Position paper submission to
ICA Workshop on Advances in Sensors and Algorithms for Topographic and Thematic Mapping

Algorithms for terrain and bathymetric sensor data

W. Randolph Franklin, Zhongyi Xie, Eddie Lau, You Li
Rensselaer Polytechnic Institute, wrf@ecse.rpi.edu, 518-276-6077, http://wrfranklin.org/

1 Abstract

We present three algorithic advances and a research topic in processing topographic and bathymetric sensor data. They are *(i)* lossy terrain compression that maintains slope accuracy, *(ii)* bathymetric surface fitting to irregular tracklines, *(iii)* lossy compression of 5D environmental data, and *(iv)* terrain modeling to maintain hydrological validity.□ The purpose is to attack several issues raised by the large amounts of data now available, with an eventual goal of a unified system.

2 Lossy terrain compression that maintains slope accuracy

Sensors such as IFSAR and LIDAR capture such quantities of terrain elevation data that it must be compressed. The data's limited accuracy justifies lossy compression. The appropriate metric to evaluate the compression should be more than simply RMS elevation error, but should consider the data's intended application. Many applications use the terrain's *slope*, which is relevant to erosion, trafficability and skyline recognition. Unfortunately, while slope is theoretically the derivative of the elevation, lossy compression amplifies errors so that the slope computed from the lossily compressed terrain may be considerably different from the slope computed from the original terrain, even though the elevations may be relatively accurate. Indeed, mathematicians are well aware that an accurate approximation $\hat{f}(x)$ to a function $f(x)$ says nothing about $|\hat{f}'(x) - f'(x)|$. Nevertheless, perhaps because of the incorrect assumption that accurate elevations imply accurate slopes, there appears to be no prior art on this topic.

This project is to lossily compress terrain so that the restored terrain's slope is also accurate, but without storing the explicit slopes. The method is an extension of our lossly terrain compression algorithm, ODETLAP[2]. ODETLAP is two things: *(i)* an algorithm to select $\mathscr{P}$ a representative set of points on a terrain, and *(ii)* an algorithm to reconstruct the terrain from those points.□ The selection algorithm is greedy, repeatedly augmenting $\mathscr{P}$ with the worst points on the existing surface and recomputing a new surface. The novel aspect here is also to select points where the reconstructed surface's slope is inaccurate. Assorted technical details must be solved to make this work. However, the results are impressive. Figures 1 and 2 show some sample results.

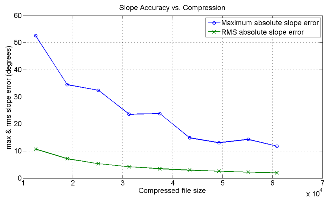

Fig. 1. Slope accuracy under lossy compression

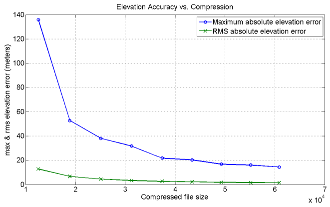

Fig. 2. Elevation accuracy under lossy compression

3 Bathymetric surface fitting to irregular tracklines

This problem is to reconstruct seafloor surface from multibeam bathymetry (MBB) data, as shown in Figure 3. This is hard because *(i)* MBB data is very unevenly spaced (dense in a swath along the ship tracklines, but then nonexistent for a long distance sideways), *(ii)* depth accuracy is a few percent, and *(iii)* there is insufficient data to infer features that are probably there.□ Current methods often *(i)* have a specific distance wired into the formula, *(ii)* do not let information flow past data points, and so *(iii)* produce artifacts (e.g., abrupt slopes, acquisition footprint), see Figure 4, and *(iv)* show details that aren't justified.□

The goal here is to spread the influence of a data point over a larger distance when the data points are more spread out. However proximity polygons are unacceptable because of their bad results. The new idea is to extend ODETLAP by varying R, its smoothness parameter over the dataset, and using smaller values when the known points are sparser. That should reduce the mean absolute error, exhibiting fine details in the generated surface, while simultaneously reducing obvious artifacts.

Using a lower R where the known points are more spread out will better preserve the depth fluctuations. The higher R where the known points are denser will smooth the surface more to reduce artifacts caused by noise in the data. That will de-correlate the trackline artifacts from how way data is collected.

The computation time can be considerably reduced by converting the overdetermined system to a normal equations problem, and then solving using Choleskey factorization. For instance, processing a 601×601 terrain with 20000 known points takes about 30 seconds in Matlab on a dual-core 2.1GHz laptop.

4 Lossy compression of 5D environmental data

Sensors, e.g., in the World Ocean Atlas 2005, are collecting multiple bands of environmental data, such as temperature, salinity, and oxygen concentration, at many times over a 3D grid, to produce a set of values over the 5D grid (x, y, z, t, b).

