

Variational Denoising of Partly Textured Images by Spatially Varying Constraints

Guy Gilboa, Nir Sochen, and Yehoshua Y. Zeevi

Abstract—Denoising algorithms based on gradient dependent regularizers, such as nonlinear diffusion processes and total variation denoising, modify images towards piecewise constant functions. Although edge sharpness and location is well preserved, important information, encoded in image features like textures or certain details, is often compromised in the process of denoising. We propose a mechanism that better preserves fine scale features in such denoising processes. A basic pyramidal structure-texture decomposition of images is presented and analyzed. A first level of this pyramid is used to isolate the noise and the relevant texture components in order to compute spatially varying constraints based on local variance measures. A variational formulation with a spatially varying fidelity term controls the extent of denoising over image regions. Our results show visual improvement as well as an increase in the signal-to-noise ratio over scalar fidelity term processes. This type of processing can be used for a variety of tasks in partial differential equation-based image processing and computer vision, and is stable and meaningful from a mathematical viewpoint.

Index Terms—Image denoising, nonlinear diffusion, spatially varying fidelity term, texture processing, variational image processing.

I. INTRODUCTION

PARTIAL differential equation (PDE) methods have been widely used over the past decade for image denoising with edge preservation. These methods are either based on the axiomatic approach of nonlinear scale space (nonlinear diffusions), or on the variational approach of energy functional minimization. Details regarding the interaction and close relations between these approaches can be found, for example, in [1] and [35].

Manuscript received September 28, 2004; revised September 5, 2005. G. Gilboa was supported in part by the National Science Foundation under Grants ITR ACI-0321917 and DMS-0312222 and in part by the National Institutes of Health under Contract P20 MH65166. N. Sochen was supported in part by MUSCLE, Multimedia Understanding through Semantics, Computation and Learning, a European Network of Excellence funded by the EC 6th Framework IST Programme; in part by the Israeli Ministry of Science; in part by the Israel Science Foundation; in part by the Tel-Aviv University fund; and in part by the Adams Center. Y. Zeevi was supported in part by the Ollendorf Minerva Center, in part by the Fund for the Promotion of Research at the Technion, and in part by the HASSIP Research Network program HPRN-CT-2002-00285, funded by the European Commission. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Stanley J. Reeves.

G. Gilboa is with the Department of Mathematics, University of California, Los Angeles, CA 90095 USA (e-mail: gilboa@math.ucla.edu).

N. Sochen is with the Department of Applied Mathematics, University of Tel-Aviv, Tel-Aviv 69978, Israel (e-mail: sochen@math.tau.ac.il).

Y. Y. Zeevi is with the Department of Electrical Engineering, Technion—Israel Institute of Technology, Haifa 32000, Israel (e-mail: zeevi@ee.technion.ac.il).

Digital Object Identifier 10.1109/TIP.2006.875247

A classical variational denoising algorithm is the total variation (TV) minimizing process of Rudin–Osher–Fatemi (ROF) [28]. This algorithm seeks an equilibrium state (minimal energy) of an energy functional comprised of the TV norm of the image I and the fidelity of this image to the noisy input image I_0

$$E_{TV} = \int_{\Omega} \left(|\nabla I| + \frac{1}{2} \lambda (I - I_0)^2 \right) dx dy. \quad (1)$$

This is further generalized by the Φ -formulation [10] with the functional

$$E_{\Phi} = \int_{\Omega} \left(\Phi(|\nabla I|) + \frac{1}{2} \lambda (I - I_0)^2 \right) dx dy. \quad (2)$$

The Euler–Lagrange (E–L) equation is

$$-F \equiv \operatorname{div} \left(\Phi'(\cdot) \frac{\nabla I}{|\nabla I|} \right) + \lambda (I_0 - I) = 0 \quad (3)$$

where $\lambda \in \mathbb{R}$ is a scalar controlling the fidelity of the solution to the input image (inversely proportional to the measure of denoising). Neumann boundary conditions are assumed. The solution is usually found by a steepest descent method

$$I_t = -F, \quad I|_{t=0} = I_0. \quad (4)$$

When the noise is approximated by an additive white process of standard deviation σ , the problem can be formulated as finding

$$\begin{aligned} \min_I \int_{\Omega} \Phi(|\nabla I|) dx dy \\ \text{subject to } \frac{1}{|\Omega|} \int_{\Omega} (I - I_0)^2 dx dy = P \end{aligned} \quad (5)$$

where $P = \sigma^2$ (note that, for noise of an impulsive type, this method is not suitable; see, e.g., [6] and [23]). In this formulation, λ can be considered as a Lagrange multiplier, computed by

$$\lambda = \frac{1}{P|\Omega|} \int_{\Omega} \operatorname{div} \left(\Phi'(\cdot) \frac{\nabla I}{|\nabla I|} \right) (I - I_0) dx dy. \quad (6)$$

The solution is attained by iteratively evolving (4) and updating (6) until convergence. As (5) uses a scalar constraint (and a scalar λ), we refer to it as the *scalar Φ problem*. The actual function with which we work in this paper is $\Phi(z) = \sqrt{1 + \beta^2 z^2}$. The process that results from this function is an approximation of TV which is easy to implement by standard discretization of the E–L equations since it has no singularity at zero gradient.

Also, when the descent is implemented by an explicit method, the time step bound is maximal (bounded by the standard CFL).

We choose it as a representative of variational denoising processes.

The performance of this, and other PDE-based methods, have shown impressive results, especially for nontextured images. The implicit assumption that underlies the formulation of these flows/equations is the approximation of images by piecewise constant functions, which are in the space of bounded variations (BV). We recall that a signal u is in BV if

$$\int_{\Omega} |Du| + \|u\|_{L^1} < \infty$$

where $|Du|$ is the distributional gradient (see definition in [11]). In some sense, they produce an approximation of the input image as the so-called “cartoon model” and, thus, naturally dispose of the oscillatory noise while preserving edges (in some cases even enhancing them, e.g., [26]).

A good cartoon model captures much of the image important information. Yet, it has several obvious drawbacks: textures are excluded, significant small details may be left out, and even large-scale features, that are not characterized by dominant edges, may be disregarded. The purpose of this paper is to show that a relatively simple modification of the above equation yields a denoising algorithm that better preserves both the structural and textural information of the image.

