Nonlinear Image Restoration Methods Based on Robust Statistics

D. Y. Jia[1], F. G. Huang[1], H. Su[1]

[1]School of Computer Science and Technology, Harbin Engineering University, China

Abstract

In this paper, a novel smoothness term of Bayesian regularization framework based on M-estimation of robust statistics is proposed, and from this term a class of fourth-order nonlinear diffusion methods are proposed. These methods attempt to approximate an observed image with a piecewise linear image, which looks more natural than piecewise constant image used to approximate an observed image by P-M [1] model.

It was shown [9] that M-estimators and W-estimators are essentially equivalent and solve the same minimization problem. We propose PL bilateral filter from equivalent W-estimator, which is designed for piecewise linear image filtering and more effective than bilateral filter.

Keywords: Bayesian regularization, M-estimation, Nonlinear diffusion, Bilateral filter.

1. Introduction

Since the elegant formulation of P-M model introduced by Perona and Malik [1] and TV-norm by Rudin *et al*. [2], nonlinear diffusion methods became more and more popular in the fields of image processing and computer vision. Both of them are PDE-based techniques that preserve edges well but have the sometimes undesirable staircase effect, see [3] [4] [5], namely the transformation of smooth regions into piecewise constant regions, and cause the processed image to look "blocky". Staircase effect is visually unpleasant and likely to cause a post processing system to falsely recognize as edges the boundaries of different blocks that actually belong to the same smooth area in the original image. To overcome this difficulty several papers have offered using fourth-order PDEs, see [4] [5].

Although P-M model and TV-norm seem very different at the first glance and originate in different mathematical theories, Mrázek *et al*. [6] demonstrated that all these two methods could be cast into a unified framework of Bayesian regularization.

In this paper, we propose a novel smoothness term of Bayesian regularization framework based on M-estimation of robust statistics and a class of fourth-order nonlinear diffusion methods. These diffusion methods attempt to approximate an observed image with a piecewise linear image, which looks more natural than piecewise constant image of P-M model and TV-norm. According to the equivalency of M-estimator and W-estimator, we then propose a filter from equivalent W-estimator for piecewise linear image filtering, denoted as PL bilateral filter, which is more effective than bilateral filter.

2. The smoothness term based on M-estimation and the relevant fourth-order nonlinear diffusion

In Bayesian regularization framework, the data and smoothness terms are combined into a single functional, thus balancing the measured data against the smoothness assumptions. The resulting functional has the form

$$\begin{aligned} E(u) &= E_D(u) + \alpha E_S(u) \\ &= \int_\Omega \left(\psi_D\left(\left|u - f\right|\right) + \alpha\, \psi_S\left(\left|\nabla u\right|\right)\right) d\Omega, \end{aligned} \tag{1}$$

where Ω is image domain, $\psi_D(\cdot)$ and $\psi_S(\cdot)$ are reasonable penalizing functions, $\alpha > 0$ is some regularization parameter, f and u denote the original and filtered image, respectively. In [6] Mrázek *et al*. demonstrated that P-M model [1] and bilateral filter [10] could be derived from the smoothness term of Bayesian regularization framework. In this paper we concentrate on the smoothness term only and propose a class of fourth-order PDEs to overcome staircase effect.

A rather detailed analysis of staircase effect associated with P-M model was carried out in [3]. Let u denote the image intensity function, t the time, and $c(\cdot)$ the diffusion coefficient, the P-M model as formulated in [1] may be presented as

$$\frac{\partial u}{\partial t} = div\left(c\left(|\nabla u|\right)\nabla u\right). \tag{2}$$

Equation (2) was associated with the following smoothness term

$$E_S(u) = \int_\Omega f(|\nabla u|)\, d\Omega, \quad (3)$$

where $f(\cdot) \geq 0$ is an increasing function associated with the diffusion coefficient as

$$c(s) = \frac{f'(s)}{s}. \quad (4)$$

As discussed in [3] the diffusion process of (2) evolves an observed image toward a piecewise constant image and is, therefore, the major reason for the diffusion process to suffer from staircase effect.

You *et al*. [4] further extended the P-M model and proposed the following smoothness term

$$E_S(u) = \int_\Omega f(|\nabla^2 u|)\, d\Omega. \quad (5)$$

The associated fourth-order nonlinear diffusion equation is

$$\frac{\partial u}{\partial t} = -\nabla^2\left[c(|\nabla^2 u|)\nabla^2 u\right]. \quad (6)$$

A related fourth-order diffusion equation and its properties are discussed in [7]. Instead of evolving an image towards a piecewise constant approximation by driving all gradients towards zero or infinity, these methods smooth and sharpen an image towards a piecewise linear approximation by driving all curvatures towards zero or infinity.

