TOTAL VARIATION BASED IMAGE RESTORATION WITH FREE LOCAL CONSTRAINTS

Leonid I. Rudin and Stanley Osher

Cognitech, Inc.
2800 - 28th St., Suite 101
Santa Monica, CA 90403

ABSTRACT

The two main plagues of image restoration are oscillations and smoothing. Traditional image restoration techniques prevent parasitic oscillations by resorting to smooth regularization. Hence, singular image features and oscillatory textures cannot be restored. The usefulness of images obtained by smooth regularization is very limited. Regularized faces and characters lose features and are unrecognizably distorted. Oscillatory patterns and textures are not allowed by regularization methods. Smoothly regularized images are of no use in criminal/civil investigations and as court evidence. [1].

A new Total Variation based approach was developed in [2] to overcome the basic limitations of all smooth regularization algorithms. The TV-based technique use the L^1 norm of the magnitude of a gradient, thus making discontinuous and nonsmooth solutions possible (if they belong to the space of functions of a bounded total variation). In TV image restoration [3], the solution is obtained by solving a time-dependent, nonlinear PDE on a manifold that satisfies the degradation constraints. In practical applications, one assumes a space-varying blurring kernel and signal-dependent (e.g. multiplicative noise) [4]. The evolution part of the TV-based PDE turned out to be related to the curve shortening equation, but scaled by an inverse $|grad|$.

However, even with the degradation constraints enforced, restoration may lose too much valuable, singular information. Most notably, adjacent features are still merged and all geometrical features (such as level sets and edges) are smoothed out. This is not surprising, since the mean curvature evolution of level sets, even scaled as in the TV method, is essentially linear dissipation in the tangential direction. Again, linearity is at the root of the problem.

Once more, we seek a solution by nonlinearly minimizing oscillations. But this time, rather than focusing on the oscillations of an image intensity (level sets) we bound the oscillation of other quantities along the "feature curves" and enforce constraints inside the "feature regions". The feature curves and regions may be formed by level sets and their interior or by a closed edge and a corresponding region.

1. INTRODUCTION

In a sequence of papers we have investigated the notion of enhancement of images which have been corrupted by noise and blur [2], [3], [4]. The Total Variation of the image is minimized subject to constraints involving the point spread function of the blurring process and the statistics of the noise. The constraints are implemented using Lagrange multipliers. The method is noninvasive and nonoscillatory.

The BV norm allows piecewise smooth functions with jumps and is the proper space for the analysis of discontinuous functions of this type, something well known in the mathematical theory of shock waves. Motivated by this, the first author devised the notion of shock filter in [12], and the two authors refined and applied this notion to the enhancement of slightly blurred images in [13].

The numerical method for performing the minimization, which involves the gradient projection method of Rosen [6], has an interesting geometric interpretation. Minimizing the total variation leads to a numerical step whereby each level set of the intensity function is shifted with velocity equal to its curvature divided by the Euclidean norm of the gradient of the intensity function. Then the result is projected onto the constraint manifold. Recent work by Alvarez, Guichard, Lions and Morel [7], proves rigorously that all morphological operators involve this procedure or a slight variant, (up to scaling by the norm of the gradient). Linear methods, such as those based on minimizing the L^2 norm of the gradient or some higher derivatives as in [8], [9], [10] cannot satisfy the axioms of morphology. The partial differential equations generated by these morphological operators were earlier used to propagate fronts in [11].

In [5], our BV principle is used within the penalty method and a fast multigrid algorithm is introduced.

In this paper we first review our TV based restoration algorithms and then develop an encyclopedia of TV related variational principles which are geometrically based. The new models include minimizing: (1) The absolute integral of the curvature along a feature curve. This allows nonsmooth curves, and (2) Total Variation of the magnitude of the gradient of the image intensity along and across feature curves. This allows for corners and a good feature separation, and (3) Total Variation of curvature along level sets to allow cusps and piece-wise convex interfaces.

Variational problems (1) and (2) yield fourth order nonlinear PDE's for the evolution part of the restoration procedure, while (3) yields a sixth order PDE. The Total Variation based image restoration techniques amount to minimizing the geometrical oscillations of 'features' in a reconstructed image, subject to image degradation constraints.

