ANISOTROPIC DIFFUSION PYRAMIDS FOR IMAGE SEGMENTATION

Scott T. Acton[1], *Alan C. Bovik*[2], *and Melba M. Crawford*[3]

[1]School of Electrical & Computer Engineering
Oklahoma State University
Stillwater, Oklahoma 74078-0321
[2]Laboratory for Vision Systems
Center for Vision and Image Science
Department of Electrical & Computer Engineering
[3]Center for Space Research and
Department of Mechanical Engineering
The University of Texas at Austin
Austin, Texas 78712

ABSTRACT

We introduce the Anisotropic Diffusion Pyramid (ADP), a structure for multiresolution image processing. We also develop the ADP for use in region-based segmentation. The pyramid is constructed using the anisotropic diffusion equations, creating an efficient scale-space representation. Segmentation is accomplished using pyramid node linking. Since anisotropic diffusion preserves edge localization as the scale is increased, the region boundaries in the coarse-to-fine ADP segmentation are accurately delineated. An application to segmentation of remotely sensed data is provided. The results of ADP segmentation are compared to Gaussian-based pyramidal segmentation. The examples show that the ADP has a superior ability to subdivide the image into integral groupings, minimizing the error in boundary localization and in pixel intensity.

1. INTRODUCTION

The aim of image segmentation is the subdivision of an image into homogeneous regions. After segmentation, the boundaries of the constituent regions may be identified, enabling classification and extraction of each aggregated region. A multiresolution or multi-scale approach to image segmentation permits the creation of homogeneous image regions at different scales. A high resolution segmentation will identify detailed, relatively small features where a low resolution segmentation produces coarser, larger features. Pyramid-based algorithms provide a economical vehicle for multiresolution segmentation. In an image pyramid, the base (or lowest level) is the original high resolution image; successive levels consist of image representations of decreasing resolution and increasing scale. Gaussian Pyramids (GPs) have been widely used in image segmentation [1]. With the GP, the initial pyramid levels are computed by applying a linear filter (Gaussian or average) to the previous level and then subsampling. For example, a 2x2 grid of pixels may be averaged and used to compute the corresponding pixel at the next higher level. As the pyramid progresses to higher levels, Gaussian smoothing is implemented and coarse features emerge as coherent regions. Unfortunately, the Gaussian smoothing distorts region boundaries through inter-region blurring. Furthermore, this region bleeding destroys the individual image region integrity by corrupting the pixel intensity (radiometric) information within each region interior.

This paper describes a multiresolution segmentation algorithm based on the *anisotropic diffusion pyramid* (ADP). The ADP is afforded by the nonlinear smoothing of anisotropic diffusion. In contrast to Gaussian smoothing which produces inter-region smoothing, anisotropic diffusion yields intra-region smoothing. The resulting pyramid preserves both boundary and radiometric information with increasing scale.

2. BACKGROUND: ANISOTROPIC DIFFUSION

A scale-space consists of a family of image descriptions that vary from fine, detailed representations to coarse representations. Perona and Malik showed that a scale-space can be represented by a progression of images computed by the heat diffusion equation [2]. The heat diffusion equation for the pixel at location (i,j) of an image $\mathbf{S}$ at iteration t is

$$S_{i,j,t} = div\left(c_{i,j,t}\nabla \mathbf{S}\right) \tag{1}$$

where ∇ is the gradient operator, *div* is the divergence operator ($div\ \mathrm{x} = \nabla\bullet\mathrm{x}$), and $c_{i,j,t}$ is the heat diffusion coefficient at image location (i,j) and iteration t. If $c_{i,j,t}$ is made constant at all image locations, the diffusion equation yields isotropic Gaussian smoothing. In an anisotropic diffusion process, the diffusion coefficient is allowed to vary according to the magnitude of the local image gradient. A discrete version of the anisotropic diffusion equation is given in [2] as follows:

0-8186-6950-0/94 $4.00 © 1994 IEEE

$$S_{i,j,t+1} = S_{i,j,t} + \lambda(c_N\nabla_N + c_S\nabla_S + c_E\nabla_E + c_W\nabla_W) \quad (2)$$

where c and ∇ are the diffusion coefficient and simple difference operator, respectively. There are four diffusion coefficients in (2) corresponding to the four (north, south, east, west) directions of diffusion at location (i,j). For example, the diffusion coefficient for the "northern" orientation is computed using $c_N = \exp[-(\nabla_{N/k})^2]$ where $\nabla_N = S_{i-1,j} - S_{i,j}$. The diffusion coefficients discourage inter-region bleeding by inhibiting neighborhood smoothing where the local image gradient is large and a region boundary is present. The parameter λ controls the rate of smoothing, and the choice of λ is addressed in [3]. In anisotropic diffusion, the scale (size) of features obtained may be increased by increasing the number of diffusions (iterations of eq. 2) as in [4] or by increasing the value of k, the scale-space parameter, as in [2].

3. ADP THEORY AND IMPLEMENTATION

3.1. ADP Structure

For an original image of dimensions NxN, the ADP contains $\log_2 N + 1$ levels. Level 0 is the original image. Level R (where $1 \le R \le \log_2 N + 1$) is the selected root level for segmentation. Each node of the original image is linked to a node of the root level. Thus, the size of the root level dictates the number of final partitions in the segmented image. For example, a 4x4 root level will produce a segmentation into at most 16 partitions. The segmentation is accomplished by first creating the initial ADP through successive anisotropic diffusion of the original image and then performing the iterative process of pyramid node linking with respect to a given root level R.

With each ascending level in the pyramid, the size is halved in each dimension. Thus, the bottom level (the original image) has NxN nodes and the final level contains one node. Node positions are denoted by (i,j) and l is the pyramid level. A node on level $l + 1$ linked to a node on level l is called the *father* of the node on level l. Similarly, a node on level $l - 1$ linked to a node on level l is called a *son* of the node on level l. For each node, five variables are maintained. The node value $p_{i,j,l}$ represents the specific image property used in segmentation. In this study, only scalar pixel intensities are used for segmentation. The variable $r_{i,j,l}$ represents the value of the node on the root level R that is linked to the node at location (i,j) on level l. The root values of the nodes on level 0 provide the final segmentation result. $s_{i,j,l}$ gives the *support* for a given node which is computed by finding the number of sons linked to the node on level $l - 1$. The variables $x_{i,j,l}$ and $y_{i,j,l}$ provide the coordinates of a node's father on level $l + 1$.

3.2. Pyramid Initialization

We present two methods for initializing the ADP. The basic method uses a single diffusion iteration to create pyramid level $l + 1$ from pyramid level l. It is therefore called a Single Diffusion ADP (ADP-SD). The advantage of the ADP-SD is the low computational expense. Since ADP-SD enacts only one diffusion iteration to create the subsequent pyramid level, intermediate anisotropic diffusion results do not need to be stored for the previous level. Thus, with the ADP-SD, subsampling can be accomplished at the same time as the computation of anisotropic diffusion; the anisotropic diffusion equation is applied at only one of four pixel locations at level l to produce level $l + 1$. When using the ADP-SD, the feature scale in segmentation is controlled by k, since the number of diffusions is fixed.

For greater control of feature scale and region boundaries, a Multiple Diffusion ADP (ADP-MD) may be employed. The ADP-MD permits multiple iterations of anisotropic diffusion at level l to compute the node values of level $l + 1$.

3.2.1. Single Diffusion ADP

The ADP-SD pyramid is initialized by creating the node values with the anisotropic diffusion update rule as follows:

$$p_{i,j,l+1} = p_{2i,2j,l} + \lambda(c_N\nabla_N + c_S\nabla_S + c_E\nabla_E + c_W\nabla_W). \quad (3)$$

The north, south, east, and west components of (3) are computed as in (2), using the node values from level l in the neighborhood of the node at $(2i,2j)$. Implementation of the ADP-SD initialization of (3) requires less than $(1/3)N^2$ anisotropic diffusion updates to create the entire initial pyramid. The computational complexity of the ADP-SD is equivalent to that of the GP.

