Inverse Problems and Imaging
Volume 3, No. 4, 2009, 693–710
doi:10.3934/ipi.2009.3.693

MULTISCALE IMAGE REPRESENTATION USING NOVEL INTEGRO-DIFFERENTIAL EQUATIONS

Eitan Tadmor

Department of Mathematics, Institute for Physical Science & Technology
and
Center of Scientific Computation and Mathematical Modeling (CSCAMM)
University of Maryland, College Park, MD 20742 USA

Prashant Athavale

Applied Mathematics and Scientific Computation (AMSC)
and
Center of Scientific Computation and Mathematical Modeling (CSCAMM)
University of Maryland, College Park, MD 20742, USA

(Communicated by Luminita Vese)

Abstract. Motivated by the hierarchical multiscale image representation of Tadmor et. al., [25], we propose a novel integro-differential equation (IDE) for a multiscale image representation. To this end, one *integrates* in inverse scale space a succession of refined, recursive 'slices' of the image, which are balanced by a typical curvature term at the finer scale. Although the original motivation came from a variational approach, the resulting IDE can be extended using standard techniques from PDE-based image processing. We use filtering, edge preserving and tangential smoothing to yield a family of modified IDE models with applications to image denoising and image deblurring problems. The IDE models depend on a user scaling function which is shown to dictate the BV^* properties of the residual error. Numerical experiments demonstrate application of the IDE approach to denoising and deblurring.

1. **Introduction.** A black and white image can be realized as a graph of a discrete function $f : \Omega \subset \mathbb{R}^2 \to \mathbb{R}$. The values of this function, $f(x)$, denote the intensity of the image at the discrete points $x \in \Omega$: the function f attains its maximum value at the brightest spots in the image and minimum value of zero at the darkest spots. The graph of an image consists of discrete pixels which for mathematical analysis, is postulated as an $L^2(\Omega)$ function.

Many problems in image processing fall under two broad categories of *image segmentation* and *image restoration*. In *image segmentation* one is interested in identifying constituent parts of a given image, whereas *image restoration* aims to denoise and deblur an observed image in order to recover its underlying "clean" image. Additive noise, denoted by η, is inadvertently added to images due to various reasons, such as limitations of the image capturing facilities or transmission losses. Besides noise, images could also be blurred due to unfocused camera lens, relative

2000 *Mathematics Subject Classification.* 26B30,65C20,68U10.

Key words and phrases. natural images, multiscale representation, total variation, denoising, deblurring, inverse scale, variational problem, integro-differential equation, energy decomposition.

This research was supported in part by NSF grant 07-07949. E.T. thanks the Centre for Advanced Study at the Norwegian Academy of Science and Letters, for hosting him as part of its international research program on Nonlinear PDEs during the academic year 2008-09.

motion between the camera and the object pictured, etc; such blurring is modeled by a linear, continuous operator, $T : L^2(\Omega) \to L^2(\Omega)$, e.g., convolution with a Gaussian kernel. Thus, the observed image, f, could be written as $f = TU + \eta$, where U is the clean image sought without blurring and noise. The recovery of the clean image from its observed blurred and noisy version f, is the problem of *image restoration*. This is an ill-posed problem which can be addressed by several inverse problems solvers. We mention in this context variational techniques using Tikhonov-like regularization, PDE-based methods, filtering, stochastic modeling and wavelets-based techniques that were developed for solving these image processing problems [1, 3, 9, 10, 11, 12, 13, 15, 17, 19, 20, 23, 24, 27].

Image restoration leads to *image decomposition*. For example, an observed image f with additive noise and no blurring is naturally decomposed into a denoised part, U_α, and a noisy part, $\eta_\alpha := f - U_\alpha$. Here, α is an algorithm-specific *scaling parameter*: in the case of Gaussian smoothing, for example, the variance of the Gaussian kernel may serve as such scaling parameter. Small scale features, categorized as noise, are then forced into η_α, resulting in a larger scale version, U_α, of the original image f. Thus, denoising of f generates *a multiscale representation*, $\{U_\alpha\}_{\alpha\in A}$ with a varying scaling parameter $\alpha \in A$. Our paper deals primarily with image restoration using PDE-based methods. Indeed, the novelty of our approach is the use of multiscale image representation based on *integro*-differential equations. The image representation is motivated by the variational-based hierarchical image decomposition of [25, 26]. Incidentally, this shows the intimate relation between the PDE-based and variational approaches in multiscale algorithms for image restoration.

We begin with some examples where denoising methods give rise to multiscale representations.

1.1. **Multiscale representations using PDE-based models.** We first discuss PDE-based models which produce multiscale representation $\{U(\cdot,t)\}_{t\geq 0}$ for a given image f. For convenience we use the time variable t as the scaling parameter. One of the earliest PDE-based methods for denoising a given image $f := U(\cdot,0)$ is the heat equation

$$\frac{\partial U}{\partial t} = \Delta U, \quad U \equiv U(x,t) : \Omega \times \mathbb{R}_+ \mapsto \mathbb{R}; \quad \frac{\partial U}{\partial \mathbf{n}}\Big|_{\partial\Omega} = 0. \tag{1.1a}$$

This yields a family of images, $\{U(\cdot,t) : \Omega \to \mathbb{R}\}_{t\geq 0}$, which can be viewed as smoothed versions of f. In this linear set up, smoothing is implemented by a convolution with the two-dimensional Gaussian kernel, $G_\sigma(x) = \frac{1}{2\pi\sigma^2}\exp\left(-\frac{|x|^2}{2\sigma^2}\right)$, with standard deviation $\sigma = \sqrt{2t}$. Hence, details with a scale smaller than $\sqrt{2t}$ are smoothed out. Here, $\lambda(t) := \sqrt{2t}$ acts as a *scaling function*. We can say that $\{U(\cdot,t)\}_{t\geq 0}$ is a multiscale representation of f, as $U(\cdot,t)$ diffuses from the small scales in f into increasingly larger scales.

Image denoising by the heat equation is based on isotropic diffusion, and consequently blurs all edges, which may contain useful information about the image. This drawback was removed by Perona-Malik (PM) model [23], which is based on *nonlinear diffusion*

$$\frac{\partial U}{\partial t} = \operatorname{div}(g(|\nabla U|)\nabla U), \quad U : \Omega \times \mathbb{R}_+ \mapsto \mathbb{R}; \quad \frac{\partial U}{\partial \mathbf{n}}\Big|_{\partial\Omega} = 0, \tag{1.1b}$$

with an initial condition $U(\cdot,0) := f$. Here, the diffusion controlling function, g, is a real valued function that vanishes at infinity, so that the amount of diffusion decreases as the gradient $|\nabla U|$ increases. Thus, g is responsible for the anisotropic nature of the PM model. The family of PM models are not well-posed. They also pose a problem for noisy images, since noise produces high gradients which can be confused with relevant edges. These shortcomings were removed by Catté et. al. [6] by replacing $g(|\nabla U|)$ with $g(|G_\sigma \star \nabla U|)$, where $G_\sigma \star \nabla U$ denotes convolution of the two-dimensional Gaussian kernel G_σ,

$$\frac{\partial U}{\partial t} = \operatorname{div}(g(|G_\sigma \star \nabla U|)\nabla U), \quad U : \Omega \times \mathbb{R}_+ \mapsto \mathbb{R}; \quad \frac{\partial U}{\partial \mathbf{n}}\Big|_{\partial\Omega} = 0, \tag{1.1c}$$

subject to $U(\cdot,0) := f$.

