Using Optical Defocus to Denoise

Qi Shan[*]   Jiaya Jia[†]   Sing Bing Kang[‡]   Zenglu Qin[†]

[*]University of Washington   [†]The Chinese University of Hong Kong   [‡]Microsoft Research

Abstract

Effective reduction of noise is generally difficult because of the possible tight coupling of noise with high-frequency image structure. The problem is worse under low-light conditions. In this paper, we propose slightly optically defocusing the image in order to loosen this noise-image structure coupling. This allows us to more effectively reduce noise and subsequently restore the small defocus. We analytically show how this is possible, and demonstrate our technique on a number of examples that include low-light images.

1. Introduction

Despite advances in camera technology, sensor noise remains a major problem in photography, especially when pictures are taken under low-light conditions or with a high ISO setting. The characterization of noise is non-trivial (see, for example, [27, 21]), as noise is a function of exposure level, photon flux, and the electron-photon conversion process. In addition, noise contains high-frequency components that are quantitatively and at times visually indistinguishable from the inherent fine structures in natural images. Denoising these images using current techniques either excessively smoothes out image detail or retains noise with the detail.

In this paper, we present a novel denoising technique that uses optical defocus to reduce signal-noise coupling from a single input image. This is based on the observation that image details are typically hard to separate from noise; we simplify the denoising algorithm by reducing these image components optically. The ISO setting of the camera can then be set high in our system to allow better shutter speed.

Contributions This paper has three main contributions. First, our imaging system uses optical defocus to hide image details, which simplifies noise reduction. Extensive experiments indicate that blurry images hide the majority of image details but does not necessarily lose all of them. Many of the image structures can be recovered. Moreover, optical defocus and noise production are two separate processes in image formation, which allows us to manipulate the former in order to simplify the reduction of the latter.
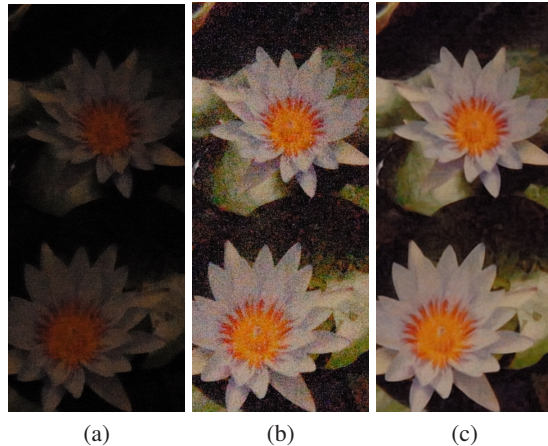

Figure 1. Effective denoising through slight optical defocus. (a) input low-light image, (b) brightness enhanced input (with noise enhanced as well), (c) restored image after noise and blur removal.

Second, we describe a new and effective method for single image noise estimation that is based on the presence of optical defocus. We analyze the relationship between defocus and noise, and propose a novel metric for defining the noise likelihood. We use *variable splitting* optimization (based on local closed-form solutions in iterations) to remove noise from the observed image. One result is shown in Figure 1; here, the underexposed image was captured in a dark room using a Nikon DSLR camera which was slightly defocused. If we merely enhance brightness, noise is amplified as well.

Third, we analyze the performance of the proposed algorithm in choosing a key parameter and quantitatively study the information gain with our new imaging technique. We show that the information loss introduced by defocus is several orders of magnitude smaller than the gain by removing noise. In addition, we are able to restore a reasonable amount of underlying image structure.

Assumptions We assume that the foreground objects are in focus after removing the defocus. Our goal is not to deblur the entire scene (especially when it has significant depth variation). Our method estimates the defocus blur PSF through a calibration process by measuring the foreground depth using cameras.
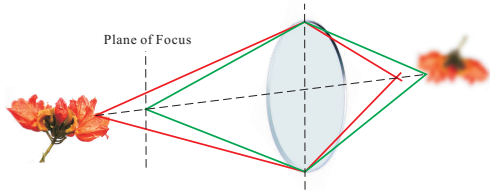
Figure 2. Optical defocusing. We adjust the lens to focus on a plane slightly *in front* of the object of interest.

2. Related Work

Representative techniques for denoising fall into multi-image and single-image approaches.

Two-image/video denoising The temporal information in video has been extensively used for denoising. For example, Chen and Tang [4] use a Spatio-Temporal Markov Random Field (MRF) approach to denoise video sequences using the motion information. The technique of Bennett and McMillan [2] enhances videos by locally deciding between spatial or temporal filtering.

There are also techniques for enhancing photos by processing two images taken with different camera settings, for example, with flash/no-flash (e.g., [25, 9]) or long/short exposure (e.g., [16, 37]). In [25], the noise in the non-flashed image is removed by bilateral filtering, where the detail is obtained from the flashed image. Eisemann and Durand [9] extracted coarse structure layer from the non-flash image and enhanced it by adding detail and color information from the flashed image. Jia *et al*. [16] captured blur/under-exposed image pairs and then transferred color from the blur image to the under-exposed counterpart. Yuan *et al*. [37] used a blur/noisy image pair; the output is constructed using deconvolution with an estimated PSF.

Methods that use multiple input images typically require pixels to be aligned over time. This is difficult to achieve in the presence of significant noise in a dynamic scene.

Single image denoising The most direct solution is to use a filter. Popular methods involve bilateral filtering [33] and anisotropic diffusion, either implemented in the form of partial differential equations (PDEs) [35, 34, 12] or derived from optimization using variational methods [30].

Single image denoising is very challenging because the problem is generally under-constrained. It requires making additional assumptions on the image or noise. For example, Roth and Black [29] modified the simple smoothness prior and introduced a high-order learning-based image prior model, which is potentially capable of better modeling natural scenes. Liu *et al*. [21] constructed a Gaussian conditional random field to infer the clean image with the piecewise smoothness assumption. [28] is recent work using total variation regularization. An interactive denoising method was proposed in [5].

