Denoising Scheme for Realistic Digital Photos from Unknown Sources

Suk Hwan Lim, Ron Maurer, Pavel Kisilev

HP Laboratories
HPL-2008-167

Keyword(s):
No keywords available.

Abstract:

This paper targets denoising of digital photos taken by cameras with unknown sensor parameters and image processing pipeline. Common noise characteristics in such images originate from camera-internal processing, such as demosaicing, tone mapping, and JPEG compression. Three of the noise characteristics that are not adequately addressed by existing denoising algorithms are spatially correlated low-frequency noise, strong signal dependency of the noise level and high levels of the chrominance noise relative to the luminance noise. We propose a generic scheme that extends existing denoisers such as the bilateral filter to account for all the problems above. Our solution combines a novel progressive pyramidal filtering scheme to address the correlated noise, filter adaptation via local noise level estimation and luminance-guided chrominance filtering to address the low-SNR of the chrominance channels. We demonstrate the effectiveness of our solution for challenging realistic noisy photos.

External Posting Date: October 21, 2008 [Fulltext] Approved for External Publication
Internal Posting Date: October 21, 2008 [Fulltext]

Submitted to International Conference on Acoustic, Speech and Signal Processing 2009

DENOISING SCHEME FOR REALISTIC DIGITAL PHOTOS FROM UNKNOWN SOURCES

Suk Hwan Lim, Ron Maurer and Pavel Kisilev
Hewlett-Packard Laboratories

ABSTRACT

This paper targets denoising of digital photos taken by cameras with unknown sensor parameters and image processing pipeline. Common noise characteristics in such images originate from camera-internal processing, such as demosaicing, tone mapping, and JPEG compression. Three of the noise characteristics that are not adequately addressed by existing denoising algorithms are spatially correlated low-frequency noise, strong signal dependency of the noise level and high levels of the chrominance noise relative to the luminance noise. We propose a generic scheme that extends existing denoisers such as the bilateral filter to account for all the problems above. Our solution combines a novel progressive pyramidal filtering scheme to address the correlated noise, filter adaptation via local noise level estimation and luminance-guided chrominance filtering to address the low-SNR of the chrominance channels. We demonstrate the effectiveness of our solution for challenging realistic noisy photos.

Index Terms— Noise filtering, denoising, correlated noise, signal-dependent noise, chromatic noise, and multi-resolution

1. INTRODUCTION

The goal of image denoising is to filter out as much noise as possible with as little degradation as possible in the underlying image. The majority of the state-of-the-art denoising methods focus their efforts on the hard task of modeling the signal statistics and assume the simplest model for the noise: Additive-White-Gaussian-Noise (AWGN), which is uncorrelated, signal-independent and position independent [1,4,5,12]. Other works on denoising target more complex noise models that are assumed to be known to the denoiser [9].

The simplistic noise models are not valid for images captured by the digital cameras or similar devices. Typically, raw image data are captured by a sensor (e.g. CCD), and then are *rendered* by several image processing steps inside the camera such as de-mosaicing, non-linear tone mapping, and lossy compression (typically JPEG). Some denoising schemes were developed for operating *inside* the camera and before compression. For example, Faraji [9] targeted CCD noise removal, assuming the knowledge of the sensor noise statistics and the tone-mapping curve, but neglected the effects of other parts of the imaging pipeline.

In this paper, we target a more difficult problem of "blind" denoising of the *rendered* photo images that were captured and compressed by cameras with unknown sensor parameters and imaging pipeline. This scenario occurs frequently in practice for the photo printing applications, where a variety of photos from different unknown sources are printed. In such a scenario, not only are the parameters of the noise model unknown, but also the noise model is untraceable, due to the spatial, chromatic and signal dependent mixing effects of demosaicing, color correction and JPEG compression [1]. In particular, we identify three common characteristics of noise in rendered photo images: Low-frequency (spatially correlated) noise, spatially-varying noise level due to the signal dependency and high levels of chrominance noise with considerably lower SNR than the luminance channel.

This paper proposes a generic denoising scheme that addresses the noise in the *rendered* realistic digital photos without relying on an explicit noise model but rather on the above common empirical noise characteristics. Section 2 details our proposed approach, while Section 3 contains results and discussion.

