VIS Capstone Address

Can I believe what I see? - Information theoretic algorithm validation

Joachim M. Buhmann
ETH Zurich

Abstract

Data Science promises us a methodology and algorithms to gain insights in ubiquitous Big Data. Sophisticated algorithmic techniques seek to identify and visualize non-accidental patterns that may be (causally) linked to mechanisms in the natural sciences, but also in the social sciences, medicine, technology, and governance. When we use machine learning algorithms to inspect the often high-dimensional, uncertain, and high-volume data to filter out and visualize relevant information, we aim to abstract from accidental factors in our experiments and thereby generalize over data fluctuations. Doing this, we often rely on highly nonlinear algorithms.

This talk presents arguments advocating an information theoretic framework for algorithm analysis, where an algorithm is characterized as a computational evolution of a posterior distribution on the output space with a quantitative stopping criterion. The method allows us to investigate complex data analysis pipelines, such as those found in computational neuroscience, neurology, and molecular biology. I will demonstrate this concept for the validation of algorithms using the example of a statistical analysis of diffusion tensor imaging data. In addition, on the example of gene expression data, I will demonstrate how different spectral clustering methods can be validated by showing their robustness to data fluctuations and yet sufficient sensitivity to changes in the data. All in all, an information-theoretical method is presented for validating data analysis algorithms, offering the potential of more trustful results in Visual Analytics.

Speaker Bio

Joachim M. Buhmann is full professor for Computer Science at ETH Zurich since 2003 representing the research area “Information Science and Engineering”. He studied physics at the Technical University of Munich and received a doctoral degree for his research work on artificial neural networks. After research appointments at the University of Southern California (1988-1991) and at the Lawrence Livermore National Laboratory (1991-1992) he served as a professor for applied computer science at the University of Bonn (1992-2003). His research interests range from statistical learning theory to applications of machine learning and artificial intelligence. Research projects are focused on topics in neuroscience, biology and medical sciences, as well as signal processing and computer vision. He has presided the German Society for Pattern Recognition (DAGM e.V.) from 2009-2015. Since 2014 he serves as Vice-Rector for Study Programmes at ETH Zurich. In 2017, he was elected as a member of the Swiss Academy of Technical Sciences SATW, as an honorary member of the German Pattern Recognition Society DAGM and as a research council member of the Swiss National Science Foundation.