From Adaptive Averaging to Accelerated Nonlinear Diffusion Filtering

Stephan Didas and Joachim Weickert

Mathematical Image Analysis Group
Faculty of Mathematics and Computer Science, Building E11
Saarland University, 66041 Saarbrücken, Germany
{didas, weickert}@mia.uni-saarland.de
http://www.mia.uni-saarland.de

Abstract. Weighted averaging filters and nonlinear partial differential equations (PDEs) are two popular concepts for discontinuity-preserving denoising. In this paper we investigate novel relations between these filter classes: We deduce new PDEs as the scaling limit of the spatial step size of discrete weighted averaging methods. In the one-dimensional setting, a simple weighted averaging of both neighbouring pixels leads to a modified Perona-Malik-type PDE with an additional acceleration factor that provides sharper edges. A similar approach in the two-dimensional setting yields PDEs that lack rotation invariance. This explains a typical shortcoming of many averaging filters in 2-D. We propose a modification leading to a novel, anisotropic PDE that is invariant under rotations. By means of the example of the bilateral filter, we show that involving a larger number of neighbouring pixels improves rotational invariance in a natural way and leads to the same PDE formulation. Numerical examples are presented that illustrate the usefulness of these processes.

1 Introduction

Adaptive averaging filters belong to the simplest and most effective tools for image processing. Since taking the average of the grey values of all pixels in a certain spatial neighbourhood is an intuitive concept, already early methods in image processing use averaging filters: For example, in the beginning of the 1980s, Lee presented an averaging filter for image denoising [1]. In the literature, there is a whole variety of methods which use the concept of averaging pixel grey values with weights depending on their tonal[1] and spatial distance. Some examples are *adaptive smoothing* [2] by Saint-Marc et al., *adaptive weights smoothing* [3] by Polzehl and Spokoiny, or the *W-estimator* [4] by Winkler et al. While many averaging filters work iteratively by applying small stencils, the *bilateral filter* of Tomasi and Manduchi [5,6] in its original form is an example for a noniterative averaging method: They proposed to use just one iteration of an averaging scheme with a large stencil. Applications for other tasks than image denoising are investigated by Smith and Brady with the *SUSAN filter* [7].

[1] The tonal difference denotes the difference of grey values.

K. Franke et al. (Eds.): DAGM 2006, LNCS 4174, pp. 101–110, 2006.

Many local adaptive filters have been introduced in an intuitive manner. Research on finding a systematic theoretical foundation for them started much more recently: Barash [8], for instance, investigated connections between bilateral filtering and nonlinear diffusion with a scalar-valued diffusivity, and Mrázek et al. [9] have introduced a common framework for a number of adaptive filters that is based on minimising suitable energy functions. An overview over several neighbourhood filtering techniques has been given by Buades et al. [10]. They start with integral formulations of neighbouring filters and relate them to methods based on PDEs [10].

In general, PDE approximations of discrete averaging filters can be useful to study the evolution of the results under iterated filtering, to prove equivalence between seemingly different methods, and to investigate why and how a discrete filter deviates from a rotationally invariant behaviour. Last but not least, these scaling limits can also lead to novel PDEs with interesting properties.

The goal of our paper is to perform novel scaling limits of a specific class of discrete adaptive averaging methods. This class includes local filters as well as more global representatives such as bilateral filtering.

This paper is organised as follows: In Section 2 we start with a fully discrete averaging filter and describe how a scaling limit of it can be related to an accelerated variant of the Perona-Malik filter. These ideas are extended to the two-dimensional case in Section 3. They motivate the use of an anisotropic filter similar to the diffusion filter in [11]. In Section 4 we extract the same filter as scaling limit of bilateral filtering. Numerical examples in Section 5 juxtapose the behaviour of the scaling limits to the averaging filters they originate from. Section 6 concludes the paper with a summary.