For both ADP-SD and ADP-MD, the node father locations are initialized to $x_{i,j,l} = i/2$ and $y_{i,j,l} = j/2$. Hence, each father node will have four sons initially. The root values are set equal to the node values at the start.

3.2.2. Multiple Diffusion ADP

Alternatively, the node values may be computed by successively applying (2) at level l of the pyramid and then subsampling (by a factor of two along the image rows and columns), to compute the node values of level $l + 1$. If n diffusions per level are used to compute the initial pyramid, the ADP-MD method requires $4n$ times the number of updates required by ADP-SD. The images and tabulated results in Section 4 show the improvement in segmentation quality gained by ADP-MD.

3.3 Pyramid Node Linking

The linking process consists of three steps: linking, recomputing node values, and propagating root values. First each node is linked to the father of its four possible fathers that has the root value closest to its node value. This linking is conducted for levels 0 to R- 1.

Then, the node values on levels 1 to R may be recomputed based on the value of a node's sons. In this study, we have investigated three different variations of the update process. The first rule enacts the standard technique of assigning the node's value to the weighted average of a node's sons. This is the update method traditionally used in the GP [1]. With the average update rule, radiometric (intensity) information is largely distorted and step edges are not retained after several updates. The second approach to node update uses the weighted median of the son node values. With the median rule, significant edges are not eliminated through successive updating, and the median rule outperforms the average rule in the presence of heavy-tailed noise. However, both the average and median updates lead to excessive blurring and yield rectangular blocking artifacts after just two or three linking iterations. A third method leaves the node values unchanged after linking. This technique provides rapid convergence to a fixed segmentation. A repeatable segmentation result can be obtained that is not dependent on the exact number of linking iterations. Without node updating, the artifact problem is also mitigated. All examples in this paper for the GP, ADP-SD, and the ADP-MD were generated without node updating.

The final step in the linking process is propagation of the root values. Starting with level R - 1, the nodes are assigned new root values based on the root values of their father. Linking, node updating, and root value propagation are repeated until convergence. The ADP linking process has converged when the root values of level 0 remain unchanged after a complete linking iteration.

4. APPLICATION AND RESULTS

We utilize a remote sensing segmentation problem to compare the performance of the GP, the ADP-SD, and the ADP-MD. Previously, a single resolution version of anisotropic diffusion has been used successfully in segmentation of remotely sensed imagery [5]. In this application, our goal is the extraction of *fire scar* regions from satellite imagery of the Great Victoria Desert in Western Australia. A fire scar is a geographical region that has been burned in a particular desert fire. The fire scars tend to exhibit homogeneous spectral properties. Once the fire scars are extracted in the segmentation process, the scar boundaries are recorded and the regions are classified as a whole by fire scar age. For this task, the properties of boundary localization and intra-region intensity statistics are of utmost importance.

For the comparative study, we segment Landsat Thematic Mapper (TM) imagery. The particular input image used is a 256x256 section of band 4 of the TM imagery, shown in Fig. 1(a). The TM image has a spatial resolution of 30m and an 8-bit dynamic intensity range. We have chosen a fire scar extracted previously (by hand), shown in Fig. 1(b), for a benchmark study. Here, we compare the ground truth segmentation of Fig. 1(b) to the regions extracted through the pyramid-based segmentation.

The segmentation quality is evaluated on two fronts: pixel localization error and pixel intensity error. Pixel localization error measures the deviation of the segmented boundary from the ground truth boundary. Since a one-to-one correspondence between the boundaries cannot be established in general, a region-based evaluation is used. First, the constant regions that have a majority of pixels overlapping the ground truth fire scar (in image position) are extracted, forming the newly segmented fire scar. When the ground truth scar is overlaid on the segmented fire scar, the number of pixels in the interior of the ground truth fire scar that are not matched with corresponding pixels in the segmented fire scar is computed (interior error). Then, the number of pixel locations in the exterior of the ground truth fire scar that have corresponding pixels in the segmented version, is computed (exterior error). These two figures are added together to provide the total pixel localization error (see the table).