Let Ω_i, $i = 1, \ldots n$, be a partition of image domain Ω. We define a piecewise linear image as

$$u(x,y) = \sum_{i=1}^{n} u_i(x,y), \quad (7)$$

where

$$u_i(x,y) = \begin{cases} linear\ \ image & (x,y) \in \Omega_i \\ 0 & otherwise \end{cases}. \quad (8)$$

Note that any two adjacent u_i and u_j must be on different planes; otherwise, we can combine them as one. Let's denote $\partial\Omega_i$, as the boundary of partition Ω_i, then $\Omega_i - \partial\Omega_i$ is the interior of Ω_i. It is obvious that

$$\nabla u_i = \text{constant}\ ,\ \nabla^2 u_i = 0\ . \quad (9)$$

Since it is required that any two adjacent u_i and u_j be on different planes, we have

$$\nabla u_i \neq \nabla u_j \quad (10)$$

for any two adjacent partitions. This indicates that the gradient is not continuous at the boundary $\partial\Omega = \bigcup \partial\Omega_i$.

Consider the image gradient difference, $|\nabla u_i - \nabla u_j|$, between pixel i and its neighboring pixel j. Within one of the piecewise linear image regions, these neighbor gradient differences will be small. Hence an optimal estimator for the "true" value of the image intensity u_i at pixel i minimizes the square of the neighbor gradient difference. In estimating the "true" intensity value at i we want to include only those neighbors that belong to the same region. If neighboring pixel j comes from different region, then gradient difference $|\nabla u_i - \nabla u_j|$ can be viewed as an outlier.

The problem of estimating a piecewise linear image from noisy data can be posed using the tools of robust statistics. We wish to find an image u that satisfies the following smoothness term

$$E_S(u) = \min_u \sum_{i\in u} \sum_{j\in N(i)} \rho\left(\left|\nabla u_i - \nabla u_j\right|, \sigma\right) \quad (11)$$

where $N(i)$ represents the spatial neighborhood of pixel i, $\rho(\cdot)$ is a robust error norm and σ is a scale parameter.

The continuous form of robust M-estimation problem (11) can be posed as

$$E_S(u) = \min_u \int_\Omega \rho(m,\sigma)\, d\Omega, \quad (12)$$

where $m = \sqrt{0.5\cdot\left(u_{xx}^2 + u_{yy}^2\right) + u_{xy}^2}$ is a descriptor from a non-directional measure of curvature magnitude, see [7].

To minimize (11), the gradient at each pixel must be close to those of its neighbors, and an appropriate choice of the $\rho-$function allows us to minimize the effect of the outliers, $|\nabla u_i - \nabla u_j|$, at the edges between piecewise linear image regions.

Compared with some other higher order smoothness terms, see [4] [5], our approach is a general framework of functional minimization, and their models represent just special cases of our work.

The equivalent Euler equation of (12) is

$$\frac{\partial^2}{\partial x^2}\left(\frac{\rho'(m)}{m}u_{xx}\right) + 2\frac{\partial^2}{\partial x\partial y}\left(\frac{\rho'(m)}{m}u_{xy}\right) + \frac{\partial^2}{\partial y^2}\left(\frac{\rho'(m)}{m}u_{yy}\right) = 0 \quad (13)$$

Equation (13) may be solved through the following gradient descent procedure

$$\frac{\partial u}{\partial t} = -\left[\frac{\partial^2}{\partial x^2}\left(\omega(m)u_{xx}\right) + 2\frac{\partial^2}{\partial x\partial y}\left(\omega(m)u_{xy}\right) + \frac{\partial^2}{\partial y^2}\left(\omega(m)u_{yy}\right)\right] \quad (14)$$

where $\omega(m)=\dfrac{\rho'(m)}{m}$. Equation (14) is the nonlinear diffusion process proposed in this paper.

