Beginnings

If someone had asked me one year ago about life after graduation, he would have found me deep within my eternal dilemma: industry or academia? Today, I am finally moving close to an answer: a belief in academic research and the advancements in life it makes possible. I hope to build on my graduate studies in Bayesian statistics and computer vision, to learn and develop new theories with practical applications. I bring an open mind and the conviction that, through research, we can improve our world.

My path through mathematics has seen countless detours, yet I am still walking it to this day. As an undergraduate in the heart of Silicon Valley, I was drawn by the wonders of programming and the seemingly instant gratification of the internet. It was fortunate, then, that in my junior summer of 1999, I attended an NSF Research Experience for Undergraduates at Cornell University. Under the direction of Professor Károly Bezdek\textsuperscript{1}, I studied computational geometry and began to understand the balance between mathematics and computer science. My work led to a paper\textsuperscript{2}, co-authored with Professor Bezdek, which appeared later that year in the journal Periodica Mathematica Hungarica.

Today, as a graduate student, I have found similar interplays between math and computer science, particularly in the fields of numerical analysis and computer vision. My work in these fields has convinced me that no matter how many times I may wander off the path, I will return to walk on it yet again.

Summary

The goal of computer vision is, to put it simply, to make a computer “see”. This seemingly innocuous definition, however, poses difficult questions. Consider image segmentation, the decomposition of a scene into its salient objects. Should a program simply find and connect edges, or should it optimize more global constraints? Does segmentation involve classification of objects, and does prior knowledge of objects influence segmentation? These philosophical and yet very physical concerns have led to a variety of approaches to computer vision, some based on neuroscience and others on pure computational power. With so many uncertainties, it is easy to see why the field is so engaging.

Professor Davi Geiger\textsuperscript{3} and I are studying applications of Bayesian belief propagation to computer vision. The graphs on which belief propagation occurs allow local information to propagate globally and are thus well suited for vision problems. Any successful application of belief propagation, however, must overcome several difficulties. For example, exact solutions on loopy graphs are NP hard to compute, and approximate variational schemes can be computationally expensive.

We propose a scheme that combines exact belief propagation on multiple sets of tree structures, with each set approximating the original graph. The algorithm is highly efficient and achieves accuracy beyond that attainable on each set alone. We have tested our solution in a controlled image segmentation environment, where we add Gaussian noise to a clean, two-level grayscale image of simple geometric shapes. The algorithm competes through extremely high noise levels with the performances of the optimal maximum-flow algorithm and standard or loopy belief propagation.

Our approach is also computationally flexible, allowing for various parameters to be learned rather than manually selected. This is particularly useful for natural image segmentation, where we would like to propagate contour boundary information in conjunction with grayscale information. The resulting parameters number in the hundreds, far too many to accurately define by hand. Thus, we train these parameters on a set of pre-segmented images, and the results are clearly competitive with those of manually constructed grayscale-based systems.

\textsuperscript{1}Director of the Centre of Computational & Discrete Geometry, University of Calgary; and head of the Department of Geometry, Eötvös Loránd University, Budapest.


\textsuperscript{3}Professor of Computer Science and Neural Science, New York University, geiger@cs.nyu.edu.
Details

Consider belief propagation on a graph with \( n \) vertices, labeled 1 through \( n \). Vertex \( i \) has the set of possible states \( X_i \) and an “observed value” \( y_i \). The “compatibility function” \( \phi_i(x_i, y_i) \) defines the statistical dependency at vertex \( i \) between its observed value \( y_i \) and the state \( x_i \in X_i \), and the function \( \psi_{ij}(x_i, x_j) \) defines how compatible two states \( x_i \in X_i \) and \( x_j \in X_j \) at neighboring vertices \( i \) and \( j \) are. The “belief” that vertex \( i \) is in state \( x_i \) is then

\[
b_i(x_i) = \frac{1}{z_i} \phi_i(x_i, y_i) \prod_{j \in N(i)} m_{ji}(x_i),
\]

where \( z_i \) normalizes the beliefs and \( N(i) \) is the set of vertex \( i \)’s neighbors. \( m_{ji}(x_i) \) is the “message” vertex \( j \) passes to vertex \( i \) indicating its belief, given the knowledge already accumulated from its ancestors, that vertex \( i \) is in state \( x_i \). It is defined as

\[
m_{ji}(x_i) = \sum_{x_j \in X_j} \phi_j(x_j, y_j) \psi_{ji}(x_j, x_i) \prod_{k \in N(j) \setminus i} m_{kj}(x_j),
\]

where \( N(j) \setminus i \) is the set of vertex \( j \)’s neighbors, excluding vertex \( i \).

On general graphs with loops, messages recurse back on themselves and become impossible to define. On trees, however, the message recursion terminates at each leaf vertex, and belief propagation is exact and, in fact, computationally efficient. With this in mind, consider a general graph with loops. We first approximate the graph with several sets of trees and perform exact belief propagation on each set. If we select these trees carefully, our solutions will already be fairly accurate compared to the exact solution.

Next, we take advantage of the fact that beliefs exist in the entire interval \([0, 1]\). Statistically, when a tree’s decisions are incorrect, the corresponding beliefs are lower than when its decisions are correct. Indeed, incorrect decisions should have weaker energy support than correct decisions. We thus propose a scheme that combines information from distinct sets of trees by selecting the decisions with the highest beliefs among all sets. This non-linear combination of information yields accuracy higher than that on each set alone and, in image segmentation tests, approaching that of the optimal result. The concept of sets of “complementary” trees reinforcing each other’s computations is the essence of our approach, which aims to be an alternative to current energy approximation schemes.

Our graph-combination approach has great flexibility in constructing and, in particular, learning new energy functions. For the problem of natural image segmentation, we start as usual by incorporating grayscale information, which requires only a handful of energy function parameters. Once we begin incorporating contour boundary information, however, the number of parameters grows to several hundred. We can accurately define these parameters by training our algorithm on a set of pre-segmented images. The local boundary interaction this learning makes possible produces segmentations competitive to those produced by purely grayscale-based systems, as evident in the segmentations of a gorilla image below.

![Image](image.png)

Figure 1: The (a) original image, (b) best hand-selected grayscale-based segmentation, and (c) trained boundary- and grayscale-based segmentation.