Design and Evaluation in Visualization Research

Panel Organizer:
Donald House, Texas A&M University

Panelists:
Victoria Interrante, University of Minnesota
David Laidlaw, Brown University
Russell Taylor, University of North Carolina
Colin Ware, University of New Hampshire

1 Introduction

Until very recently, the emphasis in Visualization research has been on methods, their algorithmic underpinnings, and their implementation in systems. Most papers have been of the proof of concept variety: describing new ideas for attacking a visualization problem and demonstrating feasibility and quality by presenting visual and performance results from a prototype implementation. Typical published evaluations might consist of 1) a compelling visual presentation on one or two data sets, 2) a comparison of computational efficiency with known algorithms, and 3) anecdotal visual comparison with other techniques. Such tests have driven creativity and advances within our community, but they do not often lead us to design principles to guide future work nor are they compelling to potential collaborators.

This panel brings together researchers who have been pioneering quite different approaches to visualization research by integrating evaluation and knowledge of visual design into their work. The panelists will present their views and experiences in using user studies for quantitative evaluation of methods, in integrating the expertise of visually trained designers into the development of methods, and in exploring the parameter space of visualization possibilities using "human-in-the-loop" experiments.

A goal of the panel is to encourage a lively and stimulating discussion by presenting challenging but highly contrasting ideas. The panel will follow the usual pattern of short position presentations, taking care to leave ample time for audience interaction, questions, and comments.

Keywords: visualization evaluation, visualization design

2 Position Statements

David Laidlaw:
Where Have We Been and Where Are We Going?

Over the last six years keynote, capstone, and panel speakers at *Visualization* and *InfoVis* have opined about various aspects of visual design and evaluation. Perceptual psychology, which links the two areas, has also been a recurring topic. I will give a historical perspective of these sessions, recognizing that they represent only the recent past, not the origin of these ideas. Some timely themes that have emerged include the value of engaging visual designers in the visualization development process; the utility, complexity, and broad gamut of evaluation methodologies; and the connections among all three areas: design, evaluation, and perception. I will explore some of these themes with examples from our work at Brown. In particular, I will describe several projects where we have been collaborating with visual designers. One of these projects was a visualization class where design students and computer science students collaborated to design interactive visualizations of flow around the wings of bats in flight; Figure 1 shows some example designs from the class

In some contexts we have found that visual designers can replace user studies in evaluating visualization methods – and visual designers are both faster and more informative. We have also tried to include visual design experts directly in the visualization tool development process. Unfortunately, this has been a somewhat frustrating process, primarily because tools are not available to let designers work, particularly for large-scale exploratory visualization. I will conclude by talking about some of the limitations that we have found.

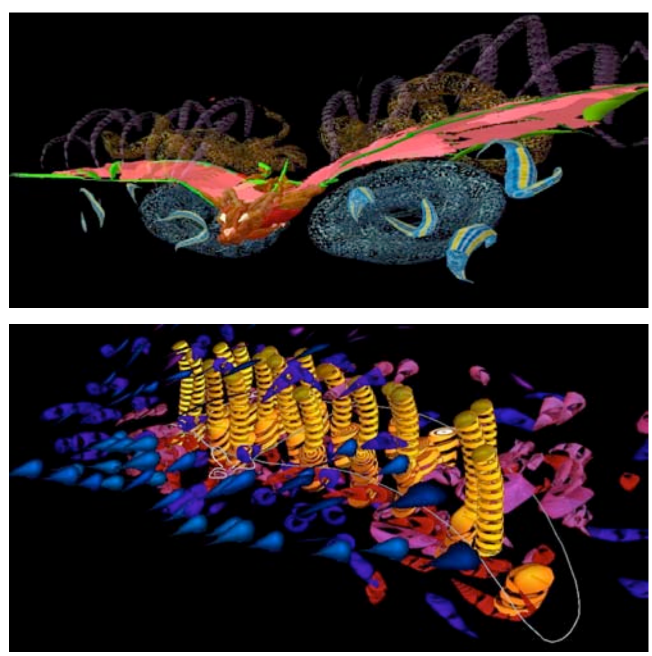

Figure 1. Designer inspired bat wing airflow visualizations

IEEE Visualization 2005
October 23-28, Minneapolis, MN, USA
0-7803-9462-3/05/$20.00 ©2005 IEEE.

Victoria Interrante:

What are Observer Experiments Good For?

