Eurographics Conference on Visualization (EuroVis) (2014)
N. Elmqvist, M. Hlawitschka, and J. Kennedy (Editors)

Preprint of Poster presented on June 2014

Interactive Oceanographic Visualization using spatially-aggregated Parallel Coordinate Plots

Richard L. S. F. George[1], Panagiotis D. Ritsos[1] and Jonathan C. Roberts[1]

[1]School of Computer Science, Bangor University, UK

Abstract

Visual Analytics interfaces allow ocean scientists to interactively investigate and compare different runs and parameterizations. However, oceanographic models are complex, temporal and the datasets that are generated are huge. Parallel Coordinate Plots can help explore multivariate data such as ocean-science data. Common issues with traditional PCPs of clutter and performance inhibit interactive spatial exploration. We describe techniques that aggregates the PCP based on the spatial nature of the data and we render the polylines as ranges.

Categories and Subject Descriptors (according to ACM CCS): H.5.2 [Information Interfaces and Presentation]: User Interfaces—Graphical User interfaces (GUI)

1. Introduction

Visual Analytics interfaces have the potential to allow ocean scientists to query, interactively investigate and quantitatively compare different runs and various parameterizations of their models. The challenge for oceanographers is that their data models are complex, temporal and huge. Parallel Coordinate Plots have been shown to be useful to help explore high dimensional datasets (e.g., [BBP08]) and can be used on ocean-science data. However, the main difficulty with traditional PCPs are clutter and performance, which inhibits interactive spatial exploration.

With a typical oceanographic dataset that we are using, when we render the data to a standard HD screen then every pixel typically represents about 400 data points. So, when all this data is naively placed into a PCP it would be unusable, with a substantial amount of overplotted polylines. Moreover, it would be impossible to query this PCP interactively due to rendering complexity and long search time on the data. Binning techniques can be used to reduce clutter, but with these simplification algorithms the spatial context of the data is lost, which is important for oceanographic investigation.

We use a PR-tree (a point-based region tree) to aggregate the PCP based on the spatial nature of the data; this now gives us a range of values in the PCP rather than one polyline. We visualize these polylines as ranges instead of individual polylines and therefore render poly-ranges.

2. Related Work

Our work focuses on coastal shelf region, which is of interest to scientists, especially with climate-change predictions suggesting that the sea level will rise between 12 to 14cm, which would effect 50% of the population [TD06, DT08]. Forecasting change is difficult and not well researched [DT08], and historic changes may not be good indicators of what will happen in the future [BJK*06]. However, challenges with oceanographic visualizations are not unique (e.g., see Lipsa et al. [LLC*11]). In addition, oceanographers still use predominantly static visualization [AAD*10]. Our motivation is to develop interactive oceanographic visualization tools.

We utilize several linked views [Rob07] in order to interact and explore the estuary model. PCPs have been used to explore similar datasets (e.g., volume rendering [TPM05]), but overplotting is a major challenge. Typically, two strategies have been used by researchers to reduce the amount of points, and to provide a rendering that is representative of the data: (1) Visual space (e.g., [JC08, GPS*11]), and (2) data space (e.g., [BTK11]). Fau et al. [FWR99] present a theoretical framework of hierarchical parallel coordinates, and use the BIRCH algorithm to group the data hierarchically. Our PR-quadtree method fits in their framework as an *implicit* hierarchical clustering strategy. While other researchers have focused on *explicit* techniques, that cluster the data in abstract spaces (through, say, learning algorithms [BH03, AA04]). Few developers use spatial structures to relate the spatial nature of the data to visualizations.

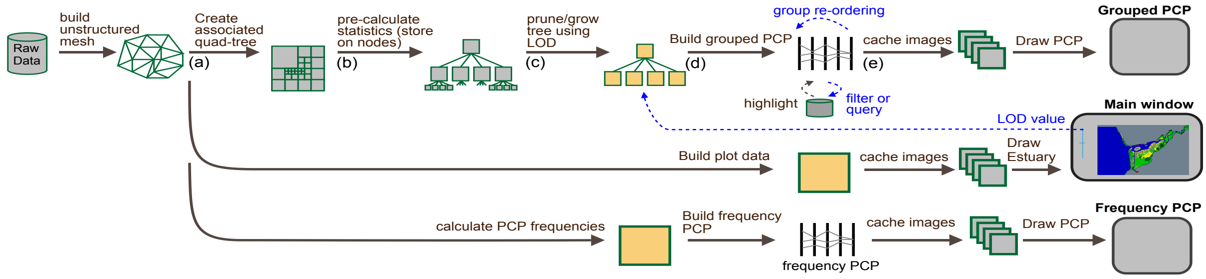

Figure 1: *The schematic demonstrates the whole process.*

3. Our methodology

Our system uses the PR-quadtree linked with a geospatial plot of the data to interact and explore the estuary simulations. We employ a data-space method, but aggregate the information in the geo-spatial domain based on the level-of-detail in the data. This is achieved through creating several associated (linked) data-structures.

