A network-theoretic approach to hyperspectral image classification

Ryan H. Lewis and Anthony A. Harkin
Rochester Institute of Technology, School of Mathematical Sciences, 85 Lomb Memorial Drive, Rochester, NY 14623, USA
Anthony.Harkin@rit.edu

Abstract. This work describes a topological and network-theoretic approach for clustering pixels in hyperspectral image data. A network structure is imposed on hyperspectral data by considering the spectral distance between pairs of pixels, which are thought of as points in a high-dimensional space. The resulting network effectively models the complex nonlinear structures in the data and it can be exploited to perform clustering by applying a community detection algorithm, such as the method of maximum modularity. The method of maximum modularity is an unsupervised, deterministic method for detecting communities in graphs where neither the number of communities nor their sizes needs to be specified in advance. Because the network structure is based on local neighborhoods within the spectral data, algorithms based on such constructions are said to be topological or geometric.

Keywords: Hyperspectral, Clustering, Classification, Community Detection

1 INTRODUCTION

Classification, or clustering, for hyperspectral imaging is the process of assigning pixels into spectrally similar groups. The most well-known data clustering technique is the $K$-means algorithm, which can be applied to group any set of $n$ data points into $K$ clusters and works by minimizing the variance within each cluster. $K$-means clustering is fast, but has several drawbacks. It requires the number of clusters to be specified in advance, and, a more serious drawback is that intra-cluster variance may not be an appropriate measure of data spread when non-linear structures are present. The realization that complex geometric structures are present in HSI data has led to a number of nonlinear approaches for dimension reduction and image analysis that go beyond $K$-means, PCA, and other linear techniques. Two well-known examples include local linear embedding [1–3] and ISOMAP [4].

In [5] a topological approach for detecting anomalies in hyperspectral images was introduced called Topological Anomaly Detection (TAD). The main idea underlying TAD is that the high dimensionality and nonlinearity of a hyperspectral data set can be effectively handled by imposing a network structure on it. Similar ideas are also employed in [6–8] to investigate topological structures contained in general data sets. HSI data can be converted into a graph by connecting pairs of pixels if their spectra are sufficiently close. In TAD, pixels belonging to large components in the resulting graph are defined to be background pixels. Non-background pixels, or anomalies, are then ranked by their distance to the background.
In this article we exploit a network-theoretic framework to perform image segmentation. By imposing a network structure on HSI data, the problem of clustering then maps to the problem of detecting communities in networks. The method of maximum modularity [9], which we use in this work, is one such choice for a community detection algorithm that has several desirable properties important for image classification.

2 NETWORK ANALYSIS & COMMUNITY DETECTION

In this section, we first describe how to convert hyperspectral image data into a network structure called a resolution graph. After describing how to construct the resolution graph, we next explain how the method of maximum modularity can be employed to cluster the data into spectral classes.

2.1 Building the Resolution Graph

A resolution graph $G_\epsilon = (V, E)$ is a mathematical structure with a set $V$ of $n$ nodes and a set $E$ of $m$ edges connecting the nodes. For a hyperspectral image with $p$ spectral bands, each pixel is considered to be a point $x_i \in \mathbb{R}^p, i = 1, \ldots, n$. We construct a resolution graph by treating each pixel as a node and connecting a pair of points with an edge if the distance between them is less than a threshold, $d(x_i, x_j) < \epsilon$. The distance metric, $d$, may be chosen from any of several options, such as euclidean, normalized euclidean, or spectral angle. In Figure 1, resolution graphs are constructed on a synthetic data set using the euclidean distance with various $\epsilon$-values. The resolution graph constructed on a small, random sample of pixels from a hyperspectral image of Buck Pond is shown in Figure 2. The resolution graph for the Buck Pond image was constructed using the euclidean distance with $\epsilon = 85$. The rule of thumb we employ for choosing a good value of $\epsilon$ is to try to select the smallest value possible such that the vast majority of pixels are in one large connected component.

The graph $G_\epsilon$ can be represented using an $n \times n$ adjacency matrix, $A$, where $A_{ij} = 1$ if there is an edge between the $i^{\text{th}}$ and $j^{\text{th}}$ nodes and $A_{ij} = 0$ otherwise. The sum of the $i^{\text{th}}$ row of $A$ is the degree $k_i$ of node $i$, and the vector $k$ contains the degrees of all the nodes. Also note that the total number of edges in the graph is $m = (1/2) \sum_i k_i$.

![Fig. 1. Examples of $G_\epsilon$ using synthetic data in $\mathbb{R}^2$ for three different values of $\epsilon$.](image-url)
Fig. 2. Constructing the resolution graph $G_\epsilon$ for a hyperspectral data set from the AVIRIS sensor. (Left) RGB image of Buck Pond along Lake Ontario. (Right) Resolution graph ($\epsilon = 85$) of 500 random pixels from the Buck Pond image projected onto two spectral bands (596.78 nm and 1135.3 nm).

