A New Framework for Hierarchical Segmentation Using Similarity Analysis *

Peter Bajcsy and Narendra Ahuja

Beckman Institute, University of Illinois at Urbana-Champaign, IL 61801, USA

Abstract. We present a new framework for hierarchical segmentation of multidimensional multivariate functions into homogeneous regions. Homogeneity is defined as constancy of n-th order derivatives (called features) of the function. The degree of similarity (measure of homogeneity) is used as a scale parameter to obtain a stack of segmentations. Hierarchical segmentation is represented as a tree which contains the geometric and topological information about the detected regions. Detected regions preserving their information in the tree over large range of scales are selected into a pyramid representation. Results showing noise robustness and computational efficiency of the proposed method are presented. Experiments to compare the method with three other segmentation techniques and applications to two- and three-dimensional images having one-, three- and six-variate data are described for the zeroth and first order region features.

1 Introduction

This paper is aimed at the fundamental problem of image segmentation. The method described next extracts the hierarchical structure present in an image. The image is viewed as a multivariate multidimensional function. The goal of segmenting multivariate multidimensional functions is to partition a regular n_s-dimensional grid of sample points in the domain of a n_f-variate function into nonoverlapping connected sets of sample points forming homogeneous regions. Homogeneity is defined as constancy of n-th order derivatives (called features) of the function [3, 5]. When the goal is to extract information at all levels of detail present in the image, multiscale approaches are required [4, 2, 1]. This segmentation problem has not been solved in general due to the following reasons: (1) The definition of features (e.g., the order of derivatives) is a priori unknown. (2) The requirements on computational resources are severe. (3) The amount and type of noise in the input data is unknown. These difficulties become more severe as the dimensionality of the data increases.

In this paper, the degree of similarity (measure of homogeneity) of two n-th order features is modeled by the mutual Euclidean distance of their n-th order differences. It is denoted throughout the paper as homogeneity δ, which limits

* This research was supported in part by Advanced Research Projects Agency under grant N00014-93-1-1167 and National Science Foundation under grant IRI 93-19038.

the maximum difference between any pair of features within a region. The order n of features depends on the application area. The segmentation method presented in this paper uses a varying homogeneity parameter which gives rise to a number of regions. The noise robustness and computational efficiency of the detected regions are characterized. A tree representation of the segmentation is extracted from these regions that retains the subset of regions corresponding to all scales present in the data. Experiments are reported with medical data, botanical data, satellite data, range data and gray level and color video sequences, having dimensionalities $n_s = 1, 2, 3$ (multidimensional sample space) and $n_f = 1, 2, 3, 6$ (multivariate function space), and using zeroth and first order region features.

2 Segmentation Using Homogeneity Analysis

Given the feature model, the proposed method consists of three major steps: (1) An image is segmented into regions having the same degree of feature homogeneity. (2) The degree of homogeneity is used as a scale parameter to obtain a number of n_s-dimensional ($n_s D$) regions. The detected regions split (merge) as the homogeneity parameter is made more (less) stringent and lead to a hierarchical region structure. (3) $n_s D$ regions from the hierarchical organization are selected to obtain a tree representation of segmentations.

Mathematical framework: An $n_s D$ image is modeled as a multidimensional multivariate function $x_i \longrightarrow f(x_i)$, where $x_i = (x_1, x_2, ..., x_{n_s})$ is a sample point and $f(x_i) = (f_0(x_i), f_1(x_i), ..., f_{n_f}(x_i))$ is the function value denoted as n_f-dimensional ($n_f D$) attribute at location x_i. An n-th order feature $f^n(x_i)$ at a sample point x_i is the n-th order derivative of f (estimated by the difference), e.g., $f^0(x_i) = f(x_i)$, $f^1(x_i) = (\frac{\partial f(x_i)}{\partial x_1}, \frac{\partial f(x_i)}{\partial x_2}, ... \frac{\partial f(x_i)}{\partial x_n})$. Image segmentation has the goal of partitioning a regular n_s-dimensional grid of sample points x_i into nonoverlapping connected sets denoted as regions $S_j = \{x_i\}$, where j is the index of the region. $n_s D$ regions are defined by homogeneity (similarity) of features within the region, and contrast (dissimilarity) with the surround. The homogeneity δ of one region S_j^δ is defined as the maximum distance between a pair of features at the sample points within the region S_j^δ; $\max\{\| f^n(x_i \in S_j^\delta) - f^n(x_k \in S_j^\delta) \|\} = \delta$. The contrast α of two neighboring regions S_1^δ and S_2^δ is defined as the minimum distance between a pair of features located across the region boundary; $\alpha = \min\{ \| f^n(x_i \in S_1^\delta) - f^n(x_k \in S_2^\delta) \| \}$.

Segmentation method: (1) Create regions $S_{x_i}^{2\delta} = \{x_k\}$ at each sample point x_i such that $\| f^n(x_i) - f^n(x_k) \| \leq \delta$. (2) Compare the pair of regions $S_{x_i}^{2\delta}$, $S_{x_l}^{2\delta}$ for every two adjacent sample points x_i, x_l. (3) Define final regions or boundaries based on the comparisons.

