Continued Development of the Look-up-table (LUT) Methodology for Interpretation of Remotely Sensed Ocean Color Data

Curtis D. Mobley
Sequoia Scientific, Inc.
2700 Richards Road, Suite 107
Bellevue, WA 98005
phone: 425-641-0944 x 109   fax: 425-643-0595   email: curtis.mobley@sequoiasci.com

LONG-TERM GOAL

The overall goal of this work is to refine and validate a spectrum-matching and look-up-table (LUT) technique for rapidly inverting remotely sensed hyperspectral reflectances to extract environmental information such as water-column optical properties, bathymetry, and bottom classification.

OBJECTIVES

My colleagues and I are developing and evaluating a new technique for the extraction of environmental information including water-column inherent optical properties (IOPs) and shallow-water bathymetry and bottom classification from remotely-sensed hyperspectral ocean-color spectra. We address the need for rapid, automated interpretation of hyperspectral imagery. The research issues center on development and evaluation of spectrum-matching algorithms, including the generation of confidence metrics for the retrieved information.

APPROACH

The LUT methodology is based on a spectrum-matching and look-up-table approach in which the measured remote-sensing reflectance spectrum is compared with a large database of spectra corresponding to known water, bottom, and external environmental conditions. The water and bottom conditions of the water body where the spectrum was measured are then taken to be the same as the conditions corresponding to the database spectrum that most closely matches the measured spectrum.

In previous LUT work, we have simultaneously retrieved water column IOPs, bottom depth, and bottom classification at each pixel from the remote-sensing reflectance R_{rs} spectra. This is much to ask from a simple R_{rs} spectrum, but we have shown that all of this information is uniquely contained in hyperspectral reflectance signatures and that the information can be extracted with considerable accuracy (Mobley et al., 2005).

Previous work has considered only retrievals based on the closest matching LUT database R_{rs} spectrum to a given image spectrum. However, exactly which database spectrum most closely matches the image spectrum can be influenced by noise in the image spectrum. Another way to do the retrievals is to find not just the closest-fitting database spectrum, but to find the k closest fitting spectra. Each of these k spectra corresponds to different environmental conditions (bottom depth,

Report Documentation Page

Public reporting burden for the collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to a penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
30 SEP 2007	**Annual**	**00-00-2007 to 00-00-2007**

4. TITLE AND SUBTITLE	5a. CONTRACT NUMBER
Continued Development Of The Look-Up-Table (LUT) Methodology For Interpretation Of Remotely Sensed Ocean Color Data	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **Sequoia Scientific, Inc,2700 Richards Road, Suite 107,Bellevue,WA,98005**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S)
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES
code 1 only

14. ABSTRACT

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**	**Same as Report (SAR)**	**12**	

bottom type, or water IOPs). The retrieval can then be taken as the mean value (or some other statistic, such as the most frequently occurring value) of the k values. If these k spectra all correspond to very nearly the same environmental conditions, then we can be confident that the retrieval is not strongly influenced by noise and is, presumably, correct to within a small error. However, if the k closest spectra correspond to widely differing environmental conditions, then we are much less confident of the correctness of the retrieval. A measure of the confidence in a depth retrieval can be based on the standard deviation of the distribution of the k retrieved depths, for example. It should be noted that even if the value of an environmental parameter as obtained from the k closest-matching spectra analysis value is the same as the value obtained for the closest-matching spectrum, the distribution of the k values can be used to compute error estimates for the retrieved quantity. Indeed, the real value of this technique often lies in the generation of confidence bounds on retrieved quantities, which in application may be as important at the retrieved value itself. In the literature this approach to classification and error estimation is known as k Nearest Neighbor (kNN) analysis.

WORK COMPLETED

This year's work centered on evaluating the kNN method for obtaining quantitative measures of the uncertainty of the depth retrievals, i.e., for putting error bars on the retrieved depths at each pixel. Examples of depth retrievals and associated error metrics are shown below. We also evaluated a number of different metrics for determining the closeness of two spectra.

