USES OF ELLIPTIC APPROXIMATIONS IN COMPUTER VISION

Jayant Shah
Mathematics Department
Northeastern University, Boston, Mass. 02115

1. INTRODUCTION

One of the problems in Computer Vision is recovery of object shapes from noisy images. Associated with this problem is the question of what is a shape and how is it to be represented. Since answers to these questions have to be ultimately tailored to the uses one has in mind, one has to bring into consideration potential applications and with it, the question of practical algorithms for implementation of the theory. Here we are concerned with mainly two-dimensional shapes. Mathematically, an object is simply an open subset in the image domain, characterized in some way. In the real world, what is an object and what is just noise or clutter depends of course on what one is looking for. For example, Figure 1a shows a noiseless synthetic image. It may be reasonable to assume that the objects in the figure are the four squares in the four corners and the two ellipses in the middle. Figure 1b shows a noisy version obtained from the image in Figure 1a by adding Gaussian noise. The signal-to-noise ratio (i.e. the ratio between the standard deviation of the image with noise removed and the standard deviation of the noise) is 1:4. The problem is to recover the original objects.

FIGURE 1a: Noise-free Image

FIGURE 1b: Noisy Version

This research was partially supported under ARO Grant, No. DAAL03–91–G-0041.

The general approach is the construction of functionals which incorporate constraints imposed by various objectives such as noise suppression, boundary detection, shape description and object matching. The motivation for building such integrated models comes from a long history of attempts to isolate, formulate and solve simpler problems in Computer Vision and the realization that no matter how sophisticated a technique is, there are inherent ambiguities which cannot be resolved in isolation. It is essential to incorporate contextual information in the model. Moreover, these methods are ideal for formulating models based on raw intensity images. By incorporating the raw image in the model, one can ensure that the input to higher level tasks such as object recognition reflects realistic ambiguities and noise, and thus ensure robustness of the solution. At the same time, integration provides further constraints which may help resolve ambiguities in the image. The disadvantage of such an approach is that as more and more realistic representations are incorporated in these functionals, they become more and more difficult to analyze and implement. In the absence of fast stochastic algorithms, an alternative approach is to construct approximations of these complex functionals, even deliberately weakening the coupling among their components and attempt to find stationary solutions by gradient descent. The purpose of this paper is to describe some numerical experiments involving such approximations.

2. SEGMENTATION PROBLEM

2.1 Segmentation Functional

Traditionally, the first step in image understanding is the detection of edges. The usual method is to smooth the image to reduce noise and then apply some kind of local edge operator such as the zero-crossings of the laplacian. Geman and Geman [GeGe] proposed a Bayesian formulation for simultaneous smoothing and boundary detection. At about the same time, Blake and Zisserman [BZ_1,BZ_2] proposed an analogous discrete model. An analytic version of these models is formulated in [MS_1] as follows:

$$E(u,B) = \frac{1}{\sigma^2}\int_R (u-g)^2 dx dy + \int_{R\setminus B} \|\nabla u\|^2 dx dy + \nu|B| \tag{1}$$

where

R is the image domain, (an open subset in $\mathbf{R}^2$);
g is the feature intensity; $g : R \rightarrow \mathbf{R}$;
B is the union of segment boundaries; thus B is the segmenting curve;
$|B|$ is the length of B;
σ and ν are the weights. (σ may be thought of as the smoothing radius in $R\setminus B$.)

The task is to find u and B which minimize $E(u,B)$. Thus the segmentation problem is viewed as the problem of finding a piecewise smooth approximation of g with minimal amount of discontinuity. The weights σ and ν control the relative degrees to which smoothness of u and its discontinuity locus are emphasized. Smoothing is explicitly prevented from extending over B. An object in this formulation is characterized by having relatively uniform feature intensity. Many important theoretical results concerning existence and regularity have been obtained for the special case where R is 2–dimensional and the general case of n-dimensional R, (see for example, [Am],[DMS],[DCL],[MS_2],[Ri],[Sh_4],[Wa]). In particular, in [MS_2] it is shown that the minimal length constraint implies that there cannot be any points where more than three objects meet and an object cannot have corners except at points where three objects meet. At such corners, the angle is always 120°. Thus the formulation implies a very special kind of shape representation and it may be necessary to alter or augment it in order to represent shapes more realistically.