Following the discussion of $\rho-$ function and scale parameter σ in [8], we choose Tukey biweight as the robust error norm and $\sigma=\sqrt{5}\sigma_e$, where

$$\rho(x,\sigma)=\begin{cases}\dfrac{x^2}{\sigma^2}-\dfrac{x^4}{\sigma^4}+\dfrac{x^6}{3\sigma^6}, & |x|\le\sigma\\ \dfrac{1}{3}, & |x|>\sigma\end{cases} \quad (15)$$

$$\omega(x,\sigma)=\begin{cases}\dfrac{1}{2}\left[1-\left(\dfrac{x}{\sigma}\right)^2\right]^2, & |x|\le\sigma\\ 0, & |x|>\sigma\end{cases} \quad (16)$$

and

$$\begin{aligned}\sigma_e &= \mathrm{MAD}(m)\\ &=1.4826\cdot\mathrm{med}\left(\left|m-\mathrm{med}(m)\right|\right).\end{aligned} \quad (17)$$

3. PL bilateral filter based on W-estimation

Equation (11) defines a M-estimator and is an instance of *iteratively reweighted least squares*. This taxonomy is extremely important because it was shown [9] that M-estimators and W-estimators are essentially equivalent and solve the same energy minimization problem.

The equivalent W-estimator of (11) is

$$u_i=\sum_{j\in N(i)}\omega\left(\left|\nabla u_i-\nabla u_j\right|,\sigma\right)u_j, \quad (18)$$

where u_i is the intensity value of pixel i.

Extend equation (18) to a larger spatial support, it becomes

$$u_i=\frac{1}{k(i)}\sum_{j\in\Omega}\omega\left(\left|\nabla u_i-\nabla u_j\right|\right)w\left(\left|i-j\right|\right)u_j \quad (19)$$

where $w(\cdot)$ is a spatial weighting function (typically a Gaussian) and $k(i)$ is a normalization term

$$k(i)=\sum_{j\in\Omega}\omega\left(\left|\nabla u_i-\nabla u_j\right|\right)w\left(\left|i-j\right|\right). \quad (20)$$

It is obvious that the structure of equation (19) is similar to that of bilateral filter, see [10], so we denote equation (19) as piecewise linear bilateral filter. The output of PL bilateral filter for a pixel i is influenced mainly by pixels that are close spatially and have a similar gradient.

It was noted by Tomasi *et al*. [10] that bilateral filter usually requires only one iteration. In the following section we will see that PL bilateral filter is more effective than bilateral filter, when applied to piecewise linear images.

4. Numerical examples

We now demonstrate the performance of the proposed fourth-order PDEs and the PL bilateral filter.

The noisy images shown in Fig.1 are degraded using Gaussian noise at a $\mathrm{SNR}=5\mathrm{dB}$, where

$$SNR=10\lg\left(\frac{std_f^2}{std_n^2}\right), \quad (21)$$

std_f and std_n are standard deviations of original image and noise respectively. It is obvious that the proposed fourth-order PDEs not only remove noise effectively but also avoid the staircase effect.

Fig.2 presents a piecewise linear test image, its noisy version and the relevant reconstructed image of PL bilateral filter. The obtained mean-square-error (MSE) gain[1] of bilateral filter after one iteration is 1.1476, while the MSE gain for the PL bilateral filter is 1.3241. It is evident that PL bilateral filter is better compared with bilateral filter.

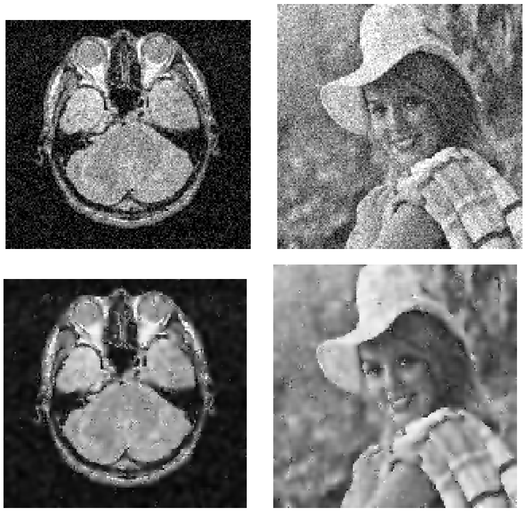
5. Conclusions

In this paper, we have introduced a novel smoothness term of Bayesian regularization framework based on M-estimation of robust statistics and a class of fourth-order nonlinear diffusion methods. These PDEs attempt to approximate an observed image with a piecewise linear image, which looks more natural than piecewise constant image used to approximate an observed image by traditional P-M model and TV-norm.

