An Adaptive Monte Carlo Approach to Nonlinear Image Denoising

Alexander Wong, Akshaya Mishra, Paul Fieguth, and David Clausi
{a28wong,akmishra,pfieguth,dclausi}@engmail.uwaterloo.ca
Systems Design Engineering, University of Waterloo, Canada

Abstract

This paper introduces a novel stochastic approach to image denoising using an adaptive Monte Carlo scheme. Random samples are generated from the image field using a spatially-adaptive importance sampling approach. Samples are then represented using Gaussian probability distributions and a sample rejection scheme is performed based on a χ^2 statistical hypothesis test. The remaining samples are then aggregated based on Pearson Type VII statistics to create a non-linear estimate of the denoised image. The proposed method exploits global information redundancy to suppress noise in an image. Experimental results show that the proposed method provides superior noise suppression performance both quantitatively and qualitatively when compared to the state-of-the-art image denoising methods.

1 Introduction

A fundamental problem faced in the field of image processing is that of image denoising, where the goal is to obtain an estimate of the original image from an image that has been contaminated by noise. Image denoising is very important for numerous applications ranging from photo enhancement and edge detection to object recognition and tracking. While many image denoising algorithms have been proposed over the years, good noise suppression remains an open challenge, particularly in situations characterized by low signal-to-noise ratios.

Image denoising algorithms can be divided into two main groups: i) transform domain denoising, and ii) spatial domain denoising. In transform domain denoising methods, the image is transformed into an alternate domain (e.g., frequency domain) and the domain coefficients are modified to suppress noise. These methods include global transform denoising approaches such as Wiener filtering [10] and local transform denoising approaches such as wavelet thresholding [4] and the Gaussian scale mixture (GSM) denoising method [7]. In spatial domain denoising methods, pixel intensities are modified directly in the spatial domain by exploiting information redundancy in images. A majority of spatial domain denoising methods rely on local image redundancy to suppress image noise. These methods will be referred to as local spatial domain denoising methods and include traditional approaches such as box filtering and Gaussian filtering, as well as newer detail-preserving approaches such as anisotropic filtering [3] and bilateral filtering [9, 2]. While computationally efficient, local spatial domain denoising methods typically perform poorly in situations characterized by low signal-to-noise ratios as the local statistics would be insufficient for providing a good estimate of the denoised image. To address this issue, more recent approaches take advantage of global information redundancy to suppress noise [5, 1]. These methods will be referred to as global spatial domain denoising methods. The main drawback to global spatial domain denoising methods is that they are computationally expensive given the amount of information processing involved to obtain estimates for the denoised image, even after various performance optimizations [5].

In this paper, we propose a novel approach to image denoising using an adaptive Monte Carlo scheme. The proposed method takes a radical departure from most existing denoising techniques in that it is stochastic in nature. As such, the denoised image estimate will vary with each run. The proposed method utilizes global information redundancy in an efficient manner and is highly robust to images characterized by low signal-to-noise ratios. Furthermore, the proposed method provides effective noise suppression while preserving image detail. This paper is organized as follows. The mathematical background behind the proposed method is presented in Section 2. Experimental results are presented and discussed in Section 3. Finally, conclusions are drawn in Section 4.

978-1-4244-2175-6/08/$25.00 ©2008 IEEE

2 Mathematical Background

Let f represent a 2-D image and n represent random noise. The noisy image g can be written as

$$g(\underline{x}) = f(\underline{x}) + n(\underline{x}), \tag{1}$$

where $\underline{x} = (x, y)$. Image denoising can be regarded as an inverse problem, where the underlying goal is to estimate f given g. In local spatial domain denoising methods, an estimate of f is defined as the weighted integral over a local neighborhood ς in g,

$$\hat{f}(\underline{x}) = \int_{\underline{x} \in \varsigma} w(\underline{x}) g(\underline{x})\, d\underline{x}, \tag{2}$$

where w represents the weighting function. Global spatial domain denoising methods extend this concept by taking the entire image g into consideration when estimating f,

$$\hat{f}(\underline{x}) = \int_{\underline{x} \in g} w(\underline{x}) g(\underline{x})\, d\underline{x} \tag{3}$$

While this approach to estimating f fully exploits global information redundancy, it is completely impractical to evaluate $\hat{f}$ deterministically from a computational perspective. To address this important issue, we propose that we find a stochastic solution for $\hat{f}$ instead in an efficient manner using a novel adaptive Monte Carlo scheme.

