Learning from the Expert: Improving Boundary Definitions in Biomedical Imagery

Stewart Crawford-Hines

Colorado State University, Fort Collins, Colorado USA
sgcraw@cs.colostate.edu

Defining the boundaries of regions of interest in biomedical imagery has remained a difficult real-world problem in image processing. Experience with fully automated techniques has shown that it is usually quicker to manually delineate a boundary rather than correct the errors of the automation. Semi-automated, user-guided techniques such as Intelligent Scissors and Active Contour Models have proven more promising, since an expert guides the process. This paper will report and compare some recent results of another user-guided system, the Expert's Tracing Assistant, a system which learns a boundary definition from an expert, and then assists in the boundary tracing task. The learned boundary definition better reproduces expert behavior, since it does not rely on the *a priori* edge-definition assumptions of the other models.

1. Background

The system discussed in this paper provides a computer-aided assist to human experts in boundary tracing tasks, through the application of machine learning techniques to the definition of structural image boundaries. Large imagery sets will usually have a repetition and redundancy on which machine learning techniques can capitalize. A small subset of the imagery can be processed by a human expert. This base can then be used by a system to learn the expert's behavior, whereafter the system can then semi-automate the task.

The biomedical domain is a rich source of large, repetitive image sets. For example, in a computed tomographic (CT) scan, cross-sectional images are generated in parallel planes typically separated by millimeters. At a 2mm separation between image planes, approximately 75 images would be generated in imaging the complete brain. Images such as this, generated along parallel planes, are called sectional imagery. Such sectional imagery abounds in medical practice: X-ray, MRI, PET, confocal imagery, electron microscopy, ultrasound, and cryosection (freezing and slicing) technologies all produce series of parallel-plane 2D images.

Generating a three dimensional polygonal model of a structure from sectional imagery requires bounding the structure across the whole image set. Currently, the reference standard for high-quality outlining tasks is an expert's delineation of the region. The state-of-the-practice is that this is done manually, which is a repetitive, tedious, and error-prone process for a human.

There has been much research directed toward the automatic edge detection and segmentation of images, from which the extraction of a boundary outline could then proceed. Systems based on this work have run into two significant problems: (1) the cost of user adjustments when the system runs into troublesome areas often exceeds the cost of manually tracing the structure from the start, and (2) the a priori assumptions implicit in these methods impact the ability of the system to match the expert’s performance on boundary definition tasks where expert judgement is called into play.

The system discussed herein, the Expert’s Tracing Assistant (ETA), provides viable assistance for such tracing tasks which has proven beneficial in several regards:

- the specialist’s time can be significantly reduced;
- errors brought on by the tedium of tracing similar boundaries over scores of similar images can be reduced; and
- the automated tracing is not subject to human variability and is thus reproducible and more consistent across images.

The thrust of this work is to learn the boundary definitions of structures in imagery, as defined by an expert, and to then assist the expert when they need to define such boundaries in large image sets. The basic methodology of this Expert’s Tracing Assistant is:

(1) in one image, an expert traces a boundary for the region of interest;

(2) that trace is used in the supervised training of a neural network;

(3) the trained network replicates the expert’s trace on similar images;

(4) the expert overrides the learned component when it goes astray.

The details of the neural network architecture and training regimen for this system have been previously documented by Crawford-Hines & Anderson [1]. This paper compares these learned boundaries to the semi-automated boundary definition methods of Intelligent Scissors by Barrett & Mortensen [2] , and Active Contour Models begun by Kass, Witkin, & Terzopoulos [3].

2. Methods for Boundary Tracing

To understand the relative merits of learning boundary contours, the Expert’s Tracing Assistant (ETA) was studied in comparison to other user-guided methods representing the current best state-of-the-practice for boundary delineation. The techniques of Active Contour Models (ACM) and Intelligent Scissors (IS) were chosen for comparison to ETA because of they have been brought into practice, they have been studied and refined in the literature, and they represent benchmarks against which other novel methods are being compared. The ground truth for the comparison of these methods is an expert's manual tracing of a structure's boundary in an image.

