Modeling Coastal Waters from Hyperspectral Imagery Using Manifold Coordinates

†Charles M. Bachmann, †Thomas L. Ainsworth, ††David B. Gillis, ††Shelia J. Maness, ‡Marcos J. Montes, ††Timothy F. Donato, ††Jeffrey H. Bowles, ††Daniel R. Korwan, †Robert A. Fusina, ††Gia M. Lamela, †† W. Joseph Rhea

Naval Research Laboratory
Remote Sensing Division, †Code 7232, ††Code 7231, ‡Code 7263
Washington, D. C. 20375, e-mail: bachmann@nrl.navy.mil

Abstract—In [1] [2], we introduced a direct data driven method of modeling nonlinear structure in hyperspectral imagery based on Isometric Mapping [15]. More recently, we have further improved the scaling of the approach [2], making it a practical method for large-scale hyperspectral scenes. The new method extracts a set of data manifold coordinates that directly parameterize nonlinearities present in hyperspectral imagery, both on land and in the water column. In the water column, this is particularly important because of the nonlinear, attenuating properties of the medium. In this paper, we model hyperspectral imagery acquired by the NRL PHILLS [5] at the Indian River Lagoon, Florida in July 2004. In our previous efforts [3] using a small subset of data derived from the surf zone outside of the lagoon, dominant manifold coordinates were shown to parameterize bathymetry directly with a high degree of correlation to a radiative transfer look-up table (LUT) approach. In the present work, we construct a full scene manifold coordinate representation and use this as the basis of a LUT for samples with known depths as determined by the SHOALS LIDAR. Sequestered test data presented to the manifold based LUT yield a mean estimated depth which differs from the LIDAR retrieved depth by less than 0.44m for depths between 0-10m with a standard deviation less than 1.2m.

I. INTRODUCTION

In [1] [2], we introduced a direct data driven method of modeling nonlinear structure in hyperspectral imagery based on Isometric Mapping [15]. More recently, we have further improved the scaling of the approach [2], making it a practical method for large-scale hyperspectral scenes. The new method extracts a set of data manifold coordinates that directly parameterize nonlinearities present in hyperspectral imagery, both on land and in the water column. The motivation for such a parameterization and its applicability to in water problems such as estimation of bathymetry, in water constituents, and bottom type, is based on the physical expectation that in shallow waters in a region that is homogeneous in bottom type and dissolved constituents, the reflectance at any particular wavelength should decay exponentially as a function of depth [10]. If the rate varies with wavelength, then the reflectance should best be described by a nonlinear sheet or manifold in spectral space [1] [6]. Other changes in the structure of the data manifold can be expected as inherent optical properties (IOP) and bottom type vary. In previous work [3], we showed that manifold coordinates were strongly correlated with the retrieved results of radiative transfer LUTs [12], which model water properties such as bathymetry, bottom type, and in water constituents. The radiative transfer LUT had been used to develop models of these in water properties for a study site at the Indian River Lagoon, Florida in July 2004 [9]. In [9], a set of look-up tables were produced by repeated execution of a radiative transfer software package known as EcoLight [11] to obtain models of the bottom type, bathymetry, CDOM, suspended sediments, and chlorophyll concentration.

In our previous efforts [3] using a small subset of data derived from the surf zone outside of the Indian River Lagoon, dominant manifold coordinates were shown to parameterize bathymetry directly with a high degree of correlation to the radiative transfer LUT approach. In particular, we found that, in the manifold coordinate representation, deep, moderate, shallow, and very shallow waters aggregated in linear segments or arcs which parameterized the progression of depth from deep to shallow water. In addition, the manifold coordinate representation was shown to extract greater details when compared with linear approaches modeling the water column in the lagoon. In the present work, we expand the scope of the comparison to take advantage of SHOALS LIDAR data collected just a few months prior to the PHILLS data acquisition. The LIDAR data was acquired in the surf zone outside of the lagoon and provides a reliable reference for bathymetry in at least a portion of the imagery acquired during the PHILLS campaign.