Since noise may be spatially correlated to the solution, imposing a set of global constraints reduces the resolution of the restoration. Namely, fine features may be grossly altered when the algorithm interprets them to be the result of the statistical variation. A restoration of superior quality is obtained by combining the above nonlinear TV evolution with local Lagrange multipliers, whose steady state solution satisfies constraints that are local to individual curves and regions. Constraints considered so far are:

- (1) Blurring kernel and local noise statistics.
- (2) Divergence free, and hence area-preserving (which assumes that the noisy feature curves oscillate around the proper mean location with prescribed variance).

Surprisingly, even standard TV restoration with Lagrange multipliers, local to segmented regions, partially restores oscillatory textures.

2. TV BASED RESTORATION ALGORITHMS

As in [2], [3] we solve the following problem. Let

$$u_0(x,y) = (Au)(x,y) + \tilde{n}(x,y) \tag{2.2}$$

where $\tilde{n}$ is Gaussian white noise. Also $u_0(x,y)$ is the observed intensity function; while $u(x,y)$ is the image to be restored. Both are defined on a region in R^1. The method is quite general – A needs only to be a compact operator on $L^2(\Omega)$.

Examples we have experimented with include motion, diffraction limited, defocus and Gaussian blur. These are all convolution type integral operators. See [3] for precise definitions.

We also deal with the more complicated problems

$$u_0(x,y) = [(Au)(x,y)][\tilde{n}(x,y)] \tag{2.3}$$

(multiplicative noise/blurring, model one)

$$u_0(x,y) = (Au)(x,y) + (u(x,y))(\tilde{n}(x,y)) \tag{2.4}$$

(multiplicative noise/blurring, model two)

Model one has been treated in [14], for example, using homomorphic filtering, i.e., basically taking the log of u_0 and treating it as a problem involving additive noise, then filtering and applying the exponential. This is not appropriate for model two.

We present below our restoration algorithms for (2.3) and (2.4). Namely, we shall discuss what the proper constraints are, apply the gradient projection method and demonstrate the results on real images.

Our constrained minimization problem involves the variation of the image, which is a direct measure of how oscillatory it is. The space of functions of bounded variation is appropriate for discontinuous functions. This is well known in the field of shock calculations – see e.g. [13] and the references therein.

Thus, our constrained minimization problem is:

$$\text{minimize} \int_\Omega \sqrt{u_x^2 + u_y^2}\,dxdy \tag{2.5}$$

subject to constraints involving u_0.

Our theoretical results involve the following two constraints involving the mean:

$$\int_\Omega u\,dxdy = \int_\Omega u_0\,dxdy \tag{2.6}$$

and the variance

$$\int_\Omega (Au - u_0)^2\,dxdy = \sigma^2. \tag{2.7}$$

Of course the x and y derivatives in (2.5) above and throughout are to be numerically approximated by letting the distance between pixels go to zero, in a standard numerical analysis fashion.

Here (2.6) indicates that the white noise $\tilde{n}(x,y)$ is of zero mean and (2.7) uses á priori information that the standard deviation of the noise $\tilde{n}(x,y)$ is σ.

Following our usual procedure [2], [3], [4] we arrive at the Euler-Lagrange equations

$$\begin{aligned} 0 \; = \; & \frac{\partial}{\partial x}\left(\frac{u_x}{\sqrt{u_x^2+u_y^2}}\right) + \frac{\partial}{\partial y}\left(\frac{u_y}{\sqrt{u_x^2+u_y^2}}\right) \\ & -\lambda(A^*Au - A^*u) \;\text{ in }\; \Omega \end{aligned} \tag{2.8}$$

with

$$\frac{\partial u}{\partial n} = 0 \;\text{ on }\; \partial\Omega. \tag{2.9}$$

Here A^* is just the adjoint integral operator.

The quantity λ is a Lagrange multiplier chosen so that the constraint (2.7) is satisfied. The first constraint (2.6) is automatically satisfied (see [3], [4]) if

$$\int Au \equiv \int u \;\text{ and }\; \int A^*u \equiv \int u \tag{2.10}$$

for each u . This is true up to normalization for a convolution. Thus, we assume (2.10) for simplicity only.

