

# Denoising-Enhancing Images on Elastic Manifolds

Vadim Ratner and Yehoshua Y. Zeevi

**Abstract**—The conflicting demands for simultaneous low-pass and high-pass processing, required in image denoising and enhancement, still present an outstanding challenge, although a great deal of progress has been made by means of adaptive diffusion-type algorithms. To further advance such processing methods and algorithms, we introduce a family of second-order (in time) partial differential equations. These equations describe the motion of a thin elastic sheet in a damping environment. They are also derived by a variational approach in the context of image processing. The new operator enables better edge preservation in denoising applications by offering an adaptive lowpass filter, which preserves high-frequency components in the pass-band better than the adaptive diffusion filter, while offering slower error propagation across edges. We explore the action of this powerful operator in the context of image processing and exploit for this purpose the wealth of knowledge accumulated in physics and mathematics about the action and behavior of this operator. The resulting methods are further generalized for color and/or texture image processing, by embedding images in multidimensional manifolds. A specific application of the proposed new approach to superresolution is outlined.

**Index Terms**—Diffusion equations, image enhancement, image resolution, image restoration, wave equations.

## I. INTRODUCTION

**I**N PHYSICS, systems are represented as a collection of discrete particles, interactions between which define the systems' dynamical behavior. These interactions are elegantly represented by partial differential equations (PDEs), which have been thoroughly explored in physics and mathematics during the last centuries. Digital images are also most commonly represented as a collection of discrete elements (pixels). This similarity to a physical system suggests that different image characteristics may also be represented and altered by appropriate PDEs, allowing the use of this powerful mathematical tool in image processing. This explains why since the introduction of the adaptive diffusion PDE to the field of image processing by Perona and Malik (PM) [18], it became the subject of extensive research. In recent years, methods inspired by the PM approach have been successfully applied to various image-pro-

cessing problems (see, for example, [3], [8], and [18]) and have been even implemented in commercial products [4].

The main advantage associated with the implementation of a PDE-based approach is that it allows space(position)-varying image processing, while retaining linear computational complexity. As a matter of fact, anisotropic diffusion-based processing is equivalent to filtering the image with a Gaussian of feature-dependent width. This allows, for example, preserving edges while smoothing flatter areas of an image.

Another advantage of PDE processing manifests itself in color image processing. Color images have been traditionally represented as three separate channels, conveying color and intensity information. Most of the research in the field of color image processing focused on decorrelating the channels, so that processing a channel does not affect information contained in other channels, but full decorrelation has not been achieved. Consequently, nonlinear processing and effects of noise result in mismatch of the 3-channel signals, yielding colored fringe artifacts. To cope with this and for other advantages in image processing, Sochen, Kimmel, and Malladi [22] proposed a new representation scheme, which enables processing of a color image as a single entity, bypassing the issue of correlation between channels. This representation, adopted in the present study, considers an image as a manifold embedded in a combined spatial-feature (color) non-Euclidean multidimensional space [23].

The purpose of the present study is to introduce a new adaptive PDE-based method that can be applied to several problems in image processing, such as denoising and deblurring. This method is based on a family of *damped wave* or *telegraphers'* equations, still unexplored in the context of image processing. These equations describe, for example, the motion of a thin elastic sheet in damping environment. This type of PDE was initially considered for its smoothing properties and for its compatibility with the manifold color image representation (a 3-D manifold being easily described as a thin elastic sheet). We will also show that the *telegraphers'* equation can be derived by variational methods. The most important advantage of using this family of equations is that it enables better edge preservation when compared with diffusion-based approaches.

This paper presents the development of the new method from continuous PDE theory to a complete, ready-to-implement, discrete superresolution algorithm. Numeric implementation issues are addressed and manifold representation is adopted for color image processing. Since the most widely used PDE image-processing scheme is based on the diffusion equation, comparisons are primarily made with diffusion-based schemes.

The paper is organized as follows. Section II derives the telegraph-diffusion (TeD) equation and explores some of its properties. Section III presents several spatially varying versions of

Manuscript received June 02, 2010; revised October 25, 2010; accepted February 11, 2011. Date of publication February 22, 2011; date of current version July 15, 2011. This work was supported in part by Philips Consumer Lifestyle, in part by a fellowship awarded to V.R., and in part by the Technion Olendorff-Minerva Center for Vision and Image Sciences. The first author was supported by an Olendorff-Minerva Fellowship. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Xilin Chen.

The authors are with the Department of Electrical Engineering, Technion-Israel Institute of Technology, Technion City, Haifa 32000, Israel (e-mail: vad@tx.technion.ac.il; zeevi@ee.technion.ac.il).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TIP.2011.2118221

TeD. Section IV introduces the elastic manifold image representation. Section V addresses numerical issues. Section VI derives a single-image super-resolution scheme based on the aforementioned methods, and experimental results (Section VII) are then followed by conclusions.

## II. TELEGRAPH-DIFFUSION EQUATION

### A. Problem Definition

Consider an image degraded by random Gaussian additive blurring noise. Improving the quality of such an image with reference to one image attribute results in a compromise in reference to other image attributes. Most filtering-based denoising schemes yield simultaneous noise reduction and image blurring. The latter occurs because image components, such as edges and textures, span similar high-frequency spectrum to that of noise, which is filtered by the denoising algorithm. For the same reason, sharpening or deblurring of an image also has an undesired side effect of noise enhancement. The main goal of the enhancing algorithms is therefore to further improve the desired processes of deblurring and denoising simultaneously. To achieve this goal, an operator is required, which allows adaptive local control of the restoration process in accordance with image structure. To identify such a process (operator), we turn to physics.

Processes that describe interactions between particles have been explored in physics for centuries. These processes have been successfully modeled by PDEs. Various mathematical tools are available for the analysis and solution of these PDEs in the continuous and discrete domains. Finding the right analogy between image processing and a physical process should, therefore, allow us to apply these tools in the solution of the image-processing problem at hand. One of these approaches, proposed by Perona and Malik in [18], considers an image to be a particle density map and the denoising process—random particle motion controlled by locally varying temperature. In this paper, we explore elastic deformation as an alternative process that exhibits smoothing qualities and allows local control. We show that it has several important advantages over diffusion in the context of image processing.

### B. Elastic Sheet Approach

Let us consider a degraded grayscale input image, represented by the function  $u_0(x, y)$ . A triplet  $(x, y, u_0)$  defines a surface, which may be treated as a thin elastic sheet placed in a damping environment. If we enforce some boundary conditions (e.g., zero spatial derivatives on the boundaries) and allow the elastic deformation process to continue for a given (short) time period  $T$ , we obtain (under some assumptions of elasticity and damping) a function (process)  $u(x, y, T)$ , which evolves to become a smoothed image  $u$ . The elastic deformation process is represented by the PDE

$$u_{tt} - \nabla \cdot (k \nabla u) + cu_t = 0 \quad (2.1)$$

where  $k$  and  $c$  are the elasticity and damping coefficients, respectively. The coefficients  $k$  and  $c$  vary across the image and over time to achieve denoising, while leaving meaningful features intact. We assume the image is rectangular for the sake of

simpler presentation and consider the case where the time  $t$  and the parameters  $k$  and  $c$  are bounded

$$u \equiv u(x, y, t), \quad 0 \leq x \leq X, \quad 0 \leq y \leq Y, \quad 0 \leq t \leq T \\ 0 \leq k \equiv k(x, y, t) < \infty, \quad 0 < c \equiv c(x, y, t) < \infty.$$

The following conditions are imposed:

$$\begin{cases} u(t \leq 0) = u_0 \\ u_t(t = 0) = 0 \\ u_x(x = 0) = u_x(x = X) = 0 \\ u_y(y = 0) = u_y(y = Y) = 0 \end{cases}$$

that is, the image is assumed to be fixed in time before the process begins. Outside the boundaries of the image, the continuation of the boundary values applies.