Wavelet-based methods make use of the observation that multiscale subbands satisfy a highly kurtotic marginal distribution (e.g., [11]). The method of Portilla et al. [26] models the wavelet coefficients at adjacent positions and scales as the product of two independent random variables and uses the Gaussian scale mixture model for denoising. Another common assumption is the existence of regular texture or repeated local appearance. In [34], "geometry tensors" were proposed for noise removal and texture preservation. Non-local spatial domain denoising methods [3, 1, 7] rely on repeated local appearance to restore the latent image.

All these single image methods work best for denoising images with little or no fine texture. Unfortunately, it is very difficult to separate camera sensor noise from subtle image structures. In this paper, we partially resolve this ambiguity by taking into account both noise and defocus blur.

Single image deblurring Another type of image artifacts caused by low lighting is motion blur. Single image deblurring methods [10, 15, 31] are capable of restoring images to a certain extent. These methods involve kernel estimation and deconvolution. Non-blind deconvolution methods [22, 20, 38] assume that noise is relatively small, so that general smoothness constraints are adequate. With the exception of [18], single image deblurring techniques usually do not work well if noise is significant.

3. Noise Analysis

An image with noise can be expressed as $\mathbf{B}' = \mathbf{x} + \mathbf{n}$, where $\mathbf{B}'$ is the observed noisy image, $\mathbf{x}$ is the latent image, and $\mathbf{n}$ is noise. $\mathbf{n}$ is typically assumed to be signal independent, i.e., it is caused by dark current, amplifier noise and the quantizer in the camera circuity [14, 21]. However, recent camera noise estimation work found that this simple formula does not sufficiently model the mechanism of an image sensor where photon flux and the uncertainty of the electron-photon conversion process produce signal (or luminance) dependent noise. The total noise variance σ^2 is dependent on the gray-level variance κ_{gray}^2. It can be written as $\sigma^2 = \kappa_{gray}^2 + C\eta^2$ [27], where η^2 is the photon noise variance and C is a weight. The existence of fine structures in $\mathbf{x}$ (typical in real images) compounds the difficulty in accurately estimating $\mathbf{n}$ from $\mathbf{B}'$.

Our solution is to slightly defocus the image. During image capture, optical defocus is relatively independent of pixel noise generation [13, 36]. Thus, the image formation process can be expressed as

$$\mathbf{B} = (\mathbf{x} \otimes \mathbf{f}) + \mathbf{n}, \tag{1}$$

where $\mathbf{B}$ and $\mathbf{x}$ are the observed and latent images. $\mathbf{f}$ and $\mathbf{n}$ are respectively defocus PSF and noise. $(\mathbf{x} \otimes \mathbf{f})$ makes the image structure less correlated with noise.

Note that equations similar to (1) were proposed in non-blind deconvolution [22, 31] as the image degradation
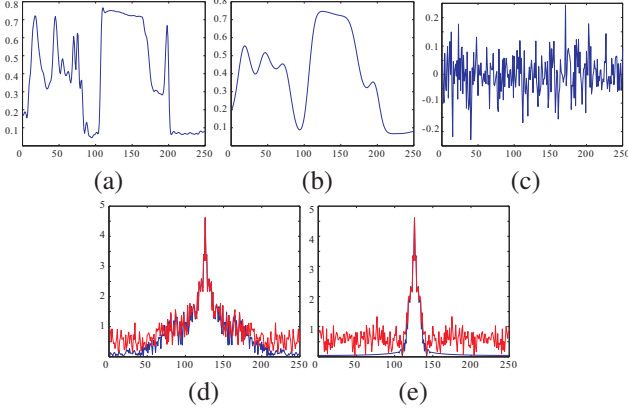
Figure 3. Reducing signal-noise coupling. (a) Input clean signal, (b) Gaussian filtered version of (a), (c) noise, (d) log-magnitudes of the unfiltered signals with (red curve) and without (blue curve) noise in the frequency domain. Both of them contain high frequency components. (e) Log-magnitudes of the filtered signals with (red curve) and without (blue curve) noise in the frequency domain. Their distributions are different.

model. However, these deconvolution methods cannot handle large noise. We have experimented with these methods to directly deblur our captured images with large regularization weights to overcome noise, and found either excessive removal of texture details or amplification of noise.

Figure 3 shows (in 1D) how blur loosens the signal-noise coupling. The red and blue curves in Figure 3(d) are plots of the log-magnitudes of the clean and noise-corrupted signals respectively in the frequency domain. They share many high frequency components that are not easily separable. However, if the signal is pre-filtered, as shown in (b), adding noise to it significantly changes its frequency distribution (shown in Figure 3(e)). The distribution difference implies a way to remove substantial noise from signal.

If noise can primarily be removed, we then digitally remove the slight optical blur through deconvolution. The recent deblurring work (e.g., [38, 19]) and our analysis (Section 5) show that deconvolving an image can recovered many details even though they are barely noticeable.

4. Noise Estimation with Focal Blur

In this section, we describe our novel method to estimate noise from a single defocused image (after which we can get the latent image by subtraction).

4.1. Noise Estimation with a Convolution Model

Following Eq. (1), a simple method to remove defocus blur while assuming small noise contamination would be to solve

$$\bar{\mathbf{x}} = F^{-1}\bar{\mathbf{B}} = \frac{F^T}{F^T F}\bar{\mathbf{B}}, \tag{2}$$

where F is the matrix form of the PSF $\mathbf{f}$. $\bar{\cdot}$ defines the vectorization operator which stacks all values in a raster-scanning order. Here $\bar{\mathbf{x}}$ and $\bar{\mathbf{B}}$ are respectively the vectorized $\mathbf{x}$ and $\mathbf{B}$. For the rest of this paper, F^{-1} and $\frac{1}{F}$ are used interchangeably to denote the inverse of F.

When used as is, Eq. (2) is very sensitive to noise because of magnification by the denominator $F^T F$. To deal with this problem, a common practice is to increase the diagonal values of $F^T F$ to stabilize the matrix inverse. Thus Eq. (2) is modified to

$$\bar{\mathbf{x}}(\lambda) = \frac{F^T}{F^T F + \lambda I}\bar{\mathbf{B}}, \tag{3}$$

where I is the identity matrix with the same dimension of $\mathbf{B}$, and parameter λ controls the regularization strength.

