IMAGE SEGMENTATION AND TOTAL VARIATION DENOISING

Erkut Erdem[*]

Hacettepe University

March 5th, 2012

CONTENTS

1 Mumford-Shah (MS) Functional 1

2 Ambrosio-Tortorelli (AT) Approximation of the MS Functional 2

3 A Common Framework for Curve Evolution, Segmentation and Anisotropic Diffusion 6

4 Total Variation (TV) Regularization 7

References 10

1 MUMFORD-SHAH (MS) FUNCTIONAL

The formulation of Mumford and Shah [8] is based on a functional minimization via which a piecewise smooth approximation of a given image and an edge set are to be recovered simultaneously. In this unified formulation, smoothing and edge detection processes work jointly to partition an image into segments. The Mumford-Shah (MS) model is:

$$E_{MS}(u,\Gamma) = \beta \int_{\Omega} (u-f)^2 dx + \alpha \int_{\Omega\setminus\Gamma} |\nabla u|^2 dx + length(\Gamma) \tag{1}$$

where

- $\Omega \subset \mathbb{R}^2$ is connected, bounded, open subset representing the image domain,
- f is an image defined on Ω,
- $\Gamma \subset \Omega$ is the edge set segmenting Ω,
- u is the piecewise smooth approximation of f,
- α, $\beta > 0$ are the scale space parameters of the model.

[*]erkut@cs.hacettepe.edu.tr

The first term in E_{MS} is the *data fidelity* term which forces u to be close to the original image f. The next two terms are the generic priors that provide certain knowledge about the solution. Specifically, the second term, the so-called *regularization* or *smoothness* term, gives preference to piecewise smooth images by penalizing high gradients. Since the integral is over $\Omega\backslash\Gamma$, this prior is turned off at image boundaries, and thus it excludes image edges to be smoothed out. The third term is a penalty term on total edge length which prevents the image to be split into many regions. Additionally, it implicitly imposes smoothness of the boundaries.

Generally, the unknown edge set Γ of a lower dimension makes the minimization of the MS model very difficult. Hence, in literature several approaches for approximating the MS model are suggested [1, 5]. In the next section, the approximation proposed by Ambrosio and Tortorelli [1] will be reviewed.

2 AMBROSIO-TORTORELLI (AT) APPROXIMATION OF THE MS FUNCTIONAL

Ambrosio and Tortorelli [1] suggested an approximation for the MS model by following the Γ convergence framework [3]. The basic idea is to introduce a smooth edge indicator function v which is more convenient than using the characteristic function χ_Γ as the edge indicator. The new function v depends on a parameter ρ, and as $\rho \to 0$, $v \to 1 - \chi_\Gamma$. That is, $v(x) \approx 0$ if $x \in \Gamma$ and $v(x) \approx 1$ otherwise. The result is the functional

$$E_{AT}(u,v) = \int_\Omega \left(\beta(u-f)^2 + \alpha(v^2|\nabla u|^2) + \frac{1}{2}\left(\rho|\nabla v|^2 + \frac{(1-v)^2}{\rho} \right) \right) dx \,. \tag{2}$$

In the Ambrosio-Tortorelli (AT) model, the continuous function v encodes the boundary information. The value of v at a point can be interpreted as a measure of *boundaryness* where the low values indicate the edge points. That is, $v \approx 0$ along the boundaries and grows rapidly away from them. Thus, the function v may be thought as a blurred version of the edge set. The parameter ρ specifies the level of blurring (Figure 1), and as $\rho \to 0$, $\frac{1}{2}\int_\Omega \left(\rho|\nabla v|^2 + \frac{(1-v)^2}{\rho} \right) dx$ approximates the cardinality of the edge set Γ.

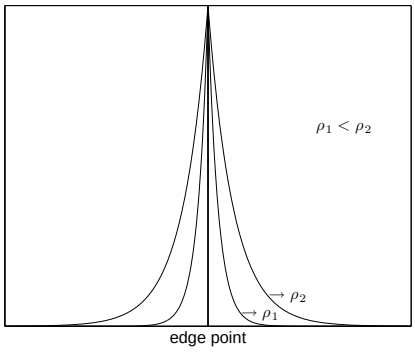


Figure 1: An example 1D edge strength function $(1 - v)$ for two different values of ρ.