One of the enduring fundamental challenges in visualization research is to determine how best to portray a set of data so that the information it represents can be accurately and efficiently understood. Both design and evaluation have key roles to play in this process. We can draw initial inspiration for the design of effective new visualization methods from the study of previous practices, including a survey of representational techniques used by skilled artists and illustrators, and from the observation of materials and phenomena in the natural world. In addition, and particularly when the space of possibilities is huge, we can use fundamental insights from the study of human visual perception to help guide us towards the most potentially promising avenues of investigation. Throughout the process of design, we especially need to constantly resist allowing our imagination space to be constrained – in subtle but potentially significant ways, of which we may not even be consciously aware – by the tools or techniques that we have available to use in exploring the vast space of possible options for information representation. Evaluation of the potential effectiveness of a visualization technique or approach is a ubiquitous process that takes a range of forms, from informal, subjective judgments that we use (consciously or unconsciously) to converge on our design decisions to formal observer experiments intended, for example, to objectively assess the extent to which the use of a particular new visualization technique (as opposed to a previous or default approach) can enhance performance on a set of specific tasks.

In this panel, my presentation will focus on the role of observer experiments in visualization design and evaluation, considering questions such as: what are they good for? when are they warranted? and how can they be used to best advantage? As well as presenting a general overview, I will draw from particular experiences in my recent research, where we have been using observer experiments in two very different projects, in each case seeking through these studies to gain fundamental insight into specific processes of visual reasoning. In the first of these projects, we have been using observer experiments to determine how peoples' interpretation of a surface's shape can be affected by various features of the surface's texture pattern, ranging from its mere presence to its directionality to its orientation with respect to the intrinsic geometry of the surface upon which it appears (Figure 2). The objective of this work is to determine how texture might be most effectively used to help convey an accurate, intuitive understanding of the 3D shape of an arbitrary, smoothly curving surface that is statically or dynamically viewed from an arbitrary vantage point. More recently, we have been using observer experiments in another project for a very different purpose: to probe the possible effects of cognitive influences on participants' judgments of egocentric distance in immersive virtual environments (Figure 3). This new work is particularly exciting because it explicitly considers perception as a process that involves top-down as well as bottom-up effects.

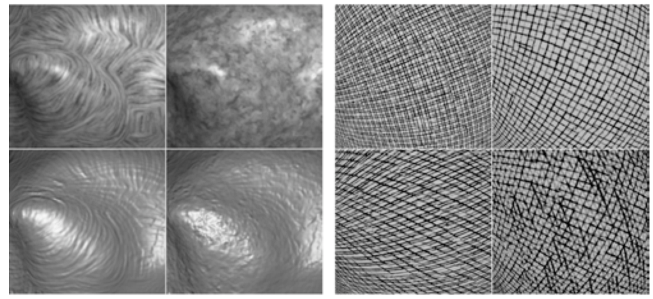

Figure 2. Presentations from experiments investigating the effects of texture on understanding of surface shape

Figure 3. Experiments probing cognitive influences on distance judgment in immersive VR

Donald House:
Can Designers and Controlled Experiments Get us Where We Need to Go?

One can view visualization as simply illustration using advanced technology, so calling on the skills of trained illustrators for visualization design makes sense. However, the assistance that traditionally trained visual designers can lend is limited by their training and experience, as computer technology permits new treatments for a number of phenomena that were very difficult to illustrate in the past. Likewise, one can view visualization as applied perception, and call on tools of the psychologist, like controlled studies to uncover basic principles. However, the complexity of the problems that we are taking on limits the usefulness of this approach. If we try to capture this complexity in controlled experiments, we are faced with a combinatorially exploding number of cases that must be considered and lifetimes of work to address a problem thoroughly. Inherently three-dimensional scenarios like multivariate layered visualizations and time-varying phenomena like unsteady flow provide two examples that go beyond the traditional training of illustrators and are too complex for direct evaluation by experiment.

The visualization community is challenged to take on complex new problems for which classical illustration methods and user studies can be limiting. Thus, our research group is proposing a third approach to the design and evaluation of effective visualizations. This is to use "human-in-the-loop" experiments designed to search the parameter space of a visualization while building large databases of evaluated visualization solutions. This experimentation is followed by data mining to extract strong examples, and to aid in the development of design guidelines and theory. I illustrate this approach by a study of the problem of optimal texturing for stereo viewing of layered surfaces. Figure 3 shows some highly rated solutions to this problem that come from our experiments.

