Iterated Nonlocal Means for Texture Restoration

Thomas Brox and Daniel Cremers

CVPR Group, University of Bonn
Römerstr. 164, 53117 Bonn, Germany
{brox,dcremers}@cs.uni-bonn.de

Abstract. The recent nonlocal means filter is a very successful technique for denoising textured images. In this paper, we formulate a variational technique that leads to an adaptive version of this filter. In particular, in an iterative manner, the filtering result is employed to redefine the similarity of patches in the next iteration. We further introduce the idea to replace the neighborhood weighting by a sorting criterion. This addresses the parameter selection problem of the original nonlocal means filter and leads to favorable denoising results of textured images, particularly in case of large noise levels.

1 From Neighborhood Filters to the Nonlocal Means Filter

In recent years, increasingly sophisticated filtering techniques have been developed in order to remove noise from a given input image $f : (\Omega \subset \mathbb{R}^2) \to \mathbb{R}$. While linear Gaussian filtering

$$u(x) = G_\rho * f(x) = \int G_\rho(x') f(x - x')\, dx' \qquad (1)$$

with a Gaussian G_ρ of width $\rho > 0$ is known to blur relevant image structures, more sophisticated nonlinear filtering techniques were developed, such as the total variation filtering [10], also known as the ROF model, which minimizes the cost functional:

$$E(u) = \int (f - u)^2\, dx + \lambda \int |\nabla u|\, dx. \qquad (2)$$

The ROF model is closely related to nonlinear diffusion filters [9], in particular to the total variation flow [1]

$$\begin{aligned} u(x, 0) &= f(x) \\ \partial_t u(x, t) &= \operatorname{div}\left(\frac{\nabla u(x,t)}{|\nabla u(x,t)|}\right). \end{aligned} \qquad (3)$$

In the space-discrete, one-dimensional setting, it was shown that the solution of this diffusion equation at a time t is equivalent to the solution of the ROF model with $\lambda = t$ as well as a certain implementation of wavelet soft shrinkage [12].

F. Sgallari, A. Murli, and N. Paragios (Eds.): SSVM 2007, LNCS 4485, pp. 13–24, 2007.

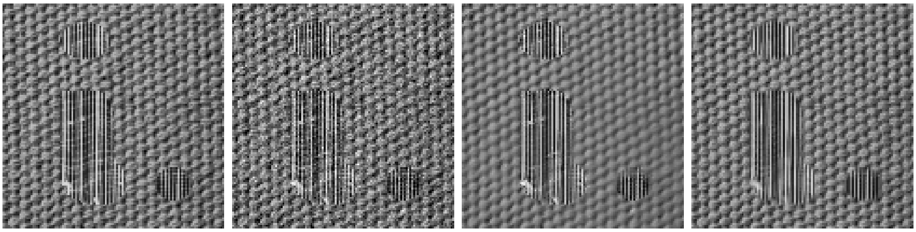

Fig. 1. Nonlocal means filter. **From left to right:** (a) Reference image of size 119×121 pixels. (b) Gaussian noise with $\sigma = 20$ added. (c) Denoising result of the nonlocal means filter ($h = 0.95\sigma$). (d) Denoising result of the changed nonlocal means filter using a sorting criterion ($n = 8$).

Despite an enormous success in image enhancement and noise removal applications, approaches like the ROF filtering remain spatially local, in the sense that at each location $x \in \Omega$ the update of u is determined only by derivatives of u at that same location x – see equation (3).

A class of image filters which adaptively takes into account intensity information from more distant locations are the Yaroslavsky neighborhood filters [16]:

$$u(x) = \frac{1}{C(x)} \int K(x, y, f(x), f(y)) \quad C(x) = \int K(x, y, f(x), f(y))\, dy. \tag{4}$$

K is a nonnegative kernel function which decays with the distances $|x - y|$ and $|f(x) - f(y)|$. Thus the application of this filter amounts to assigning to x a weighted average over the intensities $f(y)$ of all pixels y which are similar in the sense that they are close to $\{x, f(x)\}$ in space and in intensity. These filters are also known as local M-smoothers [4,15]. A similar, but iterative, filter is the bilateral filter [11,13]. Relations between such neighborhood filters and nonlinear diffusion filters have been investigated in [8].