The parabolic-hyperbolic equation (2.1), known in engineering and in physics as a *damped wave* or *telegrapher's* equation, is often encountered in various fields besides physics (e.g., in statistics [5], [11]; finance [2]; transmission of signals over telegraph wires (hence the terminology); and many more [9], [31]). It has also been thoroughly investigated from a mathematical viewpoint [1], [7], [17], [29].

We coined the term TeD for the operator presented by (2.1) because of its resemblance of the PM version of the diffusion equation (i.e., the second and the third terms)

$$u_t - \nabla \cdot (k \nabla u) = 0 \quad (2.2)$$

where  $k$  is the diffusivity coefficient and since it combines, in a way, the wave and diffusion equations. It is interesting to note that (2.1) converges to the diffusion equation after a very long time [7], [30], although this regime is of no interest in the context of image processing. Given positive and bounded coefficients, TeD converges to a unique bounded solution [17]. It thereby provides a basic form of denoising by uniformly smoothing the entire image. This quality of the elastic process is further developed in the sequel to enable edge preservation and enhancement.

### C. Process Optimization

Equation (2.1) can also be reached by means of the variational approach. Variational methods have been widely used in image processing ever since their introduction in the context of image processing by Mumford and Shah ([15]) as a mathematical method for image segmentation. In particular, total variation (TV) was used by Rudin, Osher, and Fatemi to derive the diffusion-based methods in [21], producing the required image by minimizing the surface of the resulting (grayscale) image  $u$ , over the image domain

$$\min(E(u(x, y))) = \min \iint_{\Omega} \left( \sqrt{u_x^2 + u_y^2} \right) dx dy. \quad (2.3)$$

Euler-Lagrange (EL) minimization of the measure  $E$  (energy) yields the following condition:

$$\begin{aligned} \frac{\delta E}{\delta u} = 0 &= \frac{\partial}{\partial x} \left( \frac{u_x}{\sqrt{u_x^2 + u_y^2}} \right) + \frac{\partial}{\partial x} \left( \frac{u_x}{\sqrt{u_x^2 + u_y^2}} \right) \\ &= \nabla \cdot \left( \frac{1}{\sqrt{u_x^2 + u_y^2}} \nabla u \right) \end{aligned} \quad (2.4)$$

which, when solved by gradient descent, yields the nonlinear diffusion equation.

The two minimization steps of 1) Euler-Lagrange optimization and 2) a subsequent iterative solution (by, for example, gradient descent) lend themselves to an evolutionary process that satisfies the EL condition. It is advantageous to consider the evolution as the final goal and incorporate the time axis into the minimization domain. This approach yields a process suitable for the chosen functional in a single minimization step. Let us consider the smoothing process over a finite time interval  $(0, T)$ . Since minimal surface is still desired, we require that the total surface over time will be as small as possible

$$\min (E(u(x, y, t))) = \min \int_0^T \iint_{\Omega} (\sqrt{u_x^2 + u_y^2}) dx dy dt. \quad (2.5)$$

Since the initial image  $u(x, y, t = 0)$  is fixed, this should lead to smoothing over time.

Another requirement is that the total change over the entire image should not exceed the noise level

$$\int_0^T \left( \iint_{\Omega} |u_t| dx dy \right) dt = \sigma \quad (2.6)$$

where  $\sigma$  is the noise variance. Since  $|u_t|$  is not differentiable, we use a more adequate condition

$$\int_0^T \left( \iint_{\Omega} |u_t^2| dx dy \right) dt = \tilde{\sigma}^2 \quad (2.7)$$

where  $\tilde{\sigma}$  is some positive constant.

By using Lagrange multipliers, we can derive the following minimization problem:

$$\left( \int_0^T \left( \iint_{\Omega} (\sqrt{u_x^2 + u_y^2}) - \lambda_1 u_t^2 \right) dt \right) \quad (2.8)$$

which yields the following EL equation:

$$2\lambda_1 u_{tt} - \frac{\partial}{\partial x} \left( \frac{2u_x}{\sqrt{u_x^2 + u_y^2}} \right) - \frac{\partial}{\partial y} \left( \frac{2u_y}{\sqrt{u_x^2 + u_y^2}} \right) = 0. \quad (2.9)$$

To guarantee convergence, an energy dissipation term may be added. This term can be derived from the condition

$$\int_0^T \left( \iint_{\Omega} \varphi(t) |u_t|^2 dx dy \right) dt = \text{const} \quad (2.10)$$

where  $\varphi(t)$  is a monotonic increasing function of time  $t$ , forcing the rate of change to decrease over time. The choice of  $\varphi(t) \equiv t$  transforms (2.9) to

$$2(\lambda_1 + \lambda_2 t) u_{tt} + 2\lambda_2 u_t - \frac{\partial}{\partial x} \left( \frac{2u_x}{\sqrt{u_x^2 + u_y^2}} \right) - \frac{\partial}{\partial y} \left( \frac{2u_y}{\sqrt{u_x^2 + u_y^2}} \right) = 0 \quad (2.11)$$

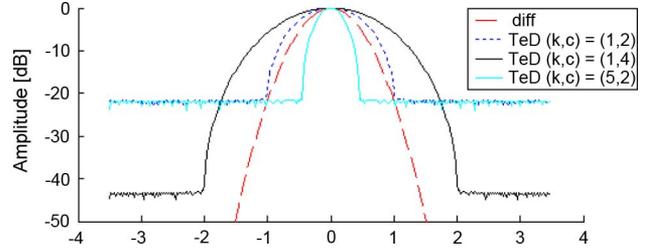


Fig. 1. Spatial frequency response of a linear 1-D diffusion (2.13) depicted for  $k = 1$  (dashed red) and telegraph-diffusion operators, depicted for 3 sets of parameters (see insert).

where  $\lambda_1, \lambda_2$  are constants. Experiments show that in discrete settings, using  $2(\lambda_1 + \alpha t)$  as the second derivative coefficient instead of  $2(\lambda_1 + \lambda_2 t)$ , for small positive or zero  $\alpha$  (i.e., decoupling the coefficients of the first and second derivatives (in time) of  $u$ ) improves the results.

We have thus shown that the TeD equation can be derived by optimizing the entire adaptive process, so that the resulting image is smooth and close to the original. The resulting equation fits the elastic model (2.1), with reduced correlation (force) between pixels that belong to different objects (i.e., lying across an edge).

#### D. Operator Analysis

To gain further insight into the behavior of the TeD equation, we first consider the linear telegrapher's equation (i.e., constant elasticity and damping coefficients). A 1-D elastic deformation can be represented as convolution with kernel  $h_{td}$ , the spatial frequency representation of which is (F denoting the Fourier transform)

$$\begin{aligned} H_{td}(\omega, t) &= F\{h_{td}\}(\omega, t) \\ &= \left( \frac{c + \sqrt{c^2 - 4k\omega^2}}{2\sqrt{c^2 - 4k\omega^2}} \right) \exp\left( \frac{-c + \sqrt{c^2 - 4k\omega^2}}{2} t \right) \\ &\quad + \left( \frac{-c + \sqrt{c^2 - 4k\omega^2}}{2\sqrt{c^2 - 4k\omega^2}} \right) \exp\left( \frac{-c - \sqrt{c^2 - 4k\omega^2}}{2} t \right) \end{aligned} \quad (2.12)$$

where all of the coefficients are non-negative, unless stated otherwise. The kernel can be derived by Fourier transforming of (2.1) in the time domain and then solving the resulting polynomial. The 1-D kernel (Fig. 1) is presented for convenience. Higher dimensional kernels are obtained by simply substituting the squared spatial frequency  $\omega^2$  with  $\omega_1^2 + \dots + \omega_n^2$ .