We shall use the gradient-projection method of Rosen [6], which, in this case, becomes the interesting "constrained" partial differential equation

$$\begin{aligned} u_t \; = \; & \frac{\partial}{\partial x}\left(\frac{u_x}{\sqrt{u_x^2+u_y^2}}\right) + \left(\frac{u_y}{\sqrt{u_x^2+u_y^2}}\right) \\ & -\lambda A^*(Au - u_0) \end{aligned} \tag{2.11}$$

for $t > 0$, $(x,y) \in \Omega$

$$\frac{\partial u}{\partial n} = 0 \;\text{ on }\; \partial\Omega \tag{2.12}$$

and $u(x,y,0)$ is given such that (2.6, 2.7) are satisfied. If (2.6) is satisfied initially, e.g. $u(x,y,0) = u_0(x,y)$, then, by conservation form of (2.11, 2.12) and by (2.10), it is always satisfied. Satisfying (2.7) can be done through a process defined in [4] from a practical point of view.

The projection step is the gradient projection method which just amounts to updating $\lambda(t)$ so that (2.7) remains true in time. This follows (see [3]) if we define

$$\lambda(t) = \frac{\int_\Omega \left(\frac{\partial}{\partial x}\left(\frac{u_x}{\sqrt{u_x^2+u_y^2}}\right) + \frac{\partial}{\partial y}\left(\frac{u_y}{\sqrt{u_x^2+u_y^2}}\right)\right)\cdot A^*(Au - u_0)\,dxdy}{\int_\Omega (A^*(Au - u_0))^2\,dxdy}. \tag{2.13}$$

We thus have a dynamic procedure for restoring the image. As $t \to \infty$, the steady state solution is the desired restoration.

Next we consider restoration involving multiplicative noise [4]. We begin by considering pure denoising. We are given an image $u_0(x, y)$ where

$$u_0 = u\eta. \tag{2.14}$$

The unknown function $u(x, y)$ is the image we wish to restore, and $\eta(x, y)$ is the noise (which we take here to be Gaussian white noise).

We are given the following information

$$\begin{array}{ll} \int \eta = 1 & \text{(mean one)} \\ \int (\eta - 1)^2 = \sigma^2 & \text{(given variance)} \end{array} \tag{2.15}$$

Thus our constrained optimization algorithm is (2.5) subject to the following constraints

$$\int \frac{u_0}{u} = 1 \tag{2.16}$$

$$\frac{1}{2}\int \left(\frac{u_0}{u} - 1\right)^2 = \frac{\sigma^2}{2} = \frac{1}{2}\int \left(\left(\frac{u_0}{u}\right)^2 - 1\right) \tag{2.17}$$

The gradient projection method leads to

$$u_t = \frac{\partial}{\partial x}\left(\frac{u_x}{\sqrt{u_x^2 + u_y^2}}\right) + \frac{\partial}{\partial y}\left(\frac{u_y}{\sqrt{u_x^2 + u_y^2}}\right) \tag{2.18}$$
$$-\lambda \frac{u_0^2}{u^3} + \mu \frac{u_0}{u^2}.$$

We now have two Lagrange multipliers which we compute by requiring

$$0 = \frac{\partial}{\partial t}\int \frac{u_0}{u} = -\int \frac{u_0}{u^2} u_t = 0 \tag{2.19}$$

$$0 = \frac{\partial}{\partial t}\frac{1}{2}\int \left(\left(\frac{u_0}{u}\right)^2 - 1\right) = -\int \frac{u_0}{u^3} u_t = 0 \tag{2.20}$$

This together with (2.18) leads us to two algebraic equations for the two unknowns and the resulting Gram determinant is nonzero.

In the following experiments we merely took $u(x, y, 0) = u_0(x, y)$ so (2.16) was satisfied while (2.17) was slightly incorrect (the right hand side was always zero). Nevertheless, the results appear to be excellent.

Figure 1 compares performance of three commonly used linear techniques vs. TV based multiplicative denoising. To be able to apply linear algorithms, we assume that the noise is additive with variance $\sigma\bar{u}$, where $\bar{u}$ is the mean of the true image u. The noisy image is generated by the model (2.15).

Next we consider images which are both blurry and noisy. Our first model is as follows. We are given an image $u_0(x, y)$ for which

$$u_0 = (Au)\eta. \quad \text{(model 1)} \tag{2.21}$$

Here we take A to be convolution with a Gaussian kernel, but the generality is as great as in section II above. The noise η is as above (2.15) and (2.15) are still satisfied.

