Exploiting Texture-Motion Duality in Optical Flow and Image Segmentation

Michael G. Ross & Paul Viola

Artificial Intelligence Laboratory
Massachusetts Institue Of Technology
Cambridge, Massachusetts 02139

http://www.ai.mit.edu

The Problem: The goal of image segmentation is to divide an image into regions that correspond with the objects present in a scene. The goal of optical flow is to examine a video sequence and calculate which motion vectors correctly describe the two-dimensional motions present. Segmentation algorithms are best at detecting the existence of low-texture regions composed of constant or smoothly-varying colors, but have difficulty dealing with complex, high-texture patterns. Conversely, optical flow algorithms function best in high-texture regions and very poorly in areas without brightness variations. Due to the complementary nature of their strengths and weaknesses, and due to the fact that both processes are extracting information about the same set of real-world objects, it is possible to improve segmentation results by incorporating optical flow information, and vice-versa.

Motivation: Optical flow algorithms are under-constrained because the traditional assumption, the brightness constancy constraint, only provides one linear equation to describe the x and y motion of an image location. Without additional information, we must make an additional assumption, usually that motion vectors vary smoothly across the image, in order to solve the flow field. This assumption is not always correct, because flow-fields can be discontinuous at object boundaries. So we can help alleviate the problem by introducing information about potential object boundaries derived from the output of an image segmentation algorithm. Similarly, the chief challenge faced by a segmentation algorithm is determining which color discontinuities represent object boundaries, and which represent variations within a single region. By adding information from optical flow, we can determine which neighboring regions are moving together and use the additional information to assist the segmentation algorithm. Thus, in both segmentation and in optical flow, we can improve performance by supporting the primary source of information with data from another algorithm.

Previous Work: The algorithms developed for this project are based on segmentation algorithms developed by Felzenszwalb and Huttenlocher [2] and an energy-minimization algorithm developed by Boykov, Veksler, and Zabih [1]. The use of segmentation information to determine smoothness boundaries of optical flow fields is similar in spirit to the use of dual Markov Random Fields in image restoration, first described by Geman and Geman [4]. A full and thorough explanation of the algorithm and background research is given by Ross' master's thesis [5].

Approach: The key theoretical component is the realization that a reasonable segmentation algorithm produces regions that have enough texture to allow for a determination of their optical flow. The Felzenszwalb-Huttenlocher algorithm can produce two types of regions. One is a region with a low-variation interior, surrounded by a higher-variation edge. The other is a region that has high interior variation. In either case, so long as the regions are reasonably sized, enough different brightness gradients are present to completely constrain the optical flow equations.

Our algorithm can be divided into three steps. First, we compute a fast image segmentation with the Felzenszwalb-Huttenlocher algorithm. Then we compute an optical flow by constraining motion to vary smoothly within a segment, but imposing no penalty for discontinuities across segment boundaries. If we assume that the image is over-segmented, that no segment contains pixels from more than one object, and that all motion can be modeled by local translations, this smoothness criteria is valid and allows the optical flow field to contain accurate motion discontinuities. The third and final step of the algorithm is combining neighboring regions that were close to being joined by the segmentation algorithm's uniformity criteria and which share identical motion-labels in the optical flow field.

Difficulty: These problems are difficult because there is no ground truth to compare results against. In segmentation, there is no widely-accepted definition of what constitutes a "region" or an "object." Optical flow is better-defined, but

Figure 1: From left to right, a frame of a video sequence of a man walking to the left, a segmentation of that frame augmented with optical flow information, and an optical flow field computed using segmentation information.

even human beings cannot determine the true flow of a natural scene with complete accuracy. Both classes of algorithms rely on pixel data as the sole information source and the information deficiencies of images must be overcome with strong prior assumptions.

Impact: Improving our ability to compute optical flows and segmentations that are reliably consistent with the objects in the imaged scenes is crucial to higher-level visual processing. Improving these processing steps will allow us to move towards accurate, object-based representations of scenes and achieve the bandwidth reductions necessary for efficient high-level visual understanding and learning algorithms.

Future Work: The current algorithm improves optical flow results more than it improves segmentation. The use of optical flow data in the segmentation algorithm can only correct over-segmentation, and has a tendency to create or increase under-segmentation. In the future we hope to develop a version of the segmentation algorithm that incorporates optical flow information in the initial determination of segments. One possibility might be to use a newer version of the Felzenszwalb-Huttenlocher algorithm [3] that determines the neighborhood relations of pixels by the distances between their (r, g, b, x, y) feature-vectors. Optical flow information could be added by moving to (r, g, b, x, y, dx, dy) feature vectors, where dx and dy represent the optical flow assigned to each location.

Research Support: This research is supported by a National Science Foundation Graduate Research Fellowship and by the NTT grants for the Image Database Project and the Variable Viewpoint Reality Project.

References:

[1] Y. Boykov, O. Veksler, and R. Zabih. Fast Approximate Energy Minimization via Graph Cuts ICCV, 1993.

[2] P. Felzenszwalb and D. Huttenlocher. Efficiently Computing a Good Segmentation. 1998 DARPA Image Understanding Workshop.

[3] P. Felzenszwalb and D. Huttenlocher. Efficiently Computing a Good Segmentation. Cornell University, 1999.

[4] S. Geman and D. Geman. Stochastic Relaxation, Gibbs Distributions, and the Bayesian Restoration of Images. PAMI 6:6, Nov. 1984.

[5] M. Ross. Exploiting Texture-Motion Duality in Optical Flow and Image Segmentation. Master's Thesis, Massachusetts Institute of Technology, August 2000.