2.2 Dirichlet Version

A simple way to preserve the actual shape singularities is to impose Dirichlet boundary conditions on u in functional (1), namely, $u^{\pm} = g^{\pm}$ along B where the superscript $\pm$ denote the values on the two sides of B. Results (mostly for the case of 1–dimensional R) concerning the existence and approximations are proved in [Sh_6]. It is shown that in dimension 1 and 2, singularities of u are then a subset of those of g; thus distortion of corners and junctions is prevented. The limitation of this modified model is that if the actual boundaries are noisy (i.e. locally the graph of a noisy function), they will not be smoothed. Moreover, if the boundaries in the image are blurred as is likely in practical situations so that g is continuous, then B would be empty. Hence the formulation must always be used with an explicit representation for blurring. Such a representation based on Ambrosio-Tortorelli approximation (see functional (6) described below) is formulated and analyzed in [Sh_6]. Note that blurring also provides smoothing of noisy boundaries. Recovery of unblurred boundaries from their blurred version is treated as a separate problem.

2.3 Extensions

An obvious extension for enriching shape representation in functional (1) is to incorporate information about the curvature of the boundary. One such extension is proposed in [NM] where $|B|$ is replaced by

$$\int_B \left(\nu_0 + \nu_1 \kappa^2\right) ds \tag{2}$$

where κ denotes the curvature. Very few theoretical results concerning the new functional have been obtained. Results for simple closed curves based on elastica for functional (2) under the assumption that $\nu_0 = 0$ and that the length of the curve is fixed have been obtained in [M,W]. In [BDP], lower semi-continuity of functional (2) is analyzed and certain pathologies that arise are exhibited. For instance, even for a simple heart shape, minimizing sequences produce regions of zero width in the limit. Practical difficulties in implementing such a functional are considerable and as yet, no one has implemented it. Allowing B to be only piecewise continuous or introduction of more shape information such as illusory contours or 3D interpretations makes the implementation even more difficult.

2.4 Elliptic Approximation

What makes the task of minimization of functional (1) very difficult is the presence of the one-dimensional segmenting curve. Ambrosio and Tortorelli [AT_1,AT_2] proposed an elliptic approximation of functional (1) so that gradient descent may be applied. Their method is as follows. Let

$$\Lambda_\rho(v) = \frac{1}{2}\int_R \left\{\rho\|\nabla v\|^2 + \frac{v^2}{\rho}\right\}dxdy \tag{3}$$

For a fixed curve B, if v_ρ minimizes $\Lambda_\rho(v)$ with the boundary condition $v = 1$ on B, then as $\rho \to 0$, v_ρ obviously tends to zero everywhere except on B where it equals one. The key observation is that $\Lambda_\rho(v_\rho) \to |B|$ as $\rho \to 0$. Values of v_ρ range between 0 and 1 and v_ρ may be viewed as blurring of B with ρ as the nominal blurring radius. The method of Ambrosio and Tortorelli is to replace the term $|B|$ in functional (1) by $\Lambda_\rho(v)$. B is now spread out over all of R and the integral in functional (1) must be modified as well since we no longer have the set $R\backslash B$. A simple choice is to

$$\text{replace} \int_{R\backslash B} \|\nabla u\|^2 dxdy \text{ by } \int (1-v)^2\|\nabla u\|^2 dxdy \tag{4}$$

The final form of the approximate functional is

$$E_\rho(u,v) = \int_R \left\{\frac{1}{\sigma^2}(u-g)^2 + (1-v)^2\|\nabla u\|^2 + \frac{\nu}{2}\left(\rho\|\nabla v\|^2 + \frac{v^2}{\rho}\right)\right\}dxdy \tag{5}$$

Ambrosio and Tortorelli prove that $E_\rho(u,v)$ converges to $E(u,B)$ in the sense of Γ–convergence. The corresponding approximate functional when Dirichlet boundary conditions are imposed is as follows [Sh6]:

$$\begin{aligned} E_{M,\rho}(u,v) = &\int_R \left[\frac{1}{\sigma^2}(u-g)^2 + (1-v)^2\|\nabla u\|^2\right]dxdy \\ &+ \int_R \left[\left\{\frac{\nu}{2} + M(u-g)^2\right\}\left\{\left(\rho\|\nabla v\|^2 + \frac{v^2}{\rho}\right)\right\}\right]dxdy \end{aligned} \tag{6}$$

where M is of order $O|\log\rho|$. When n=1, the Γ–convergence of $E_{M,\rho}(u,v)$ to $E(u,B)$ satisfying the boundary conditions $u^\pm = g^\pm$ along B is proved in [Sh6] under certain regularity conditions on g.