The stages are shown in Fig. 1, as follows: First we load and store the unstructured data (a), then link these datapoints to a tree (we have implemented both a quadtree and kd-tree, but the oceanographers prefer the quadtree variant, because we also link a plot of the quadtree that is used for selection; and the structure of the quadtree is clearer and easier to understand). In fact we use PR-quadtree, at this time we calculate statistical information on the nodes of the tree (b). Then the user can choose the LOD to display the data within the PCP by pruning or growing the tree (c). We group similar axis together such that the user can manipulate them as a unit (users can ungroup axis, or regroup as necessary). We build all the data in the PCP including the spatial dimensions (x,y and depth) (d). This enables us to aggregate the data in the PCP based on the LOD of the data, rather than the frequency of the data on the PCP. We render the images onto an off-screen buffer (e) that can then be loaded to create a fast animation of the data over time. We also create other views of the data including a frequency PCP (based on [NH06]).

We have implemented this process into our multiple-view oceanographic visualization system, Figure 2, which uses Java, OpenGL and Processing.org. Our tool has several coordinated-views, including a main plot window, several PCP variants and line graphs. Users can select a point on the main spatial view to load a detailed graph or draw a line on the window to load a line-graph to show the flux over that transect across the estuary. Brushing and time (animation) operations are coordinated across views. We have implemented different algorithms to render these poly-ranges, including sorting the ranges, using a painters algorithm to place the large ranges at the back and smaller ones at the front (Fig. 3 top left), showing the median or standard deviation in an envelope (Fig. 3 bottom left).

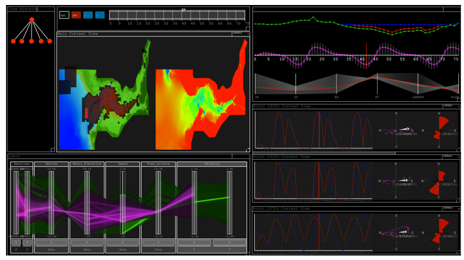

Figure 2: *The Vinca oceanographic visualization system*

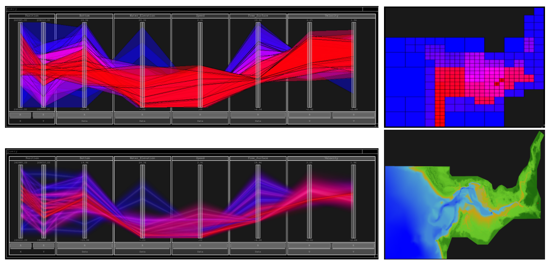

Figure 3: *Aggregated PCP, displaying the envelope using a painters algorithm (top left), and stdev (bottom left).*

4. Conclusions

We have described a PCP that is aggregated using a spatial PR tree. We have successfully implemented this work into our oceanographic visualization environment and have used a variety of rendering techniques and different transfer functions (examples shown in Fig. 3). We have visualized several different estuaries, to enable users to explore the simulation data. This ongoing work is a collaboration between expert ocean scientists and computer scientists and follows on from previous efforts to investigate big data in oceanography [GRDR12, GRD*14]. Further work will investigate how the data can be aggregated and how different PCP visualizations can help the oceanographers explore their data better; in fact one end-user said "the tool has gotten to a stage that it can be used to locate interesting features", and "we are keen to investigate flooding events and explore different parametrizations".

References

[AA04] ANDRIENKO G., ANDRIENKO N.: Parallel coordinates for exploring properties of subsets. In *Coordinated and Multiple Views in Exploratory Visualization, 2004. Proceedings. Second International Conference on* (july 2004), pp. 93 – 104. doi:10.1109/CMV.2004.1319530.

[AAD*10] ANDERSON J. C., ANDRES J. M., DAVIS M., FUJIWARA K., FANG T., NEDBAL M.: Voyager: An interactive software for visualizing large, geospatial data sets. *Marine Technology Society Journal 44*, 4 (2010), 8–19. doi:10.4031/MTSJ.44.4.2.