Piecewise smooth image u and the edge strength function v are simultaneously computed via the solution of the following system of coupled PDEs:

$$\frac{\partial u}{\partial t} = \nabla \cdot (v^2 \nabla u) - \frac{\beta}{\alpha}(u - f); \quad \left.\frac{\partial u}{\partial n}\right|_{\partial\Omega} = 0 \,, \tag{3}$$

$$\frac{\partial v}{\partial t} = \nabla^2 v - \frac{2\alpha|\nabla u|^2 v}{\rho} - \frac{(v-1)}{\rho^2}; \quad \left.\frac{\partial v}{\partial n}\right|_{\partial\Omega} = 0 \tag{4}$$

where $\partial\Omega$ denotes the boundary of Ω and n denotes the outer unit normal vector to $\partial\Omega$.

Notice that PDE for each variable can be interpreted as a biased diffusion equation that minimizes a convex quadratic functional in which the other variable is kept fixed:

Keeping v fixed, Equation (3) minimizes a convex quadratic functional given by

$$\int_\Omega \Big(\alpha v^2 |\nabla u|^2 + \beta (u-f)^2\Big) dx \,. \tag{5}$$

The data fidelity term in Equation (3) provides a bias that forces u to be close to the original image f. In the regularization term, the edge strength function v specifies the boundary points and guides the smoothing accordingly. Since $v \approx 0$ along the boundaries, no smoothing is carried out at the boundary points, thus the edges are preserved.

Keeping u fixed, Equation (4) minimizes a convex quadratic functional given by

$$\frac{\rho}{2}\int_\Omega \left(|\nabla v|^2 + \frac{1 + 2\alpha\rho|\nabla u|^2}{\rho^2} \left(v - \frac{1}{1 + 2\alpha\rho|\nabla u|^2} \right)^2 \right) dx \,. \tag{6}$$

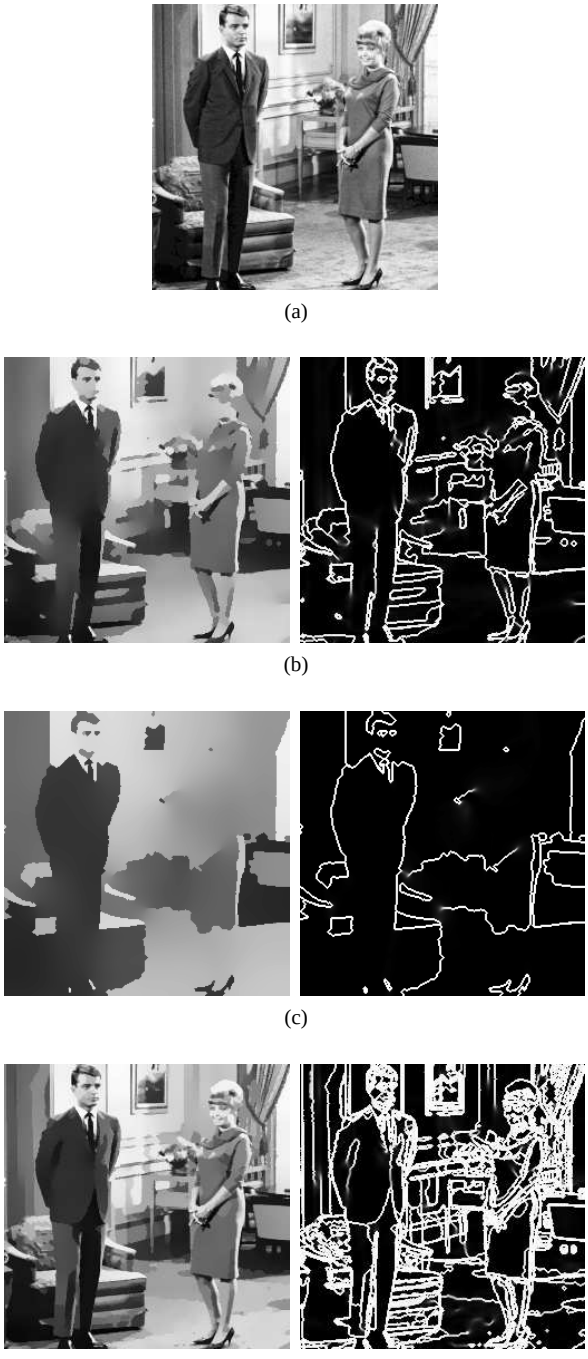
The reciprocal relationship between v and $|\nabla u|^2$ can be clearly observed in the above functional. It asserts that the function v is nothing but a smoothing of $\frac{1}{1+2\alpha\rho|\nabla u|^2}$ with a blurring radius proportional to ρ and reciprocal to $|\nabla u|$. Ignoring the smoothness term $\rho|\nabla v|^2$, which mildly forces some spatial organization by requiring the edges to be smooth, and by letting $\rho \to 0$ [4, 10], $v \approx \frac{1}{1+2\alpha\rho|\nabla u|^2}$.