The models (1.1) still suffer from a major drawback, namely, the solution $U(t)$ diffuses to the average value $\fint f$, as $t \to \infty$. Thus, a stopping criteria $t = t_c$ must be sought, so that the desired denoised image $U_c := U(t_c)$ is obtained. This raises the question of an appropriate stopping time t_c. The necessity of finding a stopping time is removed in Nordström's biased anisotropic model [22]

$$\frac{\partial U}{\partial t} = f - U + \operatorname{div}\left(g(|\nabla U|)\nabla U\right), \quad U : \Omega \times \mathbb{R}_+ \mapsto \mathbb{R}; \quad \frac{\partial U}{\partial \mathbf{n}}\Big|_{\partial\Omega} = 0. \tag{1.2}$$

In this case, the solution $U(\cdot,t)$ varies from the initial condition $U(\cdot,0) \equiv 0$ to a desired denoised image U_c, as $t \to \infty$. Thus, the family $\{U(\cdot,t)\}_{t\geq 0}$ is an inverse scale representation of U_c, with t acting as an *inverse scale* parameter, e.g., [14, 16, 5]

1.2. **Multiscale representations using variational models.** Variational approaches for image processing like Mumford-Shah segmentation [20], [21], Rudin-Osher-Fatemi (ROF) decomposition [24] etc., fall under a general category of Tikhonov regularization [27]. Here one solves the ill-posed problem of recovering u from the observed $f = Tu + \eta$. We begin by restricting our attention to the pure denoising problem seeking a faithful, noise free approximation $u \in X$ of $f = u + \eta \in L^2$, where $X \subsetneq L^2$ is an appropriate space adapted to measure edges and textures sought in u (a discussion of the deblurring problem is postponed to section 5). This leads to the following minimization problem:

$$f = u_\lambda + v_\lambda, \quad [u_\lambda, v_\lambda] := \operatorname*{arginf}_{f=u+v} \left\{ \|u\|_X + \lambda \|v\|_{L^2}^2 \right\}.$$

The term $\|u\|_X$ is a regularizing term and $u_\lambda + v_\lambda$ is a multiscale decomposition of f which varies with the positive scaling parameter, λ. In the case of the ROF model [24], for example, edges are sought in the space of bounded variations, $X = BV(\Omega)$, e.g., [2]. This yields the (BV, L^2)-decomposition of f:

$$f = u_\lambda + v_\lambda, \quad [u_\lambda, v_\lambda] := \operatorname*{arginf}_{f=u+v} \{ \|u\|_{BV} + \lambda \|v\|_{L^2}^2 \}, \tag{1.3}$$

where $\|u\|_{BV} := \int_\Omega |\nabla u|$. For small values of λ, the minimizer u_λ is only a large-scale image, consisting of only main features and prominent edges in f. On the other hand, if λ is large, then u_λ is a small-scale image which contains many details of f. Therefore, with λ viewed as a varying parameter, the ROF variational decomposition (1.3) generates a multiscale representation, $\{u_\lambda\}_{\lambda>0}$, of f, with λ serving as an *inverse-scale* parameter. The behavior of this multiscale decomposition, as a function of λ, is tied to the regularity of f, once the variational functional on the right is interpreted as an interpolation K-functional, [4].

The Euler-Lagrange equation characterizing the minimizer, u_λ, for the variational problem (1.3) reads,

$$u_\lambda = f + \frac{1}{2\lambda}\,\mathrm{div}\left(\frac{\nabla u_\lambda}{|\nabla u_\lambda|}\right). \tag{1.4a}$$

For a fixed λ, the minimizer of (1.3) can be obtained as a steady state solution of the nonlinear parabolic equation

$$\frac{\partial u}{\partial t} = f - u + \frac{1}{2\lambda}\,\mathrm{div}\left(\frac{\nabla u}{|\nabla u|}\right), \quad u \equiv u(x,t) : \Omega \times \mathbb{R}_+ \mapsto \mathbb{R}; \quad \frac{\partial u}{\partial \mathbf{n}}\Big|_{\partial\Omega} = 0. \tag{1.4b}$$

Starting with $u(\cdot,0) := f$, the PDE (1.4b) produces a multiscale representation $\{u(\cdot,t)\}_{t\geq 0}$ which approaches the ROF minimizer, u_λ, as $t \uparrow \infty$. Observe that t in (1.4b) serves as a *forward-scale* parameter for the variational ROF model (1.3). Incidentally, the variational-based PDE (1.4b) is related to Nordström model (1.2) with $g(s) := \dfrac{1}{2\lambda s}$.

1.3. **A novel multiscale integro-differential model.** In this paper, we introduce a novel *integro-differential equation* (IDE) for multiscale representation of f

$$\int_0^t u(x,s)\,ds = f(x) + \frac{1}{2\lambda(t)}\,\mathrm{div}\left(\frac{\nabla u(x,t)}{|\nabla u(x,t)|}\right), \quad u : \Omega \times \mathbb{R}_+ \mapsto \mathbb{R}; \quad \frac{\partial u}{\partial \mathbf{n}}\Big|_{\partial\Omega} = 0, \tag{1.5}$$

subject to appropriate initial condition $u(\cdot,0) = u_0(x)$ outlined in section 3 below. The integral $U(\cdot,t) := \int_0^t u(\cdot,s)\,ds$ gives a scaled version of the image f for a given t. The scaling function, $\lambda(t)$ is at our disposal. The image $U(t)$ evolves with t, from a coarse (or larger) scale images, to smaller scale images with finer details, as $\lambda(t)$ increases with time. Thus, (1.5) is an *inverse scale* method, as opposed to the *forward scale* methods such as heat equation or PM models (1.1).

The motivation behind this IDE comes from the hierarchical (BV, L^2) multiscale image decomposition of Tadmor et. al., [25, 26], which we will elaborate upon in the next section. In particular, in section 3.1 we show how the choice of scaling function $\lambda(t)$ dictates the size of the residual image $V(t) := f - U(t)$. In sections 4.1 and 4.2, we propose further extensions of our IDE approach which introduce further refinements to tangential smoothing. A final extension of the IDE model for dealing with deblurring is presented in section 5. The details of the numerical schemes used to implement the various IDE models are outlined in the Appendix.

2. **Motivation for the Integro-Differential Equation (IDE).** Rudin, Osher and Fatemi introduced a BV-based minimization functional for image denoising in [24], which in turn led to the unconstrained (BV, L^2) decomposition (1.3) in [7, 8]. The minimizer of (1.3), u_λ, is a coarse representation of the image f, containing smooth parts and prominent edges, whereas the residual v_λ contains texture and finer details, declared as "noise" of f. The parameter λ is the *inverse* scale parameter of u_λ, i.e. a small value of λ corresponds to more details in v_λ and thus, the image u_λ is more coarse and vice versa.