The structures chosen for comparison were taken from the imagery of the Visible Human dataset, from the National Library of Medicine [4], which is also becoming a benchmark set upon which many image processing and visualization methods are being exercised. Several structures were selected as representative cross-sections. For each, the IS, ACM, and ETA methods were used to define its boundary and an expert manually delineated the boundary in two independent trials.

Figure 1 shows three structures to be compared. The leg bone and skin are shown clearly, without much confusion. The leg muscle is fairly typical, surrounded mostly by highly contrasting fatty tissue, however sometimes with only a thin channel between one muscle and the next.

IS, also known as the Live-Wire tool, is a user-guided method. With an initial mouse click, the user places a starting point on a boundary of interest; the system then follows the edges in an image to define a path from that initial control point to the cursor's current screen location. As the cursor moves, this path is updated in real time and appears to be a wire snapping around on the edges in an image, hence the terminology "live wire" for this tool.

Figure 1. A transverse image of the leg, highlighting the femur (bone), the biceps femoris (muscle), and the skin.

ACM use an energy minimizing spline, which is initialized close to a structure of interest and then settles into an energy minima over multiple iterations. The energy function is defined so these minima correspond to boundaries of interest in the imagery. Since the initial contour is closed, the final result will always be a closed, continuous curve.

3. Comparing Boundaries

For a ground truth in this comparison, an expert was asked to manually trace the structures. The expert traced the structures twice, generating two independent contours for each structure, which permits a basic measure of the variation within the expert's manual tracings to be quantified. It might be argued that this ground truth is not really a truth, but one user's subjective judgement of a structural boundary. But the expert user brings outside knowledge to bear on the problem, and is dealing with more than simple pixel values when delineating a boundary. And for a system to be useful and acceptable as an assistant to an expert, it should replicate what the expert is attempting to do, rather than do what is dictated by some set of *a priori* assumptions over which the expert has no input or control.

The boundary definitions are to be quantitatively compared to each other and to the ground truth of the expert. The boundaries produced by each of these methods are basically sets of ordered points which can be connected by lines or curves or splines to create a visually continuous bound around a region. To compare two boundaries, A and B, we first connect the points of B in a piecewise linear curve, and measure the minimum distance from each point in A to the curve of B. We then flip the process around, and measure from each point in B to the piecewise linear curve of A. The collection of these measures is called a *difference set*.

Figure 2 illustrates several visualizations of this difference set. The first graph in the upper half of the figure is a plotting of distances between the curves (on the vertical axis) as a function of position on the curve (on the horizontal axis). In this example, there are perhaps three places where the curves are significantly more than one pixel apart from each other, shown in the plot by the six excursions of the graph over the distance of 1.0 (remembering the plot measures A to B and B to A, thus typically excursions show up twice). If the goal is to have the curves within one pixel of each other, this indicates that there are three places where operator intervention would be required to adjust the curves so as to meet that objective.

The lower left of Figure 2 is a simple histogram of the distance set, with the number of occurrences on the vertical axis and distance bins on the horizontal. The lower right is an empirical cumulative distribution function (CDF) over the distance set. The vertical axis measures the fraction of occurrences that are within