2 Averaging Filters and Scaling Limits in 1-D

Derivation of the Scaling Limit. We start with the consideration of an iterative weighted averaging filter of the form

$$\begin{aligned} u_i^0 &= f_i \\ u_i^{k+1} &= \frac{g\left(\left|\frac{u_{i+1}^k - u_i^k}{h}\right|\right) u_{i+1}^k + g\left(\left|\frac{u_{i-1}^k - u_i^k}{h}\right|\right) u_{i-1}^k}{g\left(\left|\frac{u_{i+1}^k - u_i^k}{h}\right|\right) + g\left(\left|\frac{u_{i-1}^k - u_i^k}{h}\right|\right)} \end{aligned} \tag{1}$$

where $f \in \mathbb{R}^n$ is an initial signal and u^k denotes the processed signal at iteration $k \in \mathbb{N}_0$. For each pixel u_i^{k+1}, the filter takes the direct neighbours u_{i-1}^k and u_{i+1}^k into account for averaging. At the boundaries, we assume mirroring boundary conditions, that means we have two dummy pixels $u_0^k := u_1^k$ and $u_{n+1}^k := u_n^k$. Typically one chooses a decreasing positive function g such that the denominator cannot be zero. This also implies that we always have convex combinations which guarantees a maximum-minimum principle for the filter. One may use e.g. the same function g as the diffusivities in nonlinear diffusion filtering [12], for instance $g(s) = \left(1 + s^2/\lambda^2\right)^{-1}$. We observe that the weights depend on the

tonal distance between the pixel and its direct neighbours divided by the spatial step size $h > 0$ between the two pixels. We introduce the abbreviations $g^k_{i+\frac{1}{2}} := g\left(\left|\frac{u^k_{i+1}-u^k_i}{h}\right|\right)$ and rewrite (1) as

$$u_i^{k+1} = \frac{g^k_{i+\frac{1}{2}} u^k_{i+1} + g^k_{i-\frac{1}{2}} u^k_{i-1}}{g^k_{i+\frac{1}{2}} + g^k_{i-\frac{1}{2}}} \tag{2}$$

$$= u_i^k + \frac{g^k_{i+\frac{1}{2}}(u^k_{i+1} - u^k_i) - g_{i-\frac{1}{2}}(u^k_i - u^k_{i-1})}{g^k_{i+\frac{1}{2}} + g^k_{i-\frac{1}{2}}} \tag{3}$$

$$= u_i^k + \frac{\frac{1}{h}\left(g^k_{i+\frac{1}{2}} \frac{u^k_{i+1}-u^k_i}{h} - g^k_{i-\frac{1}{2}} \frac{u^k_i - u^k_{i-1}}{h}\right)}{\frac{1}{h^2}\left(g^k_{i+\frac{1}{2}} + g^k_{i-\frac{1}{2}}\right)} \tag{4}$$

In this last form we notice that the iterative scheme contains finite differences which approximate spatial derivatives of u. Now we assume that u and g are sufficiently smooth to perform a Taylor expansion. For example, there appears $\frac{u^k_{i+1}-u^k_i}{h} = \partial_x u^k_{i+\frac{1}{2}} + \mathcal{O}(h^2)$ in the scheme. Together with the abbreviations introduced above this yields

$$g^k_{i+\frac{1}{2}} + g^k_{i-\frac{1}{2}} \;=\; g\left(\left|\frac{u^k_{i+1} - u^k_i}{h}\right|\right) + g\left(\left|\frac{u^k_i - u^k_{i-1}}{h}\right|\right) \;=\; 2g\left(|\partial_x u^k_i|\right) + \mathcal{O}(h^2) \tag{5}$$

and thus we can write

$$u_i^{k+1} = u_i^k + \frac{\partial_x\left(g(|\partial_x u^k_i|)\partial_x u\right) + \mathcal{O}(h^2)}{\frac{1}{h^2}\left(2g(|\partial_x u^k_i|) + \mathcal{O}(h^2)\right)} \quad . \tag{6}$$

