Novel integro-differential equations in image processing and its applications

Prashant Athavale[a] and Eitan Tadmor[b]

[a]Institute of Pure and Applied Mathematics, University of California, Los Angeles, CA 90024, U.S.A.;

[b]Department of Mathematics, Institute for Physical Science & Technology and Center of Scientific Computation and Mathematical Modeling (CSCAMM), University of Maryland, College Park, MD 20742, U.S.A.

ABSTRACT

Motivated by the hierarchical multiscale image representation of Tadmor et al.,[1] 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 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. Finally, we also propose another novel IDE based on the (BV, L^1) decomposition. We present numerical results for this IDE and its variant and examine its properties.

Keywords: natural images, multiscale representation, total variation, denoising, deblurring, inverse scale, variational problem, integro-differential equation, energy decomposition.

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 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 motion between the camera and the object pictured, etc; such blurring is modeled by a linear, continuous operator, $K : L^2(\Omega) \to L^2(\Omega)$, e.g., convolution with a Gaussian kernel. Thus, the observed image, f, could be written as $f = KU + \eta$, 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. 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.[2–15]

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*, $\{U_\alpha\}_{\alpha\in\mathcal{A}}$ with a varying scaling parameter $\alpha \in \mathcal{A}$. 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.

Computational Imaging VIII, edited by Charles A. Bouman, Ilya Pollak, Patrick J. Wolfe,
Proc. of SPIE-IS&T Electronic Imaging, SPIE Vol. 7533, 75330S · © 2010 SPIE-IS&T
CCC code: 0277-786X/10/$18 · doi: 10.1117/12.850779

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

$$\frac{\partial U}{\partial t} = \Delta U, \quad U \equiv U(x,t) : \Omega \times \mathbb{R}_+ \mapsto \mathbb{R}; \quad \frac{\partial U}{\partial \mathbf{n}}\Big|_{\partial\Omega} = 0. \tag{1}$$

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.

This isotropic diffusion blurs all edges. This drawback was removed by Perona-Malik (PM) model,[13] where they replace the linear diffusion term ΔU with a *non-linear* term $\operatorname{div}(g(|\nabla U|)\nabla U)$. 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.[16] 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, subject to $U(\cdot,0) := f$.

1.2 Multiscale representations using variational models

Variational approaches for image processing like Mumford-Shah segmentation,[12],[17] Rudin-Osher-Fatemi (ROF) decomposition[14] etc., fall under a general category of Tikhonov regularization.[15] Here one solves the ill-posed problem of recovering u from the observed $f = Ku + \eta$. 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 4.2). 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,[14] for example, edges are sought in the space of bounded variations, $X = BV(\Omega)$, e.g.[18] This yields the (BV, L^2)-decomposition of f with $X = BV$, 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 generates a multiscale representation, $\{u_\lambda\}_{\lambda>0}$, of f, with λ serving as an *inverse-scale* parameter. The Euler-Lagrange equation characterizing the minimizer, u_λ, for the ROF decomposition reads,

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

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

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

Starting with $u(\cdot,0) := f$, the PDE (2b) 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 (2b) serves as a *forward-scale* parameter for the variational ROF decomposition.

In the next section we propose an integro-differential equation (IDE) based on (BV, L^2) decomposition, and examine some of its properties. In section 4 we examine some of the extensions of the proposed IDE. Furthermore, we propose another novel IDE based on (BV, L^1) in section 5.