With this modification, we now analyze how the estimated $\mathbf{x}(\lambda)$ using Eq. (3) deviates from the ground truth latent image $\mathbf{x}^*$. This theoretical analysis will provide important insights on how image noise influences the defocused image formation, which in turn leads to a novel formula for accurate noise estimation. We write

$$
\begin{aligned}
\bar{\mathbf{x}}(\lambda) - \bar{\mathbf{x}}^* &= \frac{F^T}{F^T F + \lambda I}\bar{\mathbf{B}} - \frac{F^T}{F^T F}(\bar{\mathbf{B}} - \bar{\mathbf{n}}) \\
&= \frac{F^T\bar{\mathbf{n}}}{F^T F + \lambda I} + \frac{-\lambda I\bar{\mathbf{x}}^*}{F^T F + \lambda I}.
\end{aligned} \tag{4}
$$

$\bar{\mathbf{x}}^*$ can be expressed as $F^{-1}(\bar{\mathbf{B}} - \bar{\mathbf{n}})$ based on Eq. (1). Here the commutativity of multiplication in the denominator is not a concern because both $F^T F + \lambda I$ and $F^T F$ are symmetric. Eq. (4) contains two terms. The first denotes the influence of image noise, while the second represents the effect of structure smoothing. Note that the ground truth $\mathbf{x}^*$ is unknown. So introducing Eq. (4) is only to establish a metric to properly *measure* the influence of noise.

Our proposed metric is the partial derivative of $\mathbf{x}$ with respect to λ:

$$
\begin{aligned}
\frac{\partial \bar{\mathbf{x}}}{\partial \lambda} &= \lim_{\delta\lambda \to 0} \frac{1}{\delta\lambda}(\bar{\mathbf{x}}(\lambda + \delta\lambda) - \bar{\mathbf{x}}^*) - (\bar{\mathbf{x}}(\lambda) - \bar{\mathbf{x}}^*) \\
&= \lim_{\delta\lambda \to 0} \frac{1}{\delta\lambda}\left(\left(\frac{F^T\bar{\mathbf{n}}}{F^T F + \lambda + \delta\lambda} + \frac{-(\lambda + \delta\lambda)\bar{\mathbf{x}}^*}{F^T F + \lambda + \delta\lambda}\right) \right. \\
&\quad \left. - \left(\frac{F^T\bar{\mathbf{n}}}{F^T F + \lambda} + \frac{-\lambda\bar{\mathbf{x}}^*}{F^T F + \lambda}\right)\right) \\
&= -h(F, \lambda)\bar{\mathbf{n}} - h(F, \lambda)F\bar{\mathbf{x}}^*, \tag{5}
\end{aligned}
$$

where

$$h(F, \lambda) = \frac{F^T}{(F^T F + \lambda)^T(F^T F + \lambda)}. \tag{6}$$

We further compute the squared L_2 norm of $\frac{\partial \mathbf{x}}{\partial \lambda}$ as

$$\left\|\frac{\partial \mathbf{x}}{\partial \lambda}\right\|_2^2 = \|h(F, \lambda)\bar{\mathbf{n}} + h(F, \lambda)F\bar{\mathbf{x}}^*\|_2^2. \tag{7}$$

We now explain that $\left\|\frac{\partial \mathbf{x}}{\partial \lambda}\right\|_2^2$ is a new noise likelihood because it monotonically increases with noise variance.
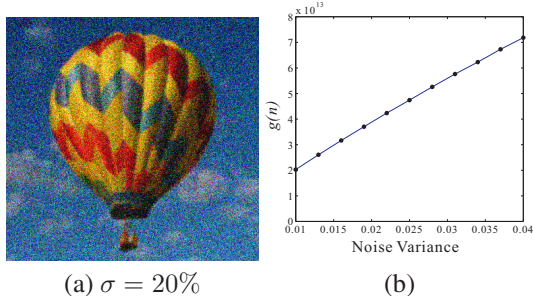
(a) $\sigma = 20\%$ (b)

Figure 4. An example to justify the noise metric $\left\|\frac{\partial \mathbf{x}}{\partial \lambda}\right\|_2^2$. (a) A blurred image with significant noise. (b) $\left\|\frac{\partial \mathbf{x}}{\partial \lambda}\right\|_2^2$ w.r.t. noise variance σ^2.

New Noise Metric Analysis We can show (details in [32]) that

$$\left\|\frac{\partial \mathbf{x}}{\partial \lambda}\right\|_2^2 \approx \sigma^2 \left\|\mathcal{H}(\mathbf{f}, \lambda)\right\|_2^2 + C, \tag{8}$$

with σ^2 being the noise variance and C being a constant. Eq. (8) has an important property: $\left\|\frac{\partial \mathbf{x}}{\partial \lambda}\right\|_2^2$ increases monotonically with σ^2. This makes it a good metric for evaluating how noisy a blurred image is. We empirically validate this property by generating 10 blurred images. They are added with different degrees of noise. One of them is shown in Figure 4(a) with noise standard deviations $\sigma = 20\%$. We then compute $\left\|\frac{\partial \mathbf{x}}{\partial \lambda}\right\|_2^2$ for each image and plot their values in Figure 4(b). As expected, $\left\|\frac{\partial \mathbf{x}}{\partial \lambda}\right\|_2^2$ monotonically increases with σ^2.

