INTEGRATING METHODOLOGIES IN IMAGE ANALYSIS*

Theo Pavlidis
Dept. of Computer Science
State University of New York
Stony Brook, NY 11794

Yuh-Tay Liow
Dept. of Electrical Engineering
State University of New York
Stony Brook, NY 11794

ABSTRACT

We discuss ways by which different methodologies for image analysis may be combined for better results. We focus on the combination of region growing and edge detection to achieve better segmentation.

Introduction

After thirty years of research, the literature of image analysis (or computer vision) contains numerous methodologies for dealing with each of the problems encountered. For example, shape analysis is addressed by Fourier descriptors, polygonal approximations of contours, skeletonization, etc. Each method has its proponents and occasionally one reads debates about their relative merits. We claim that a successful attack to image analysis problems requires the simultaneous application of more than one methodology. While there is a significant literature on integrating input from different sensors, very little has been written on integrating methodologies although is by no means new [Ya76, Gr80]. However we have found only three recent references [BML86, ABM87, FH87]. The reasons for that attitude may have to do more with an academic desire for methodological purity than the adequacy of a single methodology to completely solve a problem.

Mixing of methodologies is common in practical systems, for example in OCR (reading machines). It is well known that some pairs of characters are discriminated best by their contour features (for example D and O without serifs) while others are discriminated best by skeletal features (for example E and F with serifs). See [KPB87] for a discussion of a complete character recognition system using a mixture of methodologies. We will present here some results showing combinations of region growing and edge detection. For details of the work the reader is referred to [PL88] and [LP88]. The fundamental justification for integrating methodologies is the fact that human observers use a multiplicity of clues in analyzing an image. While one might show that under ideal conditions a single methodology is sufficient (for example, skeletonization provides complete information about shape) this is not true in the presence of noise and distortions.

* This research was supported by NSF grant IRI-8702168, by a grant from the Digital Image Processing Laboratory of the Lockheed Palo Alto Research Laboratories, and by NSF equipment grant DMC-8705290.

Improving on the Results of Region Growing with Edge Detection

Segmentation by region growing contains artifacts around regions that do not satisfy exactly the model implied by the uniformity predicate. For example, a method will produce false boundaries because the uniformity criterion may not be satisfied over a given area even if there is no clear line where a transition occurs. For example, if the light intensity varies linearly within a region R and we insist that the intensity be approximately constant within a region, then there will be artificial boundaries within R. Furthermore, it is likely that such boundaries will reflect the data structures and traversal strategies used during region growing. For example, if we traverse an image along scanlines, then artificial boundaries will tend to be parallel to scan lines. The method detailed in [PL88] uses two steps after segmentation. In the first boundaries are eliminated and in the second boundaries are modified. The boundary-elimination criterion uses a merit function of the form

$$f_1(contrast) + \beta \cdot f_2(segmentation\ artifacts) \quad (1)$$

Edge modification considers edges according to a criterion

$$f_3(contrast) + \alpha \cdot f_4(smoothness) \quad (2)$$

In the sequel we use the regions produced by a split-and-merge algorithm [HP76]. Since this method uses a quad tree for the initial traversal of the image, the false boundaries tend to consist of long vertical and horizontal segments. Therefore the merit functions is

$$f(e) = \frac{|Sum\ of\ contrast\ along\ e\,|}{length\ of\ boundary\ e} + \beta \frac{\#\ of\ direction\ changes\ along\ boundary\ e}{length\ of\ boundary\ e} \qquad (3)$$

where the first term of Equ. (3) can be interpreted as the strength of the boundary seen as an edge while the second term acts as a penalty for the artifacts of the quad tree structure which tend to be long straight lines; β is the relative weight of these two terms.

A major problem in boundary elimination is that we cannot remove a boundary arbitrarily without introducing unacceptable topologies. This may happen if the boundary between two regions is not connected and after the removal of one arc we may be left with "boundaries" that are internal to a single region. Therefore care should be taken during the traversal of boundaries to ensure that this does not happen. This means that the order of examination of boundary arcs is important since the removal of one arc may negate the removal of other arcs later.

