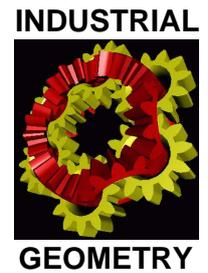


National Research Network S92

Industrial Geometry

<http://www.industrial-geometry.at>



NRN Report No. 102

A Locally Adaptive Total Variation Based Algorithm for Image Restoration

Markus Grasmair

May 2010

FWF

Der Wissenschaftsfonds.



**universität
wien**

A Locally Adaptive Total Variation Based Algorithm for Image Restoration

Markus Grasmair

Computational Science Center
University of Vienna
Nordbergstrasse 15
1090 Wien, Austria

May 10, 2010

Abstract

We introduce a total variation based, locally adaptive method for image denoising and deblurring. The algorithm iteratively updates the regularisation parameter, until a uniform, user specified regularity of the output image is reached. Here, the regularity is measured in terms of local oscillations, which are estimated using a dual formulation of total variation regularisation. Because of the chosen approach, no information on the noise level present in the corrupted images is required. Several numerical experiments are presented that indicate the suitability of the proposed algorithm for dealing with unknown noise in image restoration problems.

1 Introduction

Let $\Omega \subset \mathbb{R}^2$ be a bounded domain with Lipschitz boundary, and assume that $f \in L^2(\Omega)$ is a given noisy image. A standard method for the denoising of f is *total variation regularisation* (see [1, 3, 7, 10, 11, 14, 25, 26, 27]), which consists in minimising the functional

$$\mathcal{T}(u; \alpha) := \frac{1}{2} \int_{\Omega} (u(x) - f(x))^2 dx + \alpha |Du|(\Omega)$$

for some fixed *regularisation parameter* $\alpha > 0$. Here $|Du|$ is the *total variation* of the function $u \in L^2(\Omega)$ (see [20]). The minimiser of $\mathcal{T}(\cdot; \alpha)$ is a function $u_{\alpha} \in BV(\Omega)$ that is close to the input data f , but at the same time the number of oscillations is decreased, that is, the function u_{α} is smoother than f .

The amount of smoothing is determined by the regularisation parameter α . A small value implies that most emphasis lies on staying close to the input data. As α increases, the resulting image u_{α} will become more and more cartoon-like, mainly consisting of large, uniform regions well separated by distinct edges.

In order to find a good value for the regularisation parameter, a standard assumption is that the data f are the sum of a true image $f_0 \in BV(\Omega)$ and some random noise $n \in L^2(\Omega)$, which is assumed to have zero mean and variance $\int_{\Omega} n(x)^2 dx = \sigma^2 > 0$. In this case, it is reasonable to choose α in such a way that $\int_{\Omega} (u_{\alpha}(x) - f(x))^2 dx = \sigma^2$ (see [7, 10, 26]). In the context of regularisation theory, this parameter selection method is known as *Morozov's discrepancy principle* (see [18]).

In fact, most of the noise models assume that the function n is the realisation of an i.i.d. random variable. Then one can not only expect that $\int_{\Omega} n(x)^2 dx = \sigma^2$, but that on most subsets $E \subset \Omega$ one has that $\int_E n(x)^2 dx \approx \sigma^2 \mathcal{L}^2(E) / \mathcal{L}^2(\Omega)$. As a consequence, a reasonable parameter choice should entail that $\int_E (u_{\alpha}(x) - f(x))^2 dx \approx \sigma^2 \mathcal{L}^2(E) / \mathcal{L}^2(\Omega)$ for a suitable collection of subsets $E \subset \Omega$. This, however, is only possible, if the regularisation parameter α is allowed to vary locally, depending on the amount of smoothing needed at a point $x \in \Omega$. Consequently, one has to use a *regularisation function* $\alpha: \Omega \rightarrow (0, +\infty)$, and consider the minimisation of the space dependent total variation functional

$$\mathcal{T}(u; \alpha) := \frac{1}{2} \int_{\Omega} (u(x) - f(x))^2 dx + \int_{\Omega} \alpha(x) d|Du|. \quad (1)$$

The idea of adjusting the regularisation function α locally using statistical properties of the residuum $u_{\alpha} - f$ has first been applied in [13] to one-dimensional data $f: I \subset \mathbb{R} \rightarrow \mathbb{R}$. Starting from a large initial guess, the regularisation function α is iteratively *squeezed* locally, until the residuum resembles white noise. Independently of [13], generalisations to two-dimensional data have recently been proposed in [15, 16]. Moreover, the method presented there can also be used for the restoration of colour images and for deblurring. A very similar method has also been discussed in [22] with the goal of preserving texture in noisy images. In contrast to [13, 15, 16], the algorithm in [22] also allows the regularisation function to increase again, if the residuum becomes too small.

The application of space dependent regularisation has also been proposed independently of statistical motivations or justifications. In [21, 29], the regularisation parameter is increased locally near edges in order to further enhance the edge preserving properties of total variation regularisation. A similar enhancement of edges and even corners has also been obtained by applying anisotropic total variation regularisation [6, 24, 28] (see also [19]).

In [29, 30], a different approach to adaptive regularisation has been suggested, relying on the scale sensitivity of the total variation. The method is based on the observation that the speed at which objects in a total variation smoothed image disappear mainly depends on their *scale*, which is defined as the ratio of volume and perimeter (a precise derivation can be found in [2]). This knowledge can be used in an iterative algorithm that adapts the local regularisation function, until all features of sufficiently small scale have been removed from the noisy input image. In contrast to the other methods, this approach requires no knowledge of the noise level, as the choice of the regularisation function depends only on the regularity of the solution, not on properties of the

residual.