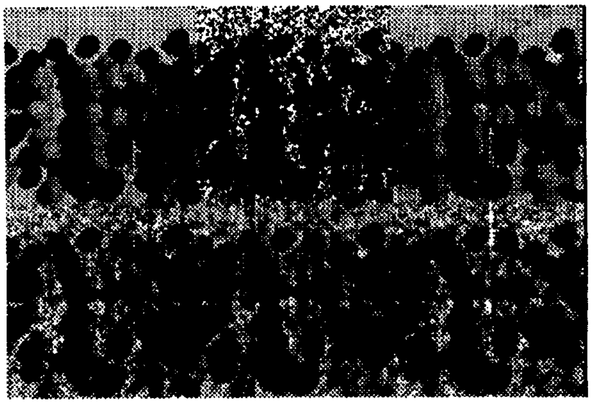

Figure 1: Top Row, Left to Right: Original, Multplicative Noise with σ=0.3, TV Restoration. Lower Row, Left to Right: Ideal Wiener Filter Restoration, Constrained Least-Square Restoration, Geometrical Mean Filter Restoration.

Our constraints are

$$\int \frac{u_0}{Au} = 1 \tag{2.22}$$

$$\frac{1}{2}\int \left(\frac{u_0}{Au} - 1\right)^2 = \frac{\sigma^2}{2} = \frac{1}{2}\int \left(\left(\frac{u_0}{Au}\right)^2 - 1\right) \tag{2.23}$$

The gradient projection method leads to

$$u_t = \frac{\partial}{\partial x}\left(\frac{u_x}{\sqrt{u_x^2 + u_y^2}}\right) + \frac{\partial}{\partial y}\left(\frac{u_y}{\sqrt{u_x^2 + u_y^2}}\right) \tag{2.24}$$
$$- \lambda A^*\left(\frac{u_0}{(Au)^2}\right)\left(\frac{u_0}{(Au)^2} - 1\right) + \mu A^*\left(\frac{u_0}{(Au)^2}\right).$$

Then two Lagrange multipliers are chosen as follows. We take as initial data $u(x, y, 0) = u_0(x, y)$. However we initially set $\lambda = 0$ and choose μ so that

$$\frac{\partial}{\partial t}\int \left(\frac{u_0}{Au}\right) = 0. \tag{2.25}$$

We monitor the quantity in (2.23), and let time evolve until the second constraint is satisfied. We then allow λ to be nonzero, continue to enforce (2.25), and set

$$\frac{\partial}{\partial t}\frac{1}{2}\left(\left(\frac{u_0}{Au}\right) - 1\right)^2 = 0. \tag{2.26}$$

This leads to a solution for μ and λ.

Our second model is as follows. We are given an image $u_0(x, y)$ for which

$$u_0 = Au + u\eta. \quad \text{(model 2)} \tag{2.27}$$

Again we chose a Gaussian blur, but the techniques are general. The noise is Gaussian white noise with mean zero and variance σ^2

$$\int \left(\frac{u_0 - Au}{u}\right) = 0 \tag{2.28}$$

$$\frac{1}{2}\int \left(\frac{u_0 - Au}{u}\right)^2 = \frac{\sigma^2}{2}. \tag{2.29}$$

Figure 2: Left: Image from Security Camera. Right: Restored image, clearly showing suspect wearing baseball cap.

The gradient projection method leads us to

$$\begin{aligned} u_t &= \frac{\partial}{\partial x}\left(\frac{u_x}{\sqrt{u_x^2+u_y^2}}\right)+\frac{\partial}{\partial y}\left(\frac{u_y}{\sqrt{u_x^2+u_y^2}}\right) \\ &- \lambda\left[A^*\left(\frac{u_0-Au}{u^2}\right)+\left(\frac{(u_0-Au)^2}{u^3}\right)\right] \quad (2.30) \\ &+ \mu\left[A^*\left(\frac{1}{u}\right)+\left(\frac{u_0-Au}{u^2}\right)\right]. \end{aligned}$$

Figure 2 is an application of restoration model 2.21 to an actual criminal justice defense case.

3. LOCALIZING THE CONSTRAINTS

TV based image restoration was developed by the authors in [2], [3], [4] to fix the inability of existing regularization algorithms to recover discontinuities. The price of using a TV variational principle is the weakening of oscillatory patterns in images. However, these oscillations can also be recovered.