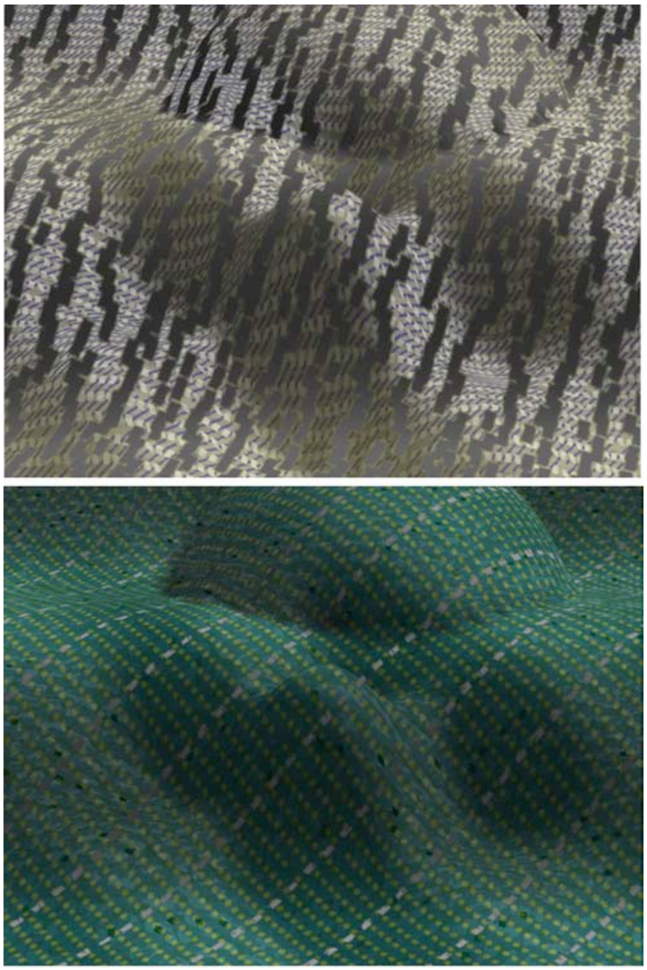

Figure 4. Layered surface texturings from "human-in-the-loop" experiments

Russell Taylor:
It is Time for Standardized Visualization Testing!

The visualization community has produced a wealth of techniques and algorithms for the display of 2D and 3D scalar, vector, and tensor fields. Scientists have provided data sets and questions to be answered. Perceptual psychology continues to provide insight into the capabilities and illusions of the human perceptual systems; enabling "rules of thumb" mapping from techniques to problems.

The field is now poised to go from art to science, by creating a set of standard data sets, questions, and evaluation user studies against which existing and new techniques can be tested. Several groups have already run such user studies in specific domains; others have developed sets of questions relevant in particular domains. For example, the various problems and user interfaces shown in Figure 3 invite the question "Which interface is 'best'?" Solid answers can only come from user studies.

The time has come to produce a set of between-users studies for different data types (2D scalar, volume scalar, 2D vector, etc) and associated IRB documents would enable each research team to test the performance of its favorite technique against well-specified task questions, producing a mapping from questions to best techniques. The mapping from domain questions to task-independent goals will require discussions with each of our science communities. The mapping from goals to experimentally-testable tasks will require iteration and significant feedback from the perceptual scientists among us. These are difficult problems that no researcher can address alone, but together we have all of the resources required. We are the community to do it.

Figure 5. Which interface is "best"?

Colin Ware:
When is Rigorous Evaluation Warranted?

With a large collection of developed algorithms and techniques at hand, visualization researchers are now coming to grips with the fact that it is not devising an algorithm that is the critical problem, but rather how to map data to display in such a way that people can see important patterns. In other words it is about perception. We inherit more than a century of human vision research and in many cases this scientific literature is a treasure trove of information as to what techniques will be effective. In other cases the methods of vision research are best used for evaluating a new display technique. But rigorous techniques of psychophysics are often difficult to get right, very time consuming, and only sometimes applicable.