In this paper, we propose a different algorithm, which also bases the choice of the regularisation function on the final solution. As opposed to the method of [29, 30], our algorithm does not remove small scale features completely, but rather attempts to reach a uniform regularity over the whole image, measured in terms of the oscillations of a dual variable. The proposed method elaborates on the algorithm presented in [23]. In addition, generalisations to the denoising of colour images and to deconvolution problems are given.

2 Local Variation

Using basic results of convex analysis, one can show that the minimisation of the functional $\mathcal{T}(\cdot; \alpha)$ defined in (1) is equivalent to the solution of a dual minimisation problem. This dual problem reads as

$$\begin{aligned} \mathcal{J}(V) &:= \int_{\Omega} (f(x) + \operatorname{div} V(x))^2 dx \rightarrow \min \\ |V(x)/\alpha(x)| &\leq 1 \text{ almost everywhere on } \Omega, \\ V(x) \cdot \nu(x) &= 0 \text{ almost everywhere on } \partial\Omega. \end{aligned} \tag{2}$$

Here $\nu(x)$ denotes the outward normal to Ω at $x \in \partial\Omega$. The solution of the dual problem is carried out on the space $L^\infty(\Omega; \mathbb{R}^2)$ of essentially bounded, vector valued functions $V: \Omega \rightarrow \mathbb{R}^2$. Thus, the divergence $\operatorname{div} V$ and the boundary conditions $V \cdot \nu = 0$ have to be interpreted in a distributional sense: $v = \operatorname{div} V$ and $V \cdot \nu = 0$ on $\partial\Omega$, if and only if

$$\int_{\Omega} \nabla \phi(x) \cdot V(x) dx = - \int_{\Omega} \phi(x) v(x) dx$$

for every $\phi \in C^\infty(\mathbb{R}^2)$.

The primal and the dual problem are equivalent in the sense that V_α solves (2) if and only if $u_\alpha := f + \operatorname{div} V_\alpha$ minimises $\mathcal{T}(\cdot; \alpha)$. Moreover, the solution V_α can be characterised by the Kuhn–Tucker conditions

$$\operatorname{div} V_\alpha = u_\alpha - f \quad \text{and} \quad V_\alpha(x) \in \alpha(x) \frac{dDu_\alpha}{d|Du_\alpha|}(x),$$

which shows that the scaled function V_α/α equals the direction of the gradient of u_α whenever it is non-zero.

This relation between V_α and u_α can be exploited for estimating the amount of noise still left in u_α . The remaining traces of noise are marked by fast oscillations of u_α , which in turn entail rapid changes of the direction of Du_α . Consequently they can be observed as variations of the variable V_α . For measuring these variations, we choose some smooth convolution kernel $\eta \in C_0^\infty(\mathbb{R}^2)$, that is, $\eta(x) \geq 0$ for every $x \in \mathbb{R}^2$ and $\int_{\mathbb{R}^2} \eta(x) dx = 1$. If η is concentrated at

zero, we can use it for defining a local mean $M_\eta(V_\alpha/\alpha): \Omega \rightarrow \mathbb{R}^2$ as

$$M_\eta(V_\alpha/\alpha)(x) = \eta * (V_\alpha/\alpha)(x) = \int_{\Omega} \frac{V_\alpha(y)}{\alpha(y)} \eta(x-y) dy,$$

and a local variance $V_\eta(V_\alpha/\alpha): \Omega \rightarrow \mathbb{R}_{\geq 0}$ as

$$\begin{aligned} \text{Var}_\eta(V_\alpha/\alpha)(x) &= \eta * (|\eta * (V_\alpha/\alpha) - V_\alpha/\alpha|^2)(x) \\ &= \int_{\Omega} \left| M_\eta(V_\alpha/\alpha)(y) - \frac{V_\alpha(y)}{\alpha(y)} \right|^2 \eta(x-y) dy. \end{aligned} \quad (3)$$

The function $\text{Var}_\eta(V_\alpha/\alpha)$ provides a rough estimate of the local regularity of u_α . If $\text{Var}_\eta(V_\alpha/\alpha)(x) = 0$, then $dDu_\alpha/d|Du_\alpha| = \zeta$ is constant near x . Then there exists a monotonous function g such that $u_\alpha(y) = g(\zeta \cdot (y-x))$ near x . Conversely, if $\text{Var}_\eta(V_\alpha/\alpha)(x) = 1$ —the largest value that can possibly be attained—, then whenever a direction ζ of dDu_α is present near x also its negative $-\zeta$ appears near x , implying that u_α is still irregular. Consequently, a large value of $\text{Var}_\eta(V_\alpha/\alpha)$ indicates that, locally, the regularisation parameter should be increased in order to achieve higher regularity of the solution, while in regions where $\text{Var}_\eta(V_\alpha/\alpha)$ oversmoothing might have occurred and therefore the regularisation parameter should be decreased.

3 The Core Algorithm

In the following we propose an algorithm that achieves the parameter adaptation suggested in the previous section. To that end, we choose some target variation $0 < \theta < 1$ and a convolution kernel η . The algorithm will iteratively adapt the function α , until the solution V_α of (2) satisfies $\text{Var}_\eta(V_\alpha/\alpha) \approx \theta$. In the numerical experiments below, the kernel η was chosen as an isotropic Gaussian of variance σ^2 . In this case, the size of σ corresponds to the size of features expected in the denoised function u_α .