In [15], the first author suggested localizing the constraints in the TV restoration procedure. This was done using Cognitech's fast version of the Morel-Koepfler multiscale segmentation [16]. This segmentation algorithm has only one free parameter (scale) and is quite robust.

An iterative procedure was developed whereby extremely local and even oscillatory features (textures) could be recovered. Figure 3 shows the recovery of the oscillatory texture from a significantly degraded image. Figure 4 compares edge detection after restoration using several methods. Combining TV with local constraints partially resolves the well known problem in image processing, namely, providing "regularization" that preserves both edges and oscillatory patterns.

4. HIGHER ORDER TV VARIATIONAL METHODS

An interesting quantity to be minimized involves the variation of the angle that the tangent to any level curve of the intensity makes around each level curve. This rapidly

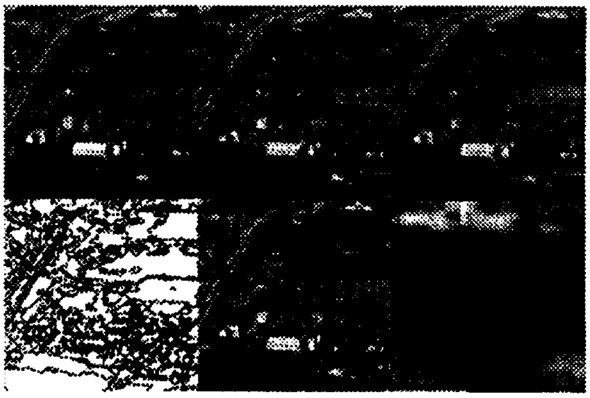

Figure 3: Top Row: Original, Noisy, TV Denoised. Bottom Row: Intermediate Segmentation, Denoised by TV with local constraints (note the appearance of the oscillatory "field texture"

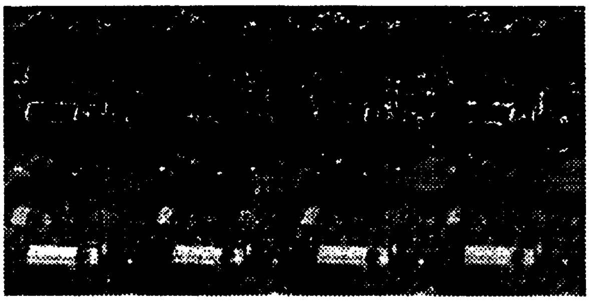

Figure 4: Bottom Row, Left to Right: Original, Multiplicative Noise Degradation, Localized TV Restoration, Constrained Least Square Restoration. Top Row, Left to Right: Edge Detection of image below

transforms into minimizing the L^1 norm of curvature

$$\text{Minimize } \int_\Omega |\nabla \cdot \left(\frac{\nabla u}{|\nabla u|}\right)| dxdy. \tag{4.31}$$

The effect of this is to "convexify" all level curves.

Another important quantity to be minimized is the variation of the logarithm of the magnitude of the gradient along the level curves. This reduces to

$$\text{Minimize } \int | \left(\frac{u_x}{|\nabla u|}\frac{\partial}{\partial y} - \frac{u_y}{|\nabla u|}\frac{\partial}{\partial x}\right) \log |\nabla u|| dxdy. \tag{4.32}$$

It is straightforward, but cumbersome to compute the Euler-Lagrange derivatives of these quantities. The main observation is that the gradient descent equations involve fourth order operators and have some interesting properties - any type of maximum principle is ruled out, in contrast with curvature dependent motion [7]. Equations (4.31) and (4.32) jointly minimize TV of the complex $\log(\nabla u)$. Observe that solutions to this variational problem admit triple points and nonsmooth interfaces. An even higher order PDE is derived by

$$\text{Minimize } \int | \left(\frac{u_x}{|\nabla u|}\frac{\partial}{\partial y} - \frac{u_y}{|\nabla u|}\frac{\partial}{\partial x}\right) \nabla \cdot \left(\frac{\nabla u}{|\nabla u|}\right) | dxdy. \tag{4.33}$$

In (4.33), the Total Variation of the level set curvature is minimized along all level sets. This would allow for cusps and other curvature singularities along internal interfaces.