In addition to the work discussed here, we performed a detailed analysis of LUT depth and bottom classification retrievals in the localized area of Horseshoe Reef, Lee Stocking Island, Bahamas, for which bottom classification information was available from underwater transects by divers. The LUT results were in good agreement with ground truth for percent coverages of sediments, corals, and mixed bottom types over the reef. A paper on that work is now in press (Lesser and Mobley, in press).

We also applied the LUT methodology to imagery of optically deep turbid waters in Puget Sound, Washington. That work (not shown here) showed the need for improved methods of atmospheric correction of hyperspectral imagery because the retrievals for a given image were sensitive to the atmospheric correction scheme used (empirical line fit or TAFKKA). It was also found that in optically deep waters the LUT-retrieved bottom depth was roughly equal to the penetration depth at the wavelengths where the water was clearest, rather than a retrieval of infinitely deep water. (The penetration depth at a given wavelength is defined as the inverse of the diffuse attenuation coefficient for downwelling plane irradiance and gives an estimate of how far a sensor can "see" into the water column at that wavelength.) The reason that LUT retrieved the penetration depth rather an infinite depth in optically deep waters is not yet understood.

During this period we also performed an analysis of several statistical measures of "best" fit of the kNN retrieved bathymetry estimates using the 2002 FERI/NAVO/NRL/USACE Joint Looe Key HyperSpectral Imaging (HSI) and LIDAR Experiment (Bissett et al, 2005; Figure 7). This analysis included both vector distance and angular separation in an attempt to determine which measure of best fit would be appropriate for use in retrieving bathymetry, IOPs, and bottom classification estimates.

RESULTS

The LUT approach to retrieving IOPs, bottom reflectance, and bottom depth information from remote-sensing reflectances has performed well in its application to various PHILLS images of optically clear

and shallow waters (e.g., Mobley, et al., 2005). This year we re-analyzed imagery from the Lee Stocking Island (LSI), Bahamas, area, for which acoustic bathymetry data were available to study the utility of kNN analysis in generating error maps corresponding to the retrieved bathymetry maps. Figure 1 shows an RGB PHILLS image taken near LSI; Fig. 2 shows the corresponding acoustic bathymetry.

When doing a depth retrieval on this image with $k = 1$, i.e., when using only the closest-matching database spectrum (with the Euclidian distance metric; see Table 1) at each image pixel, the LUT bathymetry was on average 7.0% or 0.4 m too shallow; 66% of the pixels were within ±1 m of the correct (acoustic) depth, and 87% of the pixels were within ±25% of the correct depth. When the retrievals were done with $k = 30$ and the retrieved depth was taken to be the mean of the 30 values, the LUT bathymetry was on average only 1.8% or 0.04 m too shallow. The other two statistics changed very little. Thus the kNN retrievals were on average deeper, which is correct, but the spread of retrieved vs. acoustic depths was essentially unchanged. This is likely because that spread of values in influenced by errors in geolocation of image vs. acoustic points (discussed in Mobley et al., 2005), which cannot be rectified by any analysis technique. Figure 3 shows the retrieved depths for $k = 1$, and Fig. 4 shows the retrievals defined as the mean of 30 values. The most noticeable difference is that the deeper waters at the upper right of the image are somewhat deeper for the $k = 30$ retrieval.

The k retrieved depths at each image pixel were used to generate two kinds of error maps for the depth retrievals. Figure 5 shows the map of the standard deviation of the 30 retrieved depths. This map gives an estimate of the absolute error in the retrieved depths. As would be expected, the standard deviation of the retrieved depths is greatest for the deepest water. However, some shallow areas with dark bottoms also have large standard deviations.

Figure 6 shows the map of the standard deviation divided by the mean depth, which gives a map of the relative errors in the depth retrievals. Overall for this image, this error metric is in the 0.05 to 0.15 range, although some areas with dark bottoms have larger relative errors. In general, areas with bright bottoms (ooid sands, in this image) have the smallest relative errors in the depth retrievals.

These error maps are in qualitative agreement with what is expected from signal-to-noise considerations, i.e., bathymetry for shallow or bright-bottom areas is retrieved most accurately, and deeper areas and areas with darker bottoms have more uncertainty in the retrieved bathymetry. However, the combination of LUT spectrum matching and kNN analysis allows us to generate quantitative error estimates on the retrieved bathymetry. Such error maps cannot be generated by simpler algorithms that work only with the image spectrum (e.g., band ratio algorithms).