Considering this approximation and the relation between variational regularization and diffusion equations, Equation (3) can be interpreted as a Perona-Malik type nonlinear diffusion at a specific scale. Replacing v in Equation (3) with $1/(1 + 2\alpha\rho|\nabla u|^2)$ yields

$$\frac{\partial u}{\partial t} = \nabla \cdot (g(|\nabla u|)\nabla u) - \frac{\beta}{\alpha}(u - f) \tag{7}$$

where $g(|\nabla u|) = \left(\frac{1}{1+|\nabla u|^2/\lambda^2}\right)^2$ with $\lambda^2 = 1/(2\alpha\rho)$. Thus, $\sqrt{1/(2\alpha\rho)}$ can be seen as a contrast parameter.

Since the parameters α and β define the relative importance of the regularization term, the scale is determined by the ratio α/β. Keeping the value of α fixed, decreasing the value of β results in more simplified results (Figure 2(b)-(c)). Moreover, the scale space parameters α and β also define the detail level in segmentation. With the ratio α/β fixed, the level of detail increases with the increasing α (Figure 2(b)-(d)).



(a)

(b)

(c)

(d)

Figure 2: Example segmentation results (u and $1-v$). (a) Source image. (b)-(d) Corresponding segmentations obtained with the parameters (b) $\alpha = 1$, $\beta = 0.01$, $\rho = 0.01$, (c) $\alpha = 1$, $\beta = 0.001$, $\rho = 0.01$, and (d) $\alpha = 4$, $\beta = 0.04$, $\rho = 0.01$, respectively.

Numerical Implementation

Equations (3) and (4) can be simultaneously solved for u and v using standard numerical discretization techniques such as finite differences. The coupled system is first discretized with respect to spatial variables. This leads to the following space-discrete system of equations:

$$\begin{aligned}
\frac{du_{i,j}}{dt} &= v^2_{i+\frac{1}{2},j}\cdot\left(u_{i+1,j}-u_{i,j}\right) - v^2_{i-\frac{1}{2},j}\cdot\left(u_{i,j}-u_{i-1,j}\right) \\
&+ v^2_{i,j+\frac{1}{2}}\cdot\left(u_{i,j+1}-u_{i,j}\right) - v^2_{i,j-\frac{1}{2}}\cdot\left(u_{i,j}-u_{i,j-1}\right) \\
&- \frac{\beta}{\alpha}\left(u_{i,j}-f_{i,j}\right)\,, \qquad (8)\\
\frac{dv_{i,j}}{dt} &= v_{i+1,j}+v_{i-1,j}+v_{i,j+1}+v_{i,j-1}-4v_{i,j} \\
&- \frac{2\alpha|\nabla u_{i,j}|^2 v_{i,j}}{\rho} - \frac{\left(v_{i,j}-1\right)}{\rho^2}\,. \qquad (9)
\end{aligned}$$

As in the discretization of the Perona-Malik equation, the diffusivities represented by the edge strength function v at mid-pixel points can be computed by taking averages over neighboring pixels:

$$v_{i\pm\frac{1}{2},j} = \frac{v_{i\pm1,j}+v_{i,j}}{2}, \qquad v_{i,j\pm\frac{1}{2}} = \frac{v_{i,j\pm1}+v_{i,j}}{2}\,. \qquad (10)$$