Following Meyer’s work [19], recent studies in the field suggested the use of weaker norms than L^2 , such as G and H^{-1} , for the data fidelity term [2], [3], [25], [33]. These modern techniques can better distinguish between structural and oscillatory components and tend to reduce less contrast of the structural part. However, when used for denoising (as opposed to decomposition), these procedures still recover mainly the structural image components, where textures may be over smoothed. The G -norm (and its various approximations) is low for oscillating patterns and, therefore, does not penalize much both noise and most types of textures. The TV norm, on the other hand, penalizes strongly oscillating patterns. Therefore, this type of energy minimization is still not very well adapted to capture textural parts of the image (see Fig. 6 for a comparison example).

Recent studies which perform decomposition to three categories—structure, texture and noise—by PDEs and wavelets [4], [29] seem to be an appropriate solution also for denoising of natural images. We believe that our approach can complement and improve the results presented in these papers. As the weight parameters of these methods are not spatially varying, a constant “ratio” is implicitly assumed between structures, textures and noise components throughout the image. Whereas noise can often be regarded as spatially invariant (e.g., white Gaussian noise), textures and structures are not homogeneously spread in the image (in terms of variance). Another new approach is the use of Bregman iterations proposed by Osher *et al.* [24]. In this method, as well, a scalar weight parameter λ is used. Our approach may, therefore, be introduced in the future also for these new types of more sophisticated (and complicated) denoising schemes.

In order to keep the presentation of the idea simple and focused, we retain, in this paper, the more classical variational regularization based on the gradient magnitude and L^2 fidelity term.

II. CARTOON PYRAMID MODEL

The cartoon model has been defined and investigated in the early 1980s [7], [12] was further elaborated in [20] and [21] and is widely used as the basic underlying model for many image denoising methods. In the continuous case, the cartoon has a curve Γ of discontinuities, but, everywhere else, it is assumed to have a small or a null gradient $|\nabla I|$. A multilayered image representation was suggested by [18], mainly for compression purposes. However, the wavelet compression (thresholding) technique for extracting the cartoon part, by using a very high compression factor, produces blurry results which do not preserve well edges.

The TV and other gradient-dependent filters are especially good for extracting the cartoon part of the image. We use them, therefore, as a simple pyramid (scale space) of rough image sketches at different scales. Let us define a cartoon of scale s , using the Φ process, as follows:

$$C_s \doteq I_{\Phi}|_{\lambda=\frac{1}{s}} \quad (7)$$

where I_{Φ} is the steady state of (4). See [31] and [30] for a similar definition of the scale. Let us define the residue as the difference between two scales of cartoons

$$R_{n,m} \doteq C_n - C_m \quad (n < m). \quad (8)$$

We shall refer to the noncartoon part of scale s as the residue from level zero

$$NC_s \doteq R_{0,s} = C_0 - C_s. \quad (9)$$

This cartoon and residue data structure is analogous of the pyramid of wavelet approximations. By using the definitions of (7) and (8) and integrating the E–L equation (3), we deduce the following basic properties.

Proposition 1: The cartoon pyramid model has the following scale properties.

- 1) The cartoon of scale 0 is the input image. $C_0 = I_0$.
- 2) The cartoon of scale ∞ is the mean of the input image. $C_{\infty} = \int_{\Omega} I_0(x,y) dx dy$.
- 3) The mean of any residue is zero. $\int_{\Omega} R_{n,m} dx dy = 0$.
- 4) A cartoon image can be built from residues of larger scales. $C_s = \sum_{n=s}^{\infty} R_{n,n+1} + C_{\infty}$.





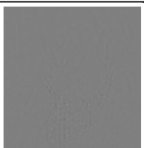
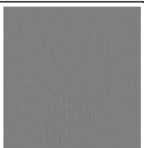
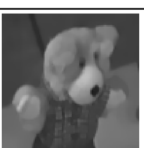








Proof: To prove Property 1, we denote $M = \int_{\Omega} \Phi(|\nabla I_0|) dx dy$. Let us assume $I \neq I_0$ in the L^2 sense for any large λ . Specifically, there exists $h \in L^2$ such that $\|h\|_2 = \epsilon > 0$ and $E_{\Phi}(I) < E_{\Phi}(I_0)$ for $\lambda > (2M)/(\epsilon^2)$, where $I + h = I_0$. Then

$$\begin{aligned} E_{\Phi}(I_0) - E_{\Phi}(I) &= \int_{\Omega} \Phi(|\nabla I_0|) - \Phi(|\nabla I|) dx dy \\ &\quad - \frac{1}{2} \lambda \|h\|_2^2 \\ &< M - \frac{1}{2} \lambda \epsilon^2 \\ &< 0 \end{aligned}$$

which contradicts our assumption.

TABLE I

EXAMPLE OF A CARTOON PYRAMID. LEFT COLUMN: SCALE $s = 1/\lambda$, SECOND COLUMN: CARTOON PART C_s , THIRD COLUMN: NONCARTOON PART NC_s , RIGHT COLUMN: RESIDUE $R_{n,m}$ [WHERE THE VALUES OF (n, m) ARE $(0, 1), (1, 10), (10, 100), (100, \infty)$, FROM SECOND ROW DOWN, RESPECTIVELY]

Scale	Cartoon	Noncartoon	Residue $R_{n,m}$
0			
1			
10			
100			
∞			

Property 2 can be proved using [1, p. 79, Prop. 3] where we deduce that C_∞ converges to the mean image value in the L^1 sense $\lim_{s \rightarrow \infty} \|C_s - \int_\Omega I_0 dx dy\|_{L^1} = 0$.

For Property 3, we use [1, p. 79, Prop. 2] to show that $\int_\Omega C_s dx dy = \int_\Omega I_0 dx dy$, $\forall s \geq 0$, and, therefore

$$\int_\Omega R_{n,m} dx dy = \int_\Omega (C_n - C_m) dx dy = 0.$$

Property 4 is verified by expanding the sum expression

$$\begin{aligned} \sum_{n=s}^{\infty} R_{n,n+1} + C_\infty &= R_{s,s+1} + R_{s+1,s+2} + \cdots \\ &\quad + R_{\infty-1,\infty} + C_\infty \\ &= C_s - C_{s+1} + C_{s+1} - C_{s+2} + \cdots \\ &\quad + C_{\infty-1} - C_\infty + C_\infty \\ &= C_s. \end{aligned}$$

In Table I, a cartoon pyramid example is shown. See [32] for an alternative pyramidal structure suggested recently.