Noise robustness: Region descriptors (sample means) of $S_{x_i}^{2\delta}$, $S_{x_l}^{2\delta}$ are used in the comparison in Step 2 to achieve noise robustness. The descriptor of $S_{x_i}^{2\delta}$ is calculated as $D(S_{x_i}^{2\delta}) = \bar{\mu}_{x_i} = \frac{1}{M_{x_i}} \sum_{k=1}^{M_{x_i}} f^n(x_k \in S_{x_i}^{2\delta})$. The comparison in Step 2 and assignment of samples to different regions in Step 3 are performed based on the following inequality: If $\| D(S_{x_i}^{2\delta}) - D(S_{x_l}^{2\delta}) \| \leq \delta$ then x_i and x_l belong to the same final region else x_i and x_l are boundary points. This grouping rule is an outcome of similarity analysis. The similarity analysis relates the values of region homogeneity and contrast statistically. Two cases are considered: $\alpha > \delta$ and $\alpha \leq \delta$. The probability of obtaining correct segmentations is obtained analytically and numerically. If $\alpha > \delta$ then $Pr(error) = 0$. If $\alpha \leq \delta$ then $Pr(error) = 1 - Pr(\| \bar{\mu}_{x_i \in S_j^\delta} - \bar{\mu}_j \| \leq \delta$, where $\bar{\mu}_j$ is the sample mean of features from an unknown region S_j^δ.

Computational efficiency: The dimensionality of computations is reduced using a separability property. Efficient segmentation method using the separability property is performed in three steps: (a) divide $n_s D$ grid of sample points into several lower-dimensional grids, which leads to lower-dimensional images. (b) compute segmentations of lower-dimensional images and (c) assemble computed lower dimensional regions or boundaries into $n_s D$ regions or boundaries of regions. The use of the separability property reduces computational complexity but decreases noise robustness. Accuracy analysis of the method using descriptors and separability shows that if $\alpha > \delta$ then $Pr(error) = 0$, but if $\alpha \leq \delta$ then $Pr(error)$ is larger than for the method using only descriptors.

Multiscale segmentation: Image segmentations are obtained for different values of a homogeneity parameter δ. *A hierarchy of regions* is created such that every region S_j^δ obtained at scale δ cannot split at $\delta + \Delta\delta$ into smaller subregions (bottom-up constraint) and cannot merge at $\delta - \Delta\delta$ with other $n_s D$ regions (top-down constraint). The hierarchy is guaranteed by modifying feature values within created regions S_j^δ at each scale δ to the sample means of created regions. Three types of *tree representations* are defined from the set of regions composing the hierarchical organization.

3 Performance Evaluation

Segmentation quality is judged based on three main criteria: (a) accuracy, (b) time and memory requirements and (c) comparisons with other segmentation methods. The tradeoff between segmentation accuracy and computational requirements is as follows: the proposed segmentation using descriptors is 3 times more noise robust than the proposed segmentation using descriptors and separability but 77 times slower and 7 times more memory intensive. The experiments were conducted for the zeroth and first order region features ($n = 0, 1$) with the worst case of boundary error equal to 7.1%. The average segmentation time for the faster method was $0.4s/seg$ if $n = 0$ and $1.54s/seg$ if $n = 1$ for $100x100$ image ($n_s = 2, n_f = 1$) on Sparc 20 workstation. Three other segmentation methods were compared based on the number of misclassified sample points using synthetic images with regions characterized by $\alpha = 1$ and $\delta = 49$. The best performance was achieved by the proposed method using descriptors followed by segmentation based on Markov Random Field, proposed method using descriptors and separability, morphological segmentation and Canny edge detector. *Multiscale segmentations* and *tree quality* were evaluated experimentally.

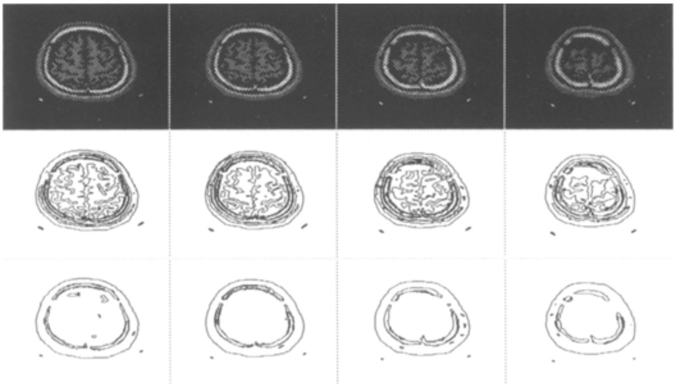

Fig. 1. Pyramid representation for 3D MR data.

Top row - original images along z axis (256x256), $n_f = 3, n_f = 1$. Cross section of pyramids at (1) fine scale δ_1 (second row from top), coarse scale δ_2 (bottom row, $\delta_1 < \delta_2$). Features with $n = 0$.

The proposed methods were used in applications involving the following types of real data: angiograms, botanical data, magnetic resonance images (see Figure 1), satellite images, range data and gray-scale and color images and video sequences.

4 Conclusions

We have presented a new framework for hierarchical image segmentation into homogeneous regions defined by constancy of the n-th order derivatives. We have formulated into the method a tradeoff between the segmentation accuracy and computational requirements. Comparative analysis and experiments with multidimensional multivariate real data were conducted to demonstrate the superb performance of the proposed methods.

References

1. N. Ahuja. A transform for detection of multiscale image structure. In *IEEE Conf. Comp. Vis. Pattern Recognition '93*, pages 780–781, New York, June 1993.
2. J. J. Koenderink. The structure of images. *Biol. Cybernetics*, 50:363–370, 1984.
3. T. Pavlidis and Y. Liow. Integrating region growing and edge detection. *IEEE Trans. on Pattern Anal. and Mach. Intel.*, 12(3):225–233, March 1990.
4. B. M. ter Haar Romeny (editor). *Geometry-driven diffusion in computer vision.* Kluwer Acamedic Publisher, 101 Philip Drive, Norwell, MA 02061, USA, 1994.
5. S. C. Zhu, T. S. Lee, and A. L. Yuille. Region competition: Unifying snakes, region growing, and Bayes/MDL for multi-band image segmentation. Technical Report CICS-P-454, Center For Intelligent Control Systems, MIT, March 1995.