Seminar ”Modern Methods in Image Analysis”
Winter Term 2008/2009

On Robust Estimation and Smoothing with Spatial and Tonal Kernels

Jan Hendrik Dithmar
jadi5003@stud.uni-saarland.de

January 27, 2009

1 Introduction

Nowadays, image denoising and simplification are well established. There exist several methods – for example local M-smoothers or bilateral filters – and new ones continue to appear. The problem is that the advantages of the different approaches and their relations are not always clear and only partially understood. It is natural to pose the question: what can be done to understand the different approaches in a better way?

The paper ”On Robust Estimation and Smoothing with Spatial and Tonal Kernels”, on which this write-up bases, looks at different methods and tries to point out their similarities.

2 Statistical estimation

For the approach of statistical estimation, let us assume there is an unknown (constant) signal u and it is observed N-times. The problem is that we actually do not measure the signal u, but a noisy variant of it

$$f_j \quad = \quad u + n_j \qquad (j = 1, ..., N) \tag{1}$$

where n stands for the noise. The goal is clear now, namely to estimate u by the use of this measured, noisy data, which leads to the theory of M-estimation. An M-estimate of u from the noisy data f_j is found by minimising

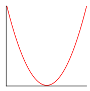
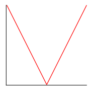
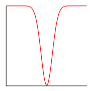
$$E(u) \quad = \quad \sum_{j=1}^{N} \Psi(|u - f_j|^2) \tag{2}$$

where Ψ is some increasing function, also called error norm or penaliser. Examples for such a penaliser are given in table 1.

Penaliser (a) from table 1 gives the arithmetic mean – which is good for removing Gaussian noise – and minimises the L^2 distance

$$E(u) \quad = \quad \sum_{j=1}^{N} |u - f_j|^2. \tag{3}$$

Table 1: Examples for error norms for M-estimators.

	error norm		*estimation result*
(a)	$\Psi(s^2) = s^2$		mean
(b)	$\Psi(s^2) = s$		median
(c)	$\Psi(s^2) = 1 - e^{-s^2/\lambda^2}$		mode approximation

Penaliser (b) gives the median – which is good for removing impulse noise – and minimises the L^1 distance

$$E(u) \quad = \quad \sum_{j=1}^{N} |u - f_j|. \tag{4}$$

The estimation result from penaliser (c) is called mode approximation where the influence of outliers is very much reduced.

But for the approach of statistical estimation, a problem arises. Remembering formula (2), we see that we minimise for only one scalar value u. This gives one estimate for the whole signal which is not desired. As a consequence, the model of statistical estimation has to be refined to overcome this problem.

3 Histogram operations

In image analysis, the data f_j (grey values) is measured at positions x_j (pixels). In other words, we usually want a solution $\mathbf{u} = (u_i)_{i=1,\ldots,N}$ where each output value u_i belongs to the position x_i instead of a single estimate u. The construction of the M-estimates u_i can be done by minimising

$$E(\mathbf{u}) \quad = \quad \sum_{i=1}^{N} \sum_{j=1}^{N} \Psi(|u_i - f_j|^2). \tag{5}$$

The minimisation of $E(\mathbf{u})$ can be done by gradient descent, where each element u_i may be processed independently. Initialising by $u_i^0 = f_i$, the gradient descent becomes

$$\begin{aligned} u_i^{k+1} &= u_i^k - \tau \frac{\partial E}{\partial u_i} \\ &= u_i^k - \tau \sum_{j=1}^{N} \Psi'(|u_i^k - f_j|^2) \cdot 2 \cdot (u_i^k - f_j) \end{aligned}$$

$$= \left(1 - 2\tau \sum_{j=1}^{N} \Psi'(|u_i^k - f_j|^2)\right) u_i^k$$
$$+2\tau \sum_{j=1}^{N} \Psi'(|u_i^k - f_j|^2) f_j \qquad (6)$$

where τ is the step size. One can speed up convergence by choosing τ adaptively to the data, e.g. making it smaller in areas of large slope. Setting

$$\tau := \frac{1}{2\sum_{j=1}^{N} \Psi'(|u_i^k - f_j|^2)}, \qquad g(s^2) := \Psi'(s^2), \qquad (7)$$

equation (6) can be rewritten into the iterative formula

$$u_i^{k+1} = \frac{\sum_{j=1}^{N} g(|u_i^k - f_j|^2) f_j}{\sum_{j=1}^{N} g(|u_i^k - f_j|^2)} \qquad (8)$$

To summarise: the model of histogram operations is an improvement in comparison to statistical estimation, since the estimation is done for each pixel x_i. But if we look closer, we can identify a further disadvantage: the estimation of the different u_i is performed by considering all pixels, which means that also pixels far away talk to each other and participate in determining the value of u_i. Such a method is called global. It is often desirable to estimate a grey value of a pixel from a local neighbourhood only.