To understand the iteration indices $k+1$ and k as discrete samples of a continuous time variable t we introduce a temporal step size $\tau > 0$. Division of both sides by τ leads to the equation

$$\frac{u_i^{k+1} - u_i^k}{\tau} = \frac{\partial_x\left(g(|\partial_x u^k_i|)\partial_x u\right) + \mathcal{O}(h^2)}{\frac{\tau}{h^2}\left(2g(|\partial_x u^k_i|) + \mathcal{O}(h^2)\right)} \tag{7}$$

where the left-hand side is an approximation for the temporal derivative $\partial_t u$ at time level k with an error in the order $\mathcal{O}(\tau)$. We set the ratio between h and τ such that $\frac{\tau}{h^2} = \frac{1}{2}$ and let h tend to zero. Then (7) approximates

$$\partial_t u = \frac{1}{g(|\partial_x u|)}\,\partial_x\left(g(|\partial_x u|)\,\partial_x u\right) \tag{8}$$

with an error in the order of $\mathcal{O}(\tau + h^2)$. This equation is similar to the nonlinear diffusion equation presented by Perona and Malik [12]:

$$\partial_t u = \partial_x\left(g(|\partial_x u|)\,\partial_x u\right) \quad . \tag{9}$$

The only difference is the factor $\frac{1}{g(|\partial_x u|)}$ on the right-hand side which acts as an acceleration of the Perona-Malik filtering process at edges. To understand this, assume that $|\partial_x u|$ is relatively small within a region. A classical Perona-Malik diffusivity is close to 1 in this case, and the factor has only a small effect. More interesting is the situation near an edge where $\partial_x u$ has large absolute value, and backward diffusion can occur for the diffusivities presented by Perona and Malik. In this case, $g(|\partial_x u|)$ is close to zero, and thus $\frac{1}{g(|\partial_x u|)}$ leads to an amplification of the backward diffusion behaviour. We can expect such equations to yield sharper results than classical Perona-Malik PDEs. One the other hand, they do not necessarily preserve the average grey value, since they cannot be written in divergence form.

Stability of an Explicit Discretisation. Since classical diffusivities g may be arbitrary close to zero, the fraction $\frac{1}{g(|\partial_x u|)}$ in (8) is not bounded. This might give rise to concerns regarding stability. However, the weighted averging scheme (1) inspires also ways how to obtain stable discretisations: An explicit Euler scheme for (8) can be written as

$$u_i^{k+1} = u_i^k + \tau \frac{2}{g_{i+\frac{1}{2}}^k + g_{i-\frac{1}{2}}^k} \frac{1}{h} \left(g_{i+\frac{1}{2}}^k \frac{u_{i+1}^k - u_i^k}{h} - g_{i-\frac{1}{2}}^k \frac{u_i^k - u_{i-1}^k}{h} \right) \tag{10}$$

$$= \frac{2\tau}{h^2} \frac{g_{i+\frac{1}{2}}^k}{g_{i+\frac{1}{2}}^k + g_{i-\frac{1}{2}}^k} u_{i+1}^k + \left(1 - \frac{2\tau}{h^2}\right) u_i^k + \frac{2\tau}{h^2} \frac{g_{i-\frac{1}{2}}^k}{g_{i+\frac{1}{2}}^k + g_{i-\frac{1}{2}}^k} u_{i-1}^k \tag{11}$$

with the same notation as above and with mirroring boundary conditions. We note that the factors in front of u_{i+1}^k, u_i^k and u_{i-1}^k sum up to 1. For $\tau \leq \frac{h^2}{2}$ all three factors are nonnegative, and thus u_i^{k+1} is a convex combination of the three pixels: the scheme is maximum-minimum-stable. Further we see that for the limit $\tau = \frac{h^2}{2}$ we obtain exactly the averaging filter (1). It should be noted that the stability of our scheme is a consequence of the arithmetic mean used in the fraction in (10) to approximate the diffusivity at the position of the pixel x_i.

