Adaptive Contextual Regularization for Energy Minimization Based Image Segmentation

by

Josna Rao

B.A.Sc., Simon Fraser University, 2007

A THESIS SUBMITED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

Master of Applied Science

in

THE FACULTY OF GRADUATE STUDIES

(Electrical & Computer Engineering)

THE UNIVERSITY OF BRITISH COLUMBIA
(VANCOUVER)

May 2010

Abstract

Image segmentation techniques are predominately based on parameter-laden optimization processes. The segmentation objective function traditionally involves parameters (i.e. weights) that need to be tuned in order to balance the underlying competing cost terms of image data fidelity and contour regularization. Conventionally, the associated parameters are set through tedious trial and error procedures and kept constant over the image. However, spatially varying structural characteristics, such as object curvature, combined with varying noise and imaging artifacts, significantly complicate the selection process of segmentation parameters.

This thesis contributes to the field of image segmentation by proposing methods for spatially adapting the balance between regularization and data fidelity in energy minimization frameworks in an autonomous manner. We first proposed a method for determining the globally-optimum spatially adaptive regularization weight based on dynamic programming. We investigated this weight with a basic minimum-path segmentation framework. Our findings indicated that the globally-optimum weight is not necessarily the best weight as perceived by users, and resulted in poorer segmentation accuracy, particularly for high noise images. We then investigated a more intuitive approach that adapts the regularization weight based on the underlying local image characteristics to more properly address noisy and structurally important regions. This method, which we termed contextual (data-driven) weighting, involved the use of a robust structural cue to prevent excessive regularization of trusted (i.e. low noise) high curvature image regions and an edge evidence measure, where both measures are gated

by a measure of image quality based on the concept of spectral flatness. We incorporated our proposed regularization weighting into four major segmentation frameworks that range from discrete to continuous methods: a minimum-path approach [9], Graph Cuts [14], Active Contours Without Edges [24], and a contextual Mumford-Shah based approach [38]. Our methods were validated on a variety of natural and medical image databases and compared against the globally-optimum weight approach and to two alternate approaches: the best possible (least-error) spatially-fixed regularization weight, and the most closely related data-driven spatially adaptive regularization method. In addition, we incorporated existing texture-based contextual cues to demonstrate the applicability of the data-driven weights.

Table of Contents

List of Tables

List of Figures

Acknowledgements

I would like to thank my supervisors, Dr. Rafeef Abugharbieh and Dr. Ghassan Hamarneh, for their guidance and expertise. I would also like to thank Miranda Poon, Shai Bagon, Yue Wu, Erkut Erdem, and Sibel Tari for providing code for segmentation methods. In addition, I would like to thank my colleagues at the Biomedical Signal and Image Computing Laboratory (BiSICL) for their help and support. Finally, I would like to thank Joe and my parents for their love and encouragement.

JOSNA RAO

The University of British Columbia

May 2010

Chapter 1

Introduction and Background

Image segmentation plays a key component in many fields, ranging from medical image analysis to popular image editing programs. Medical research studies often rely on accurate and robust segmentation techniques to provide key information about anatomical shapes. Alternately, image editing software such as the Adobe Photoshop Suite rely on segmentation methods that accurately capture object boundaries for accurate modification by users. In all these cases, the images encountered may often be corrupted by noise or imaging artefacts that often present difficulties to many image segmentation methods. For these reasons, robust automated image segmentation is a highly sought after goal that continues to defy solution.

1.1 Motivation and Problem Statement

In medical images, natural and pathological variability as well as noise often result in unpredictable image and shape features that significantly complicate segmentation tasks. For example, MR images often contain measurement noise, partial volume effects, and image nonuniformity due to magnetic field inhomogeneities and magnetic susceptibility variations in the subject [88]. Furthermore, spatially nonuniform noise can result from numerous reconstruction and postprocessing techniques on MR images to correct for intensity inhomogeneity effects [90, 84], and from images obtained with partially parallel imaging (PPI) techniques [82, 18, 19]. Spatially varying noise levels can also result from images obtained with decreased acquisition times and high speedup factors [84]. In addition, regions with missing data or occlusions are commonly encountered in medical data, such as echo dropouts in ultrasound images that leads to irregularities along the feature boundary [99]. Other such degradations in medical images are due to tissue heterogeneity ("graded decomposition" [96]) and patient motion. In general photography applications, objects containing weak boundaries are quite common from poorly focused photos or from objects where sections in the image exceed the depth of field of the camera lens. Figure 1 shows some of these examples.
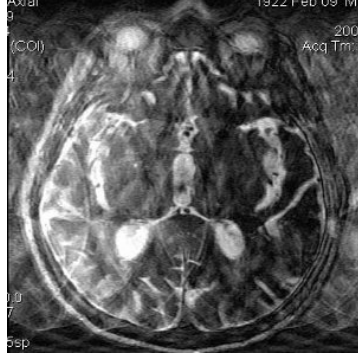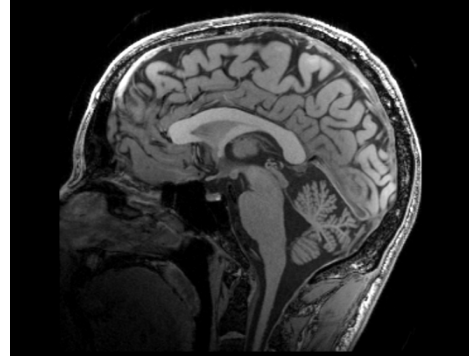
Figure 1: Examples of degradations in medical and natural images. (a) Patient movement during magnetic resonance (MR) scan [1]. (b) Intensity inhomogeneity in MR volumes from lack of bias field correction [89]. (c) Occlusions in natural images [75]. (d) Focus issues in natural images [75].

For all these cases, regularization, or smoothing, plays a crucial role in improving the robustness and accuracy of the resultant segmentations. Through the use of weighted regularization terms in conjunction with data fidelity terms, images plagued by high levels of deterioration, i.e. noise or poor edge contrast, are prevented from causing excessive irregularities and inaccuracies in the resultant segmentation. In order to increase segmentation robustness and accuracy, *more regularization is needed in less reliable image regions* which suffer from greater deterioration. One such example is shown for the noisy MR scan of a knee in Figure 2(a). A segmentation contour produced

with no regularization is erratic and inaccurate (Figure 2(b)). By including some smoothing (regularization) term, irregularities caused by the noise are reduced and a more accurate estimate of the original boundary is formed (Figure 2(c)).

(a)   (b)   (c)
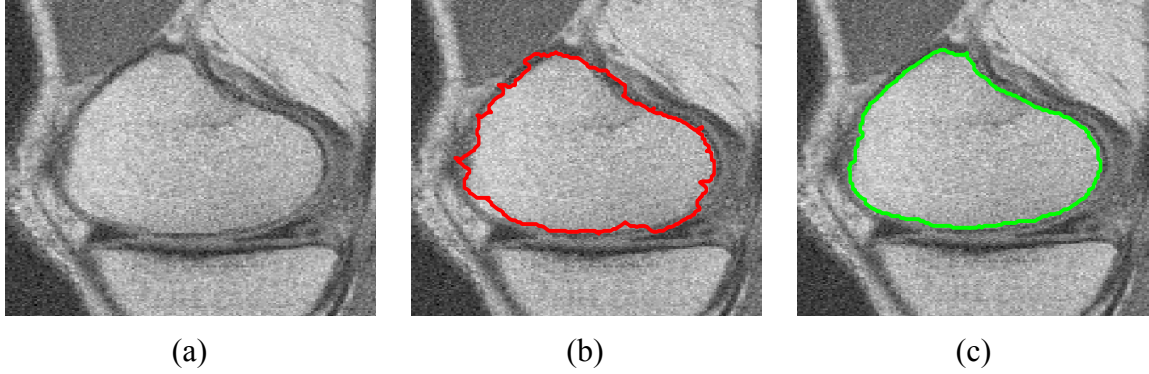
Figure 2: Example of role of regularization in degraded images. (a) Noisy MR scan of knee. (b) Contour produced with no smoothing (regularization) term (*red*). (b) Contour produced with some level of regularization (*green*). Regularization removes segmentation irregularities to form a better estimate of the true object boundary.

However, excessive regularization in regions of the image not plagued by deterioration can result in less detail and loss of key structures in the final segmentation. For example, magnetic resonance (MR) images of the brain typically feature highly detailed structures such as the cortical surface which contain many regions of high curvature, as shown in Figure 3. Excessive regularization of these high curvature regions results in segmentations which fail to capture key characteristics of the original object
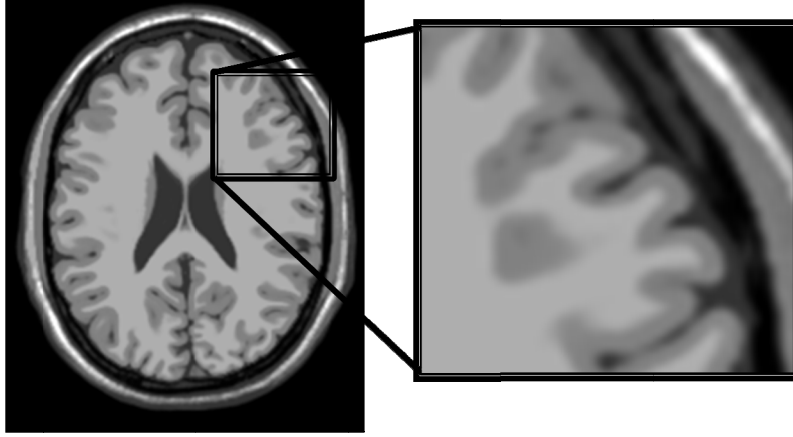
Figure 3: Example of structurally-important regions with high curvature. Cortical folds in mid-coronal MR brain scan (BrainWeb T1 image [26]) produce a highly varying object boundary.

Most reported approaches to segmentation keep a uniform level of regularization across an object contour, i.e. one that does not vary spatially and is determined empirically. This fixed approach can result in both excessive regularization and inadequate regularization for a single object depending on the object boundary variability. Spatially adapting the regularization weights across a segmentation provides greater control over the segmentation result, allowing it to adapt not only to images with spatially varying noise levels and edge strength, but also to objects with spatially-varying shape characteristics, e.g. smooth in some parts and jagged in others. In this thesis, we advocate the strong need for spatially-adaptive balancing of cost terms in an automated, robust, data-driven manner to relax the requirement on the user to painstakingly tweak these parameters. We also demonstrate how the globally optimized approach and fixed weight approaches are often inadequate for achieving accurate segmentation.

5

1.2 Energy Minimization Segmentation Methods

Our work focuses on modulating the regularization of energy-minimization segmentation methods. In this section, we will examine in detail several modern and popular segmentation approaches ranging from discrete to continuous, and we will discuss the tradeoff between data fidelity and regularization terms in these methods.

1.2.1 General Framework

The vast majority of existing segmentation methods are predominantly based on parameter-laden optimization procedures designed to produce `optimal' segmentations at their minimum. These techniques represent the energy of segmentation as a combination of smoothing terms and data fidelity terms. The 'optimum' segmentation is the segmentation which represents the minimum energy of all possible segmentations. Regularization terms are called the internal energy, and data fidelity terms are referred to as the external energy. A simplified but general form of the cost or energy, E, of a segmentation, S, of an image, I, is

$$E(S|I, \alpha, \beta) = \alpha\, E_{int}\,(S)\, +\, \beta\, E_{ext}\,(S|I) \qquad\qquad (1)$$

E_{int} is the internal energy term contributing to the regularization of the segmentation, and E_{ext} is the external energy term contributing to the contour's conformity to desired image features, e.g. edges. These terms will be discussed in more detail in Section 1.3.

The weights α (referred to as the regularization weight) and β in (1) control the highly sensitive tradeoff between the regularization terms and data fidelity terms. The

resultant segmentations can vary drastically based on how this balance is set. For example, setting the weight β to zero results in the external energy playing no role in the optimization procedure. Instead, the segmentation which minimizes the internal energy will be optimal. Similarly, setting α to zero will result in the segmentation that produces the minimum external energy to be optimal, regardless of whether the external energy reflects boundary regions or not. A demonstrative example of segmentations from different settings of the weights is shown in Figure 4.

(a)   (b)   (c)
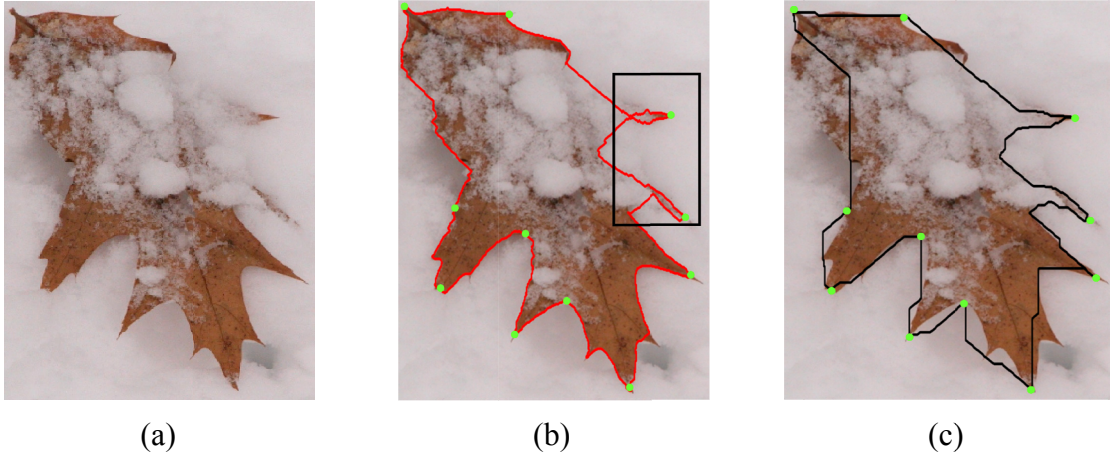
Figure 4: Importance of the regularization weight in the segmentation process. (a) Original image of an obscured-boundary leaf [75]. Segmentations (by a seeded minimum-path approach with seeds shown in *green*) produced by (b) a low regularization weight (*red*) and (c) a high regularization weight (*black*). The low regularization weight produces a segmentation with insufficient smoothing in the obscured boundary region. The high regularization weight fails to capture key regions of the object boundary.

α and β in (1) are typically set empirically by the users based on their judgment of how to best balance the requirements for regularization and adherence to image content. In most cases, this is a very difficult task and the parameters may be unintuitive for a typical non-technical end user, e.g. a clinician, who lacks knowledge of the underlying algorithm's inner working. Avoiding the practice of ad-hoc setting of such weights is the driving motivation for our work here. Addressing the issue of how to best balance

competing cost terms is of great importance to many related algorithmic formulations in computer vision. More generally, this tradeoff is seen in likelihood versus prior in Bayesian methods [2] and loss versus penalty in machine learning [106, 105]. The role of regularization is discussed in more detail in Section 1.4.

A wide variety of energy minimization segmentation methods exist, with each method differing in the type of energy functional used and in the technique used for optimizing the functional. In general, these methods can be classified into two groups based on the type of space the functional is defined on [79, 15]:

1) Functionals defined on a discrete space (discrete set of variables)

2) Functionals defined on a continuous space (continuous contour or surface)

Depending on the type of segmentation method, the segmentation S in (1) can differ in representation. In addition, segmentation techniques differ in the choice of external and internal energy terms, as will be discussed in Section 1.3

1.2.2 Discrete Methods

Discrete segmentation methods formulate the problem as a combinatorial optimization in a finite space Z^n (a finite set of integer-valued variables) [79]. These methods predominately use a graph-based representation of an image where graph vertices (nodes) correspond to image pixels. These methods can be further divided into explicit methods, such as path-based methods [9, 4, 43, 41, 55], where the object boundary is represented by a path in a graph, and implicit methods, such as graph-cuts

8

approaches [13, 15, 14] where the segmentation is represented functionally. These methods can guarantee finding the globally optimal segmentation.

1.2.2.1 Minimum Path

Minimum-path approaches formulate segmentation through modelling the image as a graph that consists of nodes and edges that link to neighbouring nodes. Image pixels represent vertices and the links from each pixel to its eight neighbours represent edges. In these approaches, the segmentation is modelled as a path along the graph consisting of a set of nodes where each node is connected by an edge to a single 'forward' node and a single 'backward' node. The subgraph formed by a closed path along the graph represents the object or region of interest. Each edge in the graph is given a cost such that the cumulative cost of the path from a start vertex to a goal vertex is the sum of the individual cost of each edge along the path. The problem of finding the optimum segmentation is modelled as a graph-searching problem where the goal is to find the optimum, or least-cost, path between two vertices. Many segmentation methods, such as [15, 17, 108, 47], use a graph-based approach to modelling the segmentation problem. Common examples of minimum path approaches are Amir *et al* [4], Geiger *et al* [4], Falcao *et al* [41], and Jerymyn and Ishikawa [55].

In particular, one popular minimum-path segmentation approach is Livewire [41, 9] where dynamic programming in conjunction with user supplied seedpoints are used to find the optimum path on a graph-based representation of the image. Livewire uses the edges costs of the graph to ensure that the optimum path represents characteristics desired in the segmentation. Each edge cost, or local cost, is a weighted sum of data fidelity and

regularization terms. The local cost, $E_{total}(p, q)$, for the directed edge from pixel p to neighbouring pixel q is as follows:

$$E_{local}(p, q) \; = \; w_G \, E_G(q) \; + \; w_D \, E_D(p, q) \tag{2}$$

where a gradient magnitude measure $E_G(q)$ acts as a data fidelity term and $E_D(p, q)$ is the regularization term that penalizes longer paths (see Section 1.3 for further discussion of energy terms). The balance between data fidelity and regularization is controlled by the weights w_G and w_D which are typically set empirically.

In order to determine the optimum contour, Livewire uses dynamic programming, which solves optimization tasks by recursively breaking down the tasks into similar but smaller subtasks which can be solved more easily. In the context of graph-based segmentation, the globally optimum path between vertices p and q is found by recursively finding the optimum path for smaller subgraphs. This process is accomplished by first initializing the vertices with the local cost at that pixel location as determined by equation (2). The target vertex q is then expanded by summing the cumulative cost of q into all adjacent vertices, and then summing the cost of those vertices into their corresponding neighbours. This expansion continuous in order of minimum cumulative cost and produces a 'wavefront' that extends over the graph, as illustrated in Figure 5 . At the end of this process, the graph is transformed into what is known as a *value function* that represents the cost of the globally optimum path from each vertex to the target vertex q, as illustrated in Figure 6. The optimum path from any vertex, including p, to the goal vertex can then be determined by a gradient descent. The benefit is that the globally optimum path from *all* vertices to the target vertex is now known. This optimization

10

procedure has been expanded in recent work to a 3D graph where the optimum path consists of pixel coordinates and a third 'scale' variable [81, 64]. For more information, refer to Chapter 2 where 3D graph searches are discussed in more detail.
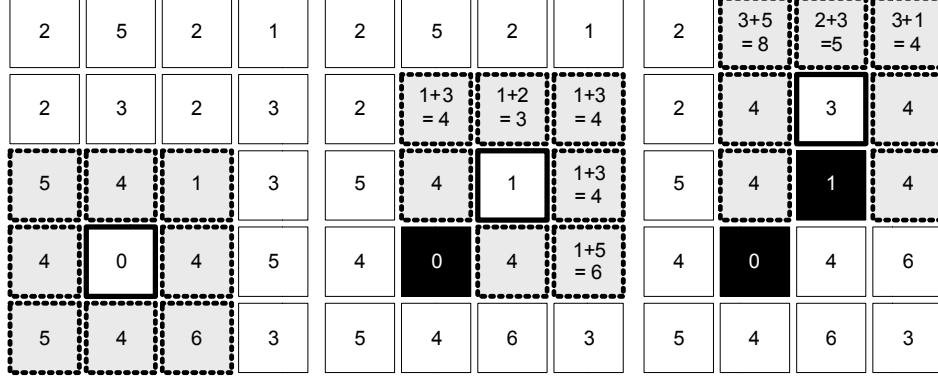
Figure 5: Diagram of wave-front expansion process in Djikstra's method. Cumulative cost from starting node is summed into unvisited neighbours, which are in turn expanded. (Reproduced with permission from [81])

| (a) | (b) | (c) |
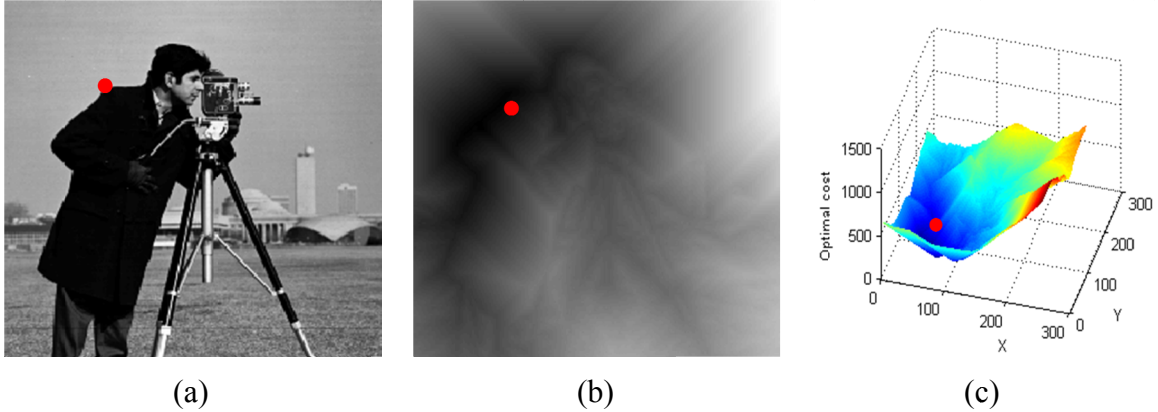
Figure 6: Value function in dynamic programming. (a) Original image with target node in red. (b) Value map formed by Djikstra's method where intensity values represent cost of optimal path from any node (pixel) to target node. Costs are lowest along pixels containing strong edge evidence. (c) 3D value function where z-axis represents cost of the optimal path from any (x, y) location to the target node. The optimal path is then determined by a gradient descent procedure.

1.2.2.2   Graph Cuts

Another popular graph-based segmentation approach are *s-t* graph cuts, first proposed for image segmentation in [17], further expanded in [13, 16, 15, 14], and

proposed outside the field of image segmentation in [48, 13, 83, 53, 60] . Our work will incorporate the implementation of graph cuts in [15].

Unlike minimum-path approaches which define the segmentation as a contour along the object boundary, graph-cuts approaches define the segmentation as a region where each pixel is represented as a binary variable, either inside or outside the object of interest. Graph-cuts defines the energy functional for a segmentation as follows:

$$E(f) = \sum_{\{p,q\}\in N} \lambda_1 V_{p,q} \left(f_p, f_q\right) + \lambda_2 \sum_{p\in P} D_p(f_p) \qquad (3)$$

$f \in \mathrm{L}$ is the labeling for all pixels $p \in P$ where L is the label space and P is the set of pixels in image I. $V_{p,q}$ is the pairwise interaction penalty between pixel pairs (i.e. the penalty of assigning labels f_p and f_q to pixels p and q), N is the set of interacting pairs of pixels, and D_p measures how well label f_p fits pixel p given the observed data. D_p and $V_{p,q}$ are discussed in more detail in Section 1.3. From the cost functional, we see that graph-cuts also involves a tradeoff between regularization and data fidelity through the weights λ_1 and λ_2 in (3).

The optimization procedure consists of generating a labeling through two types of moves: expansions and swaps, which changes the labels of large numbers of pixels simultaneously. The method divides the image into nodes, which are the pixels or voxels, terminal nodes, which consist of a source node and a sink node that represent the background and object labels, n-links, which are edges connecting neighboring nodes, and t-links connecting nodes to terminal nodes, where links are assigned a cost based on the edge. The resulting contour is defined as a 'cut' through a subset of n-links,

12

separating the sink node (background) from the source node (object). The cut cost is defined as the sum of costs from the edges that were cut.

1.2.3 Continuous Methods

Continuous segmentation methods formulate the problem in the continuous space R^∞ and use a representation of the segmentation that deforms according to external and internal forces. Additionally, these methods tend to rely on a gradient descent procedure for optimization. These types of methods are divided into explicit models and implicit models. Explicit models are based on an explicitly defined parametric curve that is evolved and deformed, and consist of active contour and snake methods [56, 28, 29, 71, 65]. Implicit models are where the contour is represented as the level-set of a higher-dimensional scalar function, as seen in [21, 86, 76, 85, 77]. Unlike discrete methods, continuous methods can only guarantee finding the local minima of the energy functional. However, continuous methods do have the advantage that pixel connectivity in the final segmentation is guaranteed. Our work will focus on two such continuous methods that are both different approximations of the Mumford-Shah segmentation framework

1.2.3.1   Mumford-Shah Model

Mumford and Shah [74] proposed a solution to the problem of image denoising by dividing an image into a smooth cartoon-like component and a noise component such that the image is smoothened but object edges are retained. This concept that has been further analyzed by [12, 44, 78]. To accomplish this, the Mumford-Shah (MS) model was proposed, where a segmentation consists of a piecewise-smooth approximation, u, of the

original image, u_0, and an edge set, Γ. In this formulation, a segmentation has the following energy:

$$E_{MS}(u, \Gamma) = \beta \int_{\Omega} \big(u(x,y) - I(x,y)\big)^2 dxdy + \alpha \int_{\Omega \backslash \Gamma} |\nabla u(x,y)|^2 dx$$

$$+ \, length(\Gamma) \tag{4}$$

where $\Omega \subset \mathfrak{R}^2$ is a connected, bounded and open subset that represents the image domain, I is the original image defined on $\mathfrak{R}$, $\Gamma \subset \Omega$ is the edge set segmenting Ω, u is the piecewise smooth approximation of I, and α and β are the scale space parameters of the model. The first term corresponds to the external energy and penalizes large differences between u and the original image. The regularization terms include $\int_{\Omega \backslash \Gamma} |\nabla u|^2 dx$ which penalizes large edge sets which would result from more erratic segmentations, and $length(\Gamma)$ which penalizes an excessively large edge set (and a low smoothened image). The weights α and β control the balance of regularization versus data fidelity.

The original MS model is difficult to minimize due to the unknown representation of the edge set Γ. We will present two different approximations of the MS problem: contextual Ambrosio and Tortorelli (AT) approximation, and the Active Contours without Edges (ACWE) approximation.

1.2.3.2   *Contextual Ambrosio-Tortorelli MS Approximation*

The AT approximation of the MS model simplifies the minimization process by introducing a smooth edge indicator function v [3], and has been utilized by many other

segmentation methods [8, 5, 37, 39, 87, 92, 94]. The original model incorporated a characteristic function χ_Γ as the edge indicator. In the AT approximation, the cardinality of Γ is approximated by:

$$\frac{1}{2} \int_\Omega \left(\rho |\nabla v|^2 + \frac{(1-v)^2}{\rho} \right) dx \tag{5}$$

where ρ is a parameter set such that as $p \to 0$, $v(x) \approx 0$ if $x \in \Gamma$ and $v(x) \approx 1$ otherwise. This approximation modifies the MS functional to the following:

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

The AT approximation allows for the partial differential equations (PDEs) that dictate the segmentation evolution to be decoupled into separate evolution equations for the image process and edge set function as follows:

$$\frac{\partial u}{\partial t} = \nabla \cdot (v^2 \nabla u) - \frac{\beta}{\alpha}(u - I); \quad \left. \frac{\partial u}{\partial n} \right|_{\partial \Omega} = 0 \tag{7}$$

$$\frac{\partial v}{\partial t} = \nabla^2 v - \frac{2\alpha |\nabla u|^2 v}{\rho} - \frac{(v-1)}{\rho^2}; \quad \left. \frac{\partial v}{\partial n} \right|_{\partial \Omega} = 0 \tag{8}$$

where $\partial \Omega$ is the boundary of Ω and n is the unit normal vector to $\partial \Omega$. The AT approximation adds the additional regularization term $\rho |\nabla v|^2$ that forces edges to be smooth. Although the parameters β and α control the data fidelity term and one of the regularization terms, respectively, $\rho |\nabla v|^2$ is not modulated by a regularization weight. From the energy functional, it also difficult to separate regularization of the image

15

process and regularization of the edge process. Additionally, it is important to note that the edge term v itself acts as a weight for the regularization of the image process. From the term $v^2 \nabla u$, if an edge exists ($v \approx 0$), no regularization occurs. Thus, we follow the Erdem and Tari [38] modification of the AT MS model that allows for proper control of regularization of both the image process and edge process through modifying the evolution equation for the image process as follows:

$$\frac{\partial u}{\partial t} = \nabla \cdot ((cv)^2 \nabla u) - \frac{\beta}{\alpha}(u - I); \quad \left.\frac{\partial u}{\partial n}\right|_{\partial\Omega} = 0 \quad\quad (9)$$

where the constant c controls how strongly the edge term v weights the level of smoothening in the image process. This segmentation approach is termed the *contextual MS* method.

The evolution equations (8) and (9) are simultaneously solved for u and v by the Finite Differences numerical discretization technique iteratively where u is updated by the evolution equation while v is kept fixed, and vice versa, and where the iterations are stopped when the solution is stationary. It is important to note that the contextual MS approach is automated in that no initial contour or seeds are used to determine u and v at the initial iterations. Instead, v is initialized as the inverse of the gradient such that regions with high edge evidence result in $v \approx 0$. Additionally, the contextual MS approach is primarily used for denoising as it produces a cartoon-like version of the original image. For use as a segmentation method, a binary mask can be formed from closed contours in the edge set.

16

1.2.3.3   Active Contours Without Edges

The active contours without edges (ACWE) segmentation approach, proposed by

Chan and Vese [24, 23, 22, 97], represents a subset of the original MS model. Here, the

original MS segmentation problem is restricted to piecewise constant functions, forming

what is known as the minimal partition problem. The original MS formulation had a data

fidelity term of $\int_\Omega |u(x,y) - I(x,y)|^2 dxdy$. In the ACWE formulation, the image

process is simplified to a binary piecewise constant function as follows:

$$u = \begin{cases} \text{average}(I) & \text{inside } \Gamma \\ \text{average}(I) & \text{outside } \Gamma \end{cases} \tag{10}$$

u is thus represented by two constants, c_1 and c_2, which represent the average of the

original image inside and outside, respectively, of the object boundary Γ. Essentially,

ACWE is a simplification of the original MS model to only segment for an object and

background and producing a binary image u with the edge set simplified to an active

contour representing the object boundary. As a binary mask is produced, ACWE is

geared more towards the application of segmentation rather than denoising. The ACWE

energy functional is:

$$E(c_1, c_2, \Gamma) = \mu \, \text{Length}(\Gamma) + v \, \text{Area}\big(inside(\Gamma)\big) \tag{11}$$

$$+ \lambda \int_{inside(\Gamma)} |I(x,y) - c_1|^2 dxdy$$

$$+ \lambda \int_{outside(\Gamma)} |I(x,y) - c_2|^2 dxdy$$

17

where the different external energy weights have been replaced with a single weight λ and the regularization weights are μ and υ (in practice, $\upsilon = 0$ typically). The weights μ and λ form the same tradeoff between data fidelity and regularization that is present in the other segmentation methods discussed.

To optimize for the energy, the functional is first transferred to a level sets formulation as follows. The segmentation contour $\Gamma \subset \Omega$ is represented by the zero level set of a Lipschitz function $\phi; \Omega \to \mathfrak{R}$ where pixels p interior to the zero-level set of ϕ are labelled as objects and exterior pixels as background. The length and area of the zero level set of ϕ is determined through the use of the Heaviside function H and the one dimensional Dirac function δ_0 as follows:

$$Length(\phi = 0) = \int_\Omega \delta_0\big(\phi(x,y)\big)|\nabla\phi(x,y)|dxdy \qquad (12)$$

$$Area(\phi \geq 0) = \int_\Omega H\big(\phi(x,y)\big)dxdy \qquad (13)$$

Through the use of H and δ_0, the energy function is written in level-sets form as follows:

$$E(c_1, c_2, \phi) = \mu \int_\Omega \delta_0\big(\phi(x,y)\big)|\nabla\phi(x,y)|dxdy \; + \; \upsilon \int_\Omega H\big(\phi(x,y)\big)dxdy \qquad (14)$$

$$+ \; \lambda \int_\Omega |I(x,y) - c_1|^2 H\big(\phi(x,y)\big)dxdy$$

$$+ \; \lambda \int_\Omega |I(x,y) - c_2|^2 \Big(1 - H\big(\phi(x,y)\big)\Big) dxdy$$

The corresponding Euler-Lagrange equation is derived by minimizing E with respect to ϕ. The optimum contour is then determined by gradient descent where the descent direction parameterized by an artificial time $t \geq 0$ is:

$$\frac{\partial \phi}{\partial t} = \delta(\phi) \left[\mu \ div \left(\frac{\nabla \phi}{|\nabla \phi|} \right) - v - \lambda(I - c_1)^2 + \lambda(I - c_2)^2 \right] = 0 \qquad (15)$$

The finite differences implicit scheme is first used to discretize ϕ. The initial ϕ_0 is then set by the user as an initial contour. The optimum contour is then determined by iteratively updating ϕ to determine ϕ^{n+1} by adding the evolution equation scaled to a step size. The process continues until the solution is stationary. Unlike the contextual MS approach, the ACWE approach is not automated and allows user control through the initial contour to segment specific objects in an image.

1.3 Energy Terms

We will next discuss the individual external and internal energy terms used in the segmentation methods of Section 1.2 and how these terms respectively enforce data fidelity and regularization of the segmentation and capture properties of the object of interest which is desired in the segmentation.

1.3.1 External Terms

The external energy terms of a segmentation method contribute to data fidelity and ensuring that the final segmentation does not greatly differ from the original image, or in some cases, ensuring that the contour is aligned with high edge characteristics of the

original image. In general, external terms can be divided into two categories: boundary terms and region terms.

1.3.1.1 Boundary terms

Methods such as Livewire and other minimum-path approaches, where the segmentation produces a path representing the object boundary, tend to use boundary terms that penalize paths, or contours, that do not consist of pixels containing high edge evidence. In the case of the simplified Livewire framework used in this thesis, the key edge evidence term is:

$$E_G = 1 - \frac{|\nabla I|}{\max(|\nabla I|)} \tag{16}$$

where the gradient magnitude of the original image I is scaled and inverted using an inverse linear ramp function such that high image gradient regions correspond to a low edge energy or cost. This also acts as a first order positioning of the contour. The original Livewire framework uses additional edge evidence terms such as the Canny edge evidence measure [20] and the Laplacian of Gaussian second order term [9], but these were omitted in the simplified representation used here for the purpose of focussing on the regularization weight balance rather than the segmentation framework itself.

1.3.1.2 Regional terms

The graph cuts segmentation method uses a region based term to reflect how well the intensity of a pixel p fits into given intensity models that are known a priori. These intensity models (or histograms) are created from seeds provided from the user representing the object labels and the background labels. The graph cuts method makes

use of a negative log likelihood function to determine the data fidelity region energy E_R as follows:

$$E_R(obj) = -\ln \Pr\left(I_p \mid {}'obj'\right) \tag{17}$$

$$E_R(bkg) = -\ln \Pr\left(I_p \mid {}'bkg'\right)$$

where the prior probability of a pixel p with intensity I_p belonging to the object label is determined by an intensity model (histogram) created from the intensities of the object seed pixels, and similarly for the background label [15, 48].

The contextual MS model uses the least square error between the segmentation (image process u) and the original image as the data fidelity term as follows:

$$E_{AT_{ext}} = \int_\Omega \beta\big(u(x,y) - I(x,y)\big)^2 dxdy \tag{18}$$

where high differences between the smoothened piecewise-constant segmentation and the original image are assigned a high energy to minimize.

The ACWE simplification of the MS model uses a similar least square error term as the external energy, but unlike the contextual MS method, the ACWE model simplifies the segmentation u to a binary image consisting of two constants, c_1 and c_2, as follows (using the level sets formulation):

$$E_{ACWE_{ext}} = \lambda \int_{\Omega} (I(x,y) - c_1)^2 H\big(\phi(x,y)\big) \, dxdy \tag{19}$$

$$+ \lambda \int_{\Omega} (I(x,y) - c_2)^2 \Big(1 - H\big(\phi(x,y)\big)\Big) dxdy$$

where

$$\begin{cases} c_1(\phi) = average(I) \; in \; \{\phi \geq 0\} \\ c_2(\phi) = average(I) \; in \; \{\phi < 0\} \end{cases} \tag{20}$$

Large differences between the segmented image and the piecewise constant original image will be assigned large external energies.