The analysis on Horseshoe Reef was completed using a traditional Euclidean distance calculation (see Table 1). However, there are a number of different statistical measures of "best" fit that could be used to complete the kNN retrievals. Table 1 lists those that we explored using the Looe Key HSI/LIDAR data.

Table 1. Metrics for measuring the closeness of two spectra i and j. Here x_{ik} means R_{rs} spectrum i at wavelength k, and the sums are over wavelength. The two spectra that give the minimum value of the metric are the closest.

Vector Distance Separation	Equation
Euclidean: Sum (over wavelength) of squared point distances	$\sqrt{\sum_{k=1}^{n}(x_{ik}-x_{jk})^2}$
Manhattan: Sum of absolute point distances	$\sum_{k=1}^{n}\lvert x_{ik}-x_{jk}\rvert$
Chebyshev: Closest absolute maximum point distance	$\max_{k}\lvert x_{ik}-x_{jk}\rvert$
Canberra: Sum of absolute point distances divided by absolute point values	$\sum_{k=1}^{n}\dfrac{\lvert x_{ik}-x_{jk}\rvert}{\lvert x_{ik}\rvert+\lvert x_{jk}\rvert}$
Bray Curtis: Sum of absolute point distances divided by sum of absolute point values	$\dfrac{\sum_{k=1}^{n}\lvert x_{ik}-x_{jk}\rvert}{\sum_{k=1}^{n}(x_{ik}+x_{jk})}$

Vector Angle Separation	Equation
Angular Separation: Cosine angle between two vectors	$\dfrac{\sum_{k=1}^{n}x_{ik}\cdot x_{jk}}{\sqrt{\sum_{k=1}^{n}x_{ik}^2\cdot\sum_{k=1}^{n}x_{jk}^2}}$
Correlation Coefficient: Cosine angle between two vectors where the coordinates are centered at the mean	$\dfrac{\sum_{k=1}^{n}(x_{ik}-\bar{x}_i)\cdot(x_{jk}-\bar{x}_j)}{\sqrt{\sum_{k=1}^{n}(x_{ik}-\bar{x}_i)^2\cdot\sum_{k=1}^{n}(x_{jk}-\bar{x}_j)^2}}$

Figure 8 shows the results over the joint HSI/LIDAR coverage area. It is evident in this data set that vector distance is the best measure for goodness of fit, and that the Manhattan distance calculation is the best calculation. There were some areas that were fit best by angular separation. These areas were primarily best fit by the Correlation Coefficient. The distance measure takes into consideration the magnitude of the R_{rs} signal, as well as the spectral shape. Its use is only appropriate where there is

high confidence in the calibration of the sensor measured water-leaving radiance and the subsequent removal of atmospheric and illumination effects. Most spectrum matching techniques avoid the use of the magnitude because of problems in sensor calibration and atmospheric correction. We demonstrate here the enhanced retrieval resulting from the distance measurements possible with R_{rs} data using spectra retrieved from a high confidence sensor and processing.

Figure 9 shows the best fit using only Manhattan and Correlation Coefficient measures. It can be seen that there are areas where the Correlation Coefficient achieves better matches than the Manhattan. This is primarily seen in shallow water areas where the magnitude of the signal is strong enough to allow for the secondary effects of spectral shape to provide a more enhanced fit. Table 2 summarizes these results.

Table 2. Depth errors for various metrics, averaged over the entire image.
RMSE = Root Mean Squared Error (meters), ABSE = Absolute Mean Squared Error (meters).

- Manhattan
  - RMSE = 2.25
  - ABSE = 1.79
- Correlation Coefficient
  - RMSE = 5.97
  - ABSE = 4.54
- Combination Man+Cor
  - RMSE = 2.31
  - ABSE = 1.68
- Best Possible Method
  - RMSE=2.07
  - ABSE = 1.43

IMPACT/APPLICATION

The problem of extracting environmental information from remotely sensed ocean color spectra is fundamental to a wide range of Navy needs as well as basic science and ecosystem monitoring and management problems. Extraction of bathymetry and bottom classification is especially valuable for planning military operations in denied access areas. The ability to simultaneously generate error estimates on retrieved values is often equally important to the ability to retrieve the environmental information itself.

