21st International Conference on Pattern Recognition (ICPR 2012)
November 11-15, 2012. Tsukuba, Japan

Bilateral Filter Evaluation Based on Exponential Kernels

Kassem Al-Ismaeil Djamila Aouada Björn Ottersten
Interdisciplinary Centre for Security, Reliability and Trust
University of Luxembourg
kassem.alismaeil@uni.lu djamila.aouada@uni.lu bjorn.ottersten@uni.lu

Bruno Mirbach
Advanced Engineering Department, IEE S.A.
bruno.mirbach@iee.lu

Abstract

The well-known bilateral filter is used to smooth noisy images while keeping their edges. This filter is commonly used with Gaussian kernel functions without real justification. The choice of the kernel functions has a major effect on the filter behavior. We propose to use exponential kernels with L^1 distances instead of Gaussian ones. We derive Stein's Unbiased Risk Estimate to find the optimal parameters of the new filter and compare its performance with the conventional one. We show that this new choice of the kernels has a comparable smoothing effect but with sharper edges due to the faster, smoothly decaying kernels.

1. Introduction

Image denoising is a common image restoration procedure. The main challenge is to find a good image denoising technique that removes noise while preserving image features such as edges and texture. Over the past three decades, many algorithms have been proposed. One common approach is to use the bilateral filter (BF) [7]. This filter is a weighted average of the local neighborhood pixels. The weighting is based on the product of two kernel functions; one spatial using the distance between the location of the center pixel and the location of the neighboring pixels. The second kernel is radiometric, and uses the distance between the intensity of the center pixel and the intensity of the neighboring pixels. Each weighting kernel is controlled by a parameter determining its width. These kernels are commonly chosen to be Gaussian functions with mean zero. Stein's Unbiased Risk Estimate (SURE) has been used to find the optimal widths of the Gaussians, i.e., their standard deviations [6], [4], [1]; the objective being to find a trade-off between image smoothing and edge preservation while minimizing SURE risk function, an estimator of the mean square error (MSE) between the noisy image and the filtered one.

As mentioned by Elad [2], as long as the kernel functions used in the BF are smoothly decaying and symmetric, they can be chosen in place of the Gaussian functions. However, very little work exists using bilateral filters with a different kernel. In [3], Farsiu et al. used an exponential kernel in their implementation of the BF, but no justification was given for this choice. It is clear that an adequate choice of the kernels may lead to a good filter performance. We further argue that a faster decaying kernel would ensure sharper edges while smoothing the rest of the image. The question is whether exponential kernels fall under this category. We propose in this paper to answer this question. We thus compare the performance of the BF using Gaussian kernels, that we refer to as $\mathrm{BF}_{\mathrm{Gauss}}$, and the BF using exponential kernels, that we refer to as $\mathrm{BF}_{\mathrm{exp}}$. We derive the SURE risk function for $\mathrm{BF}_{\mathrm{exp}}$ in order to find the filter optimal parameters. Our simulations show that for different levels of noise, $\mathrm{BF}_{\mathrm{exp}}$ consistently gives a lower or equal MSE and always provides a final image that is visually better. Given that $\mathrm{BF}_{\mathrm{exp}}$ and $\mathrm{BF}_{\mathrm{Gauss}}$ are computationally comparable, in view of our results, $\mathrm{BF}_{\mathrm{exp}}$ is at least similar to $\mathrm{BF}_{\mathrm{Gauss}}$.

The paper is organized as follows: in Section 2 we briefly review the BF and give the formulation for $\mathrm{BF}_{\mathrm{exp}}$. In Section 3, we present the corresponding parameter estimation. We give a comparison of the two filters in Section 4. We present our simulations in Section 5. Finally, we summarize this paper in Section 6.

2. Review of bilateral filtering

Let $\mathbf{x}$ be an $(m \times n)$ noise-free image degraded by added zero-mean white Gaussian noise $\mathbf{w}$ of variance σ^2 and of the same size. The observed corrupted image $\mathbf{y}$ is given by

$$\mathbf{y} = \mathbf{x} + \mathbf{w}. \tag{1}$$

The BF recovers the original image $\mathbf{x}$ by a nonlinear filtering that replaces the noisy intensity value $\mathbf{y_p}$ at each pixel location $\mathbf{p}$ with a weighted average of the neighboring pixels $\mathbf{q}$, i.e., $\mathbf{q} \in \mathcal{N}(p)$, such that:

$$\hat{\mathbf{x}}_\mathbf{p} = \frac{\sum\limits_{\mathbf{q}\in\mathcal{N}(\mathbf{p})} f_S(\mathbf{p},\mathbf{q}) f_R(\mathbf{p},\mathbf{q}) \cdot \mathbf{y_q}}{\sum\limits_{\mathbf{q}\in\mathcal{N}(\mathbf{p})} f_S(\mathbf{p},\mathbf{q}) f_R(\mathbf{p},\mathbf{q})}. \tag{2}$$

The weighting kernel $f_S(\mathbf{p},\mathbf{q})$ is based on the distance between $\mathbf{p}$ and $\mathbf{q}$, and $f_R(\mathbf{p},\mathbf{q})$ is based on a radiometric distance, i.e., the difference between the two pixel intensities $\mathbf{y_p}$ and $\mathbf{y_q}$. We write the final filtered image as $\hat{\mathbf{x}} = BF(\mathbf{y},\boldsymbol{\theta})$ with $\hat{\mathbf{x}} = [\hat{\mathbf{x}}_\mathbf{p}]_{\mathbf{p}\in[1:m]\times[1:n]}$ and $\boldsymbol{\theta}$ being the vector containing the filter parameters. The two kernels have to verify two properties: 1) symmetry, and 2) smooth decay. Conventionally, these functions are taken as Gaussians with an L^2 norm (Euclidean distance) and parameterized by (λ_g, β_g) . That is BF_{Gauss} is defined by:

$$\begin{cases} f_S(\mathbf{p},\mathbf{q}) = \exp\left(-\dfrac{\| \mathbf{p}-\mathbf{q} \|_2^2}{2\lambda_g}\right) \\ f_R(\mathbf{p},\mathbf{q}) = \exp\left(-\dfrac{|\mathbf{y_p}-\mathbf{y_q}|^2}{2\beta_g}\right) \end{cases} \tag{3}$$

Another choice for the kernels is the exponential function with an L^1 norm (Manhattan distance). The base of the exponential defines the width of the kernel and needs to be in the interval $]0,1[$ to verify Property 2). The resulting BF_{exp} is defined by:

$$\begin{cases} f_S(\mathbf{p},\mathbf{q}) = a_e^{\|\mathbf{p}-\mathbf{q}\|_1} & = \exp\left(\| \mathbf{p}-\mathbf{q} \|_1 \cdot \ln a_e\right) \\ f_R(\mathbf{p},\mathbf{q}) = b_e^{|\mathbf{y_p}-\mathbf{y_q}|} & = \exp\left(|\mathbf{y_p}-\mathbf{y_q}| \cdot \ln b_e\right) \end{cases} \tag{4}$$

with $0 < a_e, b_e < 1$. For the sake of comparison, we similarly define the bounded parameters of BF_{Gauss} as $a_g = e^{-\frac{1}{2\lambda_g}}$, and $b_g = e^{-\frac{1}{2\beta_g}}$. Comparing the two filters BF_{Gauss} and BF_{exp} passes through comparing the two parameter vectors $\boldsymbol{\theta}_g = [a_g, b_g]^T$, and $\boldsymbol{\theta}_e = [a_e, b_e]^T$. We note that the main difference between the kernels is in the square in the exponent of the Gaussian kernels, that we will see in Section 4, plays a role in the difference in performance.

3. Parameter estimation for bilateral filtering

The quality of the denoised image $\hat{\mathbf{x}}$ is very dependent on the choice of the filter parameters, $\boldsymbol{\theta}$ in general. To optimally set these parameters, we use SURE as an unbiased estimator of the MSE, obtained from the observed noisy image $\mathbf{y}$. Indeed, the quality of the denoising technique is measured by:

$$\text{MSE}(\hat{\mathbf{x}}) = \frac{1}{mn} \| \mathbf{x} - \hat{\mathbf{x}} \|_2^2 . \tag{5}$$