2. OUR DENOSING SCHEME

We describe the three challenging characteristics of noise in real digital photos that deviate significantly from the commonly used AWGN model. *First, noise in many digital photographs have significant Low Frequency* (LF) *component and large spatial correlations* whereas AWGN model assumes no spatial correlation in the image noise. Noise with significant LF component is difficult to remove, since it is difficult to separate the noise from the smooth signal features. Simply enlarging the support (kernel size) of the filter (large enough to cover such LF component of the noise) over-smoothes the image details significantly. *Second, the noise standard deviation* (STD) *is not uniform across the image.* The consequences of simplistically assuming uniform STD are under-cleaning of the noise in highly noisy areas and over-smoothing of the image features in areas with low noise. Thus, a more sophisticated model is needed to apply appropriate amount of filtering to all areas in the image. *Third, the chrominance channels typically have considerably lower signal quality than the luminance channel*. Since the chrominance channels are computed as approximate differences between the RGB color channels with correlated derivatives, they have much lower SNR than the luminance channel [1]. The SNR of the chrominance becomes even worse relative to the SNR of the luminance after JPEG compression due to subsampling and heavy quantization for the chrominance channels in typical settings. Denoising the chrominance channels becomes challenging, since the ability of denoising algorithms to separate noise from features is severely degraded when the SNR is very low, i.e. when the noise level increases to or above the typical edge-contrast levels.

We have developed a denoising scheme that addresses the three characteristics of noise (i.e., the spatially-correlated noise, spatially-varying noise level and very high chrominance noise). The scheme combines three parts, each dedicated to address each characteristic mentioned above. The principles of each part may be implemented with a variety of denoisers, but we choose to demonstrate our approaches with the simple but effective bilateral filter introduced by Tomasi and Manduchi [2], which is designed for selective denoising without blurring the edges.

2.1. A Multi-Resolution Scheme for Correlated Noise

We introduce a novel low-frequency (Gaussian-pyramid)-based Multi-Resolution (MR) framework designed for removing the spatially-correlated (low-frequency) noise and artifacts. In this framework, we first construct a Gaussian Pyramid with N+1 levels via a series of a simple smoothing low-pass filter followed by a subsampler. The pyramid construction can be summarized as

$$I_{k+1} = Subsample\left[Smooth\left(I_k\right)\right] \ (k=0,\ldots,N\text{-}1) \qquad (1)$$

,where I_k is the k^{th} level image. Note that I_0 is the original input image and the top N^{th} level image (I_N) is obtained by smoothing and downsampling the I_0 image N times. We then apply denoising to the N^{th} level image (I_N) and propagate the correction term (i.e. $Denoise(I_N) - I_N$) to the next lower level image (I_{N-1}) such that the low frequency (LF) noise in I_{N-1} is removed prior to applying the denoiser in the $N\text{-}1^{th}$ level. The propagation is performed by applying an *Upsample* operation on the correction term followed by a simple addition. Note that after the propagation step, I_{N-1} contains mostly high frequency (HF) noise with respect to the spatial resolution at that level. The denoising and propagation is repeated for each level ($k=N\text{-}1$ to 0) until the denoiser is applied to the original level. This is summarized as

$$I_k^* = Denoise\left[I_k + Upsample\left(I_{k+1}^* - I_{k+1}\right)\right] \ (k = N\text{-}1,\ldots,0), \quad (2)$$

,where I_k^* is the denoised image at kth level and the final denoised output image is I_0^*. Figure 1 illustrates the pyramid construction and propagation steps for level k. Note that the series of denoising and propagation steps allow the denoiser *at each level* to perform well since it is not required to deal with the LF noise. Before the actual denoising is applied to each level, any LF component of the noise at the level is removed by propagating the noise correction terms from the upper (lower resolution) levels.

It is worthwhile to point out that generic terminologies such as "*Smooth*", "*Downsample*", "*Upsample*" and "*Denoise*" have been used in the paragraph above. This is because our MR framework is general and robust (insensitive) enough to the variations in its building blocks. For example, *any* denoising method may be plugged into the denoise operation in Equation (2) such that prior works on noise filtering may be leveraged (the input image at each level is a natural image). Furthermore, the framework allows different noise filtering algorithms to be applied at each level and the frequency content of the residual noise at the final output image can be shaped by changing the aggressiveness of the denoiser at each level. In addition, even the simplest upscalers (for the "*Upsample*" operation) such as pixel replication or bilinear interpolation yield good results as long as the denoiser at each level performs well. Thus, the overhead in multi-resolution-related computations is small. In fact, the MR framework enables us to use smaller filter support at each level and reduce the overall computational complexity.