The time derivatives in Equations (8) and (9) can be discretized using forward differences, where regularization terms and the bias terms on the right hand side of each equation are evaluated at times k and $k+1$, respectively.

$$\begin{aligned}
\frac{u^{k+1}_{i,j}-u^k_{i,j}}{\Delta t} &= \left(v^k_{i+\frac{1}{2},j}\right)^2\cdot u^k_{i+1,j} + \left(v^k_{i-\frac{1}{2},j}\right)^2\cdot u^k_{i-1,j} \\
&+ \left(v^k_{i,j+\frac{1}{2}}\right)^2\cdot u^k_{i,j+1} + \left(v^k_{i,j-\frac{1}{2}}\right)^2\cdot u^k_{i,j-1} \\
&- \left(\left(v^k_{i+\frac{1}{2},j}\right)^2 + \left(v^k_{i-\frac{1}{2},j}\right)^2 + \left(v^k_{i,j+\frac{1}{2}}\right)^2 + \left(v^k_{i,j-\frac{1}{2}}\right)^2\right)\cdot u^k_{i,j} \\
&- \frac{\beta}{\alpha}\left(u^{k+1}_{i,j}-f_{i,j}\right), \qquad (11)\\
\frac{v^{k+1}_{i,j}-v^k_{i,j}}{\Delta t} &= v^k_{i+1,j}+v^k_{i-1,j}+v^k_{i,j+1}+v^k_{i,j-1}-4v^k_{i,j} \\
&- \frac{\alpha\left(\left(u^k_{i+1,j}-u^k_{i-1,j}\right)^2 + \left(u^k_{i,j+1}-u^k_{i,j-1}\right)^2\right)v^{k+1}_{i,j}}{2\rho} - \frac{\left(v^{k+1}_{i,j}-1\right)}{\rho^2} \qquad (12)
\end{aligned}$$

where Δt denotes the time step.

Although the suggested scheme is neither fully explicit nor fully implicit, it still allows us to compute u^{k+1} and v^{k+1} by using forward recursion as in an explicit scheme. A numerical stopping criteria for the iterative scheme can be defined in the sense that the rate of change of u is less than a threshold (Algorithm 1).

Algorithm 1 Minimization of the Ambrosio-Tortorelli Model

1: Initialize the variables with $u^0 = f$, $v^0 = \frac{1}{1+2\alpha\rho|\nabla u^0|^2}$
2: **for** $k = 0$ to $kmax$ **do**
3: Solve Equation (11) for u^{k+1}
4: **if** $|u^{k+1} - u^k| < \epsilon|u^k|$ **then**
5: stop iterations
6: **end if**
7: Solve Equation (12) for v^{k+1}
8: **end for**

3 A COMMON FRAMEWORK FOR CURVE EVOLUTION, SEGMENTATION AND ANISOTROPIC DIFFUSION

The key idea of the Ambrosio-Tortorelli approximation of the Mumford-Shah functional (Section 2) is to utilize a continuous edge strength function v. The value of v approaches to 0 at the object boundaries and grows rapidly as image gradients become small. In [10], Shah suggested a modification to the Ambrosio-Tortorelli model Equation (2), where the quadratic cost functions in both the data fidelity and the smoothing terms are replaced with L^1-functions. The modified energy is:

$$(13)\quad E_S(u,v) = \int_\Omega \left(\beta\,|u-f| + \alpha\, v^2|\nabla u| + \frac{1}{2}\left(\rho|\nabla v|^2 + \frac{(1-v)^2}{\rho}\right)\right) dx\ .$$

As $\rho \to 0$, this energy functional converges to the following functional:

$$(14)\quad E_{S2}(u,\Gamma) = \frac{\beta}{\alpha}\int_\Omega |u-f|dx + \int_{\Omega\setminus\Gamma} |\nabla u|dx + \int_\Gamma \frac{J_u}{1+\alpha J_u}ds$$

with $J_u = |u^+ - u^-|$ indicating the jump in u across Γ. u^+ and u^- denote intensity values on two sides of Γ, respectively, and thus each boundary point is weighted according to its level of contrast.