The spatial frequency response (2.12) resembles, at first, an all-pass filter. With time, the side lobes (stop-band ripples) of  $H_{td}$  decrease in amplitude, while the main lobe (pass-band) decreases in width, rendering  $H_{td}$  to become a low-pass. As  $k$  increases, the main lobe width decreases and the amplitude of the side lobes remains unchanged. For negative  $k$ , the filter enhances high frequencies in a way similar to the linear backward diffusion filter. Increasing  $c$  reduces the side-lobe amplitude and increases the main lobe width. For negative  $c$ , the side lobes become larger than the main lobe, resulting again in a high-frequency-enhancing filter. As the kernel formula indicates, nega-

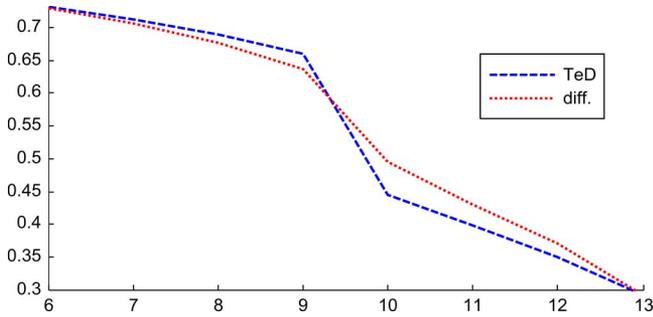


Fig. 2. Step smoothed by the adaptive TeD (dashed) and adaptive diffusion (dotted). The plots are line samples of a 2-D denoised image. The TeD edge transitioning from high to low intensity is much sharper (i.e., higher slope) and of higher contrast than that of the diffusion-processed edge.

TABLE I  
COMPARISON BETWEEN TeD AND DIFFUSION,  
A FIXED NUMBER OF ITERATIONS

| Init. PSNR | Sigma | Diffusion | var  | TeD   | var  |
|------------|-------|-----------|------|-------|------|
| 43.02      | 0.02  | 54.38     | 0.07 | 56.17 | 0.12 |
| 28.0       | 0.04  | 49.99     | 0.12 | 51.53 | 0.22 |
| 22.03      | 0.08  | 43.65     | 0.19 | 45.1  | 0.28 |
| 17.27      | 0.15  | 34.1      | 0.23 | 34.6  | 0.29 |
| 13.72      | 0.25  | 25.01     | 0.02 | 25.11 | 0.02 |
| 9.42       | 0.5   | 16.18     | 0.01 | 16.21 | 0.01 |

Columns (left to right): initial (noisy) PSNR, noise standard deviation (sigma), diffusion PSNR, diffusion PSNR variance, TeD PSNR, and its variance.

tive  $c$  is equivalent to negative time. “Slowing time,” however, is equivalent to a reduction of damping and elasticity.

The TeD equation thus presents a way to locally adjust the frequency components of an image by offering a wide range of filters, from low-pass, which is useful in denoising (if the noise is most noticeable in high frequencies), to high-frequency-enhancing filters, which are useful for the sharpening of edges.

### E. Why TeD?

The shape of the linear operator highlights one of the advantages inherent in the TeD, compared with diffusion-based methods. Linear diffusion is analogous to convolution with a Gaussian lowpass filter  $h_d$  (derived similar to TeD), represented in the spatial frequency domain by

$$H_d(\lambda, t) = F\{h_d\}(\lambda, t) = \exp(-k\lambda^2 t). \quad (2.13)$$

It is clearly observed that the TeD kernel (Fig. 1) is much closer to an ideal low-pass (a rectangle in the frequency domain) than the diffusion kernel. The pass-band in the case of TeD is flatter and the transition between the pass-band and stop-band (the “roll-off”) is steeper, with characteristics that are desirable in filter design. These qualities of the TeD kernel contribute to the preservation of high frequencies in edges, yielding sharper edges and higher edge contrast than those resulting from diffusion denoising (Fig. 2 and Table I).

Another important difference between the characteristics of the two equations is error propagation speed. The essence of the

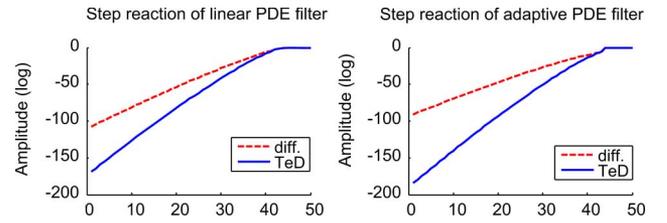


Fig. 3. Log of step response of linear TeD and diffusion (left) and adaptive TeD and diffusion (right). The  $50 \times 50$  input image consisted of a single step at  $X = 44$ . The plots depict line samples of the original 2-D functions degraded to similar values of PSNR (the  $x$ -axis being the spatial coordinate  $X$ ). These plots show that information “leaks” across the edge much faster in linear diffusion than in linear TeD (higher values of the response for  $X < 44$ , where, ideally,  $-\infty$  is expected). The difference is even more profound in the case of adaptive algorithms.

latter is the speed of a propagation of a reaction to a disturbance. A weakness of the diffusion equation is that it has unbounded propagation speed (i.e., theoretically) in any nonzero time interval, and nonzero reaction to an impulse at the origin propagates to the entire image (although not equally). For an intuitive insight into the importance of this quality, we turn again to the linear operator. The reaction of the diffusion operator to an impulse, for any time  $t > 0$  is a Gaussian function of the distance from the focal point of the disturbance, which is nonzero everywhere. The case of variable coefficients is even worse, since the diffusivity is artificially increased over flat areas ([18]), increasing the propagation effect. This quality of the diffusion equation is highly undesirable in image-processing applications, since it imposes a correlation between naturally uncorrelated features, such as different objects in an image.

This problem was dealt with a long time ago in the area of theoretical physics, where infinite information propagation speed is also undesirable. An improved model that accounts for inertia results in the telegrapher’s equation (see, for example, [2] and the references therein, or [14]). In the latter case, the speed of propagation is  $\sqrt{k}$  (Fig. 3).

Diffusion and TeD also differ in the discrete domain. The explicit time discretization scheme (approximating continuous process by dividing it into small linear time steps) requires small time steps to remain stable, which increases computational effort. The numeric stability requirement for 1-D linear discrete diffusion equation is  $\tau < h^2/2k < 1/2$ , while that of linear discrete TeD is  $\tau < h/\sqrt{k} < 1$  (CFL condition [30] with  $0 \leq k \leq 1$  being the diffusivity in diffusion and elasticity in TeD). In other words, TeD allows a larger time step.

Computational complexity of the algorithm is similar to that of diffusion (i.e.,  $O(nN)$ ,  $N$  being the number of input elements and  $n$  is the number of iterations of the PDE solver).

A discrete model of elastic surface was proposed independently of this work by Pollak *et al.* in [19], where the authors represented pixels as a grid of masses, connected by springs. Unlike the model proposed here, in their paper, Pollak *et al.* did not use a damped model, but rather reset the masses’ speed after short intervals of time  $dt$ . The PDE resulting from their model is therefore a first-order diffusion equation.



Fig. 4. Comparison of the PM-TeD and FAB-TeD performance in denoising of an image. Leftmost column: original image (top) and noisy (PSNR 28.0) image (bottom). Middle column: PM-TeD denoising result and error (PSNR  $-30.0$ ). Rightmost column: FAB-TeD denoising result and error (PSNR  $-30.7$ ). The FAB-denoising error image contains many less details than the PM denoising error.

### III. SPATIALLY VARYING TeD

#### A. Basics

By defining spatially (and temporally) varying elasticity and damping coefficients, it is possible to locally control the degree of smoothing. In this paper, we primarily address the case of the varying elasticity coefficient, although initial experiments have shown that variation of the damping coefficient can further improve performance.

