MULTILATERAL FILTERING: A NOVEL FRAMEWORK FOR GENERIC SIMILARITY-BASED IMAGE DENOISING

Irfan T. Butt, Nasir M. Rajpoot

Department of Computer Science, University of Warwick, Coventry CV4 7AL, UK
Correspondence email: `nasir@dcs.warwick.ac.uk`

ABSTRACT

We present a novel iterative nonlinear filtering framework, termed multilateral filtering, based on the idea of generic local similarity. A set of local features is computed for each pixel using its local neighborhood. Two pixels are considered to be similar if the Euclidean distance between their corresponding feature vectors is small and vice versa. Multilateral filtering results in image smoothing while preserving edge and textural features. Our experimental results show that the proposed method produces comparable and often better results than the state-of-the-art denoising methods.

Index Terms— Image denoising, Edge preservation, Bilateral filtering

1. INTRODUCTION

Noise in an image is unwanted and random variations of pixel values caused during image acquisition and/or transmission. Filtering in spatial or transform domain is one of the most fundamental tools used for noise removal in images. The vast amount of literature on filtering for noise removal in digital images can be broadly classified into two main categories: linear filtering and non-linear filtering.

Linear smoothing or filtering is widely used in those areas where noise free blur images are required. These linear filtering methods are also referred to as low-pass filtering or weighted averaging, such as Gaussian low-pass filtering whereby weights are function of the distance among pixels. In other words, these filters give larger weights to nearby pixels and smaller weights to faraway pixels in a certain neighborhood. In real-world images, the assumption that a spatial variation in pixels is quite slow fails at edges due to the non-stationary nature of real-world images. Although linear filtering is computationally tractable in spatial or frequency domain, the real dilemma with this type of methods is that they often result in over-smoothing (also known as blurrinng) of important features such as edges. To acquire noise-free edge-preserving images, many non-linear filtering techniques have been developed over the years. Among these methods, bilateral filtering [1] is one of the most popular approaches. Bilateral filtering is an extension of the Gaussian low-pass filtering approach as it contains *domain filters* that measure the spatial locality of pixels and *range filters* that measure the photometric similarities among pixels in close neighborhood. In other words, bilateral filtering computes weighted average of pixel values in the neighborhood, whereby the weights decrease as distance from the neighborhood center increases or difference between intensity value among pixels increases. Although this combination of domain and range filters is quite effective on denoising images as compared to simple domain filter like weighted averaging, bilateral filtering face a dilemma that if the pixel at the origin is corrupted with noise, larger weight will be assigned to other pixels having similar intensity to that corrupted pixel by range filter. Garnett *et al.* [2] have proposed a trilateral filter employing a local image statistic for identifying noise pixels. Another trilateral filter was presented in [3] for high contrast images and meshes.

In this paper, we present a generic framework for denoising of images using the similarity idea. We propose an extension of the non-linear bilateral filtering, termed as *multilateral filtering*, for denoising of images corrupted with noise and consisting of smooth regions, edge features, as well as textured areas. The proposed filtering operates simultaneously in three or more dimensions: *domain* filtering, *range* filtering, and K-dimensional *feature* filtering for $K \geq 1$. The first two dimensions are the same as those employed in bilateral filtering. In the larger dimensions, filtering is done based on similarity of a pixel in the neighborhood to its center pixel. A set of local features is computed for each pixel using its local neighborhood. Filtering based on features can proceed by either cascading the features into a feature vector or by cascading the filters for each feature. In this work, we have chosen the former route for simplicity and efficiency reasons. This adaptive non-linear filtering results in feature (both edge and texture) preserving images by giving larger weights to **similar** pixels and smaller weights to pixels that are different, whereby similarities among pixels are not only defined by intensity of pixels but also by features associated with these pixels.

In the following section, we describe the operation of bilateral filtering. The proposed framework of multilateral filtering is presented in Section 3. Section 4 gives experimental results on some standard greyscale images corrupted with additve Gaussian white noise. Finally, main conclusions of this paper and ideas for further exploration are summarized in Section 5.