Start with any regularisation function $\alpha_0: \Omega \rightarrow \mathbb{R}_{>0}$. For $i = 1, 2, \dots$, compute

$$V_i = \arg \min \left\{ \int_{\Omega} (f(x) + \text{div } V(x))^2 dx : \|V/\alpha_{i-1}\|_\infty \leq 1, V \cdot \nu = 0 \text{ on } \partial\Omega \right\}.$$

Compute the variance $\text{Var}_\eta(V_i/\alpha_{i-1})$ of the scaled, tentative solution V_i/α_{i-1} according to (3). Update the regularisation function by

$$\alpha_i(x) = \alpha_{i-1}(x) \left(\frac{\text{Var}_\eta(V_i/\alpha_{i-1})(x)}{2\theta} + \frac{1}{2} \right). \quad (4)$$

Repeat until convergence of $u_i := f + \text{div } V_i$.

By the choice of the update of the regularisation function α in (4), the algorithm should terminate, when $\text{Var}_\eta(V_i/\alpha_{i-1})(x) = \theta$ for every $x \in \Omega$. In

practice, however, one cannot expect that this equality will, or even can, be reached. Since the regularity will only increase by applying total variation regularisation, this can, for instance, happen, if the data f are already more regular than indicated by the parameter θ . Though in this case $\text{Var}_\eta(V_i/\alpha_{i-1})$ will always stay below θ , the algorithm will nevertheless converge: the regularisation parameter will tend to zero and the functions u_i to the data f .

Still, it is sensible to introduce a regularising element into the algorithm. To that end, we smooth the function α_i in each step by convolving it with some kernel ρ . The steps then read as

$$\alpha_i(x) = \rho * \left(\alpha_{i-1} \left(\frac{\text{Var}_\eta(V_i/\alpha_{i-1})}{2\theta} + \frac{1}{2} \right) \right)(x). \quad (5)$$

Besides stabilising the update of α , this smoothing has the side-effect of suppressing too rapid variations of the regularisation function α_i .

The method is summarised in Algorithm 1.

Data: noisy image $f \in L^2(\Omega)$;
Input: variance $\theta > 0$; convolution kernels η, ρ ; tolerance $\varepsilon > 0$;
Result: denoised function u ;
Initialization: $i = 1$; $\alpha_0 > 0$; $V_0 = 0$;
repeat
 $V_i = \arg \min_V \{ \|\text{div } V + f\|^2 : \|V/\alpha_{i-1}\|_\infty \leq 1 \}$;
 $\text{Var}_\eta(V_i/\alpha_{i-1}) = \eta * (|\eta * (V_i/\alpha_{i-1}) - V_i/\alpha_{i-1}|^2)$;
 $\tilde{\alpha}_i = \alpha_{i-1} (\text{Var}_\eta(V_i/\alpha_{i-1})/\theta + 1)/2$;
 $\alpha_i = \rho * \tilde{\alpha}_i$;
 $i \mapsto i + 1$;
until $\|\text{div } V_i - \text{div } V_{i-1}\| < \varepsilon$;
 $u = f + \text{div } V_i$;

Algorithm 1: Basic adaptation algorithm

Accelerated algorithm

In each step of Algorithm 1 an instance of total variation regularisation with parameter function α_i is solved. This is achieved by an iterative projected gradient method (see [4, 9, 17]), setting

$$\begin{aligned} \tilde{V}_i^{(k)} &= V_i^{(k-1)} + \tau \nabla (f + \text{div } V_i^{(k-1)}), \\ V_i^{(k)} &= \frac{\tilde{V}_i^{(k)}}{\max\{1, |\tilde{V}_i^{(k)}|/\alpha_{i-1}\}}. \end{aligned} \quad (6)$$

Here the parameter τ should satisfy $0 < \tau < 1/4$ in order to ensure convergence.

In the iteration (6), it is necessary to choose some initial data $V_i^{(0)}$ in each update step of α . An obvious candidate is the solution V_{i-1} of the previous

iteration step, as it already is a solution of the dual of total variation regularisation with the same data f , albeit with a different regularisation parameter. More precisely, in order to satisfy the constraint $|V_i^{(0)}| \leq \alpha_{i-1}$, we choose

$$V_i^{(0)} = \frac{V_{i-1}}{\max\{1, |V_{i-1}|/\alpha_{i-1}\}}.$$

This choice of the initial data greatly decreases computation times compared to the trivial choice $V_i^{(0)} = 0$.

As a further speed improvement, it is possible to calculate a new regularisation function α well before convergence of the iteration (6). Indeed, a sufficiently good guess on whether to increase or decrease α can in general already be obtained after a small number of projected gradient steps. Thus we propose to stop each iteration (6) after some maximal number k_{\max} of steps—in the numerical experiments in Section 5, this number was chosen as $k_{\max} = 10$.

This approach is summarised in Algorithm 2.

Data: noisy image $f \in L^2(\Omega)$;
Input: variance $\theta > 0$; convolution kernels η, ρ ; tolerance $\varepsilon > 0$; number of inner steps $k_{\max} \in \mathbb{N}$; step size $0 < \tau < 1/4$;
Result: denoised function u ;
Initialization: $i = 1$; $\alpha_0 > 0$; $V_0 = 0$;
repeat
 $V_i^0 = V_{i-1} / \max\{1, |V_{i-1}|/\alpha_{i-1}\}$;
 repeat
 $\tilde{V}_i^k = V_i^{k-1} + \tau \nabla(f + \operatorname{div} V_i^{k-1})$;
 $V_i^k = \tilde{V}_i^k / \max\{1, |\tilde{V}_i^k|/\alpha_{i-1}\}$;
 $k \mapsto k + 1$;
 until $\|\operatorname{div} V_i^k - \operatorname{div} V_i^{k-1}\| < \varepsilon$ **or** $k = k_{\max}$;
 $V_i = V_i^{k_{\max}}$;
 $\operatorname{Var}_\eta(V_i/\alpha_{i-1}) = \eta * (|\eta * (V_i/\alpha_{i-1}) - V_i/\alpha_{i-1}|^2)$;
 $\tilde{\alpha}_i = \alpha_{i-1} (\operatorname{Var}_\eta(V_i/\alpha_{i-1})/\theta + 1)/2$;
 $\alpha_i = \rho * \tilde{\alpha}_i$;
 $k = 0$; $i \mapsto i + 1$;
until $\|\operatorname{div} V_i - \operatorname{div} V_{i-1}\| < \varepsilon$;
 $u = f + \operatorname{div} V_i$;