In the case $\Phi(z) = z$ in (2) where we have the ROF model [28], one can relate the scale and residue directly with the G -norm (or star norm) presented in [19, p. 30] (see also the

discrete version defined in [3]). The G space is closely related to the dual of BV . The G norm $\|\cdot\|_G$ penalizes oscillatory and piecewise constant functions in an inverse manner to TV: whereas the TV norm of oscillatory signals is high and their G norm is low, piecewise constant signals have a low TV norm and a high G norm. For example, a signal $\sin(kx)$ in $x \in [0, 2\pi]$ and zero elsewhere has a G -norm approaching zero as $k \rightarrow \infty$, whereas its total-variation in this case approaches ∞ . For more on this subject, we refer the reader to [2]–[4], [19], [22], [25], and [33].

Proposition 2: For the functional $\Phi(z) = z$, we have the following.

- 1) $\|NC_s\|_G = s$.
- 2) $m - n \leq \|R_{n,m}\|_G \leq m + n$.
- 3) $\|R_{n,m}\|_G = 0, \forall n \geq \|C_0 - C_\infty\|_G$.

Proof: The first and third statements are direct consequences of our pyramid definitions and Theorem 3 and Lemma 4 in [19]. The second statement is validated by the relation $R_{n,m} = NC_m - NC_n$, using the triangle inequality and Statement 1. ■

Statement 1 states that the G -norm of the noncartoon part is strictly increasing (linearly), implying that larger and less oscillatory features are incorporated in this part with the growth of s . From Statement 2, one can view $R_{n,m}$ as a texture “band” of the original image with specified upper and lower bounds of the G norm. Due to the convexity of the ROF model, the solutions for similar values of λ are sufficiently close, and, therefore, we estimate that $\|R_{n,m}\|_G$ should be, in fact, closer to its lower bound $m - n$, at least when m and n are of the same order. Statement 3 shows that the telescopic buildup of the cartoon image C_s from larger texture bands, as formulated in Statement 4 of Proposition 1, is finite in practice (excluding the mean image value C_∞). For other relations connecting the scale with the G -norm, see [31].

In order to construct the pyramid, the desired scales should be specified. A simple mechanism, following Gaussian and Laplacian pyramids or wavelet decompositions to detail and approximation parts, is to use predefined scales, which grow exponentially, such as $s_k = s_0 \gamma^k$, where γ is some constant (e.g., $\gamma = 2$ for a dyadic scale). In Table I three levels of a pyramid are shown for $s_0 = 1, \gamma = 10, k = \{0, 1, 2\}$, as well as the zero and infinite scales.

In this pyramid, larger scales retain high frequencies (edges) and one does not resample or decimate the image to a smaller size. This gives more freedom for choosing any set of scale values. Specifically, the multiscale decomposition can be image driven. This topic demands more study and would not be elaborated in this paper. For some preliminary directions suggested by the authors and colleagues on how to select image-driven structure-texture splitting parameters, see [5] and [15].

A. Use for Denoising

The cartoon pyramid has a broad context and may give some more theoretical insight on issues regarding structure, texture, and scale.

For our denoising purposes, in the next section, we use only one decomposition level which should contain the noise and the textures at a similar scale or below that of the noise. In this simpler case, a good representative scale could be selected using

an estimate of the noise variance. We employ the constrained problem, similar to (5) and impose

$$P = \frac{1}{|\Omega|} \int_{\Omega} (I - I_0)^2 dx dy = \alpha \sigma^2 \quad (10)$$

where $\alpha > 0$ controls the selected scale in terms of variance. Typically, $1 \leq \alpha \leq 2$ so that most noise and the relevant textures of that scale are included in the residual part. In our natural images experiment, we set $\alpha = 1.5$.

Our model consists of three components: $I_0 = I_C + I_{NC} + I_n$, where $I_{\text{orig}} = I_C + I_{NC}$ is the original image, I_C is the Cartoon approximation, I_{NC} is the remainder noncartoon part, and I_n is an additive noise. Note that we left the definition of "noncartoon" part vague. It, typically, consists of textures, small-scale details, thin lines, etc. The only assumption we make is that it has zero mean. Under this decomposition, the residue of the filtered image I is

$$I_R \equiv I_0 - I = \tilde{I}_{NC} + \tilde{I}_n. \quad (11)$$

Note that we distinguish between the "true" nonoscillatory part and its approximation by the Φ diffusion process by the tilde superscript.

III. PROBLEM

To obtain an adaptive scheme, we generalize the Φ denoising problem by imposing a spatially varying variance constraint. Let us define first a measure to which we refer as the *local variance*

$$P_z(x, y) \equiv \frac{1}{|\Omega|} \int_{\Omega} (I_z(\tilde{x}, \tilde{y}) - \eta[I_z])^2 w_{x,y}(\tilde{x}, \tilde{y}) d\tilde{x} d\tilde{y} \quad (12)$$

where $w_{x,y}(\tilde{x}, \tilde{y}) = w(|\tilde{x} - x|, |\tilde{y} - y|)$ is a normalized ($\int_{\Omega} w_{x,y}(\tilde{x}, \tilde{y}) d\tilde{x} d\tilde{y} = 1$) and radially symmetric smoothing window, $\eta[\cdot]$ is the expected value taken with respect to the probability density $w_{x,y}(\tilde{x}, \tilde{y})/|\Omega|$ on the set $\Omega \times \Omega$ of all quadruples $(x, y, \tilde{x}, \tilde{y})$. From the definition of the local variance, it follows that $\int_{\Omega} P_z(x, y) dx dy = \mathcal{P}_z$, where

$$\mathcal{P}_z \equiv \text{var}(I_z). \quad (13)$$

We reformulate the scalar Φ problem, stated in (5), in the context of the adaptive Φ problem as follows:

$$\begin{aligned} & \min_I \int_{\Omega} \Phi(|\nabla I|) dx dy \\ & \text{subject to } P_{\hat{R}}(x, y) = S(x, y) \end{aligned} \quad (14)$$

where $I_{\hat{R}} = (I - I_0 - C)$, C is a constant and $S(x, y) \geq 0$ is assumed to be given *a priori*. We solve the optimization problem using Lagrange multipliers

$$E = \int_{\Omega} (\Phi(|\nabla I|) + \frac{1}{2} \lambda(x, y) P_{\hat{R}}(x, y)) dx dy. \quad (15)$$