As a first step, we realize that the intensity of images is quantized. If we let τ denote the small intensity quanta, then we rescale the coarse representation u_λ in τ-units. The corresponding (BV, L^2) image decomposition (1.3) takes the form

$$f = \tau u_{\lambda_0} + v_{\lambda_0}, \quad [u_{\lambda_0}, v_{\lambda_0}] := \operatorname*{arginf}_{f=\tau u+v}\left\{\|u\|_{BV} + \frac{\lambda_0}{\tau}\,\|v\|_{L^2}^2\right\}. \tag{2.1}$$

Tadmor, Nezzar and Vese observed in [25] that for a small value of the scaling parameter λ_0, the residual image v_{λ_0} may still contain important details when viewed at a finer scale. Thus, v_{λ_0} can be further decomposed using a refined scaling parameter $\lambda_1 > \lambda_0$,

$$v_{\lambda_0} = \tau u_{\lambda_1} + v_{\lambda_1}, \quad [u_{\lambda_1}, v_{\lambda_1}] := \operatorname*{arginf}_{v_{\lambda_0}=\tau u+v} \left\{ \|u\|_{BV} + \frac{\lambda_1}{\tau} \|v\|_{L^2}^2 \right\}.$$

We can continue this process for $\lambda_0 < \lambda_1 < \lambda_2 \dots$

$$v_{\lambda_{j-1}} = \tau u_{\lambda_j} + v_{\lambda_j}, \quad [u_{\lambda_j}, v_{\lambda_j}] := \operatorname*{arginf}_{v_{\lambda_{j-1}}=\tau u+v} \left\{ \|u\|_{BV} + \frac{\lambda_j}{\tau} \|v\|_{L^2}^2 \right\}. \tag{2.2}$$

Repeating this refinement step, we obtain the following *hierarchical* multiscale representation of f, [25]

$$\begin{aligned} f &= \tau u_{\lambda_0} + v_{\lambda_0} \\ &= \tau u_{\lambda_0} + \tau u_{\lambda_1} + v_{\lambda_1} \\ &= \dots\dots \\ &= \tau u_{\lambda_0} + \tau u_{\lambda_1} + \dots \tau u_{\lambda_N} + v_{\lambda_N}. \end{aligned}$$

Thus, we have

$$\sum_{j=0}^{N} u_{\lambda_j} \tau = f - v_{\lambda_N}. \tag{2.3}$$

The Euler-Lagrange equations characterizing minimizers of (2.2) are

$$v_{\lambda_{j-1}} = \tau u_{\lambda_j} - \frac{1}{2\lambda_j} \operatorname{div}\left(\frac{\nabla u_{\lambda_j}}{|\nabla u_{\lambda_j}|} \right). \tag{2.4}$$

From (2.4) and (2.2) we get

$$v_{\lambda_j} = -\frac{1}{2\lambda_j} \operatorname{div}\left(\frac{\nabla u_{\lambda_j}}{|\nabla u_{\lambda_j}|} \right),$$

and inserting this into (2.3) yields the hierarchical decomposition of f as

$$\sum_{j=0}^{N} u_{\lambda_j} \tau = f + \frac{1}{2\lambda_N} \operatorname{div}\left(\frac{\nabla u_{\lambda_N}}{|\nabla u_{\lambda_N}|} \right). \tag{2.5}$$

We consider a multiscale scaling, continuous in time, $u(x,t) : \Omega \times \mathbb{R}_+ \mapsto \mathbb{R}$ such that $u_{\lambda_j}(x) \mapsto u(x, t^j := j\tau)$. Observe that the right hand side of (2.5) is homogeneous of degree zero. Letting $\tau \to 0$, the hierarchical description (2.5) motivates a multiscale representation $u(x,\cdot)$ which is sought as a solution to our IDE (1.5),

$$\int_0^t u(x,s)\, ds = f(x) + \frac{1}{2\lambda(t)} \operatorname{div}\left(\frac{\nabla u(x,t)}{|\nabla u(x,t)|} \right), \qquad \frac{\partial u}{\partial \mathbf{n}}\Big|_{\partial\Omega} = 0. \tag{2.6}$$

The IDE (2.6) needs to be augmented with a proper choice of a scaling function $\lambda(t)$ and one needs to set the initial conditions $\lambda(0)$ and $u(x,0)$. These will be discussed in section 3.2.

An an example for the IDE multiscale representation of an image f,

$$\left\{ U(\cdot,t) := \int_0^t u(\cdot,s)\, ds \right\}_{t\geq 0},$$

is depicted in figure 3.1. Here, $u(x,t)$ denotes the *speed* at which the image $U(t)$ changes with time. The numerical scheme for its evolution using the IDE (1.5) is prescribed in section 6.

Remark 1. It is instructive to compare our IDE model (2.6) with the time dependent PDE used in solving the ROF minimization, (1.4). In contrast to the forward scale PDE realization of (1.4b), where the solution evolves from $u(\cdot,0)=0$ to a bigger scale image u_λ, our IDE model (2.6) is an 'inverse scale' model, whose solution evolves from $u(x,0)=u_0(\cdot)$ to f as $\lambda(t)\to\infty$.

3. Specifying the augmenting parameters for the IDE.

To complete the formulation of the IDE (2.6), one has to specify a scaling function, $\lambda(t)$ and the *initial conditions* $u_0(x)\equiv u(x,0)$. The function $\lambda(t)$ serves as an *inverse scaling function*: as $\lambda(t)\to\infty$, the image computed in (2.6)

$$U(t):=\int_0^t u(x,s)\,ds,$$

extracts consecutively smaller scale slices of the original image f. The residual, $V(t):=f-U(t)$ contains texture and noisy parts of f. The choices of $\lambda(t)$ and $u_0(x)$ are outlined in sections 3.1 and 3.2 below.

3.1. On the scaling function $\lambda(t)$.

It is argued in [18] that the dual norm,

$$\|w\|_*:=\sup_{\|\varphi\|_{BV}\neq 0}\frac{(w,\varphi)}{\|\varphi\|_{BV}},$$

is a proper norm to measure texture. The critical role of the scaling function $\lambda(t)$ in the IDE model (2.6) and its relationship with the star-norm is outlined in the following theorem.

Theorem 3.1. *Consider the IDE model* (2.6)

$$\int_0^t u(x,s)\,ds=f(x)+\frac{1}{2\lambda(t)}\,\mathrm{div}\left(\frac{\nabla u(x,t)}{|\nabla u(x,t)|}\right),$$

and let $V(\cdot,t)$ *be the residual*

$$V(\cdot,t):=f-U(\cdot,t).$$

Then size of the residual is dictated by the scaling function $\lambda(t)$,

$$\|V(\cdot,t)\|_*=\frac{1}{2\lambda(t)}. \tag{3.1}$$

Proof. For $\varphi\in BV(\Omega)$ we have the following

$$|\,(V(\cdot,t),\varphi)\,|=\left|\left(\frac{1}{2\lambda(t)}\,\mathrm{div}\left(\frac{\nabla u(\cdot,t)}{|\nabla u(\cdot,t)|}\right),\varphi\right)\right|\le\frac{1}{2\lambda(t)}\|\varphi\|_{BV}. \tag{3.2}$$

Thus, we have $\|V(\cdot,t)\|_*\le\frac{1}{2\lambda(t)}$. Letting $\varphi=u(\cdot,t)$, we get

$$\left|\left(\frac{1}{2\lambda(t)}\,\mathrm{div}\left(\frac{\nabla u(\cdot,t)}{|\nabla u(\cdot,t)|}\right),u(\cdot,t)\right)\right|=\frac{1}{2\lambda(t)}\|u(\cdot,t)\|_{BV}. \tag{3.3}$$

From (3.2) and (3.3) we get the desired result (3.1). □

FIGURE 3.1. The images, $U(t) = \int_0^t u(\cdot, s)\, ds$, of the IDE (1.5) at $t = 1, 4, 6, 10$. Here, $\lambda(t) = 0.002 \times 2^t$.