2. MOTIVATION FOR THE INTEGRO-DIFFERENTIAL EQUATION (IDE)

Rudin, Osher and Fatemi introduced a BV-based minimization functional for image denoising,[14] which in turn led to the unconstrained (BV, L^2) decomposition.[19,20] The minimizer of the (BV, L^2) decomposition, 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, categorized 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 takes the form

$$f = \tau u_{\lambda_0} + v_{\lambda_0}, \quad [u_{\lambda_0}, v_{\lambda_0}] := \operatorname*{arginf}_{f=\tau u+v} \left\{ \|u\|_{BV} + \frac{\lambda_0}{\tau} \|v\|_{L^2}^2 \right\}. \tag{3}$$

Tadmor, Nezzar and Vese observed[1] 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 continue this process for $\lambda_0 < \lambda_1 < \lambda_2 \ldots$

$$v_{\lambda_{j-1}} = \tau u_{\lambda_j} + v_{\lambda_j}, \quad [u_{\lambda_j}, v_{\lambda_j}] := \operatorname*{arginf}_{v_{\lambda_{j-1}}=\tau u+v} \left\{ \|u\|_{BV} + \frac{\lambda_j}{\tau} \|v\|_{L^2}^2 \right\}. \tag{4}$$

Repeating this refinement step, we obtain the following *hierarchical* multiscale representation of f,[1]

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

Thus, we have

$$\sum_{j=0}^{N} u_{\lambda_j} \tau = f - v_{\lambda_N}. \tag{5}$$

The Euler-Lagrange equations characterizing minimizers of (4) are

$$v_{\lambda_{j-1}} = \tau u_{\lambda_j} - \frac{1}{2\lambda_j} \operatorname{div} \left(\frac{\nabla u_{\lambda_j}}{|\nabla u_{\lambda_j}|} \right). \tag{6}$$

From (6) and (4) 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 (5) yields the hierarchical decomposition of f as

$$\sum_{j=0}^{N} u_{\lambda_j} \tau = f + \frac{1}{2\lambda_N} \operatorname{div} \left(\frac{\nabla u_{\lambda_N}}{|\nabla u_{\lambda_N}|} \right). \tag{7}$$

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 (7) is homogeneous of degree zero. Letting $\tau \to 0$, the hierarchical description (7) motivates a multiscale representation $u(x,\cdot)$ which is sought as a solution to our IDE,

$$\int_0^t u(x,s)\, ds = f(x) + \frac{1}{2\lambda(t)} \operatorname{div} \left(\frac{\nabla u(x,t)}{|\nabla u(x,t)|} \right), \qquad \frac{\partial u}{\partial \mathbf{n}}\Big|_{\partial\Omega} = 0. \tag{8}$$

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 1. Here, $u(x,t)$ denotes the *speed* at which the image $U(t)$ changes with time.

It is instructive to compare our IDE model (8) with the time dependent PDE used in solving the ROF minimization, (2). In contrast to the forward scale PDE realization of (2b), where the solution evolves from $u(\cdot,0)=0$ to a bigger scale image u_λ, our IDE model (8) is an 'inverse scale' model, whose solution evolves from $u(x,0)=u_0(\cdot)$ to f as $\lambda(t)\to\infty$.

3. SOME PROPERTIES OF THE (BV, L^2) IDE

To complete the formulation of the IDE (8), 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 (8)

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

extracts consecutively smaller scale slices of the original image f. The residual, $V(t) := f - U(t)$ contains texture and noisy parts of f. It is argued in[21] that the dual norm,

$$\|w\|_* := \sup_{\|\varphi\|_{BV}\neq 0} \frac{(w,\varphi)_{L^2}}{\|\varphi\|_{BV}},$$

is a proper norm to measure texture. We prove the following useful lemma.

LEMMA 3.1. *If $u\in BV$ then the star-norm of the curvature term $\kappa(u) := \operatorname{div}\left(\frac{\nabla u}{|\nabla u|}\right)$ is unity.*

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

$$|\,(\kappa(u),\varphi)_{L^2}\,| = \left|\left(\operatorname{div}\left(\frac{\nabla u}{|\nabla u|}\right),\varphi\right)_{L^2}\right| \leq \|\varphi\|_{BV}. \tag{9}$$

Thus, we have $\|\kappa(u)\|_*\leq 1$. Letting $\varphi = u$, we get

$$|(\kappa(u),u)| = \|u\|_{BV}. \tag{10}$$

From (9) and (10) we conclude that $\|\kappa(u)\|_* = 1$. □

The critical role of the scaling function $\lambda(t)$ in the IDE model (8) and its relationship with the star-norm is outlined in the following theorem, which can be proved using the above lemma.