1.3.2 Internal Terms

Internal energy terms, or smoothening/regularization terms, range from simple penalizations of long contour lengths to terms that enforce shapes or prior knowledge [10]. Here, we will focus on the terms employed by the segmentation methods introduced in Section 1.2.

The Livewire framework and the majority of minimum-path segmentation approaches use a regularization term that penalizes longer and jagged contours through estimating the contour length. In the Livewire framework, the local regularization cost of a vertex (pixel) p's link to neighbouring vertex q is an estimate of the Euclidean distance to that neighbour: $E_D = \sqrt{(p_x - q_x)^2 + (p_y - q_y)^2}$ such that diagonal neighbours incur a higher regularization cost [81].

22

In more general minimum-path frameworks, the internal energy is calculated as the piecewise length of the contour:

$$E_{int}\big(C(q)\big) = \left|\frac{\partial C(q)}{\partial q}\right| \tag{21}$$

where C is the contour parameterized by some variable q (i.e., $C(q) = C\big(x(q), y(q)\big) : [0,1] \to \Omega \subset \mathfrak{R}^2$ in image $I : \Omega \to \mathfrak{R}^2$). In addition, active contour methods, such as the classical snakes method [56], use a second order regularization term as follows:

$$E_{int} = \alpha \left|\frac{\partial C(q)}{\partial q}\right|^2 + \beta \left|\frac{\partial^2 C(q)}{\partial q^2}\right|^2 \tag{22}$$

In the graph cuts formulation, the internal energy uses a penalty term between neighbouring pixels where a certain penalty is assigned if the pixels are assigned to different labels, thereby favouring similar groupings between neighbouring pixels. This term is as follows:

$$E_{int} = \begin{cases} 1 & if \ f_p \neq f_q \\ 0 & if \ f_p = f_q \end{cases} \tag{23}$$

where pixels p and q are assigned labels f_p and f_q, respectively. Typically, the penalty term is then weighted by a term that has a high penalty when neighbouring pixels with highly dissimilar intensities are assigned the same label.

In the contextual MS approach, the internal energy consists of two regularization terms:

$$E_{int} = \alpha(v^2|\nabla u|^2) + \frac{1}{2}(\rho|\nabla v|^2) \qquad (24)$$

where the first term smoothes the image with a filter radius proportional to the values of v^2 and $\frac{\alpha}{\beta}$, which preserves edges during regularization [38, 94, 8, 44]. The second term, which comes from the AT approximation of the cardinality of the edge set Γ, enforces more smooth edges by assigning a high energy to large edge sets.

The ACWE segmentation method uses the length of the level set contour and the total area within the contour as regularization terms as follows:

$$E_{int} = \mu \, \text{Length}\{\phi = 0\} + v \, \text{Area}\{\phi \geq 0\} \qquad (25)$$

$$= \mu \int_{\Omega} \delta_0\big(\phi(x,y)\big)|\nabla\phi(x,y)|dxdy + v \int_{\Omega} H\big(\phi(x,y)\big)dxdy$$

which assigns a high energy if the zero level set of ϕ is erratic and long in length, and an additional high energy if the interior of the contour is excessively large (in practice, the area regularization term is not used however).

In additional, segmentation methods may have more specific terms designed to enforce data fidelity. For example, in addition to the standard contour length term, the classical snakes segmentation method [56] uses a curvature term for regularization, which will be further discussed in Section 1.5.2.

1.4 Regularization of Image Segmentation Methods

In all the segmentation methods that have been discussed in Section 1.2, a tradeoff exists between the external and internal energy terms. In this section, we will discuss the

role the regularization weight plays in this balance and the consequences of inadequate or excessive regularization. Additionally, we will discuss the traditional methods for setting this regularization weight.

1.4.1 Role of Regularization

Regularization plays a key role in reducing segmentation inaccuracies that arise from image degradations and from object boundary behaviour. Noise, ranging from simple impulse salt-and-pepper noise to complex average white Gaussian noise (AWGN), often arises in digital images during image acquisition. Often, environmental conditions such as low light contribute to noise, as well as image sensors, and corruption during image transfer [46]. The process of removing noise from images can often cause the object of interest to be degraded. For example, applying a low pass filter to remove the noise will often weaken the gradient of object edges. For these reasons, segmentation methods often encounter images that have not had any major noise removal preprocessing performed on them. The segmentation method must therefore by robust to noise levels. Complicating matters is the fact that noise is enhanced more than the image signal during edge detection processes [46, 42]. The end result is that noisy regions contribute high external energies and skew optimization processes to favour contours that contain these noisy regions [70]. By either penalizing longer erratic contours, or penalizing segmentations that differ in intensity or area from the original image object, a sufficient level of regularization is able to produce more robust segmentations. Unlike noise, weak gradients contribute to *low* external energy, causing optimization processes to favour contours or segmentations that contain pixels from stronger gradients in the image, even if these

strong gradients are located farther from the object of interest. Regularization by either penalizing the contour length or segmentation area, or the difference in the intensity from original object to segmentation, produces more accurate results. However, excessive regularization can remove detailed regions in the object if these are mistaken as noise or image deteriorations. We will next discuss the consequences of improper setting of the regularization weight.

1.4.2 Conventional Methods for Setting Regularization

Determining the optimum balance between regularization and adherence to image content has predominantly been done empirically and in an ad-hoc manner. Most reported approaches to segmentation keep a uniform level of regularization across an object contour, i.e. one that does not vary spatially and is determined empirically. All possible weights are tested, and the weight which produces the least error segmentation is selected as the best fixed weight. However, compensating for image deteriorations by uniformly increasing the level of regularization until the most degraded region of the image is properly regularized may result in excessive smoothing in those regions that do not require that much regularization. Subsequently, this results in a loss in segmentation accuracy, particularly for objects with highly curved boundaries. For example, consider the synthetic example in Figure 7 of an object with highly varying boundary characteristics, and where the image has been corrupted by AWGN with spatially varying levels. When the example object is segmented by a simple minimum-path approach (simplified Livewire) with a low level of regularization, the highly detailed sections of the object are accurately segmented. However, the sections deteriorated by noise are not

regularized and thus produce an inaccurate segmentation. Alternately, if the level of regularization is set high, the structurally important region of the object are excessively regularized. The result is that neither segmentation weight is suitable for the entire image.
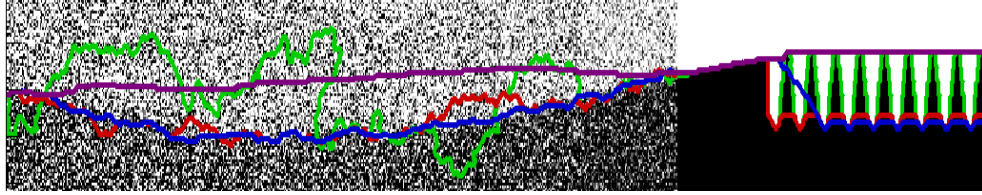
Figure 7: Segmentations on a synthetic image using spatially fixed regularization weights of 0 (*green*), 0.3 (*red*), 0.7 (*blue*), and 0.9 (*purple*). Lower regularization weights result in erratic contour behavior in the high noise left region of the image, and higher regularization weights result in poor segmentation of the high curvature right region of the image.

1.5 Related Work

1.5.1 Spatially Adaptive Regularization

As addressed in McIntosh and Hamarneh [72], adapting the regularization weights across a set of images is necessary for addressing the variability present in real image data. However, although an optimal weight can be found for a single image in a set, that weight may not be optimal for all regions of that image. In [91], a max-margin approach is used to learn the optimal parameter setting, but required the use of prior knowledge. In [60], Kolmogorov *et al.* characterized the types of energy functional that can be minimized via graph cuts and solved the optimization problem for a *range* of parameters rather than a single regularization weight. In Pluempitiwiriyawej *et al.* [80], the weights are changed as the optimization progresses, but in an ad-hoc predetermined manner. Some form of spatially adaptive regularization over a single image appeared in Dong *et al.* [35]. For segmenting an aneurysm, they varied the amount of regularization

27

based on the known surface curvature of a *pre-segmented* vessel. The results demonstrated improvements due to adaptive regularization. However, the regularization weights did not rely on the properties of the image itself, which limited the generality of the method.

The closest related work to ours is by Erdem and Tari [38] who proposed a regularizer for the MS segmentation framework with feature preserving capabilities through the use of a contextual feedback. The segmentation method itself was discussed in Section 1.2.3.2. The framework features a term that modulates the level of regularization that the edge process v has on the image process u. This term, cv in (9), depends on contextual cues through two types of feedback: negative feedback for feature smoothening, where the regularization shifts towards the maximum value of 1 when a cue measure, κ, is large, and positive feedback for feature preservation, where the regularization shifts to the minimum value of 0 when κ is large. The two forms of feedback are achieved by

$$cv = \kappa v + (1 - \kappa)V \tag{26}$$

where $V = 1$ for negative feedback, and $V = 0$ for positive feedback. The purpose of the two forms of feedback is to adjust the regularization for the edge process and the image process separately such that certain features, like texture, can be preserved (not regularized) in the image process, and other features, such as texture edges, can be eliminated (regularized) in the edge process.

Erdem and Tari focused on four data cues to modulate the regularization: (1) edge continuity, (2) edge (gradient) consistency, (3) a texture edge measure, and (4) a texture

region measure. The first cue was designed to create an edge set with linked edges by regularizing any edges that were not sufficiently linked to neighbouring edges, thus forming a coherent edge set and providing some level of weak noise detection, primarily against impulse noise. The level of directional consistency was measured through determining the angle between gradient direction vectors. The second cue estimated edge continuity to handle noise and weak edges that may cause breaking up of an edge contour. Through a support term that measures how many neighbouring pixels have an edge indicator, a measure of edge formation was calculated. As the edge is determined to be non-accidental in a region, positive feedback is used so that the diffusivity approaches 0 (feature preservation). The third cue estimated texture edges to prevent texture gradients from being included as part of the object boundary gradient set. The texture in a local window is estimated by comparing the similarity between the central patch, p^c, and patches shifted to the left, right, above, and below, as shown in Figure 8(a). The similarity metric is the Euclidean distance transform between each shifted patch and the central patch (which acts as the template), producing the similarity distributions $D_p^{up}, D_p^{down}, D_p^{left}$, and D_p^{right}. If a pixel (x, y) lies in a textured region, the central patch will differ from the shifted patches, and thus D_p^{up} will differ from D_p^{down} and D_p^{left} will differ from D_p^{right}, as illustrated in Figure 8(b). To verify this difference, the Wilcoxon Mann-Whitney test [101] (a rank-sum test) is used to produce a p-value where p-values $\ll 0.05$ indicate a significant difference. Thus the estimate of texture is determined as follows:

$$T(x,y) = 1 - \exp\left(-\gamma^t \left(\min(\rho_1(x,y), \rho_2(x,y))\right)\right) \qquad (27)$$

where γ^t is a decay rate parameter and $\rho_1(x,y)$ and $\rho_2(x,y)$ represent the p-values returned from the Wilcoxon Mann-Whitney test between D_p^{up} and D_p^{down} and between D_p^{left} and D_p^{right}. If texture exists around the pixel at (x,y), $\rho_1(x,y)$ and $\rho_2(x,y)$ will be low, indicating a significant difference between D_p^{up} and D_p^{down} and between D_p^{left} and D_p^{right}, thus producing a low $T(x,y)$. If the pixel at (x,y) lies in a piecewise constant region, the shifted patches will not differ from the central patch, and thus the distributions will not differ from one another, producing high p-values and therefore a high $T(x,y)$. We note that this method suffers from mistaking some non-texture edges as texture since non-texture edges will cause the shifted distributions to slightly differ from one another.

(a)　　　　　　　　　　　　　　(b)
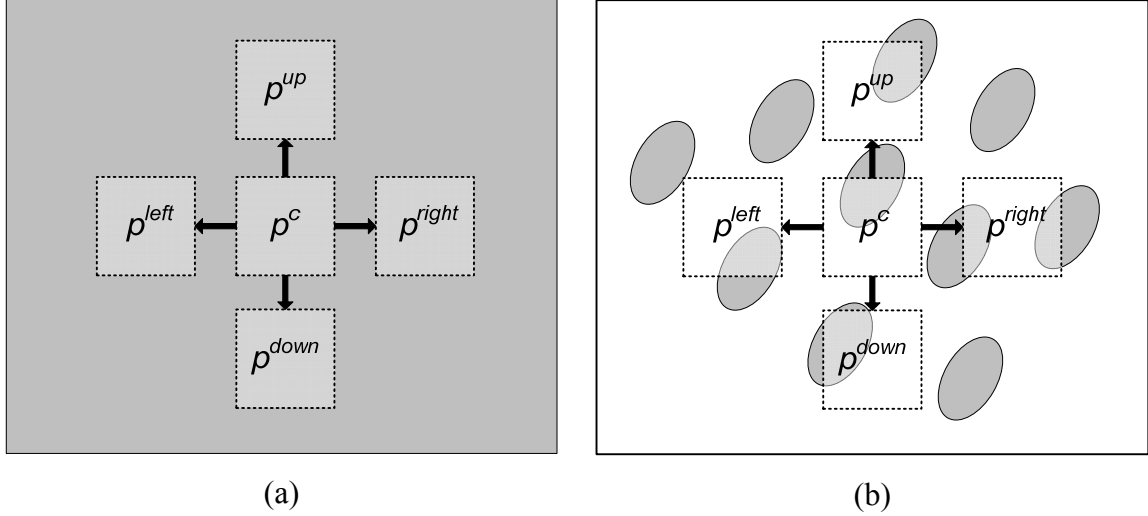
Figure 8: Texture cue devised in Erdem and Tari [38]. (a) In a piecewise-constant region of the image, the similarity distributions between the central patch, p^c, and each of the patches shifted in 4 directions are similar to one another. Thus the p-value produced by the Wilcoxon Mann-Whitney test on D^{left} and D^{right} (similarity distributions to patches on the left and right) is high and on D^{up} and D^{down} is high. (b) In a textured region of the image, the similarity distributions between p^c and each of the shifted patches differ from one another. Thus the p-value produced by D^{left} and D^{right} is low, and by D^{up} and D^{down} is low.

The fourth cue, local scale, preserved texture by preventing excessive smoothening in these regions. Using a patch around a pixel p, the median of the differences between the gradient magnitude at pixels in the patch and the median of the gradient magnitudes is determined. Texture regions will have large median gradient difference values (due to variation in the gradient magnitude in texture regions). This value is then used in an exponential decay function with a decay rate such that high values of the median difference results in the diffusivity warping towards 0 so that texture regions have feature preservation (no smoothening in these regions). After each cue is calculated, the cues are then combined by simply taking the product of the measures.

Kokkinos *et al* [59] proposed a spatially adaptive texture estimation measure through an amplitude/frequency modulation model of images that allows for a probabilistic discrimination between edges, textured and smooth regions. In [59], a texture cue, a loosely defined context-based classifier cue, and an intensity cue were used to distinguish between texture edges and edges between different objects. Only the latter edges were then used to dampen the curve evolution and define the segmentation boundary. Gilboa *et al* [45] presented a graph-cut based segmentation framework with spatially varying regularization through edge weights in the graph using a gradient magnitude-based cue. Malik *et al* [69] proposed the Normalized Cuts segmentation framework to regularize segmentation in textured regions through the use of local texture and gradient cues. Their work was also significant for the use of 'cue gating', where gradient information is suppressed in high texture regions through the following:

$$O_G(x, y) = \big(1 - T(x, y)\big)O(x, y) \qquad\qquad (28)$$

where $O(x, y)$ is the orientation cue (estimate of gradient strength), $T(x, y)$ is the texture cue, and $O_G(x, y)$ is the gated orientation cue. In addition, we note that the methods of [80, 35, 69, 38] that use continuous segmentation frameworks only modulate the regularization term and leave the data fidelity term as is.

1.5.1.1 Deficiencies in Existing Methods

The key problem with the methods described in Section 1.5.1 is that none consider object structure when controlling the regularization, particularly object curvature. As demonstrated in Section 1.1, curvature can often play a key role in distinguishing objects. For example, white matter in MR images has a highly undulating surface that requires low regularization in regions of high curvature. In addition, no method has a strong method for estimation of image noise. The method for handling noise in Erdem and Tari is a weak measure since a noise model is not considered, and instead a measure is formulated based on the connectivity of edges, which can be easily disrupted by texture edges and by stronger forms of noise (such as AWGN). One last problem with the methods of Section 1.5.1 is that each method is tied to a particular segmentation approach and often requires prior knowledge. Our goal in this thesis is to seek a more unified approach to adaptive regularization that can be generalized to many energy minimization segmentation methods with minor modifications to the energy formulation.

In addition, in Chapter 3, we will show improvements obtained from combining the texture cue of Erdem and Tari with our proposed data-driven regularization weight.

1.5.2 Estimation of Curvature

A key section of this thesis (Chapter 3) focuses on a structural measure of object curvature. Existing curvature methods focus on estimating either the curvature of segmentation contours, or estimating the curvature of iso-intensity contours in the image in an attempt to estimate object curvature. Kitchen and Rosenfeld [57] introduced a basic 2D curvature measure for iso-intensity contours in an image (posing the problem as that of corner detection) which has been used by many others [36, 56, 52]. In [57], the curvature of a 2D contour $C(x,y)$ was determined by measuring the rate of change between the gradient angle at a point in the image, $\theta = \tan^{-1}(C_y, C_x)$, and the unit vector perpendicular to the gradient direction, $\mathbf{n}_\perp = (-\sin\theta, \cos\theta)$, to form the curvature estimate K as follows

$$K = \frac{\partial \theta}{\partial \mathbf{n}_\perp} = \frac{C_{yy}C_x^2 - 2C_{xy}C_xC_y + C_{xx}C_y^2}{\left(C_x^2 + C_y^2\right)^{3/2}} \tag{29}$$

as shown in Figure 9.
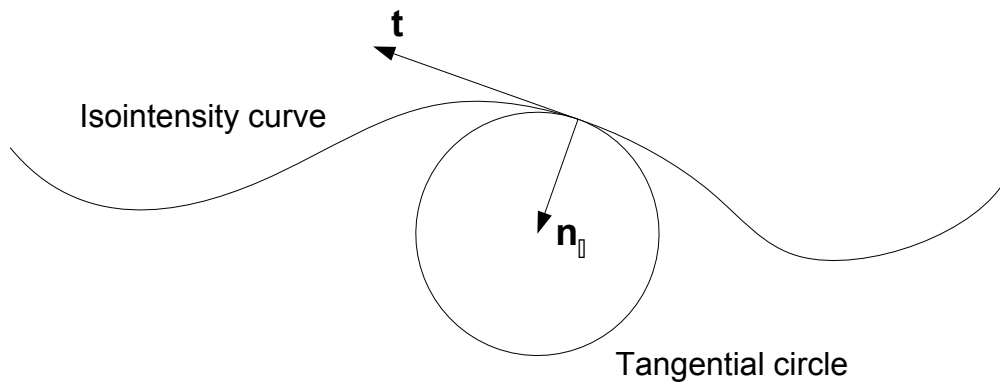
Figure 9: Curvature of an isointensity curve measured by the rate of change between the gradient angle and unit normal vector $\mathbf{n}_\perp$.

Segmentation methods such as the classical snakes active contour model [56] use this curvature term as an external energy to attract the segmentation towards high curvature regions of the image. The method calculates the curvature of the level contours and assigns a high energy for contours with high curvature. It is important to differentiate this from weighting the regularization by using the curvature. Instead of adapting how the regularization changes based on how the curvature changes, these methods instead use a curvature term in the energy functional such that the optimum contour will have low curvature if the level of regularization is high (reducing high curvature erratic regions). The methods do not verify if the high curvature regions are in fact due to noise or due to structure however.

Cohen *et al* [27] presented a method to incorporate curvature information into the motion of deformable objects, and track high curvature points through a time series of images. The motivation was that high curvature points in an image are anatomically important and should be matched correctly to the deformable curve. Hermann and Klette [52] presented a survey of various common 2D curvature estimators using differential geometry. However, these methods involve knowledge of the entire curve and were not explicitly based on image intensity information. All methods involved a preprocessing step where the longest straight line segment at each pixel that is tangent to the curve is mapped such that the curve is discretized into a series of straight line segments.

Many corner detection methods have been proposed through the use of the Curvature Scale Space (CSS), first described in [73] and expanded in [49, 103, 107]. These methods used the same basic corner detector as [57, 36, 56, 52], but also

determined the optimum scale that the curvature should be determined in. Depending on scale, the resulting curvature can change dramatically. If the goal is to highlight any sharp point of curvature in the image, the curvature must be determined in all scales. The CSS methods approached this problem by computing the curvature at the highest scale σ_{high}, thresholding the values to get possible corner candidates, and then tracking the corners to the lowest scale for localization. These methods suffer from the issue that the detection of the corner candidates is done in a single scale, σ_{high}, and involve the use of a global threshold to determine candidate corner regions.

Other methods, such as Zhang *et al* [104] and Awrangjeb *et al* [6], focus on detecting corner points, which are defined as point of extreme curvature. However, these methods do not provide an accurate measure of curvature for non-extreme regions of the image (i.e., no continuous measure of curvature is provided, only a binary measure) which is not useful for the purposes in this thesis.

The curvature estimation method most closely aligned to the method proposed in this paper is that of Lindeberg [68, 67] where a multi-scale approach is used for curvature detection and where a continuous estimate of curvature is provided for iso-intensity contours in the image. This method will be outlined in detail in Chapter 3.

1.6 Thesis Contributions

In this thesis, we present two methods for spatially adaptive regularization for general energy-minimization segmentation frameworks and present validation results on a variety of current segmentation techniques. We focus on methods that determine the

regularization in an automated manner without the need for prior information. Additionally, we focus on using a single regularization weight as opposed to multiple weights which we leave to future work (see Section 4.2). In particular, we present the following:

<u>Contribution 1</u>

In Chapter 2, we present a novel method for deriving globally optimal regularization weights for minimum-path segmentation methods. Through extensive validation with synthetic, medical, and natural scene datasets, we investigated the correlation of the corresponding results to human perception. Our results suggested that global optimums do not always result in 'correct' segmentations as determined by experts. We found the segmentations from globally optimal weights to suffer from bimodal behaviour of the weights, poor robustness to noise, and poor reflection of underlying image characteristics, which resulted in large inaccuracies. This contribution was published in the following work:

Rao, J., Hamarneh, G., Abugharbieh, R.: Adaptive contextual energy parameterization for automated image segmentation. In: ISVC. Volume 5875-I. (2009) 1089-1100.

<u>Contribution 2</u>

To address the shortcomings of our first contribution, in Chapter 3 we proposed the novel concept of locally adaptive regularization weights that incorporate both structure and noise reliability. These contextual weights are the first to adapt regularization in a manner that seeks to preserve structurally important high curvature features in images, and to estimate the reliability of image evidence through a novel and

robust signal-to-noise ratio (SNR) measure. Our work is also the first to propose general regularization weights that can be incorporated into any energy minimization segmentation method with minor changes, unlike other methods that tie regularization weights to a specific segmentation method. In addition, rather than using curvature measures as internal or external energy, we use these measures to adapt the importance of *any* general internal or external energy term in the optimization process. Through validation using synthetic, medical, and natural scene datasets, we demonstrate how the contextual weights produce significantly more accurate results than the traditional approach of least-error spatially-fixed weight, the previously discussed approach of globally optimal weights, and the existing adaptive regularization weights of Erdem and Tari [38]. We demonstrate the generality and applicability of the contextual weights by incorporation into discrete and continuous segmentation frameworks. In addition, we show how our method can easily incorporate contextual cues proposed in other works, such as texture cues. This contribution was published in the following works:

Rao, J., Hamarneh, G., Abugharbieh, R.: Adaptive contextual energy parameterization for automated image segmentation. In: ISVC. Volume 5875-I. (2009) 1089-1100.

Rao, J., Abugharbieh, R., Hamarneh, G.: Adaptive Regularization for Image Segmentation Using Local Image Curvature Cues. Submitted to: ECCV (2010).

In the concluding Chapter 4, we will discuss limitations and areas of future work in this thesis.

Chapter 2

Globally Optimal Regularization Weights

This chapter first discusses the motivation for determining the globally optimal regularization weight, and then discusses the implementation of a graph search in 3D space with references to existing techniques. This is followed by validation on standard synthetic and real databases, and concludes with an analysis of the performance of the method and an explanation of the limitations of this method.

2.1 General Framework

A theoretically appealing and intuitive approach for setting the regularization weight is to determine the globally optimal weight by incorporating the weight into an optimization process that determines the globally optimal segmentation. One such way to achieve this is to focus on graph-based minimum-path approaches that allow for

modification of the graph such that additional dimensions (to optimize along) can be added. In our work, we will in particular focus on a simple minimum path approach where we will insert our regularization weight in a convex manner. The problem of optimizing minimum-path approaches for the path spatial coordinates and a third variable has been significantly researched in the field of segmentation of tubular structures, and in particular, anatomical vessel segmentation. One such approach is Li and Yezzi [64] who optimized for the spatial path and vessel radius variable simultaneously. Additionally, Wink *et al* [100] used Djikstra's algorithm [34] to solve for the globally optimum medial (centerline of the vessel) path and vessel radius. Most recently, Poon *et al* [81] solved for the optimum vessel radius in conjunction with the median path using a modified version of Livewire, titled LiveVessel. LiveVessel focused on modifying the local cost term for vessel segmentation by adding the vessel radius as variable that played a large role in vessel cost terms.

As discussed in Section 1.2, 2D minimum-path segmentation methods optimize a path or contour C which consist of nodes p with spatial coordinates (x, y). From each node p, the path choices are to a neighbouring node $q(x', y')$. In [81], the 3D expansion of the graph search shifts the optimization space to $(x, y, scale)$ by expanding the path choices from each node $p(x, y, scale)$ to a neighbouring node $q(x', y', scale')$ with the restriction that $x \neq x'$ and $y \neq y'$ in order to form a contour with consecutively linked nodes. Through this method, the globally optimum vessel radius (scale) at each node is guaranteed along with the globally optimum path.

In our implementation, we will differ from these methods by using a convex form of the local cost functional. Our formulation employs energy-minimizing boundary-based segmentation, where the objective is to find a contour that correctly separates an object from background. We embed a parametric contour $C(q) = C(x(q), y(q)): [0,1] \to \Omega \subset \Re^2$ in image $I: \Omega \to \Re$. We use a single adaptive weight $w(q) \in [0,1]$ that varies over the length of the contour and re-write (2) as:

$$E\big(C(q), w(q)\big) = \int_0^1 w(q)E_{int}C(q) + \big(1 - w(q)\big)E_{ext}\big(C(q)\big)dq \tag{30}$$

where

$$E_{ext}\big(C(q)\big) = 1 - \big|\nabla I\big(C(q)\big)\big| / \max_\Omega \big|\nabla I\big(C(q)\big)\big| \tag{31}$$

penalizes weak boundaries and

$$E_{int}\big(C(q)\big) = \left|\frac{\partial C(q)}{\partial q}\right| \tag{32}$$

penalizes longer and jagged contours.

Our approach for determining the globally optimum regularization weight is to optimize E in equation (30) for the weight $w(q)$ *itself* in addition to optimizing the contour. In our discrete setting, this involves optimizing for $\tilde{C}(q) = \big(x(q), y(q), w(q)\big)$. The external energy term of equation (31) remains the same, but the internal energy term of equation (32) now represents the length of the contour in the 3D space (x, y, w) where large

changes in the regularization weight are penalized with a higher internal energy. We will explain the optimization process next.

2.2 Optimization Process

2.2.1 Djikstra's Method in 3D

To minimize E with respect to $C(q)$ in (30), we model a 3D graph of dimension $M \times N \times n_w$ where the image is of dimension $M \times N$ and we discretize the weight into n_w levels. Each vertex v_i represents pixel coordinates and a weight (x, y, w). Graph edges $e_{ij} = \langle v_i, v_j \rangle$ represent vertex connectedness (e.g. 24-connectedness in 3D graphs). A local cost

$$c_{ij} = w_j \, E_{int} \left(v_i, v_j \right) + \left(1 - w_j \right) E_{ext} \left(v_i \right) \tag{33}$$

is assigned to each edge e_{ij}, where $E_{int} \left(v_i, v_j \right)$ is the Euclidean distances between v_i and v_j. The contour that minimizes the total energy

$$E_{total} = \sum_{e_{ij} \in C} c_{ij} \tag{34}$$

represents the optimal solution for the segmentation. Note that the optimal path $\tilde{C}(q) = \left(x(q), y(q), w(q) \right)$ cannot pass through the same $\left(x(q), y(q) \right)$ for different w, i.e. only a single weight can be assigned per pixel. Our graph search abides by this simple and logical constraint. The optimal $C(q)$ and $w(q)$ that globally minimize (30) are calculated using dynamic programming on this (x, y, w) graph.

2.2.2 Implementation Details

The 3D graph search is implemented in a similar manner as the 2D graph search described in Section 1.2.2.1 with the additional dimension contributing to an $O(N^2)$ search process. The key difference is the expansion of a node's set of neighbouring nodes from 8 to $8n_w$ where n_w is the number of discretized weight levels. In our implementation, we selected $n_w = 11$ in the range $[0,1]$ with the distance between weight levels as $l_w = 0.1$.

1. Initialize a list of node costs to infinity. Initialize a list of visited nodes to empty. Initialize a queue of nodes to empty.

2. Set the initial node (start point) as visited, place it in the node queue, and assign it a cumulative cost of zero.

3. Retrieve the node in the queue with the smallest cumulative cost and term this the current node. Calculate the cumulative cost to each of the current node's unvisited neighbours by adding the cost of the edge e_{ij} from current node i to the neighbour node j to the cumulative cost of the current node i. If this new cost $c'_j = e_{ij} + c_i$ is less than the previously recorded cost, i.e. $c'_j < c_j$, than the old cost is overwritten with the new cost. Each neighbour is added to the node queue if it is not already there. Once all neighbours of the current node has been visited, mark the current node as visited.

4. Re-sort the node queue such that the node with the lowest cumulative cost is first out of the queue.

5. If the node queue is empty (all nodes has been visited), the algorithm is finished. Otherwise, repeat steps 3 and 4.

The process is illustrated in the flowchart in Figure 10.
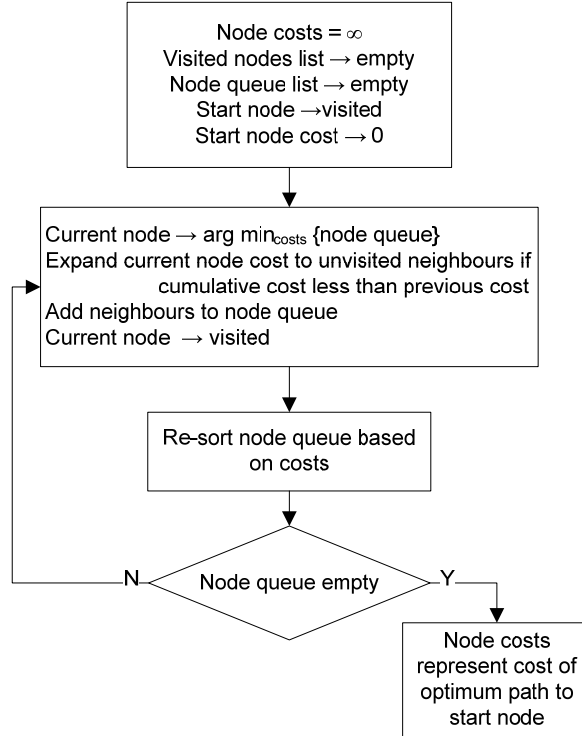
Figure 10: Graph search algorithm using Djikstra's method.

The minimum-path segmentation method described in Section 1.2.2.1 is an interactive method since it requires user-entered seedpoints around the boundary of the object of interest. Since our focus in this thesis is not on user interactivity, we selected equally spaced seedpoints around the object boundary that were automatically determined from the ground truth segmentation. The seedpoints consisted of the spatial coordinates and the regularization weight at that node. The spatial coordinates were determined from the ground truth contour, and the regularization weight was simply set to 0.5. For each pair of seedpoints, the minimum-path contour was determined from the 3D graph search.

We used the same seedpoint set for the globally optimum weight segmentation and for the least-error fixed weight segmentation for proper comparison of the results.

2.3 Validation and Results

We demonstrate the performance of the globally optimum regularization weights on a wide range of datasets, including synthetic images, medical datasets, and natural images. In addition, we validate the method quantitatively using analysis of variance (ANOVA). All our validation is performed against ground truth segmentation determined either functionally (for synthetic images) or by manual segmentation by experts.

2.3.1 Performance Criteria

Our measure of error for our synthetic dataset is to compare the vertical (along the y-axis) distance between the contour and the sinusoidal functional that is used to make the synthetic image (See Section 2.3.2 for details about how the synthetic dataset is devised), as follows:

$$H_E = \max_{q(x,y)} |y - F_S(x)| \tag{35}$$

where (x, y) are the spatial coordinates of each node q in the contour $C(q)$, and $F_S(x)$ is the sinusoidal function representing the synthetic image.

Our measure of error for the medical and natural datasets was to form a closed segmentation mask from the contour produced by the method, and to compare this mask to the ground truth segmentation by computing the Dice Similarity Coefficient (DSC)

44

measure that estimates the agreement in pixel labels between two segmentations as follows:

$$dice_{A,B} = 2\frac{(area_A \cap area_B)}{(area_A + area_B)} \tag{36}$$

where $area_A$ represents the binary segmentation mask produced from method A and $area_B$ represents the binary segmentation mask produced from method B. For the medical and natural datasets, the ground truth segmentation was that produced by manual tracing.

We benchmarked the globally optimum weights against the conventional method of determining the least-error (or maximum dice similarity) spatially-fixed regularization weight. This is accomplished by determining:

$$w_{fixed} = \arg\max_{w \in [0,1]} dice_{S_w, S_G} \tag{37}$$

where S_w is the segmentation produced by the weight w and S_G is the ground truth segmentation. The weight is discretized into 11 levels with a step size of 0.1 between each level.

We determined the segmentations for 25 trials for each image. This is of importance for synthetic images where noise was added and thus the noise profile is determined randomly for each image. We thus generated 25 noise profiles for each image and determined the segmentation for each image. The error is determined by averaging the resulting 25 segmentation accuracies. To determine if the globally optimum weights differ significantly from the least-error fixed weight, we then performed ANOVA over

the set of 25 segmentation accuracies for each image and analyzed the resulting p-value. Computationally, the 3D graph search method required less than 7 minutes for a 768×576 image when run on a Pentium 4 (3.6GHz) machine using MATLAB code.

2.3.2 Synthetic Data Validation

We first tested the method on a dataset of synthetically produced images. Through the synthetic images, we tested extreme cases of object boundary variability, noise variation, edge strength variation, and object curvature variation. We first modelled an object boundary as a sinusoidal function with spatially-varying frequency to simulate varying contour smoothness conditions. The images were produced by the functional:

$$F_S(t) = A \sin\left(\frac{t^2}{10}\right) \tag{38}$$

where the sinusoidal is parameterized by $t = [C2\,\pi f_0, C2\pi f_1]$ and will vary in frequency from f_0 at the start and f_1 at the end of the contour. A represents the amplitude (set to 2) and C represents the frequency width (set to 10). We also added spatially-varying (non-stationary) AWGN patterns of increasing variance. We also spatially varied the gradient magnitude of the object boundary across each image by applying Gaussian blurring kernels at different scales in different locations. We created 16 of these synthetic images carefully designed to cover extreme shape and appearance variations. Two synthetic images with the resulting segmentations are shown in Figure 11. The contour obtained using the globally-optimal weights method is shown in *green*, along with the contour obtained using the spatially-fixed regularization weight from (37) shown in *red*.