It should be noted that an approximation theory for functional (1) when terms containing curvature and singularities of B are introduced has yet not been found.

Applying gradient descent to (5), we get the following coupled diffusion equations:

$$\begin{aligned} \text{Smoothing: } &\frac{\partial u}{\partial t} = \nabla\cdot(1-v)^2\nabla u - \frac{1}{\sigma^2}(u-g); \quad \frac{\partial u}{\partial n}|_{\partial R} = 0 \\ \text{Boundary detection: } &\frac{\partial v}{\partial t} = \nabla\cdot\nabla v - \frac{v}{\rho^2} + \frac{2}{\nu\rho}(1-v)\|\nabla u\|^2; \quad \frac{\partial v}{\partial n}|_{\partial R} = 0 \end{aligned} \tag{7}$$

where ∂R denotes the boundary of R and n denotes the direction normal to ∂R.

The results of Ambrosio-Tortorelli diffusion (7) when applied to the example shown in Figure 1b with $\sigma = \rho = 8$ pixels are shown in Figure 2. The image in Figure 1b was represented on a 256×256 square lattice. The thin ellipse has the maximum width of 9 pixels and the distance between the two ellipses is 18 pixels. Note that long thin objects are very difficult to locate in noisy images because the smoothing radius is governed by the noise characteristics and thus may be much larger than the width of the object. Figure 2a shows the smoothed image u and Figure 2b depicts the edge strength function v. The lighter the area in the figure, the higher the value of v.

FIGURE 2a: Smoothed Image u

FIGURE 2b: Edge-strength Function v

3. SHAPE RECOVERY

The use of approximate segmentation functionals discussed above for constructing practical algorithms leads to a new difficulty, namely, the difficulty of recovering the actual boundaries from the edge-strength function v. It is apparent from Figure 2b that thresholding of v will not produce satisfactory representation of the boundaries. The trouble is that due to differing levels of contrast along the boundaries and the surrounding noise, values of v along the boundary are not constant. That is, the level curves of v are only approximately parallel to the object boundaries; in fact, they might even be perpendicular to the object boundary in places. Consequently, global thresholding of v produces representation of the boundaries by narrow strips of varying width which may not completely enclose the object if the contrast becomes sufficiently weak. Schemes for adaptive local thresholding of v are discussed in [Sh_3]. The point is that the local threshold for v should depend on the average value v in the neighborhood. What this means is that we should look at the laplacian of a smoothed version of v. This approach to local thresholding of v does produce improvement, but such further processing was found to produce further displacement of the boundary from its actual location, indicating the need for a better method for recovering the actual boundary from v.

Shape Recovery by Curve Evolution

The basic idea is borrowed from the work of Kass, Witkin and Terzopoulos [KWT] on **SNAKES**. In their framework, one introduces a simple closed curve in the image and lets it deform under forces which are determined by distance from the object boundary and by a smoothness constaint. In the framework presented here, the analogous idea [Sh_7] is to let an initial curve evolve so as to (locally) maximize the edge-strength function v along the curve in some sense. Let Γ be a simple closed curve in R. In order to move Γ to where the image intensity gradient and hence v are high, we look for the stationary points of the functional

$$L = \int_{\Gamma} (1 - v)^{\alpha} d\gamma \tag{8}$$

where γ denotes the arc-length along Γ. The evolution equation for Γ is derived by applying gradient descent to L. Let $C(p,t) : I \times [0, \infty) \rightarrow R$ be the evolving family of curves where I is the unit interval and t denotes time. We require that $C(0,t) = C(1,t)$ for all values of t and that the image of $C(p,0)$ in R coincides with Γ. Then the evolution of Γ is governed by the equation

$$\frac{\partial C}{\partial t} = [\alpha \nabla v \cdot N - (1 - v)\kappa] N \tag{9}$$

where N is the outward normal and κ is the curvature which is defined such that it is positive when Γ is a circle. Thus the points on the evolving curve move in the direction of the normal with velocity which has two components: the component depending on curvature which imposes a smoothness constraint and advection induced by v. The curve is pulled towards the object boundary by the force field ∇v induced by v. The exponent α in the expression for L serves as a weighting factor. The higher the value of α, the weaker the smoothing constraint.