THEOREM 3.2. *Consider the IDE model* (8)

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

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 $\lambda(t)$,

$$\|V(\cdot,t)\|_* = \frac{1}{2\lambda(t)}. \tag{11}$$

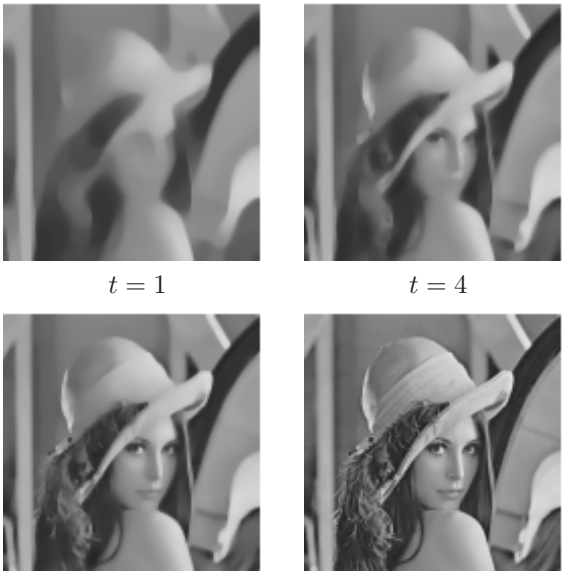


Figure 1. The images, $U(t) = \int_0^t u(\cdot, s)\, ds$, of the IDE (8) at $t = 1, 4, 6, 10$. Here, $\lambda(t) = 0.002 \times 2^t$.

The importance of Theorem 3.2 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 3.2 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 [21, §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 Tadmor et al. prove the following energy decomposition,[22] interesting in its own sake, along the lines of [1, theorem 2.2].

THEOREM 3.3. *Consider the IDE model (8) 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*

$$\int_{s=0}^{t} \frac{1}{\lambda(s)} \|u(\cdot, s)\|_{BV}\, ds + \|V(\cdot, t)\|_{L^2}^2 = \|f\|_{L^2}^2. \tag{12}$$

Using this theorem one can prove[22] that for the IDE model (8) with $f \in BV$ and a rapidly increasing scaling function $\lambda(t)$, the function f admits the multiscale representation (where equality is interpreted in L^2- sense):

$$f(x) = \int_{s=0}^{\infty} u(x, s)\, ds, \tag{13a}$$

with energy decomposition

$$\|f\|_{L^2}^2 = \int_{s=0}^{\infty} \frac{1}{\lambda(s)} \|u(\cdot, s)\|_{BV}\, ds. \tag{13b}$$

4. EXTENSIONS OF THE (BV, L^2) 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 Perona-Malik models. We will discuss such modifications in sections 4.1 below.

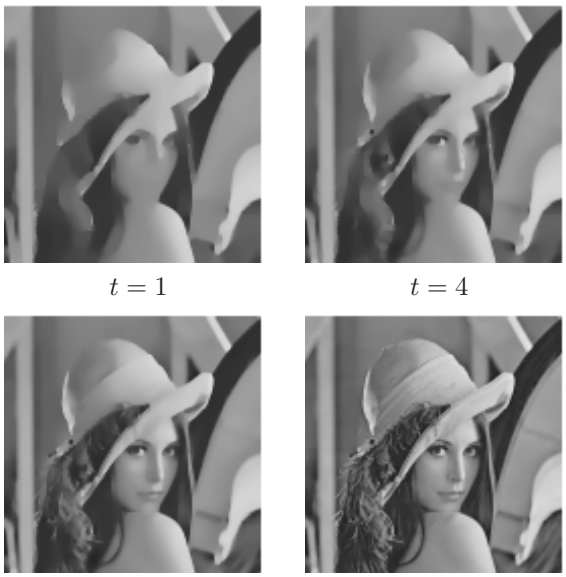


Figure 2. The images, $U(t) = \int_0^t u(\cdot, s)\, ds$, for the modified IDE (14) at $t = 1, 4, 6, 10$. Here $\lambda(t) = 0.002 \times 2^t$.