The importance of Theorem 3.1 lies in the fact that it enables us to dictate the star-norm of the residual. For small values of $\lambda(t)$, we get a significant amount of texture in the residual and thus, the image $U(t) := \int_0^t u(\cdot, s)\, ds$ will contain only features with big scale. On the other hand, as $\lambda(t)$ increases, more and more details will appear in $U(t)$. Hence, the function $\lambda(t)$ can be viewed as an '*inverse scale function*' for $U(t)$. In particular, if we choose the scaling function $\lambda(t)$, such that $\lim_{t\to\infty} \lambda(t) = c$ with a prescribed constant c, then $\lim_{t\to\infty} \|V(t)\|_* = \frac{1}{2c}$. Thus, Theorem 3.1 enables us to denoise images to any pre-determined level in the BV^* sense.

The previous theorem establishes a weak convergence in the G-topology [18, §1.14], $U(t) \rightharpoonup f$, for all L^2-images. In fact, a stronger L^2-convergence holds for slightly more regular images, e.g., $f \in BV$. To this end we first prove the following energy decomposition, interesting in its own sake, along the lines of [25, theorem 2.2].

Theorem 3.2. *Consider the IDE model (2.6) associated with an L^2- image f, and let $V(\cdot, t)$ be the residual, $V(t) = f - U(t)$. Then the following energy decomposition holds*

$$\int_{s=0}^{t} \frac{1}{\lambda(s)} \|u(\cdot, s)\|_{BV}\, ds + \|V(\cdot, t)\|_{L^2}^2 = \|f\|_{L^2}^2. \tag{3.4}$$

To verify (3.4), integrate (2.6) against $u(\cdot, t)$ in space and time to find

$$\int_{s=0}^{t} \Big(U(\cdot, s), U_s(\cdot, s)\Big) ds - \big(f, U(\cdot, t)\big) = -\int_{s=0}^{t} \frac{1}{2\lambda(t)} \|u(\cdot, s)\|_{BV}\, ds.$$

The expression on the left is then rewritten as

$$\int_{s=0}^{t}\Big(U(\cdot,s),U_s(\cdot,s)\Big)ds-\Big(f,U(\cdot,t)\Big)=\frac{1}{2}\|U(\cdot,t)\|_{L^2}^2-\Big(f,U(\cdot,t)\Big)$$
$$\equiv\frac{1}{2}\Big[\Big(U(\cdot,t)-f,U(\cdot,t)-f\Big)\Big]-\frac{1}{2}\|f\|_{L^2}^2,$$

and (3.4) follows from the last two equalities.

Remark 2. A different, equivalent way of stating Theorem 3.2 is that $(u(t),V(t))$ form a *maximal pair* in the sense that they turn the inequality $(w,\varphi)\le\|w\|_{BV}\|\varphi\|_*$ into an equality:

$$(u(\cdot,t),V(\cdot,t))=\|u(\cdot,t)\|_{BV}\|V(\cdot,t)\|_*. \tag{3.5}$$

Indeed, differentiating (3.4) with respect to time we find

$$\frac{1}{\lambda(t)}\|u(\cdot,t)\|_{BV}+2\,(V(\cdot,t),-u(\cdot,t))=0,$$

and (3.5) follows in view of (3.1), $\|V(\cdot,t)\|_*=1/2\lambda(t)$.

We now turn to upper-bound the L^2-size of the residual. Using the usual duality estimate together with (3.1) to find

$$\|V(\cdot,t)\|_{L^2}^2\le\|V(\cdot,t)\|_*\|V(\cdot,t)\|_{BV}=\frac{1}{2\lambda(t)}\|V(\cdot,t)\|_{BV}, \tag{3.6}$$

and it remains to study how fast $\|V(\cdot,t)\|_{BV}$ grows. To this end we write

$$V(x,t)=f(x)-\int_{s=0}^{t/2}u(x,s)\,ds-\int_{s=t/2}^{t}u(x,s)\,ds,$$

which implies

$$\|V(\cdot,t)\|_{BV}$$
$$\le\|f\|_{BV}+\lambda\left(\frac{t}{2}\right)\int_{s=0}^{t/2}\frac{1}{\lambda(s)}\|u(\cdot,s)\|_{BV}\,ds+\lambda(t)\int_{s=t/2}^{t}\frac{1}{\lambda(s)}\|u(\cdot,s)\|_{BV}\,ds.$$

Inserting this into (3.6) we end up with the desired upper bound,

$$\|V(\cdot,t)\|_{L^2}^2\le\frac{1}{2\lambda(t)}\|f\|_{BV}+\frac{\lambda(\frac{t}{2})}{2\lambda(t)}\|f\|_{L^2}^2+\int_{s=t/2}^{t}\frac{1}{2\lambda(s)}\|u(\cdot,s)\|_{BV}\,ds.$$

Now, the first term on the right vanishes for $f\in BV$ at the $t=\infty$-limit as $\lambda(t)\uparrow\infty$; the second term vanishes if $\lambda(t)$ increases fast enough to form a Hadamard sequence so that $\lambda(t)/\lambda(\frac{t}{2})\uparrow\infty$ (e.g., $\lambda(t)\sim 2^t$); and the third term vanishes at $t\uparrow\infty$ as the tail of the uniformly bounded time integral in the energy bound (3.4). We summarize by stating the following.

Theorem 3.3. *Given an image $f\in BV$, we consider the IDE model (2.6) with rapidly increasing scaling function $\lambda(t)$ so that*

$$\frac{\lambda(\frac{t}{2})}{\lambda(t)}\overset{t\to\infty}{\longrightarrow}0.$$

Then, f admits the multiscale representation (where equality is interpreted in L^2-sense)

$$f(x) = \int_{s=0}^{\infty} u(x,s)\, ds, \tag{3.7a}$$

with energy decomposition

$$\|f\|_{L^2}^2 = \int_{s=0}^{\infty} \frac{1}{\lambda(s)} \|u(\cdot,s)\|_{BV}\, ds. \tag{3.7b}$$

3.2. **Setting the initial conditions.** At $t = 0$, the IDE (2.6) reads

$$f + \frac{1}{2\lambda(0)} \operatorname{div}\left(\frac{\nabla u(\cdot,0)}{|\nabla u(\cdot,0)|}\right) = 0.$$

Theorem 3.1 tells us that it has no solution if the initial value of the scaling function $\lambda(0) \equiv \lambda_0$ is set such that $\lambda(0) \neq 1/(2\|f\|_*)$. To gain a better understanding for the choice of the initial parameters, $\lambda(0) = \lambda_0$ and u_0, we return to the underlying discrete version of the IDE, given by hierarchical decomposition (2.5)

$$\sum_{j=0}^{N} u_{\lambda_j} \tau = f + \frac{1}{2\lambda_N} \operatorname{div}\left(\frac{\nabla u_{\lambda_N}}{|\nabla u_{\lambda_N}|}\right).$$

This is a discrete version of the IDE (2.6), where the term on the LHS is a quadrature of the corresponding integral, $\int_0^t u(x,s)ds$, sampled at the equidistant time-steps, $t^j = j\tau$, with τ being the basic intensity quanta.