The E-L equation for the variation with respect to I is

$$\bar{\lambda}(x, y)(I - I_0 - C) - \text{div} \left(\Phi'(\cdot) \frac{\nabla I}{|\nabla I|} \right) = 0 \quad (16)$$

where, for any quantity $X(x, y)$, we define the locally averaged quantity $\bar{X}(x, y) = \int_{\Omega} X(\tilde{x}, \tilde{y}) w_{x,y}(\tilde{x}, \tilde{y}) d\tilde{x} d\tilde{y}$. We solve this equation for I by a gradient descent

$$I_t = \bar{\lambda}(x, y)(I_0 - I + C) + \text{div} \left(\Phi'(\cdot) \frac{\nabla I}{|\nabla I|} \right). \quad (17)$$

In order to compute the value of λ , we multiply the EL equation (16) by $(I - I_0 - C)$ and integrate over the domain Ω . After a change in the order of integrals in the λ term, we get

$$\int_{\Omega} (\lambda(x, y) S(x, y) - Q(x, y)) dx dy = 0 \quad (18)$$

where

$$Q(x, y) = (I - I_0 - C) \text{div} \left(\Phi'(\cdot) \frac{\nabla I}{|\nabla I|} \right).$$

A sufficient condition is

$$\lambda(x, y) = \frac{Q(x, y)}{S(x, y)}. \quad (19)$$

Finally, the constant C is obtained by solving $\partial_C E = 0$, yielding

$$C = \frac{\int_{\Omega} \lambda(x, y) (\bar{I}(x, y) - \bar{I}_0(x, y)) dx dy}{\int_{\Omega} \lambda(x, y) dx dy}. \quad (20)$$

A. Automatic Texture Preserving Denoising

In the general case, we do not have much prior knowledge on the image that can facilitate the denoising process. In our model, we assume that the noise is additive, uncorrelated with the signal (e.g., additive white Gaussian or uniform noise), and that its variance can be estimated.

Our aim is to use the Φ denoising mechanism in a more accurate and precise manner. Images which can be well represented by the large-scale cartoon model are the best candidates for successful denoising. Images with much finer texture and details will not benefit that much from the operation; while reducing most of the noise, this type of processing will inevitably degrade important image features. The first problem is to distinguish between good and bad candidates for Φ denoising. The task becomes even more complex if this is done adaptively. Many natural images exhibit a mosaic of piecewise smooth and texture patches. This type of image structure calls for position (spatial)-varying filtering operation.

The performance of the scalar Φ denoising process is illustrated in Fig. 1, using a typical cartoon-type and a textured image. The SNRs of these three processed images are summarized in Fig. 2, and plotted as a function of the residual variance

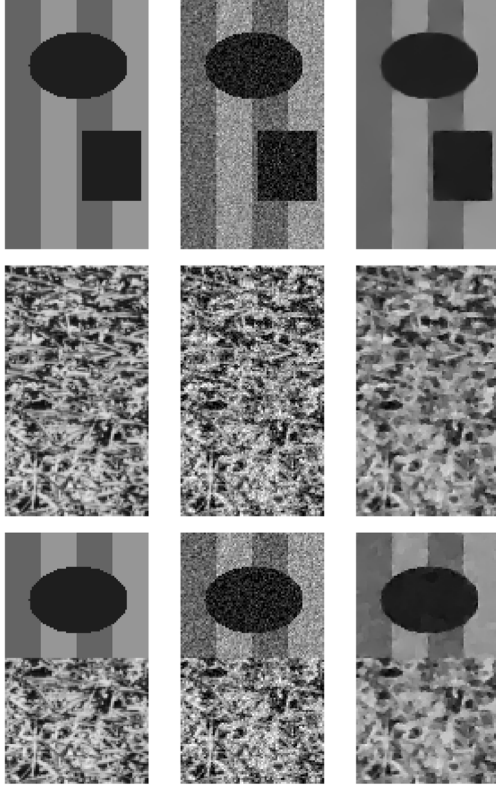


Fig. 1. Scalar Φ denoising of textured and texture-free images. Top row: Piecewise constant image. Middle row: Textured image of grass. Bottom row: Patches of the two types of images combined in one. Left column: Original images. Middle column: Noisy images. Right column: Result of scalar Φ processing (3) at convergence ($P_R = \sigma^2$).

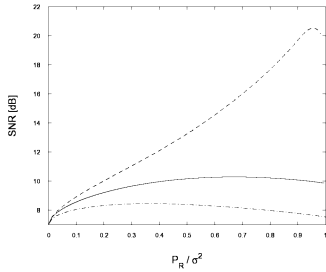


Fig. 2. SNR of scalar Φ denoising of images shown in Fig. 5. SNR is plotted as a function of the residual variance, normalized by the noise variance: P_R / σ^2 . Dashed line piecewise constant image, dash-dot line texture image, solid line combined image.

(normalized variance of the residue). Obviously, as these examples illustrate, cartoon-type images are denoised much better than textured images (both in terms of SNR and visually). Another important observation is that the maximal SNR of cartoon and noncartoon images is reached at different levels of denoising. Whereas cartoon-type images reach their peak SNR at high denoising levels ($P_R \approx \sigma^2$), noncartoon images degrade faster and require less denoising ($P_R < \sigma^2$). For deeper analysis and some bounds on the resulting SNR of Φ process denoising, see [13] and [14].

Here, we present a relatively simple method that can approximate the desired level of denoising in a region. In our above formulation (14), the problem reduces to finding $S(x, y)$.

We use the cartoon pyramid model for this purpose. Our first aim is to differentiate between the cartoon part of the image I_C and the noise and texture parts $I_{NC} + I_n$. Our splitting parameter λ (or scale $s = (1)/(\lambda)$) is selected by imposing (10) and solving (5) using (4) and (6). We assign

$$S(x, y) = \frac{\sigma^4}{P_R(x, y)} \quad (21)$$

where $P_R(x, y)$ is the local variance of the residue I_R .