Pixel intensity error is measured here by the absolute difference between the mean intensity of the ground truth fire scar and the mean intensity of the segmented fire scar (μ = 205.6). It is also important to consider the number of constant regions that compose the extracted region (given by *No. of Regions* in the table). Generally, if a region is considered to be homogeneous at a given scale, the least number of constant regions is desired for that scale (pyramid level).

In the table, we exhibit the results for all three methods for root levels 5 and 6. The actual segmented images and extracted fire scars are shown for the level 6 results in Figs. 2- 4. As shown in the table and in Fig. 2, the GP segmentation results in high pixel localization error and high intensity error, as compared to both ADP implementations. For the root level 6 (4x4), the GP-extracted fire scar has over 6000 pixels that do not match (in position) the ground truth data. The intensity error for levels 5 and 6 of the GP prohibits successful fire scar classification.

The ADP-SD and the ADP-MD provide a higher quality segmentation in terms of both pixel localization and intra-region intensity preservation. The ADP-MD gives a slightly superior performance in both categories over the faster ADP-SD implementation. Notice in Figs. 3-4 that the ADP does not suffer from the blocking artifacts common to GP segmentation. Both implementations of the ADP minimize radiometric distortion, allowing image classification from intensity information *after* segmentation. The ADP-SD pyramids were initialized using K = 50 and λ = 0.15. The ADP-MD pyramids were created using K = 15 and λ = 0.15 with 40 diffusions per level.

Algorithm	Pixel localization error			Mean Intensity	Intensity Error	No. of Regions
	Interior	Exterior	Total			
GP Level 5	201	748	949	171.6	38.8	3
GP Level 6	125	6301	6426	151.0	54.4	1
ADP-SD Level 5	343	154	497	206.0	0.6	4
ADP-SD Level 6	367	407	774	211.0	5.6	1
ADP-MD Level 5	314	151	465	206.8	1.4	3
ADP-MD Level 6	242	419	661	207.0	1.6	1

Table: Segmentation results for the image shown in Fig. 1 (a).

5. FUTURE DIRECTION

This initial study uses intensity-based segmentation to demonstrate the efficacy of the ADP. The data used were 8-bit gray-level images. Extending this experiment, we plan to segment multispectral data using ADPs. The multispectral ADPs will utilize connections between the multispectral components as well as connections between the different pyramid levels. The node values will be multidimensional vectors, instead of scalar intensities. Of course, segmentation is only one of the suitable tasks for the multiresolution ADPs. Future ADP applications will include image registration and coarse-to-fine stereo correspondence. We do not believe, however, that the ADP will be useful for image coding, as are the GP and the associated Laplacian Pyramid, since the ADP filtering is nonlinear.

REFERENCES

[1] P.J. Burt, T. Hong, and A. Rosenfeld, "Segmentation and estimation of region properties through cooperative hierarchical computation," *IEEE Trans. on Systems, Man, and Cybernetics*, vol. SMC-11, no. 12, December, 1981.

[2] P. Perona and J. Malik, "Scale-space and edge detection using anisotropic diffusion," *IEEE Trans. on Pattern Anal. and Mach. Intell.*, vol. PAMI-12, no. 6, June, 1990.

[3] S.T. Acton and A.C. Bovik, "Anisotropic edge detection using mean field annealing," *Proc. of the 1992 IEEE Int'l Conf. on Acoustics, Speech and Signal Processing*, San Francisco, March 23-26, 1992.

[4] P. Saint-Marc, J. Chen & G. Medioni, "Adaptive smoothing: a general tool for early vision," *IEEE Trans. on Pattern Anal. Mach. Intell.*, vol. PAMI-13, no. 6, June, 1991.

[5] S.T. Acton and M.M. Crawford, "A mean field solution to anisotropic edge detection of remotely sensed images," *Proceedings of the 1992 IEEE Int'l Geoscience and Remote Sensing Symposium*, Houston, Texas, May 26-20, 1992.