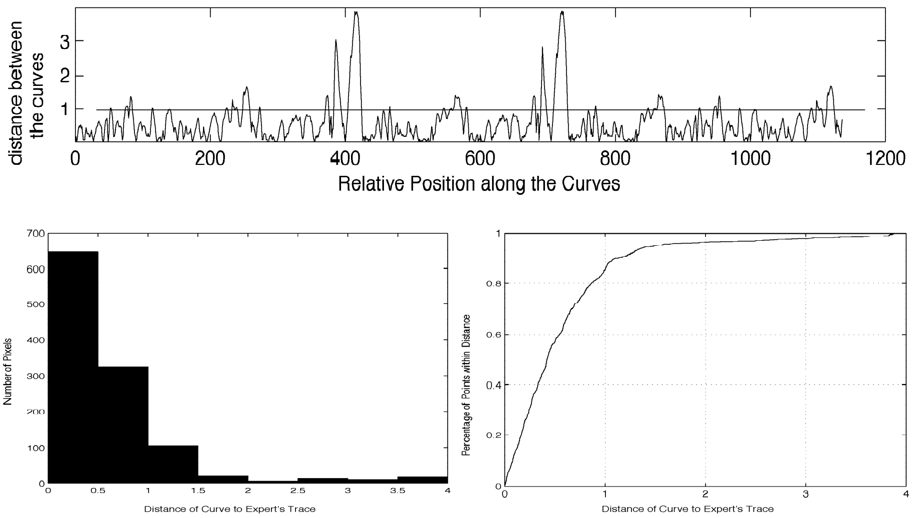

Figure 2. Some visualizations of the distance set.

a tolerance specified on the horizontal axis. The CDF allows quantification of the inter-curve distances by selecting a tolerance such as one pixel and stating, *"The curves are within one pixel of each other 86% of the time"* or by selecting a percentile such as 90% and stating, *"The curves are within 1.1 pixels of each other 90% of the time"*.

4. Key Results

We have experimented with several images, and the three structures of Figure 1 typify the range of results found so far. The expert manually outlined each structure on two independent trials. The first expert trial is used as the Ground Truth (GT), while the second expert trial (M2T) is used to provide a measure of intra-expert variation, i.e., the inherent variation a user shows in tracing a boundary at different times. Given there exists variation within what an expert might manually trace, a good boundary delineation method needn't exactly match any specific expert's boundary definition, but it should be within the range of that expert's variance.

Looking at the left-hand side of Figure 3, this is exactly what we observe. The black curve illustrates the CDF of M2T compared to the Ground Truth. Note the three methods are roughly comparable, all close to the bound of the black CDF. The right-hand side of the figure shows the performance for the muscle. The performance is consistently worse overall. Figure 4 shows a detail of the five boundaries superimposed on the original image; The expert's traces are in black, the semi-automated traces in white. Note in the lower-left of the muscle, there is no consistency of definition, even the expert made different traces at different times. All did equally poorly.

Figure 5 shows the results for the leg skin. Here the performance difference is dramatic between ETA (far left) and the IS and ACM methods (to the right). Figure 6 illustrates what is happening in this situation. The expert judgement of the skin boundary places it slightly inside what a more classically-defined "true" boundary would be; note that both IS and ACM are agreeing on where the boundary lies, and *a priori* this appears to be a sensible boundary to draw. In this case, however, the body was encased in a gel before freezing, and the expert is accounting for both gel effects and the image pre-processing in locating the actual skin boundary. The expert is consistent in this judgement, and the ETA system has learned this behavior and replicated it.

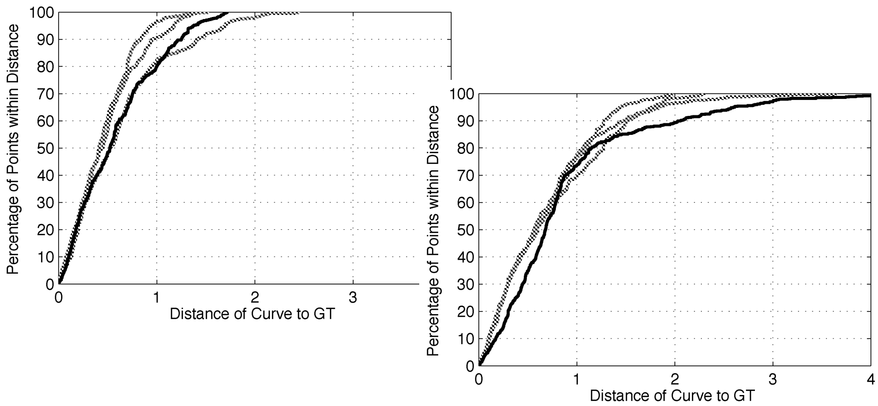

Figure 3: CDF s of M2T (black) and IS, ACM, and ETA (grey) for the bone (left) and the muscle (right) from Figure 1.