46
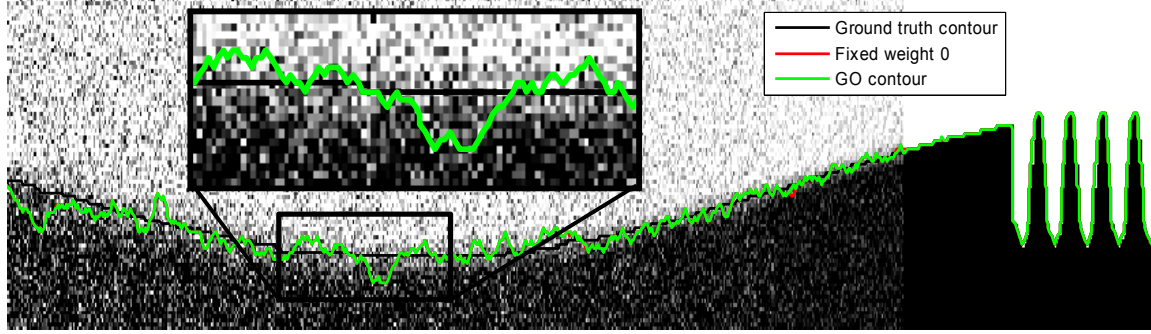
(a)
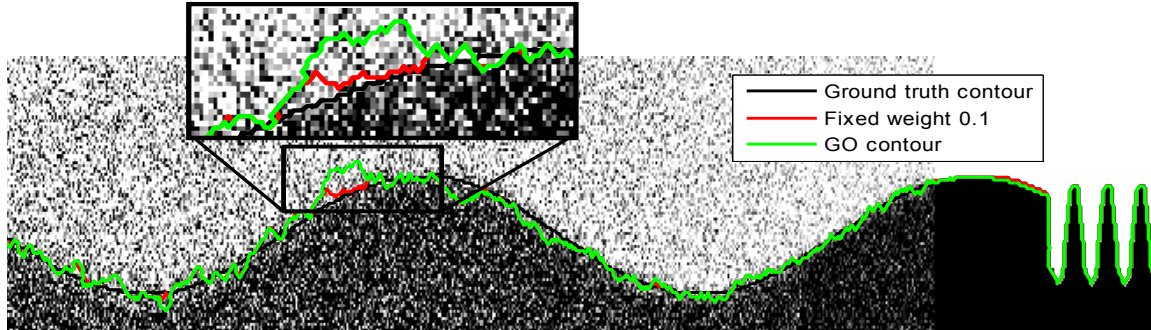
(b)

Figure 11: Segmentation of synthetic images with decreasing noise variance and decreasing Gaussian blurring from left to right. (a) Low curvature case, and (b) higher curvature case. Contours produced by globally optimal weights (*green*), least-error fixed weight (*red*), and ground truth contour (*black*). Globally optimal weights are 0 over the whole contour, resulting in little difference between the contours and poor regularization to noise.

In addition, Table 1 presents the modified Hausdorff error measure of (35) for segmentations produced by both regularization weights for all images. The error values represent the mean H_e for 25 noise realizations for each synthetic image. The least-error fixed-weight method had a mean error of 12.05 ± 1.6 while the globally optimal weight method had a mean error of 33.06 ± 3.66.

Table 1: Average error over 25 noise realizations per image produced by least-error fixed weights and globally optimal weights for synthetic set of data. Segmentations from the globally-optimal weights have significantly *less* accuracy than segmentations from the least-error fixed weight. Average *p*-value $\ll 0.05$.

Image #	Mean H_E error over 25 segmentations	
	Fixed weights	G.O. weights
1	13.02	17.68
2	15.04	45.24
3	8.92	17.4
4	13.8	38.6
5	6.8	7.16
6	12.96	42.2
7	14.4	42.4
8	13.28	40.92
9	11.8	35.72
10	13.04	47.68
11	14.84	40.04
12	6.64	7.92
13	10.62	27.24
14	12.74	41
15	11.92	44.2
16	13.02	33.64

From Figure 11, it is clear that the globally optimal (GO) weights produce a more erratic segmentation when compared to the least-error fixed weight segmentation. Although the GO weight is accurate for the noise-free regions of the image, the regions with high noise have insufficient regularization. The quantitative results in Table 1 confirm that the GO method produced significantly poorer results when compared to do the least-error fixed weight. We can further analyze the behaviour of the GO weight by plotting the weights over the contour as shown in Figure 12.
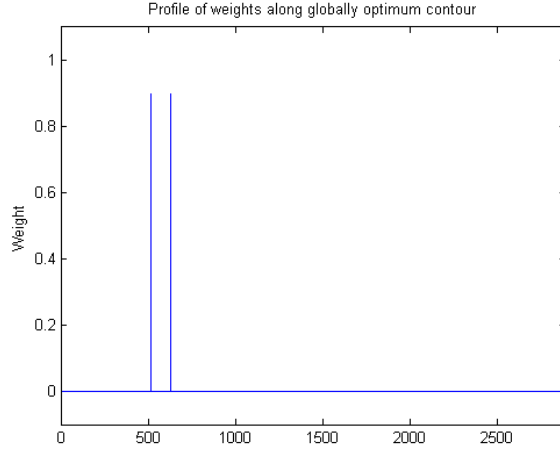
Figure 12: Profile of globally optimal weights along contour of Figure 11(a). Weights are either 0 or 1 indicating bimodal behaviour.

The weight is significantly bimodal, oscillating between 0 and 1. We will discuss the difficulties the globally optimum weight has with noise, and the cause of the bimodal weight behaviour, in Section 2.4.

2.3.3 Medical Data Validation

We also tested the globally optimal weights on various medical datasets validated with expert ground truth segmentations. We first used a dataset of sagittal slices from MR scans of the brain taken for 52 subjects. The focus of these images are on the corpus callosum (CC) structure. Segmentation of the boundary of the CC is difficult due to the weak edge separating the CC from the fornix, as labelled in Figure 13(a).

49

(a)   (b)
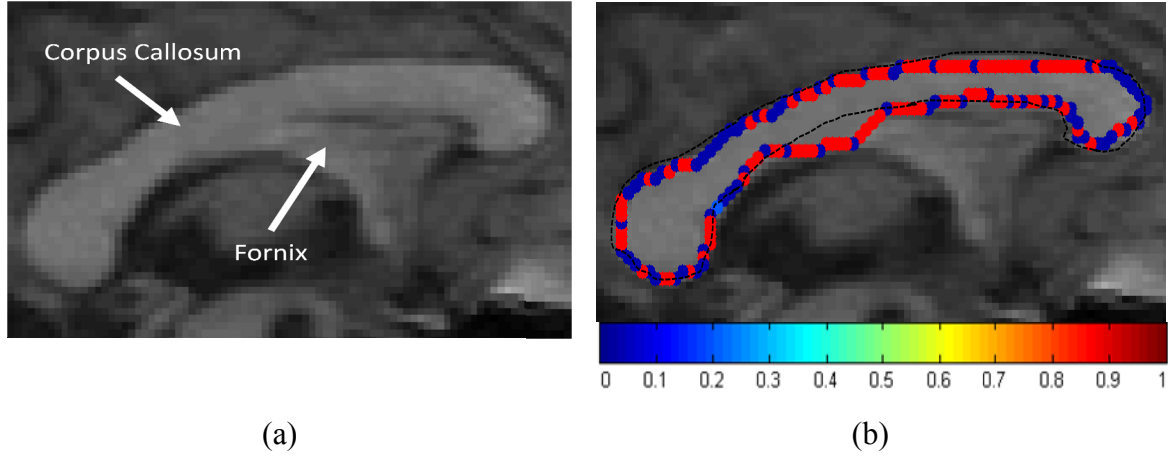
Figure 13: Segmentation of corpus callosum (CC) structure from sagittal MR slice. (a) Original image showing the weak delineation of the CC and fornix structures as labeled. (b) Resulting segmentation from globally-optimal weights with contour colour indicating weight level according to the labeled colormap, and with ground truth shown as dashed *black* contour. The bimodal behavior of the weights leads poor segmentation of the weak edge boundary.

The fornix section requires a high regularization weight to allow the contour to cut through the weak gradient rather than following the strong gradient of the fornix itself. However, the regularization weight must be low in order to accurately segment the sharp corners on both sides of the CC structure. Figure 13(b) demonstrates the performance of the globally optimum weight segmentation on one such CC image from this dataset using the colour mapping as shown where blue corresponds to a weight of 0 and red corresponds to a weight of 1. The globally optimal weights produced a poor segmentation, particularly due to the bimodal behaviour of the weights. Note how the contour obtained using globally optimal weights exhibits an optimal, yet undesirable, bimodal behaviour (either blue or red in Figure 13(b)) completely favouring only one of the terms at a time.

In addition, we tested the method on a synthetic dataset of MR brain scans taken with various modalities from the BrainWeb database [26, 63, 62, 31]. In particular, spatially varying noise can be created with this dataset by adjusting the noise level and

intensity inhomogeneity level. The segmentation of the white matter using the T1 modality is shown in Figure 14. These cases present a difficult scenario since the images contain weak edge strengths, highly varying object boundaries, and varying noise and intensity levels. The contour from the globally optimal weights, shown in *cyan*, is attracted to the strong edge boundary between the grey matter and cerebral fluid region rather than the weak edge boundary between the white matter and grey matter. The spatially-fixed weight produces a contour with excessive regularization that fails to capture cortical folds.

(a)   (b)
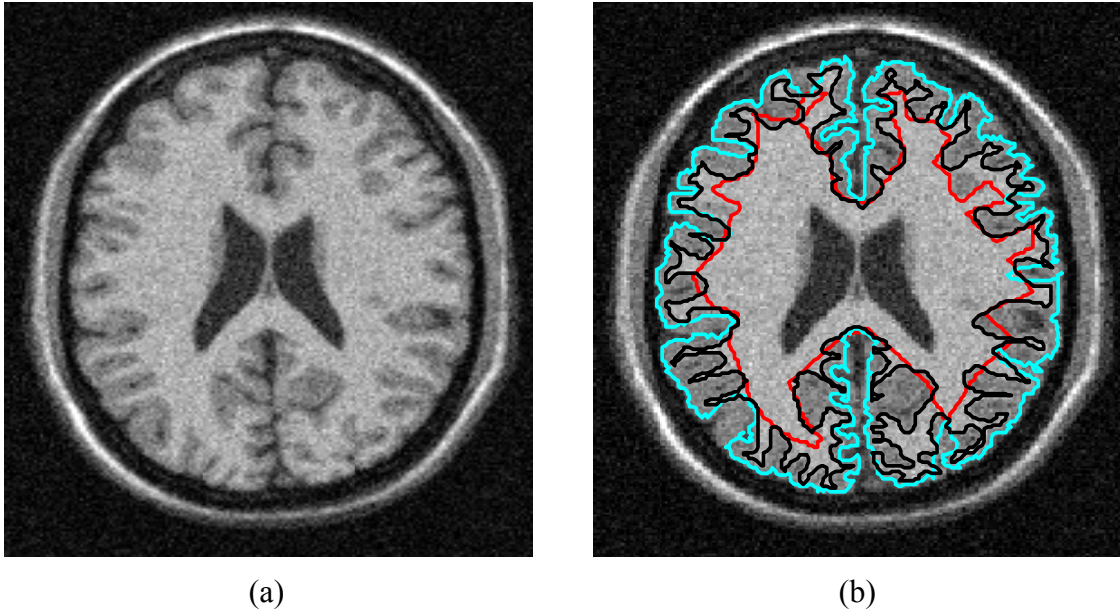
Figure 14: Segmentation of cortical boundary. (a) BrainWeb T1 mid-volume coronal slice with 9% noise. (b) Contours from globally optimal weights (*cyan*), least-error fixed weight (*red*), and ground truth (*black*). Globally optimal weights result in contour attraction to the strong edge between grey matter and cerebral fluid region rather than edge between white matter and grey matter. Fixed weight contour has excessive regularization in high curvature regions.

In Figure 15, we present the DSC of segmentations from the globally optimal weight and the spatially-fixed weight on images from the BrainWeb dataset using T1, T2, and proton density (PD) modalities with increasing noise levels (where segmentations

were determined for 25 noise realizations per noise level). As with the synthetic images, we averaged the DS over 25 segmentations for each image in each dataset. In the low noise cases, the globally optimum weight and the spatially-fixed weight produce accurate segmentations. However, as the level of noise increases, the segmentation accuracy degrades due excessive or inadequate regularization by both methods.

(a)   (b)   (c)
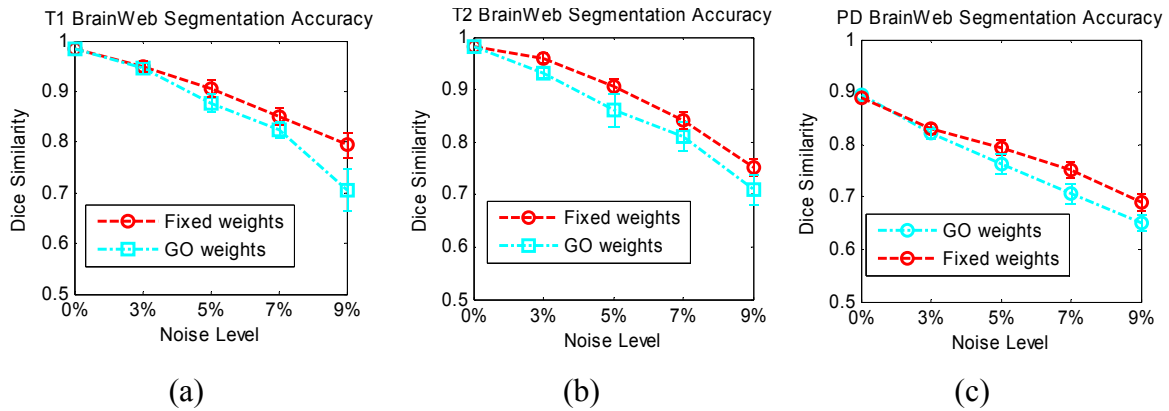
Figure 15: Dice similarity coefficient (DSC) of segmentations from globally-optimal weights (GO) and least-error fixed weights for segmentation of white matter in coronal BrainWeb slices using the (a) T1 modality, (2) T2 modality, and (3) PD modality. As the level of noise in the slices increases, the globally optimal weights perform poorer that the least-error fixed weights due to low regularization (from noise mistaken as high external energy). DSC values are averaged from segmentations of 25 noise realizations per noise level.

Our findings indicated poor performance by the globally optimal weights. Further qualitative and quantitative performance evaluation of the GO weights, in comparison with the data-driven weights (discussed in Chapter 3), are provided in Section 3.8.

2.3.4 Natural Scene Data Validation

We also tested the globally optimal weights on images of natural scenes from the following datasets: the McGill Calibrated Colour Image Database [75], the PASCAL object recognition database [40], and through the ImageNet database [32, 33] and validated through manual segmentations.

We first present the segmentation results on natural flower and leaf images taken from [75]. These images present a segmentation challenge because of the highly varying curvature properties of the object, and from the varying edge strength and occlusions from the background. Figure 16 presents segmentation results from a leaf obscured by water, and Figure 17 presents results from a flower with a highly varying object boundary. In both cases, the globally optimum method performs poorly due to inadequate regularization in regions of weak edge boundaries. The bimodal behaviour of the weights results in predominately zero levels of regularization, unlike the least-error spatially fixed weight which formed a more accurate segmentation using $w_{fixed} = 0.4$ for Figure 16 and $w_{fixed} = 0.3$ for Figure 17.

(a)   (b)
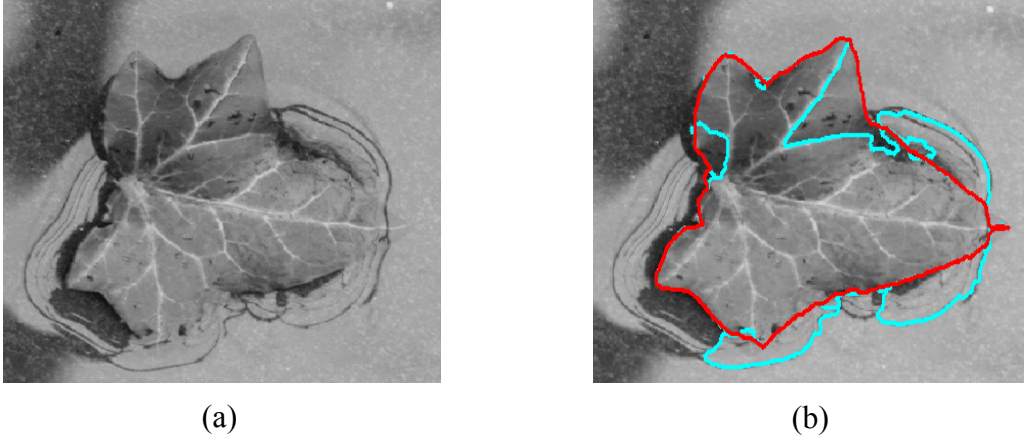
Figure 16: Segmentation of leaf image from McGill database [75]. (a) Original image, and (b) contours produced by globally optimal weights (*cyan*) and least-error fixed weights (*red*). Globally optimal weight is predominately 0 and produces segmentation with low regularization in regions of leaf submerged by water. Fixed weight cannot accurately segment the top tip of the leaf.

(a)   (b)

Figure 17: Segmentation of flower image from McGill database. (a) Original image, and (b) contours produced by globally optimal weights (*cyan*) and least-error fixed weights (*red*). Inadequate regularization by both methods results in oversegmentation and leakage into the background.

Figure 18 presents the quantitative results from images taken from the McGill, PASCAL, and ImageNet databases. As with the previous examples, the DSC represents the average over 25 segmentations for each image where each segmentation is produced from different seed placements (all equally spaced over the object boundary). Additional qualitative and quantitative results produced by the globally-optimal weights are presented in Section 3.8 and compared to the data-driven weights proposed in the next chapter.
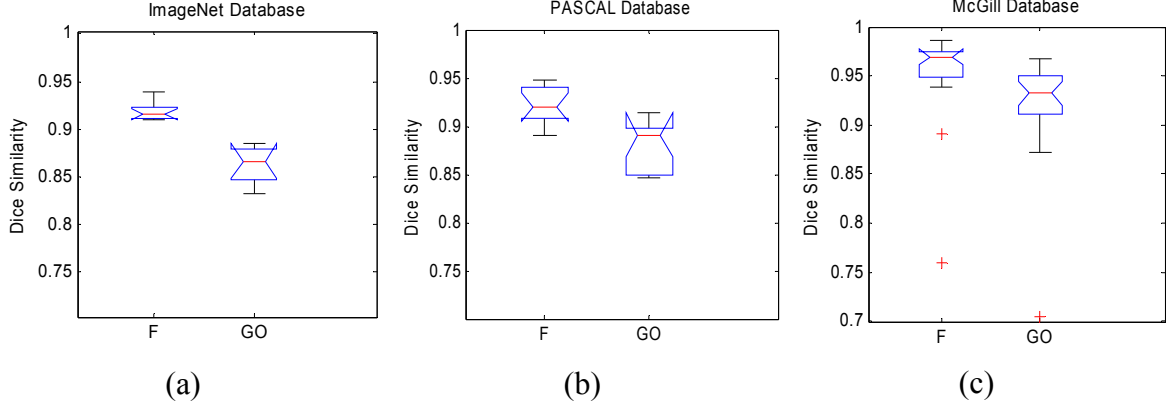
Figure 18: DSC of segmentations from least-error fixed weights (F) and globally-optimal weights (GO) on (a) 8 images from ImageNet database, (b) 11 images from PASCAL database, and (c) 24 images from McGill database. Average error over dataset (and over 25 segmentations for each image) was F = 0.9188, GO = 0.8629 for (a), F = 0.9241, GO =0.8782 for (b), and F = 0.9402, GO = 0.9024 for (c). All *p*-values $\ll 0.05$. Globally optimal weights produce segmentations significantly less accurate than the least-error fixed weight method.

2.4 Deficiencies in Globally Optimum Weight Model

From the performance of the globally optimum method on the wide variety of datasets, we can conclude that the globally optimum weight segmentation produces poorer results than the spatially-fixed weight segmentation.

The results highlight the three main drawbacks to this globally optimum (in (x, y, w)) method: (i) it encourages a bimodal behavior of the regularization weight; (ii) it does not explicitly encode image characteristics; and (iii) it combines the weight and segmentation optimization into one process, thus reducing the generality of the method.

The first drawback, that the weights are bimodal, means that the weights either take on the extreme values of $w_{GO} = 0$ or $w_{GO} = 1$ but no value in between. This bimodal behavior stems from optimizing (30) when $E_{int}(q) > E_{ext}(q)$ and vise versa. In particular, we observe that the following occurs:

$$w(q) = \begin{cases} 0 & E_{int}(q) > E_{ext}(q) \\ 1 & E_{int}(q) < E_{ext}(q) \end{cases} \qquad (39)$$

A simple proof of the bimodal behavior of the optimal weight is as follows: If the total

energy is represented as in (30) with the adaptive weight $w(q) \in [0,1]$, we can rewrite

the total energy as

$$E = wE_1 + (1 - w)E_2 = w(E_1 - E_2) + 1 \qquad (40)$$

where we assume energies $E_1 \in [0,1]$ and $E_2 \in [0,1]$. Determining the optimum energy

by gradient descent/ascent is done by first determining the rate of change of (40) with

respect to the adaptive weight, and determining maximum/minimum regions of zero rate

of change as follows:

$$\frac{dE}{dw} = E_1 - E_2 = 0 \qquad (41)$$

If $E_1 > E_2$, the slope is positive and the minimum will occur when $w = 0$. Alternately, if

$E_1 < E_2$, the slope is negative and the minimum will occur when $w = 1$. Thus, regardless

of the difference between E_1 and E_2, the optimal weight will always be bimodal, either 0

or 1. Although large changes in the weight are penalized with a higher internal energy,

the resulting behavior is still predominately bimodal since the weight shifts from 0 to 1 in

a few steps rather than in a single step (i.e., from 0 to 0.5 to 1) which results in a lower

internal energy per step.

Bimodal weights cannot address the levels of regularization required in images.

From the results in Section 2.3, the least-error spatially fixed weight was always a value

between, but not including, 0 and 1. Extreme values of regularization weights mean that

the resulting segmentation will either have the full level of regularization, which is a straight line in the case of minimum-path boundary based methods, or have no regularization, which means that the optimal contour will contain any edge evidence in the vicinity of the seedpoints. This lack of regularization results in extremely poor results for images containing regions of high external energy that do not correspond to the object boundary; in particular, for noisy images and for images containing occlusions and weak object edges.

The bimodality of the weights leads to the second drawback of this method: the weights have no relationship to the underlying image characteristics. Even though regularization is essential in regions of high image deterioration, the globally optimal weights method has no way of differentiating between non-noisy, or 'reliable'/trusted, external energy, and unreliably external energy. The result is that high noise is confused as strong edge evidence through the external energy measure, resulting in $E_{ext} < E_{int}$, which forces the globally optimal weight to 0. Essentially, the globally optimal weight is not necessary the 'correct' weight as determined by the user and as determined by the ground truth segmentation. In addition, the globally optimal weight does not encode structural information about the object of interest and does not adapt the regularization weight in accordance to regions of higher curvature or texture. This problem is likely due to deficiencies in the optimization function where the energy functional does not accurately represent the problem to be solved. The inclusion of additional energy terms specifically designed for each image segmentation problem could provide globally optimal weights that reflect correct segmentations. However, this reduces the generality of the segmentation method, and may require energy functionals with multiple

regularization weights, which is outside the scope of this thesis and is left to future work, as discussed in Section 4.2.

The final drawback of the globally optimal weight is that the minimum-path 3D graph search process used to determine the weight is intrinsically combined with the segmentation process. This combination of the weight and segmentation optimization into one process reduces the generality of the method since finding the globally optimal weights for other segmentation frameworks would require significant changes to the energy minimization process.

In conclusion, even though *optimal* with respect to E in (30), the solution proposed in this chapter is *incorrect* and, as we later demonstrate, inferior to the spatially adaptive balancing of energy cost terms based on deriving a relationship between the weights and image characteristics, as discussed next in Chapter 3.

Chapter 3

Data-driven Regularization Weights

In this chapter, we discuss our approach for spatially adapting the regularization weights based on underlying image characteristics that are measured through a series of contextual cues, which we term 'data-driven' regularization. We will first discuss the general framework of the method and how each of the data cues are determined. We will then analyze the limits and performance of each data cue before presenting the validation results on a wide variety of synthetic, natural, and medical datasets. We will demonstrate the generality of the method by incorporating the weights into four popular segmentation methods.

3.1 Overview of method

Our results from Chapter 2 motivated us to devise a relationship between the regularization weights and the underlying image characteristics. We found that the regularization weights need to vary in a manner that respects trusted edge and structurally important regions but disregards regions of high image deterioration. In particular, we will focus on three such measures: estimating the level of noise in the image, the level of trusted edge evidence, and the level of trusted curvature. We will incorporate concepts from cue gating, and we will later present results from incorporating existing data cues to show how our method is able to fit into existing techniques.

We first discuss how regularization should behave under different image characteristic conditions. For noise conditions, regularization should be increased to prevent erratic segmentation behaviour. However, regions of high curvature that correspond to object boundaries should have low regularization to allow the segmentation to capture important structural details. In addition, regions with low edge strength that belong to the object boundary should have high regularization again. Boundary regions with high texture should have high regularization to prevent these texture edges from being included as the contour itself. In each case, we must make sure that the edge, curvature, and texture measure that we devise is a 'trusted' measure; i.e. these measures should be noise-free, which we can accomplish by using the noise measure to gate these measures.

3.2 Noise Evidence Cue

In audio signal processing and compression applications, spectral flatness (SF) is a well known Fourier domain measure used to estimate noise [54, 93]. SF exploits the property that white noise exhibits similar power levels in all spectral bands and thus results in a flat power spectrum, whereas uncorrupted signals have power concentrated in certain spectral bands and thus result in a more impulse-like power spectrum [46]. This is shown in Figure 19 for two 1D signals, one clean (Figure 19(a)) and one noisy (Figure 19(c)).
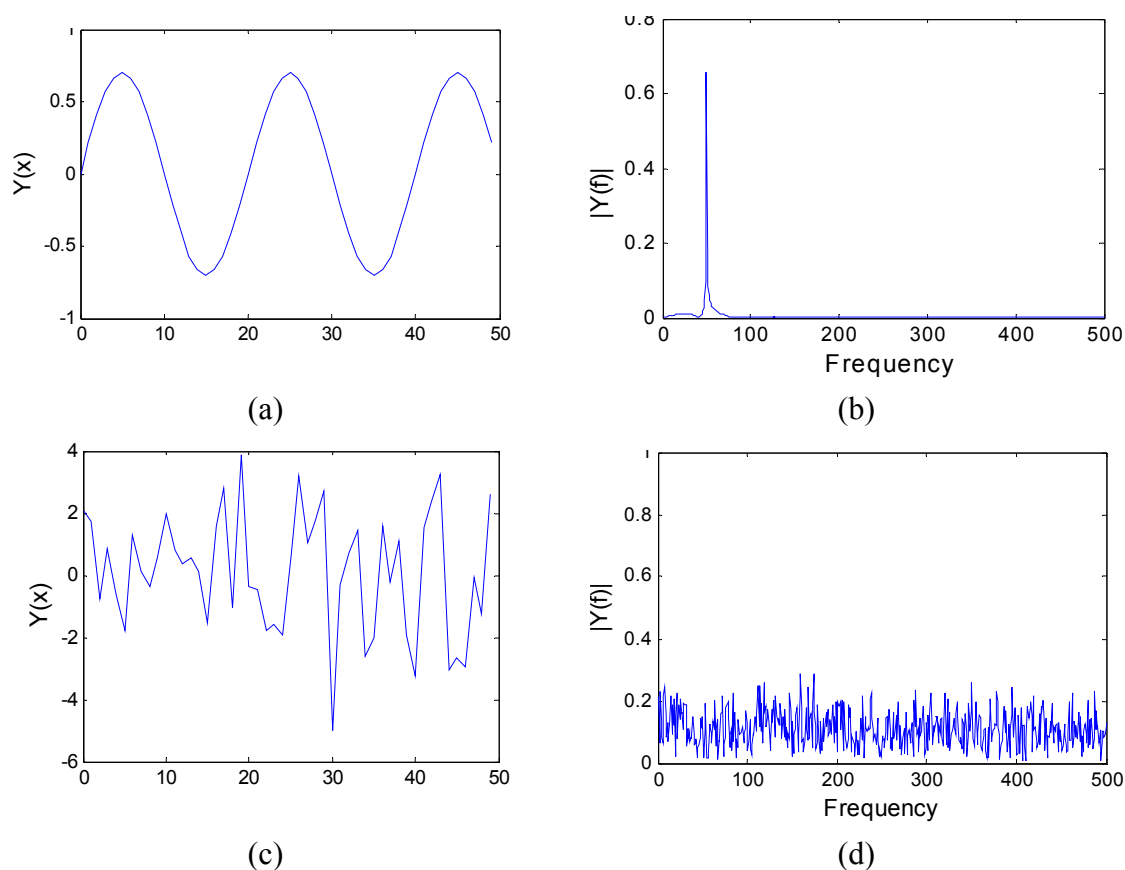
Figure 19: Spectral behavior of 1D signals. (a) Clean signal in spatial domain and (b) corresponding impulse-like behavior in spectral domain. (c) Noisy signal in spatial domain and (d) corresponding flatness in spectral domain. Noisy signals can be identified by examining flatness in the spectral domain.

61

The original SF measure in audio processing is as follows:

$$N_{1D}(x) = \frac{\exp\left(\frac{1}{2\pi}\int_{-\pi}^{\pi}\ln S_{1D}(\omega)d\omega\right)}{\frac{1}{2\pi}\int_{-\pi}^{\pi}S_{1D}(\omega)d\omega} \tag{42}$$

where $S_{1D}(\omega) = |F_{1D}(\omega)|^2$ is the power spectrum of the signal (where $F_{1D}(\omega)$ are the Fourier coordinates of the signal), and ω is the frequency. The measure N_{1D} essentially compares the ratio between the geometric mean (numerator of (42)) and the arithmetic mean (denominator of (42)), and is a general measure of flatness. If the Fourier spectrum contains an impulse, meaning that the spatial-domain signal is clean, the arithmetic mean will be much larger than the geometric mean, and $N_{1D} \to 0$. If the original signal is noisy, the Fourier spectrum will be flat, and the arithmetic mean will be very similar to the geometric mean, resulting in $N_{1D} \to 1$.

Assuming additive white noise, uncorrelated between pixels, we extend the SF measure to 2D and estimate the spatially-varying noise levels $N(x, y)$ as

$$N(x,y) = \frac{\exp\left(\frac{1}{4\pi^2}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\ln S_W(\omega_x, \omega_y)d\omega_x d\omega_y\right)}{\frac{1}{4\pi^2}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}S_W(\omega_x, \omega_y)d\omega_x d\omega_y} \tag{43}$$

where $S_W(\omega_x, \omega_y) = |F(\omega_x, \omega_y)|^2$ is the 2D power spectrum of a local window in the image, $F_W(\omega_x, \omega_y)$ is the Fourier spectrum of the image window and (ω_x, ω_y) are the two spatial radian frequencies. As with the 1D SF measure, $N(x, y) \in [0,1]$ where $N(x, y) = 0$ corresponds to low noise regions and $N(x, y) = 1$ corresponds to high noise regions. This noise measure responds best to white noise-like patterns (i.e. occupying a

very wide and flat spectrum) which are typically the most difficult type of noise. However we note that this method cannot address speckle noise (multiplicative Gaussian noise). We estimate the *local* noise level by determining $N(x, y)$ using local windows around each pixel and using those windows to calculate the Fourier spectrum. Figure 20(a) shows a synthetic image with added AWGN with increasing variance from right to left, and Figure 20(b) shows the resulting measure $N(x, y)$. As the noise variance decreases, the noise measure decreases in accordance. In addition, the noise measure is lower for the region representing the object edge.

(a)   (b)
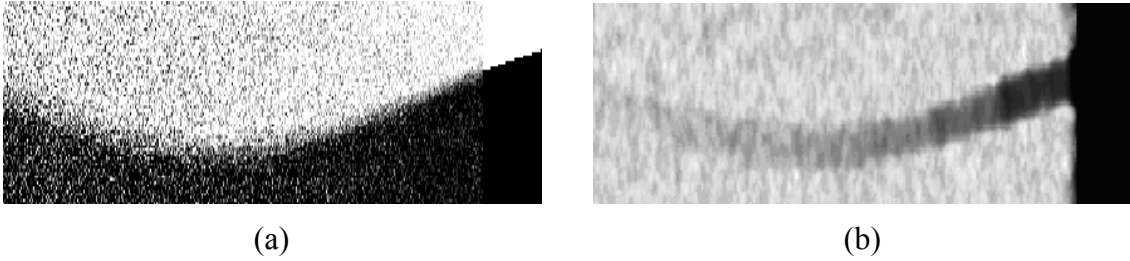
Figure 20: (a) Synthetic image with AWGN increasing in variance from left to right, and (b) resulting noise measure $N(x, y)$ where black intensities correspond to a measure of 0 and white intensities to a measure of 1. The noise measure is high for the high noise regions of the image without mistaking edges as noise.

One of the benefits of the SF measure is its consistency and robustness to changes in the image edge strength. To demonstrate this, we use the simple synthetic image of Figure 21(a) and determine the mean local spectral flatness over the image versus increasing variances of added noise (Figure 21(b)). We compare the profile of SF versus noise variance for the image of Figure 21(a) with different levels of Gaussian smoothing. As we decrease the edge strength in the image, the profile of the mean estimated noise level over the noise variance remains the same. This allows the estimation of the noise to remain decoupled from the estimation of edge strength.

(a)   (b)
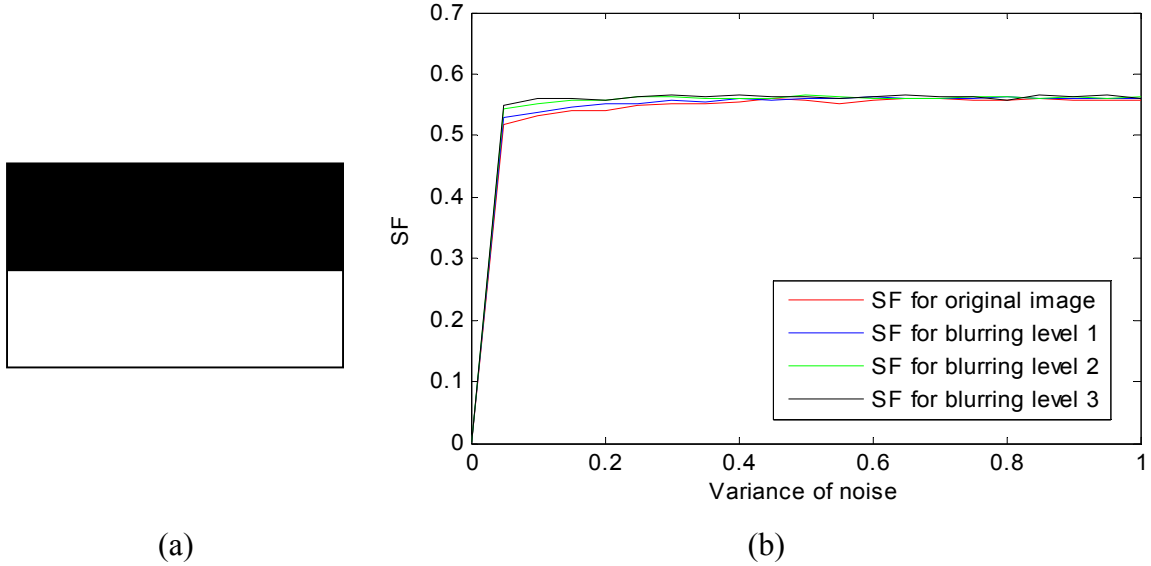
Figure 21: (a) Original synthetic image. (b) Average local spectral flatness versus variance of added AWGN for image in (a) with various levels of Gaussian blurring. As the edge strength weakens, the spectral flatness remains approximately constant, demonstrating how the noise detection is not affected by the edge strength.

The Fourier spectrum for each window region is determined through the fast Fourier transform (FFT). The size of the window for estimating the local spectral flatness plays a role in the quality of the resulting noise estimate. When using a smaller sized window, some correlation may exist between noise pixels, resulting in lowered flatness in the spectral domain and a resulting inaccurate lower noise estimate. Alternately, using a larger window size reduces the local nature of the measure. This dilemma is due to the dependence of noise on the scale of the image. As a tradeoff, we selected a window size of 15×15. In addition, we note that we have targeted AWGN in our measure as it represents one of the more difficult noise types to filter for in pre-processing steps and thus must be considered in images used for segmentation. Additional noise models will be considered in the future, as discussed in Section 4.2.

3.3 Edge Evidence Cue

We estimate the edge evidence in the image by using the measure:

$$G(x, y) = \max\left(|\nabla I_x(x, y)|, |\nabla I_y(x, y)|\right) \tag{44}$$

where $\nabla I_x(x, y)$ and $\nabla I_y(x, y)$ represent the x and y components of the image gradient. We chose this measure rather than the standard gradient magnitude for its rotational invariance in the discrete domain. With the typical gradient magnitude measure $|\nabla I| = \sqrt{\nabla I_x^2 + \nabla I_y^2}$, the result if $\nabla I_x = 1$ and $\nabla I_y = 1$ is $|\nabla I| = \sqrt{2}$ instead of $|\nabla I| = 1$.