Algorithm 2: Accelerated adaptation algorithm

Adaptation to Colour Images

The Algorithm 2 can be generalised in a straightforward manner to the denoising of vector valued data such as, for instance, colour images. The only difference to the scalar valued case is that, in the colour setting, the gradient of an image $u: \Omega \rightarrow \mathbb{R}^3$ is a tensor valued mapping $\nabla u: \Omega \rightarrow \mathbb{R}^3 \otimes \mathbb{R}^2 \sim \mathbb{R}^{3 \times 2}$. Consequently,

also the dual variable V is tensor valued, that is, $V: \Omega \rightarrow \mathbb{R}^3 \otimes \mathbb{R}^2$. Moreover, the divergence of V has to be interpreted componentwise in the sense that $\operatorname{div} V = (\operatorname{div} V_1, \operatorname{div} V_2, \operatorname{div} V_3): \Omega \rightarrow \mathbb{R}^3$. Also, the boundary condition is interpreted componentwise, that is, $V_i \cdot \nu = 0$ on $\partial\Omega$ for all i . The norm on $\mathbb{R}^3 \otimes \mathbb{R}^2$, which is needed in the formulation of the dual problem (2), is the usual Euclidean one.

4 Application to Deconvolution

In the previous section, we have only discussed the problem of denoising, where the functional to be minimised is

$$\mathcal{T}(u; \alpha) := \frac{1}{2} \|u - f\|^2 + \int_{\Omega} \alpha(x) d|Du|.$$

Now we will apply the same ideas to the problem of deconvolution. Here one has to minimise the functional

$$\mathcal{T}(u; \alpha) := \frac{1}{2} \|Ku - f\|^2 + \int_{\Omega} \alpha(x) d|Du|, \quad (7)$$

where $K: L^2(\Omega) \rightarrow X$ is a bounded, compact, linear operator (the convolution operator) mapping $L^2(\Omega)$ to the Hilbert space X .

Because of its compactness, the operator K is not continuously invertible. As a consequence, the dual problem cannot be easily formulated in a manner similar to (2). This lack of a suitable dual formulation also makes a straightforward generalisation of our adaptive parameter selection method to deconvolution problems impossible. Still, one can apply the ideas of the previous section to deconvolution, if one uses a particular iterative method for minimising (7).

In [4, 5, 12], it has been shown that one can minimise (7) using the iteration

$$\begin{aligned} w_n &= u_n + \mu K^*(f - Ku_n), \\ u_{n+1} &= \arg \min_u \left(\frac{1}{2} \|u - w_n\|^2 + \int_{\Omega} \mu \alpha(x) d|Du| \right). \end{aligned} \quad (8)$$

Here $K^*: X \rightarrow L^2(\Omega)$ denotes the adjoint of the operator K . If $\mu > 0$ is chosen in such a way that $\mu \|K^*K\| < 1$, then the sequence $(u_n)_{n \in \mathbb{N}}$ converges to a minimiser of (7). Thus the deconvolution problem can be solved by iteratively applying denoising steps to the current iterate after adding the smoothed residual.

In order to adapt the regularisation function α , it is now possible to apply after each denoising step the method presented in Section 3. Note however that, instead of α , the smaller regularisation function $\mu\alpha$ is used in the iteration steps of (8). Thus the local variance of the functions $V_i/\mu\alpha_{i-1}$ determines the update of the regularisation function. The complete adaptive deconvolution algorithm is summarised in Algorithm 3.

Data: blurred, noisy image $f \in L^2(\Omega)$;
Input: variance $\theta > 0$; convolution kernels η, ρ ; tolerance $\varepsilon > 0$; number of inner steps $k_{\max} \in \mathbb{N}$; step size $0 < \tau < 1/4$; convolution operator K ; $0 < \mu < \|K^*K\|^{-1}$;
Result: regularised solution u of $Ku = f$;
Initialization: $i = 1$; $\alpha_0 > 0$; $V_0 = 0$; $w_0 = \mu K^* f$;
repeat
 $V_i^0 = V_{i-1} / \max\{1, |V_{i-1}| / \mu \alpha_{i-1}\}$;
 repeat
 $\tilde{V}_i^k = V_i^{k-1} + \tau \nabla (w_{i-1} + \operatorname{div} V_i^{k-1})$;
 $V_i^k = \tilde{V}_i^k / \max\{1, |\tilde{V}_i^k| / \mu \alpha_{i-1}\}$;
 $k \mapsto k + 1$;
 until $\|\operatorname{div} V_i^k - \operatorname{div} V_i^{k-1}\| < \varepsilon$ **or** $k = k_{\max}$;
 $V_i = V_i^{k_{\max}}$;
 $\operatorname{Var}_\eta(V_i / \mu \alpha_{i-1}) = \eta * (|\eta * (V_i / \mu \alpha_{i-1}) - V_i / \mu \alpha_{i-1}|^2)$;
 $\tilde{\alpha}_i = \alpha_{i-1} (\operatorname{Var}_\eta(V_i / \mu \alpha_{i-1}) / \theta + 1) / 2$;
 $\alpha_i = \rho * \tilde{\alpha}_i$;
 $u_i = w_{i-1} + \operatorname{div} V_i$;
 $w_i = u_i + \mu K^*(f - Ku_i)$;
 $k = 0$; $i \mapsto i + 1$;
until $\|u_i - u_{i-1}\| < \varepsilon$;
 $u = u_i$;