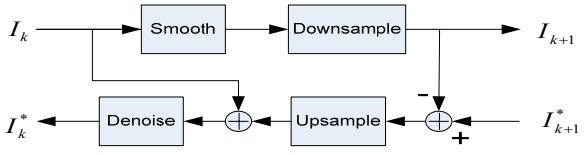


Figure 1: The block diagram of our multi-resolution framework

Our MR approach has some similarities with the conventional MR approaches (See [4][5] and refs therein). Both approaches implement MR concept by constructing a pyramid and applying the processing to the pyramid levels. In fact, the "*Smooth*" and "*Downsample*" steps in our approach correspond to the generation of the low frequency bands during the (wavelet/subband) "*decomposition*" stage in conventional MR approaches. The "*Upsample*" (and hence propagation) step in our MR approach corresponds to the "*reconstruction*" step in the conventional approach. Thus, the theory on various wavelets (e.g. Haar, Daubechies', Meyer's, sinc, spline) in the conventional MR analysis can be applied to our MR approach as well. For example, simple block averaging for the "*Smooth*" operation and pixel replication for the "*Upsample*" operation is equivalent to using Haar wavelets for subband decomposition and reconstruction.

We also point out the notable differences between our MR approach and the conventional MR approaches. Our approach constructs a Gaussian pyramid where all the images in the pyramid levels are natural images. Denoising is applied on these low-frequency bands between the propagations. On the other hand, conventional MR approaches typically construct a Laplacian pyramid and modify mid/high frequency bands (e.g. by shrinkage of coefficients), where the low-pass band is typically untouched. Note that shrinking small coefficients in the high frequency bands does not fully exploit the spatial correlations with the neighboring pixels. Also, while inter-scale dependencies in our approach are intuitive and easy to model, it is difficult for conventional MR approaches to accurately model the wavelet coefficients and their inter-scale dependencies for various pixels and images, especially when signal-dependent and spatially correlated noise are present.

In a related work, Aiazzi et. al. introduced a generalization of the Laplacian pyramid that can handle a certain class of signal dependent noise [6]. Yet, their approach is not suitable for spatially-correlated noise, since input to the denoisers at each level may still contain significant LF noise component. Furthermore, the filtered items are the mid/high frequency subbands, which are statistically different than natural images, making it difficult to apply the prior works on denoising algorithms or utilize the prior works on the statistical models of natural images.

2.2. Intensity-dependent Filtering for Non-uniform STD

Intensity-dependent noise filtering method was developed to cope with *non-uniform spatially-varying noise level*. Traditionally, denoising algorithms were developed under the assumption of uniform noise level across the image. Such algorithms are sub-optimal for photos captured by solid state image sensors such as CCDs, since the standard deviation of the noise is signal dependent. In particular, some parts of the image may be over-smoothed whereas other parts may be left with significant residual noise. Several prior works address filtering of the signal dependent noise [6,9,10] where the noise intensity is proportional to the signal raised to some fixed power or follows a piecewise power-law model. However, these methods are not general enough to handle the images from arbitrary sources of cameras with unknown imaging and rendering pipelines. Also, some of these methods require the knowledge of the tone-correction and its inverse, which are not always available or easy to estimate.

We propose a more general "blind" approach for filtering the

noise with spatially varying (intensity-dependent) standard deviation (STD). Our approach relies on estimating the noise STD as a function of the pixel intensity (i.e. $\sigma = f(I,\mathbf{p})$, where I is the pixel intensity and $\mathbf{p}$ is the noise parameters) directly from the given rendered image [8]. In our implementations, we use a simple quadratic model (i.e. $\sigma^2 = p_0+p_1I+p_2I^2$) for the noise variance since it is effective even for rendered images [1]. To estimate the parameters ($\mathbf{p}$), we bucket the image pixels into S segments (e.g., 1~51, 52~102, 103~153, 154~204 and 205~255 for S=5) and estimate a single noise STD for each segment using the method detailed in [8]. The noise parameters are estimated using S pairs of noise STD values and the average pixel intensities for each segment. Note that if this function is available from a model or off-line measurements then this step can be skipped. Since we assume that the noise STD varies only with intensity, we can compute the noise STD values for all the pixels in the image using the pixel intensity values and the estimated noise parameters.