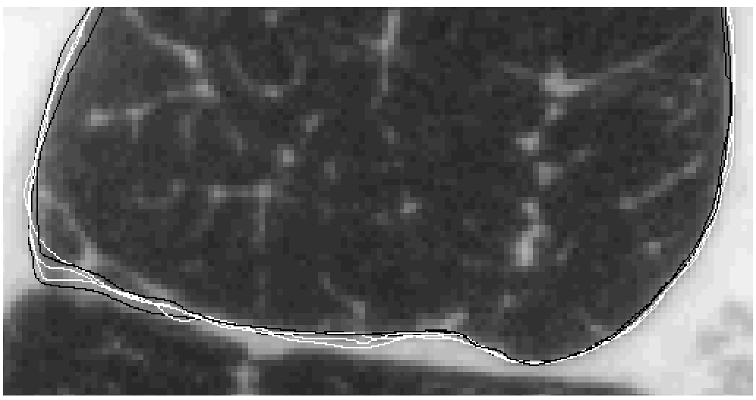

Figure 4. Detail of the five boundaries.

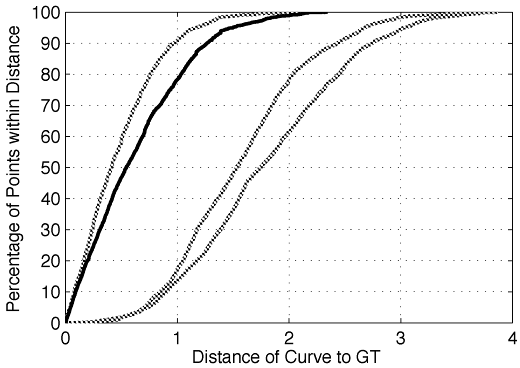

Figure 5. Results for the leg's skin: CDFs for, from left to right: ETA, M2T, ACM, and IS.

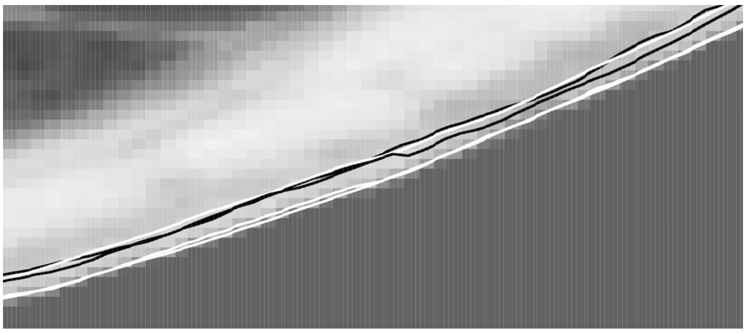

Figure 6. In this detail of the skin, the expert (in black) has traced consistently to the inside of what would be clasically considered the image's edge; ETA (white) follows the expert's lead, while IS and ACM follow more traditional edge definitions.

The results here typify what we've seen so far: the learned boundary was either consistent with the classically defined IS and ACM methods, or it did better when expert judgement was called into play.

References

[1] S.Crawford-Hines & C.Anderson, "Learning Expert Delineations in Biomedical Image Segmentation", *ANNIE 2000 - Artificial Neural Networks In Engineering,* November 2000.

[2] E.N.Mortensen & W.A.Barrett, "Interactive Segmentation with Intelligent Scissors", *Graphical Models and Image Processing*, v.60, 1998, pp.349-384.

[3] M.Kass, A.Witkin, D.Terzopoulos, "Snakes: Active Contour Models", *First International Conference on Computer Vision*, 1987, pp.259-268.

[4] National Library of Medicine, as of Jan 2001:
`http://www.nlm.nih.gov/research/visible`