Let us consider g as a 2-D random field G, X be a random variable in G, and p be an arbitrary probability density function on G. If we were to take n random samples $\underline{x}_1, \ldots, \underline{x}_n$ based on p, the Monte Carlo estimate of f can be defined as

$$\hat{f}(\underline{x}) = \sum_{i=1}^{n} w(\underline{x}_i) g(\underline{x}_i). \tag{4}$$

One problem with this "naive" Monte Carlo approach is that it generates too many samples that provide too little useful information about the point being estimated to achieve the desired solution. This can lead to poor and inconsistent noise reduction as well as an increase in computational overhead. To address this issue, we introduce a two-step approach for improving the information redundancy of samples used in estimating f.

One of the reasons that the "naive" Monte Carlo method generates many irrelevant samples is due to the fact that the samples are taken based on an arbitrary probability density function. This results in large estimation variances since the probability density distribution used in generating samples can be significantly different than the probability density distribution of samples with high information redundancy with respect to the point in f being estimated. A particular effective variance reduction technique that can be used in the Monte Carlo method is importance sampling [8]. The idea of importance sampling is that random variables with greater impact on the estimate (in this case, points with high information redundancy with respect to the point being estimated) should be sampled more frequently. This in effect reduces estimation variance and improves the consistency of the solution. In the proposed method, we utilize an importance sampling scheme that is spatially adaptive with respect to the point being estimated. The biasing density κ can be expressed as

$$\kappa(\underline{x}) = \frac{1}{(\underline{x} - \underline{x}_c)^{\alpha}}, \tag{5}$$

where $\underline{x}_c$ is the point in f being estimated and α is the importance factor (based on testing, a suitable importance factor is $\alpha = 1.3$). This biasing density is consistent with the assumption made by local spatial domain denoising methods that points that have high information redundancy are usually close to the point being estimated. The advantage of this biasing density is that while it encourages spatial locality, it also exploits global information redundancy by allowing for the possibility of distant samples with high information redundancy.

While the proposed spatially-adaptive importance sampling scheme improves the consistency of the solution by sampling more frequently in areas where high information redundancy are more likely to occur, it does not actually validate the information redundancy of the samples. Therefore, there may still be irrelevant samples that provide little useful information about the point being estimated. To address this issue, we propose the use of a statistical sample rejection scheme to further improve information redundancy of samples used in estimating f. Let each random sample $\underline{x}_1, \ldots, \underline{x}_n$ generated by the proposed importance sampling scheme be represented by a different Gaussian probability density function $q_1, \ldots, q_n$,

$$q(\underline{x}_i) \sim N(\mu_i, \sigma_i^2) \tag{6}$$

where μ_i and σ_i^2 are the mean and variance within a local neighborhood around sample $\underline{x}_i$ (based on testing, a suitable neighborhood is a circular neighborhood with radius 3). Similarly, the point being estimated is represented by a Gaussian probability function $q(\underline{x}_c) \sim N(\mu_c, \sigma_c^2)$. The information redundancy of each sample with respect to the point being estimated is evaluated using a χ^2 statistical hypothesis test,

$$\chi^2(q(\underline{x}_i), q(\underline{x}_c)) = \frac{\left(\frac{|\mu_i - \mu_c|}{\sigma_c}\right)^{-1} \exp\left[-\left(\frac{\mu_i - \mu_c}{\sqrt{2}\sigma_c}\right)^2\right]}{\sqrt{2}\Gamma(1/2)}, \tag{7}$$

where Γ is the Gamma function. Sample rejection is then performed on the set of random samples based on the following criterion,

$$\chi^2(q(\underline{x}_i), q(\underline{x}_c)) > T \tag{8}$$

where T is the rejection threshold (based on testing, a suitable rejection threshold is $T = 0.5$). Samples that do not satisfy this criterion are pruned from the set of samples used to estimate f.

Given m remaining samples $\underline{x}_1, \ldots, \underline{x}_m$ after sample rejection, where $m \leq n$ (in most cases, $m \ll n$), it is now important to determine an estimator for computing $\hat{f}$ based on information associated with each sample $g(\underline{x}_1), \ldots, g(\underline{x}_m)$. We propose a new robust nonlinear estimator based on Pearson Type VII statistics [6],

$$\hat{f}(\underline{x}) = \frac{\sum_{i=1}^{m} g(\underline{x}_i) e^{-\frac{\sum \ln\left(\sqrt{1+(g(\varsigma_i)-g(\varsigma_c))^2}\right)}{|\sigma_c^2 - \overline{\sigma^2}|+\epsilon}}}{\sum_{i=1}^{m} e^{-\frac{\sum \ln\left(\sqrt{1+(g(\varsigma_i)-g(\varsigma_c))^2}\right)}{|\sigma_c^2 - \overline{\sigma^2}|+\epsilon}}} \tag{9}$$

where ς_i and ς_c are local neighborhoods at $\underline{x}_i$ and $\underline{x}_c$ respectively, $\overline{\sigma^2}$ is the minimum local variance and σ_c^2 is the local variance at $\underline{x}$, and ϵ is a small regularization constant. It can be seen from the above formulation that the estimator adapts for each individual point in the image based on its image characteristics. This nonlinear estimation approach allows for noise suppression that preserves image detail even in situations characterized by low signal-to-noise ratio.