Thus, the function $\left\|\frac{\partial \mathbf{x}}{\partial \lambda}\right\|_2^2$ is a reasonable measure of how well noise is removed and of how the remaining image satisfies the convolution model defined with the kernel f. Let $\mathbf{n}^*$ and $\mathbf{n}'$ denote the ground truth noise and noise (somehow) estimated from $\mathbf{B}$, respectively. The remaining noise $\mathbf{n}$ in image is thus

$$\bar{\mathbf{n}} = \bar{\mathbf{n}}^* - \bar{\mathbf{n}}' = \bar{\mathbf{B}} - F\bar{\mathbf{x}}^* - \bar{\mathbf{n}}'. \tag{9}$$

Substituting $\mathbf{n}$ into Eq. (7) and denoting $E_0(\mathbf{n}') = \left\|\frac{\partial \mathbf{x}}{\partial \lambda}\right\|_2^2$ yield

$$E_0(\mathbf{n}') = \left\|h(F, \lambda)(\bar{\mathbf{B}} - \bar{\mathbf{n}}')\right\|_2^2. \tag{10}$$

$E_0(\mathbf{n}')$ indicates how good the noise estimate $\mathbf{n}'$ is by evaluating the strength of the remaining noise $\mathbf{n}^* - \mathbf{n}'$ in the image $\mathbf{B}$. $\mathbf{x}^*$ cancels out in Eq. (10) and thus does not need to be known beforehand. With the monotone property of $\left\|\frac{\partial \mathbf{x}}{\partial \lambda}\right\|_2^2$ with respect to the level of noise, we use $E_0(\mathbf{n}')$ as the likelihood in defining a new objective function.

Energy function Likelihood (10) is combined with a simple regularization term $\|\mathbf{n}'\|$ to avoid the trivial solution $\mathbf{n}' = \mathbf{B}$. The total energy $E_1(\mathbf{n}')$ is written as

$$\begin{aligned} E_1(\mathbf{n}') &= E_0(\mathbf{n}') + w\|\mathbf{n}'\| \\ &= \left\|h(F, \lambda)(\bar{\mathbf{B}} - \bar{\mathbf{n}}')\right\|_2^2 + w\|\mathbf{n}'\|, \end{aligned} \tag{11}$$
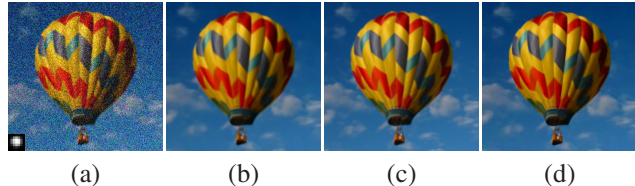

(a)   (b)   (c)   (d)

Figure 5. Denoising with slight defocus blur. (a) Input image with a high level of noise and slight out-of-focus. (b) The denoising result after convergence in 8 iterations. (c) Ground truth blurred image. (d) Our final deconvolution result.

where weight $w = 1.0$ in our experiments. To optimize $E_1(\mathbf{n}')$, we use the *variable splitting* optimization technique. The idea is to split the variable $\mathbf{n}'$ into a pair of variables ($\mathbf{n}'$ and $\mathbf{n}^s$), used respectively in the two terms in (11), such that minimizing the sum of the two terms under the constraint that $\mathbf{n}' = \mathbf{n}^s$ is essentially equivalent to solving the original problem.

Optimization Eq. (11) is thus reformulated as

$$\begin{aligned} E(\mathbf{n}', \mathbf{n}^s) &= \left\|h(F, \lambda)(\bar{\mathbf{B}} - \bar{\mathbf{n}}^s)\right\|_2^2 \\ &\quad + w\|\mathbf{n}'\| + \kappa\|\mathbf{n}' - \mathbf{n}^s\|_2^2, \end{aligned} \tag{12}$$

where the weight κ determines how similar $\mathbf{n}'$ and $\mathbf{n}^s$ are. An iterative approach is used to update $\mathbf{n}^s$ and $\mathbf{n}'$ separately.

[Updating $\mathbf{n}'$]

By removing all terms independent of $\mathbf{n}'$, we get

$$E'(\mathbf{n}') = w\|\mathbf{n}'\| + \kappa\|\bar{\mathbf{n}}' - \bar{\mathbf{n}}^s\|_2^2. \tag{13}$$

Its closed-form solution is

$$\mathbf{n}_i' = \begin{cases} \max\left(\frac{2\kappa\mathbf{n}_i^s - w}{2\kappa}, 0\right); & \mathbf{n}_i^s \geq 0 \\ \min\left(\frac{2\kappa\mathbf{n}_i^s + w}{2\kappa}, 0\right); & \mathbf{n}_i^s < 0 \end{cases} \tag{14}$$

where i indexes the pixels.

[Updating $\mathbf{n}^s$]

Similarly, removing all terms independent of $\mathbf{n}^s$ yields

$$E^s(\mathbf{n}^s) = \left\|h(F, \lambda)(\bar{\mathbf{B}} - \bar{\mathbf{n}}^s)\right\|_2^2 + \kappa\|\bar{\mathbf{n}}' - \bar{\mathbf{n}}^s\|_2^2.$$

It can be rewritten in the frequency domain as

$$E^s(\mathbf{n}^s) = \|\mathcal{H}(\mathbf{f}, \lambda) \circ (\mathcal{F}(\mathbf{B}) - \mathcal{F}(\mathbf{n}^s))\|_2^2 + \kappa\|\mathcal{F}(\mathbf{n}') - \mathcal{F}(\mathbf{n}^s)\|_2^2. \tag{15}$$

A closed-form solution also exists by computing the partial derivative with respect to $\mathbf{n}^s$ and setting it to zero:

$$\mathbf{n}^s = \mathcal{F}^{-1}\left(\frac{\kappa\mathcal{F}(\mathbf{n}') + \overline{\mathcal{H}(\mathbf{f}, \lambda)} \circ \mathcal{H}(\mathbf{f}, \lambda) \circ \mathcal{F}(\mathbf{B})}{\kappa + \overline{\mathcal{H}(\mathbf{f}, \lambda)} \circ \mathcal{H}(\mathbf{f}, \lambda)}\right). \tag{16}$$

Updating $\mathbf{n}'$ and $\mathbf{n}^s$ iterates until convergence (no more than 15 iterations generally). Because each step has a closed-form solution, the computation is much more efficient compared to conventional gradient descent. The energy is guaranteed to monotonically decrease.