A drastic improvement of these neighborhood filters is the nonlocal means filter which was recently proposed by Buades et al. [3]. Its application to video processing and surface smoothing has been demonstrated in [7,17] and a very related statistical filter was presented in [2]. The nonlocal means filter can be written as:

$$u(x) = \int w_f(x, y)\, f(y)\, dy, \tag{5}$$

with normalized weights of the form

$$w_f(x, y) = \frac{g_f(x, y)}{\int g_f(x, y)\, dy}, \tag{6}$$

where

$$g_f(x, y) = \exp\left(-\frac{d_f^2(x, y)}{h^2} \right) \tag{7}$$

and

$$d_f^2(x, y) = \int G_\rho(x') \big(f(x - x') - f(y - x') \big)^2 dx'. \tag{8}$$

In contrast to the neighborhood filters, the nonlocal means filter quantifies the similarity of pixels x and y by taking into account the similarity of whole patches around x and y. The similarity is expressed by a dissimilarity measure $d_f^2(x, y)$, which contains the size ρ of the compared patches as a parameter, and a weighting function g with parameter h, which quantifies how fast the weights decay with increasing dissimilarity of respective patches. Since the above similarity measure takes into account complete patches instead of single pixel intensities, the nonlocal means filter is able to remove noise from textured images without destroying the fine structures of the texture itself. This amazing property is demonstrated in Fig. 1.

The key idea of nonlocal means filtering is that the restoration of a destroyed texture patch is improved with support from similar texture patches in other areas of the image. The filter is, hence, based on a similar concept as the texture synthesis work of Efros and Leung [5]. In this paper, we will show that this property is further enhanced by iterated nonlocal means which shall be developed in the following.

2 Iterated Nonlocal Means

The nonlocal means filter assigns to each pixel x a weighted average over all intensities $f(y)$ of pixels y which share a similar intensity neighborhood as the point x. A trivial variational principle for this filter can be written as:

$$E(u) = \int \left(u(x) - \int w_f(x, y) f(y)\, dy \right)^2 dx. \tag{9}$$

An alternative variational formulation of nonlocal means was proposed in [6]. In this paper, we propose an iterated form of the nonlocal means filter which arises when extending the above functional by replacing w_f by w_u in the following manner:

$$E(u) = \int \left(u(x) - \int w_u(x, y) f(y)\, dy \right)^2 dx. \tag{10}$$

Thus, rather than imposing similarity of $u(x)$ to $f(y)$ for locations y where the input image $f(y)$ is similar to $f(x)$, we impose similarity to $f(y)$ for locations y where the filtered image $u(y)$ is similar to $u(x)$. This induces an additional feedback and further decouples the resulting image u from the input image f. The idea is that the similarity of patches can be judged more accurately from the already denoised signal than from the noisy input image.

2.1 Fixed Point Iteration

Due to the introduced dependence of w on u, the minimizer of (10) is no longer the result of a weighted convolution, but the solution of a nonlinear optimization problem. A straightforward way to approximate a solution of (10) is by an iterative scheme with iteration index k, where we start with the initialization $u^0 = f$.

For fixed $u = u^k$ we are in a similar situation as with the conventional nonlocal means filter. In particular, we can compute the similarity measure $w_{u^k}(x,y)$ for the current image u^k and, as a consequence, we obtain an update on u

$$u^{k+1}(x) = \int w_{u^k}(x,y)\, f(y)\, dy. \tag{11}$$

Whether this iterative process converges to a stationary solution, is subject of future investigation.

2.2 Euler-Lagrange Equation and Gradient Descent

An alternative way to find a solution of (10) is by computing its Euler-Lagrange equation, which states a necessary condition for a (local) minimum. We are seeking for the gradient

$$\frac{\partial E(u)}{\partial u} = \left.\frac{\partial E(u+\epsilon h)}{\partial \epsilon}\right|_{\epsilon \to 0}. \tag{12}$$

After evaluation and substitution of integration variables, we end up with the following Euler-Lagrange equation:

$$\begin{aligned}\frac{\partial E(u)}{\partial u} = 0 = &\left(u(x) - \int f(y) w_u(x,y) dy\right)\\
+2\iint &\left(u(z) - \left(\int f(y') w_u(z,y') dy'\right) \frac{f(y) g'(z,y)}{\int g(z,y') dy'} G_\rho(y-x)(u(z-y-x) - u(x))\right) dydz\\
+2\iiint &\left(u(z) - \left(\int f(y') w_u(z,y') dy'\right) \frac{f(y) g(z,y) g'(z,y'')}{\left(\int g(z,y') dy'\right)^2} G_\rho(z-x)(u(x) - u(y''-z-x))\right) dy'' dydz\end{aligned}$$