3 Experimental Results

The proposed adaptive Monte Carlo (AMC) denoising method was applied to four test images with different characteristics. Each test image is contaminated by white Gaussian noise with standard deviation of 26. To evaluate the performance of the proposed method in a quantitative manner, the peak signal-to-noise ratio (PSNR) of the denoised image was measured for the proposed AMC method, as well as state-of-the-art denoising methods such as bilateral filtering (BF) [9], Gaussian scale mixture (GSM) denoising [7], and nonlocal means (NLM) denoising [1] for comparison. It should be noted that the number of samples automatically selected by the AMC method to estimate each point in the denoised image is on average around 50 samples, which is substantially less than that used by global spatial domain denoising methods such as NLM. A summary of the PSNR results is shown in Table 1. The proposed method achieves noticeable PSNR gains over other test methods for all test images. The denoised images for the Lena, Boat, and Barbara images are shown in Figure 1. Upon visual inspection, it can be observed that the proposed AMC method produced denoised images with noticeably improved perceptual quality when compared to the other test methods. While noise was effectively suppressed by all of the tested methods, fine image detail such as edges are noticeably better preserved using the proposed AMC method than the other test methods.

Table 1. PSNR for Test Images

Test	PSNR (dB)			
	GSM [7]	BF [9]	NLM [1]	AMC[1]
Lena (19.53)	25.01	24.19	24.60	25.79
Boat (19.56)	25.19	24.58	24.76	25.99
House (19.51)	25.74	24.86	25.81	26.49
Barbara (19.58)	24.26	23.60	24.47	25.81

[1] PSNR is computed as the mean of 10 trials given the randomness inherent in the sampling process.

4 Conclusions

In this paper, a novel stochastic denoising method based on an adaptive Monte Carlo (AMC) scheme was introduced. The proposed method utilizes the global information redundancy in images in an efficient and adaptive manner through the use of a spatially-adaptive importance sampling scheme and a statistical sample rejection scheme. An adaptive nonlinear estimation scheme was introduced to compute estimates of the denoised image based on image characteristics to better preserve image detail while suppressing noise. Experimental results show that noticeably improved perceptual quality can be achieved using the proposed method when compared to current state-of-the-art image denoising methods. Future work involves investigating the effectiveness of alternative sampling and rejection schemes for improving noise suppression and detail preservation.

(a) Lena

(b) Boat

(c) Barbara

Figure 1. Denoising using different denoising methods (from left to right): a) Noisy image (σ=26), b) BF c) GSM, d) NLM, e) AMC

Acknowledgment

This research has been sponsored by the Natural Sciences and Engineering Research Council of Canada.

References

[1] A. Buades, B. Coll, and J. Morel. Nonlocal image and movie denoising. *International Journal of Computer Vision*, 76(2):123–139, 2008.

[2] M. Elad. On the origin of the bilateral filter and ways to improve it. *IEEE Transactions on Image Processing*, 11(10):1141–1151, 2002.

[3] S. Greenberg and D. Kogan. Improved structure-adaptive anisotropic filter. *Pattern Recognition Letters*, 27(1):59–65, 2006.

[4] Q. Li and C. He. Application of wavelet threshold to image de-noising. In *Proceedings of ICICIC*, volume 2, pages 693–696, 2006.

[5] M. Mahmoudi and G. Sapiro. Fast image and video denoising via nonlocal means of similar neighborhoods. *IEEE Signal Processing Letters*, 12(12):839–842, 2005.

[6] K. Pearson. Mathematical contributions to the theory of evolution, xix: Second supplement to a memoir on skew variation. *Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character*, 216:429457, 1916.

[7] J. Portilla, V. Strela, M. Wainwright, and E. Simoncelli. Image denoising using scale mixtures of gaussians in the wavelet domain. *IEEE Transactions on Image Processing*, 12(11):1338–1351, 2003.

[8] R. Srinivasan. *Importance sampling - Applications in communications and detection*. Springer-Verlag, Berlin, 2002.

[9] C. Tomasi and R. Manduchi. Bilateral filtering for gray and color images. In *Proceedings of ICCV*, pages 836–846, 1998.

[10] N. Wiener. *Extrapolation, Interpolation, and Smoothing of Stationary Time Series*. Wiley, New York, 1949.