5-2004

Quantitative Inter-channel Calibration of SHOALS Signals for Consistent Bottom Segmentation and Characterization

Semme J. Dijkstra
University of New Hampshire, Durham, Semme.Dijkstra@unh.edu

G. Elston
University of New Hampshire, Durham

Quantitative Inter-Channel Calibration of SHOALS Signals for Consistent Bottom Segmentation and Characterization

Semme J. Dijkstra, Gareth R. Elston
Center for Coastal and Ocean Mapping
University of New Hampshire
24 Colovos Road, Durham. NH 03824
s.dijkstra@unh.edu

ACKNOWLEDGEMENT

We would like to acknowledge the support of NOAA, through grant NA17OG2285, and the USGS. Furthermore we thank Optech for help on any questions we have and the USACE for giving us access to their SHOALS Post Flight Processing System.

ABSTRACT

This poster presents a model based clustering method for the segmentation and subsequent calibration of SHOALS data. As features for the segmentation we make use of the estimated maximum power and pulse-width of each bottom return.

We describe the feature data by parameterized Gaussian mixture models and use EM (Expectation-Maximization) methods for clustering. The Bayesian information criterion is used to estimate the number of clusters and therefore the underlying number of individual classes.

To concurrently process data from both the deep PMT (Photo Multiplier Tube) and shallow APD (Avalanche Photo Diode) channels, it is imperative that the features extracted from each are consistent. To this end we developed a quantitative inter-channel calibration procedure that adjusts the data from the PMT to match the APD data.

The SHOALS data used in this project was obtained from Lake Tahoe in July 2000 to supplement a USGS multibeam sonar survey.

We would like to acknowledge the support of NOAA, through grant NA17OG2285, and the USGS. Furthermore we thank Optech for help on any questions we have and the USACE for giving us access to their SHOALS Post Flight Processing System.

INTRODUCTION

The SHOALS LIDAR system was originally developed as a tool to improve speed and efficiency of bathymetric data collection for the production of nautical charts (Guenther et al., 1989). This system uses four channels for data collection, for each of which a time series is stored. Two of the channels are used to determine the water surface return and two more are used for mapping the bottom return: The Avalanche Photo Diode (APD) channel for shallow water and the Photo Multiplier Tube (PMT) channel for deeper water. A number of efforts are underway to extract bottom type information from LIDAR data as an added data product. Some promising results have been achieved mapping pseudo-reflectance values through inversion of the bathymetric LIDAR equation (Tuell et al., 2004). Philpot and Wang have presented work to segment bottom types based on pseudo-reflectance values (Philpot et al., 2002).

The work presented here is part of our initial effort to automatically segment bottom types based features extracted from LIDAR data. Our effort consists of several stages: 1) return characterization; 2) feature extraction 3) feature conditioning; 4) Feature segmentation. In the future we will need to add a classification and verification stage.

Return characterization is achieved through a parameter-estimation approach that in the near future will be combined with a curve-fitting approach (Elston et al, 2004). The segmentation procedure is complicated by the fact that the observed intensities from the APD and PMT channel are not consistent (Tuell, 2003). This poster focuses on our approaches to the relative alignment of data from both the APD and PMT channels.

The data presented are from a SHOALS data set obtained in July of 2000 at Emerald Bay as part of a USGS survey of Lake Tahoe, California.