As stated previously, we demonstrate our approach with the simple but effective bilateral filter [2]. Its key idea is to modify the weights of a convolution mask in an image dependent manner, based on intensity differences between the pixel under consideration and its neighbors. The signal $I(i)$ is filtered with the following formula

$$I^*(i) = \frac{1}{\sum_j g\big(I(i)-I(i-j)\big)\cdot h(j)} \sum_j I(i-j)\cdot g\big(I(i-j)-I(i)\big)\cdot h(j) \quad (3)$$

,where $h(j)$ is the convolution kernel, $I^*(i)$ is the denoised output at pixel location i and $g(I(i) - I(i-j))$ is the photometric distance term used for selectively denoising the pixels without blurring the edges. Essentially, $g(\ )$ is the function that determines whether the differences between the neighboring pixels are due to the actual image contents (e.g. edges) or noise. In the conventional bilateral filter [2], $g(\ )$ does not depend on the pixel intensity and is often a fixed function. An example of $g(\ )$ is a Gaussian function with a fixed cut-off T where

$$g(\Delta I) = \exp\left(-\frac{\Delta I^2}{2T^2}\right). \quad (4)$$

In our approach, we let T vary and tie it with the estimated noise STD function such that the filter can perform denoising "evenly" on both the high and the low noise areas. Our new spatially-varying photometric function $g_1(\ )$ is

$$g_1(\Delta I, I, p) = \exp\left(-k\cdot\frac{\Delta I^2}{f(I,p)^2}\right) \quad (5)$$

,where k is an optional constant which lets users choose the aggressiveness of the noise filter.

2.3. Luminance-guided Filtering for Chrominance Channels with Very Low SNR

A chrominance filtering method with luminance-guidance was developed to remove *high levels of chrominance noise* effectively. Due to the very low SNR in the chrominance channels, it is very difficult to remove the noise effectively and more sophisticated techniques are needed. Our approach is to filter the reliable luminance first and then use it to guide the process of filtering the chrominance. Note that the luminance signal is significantly less noisy and even sharper than the chrominance signal. The presence of the image details in the luminance channel is first detected, which guides how filtering is applied to the co-located pixels in the chrominance channels. A similar idea was presented earlier by Netraveli [11] for denoising TV signals. In our approach, we modified the conventional bilateral filter [2] such that the filtered luminance signal is also used for computing the taps of the chrominance filter. The luminance-guidance term allows us to apply more aggressive filtering without over-smoothing the details because the noise and the signal can be better separated with the additional information from the luminance channels. A specific implementation form for the luminance-guided filtering is shown in Equation (6) for the example of denoising $cb(i)$

$$cb^*(i) = \frac{\sum_j cb(i-j)\cdot g_2\big(cb(i-j)-cb(i), y^*(i-j)-y^*(i), \lambda\big)\cdot h(j)}{\sum_j g_2\big(cb(i-j)-cb(i), y^*(i-j)-y^*(i), \lambda\big)\cdot h(j)} \quad (6)$$

,where $y^*(i)$ is the denoised luminance value at pixel location i and λ is the weighting between the luminance guidance term and the self-guidance term. The equation can be obtained by replacing the generic photometric distance function $g(\ I(i) - I(i-j)\ )$ in Equation (3) with a luminance-guided photometric distance function $g_2(\ cb(i)-cb(i-j)\ ,\ y^*(i)-y^*(i-j)\ ,\lambda)$. In particular, when a Gaussian function is chosen for the photometric distance function, we obtain

$$g_2(\Delta cb, \Delta y^*, \lambda) = \exp\left(-\frac{\Delta cb^2 + \lambda\cdot(\Delta y^*)^2}{2T_2^{\ 2}}\right). \quad (7)$$