A Weighted Averaging Variant Involving the Central Pixel. The filter (1) does not involve the central pixel u_i itself in the average. This might cause problems for certain initial signals: If we choose f to be an alternating signal with two different values, then applying the filter will simply exchange the grey values. To avoid this problem one can give the central pixel a nonnegative weight and involve it in the averaging process. For example, such a modified scheme could look like

$$u_i^{k+1} = \frac{g\left(\left|\frac{u_{i+1}^k - u_i^k}{h}\right|\right) u_{i+1}^k + \alpha u_i^k + g\left(\left|\frac{u_{i-1}^k - u_i^k}{h}\right|\right) u_{i-1}^k}{g\left(\left|\frac{u_{i+1}^k - u_i^k}{h}\right|\right) + \alpha + g\left(\left|\frac{u_{i-1}^k - u_i^k}{h}\right|\right)} \tag{12}$$

where we have given the central pixel a fixed weight $\alpha > 0$. The same reasoning as presented above relates this averaging filter to the PDE

$$\partial_t u = \frac{1}{\frac{\alpha}{2} + g(|\partial_x u|)} \partial_x \left(g(|\partial_x u|)\, \partial_x u \right) \quad . \tag{13}$$

Here we see that there is still some factor influencing the velocity of the diffusion process, but this factor now is bounded from above to $\frac{2}{\alpha}$. Compared to (8), this slows down the evolution in regions with small derivatives of u.

3 Averaging Filters and Scaling Limits in 2-D

In this section we consider filtering of images with a two-dimensional domain with weighted averaging over the direct neighbouring pixels. Let $\mathcal{N}(i)$ be the set of indices of the maximal four direct neighbours of the pixel with index i. Then an equivalent of the weighted averaging filter (1) in two dimensions can be written as

$$u_i^{k+1} = \frac{\sum\limits_{j \in \mathcal{N}(i)} g\left(\left| \frac{u_j^k - u_i^k}{h} \right| \right) u_j^k}{\sum\limits_{j \in \mathcal{N}(i)} g\left(\left| \frac{u_j^k - u_i^k}{h} \right| \right)} \quad . \tag{14}$$

Numerator and denominator of this scheme can be understood as the sum of numerators and denominators of two one-dimensional schemes in x- and y-direction. Thus the reasoning described in the last section shows that (14) is a consistent approximation for

$$\partial_t u = \frac{\partial_x \left(g(|\partial_x u|) \partial_x u \right) + \partial_y \left(g(|\partial_y u|) \partial_y u \right)}{g(|\partial_x u|) + g(|\partial_y u|)} \quad . \tag{15}$$

This equation is not rotationally invariant, and thus will lead to artifacts in images with rotational invariant objects. This indicates that also the weighted averaging method (14) leads to such artifacts which is shown with a practical example in Fig. 2.

To circumvent this shortcoming, we understand equation (15) as a crude approximation of the rotationally invariant equation

$$\partial_t u = \frac{1}{\int\limits_0^\pi g(|\partial_{e_\varphi} u|)\, d\varphi} \cdot \int\limits_0^\pi \partial_{e_\varphi} \left(g(|\partial_{e_\varphi} u|) \partial_{e_\varphi} u \right)\, d\varphi \tag{16}$$

where we write $e_\varphi = (\cos(\varphi), \sin(\varphi))^T$ for the unit vector in direction φ. In (15) the integrals are approximated as trapez sums where only two evaluation points of the integrands are used. Similar to [13] we introduce a smoothing of the argument of the diffusivity by the convolution of u with a Gaussian kernel of standard deviation σ, and we write $u_\sigma = K_\sigma * u$. This convolution can also

simply be introduced in the arguments of the weights used in the averaging process (14). It does not affect the reasoning leading to the PDE (16).

