Calvin Zhao
4/8/15
CIS II: Paper Seminar Critical Review

Critical Review: Contour Detection and Hierarchical Image Segmentation

Arbelaez, P., Maire M., Fowlkes C., Malik, J., Contour Detection and Hierarchical Image Segmentation. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2011

Project Review

Our project is titled "Image Processing for Video-CT Registration in Sinus Surgery." We aim to develop an occluding contour extraction algorithm that can reliably and efficiently find edges in sinus surgery endoscopic videos. Our project is a critical component of a larger one that aims to enhance magnetic tracker resolution in sinus surgery. The proposed registration algorithm is being developed by Seth Billings, a graduate student, and it will map the extracted contours to patient CT data in order to enhance registration. We hope that ultimately this enhanced registration will enable an augmented reality overlay to be projected over the endoscopic data so that surgeons can perform faster and better surgeries.

Occluding contour extraction is not trivial. One of the largest hurdles is filtering out noise. We currently are experimenting with Canny edge detection, which is a basic algorithm that mainly analyzes pixel intensity values. While Canny edge detection is very comprehensive and can identify a large number of contours, it has a large amount of noise. Edges from lighting artifacts and textures show up prominently and they are hard to remove. Canny edge detection relies on 3 main methods for reducing noise. The first layer is a Gaussian blur to remove small texture artifacts, the second is non maximum suppression, which thins out edges to make them more precise, and the third is a double threshold hysteresis method that categorizes edges into weak and strong. Weak edges are eliminated if they do not intersect any strong ones. From our results, it is clear more advanced edge detection methods are needed for us to achieve an accuracy level sufficient for Seth Billing's registration algorithm.

Paper Selection

Figure 1

The paper reviewed herein is titled *Contour Detection and Hierarchical Image Segmentation*. It was published in 2011 by 3 researchers at the University of California at Berkeley. In this paper, the authors explore a novel image segmentation and contour detection method. Their results look promising, as they are able to accurately segment high-noise, real world images, as seen in Figure 1. Using various benchmarks, the paper empirically proves that their methodology is superior to existing ones. The paper utilizes multiple methods to deliver such results. Their key advantages are that they use a multi-directional gradient to evaluate and weigh contours and examine the picture globally to apply appropriate contour weights. Their image segmentation methodology also guarantees closed contours and is easily tuned to allow for more or less detail.

The paper starts with a lengthy overview of existing contour and image segmentation methodologies. It goes into length about their flaws and speculates on the thought processes behind the methods. The paper then discusses benchmarks and quantitative methods of evaluating performance both for segmentation and contour detection. The paper then goes into the work done, starting with contour detection and then segmentation. It concludes with benchmarks and an appendix that talks about the more technical details with regards to performance and efficiency.

Contour Detection

Probability of Boundary (Pb)

The basis for the author's contour detection algorithm is a method coined *probability of boundary*. This method uses an oriented gradient signal to examine the strength of a contour through a set of pixels. The method works by examining each pixel locally. It locates a region of pixels in a radius around the target pixel and bifurcates the region with a straight line. The line is set at an angle θ . The two halves of the region are then examined independently. A pixel intensity histogram is constructed from reach region and the χ^2 distance between these two histograms is calculated and the result is termed the *gradient magnitude*. This value is calculated for each pixel at each particular θ .

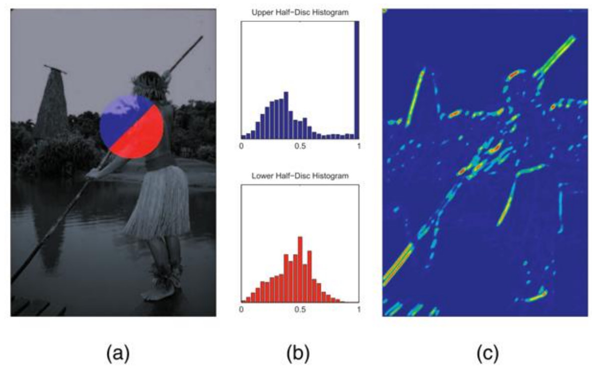

Figure 2. C is the gradient magnitude

Multiscale Probability of Boundary (mPb)

The authors then extrapolate from Pb. They use 4 different channels and 8 different θ to perform Pb on. The 4 channels they examined were intensity, color A, color B and Texton. Texton is a channel that is formed by convolving the image with various Gaussian derivative filters. The pixels are then clustered using K-means and the resulting cluster assignments replace the pixel intensity information to form a new image. This new Texton image highlights the strongest edges. Pb is performed considering each of the 4 channels. The results are summed to form the mPb.

$$mPb(x, y, \theta) = \sum_i \alpha_i G_{i,\alpha(i)}(x, y, \theta)$$

The multiscale probability of boundary is defined above where i is the channel, α is a constant, and $G(x, y, \theta)$ is the gradient magnitude. The authors further compress their algorithm to take

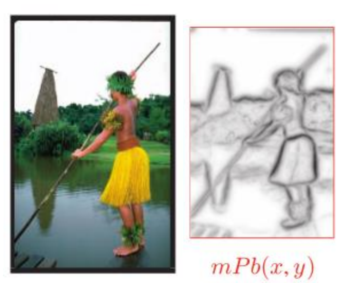

Figure 3.

only the maximum θ at a particular pixel. They define the following as the mPb.

$$mPb(x,y) = \max_{\theta}\{mPb(x,y,\theta)\}$$

This final result can be seen in figure 3. It is the aggregate of 4 different channels across 8 different θ ranging within $[0,\pi]$.

