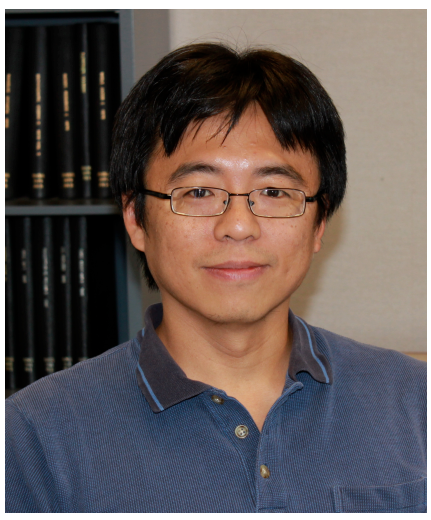
UNCLASSIFIED

Special Information Science and Technology Seminar Speaker



Han-Wei Shen
Ohio State University

Uni- and Multi-variate Scalar Data Analytics Using Information Theory

Monday, August 19, 2013

2:00 - 3:00 PM

TA-3, Bldg. 0207, Room 216, Jemez/Cochiti Conference Rooms (Study Center)

Abstract: As the size of data available to us continues to rise, it becomes more and more challenging to know how to find the most salient information from the data, and how differen sources of information are related to each other. In this talk, I will discuss an information-theroretic framework for big data analytics. Specifically, I will discuss two of our most recent works on applying information theory for uni- and multi-variate scalar data analysis and visualization. While isosurfaces have been widely used for scalar data visualization, it is often difficult to determine if the selected isosurfaces for visualization can sufficiently represent the entire scalar field. To analyze the representativeness of given isosurfaces, our basic idea is to determine if the intermediate isosurfaces enclosed in a subvolume bounded by two pre-selected isosurfaces can be easily derived. We use a level-set approach to peform surface morphing and devise an entropy-based quantitative model called information map to comopare the morph surfaces and the ground truth isosurfaces. This information map can be used as an indicator for the representativeness of the boundary isosurfaces, allowing a quantitative measurement of their information content. In the second part of my talk, I will present a new approach towards building an exploration framework based on information theory to guide the users through the multivariate data exploration process. In our framework, we compute the total entropy of the multivariate data set and identify the contribution of individual variables to the total entropy. The variables are classified into groups based on a novel graph model where a node represents a variable and the links encode the mutual information shared between the variables. The variables inside the groups are analyzed for their representativeness and an information based importance is assigned. We exploit specific information metrics to analyze the relationship between the variables and use the metrics to select isosurfaces of selected variables.

Biography: Han-Wei Shen is a full professor with the Department of Computer Science and Engineering at The Ohio State University, USA. He received his BS degree from Department of Computer Science and Information Engineering at National Taiwan University in 1988, the MS degree in computer science from the State University of New York at Stony Brook in 1992, and the PhD degree in computer science from the University of Utah in 1998. From 1996 to 1999, he was a research scientist at NASA Ames Research Center in Mountain View California. His primary research interests are scientific visualization and computer graphics. Professor Shen is a winner of National Science Foundation's CAREER award and US Department of Energy's Early Career Principal Investigator Award. He also won the Outstanding Teaching award twice in the Department of Computer Science and Engineering at the Ohio State University.

For more information contact the technical host Curt Canada, cvc@lanl.gov, 665-7453 or James Ahrens, ahrens@lanl.gov, 667-5797.

Los Alamos NATIONAL LABORATORY EST. 1943

Hosted by the Information Science and Technology Institute (ISTI)

ISTI INFORMATION SCIENCE & TECHNOLOGY INSTITUTE