An unbiased estimator of (3) is given in [6], and defined as the following SURE risk function:

$$R_{\boldsymbol{\theta}} = \frac{1}{mn} \| \mathbf{y} - \hat{\mathbf{x}} \|_2^2 - \sigma^2 + 2\frac{\sigma^2}{mn} \text{div}_\mathbf{y}\{\hat{\mathbf{x}}\}, \tag{6}$$

where $\text{div}_\mathbf{y}\{\hat{\mathbf{x}}\}$ is the divergence of the denoising filter BF (e.g., BF_{Gauss} or BF_{exp}) with respect to the observed image such that:

$$\text{div}_\mathbf{y}\{\hat{\mathbf{x}}\} = \sum_{\mathbf{l}\in[1:m]\times[1:n]} \frac{\partial \hat{\mathbf{x}}_\mathbf{l}}{\partial \hat{\mathbf{y}}_\mathbf{l}}. \tag{7}$$

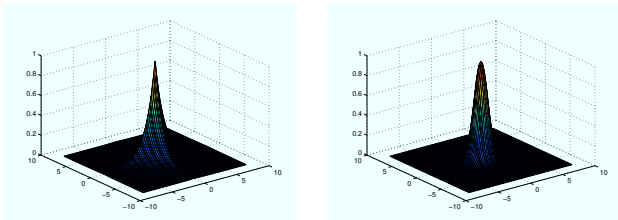
Finding the optimal $\boldsymbol{\theta}$ follows as: $\hat{\boldsymbol{\theta}} = \underset{\boldsymbol{\theta}}{\text{argmin}}\ \hat{R}_{\boldsymbol{\theta}}$. In practice, the noise variance σ^2 can easily be estimated from the observed data.

In case of BF_{Gauss}, (3) is given in [6]. We herein give the derivation for the case of the proposed BF_{exp}. We first define $f_{SR}(\mathbf{p},\mathbf{q}) = a_e^{\|\mathbf{p}-\mathbf{q}\|_1} b_e^{|\mathbf{y_p}-\mathbf{y_q}|}$, then:

$$\begin{aligned} \frac{\partial \hat{\mathbf{x}}_\mathbf{p}}{\partial \mathbf{y_p}} &= \frac{\partial \left[\dfrac{\sum\limits_{\mathbf{q}\in\mathcal{N}(\mathbf{p})} f_{SR}(\mathbf{p},\mathbf{q})\mathbf{y_q}}{\sum\limits_{\mathbf{q}\in\mathcal{N}(\mathbf{p})} f_{SR}(\mathbf{p},\mathbf{q})} \right]}{\partial \mathbf{y_p}} \\ &= \ln b_e \left[\frac{\sum\limits_{\mathbf{q}\in\mathcal{N}(\mathbf{p})} f_{SR}(\mathbf{p},\mathbf{q})\text{sign}(\mathbf{y_p}-\mathbf{y_q})\mathbf{y_q}}{\sum\limits_{\mathbf{q}\in\mathcal{N}(\mathbf{p})} f_{SR}(\mathbf{p},\mathbf{q})} \right] \\ &- \ln b_e \left[\frac{\sum\limits_{\mathbf{q}\in\mathcal{N}(\mathbf{p})} f_{SR}(\mathbf{p},\mathbf{q})\text{sign}(\mathbf{y_p}-\mathbf{y_q})}{\sum\limits_{\mathbf{q}\in\mathcal{N}(\mathbf{p})} f_{SR}(\mathbf{p},\mathbf{q})} \right] \\ &\times \left[\frac{\sum\limits_{\mathbf{q}\in\mathcal{N}(\mathbf{p})} f_{SR}(\mathbf{p},\mathbf{q})\mathbf{y_q}}{\sum\limits_{\mathbf{q}\in\mathcal{N}(\mathbf{p})} f_{SR}(\mathbf{p},\mathbf{q})} \right], \end{aligned}$$

where sign$(\cdot)$ is the sign function. We thus find the optimal $\boldsymbol{\theta}_e$ that ensures the best possible denoising using $\mathrm{BF}_{\mathrm{exp}}$. Similarly we find the optimal $\boldsymbol{\theta}_g$ that ensures the best possible denoising using $\mathrm{BF}_{\mathrm{Gauss}}$.

4. Comparison of the two bilateral filters



(a) Exponential kernel (b) Gaussian kernel

Figure 1. Exponential and Gaussian kernels.

