$ 

(3.7)

We gain additional intuition from rewriting (3.6) in the following manner:

$$\tilde{f}_M(h_M; \alpha, \beta) = \hat{\pi}(h_M) \cdot z_1 \cdot \mathbb{E} \left[ \frac{f_M}{z_1} \bigg| h_M, z = z_1 \right] + (1 - \hat{\pi}(h_M)) \cdot z_2 \cdot \mathbb{E} \left[ \frac{f_M}{z_2} \bigg| h_M, z = z_2 \right].$$

(3.8)

Recalling (3.2) and noting that $f_M/ z = \mathbb{E}[g_M|z]$, the minimum mean squared error estimate $\mathbb{E} \left[ \frac{f_M}{z} \bigg| h, z \right]$ that appears in (3.8) is actually a Poisson image denoising method.
\( \hat{f}_P(h; \alpha z, \beta) \) corresponding to the noise model in (2.1) (note that \( \alpha \) was replaced by \( \alpha z \)). Fortunately, it implies that existing state-of-art Poisson image denoising methods can be used in a modified manner in real-world application by following the steps:

1. perform Poisson image denoising: \( \hat{f}_P(h; \alpha z_1, \beta) \)
2. perform Poisson image denoising: \( \hat{f}_P(h; \alpha z_2, \beta) \)
3. compute weights \( \hat{\pi}(h_M) \) via (3.7)
4. combine:

\[
\hat{f}_M(h_M; \alpha, \beta) = \hat{\pi}(h_M) \cdot z_1 \cdot \hat{f}_P(h_M; \alpha z_1, \beta) + (1 - \hat{\pi}(h_M)) \cdot z_2 \cdot \hat{f}_P(h_M; \alpha z_2, \beta).
\]

(3.9)

We conclude that “Poisson mixture image denoising” problem can be solved by “a mixture of Poisson image denoising.”

The solution in (3.9) has an intuitive non-Bayesian interpretation as well. The noise model of (3.1) informs us that the pixel values are conditionally Poisson counts. If the hidden Bernoulli variable \( z \) is known, then the Poisson image denoising method of \( z \cdot \hat{f}_P(h; \alpha z, \beta) \) is undoubtedly a “good” solution, regardless of whether \( \hat{f}_P \) is a Bayesian or non-Bayesian. From this perspective, \( \hat{\pi}(h_M) \) in (3.9) plays the role of a hypothesis testing, where the magnitude of \( \hat{\pi}(h_M) \) increases with our confidence of the event \( z = z_1 \).

We illustrate this point by example. Figure 3.2 show the result of Poisson denoising algorithms \( z \hat{f}_P(h; \alpha z, \beta) \) using the method in [1]. Neither of these solutions are entirely satisfactory. Denoising result with \( z = z_1 \) shown in Figure 3.2(b) preserves image de-
tails such as edges and textures (see house), but the undersmoothing is unacceptable in homogeneous regions. The impulse-noise-like artifacts evidenced in Figure 1.1(c) is also present here as well (although more obvious in the homogeneous regions, the artifacts are also present on house). On the other hand, although denoising results with \( z = z_2 \) shown in Figure 3.2(c) oversmoothes edges and textures, the absence of impulse-noise-like artifacts gives additional credibility to the Poisson mixture model in (3.1).

A sensible “mixture of Poisson image denoising” strategy is to combine the two denoising outputs via \( \hat{\pi}(h_M) \) with the end goal to make the image denoising method robust to the artifacts evidenced in Figure 1.1(c) and Figure 3.2(b), but without losing image details like the oversmoothed image in Figure 3.2(c). To this end, we focus on the fact that differences between \( z_1 \cdot \hat{f}_P(h; \alpha z_1, \beta) \) and \( z_2 \cdot \hat{f}_P(h; \alpha z_2, \beta) \) are small unless the impulse-noise-like artifacts are present. Although a rigorous treatment such as (3.7) would also be appropriate, an ad-hoc convex combination of the form

\[
\hat{\pi}(h_M) = \exp \left( -\frac{\left( z_1 \cdot \hat{f}_P(h_M; \alpha z_1, \beta) - z_2 \cdot \hat{f}_P(h_M; \alpha z_2, \beta) \right)^2}{\sigma^2} \right)
\]

for some user-specified parameter \( \sigma \) was surprisingly effective. Specifically, \( \hat{\pi} \) approaches 0 when the differences between the two denoised images are large, attenuating the contribution of \( z_1 \cdot \hat{f}_P(h; \alpha z_1, \beta) \) that is likely suffering from denoising artifacts.