To introduce our method, we consider an image as an elastic sheet (3-D manifold). The initial structure of the sheet represents the noisy input image  $u_0$ . To improve SNR and preserve edges, the sheet is made to be more elastic (higher  $k$ ) over smooth areas and more rigid (lower  $k$ ) around edges. An example of this elasticity coefficient is

$$k(s) = \frac{1}{1 + \frac{|s|^2}{\kappa}} \quad 0 < \varepsilon^2 \equiv \text{const} \ll 1, \quad 0 \leq \kappa \equiv \text{const} \quad (3.1)$$

similar to coefficients proposed for the diffusion operator (see, for example, [8]), where  $s$  is an edge detector (absolute gradient of the processed image throughout this paper). An image that results from the application of the basic TeD operator (Fig. 4) consists mostly of flat surfaces separated by sharp edges; a cartoon-like representation of the original image. This type of representation is useful, for example, in image segmentation.

Elasticity functions and their properties are presented in [20]. In this paper, we focus our analysis and application on *forward-and-backward* (FAB) coefficients, introduced by Gilboa *et al.* in [8] and on directional coefficients (Weickert *et al.* [25], [26]).

#### B. Forward-and-Backward (FAB) Processes

Edge preservation or even enhancement by adaptive TeD may be further improved by allowing local sharpening processes. These can be achieved by reversing the smoothing process around edges by means of negative coefficients. When selective sharpening is required, a combined FAB-TeD is desirable. It can be implemented by varying either the elasticity or the damping coefficients. The latter will be addressed elsewhere. An interesting option to explore is the combination of smooth positive elasticity and discrete positive/negative damping (positive over smooth areas, negative around edges).

Unlike the case of the damping coefficient, negative elasticity is not equivalent to negative time. Nevertheless, it still sharpens the image, increasing slope and contrast of edges, and even creating Mach-bands around the edges (overshoots in step-like edges, further increasing edge contrast). Fig. 4 presents the results of applying FAB-TeD with elasticity defined by

$$k = \frac{1}{1 + \left(\frac{|\nabla I|}{k_f}\right)^n} - \frac{a}{1 + \left(\frac{|\nabla I| - k_b}{\omega}\right)^{2m}} \quad (3.2)$$

where  $k_f$  controls the amplitude of the maximum,  $\omega$  is the width of the negative lobe,  $k_b$  is the position of the minimum, and  $a$  is the amplitude of the minimum. Good choices (empirically found [8]) for  $(m,n)$  are  $(4,2)$ . It is important to note that negative coefficients introduce instability that should degrade the image after long periods of time. The instability, however, is insignificant in short-time evolutions, which are of interest in an image denoising enhancement.

#### C. Coherence Enhancing Filtering

Nonlinear PDE-based methods allow space-varying processing, but they are still isotropic (i.e., they display no directional dependency when viewed on a small enough scale). For every pixel, the strength of smoothing changes according to its proximity to an edge, but the smoothing is symmetric in all orientations. When dealing with oriented structures, not only does the strength of smoothing need to be controlled by image features, but also its orientation. A method proposed by Weickert [25]–[28] achieves that by manipulating eigenvalues of a smoothed structure tensor (3.5), enhancing diffusion along the edges and reducing, or even reversing it across edges (i.e., perpendicular to the edge orientation) [25]. This is obtained by replacing the spatial term  $\nabla \cdot (k\nabla I)$  in diffusion and TeD equations by

$$\nabla \cdot (D\nabla I). \quad (3.3)$$

In the case of a grayscale image,  $D$  is a  $2 \times 2$  matrix, which contains structural information about the image. It is defined as follows:

$$D = (\omega_1; \omega_2) \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \begin{pmatrix} \omega_1 \\ \omega_2 \end{pmatrix} = \begin{pmatrix} a & b \\ b & c \end{pmatrix} \quad (3.4)$$

where  $\omega_i$  are eigenvectors of the smoothed structure tensor  $J_\rho$ , pointing in the direction of the steepest descent and its orthogonal counterpart

$$\begin{aligned} J_\rho (\nabla I_\partial \nabla I_\partial^T) &= \begin{pmatrix} j_{11} & j_{12} \\ j_{12} & j_{22} \end{pmatrix} = K_\rho * (\nabla I_\partial \nabla I_\partial^T) \\ I_\partial &= K_\partial * I(\cdot, t) \\ K_\chi(x) &= \frac{1}{2\pi\chi^2} \exp\left(-\frac{|x|^2}{2\partial^2}\right). \end{aligned} \quad (3.5)$$

The eigenvalues  $\lambda_i$  are as follows:

$$\begin{aligned} \lambda_1 &= \alpha \\ \lambda_2 &= \alpha + (1 - \alpha) \exp\left(\frac{-1}{(j_{11} - j_{22})^2 + 4j_{12}^2}\right) \end{aligned} \quad (3.6)$$

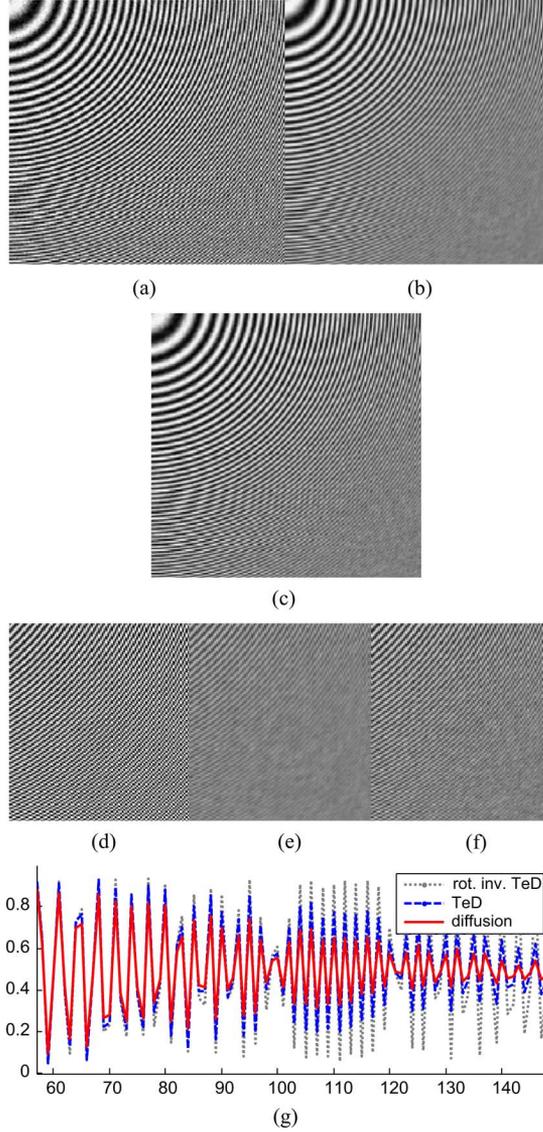


Fig. 5. Tensor processing. (a) Original image. (b) Tensor-diffusion denoising. (c) Tensor TeD denoising. (d)–(f) zoomed lower-right corners of (a)–(c) (g) gray-level values obtained along the crosscut of the main diagonal of (a) (solid gray), (b) (solid red), and (c) (dashed blue). The original image was denoised by the two methods to achieve similar PSNR of the upper left quadrant (lower frequencies). While both methods reduced the quality of the lower right quadrant (higher frequencies), TeD has preserved a much better structure detail than the diffusion-based algorithm (compare (e) and (f)). This is, in particular, noticeable in (g), where one can easily see that the diffusion-related signal deteriorates faster than the TeD-processed signal.

where  $\alpha$  is a small parameter controlling diffusion in the direction of  $\nabla I(\omega_1)$  (i.e., across edges). The eigenvalue  $\lambda_2$  controls smoothing in the direction of strongest coherence, i.e.,  $\omega_2$ . It is

also possible to use negative  $\alpha$  [8], applying inverse TeD across edges. The negative  $\alpha$  is useful in the enhancement of edges.

This scheme yields the following  $3 \times 3$  convolution mask for the computation of (3.3) [26], where  $a, b, c$ , and  $d$  are the components of the matrix  $D$ , as shown in the equation at the bottom of the page.

Comparing the performance of the tensor-diffusion and the tensor-TeD (Fig. 5), one observes that the latter yields better results. One reason is the better edge preservation. Orientation sensitivity of the TeD denoising scheme also contributes to its performance.

