VisiGeek: Simultaneous Comparative Visualization

Jim Youll
MIT Media Laboratory
20 Ames Street, Cambridge, MA 02139 U.S.A.
http://www.media.mit.edu/~jim, jim@media.mit.edu

ABSTRACT

VisiGeek is a display tool enabling the simultaneous visual comparison of datasets which form electronic portraits of individuals, groups, or social spaces. The visualizer contrasts differences in the datasets through animations of geometric patterns, letter-sizes and motion. Similarity between the elements of two sets is correlated with stability in the animation, and differences with instability. Applications of simultaneous comparative visualization include displays of affinity between individuals or groups, and as guides in exploration of electronic social spaces. The goal of VisiGeek is to graphically expose differences between datasets while allowing each element to directly express its contribution to the composite result.

This project may be accessed online at http://www.media.mit.edu/~jim/projects/visigeek/

KEYWORDS

Visualization, digital portrait, Geek Code

INTRODUCTION

As individual brushstrokes together make a painted portrait, digitally-recorded preferences, profiles and histories may form an electronic "portrait" of a person, group, or virtual place. VisiGeek creates portraits of these things by assembling the data points into images in which all elements are expressed in an animated, graphical form, allowing the behaviors of the elements – changes in motion and color – to facilitate comprehension of a large field of data points as viewers literally "look at the data."

Personal data which could form electronic portraits permeates the Internet, in explicit coding systems, in explicit preference and profile settings bound to web sites, and in implicit form in the content of personal websites and public newsgroups. Expression of the "similarity" between two such portraits, in community- or friend-finding tasks, for example, is difficult due to the discrete nature of the data elements, dissimilarities in the information being compared, and the need to compare them as both individual data points and in the aggregate as a composite "portrait".

For example, Visual Who [2] is a tool for visualizing an electronic community using mailing list subscriptions to model patterns of affiliation between community members and to encourage interactive exploration of the underlying data. Other approaches to comparative digital portraits have included portraits based on explicit preference declarations, such as Ringo [5], or the live samples of chat room conversations made by Butterfly [6].

VisiGeek takes a different approach, not combining the data into a composite summary, but exposing the discrete elements of the dataset in a common visual space and allowing the patterns which emerge to reveal the nature of the similarities and differences in the underlying portraits.

THE GEEK CODE

The "Geek Code" [4] dates to a time when the Internet was text-driven, before web pages and personal photos. It is a dense self-representation of a person which mimics formal descriptions of stars [3]. Personally-descriptive encodings first appeared in the footers of newsgroup postings and are now borne in e-mail signatures and on web pages. The Geek Code uses standard formatting similar to PGP signatures, to simplify mechanical discovery and parsing.

Figure 1. Geek codes in a newsgroup.

The Geek Code allows people to classify themselves in several categories such as appearance, computers, politics, entertainment, and lifestyle, each of which is represented by a single letter or combination of letters followed by zero or more "+" and "-" signs referencing a descriptive, hierarchical taxonomy. For example, the description of "dress" ranges from "d++ I tend to wear conservative dress…" to "d--- Punk dresser, including but not limited to, torn jeans and shirts, body piercings…" A Geek Code "portrait" may be modified by symbols signifying variability of the represented values.

Geek Codes are used here because they are personal, readily available, reasonably well-evolved, and stable. We do not suggest that Geek Codes constitute a comprehensive portrait of an individual, only that they provide a ready, refined source of sample data.

IMPLEMENTATION

We may say that two individuals who share similar traits or ideas are "in sync". VisiGeek adopts this analogy by

considering two profiles to be "in phase" when generally similar, and "out of phase" when different. It correlates "similarity in the data" with stability and harmony in the animated visual expression of that data. Less motion, fewer color shifts, and a more visually- coherent structures signify datasets with generally-similar traits, or in the case of our sample data, two individuals with generally-similar profiles and preferences. Large sweeps of motion, broad color shifts and increased activity draw attention to those features which are different.

Whyte [8] observed that those who behave differently from the majority in a crowd stand out. In a visualization which attempts to highlight differences, the differences may be made to stand out in a field of otherwise-similar elements through divergent "behavior". Motion though animation is emerging as a "perceptually efficient display dimension" [1], exploiting the human vision system's preattentive sensitivity to motion [7]. VisiGeek differentiates individual behavior through the unique motion of each data element, exploiting this feature of human perception.