4 Local M-smoothers

There exists a natural way to overcome the disadvantage mentioned in the previous paragraph. In the case of M-estimation, the introduction of a second weighting term, which depends on the spatial distance between the position of the restored pixel u_i and the input sample f_j, is sufficent to estimate the grey value of a pixel in a local neighbourhood only. For the local M-smoothers, the functional to minimise reads as follows:

$$E(\mathbf{u}) = \sum_{i=1}^{N} \sum_{j \in \mathcal{B}(i)} \Psi_D(|u_i - f_j|^2) w_D(|x_i - x_j|^2) \qquad (9)$$

where $\mathcal{B}(i)$ is introduced for computational convenience only to make the index j run through the neighbourhood of x_i. The spatial weights w represent for example a hard disk-shaped window

$$w(s^2) = \begin{cases} 1 & s^2 < \theta \\ 0 & \text{otherwise} \end{cases} \qquad (10)$$

or a soft window

$$w(s^2) = e^{-s^2/\theta^2}. \qquad (11)$$

In section 3, an example was given how to compute an iterative formula. Similar to that, the following iterative formula can be computed:

$$u_i^0 = f_i$$
$$u_i^{k+1} = \frac{\sum_{j \in \mathcal{B}(i)} g(|u_i^k - f_j|^2) w(|x_i - x_j|^2) f_j}{\sum_{j \in \mathcal{B}(i)} g(|u_i^k - f_j|^2) w(|x_i - x_j|^2)} \qquad (12)$$

where g is called the tonal weight and w is called the spatial weight. This procedure is called *W-estimator* and is a possibility to obtain a solution to the local M-estimation problem. Please note, that we are only interested in the steady state for $k \to \infty$, not in the evolution towards the minimiser.

5 Bilateral filtering

Let us focus on the smoothness term $E_S(\mathbf{u}) = \Psi_S(|\nabla u|^2)$ which we want to estimate in the discrete world as a sum of squared differences from a pixel to its neighbours:

$$E_S(\mathbf{u}) = \sum_{i=1}^{N} \Psi_S \left(\sum_{j \in \mathcal{N}(i)} |u_i - u_j|^2 \right) \tag{13}$$

where $\mathcal{N}(i)$ is the set of 4-neighbours of a pixel i. In this case, an interesting fact pops up. Just by exchanging the order of summation and penalisation in the last term, we can express (13) as an anisotropic smoothness measure:

$$E_S(\mathbf{u}) \quad = \quad \sum_{i=1}^{N} \sum_{j \in \mathcal{N}(i)} \Psi_S(|u_i - u_j|^2). \tag{14}$$

If we now increase the size of the neighbourhood from which the expression (14) is estimated, we end up in

$$E_S(\mathbf{u}) = \sum_{i=1}^{N} \sum_{j \in \mathcal{B}(i)} \Psi_S(|u_i - u_j|^2) w(|x_i - x_j|^2) \tag{15}$$

where $\mathcal{B}(i)$ is the larger neighbourhood set, and the summation is additionally weighted by a function w of the spatial distance between pixels. Minimising this smoothness term can be done by the iterative procedure

$$\begin{aligned} u_i^0 &= f_i \\ u_i^{k+1} &= \frac{\sum_{j \in \mathcal{B}(i)} g(|u_i^k - u_j^k|^2) w(|x_i - x_j|^2) u_j^k}{\sum_{j \in \mathcal{B}(i)} g(|u_i^k - u_j^k|^2) w(|x_i - x_j|^2)} \end{aligned} \tag{16}$$

The above formula is called iterative bilateral filter. This stands in contrast to the original proposal of bilateral filtering as a heuristic algorithm. Note, that in this context the evolving image is of interest since we would obtain a flat image for $k \to \infty$.