After initializing $u^0 = f$, the gradient descent equation

$$u^{k+1} = u^k - \tau \frac{\partial E(u)}{\partial u} \tag{13}$$

yields the next local minimum for some sufficiently small step size τ and $t \to \infty$. Obviously, gradient descent with the gradient being reduced to the first term in (2.2) leads for $\tau = 1$ to the same iterative scheme as in Section 2.1. The additional two terms take the variation of w_u into account and ensure convergence for sufficiently small step sizes τ. However, these terms induce a very large computational load in each iteration. In particular, the time complexity of the third term in each iteration is $O(MN^4)$, where N is the number of pixels in the image and M the number of pixels in the compared patch. For comparison, the first term only has a time complexity of $O(MN^2)$ and a nonlinear diffusion filter like TV flow has a time complexity of $O(N)$ in each iteration. Hence, in our experiments, we took only the first term into account.

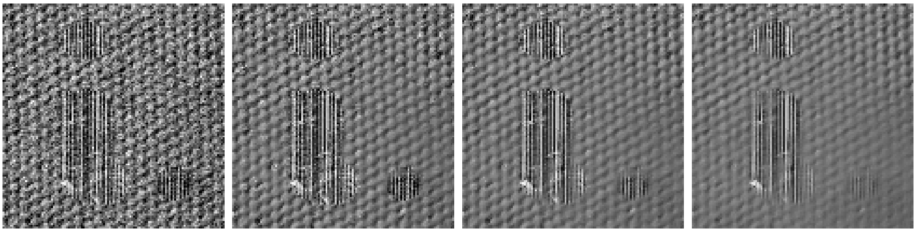

Fig. 2. Parameter sensitivity of nonlocal means filter in case of high noise levels. **From left to right:** (a) Gaussian noise with $\sigma = 30$ added to reference image in Fig.1a. (b,c,d) Denoising result of nonlocal means filter for $h = 0.65\sigma$, $h = 0.7\sigma$, and $h = 0.75\sigma$. The fact that the restored image contains both unfiltered, noisy areas, as well as over-smoothed areas shows that there exists no ideal choice of the parameter h.

3 A Robust Threshold Criterion

While the nonlocal means filter can yield astonishing denoising results, a deeper experimental investigation reveals a large sensitivity to the parameter h in

$$g_f(x, y) = \exp\left(-\frac{d_f^2(x, y)}{h^2}\right), \tag{14}$$

which is responsible for steering the decay of weights for decreasing similarity of patches. This parameter sensitivity increases with the noise variance σ in the image. Moreover, if the noise level exceeds a certain value, it is no longer possible to choose a global h such that the noise is removed everywhere without destroying repetitive structure somewhere else in the image. This is demonstrated in Fig. 2.

The reason for this effect is the weighting function g. By definition we have $g_f(x, x) = 1$. Suppose there is a highly repetitive patch and the noise level is rather small. In this case, there will be many similar patches with $g(x, y) \approx 1$ and the smoothing between these patches works well. Now suppose a patch that is hardly similar to other patches in the image, or only very few of them. Consequently, there will be almost no change at x since $g(x, x) = 1$ and $g(x, y) \approx 0$ almost everywhere. In this case, one has to increase h such that there are enough y with $g(x, y) > \epsilon$ in order to see a smoothing effect.

Buades et al. have been aware of this problem and suggested to set $g(x, x)$ to $\max_{y \neq x} g(x, y)$. Although this attenuates the problem, it does not resolve it, as it only ensures the averaging of at least two values. The results shown in Fig. 2, where we implemented this idea, reveal that the averaging of at least two values is in many cases not sufficient.

Here, we suggest to approach the problem from a different direction. Instead of defining a function g that assigns weights to positions y, we choose the number n of positions that is appropriate to remove a certain noise level. We then simply take those n patches with the smallest dissimilarity $d^2(x, y)$.