Each Geek Code element occupies space in the display window. The display manager steps between the minimum and maximum values for an element and places a visual representation of the current display value in the given code's space. If two people are being compared, it steps between the two individuals' maximum and minimum values for each element. When the values are identical, the image for that element is stationary and stable. As the values diverge, the amount of motion or change increases proportionally to the difference. Several different visualizations are being considered, ranging from a literal representation of the codes using letters of varying sizes to visual representations in which the motion or color of shapes such as rectangles and circles varies to show the range of data being represented.

VisiGeek has three components: a spider to gather codes, a parser to translate the codes into standard form, and a display manager to control interactive visualization. Two iterators drive the visualization: one modulo the size of the space like a rotary timer that resets to zero after reaching a maximum value, and another which ramps gradually between the two values.

ASSIGNING MEANING TO DATA

Considerations included representation of variant traits, absent traits, and wild traits not in the official taxonomy. The representation of "self" could symbolize either a user's personal traits, for comparisons, or the traits sought in others, as a measure of affinity rather than similarity.

FUTURE WORK

User testing should compare the system's current effectiveness in conveying a sense of "difference" with enhancements including better-evolved pattern generators,

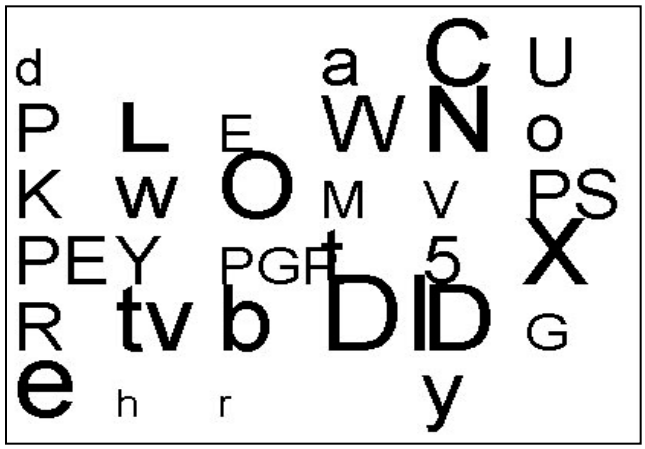

Figure 2. Freeze-frame display of VisiGeek

pattern generation employing more state information, drag and drop rearrangement of elements in conjunction with pattern sets optimized for visual sorting, and novel representations of modifiers that may be appended to codes.

ACKNOWLEDGMENTS

Thanks to Professor Judith Donath, Professor Pattie Maes, Nelson Minar, Dana Spiegel and Tim Bickmore for ideas and assistance.

REFERENCES

1. L. Bartram, Perceptual and Interpretive Properties of Motion for Information Visualization, *Proceedings of the Workshop on New Paradigms in Information Visualization and Manipulation*, 3-7 (1997).
2. J. Donath, Visual Who: Animating the affinities and activities of an electronic community, *ACM Multimedia 95*
3. GCVS Research Group, Sternberg Astronomical Institute, Moscow, General Catalogue of Variable Stars: International Service of Variable Stars: Naming, Classification, Identification *http://www.sai.msu.su/groups/cluster/gcvs/gcvs.html*
4. R. Hayden, The Code of the Geeks v3.12, *http://www.geekcode.com*.
5. U. Shardanand and P. Maes, Social Information Filtering: Algorithms for Automating "Word of Mouth," *Conference Proceedings on Human Factors in Computing Systems*, 210 (1995).
6. N. Van Dyke, H. Lieberman and P. Maes, Butterfly: A Conversation-Finding Agent for Internet Relay Chat, *Proceedings of the 1999 International Conference on Intelligent User Interfaces* , 39-41 (1999).
7. C. Ware, J. Bonner, W. Knight and R. Cater, Moving icons as a human interrupt, *International Journal of Human-Computer Interaction*, 4(4):341-348 (1992).
8. W. Whyte, City: Rediscovering the Center, New York: *Doubleday* (1988).