TEXTURE-BASED INFRARED IMAGE SEGMENTATION BY COMBINED MERGING AND PARTITIONING

W. Brendan Blanton and Kenneth E. Barner

University of Delaware, Electrical and Computer Engineering, Newark, Delaware

ABSTRACT

This paper describes a method of image segmentation using recursive splitting and merging based on texture similarity measures. This technique addresses the problem of segmenting image regions of varying texture with limited intensity edges. The proposed technique provides a framework for texture based image segmentation that is shown to be applicable across a wide variety of image content. The primary motivation for this work is the segmentation of infrared images. Infrared imagery is characterized by narrow histograms corresponding to ambient scene temperature. Results illustrate that using texture signatures for infrared imagery yields enhanced segmentation performance over luminance features. Additional benefit for infrared imagery and better generality to other images types is obtained when luminance and texture are both applied to the segmentation criteria. A method for the quantitative comparison of segmentation results is presented and benchmarks are provided against serval recent segmentation algorithms.

Index Terms—
Image Segmentation, infrared imaging, wavelet transforms

1. INTRODUCTION

Despite the volumes of research on image segmentation, it continues to be a challenging problem in image processing. Segmentation is the process of dividing an image into a subset of connected regions based on a application defined criteria. For natural scenery images, this usually means the separation of the image into its constituent objects or feature types (e.g. sky, trees, buildings, etc). In the case of infrared sensors, an image is produced by mapping thermal emissivity to intensity levels. This paper overviews a technique applicable to the segmentation of infrared imagery; however, we illustrate its general applicability to a wide variety of image types.

Infrared imagery often is characterized by a narrow histogram concentrated around ambient temperature in which detector and atmospheric noise are often present. These factors result in a difficulties when a luminance and luminance edges are applied directly to segmentation. We illustrate that a combined texture and luminance based approach, using a wavelet-based watershed, results in improvements in initial segmentation. We propose a combination of splitting and merging to reduce the over segmentation of the watershed and capture texture edges. We also apply a unique cost function for automatic merging termination.

1.1. Related Works

Three main classes of methods for texture based segmentation have emerged in recent years: model based, global cuts/ graph partitioning, and split-merge based methods. The model based methods strive to capture the underlying structure of the texture and classify image pixels in to different segmented regions. A variety of different models are used for this method including Hidden Markov Models [1], anisotropic diffusion models [2], and general statistical models [3] based on filter responses. The global cut method for image segmentation uses a criteria to find the lowest cost cut in an image [4]. Similar to the model based approach, the underlying image structure is ignored in the graph partitioning approach. Also, once a cut is made it is used in the final segmentation, often leading to over-segmentation. Finally, there are various methods of image segmentation that are so called top down splitting approaches. Many of these methods are based on the popular watershed transform [5]. The watershed is popular because it easily produces connected regions. The drawback of these methods is the potential for over segmentation.

1.2. Paper Organization

Section 2 contains an overview and details of the proposed algorithm. Sections 3 and 4 provide implementation details for the selection of texture features and termination criteria respectively. Section 5 contains quantitative results of the proposed method compared with three recent segmentation algorithms. Quantitative segmentation performance is presented for grayscale and infrared imagery . Section 6 contains a general discussion of results.

2. PROPOSED METHOD

This section contains an overview of the proposed algorithm depicted in Fig. 1. A edge enhancing noise reduction filter is applied as a preprocessing step to remove gradient noise and enhance weak edges. The initial segmented edge image is extracted using a combination luminance and wavelet based gradient method. A watershed is then applied to the edge image to obtain an initial connected region segmentation map. Iterative cuts are formed from this segmentation image to capture remaining texture edges that may have been lost in the initial watershed. Iterative merging based on a similar texture and luminance criteria is then applied to reduce over segmentation.

2.1. Watershed Preprocessing for Edge Enhancement

As a first stage of the segmentation problem it is often necessary to enhance the main edges of the image while rejecting much of the gradient noise that leads to over segmentation. In this regard, a lower-upper-middle (LUM) filter was chosen [6]. The LUM filter has several advantages including its ability to perform edge enhancing and noise rejection simultaneously.