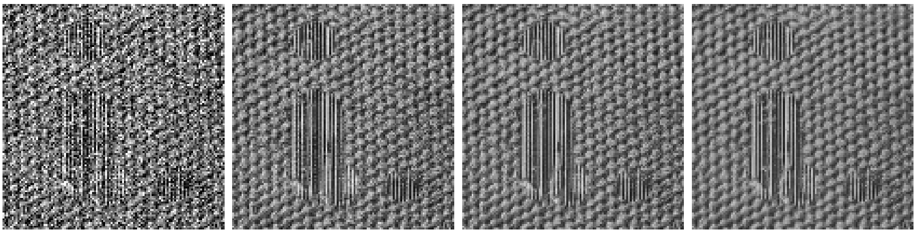

Fig. 3. From left to right: (a) Gaussian noise with $\sigma = 50$ added to the reference image in Fig.1a. (b,c,d) Denoising result of nonlocal means filter with the new sorting criterion for $n = 10$, $n = 20$, and $n = 40$. Although the noise level is significantly higher than in Fig. 2, the exact parameter choice is less critical and the results look favorable.

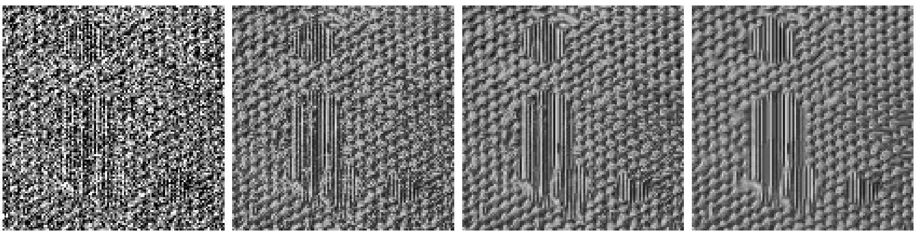

Fig. 4. Iterated nonlocal means. **From left to right:** (a) Gaussian noise with $\sigma = 70$ added to reference image in Fig.1a. (b,c,d) Denoising result of the iterated nonlocal means filter with the sorting criterion after 1, 2, and 5 iterations ($n = 20$). Iterations improve the regularity of the texture.

By considering for any pixel x the n most similar pixels rather than all those pixels of similarity above a fixed threshold, we allow for denoising which does not depend on how repetitive the respective structure at x is in the given image.

4 Experimental Results

Figure 3 shows the effect of the new sorting criterion. Although the noise level in the input image has been chosen much higher than in Figure 2, the result looks more appealing. The noise has been removed while the repetitive texture patterns have been preserved. Even the contrast did not suffer severely. Moreover, the sensitivity with respect to the parameter choice has been attenuated considerably: All three results depicted in Figure 3 are satisfactory, although the parameter n has been varied by a factor 4. For all experiments, also in the previous sections, we fixed $\rho = \sqrt{8}$ and used a 9×9 window for implementation.

Figure 4 demonstrates the impact of iterating the nonlocal means filter. Again we increase the noise level. Note that due to the high amount of noise and clipping intensities that exceed the range of $[0, 255]$, the noise is not fully Gaussian anymore. Nevertheless, the results that can be obtained with the modified

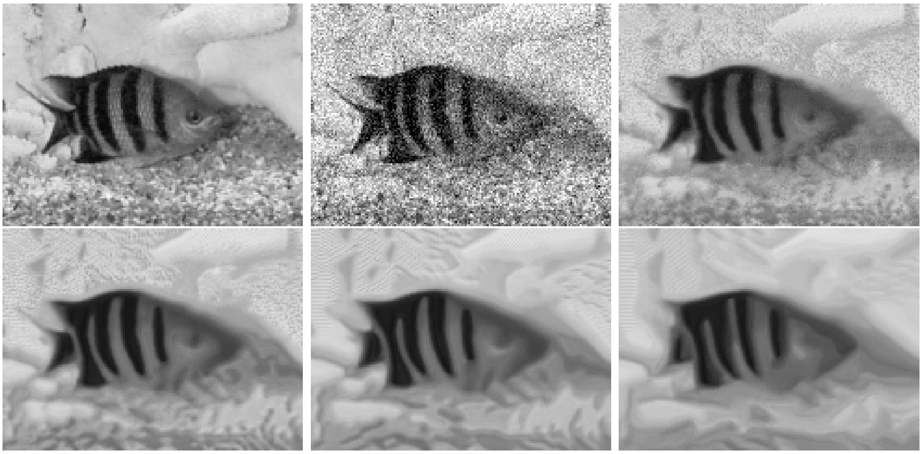

Fig. 5. Denoising of a natural, non-regular image. **From left to right, top to bottom:** (a) Reference image of size 162×120 pixels. (b) Gaussian noise with $\sigma = 30$ added. (c,d,e,f) Denoising result of the iterated nonlocal means filter with the sorting criterion after 1, 2, 4, and 10 iterations ($n = 100$).

nonlocal means filter, in particular with its iterated version, look quite satisfactory. Clearly, iterating the filter improves the regularity of the texture pattern.

