Measuring Exploration Coverage and Evaluating Refinding Mechanisms

Celine Latulipe University of North Carolina at Charlotte
Charlotte, NC, USA
clatulip@uncc.edu

ABSTRACT
I discuss the issues of measuring exploration coverage to evaluate exploratory interaction techniques. I also discuss the 'refinding' problem in visual data exploration and how refinding techniques can be evaluated.

Author Keywords
visualization, interaction, evaluation

ACM Classification Keywords
H5.m. Information interfaces and

INTRODUCTION

Very large sets of data with many interacting parameters are the basis for the field of study of information visualization. In visual analytics research, the essential problem is finding ways for people to explore large-scale data sets. The human-intervention aspect is inherent in the problem. House et al. have proposed the 'human-in-the-loop' model for evaluating the optimum settings for a given type of visualization, which uses human interaction with the data to populate a genetic algorithm for finding optimal settings [1]. My research takes one step back and asks: what are the best techniques for users interacting with multi-variate visual data?

In my research in human-computer interaction, I have been working on what is essentially a 'small' information visualization problem that has forced me to confront issues of evaluating interactive visualization techniques. I have two main areas of interest in information visualization evaluation: I want to determine best practices for measuring exploration coverage and I want to determine how to evaluate mechanisms for 'refinding', where refinding is the act of a user returning to a configuration of interest that has already been visited. I am also interested in investigating the cognitive load imposed on the user during both exploration and refinding.

THE TONEZONE EXAMPLE

I have developed a novel, two-handed interaction technique for allowing users to manipulate the tones in digital photos. The ToneZone technique (see Figure 1) allows the two hands to simultaneously control the minimum and maximum input and output tones, which is identical to the users controlling four different sliders in Adobe Photoshop's Levels tool. Each of these four factors (minimum and maximum input tone and minimum and maximum output tone) can vary in value from 0 to 255 for a typical digital photo. With four variables, each with 255 possible values, the combination of options is 255^4. In fact, the maximum input tone is constrained with respect to the minimum output tone, so the combination is slightly smaller. Regardless, the number of possible combinations of these four simple settings is large. It is unlikely that a user will reach many of these combinations with one mouse controlling four separate sliders. A user controlling all four variables at once using symmetric, two-handed exploration, is likely to reach many more of these variables.

Figure 1. The ToneZone interaction technique allows users to simultaneously explore four different tone parameters, and presumably increase exploration coverage compared to traditional tone level tools.

The ToneZone example is a small version of a common problem in visual analytics: that of combinatorial explosion. The important question to researchers concerned with evaluation in this area is how many of those combinations are visited by the user, and more generally, what is the best way to measure coverage of variable factorizations?

EXPLORATION COVERAGE

Interaction techniques that allow a user to vary multiple parameters at once, in order to see the interaction of the variables on a given data set, are prevalant in information visualization. The number of parameters that can be varied

simultaneously depends on the richness of the input devices and the interaction techniques. In my research, I have shown that two standard USB mice can be used to vary four parameters at a time. With 6 DOF devices, it's possible to develop techniques that allow more parameters to be controlled simultaneously. The power in these techniques arises from the ability of the user to explore many more combinations of the input parameters than would be possible using a single, standard input device. This is 'expressive exploration'.

While random exploration is unlikely to be as useful as 'expressive exploration' that is carefully attended to and controlled by a user, there is some possibility of finding interesting configurations of the input parameters through random exploration. This means that a simple measure of coverage may be a good first indicator of the utility of an exploration interaction technique. Coverage is the number of possible configurations of the input parameters that are visited, even momentarily, by the user. It is expressed as a percentage of the total number of possible configurations. For many information visualizations, the number of possible input parameter configurations is huge. There are a number of issues to consider in measuring coverage:

- What constitutes a visit to a particular configuration? (If the user simply passes through the configuration, should that count, or should it only count if the user pauses?)
- Should configurations that are almost identical be grouped?
- If configurations are grouped, how big should the groups be, and should the boundaries be fixed or overlapping?

Experimental Setup

From an HCI perspective, coverage is an important concept because it allows us to compare how much exploration can be accomplished in a given time using different interaction techniques. However, if coverage is the main metric being used, the researcher must then think about an appropriate experimental setup. In particular, if we wish to measure the ease of exploration, the task needs to encourage exploration, but still needs to be focused on some end goal in order to ensure that users don't adjust variables in a strictly random fashion. To evaluate exploration coverage with the ToneZone interface, we ask our users to adjust digital photos to make a variety of 'interesting images' with them, which will then be judged on aesthetic value. We explain that we are not looking for the most realistic photos, but we wish them to try to come up with three or four different versions of each photo that might make interesting wall art. This setup encourages the users to explore the space, while working towards a goal.

Data Analysis

Once a suitable task is in place and users can be run through a data exploration experiment, the researcher must then analyze the exploration data to mine the coverage information. Clustering techniques are the obvious choice for such analysis. However, many HCI researchers may not be familiar with clustering algorithms. In our experimental setup we use scripts that simply divide the four-dimensional space into discrete configuration groups and count group visits.

TONEZONE SPATIAL MEMORY CUES

The second issue that I am dealing with in analyzing the ToneZone interaction relates to the use of spatial memory cues. In the ToneZone example, the users are controlling four variables through the position of two cursors. The rectangle joining these two cursors is a visual aid to help perceptually unify the points, so that the users can control the two separate points as a joint task. However, it seems likely that the rectangle also acts as a spatial memory cue. Once a user has found a good combination of the variables (and likes the way the underlying image looks), he may want to continue exploring. If he goes off to do some exploration, he can still easily return to the 'sweet spot' he found earlier, simply by putting the rectangle back into the same configuration. He is likely to remember the size and position of the ToneZone rectangle because it is super-imposed on the image.

THE REFINDING PROBLEM

The ToneZone illustration is an example of the 'refinding problem' where a user needs to revisit a configuration that he or she has already tried, whether the configurations consist of manipulations of an object, a document or a finding in a scientific visualization tool. There are three different possible approaches to the refinding problem. First, many digital tools allow users to explicitly save various configurations of data, However, explicit 'save' functions need to be invoked to be useful and are an interruption to the workflow. An alternative approach is for the system to use intelligence to figure out which configurations a user might wish to 'refind'. A third approach is to keep a full history of interaction, allowing users to rewind through an entire session in order to refind. A fourth approach is to attempt to build into the interaction mechanisms that will naturally allow the user to easily refind a configuration of interest. This is where the ToneZone fits in: the spatial memory cues act as a built-in mechanism to help the users refind configurations of interest.

The question of interest to workshop attendees, is how to measure solutions to the refinding problem. What types of tasks should be used to evaluate refinding mechanisms? Our approach is to ask users to find their own 'solution' for a given data set, and then later (after doing some other tasks), ask them to refind the same solution with the same data set and see how quickly and accurately they can do this. This approach uses the traditional time and error metrics, which may be the most relevant for the refinding problem. I am also considering an investigation of cognitive load issues by giving users a secondary task to complete at the same time.

WORKSHOP CONTRIBUTIONS

I am currently running one experiment to test exploration with the ToneZone interface and I am preparing to run an experiment to test refinding. I plan to share these experiences with workshop participants.

REFERENCES

1. D. H. House, A. S. Bair, and C. Ware. An approach to the perceptual optimization of complex visualizations. *IEEE Transactions on Visualization and Computer Graphics*, 12(4):509–521, 2006.