Algorithm 3: Adaptation algorithm for deconvolution problems

5 Numerical Results

In this section we apply the adaptive algorithm to different test examples. For the numerical results we have used a discretisation by finite differences proposed in [8] (see also [4]). The rectangular domain Ω is partitioned into a uniform grid of size $N \times M$. The discrete gradient of a vector $u \in \mathbb{R}^{N \times M}$ is then defined as

$$(\nabla u)_{i,j}^{(1)} = \begin{cases} u_{i+1,j} - u_{i,j}, & \text{if } i < N, \\ 0, & \text{if } i = N, \end{cases}$$

$$(\nabla u)_{i,j}^{(2)} = \begin{cases} u_{i,j+1} - u_{i,j}, & \text{if } j < M, \\ 0, & \text{if } j = M, \end{cases}$$

its norm as

$$\|\nabla u\| = \sum_{i,j} |(\nabla u)_{i,j}| = \sum_{i,j} \sqrt{((\nabla u)_{i,j}^{(1)})^2 + ((\nabla u)_{i,j}^{(2)})^2}.$$

The adjoint div of $-\nabla$ with respect to the Euclidean scalar product is defined for $V \in (\mathbb{R}^2)^{N \times M}$ as

$$(\text{div } V)_{i,j} = \begin{cases} V_{i,j}^{(1)} - V_{i-1,j}^{(1)}, & \text{if } 1 < i < N, \\ V_{i,j}^{(1)}, & \text{if } i = 1, \\ -V_{i-1,j}^{(1)}, & \text{if } i = N, \end{cases} + \begin{cases} V_{i,j}^{(2)} - V_{i,j-1}^{(2)}, & \text{if } 1 < j < M, \\ V_{i,j}^{(2)}, & \text{if } j = 1, \\ -V_{i,j-1}^{(2)}, & \text{if } j = M. \end{cases}$$

For the kernels η and ρ , which are used in the definition of the variance and also for the smoothing of the regularisation function, we have chosen (discrete approximations to) symmetric Gaussian kernels of variance σ_η and σ_ρ , respectively. In order to reduce boundary artifacts, the convolutions have been computed assuming mirrored boundary conditions.

Gray level images

Our first test example is an image distorted by increasing levels of i.i.d. Gaussian random noise (see Figure 1, first row). We have applied our algorithm both to the original, noise-free image and the different noisy versions (see Figure 1, last row). The variances of the Gaussian convolution kernels η and ρ used in the adaptive method were chosen as $\sigma_\eta = 2$ and $\sigma_\rho = 0.5$, the total image measuring 256×256 pixels. The target variation was set to $\theta = 0.7$, allowing for larger details to remain in the image while removing small scale texture.

For comparison, we have also applied standard total variation regularisation with a constant, non-adaptive regularisation parameter to the same test images. The regularisation parameters were determined in such a way that they yield optimal results in terms of signal to noise ratio for one specific noise level: in the case of a low noise level, α was set to 7 (see Figure 1, second row); for the highest noise level, the parameter choice $\alpha = 45$ has been used (see Figure 1, third row). As a consequence, good denoising results are obtained for the optimal noise levels.

If, however, one uses the non-adaptive method for the denoising of images with a non-optimal noise level, the results deteriorate significantly. If the regularisation parameter is chosen too small, then almost no smoothing is obtained; there is hardly any difference between the noisy and the smoothed image. On the other hand, a too large regularisation parameter yields to oversmoothing, removing all but the largest scale features.

The situation is notably different for our adaptive method. While the images are not as good as standard total variation regularisation with optimal parameter choice, it yields useful results over the whole range of noise levels. Note moreover, that the optimal regularisation parameters for the standard methods have been chosen with knowledge of the original image, whereas no knowledge even of the noise level is assumed for the adaptive algorithm.

In order to illustrate the influence of the parameter θ , the target variation of the denoised image, we have applied our method to a noise-free image for varying θ . The results are shown in Figure 2. One sees that with decreasing θ ,

