Data Visualization and Analytics for Geoscience

IIS-GEO Workshop
Chris North, Virginia Tech

Ubiquitous geo-visualization: Geoscience data visualization is a mainstay in the data visualization field. Many relevant methods have been developed and initial basic theories established for visual representations and interaction techniques. A practical impact is the near ubiquitous availability of very detailed high-speed geographic visualization capabilities, such as Google Maps and Google Earth, and the corresponding APIs that support integration with other user data.

Interactive visual analytics: To support the process of gaining insight from big data, this growing research area emphasizes the integration of human cognition (via data visualization) and computational horsepower (via algorithmic analytical models). A critical challenge is to create usable forms of interaction between the two that support human-in-the-loop analysis, in which human and algorithm co-operate to iteratively develop and evaluate hypotheses from big data. Interactivity can enable geoscience experts to inject their expertise and intuition to explore alternative models and guide algorithms towards interesting findings in big geoscience data.

Semantic interaction: Semantic interaction in visual analytics emphasizes usability for analysts who are domain (geoscience) experts, without needing to also become algorithmic (computer science) experts. This is important as the data size in geosciences explodes and requires the ability to steer models of greater complexity at multiple levels of scale. The essential idea is shield domain experts from the complexity and pre-mature formality of the input parameters of analytical models. Instead, enable them to directly manipulate the meaningful output of the models in visual form, which then via computational machine learning methods, is transformed into operations on the inputs. A broader challenge of this approach is to develop systems that observe the human expert's sensemaking activity, learn about the expert's interests, and steer models to respond in meaningful ways.

Large-scale data visualization: Significant progress has been made in enabling interactive visualization of very large-scale data, along two research fronts. One front is the back-end computational support for steerable online massive data processing, simulation, and graphics rendering, such as GPU computing techniques for in-situ visualization. The second front is the front-end interactive visual support, using large-scale high-resolution display systems combined with usable novel interactive techniques, enabling multi-scale views of geo-data that effectively exploit human capabilities for visual processing, spatial cognition, and physical navigation.

Geo-referenced visual data integration: The increased availability of geo-tagged information has led to a new era of big data for geo-visualization applications. Integrating diverse data types enables novel visualization opportunities. For example, Microsoft's PhotoSynth integrates photos with geo-spatial information to produce realistic 3D walk-throughs of world locations.

Crowd sourced data collection and analysis: Advances in human-computer interaction (HCI) and social computing, enable vast data collection through involvement of citizens. Cell phones combined with social media applications such as Flicker have produced massive collections of geo-tagged visual media, ripe for processing. Such data is complementary to government and industry sponsored collections such as satellite imagery and Google's street view. Open challenges remain in putting the citizen person-power to work on the analysis of such data, and efforts are underway using crowd-sourcing tools such as Amazon's Mechanical Turk.