Figure 5 shows one example. The input image shown in (a) is defocus blurred with significant noise ($\sigma = 20\%$). The final denoising result is shown in Figures 5(b). Its PSNR is as high as 33.83 (or MSE 5.19) compared to the ground truth blurred image shown in Figure 5(c). The denoised image will be further deconvolved to remove the slight blur (details in Sections 5 and 6). The final result shown in Figure 5(d) contains many details. It is notable that the original PSNR is only 14 before denoising.

4.2. Determining λ

The value of λ in the above formulas has a significant impact in denoising. We can show that the following condition must be enforced:

$$\|\mathcal{H}(\mathbf{f}, \lambda)\|_2^2 > w \|\Delta \mathbf{n}\|_1 / \|\Delta \mathbf{n}\|_2^2, \qquad (17)$$

where $\Delta \mathbf{n} = \mathbf{n}^* - \mathbf{n}'$, $\mathbf{n}^*$ is the ground truth noise, and $\mathbf{n}'$ is the noise estimate. The details of derivation can be found in [32]. (17) indicates that a large value of λ adversely affects convergence and *the maximum value of λ depends on w and the noise estimation error*. We can show that the upper bound of λ is approximately $w/\|\mathbf{n}^*\|_2$ (see the derivation in [32]). In our experiments, we assign λ a fixed small value 10^{-4} to inhibit its negative influence.

5. Deconvolution and Error Analysis

After noise removal, we deconvolve the image. In this section, we show that this process introduces error that is insignificant compared to image noise.

We analyze a naïve algorithm characterized by Eq. (3), which simply inverts the blur process through blur matrix division. Its error estimate can generally be regarded as the *upper bound* of errors produced by various deconvolution methods because almost all deconvolution methods, such as [38, 20, 31, 19], use more advanced techniques to regularize deblurring, and are thus capable of producing much higher quality results. We show, compared to the scale of image noise, even the upper bound of the deconvolution error is sufficiently small.

The deconvolution error is the difference between the recovered and ground truth latent images. For Eq. (3), the error can be expressed as

$$\bar{\mathbf{x}}(\lambda) - \bar{\mathbf{x}}^* \quad = \quad \frac{F^T \bar{\mathbf{n}}}{F^T F + \lambda I} + \frac{-\lambda I \bar{\mathbf{x}}^*}{F^T F + \lambda I}, \qquad (18)$$

where $\mathbf{n}$ is mostly quantization error after noise removal. Note that λ is used to stabilize deconvolution, and is usually with very small value (10^{-4} in our experiments). We can therefore ignore the second term in Eq. (18), yielding

$$\|\bar{\mathbf{x}}(\lambda) - \bar{\mathbf{x}}^*\|_2^2 \quad \approx \quad \left\| \frac{F^T \bar{\mathbf{n}}}{F^T F + \lambda I} \right\|_2^2.$$

Expressing it in the frequency domain gives

$$\left\| \frac{F^T \bar{\mathbf{n}}}{F^T F + \lambda I} \right\|_2^2 = \left\| \frac{\overline{\mathcal{F}(\mathbf{f})} \circ \mathcal{F}(\mathbf{n})}{\overline{\mathcal{F}(\mathbf{f})} \mathcal{F}(\mathbf{f}) + \lambda} \right\|_2^2 = \left\| \frac{\overline{\mathcal{F}(\mathbf{f})}}{\overline{\mathcal{F}(\mathbf{f})} \mathcal{F}(\mathbf{f}) + \lambda} \right\|_2^2 \sigma^2,$$

where σ is the standard deviation of noise $\mathbf{n}$. This shows that $\mathbf{n}$ is magnified by a factor

$$\mathcal{J} = \sqrt{\left\| \frac{\overline{\mathcal{F}(\mathbf{f})}}{\overline{\mathcal{F}(\mathbf{f})} \mathcal{F}(\mathbf{f}) + \lambda} \right\|_2^2}. \qquad (19)$$

We now estimate the magnitude of $\mathcal{J}$ using different out-of-focus PSFs. In our experiments, the value varies from 17.7 to 28.6. The quantization error $\mathbf{n}$ is generally modeled as a uniform distribution in the range $[-0.5, 0.5)$, with a standard deviation of 0.29. With these quantities, it can be estimated that the reconstruction PSNR of the naïve deconvolution algorithm easily exceeds 33. The MSE ranges from 24 to 66. These quantities indicate that the error introduced even using this simple deconvolution algorithm is very small if the convolution model is satisfied. In comparison, the input image has significant noise where the PSNR is 14 and the MSE is 2.2×10^3. It is two orders of magnitude larger than the deconvolution error. In our experiments, the error introduced only from deconvolution is small enough compared to the contribution of defocus blur to intensive noise removal, as illustrated in Figure 5.

6. Implementation Details

Empirically, we first perform photometric calibration [8] and then produce the defocus blur using the camera manual focusing function to a slightly near point instead of the ideal object plane, as depicted in Figure 2. To estimate the PSF, we first record the manual focusing distance u_1 from the camera lens and the ideal object distance u_2 using the rangefinder attached to the camera. We then apply the calibration technique of [20] on u_1 and u_2 to estimate the PSF.

We constrain the size of the defocus blur kernel to be at most 11×11 (pixels). We initialize the noise layer using the method of Dabov et al. [6]. Because this method does not take the blur model into consideration, the noise estimate contains errors. We then apply our method, as described in Section 4, to optimize the noise map. After denoising, we remove the slight defocus blur introduced optically using the executable for non-blind deconvolution [31]. The denoising and deblurring steps alternate. Typically at most 10 iterations are enough to produce a visually compelling result. The running time is about 3 minutes for an image with 800×600 pixels on a desktop PC with a Core2Duo 2.8GHz CPU.