4.1 IDE with filtered diffusion

Recall that one of the drawbacks in using the heat equation for denoising is that it results in an isotropic diffusion. The Perona-Malik models 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

$$\int_0^t u(x,s)\, ds = f(x) + \frac{g(|G_\sigma \star \nabla u(x,t)|)}{2\lambda(t)} \operatorname{div}\left(\frac{\nabla u(x,t)}{|\nabla u(x,t)|}\right); \quad \frac{\partial u}{\partial \mathbf{n}}\Big|_{\partial\Omega} = 0. \tag{14a}$$

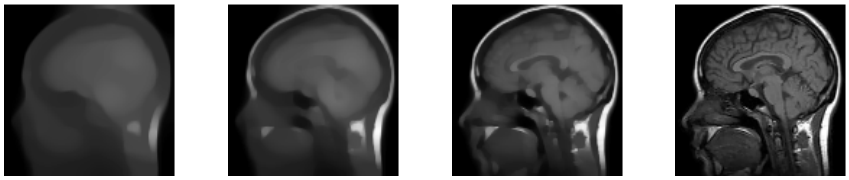
Similar to the Perona Malik models, we can choose the pre-factor function g so that it vanishes at infinity to control the diffusion at prominent edges in the image. Thus, the function g acts here akin to a high-pass filter which retains prominent edges in the image $\int_0^t u(x,s)\, ds$ without diffusing them. As choices for such a g-filter, figure 2 displays the results of the modified IDE (14a) with

$$g(s) = \frac{1}{1 + (s/\beta)^2}, \tag{14b}$$

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 (14) is given in section 6. Comparing the results of the filtered IDE (14a) shown in figure 2, we observe that edges, which are diffused by the basic IDE (8) as depicted in figure 1, are preserved in figure 2. 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 3, the edges are blurred for smaller values of t with the standard IDE (8), but with the filtered diffusion IDE (14a) we retain relevant edges, as shown in figure 4.

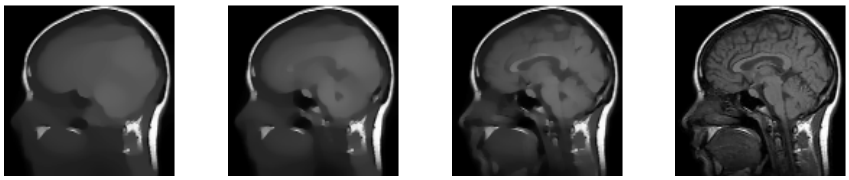
4.2 The IDE model for deblurring

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



$t = 1$ $t = 4$ $t = 6$ $t = 10$

Figure 3. The images, $U(t) = \int_0^t u(\cdot, s)\, ds$, of the standard IDE (8) at $t = 1, 4, 6, 10$. Here, $\lambda(t) = 0.002 \times 2^t$.



$t = 1$ $t = 4$ $t = 6$ $t = 10$

Figure 4. The images, $U(t) = \int_0^t u(\cdot, s)\, ds$, of the filtered IDE (14a) at $t = 1, 4, 6, 10$. Here, $\lambda(t) = 0.002 \times 2^t$.

averaging etc. Thus, an observed image is expressed as $f = KU$, where U is the original unblurred image which we aim to recover. Hierarchical decomposition of blurred images was discussed by Tadmor et al.[23] 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

$$v_{\lambda_{j-1}} = \tau K u_{\lambda_j} + v_{\lambda_j}, \quad \underset{v_{\lambda_{j-1}} = \tau K u + v}{\operatorname{arginf}} \{ \|u\|_{BV} + \frac{\lambda_j}{\tau} \|v\|_{L^2}^2 \}. \tag{15}$$