2. BILATERAL FILTERING

The bilateral filters proposed by Tomasi & Manduchi [1] belong to a class of non-linear filters designed for edge-
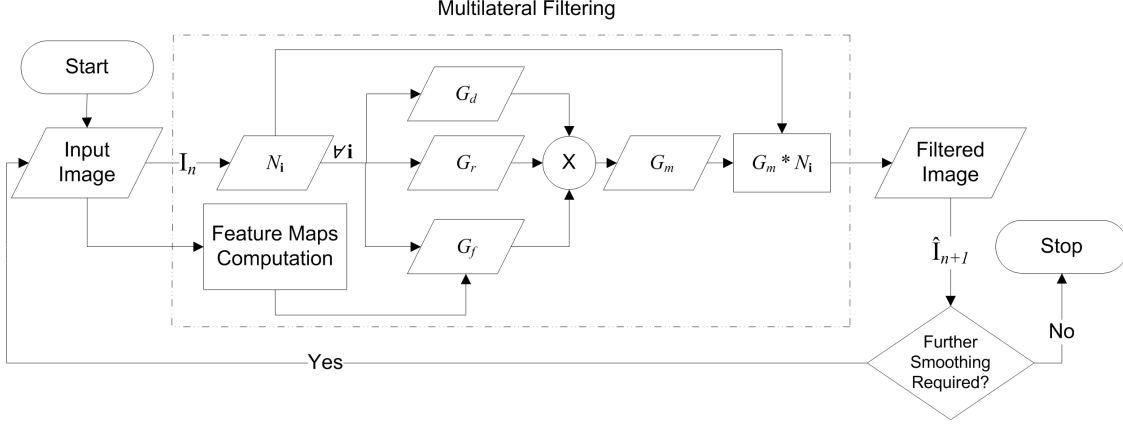
Fig. 1. The Proposed Iterative Multilateral Filtering Framework

preserving image denoising. They operate by convolving a given image with a combination of domain and range filters. The domain filter contains the Gaussian domain weights computed by the geometric closeness function $D(\mathbf{i}, \mathbf{j})$, where $D(\mathbf{i}, \mathbf{j})$ represents the Euclidean distance between the origin pixel at $\mathbf{i}$ and a nearby pixel at $\mathbf{j}$. The range filter, on the other hand, contains Gaussian range weights computed using an intensity difference function $R(\mathbf{i}, \mathbf{j})$, where $R(\mathbf{i}, \mathbf{j})$ gives the absolute difference between intensity values $I(\mathbf{i})$ and $I(\mathbf{j})$ corresponding to locations $\mathbf{i}$ and $\mathbf{j}$. The range filter is such that the larger the difference between two intensity values at $\mathbf{i}$ and $\mathbf{j}$, the smaller the corresponding weight is and vice versa. Let $G_{d,\mathbf{i}}$ and $G_{r,\mathbf{i}}$ respectively denote the domain and range filters, as defined below:

$$G_{d,\mathbf{i}} = \exp\left(-0.5[D(\mathbf{i}, \mathbf{j})/\sigma_d]^2\right), \quad (1)$$

$$G_{r,\mathbf{i}} = \exp\left(-0.5[R(\mathbf{i}, \mathbf{j})/\sigma_r]^2\right), \quad (2)$$

$\forall \mathbf{j} \in N_\mathbf{i}$, where $N_\mathbf{i}$ denotes the set of pixel coordinates in the local neighborhood of the center location $\mathbf{i}$, and σ_d and σ_r denote widths of the two Gaussian kernels. Bilateral filter is simply a product of the domain filter $G_{d,\mathbf{i}}$ and the range filter $G_{r,\mathbf{i}}$. A given image $\mathbf{I}$ is then convolved with the bilateral filter to obtain the denoised image $\hat{\mathbf{I}}$. For a particular pixel at location $\mathbf{i}$, $\hat{I}(\mathbf{i})$ is obatined as follows,

$$\hat{I}(\mathbf{i}) = (G_{d,\mathbf{i}} \cdot G_{r,\mathbf{i}}) \; * \; N_\mathbf{i} \quad (3)$$

$\forall \mathbf{i} \in \mathcal{C}$, where $\mathcal{C}$ denotes a set of all image coordinates and $N_\mathbf{i}$ denotes the neighborhood of pixel at location $\mathbf{i}$ as defined above.

3. MULTILATERAL FILTERING

One of the major limitations of bilateral filtering is that the range filter coefficients rely heavily on actual pixel intensity values which may in turn have been influenced by noise. The range filter computed in this way may consider two noisy pixels to be similar when they happen to have similar intensity values only because of the influence of noise. It assumes that the similarity of two pixels can be determined just by analyzing their intensity values. Furthermore, bilateral filtering does not take into account any regional (e.g., textural) characteristics which can sometimes be inferred from analysis of the local statistical or structural properties.