This can lead to interesting effects, if the filter is applied to an image that is mainly non-regular. Such a case is shown in Fig. 5. For an increasing number of iterations, the filter acts more and more coherence enhancing reminiscent of curvature motion or coherence enhancing anisotropic diffusion [14].

While the denoising in Fig. 5 has been achieved with a quite large number of partners, namely $n = 100$, Fig. 6 shows what happens if one decreases n and instead increases the number of iterations. Due to the small number of neighbors, the filter is not able to fully remove the noise. Interestingly, with further iterations, the filter detects structures in the noise that have not been present in the image and starts enhancing these.

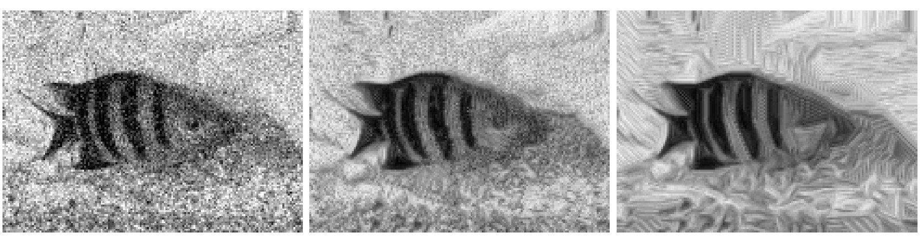

Fig. 6. Hallucination of regular patterns in noise for many iterations and small mask size. **From left to right:** (a) Gaussian noise with $\sigma = 30$ added to reference image in Fig.5a. (b,c) Denoising result of the iterated nonlocal means filter with the sorting criterion after 1 and 300 iterations ($n = 10$). For small n, the iterated filter creates structures from the noise.

In Fig. 7 we show a comparison of the modified, iterated nonlocal means filter to the ROF model. In order to demonstrate the robustness of the parameter settings, the image contains various different textures of different scales. The results, in particular the two closeups in Fig. 8, reveal a very precise reconstruction of all textures despite the fixed parameter setting. Even very fine texture details are preserved. This is in contrast to the ROF model, which preferably removes small scale structures. In most cases such structures are noise pixels, but they may also be important parts of the texture.

Finally, we performed a quantitative evaluation of various filters, including the ROF model, conventional nonlocal means, nonlocal means with the suggested