This gives us a *hierarchical multiscale representation*[23] of the blurred image f,

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

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

$$\tau \sum_{j=0}^{N} K^* K u_{\lambda_j} = K^* f - K^* v_{\lambda_N}. \tag{16}$$

Using the Euler-Lagrange characterization of the minimizer in (15),

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

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

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

Using the above expression we can rewrite (16) as

$$\sum_{j=0}^{N} K^* K u_{\lambda_j} \tau = K^* f + \frac{1}{2\lambda_N} \operatorname{div} \left(\frac{\nabla u_{\lambda_N}}{|\nabla u_{\lambda_N}|} \right). \tag{17}$$

(a) (b)

Figure 5. 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 (18), as $t \to \infty$.

As $\tau \to 0$, the expression (17) motivates the following integro-differential equation (IDE) for deblurring, where $u(x,t) : \Omega \times \mathbb{R}_+ \mapsto \mathbb{R}$ is sought such that

$$\int_0^t K^*Ku(x,s)\,ds = K^*f(x) + \frac{1}{2\lambda(t)}\,\mathrm{div}\left(\frac{\nabla u(x,t)}{|\nabla u(x,t)|}\right); \quad \frac{\partial u}{\partial \mathbf{n}}\Big|_{\partial\Omega} = 0. \tag{18}$$

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 K is in general non-invertible for general L^2 images, but it is assumed to be invertible on the restricted set of multiscale representations $\int_0^t K^*Ku(x,s)\,ds$. Thus, the deblurring IDE (18) gives us a recipe to extract the unblurred image U from its blurred version f.

We can see the deblurring result of (18) in figure 5. 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

$$K^*KU(x,t) = K^*f(x) + \frac{g(|G_\sigma \star u(x,t)|)}{2\lambda(t)}\,\mathrm{div}\left(\frac{\nabla u(x,t)}{|\nabla u(x,t)|}\right); \quad \frac{\partial u}{\partial \mathbf{n}}\Big|_{\partial\Omega} = 0. \tag{19}$$

5. INTEGRO-DIFFERENTIAL EQUATIONS BASED ON (BV, L^1) MINIMIZATION

Chan and Esedoḡlu[24] modify the L^2-fidelity term in the ROF minimization problem with an L^1 fidelity term, to get the following (BV, L^1) minimization problem,

$$f = u_\lambda + v_\lambda, \quad [u_\lambda, v_\lambda] := \underset{f=u+v}{\mathrm{arginf}}\{\, \|u\|_{BV} + \lambda\,\|v\|_{L^1}\}, \tag{20}$$

The (BV, L^1) minimization has very different properties from the (BV, L^2) minimization. The (BV, L^1) minimization is contrast invariant, as opposed to the (BV, L^2) minimization. Chan and Esedoḡlu[24] also show that the (BV, L^1) minimization is geometric in nature. To this effect they prove that for $f \in L^1(\Omega)$ with $\mathrm{supp}(f) \in B_R(0)$, then the (BV, L^1) decomposition (20) yields a trivial minimizing pair, $u_\lambda = 0$ and $v_\lambda = f$, for all λ which are less than a critical threshold $\lambda_L \propto \frac{1}{R}$.

Secondly, if the f is a characteristic function χ_Σ then there exists a λ_H such that for all $\lambda > \lambda_H$, the minimization (20) yields a unique minimizing pair $u_\lambda = f$, $v = 0$.

Recall, in the case of the standard ROF model, a similar critical value of the scale parameter exists, but it depended on the star-norm of the function f. Chan and Esedoḡlu demonstrate[24] that the scale space generated by λ is essentially different than the scale space generated in case of the standard ROF model.