A resolution graph may have multiple connected components, as illustrated by the networks shown in Figures 1 and 2. In turn, each component may have a complex community structure. A connected component is considered to be a background component if it contains more than $\delta$ percent of the total pixels in the image, otherwise we say it is an anomalous component. An anomalous component may consist of a very small number of pixels, or even a single, isolated pixel.

2.2 The Method of Maximum Modularity

For each background component in the resolution graph we wish to determine whether there exists a community structure which naturally divides the pixels (nodes) into smaller clusters. A schematic diagram of a single connected component containing what appears to an observer to be three communities is shown in Figure 3. In the seminal work [9], the intuitive notion of community structure is made more precise. The crucial observation is that a good division of the network might exist if there are fewer than

Fig. 3. A single connected component containing three communities.
the expected number of edges between two communities than one would find based on random chance. Meaning, a strategy for clustering nodes into two communities is to consider a quantity \( Q \), called modularity, which measures the distance of a network from a corresponding random network with the same degree distribution.

To derive a mathematical expression for the modularity \( Q \), suppose we have a network with \( n \) nodes that belong in one of two communities, either \( C_1 \) or \( C_2 \). The expected number of edges between nodes \( i \) and \( j \), if edges are placed at random, is \( k_i k_j / 2m \), where we recall that \( k_i \) is the degree of node \( i \). The modularity of the network is now defined as

\[
Q = \frac{1}{2m} \sum_{i,j \in C_1} \left( A_{ij} - \frac{k_i k_j}{2m} \right) + \frac{1}{2m} \sum_{i,j \in C_2} \left( A_{ij} - \frac{k_i k_j}{2m} \right).
\]

The problem of community detection now boils down to finding the grouping of nodes into the two communities such that \( Q \) is maximized.

Let \( s \) be a vector where \( s_i = 1 \) if node \( i \) is in group \( C_1 \) and \( s_i = -1 \) if it is in group \( C_2 \). Since the quantity \( \frac{1}{2} (s_i s_j + 1) \) is 1 if \( i \) and \( j \) are in the same group and 0 otherwise, the modularity can be expressed as

\[
Q = \frac{1}{4m} \sum_{i,j} \left( A_{ij} - \frac{k_i k_j}{2m} \right) (s_i s_j + 1) = \frac{1}{4m} \sum_{i,j} \left( A_{ij} - \frac{k_i k_j}{2m} \right) s_i s_j.
\]

Therefore, the modularity can be written in the matrix form

\[
Q = \frac{1}{4m} s^T Bs
\]

where the modularity matrix is defined by

\[
B = A - \frac{kk^T}{2m}.
\]

The modularity matrix \( B \) is a real, symmetric matrix whose row and column sums are zero, and it has an eigenvector \( \langle 1, 1, \ldots, 1 \rangle \) with a corresponding zero eigenvalue.

For a network with \( n \) nodes that we wish to split into two communities, there are \( 2^n \) possible choices for the vector \( s \) and it is clearly not feasible for large graphs to maximize \( Q \) by evaluating all of them. By expanding \( s \) in terms of the unit eigenvectors \( \{v_1, v_2, \ldots, v_n\} \) of \( B \), it is not difficult to show [9] that an approximate maximum for the modularity can be obtained by assigning nodes into two groups according to the signs of the elements of the leading eigenvector, namely \( s = \text{sign}(v_1) \). In fact, the magnitudes of the elements of \( v_1 \) provide an indication of how firmly each pixel (node) belongs to its cluster. Moreover, if the entries of the leading eigenvector are all of the same sign then this indicates that the network is indivisible.

Once the nodes have been divided into two communities, the process may be repeated recursively to further subdivide the network until all clusters are found. Figure 4 shows the first two levels of the tree hierarchy produced by clustering the Buck Pond scene. The resolution graph for this image began with a single connected component, which was recursively divided using the method of maximum modularity. At each level
of the recursion, the adjacency matrix $A$ and the degree vector $k$ were recalculated to reflect the removal of edges. In Figure 4, the image was first divided into water and land. At the next level of recursion, each of these two classes was further partitioned without consideration of the other class. In other words, water pixels were not considered at all when subdividing the land into mud and grass. This type of local refinement is fundamentally different than clustering algorithms such as $K$-means.

**Fig. 4.** Two levels of recursion for the Buck Pond chip. The $K$-means result with four classes is shown for comparison.

### 3 RESULTS

For all of the results in this section, resolution graphs were constructed from the hyperspectral images using the normalized euclidean distance. On each background component, modularity was performed recursively until all communities within that component were detected. It should be emphasized that, because the modularity algorithm was used to perform the clustering, neither the sizes nor the number of spectral classes were specified in advance.