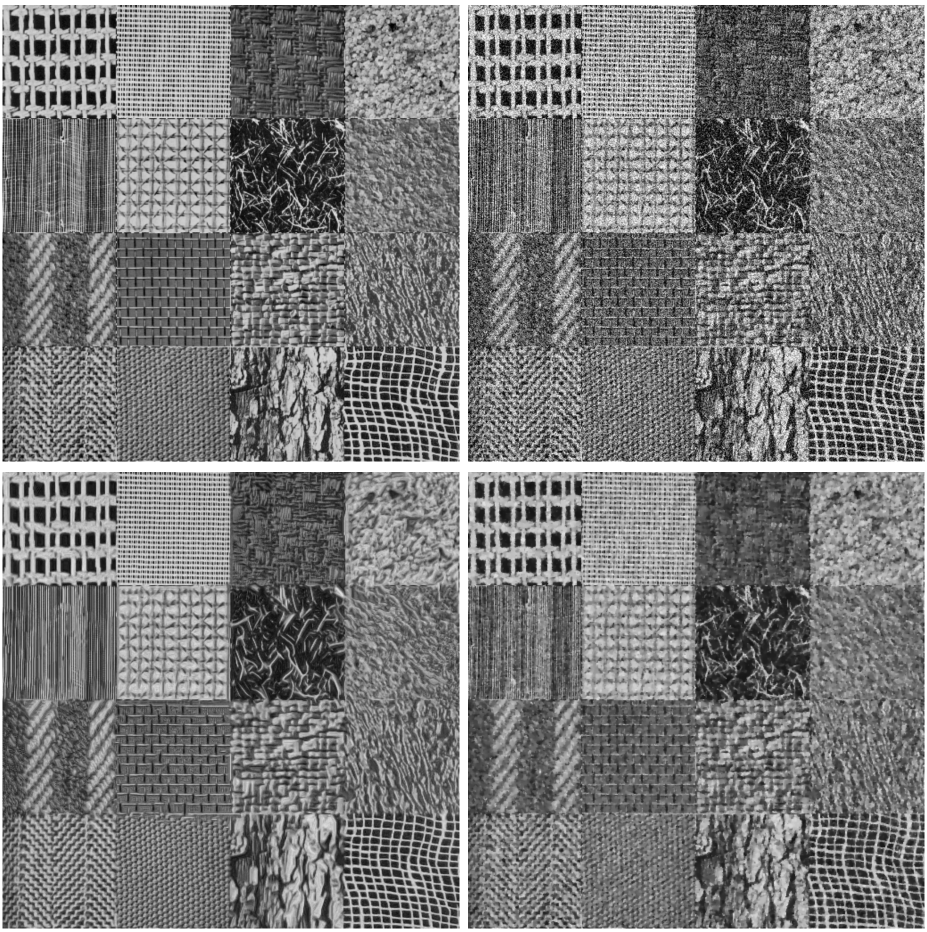

Fig. 7. Comparison to ROF denoising model. **Top left:** (a) Reference image of size 512×512 pixels. **Top right:** (b) Gaussian noise with $\sigma = 40$ added. **Bottom left:** (c) Denoising result with iterated nonlocal means and the sorting criterion ($n = 20$) after 2 iterations. **Bottom right:** (d) ROF model for $\alpha = 20$.

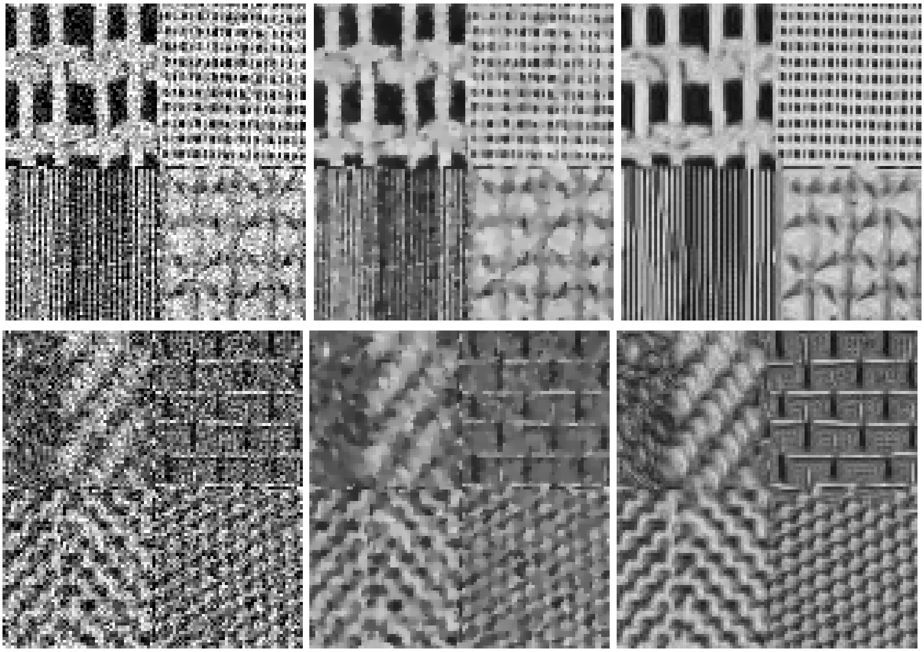

Fig. 8. Zoom into two regions in Fig. 7. **Each row from left to right:** (a) Noisy input image. (b) ROF model. (c) Iterated nonlocal means.

Fig. 9. Undisturbed reference images for the quantitative comparison in Table 1

Table 1. Root mean square error for the input images in Fig. 9 and different denoising techniques. See Fig. 10 for the resulting images. The value in brackets is the optimized setting of the free parameter(s).

	texture ($\sigma = 70$)		owl ($\sigma = 40$)		Lena ($\sigma = 30$)	
Input image	62.28		38.06		29.16	
ROF model	34.28	(28)	22.78	(16)	11.33	(18)
Nonlocal means	38.73	(0.45σ)	27.29	(0.6σ)	13.69	(0.6625σ)
Sorting criterion	32.02	(175)	23.95	(40)	13.38	(100)
Iterated sorting criterion	28.16	(75/3)	23.95	(40/1)	12.57	(65/2)