The implementation of curve evolution is no longer a straightforward matter of using finite differences. Moving the points on an evolving curve directly by discretizing the curve leads to many difficulties. For example, the chosen points might bunch up causing numerical instabilities. The curve may also undergo topological changes. An alternate method is the one proposed by Osher and Sethian [OS]. In their approach, the initial curve is embedded in a surface as a level curve and then evolution is applied to the surface so that all of its level curves evolve simultaneously. Assume that Γ is embedded in a surface $f_0 : R \rightarrow \mathbf{R}$ as a level curve. Let $f(t,x,y)$ denote the evolving surface such that $f(0,x,y) = f_0(x,y)$. Then, in order to let all the level curves of f_0 evolve simultaneously, we consider the functional

$$G(f) = \int_{-\infty}^{\infty} \int_{\Gamma_c} (1 - v)^{\alpha} d\gamma_c dc \tag{10}$$

where $\Gamma_c = \{(x,y)|f(t,x,y) = c\}$. But by the coarea formula[2],

$$G(f) = \int\int_R (1-v)^\alpha \|\nabla f\| dxdy \tag{11}$$

By calculating the first variation of the last functional, we get the gradient descent equation as

$$\begin{aligned} \frac{\partial f}{\partial t} &= -\alpha \nabla v \cdot \nabla f + (1-v)\|\nabla f\| \nabla \cdot \left(\frac{\nabla f}{\|\nabla f\|} \right) \\ &= -\alpha \nabla v \cdot \nabla f + (1-v) \frac{f_y^2 f_{xx} - 2 f_x f_y f_{xy} + f_x^2 f_{yy}}{\|\nabla f\|^2} \\ \frac{\partial f}{\partial n}|_{\partial R} &= 0; \ f = f_0 \text{ at } t = 0. \end{aligned} \tag{12}$$

An important question now is: How to specify the intial curve Γ? An automatic specification of the initial curve is a difficult problem because its solution implies that the object boundaries have already been found, at least approximately. The strategy used in [Sh_7] for specifying the initial curve is based on the following considerations. If we set $v = 1$ and $f_0 = g$ in equation (12), then it reduces to the diffusion equation proposed in [ALM] for smoothing images. In [ALM], diffusion is allowed to occur only in the direction of the level curves, the expectation being that in this way, the object boundaries will be smoothed without being blurred. The difficulty with this approach is of course that the object boundaries in general are not level curves in an image. Traditional use of the zero-crossings of the laplacian of smoothed images to detect edges may be thought of as an attempt to remedy this situation by representing object boundaries by the level curves of the smoothed laplacian instead of the level curves of the image itself. Hence, the strategy in [Sh_7] is to use the zero-crossings of the laplacian as the initial approximation Γ for the object boundaries. A fundamental difficulty with the use of level curves as object boundaries is that level curves cannot represent triple points which are very important features in image analysis. This problem is not considered here. Another difficulty is that the theory of curve evolution is based on the assumption that Γ is a simple closed curve, an assumption which is violated at the saddle points of f. If the laplacian is too noisy, the evolution is dominated by what happens at its saddle points and becomes unpredictable. This is because at a saddle point, the first term on the right hand side of the evolution equation (12) vanishes and both the numerator and the denominator vanish in the second term. Hence, the behaviour of the second term becomes very sensitive to noise near a saddle point. In order to demonstrate the feasibility of this approach, an ad hoc solution to this problem has been adopted in [Sh_7]. The Ambrosio-Tortorelli solution u is further smoothed by non-uniform smoothing until the zero-crossings of the laplacian of the smoothed version have relatively few self-intesections (see [Sh_7] for more details). Then we set f_0 equal to the laplacian of the smoothed u and let it evolve according to equation (12). The zero-crossings of the evolving f are taken as the successive refinements of the object boundaries.

For numerical implementation of equation (12), the usual finite-difference schemes (say, central differences) are exactly the wrong thing to apply and must never be used. The main point of

[2] The functional $G(f)$ has been extended to a segmentation functional analogous to functional (1), see [Sh_8

Osher and Sethian is that since we expect the evolving surface to develop discontinuities or 'shocks' where the object boundaries are, the directions of the finite differences must chosen adaptively. Once this principle is incorporated in the numerical scheme, nonlinear diffusion (12) behaves very robustly.