In the case where $I_R \approx I_n$ (basic cartoon model without textures or fine scale details), this scheme is similar to the scalar Φ process. In this case, α should be close to 1. The local variance of the residue is almost constant ($P_R(x, y) \approx \sigma^2$) and, hence, $S(x, y) \approx \sigma^2$. We get a high-quality denoising process where $I \approx I_C = I_{\text{orig}}$. In the case of most natural images, however, textures will also be filtered and included in the residue part. As the noise is uncorrelated with the signal, we can approximate the total variance of the residue as $P_{NC}(x, y) + P_n(x, y)$, the sum of local variances of the noncartoon part and the noise, respectively. Thus, textured regions are characterized by high local variance of the residue. In order to preserve the detailed structure of such regions, the level of filtering there should be reduced over these regions.

Let us recall the classical Wiener filter (optimal linear filter in the mean squared-error sense). Its formulation in the frequency domain is

$$G(\omega) = \frac{Ps(\omega)}{Ps(\omega) + Pn(\omega)} \quad (22)$$

where $Ps(\omega)$ and $Pn(\omega)$ are the power spectrum of the signal and noise, respectively. The basic concept amounts to a reduction in the extent of filtering ($G \rightarrow 1$) at frequencies where the signal power exceeds that of the noise.

In our case, we have a similar principle, whereby reduction in the extent of filtering (i.e., $S \rightarrow 0$) is called for in regions where signal power exceeds that of the noise. The signal is in this case that portion of the image accounting for the texture and fine details that may be filtered out by the Φ process. Formally, substituting for $P_R(x, y)$ in (21), the relation $P_R(x, y) \approx P_{NC}(x, y) + P_n = P_{NC}(x, y) + \sigma^2$, we get

$$S(x, y) \approx \sigma^2 \frac{1}{1 + P_{NC}(x, y)/\sigma^2}. \quad (23)$$

B. Algorithm

- 1) Separate the noise and relevant textures by minimizing $\int \Phi(|\nabla I|)$ subject to (10) and setting $I_R = I_0 - I$.
- 2) Compute the local variance of I_R by (12) and then compute the local constraints $S(x, y)$ by (21).
- 3) Solve (14) by iteratively evolving (17) and update $\lambda(x, y)$ and C according to (19) and (20).

C. Denoising With Prior Information

In cases where more information regarding the structure of the original signal is available, the performance of denoising process incorporating a spatially varying fidelity constraint can

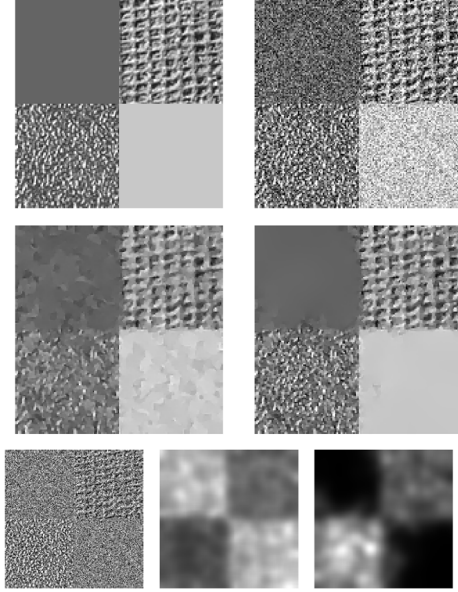


Fig. 3. Processing of a noisy mosaic of textures (fabric and metal) and smooth areas. From top: Original mosaic made of patches of fabric and metal textures, (left) juxtaposed with two constant patches. Noisy version I_0 of the original with (right) $\text{SNR} = 2.4$ dB, $\sigma = 40$. Result of processing with (left) scalar λ – $\text{SNR} = 6.4$ dB, result with adaptive (right) λ – $\text{SNR} = 7.6$ dB. Bottom row: (left) Residue I_R ; (middle) $S(x, y)$ calculated according to the residue; (right) $\lambda(x, y)$ at the convergence of the process.

be substantially ameliorated. The specifics are application-dependent and heuristic in nature. We, therefore, mention here only a few related ideas. To preserve specific features in the denoising process, such as long thin line or known types of textures, one can preprocess with the corresponding feature detector (Hough transform, texture detector). The value of $S(x, y)$ depends, then, locally on the feature detector response. Cases of spatially varying noise also fit the model. For example, in low-quality JPEG images, the boundaries between 8×8 pixel blocks are often more noisy and the fidelity to the original data on these block boundaries should, therefore, be decreased (S increased). See [17] for a different solution by an adaptive window approach.

IV. EXAMPLES

The effects of adaptive- versus scalar-fidelity denoising are illustrated using a synthetic mosaic comprised of two textured patches juxtaposed with two smooth patches (Fig. 3). The scalar fidelity term requires that a global variance, equal to the noise variance, be filtered. As the Φ process is smoothing both texture and noise, more variance is filtered in the textured regions than in the originally smooth ones. This results in over smoothing of textured regions, whereas smooth regions are not sufficiently denoised (Fig. 3, left side second row from top). The adaptive fidelity term process (second row right) applies different levels of denoising in different regions. This improves the result both visually (texture is better preserved, smooth regions are better denoised) and in terms of signal-to-noise ratio. In the third row of Fig. 3, we show how the required spatially varying variance $S(x, y)$ (middle) depends on the value of the residue, I_R (left). The value of the adaptive fidelity term, $\lambda(x, y)$ (right), is shown for the converged process (lighter regions indicate higher



Fig. 4. Example of processing results obtained with a natural image. From top: (Left) Original “Barbara image.” (Right) Noisy version of the original image, I_0 , with $\text{SNR} = 8.7$ dB, $\sigma = 20$. (Left) Result of processing with scalar λ ($\text{SNR} = 12.6$ dB). (Right) Result of processing with adaptive λ ($\text{SNR} = 14.2$ dB). (Left) Residue I_R . (Right) $S(x, y)$ calculated according to residue (middle) $\lambda(x, y)$ at convergence of process.