DATA COLLECTION

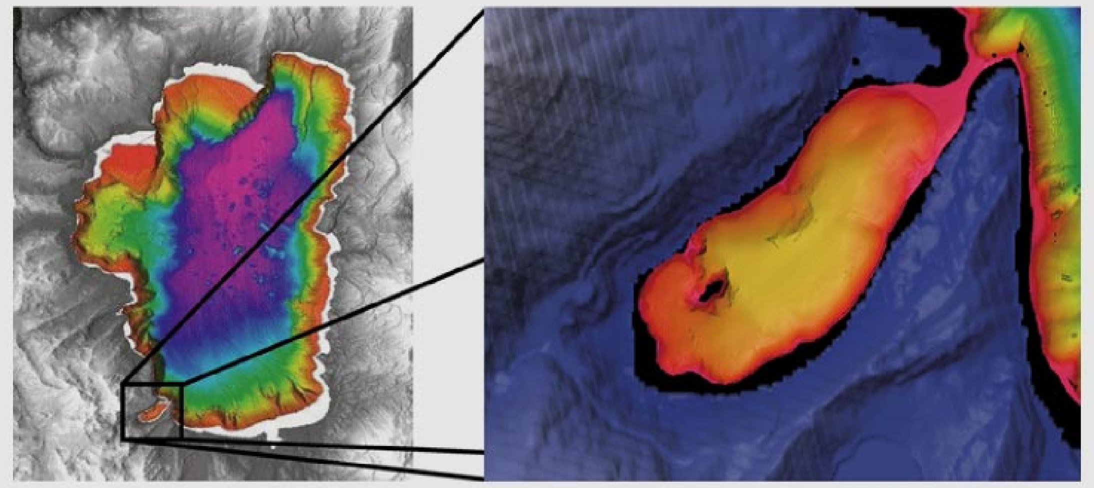


Figure 2. Emerald Bay is located at the Southwest end of Lake Tahoe, CA (left). All the data presented here were obtained in Emerald Bay (right).

The SHOALS lidar data we analyze was collected for the USGS to complete the near-shore regions of a multibeam sonar survey of Lake Tahoe (Gardner, 1998). Figure 2 shows the land elevation and sonar bathymetry DTMs for Lake Tahoe with a more detailed view of Emerald Bay. This is a challenging data set because the bathymetry contains areas of steep slope, rocky outcrops, and possibly a number of bottom types; in addition, our analysis reveals the presence of regions of water with different optical properties.

Peak power and pulse width features are extracted (Elston, 2004) from the normalized waveforms stored in the optional diagnostic output files from the SHOALS Post-Flight Processing System (PFPS).

To ensure that full pulses are analyzed, the waveforms to be processed are chosen from the APD when the depth is < 11 m, and from the PMT when the depth is > 8 m. This provides a region of overlap between the channels to allow their peak intensities to be matched before the clustering and bottom segmentation algorithms are applied.

FEATURE SELECTION AND IMPULSE RESPONSE MODELING

The model for round trip bottom impulse response may be described by a gamma function (Thomas et al, 1979). These functions are well described by the location of the peak and inflection points. We therefore selected peak power and the distance between the inflection points as the primary features for the bottom return characterization and subsequent segmentation (Figure 1 and 4). In this initial study we use a parameter estimation approach to determine the location of the points; in the near future this will be augmented with a gamma function fitting approach (Elston et al, 2004).