TRANSITIONS

Various databases of water IOPs, bottom reflectances, and the corresponding R_{rs} spectra, along with the specialized Hydrolight code and spectrum-matching algorithms have been transitioned to Dr. Paul Bissett at the Florida Environmental Research Institute for processing his extensive collection of SAMPSON imagery acquired in coastal California waters, and for use in comparisons of LUT and LIDAR bathymetry. The LUT software was also used in processing imagery acquired by the Naval Research Laboratory and Battelle Northwest National Laboratory in Sequim Bay, Washington.

RELATED PROJECTS

This work is being conducted in conjunction with Dr. Paul Bissett of FERI, who is separately funded for this collaboration. I have also received separate funding for the application of the LUT methodology to imagery obtained in Puget Sound.

REFERENCES

Bissett, W.P., DeBra, S., Kadiwala, M., Kohler, D.D.R., Mobley, C.D., Steward, R.G., Weidemann, A.D., Davis, C.O., Lillycrop, J. and Pope, R.L., 2005. Development, validation, and fusion of high-resolution active and passive optical imagery. In: I. Kadar (Editor), Signal Processing, Sensor Fusion, and Target Recognition XIV. Proceedings of SPIE Vol. 5809. SPIE, Bellingham, WA, pp. 341-349.

Mobley, C. D., L. K. Sundman, C. O. Davis, T. V. Downes, R. A. Leathers, M. J. Montes, J. H. Bowles, W. P. Bissett, D. D. R. Kohler, R. P. Reid, E. M. Louchard, and A. Gleason, 2005. Interpretation of hyperspectral remote-sensing imagery via spectrum matching and look-up tables. *Applied Optics* 44(17), 3576-3592.