Note that the photometric distance function can be made strictly luminance-guided (i.e., without any self-guidance term Δcb^2) but occasionally produce undesired artifacts as the chrominance variations can occur with little or no luminance variations. Furthermore, we can let T_2 vary with the pixel intensities as we described in Section 2.2. The photometric distance function with both the luminance-guidance and the intensity dependence term is

$$g_3(\Delta cb, \Delta y^*, y^*, \lambda) = \exp\left(-k\cdot\frac{\Delta cb^2 + \lambda\cdot(\Delta y^*)^2}{f(y^*, \lambda, p_{Cb}, p_{y^*})^2}\right). \quad (8)$$

Note that the function $f(\ )$ is a function of the luminance intensities instead of the chrominance intensities. As the noise STD varies with the amount of photons captured at each pixel in the image sensor, the chrominance STD is better modeled as a function of the luminance intensities than the chrominance intensities.

2.4. A Filter Combining All Three Methods

In the final implementation, the three new approaches discussed in previous sections are implemented with the bilateral filter as a baseline. Prior to performing the denoising, we first transform the RGB image into YCbCr space and apply the noise estimator [8] to estimate the noise STD profile directly from the image. Three separate image pyramids are generated for Y, Cb and Cr channels and the methods described in Section 2.2 and 2.3 are applied in the Multi-resolution (MR) fashion (as described in Section 2.1). Less than three levels of pyramid (N+1<3) are used for the luminance channel while less than five levels of pyramid (N+1<5) are used for the chrominance channels, since the noise in the chrominance channels have higher spatial correlations. Note that the chrominance channels are typically coarsely quantized (especially for high-frequencies) and spatially subsampled during JPEG compression.

3. EXPERIMENTAL RESULTS

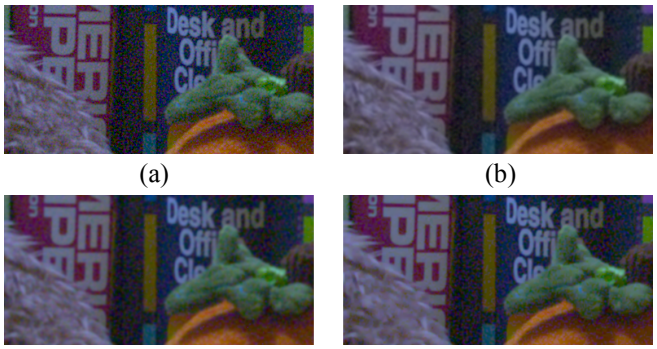
The objective of our work is to remove the noise in realistic digital photographs rendered by real digital cameras. In order to evaluate the denoising methods for such scenarios, the test images were obtained by accurately simulating the process of degradation in digital cameras. Each "noise-free" image was carefully generated by capturing 160 raw frames of the same scene and averaging those frames to virtually eliminate the noise. Noise was then added to the raw "noise-free" images by using a realistic model of solid-state image sensors [1]. The raw noisy images were then rendered by the generic pipeline of a digital camera. The rendering pipeline includes steps such as demosaicing, color correction, gamma correction, smoothing and sharpening. Optionally, the rendered images were compressed with the JPEG standard. Finally, several denoising methods were applied to these rendered noisy images. For quantitative comparisons, we also rendered the raw "noise-free" images by passing them through the same imaging pipeline. The rendered "noise-free" images served as the reference images (i.e., the ground truth) when evaluating the performances of the denoising methods. The performances were measured by computing the mean-squared-error (and hence the PSNR) between the denoised images and the rendered "noise-free" images which served as the ground truths.

We chose to compare our method with three nonlinear denoising methods (bilateral filter [2], BLS-GSM by Portilla et al. [12] and the anisotropic diffusion by Perona and Malik [13]) that have demonstrated good denoising results and have their code available publicly. Each method has a parameter (*T*, sigma and kappa respectively) that relate to the noise level of the input image and primarily control the aggressiveness of the denoising relative to preservation of the image features. Since these methods do not provide means to choose the parameters optimally without the prior knowledge of the noise power, we applied many sets of these parameters for fair comparisons. For all other parameters, default values and configurations provided by the authors were chosen (e.g., 5 by 5 Gaussian kernel for bilateral filter and 6 iterations for Anisotropic diffusion). See [2][12][13] for details. Since our method estimates the noise profile directly from the images, it was not necessary to apply multiple sets of parameters for our method. Table I summarizes the improvement in PSNR (relative to that of the input image with no denoising) of various methods with select parameters (*T*, sigma or kappa) denoted by the numbers in the parenthesis. Note that the images 1 and 2 are uncompressed images while the images 3 and 4 are compressed with JPEG (Q=75) and have compression artifacts in addition to the noise. It can be seen that our method gives the highest PSNR improvement. Note the gap between the conventional bilateral filter and our method which uses bilateral filter as a base method.