We present a generic framework for image denoising based on a weighted averaging of image pixels using the idea of **regional similarity** in order to overcome the above limitations. The overall concept of our method is illustrated in Figure 1. First, we compute a set of suitable features for the input image $\mathbf{I}$. These features should characterize the regional similarities of a pixel's neighborhood such that two pixels belonging to the same region have a high value of the similarity function and vice versa. Similarly, if two pixels belong to a similar kind of edge feature in the image, this should be reflected by a relatively high value of the similarity function. Once a suitable set of features has been computed, the algorithm proceeds by computing the domain and range filters, as described in the previous section, as well as the *feature filter*. As discussed earlier in Section 1, the computed features are cascaded in the form of a feature vector $\mathbf{f_i} = \{f_1(\mathbf{i}), f_2(\mathbf{i}), \ldots, f_K(\mathbf{i})\}$, where $f_k(\mathbf{i})$ denotes the kth feature for the pixel at $\mathbf{i}$ and K denotes the total number of features computed per pixel. The adaptive feature filter is then defined as follows,

$$G_{f,\mathbf{i}} = \exp\left(-||\mathbf{f_i} - \mathbf{f_j}||^2/2\sigma_f^2\right), \quad (4)$$

$\forall \mathbf{j} \in N_\mathbf{i}$, where $N_\mathbf{i}$ is as defined above and σ_f is width of the Gaussian kernel associated with the feature filter. The filtered image $\hat{\mathbf{I}}$ is computed by taking a convolution of the input image $\mathbf{I}$ with the product $G_{m,\mathbf{i}}$ of the three Gaussian kernels,

$$G_{m,\mathbf{i}} = (G_{d,\mathbf{i}} \cdot G_{r,\mathbf{i}} \cdot G_{f,\mathbf{i}}). \quad (5)$$

For a particular pixel at location $\mathbf{i}$, its denoised intensity value is computed as below,

$$\hat{I}(\mathbf{i}) = G_{m,\mathbf{i}} \; * \; N_\mathbf{i}. \quad (6)$$

$\forall \mathbf{i} \in \mathcal{C}$. This completes one iteration of multilateral filtering.

3.1. Feature Normalization

It is worth noting that for different features put together in the form of a feature vector, all individual features $f_k(\mathbf{i})$, $\forall k = 1, 2, \ldots, K$ are normalized to have a zero mean and a unit variance as follows,

$$f_k(\mathbf{i}) \longleftarrow \left(\frac{f_k(\mathbf{i}) - \mu_k}{\sigma_k} \right) \qquad (7)$$

where μ_k and σ_k respectively denote mean and standard deviation of values of the kth feature for all pixels. Normalization of all features in this way ensures that equal weightage is given to all features and no single feature is allowed to dominate other features.

3.2. Iterative Multilateral Filtering

Although the original idea of bilateral filtering was non-iterative, Barash [4] showed that an iterative application of bilateral filtering may be required in images with high levels of noise. Using the robust median filter (RMF) estimator [5] for noise standard deviation σ_n in the smoothed image, we can determine if further smoothing is required. If so, another iteration of multilateral filtering is performed on the result $\mathbf{I}_n$ of the previous iteration to obtain $\hat{\mathbf{I}}_{n+1}$. Needless to say, we start with $\mathbf{I}_0 = \mathbf{I}$.

4. EXPERIMENTAL RESULTS

We conducted a large number of experiments with both synthetic and real-world images corrupted with additive Gaussian white noise (AWGN). Two aspects of the proposed framework were subject of our preliminary investigation reported in this paper. First, which type of features are useful in denoising? Second, how does the proposed method fare as compared to the original bilateral filtering and other state-of-the-art methods?

