

Human Perception

for Multidimensional Data Projection Tasks

Dr. Ronak Etemadpour
Prospective Faculty Candidate
Oklahoma State University
December 12, 2014
3:30 pm
Engineering Center Room 110

Different areas of data resource such as science, business, and engineering produce large amount of data with increasingly complexity and dimensionality. However, given data, users often have little or no idea as to what problems to solve. Big Data cannot be accommodated due to the limited scalability in terms of human perception and computer screen space. Visualization as an interactive visual representation of data turns things that we want to find out in Big Data to something that we can see in order to amplify human cognition. Multidimensional data sets with very large numbers of objects and/or a very large number of dimensions, are common that a variety of visualization methods can be employed in order to represent the data effectively and to enable the user to explore the data at different levels of detail. A common strategy for encoding multidimensional data for visual analysis is to use dimensionality reduction techniques that project data from higher dimensions onto a lower-dimensional space. I focus on projection techniques that output 2D or 3D scatterplots which can then be used for a range of data analysis tasks. Existing taxonomies for multidimensional data projections focus primarily on tasks in order to evaluate the human perception of class or cluster separation and/or preservation. However, real-world data analysis of complex data sets often includes other tasks besides cluster separation, such as: cluster identification, similarity seeking, cluster ranking, comparisons, counting objects, etc. A contribution of my work is the identification of subtasks grouped into four main categories of data analysis tasks. A user-centric task categorization can be used to guide the organization of multidimensional data projection layouts. Moreover, it can be used as a guideline for visualization designers when faced with complex data sets requiring dimensionality reduction. The effectiveness of a visualization system can be evaluated by providing an expanded range of relevant tasks. These tasks are gathered from an extensive study of visual analytics projects across real-world application domains, all of which involve multidimensional projection.