January 01, 2007

Fast Multi-Scale Regularization and Segmentation of Hyperspectral Imagery Via Anisotropic Diffusion and Algebraic Multigrid Solvers

Julio Martin Duarte-Carvajalino
University of Puerto Rico - Mayaguez

Miguel Vélez-Reyes
University of Puerto Rico - Mayaguez

Guillermo Sapiro
University of Minnesota - Twin Cities

Paul E. Castillo
University of Puerto Rico - Mayaguez

Fast Multi-Scale Regularization and Segmentation of Hyperspectral Imagery via Anisotropic Diffusion and Algebraic Multigrid Solvers

Mr. Julio Martin Duarte-Carvajalino[a] e-mail: jmartin@ece.uprm.edu
Dr. Miguel Velez-Reyes[a] e-mail: mvelez@ece.uprm.edu
Dr. Guillermo Sapiro[b] e-mail: guille@umn.edu
Dr. Paul Castillo[a] e-mail: paul.castillo@math.uprm.edu

[a]University of Puerto Rico-Mayaguez, Mayaguez, PR 00681-9042, USA
[b]University of Minnesota, Minneapolis, MN 55455-0436, USA

ABSTRACT

We introduce a fast algorithm for anisotropic smoothing and segmentation of hyperspectral imagery. Anisotropic smoothing reduces the spatial and spectral variability within uniform regions in the image, while preserving the sharp discontinuities on the image boundaries, which in turn improves the segmentation by increasing the separability between the different regions in the image. The algorithm solves a discretized Partial Differential Equation (PDE) that generates a discrete scale-space, represented by an irregular pyramid of coarser versions of the image. As the image is coarsened, representative pixels are selected at each scale, enabling a multi-scale segmentation of the image. The segmentation is performed in a top-down process that uses the representative pixels identified on each scale, as seeds. The PDE is solved using Algebraic Multigrid (AMG), a numerical analysis technique useful for boundary value problems on highly unstructured grids, with greater accuracy and speed than traditional relaxation techniques. The coarsening step in AMG is based on a modified version of the Iterated Weighted Aggregation method (IWA), tailored to exploit the discrimination power of high dimensional spaces such as those represented by hyperspectral data.

STATE OF THE ART

1. **Anisotropic Diffusion (milestones)**

- Perona-Malik proposed an anisotropic diffusion PDE to smooth grayscale images.
- Alvarez et al. showed that the Perona-Malik PDE was ill-posed, tough it can be made well-posed by regularization. They also showed that the regularized PDE generated a scale-space representation of the image, with some nice architectural, information-reducing and invariance properties.
- Weickert presented a unified view of anisotropic diffusion for scalar, vector and matrix-valued images.

2. **Anisotropic Diffusion on Hyperspectral images**

- Lennon et al. extended the Perona-Malik equation to be used with an hyperspectral image taken using the CASI sensor. Nevertheless, they selected 17 bands from the 450-950 nm range and they used the original ill-posed Perona-Malik PDE.
- Duarte et al. extended a regularized version of the Perona-Malik PDE to Hyperspectral images using semi-implicit numerical schemes that allows smoothing of hyperspectral imagery in linear time. They also showed[7] that anisotropic diffusion increases class separability and therefore classification accuracy.

3. **Fast Multi-Scale Image Segmentation**

- Sharon et al. proposed an algorithm for segmenting grayscale images using a scale-space representation, based on AMG and a new proposed coarsening method called Iterated Weight Aggregation (IWA). The running time of the algorithm is O(*n*), being *n* the number of pixels in the image).
- Galun et al. extended Sharon' segmentation algorithm to include statistical moments and texture, while keeping the running time linear in the size of the image.
- Galli-Candia performed a multi-spectral extension of Sharon's algorithm to segment remote sensing images. The extension consists into use Euclidean distance on IWA and affect the weights with the Bhattacharyya distance, when the regions have enough number of pixels.

Sharon's segmentation algorithm is based on hierarchical clustering, rather than on the scale-space representation of the image using parabolic (geometric) PDEs. We propose here to integrate the well-founded geometric scale-space representation of an image using PDES with a modified version of Sharon's segmentation algorithm that fits perfectly well within this framework.

The main contribution of this work is the integration of anisotropic diffusion using AMG as a solver and the Sharon's segmentation algorithm, both extended to smooth and segment hyperspectral imagery. We also modified Sharon's algorithm to make it more efficient for hyperspectral imagery.