Figures 3 and 4 show the results of evolution when applied to the example shown in Figure 1b. Figure 3 shows the zero-crossings of the evolving laplacian at $t = 0, 40, 80, 160, 320, 640, 1280$. Thus, the zero-crossings at $t = 0$ are the zero-crossings of the laplacian of of a smoothed version of u. The example illustrates that the boundaries found by this method are at least metastable, that is, they persist for a long time. Figure 4 shows the superposition of the zero–crossings smoothed by evolution ($t = 1280)$ on the noisefree image (Figure 4a) as well as on the smoothed image u (Figure 4b). Smoothed image u was used for superposition rather than g because the evolution is governed by v which, in turn, is determined by the boundaries in u. Note the accuracy of placement of the smoothed zero-crossings on the noiseless image, including the thin ellipse. The corners are also fairly well represented. The boundary deviates in places from the boundary in the noiseless image when it follows some accidental feature introduced by the noise. (It is possible to discern these features by close inspection of Figure 2a.) It is interesting that although the four corner squares are identical in the noiseless image, the final boundaries are different in all four cases because the accidental noise features are different. The worst deviation occurs in the case of the lower right corner square. The very thin ends of the thin ellipse are lost because the boundary representation v is too coarse ($\rho = 8$ pixels). Some portions towards the ends of the thin ellipse are lost because they are obscured by noise as one can see in Figure 2a.

FIGURE 3: EVOLUTION OF ZERO_CROSSINGS

FIGURE 4a **FIGURE 4b**

4. DIFFUSION SYSTEMS

Use of curve evolution for shape recovery as described above exemplifies an approach for getting around the difficulties in finding the minimizing solutions for functionals in Computer Vision. The basic ingredient was to incorporate the edge-strength function v in the algorithms designed to solve the next level of tasks in Computer Vision. If these tasks are also governed by evolution equations, then the idea is to let this evolution be controlled by v. In fully integrated models, one would try to construct a coupled system of diffusion equations in which each equation controls one particular process or a task like smoothing or edge detection or matching and outputs of these equations provide controls for each other's evolution. The diffusion system (7) of Ambrosio and Tortorelli may be viewed as a system of this type. Other examples are nonlinear smoothing [Sh_3,PPGO], stereo [Sh_5] and optical flow [PGO]. The difficulty in designing nonlinear systems of diffusion equations for problems in Computer Vision is that such systems are one of the most difficult mathematical objects to analyze and no general mathematical principles are available to guide the design. The examples cited above are based on purely heuristic arguments. For illustration, the case of stereo is described below in some detail.

A diffusion system for the stereo problem

In stereo vision, the basic problem is the correspondence problem, that is, the problem of matching the two views seen by the two eyes (or cameras) and thus computing the disparity between the two images. Let R_l and R_r be the left and right image planes respectively, assumed to be coplanar. Let $I_l : R_l \to \mathbf{R}$ and $I_r : R_r \to \mathbf{R}$ be the respective image intensities. A major complication in matching the two images is the presence of discontinuities. There are sudden jumps in the disparity at the boundaries between objects which are at differing distances from the eyes. Another consequence of this is the phenomenon of half-occlusions which are the areas in the scene seen by one eye and not the other. Thus it is necessary to prevent matching of the two images in areas of half-occlusions. In the one-dimensional case, these considerations lead to the following functional for the map $f_l : R_l \to R_r$ matching the left image with the right image:

$$\begin{aligned} E_l(f_l) = & \int\limits_{R_\ell \setminus (B_l \cup O_l)} \left\{ (f_l')^2 + \left(\frac{1}{f_l'}\right)^2 + \frac{4}{\sigma^2}[I_l - I_r \circ f_l]^2 \right\} dx_l \\ & + 4 \sum_{x_l \in B_l} \left(\nu + \gamma \left| f_l^+(x_l) - f_l^-(x_l) \right| \right) \\ & + 4\nu(number\ of\ components\ of\ O_l) + 4\,\gamma\, length(O_l) \end{aligned} \tag{13}$$

where $B_\ell = \{x_l : f_l'(x_l) = \infty\}$ and occlusion $O_l = \{x_l : f_l'(x_l) = 0\}$