3.4 Reliability Measure Formulation

The edge evidence cue of (44) can be highly unreliable if the image is deteriorated by noise. Therefore, we use the noise measure $N(x, y)$ of Section 3.2 to 'gate' the cue such that gradient information is suppressed in high noise regions. This concept of cue gating was first proposed in Malik *et al* [69] for the purposes of suppressing gradient information in high texture regions (described in further detail in Section 1.5.1). With this noise-gating, our edge evidence cue is as follows,

$$E_G(x, y) = \left(1 - N(x, y)\right)G(x, y) \tag{45}$$

where $E_G(x, y) \in [0,1]$. We term this noise-gated edge evidence cue as the *image reliability* as it represents how reliable an edge is.

3.5 Curvature Cue

We will next discuss how we estimate the level of curvature of the object boundary through a measure termed the curvature cue.

3.5.1 Curvature Formulation

We determine the curvature by a method similar to that of Section 1.5.2 [57, 36]. Let $I(x, y; \sigma)$ be a smoothened image such that

$$I(x, y; \sigma) = G_\sigma(x, y) * I_s(x, y) \tag{46}$$

where $I_s(x, y)$ is the original image and σ is the Gaussian scale parameter. The unit vector tangent to the isointensity contour, $C_I(x, y; \sigma)$ passing through a point (x, y) is given as:

$$\mathbf{t}(x, y; \sigma) = \frac{1}{\sqrt{I_{x,\sigma}^2(x, y) + I_{y,\sigma}^2(x, y)}} \begin{bmatrix} I_{y,\sigma}(x, y) \\ -I_{x,\sigma}(x, y) \end{bmatrix} \tag{47}$$

where $I_{x,\sigma}$ and $I_{y,\sigma}$ are the image derivatives along x and y, respectively, at scale σ.

Denoting the Hessian matrix of $I(x, y; \sigma)$ by $H_\sigma(x, y)$ as follows

$$H_\sigma(x, y) = \begin{bmatrix} I_{xx,\sigma}(x, y) & I_{xy,\sigma}(x, y) \\ I_{xy,\sigma}(x, y) & I_{yy,\sigma}(x, y) \end{bmatrix} \tag{48}$$

the local image curvature $K(x, y; \sigma)$ can be calculated as:

$$K(x,y;\sigma) \;=\; \frac{|\mathbf{t}^T H_\sigma \mathbf{t}|}{\sqrt{I_{x,\sigma}^2(x,y) + I_{y,\sigma}^2(x,y)}}. \tag{49}$$

Note that the absolute value of $K(x,y;\sigma)$ is used since we are not concerned with differentiating between convex and concave curvature. (49) expands to

$$K(x,y;\sigma) = \left| \frac{I_{y,\sigma}^2 I_{xx,\sigma} - 2I_{x,\sigma} I_{y,\sigma} I_{xy,\sigma} + I_{x,\sigma}^2 I_{yy,\sigma}}{\left(I_{x,\sigma}^2 + I_{y,\sigma}^2\right)^{3/2}} \right|. \tag{50}$$

We follow the method in [67] where $K(x,y;\sigma)$ is enhanced to have a stronger response near edges by multiplication with the gradient magnitude raised to some power, which we chose as 2. The edge enhanced curvature is then

$$\widetilde{K}(x,y;\sigma) = \left| \frac{I_{y,\sigma}^2 I_{xx,\sigma} - 2I_{x,\sigma} I_{y,\sigma} I_{xy,\sigma} + I_{x,\sigma}^2 I_{yy,\sigma}}{\sqrt{I_{x,\sigma}^2 + I_{y,\sigma}^2}} \right|. \tag{51}$$

3.5.2 Normalized Rescaled Curvature

The selection of the Gaussian kernel size σ, also referred to as the *scale* of the image, when smoothening the image in (46) plays a role in determining what sized structures in the image we will obtain meaningful curvature values for. Generally, larger structures will have meaningful curvature values when the original image is smoothened by a larger Gaussian kernel, and vice versa for smaller structures. However, curvature values cannot be compared across scales since the amplitude of the image spatial derivatives decreases with increasing scale. Thus, to compare curvature values across different scales, the curvature must be scale normalized, which is accomplished through the normalized scale coordinates of [67]. Once the normalized rescaled curvature, $\widetilde{K}_{norm}$,

is determined, Lindeberg approaches automated scale space selection by selecting scales at which the normalized rescaled curvature assumes maximum values.

$\widetilde{K}_{norm}$ is determined through scale-normalized coordinates, ξ, where

$$\xi = \frac{x}{\sigma} \tag{52}$$

and where the normalized derivative operators for these coordinates are

$$\partial_\xi = \sigma \partial_x, \tag{53}$$

$$\partial_\xi^2 = \sigma^2 \partial_x.$$

We substitute the scale normalized coordinates into (51) as follows:

$$\widetilde{K}_{norm}(x, y; \sigma) \tag{54}$$

$$= \left| \frac{\left(\sigma I_{y,\sigma}\right)^2 \left(\sigma^2 I_{xx,\sigma}\right) - 2\left(\sigma I_{x,\sigma}\right)\left(\sigma I_{y,\sigma}\right)\left(\sigma^2 I_{xy,\sigma}\right) + \left(\sigma I_{x,\sigma}\right)^2 \left(\sigma^2 I_{yy,\sigma}\right)}{\sqrt{\left(\sigma I_{x,\sigma}\right)^2 + \left(\sigma I_{y,\sigma}\right)^2}} \right|$$

$$= \left| \frac{\sigma^4 \left(I_{y,\sigma}^2 I_{xx,\sigma} - 2 I_{x,\sigma} I_{y,\sigma} I_{xy,\sigma} + I_{x,\sigma}^2 I_{yy,\sigma}\right)}{\sigma \sqrt{I_{x,\sigma}^2 + I_{y,\sigma}^2}} \right|$$

which can be simplified as:

$$\widetilde{K}_{norm}(x, y; \sigma) = \sigma^3 \, \widetilde{K}(x, y; \sigma). \tag{55}$$

After the curvature values at each scale have been normalized, the final curvature cue at every pixel is determined by selecting the scale at which $\widetilde{K}_{norm}$ assumes a maximum value as follows:

$$K_n(x, y) = \max_{\sigma} \widetilde{K}_{norm}(x, y; \sigma). \qquad (56)$$

3.5.3 Noise-Gated Curvature

Although robust to signal variations, the curvature measure $K_n(x, y)$ is sensitive to noise and might inaccurately give rise to a strong response at non-structure, high-noise regions of the image. Following the concept of cue gating, as we have previously used to gate the edge evidence cue in Section 3.4, we define a noise-gated curvature cue, $K_G(x, y)$, that suppresses our curvature cue in high noise regions as follows:

$$K_G(x, y) = \big(1 - N(x, y)\big) K_n(x, y). \qquad (57)$$

We demonstrate the curvature cue on a series of synthetic images shown in Figure 22. Figure 22(a) to (c) are synthetic images of objects with high curvature shapes, and Figure 22(d) to (f) are the resultant $K_G(x, y)$ measures. The spiral image of Figure 22(b) results in increased detected curvature towards the center region (Figure 22(e)), and the noisy regions of Figure 22(c) do not disrupt the curvature measure (Figure 22(f)).
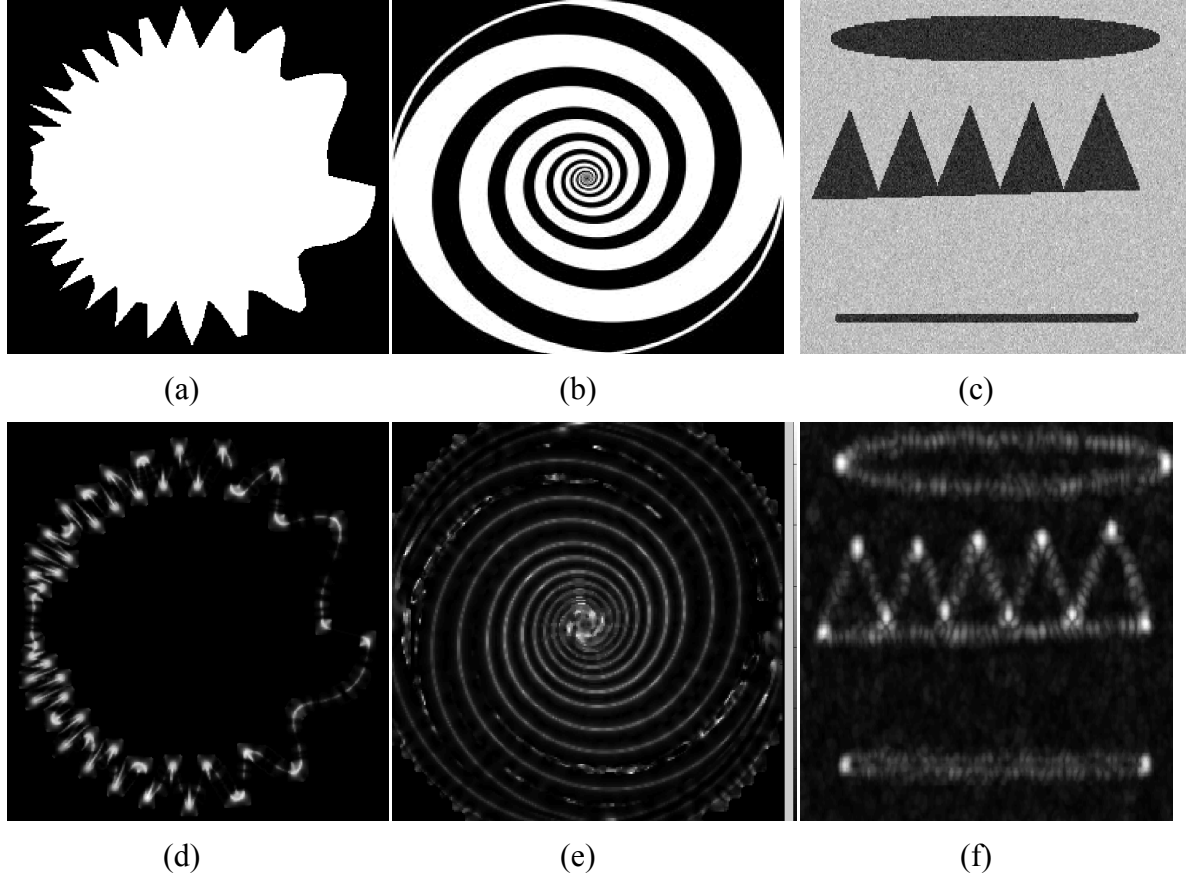
Figure 22: (a), (b), (c) Synthetic images and (d), (e), (f) resulting gated curvature measure $K_G(x, y)$ where black regions correspond to a measure output of 0 and white regions to 1. In the spiral of (b), the curvature measure increases towards the center. In the noisy image of (c), the resulting gated-curvature is high for the extremities of the shape only.

In addition, we demonstrate the need for noise gating on the synthetic image of Figure 23(a) which has added AWGN of variance 0.4. Figure 23(b) shows the non-gated curvature cue which poorly reflects the actual object curvature and is falsely high for noise regions. The noise-gated curvature in Figure 23(c) removes the false positives and enforces higher regularization for the non-edge regions.
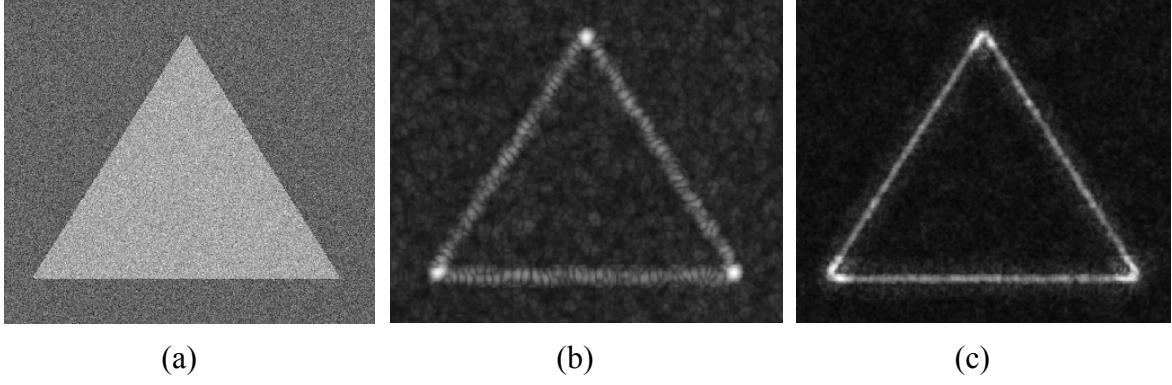
Figure 23: (a) Original synthetic image with AWGN of variance 0.4. (b) Curvature cue $K_n(x,y)$ (without noise gating) where intensities of 1 correspond to high curvature and intensities of 0 correspond to low curvature. High noise regions are incorrectly detected as high curvature levels. (c) Noise-gated curvature cue $K_G(x,y)$ where the sharp corners of the triangle are assigned the highest curvature and noisy regions are disregarded.

3.6 Data-Driven Regularization Weight Formulation

We next discuss how the data cues are combined and mapped to the regularization weight such that the weight behaves as discussed in Section 3.1.

3.6.1 Cue Combination and Mapping to Weight

To combine our noise-gated local image cues in a meaningful way, we define a mapping of those cues into our single adaptive weight, $w(x,y)$, that satisfies the following requirements: (1) In high trusted (noise-gated) edge evidence, little regularization is needed, regardless of the curvature strength. (2) In regions with low edge evidence, we set the regularization to be inversely proportional to the trusted (noise-gated) curvature such that high curvature regions are not overly regularized. Note that `high' curvature or edge evidence means a value close to 1 as all our cues are normalized. Thus we form the adaptive weight as follows:

71

$$w(x,y) = 1 - E_G(x,y)^{\gamma_e\left(1-K_G^{\gamma_c}(x,y)\right)}. \tag{58}$$

If $E_G(x,y)$ is large (approaching maximum value of 1), the exponent has little effect on the resulting weight, and requirements (1) is satisfied. If $E_G(x,y)$ is low and $K_G(x,y)$ is non-zero, the noise-gated edge evidence will be raised to a power $\left(1 - K_G(x,y)\right) \approx 0$, resulting in a lower $w(x,y)$, satisfying requirement (2). Note that the detrimental effects from noise are handled by this model through the noise-gating of the cues. We refer to $E_G(x,y)^{\gamma_e\left(1-K_G^{\gamma_c}(x,y)\right)}$ as the curvature-modulated image reliability measure. We include the parameters γ_e and γ_c to allow minor adjustments of how strongly the edge evidence and curvature term should affect the regularization weight. We stress that these parameters are to allow for user tweaking and are set to a constant value over the image (i.e., the weights still vary spatially in an automated manner).

Figure 24(a) shows a surface plot of the regularization weight $w(E_G, K_G)$ as $E_G(x,y)$ and $K_G(x,y)$ vary from 0 to 1, and where the parameters $\gamma_e = 1$ and $\gamma_c = 1$. As both of the cues increase, the regularization weight decreases since the local region of the image is considered more 'reliable' and regularization is not needed. However, as the cues decrease, the regularization weight increases since the edge evidence is no longer present or is not trusted (i.e. high noise is present). In addition, Figure 24(b) to (d) shows $w(E_G, K_G)$ for parameters settings of $(\gamma_e, \gamma_c) = (0.8, 0.4)$, $(\gamma_e, \gamma_c) = (1, 0.2)$, and $(\gamma_e, \gamma_c) = (0.2, 1)$, respectively. The parameters change the level of concavity of the function, and the maximum weight produced.
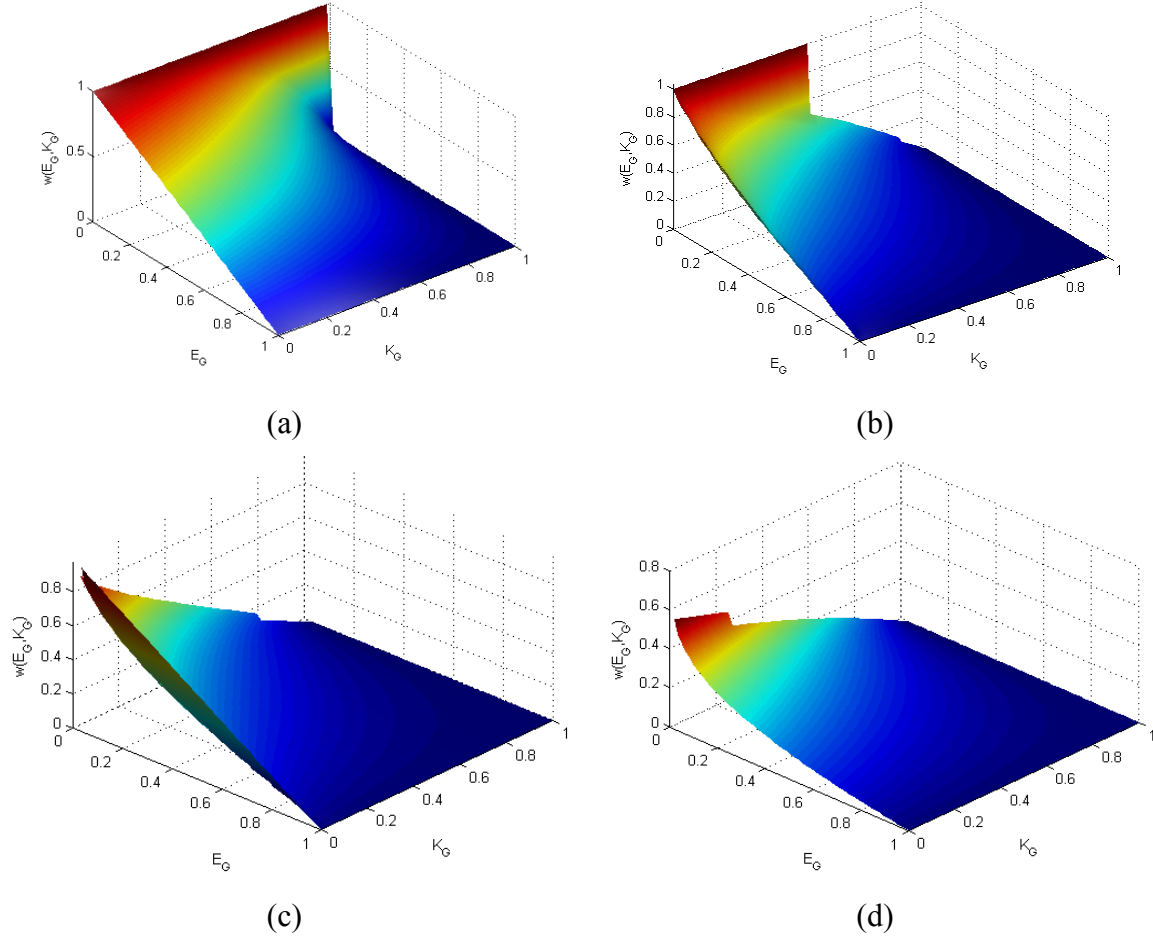
Figure 24: Surface plot of $w(E_G, K_G)$ based on the gated edge evidence (E_G) and gated curvature cue (K_G), and on different selections of the parameters γ_e and γ_c. (a) $\gamma_e = 1, \gamma_c = 1$. (b) Default setting $\gamma_e = 0.8, \gamma_c = 0.4$. (c) $\gamma_e = 1, \gamma_c = 0.2$, (d) $\gamma_e = 0.2, \gamma_c = 1$. The parameters change the level of concavity of the function and the maximum weight produced

We demonstrate the cues and resulting weight measure for two synthetic images with a sinusoidal varying boundary (produced by (38) in Section 2.3.2), spatially varying AWGN and Gaussian smoothening with spatially varying kernel sizes, as shown in Figure 25(a) and (b). The resulting measures (Figure 25(c) to (j)) demonstrate robustness to noise and a final weighting that regularizes high noise regions but lower regularization in high curvature and strong trusted edge regions.
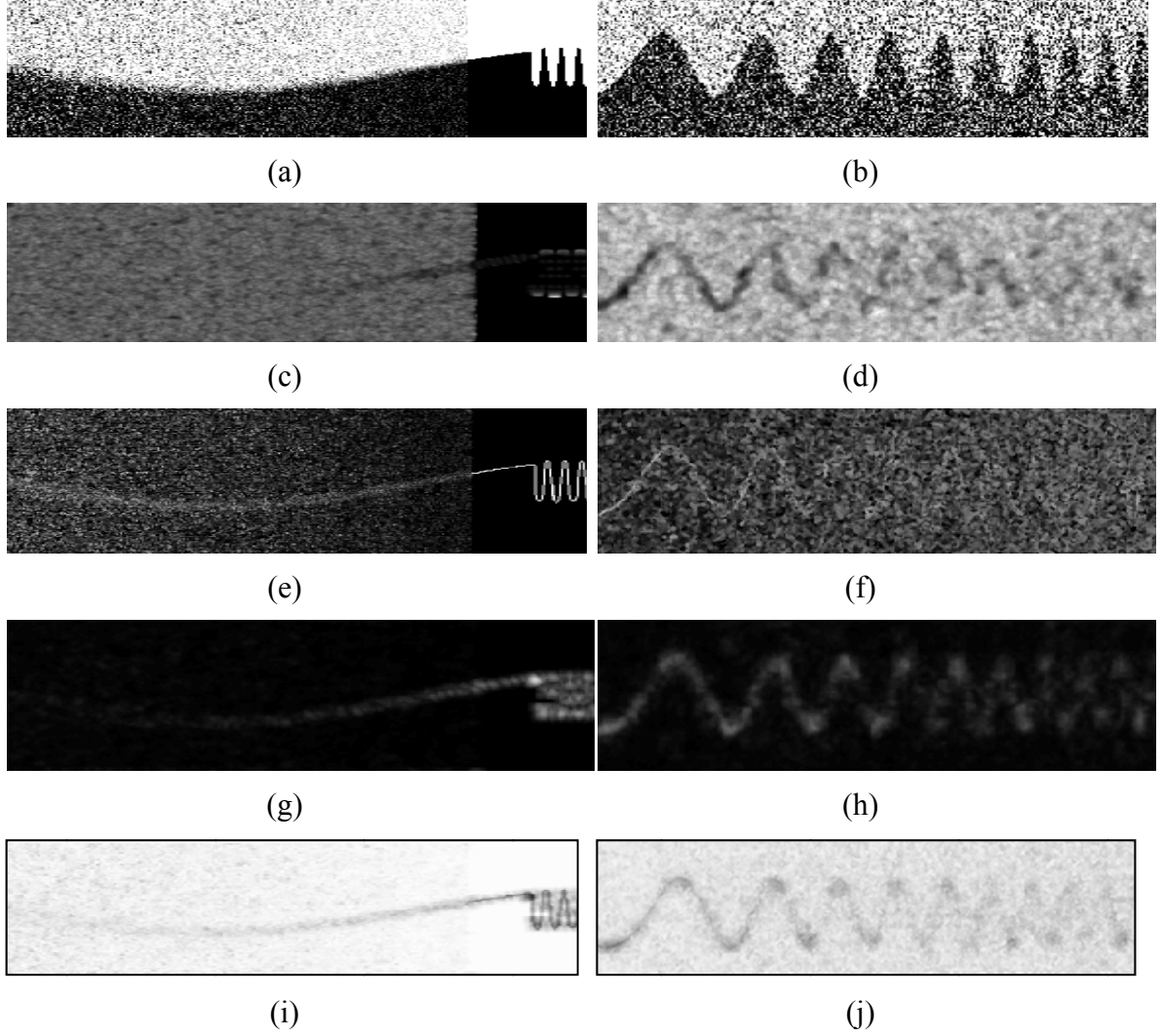
Figure 25: (a), (b) Synthetic images with varying characteristics. (c), (d) Noise measure $N(x, y)$. (e), (f) Noise-gated edge cue $E_G(x, y)$. (g), (f) Noise-gated curvature cue $K_G(x, y)$. (i), (f) Reliability weight $w(x, y)$. Black intensities correspond to measures of 0 and white intensities to 1. The left column image has spatially varying noise and blurring (increasing from right to left) and with changing boundary smoothness (smooth on the left and jagged on the right). The right column image has higher curvature and noise levels. The results confirms the desired behavior of the reliability measure.

3.6.2 Incorporation of Texture and Additional Cues

In many natural images, large gradients and large curvature values can arise from texture edges rather than from edges representing object boundaries. To prevent texture edges from being included in the final edge set of the image, we must ensure greater

regularization occurs in textured regions. We employ a texture measure from Erdem and Tari [38] that estimates the probability of a pixel being near a texture edge, discussed in Section 1.5.1 and shown again here for clarity:

$$T(x,y) = 1 - \exp\left(-\gamma^t \left(\min(\rho_1(x,y), \rho_2(x,y)) \right) \right) \tag{59}$$

We incorporate the texture cue into our framework by modifying E_G in (45) to form the texture-gated and noise-gated edge evidence term as follows:

$$E_{G,T}(x,y) = T(x,y)\left(1 - N(x,y) \right) |\nabla I(x,y)|. \tag{60}$$

Incorporating $E_{G,T}$ with our spatially adaptive weight produces the texture-dependent regularization weight, $w_t(x,y)$, as follows:

$$w_t(x,y) = 1 - E_{G,T}^{\gamma_e}(x,y)^{\left(1 - K_G^{\gamma_c}(x,y)\right)}. \tag{61}$$

In addition to texture, any other data cue can be combined into our regularization framework through multiplication into the gated edge evidence term of (60).

3.7 Incorporation of Regularization Weights into Segmentation Frameworks

The regularization weight mapping of (58) produces a 2D map of weights over the image. This weight map can be input into any energy minimization segmentation framework with only minor modifications to allow for convex adaptive weighting. To demonstrate this, we will now incorporate the data-driven regularization weight of (58) into four existing segmentation frameworks with minor changes to the energy functional.

3.7.1 Minimum-path Frameworks

We incorporate the data-driven regularization weight into a basic minimum-path framework that was described in Section 2.1 and repeated here as:

$$E\big(C(q), w(q)\big) = \int_0^1 w(q)E_{int}C(q) + \big(1 - w(q)\big)E_{ext}\big(C(q)\big)dq \qquad (62)$$

The external energy is the same as in Section 2.1, however the internal energy now reflects the length of the contour in the 2D space (x, y). The optimization process is through a 2D graph search using Djikstra's method as described in Section 1.2.2.1. By using this framework, we are able to first test the data-driven weight against the globally optimal weight for validation.

3.7.2 Graph Cuts

We incorporated our adaptive weights, $w(p)$, into a graph cuts (GC) based segmentation process [7, 17]. The segmentation energy in this case becomes:

$$E(f) = \sum_{p,q \in N} w(p)E_{int}\big(f_p, f_q\big) + \sum_{p \in P} \big(1 - w(p)\big)E_{ext}\big(f_p\big) \qquad (63)$$

where $f \in$ L is the labelling for all pixels $p \in P$, where L is the space of all possible labellings, and P is the set of pixels in image I. E_{int} is the interaction penalty between pixel pairs (i.e. the penalty of assigning labels f_p and f_q to pixels p and q), E_{ext} measures how well label f_p fits pixel p given the observed data, and N is the set of interacting pairs of pixels. Refer to Section 1.3 for more details about E_{int} and E_{ext}.

3.7.3 Active Contours without Edges

We next present the ACWE segmentation framework with spatially adaptive regularization. Although we have used a convex weighting scheme in the discrete segmentation frameworks, we will only weight the regularization term itself for the continuous frameworks. This follows the formulation used by the majority of existing spatially adaptive regularization methods for continuous frameworks [38, 80, 35, 69]. We found through testing that a convex weighting results in impedance of curve evolution when the external terms are weighted to zero. However we will leave the analysis of curve evolution and curve stalling in level sets to future work (Section 4.2). We modified (14) in Section 1.2.3.3 to incorporate spatially adaptive regularization by replacing μ with an adaptive weight, as follows:

$$
E\big(\phi(x,y)\big) = \int_\Omega w(x,y)\delta\big(\phi(x,y)\big)\,|\nabla\phi(x,y)|dx\,dy \tag{64}
$$

$$
+ \int_\Omega |I(x,y) - c_1|^2\, H\big(\phi(x,y)\big)
$$

$$
+ |I(x,y) - c_2|^2\Big(1 - H\big(\phi(x,y)\big)\Big)\,dxdy
$$

where the segmentation is represented here via a Lipschitz function, $\phi(x,y)\colon \Omega \to \Re$, where pixels p interior to the zero-level set of $\phi(x,y)$ are labelled as objects and exterior pixels as background. Also note that we use the notation

$$
|\nabla\phi(x,y)| = \sqrt{\left(\frac{\partial\phi}{\partial x}\right)^2 + \left(\frac{\partial\phi}{\partial y}\right)^2}. \tag{65}
$$

77

We determine $\phi(x, y)$ that minimizes (64) by using the Euler-Lagrange equation to solve

the gradient descent PDE:

$$\frac{\partial \phi}{\partial t} = -\frac{\partial E}{\partial \phi} = -\left[\frac{\partial L}{\partial \phi} - \frac{d}{dx}\frac{\partial L}{\partial \phi_x} - \frac{d}{dy}\frac{\partial L}{\partial \phi_y} \right] = 0 \qquad (66)$$

where

$$L = w(x, y)\delta\big(\phi(x, y)\big)|\nabla\phi(x, y)| \qquad (67)$$

$$+ |I(x, y) - c_1|^2 H\big(\phi(x, y)\big) + |I(x, y) - c_2|^2 \big(1 - H\big(\phi(x, y)\big)\big)$$

and where we use the notation $\phi_x = \frac{\partial \phi}{\partial x}$. We first determine the partial derivative $\frac{\partial L}{\partial \phi}$ as

follows:

$$\frac{\partial L}{\partial \phi} = \frac{\partial}{\partial \phi} w(x, y)\delta\big(\phi(x, y)\big)|\nabla\phi(x, y)| \qquad (68)$$

$$+ \frac{\partial}{\partial \phi}\left[|I(x, y) - c_1|^2 H\big(\phi(x, y)\big) + |I(x, y) - c_2|^2 \big(1 - H\big(\phi(x, y)\big)\big) \right]$$

which in expanded form is:

$$\frac{\partial L}{\partial \phi} = \frac{\partial}{\partial \phi} w(x,y)\delta\big(\phi(x,y)\big)|\nabla\phi(x,y)| \tag{69}$$

$$+ \frac{\partial}{\partial \phi}|I(x,y) - c_1|^2 H\big(\phi(x,y)\big) + \frac{\partial}{\partial \phi}|I(x,y) - c_2|^2$$

$$- \frac{\partial}{\partial \phi}|I(x,y) - c_2|^2 H\big(\phi(x,y)\big).$$

We use the property

$$\frac{\partial}{\partial \phi}H\big(\phi(x,y)\big) = \delta\big(\phi(x,y)\big) \tag{70}$$

and the fact that

$$\frac{\partial}{\partial \phi}|I(x,y) - c_2|^2 = 0 \tag{71}$$

to simplify (69) as follows:

$$\frac{\partial L}{\partial \phi} = w(x,y)\delta_\phi\big(\phi(x,y)\big)|\nabla\phi(x,y)| \tag{72}$$

$$+ \delta\big(\phi(x,y)\big)[|I(x,y) - c_1|^2 - |I(x,y) - c_2|^2]$$

where we use the notation $\frac{\partial}{\partial \phi}\delta\big(\phi(x,y)\big) = \delta_\phi\big(\phi(x,y)\big)$.