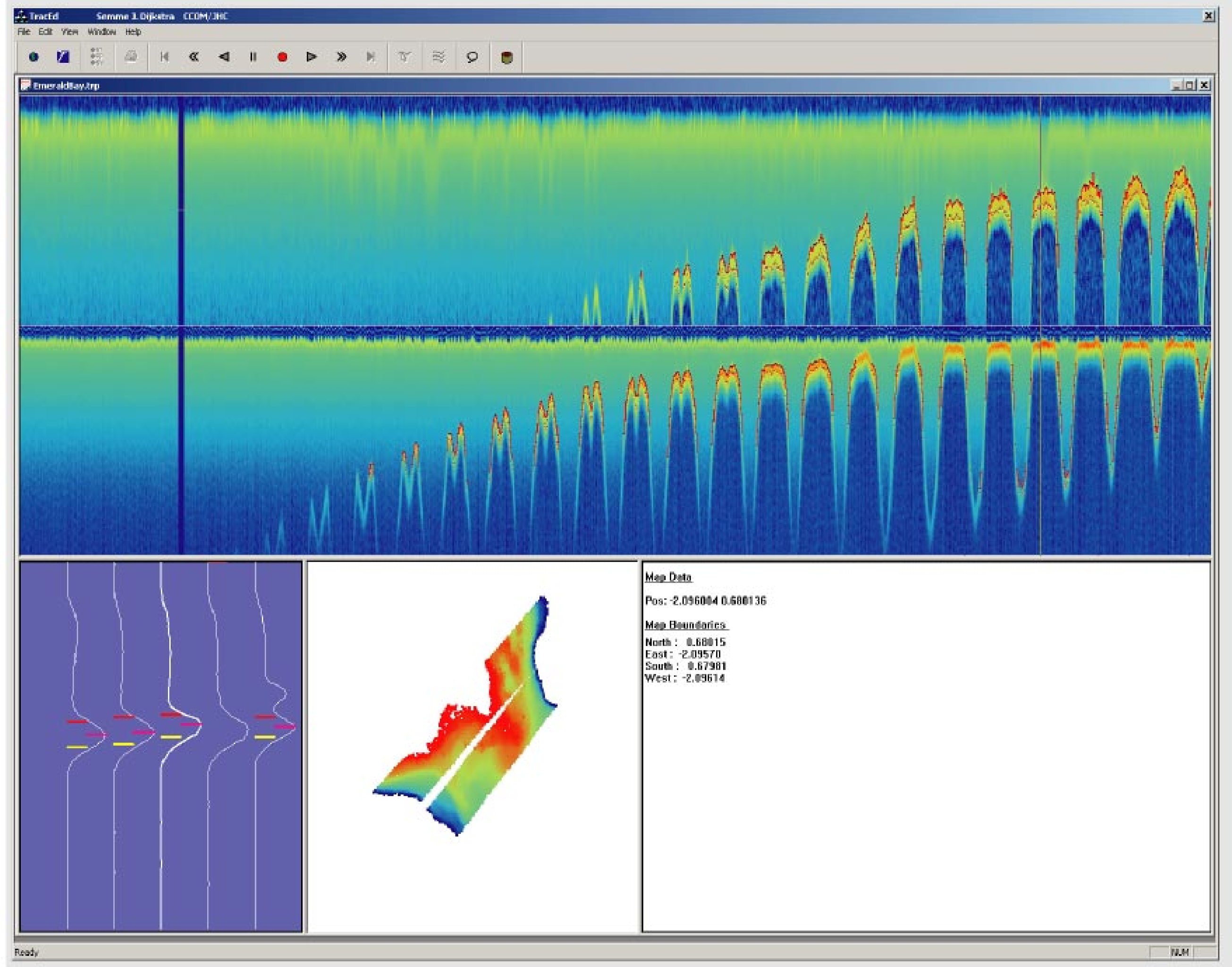
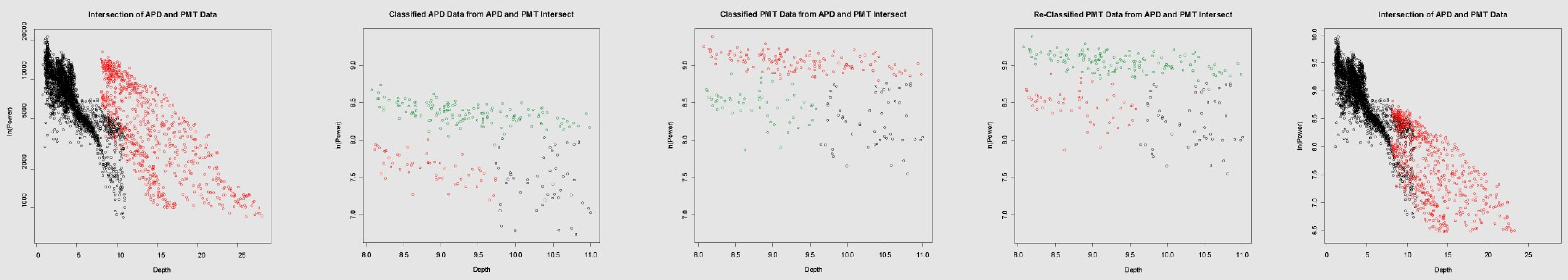


Figure 1. TracEd (Dijkstra, 2000) visualization of SHOALS data. Normalized time series from the APD and PMT channel are shown in the top half of the display. Estimated locations for peak power and inflection points are shown superimposed over individual traces in the lower left corner display. The lower middle display shows color-coded bathymetry imported from the SHOALS processed data.

APD and PMT Inter-Channel Calibration



Shown is the mismatch in the natural log of peak power vs. depth data from the APD (black) and PMT (red) channel.

A segmented subset of data from the APD channel obtained from the overlapping region.

Segmented subset of the co-occuring data points from the PMT channel

The PMT channel data is re-classified based on the correlation of the clusters to those from the APD channel

Data from both the APD (black) and adjusted PMT (red) channel. Note the good overlap of the data and the two distinctive decaying trends caused by the presence of two water masses with varying attenuation coefficients

Figure 3.

In the data that we have available to us there is a mismatch between the data from the APD and PMT channels (Figure 3.), which has also been observed by others (Tuell, 2003). The first step in the data conditioning sequence applied by us is to match the data from the APD channel to the PMT channel using in-sample statistics. An option would be to simply perform a linear least squares fit to the data in the overlapping region in semi-log space. However, in this case the residuals would be significant due to the presence in the overlapping region of two distinct water-masses with varying attenuation coefficients k and distinct bottom types in terms of the reflection coefficient.