#### 3.1 Braddock Bay

The first scene analyzed was taken near Braddock Bay along Lake Ontario in western New York by the AVIRIS sensor. This image has 152 spectral bands and contains 223 samples and 201 lines. The image was converted into a resolution graph using $\epsilon = 0.08$, which resulted in a network containing $n = 44,823$ nodes and $m = 198,472,864$ edges. Note that the number of edges is only 9.8% of the number of edges there would be if the resolution graph were fully complete, meaning an edge between every pair of
nodes. To get a sense of the structure of the Braddock Bay resolution graph, the degree distribution is shown in Figure 5.

![Degree distribution of the Braddock Bay resolution graph with $\epsilon = 0.08$.](image)

**Fig. 5.** Degree distribution of the Braddock Bay resolution graph with $\epsilon = 0.08$.

With the value of $\epsilon = 0.08$ that was used, the Braddock Bay graph has one large connected component, that is to say, one can find a path between any two nodes in the graph. The method of maximum modularity was run on the graph and detected 52 spectral classes (Figure 6).

![RGB image taken near Braddock Bay along Lake Ontario.](image)  ![Classification of the scene with $\epsilon = 0.08, \delta = 2\%$. The modularity algorithm detected 52 spectral classes.](image)

**Fig. 6.** (Left) RGB image taken near Braddock Bay along Lake Ontario. (Right) Classification of the scene with $\epsilon = 0.08, \delta = 2\%$. The modularity algorithm detected 52 spectral classes.

Recall that the algorithm first splits the graph into two communities, then each of those communities is divided in two, and so forth in a recursive fashion. The modularity algorithm automatically terminates when the signs of the all the elements in the leading eigenvector of the modularity matrix are identical. For processing the Braddock Bay scene, we decided to include an optional stopping criterion for the recursion by not allowing communities which contained less than $\delta = 2\%$ percent of the total pixels in the image to be further subdivided. One might also consider terminating the subdivision at a predefined maximum recursion depth as another optional stopping criterion. The choice of adding optional stopping criterion to terminate the recursive subdivisions of the communities allows the user to control the refinement of the classification.
3.2 Copperas Cove Walmart

We next analyze a chip taken from an urban scene near a Walmart in Copperas Cove, Texas using the HYDICE sensor. This image has 162 spectral bands and contains 183 samples and 110 lines. It was converted into a resolution graph using $\epsilon = 0.28$, which yielded a network containing $n = 20,130$ nodes and $m = 72,028,116$ edges. The number of edges is 8.8% of the total possible for the corresponding complete graph. As in the Braddock Bay graph, the Walmart resolution graph has one large connected component. The method of maximum modularity was run and detected 66 spectral classes (Figure 7).

![RGB image of an urban scene near Walmart in Copperas Cove, Texas. Classification of the Walmart scene with $\epsilon = 0.28$, $\delta = .5\%$. The modularity algorithm detected 66 spectral classes.](image)

The sizes of the clusters for the Copperas Cove Walmart scene are displayed in Figure 8. Of the 66 spectral classes detected, 39 of the classes encountered the optional stopping criterion to not further divide clusters containing less than $\delta \times n \approx 101$ pixels. The other 27 spectral classes each contain more than 101 pixels and stopped subdividing when the modularity algorithm itself refused to partition them any further.

![Number of pixels in each of the 66 spectral classes for the Walmart scene.](image)
4 CONCLUSION

We have presented a topological and network theoretic framework for the unsupervised classification of hyperspectral image data which we refer to as Topological Image Classification (TIC). Converting the image into a resolution graph is not only an effective way of modeling the complex, nonlinear structures embedded in the data, but also permits the data to be processed by community detection algorithms such as the method of maximum modularity. The maximum modularity algorithm employs a course grain approach by recursively partitioning the image into smaller and smaller clusters and halts when it automatically determines no further subdivision is possible, or until an optional stopping condition is satisfied. It should also be emphasized that the magnitudes of the leading eigenvector of the modularity matrix convey how firmly each pixel belongs in its respective spectral class.

Enhancements and modifications to the TIC paradigm are possible. For example, numerous other community detection algorithms could be investigated for HSI data, such as the local methods described in [10–12]. In future studies, the sensitivity of the classification to changes in the parameters $\epsilon$ and $\delta$ will be further explored. Other lines of study will include examining whether the clusters found by TIC can be used as the basis for a compression scheme for hyperspectral image data.

Acknowledgments

The authors of this paper are grateful for the helpful discussions with Dr. William Basener (Math, RIT) and Dr. Dave Messinger (Imaging Science, RIT). The authors would also like to acknowledge the staff and resources of RIT Research Computing.

References