7. Quantitative Evaluation

The first example shown in Figure 6 is to quantitatively evaluate the effectiveness of our method given significant

(a) Input with blur   (c) Ground truth of (a)   (d) Our result of (a)   (g)   (h)

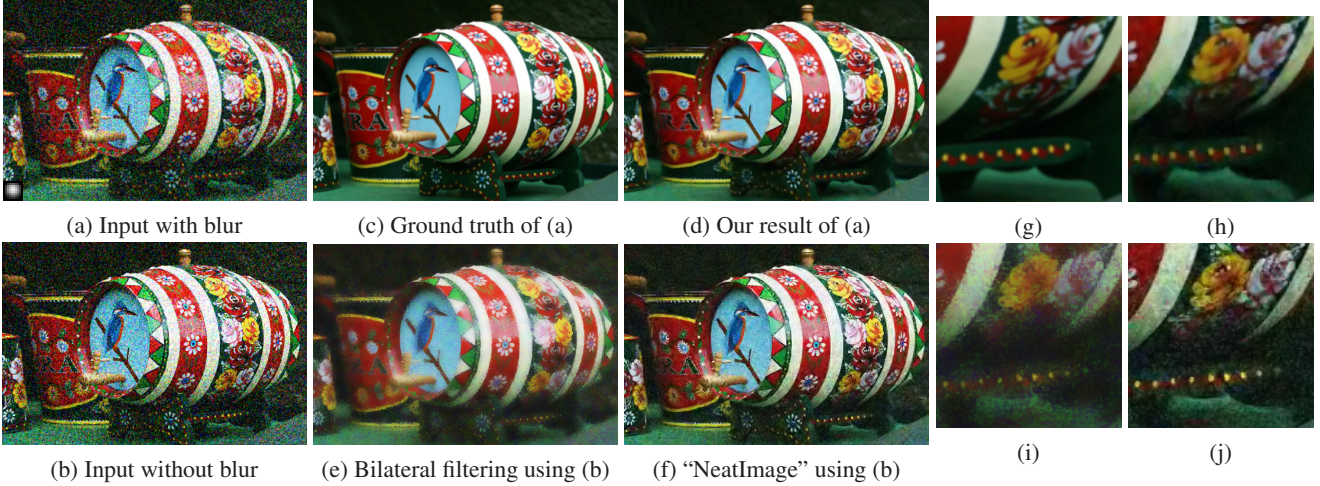(b) Input without blur   (e) Bilateral filtering using (b)   (f) "NeatImage" using (b)   (i)   (j)

Figure 6. Quantitative evaluation. The input image (shown in (a)) is blurred with the PSF shown on bottom left. Significant CCD noise ($\sigma = 20\%$) is also added. (b) Another input noisy image that is not focal blurred. (c) The ground truth sharp image without noise. (d) Our restoration result of (a), with PSNR 28.8. (e)-(f) Results of bilateral filtering [9] and "NeatImage" [24] with PSNRs 21.3 and 22.5, respectively. (g)-(j) Close-ups of (c)-(f).

| PSNR	$\sigma = 10\%$				$\sigma = 15\%$			
File name	bilat	PDE	wavelet	Ours	bilat	PDE	wavelet	Ours
100080.jpg	26.22	32.44	32.65	34.55	28.51	31.15	31.24	34.39
103041.jpg	25.35	29.90	29.91	31.53	26.66	28.00	28.39	31.07
108041.jpg	24.51	28.20	28.29	29.38	24.99	25.90	26.99	28.52
134008.jpg	25.30	29.33	29.73	31.25	26.26	27.55	28.27	30.89
161062.jpg	25.73	29.58	30.36	31.31	27.04	28.11	28.86	30.85
166081.jpg	24.57	27.79	28.24	29.53	25.74	26.34	27.03	29.31
176039.jpg	25.69	28.67	29.06	27.70	25.36	26.46	27.22	29.23
209070.jpg	24.74	29.06	29.64	30.24	25.69	27.32	27.99	29.77
22090.jpg	25.59	29.46	29.97	31.06	26.22	27.51	28.30	30.25
246053.jpg	26.77	32.45	31.53	34.34	27.85	30.07	29.67	33.19
247085.jpg	24.76	28.51	28.92	29.99	25.98	26.80	27.50	29.69
353013.jpg	25.29	28.35	27.29	30.12	24.90	25.97	25.77	29.13
mean	25.43	29.33	29.54	31.06	26.22	27.43	28.05	30.42
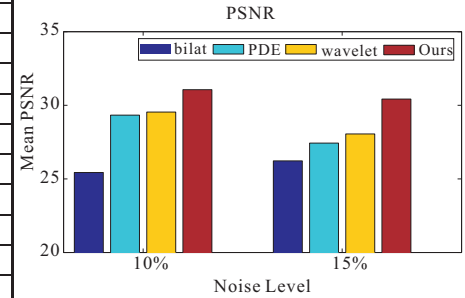
Table 1. Left: PSNRs of a set of the processed images for comparison. Red-title images are shown in [32]. Our PSNRs are calculated based on the final deconvolution results. Right: histogram of the mean PSNRs.

(a)   (c)   (d)   (g)   (h)
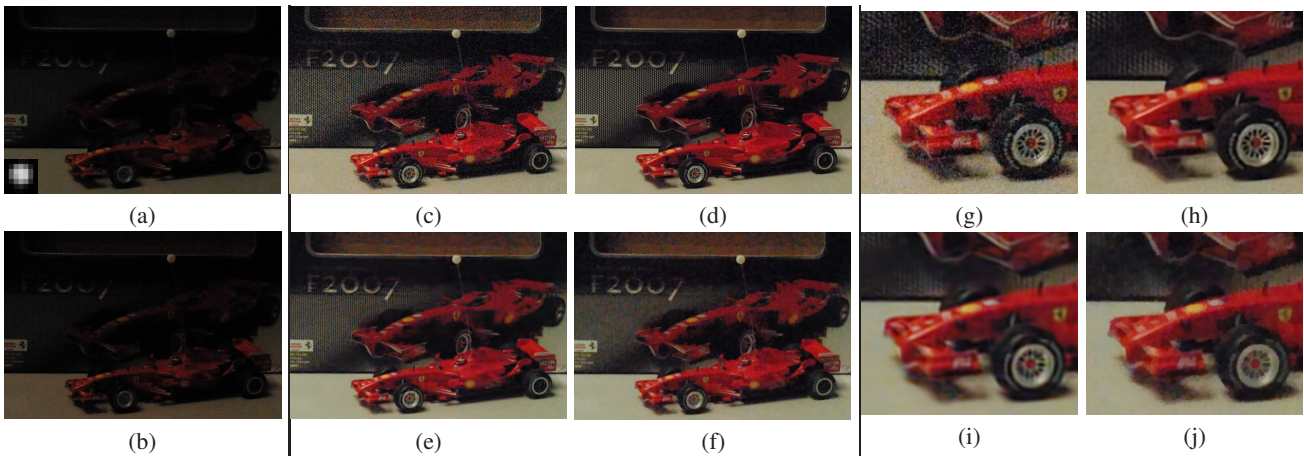
(b)   (e)   (f)   (i)   (j)

