3-D Visualization and Cluster Analysis of Seismic Events

2004 Spring UROP Proposal
Benjamin J Kadlec

Summary:

The size of datasets in the geosciences are growing at an exponential rate due to improved instrumentation, larger numerical grids, and increasingly large storage devices. Analyzing this data is increasingly complicated as it becomes more detailed and multi-dimensional in space and time. This poses a problem for researchers who must sift through terabytes of data in order to discover useful relationships inside this information. New methods need to be applied to data-mining in order to simplify the analysis of this deluge of information. Cluster analysis has been used for a long time in pattern recognition and is now being used for investigating spatio-temporal relationships in earthquake data. The usefulness of this tool has demonstrated that small magnitude events produce local earthquake clusters which correspond to neighboring large events (Dzwinel et. al., 2003). Statistics on clustered earthquake data can only provide limited information, because relations between clusters are not easily seen numerically. In order to fully investigate the structures inherent in earthquake clusters, 3-d visualization of this data needs to take place. This visualization should recognize the intrinsic patterns and relationships between clusters while detecting the boundaries and spatial distribution of individual events.

I will continue my current research under Dr. David Yuen in visualizing geophysical data, specifically earthquake events, from both synthetic and real datasets. The synthetic data are obtained from four numerically simulated models taken from a catalog shown in (Ben-Zion, 1996) with earthquake events distributed in 2-dimensions of space. The real earthquake data represents actual events on the island of Japan, provided by the Japanese Meteorological Agency, in 3-dimensions of space spanning a range of 6 years. These and other data sets will be analyzed using clustering schemes from both self-made software and a packaged clustering toolkit. Different clustering and feature extraction procedures will be considered in order to determine the most appropriate technique for finding spatio-temporal relationships between earthquake events. Using the powerful visualization package Amira, these datasets will be interrogated working hand-in-hand with data clustering by visually extracting and clarifying earthquake cluster structures. For visualizing 3-D data, a novel technique will be developed to create a dynamic view in order to incorporate the time-dependencies of this data at different scales. These results will be applied to an
ongoing research project WEB-IS (Web-based interrogative system), which allows remote, interactive visualization of large-scale 3-D data over the internet, along with data analysis and data mining capabilities (Garbow 2003). This ongoing work is the result of collaborations between both graduate and undergraduate students and researchers at the University of Minnesota and abroad. I have recently helped present some results during a Data Mining Workshop at the Minnesota Supercomputing Institute (Yuen, 2003), and will be presenting additional results in December at the Fall Meeting of the American Geophysical Union (AGU) (Kadlec, 2003). Results obtained during the Spring research period may be presented in May at the 2004 Joint Assembly of AGU. Referenced papers have been written and submitted for publication on past research and results from this future work will be summarized in a paper as well.

Timetable:
I am currently taking part in ongoing research on this subject, so this work will remain in effect from January 1, 2004 through June 15, 2004 while under the UROP program. My work will be spread across this time span, mainly during the 2004 Spring semester. I will spend my time directly working on this topic, while at times embarking on other related research opportunities that arise. The 130 hours of work allotted for this project will easily be fulfilled within this time period and additional work will be supported with resources from my sponsor.

Sponsor:
This project relates to the research of my faculty sponsor, Dr. Yuen, as I have been working directly with him on this and other projects for over a year. His interests that relate to this project include scientific visualization, cluster analysis of multi-resolution datasets and geoinformatic infrastructures. I will be in constant communication with Dr. Yuen providing updates on the status of my work and receiving his insight on this research.

Education:
My past research experiences in this field have prepared me for undertaking the project I propose here, and I am anxious for my education to benefit from this program. I am currently studying Computer Engineering in the Institute of Technology and hope to attend a graduate school after my undergraduate education. Working under my faculty sponsor for some time now, I have obtained this redefined the goal for continuing education after being an undergraduate. I am interested in pursuing a PhD program emphasizing on applied mathematics or computer science which will allow me to continue research of this type.
References:


