Learning Based Interactive Image Segmentation

Bir Bhanu and Stephanie Fonder
Center for Research in Intelligent Systems
University of California, Riverside, California 92521
Email: {bhanu, steph}@cris.ucr.edu

Abstract

In this paper we present an approach to image segmentation in which user selected sets of examples and counter-examples supply information about the specific segmentation problem. Image segmentation is guided by a genetic algorithm which learns the appropriate subset and spatial combination of a collection of discriminating functions, associated with image features. The genetic algorithm encodes discriminating functions into a functional template representation, which can be applied to the input image to produce a candidate segmentation. The quality of each segmentation is evaluated within the genetic algorithm, by a comparison of two physics-based techniques for region growing and edge detection. Experimental results on real SAR imagery demonstrate that evolved segmentations are consistently better than segmentations derived from the Bayesian best single feature.

1 Introduction

Segmentation is a low-level process that is a first step to many computer vision tasks. The problem involves partitioning the image into several regions which are homogeneous within themselves and distinct from each other, according to some set of criteria. There exist a variety of approaches to image segmentation, including edge detection, region splitting, region merging and clustering. Each of these approaches suffers from sensitivity to parameters for thresholding, and/or termination conditions. Still other approaches combine a few of these methods in an attempt to gain the strength of more than one technique and overcome some of the weaknesses of each. However, whichever of these weaknesses causes the algorithm to fail, the underlying cause is the inability to specify how homogeneous a region should be and how distinct bordering regions should be in an application dependent manner. For the approach presented in this paper, application dependency is overcome by allowing the user to interactively train the segmentation tool for his/her application. The contributions of this research include: genetic learning of functional template design, physics-based segmentation evaluation, novel crossover operator and fitness function, as well as a system prototype and experiments on real synthetic aperture radar (SAR) imagery.

2 Technical Approach

In an interactive session, the user selects a set of examples and a set of counter-examples. The example and counter-example sets are then used to scale the data and create histograms, which are useful both to visualize and quantify the class separation for a variety of features. Based on histogram overlap, a set of discriminating functions is designed to perform discrimination between the example class and counter-example class.

A genetic algorithm encodes these functions into a functional template representation and produces a population of initial functional templates. These functional templates are applied to the input image to produce segmentations. The results of the segmentations are quantified by an evaluation process (the fitness function), and the population of functional templates are combined and modified via a set of operations based on genetic evolution, in an effort to evolve an optimized segmentation. The process of segmentation, evaluation, and recombination, is repeated for a given number of generations and the best result of the final generation is presented as output to the user.

Although the genetic algorithm evaluates candidate segmentations via a comparison to region-based and edge-based techniques, these are only used as a guide for the process of searching through possible segmentation outcomes produced by the combinations of the discriminating functions and their spatial arrangement.

299

Since the discrimination functions inherently contain classification information, it is possible to outperform the region-based and edge-based approaches that are used to evaluate the segmentation quality. Novel features of our approach are given below.

Functional Template Design - Templates are commonly used in computer vision and pattern recognition for image enhancement, image segmentation and image recognition. Unlike these traditional templates, functional templates, where each element of the template is an index to a function, have also been used recently for segmentation and classification [1,2]. However, the prior manually designed functional templates could take months of effort. Previous work combined features in prespecified ways, resulting in a cumbersome design process and a single discriminating function based on the Bayesian best features, was used in every position of the functional template. Unlike the previous work, we automate the design of the functional template. In this research discriminating functions based on the features are combined into the functional template using a genetic algorithm. In addition to the selection of a subset of features, the genetic algorithm determines the spatial placement of functions within the template. This means that the combinations of features are not limited to prespecified combinations. Furthermore, the approach does not rely on the Bayesian best feature. The Bayesian performance is calculated solely on example and counter-example sets, which may not characterize their respective classes entirely.

The design of the functional template of a given size requires a solution of the combinatorics problem. For example, for 20 functions, the size of the search space for a 3x3 template is 512 billion. We use GAs (genetic algorithms) as function optimizers since they allow the possibility of achieving the global maximum without exhaustive search.

Physics-Based Segmentation Evaluation - Segmentation evaluation, during the learning process, is performed by a comparison of two segmentation techniques: edge-base and region-based. These techniques incorporate SAR-specific information to produce a segmentation. The physics-based algorithms are performed on an image that has been denoised. Denoising was performed using wavelets, which has been shown to significantly enhance a SAR image [3]. A region is grown from the set of examples, and the edges in the image are detected.

Both of these techniques use a log likelihood ratio test. The distributions used in the test are specifically developed for SAR imagery. The physics-based results of region growing and edge-detection are used within

the genetic algorithm for segmentation evaluation. A candidate segmentation is evaluated quantitatively by comparison to the physics-based results, in a novel fitness function.

Improved Crossover Operator and Novel Fitness Function- Every genetic algorithm has a crossover operator. In this work, the crossover operator is an extension of the operator presented in [4]. That operator preserves 2D spatial information by exchanging the information from identical rectangular areas of the parent functional templates. The size and position of the rectangular area to be crossed over is chosen randomly each time the operator is executed. Although the prior operator's templates preserve the 2D spatial information, it favors the exchange of center element. Such a crossover operator would favor spatial information near the center, over such information near the borders of the functional template.

In this work, the crossover operator is designed to remove this bias. This is accomplished by allowing the rectangle used in crossover to (conceptually) wrap around the template in both the horizontal and vertical directions. The result is an operator which allows unbiased evolution of spatial information.

The novel fitness function, compares a candidate segmentation to portions of the physics-based region and edge estimates. The fitness function is the average of two terms: a region term and an edge term. The region term encourages a segmentation to correctly classify pixels within the region from which the examples are selected. The edge term encourages regions classified as example regions to have edges coinciding with image edges.

3 Experimental Results

A prototype of the system has been developed and tested on synthetic and SAR imagery. Segmentations produced with the system are compared with three default single feature templates. The first default template contains the Bayesian best feature. The second default template maximizes the percent of the pixels classified correctly (PCC). The third default template maximizes a normalized version of PCC, which averages normalized example accuracy and normalized counter-example accuracy (NPCC).

An example experiment for paved road vs. grass is shown in Figures 1-3. The original SAR image (with the users examples and counterexamples), results of denoising, region growing, and edge detection, as well as ground truth (for evaluation of results only) are shown in Figure 1. We are learning a 3x3 template. The elements of a template contain the function number. We

have used 23 functions of intensity, local mean, local standard deviation and Gabor wavelet filters of various scales and orientations. As shown in Figure 2 the single function corresponding to the best (primitive) feature which gave the lowest Bayesian error (pixel classification) was #5 (7x7 local mean). The best function corresponding to the single feature template was based on feature #0 (image intensity). EA is the % of example accuracy, CA is the % of counter-example accuracy. PCC is probability of correct pixel classification. NPCC is normalized probability of correct classification, it takes into account the unequal number of example and counter-example pixels.

The fitness function used to evaluate the quality of image segmentation is obtained by a comparison of physics-based edge detection and region growing to compute fitness of an individual (template). A region is grown from a set of examples and the edges in the image are detected. Both edge detection and region growing use a log likelihood ratio test and SAR specific distributions. The candidate segmentation of the example region is compared with (1) region growing results to obtain the normalized Region Term in the example, and (2) it is also compared with the edges that have been obtained to get an evaluation based on normalized edge-border coincidence (Edge Term). The fitness function is one-fourth of these two evaluations. GAs have a population of 100, tournament size of 10%, cross-over rate of 25%, mutation rate of 1% and they were evolved for 10 generations. GA results correspond to the result with the highest fitness at the end of 10 generations. The last column in Table 2 shows the average results with 10 different random seeds. Note that the GA results are better than the Bayesian or the default template based on a single feature. The learned template used functions based on features (0, 12, 17, 18 and 22).

Region Growing/Edge Detection - The region growing and edge detection results are shown in Fig. 1. Region growing found just over half of the road, as the region only grew downward. Edge detection found the edges of the road particularly well, as well as some texture edges in the grass class. Although a couple of the extra texture edges were included in the fitness term, the majority were excluded.

Segmentation Quality - The evolved segmentation results perform better than default results, as shown in Figure 2, for both the region term and the edge term. It is not surprising then, that evolved results have significantly better example accuracy performance and have nearly perfect counter-example accuracy, which is consistent with the defaults. Thus, the evolved results are significantly better than default
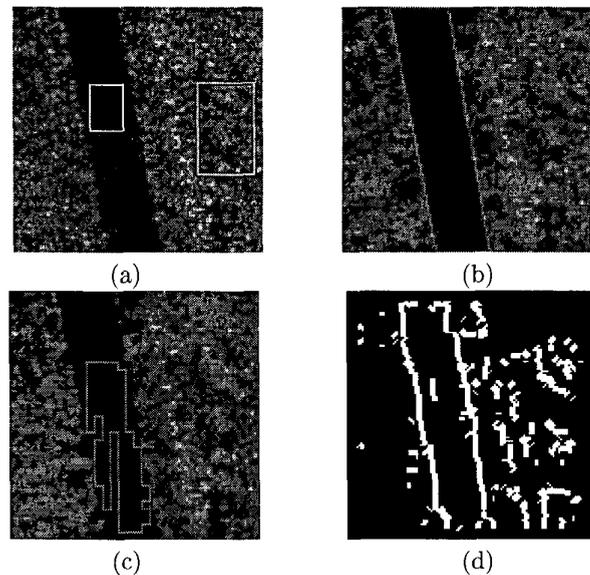
Figure 1. Paved Road vs. Grass: (a) Original Image with Examples (yellow/blue) and Counter-examples (green/purple) (b) Denoised Image with Ground Truth (red) (c) Region Growing (Red) and (d) Edge Results.

segmentations according to the NPCC. (The PCC does not show this improvement as dramatically since example pixels represent only 23% of the image.)