#### IV. COLOR MANIFOLD FRAMEWORK

The TeD is now generalized to the consideration of color (or multispectral) image processing, to provide a unified framework for the spatially varying methods presented before. One way to adopt the TeD method and use it on color images is to process each color channel separately. This method, although simple and easy to implement, disregards correlation between color channels. The result is, therefore, far from optimal and is affected by the color representation scheme that one implements. For example, in RGB, color-space color artifacts are created and in YCrCb, color-space overall saturation is decreased [20]. Further, due to nonlinear processing of the three channels, the three images do not cohere into one unified image and color artifacts appear around the edges.

An alternative approach is to embed the three color channels of the image in a manifold. Instead of separate channels, the image is then represented as a 2-D surface  $(\sigma^1, \sigma^2, C^1(\sigma^1, \sigma^2), C^2(\sigma^1, \sigma^2), C^3(\sigma^1, \sigma^2))$  embedded in a 5-D spatial-color manifold [16], [23]. This allows simultaneous processing of the three channels in a unified manner and lends itself to the application of a variety of tools offered by manifold representation and by the geometric approach. In this case, the variables  $(\sigma^1, \sigma^2)$  represent spatial information and  $(C^1, C^2, C^3)$  represent color and intensity. To further highlight this approach, several important relevant issues are adopted from Sochen *et al.* [22], [23]. A more detailed theoretical background is available in [16].

##### A. Metric and Distance

Let us first define the terminology. A general definition of distance element over manifold  $M$  is

$$ds^2 = \sum_{\mu, \nu=1}^{\dim(M)} g_{\mu\nu} dx^\mu dx^\nu, \quad (4.1)$$

$$\begin{pmatrix} \frac{b_{i-1,j} - b_{i,j+1}}{4} & \frac{c_{i,j+1} + c_{i,j}}{2} & \frac{b_{i+1,j} + b_{i,j+1}}{4} \\ \frac{a_{i-1,j} + a_{i,j}}{2} & -\frac{a_{i-1,j} + 2a_{i,j} + a_{i+1,j}}{2} & \frac{a_{i+1,j} + a_{i,j}}{2} \\ \frac{b_{i-1,j} + b_{i,j-1}}{4} & \frac{c_{i,j-1} + c_{i,j}}{2} & \frac{b_{i+1,j} - b_{i,j-1}}{4} \end{pmatrix}$$

where  $x^i$  is the  $i$ th element of a vector  $x \in M$  and  $g_{\mu\nu}$  is a metric defined in  $M$ . In the example of  $M = \mathbb{R}^2$  with Euclidean distance,  $g_{\mu\nu}$  is a  $2 \times 2$  unity matrix, resulting in the familiar Euclidean distance  $ds^2 = (dx^1)^2 + (dx^2)^2$ . The metric  $g_{\mu\nu}$  defines, in this case, an inner product on  $M$

$$(v, w)_{(M,g)} = v^T g w, \quad v, w \in M \quad (4.2)$$

where  $g$  is a matrix with elements  $g_{\mu\nu}$ , and  $(\cdot, \cdot)_{(M,g)}$  is the inner product of vectors in  $M$  with metric  $g$ .

Let  $u$  be an embedding of a submanifold  $\Sigma$  in a manifold  $M$

$$u: \Sigma \rightarrow M, \text{ or } u: (\sigma^1, \dots, \sigma^n) \rightarrow (u^1, \dots, u^m), \\ n = \dim(\Sigma), \quad m = \dim(M)$$

(we will mostly use  $\Sigma \equiv \mathbb{R}^2$  for the image coordinates and  $M \equiv \mathbb{R}^5$  for the space-feature manifold). The distance element on the image is calculated as before, with the following metric on  $\Sigma$ :

$$g_{\Sigma\mu\nu} = \sum_{\alpha, \beta=1}^{\dim(M)} g_{M\alpha\beta} \frac{\partial u^\alpha}{\partial \sigma^\mu} \frac{\partial u^\beta}{\partial \sigma^\nu}, \quad \alpha, \beta \in [1.. \dim(\Sigma)] \quad (4.3)$$

where  $g_{M\alpha\beta}$  is the metric on  $M$ , and  $g_{\Sigma\mu\nu}$  is called the *induced metric*.

### B. Manifold Image Representation

The embedding manifold  $M \equiv \mathbb{R}^5$  with Euclidean metric was used in [22] and [23]. Choosing  $u$  as the spatial and color components,  $u(\sigma^1, \sigma^2) = (\sigma^1, \sigma^2, R(\sigma^1, \sigma^2), G(\sigma^1, \sigma^2), B(\sigma^1, \sigma^2))$ , yields good results. The aforementioned choices determine the following structures of the induced metric (note that from now on  $\Sigma$  is omitted from  $g_{\Sigma\mu\nu}$ , since it is the only metric that is of interest to us):

$$g_{\mu\nu} = \sum_{\alpha=1}^5 \frac{\partial u^\alpha}{\partial \sigma^\mu} \frac{\partial u^\alpha}{\partial \sigma^\nu}. \quad (4.4)$$

Having defined the distance element, the following surface minimization functional is defined:

$$S = \int d\sigma^1 d\sigma^2 \sqrt{\det(g)} \sum_{\mu, \nu=1}^2 (g^{-1})_{\mu\nu} \sum_{\alpha=1}^5 \frac{\partial u^\alpha}{\partial \sigma^\mu} \frac{\partial u^\alpha}{\partial \sigma^\nu} \\ = \int d\sigma^1 d\sigma^2 \sqrt{\det(g)} \left( (g^{-1})_{11} (1 + R_x^2 + G_x^2 + B_x^2) \right. \\ \left. + (g^{-1})_{22} (1 + R_y^2 + G_y^2 + B_y^2) \right. \\ \left. + 2(g^{-1})_{12} (R_x R_y + G_x G_y + B_x B_y) \right) \quad (4.5)$$

where, for convenience, we have replaced  $(\sigma^1, \sigma^2)$  by  $(x, y)$ .

The minimization of the aforementioned functional (also known as *Polyakov action*) with respect to the image components  $(R, G, B)$  yields the Beltrami flow [22]

$$C_t = \frac{1}{\sqrt{\det(g)}} \sum_{\mu, \nu=1}^2 \partial_\mu \left( \sqrt{\det(g)} (g^{-1})_{\mu\nu} \partial_\nu C \right) \quad (4.6)$$

where  $C$  is one of the color channels  $R, G$ , or  $B$ . For gray-scale images, the process degenerates to

$$C_t = \frac{1}{\sqrt{\det(g)}} \nabla \left( \sqrt{\det(g)} (g^{-1}) \nabla C \right) \quad (4.7)$$

which is very similar to the tensor diffusion of Weickert [25], except that the eigenvalues of the Weickerts structure tensor (smoothed induced metric) are replaced by general functions, while here they are replaced by their inverses. In fact, using inverse eigenvalues may cause instability in numerical solutions. This may be overcome by merging the two methods by using the induced metric  $g_{\mu\nu}$  as a structure tensor for color images.

By exploiting the similarities between TeD and diffusion equations, one can derive the following TeD-based smoothing operator (by adding a second derivative in the time of  $C$  to 4.7, similar to the way TeD 2.1 can be seen to be derived from diffusion 2.2)

$$C_{tt} = \frac{1}{\sqrt{\det(g)}} \nabla \left( \sqrt{\det(g)} (g^{-1}) \nabla C \right) - c C_t, \quad (4.8)$$

where “det” marks the matrix determinant. The experimental comparison between manifold TeD image processing and separate channel processing can be found in [20].

## V. NUMERIC SCHEMES

Discretization schemes and their convergence so far have received limited attention in the context of PDE image processing. Weickert *et al.* in [27], presented one of these schemes, which differs from the basic approach in that it offers unconditional stability. Its application to the TeD method will be presented.