### 3.3 Denoising Results

To verify the effectiveness of the mixture Poisson denoising approach, we conduct experiments using real sensor data from Nikon D90 captured in raw sensor mode with
Figure 3.2: Denoised results of Poisson mixture image denoising method. A 5-level wavelet transform is performed for denoising. (d) is the difference between $z_1 \cdot \hat{f}_P(h_M; \alpha z_1, \beta)$ and $z_2 \cdot \hat{f}_P(h_M; \alpha z_2, \beta)$ in wavelet domain.
Figure 3.3: (a)-(c) are results of [36, 24, 37] denoising algorithms. (d)-(f) are results of Poisson mixture image denoising.
all manual setting. For each low light scene (1.2 lux), we captured two additional images—one of the X-Rite ColorChecker placed in the camera’s optical path to learn the noise parameters $\alpha$ and $\beta$; and of a “blank” image captured with lens cap on to assess the level of electronic/signal-independent noise. (Although noise parameters $\alpha$ and $\beta$ can be computed automatically using methods in [13] also, we take the calibration approach to so that denoising performance isn’t confounded by noise parameter estimation.) Denoising was performed on the red pixels only to avoid complications of the color filter array (red image is also most noisy). We are unable to provide a numerical score of performance based on simulation experiments, with the discrepancy between simulation and real image sensor data results being precisely the point of this paper. Hence our assessment of denoising performance in this paper is limited to visual inspection, unfortunately.

Figures 3.3(a)-(c) show the results of DWT-based Poisson image denoising algorithms in [36, 24] and DCT-based Poisson-Gaussian hybrid image denoising method in [37], respectively. Though there are major differences between Figures 3.3(a)-(b), both DWT-based denoising methods suffer from impulse-noise-like artifacts. Unsurprisingly, DCT-based denoising method shown in Figure 3.3(c) is largely free of the artifacts in question, thanks to the agreement between DCT noise distribution in simulation and in real sensor data. Though the reconstructed edge contrast was higher, there were a number of false edges and the output has a waxy appearance overall.

Figures 3.3(d)-(f) show the results of Poisson-mixture image denoising implemented as mixture of Poisson image denoising. The DWT-based Poisson image denoising method of [36] improves significantly with the convex combination weights of (3.10). The artifacts were successfully attenuated, even while retaining image details such as
edges and textures largely intact. Although artifacts were removed from DWT-based
denoising method of [24] also, noise suppression itself wasn’t sufficient enough to
achieve an acceptable result. For the sake of completeness, the DCT-based denoising
method was also treated by the mixture image denoising procedure. As expected,
there were no significant differences between Figures 3.3(c) and 3.3(f).
CHAPTER IV
CONCLUSION

We proposed a new camera sensor noise model in Haar wavelet domain aimed at rectifying the mismatch of ordinarily assumed noise models in image denoising. First, we designed a mixture Poisson model based on the canonical noise model. The Quantile analysis is provably better match with empirical image sensor noise than existing alternatives (2.1)-(2.3). Second, grounded on Poisson mixture noise model, we developed a mixture Poisson image denoising technique which can be implied to existing state-of-art Poisson denoising methods for real image sensor. Experimental results confirm that the mixture Poisson method evidently improves the denoising effects in real camera data.

The significance of this work extends the acknowledgement of real camera sensor noise distribution. The mixture Poisson model fits closely to real data and diminishes the mismatch of heavy tail behavior in wavelet domain. This is useful for developing effective denoising methods for real sensor image. The application proposed mixture Poisson image denosing technique efficiently attenuated the artifacts which caused by model mismatch and preserved image details mostly. The future work would likely apply the image denoising method to demosaicing algorithm for reconstructing full three-color denoised images.
BIBLIOGRAPHY


[37] Kostadin Dabov, Alessandro Foi, Vladimir Katkovnik, and Karen Egiazarian,