Spectral Probability Boundary (sPb)

A different approach that the paper considered was sPb. This method incorporates global information to gauge the strength of each contour. It is calculated by examining local pixels in a radius around a target pixel. Only Pixels which have a strong contour, as measured by Pb, across them are considered.

$$W_{ij} = \exp(-\max_{p \in \overline{ij}}\{mPb(p)\}/\rho)$$

W is a sparse symmetry affinity matrix where ρ is a constant, and $\overline{ij}$ is the segment which connects pixels i and j. In order to consider global information the authors consider $D_{ii} = \sum_j W_{ij}$. $(D-W)v = \lambda Dv$ is solved and applied into the final definition for sPb:

$$sPb(x,y,\theta) = \sum_{k=1}^{n} \frac{1}{\sqrt{\lambda_k}} \nabla_\theta v_k(x,y)$$

This equation highlights the most salient contours and provides a weight of a contour based off the entire image and not just localized areas. This is a key component that differs from other edge detectors and allows the authors to get such good results. Figure 4 displays the final result. Thickerlines indicate higher weighted edges.

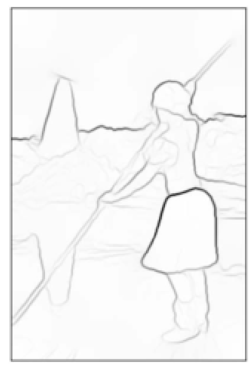

Figure 4.

Global Probability Boundary (gPb)

In the previous sections, we discussed mPb and sPb as two methods for extracting contours. mPb is capable of extracting many contours while sPb highlights the most important ones. The authors combined both methods into an algorithm coined gPb. This is simply a linear combination of both methods to get the best of both worlds. gPb is defined as

$$gPb(x,\mathrm{y},\theta) = \sum_s \sum_i \beta_{i,s} G_{i,\sigma(i,s)}(x,y,\theta) + \gamma * sPb(x,y,\theta)$$

where β and γ are constants. The authors weighted both components equally in the paper to yield their results.

Calvin Zhao
4/8/15
CIS II: Paper Seminar Critical Review

Segmentation

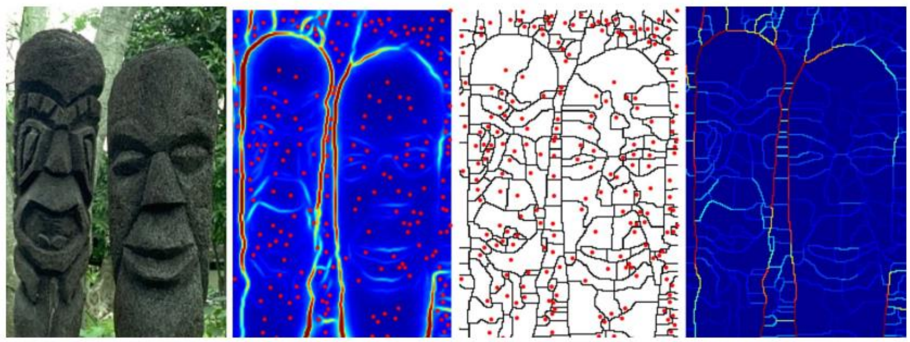

Figure 5.

The authors performed segmentation separately from contour detection. Though they used their contour detection algorithm in their segmentation algorithm, any contour detection algorithm would have sufficed.

In their segmentation algorithm, they performed contour detection across the entire image. They weighted each pixel in an image with the probability that that pixel was an edge. Local minima were then highlighted. In Figure 5, the second image to the left shows these results. The heatmap corresponds to the likelihood that a pixel represents an edge, and the red dots indicate local minima.

These minima points were used in an oriented watershed transform. To understand an oriented watershed transform, you can imagine that the heat map represents a topological height map. The red dots are basins where there is a dip in the topology. If these basins are filled from the bottom up, there would be distinct lines at which water from 2 different basins touch. In figure 5, the third image from the left displays these contours where these “bodies of water” would collide. The sooner a body of water collides, the less weight that contour is assigned. In the last image of figure 5, the places where the collisions occur last are assigned a higher value and are depicted in red. In this manner, an entire image can be segmented into areas which have different weights assigned to their edges.