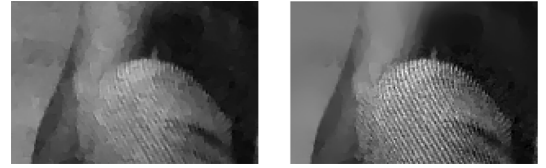


Fig. 5. Enlargement of Barbara's right knee (full images are in Fig. 3). Top: Result of scalar process. Bottom: Result of adaptive process.

value). Naturally, the values of $\lambda(x, y)$ are inversely related to the residual variance measure $S(x, y)$.

Processing a noisy version of the Barbara image (Fig. 4), it is demonstrated how the adaptive Φ method performs well on natural images. Our simple local variance criterion seems to be sufficient to differentiate textured from smooth regions, even in relatively complex images. Accordingly, appropriate local requirements on the variance to be filtered are applied. In Fig. 5, Barbara's right knee is enlarged to highlight similar phenomena to those obtained in the case of the synthetic example, where textures are preserved and the denoising of smooth regions is stronger.

Fig. 6 shows the Teddy Bear from the Toys image. A comparison is performed also to the $\text{TV} - H^{-1}$ model of [25] and to [3] which implements Meyer's $\text{TV} - G$ model [19] (with a small L^2 residual). The scalar fidelity terms are chosen such that the variance of the residual is σ^2 (here $\sigma = 10$). Our algorithm diminishes the denoising in the textural parts of the bear. The $\text{TV} - G$ model is quite competitive, but still degrades the shirt textures. Similar effects can be seen in Fig. 7, where the process is compared with the regularized version of P-M [26]

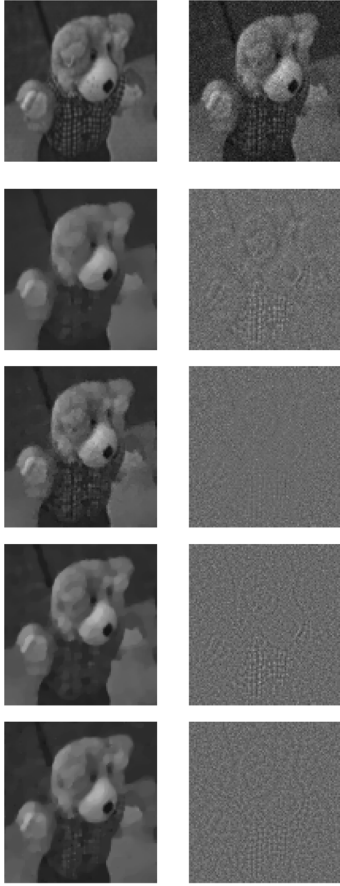


Fig. 6. Part of the Toys image. Top: (left) original; (right) noisy ($\sigma = 10$). Each row depicts (left) the denoised image I and (right) the residual $I_0 - I$ of the following models: $TV - L^2$ (scalar), our proposed adaptive $TV - L^2$, scalar $TV - H^{-1}$ [25], and scalar $TV - G$ [19], [3], from second row to bottom, respectively.

proposed by Catte *et al.* [8]. Relatively small regularization is used for the gradient computation of the diffusion coefficient (variance of Gaussian is 0.1). This causes some isolated points to remain. Stronger regularization in our experiments resulted in extensive over smoothing of the textures. One may observe that the textural snow background is better preserved by our proposed method, while the smooth coat parts are well denoised.

In Table II, we show the comparison between scalar and adaptive processes in terms of SNR. In the scalar process, we show two cases. The “Standard Scalar” column refers to the λ chosen according to the constrained problem (5). The “Optimal Scalar” refers to choosing the parameter λ such that the maximal SNR of the recovered image is reached (out of a finite set of 30 optional values). Naturally, this result can be achieved only in simulations when the original clean image is at hand (see [14] for a way to approximate this parameter). Nevertheless, our algorithm consistently achieves better SNR than the optimal scalar. Note that with respect to the SNR criterion modern multiscale wavelet-based techniques achieve better performance (see, e.g., [27]). This is due to their ability to denoise well also the textural parts. However, in general, wavelet denoising produces less sharp results near edges and may have some oscillations. This affects the denoising quality but is less reflected by the SNR criterion. Convex gradient-based



Fig. 7. Comparison between regularized Perona–Malik (P–M) and our adaptive scheme. Top: (left) original image, (right) contaminated by additive white Gaussian noise ($\sigma = 15$). Bottom: Image denoised using (left) regularized P–M and processing with adaptive λ . Textures and small scale features are kept better in our scheme.

TABLE II
DENOISING RESULTS OF A FEW CLASSICAL IMAGES. FROM LEFT, SNR OF THE NOISY IMAGE (SNR_0), SNRS OF SCALAR λ DENOISING [“OPTIMAL” AND “STANDARD” (SEE EXPLANATIONS FOR THIS TABLE FOR DETAILS)], SNR OF OUR ADAPTIVE λ DENOISING (“OURS—ADAPTIVE”). ALL EXPERIMENTS WERE DONE ON IMAGES DEGRADED BY ADDITIVE WHITE GAUSSIAN NOISE ($\sigma = 10$)

Image	SNR_0	Optimal Scalar	Standard Scalar	Ours – Adaptive
Cameraman	15.86	19.56	19.32	20.81
Lena	13.47	18.19	17.65	18.59
Boats	15.61	20.23	19.83	20.62
Sailboat	10.36	15.51	15.16	16.30
Toys	10.00	17.69	17.29	17.72

variational denoising methods admit the maximum principle and do not produce oscillatory solutions. Our algorithm retains these desired qualities.

Implementation Details

We used explicit Euler schemes to implement the iterative processes. The averaging window $w(x, y)$ was selected to be a Gaussian of standard deviation $\sigma_w = 5$. The potential in all images was $\Phi(s) = \sqrt{1 + s^2}$ ($\beta = 1$). As we used gray level images with values in the range $[0, 255]$ the results are similar to TV denoising. We observed that the calculation of the constant C gives very little improvement. Therefore, we set $C = 0$. In the experiment on natural images (results shown in Table I), we set a constant residue variance $\mathcal{P}_R = 1.5\sigma^2$ ($\alpha = 1.5$ in (10)). Texture patches were taken from the VisTex archive [34]. All images were processed automatically with the same parameters (no tuning of parameters was performed for each image).