At $t = 0$, $u(0) \equiv u_{\lambda(0)}$ is determined as the solution of

$$u(0)\tau = f + \frac{1}{2\lambda(0)} \operatorname{div}\left(\frac{\nabla u(\cdot,0)}{|\nabla u(\cdot,0)|}\right).$$

This is the Euler-Lagrange equation associated with the ROF variational decomposition (2.1), and according to [18, Theorem 3], [25, Corollary 2.5], the minimizer of the latter vanishes, $u(0) \equiv 0$, if

$$\lambda(0) < \frac{1}{2\|f\|_*}. \tag{3.8}$$

Assuming that (3.8) holds, then at the solution at the next hierarchical step, $u(t = \tau) \equiv u_{\lambda_1}$ is determined by

$$u(\tau)\tau = f + \frac{1}{2\lambda(\tau)} \operatorname{div}\left(\frac{\nabla u(\cdot,\tau)}{|\nabla u(\cdot,\tau)|}\right).$$

Viewed as the corresponding minimizer, $u(\tau)$ will vanish if $\lambda(\tau) \equiv \lambda_1 < 1/(2\|f\|_*)$. This process will continue to produce vanishing solutions $u(j\tau) \equiv u_{\lambda_j} = 0$ until the first time, $t_0 := j_0\tau$, when the scale $\lambda(t_0) = \lambda_{j_0}$ becomes *large enough* so that

$$\lambda(t_0) \geq \frac{1}{2\|f\|_*}. \tag{3.9a}$$

At this scale, the IDE picks up the first large features of the image f with non-trivial initial conditions, $u(t_0)$,

$$u(t_0)\tau = f + \frac{1}{2\lambda(t_0)} \operatorname{div}\left(\frac{\nabla u(\cdot,t_0)}{|\nabla u(\cdot,t_0)|}\right). \tag{3.9b}$$

The IDE (2.6),(3.9) then can be equivalently written as

$$\int_{t_0}^{t} u(x,s)\,ds = f(x) + \frac{1}{2\lambda(t)}\,\mathrm{div}\left(\frac{\nabla u(x,t)}{|\nabla u(x,t)|}\right), \qquad \frac{\partial u}{\partial \mathbf{n}}\Big|_{\partial\Omega} = 0, \qquad t \geq t_0, \tag{3.10}$$

where the initial time t_0 is determined as the first scale such that (3.9) holds. This setup is in complete analogy with the initialization process of the hierarchical decomposition in [26, section 2.3].

4. Extensions of the IDE model.

Our IDE model is motivated by a variational formulation. An important advantage of the IDE model, however, is that it is no longer limited to a variational formulation and we can therefore extend it using PDE-based modifications similar to (1.1b) and (1.1c). We will discuss such modifications in sections 4.1 and 4.2 below.

4.1. IDE with filtered diffusion.

Recall that one of the drawbacks in using the heat equation (1.1a) for denoising is that it results in an isotropic diffusion. The PM model (1.1b) removes this drawback by introducing a diffusion controlling function, that controls the diffusion near prominent edges in a given image. We propose a similar modification to our IDE model, seeking $u(x,t) : \Omega \times \mathbb{R}_+ \mapsto \mathbb{R}$ such that

$$\int_{0}^{t} u(x,s)\,ds = f(x) + \frac{g(|G_\sigma \star \nabla u(x,t)|)}{2\lambda(t)}\,\mathrm{div}\left(\frac{\nabla u(x,t)}{|\nabla u(x,t)|}\right); \quad \frac{\partial u}{\partial \mathbf{n}}\Big|_{\partial\Omega} = 0, \tag{4.1a}$$

subject to $u_0(x) \equiv u(\cdot,0)$ such that

$$u_0 = f + \frac{1}{2\lambda(0)} g(|G_\sigma \star \nabla u_0|)\,\mathrm{div}\left(\frac{\nabla u_0}{|\nabla u_0|}\right), \qquad \lambda(0) > \frac{g(|G_\sigma \star \nabla u_0|)}{2\|f\|_*}.$$

Similar to the PM models (1.1b), we can choose the pre-factor function g so that it vanishes at infinity to control the diffusion at prominent edges in the image. Thus, the function g acts here as a high-pass filter which retains prominent edges in the image $\int_0^t u(x,s)\,ds$ without diffusing them. As choices for such a g-filter, figure 4.1 displays the results of the modified IDE (4.1a) with

$$g(s) = \frac{1}{1 + (s/\beta)^2}, \tag{4.1b}$$

Here, the constant β determines the extent to which edges are preserved: for small β's, relevant edges are preserved whereas for large β's, they are diffused. Detailed discussion of the numerical scheme for the filtered diffusion model (4.1) is given in section 6. Comparing the results of the filtered IDE (4.1a) shown in figure 4.1, we observe that edges, which are diffused by the basic IDE (2.6) as depicted in figure 3.1, are preserved in figure 4.1. This phenomenon is more apparent for smaller values of t due to the fact that as t increases, $U(\cdot,t)$ in both models approaches f, and consequently, suffer from less diffusion of the edges. The usefulness of the filtered diffusion IDE model becomes apparent when certain edges are required in the scale-space for smaller values of t. For example, in figure 4.2, the edges are blurred for smaller values of t with the standard IDE (1.5), but with the filtered diffusion IDE (4.1a) we retain relevant edges, as shown in figure 4.3.

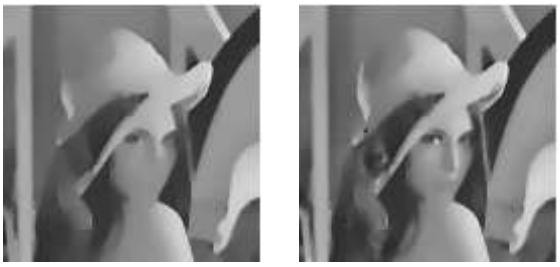


FIGURE 4.1. The images, $U(t) = \int_0^t u(\cdot, s)\, ds$, for the modified IDE (4.1) at $t = 1, 4, 6, 10$. Here $\lambda(t) = 0.002 \times 2^t$.

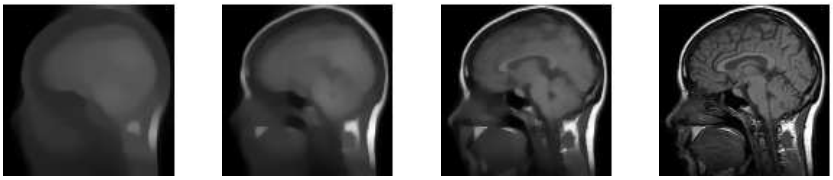


FIGURE 4.2. The images, $U(t) = \int_0^t u(\cdot, s)\, ds$, of the standard IDE (1.5) at $t = 1, 4, 6, 10$. Here, $\lambda(t) = 0.002 \times 2^t$.

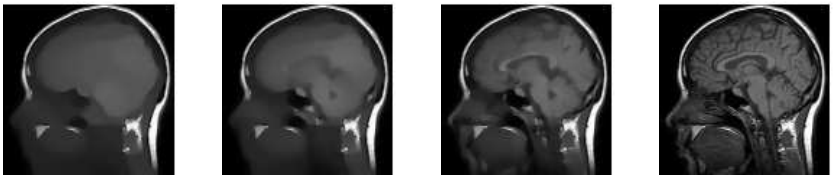
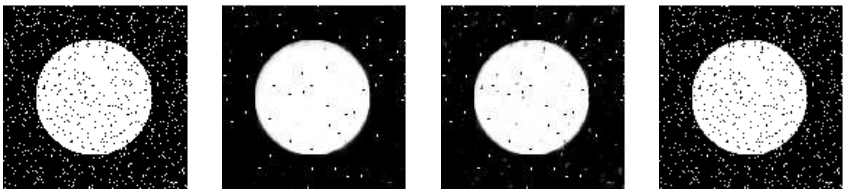


FIGURE 4.3. The images, $U(t) = \int_0^t u(\cdot, s)\, ds$, of the filtered IDE (4.1a) at $t = 1, 4, 6, 10$. Here, $\lambda(t) = 0.002 \times 2^t$.



f at $t=0$ $\quad$ $t=1$ $\quad$ $t=4$ $\quad$ $t=7$

FIGURE 4.4. A given noisy image f and the IDE images, $\int_0^t u(\cdot,s)\,ds$, of (1.5) at $t=1,4,7$. Here, the scaling function is $\lambda(t)=0.002\times 2^t$. Most of the noise is present at scale $t=7$.