The first requirement for success is a simple-pattern identification task that is basic or representative. Examples are: perceiving short paths in network diagrams, finding advection paths in flow data, or judging which of two displayed values is the larger. Given such a pattern it is usually easy to devise a metric – we can then measure the number of steps that can be resolved using a particular method, the data density that can be achieved, the number of data dimensions that can be independently displayed, or the error rate. For example, if a technique is claimed to show five variables on a continuous surface then it is useful to know a) it they are perceptually independent b) the data density supported and c) the resolving power of the method.

Rigorous methods are usually not applicable to elaborate visualizations having many design variables. The result is a combinatorial explosion of possibilities that would take many lifetimes to formally evaluate. There is however a way for psychophysical methods to help designers of complex visualizations and this is in the development and evaluation of design principles based on scientific theory. My bottom line is that visualization researchers are in the applied perception business whether they know it or not. To explore the relevant issues can require instrumentation specially designed for the purpose, such as the stereoscope shown in Figure 4 that operates at the resolution of the human retina.

Novel techniques can and should be evaluated as to their effectiveness. The techniques of psychophysics are useful to provide a basic foundation.

Figure 6. A retinal resolution digital stereoscope used for evaluating visualization methods.

3 Biographical Sketches

Donald House

Donald House is Professor of Visualization Sciences at Texas A&M University, where he teaches and conducts research in an interdisciplinary environment in the College of Architecture. He completed his Ph.D. in Computer and Information Sciences at UMass, Amherst, and he holds an M.S. in Electrical and Systems Engineering from Rensselaer. He is best know for his work in cloth modeling and animation, but early in his research career he studied computational brain mechanisms for depth vision in anuran amphibians (frogs and toads). Having failed to escape the complex issues of depth vision, he has recently become fascinated with the problem of perceptually optimizing the visualization of multiple overlapping layers on computer displays.

Victoria Interrante

Victoria Interrante is an Associate Professor in the Department of Computer Science and Engineering at the University of Minnesota. Her research focuses on the application of insights from visual perception, art and illustration to the design of more effective techniques for conveying information through images. She enjoys a variety of interdisciplinary collaborations with colleagues from the departments of Aerospace Engineering, Mechanical Engineering, Astrophysics, Geology and Architecture. Her present projects include: the study of texture and texture synthesis for representing 3D shape, visualizing multivariate data, and representing data uncertainty; the study of spatial perception in immersive virtual environments; and the development of VR tools to enhance the process of conceptual design in architecture. She received her PhD in 1996 from the University of North Carolina at Chapel Hill, and from 1996-1998 was a staff scientist at ICASE/NASA Langley Research Center.

David Laidlaw

David Laidlaw is an Associate Professor in the Computer Science Department at Brown University. His research centers around applications of visualization, modeling, computer graphics, and computer science to other scientific disciplines including archaeology, developmental neurobiology, medical imaging, orthopedics, art, cognitive science, remote sensing, and fluid mechanics to develop new computational applications and to understand their strengths and weaknesses. Particular interests include visualization of multi-valued multidimensional imaging data, comparisons of virtual and non-virtual environments for scientific tasks, and applications of art and perception to visualization. His PhD in Computer Science is from Caltech, where he also did post-doctoral work in the Division of Biology

Russell Taylor

Russell Taylor is a Research Associate Professor of Computer Science, Physics & Astronomy, and Applied & Materials Sciences at the University of North Carolina at Chapel Hill. He is the co-director of the UNC NIH National Research Resource for Computer Integrated Systems for Microscopy and Manipulation. His research interests include Scientific Visualization, Distributed Virtual Worlds, Haptic Display, and Interactive 3D Computer Graphics. All of these come together in his role as the director of the Computer Science team in the UNC Nanoscale Science Research Group, which also includes Physicists, Chemists, Gene Therapists, Biologists, Library Scientists, and Perceptual Psychologists working together to develop improved interfaces for scanned-probe and other microscopes. Russell teaches a course on Visualization in the Sciences for computer scientists and natural scientists.

Colin Ware

Colin Ware is director of the Data Visualization Research Lab, which is part of the Center for Coastal and Ocean Mapping at the University of New Hampshire. He holds a PhD in Experimental Psychology (Toronto) and an MMath in Computer Science (Waterloo). Most of Ware's research involves applying perception to problems in visualization and he is the author of the leading textbook in this area, Information Visualization: Perception for Design now in its second edition. He has also has been the lead designer of a number of visualization software products including Fledermaus, GraphVisualizer3D and GeoZui4D. Right now he is having most fun with whale tracking data and with ocean flow models.