Let e_i denote a connected arc of a boundary. We select a threshold Th and define an overall merit function as:

$$F(e_1, e_2, \ldots, e_n) = \sum_{i, e_i\ elim} (Th - f(e_i)) + \sum_{j, e_j\ pres} (f(e_j) - Th) \qquad (4)$$

which is to be maximized. This is trivially achieved by removing all boundaries with $f(e_i) < Th$ and keeping all others. Such a solution may not be topologically valid though. In considering different sets of boundaries to be eliminated, we must avoid an exhaustive search whose cost will grow exponentially with the number of boundaries. We expect a suboptimal result can be attained by carrying out the boundary-elimination process sequentially by starting with the weakest boundary and by the prediction of the effect on $F(.)$ if a single boundary is eliminated.

After the removal of boundaries which appear to be due to segmentation artifacts we modify the rest by maximizing the following merit function which similar to that of Equ. (3) but with one additional term.

$$\int |\nabla I(\mathbf{W}(t))| - \alpha \cdot curvature(\mathbf{W}(t)) - \gamma \cdot |\phi'(\mathbf{W}(t))|\, dt \qquad (5)$$

The whole expression is both normalized with respect to $\int length(\mathbf{W}(t))dt$. $\mathbf{W}(t) = (x(t), y(t))^t$ is the position of the contour, $|\nabla I(\mathbf{W}(t))|$, the magnitude of image gradient, is the intensity contrast along the contour, $curvature(.)$ is the evaluated curvature along the contour, $\phi(\mathbf{W}(t))$ is the phase part of the image gradient at point $\mathbf{W}(t)$ and "$'$" mean first order derivate. The first term, $|\nabla I(\mathbf{W}(t))|$, is denoted as "fidelity" while the second and the third terms, $curvature(.)$ and $|\phi'(\mathbf{W}(t))|$, are denoted as "smoothness," *i.e.*, we edit the contour to achieve "smoothness" without sacrificing much "fidelity;" the relative weights among these three terms are controlled by coefficients α and γ.

The introduction of the $curvature()$ term in Equ. (5) is to make the resultant contour smooth while that of the $|\phi'(\mathbf{W}(t))|$ term is to avoid the problem that while searching a point P_t in the neighborhood of P to substitute P, we might mislocate P_t at a strong edge nearby. Minimizing $\sum |\phi'(\mathbf{W}(t))|$ maintains a contour with the same gradient direction so that the resultant contour will not be influenced by strong edges nearby. While most researchers used only the magnitude part of the image gradient, $|\nabla I(\mathbf{W}(t))|$, to locate (or modify) contours, we find that the phase information, $\phi(\mathbf{W}(t))$, is as important as the magnitude part. The phase term of gradient was also applied in [Ba76][BHR86] to locate contours and to extract straight lines. Similar schemes to Equ.(5) were also employed by Montaneri [Mo71], Martelli [Ma72], and Ballard [Ba76] to detect lines, edges, and contours. More sophisticated examples can be found in [KWT87] where Kass, Witkin and Terzopoulos use a similar idea to modify tentative contours. Figure 1 shows the result after the split-and-merge algorithm (top) and the final result after boundary elimination and modification (bottom).

Detecting Buildings from their Shadows

It is possible to integrate context in the early stages of processing in addition to integrating region growing and edge detection. We shall deal with the interpretation of aerial photographs. Under many conditions of illumination a region corresponding to a building will be next to a region corresponding to its shadow. Shadows tend to be much darker than anything else in the image so they are easier to detect. In addition shadows of buildings are likely to have certain regularity in their shape. Two methods are described in [LP88] where the reader is referred to for details. In the first method we start with edge detection then combine the result with region growing; while in the second method we use the segmentation of region growing first then apply the idea of edge detection to enhance the result.