PUBLICATIONS

Lesser, M. P. and C. D. Mobley. Bathymetry, optical properties, and benthic classification of coral reefs using hyperspectral remote sensing imagery. *Coral Reefs* [Refereed, in press]
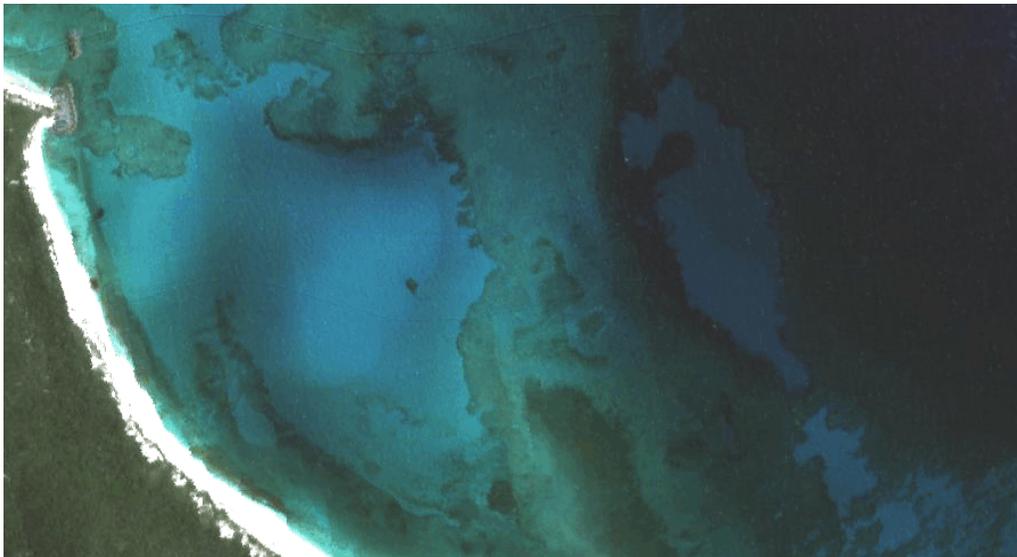
Figure 1. An RGB image of the Horseshoe Reef area made from a PHILLS hyperspectral image taken May 20, 2000. The bottom includes areas of highly reflecting ooid sands, low reflecting, dense sea grass beds, and low to intermediate reflecting areas of mixed sediments, corals, sea grass, turf algae, and macrophytes.
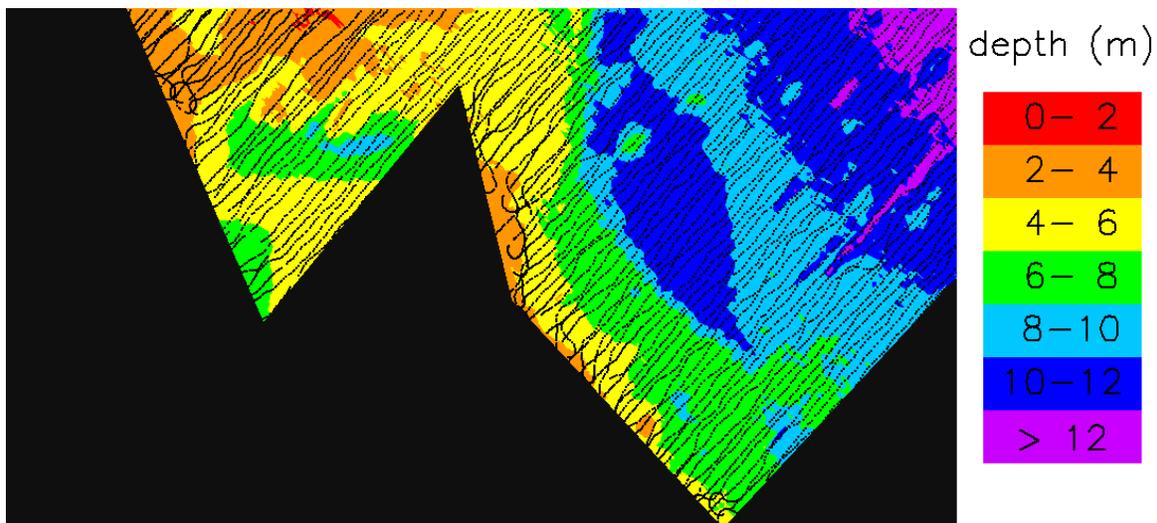
Figure 2. Acoustic bathymetry coverage for the area corresponding to Fig. 1. The black dots show the locations of the acoustic pings; the solid black area had no acoustic coverage. The acoustic depths are used for validation of the LUT-retrieved depths at the corresponding pixels. [The color coding identifies the depth, binned into 2 m bins for convenient viewing.]
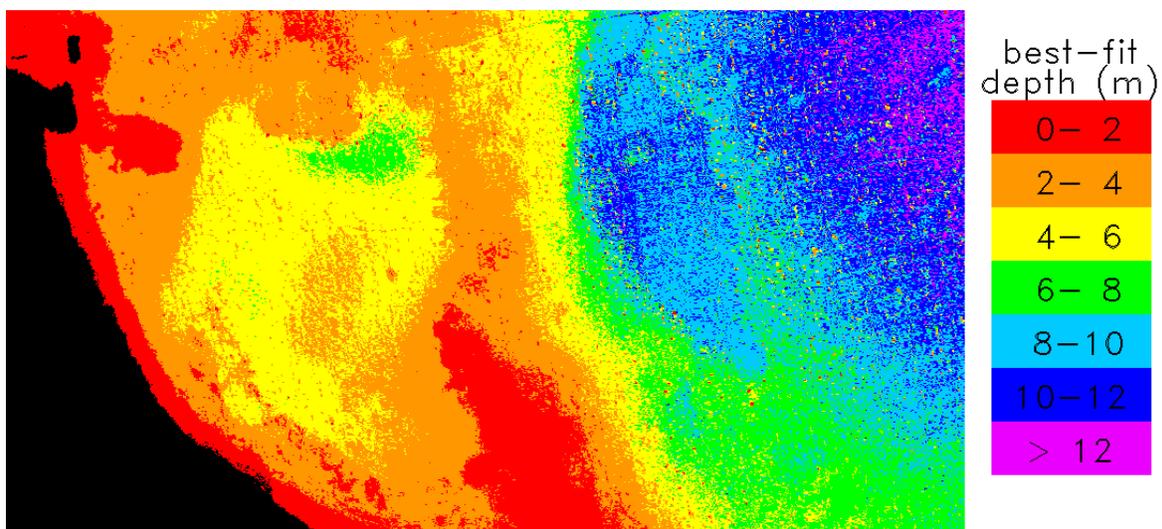