II. DATA AND STUDY AREA

A. The Indian River Lagoon, Florida

The Indian River Lagoon (IRL) is a 156-mile estuary located on Florida’s eastern seaboard [8]. The IRL is among the most biologically diverse estuaries in North America, containing habitats for over 4000 species of plants and animals in mangrove forests, saltmarshes, and seagrass meadows. The ecological diversity is attributable to the fact that the IRL borders both temperate and subtropical climates located approximately...
between 27 and 29 North latitude. Water circulation patterns within the estuary are wind-driven rather than influenced by gravity or tide as compared to rivers or coastal ocean areas respectively. Five inlets connect the IRL to the Atlantic Ocean, but the work in this study focuses only on the area between the Fort Pierce and St. Lucie inlets. Stressful influences from both agricultural and urbanization developments have necessitated pro-active measures to determine the welfare of estuary water quality [14]. The State of Florida has mandated a Surface Water Improvement and Management Plan (SWIM) to monitor the IRL by using submerged aquatic vegetation (SAV) as the primary indicator of lagoon health.

B. PHILLS Airborne Hyperspectral Data Collections

Airborne hyperspectral imagery of the Indian River Lagoon was collected by the Naval Research Laboratory’s PHILLS [5] sensor on July 14-15, 2004. The PHILLS collected data in 64 spectral channels between 0.39 and 1.0 $\mu$m. Twenty-five hyperspectral flightlines were obtained at a GSD of 4m. The data were atmospherically corrected using an in-house package known as Tafkaa [13]. The location of the flights is shown in Figure 1 along with a mosaic of the imagery collected on July 14, 2004. In Section IV, we focus primarily on flightline number 3, shown in Figure 2.

III. METHODOLOGY

In this paper, we use an enhanced version of the ISOMAP algorithm [15] [4] to model the nonlinear structure of our coastal hyperspectral imagery of the Indian River Lagoon. Our approach incorporates several features designed to allow improved scaling of this algorithm which, in its original form, could not be applied to data sets of the size found in remote sensing, where typically we need to process $O(10^6) - O(10^7)$ samples or greater. The enhanced ISOMAP (ENH-ISOMAP) [2] incorporates several features, including: (1) the use of “landmarks” [7] to reduce the memory and computational burden associated with the geodesic distance calculation, (2) a fast, exact neighborhood calculation; (3) an improved landmark selection algorithm; (4) a fast algorithm for attaching disconnected pockets of data separated from the main distributions; and (5) a flexible neighborhood definition to accommodate the large variations in spectral density found in hyperspectral scenes (especially coastal imagery). The fast neighborhood calculation cited above is particularly important for several reasons. For one, even with the use of landmarks to streamline processing, the neighborhood calculation, used to initialize the sparse graph used in the geodesic distance calculations, rapidly becomes the dominant calculation as the sample size increases. The use of ENH-ISOMAP allows for solutions up to $O(10^5)$ samples on conventional single-CPU computing architectures; however to scale up to larger scenes, we have had to develop a method for scaling up to the $O(10^6) - O(10^7)$ sample range and beyond [2]. The basis of the method involves building a representative “backbone” cube of size $O(10^5)$ samples, obtaining manifold coordinates with ENH-ISOMAP and then inserting the remaining samples in the scene via a reconstruction algorithm that yields their coordinates in the same manifold coordinate system. The fast neighborhood search algorithm is also

![Figure 1](image1.png)

Fig. 1. (Top) The twenty-five flightlines flown by the PHILLS hyperspectral sensor during July 14-15, 2004. (Bottom) An RGB mosaic of the flightlines flown on July 14, 2004.
used in the reconstruction algorithm. Further details are described in [2].