We then determine $\frac{\partial L}{\partial \phi_x}$ from (67) as follows:

$$\frac{\partial L}{\partial \phi_x} = \frac{\partial}{\partial \phi_x} w(x, y) \delta(\phi(x, y)) |\nabla \phi(x, y)| \tag{73}$$

$$+ \frac{\partial}{\partial \phi_x} \Big[|I(x, y) - c_1|^2 H(\phi(x, y))$$

$$+ |I(x, y) - c_2|^2 \left(1 - H(\phi(x, y))\right) \Big].$$

We note that

$$\frac{\partial}{\partial \phi_x} \Big[|I(x, y) - c_1|^2 H(\phi(x, y)) + |I(x, y) - c_2|^2 \left(1 - H(\phi(x, y))\right) \Big] = 0. \tag{74}$$

and that

$$\frac{\partial}{\partial \phi_x} w(x, y) \delta(\phi(x, y)) |\nabla \phi(x, y)| = w(x, y) \delta(\phi(x, y)) \frac{\partial}{\partial \phi_x} |\nabla \phi(x, y)| \tag{75}$$

$$= w(x, y) \delta(\phi(x, y)) \frac{2\phi_x}{2 \sqrt{\left(\frac{\partial \phi}{\partial x}\right)^2 + \left(\frac{\partial \phi}{\partial y}\right)^2}}$$

$$= w(x, y) \delta(\phi(x, y)) \frac{\phi_x}{|\nabla \phi(x, y)|}$$

Thus we can simplify (73) as follows:

$$\frac{\partial L}{\partial \phi_x} = w(x, y) \delta(\phi(x, y)) \frac{\phi_x}{|\nabla \phi(x, y)|}. \tag{76}$$

Similarly we obtain the following for $\frac{\partial L}{\partial \phi_y}$:

$$\frac{\partial L}{\partial \phi_y} = w(x,y)\delta(\phi(x,y))\frac{\phi_y}{|\nabla\phi(x,y)|}. \tag{77}$$

We then take the derivative of (76) with respect to x as follows:

$$\frac{d}{dx}\left(\frac{\partial L}{\partial \phi_x}\right) = \frac{d}{dx}\left(w(x,y)\delta(\phi(x,y))\frac{\phi_x}{|\nabla\phi(x,y)|}\right). \tag{78}$$

Using the product rule

$$\frac{d(AB)}{dx} = A\frac{dB}{dx} + B\frac{dA}{dx}, \tag{79}$$

and noting that due to the chain rule, we can make the expansion

$$\frac{d}{dx}\delta(\phi(x,y)) = \delta_\phi(\phi(x,y))\phi_x, \tag{80}$$

we expand (78) as follows:

$$\begin{aligned}
\frac{d}{dx}\left(\frac{\partial L}{\partial \phi_x}\right) &= \frac{d}{dx}[w(x,y)\delta(\phi(x,y))]\frac{\phi_x}{|\nabla\phi(x,y)|} \\[2mm]
&\quad + w(x,y)\frac{d}{dx}[\delta(\phi(x,y))]\frac{\phi_x}{|\nabla\phi(x,y)|} \\[2mm]
&\quad + w(x,y)\delta(\phi(x,y))\frac{d}{dx}\left[\frac{\phi_x}{|\nabla\phi(x,y)|}\right] \\[3mm]
&= w_x(x,y)\delta(\phi(x,y))\frac{\phi_x}{|\nabla\phi(x,y)|} + w(x,y)\delta_\phi(\phi(x,y))\phi_x\frac{\phi_x}{|\nabla\phi(x,y)|} \\[2mm]
&\quad + w(x,y)\delta(\phi(x,y))\frac{d}{dx}\left[\frac{\phi_x}{|\nabla\phi(x,y)|}\right]
\end{aligned} \tag{81}$$

$$= w_x(x,y)\delta(\phi(x,y))\frac{\phi_x}{|\nabla\phi(x,y)|} + w(x,y)\delta_\phi(\phi(x,y))\frac{\phi_x^2}{|\nabla\phi(x,y)|}$$

$$+ w(x,y)\delta(\phi(x,y))\frac{d}{dx}\left[\frac{\phi_x}{|\nabla\phi(x,y)|}\right].$$

Similarly, we take the derivative of (77) with respect to y as follows:

$$\frac{d}{dy}\left(\frac{\partial L}{\partial \phi_y}\right) = w_y(x,y)\delta(\phi(x,y))\frac{\phi_y}{|\nabla\phi(x,y)|} \tag{82}$$

$$+ w(x,y)\delta_\phi(\phi(x,y))\frac{\phi_y^2}{|\nabla\phi(x,y)|}$$

$$+ w(x,y)\delta(\phi(x,y))\frac{d}{dy}\left[\frac{\phi_y}{|\nabla\phi(x,y)|}\right].$$

We then combine (81) and (82). We first note that

$$\frac{\phi_x^2 + \phi_y^2}{|\nabla\phi(x,y)|} = |\nabla\phi(x,y)| \tag{83}$$

and thus that

$$w(x,y)\delta_\phi(\phi(x,y))\left(\frac{\phi_x^2}{|\nabla\phi(x,y)|} + \frac{\phi_y^2}{|\nabla\phi(x,y)|}\right) \tag{84}$$

$$= w(x,y)\delta_\phi(\phi(x,y))|\nabla\phi(x,y)|.$$

We also note that

$$w_x(x,y)\frac{\phi_x}{|\nabla\phi(x,y)|} + w_y(x,y)\frac{\phi_y}{|\nabla\phi(x,y)|} = \nabla w(x,y)\cdot\frac{\nabla\phi(x,y)}{|\nabla\phi(x,y)|}. \tag{85}$$

Additionally, using the product rule (79), we note that:

$$\frac{d}{dx}\left[\frac{\phi_x}{|\nabla\phi(x,y)|}\right] + \frac{d}{dy}\left[\frac{\phi_y}{|\nabla\phi(x,y)|}\right] \tag{86}$$

$$= \phi_x\frac{d}{dx}\left(\frac{1}{|\nabla\phi(x,y)|}\right) + \frac{1}{|\nabla\phi(x,y)|}\frac{d}{dx}\phi_x$$

$$+ \phi_y\frac{d}{dy}\left(\frac{1}{|\nabla\phi(x,y)|}\right) + \frac{1}{|\nabla\phi(x,y)|}\frac{d}{dy}\phi_y$$

$$= \left[\phi_x\frac{d}{dx}\left(\frac{1}{|\nabla\phi(x,y)|}\right) + \phi_y\frac{d}{dy}\left(\frac{1}{|\nabla\phi(x,y)|}\right)\right]$$

$$+ \frac{1}{|\nabla\phi(x,y)|}\left[\frac{\partial^2\phi}{\partial x^2} + \frac{\partial^2\phi}{\partial y^2}\right]$$

$$= \nabla\phi \cdot \nabla\frac{1}{|\nabla\phi(x,y)|} + \frac{1}{|\nabla\phi(x,y)|}\,\text{div}\,(\nabla\phi).$$

Since the divergence of a scalar function ς and a vector $\mathbf{F}$ is as follows:

$$\text{div}(\varsigma\mathbf{F}) = (\nabla\varsigma)\cdot\mathbf{F} + \varsigma\,\text{div}\,(\mathbf{F}), \tag{87}$$

we can simplify (86) by using the substitution $\varsigma = \frac{1}{|\nabla\phi(x,y)|}$ and $\mathbf{F} = \nabla\phi$ as follows:

$$\left[\phi_x\frac{d}{dx}\left(\frac{1}{|\nabla\phi(x,y)|}\right) + \phi_y\frac{d}{dy}\left(\frac{1}{|\nabla\phi(x,y)|}\right)\right] \tag{88}$$

$$= \left[\frac{d}{dx}\left(\frac{\phi_x}{|\nabla\phi(x,y)|}\right) + \frac{d}{dy}\left(\frac{\phi_y}{|\nabla\phi(x,y)|}\right)\right]$$

$$= \text{div}\left(\frac{\nabla\phi(x,y)}{|\nabla\phi(x,y)|}\right).$$

We use (88) to derive:

$$w(x,y)\delta(\phi(x,y))\left(\frac{d}{dx}\left[\frac{\phi_x}{|\nabla\phi(x,y)|}\right] + \frac{d}{dy}\left[\frac{\phi_y}{|\nabla\phi(x,y)|}\right]\right) \tag{89}$$

$$= w(x,y)\delta(\phi(x,y))\text{div}\left(\frac{\nabla\phi(x,y)}{|\nabla\phi(x,y)|}\right)$$

From (84), (85), and (89), we combine (81) and (82) as follows:

$$\frac{d}{dx}\left(\frac{\partial L}{\partial\phi_x}\right) + \frac{d}{dy}\left(\frac{\partial L}{\partial\phi_y}\right) = w(x,y)\delta_\phi(\phi(x,y))|\nabla\phi(x,y)| \tag{90}$$

$$+\delta(\phi(x,y))\nabla w(x,y)\cdot\frac{\nabla\phi(x,y)}{|\nabla\phi(x,y)|} + w(x,y)\delta(\phi(x,y))\text{div}\left(\frac{\nabla\phi(x,y)}{|\nabla\phi(x,y)|}\right)$$

We substitute (72) and (90) into the Euler-Lagrange (66) to obtain the final evolution equation that corresponds to the spatially adaptive ACWE function (64) as follows:

$$\frac{\partial\phi}{\partial t} = \delta(\phi(x,y))\nabla w(x,y)\cdot\frac{\nabla\phi(x,y)}{|\nabla\phi(x,y)|} \tag{91}$$

$$+ w(x,y)\delta(\phi(x,y))\text{div}\left(\frac{\nabla\phi(x,y)}{|\nabla\phi(x,y)|}\right)$$

$$- \delta(\phi(x,y))[|I(x,y) - c_1|^2 + |I(x,y) - c_2|^2]$$

When we compare this new form of the evolution equation to the original form in (15) from Section 1.2.3.3, we see that we obtain the new term $\delta(\phi(x,y))\nabla w(x,y)\cdot\frac{\nabla\phi(x,y)}{|\nabla\phi(x,y)|}$.

We optimize the spatially adaptive ACWE energy functional by first using finite differences to discretize the system, then initializing $\phi(x,y)$ to a user-entered contour in the vicinity of the object to segment, and then iteratively adding the image evolution term of (91) to $\phi(x,y)$ until $\phi(x,y)$ is stationary.

3.7.4 Contextual Mumford-Shah Framework

In order to compare our method against the closest existing adaptive regularization technique, we incorporated our weights into the Mumford-Shah based segmentation framework of Erdem and Tari [38] as described in Section 1.5.1 where contextual regularization is used. We note that the regularization weights of the Erdem and Tari (ET) method is not general and can only be used and compared against within this framework. We incorporated our data cues into the ET method by using the negative feedback method (described in Section 1.5.1 and reprinted here for clarity)

$$cv = \kappa v + (1 - \kappa) \tag{92}$$

where κ represents the combined cues used in the ET method. We replace κ with our curvature-modulated reliability term (see Section 3.6.1) as follows:

$$cv = E_G^{\gamma_e}(x,y)^{\left(1 - K_G^{\gamma_c}(x,y)\right)} v + \left(1 - E_G^{\gamma_e}(x,y)^{\left(1 - K_G^{\gamma_c}(x,y)\right)}\right) \tag{93}$$

which simplifies to

$$cv = E_G^{\gamma_e}(x,y)^{\left(1 - K_G^{\gamma_c}(x,y)\right)} (v - 1). \tag{94}$$

In addition, we also tested incorporating the texture edges term of [38] such that

$$cv = E_{G,T}^{\gamma_e}(x,y)^{\left(1 - K_G^{\gamma_c}(x,y)\right)} (v - 1). \tag{95}$$

where $E_{G,T}(x,y)$ is described in Section 3.6.2. The cv term modifies the evolution equations of the functional (as described in Section 1.5.1) which are then used to

iteratively update the image process and edge process to form the segmentation (as described in Section 1.2.3.2).

3.8 Validation and Results

We demonstrate the performance of the data-driven regularization weights on the synthetic, medical, and natural datasets that we first introduced in Section 2.3 with the globally optimal weight. In addition to the least-error spatially fixed weight, which we will compare our results against for all methods, we will compare our method to the globally optimal weight for the minimum-path segmentation framework, and to the ET method for the contextual Mumford-Shah segmentation framework.

3.8.1 Performance Criteria

We tested each segmentation method of Section 3.7 with databases suitable for the method. For example, natural images with object comprising of a wide intensity range cannot be used for segmentation methods with regions-based external terms since these methods assume that the object of interest is approximately piecewise constant. However, these non-piecewise constant objects can be segmented by methods with boundary external terms, such as the minimum-path framework of Section 3.7.1.

In our tests using the minimum-path segmentation framework, we first analyze results on the sinusoidal synthetic dataset described in Section 2.3.2 that was used for the globally optimal weight validation. We measure our error for this dataset using the same H_E error as we did for the globally-optimal weight validation (see (35) in Section 2.3.1). We also tested the minimum-path segmentation framework on medical and natural

datasets where we used the DSC error metric from (36) in Section 2.3.1. For the graph cuts framework, we first analyze results on a synthetic dataset testing detection of objects with decreasing signal-to-noise ratio (SNR). Since the process involves matching k labels between the ground truth (original noise-free shape) and the segmentation, we determine the error by using the multi-labelling Hungarian method [61] to solve the label assignment problem, and then using the DSC measure of (36) in Section 2.3.1. The remainder of the testing on the medical and natural datasets using all four of the aforementioned method used the DSC measure.

We compared our method against existing techniques suitable for each segmentation framework. For all frameworks, we compared our data-driven weight against the least-error spatially fixed weight w_{fixed} as described in Section 2.3.1. For the minimum-path framework, we additionally compared our method against the globally optimal weight as discussed in Chapter 2. For the contextual Mumford-Shah framework, we additionally compared our method against the ET regularization cues. Within our method, we also compared the weights produced by only the noise-gated edge cue to the weights produced by the curvature-and-noise modulated edge cue, and by addition of the texture cue into our framework.

For each set of results, we determined 25 segmentations for each image and averaged to produce the quantitative results. In addition, we performed ANOVA to determine if the data-driven weights produced significantly more accurate results than w_{fixed}, w_{GO}, and the ET method.

87

3.8.2 Parameter Selection and Implementation Details

For the segmentation approaches, we used a graph cuts wrapper from [7], an implementation of ACWE from [102], and an implementation of the contextual MS method from [38], all of which were modified as proposed in Section 3.7.

The minimum-path segmentation framework determines the optimal contour between seedpoints, which we entered as equidistant points determined automatically around the object boundary as determined by the ground truth segmentation, and where each seedpoint consists of the pixel coordinates (x, y). For the graph cuts framework, we selected a low number of random seeds (0.3% of image pixels for each label) automatically by using the ground truth. For the ACWE and contextual MS frameworks, we used an initial contour of a $M/4 \times N/4$ square placed in the center of the object of interest, where $M \times N$ is the image size. For all frameworks, we used the same initializations/seedpoints for the comparison methods. For our data-driven weights, we set the parameters $\gamma_e = 0.8$ and $\gamma_c = 0.4$.

3.8.3 Computational Performance

In our tests, we used un-optimized MATLAB code on a PC with 3.6 GHz Intel Core Duo processor and 2GB of RAM. Computationally, the proposed method required less than 3 minutes to calculate the regularization weights for a 768×576 image. The most computationally intensive aspect of our work in the noise measure as determined through spectral flatness (see Section 3.2) because the fast Fourier transform (FFT) must be determined for each local window region in the image.

3.8.4 Synthetic Data Validation

3.8.4.1 Noise limitation tests

We first analyzed the robustness of our data-driven weights against increasing levels of noise (AWGN) to determine the limit at which the weights are no longer meaningful. Figure 26(a), (b), and (c) show synthetic images corrupted by increasing levels of AWGN. The graph cuts adaptive weight segmentation for the images corrupted by noise levels of 0.05 and 1.05 std. dev. (Figure 26(d) and (e), respectively) adheres to the corners of the object and does not leak outside of the object, unlike the fixed weight segmentation in red. At an extremely high noise level of 1.90 std. dev. shown in Figure 26(c), the resulting adaptive weight segmentation Figure 26(f) begins to show holes and degradation. Analysis of the DSC of adaptive weight graph cuts segmentations for the synthetic image of Figure 27 over various noise levels (25 noise realizations for each noise level) found that segmentation accuracy begins to drop at noise levels greater than 1.75 std. dev.
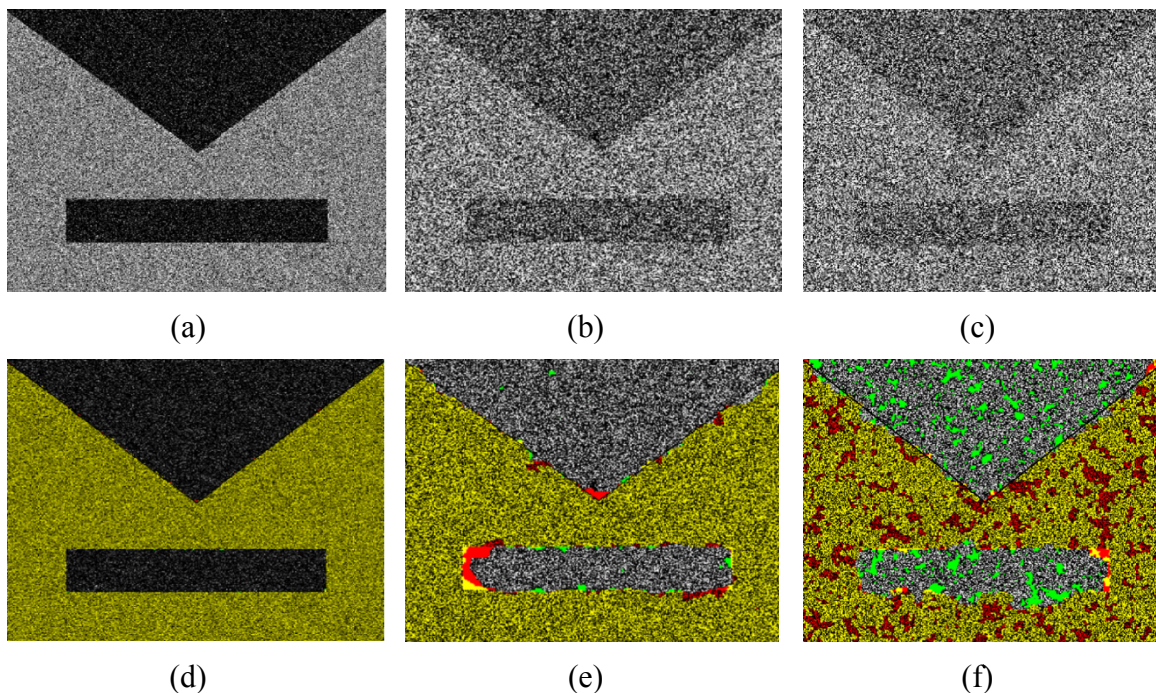
Figure 26: Segmentation of grey object in synthetic image corrupted by AWGN with increasing standard deviation. (a), (b), (c) Original images with std. dev. of 0.05, 1.05, and 1.90, respectively. (d), (e), (f) Corresponding segmentations from the proposed adaptive weight *green* and the least-error fixed weight *red* where *yellow* regions are where the segmentations overlap. At the high noise level of (c), the segmentation (f) begins to form holes and inaccuracies.
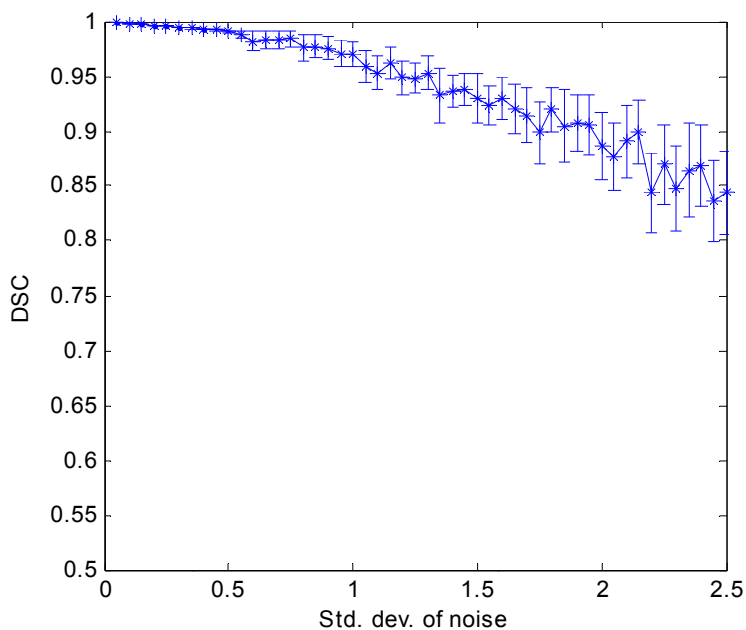
Figure 27: Segmentation accuracy (DSC) as increasing levels of noise are added to the synthetic image of Figure 26(a). Accuracy begins to drop at noise levels greater than 1.75 std. deviation.

90

We tested the performance of the minimum-path segmentation framework for the same sinusoidal synthetic dataset of Section 2.3.2 that is designed to cover extreme shape and appearance variations with AWGN of spatially varying variance and Gaussian blurring with spatially varying kernel sizes. We compared our results against the globally-optimal weight segmentation and the least-error fixed weight segmentation, as shown in Figure 28(a) and (b) for the two synthetic images of Figure 25(a) and (b) (see Section 3.6.1):

(a)

(b)
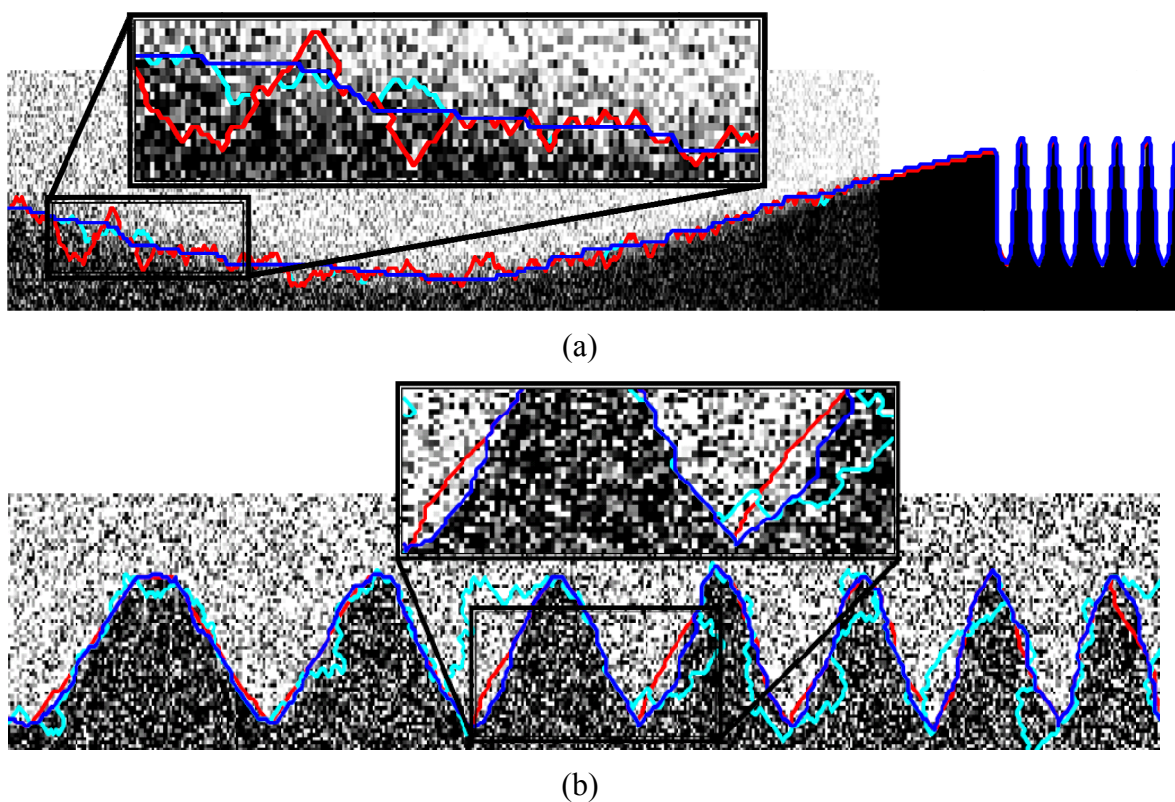
Figure 28: Synthetic dataset testing with minimum-path approach. (a) Image with spatially varying noise and blurring (increasing from right to left) and with changing boundary smoothness (smooth on the left and jagged on the right). (b) Image with higher curvature and noise levels. Contours obtained from: (*blue*) data-driven weights, (*red*) best fixed weight, and (*cyan*) globally optimum weight. The data-driven weights produce the only segmentation that accurately regularization in high noise regions and accurately captures high curvature regions of the boundary.

91

In Figure 28(a), the least-error fixed weight is low (0.2) in order to segment the high curvature region of the image, and thus is not able to accurately segment the high noise region of the object. The data-driven weight produces an accurate segmentation by lowering regularization in the high curvature and reliable edge region of the image, and increasing regularization is the unreliable high noise regions. The curvature-modulated reliability measure (see Figure 25(g) and (h) in Section 3.6.1) confirms this behaviour.

We quantitatively examined our method's performance using ANOVA testing on 25 noise realizations of each image in the dataset, where the error was determined by H_E. Our method resulted in a mean error (in pixels) of 6.33 ± 1.36, whereas the best fixed-weight method had a mean error of 12.05 ± 1.61, and the globally-optimum weight method had a mean error of 33.06 ± 3.66. Furthermore, for each image, we found our method to be significantly more accurate with all p-values $\ll 0.05$. Our error results are presented in Table 2.

Table 2: Average error over 25 noise realizations per image produced by data-driven weights, least-error fixed weights, and globally optimal weights for synthetic set of data. Segmentations from the data-driven weights have errors less than alternate methods with *p*-values from ANOVA testing $\ll 0.05$.

Image #	Mean H_E error over 25 segmentations		
	Fixed weights	Data-driven weights	G.O. weights
1	13.02	7.2	17.68
2	15.04	7.08	45.24
3	8.92	6.04	17.4
4	13.8	8.64	38.6
5	6.8	3.64	7.16
6	12.96	5.92	42.2
7	14.4	7.4	42.4
8	13.28	7.56	40.92
9	11.8	3.12	35.72
10	13.04	4.76	47.68
11	14.84	8.16	40.04
12	6.64	3.32	7.92
13	10.62	5.4	27.24
14	12.74	7.48	41
15	11.92	6.32	44.2
16	13.02	9.28	33.64

3.8.4.3 Graph Cuts Synthetic Tests

We validated graph cuts with our proposed method on simulated noisy images of variably-sized ellipses with complicated background patterns, e.g. with image contrast decreasing from right to left, as in Figure 29(a). The leftmost ellipses with lower contrast have a lower SNR than the rightmost ellipses and thus require greater regularization. Note how our resulting reliability measure Figure 29(b) indicates lower image reliability for low contrast ellipses. When comparing our segmentation results to those from the spatially-fixed weight, as shown in Figure 29(c), 6 ellipses out of 14 were mislabelled, whereas GC with adaptive regularization correctly labelled 12 ellipses (Figure 29(d)). To quantify the advantage of our approach, we tested a synthetic dataset of images containing 2 to 40 ellipses at various noise levels. We calculated the DSC of the

segmentation to the ground truth for each ellipse and averaged over all the ellipses in the image. Figure 30 plots the difference in average DSC between adaptive regularization GC and standard GC for images of increasing ellipse numbers. The same images were also tested at various noise levels. Note that a positive difference in the DSC indicates that our proposed regularization method with GC had greater success detecting low contrast ellipses.

(a)   (b)

(c)   (d)

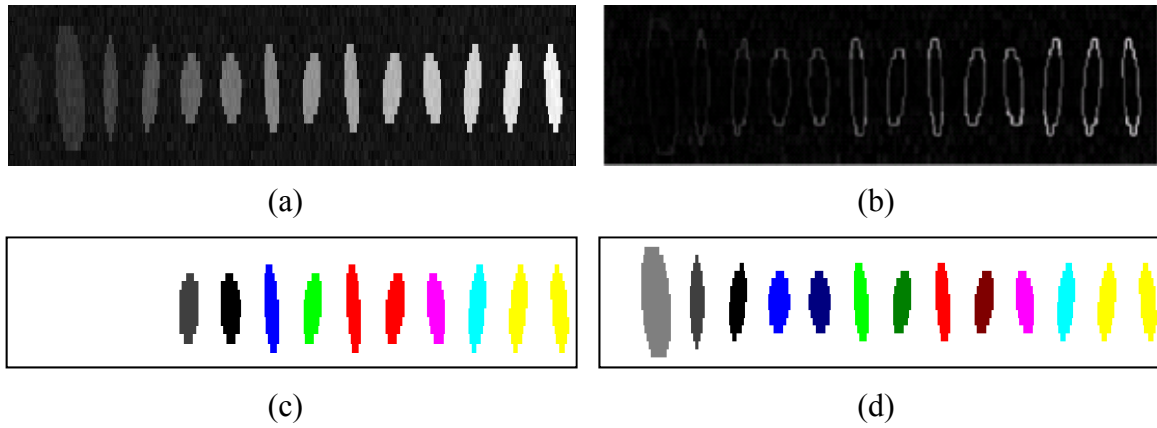**Figure 29: Segmentation of a synthetic image using GC with adaptive regularization. (a) Synthetic image of 14 ellipses with image contrast increasing from left to right. (b) Reliability calculated by our data-driven weights. (c) Segmentation from fixed weights, where each color represents a separate label. (d) Segmentation from data-driven weights, which result in greater numbers of low-contrast ellipses successfully segmented.**
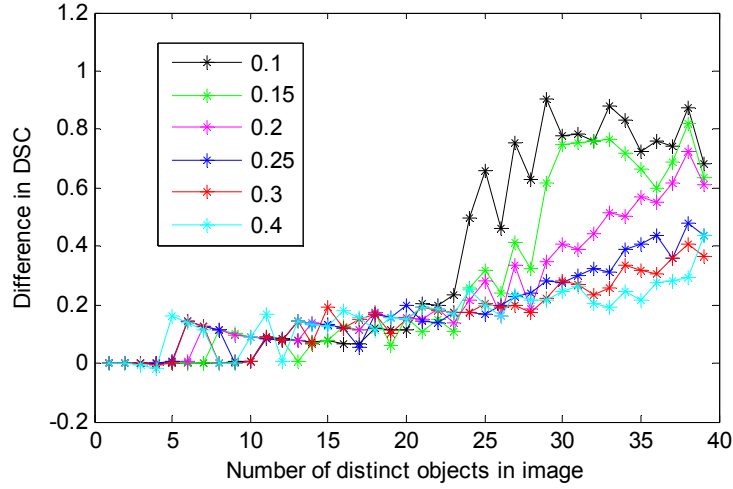
Figure 30: Difference in average DSC between adaptive regularization GC and fixed regularization GC for images with increasing numbers of ellipses. Different curves represent different noise standard deviations as shown in legend. Positive DSC difference indicates segmentations from data-driven weights are more successful than fixed weights at labelling ellipses with low image quality.

3.8.5 Medical Data Validation

We performed tests on the following medical datasets: set of 52 sagittal slices centered around the CC structure, 8 mammography images from the Digital Database for Screening Mammography (DDSM) [51, 50], 15 images from the BrainWeb database [63, 62, 31] using the T1, T2, and PD modalities with varying amounts of noise and intensity inhomogeneity, and coronal and transverse slices from the 18-subject Internet Brain Segmentation Repository (IBSR) database (provided by the Center for Morphometric Analysis at Massachusetts General Hospital and available at http://www.cma.mgh.harvard.edu/ibsr/).

3.8.5.1 Minimum-Path

Sagittal MR images of the CC exhibit the known problem of a weak boundary where the CC meets the fornix (as discussed and labelled in Figure 13(a) of Section 2.3.3). Unlike the globally optimal weights which produced bimodal behaviour (either

blue or red in Figure 13(b)), our method automatically boosts up the regularization (stronger red in Figure 31(a)) at the CC-fornix boundary producing a better delineation, as seen in the segmentation results (Figure 31(b)).

(a)

(b)
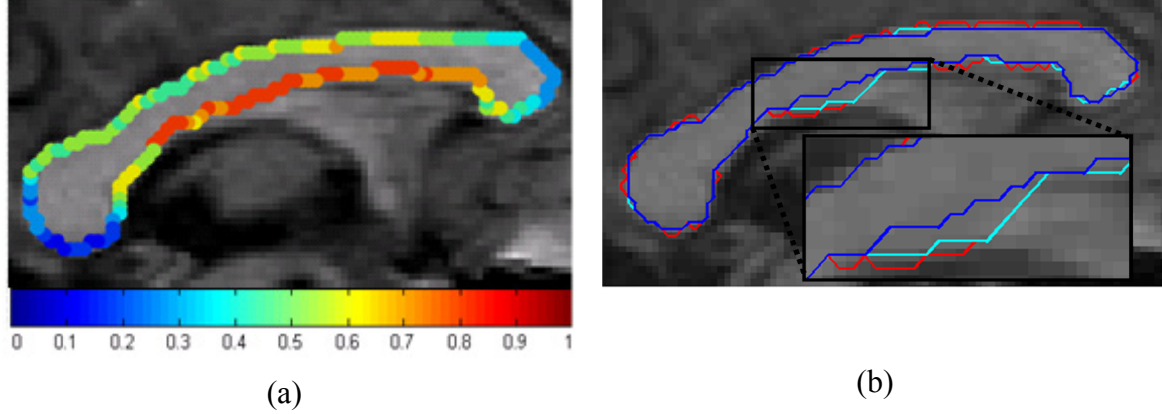
Figure 31: Segmentation from (a) proposed data-driven weight method for a corpus callosum MR image. The coloring of the contours reflects the value of the spatially-adaptive weight with pure blue corresponding to $w = 0$ and pure red to $w = 1$. The data-driven weights results in greater regularization in the difficult fornix region and has a smooth transition between weights. (b) Contours produced by using the proposed adaptive weight (*blue*), best fixed-weight (*red*), and the globally-optimum weight (*cyan*).

The quantitative results of the segmentations using the minimum-path framework on images from the CC dataset, IBSR dataset, BrainWeb dataset, and DDSM dataset are presented Figure 32, Figure 33, Figure 34, and Figure 35, where the DSC for each image was averaged over 25 segmentations. The segmentations of the IBSR dataset (Figure 33) were for the boundary of the cortical white matter in transverse and coronal slices. Figure 34 shows how the segmentations from the data-driven weights do not degrade at higher noise levels for the BrainWeb database using T1, T2, and PD modalities, where we performed the segmentations for 25 noise realizations per noise level. Figure 35 presents results for segmenting cancer tissue in mammography images from the DDSM dataset.

Figure 32: CC 52-image dataset DSC results with minimum-path segmentation method. Graph shows average DSC over dataset for segmentations from DD = data-driven weights, F = least-error fixed weights, and GO = globally optimal weights. DSC for each image was averaged from 25 segmentations with different seeds. DD average DSC (over total dataset) was 0.9224, F average DSC was 0.8984, GO average DSC was 0.8657. Average *p*-values for DD vs. F and for DD vs. GO were << 0.05. Segmentations from the DD weights were significantly more accurate.

(a)   (b)
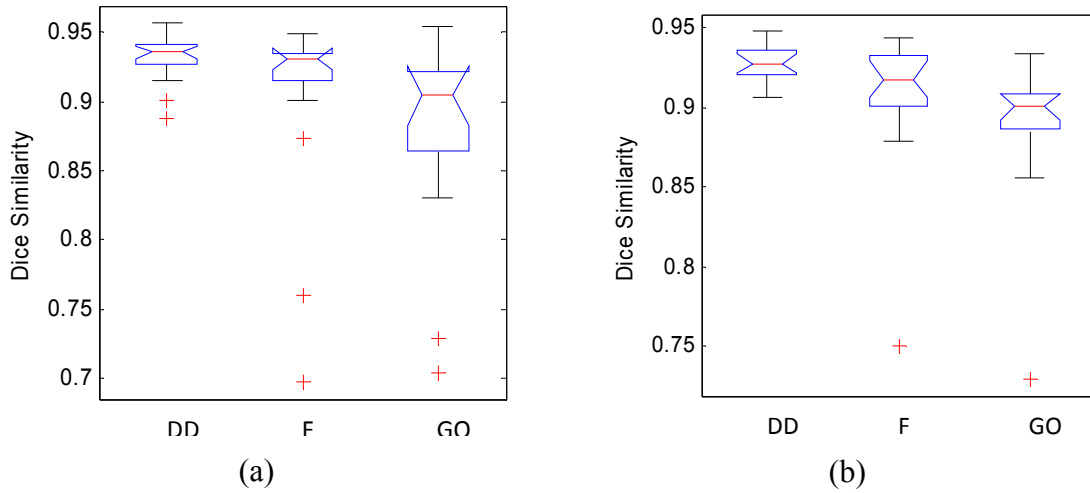
Figure 33: DSC results for IBSR dataset of (a) coronal and (b) transverse MR slices from 18 subjects when segmenting for white matter using minimum-path approach. Graph shows mean and variance for segmentations produced by data-driven weights (DD), least-error fixed weight (F), and the globally optimal weights (GO). The DD average DSC (over 18 image dataset) was 0.9323 for the coronal dataset and 0.9278 for the transverse dataset. The F average DSC was 0.9036 and 0.9094 for the coronal and transverse datasets, respectively, and the GO average DSC was 0.8806 and 0.8886 for the coronal and transverse. All *p*-values were $\ll 0.05$. Although the difference between the segmentations from the DD weights and the fixed weights were low, the DD weight segmentations were still significantly more accurate.

97
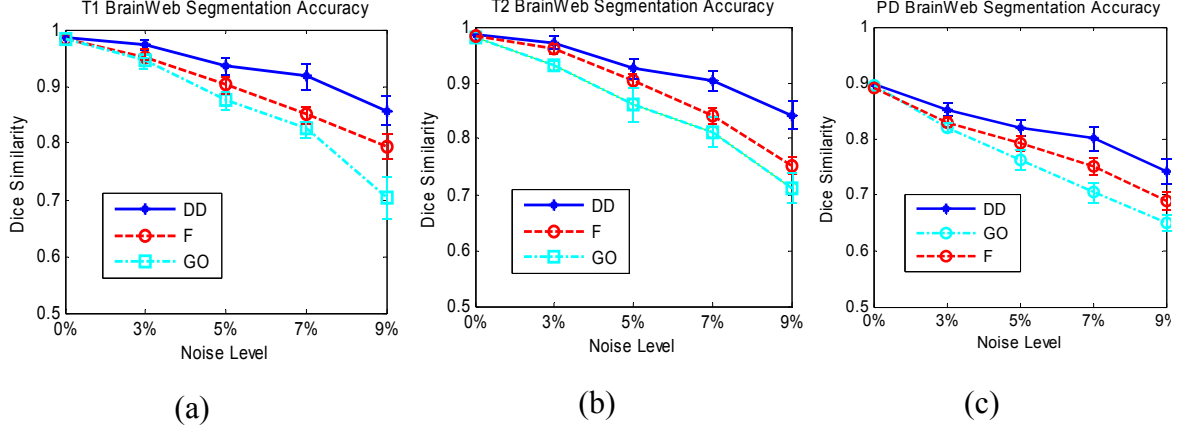
Figure 34: DSC results for slices from coronal BrainWeb dataset when segmenting for white matter using minimum-path approach. (a) T1 data, (b) T2 data, and (c) PD data. Data-driven weights (DD) produce segmentation accuracies that decrease less than fixed weights (F) and globally optimum weight (GO) as the level of noise in the image increases from 0% to 9% (averaged over 25 noise realizations per noise level).