6 A Unifying Framework

The formulas (12) and (16) look quite similar. And indeed, the main difference lies in the arguments used. We can think of local M-estimators as a data term and of bilateral filtering as a smoothness term. With this in mind, let's recall the data term contribution

$$E_D(\mathbf{u}) \quad = \quad \sum_{i=1}^{N} \sum_{j=1}^{N} \Psi_D(|u_i - f_j|^2) w(|x_i - x_j|^2) \tag{17}$$

Figure 1: Input images for the filtering examples. **Left:** Noise-free version. **Right:** Noisy version. **Authors:** P. Mrazek, J. Weickert, A. Bruhn

and the smoothness term contribution

$$E_S(\mathbf{u}) \quad = \quad \sum_{i=1}^{N}\sum_{j=1}^{N} \Psi_S(|u_i - u_j|^2) w(|x_i - x_j|^2). \tag{18}$$

These two can be combined in the fashion of a convex combination:

$$E(\mathbf{u}) \quad = \quad \sigma E_D(\mathbf{u}) + (1-\sigma) E_S(\mathbf{u}) \qquad 0 \leq \sigma \leq 1. \tag{19}$$

This unifying framework gives – among other advantages – freedom in selecting the penaliser type for Ψ_D and Ψ_S or freedom in the balance between data term and smoothness term.

7 Experiments

Two images are used as inputs for the experiments presented in this section. These images are given in figure 1. It is important to mention that this section is not intended to claim that a method A performs better than a method B. However, it should demonstrate the effect of individual filter components to the final result.

Figure 2 shows the result when changing the window size of the spatial window ($\theta = 1$, $\theta = 3$, $\theta = 10$), cropped circularly into windows of size 3×3, 7×7 and 21×21, respectively. One can see in the top row that the image becomes smoother when increasing the window size. For the bottom row, the effect of the different window sizes is small if the same number of iterations is used (200 iterations in this case).

In figures 3 and 4, all images were created using a 7×7 soft spatial window. We see that the L^2 penaliser (left column) blurs the image most and removes noise very well. The mode approximations on the right column of the two figures preserve the edges very well while being very sensitive to noise. The L^1 penalisation in the middle of figures 3 and 4 seems to represent a good compromise between blurring too much and being too sensitive to noise.

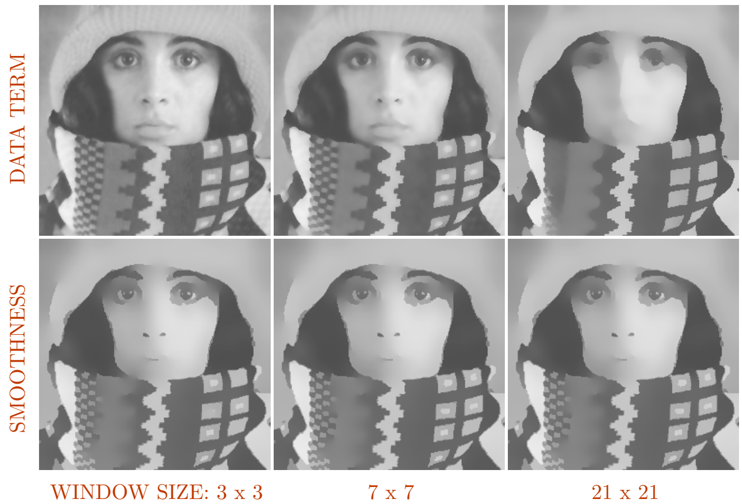


Figure 2: Filtering of the noise-free image using the penaliser $\Psi(s^2) = 1 - e^{-s^2/\lambda^2}$ with soft window $w(s^2) = e^{-s^2/\theta}$ of varied size of the spatial neighbourhood ($\theta = 1, \theta = 3, \theta = 10$). **Top:** local M-smoothers (data term, steady state after iteration). **Bottom:** bilateral filtering (smoothness term, 200 iterations). **Authors:** P. Mrázek, J. Weickert, A. Bruhn.