5. INCOMPRESSIBLE MOTION

A variant of our algorithms comes from the realization that the gradient descent approach can be interpreted as moving each level set of u via a certain normal velocity. As mentioned above, the TV restoration corresponds to normal velocity which is curvature of level sets divided by the magnitude of the gradient. Borrowing from fluid dynamics, we may project the velocity vector onto its divergence free component. This projection is defined by

$$\Pi = I - \nabla(\Delta^{-1})\nabla \cdot \tag{5.34}$$

where ∇ is gradient, $\nabla \cdot$ is divergence, I is the identity, and Δ^{-1} is the inverse Laplacian, with zero Neumann data. The latter can be easily obtained via an FFT.

Thus, for incompressible TV denoising (without constraints), we solve

$$u_t = \left(\Pi\left(\left(\nabla \cdot \frac{\nabla u}{|\nabla u|}\right)\frac{1}{|\nabla u|}\right)\right) \cdot \nabla u. \tag{5.35}$$

The main effect is to preserve the integral of all functions of u over any interior region as the level set moves, yet still diminish the total variation.

In a forthcoming paper [17], the authors will demonstrate finite difference schemes for (4.31) - (5.35).

6. REFERENCES

[1] L. Rudin, S. Jensen, *Feature-Accurate Image/Video Processing Techniques for Criminal Investigation*, SPIE Proc. on Automatic Systems for the Identification and Inspection of Humans, 1994, San Diego, CA.

[2] L. Rudin, S. Osher, and E. Fatemi, *Nonlinear Total Variation Based Noise Removal Algorithms*, Physica D., Vol. 60, (1992), pp. 259-268.

[3] L. Rudin, S. Osher, and C. Fu, *Total Variation Based Restoration of Noisy, Blurred Images*, SINUM, to appear.

[4] P.-L. Lions, S. Osher, and L.I. Rudin, *Denoising and Deblurring Images with Constrained Nonlinear Partial Differential Equations*, submitted to SINUM.

[5] C.R. Vogel, M.E. Oman, *Iterative Methods for Total Variation Denoising*, preprint.

[6] J.G. Rosen, *The Gradient Projection Method for Nonlinear Programming, Part II, Nonlinear Constraints*, J. Soc. Indust. Appl. Math., Vol. 9, (1961), pp. 514-532.

[7] L. Alvarez, F. Guichard, P.-L. Lions, and J.-M. Morel, *Axioms and Fundamental Equation of Image Processing*, Archive Ration. Mech. and Analysis, **123**, (1993), pp. 199-258.

[8] B.R. Hunt, *The Application of Constrained Least Squares Estimation to Image Restoration by Digital Computer*, IEEE Trans. Comput., Vol. C-22, (1973), pp. 805-812.

[9] S. Twomey, *On the Numerical Solution of Fredholm Integral Equations of the First Kind by the Inversion of the Linear System Produced by Quadrature*, J. ACM, Vol. 10, (1963), pp. 97-107.

[10] S. Twomey, *The Application of Numerical Filtering to the Solution of Integral Equations Encountered in Indirect Sensing Measurements*, J. Franklin Inst., Vol. 297, (1965), pp. 95-109.

[11] S. Osher and J.A. Sethian, *Fronts Propagating with Curvature Dependent Speed. Algorithms Based on a Hamilton-Jacobi Formulation*, J. Comp. Phys , Vol. 79, (1988), pp. 12-49.

[12] L. Rudin, *Images, Numerical Analysis of Singularities, and Shock Filters*, Caltech, Comp. Sc. Dept., Report #TR:5250:87 (1987).

[13] S. Osher and L.I. Rudin, *Feature-Oriented Image Enhancement Using Shock Filters*, SIAM J. Num. Analysis, Vol. 27, (1990), pp. 919-940.

[14] E.R. Cole, *The Removal of Unknown Image Blurs by Homomorphic Filters*, Computer Science Report UTEC-CSC-74-2029, (1973), University of Utah.

[15] L. Rudin, *Segmentation and Restoration Using Local Constraints*, Cognitech Report #36, (1993).

[16] G. Koepfler, C. Lopez, and J.M. Morel, *A Multiscale Algorithm for Image Segmentation by Variational Methods*, SINUM, Vol. 31, (1994), pp. 282-299.

[17] L. Rudin, S. Osher, *Image Restoration with Higher Order TV-based Methods*, Cognitech Report #41(5/1994).