Minimizing E_S corresponds to the gradient descent equations:

$$(15)\ \frac{\partial u}{\partial t} = 2\nabla v\cdot\nabla u + v|\nabla u|\ curv(u) - \frac{\beta}{\alpha v}|\nabla u|\frac{(u-f)}{|u-f|};\quad \left.\frac{\partial u}{\partial n}\right|_{\partial\Omega} = 0\ ,$$

$$(16)\ \frac{\partial v}{\partial t} = \nabla^2 v - \frac{2\alpha|\nabla u|v}{\rho} - \frac{(v-1)}{\rho^2};\quad \left.\frac{\partial v}{\partial n}\right|_{\partial\Omega} = 0$$

with $curv(u) = \nabla\cdot\left(\frac{\nabla u}{|\nabla u|}\right)$.

The Equation (16) is very similar to the evolution equation of v Equation (4) in the AT model; only $|\nabla u|^2$ is replaced with $|\nabla u|$. The determining factor of the model is the new evolution equation of u Equation (15). Replacing L^2-norms in both the data fidelity and the smoothness terms by their L^1-norms generates shocks in u and thus object boundaries are recovered as actual discontinuities. As it can be clearly seen from Figure 3, the suggested smoothing process of u gives rise to more cartoon-like, piecewise constant images (these results are obtained by using a half-quadratic approx-

imation of Shah's modified energy proposed in [7]). However, the robust norms utilized attract the image towards the cartoon limit and catch unintuitive regions such as the one at the man's right shoulder and the ones on the floor. It is important to remark that the effect of the scale space parameters α and β on segmentation results is similar compared to the one in AT model (cf. Figure 2). The amount of smoothing is determined by the ratio α/β, and increasing the value of α while keeping α/β fixed leads to more detailed segmentations.

One of the underlying assumptions of the original MS model and AT approximation is that the filtered image varies from the observed image by Gaussian noise. Hence, when a source image is corrupted by impulse noise, the corresponding smoothing process produces inadequate results. However, replacing the L^2-norm with the L^1 in the modified model yields to a robust data fidelity term that can cope with impulse noise. For example, consider the noisy image given in Figure 4(a), which is degraded with 5% salt and pepper noise. Figure 4(b) and (c) depict the outcomes of the AT approximation and the modified model, respectively. As they demonstrate, the modified model eliminates the impulse noise during smoothing, however, noise still present in the AT result.

4 Total Variation (TV) Regularization

Rudin et al. [9] formulated image restoration as minimization of the total variation (TV) of a given image under certain assumptions on the noise. The Total Variation (TV) regularization model is generally defined as:

$$E_{TV}(u) = \int_{\Omega} \left(\frac{1}{2}(u-f)^2 + \alpha|\nabla u| \right) dx \tag{17}$$

where

- $\Omega \subset \mathbb{R}^2$ is connected, bounded, open subset representing the image domain,
- f is an image defined on Ω,
- u is the restored version of g,
- $\alpha > 0$ is a scalar.

The gradient descent equation for Equation (17) is defined by

$$\frac{\partial u}{\partial t} = \nabla \cdot \left(\frac{\nabla u}{|\nabla u|} \right) - \frac{1}{\alpha}(u-f); \quad \left.\frac{\partial u}{\partial n}\right|_{\partial\Omega} = 0\,. \tag{18}$$

Since the value of α specifies the relative importance of the fidelity term, it can be interpreted as a scale parameter that determines the level of smoothing. In Figure 5, the role of this parameter is illustrated. The restored images are given in Figure 5(b)-(d). As it can be clearly seen, level of smoothing increases with increasing α.

In the original formulation of Rudin et al., the observed image f was assumed to be degraded by additive Gaussian noise with zero mean and known variance σ^2. Hence, in order to restore a given image, they propose to solve the constrained optimization

(a)

(b)

(c)

(d)

Figure 3: Example segmentation results (u and $1-v$). (a) Source image. (b)-(d) Corresponding segmentations obtained with parameters (b) $\alpha = 1$, $\beta = 0.01$, $\rho = 0.01$, (c) $\alpha = 1$, $\beta = 0.001$, $\rho = 0.01$, and (d) $\alpha = 4$, $\beta = 0.04$, $\rho = 0.01$.