An equation similar to (16) has been studied in [11] in the context of anisotropic diffusion filtering:

$$\partial_t u = \frac{2}{\pi} \int_0^{\pi} \partial_{e_\varphi} \left(g(|\partial_{e_\varphi} u_\sigma|) \partial_{e_\varphi} u \right) \, d\varphi. \tag{17}$$

The proofs in [11] can be applied to show that (16) can be transformed into

$$\partial_t u = \frac{1}{\operatorname{trace}(D(\nabla u_\sigma))} \operatorname{div} \left(D(\nabla u_\sigma) \cdot \nabla u \right) \tag{18}$$

with the diffusion tensor $D(\nabla u_\sigma) = \int_0^{\pi} e_\varphi e_\varphi^\top \, g(|\partial_{e_\varphi} u_\sigma|) \, d\varphi$. In [11] the eigenvectors of this diffusion tensor $D(\nabla u_\sigma)$ are calculated as $v_1(\psi) = (-\sin(\psi), \cos(\psi))^T$ and $v_2(\psi) = (\cos(\psi), \sin(\psi))^T$ where $\nabla u_\sigma \neq 0$ and (r, φ) are the polar coordinates of ∇u_σ. That means v_1 is the direction of the isophote of u_σ (along an edge), while v_2 is the direction across the edge. The corresponding eigenvalues are given by

$$\lambda_1(\nabla u_\sigma) = \int_0^{\pi} \sin^2(\varphi) \, g(|\partial_{e_\varphi} u_\sigma|) \, d\varphi \quad \text{and} \tag{19}$$

$$\lambda_2(\nabla u_\sigma) = \int_0^{\pi} \cos^2(\varphi) \, g(|\partial_{e_\varphi} u_\sigma|) \, d\varphi \ . \tag{20}$$

Equation (18) is the relevant formulation for practical implementations. This equation is rotationally invariant, since the eigenvectors follow a rotation of the input image, and the eigenvalues are invariant under image rotations. Since $\operatorname{trace}(D(\nabla u_\sigma))$ is always significantly larger than zero, the sharpening of the edges will be less pronounced in this anisotropic case (similar to (13)). Nevertheless, we are going to see with numerical examples that not only preservation of edges, but also sharpening is possible with this filter.

4 Larger Neighbourhood and Rotational Invariance

In the last section we have derived an anisotropic PDE filter from a weighted averaging of the direct neighbouring pixels. To circumvent the lack of rotational invariance in (15) we have understood it as a very crude approximation of the rotational invariant approach (16). Nevertheless, there are discrete filters which address the problem of lacking rotational invariance by involving information from pixels in a larger neighborhood.

We consider here the prominent example of the bilateral filter [5,8,6]. Even though this filter is proposed as a noniterative method, it can make sense to perform several filtering steps; thus we understand it as an iterative averaging filter. In one filtering step, not only the direct neighbouring pixels are involved

in the averaging, but an extended neighbourhood $i + \mathcal{B}_R$. Here $\mathcal{B}_R = \{j \in \mathbb{R}^2 : |j| \le R\} \cap G_h$ denotes the intersection of the disc of radius R in $\mathbb{R}^2$ with the pixel grid G_h. A variant of the bilateral filter then looks like this:

$$u_i^{k+1} = \frac{\sum\limits_{j \in \mathcal{B}_R} g\left(\frac{|u_{i+j}^k - u_i^k|}{|j|}\right) \frac{w(|j|)}{|j|^2} u_{i+j}^k}{\sum\limits_{j \in \mathcal{B}_R} g\left(\frac{|u_{i+j}^k - u_i^k|}{|j|}\right) \frac{w(|j|)}{|j|^2}} \quad . \tag{21}$$

The spatial distance between u_i and u_{i+j} results in a usually smaller weight $w(|j|)/|j|^2$, where an example for w is $w(h) = h^2 \exp(-h^2)$. In this special example, w leads to a Gaussian weight depending on the distance of the two pixels.