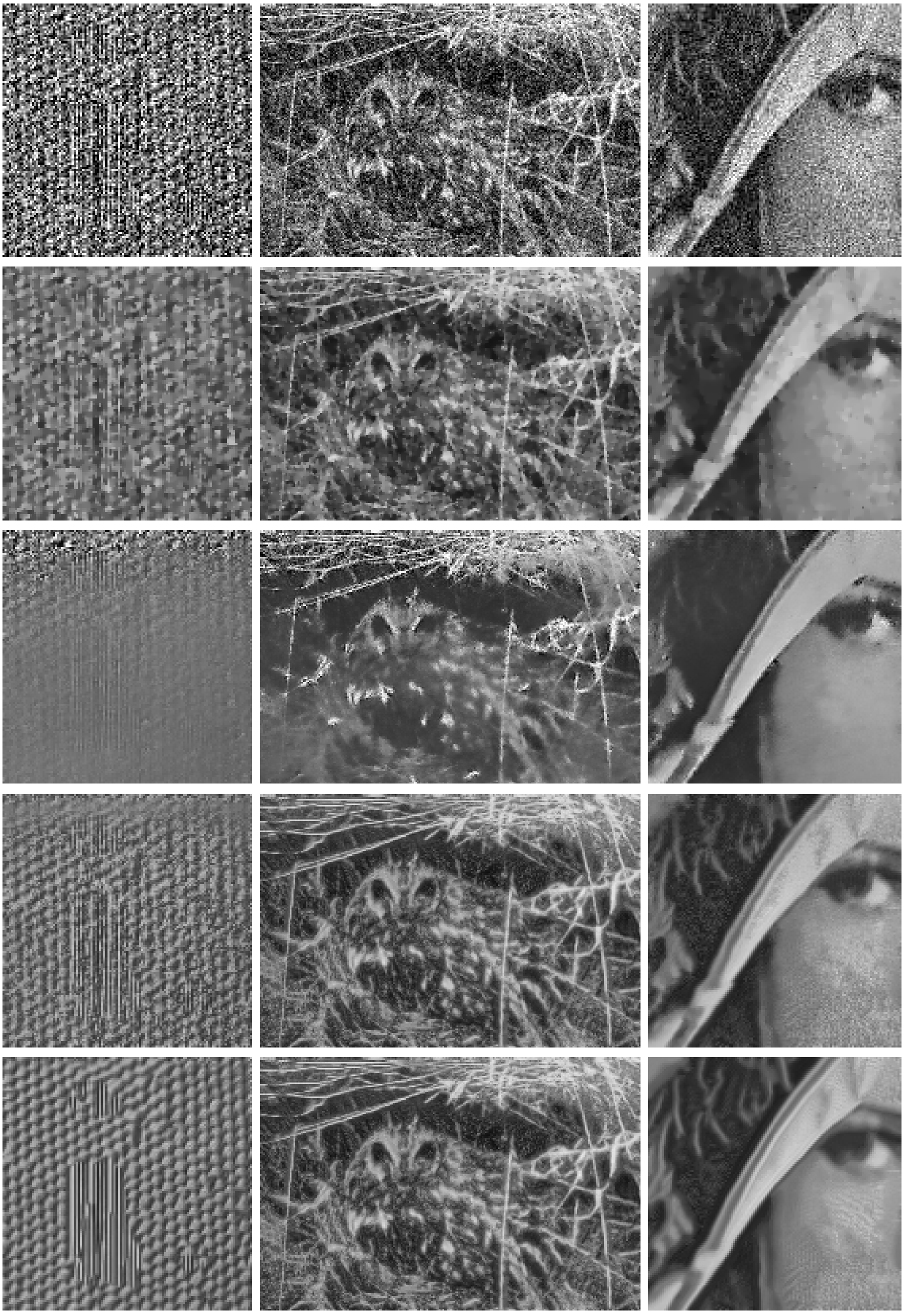

Fig. 10. Images corresponding to the result in table 1. **Each column from top to bottom:** Noisy input image, TV flow, nonlocal means filter, nonlocal means with sorting criterion, iterated nonlocal means. In case of the owl image, the iterated and non-iterated results are identical.

sorting criterion, as well as its iterated version. Table 1 lists the root mean square (RMS) error

$$e_{RMS} := \sqrt{\frac{1}{N}\sum_{i=1}^{N}(r(i) - u(i))^2} \tag{15}$$

between the outcome u of the filter and the undisturbed reference image r. The test images and the filtering results are shown in Fig. 9 and Fig. 10, respectively. For a fair comparison, we ensured that exactly the same noise was added in all test runs and optimized the free parameters.