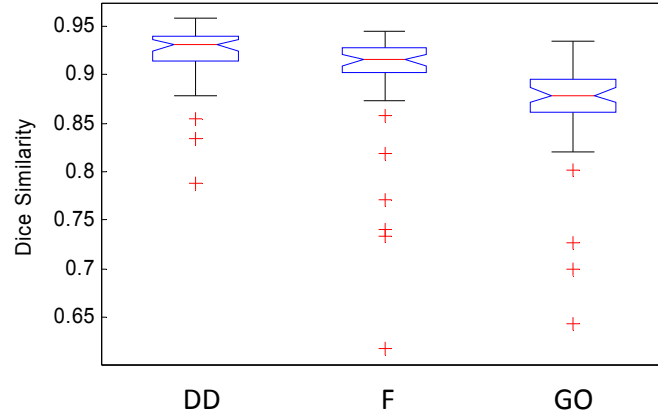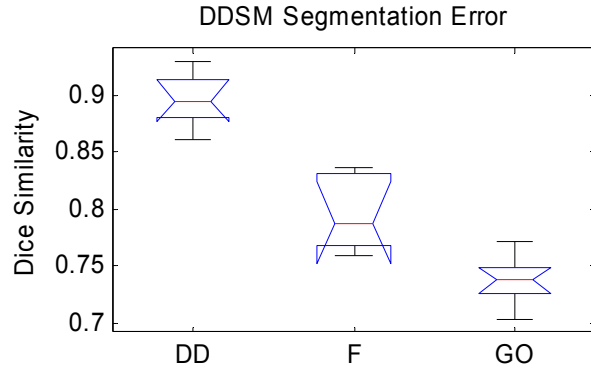**Figure 35: DSC results for segmentation of cancer tissue in 8 mammography images from DDSM database using minimum-path approach. The data-driven weights (DD) resulted in a mean error of 0.8958, the fixed-weight (F) in a mean error of 0.7961 and the globally-optimal weights (GO) in 0.7370 with all p-values $\ll 0.05$. Segmentations from the DD weights show a clear improvement over segmentations from F and GO weights.**

3.8.5.2 Graph Cuts

We next present qualitative results using graph cuts with our data-driven regularization weights on images from BrainWeb. Figure 36(a) shows a T1 image with an intensity inhomogeneity of 20%. High curvature-modulated reliability in the cortical folds (Figure 36(b)) results in lower regularization in these regions. The overlayed GC segmentations (Figure 36(c)) using the adaptive regularization weight versus the least-

98

error fixed weight shows greater segmentation accuracy in high curvature regions. Additionally, the curvature-modulated regularization weights show improvements over the non-curvature weights of Section 3.4 (Figure 36(d)). Figure 37(a) shows the same T1 image of Figure 36(a) but with a noise level of 7%. The resulting curvature-modulated reliability map (Figure 37(b)) is not corrupted by the noise and still enforces greater regularization in high curvature cortical folds, as seen in the resultant segmentation comparisons of Figure 37(c) and Figure 37(d). At higher noise levels, our data-driven weights results in a more accurate segmentation than the standard least-error uniform weight, and even more accurate than our non-curvature image reliability system, thus verifying the importance of the noise-gated curvature cue.
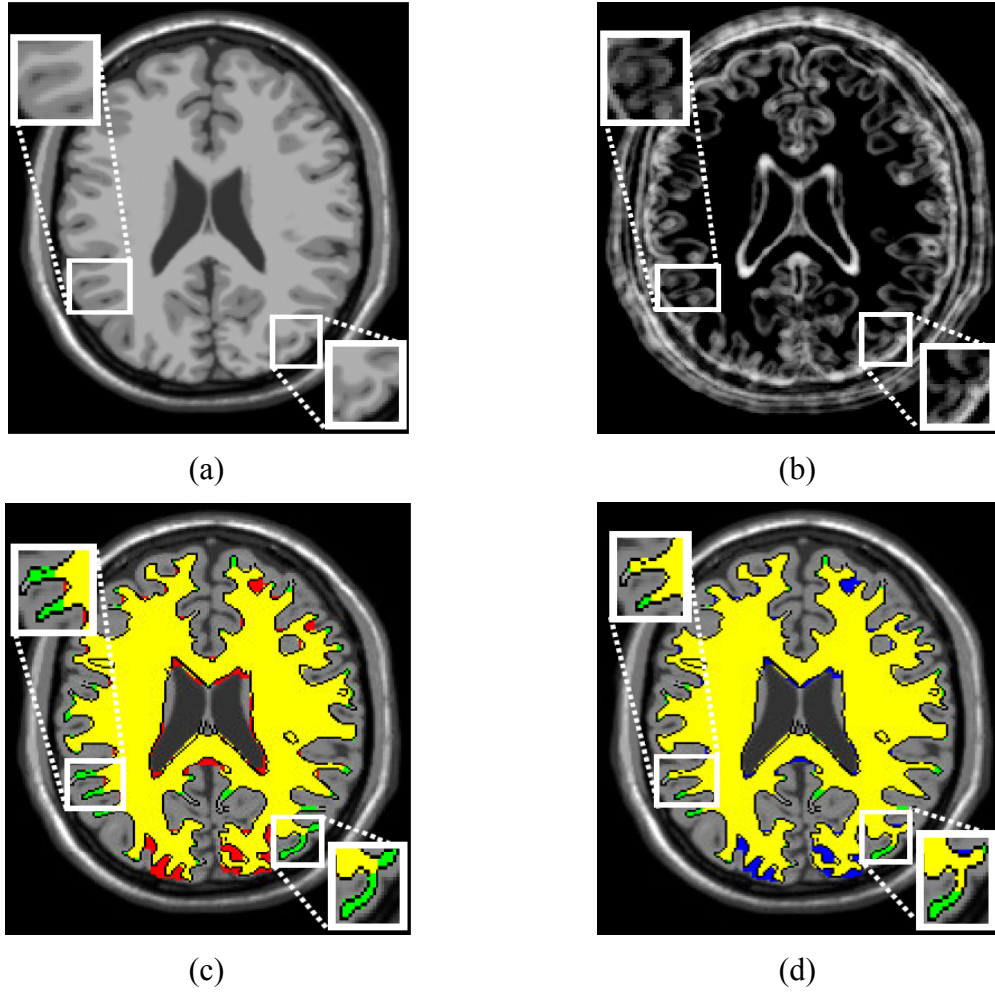
Figure 36: Segmentation of MR data from BrainWeb using GC with curvature-modulated regularization. (a) T1 slice with 20% intensity non-uniformity. (b) Curvature-modulated reliability. *Black* **intensities corresponds to 0 (low reliability/high regularization) and** *white* **to 1. Note higher reliability in cortical folds. (c) Comparison of segmentations from the curvature-modulated weight** (*green*) **to the least-error fixed weight** (*red*)**, and (d) to the non-curvature modulated image reliability weight** (*blue*)**. Yellow regions are where the segmentations overlap, and ground truth contour is shown in black. Segmentations from the curvature-modulated DD weights result in no leakage into the background and accurately capture the high-curvature cortical folds.**
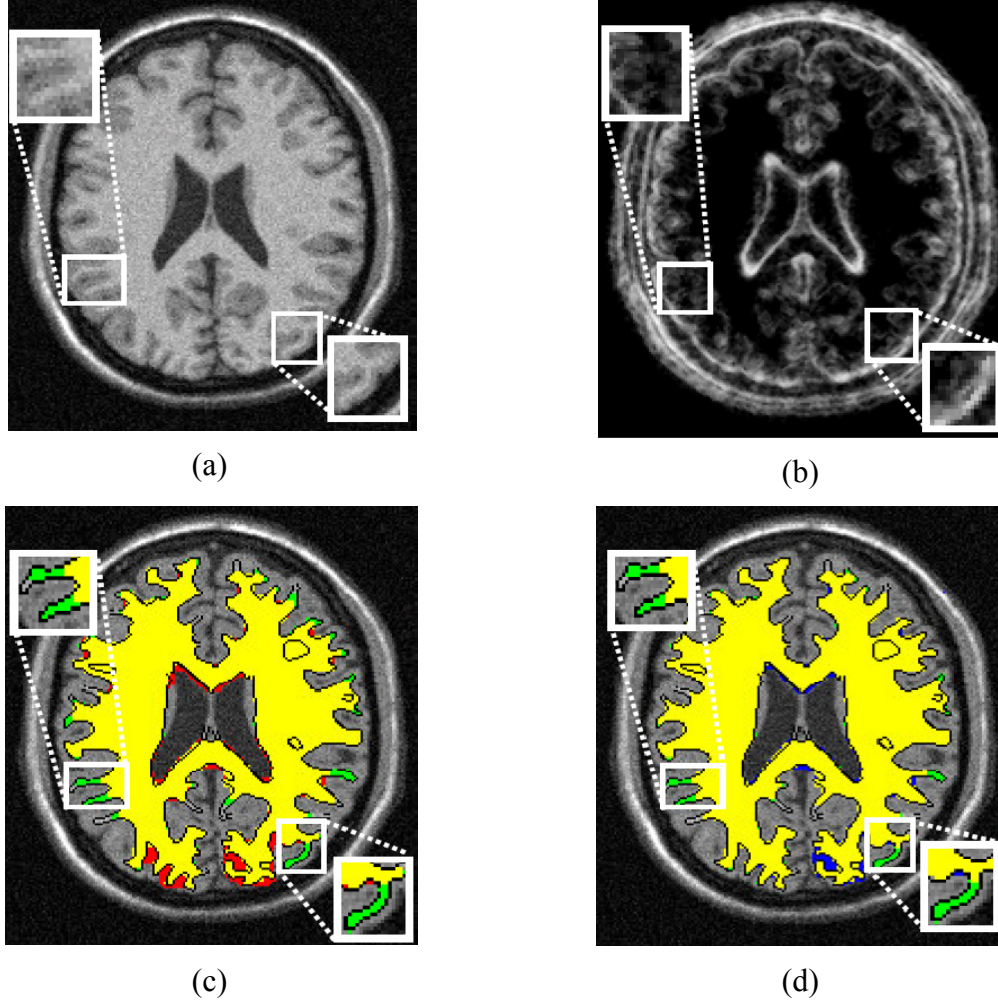
Figure 37: Segmentation of noisy MR data from BrainWeb using GC with curvature-modulated regularization. (a) T1 slice with 7% noise level. (b) Curvature-modulated reliability. (c) Comparison of segmentations from the curvature-modulated adaptive weight (*green*) to the least-error fixed weight (*red*), and (d) to the non-curvature modulated image reliability weight (*blue*).

We present quantitative results for our graph cuts tests with the BrainWeb, DDSM, and IBSR datasets in Figure 38, Figure 39, and Figure 40. On average, these results are lower than the results from the minimum-path segmentation approach because the lack of seedpoints reduces the ability to target the object of interest for each image. For the BrainWeb dataset (Figure 38), as the level of noise increases (where we performed the segmentations for 25 noise realizations per noise level), the data-driven weights produce more accurate segmentations than the least-error fixed weights. For the

101

DDSM dataset (Figure 39), both methods produce segmentations with low DSC which is due the complex background in the mammography images, where the non-cancerous tissue has similar intensity values to the cancerous tissue. Unlike minimum-path approaches where we can target the tissue object of interest with seedpoints, the seeding used for graph cuts is only for intensity profile estimation. The result is that many objects in the background (breast tissue) are mistaken as cancerous tissue due to similarities in intensity values. For the IBSR segmentations (Figure 40), the data-driven weights produced significantly more accurate segmentations for the coronal and transverse slices (when segmenting for white matter). When segmenting for the CC structure in the sagittal slices (Figure 40(c)), the data-driven weights were not significantly more accurate due to similarities in intensity between the CC structure and the fornix.

(a)   (b)   (c)
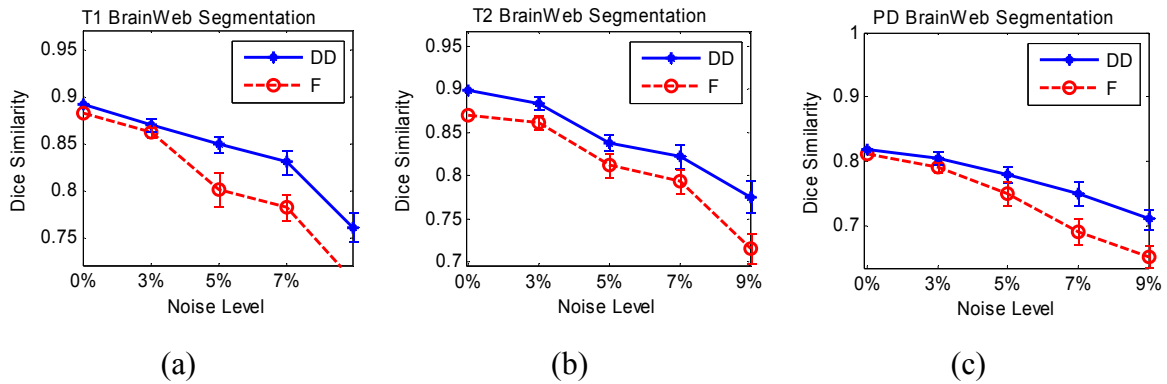
Figure 38: DSC results for BrainWeb segmentation of white matter using graph cuts approach for (a) T1, (b) T2, and (c) PD modalities with increasing noise levels. Data-driven weights (DD) produce less degradation in results at high noise levels when compared to least-error fixed weights (F). DSC values are averaged over segmentations for 25 noise realizations per noise level. As the noise level increases, the segmentations from the DD weights result in greater DSC than the fixed weight segmentations.

DDSM Segmentation by Graph Cuts
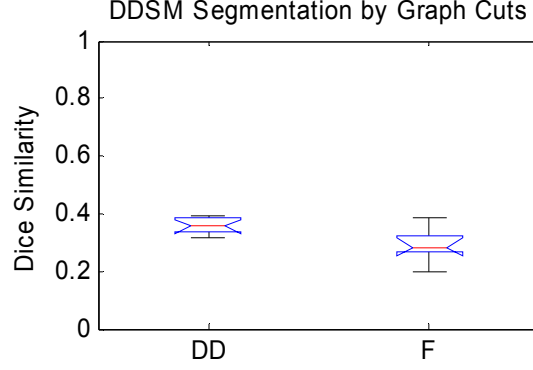
Figure 39: DSC results for segmentation of cancer tissue in 8 mammography images from DDSM database using the graph cuts approach. As these images represent difficult scenarios due to noise, weak edges, and non-piecewise constant objects, both methods for regularization weights produce poorer results. However, the data-driven weights (DD) produce significantly improved results over the fixed weights (F) segmentation. DD average DSC = 0.3576, F average DSC = 0.2909.

(a)   (b)   (c)
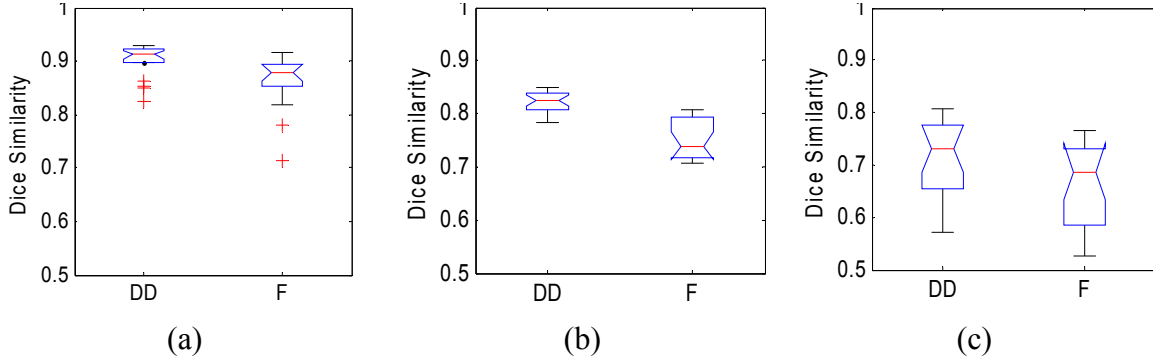
Figure 40: DSC results for IBSR dataset of (a) coronal, (b) transverse and (c) sagittal MR slices from 18 subjects when segmenting for white matter (for (a) and (b)) and CC (for (c)) using graph cuts. The DD average DSC (over 18 image dataset) was 0.8999 for the coronal dataset, 0.8199 for the transverse dataset, and 0.7149 for the sagittal dataset. The F average DSC was 0.8627, 0.7462, and 0.6611 for the coronal, transverse, and sagittal datasets, respectively. Average *p*-values were 0.0118, 9.89E-9, 0.052. For the CC segmentations in the sagittal plane, the proposed method did not produce significantly improved results. This is since the fornix and CC have the same intensity and region-based external terms are not successful in differentiating the structures.

3.8.5.3   Active Contours Without Edges

We next present results using the ACWE segmentation framework on medical datasets. Figure 41(a) shows a PD coronal slice with 5% noise from BrainWeb. The resulting curvature-modulated reliability (Figure 41(b)) is bright for the high curvature tips of the central ventricles and for the cortical folds of the white matter. The resulting segmentation (Figure 41(c)) from the data-driven weights (*green*) for the white matter is

103

able to capture the cortical folds while not oversegmenting into the ventricle region, unlike the least-error fixed weight result (*red*).

<center>(a)   (b)   (c)</center>
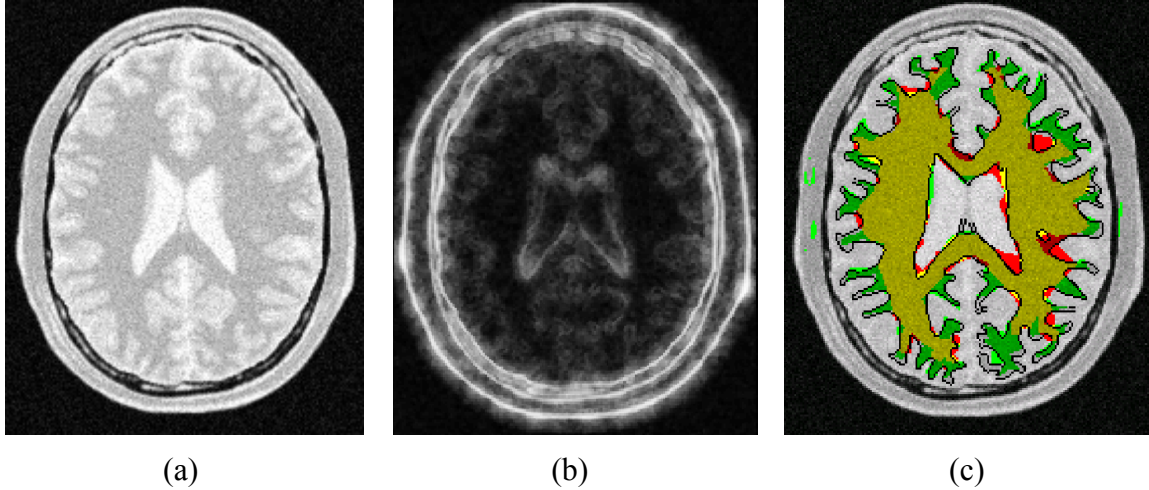
Figure 41: Segmentation of white matter in PD coronal slice from BrainWeb with 5% noise. (a) Original image, (b) curvature-modulated reliability, and (c) comparison of segmentations from data-driven weight (*green*), least-error fixed weight (*red*), where overlapping regions are in *yellow* and ground truth contour shown in *black*. Data-driven segmentation results in less over-segmentation into ventricle region and better segmentation of high curvature tips.

In addition, we present segmentations of cancer tissue from mammography images (DDSM database) such as in Figure 42, where the ACWE contour was initialized as a square shown in Figure 42(a) and iterations were run until the contour evolution converged (at most 700 iterations). The segmentation from the data-driven weights more accurately grows into the left region of the tissue. Neither segmentation accurately contains the right region of the tissue which is due to the cancer object containing differing intensities.

<center>(a)   (b)   (c)</center>
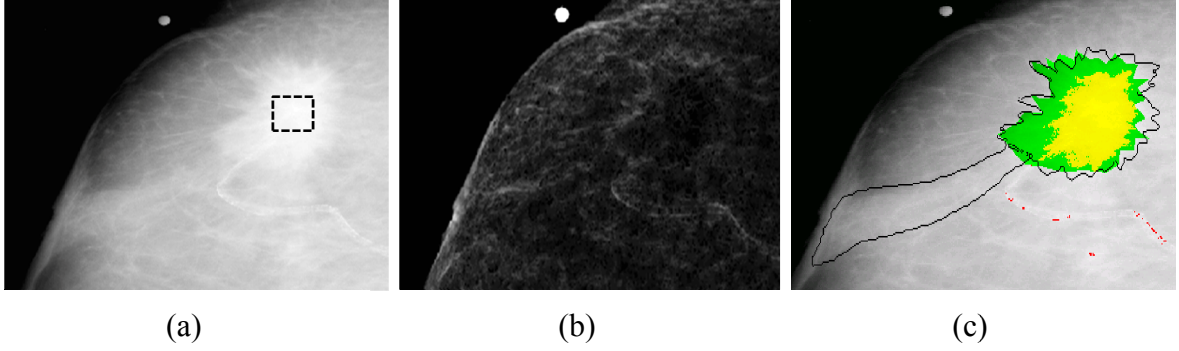
Figure 42: ACWE segmentation of mammography image from DDSM database [50, 51]. (a) Original image with initial ACWE contour, (b) curvature-modulated reliability, and (c) comparison of segmentations from data-driven weights (*green*) and fixed-weights (*red*) where *yellow* represents overlap and ground truth is shown in *black*. Segmentation from DD weights more successfully capture the white-intensity region of the cancer tissue. Neither segmentation correctly captures the low-intensity region of the tissue which represents a difficult case.

We present quantitative results for the BrainWeb, DDSM, and IBSR datasets in Figure 43, Figure 44, and Figure 45. As with graph cuts, the ACWE results on the DDSM database were low for both the fixed weight and data-driven weight due to greatly different intensity values in the cancer masses as shown in Figure 42(a). The data-driven weights provide significantly more accurate segmentations for the IBSR coronal and transverse datasets (Figure 45) when segmenting for white matter, but insignificant results for the IBSR sagittal dataset when segmenting for the CC structure due to the similarities in intensity between the CC and fornix.

<center>105</center>
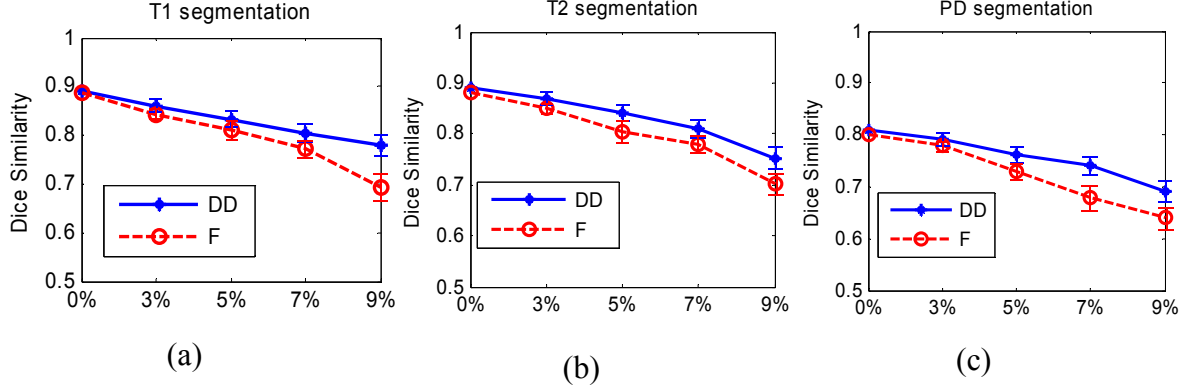
Figure 43: DSC results for BrainWeb segmentation of white matter using ACWE approach for (a) T1, (b) T2, and (c) PD modalities with increasing noise levels. Data-driven weights (DD) produce less degradation in results at high noise levels when compared to least-error fixed weights (F). DSC values averaged over segmentations from 25 noise realizations per noise level. Segmentations from the DD weights are more accurate than fixed weight segmentations at high noise levels.
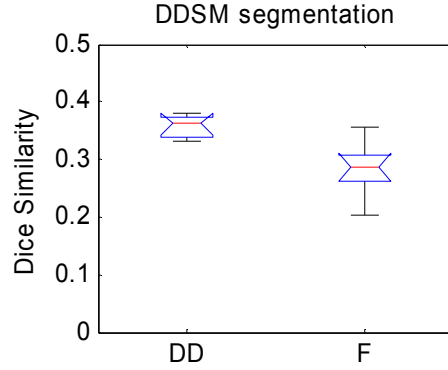
Figure 44: DSC results for segmentation of cancer tissue in 8 mammography images from DDSM database using the ACWE approach. The data-driven weights (DD) produce significantly improved results over the fixed weights (F) segmentation. DD average DSC = 0.3565, F average DSC = 0.2833. Average *p*-value was 6E-4. Although the DD weight segmentations are more accurate, both methods produce low accuracies since neither method could segment regions of the cancer tissue with greatly differing intensity values.

106
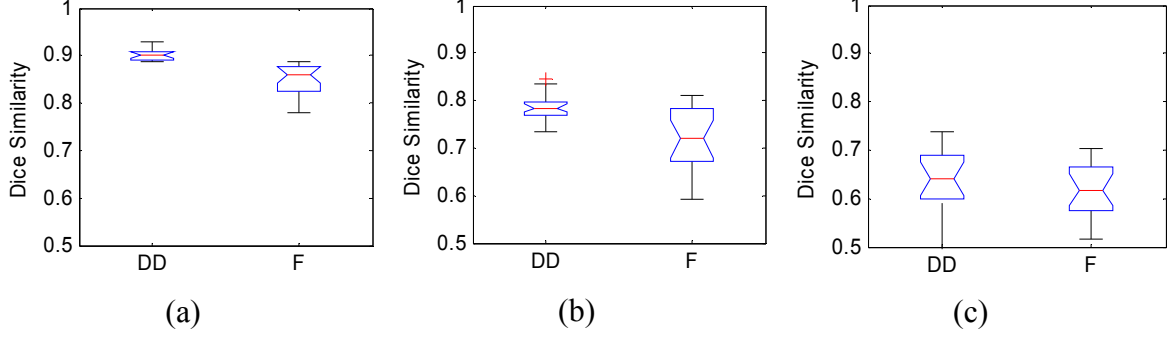
Figure 45: DSC results for IBSR dataset of (a) coronal, (b) transverse and (c) sagittal MR slices from 18 subjects when segmenting for white matter (for (a) and (b)) and CC (for (c)) using ACWE. The DD average DSC (over 18 image dataset) was 0.9019 for the coronal dataset, 0.7895 for the transverse dataset, and 0.6345 for the sagittal dataset. The F average DSC was 0.8466, 0.7196, and 0.6196 for the coronal, transverse, and sagittal datasets, respectively. Average *p*-values were 3.303E-7, 3E-4, 0.482. The segmentation from the DD weights was significantly more accurate for the coronal and transverse datasets, but was not significantly more accurate for the sagittal dataset when segmenting for the CC structure. This is due to the intensity similarities between the CC and the fornix structures which produces problems for region-based external terms.

3.8.5.4 Contextual Mumford-Shah Method

We validated the contextual Mumford-Shah method by comparing results to both the least-error fixed weight segmentation and to the segmentation produced the ET adaptive regularization method. The contextual MS method is automated and does not allow for user input to focus the segmentation on certain objects. As such, we chose to demonstrate segmentations of the central ventricle structure using the BrainWeb and IBSR databases rather than segmentations of the white matter which is a complex structure encompassing the entire image. We targeted the ventricles by cropping the image around the structure. Figure 46(a) and Figure 47(a) show the central ventricle region taken from coronal slices from the BrainWeb dataset, where Figure 46(a) depicts a PD image with a noise level of 5% and Figure 47(a) depicts a T1 image with a noise level of 9%. For both cases, the curvature-modulated reliability (Figure 46(b) and Figure 47(b)) is higher for the tips of the ventricle structure, and the resulting segmentations

107

(Figure 46(c) and (d), and Figure 47(c) and (d)) are more accurate against both alternate

methods due to lower regularization in the tip regions.

(a)

(b)

(c)

(d)
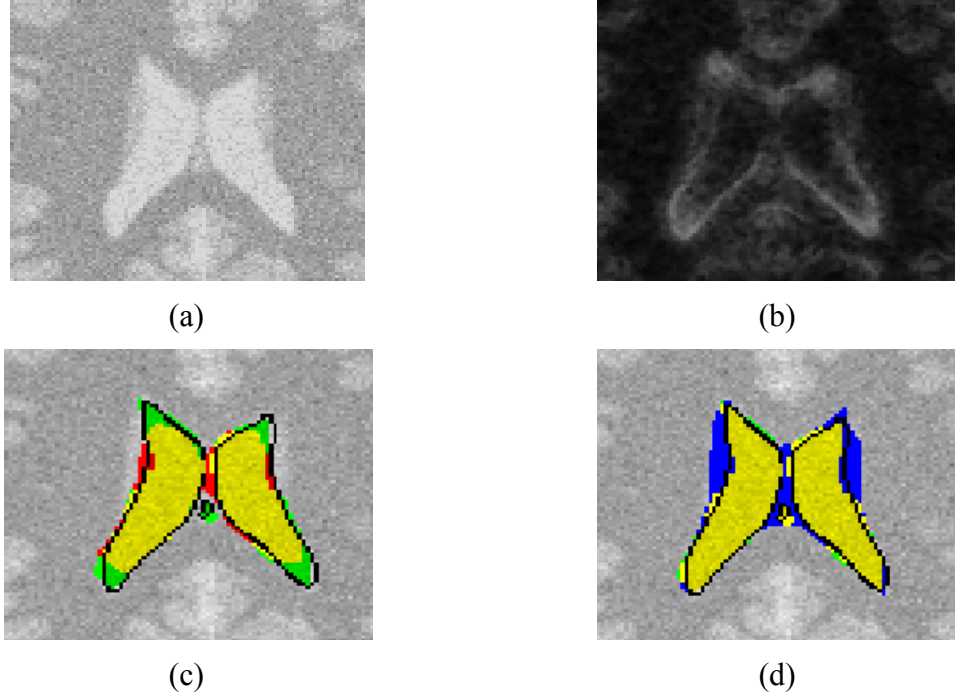
Figure 46: Segmentation of central ventricle structure from mid-volume coronal slices (BrainWeb) (a) PD image with 5% noise, and (b) curvature-modulated reliability. (c) Comparison of segmentations between data-driven weights (*green*) and fixed weights (*red*), and (d) between data-driven weights and Erdem-Tari (ET) weights (*blue*) where *yellow* regions represent segmentation overlap and *black* contour represents the ground truth. The data-driven weights provide lower regularization in the tips of the central ventricle structure and the weak edge sides.

(a)   (b)

(c)   (d)
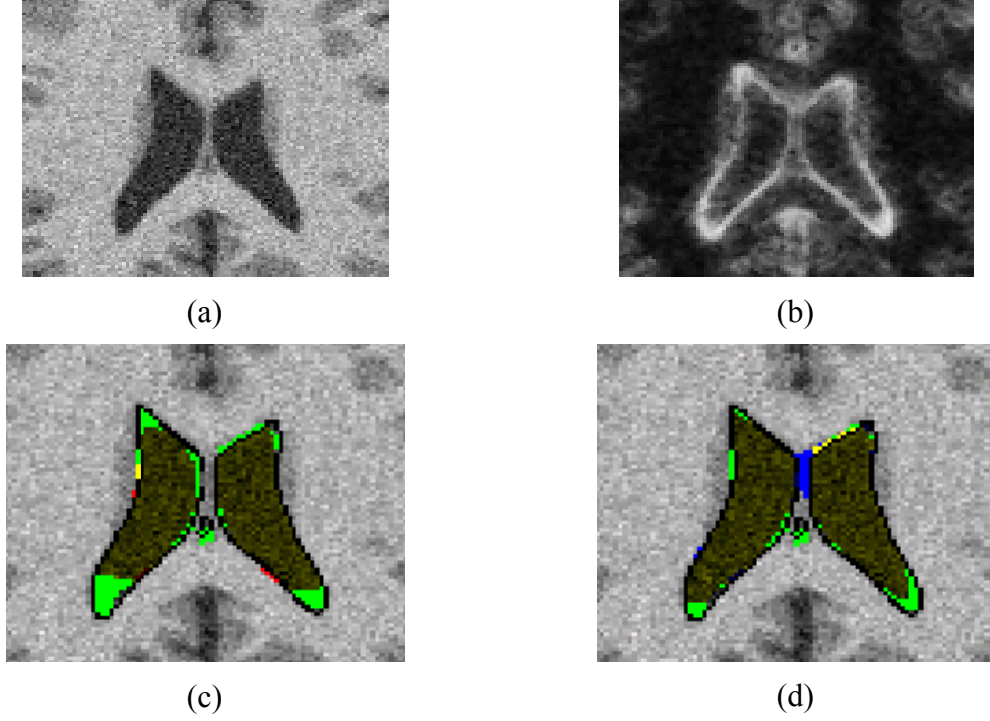
Figure 47: (a) T1 image with 9% noise (BrainWeb), and (b) curvature-modulated reliability. (c) Comparison of segmentations between data-driven weights (*green*) and fixed weights (*red*), and (d) between data-driven weights and the Erdem-Tari weights (*blue*) where *yellow* regions represent segmentation overlap and *black* contour represents the ground truth. Only the data-driven weights provide a segmentation that captures the high curvature tips.

We present the quantitative results for the BrainWeb, DDSM, and ISBR (coronal) datasets in Figure 48, Figure 49, and Figure 50. The data-driven weights provide significantly more accurate segmentations for the coronal IBSR ventricle application, and do not degrade over increased levels of noise for the BrainWeb segmentations.
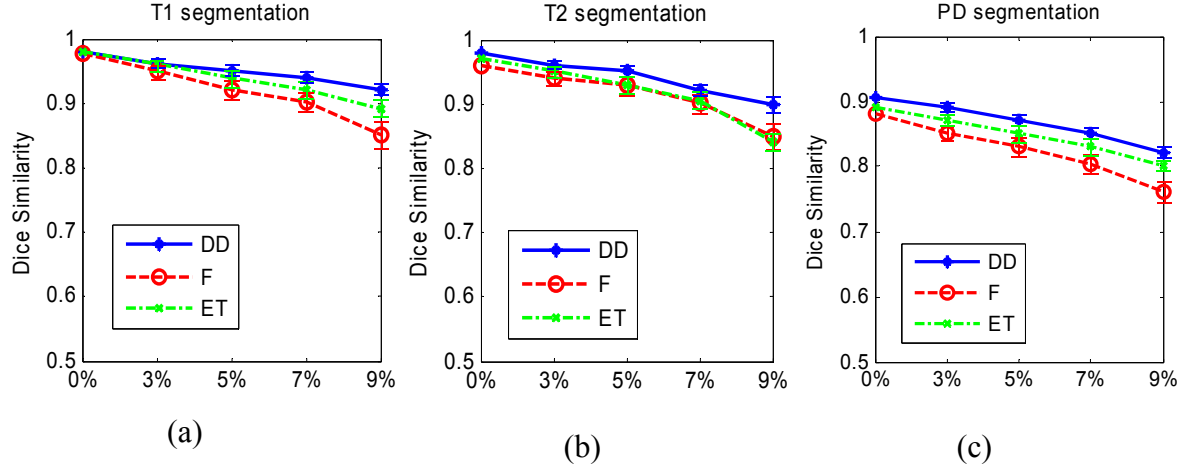
Figure 48: DSC results for BrainWeb segmentation of central ventricle structure using contextual MS approach for (a) T1, (b) T2, and (c) PD modalities with increasing noise levels. Data-driven weights (DD) produce less degradation in results at high noise levels when compared to least-error fixed weights (F) and Erdem-Tari weights (ET). DSC values are averaged over segmentations from 25 noise realizations per noise level.
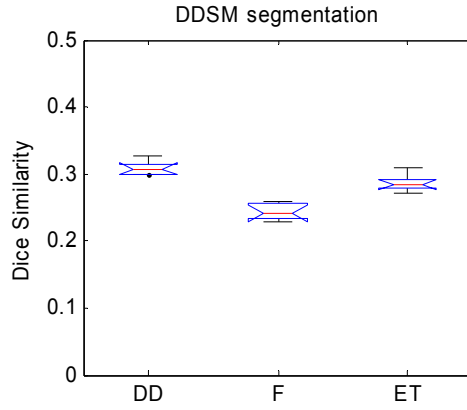
Figure 49: DSC results for segmentation of cancer tissue in 8 mammography images from DDSM database using the contextual MS approach. The data-driven weights (DD) produce significantly improved results over the fixed weights (F) segmentation. DD average DSC = 0.3092, F average DSC = 0.2433, ET average DSC = 0.2863. All results are poor due to the non-piecewise constant cancer object.

110
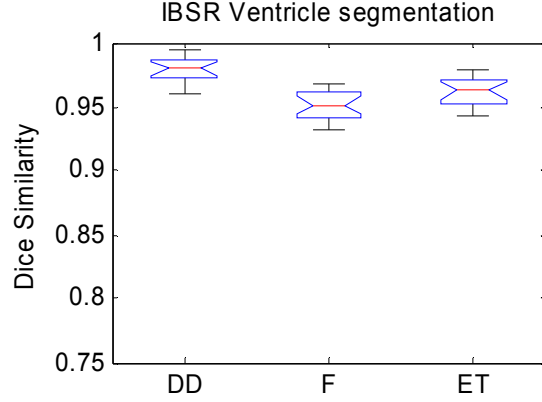
Figure 50: DSC results for IBSR dataset of coronal MR slices from 18 subjects when segmenting for the central ventricle structure using contextual MS approach. Graph shows mean and variance for segmentations produced by data-driven weights (DD), least-error fixed weight (F), and the Erdem-Tari weights (ET). The DD average DSC (over 18 image dataset) was 0.9783, the F average DSC was 0.9509, and the ET average DSC was 0.9617. All averaged *p*-values were $\ll 0.05$. Segmentation from the DD weights are significantly more accurate than the segmentations from the F and ET weights.