To perform image segmentation, the authors iteratively combined similar areas. They performed these iterations until the image was completely combined and then backtracked through their iterations to find an appropriate segmentation. This allows for easy calibration and adjustment of the segmentation specificity.

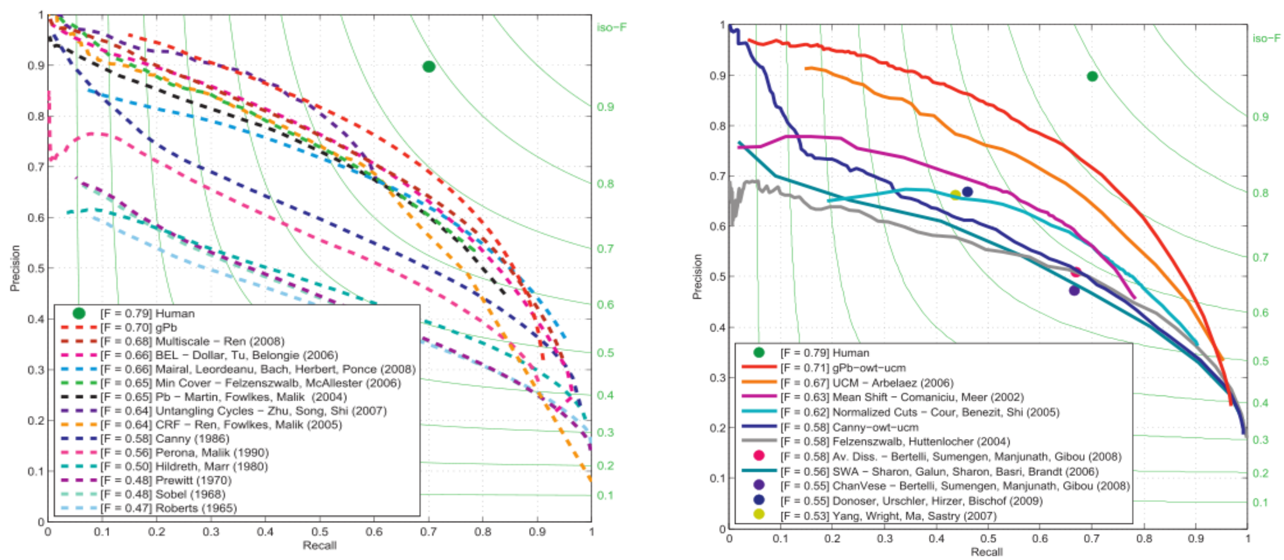

Figure 6. The left side is for Contour Detection and the right is Segmentation

Benchmarks

The authors ran their code alongside previous publications on a dataset of everyday images they called BSDS (Berkley Segmentation Data Set). They plotted the results in a precision and recall graph to show that their algorithm performed better than existing ones. The results can be seen in figure 6.

Critique

The paper overall was well written and testing was very through. There clearly and transparently showed advantages over previous methods. However there were a few parts that made the paper confusing. For one, terminology was never clearly defined. mPb, Pb, sPb and gPb were used and were often left undefined, only to be explained later in the paragraph.

In addition, I found the introduction tedious and unnecessary. Though background information was very through, it did little to support the authors work. Benchmark comparisons would have sufficed to show improvement.

Other areas of the paper had insufficient explanation, particularly the Texton and sPb explanations. I imagine the math would have been tedious, but the authors skipped multiple steps to draw their conclusions. The origins of the sPb and the rationale behind the final equation were never properly explained. Textons, which was a critical component in the author's mPb algorithm, was merely glossed over.

The paper also seemed disjoint. The segmentation and contour sections were seemingly unrelated. Though the segmentation algorithm utilized the author's contour algorithm, it could have also used any number of edge detection methods. There was very little linking the two sections together. In fact, there are two conclusion sections (one for each part). I almost felt like I was reading 2 different papers.

Conclusion

It is clear this paper provides advantages over existing algorithms. This can be verified both visually and empirically. However, use of these algorithms may prove difficult for our specific project. One of the issues is that these algorithms are very slow. Computation of multiple eigenvalues and vectors for each pixel in the sPb algorithm as well as each angle in the mPb algorithm slow down iterations. In applications such as surgery, where real time contour detection is imperative, cannot tolerate such high computational overhead.

Since our videos are mostly monochromatic and feature mainly red and pink, use of multiple color channels in the mPb algorithm may not be beneficial. Our data differs from everyday life photos in that there is little color diversity and there is high glare from the endoscopic light source. These two issues make direct implementation of these algorithms hard.

In addition, their segmentation algorithm may not be easily implementable. The paper focuses on closed contour segmentation, in which each image area is bound by a clearly defined contour. In our videos this is not the case. Human anatomy often as segments that branch off of others, and these new branches cannot be easily confined to a closed contour. Additional experimentation is needed to adapt these segmentation algorithms for our purposes.