To simplify matters, we first discuss a 1-D case, where  $u^{i,j}$  is the  $i$ th element (spatial coordinate  $x = i * h$ ) of vector  $u^j$  (input at time  $t = j * \tau$ ). We also assume constant damping  $c$ . Basic discrete representation of (2.3) is as follows ( $\tau$  and  $h$  being the temporal and spatial steps):

$$\frac{u^{x,j+1} - 2u^{x,j} + u^{x,j-1}}{\tau^2} + c \frac{u^{x,j+1} - u^{x,j}}{\tau} \\ - \frac{k^{x,j} [u^{x+1,j} - u^{x,j}] - k^{x-1,j} [u^{x,j} - u^{x-1,j}]}{h^2} = 0. \quad (5.1)$$

This representation defines the following iterative, explicit (Forward Euler), update scheme of  $u$ :

$$(1 + c\tau) \underline{u}^{j+1} = (2 + c\tau) \underline{u}^j - \underline{u}^{j-1} + \tau^2 A^j \underline{u}^j \quad (5.2)$$

where  $\underline{u}^j$  is the input vector and  $A^j$  is a matrix corresponding to the spatial derivative operators and coefficients  $k$  at time  $j * \tau$ .

An advantage of the forward Euler scheme is that it is straightforward and can well approximate the continuous process with any given accuracy by using small enough time steps. This is also the major drawback of the explicit scheme—it requires small time steps to converge to a stable solution and is therefore very demanding computationally. To compare this, the most stable solution is the implicit solution. It uses data from time

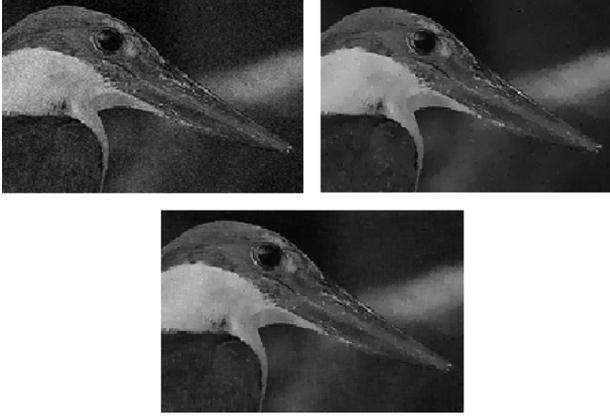


Fig. 6. Comparison of explicit (upper right) and two iterations of semi-implicit (bottom) denoising of a noisy image (upper left).

step  $j + 1$  to perform the spatial calculations (coefficients and derivatives) in the right-hand side of the equation

$$(1 + c\tau) \underline{u}^{j+1} = (2 + c\tau) \underline{u}^j - \underline{u}^{j-1} + \tau^2 A^{j+1} \underline{u}^{j+1}. \quad (5.3)$$

It is numerically stable for any time step, allowing a single iteration solution. However, it requires the solution of  $N$  nonlinear equations, where  $N$  is the number of elements in  $u$ , owing to the fact that  $A^{j+1}$  depends on  $\underline{u}^{j+1}$ . An intermediate scheme closely related to backward Euler (BE) discretization was proposed in the context of image processing by Weickert *et al.* in [27], resulting in the following discrete equation:

$$(1 + c\tau) \underline{u}^{j+1} = (2 + c\tau) \underline{u}^j - \underline{u}^{j-1} + \tau^2 A^j \underline{u}^{j+1} \quad (5.4)$$

which can be rewritten, assuming pixel dimension  $h = 1$  and unity matrix  $I$ , as

$$\underline{u}^{j+1} = ((1 + \tau c)I - \tau^2 A^j)^{-1} [(\tau c + 2) \underline{u}^j - \underline{u}^{j-1}]. \quad (5.5)$$

In this case,  $A^j$  is a tridiagonal matrix, which enables efficient (linear complexity) inversion of  $[(1 + \tau c)I - \tau^2 A^j]$  (which is a tridiagonal matrix) by using the Thomas algorithm [27]. Although it requires additional calculations, the scheme retains linear complexity and is more stable than the standard approach, allowing larger time step and fewer iterations (reduction by order of magnitude). Although the tradeoff between quality and processing time (number of iterations, time-step size) still exists (i.e., for larger time steps), there is a noticeable reduction in image quality (Fig. 6), even a single iteration produces visually acceptable results.

The semi-implicit scheme works well with vector signals. The generalization of the algorithm to higher dimensions is, however, not trivial. The naïve approach of vectorizing (raster scan, Hilbert scan) the input image and applying the BE scheme results in nonzero values outside the main three diagonals of  $A$ , increasing the matrix inversion complexity to  $O(n^2)$

$$\left( (1 + \tau c)I - \tau^2 \sum_{d=1}^m A_d \right) \underline{u}^{j+1} = (\tau c + 2) \underline{u}^j - \underline{u}^{j-1} \quad (5.6)$$

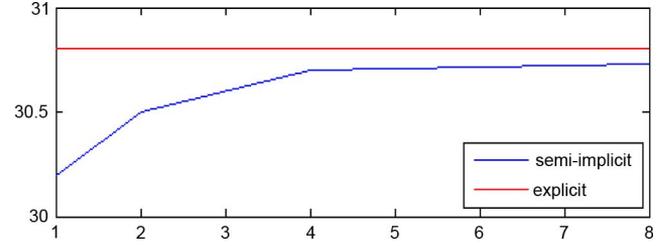


Fig. 7. PSNR ( $y$  axis) obtained by semi-implicit TeD as a function of the number of iterations ( $x$  axis), based on a number of natural images. These data demonstrate that for larger time-steps (fewer iterations), the PSNR of the denoised image decreases. This implies that larger time steps yield less consistent semi-implicit schemes.

where  $A_d$  corresponds to the partial spatial derivative operator along dimension  $d$ . The matrix  $((1 + \tau c)I - \tau^2 \sum_{d=1}^m A_d)$  is obviously not tridiagonal, a structure which does not allow using efficient inversion methods. Weickert *et al.* [27] proposed a solution that approximates the  $m$ -dimensional semi-implicit scheme by performing the update separately along each axis. The 2-D input image is raster scanned twice per iteration—along the  $x$  and  $y$  axes, and the two resulting vectors are updated

$$\underline{u}^{j+1} = \frac{1}{m} \sum_{d=1}^m \left[ ((1 + \tau c)I - \tau^2 A_d)^{-1} \left( (\tau c + 2) \underline{u}_d^j - \underline{u}_d^{j-1} \right) \right] \quad (5.7)$$

where the vector  $\underline{u}_d$  is generated by raster scanning the image  $u$  along dimension  $d$ . This approximation of the BE scheme is unconditionally stable ([27]) (i.e., small changes in the input produce small changes in the output). It does, however, become inconsistent for very large timesteps (Fig. 7) (i.e., when large time steps are used, no matter how fine is the grid (pixel size), the continuous solution does not fulfill the discrete scheme). Our experiments have shown that the semiimplicit scheme achieves results similar to explicit for as few as two iterations. For a single iteration, however, artifacts of inconsistency begin to appear, decreasing the PSNR (Fig. 7).

## VI. SUPERRESOLUTION

Let us consider an image  $u$  that is obtained by blurring/sub-sampling of an original image  $u_{\text{orig}}$ . The essence of the super-resolution problem consists of restoring  $u_{\text{orig}}$  based on  $u$  and/or enhancing certain image attributes (e.g., edges). Most existing superresolution algorithms require, as an input, multiple images  $u_i$  obtained from different sources or a video sequence. Some require prior learning from a sequence of (similar) images [6]. There are, however, cases where only a single image is available, which one would still wish to enhance. Since blurring/sub-sampling results in loss of high-frequency information, the super-resolution problem poses an impossible task. A related task can, however, be dealt with efficiently (i.e., the problem of increasing the resolution of  $u$  in such a way that the processing will yield a natural-looking enhanced image ([8]). This is closely related to recently proposed methods of smart interpolation [13], [32]. Although no image-specific information is available that will replace the lost high-frequency components, we can make several assumptions based on the current understanding of human vision, on how a general natural image should look like. For



Fig. 8. Example of a single image superresolution result. Note the general impression of enhancement as well as the fine detail that emerges in various regions of the enlarged image by the applications of the FAB TeD processing scheme. Careful inspection of the image reveals that this subjective impression of enhancement is primarily due to sharpening of edges and thin lines. Below are samples of the 2-D Fourier transforms of the subsampled image (green) and the result (blue). The resulting image has a wider bandwidth than the subsampled image, indicating increased sharpness .

example, it should have sharp edges, which separate relatively smooth regions. These assumptions may be considered, in a way, as being equivalent to prior information (Fig. 8).