Using these properties as a basis we propose a hierarchical image decomposition[25] as follows. Starting with a small value of λ_0, we can decompose the given image f using the (BV, L^1) scheme as follows:

$$f = \tau u_{\lambda_0} + v_{\lambda_0}, \quad [u_{\lambda_0}, v_{\lambda_0}] := \underset{f=\tau u+v}{\operatorname{arginf}} \{ \|u\|_{BV} + \frac{\lambda_0}{\tau} \|f-u\|_{L^1} \}.$$

The image v_{λ_0} can further be decomposed into smaller scale with $\lambda_1 > \lambda_0$,

$$v_{\lambda_0} = \tau u_{\lambda_1} + v_{\lambda_1}, \quad [u_{\lambda_1}, v_{\lambda_1}] := \underset{v_{\lambda_0}=\tau u+v}{\operatorname{arginf}} \{ \|u\|_{BV} + \frac{\lambda_1}{\tau} \|v_{\lambda_0}-u\|_{L^1} \}.$$

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

$$v_{\lambda_{k-1}} = \tau u_{\lambda_k} + v_{\lambda_k}, \quad [u_{\lambda_k}, v_{\lambda_k}] := \underset{v_{\lambda_{k-1}}=\tau u+v}{\operatorname{arginf}} \{ \|u\|_{BV} + \frac{\lambda_k}{\tau} \|v_{\lambda_{k-1}}-u\|_{L^1} \}. \tag{21}$$

Repeating this refinement step, we obtain the following *hierarchical* (BV, L^1) *decomposition* of f:

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

This yields a *hierarchical* (BV, L^1) *multiscale image decomposition*,

$$f = \sum_{k=0}^{N} \tau u_{\lambda_k} + v_{\lambda_N}. \tag{22}$$

The N^{th} step in (BV, L^1) scheme $\tau u_{\lambda_k} + v_{\lambda_k} = v_{\lambda_{k-1}}$

$$[u_{\lambda_N}, v_{\lambda_N}] = \underset{\{v_{\lambda_{N-1}}=\tau u+v\}}{\operatorname{arginf}} \left(\int_\Omega |\nabla u| + \frac{\lambda_N}{\tau} \int_\Omega |v_{\lambda_{N-1}} - \tau u| \right)$$

The Euler-Lagrange differential equation for the above minimization reads

$$\operatorname{sgn}(\tau u_{\lambda_N} - v_{\lambda_{N-1}}) = \frac{1}{\lambda_N} \operatorname{div} \left(\frac{\nabla u_{\lambda_N}}{|\nabla u_{\lambda_N}|} \right) \tag{23}$$

From (22) we have

$$v_{\lambda_{N-1}} = f - \sum_{k=0}^{N-1} \tau u_{\lambda_k}.$$

Using the above expression in (23) we get

$$\operatorname{sgn} \left(\sum_{k=0}^{N} u_{\lambda_k} \tau - f \right) = \frac{1}{\lambda_N} \operatorname{div} \left(\frac{\nabla u_{\lambda_N}}{|\nabla u_{\lambda_N}|} \right).$$

This motivates the following IDE:

$$\operatorname{sgn}\left(\int_{s=0}^{t} u(x,s)\,dx - f(x)\right) = \frac{1}{\lambda(t)}\operatorname{div}\left(\frac{\nabla u(x,t)}{|\nabla u(x,t)|}\right), \tag{24}$$
$$u:\Omega\times\mathbb{R}_+\mapsto\mathbb{R};\quad \frac{\partial u}{\partial \mathbf{n}}\Big|_{\partial\Omega}=0.$$

Using lemma 3.1, we obtain the following theorem regarding the IDE (24).

THEOREM 5.1. *Consider the IDE model* (24)

$$\operatorname{sgn}\left(\int_{0}^{t} u(x,s)\,ds - f(x)\right) = \frac{1}{\lambda(t)}\operatorname{div}\left(\frac{\nabla u(x,t)}{|\nabla u(x,t)|}\right),$$