Template Design - The intensity feature was the most common in the evolved results in this experiment. This is not surprising as it was the selected feature for both NPCC and PCC defaults. However, most of the example area can be classified correctly with an average of four instances of this feature, leaving the remaining template positions for other features to improve performance. Boundary accuracy and edge term improve with the remaining functions, typically 90° and 135° orientations of the small scale (11x11) Gabor standard deviation feature (functions 17 and 22). These features make intuitive sense for edge performance in this image as they are aligned with the orientation of the road.

Evolution - The evolution of segmentation quality is shown in Figure 3. The region term is optimized in the first few generations by the intensity function. Later generations fine tune edge performance further with the selection of appropriate Gabor standard deviation features. Meanwhile, the region term levels off. The fitness function is effectively improving the

301

| Bayesian Default | NPCC/PCC Default | GA Result |

5	5	5
5	5	5
5	5	5

0	0	0
0	0	0
0	0	0

0	0	17
17	0	18
22	22	12
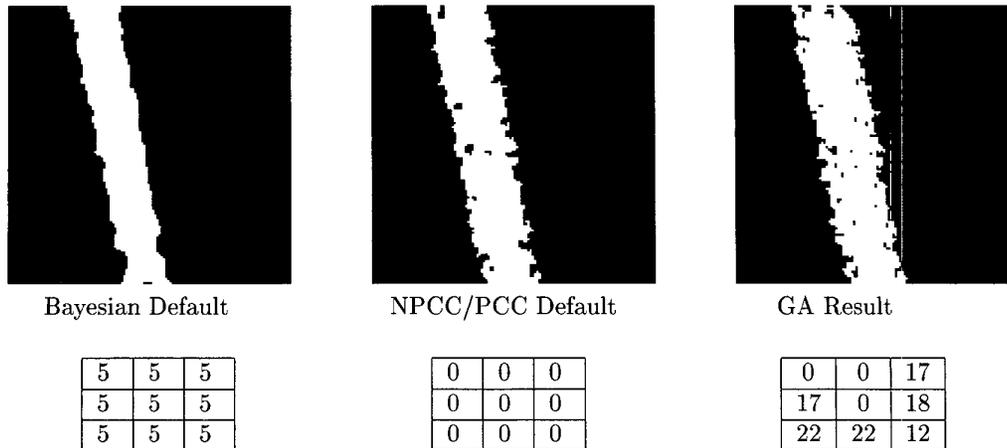
	Bayes	NPCC/PCC	GA Result	Ave. (10 seeds)
Region Term	0.861	0.953	0.970	0.963
Edge Term	0.099	0.513	0.599	0.554
Fitness	0.240	0.367	0.393	0.380
EA	0.672	0.809	0.846	0.827
CA	1.000	1.000	0.999	0.999
PCC	0.923	0.955	0.963	0.958
NPCC	0.836	0.904	0.923	0.914

Figure 2. Paved Road vs. Grass Results
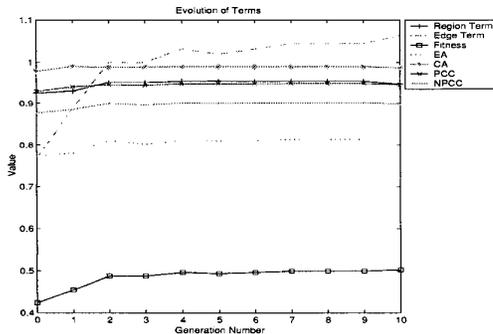
Figure 3. Evolution of Segmentation Quality.

example accuracy term and counter-example accuracy remains quite constant throughout. Thus both PCC and NPCC are increasing, although the improvement of NPCC is more significant.

4 Conclusions

Our research on many real SAR examples, like the one illustrated here, shows that the system can efficiently learn the functional template and that the seg-

mentation results are superior to a template designed using a single best Bayesian feature.

References

1. J.G. Verly and R.T. Lacoss, "Automatic target recognition of LADAR imagery using functional templates derived from 3-D CAD models. In Reconnaissance, Surveillance, and Target Acquisition for Unmanned Ground Vehicle. O Firschein and T.M. Strat (Ed.), pp. 195-218, Morgan Kaufmann Publishers, Inc.

2. R.L. Delanoy and S.W. Troxel, "Toolkit for image mining: user trainable search tools, *The Lincoln Lab Journal*, 8(2), 145-160, 1995.

3. J. E. Odegard, "Moments, smoothness and optimization of wavelet systems, " Ph.D. Thesis, Rice University, 1996.

4. A. J. Katz and P.R. Thrift, Generating image filters for target recognition, *IEEE Trans. on PAMI*, 16(9), 906-910, 1994.