2.2. Edge Image Computation and Watershed

The proper computation of the edge image is essential to obtaining acceptable performance during the merging. The initial partition-
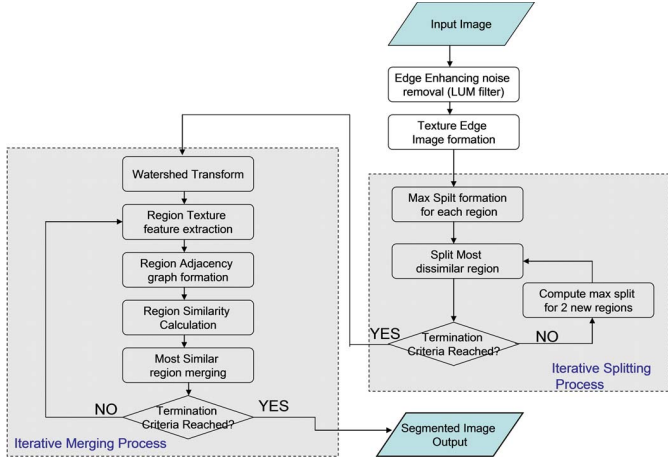
Fig. 1. Block Diagram of Proposed Segmentation Method

Fig. 2. Texture performance measure comparisons

ing of the image into regions is accomplished using the watershed transform as proposed by Vincent and Soile [5]. In our case, the boundaries of these water sheds (or catchment basins) are formed based on the wavelet texture gradient computation as formulated by Mallat [7].

2.3. Wavelet Transform

Various methods have been employed for texture analysis. These techniques include frequency, model-based, statistical, and geometric approaches to texture feature extraction. The wavelet approach is a very desirable form to extract texture features since it contains separated frequency, scale and direction information [8]. The undecimated wavelet transform was chosen for this application since it allows a direct feature extraction without scaling. Since all subbands share a common resolution, they can be used directly without constant resolution changes for each region. Several wavelets were used for the transform to compare the relative performance of the texture extraction. Consistent with [9] the results did not vary significantly with the wavelet used.

2.4. Region Adjacency and Similarity Merging

An initial region adjacency graph (RAG) is formed that provides a table of all regions that share a border in the initial watershed segmentation by using the method outlined by Haris *et al.* [9]. For each entry in the RAG a dissimilarity measure was formed

$$\delta(R_g, R_h) = \frac{\|R_g\| \cdot \|R_g\|}{\|R_g\| + \|R_g\|} \cdot \sum |M(R_g) - M(R_h)| \quad (1)$$

where, $\delta(R_g, R_h)$ is the dissimilarity between regions R_g and R_h, $M(R_h)$ is the texture measure for R_h, and $\|R_h\|$ is the number of pixels in region R_h.

The RAG is scanned for the minimum dissimilarity (δ) and the two regions corresponding to lowest dissimilarity are combined.

2.5. Region Partitioning

The process of region splitting can be thought of as the inverse of the iterative merging process. Each region in the image is analyzed for the cut that maximizes the dissimilarity measure provided in section
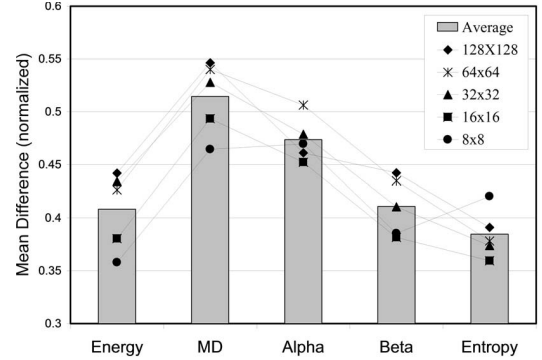
3. A cut is a segmentation of the region into two distinct regions R_g and R_h such that $R_g \cap R_h = \emptyset$ and $R_g \cup R_h = R^*$ where R^* is the original image region.

The maximum cut is determined for each region in the image. The global maximum cut for all regions is selected. The regions that correspond to this cut are segmented. Next, new cuts are determined for these two new regions. The global maximum cut is then selected. This process of cutting repeats iteratively for a specified number of iterations or until some minimum dissimilarity is reached.

3. TEXTURE FEATURES FOR SEGMENTATION

Following the initial watershed segmentation, Iterative merging of adjacent regions is applied based on a similarity measure. Several texture-based similarity measures were chosen for consideration. This selection was primarily based on the observation that the initial segmented image contains many small regions. Many texture methods rely on accumulating statistical feature over a larger area and provide unstable texture measures as the region size is reduced. In addition, the methods chosen must be readily applied to irregularly shaped regions. The texture measures chosen were wavelet sub-band energy, wavelet sub-band mean difference, wavelet sub-band histogram signatures, and Renji entropy [10] [11]. All of these measures can be applied to an irregularly shaped region and are potentially scale and rotation invariant.

The evaluation of these texture measures was performed on a typical texture test image from the Brodatz texture library containing 16 textures [12]. The texture feature for each of the test image regions was formed and normalized. A difference matrix is formed showing the difference in the texture measure between each test region. The mean values of the difference graphs are then extracted for each measure to provide an overall dissimilarity performance measure. Fig. 2 shows these measures which were produced with several sizes to obtain a determination of robustness across region area. The Mean difference was selected, because it provides a stable texture difference measure across region size. This is crucial in the merging process, since the regions start out with small watershed segmented regions and combine to form larger homogenous texture regions. In addition to the Brodatz test pattern, a similar texture test pattern was extracted from various infrared images. The results again indicated that the mean difference is the preferred discriminating measure across a wide variety of region sizes.

4. MERGING TERMINATION CRITERIA

A typical problem in iterative based segmentation techniques is the determination of the appropriate iteration to terminate the merging process. This is often performed manually for a given image set by running a series of experiments. We have developed a method to automatically terminate merging based on a cost function consisting of region number and edge coincidence.

Edge-Border coincidence is a measure of the overlap of the final merged region borders to the edges found by an edge operator. The coincidence measure is determined as follows: Let E be the set of pixels extracted from the edge operator after thresholding. E is a binary edge image. Let S be the binary image that contains the edges from the segmentation and merging procedure. Then the edge-border coincidence (C_{EB}) is given by:

$$C_{EB} = 1 - \frac{n(S \cap E)}{n(E)} \qquad (2)$$

$n(A)$ is the number of elements of the set A.

Boundary Consistency/Edge Penalty is a performance measure similar to the edge-border coincidence except that the region borders that do not coincide with edges are used to penalize the segmentation quality:

$$C_{BC} = 1 - \frac{n(S \cap \hat{E})}{n(E)} \qquad (3)$$

Finally, the a simple cost is formed as the number of regions. This cost decreases linearly with decreasing number of regions in the merged image. This region cost, C_R, is computed as the ratio of the number of segmented regions to the total number of regions in the initial segmentation.

The merging costs functions described are combined to form a single termination cost, C_T, as a simple linear combination given by

$$C_T = C_R + C_{BC} + C_{EB} \qquad (4)$$

Merging is terminated when the cost reaches a minimum value. The termination criteria was tested against a variety of images and will be represented in the examples shown in later experimental results.

5. SIMULATION PERFORMANCE RESULTS

The performance of segmentation algorithms are often difficult to characterize due to the subjective nature of the results. This section provides a quantitative description of the segmentation performance based on the Berkeley Segmentation Database [13]. Martin *et al.* developed a framework in which segmentation performance can be quantified and benchmarked across many methods. They assert that a generalized approach to ground truth for a segmentation result is that developed by a human subject.

5.1. Segmentation Performance measures

In addition to the library of segmentation ground truth, Martin *et al.* also suggested two measure for segmentation performance called the Global Consistency Error (GCE) and Local Consistency Error (LCE). These measures seek to determine the consistency between regions formed in the benchmark set and the results of the segmentation routines. Both measures start by forming a set of error images,

$$E(S_1, S_2, p_{ij}) = \frac{\| R(S_1, p_{ij}) \setminus R(S_2, p_{ij}) \|}{\| R(S_1, p_{ij}) \|} \qquad (5)$$

where S_1 and S_2 are the segmentations to be compared, p_{ij} the pixel location corresponding to i and j the row and columns indices, and R the segmented region. The GCE is formed by the global minimum of the error image sum scaled by the number of pixels,

$$GCE(S_1, S_2) = \frac{1}{i \cdot j} min \left\{ \sum_{ij} E(S_1, S_2, p_{ij}), \sum_{ij} E(S_2, S_1, p_{ij}) \right\} \qquad (6)$$

while the LCE is sum of the pixel-wise minimum of the error images scaled by the number of pixels.

$$LCE(S_1, S_2) = \frac{1}{i \cdot j} \sum_{ij} min \left\{ E(S_1, S_2, p_{ij}), E(S_2, S_1, p_{ij}) \right\} \qquad (7)$$

Although GCE and LCE provide measures of the overall region matching between the intended and test segmentation; they do not provide a direct measure of the accuracy of the region boundaries or the amount of erroneous edge content. This factor is important in many applications thus two additional measures are introduced; Edge Detection Rate

$$E_D(S_1, S_2) = \frac{\| E_I(S_1) \cap E_I(S_2) \|}{\| E_I(S_1) \|} \qquad (8)$$

and edge false alarm rate.

$$E_{FA}(S_1, S_2) = \frac{\| \overline{E_I(S_1)} \cap E_I(S_2) \|}{\| E_I(S_2) \|} \qquad (9)$$

5.2. Segmentation Performance Comparison

The quantitative segmentation performance has a context only in comparative benchmarks of the same image with different segmentation schemes. Three methods were chosen for comparison with the proposed method: Shi and Malik's [14] method of ratio cuts, the local variation segmentation of Felzenszwalb and Huttenlocher [15], and the mean shift based segmentation method of Comanicu and Meer [16]. A cross section of images were chosen for testing that contained high, moderate and low scores in terms of the provided benchmarks. In order to summarize the results, the the four quantitative performance measures are determined and normalized for each image and averaged; resulting in a single number for each measure (1=best, 0=worst) summarized in table 1.

Table 1. Berkeley Data Set Normalized Performance

Measure	Proposed	Ratio Cut	Local Var.	Mean Shift
GCE	0.796	0.337	0.480	0.387
LCE	0.524	0.125	0.581	0.770
E_D	0.885	0.021	0.259	0.836
E_{FA}	0.953	0.172	0.270	0.605

Unfortunately, the Berkeley dataset does not contain reference infrared images for benchmarking, so set of hand labeled infrared imagery was created by human subjects in a fashion similar to the Berkeley dataset. Fig. 3 shows the images used, their hand segmentations, and corresponding segmentations applying the four methods. Table 2 provides the normalized results for the infrared images. As expected, the results achieved on infrared imagery are superior since the proposed method has been optimized for infrared imagery. The proposed technique far exceeds all others in every measure with the exception of the mean shift in edge detection performance. The high
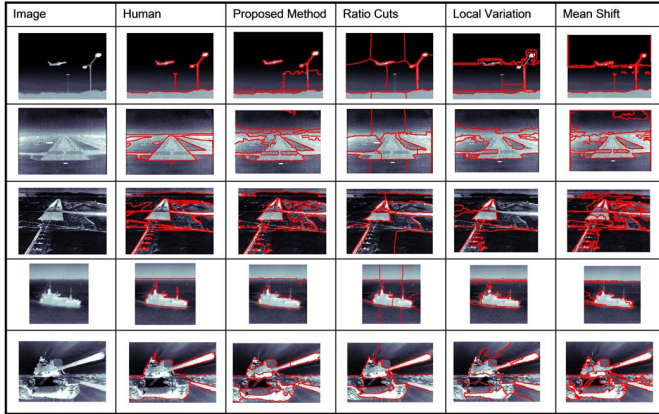
Fig. 3. Infrared Image Segmentation Results

Table 2. Infrared Normalized Performance

Measure	Proposed	Ratio Cut	Local Var.	Mean Shift
GCE	0.762	0.478	0.431	0.328
LCE	0.801	0.391	0.322	0.486
E_D	0.816	0.218	0.017	0.948
E_{FA}	0.993	0.559	0.201	0.247

performance in edge detection in the mean shift is largely due to the many superfluously edges created in the segmentation; thus its likewise poor performance in edge false alarm rate.