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

$$V(\cdot,t) := f - \int_0^t u(\cdot,s)\,ds.$$

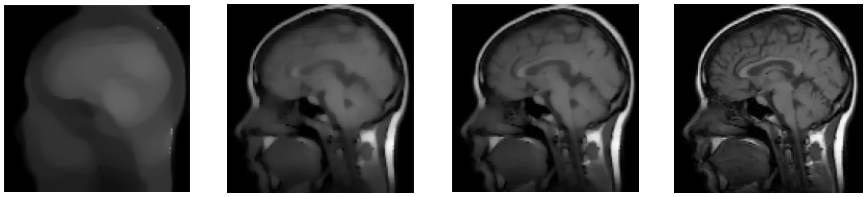
Then size of the residual is dictated by the scaling function $\lambda(t)$*,*

$$\|\operatorname{sgn}(V(\cdot,t))\|_* = \frac{1}{\lambda(t)}. \tag{25}$$

Finally, we propose edge enhancing modification to (24) along the lines of (14b), seeking $u(x,t):\Omega\times\mathbb{R}_+\mapsto\mathbb{R}$ such that

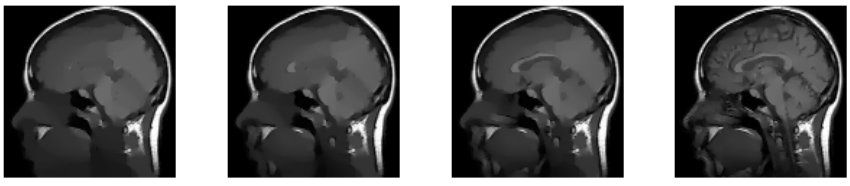
$$\operatorname{sgn}\left(\int_{s=0}^{t} u(x,s)\,dx - f(x)\right) = \frac{g(|G_\sigma\star\nabla u(x,t)|)}{2\lambda(t)}\operatorname{div}\left(\frac{\nabla u(x,t)}{|\nabla u(x,t)|}\right);\quad \frac{\partial u}{\partial \mathbf{n}}\Big|_{\partial\Omega}=0, \tag{26}$$

The numerical results of (26) are depicted in figure 7.



$t=1$ $t=6$ $t=7$ $t=10$

Figure 6. The images, $U(t)=\int_0^t u(\cdot,s)\,ds$, of the (BV,L^1) IDE (24) at $t=1,6,7,10$. Here, $\lambda(t)=0.01\times 2^t$.



$t=1$ $t=6$ $t=7$ $t=10$

Figure 7. The images, $U(t)=\int_0^t u(\cdot,s)\,ds$, of the (BV,L^1) IDE (26) at $t=1,6,7,10$. Here, $\lambda(t)=0.001\times 2^t$.

6. NUMERICAL DISCRETIZATIONS

In this section we outline numerical implementations for the IDEs proposed in this paper. First let us concentrate on the basic IDE model (8), rewritten here for convenience:

$$\int_0^t u(x,s)\,ds = f(x) + \frac{1}{2\lambda(t)}\,\mathrm{div}\left(\frac{\nabla u(x,t)}{|\nabla u(x,t)|}\right). \tag{27}$$

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 $t^{n+1} = (n+1)\Delta t$:

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

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 (8) is discretized[1, 22, 23] at $t = t^{n+1}$. The resulting nonlinear system is solved using Jacobi iterations which leads to the fixed-point iterations[22] to compute $\omega_{i,j}^{n+1}$.

Next, we consider the filtered IDE (14a), which is rewritten here for convenience as

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

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

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

and we end up with a similar discrete IDE scheme with $\lambda^{(n)} \mapsto \lambda^{(n)}/g\left(\left|G_\sigma \star \nabla \omega_{i,j}^n/\Delta t\right|\right)$.

For the deblurring IDE (18) we employ a time marching to compute $\omega_{i,j}^{n+1}$. For the (BV, L^1) induced IDE (24) we replace $\lambda(t)$ with $\frac{\lambda(t)}{2|U(t)-f|}$ in the numerical discretization of (8).

REFERENCES

[1] Tadmor, E., Nezzar, S., and Vese, L., "A multiscale image representation using hierarchical (BV, L^2) decompositions," *Multiscale Model. Simul.* **2**(4), 554–579 (2004).