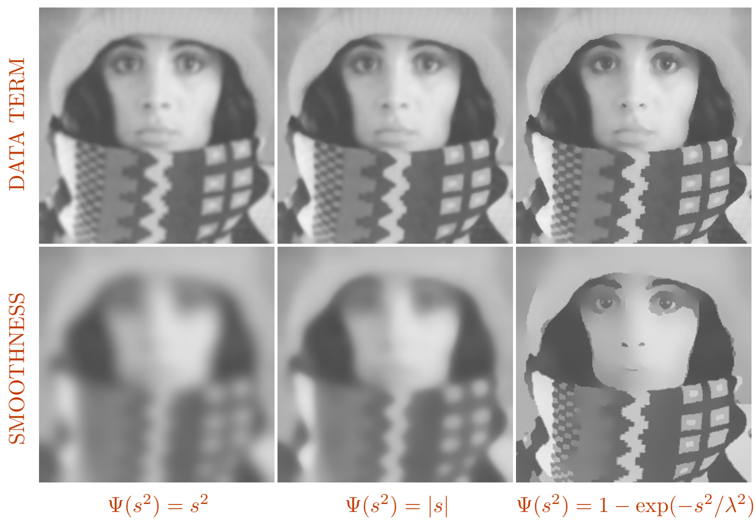


Figure 3: Effect of the penaliser type on the filtering result, starting with the noise-free image (with soft spatial weighting $w(s^2) = e^{-s^2/\theta^2}$, $\theta = 3$). **Authors:** P. Mrázek, J. Weickert, A. Bruhn.

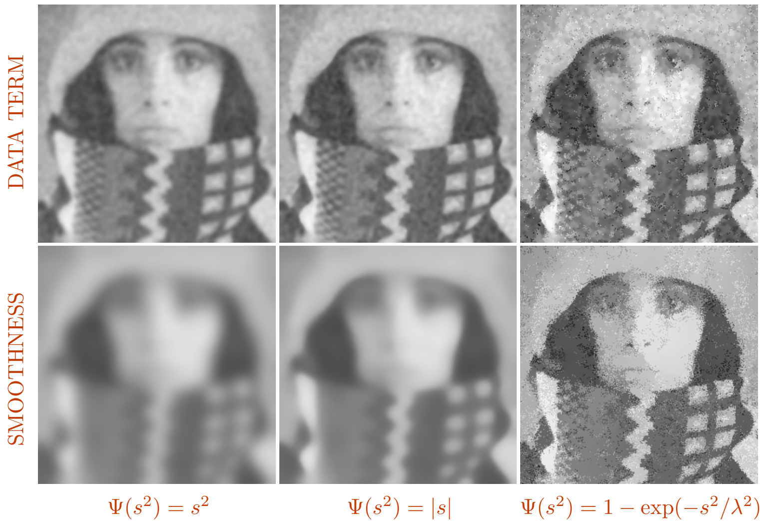


Figure 4: Effect of the penaliser type on the filtering result, starting from the noisy image (filtering in a soft window $w(s^2) = e^{-s^2/\theta^2}$, $\theta = 3$). **Authors:** P. Mrázek, J. Weickert, A. Bruhn.

8 Summary

We have seen that local M-smoothers minimise an energy with a windowed data term. Performing a fixed-point iteration results in W-estimators. In contrast to that we investigated iterative bilateral filtering which minimises an energy with a windowed smoothness term. Since both approaches are pretty similar in the discrete world, it was possible to formulate these in a unifying framework that minimises functionals, combining nonlocal data term and nonlocal smoothness term.

The mutual influence of image pixels is controlled by weighting functions depending on the spatial and tonal distances. The unifying framework helps for a better understanding of the different methods and allows to combine the advantages of known filters to make way for new techniques.

9 References

- R. Klette, R. Kozera, L. Naakes, J. Weickert (Eds.): Geometric Properties for Incomplete Data, 335 – 352. Springer Dordrecht, 2006.
- Lecture ”Differential Equations in Image Processing and Computer Vision”. SS 2008.
  `http://www.mia.uni-saarland.de/Teaching/dic08.shtml`.
- C. K. Chu, I. Glad, F. Godtliebsen, J. S. Marron: Edge-preserving smoothers for image processing. *Journal of the American Statistical Association*, Vol. 93, No. 442, 526–556, 1998.
- G. Winkler, V. Aurich, K. Hahn, A. Martin: Noise reduction in images: some recent edge-preserving methods. *Pattern Recognition and Image Analysis*, Vol. 9, No. 4, 749–766, 1999.
- C. Tomasi, R. Manduchi: Bilateral filtering for gray and color images. *Proc. Sixth International Conference on Computer Vision* (Bombay, India, Jan. 1998), 839–846, Narosa, 1998.