The approach taken here is to cluster the data for the APD and PMT channel independently for the overlapping region. Next the resulting clusters from the APD and PMT channels are matched and if they correlate well (R > .9) a linear least-squares fit is performed for each of the matched clusters. A weighed mean of the coefficients is determined for each channel where the number of samples populating a cluster determines weight. Subsequently all the PMT data are adjusted in an iterative loop to match the APD data using the convergence of the fitted coefficients as a criterion.

DATA CONDITIONING AND MODEL BASED CLUSTERING

Before submitting the derived peak power and pulse width to a mixture-modeling algorithm for segmentation we need to remove any trends i.e., systematic dependencies and artifacts. In the data discussed here this pre-conditioning includes the removal of a mismatch between the APD and PMT channel data (Figure 4). The process of estimating the mismatch between the channels includes a segmentation of the data in the overlapping region to correct for the effect of varying bottom type and, notably in the data presented here, varying optical properties of the water. For this purpose we use an unsupervised segmentation of power versus depth data which we describe by finite, parameterized Gaussian mixture models. They can easily be fit by likelihood maximization using an EM (Expectation-Maximization) method (Dempster et al., 1977).

For bottom type classification it is typically difficult to obtain both an a-priori estimate of the model dimension and ground-truth for training purposes. Therefore we use an unsupervised classification approach. A significant problem is then to estimate the ideal model dimension using an in-sample estimate. For this purpose there are several options available, prime among which are the Akaike Information Criterion (AIK) and the Bayesian information criterion (BIC) (Schwartz, 1984). We choose the BIC because the AIK has been proven to be inconsistent (Kashyap, 1980), while the BIC has been shown to be consistent for mixture models (Keribin, 1997). A disadvantage is that the BIC carries a penalty for model complexity and therefore tends to underestimate the model dimension - In the future we would like to further enhance the approach by using an iterative process that includes similarity analysis.

For all the processing described in the following paragraphs we have used "MCLUST" (Fraley, 2002) software routines called from the Lassoo visualization and segmentation software (Dijkstra, 1996). Note that we have used MCLUST's option of determining the BIC for various possible covariance matrices and in each case have ended up using ellipsoidal distributions of varying volume, shape and orientation.