Both exponential and Gaussian kernel functions are symmetric and smoothly decaying functions as depicted in Figure 1. However, the decay of the exponential kernel is faster which should achieve sharper edges.
BF is about finding a trade-off between the parameters; spacial and radiometric. These parameters, $\boldsymbol{\theta}_g$ and $\boldsymbol{\theta}_e$, control the kernels decay. Small parameter values give a simple uniform non-adaptive filtering which is known to degrade the image edges, and large values reduce the smoothing effect. As illustrated in Figure 2(b), the optimized radiometric parameters, both b_g and b_e, are almost the same for both kernels. On the other hand, the spatial parameters shown in Figure 2(a) of the Gaussian kernel decrease by increasing the noise level compared to the exponential. Thus, the exponential kernel leads to sharper edges (Figure 4(d)) than the Gaussian kernel illustrated in Figure 4(c).

5. Experimental results

In our experiments we illustrate the performance of the BF using the proposed kernel compared to the standard Gaussian kernel. First, we ran a Monte-Carlo simulation over 50 normalized noisy images by adding white Gaussian noise with a noise variance varying from 1% to 10%. At each noise level, we denoise the images by a BF with the proposed exponential kernel and the standard Guassian kernel. The spatial and radiometric smoothing parameters for both kernels were optimized based on the SURE approach. In Figure. 3, the average root mean square error (RMSE) for both kernels is illustrated, where we can see that the proposed

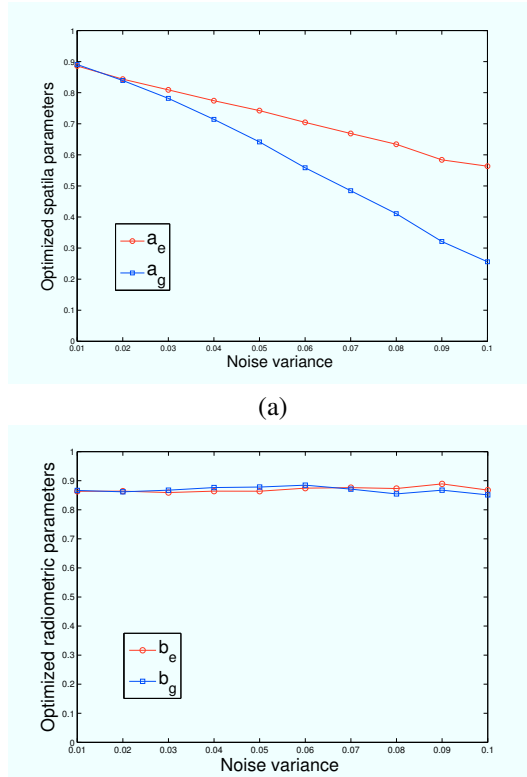


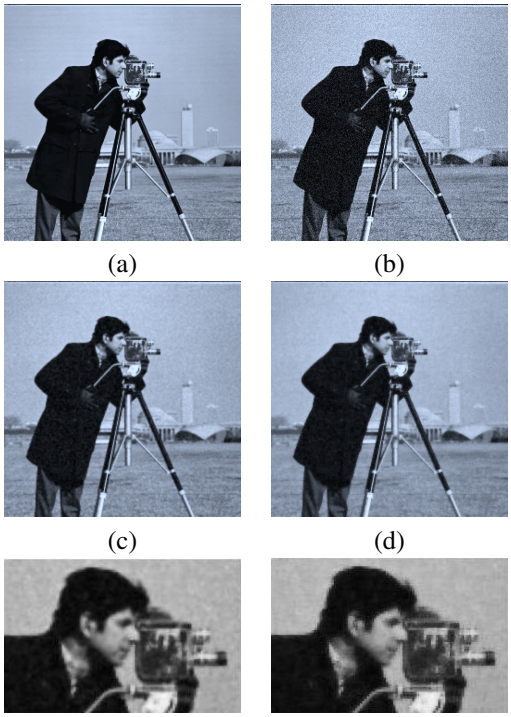
Figure 2. Optimized parameters: (a) spatial, (b) radiometric.