4.2. **IDE with tangential smoothing.** The approach of using the diffusion controlling function works well with natural images with moderate gradients. With other images, however, such as those which often arise in computer vision and industrial applications, the boundaries of their internal objects are marked with large, sharp gradients; for example, characteristic function χ_D, where $D\subset\Omega$. In such cases, we can choose to smooth only in the tangential direction to the boundaries of the objects, e.g., [1]. To this end, write $\Delta u := u_{TT}+u_{NN}$, where u_{TT} and u_{NN} are the tangential and normal diffusion components, i.e.

$$u_{TT}=\Delta u-u_{NN}=|\nabla u|\,\mathrm{div}\left(\frac{\nabla u}{|\nabla u|}\right),\qquad u_{NN}=\left\langle\frac{\nabla u}{|\nabla u|},\nabla^2 u\frac{\nabla u}{|\nabla u|}\right\rangle.$$

Restricting the diffusion in our IDE model to tangential directions, this leads to modified IDEs with tangential smoothing,

$$\int_0^t u(x,s)\,ds=f(x)+\frac{1}{2\lambda(t)}|\nabla u(x,t)|\,\mathrm{div}\left(\frac{\nabla u(x,t)}{|\nabla u(x,t)|}\right);\quad \frac{\partial u}{\partial\mathbf{n}}\Big|_{\partial\Omega}=0, \tag{4.2}$$

and with tangential smoothing and filtering,

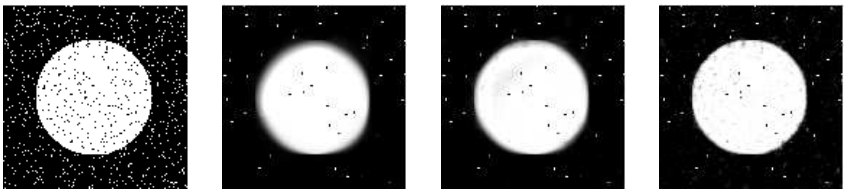
$$\begin{aligned}&\int_0^t u(x,s)\,ds\\ &=f(x)+\frac{g(|G_\sigma\star\nabla u(x,t)|)}{2\lambda(t)}|\nabla u(x,t)|\,\mathrm{div}\left(\frac{\nabla u(x,t)}{|\nabla u(x,t)|}\right);\quad \frac{\partial u}{\partial\mathbf{n}}\Big|_{\partial\Omega}=0.\end{aligned} \tag{4.3}$$

As before, $u:\Omega\times\mathbb{R}_+\mapsto\mathbb{R}$ evolves in inverse scale space starting with $\lambda(0)$ and $u_0(x)\equiv u(\cdot,0)$,

$$u_0\tau=f+\frac{1}{2\lambda(0)}g(|G_\sigma\star\nabla u_0|)|\nabla u_0|\,\mathrm{div}\left(\frac{\nabla u_0}{|\nabla u_0|}\right),\qquad \lambda(0)\geq\frac{g(|G_\sigma\star\nabla u_0|)|\nabla u_0|}{2\|f\|_*}.$$

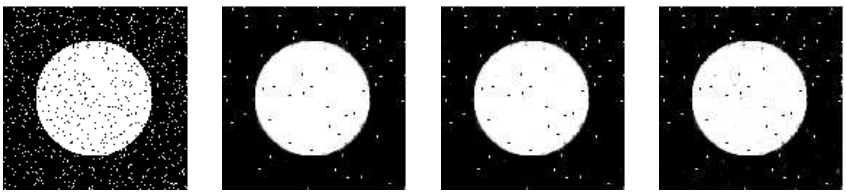
Numerical experiments are shown in figures (4.4)-(4.6). Compare the standard IDE results (1.5) shown in figure 4.4 with the tangential smoothing (4.2) shown in figure 4.5 and with additional filtering, (4.3), in figure 4.6: the point here is that tangential diffusion model preserves the edges, while denoising the rest of the image in a much faster rate than in the standard IDE model.

5. **The IDE model for deblurring.** We now extend our IDE model to deblurring of images. Blurring is modeled by a continuous, linear operator $T:L^2(\Omega)\to L^2(\Omega)$. Examples of a blurring operator include convolution with a Gaussian kernel,



f at $t = 0$ $t = 1$ $t = 4$ $t = 7$

FIGURE 4.5. The same noisy image f and the corresponding $\int_0^t u(\cdot, s)\, ds$, of the IDE with tangential smoothing (4.2) at $t = 1, 4, 7$. The same scaling function as before, $\lambda(t) = 0.002 \times 2^t$. Large portion of the noise is suppressed at $t = 7$ but there is normal diffusion of edges.



f at $t = 0$ $t = 1$ $t = 4$ $t = 7$

FIGURE 4.6. The same noisy image and the images, $\int_0^t u(\cdot, s)\, ds$, of IDE with tangential smoothing and filtering (4.3) at $t = 1, 4, 7$. Here, $\lambda(t) = 0.002 \times 2^t$ and $g(s) = 1/(1 + (s/5)^2)$. Noise is suppressed with minimal normal edge diffusion.

directional averaging etc. Thus, an observed image is expressed as $f = TU$, where U is the original unblurred image which we aim to recover. Hierarchical decomposition of blurred images was discussed in [26]. To this end, one sets a sequence of increasing scaling parameters $\lambda_0 < \lambda_1 < \lambda_2 \ldots$. Starting with $v_{-1} = f$, we get the following iteration

$$v_{\lambda_{j-1}} = \tau T u_{\lambda_j} + v_{\lambda_j}, \quad \operatorname*{arginf}_{v_{\lambda_{j-1}} = \tau T u + v} \{ \|u\|_{BV} + \frac{\lambda_j}{\tau} \|v\|_{L^2}^2 \}. \tag{5.1}$$

This gives us a *hierarchical multiscale representation* of the blurred image f presented in [26],

$$\begin{aligned} f &= \tau T u_{\lambda_0} + v_{\lambda_0} \\ &= \tau T u_{\lambda_0} + \tau T u_{\lambda_1} + v_{\lambda_1} \\ &= \ldots\ldots \\ &= \tau T u_{\lambda_0} + \tau T u_{\lambda_1} + \ldots \tau T u_{\lambda_N} + v_{\lambda_N}. \end{aligned}$$

Thus, after applying the conjugate T^* to the above equation we obtain,

$$\tau \sum_{j=0}^{N} T^* T u_{\lambda_j} = T^* f - T^* v_{\lambda_N}. \tag{5.2}$$

Using the Euler-Lagrange characterization of the minimizer in (5.1),

$$T^*v_{\lambda_{j-1}} = \tau T^*Tu_{\lambda_j} - \frac{1}{2\lambda_j}\,\mathrm{div}\left(\frac{\nabla u_{\lambda_j}}{|\nabla u_{\lambda_j}|}\right),$$

which, in view of $T^*v_{\lambda_{j-1}} = \tau T^*Tu_{\lambda_j} + T^*v_{\lambda_j}$ implies

$$T^*v_{\lambda_j} = -\frac{1}{2\lambda_j}\,\mathrm{div}\left(\frac{\nabla u_{\lambda_j}}{|\nabla u_{\lambda_j}|}\right).$$