(a) (b) (c)

Figure 4: Segmentation of a noisy image degraded with 5% salt and pepper noise. (a) Source image. (b) Reconstruction using AT model. (c) Reconstruction using Shah's modified functional (both results are obtained with $\alpha = 1$, $\beta = 0.01$, $\rho = 0.01$).

problem

$$\min_u \int_\Omega |\nabla u| dx \tag{19}$$

subject to

$$\int_\Omega (u - f)^2 dx = \sigma^2 \,. \tag{20}$$

When TV regularization is defined as a constrained optimization problem, $\frac{1}{\alpha}$ can be considered as a Lagrange multiplier, which has to be determined by taking the given constraint into account.

TV regularization can be associated with a nonlinear diffusion filter, the so-called *TV flow* [2, 6]. Ignoring the fidelity term in Equation (18) leads to the PDE

$$\frac{\partial u}{\partial t} = \nabla \cdot (g(|\nabla u|)\nabla u) \tag{21}$$

with $u^0 = f$ and the diffusivity function $g(|\nabla u|) = \frac{1}{|\nabla u|}$. Notice that this diffusivity function has no additional contrast parameter as compared with the Perona-Malik diffusivities. Figure 6 and Figure 7 depict TV scale space examples sampled at different diffusion times for two different images. It is evident from these images that the corresponding smoothing process yields segmentation-like, piecewise constant images.

Numerical Implementation

The evolution equation of u (18) can be discretized by using standard finite differences. An important point is that the solution of TV regularization or equivalently TV flow leads to singular diffusivities as shown in Equation (21). In numerical implementations based on standard discretization, this leads to stability problems as the image gradient tends to zero. A common solution to this problem is to add a small positive constant ϵ to image gradients.

After ϵ-regularization, the space-discrete version of Equation (18) can be written as:

$$\begin{aligned}
\frac{\partial u}{\partial t} &= \frac{\partial}{\partial x}\left(\frac{u_x}{\sqrt{|\nabla u|^2+\epsilon^2}}\right)+\frac{\partial}{\partial y}\left(\frac{u_y}{\sqrt{|\nabla u|^2+\epsilon^2}}\right)-\frac{1}{\alpha}(u-f) \\
&= \frac{u_{xx}\left(u_y^2+\epsilon^2\right)-2u_xu_yu_{xy}+u_{yy}\left(u_x^2+\epsilon^2\right)}{\left(u_x^2+u_y^2+\epsilon^2\right)^{\frac{3}{2}}}-\frac{1}{\alpha}(u-f)\,, \\
\frac{du_{i,j}}{dt} &= \frac{\frac{d^2u_{i,j}}{dx^2}\left(\left(\frac{du_{i,j}}{dy}\right)^2+\epsilon^2\right)-2\left(\frac{du_{i,j}}{dx}\right)\left(\frac{du_{i,j}}{dy}\right)\left(\frac{d^2u_{i,j}}{dxdy}\right)+\frac{d^2u_{i,j}}{dy^2}\left(\left(\frac{du_{i,j}}{dx}\right)^2+\epsilon^2\right)}{\left(\left(\frac{du_{i,j}}{dx}\right)^2+\left(\frac{du_{i,j}}{dy}\right)^2+\epsilon^2\right)^{\frac{3}{2}}} \\
&- \frac{1}{\alpha}\left(u_{i,j}-f_{i,j}\right)
\end{aligned} \tag{22}$$