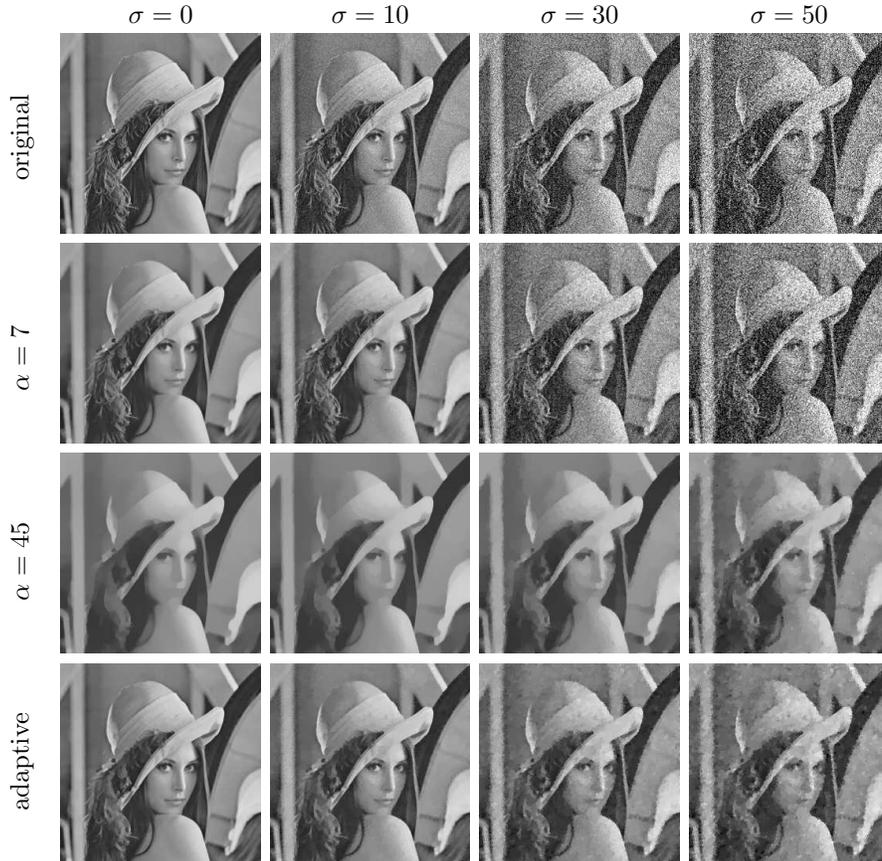


Figure 1: Comparison of the adaptive algorithm with standard total variation regularisation. *First row:* Original image, distorted with Gaussian random noise (variance $\sigma = 0, 10, 30, 50$). *Second row:* Result of standard total variation regularization; the regularisation parameter was chosen optimal for the noise level $\sigma = 10$. *Third row:* Result of standard total variation regularization; the regularisation parameter was chosen optimal for the noise level $\sigma = 50$. *Fourth row:* Result of the adaptive algorithm with $\theta = 0.7$.



Figure 2: Influence of the regularity parameter θ . *Upper row:* Original image and adaptively denoised image with $\theta = 0.8, 0.7$. *Lower row:* Denoised image with $\theta = 0.6, 0.4, 0.2$.

more and more details are lost in the image, starting with texture and the finest scale structures, until only the largest scale structures in the image are left. In some sense, the parameter θ acts therefore in a similar manner as the (inverse of the) regularisation parameter in standard total variation regularisation.

Colour images

In Figure 3 we have applied our denoising algorithm to a colour image, which is distorted by i.i.d. Gaussian random noise of increasing variance (see Figure 3). Again, the adaptivity of the algorithm is clearly visible in the result. In the case of the original image, the fur is treated as noise, while the rest of the image is interpreted as already being clean. As the noise level increases, the amount of regularisation applied to the different parts of the image also increases uniformly (see Figure 3, third column). Only at the most prominent edges between the different colours present in the image, the increase in the regularisation function is slower. Thus one sees that the adaptive algorithm adds a further edge preserving effect to total variation regularisation.

Deconvolution

The final example concerns the application of the proposed algorithm to deblurring. Because of the ill-posedness of the problem, it is necessary to choose smaller target variations θ than for the task of denoising in order to obtain good results. For the numerical tests in Figure 4, the target variation was chosen as

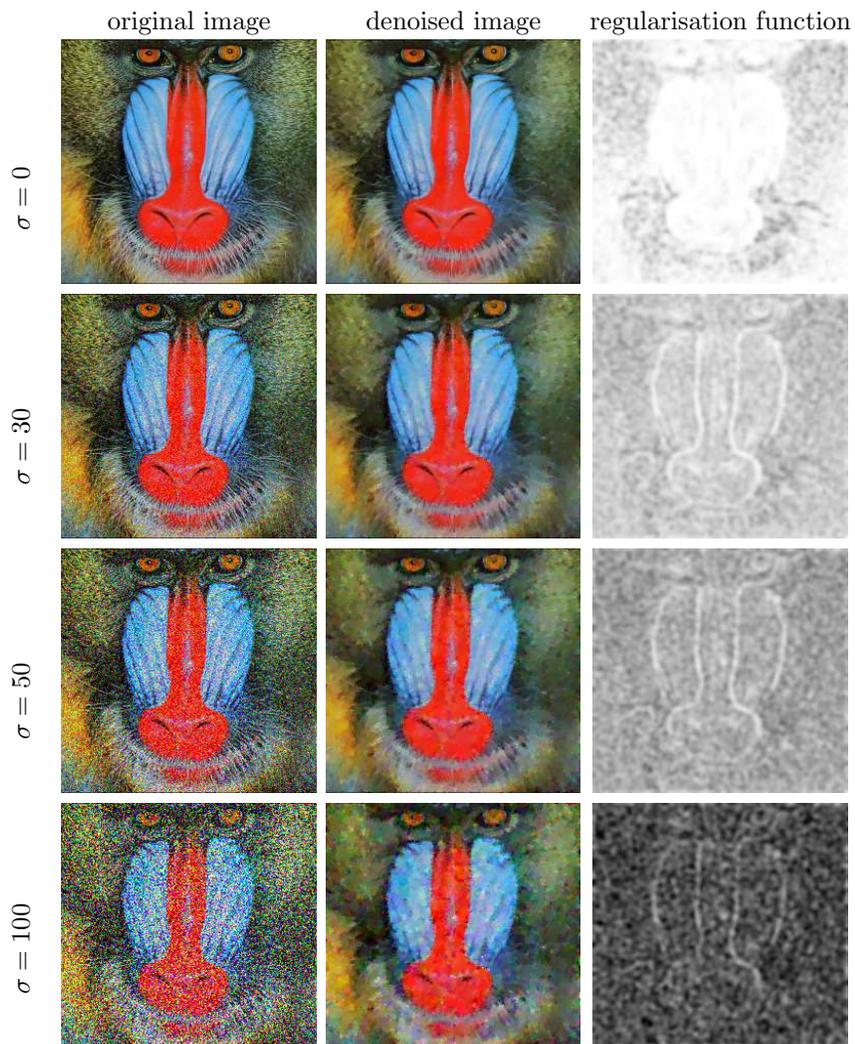


Figure 3: Adaptive total variation regularisation for denoising of colour images. *Left column:* Original image, distorted with Gaussian random noise (variance $\sigma = 0, 30, 50, 100$). *Middle column:* Result of the adaptive algorithm with $\theta = 0.8$. *Right column:* Finally chosen regularisation function; darker pixels indicate higher values of α .