Figure 3. LUT depth retrieval obtained from the closest matching database spectrum (k = 1). [The color coding identifies the depth, binned into 2 m bins for convenient viewing.]
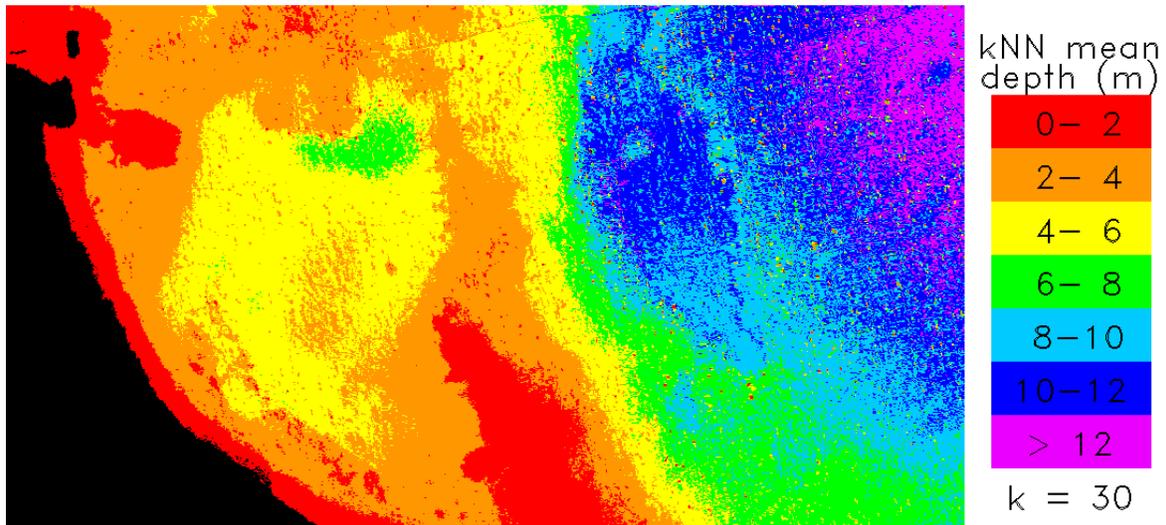
Figure 4. LUT depth retrieval computed as the mean of the k = 30 closest matching database spectra. Note that the retrievals are somewhat deeper than those for k = 1. [The color coding identifies the depth, binned into 2 m bins for convenient viewing.]
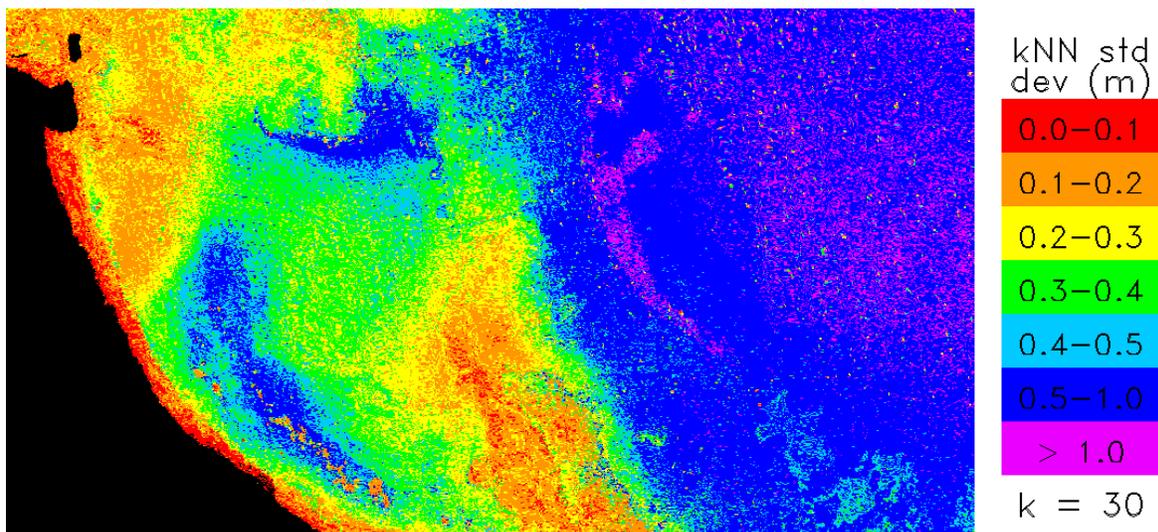