CHALLENGES AND SIGNIFICANCE

The high spectral resolution provided by hyperspectral sensor systems enables higher discrimination capabilities than multispectral sensors such as IKONOS, QUICKBIRD, and LANDSAT, currently used to monitor the earth surface from remote sensing platforms. However, proper exploitation of the higher information content provided by hyperspectral imagery has been damped by the so called curse of the dimensionality. The usual approach in hyperspectral imagery has been to project the data onto a space of lower dimensionality using for example principal component techniques, where traditional statistical methods can be used to extract valuable information from the images. Nevertheless, dimension reduction techniques have currently significant time and memory complexity making them computationally expensive and hardly scalable for large data sets, as it is the case of hyperspectral imagery. In addition, traditional image processing algorithms in remote sensing work entirely on pixel by pixel basis, ignoring the high spatial and spectral correlation within neighboring pixels.

The unsatisfying results of pixel-based processing algorithms for both high spatial and spectral imagery has led to the employment of region-based algorithms that segments the image into a disjoint set of homogeneous regions (structures), before applying any higher level process such as classification, registration, target detection, change detection, impainting, etc. Nevertheless, segmentation algorithms have been introduced relatively late in remote sensing given the complexity and dimensionality of the data, the high heterogeneity (spatial and spectral) of the objects, and the difficulty of using model-based methods, such as the classic background-foreground model used for several grayscale and color segmentation algorithms.

Region based methods had evolved in remote sensing, within the scale-space framework, since, different image structures appears on an image at different image scales (spatial resolutions). Less aware is the remote sensing community of the use of parabolic PDEs for the generation of nonlinear scale-spaces that have well-sounded mathematical and numerical foundations. This scale-space formulation provides a framework for image segmentation and higher level image processing and understanding, as some previous work had shown.

This work introduces a fast geometric scale-space representation and segmentation of Hyperspectral imagery using Algebraic Multigrid (AMG), a state of the art numerical method to solve boundary value problems on highly unstructured grids. The main contribution of this work consists in the integration of AMG as a solver for the anisotropic diffusion PDE, and a fast segmentation algorithm for grayscale and color images proposed by Sharon et al. which we modify here to make it computationally more efficient to segment Hyperspectral imagery.

BACKGROUND

Anisotropic Diffusion for Vector-Valued Images

Let $\mathbf{U}^0(\mathbf{x}) = (\mathbf{u}_1(\mathbf{X}), \mathbf{u}_2(\mathbf{X}), \ldots, \mathbf{u}_m(\mathbf{X}))^T$, $\mathbf{x}=(x,y)$, be a multi-spectral image with *n* pixels and *m* bands and domain $\Omega \subseteq R^2 x R^m$. The multi-scale representation of the image is given by the evolution of the following PDE [Weickert02],

$$\frac{\partial \mathbf{u}_i}{\partial t} = \nabla \bullet \left(D\left(\sum_{j=1}^{m} \nabla \mathbf{u}_j \nabla \mathbf{u}_j^T \right) \nabla \mathbf{u}_i \right), \quad i = 1, \cdots, m \qquad (1)$$

where, D is a positive definite diffusion matrix, *t* is the scale parameter and the initial condition for the PDE is the original image. The diffusion matrix is designed such that diffusion across edges is inhibited.

The key point here is that edge strength in vector-valued images is measured using the Di Zenzo's matrix $\Sigma_j \nabla \mathbf{u}_j \nabla \mathbf{u}_j^T$, which eigenvectors describe the directions of highest and lowest change. Notice that, since *D* is the same for all image bands, the evolutions between channels are synchronized, avoiding that an edge would appear on different locations in the image.

The semi-implicit discretization of (1), in matrix-vector notation is given by,

$$\left(\mathbf{I} - \mu \mathbf{G}^n\right)\mathbf{U}^{n+1} = \mathbf{U}^n, \qquad (2)$$

where **U** is the hyperspectral image as defined earlier, **I** is the identity matrix, **G** the matrix of diffusion coefficients, *n* is shortcut notation for the discrete scale $n\Delta t$, and $\mu = \Delta t/(\Delta x \Delta y)$. Equation (2) provide us a way of computing recursively $\mathbf{U}^{n+1}$, the image at scale $(n+1)\Delta t$, knowing the image at the previous scale $\mathbf{U}^n$, where $\mathbf{U}^0$ is the original image. Explicit discretizations of (1) are numerically stable only for $\mu \le 1/4$, which limit their application. On the other hand, semi-implicit schemes are numerically stable for all values of μ. Nevertheless, the numerical stability of semi-implicit schemes, come at a price, we have to solve a set of linear equations, at each iteration step.