3.8.6 Natural Scenes Data Validation

We next present results from using the following datasets of natural scenes: the McGill Calibrated Colour Image Database [75], the PASCAL object recognition database [40], and through the ImageNet database [32, 33].

3.8.6.1 *Minimum-Path*

We first demonstrate results on an image from the McGill database, the tree leaf on a complicated background shown in Figure 51(a). The resulting curvature-modulated reliability measure (Figure 51(b)) indicates higher regularization at the regions of the leaf obscured by snow and lower regularization in the leaf tip regions. The resulting segmentation from the data-driven segments the obscured shape of the leaf accurately (Figure 51(c)).

111
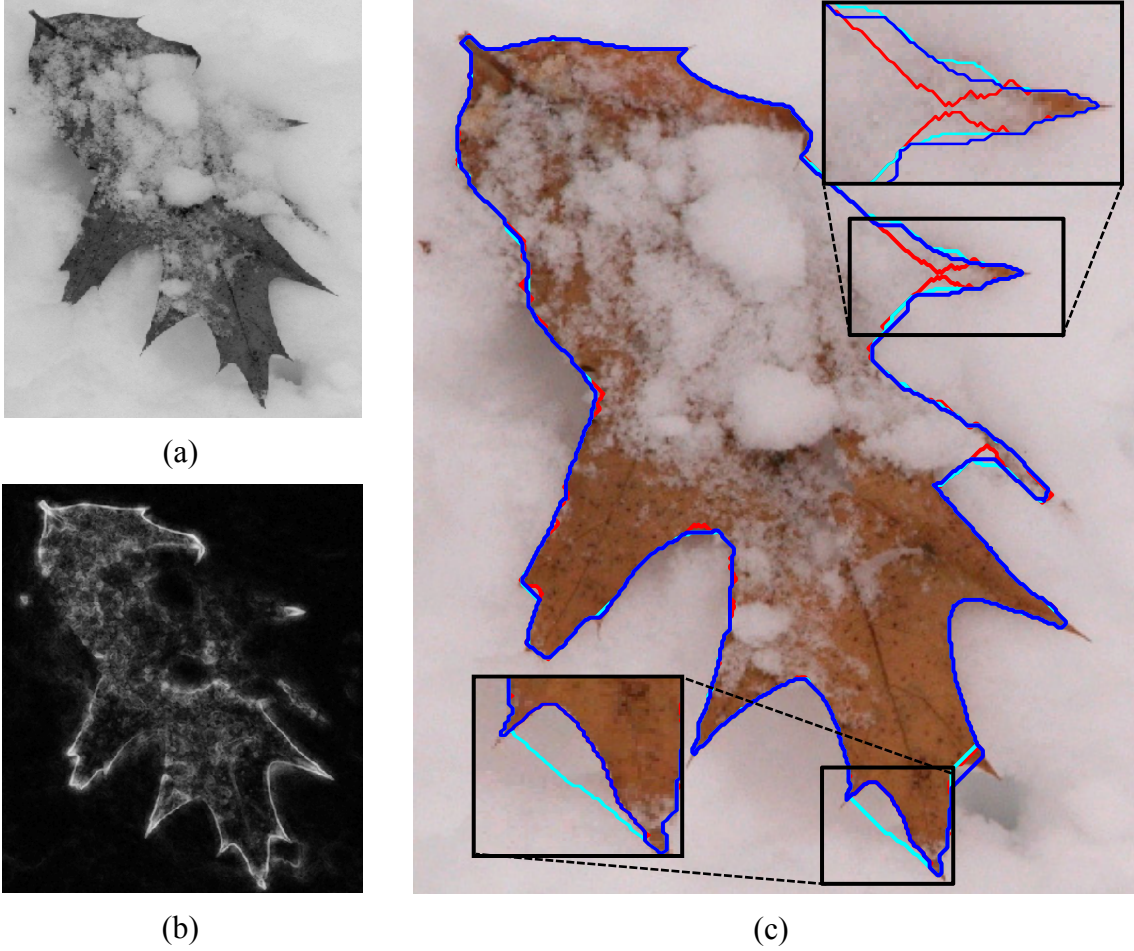
Figure 51: (a) Original leaf image (McGill dataset [75]). (b) Reliability calculated by our proposed method. Contours produced by using (c) data-driven weights (*blue*), least-error fixed-weight (*red*), and globally optimal weight (*cyan*). The GO contour fails to capture the leaf tip region, and the fixed weight contour fails to regularize in the obscured edge region, as highlighted.

Additionally, we segmented the airplane image of Figure 52(a) (from the PASCAL database) which consists of high curvature structures that require lower regularization, and weak-edged regions that require high regularization. The curvature-modulated reliability measure (Figure 52(b)) shows that the plane tips in the image are assigned low regularization weights (high reliability). The resulting segmentation using the data-driven weights correctly captures these regions (Figure 52(c)) whereas the least-error fixed weight, 0.5 in this case, correctly regularizes the weak edge regions but is too excessive to capture the airplane tips. The globally optimum weight, which was

predominately zero for this segmentation, correctly segments the tips but inaccurately includes the weak edge regions.

(a)　　　　　　　　　　　　　　　　　　　(b)

(c)
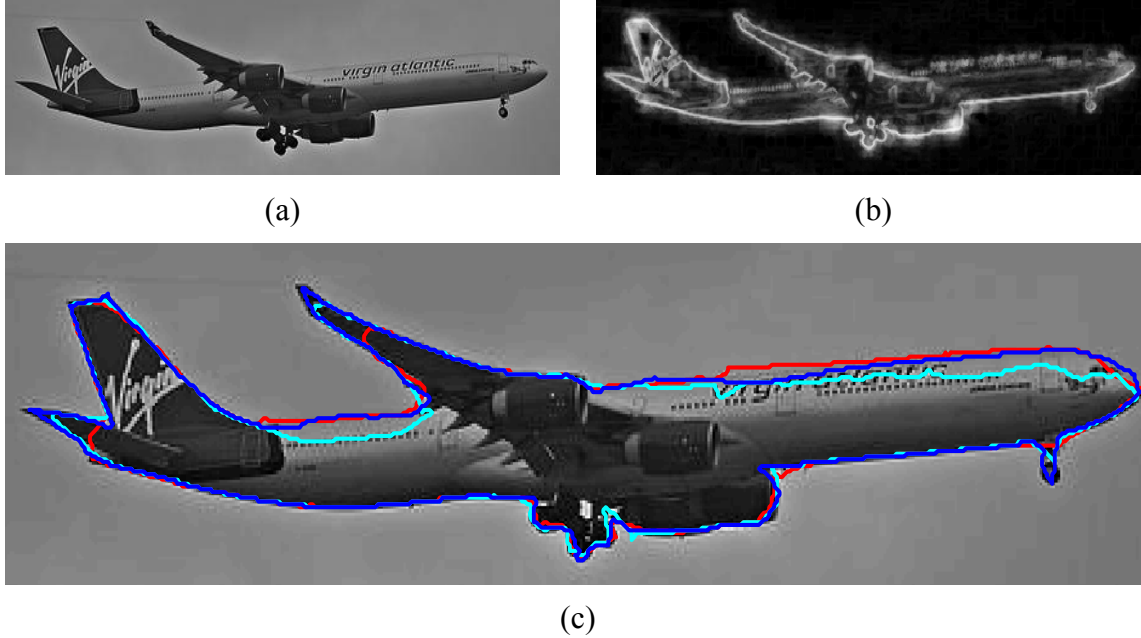
Figure 52: Airplane image (PASCAL dataset [40]) segmented by minimum-path approach. (a) Original image, (b) curvature-modulated reliability, and (c) contours from data-driven weight (*blue*), least-error fixed weight (*red*) and globally optimal weight (*cyan*). Segmentation from the GO weights inadequately regularizes in the weak edge upper region of the plane. Segmentation from the fixed weight fails to capture the wing tips.

We also present the contours produced by the data-driven weights for the examples shown in Section 2.3.4 of Chapter 2. Figure 53 shows segmentations of a complex-boundary flower and a submerged leaf. In Figure 53(a), only the data-driven weight in *blue* prevents oversegmenting of the image and background leakage. In Figure 53(b), the data-driven weight segmentation accurately captures the weak-edge leaf tip region.

<center>(a)   (b)</center>
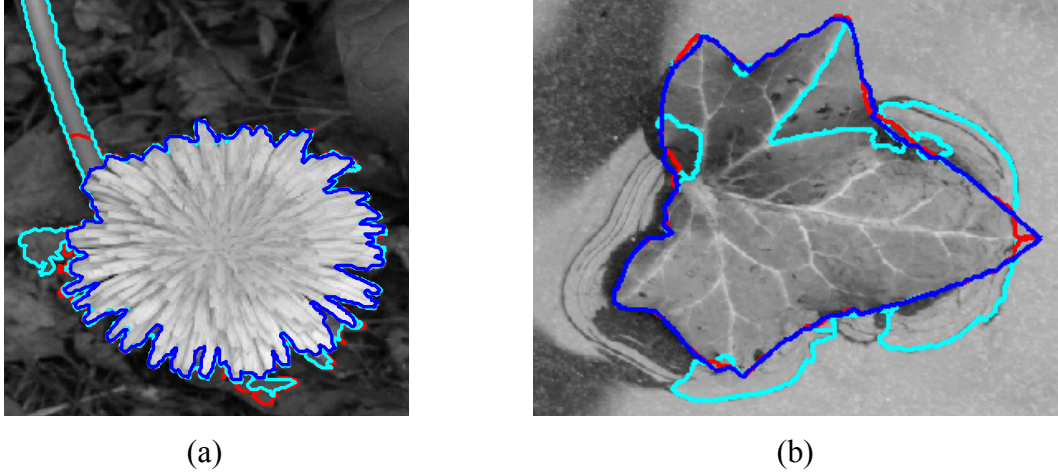
Figure 53: Segmentations from McGill database images of (a) a flower with a complex background and (b) a leaf with regions of the boundary obscured by water. The segmentation from the data-driven weights (*blue*) provide greater regularization in weak edge regions and lower regularization in high curvature regions, unlike the segmentations from the globally-optimal weight (*cyan*) and the spatially fixed weight (*red*).

We present the quantitative results on the McGill, PASCAL, and ImageNet databases in Figure 54. The ANOVA results indicate the significant improvement the data-driven weights provide when compared to existing methods. We note that these databases do not contain texture images, which we will address in our graph cuts and contextual MS tests.
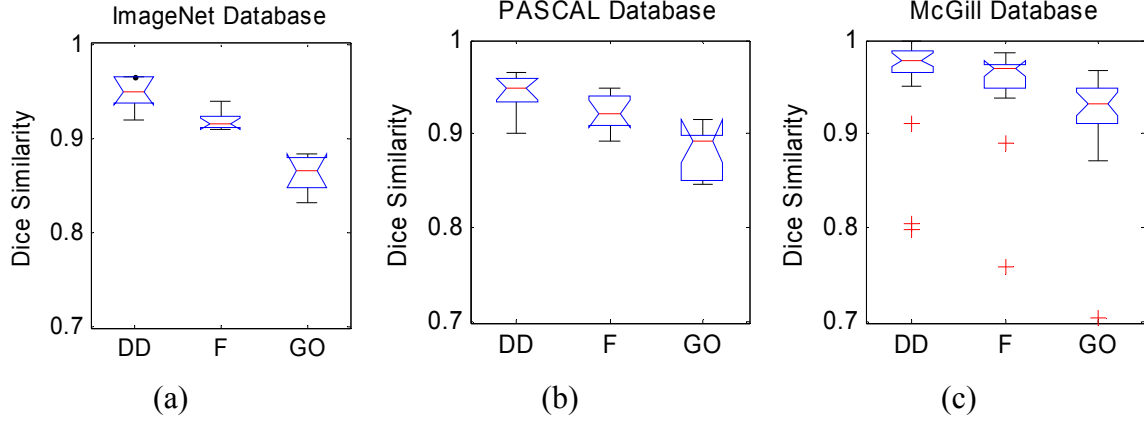
Figure 54: DSC of segmentations from data-driven weights (DD), least-error fixed weights (F) and globally-optimal weights (GO) on (a) 8 images from ImageNet database, (b) 11 images from PASCAL database, and (c) 24 images from McGill database. Average error over dataset (and over 25 segmentations for each image) was DD = 0.9489, F = 0.9188, GO = 0.8629 for (a), DD = 0.9455, F = 0.9241, GO =0.8783 for (b), and DD = 0.9619, F = 0.9402, GO = 0.9025 for (c). All *p*-values $\ll 0.05$. For all datasets, the segmentations from the DD weights produce significantly higher DSC than the segmentations from the GO weights and fixed weights.

3.8.6.2 *Graph Cuts*

We present results of the graph cuts segmentation method with our proposed regularization framework on a flower image from the McGill database, as shown in (Figure 55(a)) where this image has been corrupted by AWGN with a standard deviation of 0.3. From this image, we produced the curvature-modulated reliability mapping in (Figure 55(b)). The higher curvature-modulated reliability in the petal tip regions allows for a more accurate segmentation when compared to the least-error fixed weight segmentation (Figure 55(c)). In addition, we investigate the key role of the curvature cue by comparing the curvature-modulated segmentation to the non-curvature reliability weight segmentation (Figure 55(d)) which, as expected, required higher regularization in the detailed petal tip regions, resulting in leakage into the background.

115

(a)

(b)

(c)

(d)
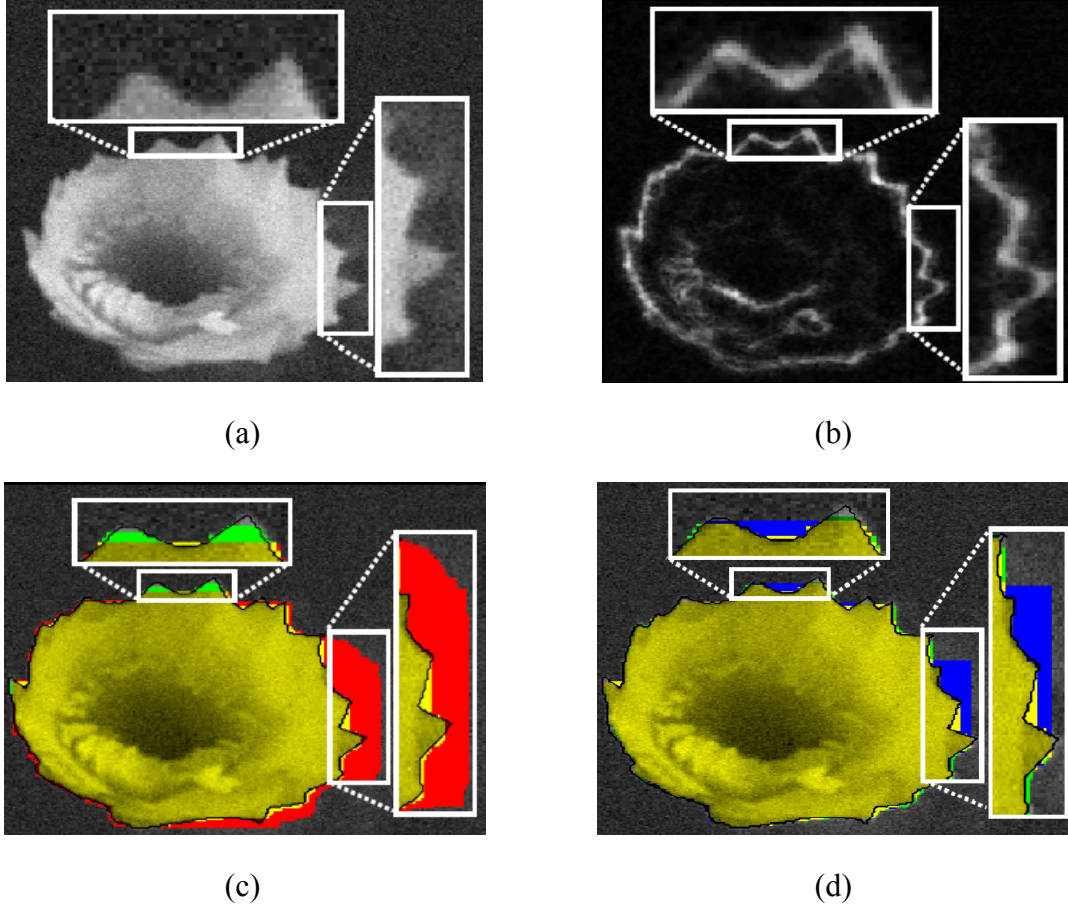
Figure 55: Graph cuts segmentation of flower image from ImageNet dataset [32]. (a) Original image with AWGN of standard deviation 0.3. (b) Curvature-modulated reliability (higher in petal tip and crevice regions) (c) Comparison of segmentations from the curvature-modulated reliability weight (*green*) to the least-error fixed weight (*red*), and (d) to the non-curvature reliability weight (*blue*) with overlapping regions in *yellow*. The curvature-modulated DD weights provided the best segmentation of the petal tips and had the least amount of leakage.

We segmented another flower image from the McGill database with corruption by AWGN of standard deviation 0.3 (image values normalized to range between 0 and 1), as shown in Figure 56(a). The curvature-modulated reliability (Figure 56(b)) produces lower regularization weights in the petal tips and petal crevices. In Figure 56(c), the fixed-weight segmentation excessively regularizes in the petal region, resulting in leakage (shown in *red*). Our method does not leak into the background and is able to capture the petal tips (shown in *green*).

(a)   (b)

(c)   (d)
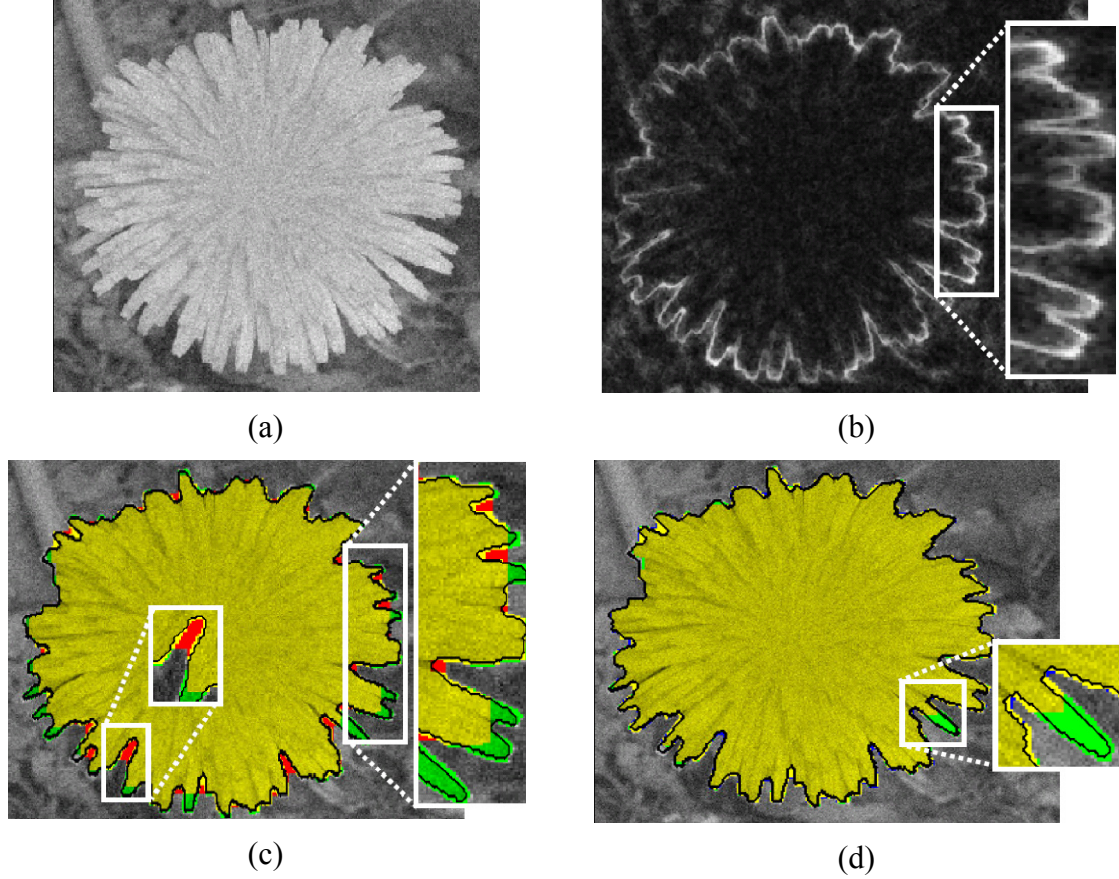
Figure 56: Graph cuts segmentation of natural image (a) Original image corrupted by AWGN with standard deviation of 0.3. (b) Curvature-modulated reliability. (c) Comparison of segmentations from the data-driven weight (*green*) to the least-error fixed weight (*red*) and to non-curvature reliability weight (*blue*) with overlapping regions in *yellow*. In (c), high regularization in the background prevents the segmentation from the curvature-modulated DD weights from leaking, unlike the fixed-weight method in *red*. In (d), the non-curvature modulated weights (see overlapped region in *yellow*) fail to capture all petals.

We demonstrate the ability of the texture-modulated weight $w_t(x, y)$ from (61) (see Section 3.6.2) to segment the textured image of Figure 57(a) (from ImageNet) where we set the parameter γ_t in (59) to 0.1. The curvature modulated reliability shown in Figure 57(b) is erroneously large for regions with texture. The curvature-and-texture modulated reliability shown in Figure 57(c) is lower for the texture edges. The resulting GC segmentation from using $w_t(x, y)$ is shown in Figure 57(d). Incorporation of a texture cue reduces leakage into the background, and the curvature cue reduces

117

regularization in the protrusion region of the plant. However, this example does highlight how the curvature measure can mistake texture regions for structurally important edges when in fact these regions should be regularized to prevent inclusion into the segmented contour. We are able to solve the issues by incorporating existing texture cues into our method.

(a)

(b)

(c)

(d)
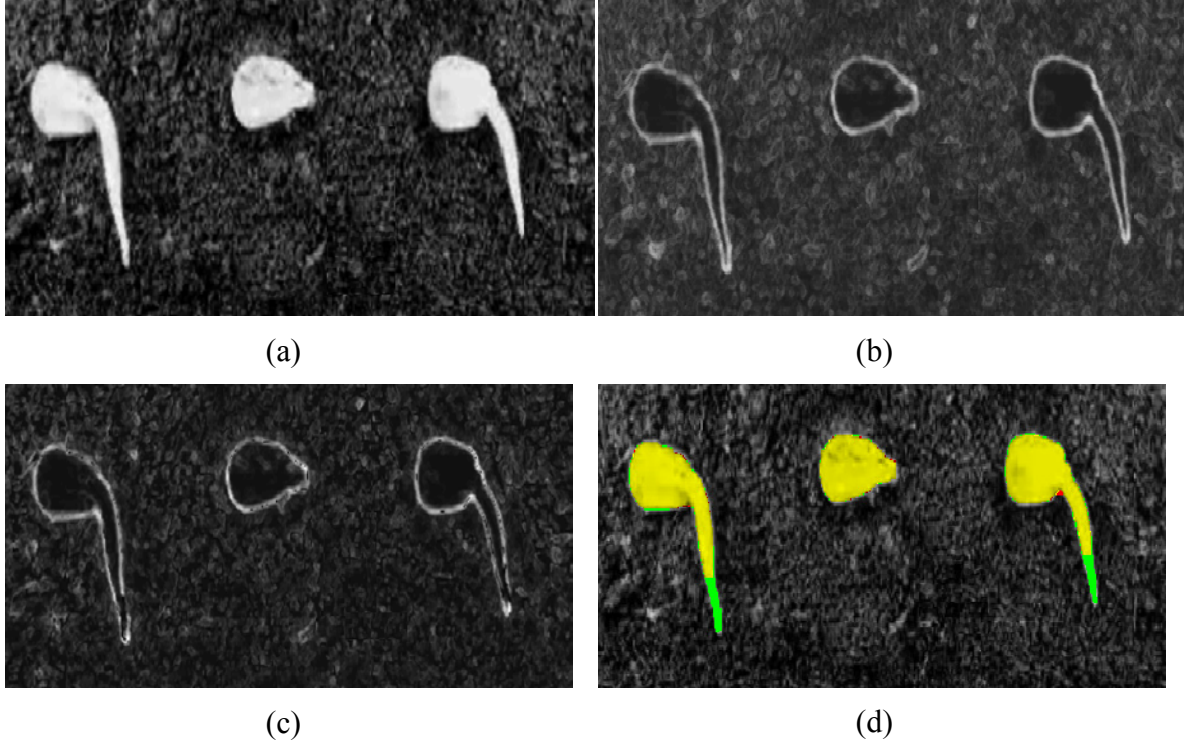
Figure 57: Graph cuts segmentation of textured natural image (from ImageNet dataset [32]) using the curvature-and-texture modulated weight. (a) Original image. (b) Curvature-modulated reliability calculated by our method with no texture gating. (c) Curvature-and-texture modulated reliability. (d) Comparison of segmentations from the proposed adaptive weight (*green*) to the least-error fixed weight (*red*) with overlapping regions in *yellow*. Reduced reliability in high-texture regions prevents these regions from inclusion into the segmentation result. Only the segmentation from the DD weights in *green* captures plant tips.

We present quantitative results on non-textured images from the ImageNet and McGill databases in Figure 58(a) and (b) where the error is the DSC averaged over 25 trials. In addition, we included results from a set of textured images in Figure 58(c). From the results, the data-driven approach is significantly more accurate when compared to the

118

fixed-weight approach for all datasets except for the textured dataset (see DD result in Figure 58(c)). However, the inclusion of the texture cue (DD-Text in Figure 58(c)) leads to significantly improved results.
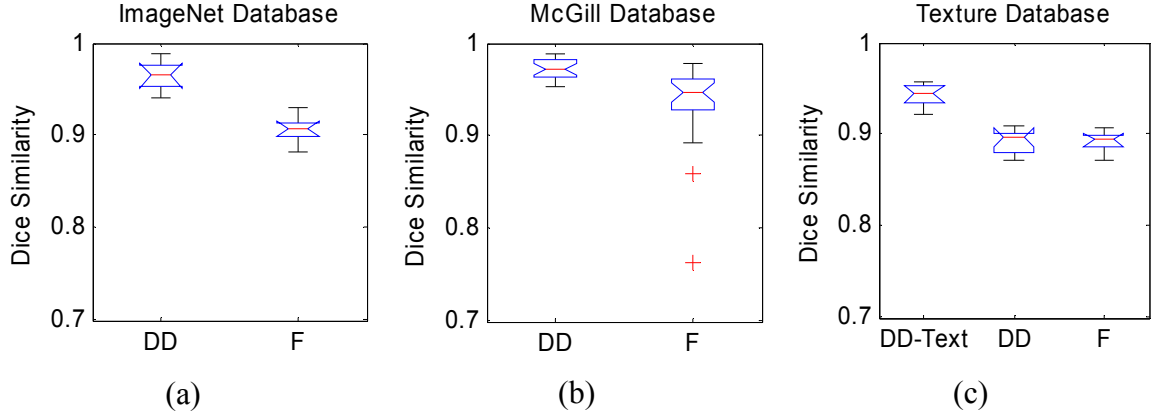
Figure 58: Segmentation DSC from using data-driven weights (DD), least-error fixed weights (F), and data-driven weights with the texture cue (DD-Text) on natural images from (a) the ImageNet database (8 images), (b) McGill dataset (24 images), and (c) selected textured images taken from ImageNet database (10 images). Average DSC over dataset (and over 25 segmentations per image) are as follows: DD = 0.9651, F = 0.9066 for (a), DD = 0.9727, F = 0.9236 for (b), and DD = 0.8916, F = 0.8922, DD-Text = 0.9421 for (c). All *p*-values $\ll 0.05$. Addition of texture cue in (c) resulted in improved performance over standard DD weights and fixed weights.

3.8.6.3 *Active Contours Without Edges*

We demonstrate the ACWE segmentation method with our regularization framework on the octopus image of Figure 59(a) (from the ImageNet database). Iterations were run until the contour evolution converged (1000 iterations). The low curvature-modulated reliability (Figure 59(b)) in regions outside the octopus prevents the resulting segmentation from including shadows in the background, unlike the fixed weight segmentation (Figure 59(c)) which leaked into the background of the image (see red region).

119

(a)   (b)

(c)
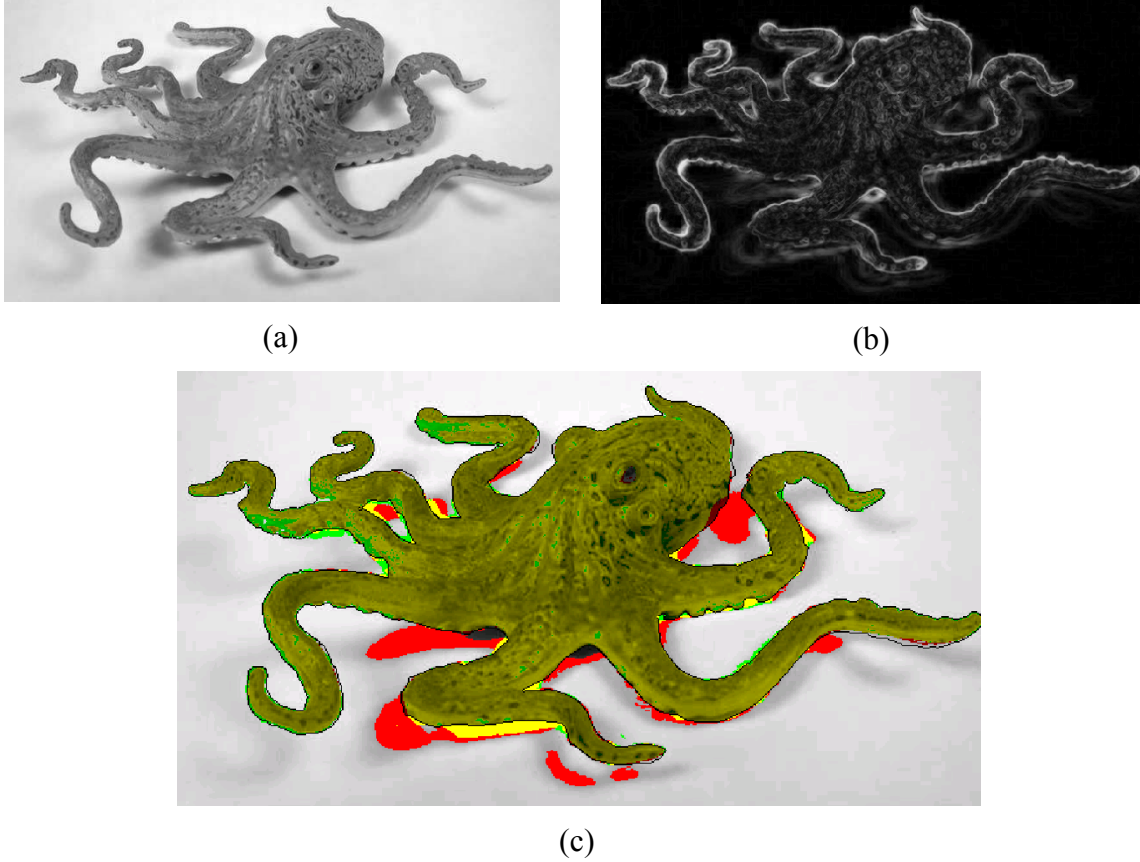
Figure 59: Active Contours segmentation of a natural image. (a) Original image. (b) Curvature-modulated reliability calculated by our method. (c) Comparison of segmentations from the data-driven weight (*green*) to the least-error fixed weight (*red*) where *yellow* regions are where segmentations overlap. Segmentation from fixed weight (*red*) leaks into the background, unlike segmentation from DD weights (see overlapped region in *yellow*).

Another such example from the McGill database is an image of a flower containing shadow regions as shown in Figure 60(a). The shadow inner folds of the flower have a low reliability measure as shown in Figure 60(b) due to the weak edge and low curvature. The resulting fixed-weight segmentation in Figure 60(c) has inadequate regularization in the weak edge shadow regions, whereas the data-driven weights produce a more accurate segmentation that regularizes in the shadows but reduces regularization in the high curvature region where the flower is obscured.

120

(a)   (b)   (c)
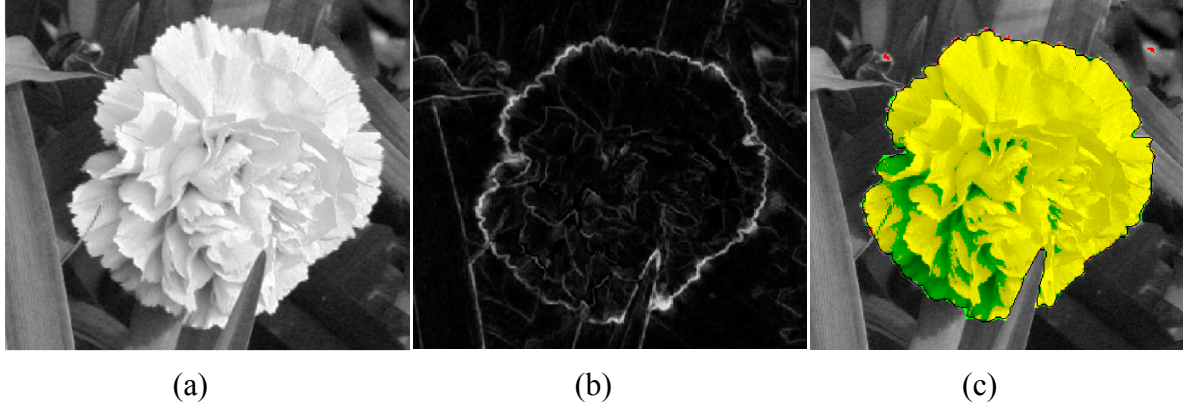
Figure 60: ACWE segmentation of flower. (a) Original image, (b) curvature-modulated reliability measure, and (c) comparison of segmentation from data-driven weights (*blue*) to the least-error fixed weight (*red*) where (*yellow*) regions represent segmentation overlap and the ground truth is shown as the *black* contour. Segmentation from fixed weight fails to capture shadow regions of flower, unlike segmentation from DD weights (see *green* region).

We present the quantitative results for the ImageNet and McGill databases in Figure 61. As seen in Figure 61(b), the ACWE method with data-driven weights did not produce significantly more accurate segmentations than the fixed weights on the McGill database (*p*-value was 0.0514). This is likely because for certain images in the dataset, both methods performed poorly due to non-optimal placement of the initial contour. The ACWE segmentation method is sensitive to initial contour placement which is not related to our proposed work.
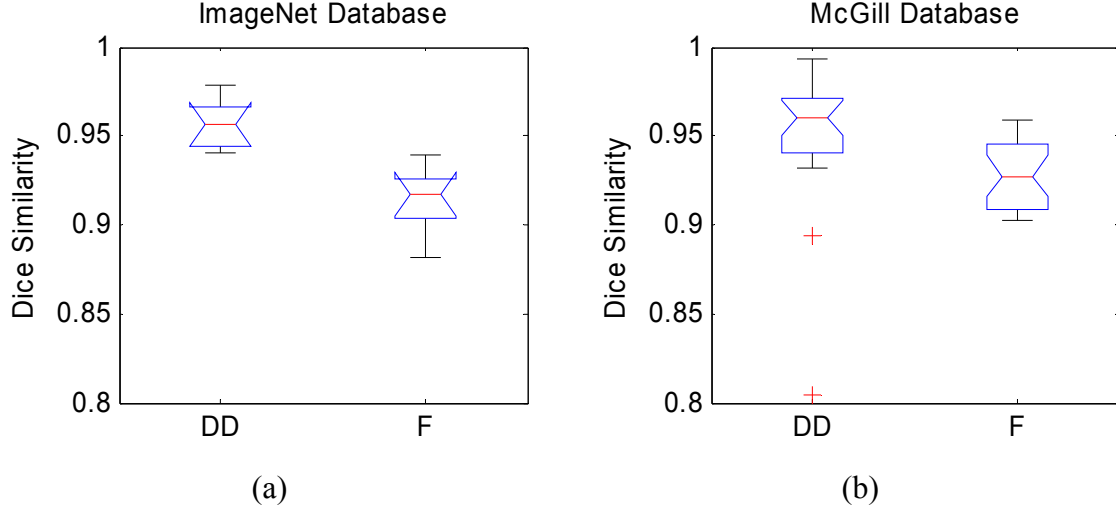
Figure 61: Segmentation DSC from using data-driven weights (DD) and least-error fixed weights (F) on natural images from (a) the ImageNet database (8 images) and (b) McGill dataset (24 images). Average DSC over dataset (and over 25 segmentations per image) are as follows: DD = 0. 9563, F = 0. 9144 for (a) and DD = 0. 9435, F = 0. 9044 for (b). Average p-value for (a) is 2E-4, but average p-value for (b) is 0.0514. Segmentations on the McGill database using the DD weights were not significantly more accurate due to non-optimal placement of the initial contour.