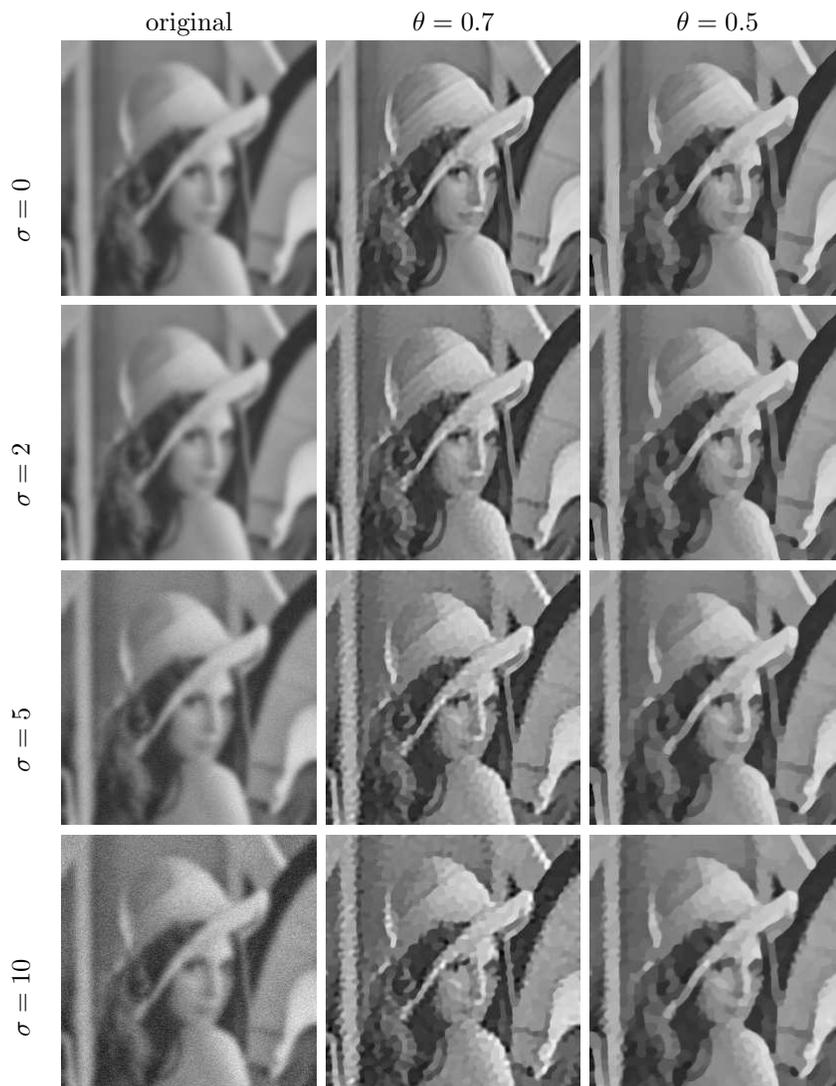


Figure 4: Adaptive total variation regularisation for deconvolution. *Left column:* Blurred image, distorted with Gaussian random noise (variance $\sigma = 0, 2, 5, 10$). *Middle column:* Result of the adaptive algorithm with $\theta = 0.7$. *Right column:* Result of the adaptive algorithm with $\theta = 0.5$.

$\theta = 0.7$ and $\theta = 0.5$. The convolution kernel defining the blurring was a symmetric Gaussian kernel of variance 3, the image measuring 256×256 pixels. In addition, i.i.d. Gaussian random noise of variance up to 10 was added to the blurred image. For small noise levels or a noise free image, the reconstruction is reasonable for both target variations, $\theta = 0.7$ and $\theta = 0.5$. For higher noise levels, however, the resulting image tends to be too irregular in case of the larger target variation; artifacts stemming from noise are retained. In contrast, the reconstruction with the smaller target variation remains quite good.

6 Conclusion

For the denoising of images with unknown or varying noise levels by means of total variation regularisation, it is necessary to adapt the regularisation parameter locally. The adaptation method proposed in this paper is based on the ideas presented in [23]. It bases the choice of the ensuing regularisation function on the regularity of the final image, which is measured using the variations of a dual variable. These variations correspond to local changes in the direction of the gradient of the output image, and therefore can indicate, whether noise is still present in the solution. The proposed algorithm can be directly applied to the denoising of both gray value and colour images. In addition, a generalisation to the inversion of linear operators, as for instance blurring operators, has been presented. The numerical examples support the claim that the algorithm is suited for the reconstruction of noise images without knowledge of the noise level.

Acknowledgement

This work has been supported by the Austrian Science Fund (FWF) within the national research network *Industrial Geometry*, project 9203-N12.