[2] Alvarez, L., Lions, P.-L., and Morel, J.-M., "Image selective smoothing and edge detection by nonlinear diffusion. II," *SIAM J. Numer. Anal.* **29**(3), 845–866 (1992).

[3] Aubert, G. and Vese, L., "A variational method in image recovery," *SIAM J. Numer. Anal.* **34**(5), 1948–1979 (1997).

[4] Chan, T. and Shen, J., [*Image processing and analysis*], Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA (2005). Variational, PDE, wavelet, and stochastic methods.

[5] Chan, T., Shen, J., and Vese, L., "Variational PDE models in image processing," *Notices Amer. Math. Soc.* **50**(1), 14–26 (2003).

[6] Esedoglu, S. and Osher, S. J., "Decomposition of images by the anisotropic Rudin-Osher-Fatemi model," *Comm. Pure Appl. Math.* **57**(12), 1609–1626 (2004).

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

[8] Gonzales, R. and Woods, R., [*Digital image processing*], Prentice Hall, 3 ed. (2008).

[9] Li, S., [*Markov Random Field Modeling in Computer Vision*], Springer-Verlag (1995).

[10] Mallat, S., [*A wavelet tour of signal processing*], Academic Press Inc., San Diego, CA (1998).

[11] Morozov, V., [*Regularization methods for ill-posed problems*], CRC Press, Boca Raton, FL (1993). Translated from the 1987 Russian original.
[12] Mumford, D. and Shah, J., "Boundary detection by minimizing functionals," in [*in Proceedings of the IEEE Computer Vision Pattern Recognition Conference, San Francisco, CA*], 22–26 (1985).
[13] Perona, P. and Malik, J., "Scale-space and edge detection using anisotropic diffusion," tech. rep., Berkeley, CA, USA (1988).
[14] Rudin, L., Osher, S., and Fatemi, E., "Nonlinear total variation based noise removal algorithms," *Phys. D* **60**, 259–268 (1992).
[15] Tikhonov, A. and Arsenin, V., [*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.
[16] Catté, F., Lions, P.-L., Morel, J.-M., and Coll, T., "Image selective smoothing and edge detection by nonlinear diffusion," *SIAM J. Numer. Anal.* **29**(1), 182–193 (1992).
[17] Mumford, D. and Shah, J., "Optimal approximations by piecewise smooth functions and associated variational problems," *Comm. Pure Appl. Math.* **42**(5), 577–685 (1989).
[18] Ambrosio, L., Fusco, N., and Pallara, D., [*Functions of bounded variation and free discontinuity problems*], Oxford Mathematical monographs, Oxford University Press, New York (2000).
[19] Chambolle, A. and Lions, P.-L., "Image restoration by constrained total variation minimization and variants," *SPIE Electronic Imaging Proceedings* **2567** (1995).
[20] Chambolle, A. and Lions, P.-L., "Image recovery via total variation minimization and related problems," *Numer. Math.* **76**(2), 167–188 (1997).
[21] Meyer, Y., [*Oscillating patterns in image processing and nonlinear evolution equations*], vol. 22 of *University Lecture Series*, American Mathematical Society, Providence, RI (2001). The fifteenth Dean Jacqueline B. Lewis memorial lectures.
[22] Tadmor, E. and Athavale, P., "Multiscale image representation using integro-differential equations," *Inverse problems and imaging* **3**, 693–710 (November 2009).
[23] Tadmor, E., Nezzar, S., and Vese, L., "Multiscale hierarchical decomposition of images with applications to deblurring, denoising and segmentation," *Commun. Math. Sci.* **6**(2), 1–26 (2008).
[24] Chan, T. and Esedoḡlu, S., "Aspects of total variation regularized L^1 function approximation," *SIAM J. Appl. Math.* **65**(5), 1817–1837 (electronic) (2005).
[25] Tadmor, E. and Athavale, P., "Integro-differential equations based on (BV, L^1) minimization," *(under preparation)* .