A similar functional may be obtained for the map $f_r : R_r \to R_l$ matching the right image with the left image. Maps f_l and f_r are inverses of each other. In particular, there is a correspondence between the components of B_l and the components of O_r. Similarly, there is a correspondence between the components of B_r and the components of O_l. Complexity of functional (13) forces us to look for approximate solutions. For this purpose, orient coordinates x_l in R_l and x_r in R_r from left to right and define the left disparity d_l by the equation

$$f_l = x_l + d_l \tag{14}$$

Then,

$$(f_l')^2 + \left(\frac{1}{f_l'}\right)^2 \approx 2 + 4(d_l')^2 \text{ if } |d_\ell'| << 1 \tag{15}$$

Define a functional $\hat{E}_l(d_l)$ over R_l as follows:

$$\begin{aligned} \hat{E}_l(d_l) = & \int\limits_{R_\ell \setminus (B_l \cup O_l)} \left\{ (d_l')^2 + \frac{1}{\sigma^2}[I_l - I_r \circ f_l]^2 \right\} dx_l \\ & + \sum_{x_l \in B_l} \left(\nu + \gamma \left| f_l^+(x_l) - f_l^-(x_l) \right| \right) \\ & + \nu(number\ of\ components\ of\ O_l) + \ \gamma\, length(O_l) \end{aligned} \tag{16}$$

Similarly, define d_r and $\hat{E}_r(d_r)$ after orienting x_l and x_r from right to left. Then

$$\hat{E}(d) \approx C + 4\hat{E}_l(d_l) \approx C + 4\hat{E}_r(d_r) \tag{17}$$

where C is a constant. Hence we consider the functionals $\hat{E}_l(d_l)$ and $\hat{E}_r(d_r)$ in order to find approximate solutions. Except for the presence of O_l and O_r, functionals $\hat{E}_l$ and $\hat{E}_r$ are very similar in form to the functionals used in segmentation problems. Since $B_l \cap O_l$ and $B_r \cap O_r$ are empty, a strategy to find approximate solutions would be as follows: Apply gradient descent alternately to $\hat{E}_l$ and $\hat{E}_r$. That is, alternately minimize $\hat{E}_l$ with respect to d_l and B_l keeping O_l fixed and minimize $\hat{E}_r$ with respect to d_r and B_r keeping O_r fixed. Determination of the discontinuity locus B_l when $\hat{E}_l$ is minimized also determines the occlusion O_r in R_r since $O_r = R_r \backslash$(range of f_l). Similarly the discontinuity locus B_r determines the occlusion O_l in R_l.

We make three further simplifications before implementing this strategy. First, we follow the example of segmentation functionals and simplify the penalty term for the discontinuity locus by setting $\gamma = 0$. Next, we drop the constraint that f_l and f_r have the same graph. This is based on the heuristic that minimization of $\hat{E}_l$ and $\hat{E}_r$ without this constraint will still produce comparable optimal values for disparities. Finally we simplify the constraint that f_l and f_r must extend continuously across the occlusion sets O_l and O_r respectively such that they are constant over the occlusion set. We relax this requirement by merely extending d_l and d_r across the occlusion set and including the cost of $|d_l'|^2$ and $|d_r'|^2$ over the occlusion set in their respective functionals.

The 2–dimensional generalization is straight-forward. The final result is a pair of coupled functionals, E_l and E_r defined over R_l and R_r respectively as follows:

$$
\begin{aligned}
E_l(d_l) &= \int\int_{R_l \backslash B_l} \|\nabla d_l\|^2 + \frac{1}{\sigma^2}\int\int_{R_l \backslash O_l} \left\{[I_l - I_r \circ f_l]^2\right\} + \nu|B_l| \\
E_r(d_r) &= \int\int_{R_r \backslash B_r} \|\nabla d_r\|^2 + \frac{1}{\sigma^2}\int\int_{R_r \backslash O_r} \left\{[I_r - I_l \circ f_r]^2\right\} + \nu|B_r| \\
&\text{where } O_l = R_l \backslash (\text{range of } f_r) \text{ and } O_r = R_r \backslash (\text{range of } f_l)
\end{aligned}
\tag{18}
$$

For a practical implementation by gradient descent, we now mimick the method of Ambrosio and Tortorelli and approximate the discontinuity loci B_l and B_r by continuous functions w_l and w_r. To calculate w_l corresponding to given d_l, apply gradient descent to the functional

$$
E_D(w_l) = \int\int \{\rho\|\nabla w_l\|^2 + \frac{w_l^2}{\rho} + 2(1-w_l)^2\|\nabla f_\ell\|^2\} \tag{19}
$$