In our previous work [Duarte et al.-06,07], we use Alternating Direction Implicit (ADI) and Additive Operator Splitting (AOS) approximated methods to solve (2) approximately, we also use Preconditioned Conjugated Gradient (PCG) methods to obtain more accurate solutions to (2). We found, that PCG methods provide solutions of higher quality than AOS and ADI methods, which in turn provides better classification accuracies over the smoothed images. Nevertheless, PCG methods converges very slow to the exact solution of (1), making them not scalable.

Algebraic Multigrid (AMG)

Multigrid methods [Briggs et al.-00] surge from the analysis of classic iterative (relaxation) methods for the solution of linear systems of equations. Classic iterative methods reduce efficiently the high frequency components of the error; but, they are extremely inefficient to reduce the low frequency components. Multigrid methods aim to reduce the error components in all frequencies, in linear time and independently of the size of the data, making them algorithmically scalable. Multigrid methods use two complementary processes: smoothing of the error (relaxation) and coarse-grid correction. Coarse-grid correction involves transferring information from a fine to a coarser grid via a restriction operation. The coarsening process is continued until a relatively small grid is reached, where the linear system can be solved exactly, with little computational cost. The solution is then propagated back to the finer level via interpolation. The success of multigrid resides in the coarsening operation that displaces the low frequency components of the error to high frequencies in the coarse grid, where classical relaxation methods are used to reduce the high frequency components of the error. The relaxation can be accomplished by a simple iterative method such as Jacobi or Gauss-Seidel.

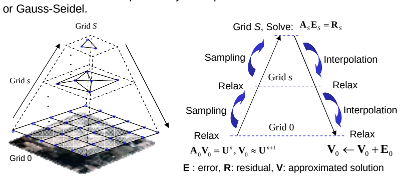


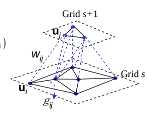
Figure 1 a) Schematic of AMG grid structure, b) Schematic of a V-cycle in AMG algorithm

Fast Multi-scale Segmentation (grayscale, [Sharon et al.-01])

In AMG, an image can be viewed as a graph, $G^0 = (V^0,E^0)$, where the superscript indicates the scale. V^0 is the set of vertices at scale 0 and it coincides with the pixels in the image. E is the set of edges connecting the image pixels to their nearest neighbors. Associated to E, we have a weight function *w*(E) that measures the affinity or similarity between two vertices.

AMG provides implicitly a soft-segmentation of the image by the dependences *w*ij between vertices in the fine and coarser levels as Figure 2 shows,

$$w_{ij}^s = w_{ji}^s = g_{ij}^s \Big/ \sum_{(i,k)\in s} g_{ik}^s, \quad g_{ij}^s = e^{-\alpha\theta(\mathbf{u}_i,\mathbf{u}_j)} \qquad (3)$$



The segmentation algorithm consists into obtain a hard-segmentation at the finer level using the representatives found at the coarser level and performing a top-down sharpening. Let us say grid s+1 has been segmented, we want to know how to segment grid s,

- Since all vertices in s+1 are segmented, p_i^r i.e. the probability that vertex *i* $\in$ s+1 belongs to representative *r* is known.
- Set $p_i^r = 1$ if $p_i^r \ge 1-\delta$, $p_i^r = 0$ if $p_i^r < \delta$.
- Update the probabilities on level s by performing υ Gauss-Seidel relaxations as $p_i^r = \sum_{(i,j)\in s} g_{ij}^s p_j^r \Big/ \sum_{(i,j)\in s} g_{ij}^s$

TECHNICAL APPROACH

We modify Sharon's algorithm in order to improve its efficiency for segmenting hyperspectral images. Briefly,

