I. SEGMENTATION PROCESSES IN THE VISIONS SYSTEM[1]

John Prager, Paul Nagin, Ralf Kohler, Allen Hanson , and Edward Riseman
Computer and Information Science Department
University of Massachusetts
Amherst, Massachusetts 01003

Abstract

In VISIONS the segmentation algorithms receive input from a digitized color image as an array of pixel values, and generate output as a layered graph of symbolically named regions, boundaries, and endpoints — each with their descriptive attributes. This processed sensory data provides the initial input to the model-building system of VISIONS.

Introduction

This set of three short papers is a description of the first pass at implementing VISIONS, a computer system designed to understand digitized images of natural outdoor scenes (e.g., Figure 1). VISIONS [Hanson and Riseman, 1976] is divided into two major subsystems: the low-level system is responsible for segmenting the image; the high-level system is responsible for interpreting the image. This paper discusses segmentation; the second paper outlines the multi-level representations, the control structure, and the search space for model-building; and the third paper discusses shape and the linkages between regions, surfaces, volumes, and objects.

Two approaches to image segmentation are presented here: an edge/line segment process [Prager, Hanson and Riseman, 1977] and a region analysis process [Nagin, Hanson and Riseman, 1977].

Boundary Analysis

The line/boundary analysis process is based on an interpixel representation of only horizontal and vertical edges. An advantage of this approach is that decisions concerning orientations of boundary segments (i.e., groups of connected edges) are delayed until the local information has aggregated and global information from the whole segment is available. The algorithm involves two preprocessing operations, a simple differentiation pass using a 1x2 edge mask and suppression of non-maxima (parallel) edges which have the same direction (sign) of contrast. Next relaxation processes [Rosenfeld, Hummel and Zucker, 1976] are used to organize information. These processes, based on a paradigm of local cooperation and competition, update the strength of edges as a function of local context. They employ a probabilistic case analysis of line continuations, line endings, and vertices; sample results are shown in Figure 2.

Since, the properties of a boundary can be expected to remain invariant only for segments which lie between at most two regions, the goal is to extract and symbolically label them. Therefore, the next step of the algorithm uses the similarity of properties to either side of a pair of edges to bind them into segments, thus integrating both edge and region information at a local level. In order to form these segments, binding must not take place across vertices where more than 2 line segments meet. After extracting boundary segments, properties such as confidence, straightness or curvature, contrast, etc. are computed. Elimination of low confidence lines can take place during a postprocessing stage, if desired.

Region Analysis

We have chosen to develop a region formation process using global information as a complementary approach to the locally-directed line analysis. Here the properties of points and small areas are analyzed via global histograms (one- and two-dimensional) to focus on the internal structure of regions [Hanson, Riseman and Nagin, 1975; Ohlander, 1975; Price, 1976]. The features selected to provide the basis for the segmentation are first used to form plans [Nagin, Hanson and Riseman, 1977; Price, 1976; Kelly, 1971]. The plans drastically reduce the data while organizing a gross decomposition of the image into several large areas. Each of these areas can then be further refined without interference in the histogram analysis from other areas in the image.

Plans are formed using two techniques: a) quantization compression, and b) spatial compression. Compressing the quantization scale of a given feature to one or two bits produces a plan which emphasizes only the strongest differences in global feature activity. This technique can be generalized by intersecting the results obtained from a small set of features (Figure 3). The second stage of planning involves reducing the spatial resolution of the plan data by replacing non-overlapping local windows of pixel values with scalar features extracted from the local window. The plans can now be more carefully refined by splitting each overmerged area into regions using similar techniques. The goal of the refinement process is to eventually provide a region segmentation of the image together with associated descriptive region properties.

Merging Representations

Techniques for merging these two segmentations to form more reliable output are currently under examination. By converting the region plans to boundaries, the line and region information can be intersected and confidences of regions and boundaries can be improved. In Figure 4, the boundary segmentation of Figure 2 has been improved using region plans.

This work was supported by ONR grant N00014-75-C-0459.

[2] A. Hanson is affiliated with Hampshire College, Amherst, Massachusetts 01002.

The resulting segmentation is passed to the high-level system via the RSE data structure. The RSE is a three-level directed graph of regions, boundary segments, and endpoints. Each node in the graph has a symbolic name and an associated list of properties, while arcs encode the relationship between nodes.

References

Hanson, A., Riseman, E. and Nagin, P. [1975] "Region Growing in Textured Outdoor Scenes," Proc. 3rd Milwaukee Symposium on Automatic Computation and Control, 407-417.

Hanson, A. and Riseman, E. [1976] "A Progress Report on VISIONS: Representation and Control in the Construction of Visual Models," COINS TR 76-9, Univ. of Mass., July 1976.