The first method can be divided into three phases. (1) We use edge detection [BHR86] to locate shadows and thus to confirm the presence of a building since they provide very reliable features with high signal to noise (S/N) ratios. (2) In the low-contrast areas where boundary features can not be reliably obtained by edge detection, we use a region growing operator to delineate the boundaries. The "leaking" problem that occurs with applying region growing between two low-contrast regions is effectively solved by using morphological operations with a directional operator related to the light orientation. (3) A beautification phase is used to enhance the delineated contours. The idea of beautification was first proposed in the context of computer graphics to modify a rough technical drawing [PVW85]. However, we found that a similar idea can also be applied to enhance the result of the building detector if the orientation of the building is known.

The second method can also be divided into two phases. (1) We follow the paradigm of hypothesis testing on each region of the segmentation. Two features, shadow and contour shape, are used to verify the hypotheses. (2) To beautify the contours of the detected buildings, we use the shape and edge information to remove artifacts and to modify the contours. Since we combine the best parts of the features obtained by edge detection and region growing to find the shape description of each building, the results turn out to be accurate and reliable. Figures 2 and 3 show examples obtained from each method.

Conclusions

We have shown how the results of segmentation can be improved with an integrated approach. It is tempting to cast the whole approach as a large optimization problem. Given an image, partition it into regions R_i by minimizing an expression of the form

$$\sum_i \sum_k U_k(R_i) + \sum_j \sum_k V_k(e_i)$$

where $U()$ are region costs (lack of uniformity, unexpected shape, etc.) and $V()$ boundary costs (lack of contrast, sharp turns, etc.). The challenge is to find functions $U()$ and $V()$ which are realistic and also result in a tractable optimization problem. Very special cases of this approach have been tried in the context of regularization [MS85].

References

[ABM87] Anderson, H.L., R. Bajcsy, and M. Mintz, "A Modular Feedback System for Image Segmentation," *University of Pennsylvania GRASP Laboratory Technical Report 110*, 1987.

[BML86] Bajcsy M., M. Mintz, and E. Liebman, "A Common Framework for Edge Detection and Region Growing," *University of Pennsylvania GRASP Laboratory Technical Report 61*, 1986.

[Ba76] Ballard, D. H. "Hierarchic Detection of Tumors in Chest Radiographs," *ISR-16* January 1976.

[BHR86] Burns, J.B., A.R. Hanson and E.M. Riseman, "Extracting Straight Lines," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, PAMI-8 (1986), pp. 425-455.

[FH87] Fua, P. and T. Hanson, "Using Generic Geometric Models for Intelligent Shape Extraction," *Proc. DARPA Image Understanding Workshop (1987)*, pp. 227-233.

[Gr80] Grinaker, S, "Edge Based Segmentation and Texture Separation," *Proc. of 5th Int. Conference on Pattern Recognition, Miami Beach*, (1980), pp. 554-557.

[HP76] Horowitz, S. L. and T. Pavlidis, "Picture Segmentation by a Tree Traversal Algorithm," *Journal of ACM* 23 (1976), pp. 368-388.

[KPB87] Kahan, S., T. Pavlidis, and H. S. Baird, "Building a font and size independent character recognition system", *IEEE Trans. on Pattern Analysis and Machine Intelligence*, PAMI-9 (1987), pp. 274-288.

[KWT87] Kass, M., A. Witkin, and D. Terzopoulos, "Snakes: Active Contour Models," *1st International Conference on Computer Vision* (1987), pp. 259-267.

[LP88] Liow, Y. T. and T. Pavlidis, "Use of Shadows for Detecting Buildings in Aerial Images" submitted for publication.

[Ma72] Martelli, A. "Edge Detection using heuristic search methods," *Computer Graphics and Image Processing*, vol 1, (1972), pp. 169-182.