- Sharon's algorithm uses state vectors of the same size as the number of pixels in the image. A better approach for large sparse graphs is to use Red-Black trees to store the neighborhood of each vertex. Red-Black trees have also the additional advantage of providing fast searches within the tree.
- The running time of the algorithm is linear in time, but the constant of linearity grows exponentially with the size of the neighborhood. We eliminate on each vertex, the weights with values lower than 0.1. Additionally, we limit the number of neighbors to 10, reducing considerably the running time of the algorithm, without affecting negatively the accuracy of the AMG solver, or the segmentation algorithm.
- Sharon's segmentation algorithm might leave large regions of the image without label. We solve the problem by not allowing that vertices with very low probability of belonging to any segment, on a given scale, affect the vertices in the next finer scale. We simply use the global measures accumulated to assign vertices with low probabilities to the closest segment, in terms of similarity.
- Sharon's algorithm assigns pixels to each segment, one at a time. This segmentation is time consuming (especially with many segments). Instead, we perform the segmentation with all the representatives at the same time. The results are the same in terms of segmenting the same regions, but the improvement in running time is significant.

We showed in [Duarte et al.-06,07] that thanks to the smoothing properties of anisotropic diffusion a more selective diffusion coefficient than the one used by Sharon et al can be used. We use the diffusion coefficient proposed by J. Weickert, which produce segmentation-like smoothed images, given by

$$g(\theta) = \begin{cases} 1 & \theta = 0 \\ 1 - e^{-\frac{3.31488}{(\alpha\theta)^8}} & \theta > 0 \end{cases}, \quad \theta = \sqrt{\frac{1}{m}\sum_{k=1}^{m} \left\| \nabla \mathbf{u}_k \right\|^2} \qquad (4)$$

We also showed [Duarte et al.-07] that the diffusion coefficient can be computed from previous grid as

$$\mathbf{G}^{s+1} = \left[g_{kl}^{s+1}\right], \quad g_{kl}^{s+1} = \frac{1}{1 + \sum_{j \in V^s \setminus V^{s+1}} w_{ij}^s} \sum_{k,l \in V^s} w_{ik}^s g_{kl}^s w_{lj}^s. \qquad (5)$$

Which coincides with the Iterated Weighted Aggregation (IWA) algorithm proposed by [Sharon et al-01] using a variational approach, but with a normalizing term in the denominator, in our case.

ACCOMPLISHMENTS

We improve the accuracy of the computed solution of (1), with respect to our previous work [Duarte et al.-06,07] and achieved scalability for hyperspectral imagery using AMG (see Figure 2). We also extend and improve a segmentation algorithm, based on Sharon's algorithm and the structure obtained with AMG (Figure 3). In [Duarte et al-07] we showed that classification accuracy improves by smoothing and then segmenting hyperspectral imagery (Figure 3 and Table 1). The code was implemented in C++ under the CYGWIN environment (http://www.cygwin.com/). We use LAPACK to obtain the exact solution of (1) on the coarser grid and on the finer grid, in order to test the accuracy of AMG as a solver. We also use the Geospatial Data Abstraction Library (GDAL, http://www.gdal.org/) which supports more than 50 raster image formats, including BIL, BSQ and BIP formats, commonly used in hyperspectral imagery, without limit in the size of the image.

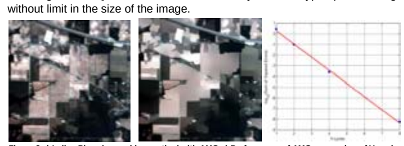


Figure 2 a) Indian Pines image, b) smoothed with AMG c) Performance of AMG vs. number of V-cycles

Figure 3 Segmented Indian Pines image, a) without smoothing, b) smoothed using Euclidean Distance, c) Smoothed using the Spectral Angle

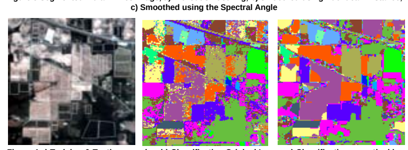


Figure 4 a) Training & Testing samples, b) Classification Original Image, c) Classification smoothed image

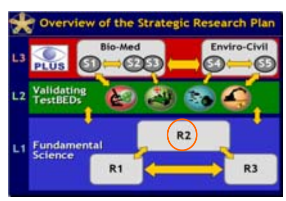
Figure 5 Classification of the segmented images a) without smoothing , b) Smoothing using Euclidean Distance, c) Smoothing using Spectral Angle

Table 1 Summary Classification Accuracies

Image	Classification Accuracy (%)	
	Training	Testing
Original	90.8	71.8
Smoothed with AMG	99.8	80.6
Segmented with Sharon's algorithm	97.4	84.8
Segmented with AMG-Euclidean Distance	95.1	88.3
Segmented with AMG-Spectral Angle	98.5	90.1

FUTURE PLANS

- Improve segmentation by including similarity metrics that consider the spatial correlation between neighboring regions.
- Consider non-flat manifolds to smooth and segment Hyperspectral images.