To answer the first question, we conducted experiments with a number of local statistical features such as energy, variance, entropy and several combinations of these features. Here we present a subset of our experimental results with three of the best statistical features: local energy (E), local variance (var), and a combination of the two. Experimental results for four 256-level greyscale images, each of a 512×512 resolution, are given in Table 1. Two of these are images containing textural patterns: *Cosine Grating*, a synthetic image containing two cosine waves, and a *Fingerprint* image. The other two images are taken from the standard real-world test image databases: *Lena* image consisting of relatively smooth regions and *Barbara* containing fine periodic textures at different orientations. Denoising performance is measured in terms of the Peak Signal to Noise Ratio (PSNR). For each greyscale test image, five noisy versions were created by adding white Gaussian noise with standard deviations 10, 20, 30, 40, and 100. The parameters of bilateral filtering were set as follows: the window size is 11×11, $\sigma_d = 1.8$, whereas σ_r is calculated in a way that PSNR computed from

Input Image	σ_n	Bilateral Filtering [1]	Multi-lateral var	Multi-lateral E	Multi-lateral E, var
Cosine Grating	10	30.23	33.50	33.10	**33.70**
	20	23.61	25.53	29.45	**29.83**
	30	21.81	23.60	**26.79**	26.54
	40	20.61	22.00	24.46	**24.55**
	100	17.80	18.10	**19.20**	19.00
Finger-Print	10	32.12	32.68	**33.10**	32.93
	20	28.14	28.14	**29.61**	29.10
	30	24.73	24.73	**26.79**	26.10
	40	22.81	22.81	**24.97**	24.30
	100	19.29	19.29	**19.64**	19.47
Lena	10	33.75	34.00	**34.34**	34.10
	20	30.57	30.57	**31.13**	30.72
	30	28.82	28.82	**29.32**	28.90
	40	27.64	27.64	**28.03**	27.70
	100	22.88	22.88	22.88	22.88
Barbara	10	31.45	31.56	**31.71**	31.64
	20	27.19	27.19	**27.52**	27.33
	30	25.12	25.12	**25.36**	25.18
	40	23.98	23.98	**24.10**	23.98
	100	20.87	20.87	20.87	20.87
Average	-	25.67	26.15	**27.12**	26.94

Table 1. Results of Multilateral Filtering using Local Statistical Features in terms of PSNR (dB)

output image and original image is maximized. For our proposed method, we set the same window size, σ_d, σ_r to give a fair comparison. The value of σ_f is adjusted to maximize the PSNR. It can be seen from Table 1 that while the proposed multilateral filtering with a combination of local energy and variance achieves a PSNR gain of up to 6.22dB over the best bilateral filtering result in case of the *Cosine Grating* image at $\sigma_n = 20$, local energy (E) is the overall winner. This is due to the fact that measures of local variation, such as local entropy or local variance, are more sensitive to noise than local energy which has a low-pass effect. It is worth noting that in the worst case, multilateral filtering behaves at least as well as a bilateral filter when the noise level is high, e.g., when $\sigma_n = 100$ for *Lena* and *Barbara*. In the presence of heavy noise, local statistical features do not provide much extra information about the regional characteristics of a pixel. It is worth noting that multilateral filtering does not only produce better results in terms of objective PSNR but also in terms of subjective visual quality. Visual results for both *Cosine Grating* and *Fingerprint* images are shown in Figs. 2 and 3. As can be seen in these results, the proposed method produces smoother ridges and sharper edges as compared to the standard bilateral filtering.

In Table 2, we present a comparison of denoising results for the proposed method using E with three methods: the original bilateral filtering [1], a Bayesian approach to wavelet shrinkage (**BayesShrink**) [6], and a more recent denoising method using a contourlet transform (**Contourlet-MD**) with sharp localization in frequency and using a multiscale decomposition [7]. Resuls are given for three standard images for which published results using the other three methods are available. It can be seen that the proposed method outper-
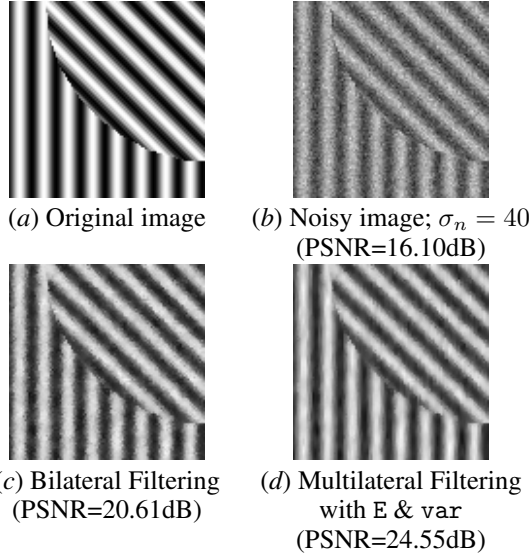
(*a*) Original image

(*b*) Noisy image; $\sigma_n = 40$
(PSNR=16.10dB)