In the present work, we obtain full scene manifold coordinates using the method described above [2] for PHILLS HSI flightlines in the Indian River Lagoon. In the Results Section IV, we provide an example of this approach, showing manifold coordinates obtained for the third HSI line of Figure 1. At a set of randomly chosen sample points, we then constructed a table in which the manifold coordinates were associated with a known depth obtained from the SHOALS LIDAR. The entire scene was then mapped by querying this manifold based LUT as described below.

IV. RESULTS

Figure 2 shows an RGB derived from the third PHILLS HSI flightline of the set portrayed in Figure 1 after atmospheric correction and georectification. The area overlapped by the SHOALS LIDAR is a subset of this scene, covering the ocean region to the East (right) of the scene and a portion of the channel feeding into the ocean. The LIDAR and HSI were both georectified. Manifold coordinates were derived for the entire HSI scene (1,019,500 pixels). Figure 2 shows two of the many possible combinations of manifold coordinates derived from the full manifold coordinate product; we show just the subsets associated with water pixels in the vicinity of the coincident LIDAR data. As we have observed before [3], the manifold representation reveals a significant amount of structure in the water column and is potentially related to many in water properties such as bottom type, CDOM, chlorophyll concentration, suspended sediments, and bathymetry. Indeed the subset of manifold coordinates portrayed in Figure 2 reveals a significant amount of structure.

From these set of water pixels represented in manifold coordinates a set of samples was chosen randomly and the corresponding SHOALS lidar measurement was assigned to the same position in a look-up table. In all, 75,000 samples were chosen from the water portion of the scene. Next each manifold coordinate sample in the scene was then used to query the manifold based LUT in the following manner. For each water pixel in the scene, the closest neighbors of the sample in manifold coordinates were identified using the fast neighbor search algorithm described in [2], and a reconstruction of the sample’s manifold coordinates in terms of the closest neighbors in the LUT was then performed. The same weights associated with the reconstruction of the query’s manifold coordinates were then applied to the associated depths for each of the neighbors and the results summed to obtain the estimated depth of the query sample. The resulting bathymetry map is shown in Figure 3. All of the test points not originally in the LUT were then used to assess the accuracy of the bathymetric map by comparing the estimated result with the result obtained from the SHOALS LIDAR. Figure 3 shows the statistical result for the test samples (133,090) not in the original LUT, comparing the depth estimated by the manifold LUT vs the retrieved depth obtained by the SHOALS LIDAR. We note that for depths between 0-10m, the predicted depth from the manifold based LUT differed from the retrieved LIDAR depth by < 0.44m, with a standard deviation < 1.2m. Even for depths in the range between 10-14m, the difference between the manifold LUT estimate and the LIDAR differs by < 0.84m, but the bound on the standard deviation has doubled by the time we reach 14m depth.

V. CONCLUSIONS

In earlier work [3], we had seen concrete evidence that manifold coordinate representations of hyperspectral imagery are strongly correlated with retrievals of in water properties obtained from radiative transfer LUTs. Results suggested that manifold coordinates could be used to provide an intrinsic parameterization of the nonlinear structure of hyperspectral data for modeling in water properties. In the present work, we focussed on modeling bathymetry with manifold coordinates, taking advantage of our recent algorithm development efforts which have allowed us to make this methodology practical for remote sensing scales of $O(10^6) - O(10^7)$ samples or
greater [2]. The latter approach was an improved version of work we presented earlier [1]. With the new method for obtaining manifold coordinates [2], we achieve better scaling and removed artifacts which appeared in the original scaling methods we presented in [1]. In this paper, we showed manifold coordinates for one of the hyperspectral scenes obtained by the NRL PHILLS hyperspectral ocean data, Proc. SPIE Int. Soc. Opt. Eng. 5806, 342-351, 2005.

Future efforts will focus on extending the results beyond single scenes using the reconstruction principle. Likewise, the strong correlations previously observed with many of the retrieved water properties in radiative transfer LUTs suggest the possibility of combining manifold representations with radiative transfer LUTs in future work, as we proposed in [3].

REFERENCES