PUBLICATIONS ACKNOWLEDGING NSF SUPPORT

- J. Duarte-Carvajalino, P. Castillo, Paul, M. Vélez-Reyes, "Nonlinear diffusion using semi-implicit schemes in hyperspectral imagery," ADMI 2005 –Modeling Diversity in Computing and Engineering, The Symposium on Computing at Minority Institutions, October 13 - 15, Rincon, PR, 2005.
- J. Duarte-Carvajalino, P. Castillo, and M. Velez-Reyes, "Comparative Study of Semi-implicit Schemes for Anisotropic Diffusion in Hyperspectral Imagery," accepted for publication on *IEEE Trans. Image processing*, vol. 16, no. 5, May 2007.
- J. Duarte-Carvajalino, M. Velez-Reyes, and P. Castillo, "Scale-space in Hyperspectral Image Analysis," *SPIE Defense and Security Symposium,* vol. 6233, pp. 334-345, 2006.
- J. Duarte-Carvajalino, G. Sapiro, M. Velez-Reyes, and P. Castillo, "Fast Multi-Scale Regularization and Segmentation of Hyperspectral Imagery via Anisotropic Diffusion and Algebraic Multigrid Solvers," *SPIE Defense and Security Symposium* to be held in Orlando, vol. 6535, April 9-13, 2007.

OTHER REFERENCES

- L. Alvarez, F. Guichard, P. L. Lions, and J. M. Morel, "Axioms and fundamental equations of image processing," *Arch. Rational Mech. Anal.*, **123**:199-257, 1993.
- A. Akselrod-Ballin, M. Galun, M. J., Gomori, M., Filippi, P., Valsasina, R., Basri, and A., Brandt "An Integrated Segmentation and Classification Approach Applied to Multiple Sclerosis Analysis," *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, **1**:1122-1129, 2006.
- W. L. Briggs, V. E. Henson, and S. F. McCormick, A Multigrid Tutorial, 2nd Ed. SIAM Edt., 193 p., 2000.
- S. Di Zenzo, "A note on the gradient of a multi-image," Computer Vision, Graphics, and Image Processing, **33**: 116-125, 1986.
- M. Galun, E. Sharon, R. Basri, and A. Brandt, "Texture Segmentation by Multiscale Aggregation of Filter Responses and Shape Elements," *Proc. IEEE International Conference on Computer Vision,* **1**: 716-723, 2003.
- L. Galli and D. de Candia, "Multispectral Image segmentation via Iterated Weighted Aggregation Method," *SPIE Image and Signal Processing for Remote Sensing Symposium*, **5982**:74-81, 2005.
- M. Lennon, G. Mercier, and L. Hubert-Moy, "Nonlinear filtering of hyperspectral images with anisotropic diffusion," *IEEE Intl. Geoscience and Remote Sensing Symposium,* 4:2477-2479, June 2002.
- P. Perona and J. Malik, "Scale-space and edge detection using anisotropic diffusion, "*IEEE Trans. Pattern Analysis and Machine Intelligence*, **12**(7):629-639, July 1990.
- E. Sharon, A. Brandt, and R. Basri, "Fast Multiscale Image Segmentation," *Proceedings IEEE Conference on Computer Vision and Pattern Recognition*, **1**:70-77, 2000.
- J. Weickert and T. Bronx, "Diffusion and Regularization of Vector- and Matrix-valued Images," In M. Z. Nashed and O. Scherzer, editors, *Inverse Problems, Image Analysis, and Medical Imaging*, *Contemporary Mathematics*, **313**: 251–268, 2002.

This work was partially supported by National Geospatial Agency and use facilities supported by the Engineering Research Centers Program of the National Science Foundation under award EEC-9986821. Julio M. Duarte-Carvajalino is partially supported by a Fellowship from the Puerto Rico NSF EPSCoR Program.