Figure 7. Image example. (a) The captured out-of-focus under-exposed image. (b) Another captured in-focus noisy image with the same exposure setting. (c) The intensity enhanced input (4X brightness). It contains significant noise. (d) Our final result from (c) after denoising and removing the blurriness. (e) The denoising result of the Gaussian mixture wavelet method [26] with (b) as input. (f) The denoising result of Rodriguez and Wohlberg [28] from an intensity enhanced version of (b). (g)-(j) Close-ups of (c)-(f).

PSNR	$\sigma = 10\%$		
File name	0th	1st	Ours
100075.jpg	28.14	28.96	30.34
105053.jpg	30.63	31.95	33.27
106025.jpg	32.03	34.22	35.47
15088.jpg	27.71	28.76	30.80
22013.jpg	27.14	28.84	29.06
314016.jpg	26.81	27.83	29.89
mean	28.34	29.95	30.75

Table 2. Part of the PSNRs compared to those in Table 2 of [21]. Columns "0th" and "1st" show the statistics obtained using orders zero and one models (described in [21]), respectively.
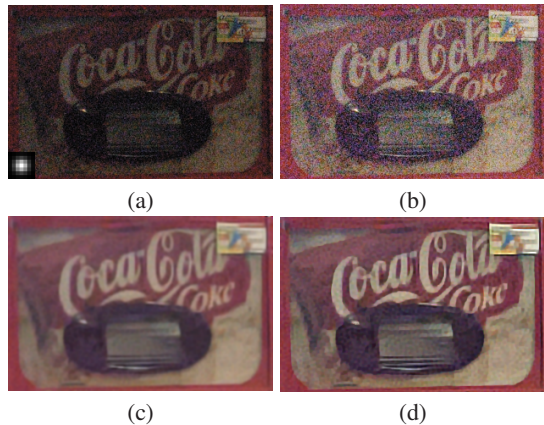

(a)　　　　　　　　　(b)

(c)　　　　　　　　　(d)

Figure 8. Coke example. (a) Image captured under low light. (b) Image after intensity enhancement, with noise proportionally amplified. (c) Result of noise removal. (d) Final result after removing defocus blur.

image noise. The input image (shown in Figure 6(a)) is blurred (PSF shown on bottom left) followed by adding large CCD noise [21] ($\sigma = 20\%$). Our image restoration result is shown in Figure 6(d). In (e)-(f), we show the results from two other denoising algorithms with an unblurred noisy image (shown in Figure 6(b)) as input.

We then collect the statistics of our denoising method using a set of image examples. In this experiment, 16 images containing different types of objects and scenes were selected from the Berkeley segmentation data set [23]. We blurred these input images by convolving them with small defocus kernels estimated by Joshi et al. [17]. This process was followed by adding white Gaussian noise, respectively at 10% and 15% levels. For final result comparison, several other denoising methods were also tested on the images with the same amount of noise (but without defocus blur). Details and visual comparisons are given in [32]. Part of the PSNRs are listed in Table 1.

In addition, we compare our method with that of Liu *et al.* [21]. The PSNRs extracted from [21] are used for comparison. Again, our method uses blurred images with additive 10% AWGN while the noisy images in [21] do not undergo the blurring process. Part of the PSNRs are pre-
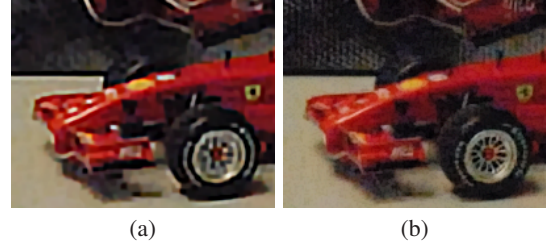

(a)　　　　　　　　　(b)

Figure 9. Image reconstructed with miscalibrated kernels. The input image is in Fig. 7(c) with kernel size of 7×7. (a) Using 11×11 Gaussian kernel (close-up view). (b) Using 3×3 Gaussian kernel (close-up view).

sented in Table 2.

8. More Experimental Results

We now show two more examples where the input images are captured under low light and with the camera slightly defocused. Several other examples are included in our technical report [32]. In Figure 7, we show an image captured by a Nikon D200 camera. It is severely underexposed with significant noise. For comparison, we also took a corresponding in-focus image (shown in Figure 7(b)) and enhance it as input to other denoising methods. The results are shown in (d)-(f), with close-ups in (h)-(j). It is noticeable that our result retains the most details. Figure 8 shows another example where the input image was taken under the similar condition.

9. Concluding Remarks

We have presented a new denoising technique based on optical defocus to reduce the signal-noise coupling. We showed that the gains in noise reduction more than offset the degradation in signal due to the defocus. We also introduced a new metric for evaluating how noisy a blurred image is. Experimental results were shown to validate our technique and analysis.

Limitations Our technique is less effective in cases where depth cannot be quantified (e.g., in macro photography). Our method tends to work best when both foreground and background are in the field of view. Our technique also assumes a certain style of photography where the nearest object is originally in focus (which is common); it removes only a small amount of defocus. Finally, we assume the blur PSF is spatially-invariant. From our experiments, our technique is tolerant towards moderate changes in the PSF (distortions of less than 4 pixels). Fig. 9, we show examples of what happens when the kernel is significantly misestimated. A larger-than-correct kernel makes the result over-sharpened but with less details (a) while a smaller-than-correct kernel makes the result look noisier (b). We will address these issues as part of our future work.