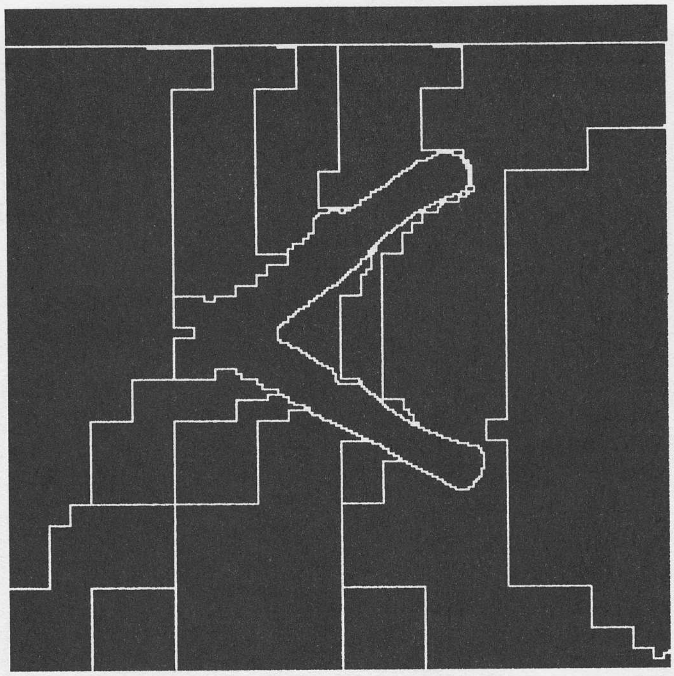
[Mo71] Montaneri, U. "On the Optimal detection of Curves in Noisy Picture," *Commun. ACM 14.5*, May 1971, pp. 335-345.

[MS85] Mumford, D. and J. Shah, "Boundary Detection by Minimizing Functionals," *Proc. IEEE Computer Vision and Pattern Recogn. Conf.* San Francisco, June 19-23, 1985, pp. 22-26.

[PL88] Pavlidis, T. and Y. T. Liow "Integrating Region Growing and Edge Detection," *Proc. IEEE Computer Vision and Pattern Recogn. Conf.* Ann Arbor, Michigan June 6-10, 1988, pp. (in press).

[PVW85] Pavlidis, T., and C. J. Van Wyk, "Beautifier for Drawings and Illustrations," *Proceedings of SIGGRAPH'85, San Francisco*, (1985), pp. 225-234.

[Ya76] Yakimovsky, Y., "Boundary and Object Detection in Real World Images," *JACM*, 23 (1976), pp. 599-618.



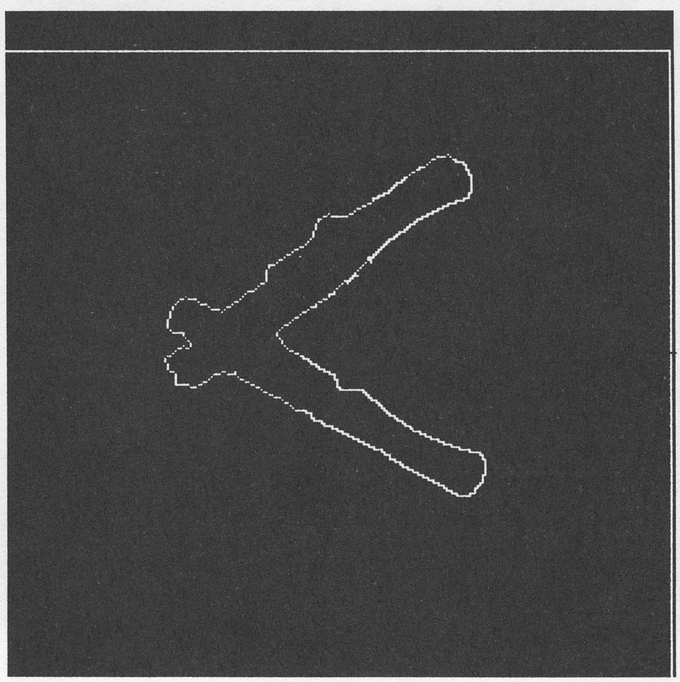


Figure 1: Results of segmentation after the split-and-merge algorithm (top) and after the application of boundary elimination, editing, and modification (bottom).

Figure 2: Buildings overlayed on the original image. They were obtained by edge detection followed by region growing.