To calculate d_l for given w_l, f_r and w_r, apply gradient descent to the functional

$$
E_M(d_l) = \int\int \left\{(1-w_l)^2\|\nabla d_l\|^2 + \frac{1}{\sigma^2}(I_l - I_r \circ f_l)^2\left(1 - w_r \circ f_r^{-1}\right)^2\right\} \tag{20}
$$

Similarly, define $E_D(w_r)$ and $E_M(d_r)$ to calculate w_r and d_r. The result is a coupled system of 4 diffusion equations as follows:
With x_l oriented from left to right, equations for d_l and w_l are

$$
\begin{aligned}
\frac{\partial d_l}{\partial t} =& \nabla \cdot \left((1-w_l)^2 \nabla d_l\right) \\
&+ \frac{1}{\sigma^2}[I_l - I_r \circ f_l]\left(\frac{\partial I_r}{\partial x_r} \circ f_l\right)\left(1 - w_r \circ f_r^{-1}\right)^2 \\
\frac{\partial w_l}{\partial t} =& \rho\nabla^2 w_l - \frac{w_l}{\rho} + \frac{2}{\nu}(1-w_l)\|\nabla f_l\|^2
\end{aligned}
\tag{21}
$$

With x_r oriented from right to left, equations for d_r and w_r are

$$
\begin{aligned}
\frac{\partial d_r}{\partial t} =& \nabla \cdot \Big((1-w_r)^2 \nabla d_r \Big) \\
& + \frac{1}{\sigma^2} [I_r - I_l \circ f_r] \left(\frac{\partial I_l}{\partial x_l} \circ f_r \right) \big(1 - w_l \circ f_l^{-1}\big)^2 \\
\frac{\partial w_r}{\partial t} =& \rho \nabla^2 w_r - \frac{w_r}{\rho} + \frac{2}{\nu}(1-w_r) \|\nabla f_r\|^2
\end{aligned}
\tag{22}
$$

The boundary conditions are as follows. At the ends of each epipolar line, we require that d_l and d_r must be zero. There is no loss of generality by this assumption. If the disparities in fact are not zero at the end-points, then the situation is represented by making disparity discontinous at such end-points and having a corresponding occlusion area touching the end-points in the image of the other eye. The remaining boundary conditions are the homogeneous Neumann boundary conditions; that is, the normal derivative of the variables is set equal to 0.

For illustration, we present an example of a random dot stereogram shown in Figure 5. Each "dot" is a 4×4 pixel in size. Each image consists of 256×256 pixels. The left image is obtained from a random-dot image by shifting the central 96×96 pixel square to the left by 8 pixels and shifting the 176×176 square surrounding it by 4 pixels to the left. The right image is obtained by shifting the same squares to the right by the same amounts. Thus the reconstructed object is a square "wedding cake". Figure 6 portrays the numerically computed disparity values. The brighter areas indicate the higher disparity values. Figure 7 depicts the values of w_l and w_r. The bright areas indicate the blurred version of the discontinuity locus of d_l and d_r while the dark areas correspond to occlusion. The uniform gray indicates approximately constant disparity.

FIGURE 5: LEFT AND RIGHT IMAGES

FIGURE 6: LEFT AND RIGHT DISPARITIES d_l, d_r

FIGURE 7: w_l, w_r

BIBLIOGRAPHY

[ALM] L. Alvarez, P.L. Lions and J.M. Morel: ”Image Selective Smoothing and Edge Detection by Nonlinear Diffusion II”, SIAM J.Num.Anal, (June,1992).

[AT_1] L. Ambrosio and V.M. Tortorelli: ”Approximation of Functionals Depending on Jumps by Elliptic Functionals via Γ-convergence”, Arch.Rat.Mech.Anal. 111, pp.291-322, (1990).

[AT_2] L. Ambrosio and V.M. Tortorelli: ”On the Approximation of Functionals depending on Jumps by Quadratic, Elliptic Functionals”, Boll. Un. Mat. Ital. (1992).

[BDP] G. Bellettini, G. Dal Maso and M. Paolini: ”Semicontinuity and Relaxation Properties of a Curvature depending Functional in 2d”, SISSA Report, 17/92/MA, (Feb. 1992).