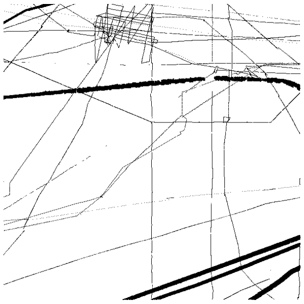

Fig. 3. Uneven bathymetry trackline data

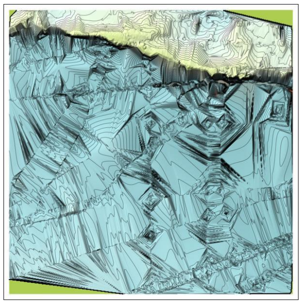

Fig. 4. Poor fitting with a second-order spline

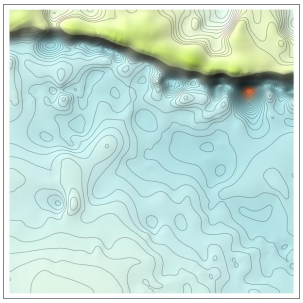

Fig. 5. Surface computed with variable-R (10 to 30) ODETLAP

We wish to compress this data, but see little prior art. Our research is based on the following principles: *(i)* to assume that if one band has a large derivative at a particular (x,y,z,t) then likely the other bands also will, *(ii)* to treat the data as one 5-D dataset (rather than as many unrelated 3D datasets), and *(iii)* to compress lossily since the data is imprecise.

Our technique is an extension of ODETLAP to 5D[1]. The major challenge is that everything is harder in higher dimensions. We are already seeing compression ratios of 100:1 with mean errors of under 1.5%. ODETLAP is an extension of a Laplacian partial differential equation to an overdetermined sparse linear system. Our innovation of using *overdeterminism*, which makes the solution a nonlinear function of the input, completely changes the system's mathematical properties, e.g., by now allowing the formation of local extrema. These new properties allow an unprecedented compression level. Table 1 shows some sample results.

	3D-ODETLAP				3D-SPIHT		
Variable	Mean Error(%)	Max Error(%)	Compressed File(bytes)	Compression Ratio	Mean Error(%)	Max Error(%)	Compression Ratio
Salinity	0.0532	0.2174	27,377	77:1	0.0530	0.4946	11:1
Temperature	0.4993	2.0673	21,508	98:1	0.50	17.91	135:1
Dissolved O_2	0.9993	4.4145	21,040	100:1	1.002	24.9965	71:1
Apparent O_2 util.	0.9999	4.0170	24,775	85:1	0.9991	20.3609	81:1
Percent O_2 satur.	0.9985	4.5672	26,743	78:1	0.9969	20.3610	65:1
Phosphate	0.9993	4.5241	24,493	86:1	0.99784	15.6922	65:1
Nitrate	1.0242	4.6946	31,905	66:1	1.0006	18.5360	59:1
Silicate	0.9996	5.1437	23,076	91:1	1.0018	21.6457	81:1

Table 1. ODETLAP's smaller compression error than SPIHT

5 Terrain modeling to maintain hydrological validity

Our goal is to model "valid" terrain, whatever that means, from the sensors' data. One measure of validity, for terrain above sea level, is that it look like it was formed by surface water flow. This requires the study of nonlinear terrain formation operators that produce hydrologically-valid terrain. Nonlinearity is powerful, but difficult to study. The strategy will be to design operators that are based on how the terrain was formed. They will then be validated by tests on large terrain databases.

Any research into hydrological representations must acknowledge the properties of terrain including the following. *(i)* In places, the terrain is C^{-1}, i.e., discontinuous. Indeed, although techniques such as contour lines have difficulty representing them, these may be the most important features for many users. They certainly affect mobility and erosion. *(ii)* Terrain is not symmetric in Z. Above sea level, there are many local maxima but few local minima, since they usually fill in and become lakes. *(iii)* Water and land are different. Given the polylines for a group of adjacent coastlines, without knowing which region is the land and which water, it is often possible to tell.

References

[1] Tsz-Yam Lau, You Li, Zhongyi Xie, and W. Randolph Franklin. Sea floor bathymetry trackline surface fitting without visible artifacts using ODETLAP. In *17th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems (ACM SIGSPATIAL GIS 2009)*, Seattle WA USA, 4–6 Nov 2009.

[2] Jared Stookey, Zhongyi Xie, Barbara Cutler, W. Randolph Franklin, Daniel M. Tracy, and Marcus V.A. Andrade. Parallel ODETLAP for terrain compression and reconstruction. In Walid G. Aref et al., editors, *16th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems (ACM GIS 2008)*, Irvine CA, 5–7 Nov 2008. `http://acmgis08.cs.umn.edu/`.