Based upon the aforementioned assumptions, we propose a single image superresolution, or image fidelity enhancement scheme. It consists of two steps:

- Step 1) increasing an image to a desired size;
  - Step 2) improving the quality of the resulting image.
- The first step requires simple interpolation (bicubic or spline since both achieve good results). It results in a larger number of pixels, but a blurry image. The second step incorporates two tasks: 1) selective sharpening of meaningful edges and thin lines

and 2) reduction of noise-like artifacts caused by the interpolation (e.g., ringing effects). We have already observed that these conflicting demands are accommodated well by the FAB TeD operator. The proposed scheme is as follows:

$$u \longrightarrow \boxed{\text{interpolation}} \xrightarrow{\text{(I)}} \boxed{\text{FAB TeD}} \xrightarrow{\text{(II)}} \tilde{u}_{\text{orig}}$$

Step 1) consists of possible preprocessing (such as TeD denoising) and interpolation. Step 2) requires inclusion of additional information that may be available (e.g., edge information, segmentation, areas of increased sharpness, or texture). This information can be incorporated in the calculation of coefficients.

Note that we do not pay special attention to noise in the initial, low-resolution image for two reasons: 1) TeD-based denoising can be applied before Step 1) and 2) low-resolution noise is further reduced along with the ringing effects by processing the interpolated image.

## VII. EXPERIMENTS

Since the performance of diffusion-based methods has been well established in the literature ([3], [8], [25]) and since the diffusion and TeD equations are very closely related, we first compare the performance of basic nonlinear diffusion and TeD schemes, implementing the elasticity coefficients according to (3.1). We then test the performance of the semi-implicit numeric scheme in the context of TeD.

There are several difficulties in comparing even such similar denoising methods as diffusion and TeD, since one can obtain less stable and more accurate schemes even for the same coefficient functions. Further, similar coefficients affect the two equations differently. We therefore compare the performances as follows. First, we compare the best results that can be achieved by each of the two methods given similar resources (Table I); we use similar coefficient schemes, with parameters adjusted so that the best peak signal-to-noise ratio (PSNR) is achieved in a fixed number of iterations (the same number for TeD and diffusion), with the largest time-step size that yields a stable solution for each algorithm (e.g., for the initial noise PSNR of  $\sim 20$  dB, the diffusion parameters used were  $\kappa = 50$ ,  $dt = 0.25$ , and for TeD  $\kappa = 25$ ,  $c = 1.5$ ,  $dt = 0.5$ ). Second, we compare the number of iterations required to reach a certain PSNR by the two algorithms (Table II). We find the parameters  $\kappa$  (3.1) and  $c$  (2.1) that result in the minimal number of iterations required to reach some fixed PSNR values, thus comparing the resources required by each method. We can compare the performance of TeD and diffusion by comparing the number of iterations that each algorithm requires since the amount of calculations per pixel during an iteration of the algorithms is practically the same (the difference is one multiplicative and two additive operations per pixel, which is negligible compared to the total number of operations per pixel, which is around 20 additive and 20 multiplicative operations).

The results (Table I) indicate that for lower initial noise levels (PSNR above 20 dB), TeD outperforms diffusion by more than 2 dB. This is a significant difference, considering the fact that

TABLE II  
COMPARISON OF TED AND DIFFUSION FOR A SET FINAL VALUE OF PSNR

| Init. PSNR | Sigma | PSNR  | diffusion iterations | TeD iterations |
|------------|-------|-------|----------------------|----------------|
| 43.02      | 0.02  | 60.74 | 281                  | 133            |
| 28.0       | 0.04  | 48.54 | 459                  | 206            |
| 22.03      | 0.08  | 35.06 | 1783                 | 679            |
| 17.27      | 0.15  | 26.69 | 653                  | 217            |
| 13.72      | 0.25  | 20.88 | 400                  | 108            |
| 9.42       | 0.5   | 14.05 | 590                  | 158            |

Columns (left to right): initial (noisy) PSNR, noise standard deviation (sigma), final PSNR, number of diffusion iterations and number of TeD iterations required to reach the PSNR.

the PSNR is measured across the entire image, while the advantage of TeD is noticeable mainly around the edges (which also convey most of the image information). For higher noise levels, the advantage of TeD is less significant. This may be attributed to the fact that high noise alters the high-frequency components of edges to such an extent that they become irrecoverable by the proposed methods. Some realizations of the random noise make some details irrecoverable. While diffusion removes these details altogether, TeD recovers them when possible. This may explain the increased variance of the TeD results for lower noise levels. In general, we have seen that the worst results of TeD on synthetic and natural images are still the same or better than those of diffusion. The results also indicate (Table II) that TeD reaches a set PSNR value faster (in fewer iterations).

To test the performance of semi-implicit TeD, we restore a noisy (PSNR 26.4) kingfisher image. To obtain a reference result, we perform 50 iterations of explicit TeD denoising for a time step of 0.1 and different values of  $c$  and  $k$ . The best set of parameters (determined experimentally) is  $c = 4$ ,  $k = 100$  (for gray levels between 0 and 255). These parameters yield a PSNR of 30.8 (Fig. 6). We then proceed to denoise the image by semi-implicit methods by using  $c = 4$ ,  $k$  between 1 and 100, for different time steps and max iterations (Figs. 6 and 7). It is observed that there is a slight decrease in PSNR as the time step increases. As expected, semiimplicit results approach those achieved by the explicit scheme for smaller time steps.

Additional experimental results and images are available at <http://visl.technion.ac.il/TED-page/>.

## VIII. CONCLUSION

The new family of PDE image-processing operators, which differs from existing ones by the incorporated second-order time derivative, enables better preservation of edges and small details, by offering an adaptive bandpass filter with better characteristics than those of existing PDE-based methods. The new operators were derived by the intuitive approach (smoothing properties of damped elastic motion) and by variational methods. Most of the research on existing, diffusion-based operators applies also to the new family of TeD equations. Considering the variational basis of the TeD equation, other applications in various areas of image processing and computer vision, where diffusion equation is currently used, such as segmentation, become attractive.

Convergence of the continuous TeD was proven by Nakao [17]. Initial analysis, supported by computational results, indicates that the proposed discrete image-processing scheme is stable. However, rigorous stability and consistency analysis still remain to be performed on the new family of equations. The proposed enhancement scheme, including superresolution, can be generalized to deal with (medical and other) volumetric data as well as video. These results will be presented elsewhere, along with the application of variable damping coefficients.

## ACKNOWLEDGMENT

The authors would like to thank Prof. A. Nepomnyashchy, Prof. Y. Pinchover, and Prof. G. Wolansky from the Mathematics Department and Prof. M. Rubin from the Mechanical Engineering Department for fruitful discussions which contributed greatly to this paper. The authors would also like to thank Prof. M. Nakao for indicating the mathematical basis of the presented algorithms and they offer special thanks to Dr. B. Ratner, whose continuing contribution to this paper has been invaluable, both in discussions and in proposing physics-based ideas. The authors thank tsuntsun3 and citrit1 from flickr.com for kindly allowing us to use their photographs in this paper and, last but not least, A. Faktor and T. Faktor for their assistance with computer simulations.