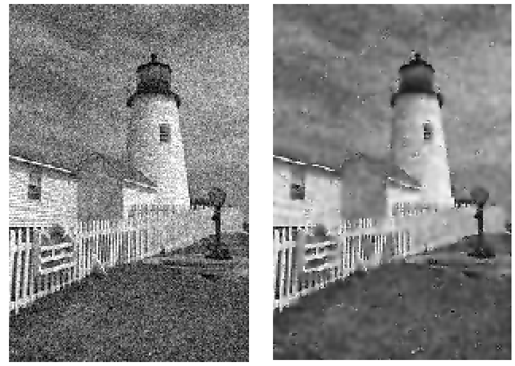
Cause M-estimators and W-estimators are essentially equivalent and solve the same minimization problem, we propose a noniterative and simple filter from equivalent W-estimator, denoted as PL bilateral filter, for piecewise linear image filtering, which is more effective than traditional bilateral filter.

Finally, the effectiveness of methods is emphasized by various experimental results for real and synthetic images.

[1] This gain is defined as the ratio between the MSE before and after the filtering.

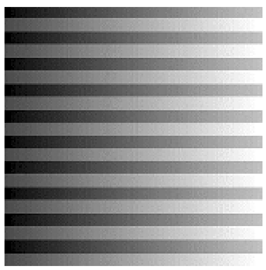


(a) Top-noisy images, bottom-filtered images

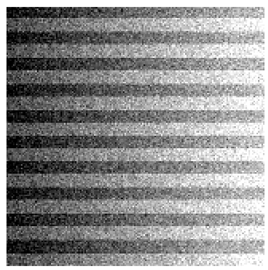


(b) Left-noisy image, right-filtered image

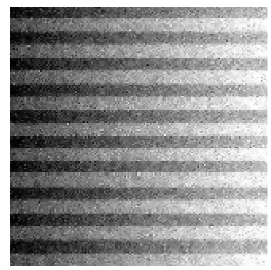
Fig. 1: Results of fourth-order PDEs using the Tukey biweight.



(a) Original image



(b) Noisy image



(c) Filtered image

Fig. 2: Results of PL bilateral filter applied to piecewise linear image.

6. References

[1] P. Perona and J. Malik, "Scale-Space and Edge Detection Using Anisotropic Diffusion," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 12, no. 7, pp. 629-639, 1990.

[2] L. Rudin, S. Osher, E. Fatemi, "Nonlinear Total Variation Based Noise Removal Algorithms," *Physica D: Nonlinear Phenomena*, vol. 60, no. 11, pp. 259-268, 1992.

[3] Y. L. You, W. Xu, A. Tannenbaum, M. Kaveh, "Behavioral Analysis of Anisotropic Diffusion in Image Processing," *IEEE Transactions on Image Processing*, vol. 5, no. 11, pp. 1539-1553, 1996.

[4] Y. L. You and M. Kaveh, "Fourth-Order Partial Differential Equations for Noise Removal," *IEEE Transactions on Image Processing*, vol. 9, no. 10, pp. 1723-1730, 2000.

[5] M. Lysaker, A. Lundervold, X. C. Tai, "Noise Removal Using Fourth-Order Partial Differential Equation with Applications to Medical Magnetic Resonance Images in Space and Time," *IEEE Transactions on Image Processing*, vol. 12, no. 12, pp. 1579-1590, 2003.

[6] P. Mrázek, J. Weickert, A. Bruhn, "On Robust Estimation and Smoothing with Spatial and Tonal Kernels," Universität Bremen, http://www.math.uni-bremen.de/zetem/DFG-Schwerpunkt/SP.preprints.html, 2004.

[7] J. Tumblin and G. Turk, "LCIS: A Boundary Hierarchy for Detail-Preserving Contrast Reduction," *Proc. of SIGGRAPH 1999 Annual Conference on Computer Graphics*, pp. 83-90, Los Angeles, 1999.

[8] M. J. Black, G. Sapiro, D. H. Marimont, D. Heeger, "Robust Anisotropic Diffusion," *IEEE Transactions on Image Processing*, vol. 7, no. 3, pp. 421-432, 1998.

[9] F. R. Hampel, E. M. Ronchetti, P. J. Rousseeuw, W. A. Stahel, "Robust Statistics: The Approach Based on Influence Functions," *Wiley*, New York, pp. 116, 1986.

[10] C. Tomasi, R. Manduchi, "Bilateral Filtering for Gray and Color Images," *Proc. of International Conference on Computer Vision*, pp. 839-846, Washington, 1998.