[BBP08] BLAAS J., BOTHA C., POST F.: Extensions of parallel coordinates for interactive exploration of large multi-timepoint data sets. *Visualization and Computer Graphics, IEEE Transactions on 14*, 6 (2008), 1436–1451. doi:10.1109/TVCG.2008.131.

[BH03] BERTHOLD M., HALL L.: Visualizing fuzzy points in parallel coordinates. *Fuzzy Systems, IEEE Transactions on 11*, 3 (june 2003), 369 – 374. doi:10.1109/TFUZZ.2003.812696.

[BJK*06] BROWN I., JUDE S., KOUKOULAS S., NICHOLLS R., DICKSON M., WALKDEN M.: Dynamic simulation and visualisation of coastal erosion. *Computers, environment and urban systems 30*, 6 (2006), 840–860. doi:10.1016/j.compenvurbsys.2005.08.002.

[BTK11] BERTINI E., TATU A., KEIM D.: Quality metrics in high-dimensional data visualization: An overview and systematization. *Visualization and Computer Graphics, IEEE Transactions on 17*, 12 (dec. 2011), 2203–2212. doi:10.1109/TVCG.2011.229.

[DT08] DAVIES A. G., THORNE P. D.: Advances in the study of moving sediments and evolving seabeds. *Surveys in Geophysics 29* (2008), 1–36. doi:10.1007/s10712-008-9039-x.

[FWR99] FUA Y.-H., WARD M. O., RUNDENSTEINER E. A.: Hierarchical parallel coordinates for exploration of large datasets. In *Proceedings of the conference on Visualization '99: celebrating ten years* (Los Alamitos, CA, USA, 1999), VIS '99, IEEE Computer Society Press, pp. 43–50. doi:10.1109/VISUAL.1999.809866.

[GPS*11] GENG Z., PENG Z., S.LARAMEE R., RICK W., C. ROBERTS J.: Angular histograms: Frequency-based visualizations for large, high dimensional data. *Visualization and Computer Graphics, IEEE Transactions on 17*, 12 (dec. 2011), 2572 –2580. doi:10.1109/TVCG.2011.166.

[GRD*14] GEORGE R. L., ROBINS P., DAVIES A. G., RITSOS P. D., ROBERTS J. C.: Interactive visual analytics of hydrodynamic flux for the coastal zone. *Environmental Earth Sciences* (2014). to be published.

[GRDR12] GEORGE R. L., ROBINS P., DAVIES A. G., ROBERTS J. C.: Interactive visual analytics of coastal oceanographic simulation data. In *Large Data Analysis and Visualization (LDAV), 2012 IEEE Symposium on* (2012), p. 2.

[JC08] JOHANSSON J., COOPER M.: A screen space quality method for data abstraction. *Computer Graphics Forum 27*, 3 (2008), 1039–1046. doi:10.1111/j.1467-8659.2008.01240.x.

[LLC*11] LIPSA D. R., LARAMEE R. S., COX S. J., ROBERTS J. C., WALKER R.: Visualization for the Physical Sciences. In *EG 2011 - State of the Art Reports* (Llandudno, UK, 2011), John N., Wyvill B., (Eds.), Eurographics Association, pp. 49–73. doi:10.1111/j.1467-8659.2012.03184.x.

[NH06] NOVOTNY M., HAUSER H.: Outlier preserving focus+content visualization in parallel coordinates. *IEEE Transacations on Visualization and Computer Graphics 12 (5)* (2006), 893–900. doi:10.1109/TVCG.2006.170.

[Rob07] ROBERTS J. C.: State of the art: Coordinated multiple views in exploratory visualization. In *Coordinated and Multiple Views in Exploratory Visualization, 2007. CMV '07. Fifth International Conference on* (July 2007), pp. 61–71. doi:10.1109/CMV.2007.20.

[TD06] THORNE P. D., DAVIES A. G.: Sounding out sediment transport: Observations and modelling of ripple generated vortices. *Ocean Challenge 14*, 3 (2006), 15–21.

[TPM05] TORY M., POTTS S., MOLLER T.: A parallel coordinates style interface for exploratory volume visualization. *Visualization and Computer Graphics, IEEE Transactions on 11*, 1 (jan.-feb. 2005), 71 –80. doi:10.1109/TVCG.2005.2.