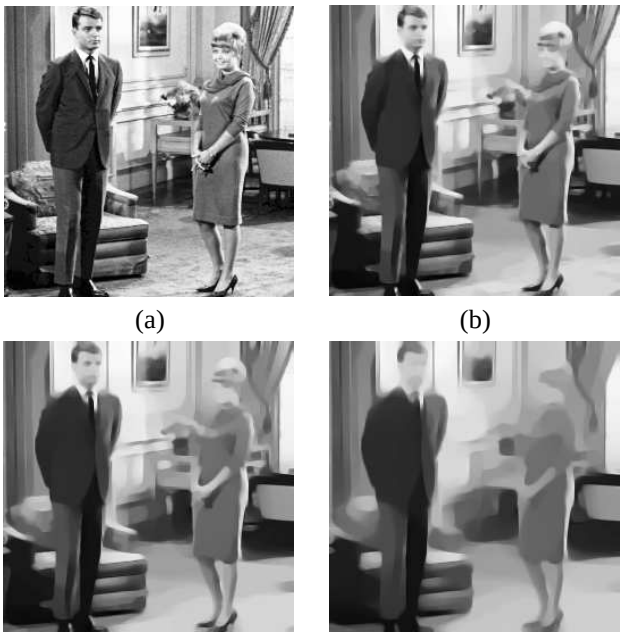
where

$$\frac{d^2u_{i,j}}{dxdy}\approx\frac{u_{i+1,j+1}-u_{i+1,j-1}-u_{i-1,j+1}+u_{i-1,j-1}}{4}\,. \tag{23}$$

When explicit time discretization is employed, numerical solution to Equation (22) can be computed by the following iterative scheme, where homogeneous Neumann boundary condition is imposed along the image boundary:

$$\begin{aligned}
\frac{u_{i,j}^{k+1}-u_{i,j}^k}{\Delta t} &= \left(\left(\frac{u_{i+1,j}^k-u_{i-1,j}^k}{2}\right)^2+\left(\frac{u_{i,j+1}^k-u_{i,j-1}^k}{2}\right)^2+\epsilon^2\right)^{-\frac{3}{2}} \\
&\cdot\Bigg[\left(u_{i+1,j}^k-2u_{i,j}^k+u_{i-1,j}^k\right)\left(\left(\frac{u_{i,j+1}^k-u_{i,j-1}^k}{2}\right)^2+\epsilon^2\right) \\
&- \frac{1}{8}\left(u_{i+1,j}^k-u_{i-1,j}^k\right)\left(u_{i,j+1}^k-u_{i,j-1}^k\right) \\
&\quad\left(u_{i+1,j+1}^k-u_{i+1,j-1}^k-u_{i-1,j+1}^k+u_{i-1,j-1}^k\right) \\
&+ \left(u_{i,j+1}^k-2u_{i,j}^k+u_{i,j-1}^k\right)\left(\left(\frac{u_{i+1,j}^k-u_{i-1,j}^k}{2}\right)^2+\epsilon^2\right)\Bigg] \\
&- \frac{1}{\alpha}\left(u_{i,j}^k-f_{i,j}\right)
\end{aligned} \tag{24}$$

where Δt denotes the time step. ϵ-regularization requires the stability condition $\Delta t \leq 0.25\epsilon$, and thus a sufficiently large number of iterations is needed for small values of ϵ.



(a) (b)

Figure 5: Example TV restoration results. (a) Source image. (b)-(d) Corresponding segmentations obtained with (b) $\alpha = 50$, (c) $\alpha = 100$, and (d) $\alpha = 200$.

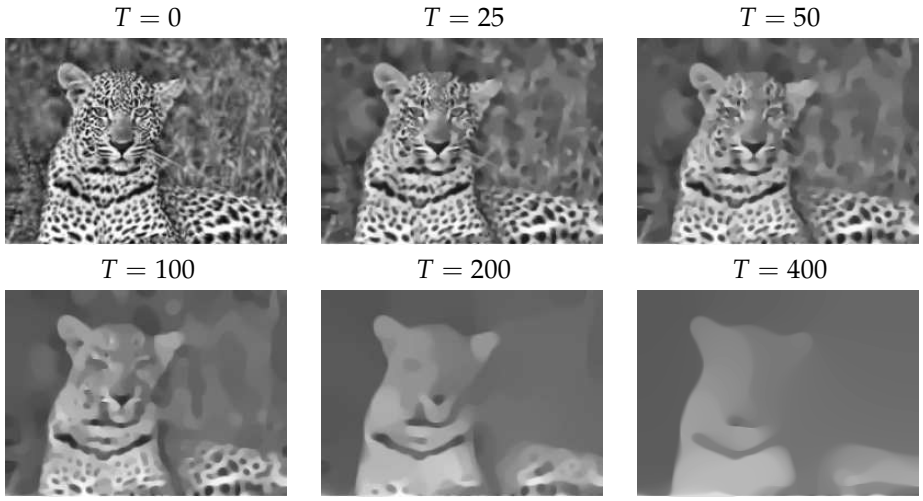


Figure 6: TV flow results for different diffusion times.