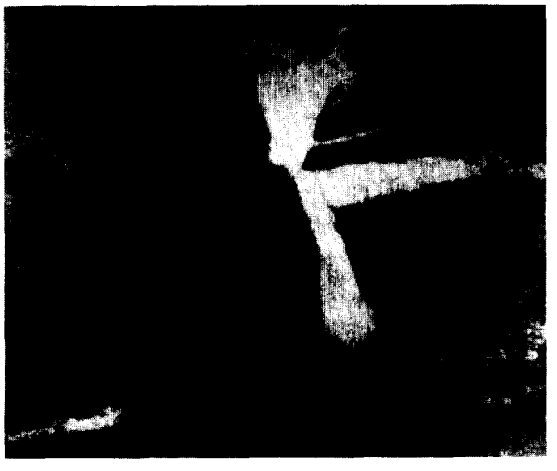


Fig. 1 (a) The original image (TM Band 1).

(b) The ground truth fire scar region.

Fig. 2 (a) GP segmentation (level 6).

(b) The extracted fire scar region.

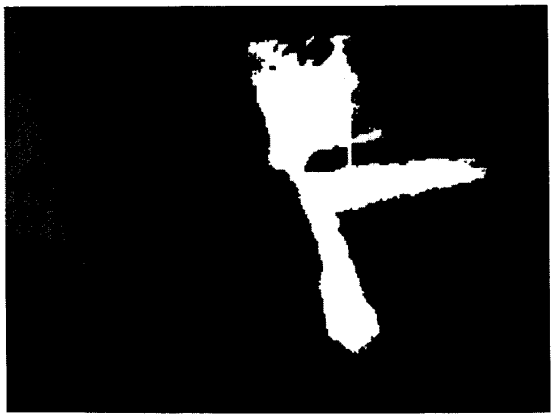
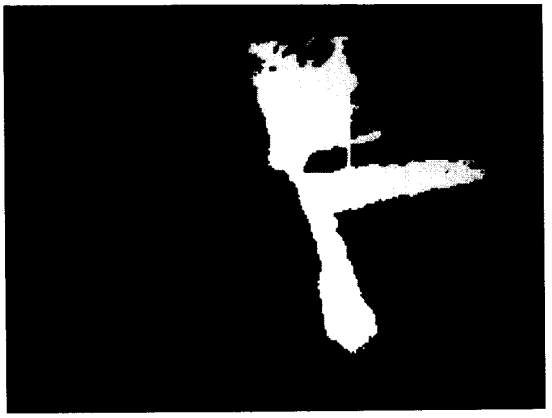


Fig. 3 (a) ADP-SD segmentation (level 6).



(b) The extracted fire scar region.

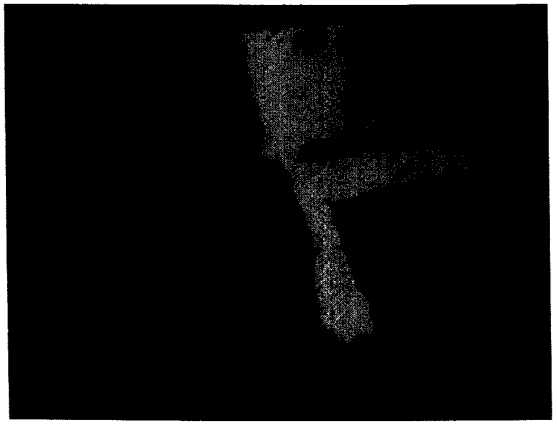
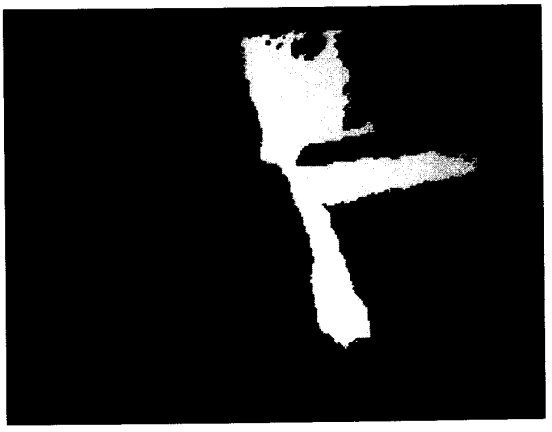


Fig. 4 (a) ADP-MD segmentation (level 6).



(b) The extracted fire scar region.