[BZ_1] A. Blake and A. Zisserman: ”Using weak continuity constraints”, Report CSR-186-85, Dept. of Comp. Sci., Edinburgh Univ. (1985)

[BZ_2] A. Blake and A. Zisserman: Visual Reconstruction, M.I.T. Press (1987)

[DMS] G. Dal Maso, J.M. Morel and S. Solimini: ”A Variational Method in Image Segmentation: Existence and Approximation Results”, Acta Math. 168, pp.89-151, (1989).

[DCL] E. De Giorgi, M. Carriero and A. Leaci: ”Existence Theorem for a Minimum Problem with Free Discontinuity Set”, Arch.Rat.Mech.Anal. 108, pp.195-218, (1990).

[GeGe] S. Geman and D. Geman: ”Stochastic relaxation, Gibbs’ distributions, and the Bayesian restoration of images”, IEEE Trans., PAMI 6, pp. 721-741 (1984)

[KWT] M. Kass, A. Witkin and D. Terzopoulos: ”Snakes: Active Contour Models”, First International Conf. on Computer Vision, (1987).

[M] D. Mumford: ”Elastica and Computer Vision”, Preprint.

[MS_1] D. Mumford and J. Shah: ”Boundary detection by minimizing functionals, I”, Proc. IEEE Conf. on Computer Vision and Pattern Recognition, San Francisco (1985)

[MS_2] D. Mumford and J. Shah: ”Optimal approximations by piecewise smooth functions and associated variational problems”, Comm. on Pure and Appl. Math., v. XLII, n.5, pp.577-684 (July, 1989)

[NM] M. Nitzberg and D. Mumford: ”The 2.1D Sketch”, Third Int'l Conf. on Comp. Vision, (December, 1990).

[OS] S. Osher and J. Sethian: ”Fronts Propagating with Curvature Dependent Speed: Algorithms based on the Hamilton-Jacobi Formulation”, J. Comp. Physics, 79, (1988).

[PPGO] M. Proesmans, E.J. Pauwels, L.J. Van Gool and A. Oosterlinck: ”Image Enhancement using Non-Linear Diffusion”, Proc. IEEE Conf. on Comp. Vision and Pattern Recognition, (June, 1993).

[PGO] M. Proesmans, L.J. Van Gool and A. Oosterlinck: ”Determination of Optical Flow and its Discontinuities using Non-linear Diffusion”, Preprint, ESAT-MI2, Katholieke Univ. Leuven, Leuven, Belgium, ((1993).

[Ri] T. Richardson: Ph.D. Thesis, Department of Electrical Engineering and Computer Science, MIT (1990)

[Sh_1] J. Shah: ”Parameter Estimation, Multiscale Representation and Algorithms for Energy-Minimizing Segmentations”, Tenth International Conference on Pattern Recognition, (June, 1990)

[Sh_2] J. Shah: ”Segmentation by Nonlinear Diffusion”, Proc. IEEE Conf. on Comp. Vision and Pattern Recognition, (June, 1991).

[Sh_3] J. Shah: ”Segmentation by Nonlinear Diffusion, II”, Proc. IEEE Conf. on Comp. Vision and Pattern Recognition, (June, 1992).

[Sh_4] J. Shah: ”Properties of Energy-Minimizing Segmentations”, SIAM J. on Control and Optim. v.30, no.1, 99–111, (1992).

[Sh_5] J. Shah: “A Nonlinear Diffusion Model for Discontinuous Disparity and Half-Occlusions in Stereo”, Proc. IEEE Conf. on Comp. Vision and Pattern Recognition, (June, 1993).

[Sh_6] J. Shah: “Piecewise Smooth Approximations of Functions”, Calculus of Variations and Partial Differential Equations, 2, pp.315-328, (1994).

[Sh_7] J. Shah: ”Shape Recovery from Noisy Images by Curve Evolution”, IASTED International Conf. on Signal and IMage Processing, (November, 1995).

[Sh_8] J. Shah: ”A Common Framework for Curve Evolution, Segmentation and Anisotropic Diffusion”, Proc. IEEE Conf. on Comp. Vision and Pattern Recognition, (June, 1996).

[Wa] Y. Wang: Ph.D. Thesis, Dept. of Math, Harvard Univ, (1989).

[W] Y. Wen: ”L^2 Flow of Curve Straightening in the Plane”, Duke Math. J. v.70, n.3, pp.683-398, (June, 1993).