Using the above expression we can rewrite (5.2) as

$$\sum_{j=0}^{N} T^*Tu_{\lambda_j}\tau = T^*f + \frac{1}{2\lambda_N}\,\mathrm{div}\left(\frac{\nabla u_{\lambda_N}}{|\nabla u_{\lambda_N}|}\right). \tag{5.3}$$

As $\tau \to 0$, the expression (5.3) motivates the following integro-differential equation (IDE) for deblurring, where $u(x,t) : \Omega \times \mathbb{R}_+ \mapsto \mathbb{R}$ is sought such that

$$\int_0^t T^*Tu(x,s)\,ds = T^*f(x) + \frac{1}{2\lambda(t)}\,\mathrm{div}\left(\frac{\nabla u(x,t)}{|\nabla u(x,t)|}\right); \quad \frac{\partial u}{\partial \mathbf{n}}\Big|_{\partial\Omega} = 0. \tag{5.4}$$

In this IDE, $\int_0^t u(\cdot,s)\,ds$ provides a multiscale representation of the *unblurred image* $U(x,t) := \int_0^t u(x,s)\,ds$. Note that the blurring operator T is in general non-invertible for general L^2 images, but it is assumed to be invertible on the restricted set of multiscale representations $\int_0^t T^*Tu(x,s)\,ds$. Thus, the deblurring IDE (5.4) gives us a recipe to extract the unblurred image U from its blurred version f.

We can see the deblurring result of (5.4) in figure 5.1. Furthermore, we can modify the deblurring integro-differential equation using edge enhancing filtering, where a $U(x,t) = \int_0^t u(x,s)\,ds : \Omega \times \mathbb{R}_+ \mapsto \mathbb{R}$ is sought as a solution of

$$T^*TU(x,t) = T^*f(x) + \frac{g(|G_\sigma \star u(x,t)|)}{2\lambda(t)}\,\mathrm{div}\left(\frac{\nabla u(x,t)}{|\nabla u(x,t)|}\right); \quad \frac{\partial u}{\partial \mathbf{n}}\Big|_{\partial\Omega} = 0. \tag{5.5}$$

Conclusions. We introduced a novel integro-differential equation (IDE) for multiscale decomposition of images. This is a continuous analogue of the hierarchical decomposition in [25, 26] with the same computational complexity of one ROF solver per time step. The basic IDE evolves in inverse time scale. Its continuous formulation enables us to incorporate related techniques from PDE-based methods of filtering, anisotropic tangential smoothing and deblurring. The resulting family of IDE models depend on a scaling function, $\lambda(t)$, at our disposal, which is shown to dictate the size of the error measured in the weak BV^*-norm. Numerical simulations show the utility of our IDE model as a promising alternative for PDE-based models. In particular, the deblurring results based on the IDE model are particularly striking.

6. Appendix: Numerical discretizations.

In this appendix we describe the numerical implementation of (1.5) and (5.4). First let us concentrate on the basic IDE model (1.5), rewritten here for convenience:

$$\int_0^t u(x,s)\,ds = f(x) + \frac{1}{2\lambda(t)}\,\mathrm{div}\left(\frac{\nabla u(x,t)}{|\nabla u(x,t)|}\right). \tag{6.1}$$

(a) (b)

FIGURE 5.1. Image (a) shows a blurred image of Lenna blurred using a Gaussian kernel with $\sigma = 1$. Image (b) shows the result of the deblurring IDE model (5.4), as $t \to \infty$.

As usual, $U(t) := \int_0^t u(x,s)\,ds$ is the exact solution. Let Δt be the time step and U^{n+1} will denote the corresponding computed solution at $t^{n+1} = (n+1)\Delta t$:

$$U^{n+1} = U^n + W^{n+1}, \qquad W^{n+1} \equiv W_{i,j}^{n+1} := u_{i,j}^{n+1}\Delta t,$$

where $u_{i,j}^{n+1} \equiv u^{n+1}(ih, jh)$ is the approximate solution of the IDE at grid point (ih, jh). With this, the IDE (6.1) is discretized at $t = t^{n+1}$:

$$\begin{aligned} U_{i,j}^n + \omega_{i,j}^{k+1} &= f_{i,j} \\ &+ \frac{1}{2\lambda^{(n+1)}h^2}\left[\frac{\omega_{i+1,j}^k - \omega_{i,j}^{k+1}}{\sqrt{\varepsilon^2 + (D_{+x}\omega_{i,j}^k)^2 + (D_{0y}\omega_{i,j}^k)^2}}\right. \\ &\qquad \left. - \frac{\omega_{i,j}^{k+1} - \omega_{i-1,j}^k}{\sqrt{\varepsilon^2 + (D_{-x}\omega_{i,j}^k)^2 + (D_{0y}\omega_{i-1,j}^k)^2}}\right] \\ &+ \frac{1}{2\lambda^{(n+1)}h^2}\left[\frac{\omega_{i,j+1}^k - \omega_{i,j}^{k+1}}{\sqrt{\varepsilon^2 + (D_{0x}\omega_{i,j}^k)^2 + (D_{+y}\omega_{i,j}^k)^2}}\right. \\ &\qquad \left. - \frac{\omega_{i,j}^{k+1} - \omega_{i,j-1}^k}{\sqrt{\varepsilon^2 + (D_{0x}\omega_{i,j-1}^k)^2 + (D_{-y}\omega_{i,j}^k)^2}}\right]. \end{aligned} \tag{6.2}$$

The nonlinear system (6.2) is solved using Jacobi iterations which leads to the fixed-point iterations for computing ω^{k+1}:

(6.3a)

$$\omega_{i,j}^{k+1} = \frac{2\lambda^{(n+1)}h^2(f_{i,j} - U_{i,j}^n) + c_E\omega_{i+1,j}^k + c_W\omega_{i-1,j}^k + c_S\omega_{i,j+1}^k + c_N\omega_{i,j-1}^k}{2\lambda^{(n+1)}h^2 + c_E + c_W + c_S + c_N}.$$

Here, $\lambda^{(n+1)} = \lambda(t^{n+1})$ are the discrete scaling parameters and c_E, c_W, c_S, c_N are the discrete coefficients

$$\begin{aligned} c_E &:= \frac{1}{\sqrt{\varepsilon^2 + (D_{+x}\omega^k_{i,j})^2 + (D_{0y}\omega^k_{i,j})^2}}, \\ c_W &:= \frac{1}{\sqrt{\varepsilon^2 + (D_{-x}\omega^k_{i,j})^2 + (D_{0y}\omega^k_{i-1,j})^2}}, \\ c_S &:= \frac{1}{\sqrt{\varepsilon^2 + (D_{0x}\omega^k_{i,j})^2 + (D_{+y}\omega^k_{i,j})^2}}, \\ c_N &:= \frac{1}{\sqrt{\varepsilon^2 + (D_{0x}\omega^k_{i,j-1})^2 + (D_{-y}\omega^k_{i,j})^2}}, \end{aligned}$$

In the computations above we set $h = 1$. To minimize the grid effects, we alternate the directions in which the above iterations were carried out, starting at the top-left corner position $(1,1)$, fixing $i = 1$ we vary $j = 1$ to j_{max} (East-South direction), initiating the next iteration at the top-right corner, and so on. This fixed point iterations (6.3a) yield $\omega^k \overset{k\to\infty}{\longrightarrow} W^{n+1} \equiv u^{n+1}\Delta t$ and we can update the computed image U:

$$U^{n+1} = U^n + W^{n+1}. \tag{6.3b}$$

Next, we consider the filtered IDE (4.1a), which is rewritten here for convenience as

$$\int_0^t u(x,s)\,ds = f(x) + \frac{g(|G_\sigma \star \nabla u(x,t)|)}{2\lambda(t)} \operatorname{div}\left(\frac{\nabla u(x,t)}{|\nabla u(x,t)|}\right).$$

The only difference here is the additional diffusion controlling function $g(|G_\sigma \star \nabla u(x,t)|)$, where G_σ is the two-dimensional Gaussian smoothing with standard deviation σ. The function $g(s) = \frac{1}{1+(s/\beta)^2}$ with $\beta = 5$ is used in our numerical experiments. We approximate

$$g(|G_\sigma \star \nabla u(x,t)|) \approx g\left(\left|G_\sigma \star \frac{\nabla \omega^n_{i,j}}{\Delta t}\right|\right),$$

and the expression on the right enters into the RHS of (6.2). We end up with the same discrete IDE scheme (6.3) with $\lambda^{(n)} \mapsto \lambda^{(n)}/g\left(\left|G_\sigma \star \nabla \omega^n_{i,j}/\Delta t\right|\right)$.