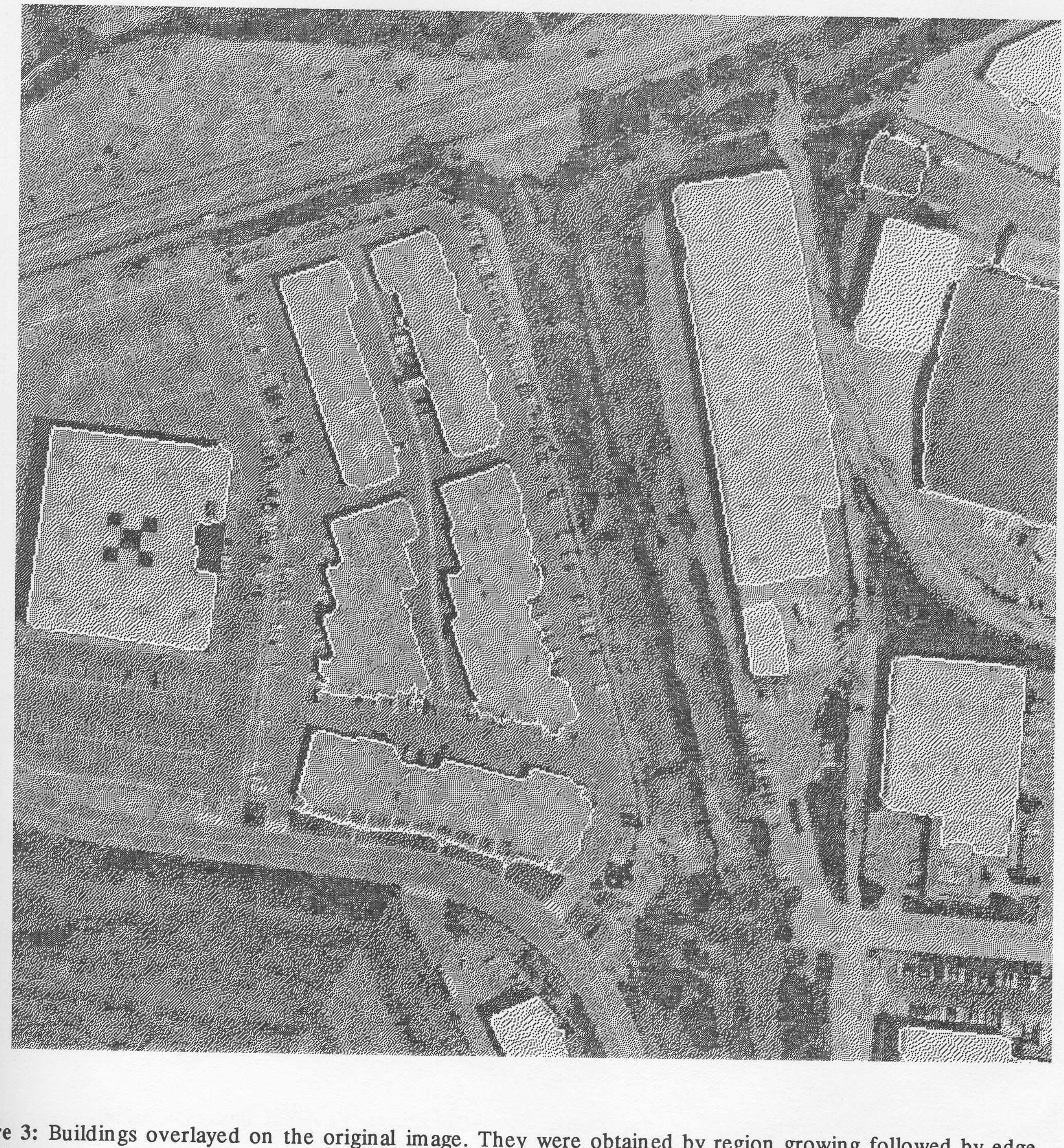


Figure 3: Buildings overlayed on the original image. They were obtained by region growing followed by edge detection.