Figure 5. Standard deviation of the LUT depths for the k = 30 closest matching spectra. The standard deviations of the retrieved depths are greatest in deeper waters and in areas with dark bottoms.
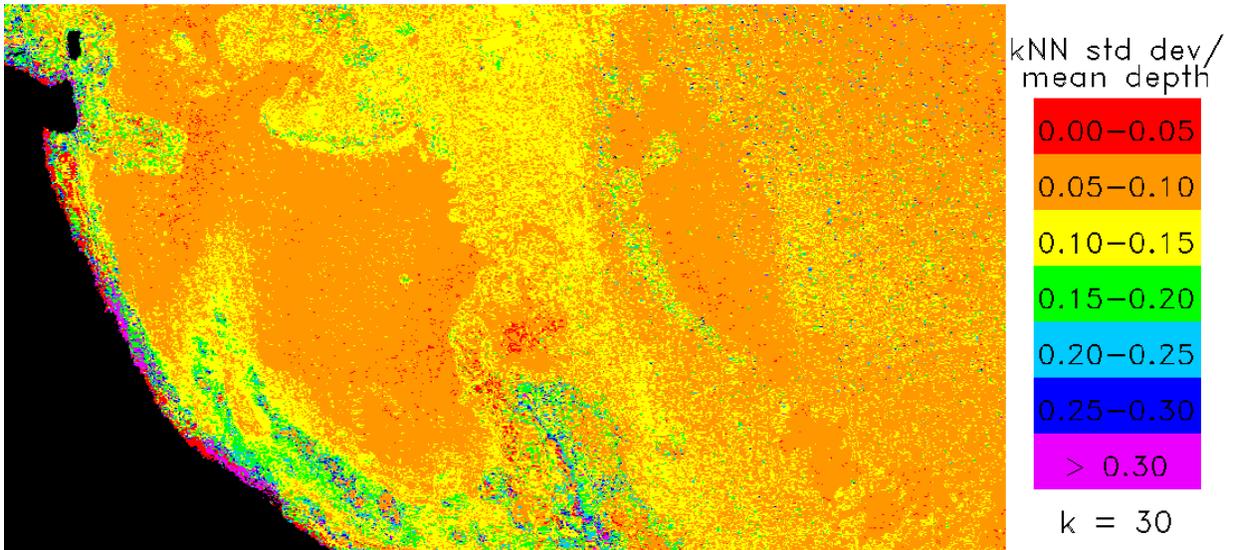
Figure 6. Ratio of the standard deviation (Fig. 5) to the mean depth (Fig. 4) for k = 30. This figure shows that the relative depth error is generally in the 0.05 to 0.15 range over most of the image. Greater relative errors occur over some areas of darker bottoms.
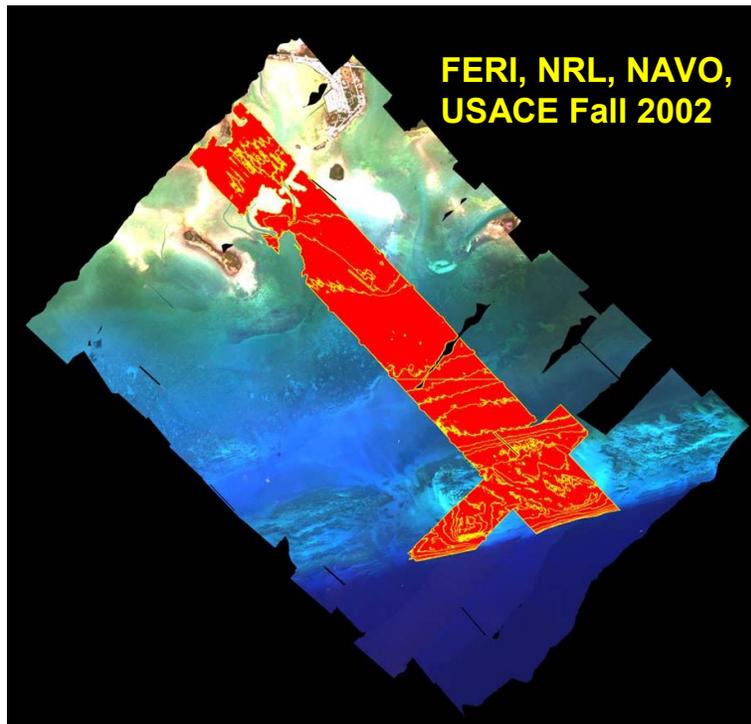