Finally, we describe the numerical implementation of the deblurring IDE models (5.4) and its filtered version (5.5). The equation (5.4) is rewritten here for convenience.

$$T^*T \int_0^t u(x,s)\,ds = T^* f(x) + \frac{1}{2\lambda(t)} \operatorname{div}\left(\frac{\nabla u(x,t)}{|\nabla u(x,t)|}\right). \tag{6.4}$$

Let $U(t) := \displaystyle\int_0^t u(x,s)\,ds$. As before, the left hand side of the above equation is approximated as follows

$$U^{n+1} = U^n + W^{n+1}, \qquad W^{n+1} \equiv W^{n+1}_{i,j} := u^{n+1}_{i,j}\Delta t, \tag{6.5}$$

and time-marching to a steady solution of (6.4) yields the following iteration for computing W^{n+1} as $W^{n+1} = \lim \omega^k$,

$$\begin{aligned}\frac{\omega_{i,j}^{k+1}-\omega_{i,j}^{k}}{\delta t} =& T^* f_{i,j} - T^* T U_{i,j}^{n} \\ &+ \frac{1}{2\lambda^{(n+1)}h^2}(c_E\omega_{i+1,j}^{k} + c_W\omega_{i-1,j}^{k} + c_S\omega_{i,j+1}^{k} + c_N\omega_{i,j-1}^{k}) \\ &- \frac{1}{2\lambda^{(n+1)}h^2}\omega_{i,j}^{k+1}(c_E + c_W + c_S + c_N),\end{aligned}$$

where c_E, c_W, c_S, c_N are defined as before.

REFERENCES

[1] L. Alvarez, P-L. Lions and J-M. Morel, *Image selective smoothing and edge detection by nonlinear diffusion II*, SIAM J. Numer. Anal., **29** (1992), 845–866.

[2] L. Ambrosio, N. Fusco and D. Pallara, "Functions of Bounded Variation and Free Discontinuity Problems," Oxford Mathematical monographs, Oxford University Press, New York, 2000.

[3] G. Aubert and L. Vese, *A variational method in image recovery*, SIAM J. Numer. Anal., **34** (1997), 1948–1979.

[4] C. Bennett and R. Sharpley, "Interpolation of Operators," Academic Press, 1988.

[5] M. Burger, G. Gilboa, S. Osher and J. Xu, *Nonlinear inverse scale space methods*, Comm. Math. Sci., **4** (2006), 179–212.

[6] F. Catté, P-L. Lions, J-M. Morel and T. Coll, *Image selective smoothing and edge detection by nonlinear diffusion*, SIAM J. Numer. Anal., **29** (1992), 182–193.

[7] A. Chambolle and P-L. Lions, *Image restoration by constrained total variation minimization and variants*, in "SPIE Electronic Imaging Proc.," **2567** (1995), 50–59.

[8] A. Chambolle and P-L. Lions, *Image recovery via total variation minimization and related problems*, Numer. Math., **76** (1997), 167–188.

[9] T. Chan and J. Shen, "Image Processing and Analysis, Variational, PDE, Wavelet, and Stochastic Methods," SIAM, Philadelphia, PA, 2005.

[10] T. Chan, J. Shen and L. Vese, *Variational PDE models in image processing*, Notices Amer. Math. Soc., **50** (2003), 14–26.

[11] S. Esedoglu and S. J. Osher, *Decomposition of images by the anisotropic Rudin-Osher-Fatemi model*, Comm. Pure Appl. Math., **57** (2004), 1609–1626.

[12] S. Geman and D. Geman, *Stochastic relaxation, gibbs distributions, and the bayesian restoration of images*, IEEE-PAMI, **6** (1984), 721–741.

[13] R. Gonzales and R. Woods, "Digital Image Processing," 3^{rd} edition, Prentice Hall, 2008.

[14] C. Groetsch and O. Scherzer, *Inverse scale space theory for inverse problems*, in "Lecture Notes in Computer Science, Proc. Third Int'l Conf. on Scale-Space and Morphology in Computer Vision", **2106**, (2001), 317–325.

[15] S. Li, "Markov Random Field Modeling in Computer Vision," Springer-Verlag, 1995.

[16] J. Lie and J. Nordbotten, *Inverse scale spaces for nonlinear regularization*, J. Math. Imaging and Vision, **27** (2007), 41–50.

[17] S. Mallat, "A Wavelet Tour of Signal Processing," Academic Press, San Diego, CA, 1998.

[18] Y. Meyer, "Oscillating Patterns in Image Processing and Nonlinear Evolution Equations," University Lecture Series, vol. **22**, American Mathematical Society, Providence, RI, 2001, The fifteenth Dean Jacqueline B. Lewis memorial lectures.

[19] V. Morozov, "Regularization Methods for Ill-Posed Problems," CRC Press, Boca Raton, FL, 1993, Translated from the 1987 Russian original.

[20] D. Mumford and J. Shah, *Boundary detection by minimizing functionals*, in "Proc. IEEE Computer Vision Pattern Recognition Conference", San Francisco, CA, (1985), 22–26.

[21] D. Mumford and J. Shah, *Optimal approximations by piecewise smooth functions and associated variational problems*, Comm. Pure Appl. Math., **42** (1989), 577–685.

[22] N. Nordström, *Biased anisotropic diffusion - a unified regularization and diffusion approach to edge detection*, Image and Vision Computing, **8** (1990), 318–327.

[23] P. Perona and J. Malik, *Scale-space and edge detection using anisotropic diffusion*, Tech. report, Berkeley, CA, USA, 1988.

[24] L. Rudin, S. Osher and E. Fatemi, *Nonlinear total variation based noise removal algorithms*, Phys. D, **60** (1992), 259–268.
[25] E. Tadmor, S. Nezzar and L. Vese, *A multiscale image representation using hierarchical (BV, L^2) decompositions*, Multiscale Model. Simul., **2** (2004), 554–579.
[26] E. Tadmor, S. Nezzar and L. Vese, *Multiscale hierarchical decomposition of images with applications to deblurring, denoising and segmentation*, Commun. Math. Sci., **6** (2008), 281–307.
[27] A. Tikhonov and V. Arsenin, "Solutions of Ill-Posed Problems," V. H.Winston & Sons, Washington, D.C.: John Wiley & Sons, New York, 1977, Translated from the Russian, Preface by translation editor Fritz John, Scripta Series in Mathematics.

Received July 2009; revised August 2009.

E-mail address: tadmor@cscamm.umd.edu
E-mail address: prashant@math.ucla.edu