Acknowledgement

References

[1] N. Azzabou, N. Paragios, F. Guichard, and F. Cao. Variable bandwidth image denoising using image-based noise models. In *CVPR*, 2007.

[2] E. P. Bennett and L. McMillan. Video enhancement using per-pixel virtual exposures. *ACM Trans. Graph.*, 24(3):845–852, 2005.

[3] A. Buades, B. Coll, and J.-M. Morel. A non-local algorithm for image denoising. In *CVPR*, pages 60–65, 2005.

[4] J. Chen and C.-K. Tang. Spatio-temporal markov random field for video denoising. In *CVPR*, 2007.

[5] J. Chen, C.-K. Tang, and J. Wang. Noise brush: Interactive high quality image-noise separation. *ACM Trans. Graph.*, 2009.

[6] K. Dabov, A. Foi, and K. Egiazarian. Image restoration by sparse 3D transform-domain collaborative filtering. In *SPIE Electronic Imaging*, 2008.

[7] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian. Color image denoising via sparse 3D collaborative filtering with grouping constraint in luminance-chrominance space. In *ICIP*, 2007.

[8] P. Debevec and J. Malik. Recovering high dynamic range radiance maps from photographs. *ACM Trans. Graph.*, 1997.

[9] E. Eisemann and F. Durand. Flash photography enhancement via intrinsic relighting. *ACM Trans. Graph.*, 23(3):673–678, 2004.

[10] R. Fergus, B. Singh, A. Hertzmann, S. T. Roweis, and W. T. Freeman. Removing camera shake from a single photograph. *ACM Trans. Graph*, 25:787–794, 2006.

[11] D. J. Field. Relations between the statistics of natural images and the response properties of cortical cells. *Journal of the Optical Society of America A*, 4:2379–2394, 1987.

[12] G. Gilboa, N. A. Sochen, and Y. Y. Zeevi. Image enhancement and denoising by complex diffusion processes. *TPAMI*, 26(8):1020–1036, 2004.

[13] G. Healey and R. Kondepudy. Radiometric CCD camera calibration and noise estimation. *TPAMI*, 16(3):267–276, 1994.

[14] S. Ioué and K. R. Spring. *Video Microscopy, 2nd ed.* Plenum Press, 1997.

[15] J. Jia. Single image motion deblurring using transparency. In *CVPR*, 2007.

[16] J. Jia, J. Sun, C.-K. Tang, and H.-Y. Shum. Bayesian correction of image intensity with spatial consideration. In *ECCV*, pages 342–354, 2004.

[17] N. Joshi, R. Szeliski, and D. J. Kriegman. PSF estimation using sharp edge prediction. In *CVPR*, 2008.

[18] N. Joshi, C. L. Zitnick, R. Szeliski, and D. J. Kriegman. Image deblurring and denoising using color priors. In *CVPR*, 2009.

[19] D. Krishnan and R. Fergus. Fast image deconvolution using hyper-laplacian priors. In *NIPS*, 2009.

[20] A. Levin, R. Fergus, F. Durand, and B. Freeman. Image and depth from a conventional camera with a coded aperture. *ACM Trans. Graph.*, 2007.

[21] C. Liu, R. Szeliski, S. B. Kang, C. L. Zitnick, and W. T. Freeman. Automatic estimation and removal of noise from a single image. *TPAMI*, 30(2):299–314, 2008.

[22] L. Lucy. Bayesian-based iterative method of image restoration. *Journal of Astronomy*, 79:745–754, 1974.

[23] D. R. Martin, C. Fowlkes, D. Tal, and J. Malik. A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics. Technical report, University of California, Berkeley, 2001.

[24] NeatImage©. http://www.neatimage.com/. 2009.

[25] G. Petschnigg, R. Szeliski, M. Agrawala, M. F. Cohen, H. Hoppe, and K. Toyama. Digital photography with flash and no-flash image pairs. *ACM Trans. Graph.*, 23(3):664–672, 2004.

[26] J. Portilla, V. Strela, M. J. Wainwright, and E. P. Simoncelli. Image denoising using scale mixtures of gaussians in the wavelet domain. *TIP*, 12(11):133–1351, 2003.

[27] N. Ratner and Y. Y. Schechner. Illumination multiplexing within fundamental limits. In *CVPR*, 2007.

[28] P. Rodriguez and B. Wohlberg. Efficient minimization method for a generalized total variation functional. *TIP*, 18(2):322–332, 2009.

[29] S. Roth and M. J. Black. Fields of experts: A framework for learning image priors. In *CVPR*, pages 860–867, 2005.

[30] H. Scharr, M. J. Black, and H. W. Haussecker. Image statistics and anisotropic diffusion. In *ICCV*, pages 840–847, 2003.

[31] Q. Shan, J. Jia, and A. Agarwala. High-quality motion deblurring from a single image. *ACM Trans. Graph.*, 27(3), 2008.

[32] Q. Shan, J. Jia, S. B. Kang, and Z. Qin. Using optical defocus to denoise. Technical report, 2010. www.cs.washington.edu/homes/shanqi/work/denoise10/.

[33] C. Tomasi and R. Manduchi. Bilateral filtering for gray and color images. In *ICCV*, pages 839–846, 1998.

[34] D. Tschumperlé. Fast anisotropic smoothing of multi-valued images using curvature-preserving PDE's. *IJCV*, 68(1):65–82, 2006.

[35] D. Tschumperlé and R. Deriche. Vector-valued image regularization with PDEs: A common framework for different applications. *TPAMI*, 27(4):506–517, 2005.

[36] Y. Tsin, V. Ramesh, and T. Kanade. Statistical calibration of the CCD imaging process. In *ICCV*, pages 480–487, 2001.

[37] L. Yuan, J. Sun, L. Quan, and H.-Y. Shum. Image deblurring with blurred/noisy image pairs. *ACM Trans. Graph.*, 26(3):1, 2007.

[38] L. Yuan, J. Sun, L. Quan, and H.-Y. Shum. Progressive inter-scale and intra-scale non-blind image deconvolution. *ACM Trans. Graph.*, 27(3), 2008.