The numbers only partially support the visual impression, as the results of the ROF model reveal more unpleasant artifacts than its good RMS errors indicate. However, the numbers are in line with the impression that the nonlocal means filter using the sorting criterion, in particular its iterated version, performs best for regular texture patterns. In case of images that are dominated by piecewise homogeneous areas, the conventional nonlocal means filter still yields the most appealing results.

5 Conclusions

We proposed a variational formulation of the recently developed nonlocal means filter and introduced an additional feedback mechanism at the variational level. We showed that the solution by a fixed point iteration gives rise to an iterated version of the nonlocal means filter. Moreover, we proposed to replace the neighborhood weighting in the original formulation by a sorting criterion which assures that the amount of filtering no longer depends on how repetitive respective image structures are in the given image. Experimental results demonstrate that the iterated nonlocal means filter outperforms both nonlocal means and total variation filtering when applied to the restoration of regular textures. At the same time, our experiments indicate that the increased feedback may lead to a hallucination of regular patterns in noise for large iteration numbers.

References

1. F. Andreu, C. Ballester, V. Caselles, and J. M. Mazón. Minimizing total variation flow. *Differential and Integral Equations*, 14(3):321–360, Mar. 2001.
2. S. Awate and R. Whitaker. Unsupervised, information-theoretic, adaptive image filtering for image restoration. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28(3):364–376, Mar. 2006.
3. A. Buades, B. Coll, and J. M. Morel. A review of image denoising algorithms, with a new one. *SIAM Interdisciplanary Journal*, 4(2):490–530, 2005.
4. C. K. Chu, I. Glad, F. Godtliebsen, and J. S. Marron. Edge-preserving smoothers for image processing. *Journal of the American Statistical Association*, 93(442): 526–556, 1998.
5. A. Efros and T. Leung. Texture synthesis by non-parametric sampling. In *Proc. International Conference on Computer Vision*, pages 1033–1038, Corfu, Greece, Sept. 1999.

6. S. Kindermann, S. Osher, and P. W. Jones. Deblurring and denoising of images by nonlocal functionals. *SIAM Interdisciplinary Journal*, 4(4):1091–1115, 2005.
7. M. Mahmoudi and G. Sapiro. Fast image and video denoising via nonlocal means of similar neighborhoods. *Signal Processing Letters*, 12(12):839–842, 2005.
8. P. Mrázek, J. Weickert, and A. Bruhn. On robust estimation and smoothing with spatial and tonal kernels. In R. Klette, R. Kozera, L. Noakes, and J. Weickert, editors, *Geometric Properties from Incomplete Data*, pages 335–352. Springer, Dordrecht, 2006.
9. P. Perona and J. Malik. Scale space and edge detection using anisotropic diffusion. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12: 629–639, 1990.
10. L. I. Rudin, S. Osher, and E. Fatemi. Nonlinear total variation based noise removal algorithms. *Physica D*, 60:259–268, 1992.
11. S. M. Smith and J. M. Brady. SUSAN: A new approach to low-level image processing. *International Journal of Computer Vision*, 23(1):45–78, May 1997.
12. G. Steidl, J. Weickert, T. Brox, P. Mrázek, and M. Welk. On the equivalence of soft wavelet shrinkage, total variation diffusion, total variation regularization, and SIDEs. *SIAM Journal on Numerical Analysis*, 42(2):686–713, May 2004.
13. C. Tomasi and R. Manduchi. Bilateral filtering for gray and color images. In *Proc. Sixth International Conference on Computer Vision*, pages 839–846, Bombay, India, Jan. 1998. Narosa Publishing House.
14. J. Weickert. Coherence-enhancing diffusion of colour images. *Image and Vision Computing*, 17(3/4):199–210, Mar. 1999.
15. G. Winkler, V. Aurich, K. Hahn, and A. Martin. Noise reduction in images: some recent edge-preserving methods. *Pattern Recognition and Image Analysis*, 9(4):749–766, 1999.
16. L. P. Yaroslavsky. *Digital Picture Processing - An Introduction.* Springer, 1985.
17. S. Yoshizawa, A. Belyaev, and H.-P. Seidel. Smoothing by example: mesh denoising by averaging with similarity-based weights. In *Proc. International Conference on Shape Modeling and Applications*, pages 38–44, June 2006.