3.8.6.4   Mumford-Shah Method

For the contextual Mumford-Shah method, we tested textured natural images, such as the image of an amoeba shown in Figure 62 (from the ImageNet database). The curvature-modulated reliability is falsely high for the inner textured region of the amoebas, resulting in a disconnected segmentation as shown in Figure 62(b) in green. However, by addition of the texture term from the ET framework, the texture-and-curvature modulated reliability is smooth for these sections of the image, and the resulting segmentation is more accurate than the fixed weight and ET weights (Figure 62(c) and (d)). Although the segmentation produced by the ET method correctly regularizes the textured region, it is unable to segment the protrusions of the amoeba which is due to the framework lacking a structural curvature cue.
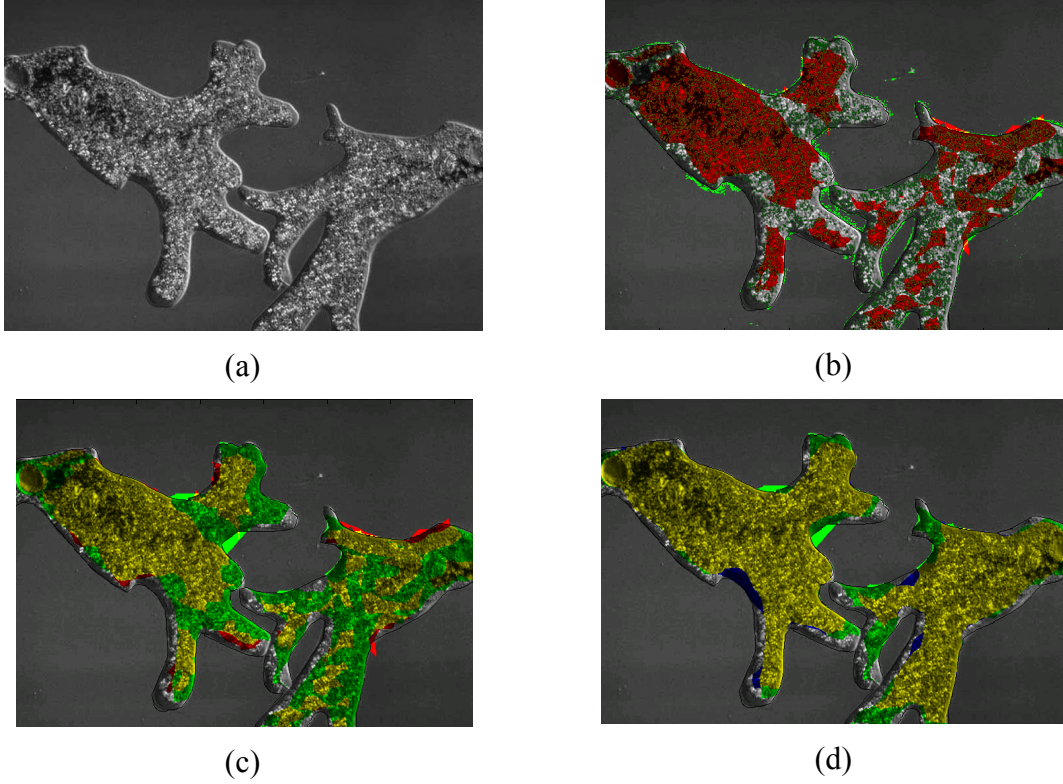
Figure 62: Segmentation of amoeba image from ImageNet. (a) Original image. (b) Comparison of segmentation from non-texture data-driven weight (*green*) to segmentation from fixed weight (*red*). (c) Comparisons of segmentation from texture data-driven weight (*green*) to fixed weight (*red*), and (d) to segmentation from ET cues (*blue*). Segmentation from the DD weights without texture mistakes textured regions as separate objects, resulting in a fragmented segmentation (see *green* segmentation in (b)). Segmentation from the DD weights *with* texture correctly captured the full objects, including protrusions, unlike the fixed weights (see overlapped *yellow* in (c)) and the ET weights (see overlapped *yellow* in (d)).

In addition, we segmented the texture image of a cheetah shown in Figure 63(a). The non-texture reliability in Figure 63(b) is falsely high in the spotted region of the cheetah. However, by incorporating the ET texture cue, the resulting texture-and-curvature modulated reliability is lower in the spotted region (Figure 63(c)). The resulting segmentation from the texture data-driven weight in green is more accurate than the segmentations from the fixed weight in red (Figure 63(d)) and the ET weight in blue (Figure 63(e)).

(a)

(b)

(c)

(d)

(e)
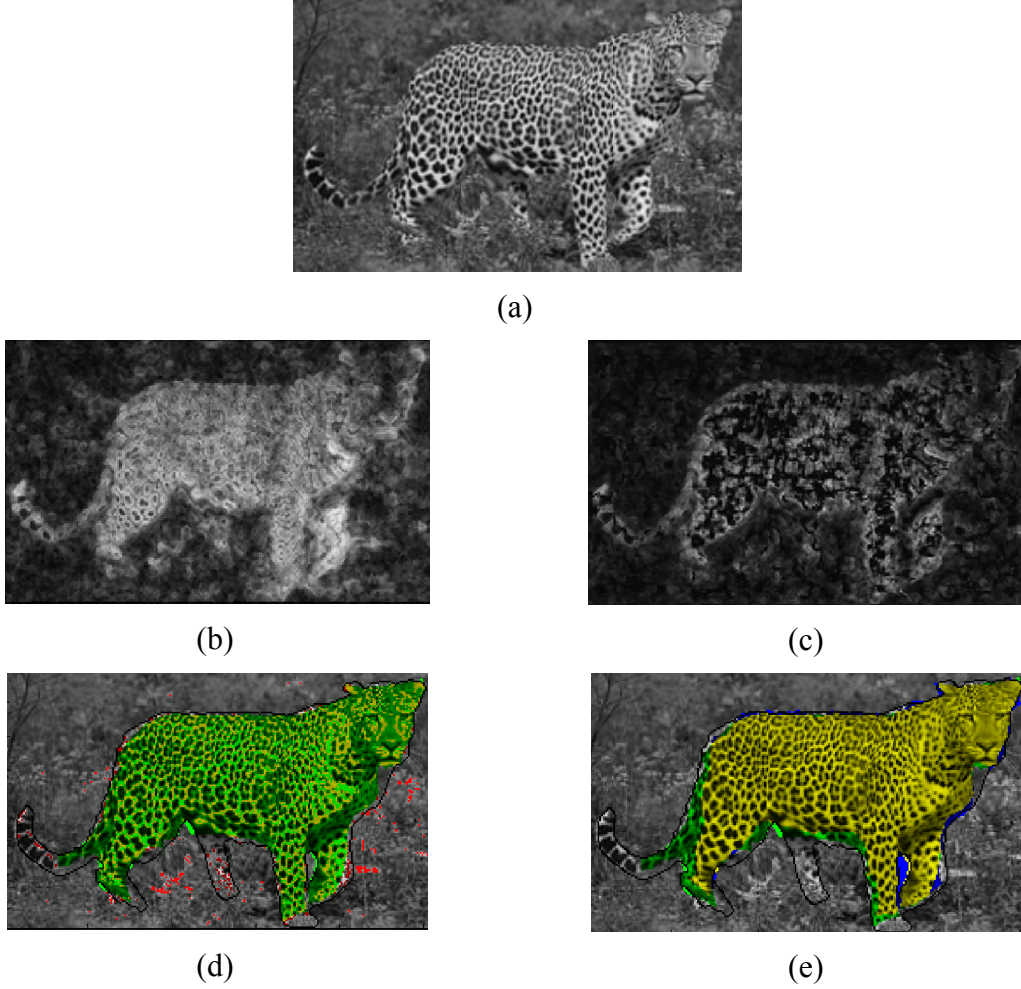
Figure 63: Segmentation of cheetah image presented in [38]. (a) Original image. (b) Curvature-modulated reliability. (c) Texture-and-curvature modulated reliability from incorporating texture cue into our framework. (d) Comparison of segmentations results from data-driven weight (*green*), fixed weight (*red*), and (e) ET weight (*blue*) where *yellow* regions represent overlap. Reliability without texture in (b) inaccurately detects textured regions as high curvature regions. This undesirable behavior is removed in (c) with the addition of the texture cue. The resulting segmentation (*green*) more accurately captures the cheetah than the fixed weights (*red*) and the ET weights (*blue*) which result in leakage.

The quantitative results are presented in Figure 64 where we measure the error as the DSC between the binary mask created from the edge-process produced by the MS method and the ground truth segmentation. We find that the data-driven weights produce more accurate segmentations for the non-textured dataset (Figure 64(a)) but produce less accurate segmentations for the textured set (DD in Figure 64(b)). However, the addition

124

of the textured cue from the ET framework results in more accurate segmentation than

the non-curvature-modulated ET framework (DD-Text in Figure 64(b)).

(a)   (b)
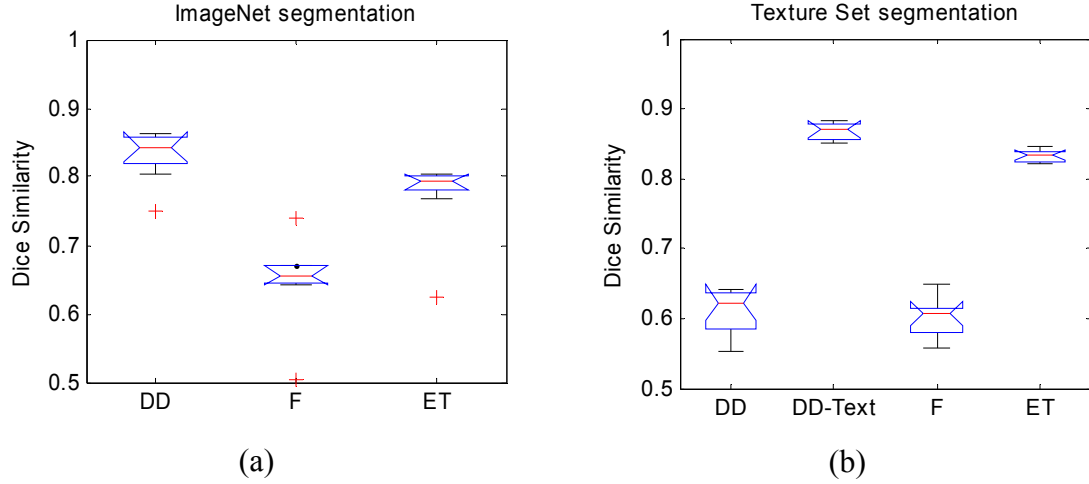
Figure 64: Segmentation DSC from using data-driven weights (DD), least-error fixed weights (F), data-driven weights with the texture cue (DD-Text), and Erdem-Tari weights (ET) on natural images from (a) the ImageNet database (8 non-textured images), and (b) selected textured images taken from ImageNet database (10 images). Average DSC over dataset (and over 25 segmentations per image) are as follows: DD = 0.8311, F = 0.6482, and ET = 0.7719 for (a), and DD = 0.6107, DD-Text = 0.8681, F = 0.6004, and ET = 0.8323 for (b). Addition of texture cue in (b) resulted in improved performance over standard DD weights, fixed weights, and ET weights. All *p*-values of DD versus fixed and ET methods in (a) and DD-text versus fixed method and ET method in (b) were $\ll 0.05$. Segmentations from the DD weights on the non-textured ImageNet dataset were significantly more accurate than the fixed weight and ET weights. Addition of the texture cue resulted in significantly more accurate performance for the textured set in (b).

Chapter 4

Conclusions

In this chapter, we discuss conclusions for this thesis and how our work fits into the general field of image segmentation. We discuss the limitation of this method and present an overview of the areas for future research in this work.

4.1 Discussion

In this thesis, we presented novel approaches for addressing a ubiquitous problem that plagues most energy minimization-based segmentation techniques; how to properly balance the weights of competing data fidelity and regularization energy terms. We focused on automated methods to determine a single regularization weight for convex energy functionals. We first presented a method for determining the globally optimized values of the object function parameters. During validation, we found this method to suffer from bimodal weights, poor response to image characteristics, and decreased

generality to other segmentation methods. We then proposed spatially adaptive weights that depended on contextual cues that gauge image reliability and structural evidence. This method employed a novel and robust local measure of signal reliability through estimating the spectral flatness of the image. In addition, our method used a local scale-invariant curvature cue for modulating regularization in conjunction with edge evidence, where both cues were made robust to noise through gating by the local signal reliability. We demonstrated the applicability of our contextual weights by incorporating the weights into a variety of continuous and discrete segmentation frameworks, including minimum-path, graph cuts, and two forms of the Mumford-Shah model: active contours without edges, and the AT approximation of the Mumford-Shah model with feedback regularization. We validated the contextual weights on a wide variety of datasets, the majority of which are publically available: functionally-derived images, DDSM mammography dataset, 52-subject sagittal slices of CC structure, 3-modality BrainWeb data with varying noise and intensity inhomogeneity, 18-subject IBSR dataset, images from the McGill Colour Calibration database, PASCAL object recognition database, and the ImageNet database. We compared the contextual weights against the least-error spatially fixed weight, the globally optimal weights, and the spatially adaptive weights proposed by Erdem and Tari [38]. In addition, we incorporated existing texture-based contextual cues into our method to show how it can be incorporated with current techniques in the field.

4.2 Limitations and Future Directions

Although we have demonstrated the significant improvements in segmentation accuracy our contextual weights provide, there are limitations to our methods and areas for future research, which we list here:

Energy Functionals with Multiple Weights

In our work, we have focused on using a single regularization weight and adapting this weight accordingly. However, many energy minimization segmentation techniques use multiple weights, i.e. as follows:

$$E_{total} = \alpha_1 E_1 + \alpha_2 E_2 + \alpha_3 E_3 + \cdots \tag{96}$$

Our work is limited to single weight energy formulations, which we form by grouping the terms into either E_{int} or E_{ext} and using a single weight in a convex combination for graph cuts and minimum-path approaches, and non-convex for ACWE and MS. However, single weights are limited since many segmentation techniques give greater importance to certain energy terms E_i through assigning that term a higher α_i. Determining how to best set multiple regularization weights will greatly increase the generality and usability of this method.

Additional Cues for Non-Spectrally Flat Noise Models

Our method for noise estimation is limited to types of noise that are spectrally flat, i.e. white noise. However, many images may be corrupted by other noise models, particularly in medical imaging. For example, ultrasound images are often corrupted by

multiplicative noise (speckle noise) which our method is not able to estimate. A future area of research would be to devise new noise cues to handle various noise models and combine them with the existing noise cue.

Expansion to 3D

Our current method is limited to 2D images. However, in medical imaging, 3D data is far more common and is segmented by various 3D energy minimization techniques, predominately deformable model methods. Thus our contextual regularization method must be modified by (a) expanding the spectral flatness, edge evidence, and curvature cues to the 3D space, and (b) incorporating these cues into 3D segmentation frameworks.

Convex Functionals for Continuous Segmentation Methods

In our tests with continuous segmentation frameworks, such as ACWE and the MS model, we had difficulties using a convex energy functional with respect to the adaptive weight as we had done for the discrete frameworks; instead, we only weighted the regularization term. We found that weighting the external terms resulted in the stalling of the curve evolution. However, we were unable to determine exactly why that occurred. Further research would be beneficial in this area so that our method could be incorporated into segmentation frameworks in a more unified manner.

Investigating Globally Optimal Weights for Additional Segmentation Methods

We presented a method for determining the globally-optimal weight for the minimum-path segmentation framework, and showed that it did not provide accurate

results. However, we did not determine the globally-optimal weights for additional segmentation frameworks, such as continuous methods. A key area of future research is to determine if the globally optimum weight for other segmentation frameworks are also insufficient in addressing regularization needs. This would be important in proving that the globally optimal weights do not necessarily reflect correct segmentations.

Further Validation with Additional Segmentation Frameworks and Comparison Methods

We presented results from validation with four segmentation techniques, and comparisons against two existing methods (spatially-fixed, and the ET weights). Future work should focus on exploring the use of our weights with additional segmentation frameworks, such as the robust higher order potentials method by Kohli and Ladicky [58], normalized cuts [69], and shape priors and discrete MRFs by Besbes *et al* [10]. In addition, future work should focus on expanding the number of methods we compared our work against. The contextual texture-based regularization method of Malik *et al* [69] using Normalized Cuts should be validated, along with the texture-based regularization of Kokkinos *et al* [59] using segmentation by weighted curve evolution. In addition, a key family of segmentation approaches which our existing validation work does not cover are clustering techniques, such as K-means clustering [30, 95, 98], fuzzy c-means clustering [11, 25], spectral clustering (i.e., Normalized Cuts [69]), and expectation-maximization (EM) clustering [66].

Additional Contextual Cues

Incorporation of additional contextual cues would benefit our work, particularly in the area of texture estimation, additional noise estimation methods, and shape information. These additional cues can be incorporated into our method through multiplication with the noise-gated edge evidence as we have done with the Erdem and Tari texture cue in Section 3.6.2. In particular, the structural curvature cue would benefit from use of a structure tensor matrix where the eigen-decomposition of this matrix may capture more precise gradient features.

Bibliography

[1]   Propeller MRI. *Southern MRI Research Institute.* [Online]. Available:

        http://southernmri.org/?q=node/13. Accessed: May 1, 2010.

[2]   A. Akselrod-Ballin, M. Galun, M. John Gomori, A. Brandt, and R. Basri. Prior

        knowledge driven multiscale segmentation of brain MRI. In *MICCAI (2)*, volume

        4792 of *Lecture Notes in Computer Science*, pages 118–126. Springer, 2007.

[3]   L. Ambrosio and V.M. Tortorelli. Approximation of functional depending on

        jumps by elliptic functional via t-convergence. *Communications on Pure and*

        Applied Mathematics, 43(8):999–1036, 2006.

[4]   A. Amini, T. Weymouth, and R. Jain. Using dynamic programming for solving

        variational problems in vision. *IEEE Transactions on pattern analysis and*

        machine intelligence, pages 855–867, 1990.

[5]   C. Aslan and S. Tari. An axis-based representation for recognition. In

        International Conference on Computer Vision, pages II: 1339–1346, 2005.

[6]   M. Awrangjeb and G. Lu. Robust image corner detection based on the chord-to-

        point distance accumulation technique. *IEEE Transactions on Multimedia*,

        10(6):1059–1072, October 2008.

[7]   S. Bagon. MATLAB wrapper for graph cuts.

        http://www.wisdom.weizmann.ac.il/bagon, 2006.

[8]   L. Bar, N. Kiryati, and N. Sochen. Image deblurring in the presence of impulsive

        noise. *International Journal of Computer Vision*, 70(3):279–298, 2006.

[9]   W. A. Barrett and E. N. Mortensen. Interactive live-wire boundary extraction.

        Medical Image Analysis, 1:331–341, 1997.

[10]   A. Besbes, N. Komodakis, G. Langs, and N. Paragios. Shape priors and discrete MRFs for knowledge-based segmentation. In *Computer Vision and Pattern Recognition*, pages 1295-1302, 2009.

[11]   J. C. Bezdek, L. O. Hall, and L. P. Clarke. Review of MR image segmentation techniques using pattern recognition. *Medical Physics*, 20(4):1033–1048, 1993.

[12]   A. Blake and A. Zisserman. *Visual Reconstruction*. MIT Press, Cambridge, MA, 1987.

[13]   Y. Boykov and M.P. Jolly. Interactive graph cuts for optimal boundary and region segmentation of objects in ND images. In *International Conference on Computer Vision*, volume 1, pages 105–112. Citeseer, 2001.

[14]   Y. Boykov and O. Veksler. Graph cuts in vision and graphics: Theories and applications. *The Handbook of Mathematical Models in Computer Vision*. Springer, 2006.

[15]   Y. Boykov and G. Funka-Lea. Graph cuts and efficient N-D image segmentation. *International Journal of Computer Vision*, 70(2):109–131, 2006.

[16]   Y. Boykov and V. Kolmogorov. An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision. *IEEE transactions on Pattern Analysis and Machine Intelligence*, 26(9):1124–1137, 2004.

[17]   Y. Boykov, O. Veksler, and R. Zabih. Fast approximate energy minimization via graph cuts. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(12):1222–1239, 2001.

[18]   M. Bydder, D.J. Larkman, and J.V. Hajnal. Simultaneous acquisition of spatial harmonics (SMASH): ultra-fast imaging with radiofrequency coil arrays. *Magnetic Resonance in Medicine*, 38(1):591–603, 1997.

[19]   M. Bydder, D.J. Larkman, and J.V. Hajnal. Generalized SMASH imaging. *Magnetic Resonance in Medicine*, 47(1):160–170, 2002.

[20]   J. Canny. A computational approach to edge detection. In *RCV87*, pages 184–203, 1987.

[21]   V. Caselles, R. Kimmel, and G. Sapiro. Geodesic active contours. *International Journal of Computer Vision*, 22(1):61–79, 1997.

[22]   T. F. Chan, B. Y. Sandberg, and L. A. Vese. Active contours without edges for vector-valued images. *Journal of Visual Communication and Image Representation*, 11(2):130–141, 2000.

[23]   T. F. Chan and L. A. Vese. An active contour model without edges. In *Scale Space*, pages 141–151, 1999.

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

[25]   K. S. Chuang, H. L. Tzeng, S. Chen, J. Wu, and T. J. Chen. Fuzzy c-means clustering with spatial information for image segmentation. *Computerized Medical Imaging and Graphics*, 30(1):9–15, 2006.

[26]   C. A. Cocosco, V. Kollokian, R. Kwan, and A. Evans. BrainWeb: Online interface to a 3D MRI simulated brain database. In *NeuroImage*, volume 5. Academic Press, 1997.

[27]   I. Cohen, N. Ayache, and P. Sulger. Tracking points on deformable objects using curvature information. In *ECCV*, pages 458–466, 1992.

[28]   L. D. Cohen. On active contour models and balloons. *CVGIP: Image Understanding*, 53(2):211–218, 1991.

[29]   L. D. Cohen and R. Kimmel. Global minimum for active contour models: A minimal path approach. *International Journal of Computer Vision*, 24(1):57–78, 1997.

[30]   G. B. Coleman and H. C. Andrews. Image segmentation by clustering. *Proceedings of the IEEE*, 67(5):773–785, 1979.

[31]   D.L. Collins, A.P. Zijdenbos, V. Kollokian, J.G. Sled, N.J. Kabani, C.J. Holmes, and A.C. Evans. Design and construction of a realistic digital brain phantom. *IEEE Transactions on Medical Imaging*, 17(3):463–468, 1998.

[32]   J. Deng, W. Dong, R. Socher, L.J. Li, K. Li, and L. Fei-Fei. ImageNet: A Large-Scale Hierarchical Image Database. In *CVPR*, 2009.

[33]   J. Deng, K. Li, M. Do, H. Su, and L. Fei-Fei. Construction and Analysis of a Large Scale Image Ontology. Vision Sciences Society, 2009.

[34]   E. Dijkstra. A note on two problems in connection with graphs. *Numerical Mathematics*, 1(5):269–271, 1959.

[35]   B. Dong, A. Chien, Y. Mao, J. Ye, and S. Osher. Level set based surface capturing in 3D medical images. In *MICCAI (1)*, volume 5241 of *Lecture Notes in Computer Science*, pages 162–169. Springer, 2008.

[36]   M. Donias, P. Baylou, and N. Keskes. Curvature of oriented patterns: 2-D and 3-D estimation from differential geometry. In *ICIP*, pages I: 236–240, 1998.

[37] E. Erdem, A. Erdem, and S. Tari. Edge strength functions as shape priors in image segmentation. In *EMMCVPR*, pages 490–502, 2005.

[38] E. Erdem and S. Tari. Mumford-Shah regularizer with contextual feedback. *Journal of Mathematical Imaging and Vision*, 33(1):67–84, 2009.

[39] S. Esedoglu and J. H. Shen. Digital inpainting based on the Mumford-Shah-Euler image model. *European Journal of Applied Mathematics*, 13:353–370, 2002.

[40] M. Everingham, L. Van Gool, C. K. I. Williams, J. Winn, and A. Zisserman. The PASCAL Visual Object Classes Challenge 2009 (VOC2009) Results. http://www.pascal-network.org/challenges/VOC/voc2009/workshop/index.html.

[41] A.X. Falcão, J.K. Udupa, S. Samarasekera, S. Sharma, B.E. Hirsch, and R.A. Lotufo. User-Steered Image Segmentation Paradigms: Live Wire and Live Lane. *Graphical Models and Image Processing*, 60(4):233–260, 1998.

[42] O. Faugeras. *Three-Dimensional Computer Vision: A Geometric Viewpoint*. MIT Press, Cambridge, Massachusetts, 1993.

[43] D. Geiger, A. Gupta, LA Costa, and J. Vlontzos. Dynamic programming for detecting, tracking, and matching deformable contours. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 17(3):294–302, 1995.

[44] S. Geman and D. Geman. Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 6(6):721–741, 1984.

[45] G. Gilboa, J. Darbon, S. Osher, and T. Chan. Nonlocal convex functionals for image regularization. *UCLA CAM-report*, pages 06–57, 2006.

[46] R. C. Gonzales and R. Woods. *Digital Image Processing*. Addison Wesley, 1992.

[47]   L. Grady. Random walks for image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28(11):1768–1783, 2006.

[48]   D. Greig, B. Porteous, and A. Seheult. Exact maximum a posterori estimation for binary images. *Journal Royal Statistical Society*, B: 51(2):271–279, 1989.

[49]   X. C. He and N. H. C. Yung. Curvature scale space corner detector with adaptive threshold and dynamic region of support. In *ICPR*, pages II: 791–794, 2004.

[50]   M. Heath, K. Bowyer, D. Kopans, P. Kegelmeyer, R. Moore, K. Chang, and S. Munishkumaran. Current status of the digital database for screening mammography. *Computational Imaging and Vision*, 13:457–460, 1998.

[51]   M. Heath, K. Bowyer, D. Kopans, R. Moore, and P. Kegelmeyer. The digital database for screening mammography. In *Proceedings of the 5th International Workshop on Digital Mammography*, pages 212–218, 2000.

[52]   S. Hermann and R. Klette. A comparative study on 2D curvature estimators. In *ICCTA*, pages 584–589. IEEE Computer Society, 2007.

[53]   H. Ishikawa and D. Geiger. Occlusions, discontinuities, and epipolar lines in stereo. In *ECCV*, page I: 232, 1998.

[54]   N. S. Jayant and Peter Noll. *Digital Coding of Waveforms*. Prentice-Hall, Englewood Cliffs, 1984.

[55]   I.H. Jermyn and H. Ishikawa. Globally optimal regions and boundaries. In *Computer Vision, 1999. The Proceedings of the Seventh IEEE International Conference on*, volume 2, 1999.

[56]   M. Kass, A. Witkin, and D. Terzopoulos. Snakes: Active contour models. *International Journal of Computer Vision,* 1(4):321–331, 1988.

[57]   L. Kitchen and A. Rosenfeld. Gray-level corner detection. *Pattern Recognition Letters*, 1(2):95 – 102, 1982.

[58]   P. Kohli, L. Ladicky, and P.H.S. Torr. Robust higher order potentials for enforcing label consistency. *International Journal of Computer Vision*, 82(3):302–324, 2009.

[59]   I. Kokkinos, G. Evangelopoulos, and P. Maragos. Texture analysis and segmentation using modulation features, generative models, and weighted curve evolution. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 31(1):142–157, 2009.

[60]   V. Kolmogorov, Y. Boykov, and C. Rother. Applications of parametric maxflow in computer vision. *ICCV*, 8, 2007.

[61]   H.W. Kuhn. The Hungarian method of solving the assignment problem. *Naval Res. Logistics Quart.*, 2:83–97, 1955.

[62]   R. Kwan, A. Evans, and G. Pike. An extensible MRI simulator for post-processing evaluation. In *Visualization in Biomedical Computing*, pages 135–140. Springer, 1996.

[63]   R. Kwan, A. Evans, and G. Pike. MRI simulation-based evaluation of image-processing and classification methods. *IEEE Transactions on Medical Imaging*, 18(11):1085–1097, 1999.

[64]   H. Li and A. Yezzi. Vessels as 4-D curves: Global minimal 4-D paths to extract 3-D tubular surfaces and centerlines. *IEEE Transactions on Medical Imaging*, 26(9):1213–1223, 2007.

[65] J. Liang, T. McInerney, and D. Terzopoulos. United snakes. *Medical Image Analysis*, 10(2):215–233, 2006.

[66] Z. Liang, R.J. Jaszczak, and R.E. Coleman. Parameter estimation of finite mixtures using the EM algorithm and information criteria with application to medical image processing. *IEEE Transactions on Nuclear Science*, 39(4):1126–1133, 1992.

[67] T. Lindeberg. On scale selection for differential operators. In *Scan. Conf. on Image Analysis*, pages 857–866, 1993.

[68] T. Lindeberg. Edge detection and ridge detection with automatic scale selection. *International Journal of Computer Vision*, 30(2):117–154, 1998.

[69] J. Malik, S. Belongie, T. K. Leung, and J. Shi. Contour and texture analysis for image segmentation. *International Journal of Computer Vision*, 43(1):7–27, 2001.

[70] P. Martin, P. Réfrégier, F. Goudail, and F. Guérault. Influence of the noise model on level set active contour segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26(6):799–803, 2004.

[71] T. McInerney and D. Terzopoulos. Deformable models in medical image analysis: A survey. *Medical Image Analysis*, 1(2):91–108, 1996.

[72] C. McIntosh and G. Hamarneh. Is a single energy functional sufficient? Adaptive energy functionals and automatic initialization. In *MICCAI (2)*, volume 4792 of *Lecture Notes in Computer Science*, pages 503–510. Springer, 2007.

[73] F. Mokhtarian and R. Suomela. Robust image corner detection through curvature scale space. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(12):1376–1381, 1998.

[74]   D. Mumford and J. Shah. Optimal approximations by piecewise smooth functions and variational problems. *Comm. on Pure and Applied Math.*, XLII(5):577–685, 1988.

[75]   A. Olmos and F. Kingdom. McGill calibrated colour image database. http://tabby.vision.mcgill.ca. 2004.

[76]   S. J. Osher and R. Fedkiw. *Level Set Methods and Dynamic Implicit Surfaces*. Springer, 2002.

[77]   S. J. Osher and N. Paragios. *Geometric Level Set Methods in Imaging, Vision, and Graphics*. Springer-Verlag, 2003.

[78]   P. Perona and J. Malik. Scale space and edge detection using anisotropic diffusion. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12(7):629–639, 1990.

[79]   D.L. Pham, C. Xu, and J.L. Prince. Current methods in medical image segmentation. *Annual Review of Biomedical Engineering*, 2(1):315–337, 2000.

[80]   C. Pluempitiwiriyawej, J. M. F. Moura, Y. Lin Wu, and C. Ho. STACS: New active contour scheme for cardiac MR image segmentation. *IEEE Transactions on Medical Imaging*, 24(5):593–603, 2005.

[81]   M. Poon, G. Hamarneh, and R. Abugharbieh. Live-vessel: Extending livewire for simultaneous extraction of optimal medial and boundary paths in vascular images. In *MICCAI (2)*, volume 4792 of *Lecture Notes in Computer Science*, pages 444–451. Springer, 2007.

[82] K. P. Pruessmann, M. Weiger, M. B. Scheidegger, and P. Boesiger. SENSE: sensitivity encoding for fast MRI. *Magnetic Resonance in Medicine*, 42(5):952–962, 1999.

[83] S. Roy and I. J. Cox. A maximum-flow formulation of the N-camera stereo correspondence problem. In *ICCV*, pages 492–502, 1998.

[84] A. A. Samsonov and C. R. Johnson. Noise-adaptive nonlinear diffusion filtering of MR images with spatially varying noise levels. *Magnetic Resonance in Medicine*, 52(4):798–806, 2004.

[85] G. Sapiro. *Geometric partial differential equations and image analysis*. Cambridge Univ Pr, 2001.

[86] J. A. Sethian. *Level Set Methods and Fast Marching Methods*. Cambridge University Press, 1999.

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

[88] D. W. Shattuck, S. R. Sandor-Leahy, K. A. Schaper, D. A. Rottenberg, and R. M. Leahy. Magnetic resonance image tissue classification using a partial volume model. *NeuroImage*, 13(5):856–876, 2001.

[89] J. G. Sled, A. P. Zijdenbos, and A. C. Evans. A nonparametric method for automatic correction of intensity nonuniformity in MRI data. *IEEE Transactions on Medical Imaging*, 17(1):87–97, 1998.

[90] M. Styner, C. Brechbuhler, G. Szckely, and G. Gerig. Parametric estimate of intensity inhomogeneities applied to MRI. *IEEE Transactions on Medical Imaging*, 19(3):153–165, 2000.

[91]   M. Szummer, P. Kohli, and D. Hoiem. Learning CRFs using Graph Cuts. In *ECCV*, pages 582–595. Springer, 2008.

[92]   Z. S. G. Tari, J. Shah, and H. Pien. Extraction of shape skeletons from grayscale images. *Computer Vision and Image Understanding*, 66(2):133–146, 1997.

[93]   D. S. Taubman and M. W. Marcellin. *JPEG 2000: Image Compression Fundamentals, Standards and Practice*. Kluwer Academic Publishers, Norwell, MA, USA, 2001.

[94]   S. Teboul, L. Blanc Feraud, G. Aubert, and M. Barlaud. Variational approach for edge-preserving regularization using coupled PDEs. *IEEE Transactions on Image Processing*, 7(3):387–397, 1998.

[95]   J. Theiler and G. Gisler. A contiguity-enhanced k-means clustering algorithm for unsupervised multispectral image segmentation. In *Proc. SPIE*, volume 3159, pages 108–118. Citeseer, 1997.

[96]   J. K. Udupa and G. J. Grevera. Go digital, go fuzzy. *Pattern Recognition Letters*, 23(6):743–754, 2002.

[97]   L. A. Vese and T. F. Chan. A multiphase level set framework for image segmentation using the Mumford and Shah model. *International Journal of Computer Vision*, 50(3):271–293, 2002.

[98]   K. Wagstaff, C. Cardie, S. Rogers, and S. Schroedl. Constrained k-means clustering with background knowledge. In *Proc. 18th International Conf. on Machine Learning*, pages 577–584. Morgan Kaufmann, San Francisco, CA, 2001.

[99] I. Wendelhag, Q. Liang, T. Gustavsson, and J. Wikstrand. A new automated computerized analyzing system simplifies readings and reduces the variability in ultrasound measurement of intima-media thickness. *Stroke*, 28(11):2195, 1997.

[100] O. Wink, WJ Niessen, and MA Viergever. Multiscale vessel tracking. *IEEE Transactions on Medical Imaging*, 23(1):130–133, 2004.

[101] L. Wolf, X. Huang, I. Martin, and D. Metaxas. Patch-based texture edges and segmentation. In *ECCV*, pages 481–493, 2006.

[102] Y. Wu. Chan Vese active contours without edges. http://www.mathworks.com/matlabcentral/fileexchange/23445. 2009.

[103] X. H. Zhang, M. Lei, D. Yang, Y. Wang, and L. T. Ma. Multi-scale curvature product for robust image corner detection in curvature scale space. *Pattern Recognition Letters*, 28(5):545–554, 2007.

[104] X. H. Zhang, H. Wang, M. Hong, L. Xu, D. Yang, and B. C. Lovell. Robust image corner detection based on scale evolution difference of planar curves. *Pattern Recognition Letters*, 30(4):449–455, 2009.

[105] P. Zhao and B. Yu. Boosted lasso. *Journal of Machine Learning Research, to appear*, 2004.

[106] Peng Zhao and Bin Yu. Stagewise lasso. *Journal of Machine Learning Research*, 8:2701–2726, 2007.

[107] B. J. Zhong and W. H. Liao. Direct curvature scale space: Theory and corner detection. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 29(3):508–512, 2007.

[108] S.C. Zhu and A. Yuille. Region competition: Unifying snakes, region growing, and Bayes/MDL for multi-band image segmentation. In *ICCV*, pages 416-423, 1995.