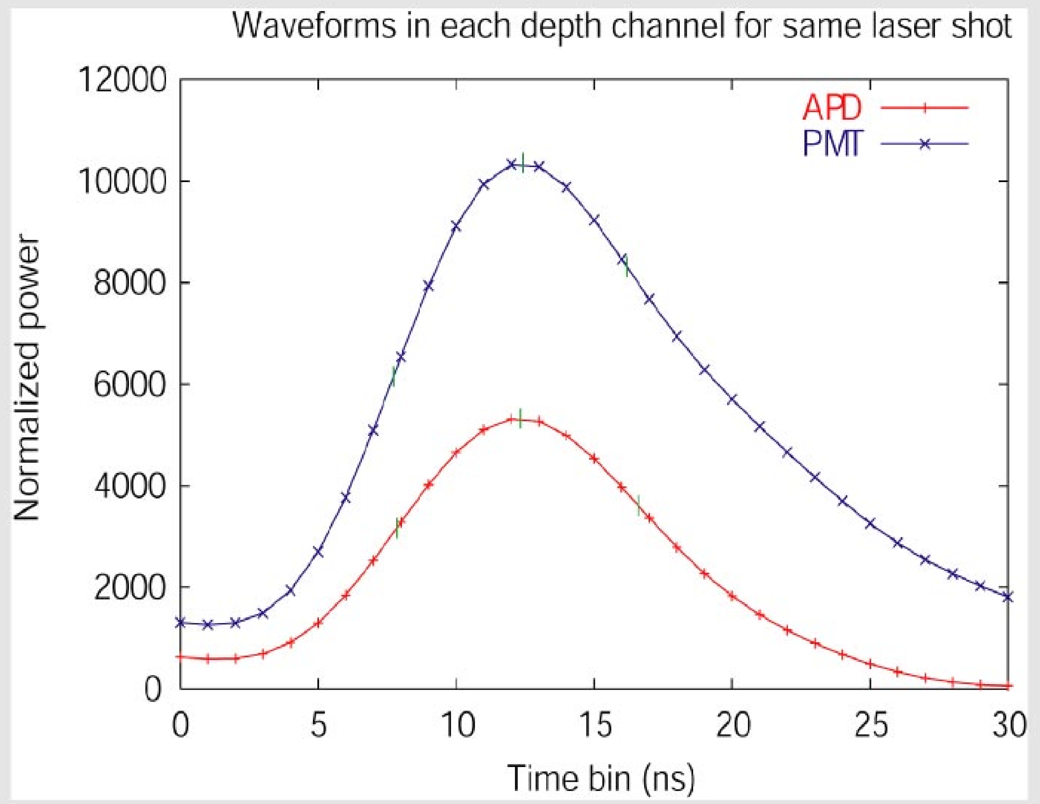


Figure 4. Shown is the normalized power vs. depth data from both the APD and PMT channel. Note the mismatch in the observed intensities between the channels. Also shown (in green) are estimates for the location of peak power and inflection points.

REFERENCES

Dijkstra, S. J., Mayer, L.A. (1996), Lassoo! An interactive sediment classification system. Proceedings IEEE Oceans '96, Ft. Lauderdale, FL-USA

Dijkstra, S. J., Mayer, L.A. (2000), TracEd: a remote acoustic seafloor characterization system for use with vertical incidence echosounders, Proceedings IEEE/MTS Oceans 2000, Providence, RI-USA

Dempster, A.P., Laird, N.M., Rubin, D.B. (1997). Maximum likelihood from incomplete data via the EM algorithm (with discussion). Journal of the Royal Statistical Society B 39, 1-38.

Elston, G.R., Dijkstra, S.J. (2004). Robust characterization of SHOALS LIDAR Signals for Bottom Segmentation and Classification: A Combined Parameter-Estimation and Curve-Fitting Approach, Proceedings ASPRS 2004 Annual Conference, Denver CO.-USA

Fraley, C., Raftery, A.E. (2002). MCLUST: Software for model-based clustering, density estimation and Discriminant Analysis. Technical Report No. 415, Dept. of Statistics, University of Washington, Seattle, WA.-USA

Gardner, J.V., Mayer, L.A., Hughes-Clarke, J. (1998). The bathymetry of Lake Tahoe, California-Nevada. Open-File Report 98-509, USGS. http://blt.wr.usgs.gov/openfile.html

Guenther, G.C. (1985). Airborne laser hydrography: system design and performance factors. Library of Congress Catalog number 85-600602, NOAA professional paper series, Springfield, VA, USA

Guenther, G.C. (1989). Airborne laser hydrography to chart shallow coastal waters. In: Sea Technology, vol. 30, no. 3, pp. 55-59.

Kashyap, R. (1980). Inconsistency of the AIC rule for estimating the order of autoregressive models, IEEE Trans. Auto. Control, Vol. AC-25, No. 5, pp. 996-998.

Keribin, C. (1997). Consistent estimation of the order of mixture models." Tech. rep., Universite d'Evry-Val d'Essonne, Laboratoire Analyse et Probabilite, France

Philpot, W.D., Wang, C. K. (2002). Using SHOALS LIDAR system to detect bottom material change, Final Report, Hydrographic Science Research Center, Univ. of Southern Mississippi, Agreement # USM-GR00662-07

Schwartz, G. (1978). Estimating the dimension of a model. The Annals of Statistics, Vol. 5, No. 2, pp 461-464.

Steinvall, O. K., Koppari, K. R. (1996). Depth sounding lidar: an overview of Swedish activities and future prospects. Laser Remote Sensing of Natural Waters: From Theory to Practice: CIS Selected Papers, SPIE, Vol. 2964

Thomas, R.W.L., Guenther, G.C. (1979). Theoretical characterization of bottom returns for bathymetric LIDAR. Proceedings International Conference on Lasers '78, Orlando, FL.-USA

Tuell, G. (2003). Data Fusion of airborne laser data with passive spectral data. Proceedings 4th Annual Airborne Hydrography Workshop, Fort Lauderdale, FL-USA (In press)

Tuell, G., Guenther, G., Feygels, V., Park, J.Y. (2004). Estimations of bottom reflectance in SHOALS data through inversion of the bathymetric laser power equation, Proceedings International LIDAR Mapping Forum 2004, Orlando FL-USA (In press)