We also show qualitative results since PSNR cannot fully capture human visual preferences. Figure 2 shows cropped versions of the noisy input image 2 and the results from various denoising methods. Note that the noise in the input image has significant low-frequency chrominance component (visible as large and colorful grains) which has vastly different characteristics from AWGN. Our method was able to remove high levels of spatially-correlated noise while the details are maintained. Other methods, however, failed to remove the spatially-correlated noise or blur the image excessively.

Input image	Image1	Image2	Image3	Image4
Bilat (60)	1.77	3.81	2.13	2.45
Bilat (120)	3.13	4.40	2.55	2.53
AnisoDiff (60)	1.97	5.05	2.95	3.10
AnisoDiff (120)	4.76	5.41	3.36	3.05
BLS_GSM (60)	4.30	6.56	3.80	3.58
BLS_GSM (120)	7.38	4.78	5.08	1.73
Our MR method	7.59	7.36	5.10	4.51

Table I: PSNR improvement in dB for various denoising methods



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

Figure 2: Denoising results for (a) input image, (b) our method, (c) anisotropic diffusion (kappa=120), (d) BLS_GSM (sig=20)

4. REFERENCES

[1] S. H. Lim, "Characteristics of Noise in Digital Photographs for Image Processing", *Proceedings of SPIE - The International Society for Optical Engineering* 6069, article no. 60690O, January 2006

[2] C. Tomasi and R. Manduchi, "Bilateral filtering for gray and color images", in *Proceedings of the IEEE International Conference on Computer Vision*, Bombay, India, 1998

[3] M. Elad, "On the Origin of the Bilateral Filter and Ways to Improve It", *IEEE Transactions of Image Processing*, v 11, p. 1141~1151, October 2002

[4] D. L. Dononho, "Denoising by soft-thresholding", *IEEE Transactions on Information Theory*, p.613~622, December 1995

[5] A. S. Wilsky, "Multiresolution Markov Models for Signal and Image Processing", *Proceedings of the IEEE*, v. 90, p. 1396~1458, August 2002

[6] B. Aiazzi, L. Alparone, S. Baronti and G. Borri, "Pyramid-Based Multiresolution Adaptive Filters for Additive and Multiplicative Image Noise", *IEEE Transactions on Circuits and Systems – II: Analog and digital signal processing*, v. 45, n. 8, p. 1092~1097, August 1998

[7] S. Gezici, I. Yilmaz, O.N. Gerek, E. Cetin, "Image Denoising using adaptive subband decomposition", *IEEE International Conference on Image Processing*, v. 1, p. 261~264, Oct 2001

[8] P. Kisilev, D. Shaked, S. H. Lim, "Noise and signal activity maps for better imaging algorithms," *IEEE International Conference on Image Processing* 2007

[9] H. Faraji, W. J. MacLean, "CCD noise removal in digital images," *IEEE Transactions of Image Processing*, vol.15, no.9, p. 2676~2685, Sept. 2006

[10] Argenti, F.; Torricelli, G.; Alparone, L., "Signal-dependent noise removal in the undecimated wavelet domain," *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing, 2002,* vol.4, no.pp. IV-3293~3296 vol.4, 2002

[11] A. Netravali, "Noise Removal from Chrominance Components of a Color Television Signal," *IEEE Transactions on Communications*, vol.26, no.8, p. 1318~1321, Aug 1978

[12] J. Portilla, V. Strela, M. Wainwright, and E. Simoncelli, "Image denoising using scale mixtures of Gaussian in the wavelet domain," *IEEE Trans. on Image Processing*, vol. 12, no. 11, p. 1338~1351, Nov. 2003.

[13] P. Perona and J. Malik, "Scale-space and edge detection using anisotropic diffusion," *IEEE Pattern Analysis and Machine Intelligence*, vol. 12, no. 7, p. 629~639, Jul. 1990.