REFERENCES

[1] L. Ambrosio and V. Tortorelli. On the approximation of functionals depending on jumps by elliptic functionals via Γ-convergence. *Commun. Pure Appl. Math.*, 43(8):999–1036, 1990.

[2] F. Andreu, C. Ballester, V. Caselles, and J. M. Mazón. Minimizing total variation

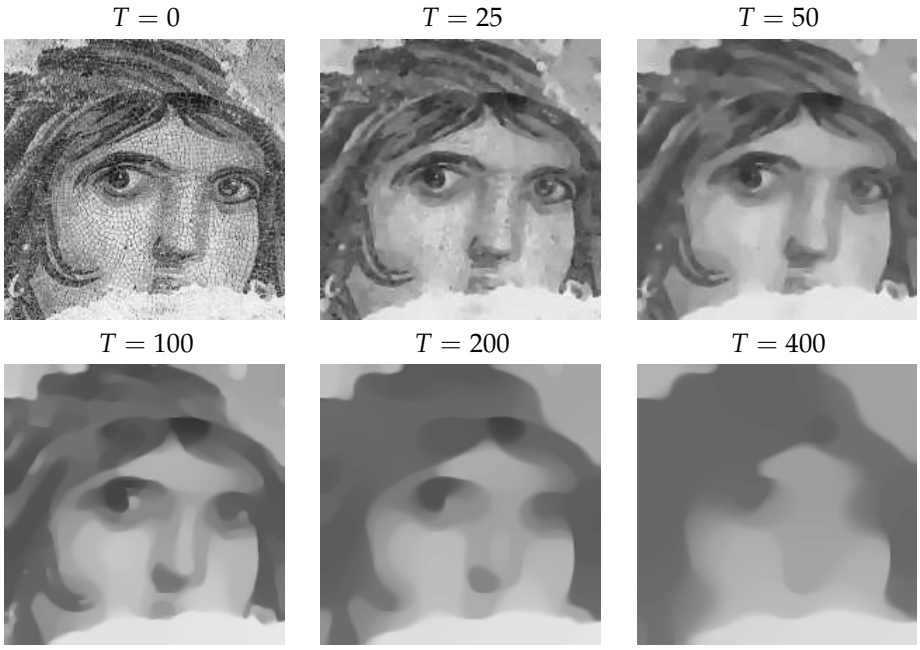


Figure 7: TV flow results for different diffusion times.

flow. *Differential and Integral Equations*, 14(3):321–360, 2001.

[3] A. Braides. *Approximation of Free-Discontinuity Problems.* Lecture Notes in Mathematics, Vol. 1694. Springer-Verlag, 1998.

[4] T. Chan and L. Vese. Variational image restoration and segmentation models and approximations. *UCLA CAM-report 97-47*, September 1997.

[5] T. F. Chan and L. A. Vese. Active contours without edges. *IEEE Trans. Image Processing*, 10(2):266–277, 2001.

[6] F. Dibos and G. Koepfler. Global total variation minimization. *SIAM J. Numer. Anal.*, 37(2):646–664, 2000.

[7] J. P. Kaufhold. *Energy Formulations of Medical Image Segmentations.* PhD thesis, Boston University College of Engineering, Department of Biomedical Engineering, Boston, MA, 2000.

[8] D. Mumford and J. Shah. Optimal approximations by piecewise smooth functions and associated variational problems. *Commun. Pure Appl. Math.*, 42(5):577–685, 1989.

[9] L. Rudin, S. Osher, and E. Fatemi. Nonlinear total variation based noise removal algorithms. *Phys. D.*, 60:259–268, 1992.

[10] J. Shah. A common framework for curve evolution, segmentation and anisotropic diffusion. In *CVPR*, pages 136–142, 1996.