We now want to imitate the approach presented in Section 2. To this end we only consider one half of the disc $\mathcal{B}_R^+ = \{(x, y) \in \mathcal{B}_R | x \ge 0\}$ and rewrite the sums in (21) as

$$u_i^{k+1} = \frac{\sum\limits_{j \in \mathcal{B}_R^+} \frac{w(|j|)}{|j|^2} \left(g\left(\frac{|u_{i+j}^k - u_i^k|}{|j|}\right) u_{i+j}^k + g\left(\frac{|u_i^k - u_{i-j}^k|}{|j|}\right) u_{i-j}^k \right)}{\sum\limits_{j \in \mathcal{B}_R^+} \frac{w(|j|)}{|j|^2} \left(g\left(\frac{|u_{i+j}^k - u_i^k|}{|j|}\right) + g\left(\frac{|u_i^k - u_{i-j}^k|}{|j|}\right) \right)} \quad . \tag{22}$$

The novelty in this two-dimensional case is that we have to consider several directional derivatives. We see that there appear directional finite differences in (22). Let $e_\varphi = \frac{j}{|j|}$ be the unit vector pointing in the direction of $j \neq 0$, and $h = |j|$ be the length of the vector j. A Taylor expansion of u around the pixel i yields

$$u_{i+j} = u_i + \langle \nabla u, j \rangle + \mathcal{O}(h^2) = u_i + (\partial_{e_\varphi} u) \cdot h + \mathcal{O}(h^2)$$

which will be useful in the form

$$\frac{u_{i+j} - u_i}{h} = \partial_{e_\varphi} u + \mathcal{O}(h) \quad . \tag{23}$$

Applying the Taylor formula (23) to (22) allows us to write

$$u_i^{k+1} - u_i^k = \frac{\sum\limits_{j \in \mathcal{B}_R^+} w(h) \left(\partial_{e_\varphi} \left(g(|\partial_{e_\varphi} u|) \partial_{e_\varphi} u \right) + \mathcal{O}(h^2) \right)}{\sum\limits_{j \in \mathcal{B}_R^+} \frac{2w(h)}{h^2} \left(g(|\partial_{e_\varphi} u|) + \mathcal{O}(h^2) \right)} \tag{24}$$

At this point we investigate the scaling limit if we let the spatial step sizes in x- and y-direction tend to zero while we keep the size R of the neighbourhood fixed. This means that the number of grid points in our neighbourhood $\mathcal{B}_R$ is tending to infinity. Thus we can consider the sums in (24) as Riemann sums which approximate integrals over the set $\mathcal{B}_R^+$:

$$u^{k+1} - u^k = \frac{1}{\int\limits_0^R \frac{2w(h)}{h^2} \int\limits_0^\pi g(|\partial_{e_\varphi} u^k|) \, d\varphi \, dh} \cdot \int\limits_0^R w(h) \int\limits_0^\pi \partial_{e_\varphi} \left(g(|\partial_{e_\varphi} u^k|) \partial_{e_\varphi} u^k \right) \, d\varphi \, dh \tag{25}$$

Since the inner integrals do not depend on the radius r, the outer ones are just a scaling factor, that means (25) is corresponding up to a constant factor to

$$u^{k+1} - u^k = \frac{1}{\int_0^\pi g(|\partial_{e_\varphi} u^k|)\, d\varphi} \cdot \int_0^\pi \partial_{e_\varphi} \left(g(|\partial_{e_\varphi} u^k|) \partial_{e_\varphi} u^k \right) d\varphi \ . \tag{26}$$

If we understand the right-hand side as temporal forward difference we can see (26) as an approximation to (16). This provides a novel interpretation of bilateral filtering as an anisotropic PDE.

5 Experiments

Now we show some numerical examples to illustrate the practical behaviour of averaging methods and our novel PDE methods. As weight function or diffusivity we use the classical diffusivity $g(s) = \left(1 + \frac{s^2}{\lambda^2}\right)^{-1}$ by Perona and Malik [12].

First we display an experiment in the one-dimensional case in Fig. 1. We see that the presence of the acceleration factor allows for sharper edges. With the same evolution time we can achieve a stronger edge enhancement than with a classical nonlinear diffusion equation of Perona-Malik type.

Figure 2 visualises the lack of rotational invariance of local averaging filters and how it can be improved with a larger neighbourhood in the bilateral filter. Even a better effect than extending the neighbourhood can be achieved with the anisotropic nonlinear diffusion equation (16).