References

- [1] R. Acar and C. R. Vogel. Analysis of bounded variation penalty methods for ill-posed problems. *Inverse Probl.*, 10(6):1217–1229, 1994.
- [2] W. K. Allard. Total variation regularization for image denoising. I. Geometric theory. *SIAM J. Math. Anal.*, 39(4):1150–1190, 2007/08.
- [3] G. Aubert and P. Kornprobst. *Mathematical problems in image processing*, volume 147 of *Applied Mathematical Sciences*. Springer, New York, second edition, 2006. Partial differential equations and the calculus of variations, With a foreword by Olivier Faugeras.
- [4] J.-F. Aujol. Some first-order algorithms for total variation based image restoration. *J. Math. Imaging Vision*, 34(3):307–327, 2009.
- [5] J. Bect, L. Blanc-Féraud, G. Aubert, and A. Chambolle. A l^1 -unified variational framework for image restoration. In *Proceedings of the 8th European conference on*

- computer vision*, volume 3024 of *Lecture Notes in Computer Sciences*. Springer, Berlin, 2004.
- [6] B. Berkels, M. Burger, M. Droske, O. Nemitz, and M. Rumpf. Cartoon extraction based on anisotropic image classification. In *Vision, Modeling, and Visualization Proceedings*, pages 293–300, 2006.
 - [7] M. Burger and S. Osher. Convergence rates of convex variational regularization. *Inverse Probl.*, 20(5):1411–1421, 2004.
 - [8] A. Chambolle. An algorithm for total variation minimization and applications. *J. Math. Imaging Vision*, 20(1–2):89–97, 2004.
 - [9] A. Chambolle. Total variation minimization and a class of binary MRF models. In *Energy Minimization Methods in Computer Vision and Pattern Recognition*, volume 3757 of *Lecture Notes in Computer Vision*, pages 136–152. Springer Berlin/Heidelberg, 2005.
 - [10] A. Chambolle and P.-L. Lions. Image recovery via total variation minimization and related problems. *Numer. Math.*, 76(2):167–188, 1997.
 - [11] G. Chavent and K. Kunisch. Regularization of linear least squares problems by total bounded variation. *ESAIM Control Optim. Calc. Var.*, 2:359–376, 1997.
 - [12] I. Daubechies, M. Defrise, and C. De Mol. An iterative thresholding algorithm for linear inverse problems with a sparsity constraint. *Comm. Pure Appl. Math.*, 57(11):1413–1457, 2004.
 - [13] P. L. Davies and A. Kovac. Local extremes, runs, strings and multiresolution. *Ann. Statist.*, 29(1):1–65, 2001.
 - [14] D.C. Dobson and O. Scherzer. Analysis of regularized total variation penalty methods for denoising. *Inverse Probl.*, 12:601–617, 1996.
 - [15] Y. Dong and M. Hintermüller. Multi-scale vectorial total variation with automated regularization parameter selection for color image restoration. In [31], pages 271–281, 2009.
 - [16] Y. Dong, M. Hintermüller, and M.M. Rincon-Camacho. A multi-scale vectorial L^T -TV framework for color image restoration. IFB-Report 28, Institute of Mathematics and Scientific Computing, University of Graz, 2009.
 - [17] V. Duval, J.-F. Aujol, and L. Vese. Projected gradient based color image decomposition. In [31], pages 295–306, 2009.
 - [18] H. W. Engl, M. Hanke, and A. Neubauer. *Regularization of inverse problems*, volume 375 of *Mathematics and its Applications*. Kluwer Academic Publishers Group, Dordrecht, 1996.
 - [19] S. Esedoğlu and S. Osher. Decomposition of images by the anisotropic Rudin-Osher-Fatemi model. *Comm. Pure Appl. Math.*, 57(12):1609–1626, 2004.
 - [20] L. C. Evans and R. F. Gariepy. *Measure theory and fine properties of functions*. Studies in Advanced Mathematics. CRC Press, Boca Raton, FL, 1992.
 - [21] I. A. Frigaard and O. Scherzer. Herschel–Bulkley diffusion filtering: non-Newtonian fluid mechanics in image processing. *Z. Angew. Math. Mech.*, 86(6):474–494, 2006.
 - [22] G. Gilboa, N. Sochen, and Y.Y. Zeevi. Variational denoising of partly-textured images by spatially varying constraints. *IEEE Trans. Image Process.*, 15(8):2281–2289, 2006.

- [23] M. Grasmair. Locally adaptive total variation regularization. In [31], pages 331–342, 2009.
- [24] M. Grasmair and F. Lenzen. Anisotropic total variation filtering. Reports of FSP S092 - "Industrial Geometry" 84, University of Innsbruck, Austria, 2009. Submitted.
- [25] M. Z. Nashed and O. Scherzer. Least squares and bounded variation regularization with nondifferentiable functional. *Numer. Funct. Anal. Optim.*, 19(7&8):873–901, 1998.
- [26] L. I. Rudin, S. Osher, and E. Fatemi. Nonlinear total variation based noise removal algorithms. *Phys. D*, 60(1–4):259–268, 1992.
- [27] O. Scherzer, M. Grasmair, H. Grossauer, M. Haltmeier, and F. Lenzen. *Variational Methods in Imaging*, volume 167 of *Applied Mathematical Sciences*. Springer, New York, 2009.
- [28] G. Steidl and T. Teuber. Anisotropic smoothing using double orientations. In [31], pages 477–489, 2009.
- [29] D.M. Strong. Adaptive Total Variation Minimizing Image Restoration. CAM Report 97-38, University of California, Los Angeles, 1997.
- [30] D.M. Strong, J.-F. Aujol, and T.F. Chan. Scale recognition, regularization parameter selection, and Meyer's G norm in total variation regularization. *Multiscale Model. Simul.*, 5(1):273–303 (electronic), 2006.
- [31] X. Tai, K. Mørken, M. Lysaker, and K.-A. Lie, editors. *Scale Space and Variational Methods in Computer Vision*, volume 5567 of *Lecture Notes in Computer Science*, Berlin/Heidelberg, 2009. Springer. Second International Conference, SSVM 2009, Voss, Norway, June 1-5, 2009. Proceedings.