Figure 7. Joint HSI/LIDAR Experiment coverage area. The red area denotes region of joint data coverage.
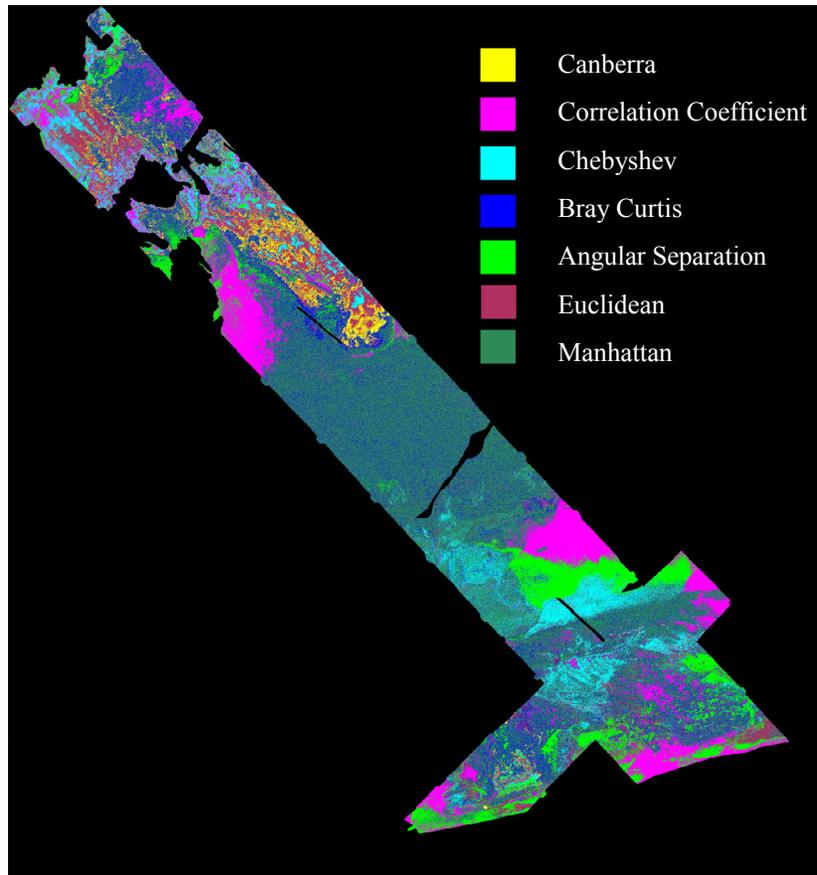
Figure 8. Visual results of statistical measure analysis of best fit. The color coding shows which distance metric gave the best agreement between LUT and LIDAR bathymetry. Overall, vector distance was better than vector angular separation in determining best fit. The Manhattan distance calculation was better than Euclidean. There were some areas where angular separation was better.

Figure 9. The same area as Figure 8, with only the best vector distance (Manhattan) and angle separation (Correlation Coefficient). This image shows that over shallow waters the Correlation Coefficient sometimes retrieved better matches than the Manhattan calculation. This results from the fact that the spectra magnitudes are large enough from the bright shallow water signal to allow for better certainty from the angular separation.