Figure 3 shows the denoising capabilities of the anisotropic diffusion equation (16) for real-world data. The anisotropic behaviour is clearly visible.

6 Conclusions

We have described the close relationship between weighted averaging processes and filters based on partial differential equations with an acceleration factor. In

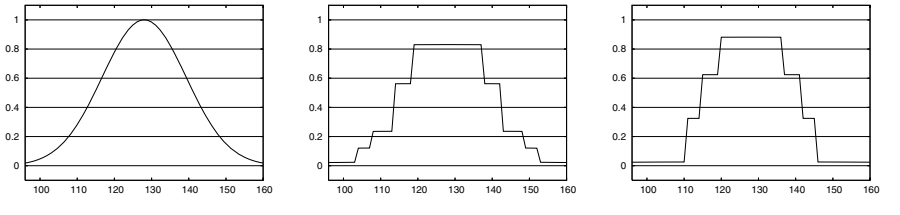

Fig. 1. Accelerated Perona-Malik diffusion in 1-D. Left: Original signal (64 point width section of a signal with 256 pixels). Middle: Perona-Malik diffusion ($\lambda = 0.005$, $t = 5000$). Right: Perona-Malik diffusion with additional factor (8) and the same parameters.

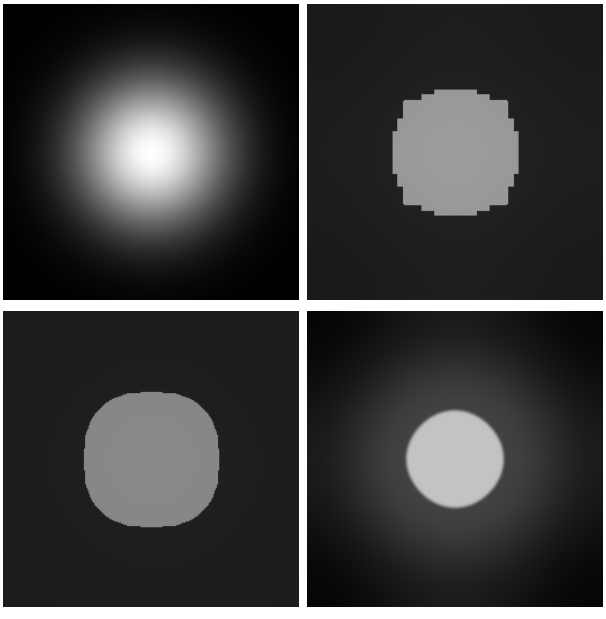

Fig. 2. Weighted averaging and accelerated diffusion. Top Left: Original image (size: 256 x 256 pixels). Top Right: Weighted averaging (equation (14), $\lambda = 3.0$, 15000 iterations). Bottom Left: Iterated bilateral filtering ($\lambda = 3.0$, window size 5×5 pixels, $w(h) = h^2 \exp\left(-h^2/4\right)$, 5000 iterations). Bottom Right: Accelerated anisotropic diffusion ($\lambda = 10$, $\sigma = 2$, $t = 1660$).

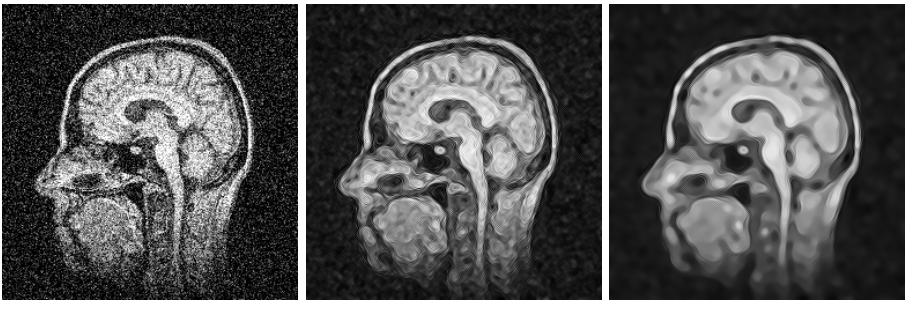

Fig. 3. Accelerated diffusion. Left: Original image (size: 256 x 256 pixels) and additive Gaussian noise with standard deviation 50. Middle: Accelerated anisotropic diffusion ($\lambda = 2$, $\sigma = 3$, $t = 2$). Right: Same, but with $t = 10$.