V. CONCLUSION

The widely used variational denoising algorithms with global variance constraints perform well on simple cartoon-type

images, where most of the information is represented by the simple structural approximation of the image. However, in order to preserve texture and small scale details, more subtle constraints are called for. We developed an adaptive variational scheme that controls the level of denoising by local variance constraints.

A pyramidal model of structure and texture was presented in which the structural component at any scale could be built in a telescopic manner by texture bands of subsequent higher scales. Some insight on the decomposition was given also with relation to Meyer's G norm [19].

Following this image model, we use the scalar process to separate the noise and the relevant textures of the image which could be degraded in the denoising process. Regions of the residual part with higher local variance than that of the noise are treated as textured regions where denoising should be inhibited (in a soft manner). This is accomplished by introducing a new variational formulation with local variance constraints. *A priori* knowledge on the details to be preserved can further enhance this method.

We have shown that the proposed scheme can filter noise better than the scalar constraint process over a variety of synthetic and natural images. Visually, the processed images look more natural and less "cartoon like." With respect to SNR, our algorithm consistently achieves higher SNR than the optimal that could be achieved with a single scalar value of λ . This study assumed a simple regularizing model based on the gradient magnitude and L^2 fidelity. The ability to effectively reduce noise from textural parts is, therefore, limited. Further improvement may be gained by combining PDE-based and wavelet-based methods in a spatially varying manner, for structures and textures, respectively.

Local variance constraints can be used in almost any variational denoising schemes including ones with more sophisticated fidelity terms that are better adequate for oscillatory patterns [2], [3], [19], [25], [33].

REFERENCES

- [1] G. Aubert and P. Kornprobst, "Mathematical problems in image processing," in *Applied Mathematical Sciences*. New York: Springer-Verlag, 2002, vol. 147.
- [2] J. F. Aujol and G. Gilboa, "Constrained and SNR-based solutions for TV-Hilbert space image denoising," *J. Math. Imag. Vis.*, to be published.
- [3] J. F. Aujol, G. Aubert, L. Blanc-Féraud, and A. Chambolle, "Image decomposition into a bounded variation component and an oscillating component," *J. Math. Imag. Vis.*, vol. 22, no. 1, Jan. 2005.
- [4] J. F. Aujol and A. Chambolle, "Dual norms and image decomposition models," *Int. J. Comput. Vis.*, vol. 63, no. 1, pp. 85–104, Jun. 2005.
- [5] J. F. Aujol, G. Gilboa, T. Chan, and S. Osher, "Structure-texture image decomposition—modeling, algorithms, and parameter selection," *Int. J. Comput. Vis.*, vol. 67, no. 1, pp. 111–136, 2006.
- [6] L. Bar, N. Sochen, and N. Kiryati, "Image deblurring in the presence of salt-and-pepper noise," in *Proc. Scale-Space Conf.*, 2005, pp. 107–118, LNCS 3459.
- [7] A. Blake and A. Zisserman, *Visual Reconstruction*. Cambridge, MA: MIT Press, 1987.
- [8] F. Catte, P. L. Lions, J. M. Morel, and T. Coll, "Image selective smoothing and edge detection by nonlinear diffusion," *SIAM J. Numer. Anal.*, vol. 29, no. 1, pp. 182–193, 1992.
- [9] T. F. Chan and J. Shen, "A good image model eases restoration—On the contribution of Rudin-Osher-Fatemi's BV image model," *IMA Preprints* 1829, Feb. 2002.
- [10] R. Deriche and O. Faugeras, "Les EDP en traitement des images et vision par ordinateur," *Trait. Signal*, vol. 13, no. 6, 1996.
- [11] L. C. Evans and R. F. Gariepy, *Measure Theory and Finite Properties of Functions*. Boca Raton, FL: CRC, 1992.
- [12] S. Geman and D. Geman, "Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. PAMI-6, no. 6, pp. 721–741, Nov. 1984.
- [13] G. Gilboa, N. Sochen, and Y. Y. Zeevi, "Estimation of the optimal variational parameter via SNR analysis," in *Proc. Scale-Space Conf.*, 2005, pp. 230–241, LNCS 3459.
- [14] G. Gilboa, N. Sochen, and Y. Y. Zeevi, "Estimation of optimal PDE-based denoising in the SNR Sense," *IEEE Trans. Image Process.*, vol. 15, no. 8, pp. 2269–2280, Aug. 2006.
- [15] G. Gilboa, N. Sochen, and Y. Y. Zeevi, "Texture preserving variational denoising using an adaptive fidelity term," in *Proc. Variational and Level-Set Methods*, Nice, 2003, pp. 137–144.
- [16] G. Gilboa, N. Sochen, and Y. Y. Zeevi, "PDE-based denoising of complex scenes using a spatially-varying fidelity term," in *Proc. Int. Conf. Image Process.*, Barcelona, Spain, 2003, pp. 865–868.
- [17] C. Kervrann, "An adaptive window approach for image smoothing and structures preserving," in *Proc. ECCV*, 2004, pp. 132–144.
- [18] F. G. Meyer, A. Averbuch, and R. R. Coifman, "Multi-layered image representation: Application to image compression," *IEEE Trans. Image Process.*, vol. 11, no. 9, pp. 1072–1080, Sep. 2002.
- [19] Y. Meyer, "Oscillatory patterns in image processing and nonlinear evolution equations," *AMS*, ser. University Lecture, vol. 22, 2001.
- [20] D. Mumford and J. Shah, "Optimal approximations by piece-wise smooth functions and associated variational problems," *Commun. Pure Appl. Math.*, vol. LII, pp. 577–685D, 1989.
- [21] D. Mumford, "The Bayesian rationale for energy functionals," in *Geometry Driven Diffusion in Computer Vision*, B. M. ter Haar Romeny, Ed. Norwell, MA: Kluwer, 1994.
- [22] A. Obereder, S. Osher, and O. Scherzer, "On the use of dual norms in bounded variation type regularization," Jun. 2004, Univ. California, Los Angeles, CAM Rep. 04-35.
- [23] M. Nikolova, "A variational approach to remove outliers and impulse noise," *JMIV*, vol. 20, no. 1–2, pp. 99–120, 2004.
- [24] S. Osher, M. Burger, D. Goldfarb, J. Xu, and W. Yin, "An iterative regularization method for total variation based image restoration," *SIAM J. Multiscale Model. Simul.*, vol. 4, pp. 460–489, 2005.
- [25] S. Osher, A. Sole, and L. Vese, "Image decomposition and restoration using total variation minimization and the H^{-1} norm," *SIAM Multiscale Model. Simul.*, vol. 1, no. 3, pp. 349–370, 2003.
- [26] P. Perona and J. Malik, "Scale-space and edge detection using anisotropic diffusion," *Pattern Anal. Mach. Intell.*, vol. 12, no. 7, pp. 629–639, 1990.
- [27] J. Portilla, V. Strela, M. Wainwright, and E. P. Simoncelli, "Image denoising using scale mixtures of Gaussians in the wavelet domain," *IEEE Trans. Image Process.*, vol. 12, no. 11, pp. 1338–1351, Nov. 2003.
- [28] L. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Phys. D*, vol. 60, pp. 259–268, 1992.
- [29] J. L. Starck, M. Elad, and D. L. Donoho, "Image decomposition via the combination of sparse representations and a variational approach," *IEEE Trans. Image Process.*, vol. 14, no. 10, pp. 1570–1582, Oct. 2005.
- [30] D. M. Strong and T. F. Chan, "Edge-preserving and scale-dependent properties of total variation regularization," *Inv. Probl.*, vol. 19, no. 6, pp. S165–S187, 2003.
- [31] D. M. Strong, J. F. Aujol, and T. F. Chan, "Scale recognition, regularization parameter selection, and Meyer's G norm in total variation regularization," 2005 Univ. California, Los Angeles, CAM Rep. 05-02 [Online]. Available: <ftp://ftp.math.ucla.edu/pub/camreport/cam05-02.pdf>
- [32] E. Tadmor, S. Nezzar, and L. Vese, "A multiscale image representation using hierarchical (BV, L^2) decompositions," *SIAM Multiscale Model. Simul.*, vol. 2, no. 4, pp. 554–579, 2004, 05-02.
- [33] L. A. Vese and S. J. Osher, "Modeling textures with total variation minimization and oscillating patterns in image processing," *J. Sci. Comput.*, vol. 19, no. 1–3, pp. 553–572, 2003.
- [34] VisTex Vision Texture Archive of the MIT Media Lab [Online]. Available: <http://www-white.media.mit.edu/vismod/imagery/Vision-Texture/vistex.html>
- [35] J. Weickert, "A review of nonlinear diffusion filtering," in *Proc. Scale-Space Theory Computer Vision Conf.*, B. ter Haar Romeny, L. Florack, J. Koenderink, and M. Viergever, Eds., 1997, pp. 3–28, LNCS 1252.
- [36] J. Weickert, "Coherence-enhancing diffusion filtering," *Int. J. Comput. Vis.*, vol. 31, no. 2/3, pp. 111–127, 1999.