$\mathrm{BF}_{\mathrm{exp}}$ performs better than the standard $\mathrm{BF}_{\mathrm{Gauss}}$ for this 'cameraman' example. Moreover, the proposed kernel shows it superiority over the standard Gaussian where it leads to a visual improvement in denoising results as shown in Figure. 4.

Next, we tested our algorithm on a 1D signal by adding a noise of variance $\sigma = 0.05$. As illustrated in Figure 5, the exponential kernel $\mathrm{BF}_{\mathrm{exp}}$ illustrated in blue, gives a result that is closer to the original noise-free signal, confirming its better performance.

6. Conclusion

Tomasi and Manduchi have proposed the bilateral filter as a noise removal algorithm for images. In this work we have proposed to use the exponential kernel as an alternative to the standard Gaussian commonly used by the community. We verified that the proposed kernel is numerically better than the standard Gaussian for image denoising. Moreover, we showed that the optimum spatial and radiometric parameters provided by the exponential kernel lead to a better trade-off between blur-



(a) (b) (c) (d) (e) (f)

Figure 4. Denoising example: (a) original image, (b) noisy image(σ=0.08), (c)-(d) denoised images using exponential and Gaussian kernels, respectively.(e)-(f) zoomed patch for BF_{Gauss} and BF_{exp}, respectively .

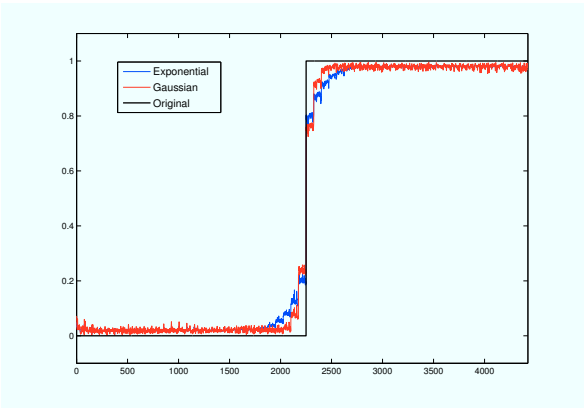


Figure 5. Illustration on denoising a 1-D signal. See the text for explanation.

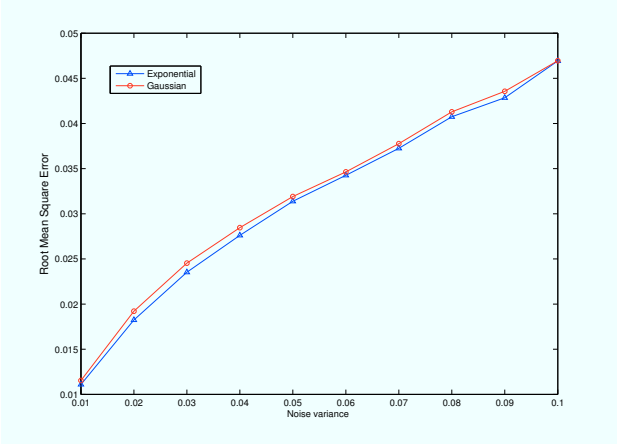


Figure 3. RMSE of bilateral filter using exponential and Gaussian kernels .

ring and denoising, thus suppressing noise while preserving edges.

7 Acknowledgment

This work was supported by the National Research Fund (FNR), Luxembourg, under the CORE project C11/BM/1204105/FAVE/Ottersten.

References

[1] M. K. D. V. Ville. Sure-based non-linear means. *Transactions on image processing*, 20:2683–2690, 2011.

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

[3] S. Farsiu, D. Robinson, M. Elad, and P. Milanfar. Fast and robust multi-frame super-resolution. *IEEE TIP*, 13:1327–1344, 2003.

[4] M. U. S. Ramani, T. Blu. Monte-carlo sure: a black-box optimization of regularization parameters for general denoizing algorithm. *Transactions on image processing*, 17:1540–1554, 2008.

[5] V. Solo. Sure-fired way to choose smoothing parameters in ill-conditioned inverse problems. *Int. Conf. Image Processing*, III:89–92, 1996.

[6] C. Stein. Estimation of the mean of multivariate normal distribution. *Ann. Statist.*, 9:1135–1151, 1981.

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

[8] C. Yonina. Generalized sure for exponential families applications to regularization. *Transactions on signal processing*, 57:471–481, 2009.