the 1-D setting we have shown that a suitable scaling limit leads to a modification of the nonlinear diffusion filter of Perona and Malik [12]. The modification consists of a factor that accelerates the sharpening of edges and may give an improved edge enhancement. In the two-dimensional setting, choosing only a small neighbourhood for the averaging can lead to lack of rotational invariance.

However, it can be regarded as a crude approximation of a rotationally invariant PDE that resembles the anisotropic diffusion filter of Weickert [11]. We have also derived the same PDE as a scaling limit of bilateral filtering. This provides additional insights in the behaviour of the widely-used bilateral filter, and shows a way how to improve its invariance under rotations. It is our hope that these examples will motivate more people to analyse the fruitful connections between averaging filters and PDE-based methods in the future.

Acknowledgements. We gratefully acknowledge partly funding by the *Deutsche Forschungsgemeinschaft (DFG)*, project WE 2602/2-2.

References

1. Lee, J.S.: Digital image smoothing and the sigma filter. Computer Vision, Graphics, and Image Processing **24** (1983) 255–269
2. Saint-Marc, P., Chen, J.S., Medioni, G.: Adaptive smoothing: A general tool for early vision. IEEE Transactions on Pattern Analysis and Machine Intelligence **13** (1991) 514–529
3. Polzehl, J., Spokoiny, V.: Adaptive weights smoothing with applications to image restoration. Journal of the Royal Statistical Society, Series B **62** (2000) 335–354
4. Winkler, G., Aurich, V., Hahn, K., Martin, A.: Noise reduction in images: Some recent edge-preserving methods. Pattern Recognition and Image Analysis **9** (1999) 749–766
5. Tomasi, C., Manduchi, R.: Bilateral filtering for gray and colour images. In: Proc. of the 1998 IEEE International Conference on Computer Vision, Bombay, India, Narosa Publishing House (1998) 839–846
6. Elad, M.: On the origin of the bilateral filter and ways to improve it. IEEE Transactions on Image Processing **11** (2002) 1141–1151
7. Smith, S.M., Brady, J.M.: SUSAN – A new approach to low lewel image processing. International Journal of Computer Vision **23** (1997) 43–78
8. Barash, D.: A fundamental relationship between bilateral filtering, adaptive smoothing and the nonlinear diffusion equation. IEEE Transactions on Pattern Analysis and Machine Intelligence **24** (2002) 844–847
9. Mrázek, P., Weickert, J., Bruhn, A.: On robust estimation and smoothing with spatial and tonal kernels. In Klette, R., Kozera, R., Noakes, L., Weickert, J., eds.: Geometric Properties for Incomplete Data. Volume 31 of Computational Imaging and Vision. Springer, Dordrecht (2006) 335–352
10. Buades, A., Coll, B., Morel, J.M.: A review of image denoising algorithms, with a new one. Multiscale Modeling and Simulation **4** (2005) 490–530
11. Weickert, J.: Anisotropic diffusion filters for image processing based quality control. In Fasano, A., Primicerio, M., eds.: Proc. Seventh European Conference on Mathematics in Industry. Teubner, Stuttgart (1994) 355–362
12. Perona, P., Malik, J.: Scale space and edge detection using anisotropic diffusion. IEEE Trans. Pattern Anal. Mach. Intell. **12** (1990) 629–639
13. Catté, F., Lions, P.L., Morel, J.M., Coll, T.: Image selective smoothing and edge detection by nonlinear diffusion. SIAM Journal on Numerical Analysis **29** (1992) 182–193