(*c*) Bilateral Filtering
(PSNR=20.61dB)

(*d*) Multilateral Filtering
with `E` & `var`
(PSNR=24.55dB)

Fig. 2. Denoising Results for a Part of *Cosine Grating*

(*a*) Original image

(*b*) Noisy image; $\sigma_n = 40$
(PSNR=16.10dB)

(*c*) Bilateral Filtering
(PSNR=22.81dB)

(*d*) Multilateral Filtering
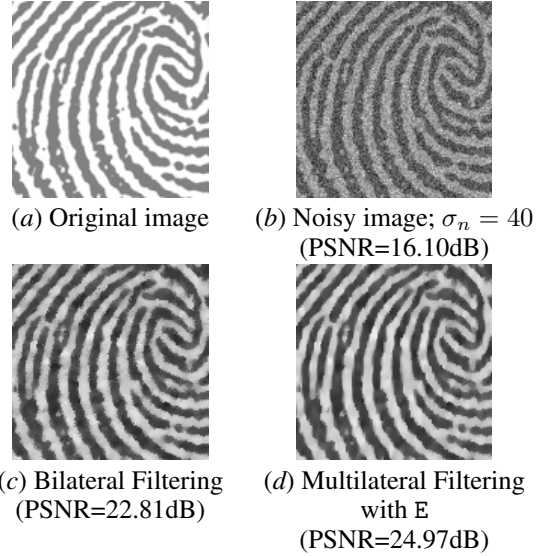with `E`
(PSNR=24.97dB)

Fig. 3. Denoising Results for Central Part of *Fingerprint*

Input Image	σ_n	Bilateral Filtering [1]	Bayes Shrink [6]	Counterlet MD [7]	The Proposed Method
Lena	10	33.75	33.38	32.59	**34.43**
	20	30.57	30.27	**31.20**	31.13
	30	28.82	28.60	28.80	**29.32**
Barbara	10	31.45	31.25	30.16	**31.71**
	20	27.19	27.32	27.02	**27.52**
	30	25.12	25.34	25.30	**25.36**
House	10	33.81	33.07	32.19	**34.45**
	20	30.37	29.83	29.57	**31.11**
	30	28.35	27.12	28.00	**29.08**
Average	-	29.94	29.58	29.43	**30.45**

Table 2. Comparative Image Denoising Results in terms of PSNR (dB)

forms all the other methods in almost all the cases.

5. CONCLUSIONS

In this paper, we have presented a novel framework for adaptive nonlinear image smoothing based on featuer similarity. The proposed filtering method is aimed at smoothing images while preserving edge as well as regional features. We showed that local statistical features such as energy and variance yield plausible denoising results when employed in a multilateral filtering framework. In the worst case, multilateral filtering performs as well as bilateral filtering due to the sensitivity of local statistical features when the noise level is relatively high. Our future work will look into ways to compute robust local features and improved texture descriptors.

Acknowledgements

6. REFERENCES

[1] C. Tomasi and R. Manduchi, "Bilateral filtering for gray and color images," in *Sixth International Conference on Computer Vision (ICCV)*, 1998, pp. 839–846.

[2] R. Garnett, T. Huegerich, C. Chui, and W. He, "A Universal Noise Removal Algorithm With an Impulse Detector," *Image Processing, IEEE Transactions on*, vol. 14, no. 11, pp. 1747–1754, 2005.

[3] P. Choudhury and J. Tumblin, "The trilateral filter for high contrast images and meshes," in *Proceedings of the 14th Eurographics workshop on Rendering*. Eurographics Association Aire-la-Ville, Switzerland, Switzerland, 2003, pp. 186–196.

[4] D. Barash, "Fundamental relationship between bilateral filtering, adaptivesmoothing, and the nonlinear diffusion equation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 6, pp. 844–847, 2002.

[5] D.L. Donoho and J.M. Johnstone, "Ideal spatial adaptation by wavelet shrinkage," *Biometrika*, vol. 81, no. 3, pp. 425–455, 1994.

[6] SG Chang, B. Yu, and M. Vetterli, "Adaptive wavelet thresholding for image denoising and compression," *IEEE Transactions on Image Processing*, vol. 9, no. 9, pp. 1532–1546, 2000.

[7] Y. Lu and M.N. Do, "A New Contourlet Transform with Sharp Frequency Localization," in *IEEE International Conference on Image Processing*, 2006, pp. 1629–1632.