Kelly, M.D. [1971] "Edge Detection in Pictures by Computer Using Planning," Machine Intelligence 6, 379-409.

Nagin, P.A., Hanson, A.R. and Riseman, E.M. [1977] "Region Extraction and Description Through Planning," COINS TR 77-8, Univ. of Mass., May 1977.

Ohlander, R. [1975] "Analysis of Natural Scenes," Ph.D. Thesis, Carnegie-Mellon Univ., April 1975.

Prager, J.M., Hanson, A.R. and Riseman, E.M. [1977] "Extracting and Labelling Boundary Segments in Natural Scenes," COINS TR 77-7, Univ. of Mass., May 1977.

Price, K.E. [1976] "Change Detection and Analysis in Multi-Spectral Images," Ph.D. Thesis, Carnegie-Mellon Univ., Dept. of Computer Science.

Rosenfeld, A., Hummel, R.A. and Zucker, S.W. [1976] "Scene Labelling by Relaxation Operations," IEEE Trans. Systems, Man and Cybernetics, SMC-6, 420-433.

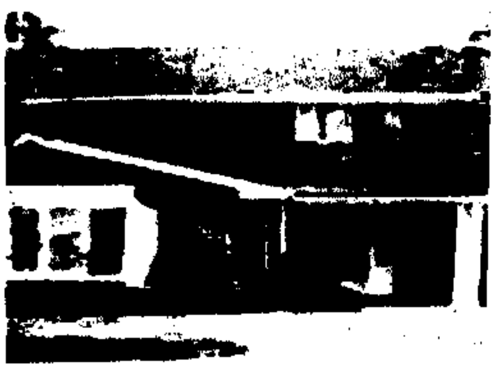

Figure 1 Intensity image obtained from color data. The central 128^2 pixels of a 256^2 image are shown.

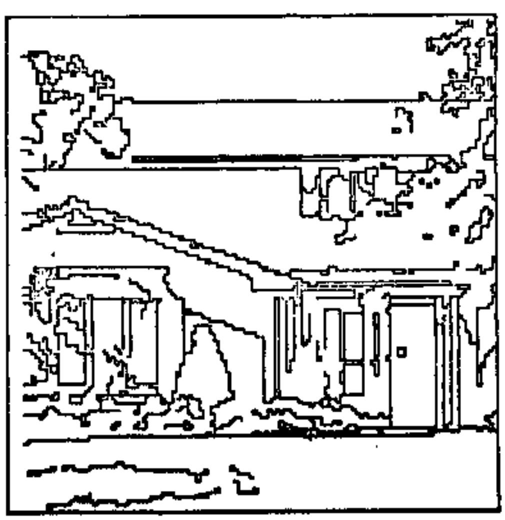

Figure 2 Results obtained from the boundary analysis. The image of Figure 1 was preprocessed and then differentiated using a 1×2 mask. A relaxation process, employing contextual information, was used to organize the edge data into boundaries.

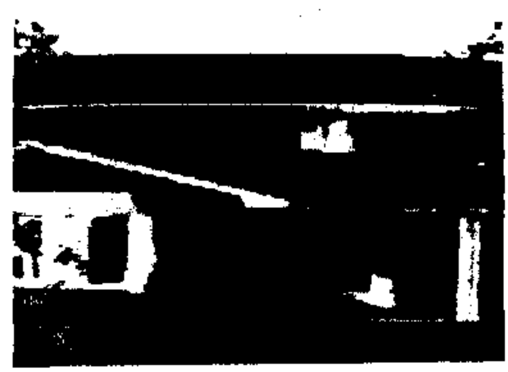

Figure 3 Results from the region analysis. Histogram-guided quantization was applied to each of 3 features independently. The individual plans were then intersected to obtain the results shown.

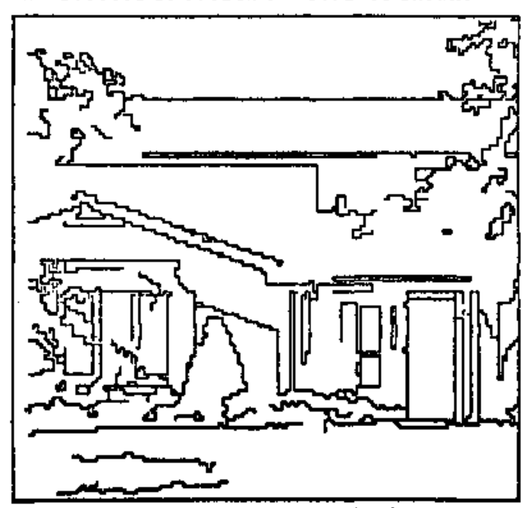

Figure 4 Combined results from both processes. Low-confidence lines which are unsupported by boundaries of region plans are suppressed, yielding an improved boundary segmentation.