Guy Gilboa received the B.Sc. degree in electrical engineering from the Ben-Gurion University, Israel, in 1997, and the Ph.D. degree in electrical engineering from the Technion—Israel Institute of Technology, Haifa, in 2004.

He is currently a Postdoctoral Fellow at the Department of Mathematics, University of California, Los Angeles, hosted by Prof. S. Osher. Previously, he was with the Intel Development Center, Haifa, for three years, where he worked on the design of processors. His main research interests are related to variational and PDE-based processes applied to image enhancement, denoising, and decomposition.



Nir Sochen received the B.Sc. degree in physics and the M.Sc. degree in theoretical physics from the University of Tel-Aviv, Tel-Aviv, Israel, in 1986 and 1988, respectively, and the Ph.D. degree in theoretical physics from the Université de Paris-Sud, Paris, France, in 1992, while conducting his research at the Service de Physique Théorique, Centre d'Etude Nucleaire, Saclay, France.

He continued with one year of research at the Ecole Normale Supérieure, Paris, on the Haute Etude Scientifique fellowship, and a three years National Science

Foundation fellowship with the Physics Department, University of California, Berkeley, where his interest shifted from quantum field theories and integrable models, related to high-energy physics and string theory, to computer vision and image processing. He spent one year with the Physics Department, University of Tel-Aviv, and two years with the Faculty of Electrical Engineering, Technion—Israel Institute of Technology, Haifa. Since 1999, he has been a Senior Lecturer with the Department of Applied Mathematics, University of Tel-Aviv, where he is also a member of the Ollendorff Minerva Center. His main research interests include the applications of differential geometry and statistical physics in image processing and computational vision.

nion—Israel Institute of Technology, Haifa. Since 1999, he has been a Senior Lecturer with the Department of Applied Mathematics, University of Tel-Aviv, where he is also a member of the Ollendorff Minerva Center. His main research interests include the applications of differential geometry and statistical physics in image processing and computational vision.



Yehoshua Y. Zeevi received the Ph.D. degree from the University of California, Berkeley.

He is the Barbara and Norman Seiden Professor of computer science with the Department of Electrical Engineering, Technion—Israel Institute of Technology, Haifa. He is the Founder of the Jacobs Center for Communication and Information Technologies (CCIT) and the Ollendorff Minerva Center, University of Tel-Aviv, Tel-Aviv, Israel, and served as the Head of these centers. He was also the Dean of the Faculty of Electrical Engineering (1994 to

1999). He was a Vinton Hayes Fellow at Harvard University, Cambridge, MA, and has been a regular visitor there. He was also a Visiting Professor at the Massachusetts Institute of Technology, Cambridge, and Rutgers University, Piscataway, NJ, a Senior Visiting Scientist at the NTT Research Center, Yokoska, Japan, and a Visiting Scientist at Lawrence Berkeley Laboratory. He is the coinventor of many patents and the author of over 200 publications related to vision and image sciences. He is the Editor of three books

Dr. Zeevi is a Fellow of the SPIE and the Rodin Academy and the Editor-in-Chief of the *Journal of Visual Communication and Image Representation*. He is one of the founders of i Sight, Inc., a company that developed digital video cameras that mimic the eye, and of UltraGuide, Ltd., a medical technology company that develops guidance systems under ultrasound imaging.