6. CONCLUSIONS

The method of image segmentation presented is based applying texture based watersheds, splitting, and merging. In particular, a texture based method is chosen to compensate for the often poor areas of contrast in infrared imagery. Experimental results when these techniques are applied to a variety of imagery are given. In particular, test patterns from the Berkeley segmentation database where tested against three recently developed segmentation algorithms. The quantitative results based on four segmentation performance measures showed that the proposed method resulted in greater performance. In addition, when the infrared imagery was tested using the same method the results were even more disparate with the proposed algorithm clearly outperforming the three test methods. It should also be noted that during testing the other algorithms had to be executed numerous times to obtain valid results. This is due to the fact that there is no automatic stopping criteria for the other methods. In addition, the other methods tested had numerous parameters that needed to be adjusted for the different image types to obtain a result that was comparable in terms of the number of segmented regions. The proposed algorithm not only outperformed these methods, but obtained all parameters directly from the input imagery. Also, the automatic stopping criteria used resulted in overall good segmentation outputs for a wide variety of input imagery. The combination of the unsupervised nature and high segmentation performance across a wide variety of image types show the desirability of this algorithm not only to infrared imagery, but to the general class of segmentation applications.

7. REFERENCES

[1] G. Fan and X.-G. Xia, *Nonlinear Signal and Image Processing*, K. E. Barner and G. R. Arce, Eds. CRC Press, 2004.

[2] S. Manay and A. Yezzi, "Anti-geometric diffusion for adaptive thresholding and fast segmentation," *IEEE Transactions on Image Processing*, vol. 12, no. 11, pp. 1310–1323, November 2003.

[3] M. Heiler and C. Schnorr, "Natural image statistics for natural image segmentation," *Proceedings of the Ninth IEEE International Conference on Computer Vision (ICCV03)*, pp. 1–8, 2003.

[4] S. Wang and J. M. Siskind, "Image segementation with ratio cut," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 25, no. 6, pp. 675–690, June 2003.

[5] L. Vincent and P. Soille, "Watersheds in digital spaces: An efficient algorithm based on immersion simulations," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 13, no. 6, pp. 583–598, June 1991.

[6] R. C. Hardie and C. G. Boncelet, "Lum filters: A class or rank-order-based filters for smoothing and sharpening," *IEEE Transactions on Signal Processing*, vol. 41, no. 3, pp. 1061–1076, 1993.

[7] S. G. Mallat and S. Zhong, "Characterization of signals from multiscale edges," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 14, pp. 710–732, 1992.

[8] S. G. Mallat, "A theory for multiresolution signal decomposition: The wavelet representation," *IEEE Trans. On Pattern Anal. And Machine Intel*, vol. 11, no. 7, pp. 674–693, July 1989.

[9] K. Haris, S. N. Efstratiadis, N. Maglaveras, and A. K. Katsaggelos, "Hybrid image segmentation using watersheds and fast region merging," *IEEE Transactions on Image Processing*, vol. 7, no. 12, pp. 1684–1699, December 1998.

[10] S. Grigorescu and N. Petkov, "Texture analysis using rnyi's generalized entropies," *Proc. IEEE Int. Conf. on Image Processing (ICIP)*, pp. 241–244, 2003.

[11] G. V. de Wouwer, P. Scheunders, and D. V. Dyck, "Statistical texture representations for discrete wavelet representations," *IEEE Transactions on Image Processing*, vol. 8, no. 4, pp. 592–598, 1999.

[12] P. Brodatz, *Textures: A Photographic Album for Artists and Designers*. Dover Publications, 1966.

[13] D. Martin, C. Fowlkes, D. Tal, and J. Malik, "A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics," *IEEE Proc. 8th International Conference on Computer Vision*, vol. 2, pp. 416–423, July 2001.

[14] J. Shi and J. MalikGrau, "Normalized cuts and image segmentation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 8, August 2000.

[15] P. F. Felzenszwalb and D. P. Huttenlocher, "Image segmentation using local variation," *IEEE Proceedings on Computer Vision and Pattern Recognition*, June 1998.

[16] D. Comanicu and P. MeerGrau, "Mean shift: A robust approach toward feature space analysis," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, pp. 603–619, May 2002.