## REFERENCES

- [1] M. Aassila, "Global existence and energy decay for a damped quasi-linear wave equation," *Math. Meth. Appl. Sci.*, vol. 21, pp. 1185–1194, 1998.
- [2] Ahmed and Abdusalam, "On modified Black-Scholes equation," *Chaos, Solitons Fractals*, vol. 22, no. 3, pp. 583–587, 2004.
- [3] L. Alvarez, P. L. Lions, and J. M. Morel, "Image selective smoothing and edge detection by nonlinear diffusion," *SIAM J. Num. Anal.*, vol. 29, no. 3, pp. 845–866, Jun. 1992.
- [4] Analyze is a Registered Trademark of Mayo Medical Ventures. Rochester, MN.
- [5] S. R. Dunbar, "A branching random evolution and a nonlinear hyperbolic equation," *SIAM J. Appl. Math.*, vol. 48, no. 6, pp. 1510–1526, Dec. 1988.
- [6] W. T. Freeman, T. R. Jones, and E. C. Pasztor, "Example-based super-resolution," *IEEE Comput. Graphics Appl.*, vol. 22, no. 2, pp. 56–65, Mar./Apr. 2002.
- [7] T. Gallay and G. Raugel, "Scaling variables and asymptotic expansions in damped wave equations," *J. Diff. Equations*, vol. 150, no. 1, pp. 42–97, 1998.
- [8] G. Gilboa, N. Sochen, and Y. Y. Zeevi, "Forward-and-backward diffusion processes for adaptive image enhancement and denoising," *IEEE Trans. Image Process.*, vol. 11, no. 7, pp. 689–703, Jul. 2002.
- [9] S. Godoy and L. S. Garcia-Colin, "Nonvalidity of the telegrapher's diffusion equation in two and three dimensions for crystalline solids," *Phys. Rev. E.*, vol. 55, no. 3, Mar. 1997.
- [10] R. Kimmel, R. Malladi, and N. Sochen, "Images as embedded maps and minimal surfaces: Movies, color, texture and volumetric medical images," *Int. J. Comput. Vis.* 39, vol. 2, pp. 111–129, 2000.
- [11] A. D. Kolesnik and E. Orsingher, "A planar random motion with an infinite number of directions controlled by the damped wave equation," *J. Appl. Prob.*, vol. 42, pp. 1168–1182, 2005.
- [12] P. Maragos and F. Meyer, "Nonlinear PDEs and numerical algorithms for modeling levelings and reconstruction filters," in *Proc. Scale-Space, Lecture Notes Comput. Sci.*, 1999, pp. 363–374.
- [13] K. Mishiba, T. Suzuki, and M. Ikehara, "Edge-adaptive image interpolation using constrained least squares," in *Proc. ICIP*, Sep. 2010, pp. 2837–2840.
- [14] P. M. Morse and H. Feshbach, *Methods of Theoretical Physics*. New York: McGraw-Hill, 1953, vol. 1.

- [15] D. Mumford and J. Shah, "Optimal approximations by piecewise smooth functions and associated variational problems," *Commun. Pure Appl. Math.*, vol. 42, pp. 577–685, 1989.
- [16] M. Nakahara, *Geometry, Topology and Physics*, 2nd ed. Bristol, U.K.: IOP, 2003.
- [17] M. Nakao, "Decay and global existence for nonlinear wave equations with localized dissipations in general exterior domains," *Operator Theory, Adv. App.*, vol. 159, pp. 213–299, 2007.
- [18] P. Perona and J. Malik, "Scale-space and edge detection using anisotropic diffusion," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 12, no. 7, pp. 629–639, Jul. 1990.
- [19] I. Pollak, A. S. Willsky, and H. Krim, "Image segmentation and edge enhancement with stabilized inverse diffusion equations," *IEEE Trans. Image Process.*, vol. 9, no. 2, pp. 256–266, Feb. 2000.
- [20] V. Ratner and Y. Y. Zeevi, "Image enhancement using elastic manifolds," *Proc. ICIP*, pp. 769–774, Sep. 2007.
- [21] L. I. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Phys. D*, vol. 60, pp. 259–268, 1992.
- [22] N. Sochen, R. Kimmel, and R. Malladi, "General framework in low level vision," *IEEE Trans. Image Process.*, vol. 7, no. 3, pp. 310–318, Mar. 1998.
- [23] N. Sochen and Y. Y. Zeevi, "Images as manifolds embedded in spatial-feature non-Euclidean space," Israel Inst. Technol, Haifa, Israel, CCIT Rep., 1998.
- [24] I. Vitsnudel, Y. Y. Zeevi, and R. Ginosar, "Neural network aided design for image processing," *Proc. SPIE*, vol. 1606, pp. 1086–1091, 1991.
- [25] J. Weickert, "Coherence-enhancing diffusion filtering," *Int. J. Comput. Vis.*, vol. 31, pp. 111–127, 1999.
- [26] J. Weickert, "Nonlinear diffusion filtering," in *Handbook of Computer Vision and Applications*, B. Jahne, H. Haussecker, and P. Geissler, Eds. London, U.K.: Academic Press, 1999, vol. 2.
- [27] J. Weickert, B. M. t. H. Romeny, and M. A. Viergever, "Efficient and reliable schemes for nonlinear diffusion filtering," *IEEE Trans. Image Process.*, vol. 7, no. 3, pp. 398–410, Mar. 1998.
- [28] J. Weickert and H. Schar, "A scheme for coherence-enhancing diffusion filtering with optimized rotation invariance," *J. Vis. Commun. Image Represent.*, vol. 13, pp. 103–118, 2002.
- [29] Z. Yang and G. Chen, "Global existence of solutions for quasi-linear wave equations with viscous damping," *J. Math. Anal. App.*, pp. 604–618, 2003.
- [30] E. Zauderer, *Partial Differential Equations of Applied Mathematics*, 2nd ed. New York: Wiley, 1998.
- [31] X. Zhang and Z. Q. Zhang, "Wave transport through thin slabs of random media with internal reflection: Ballistic to diffusive transition," *Phys. Rev. E*, vol. 66, no. 016612, 2002.
- [32] W. Zhe, K. Ma, and C. Cai, "Edge-contrast-guided image interpolation using directional variation field diffusion," in *Proc. ICIP*, Sep. 2010, pp. 2005–2008.



**Vadim Ratner** received the B.Sc. degree in computer sciences from the Technion-Israel Institute of Technology, Haifa, Israel, in 2003, where he is currently pursuing the Ph.D. degree in electrical engineering from the Technion-Israel Institute of Technology.

Previously, he worked for five years in software development. His main research interests are related to variational methods and PDE-based processes applied to image enhancement and denoising.



**Yehoshua Y. Zeevi** received the Ph.D. degree from the University of California, Berkeley, where he was subsequently a Visiting Scientist at the Lawrence-Berkeley Laboratory.

He was a Vinton Hayes Fellow at Harvard University, Cambridge, MA, where he was a Visiting Professor for many years. He was also a Visiting Professor at the Massachusetts Institute of Technology, Cambridge; Rutgers, Piscataway, NJ; and Columbia University, New York; as well as a Distinguished Visiting Scientist at the NTT Research Center, Yokosuka, Japan.

He was the Barbara and Norman Seiden Professor of Computer Sciences in the Department of Electrical Engineering, Technion-Israel Institute of Technology, Technion City, Haifa, Israel, where he served as the Dean from 1994 to 1999. He was also a member of the Board of Governors of the Technion. Currently, he is the Head of the Ollendorff Center for Vision and Image Sciences. He is the Co-Inventor of many patents, including the high-dynamic-range camera and the single-lens 3-D camera and the author of many publications related to vision and image sciences. He is the Editor-in-Chief of the *Journal of Visual Communication and Image Representation*. He was one of the founders of i Sight, Inc., a company that developed digital video cameras that mimics the eye and of UltraGuide Ltd. and CByonda medical technology companies. He is a Co-Founder and Chairman of the board of Cortica Ltd., a company devoted to search by content.