

M3: Model data representation using virtual reality [Model Qualify, Model Verify, Model Validity]

Gagandeep Singh Kahlon

*Department of Computer Science Engineering
CT Group, Jalandhar, Punjab, India
Email- gagandeepsinghkahlon@gmail.com*

Sukhwinder kaur

*Department of Computer Science Engineering
CT Group, Jalandhar, Punjab, India
Email- multani.sukhi@gmail.com*

Dr.R C Gangwar

*Department of Computer Science Engineering
Beant College of Engg & Tech
Email- rakeshgangwar@gmail.com*

Abstract- Researchers have argued that perceptual issues are important in determining what makes an effective visualization, but generally only provide descriptive guidelines for transforming perceptual theory into practical designs. In order to bridge the gap between theory and practice in a more rigorous way, a computational model of the primary visual cortex is used to explore the perception of data visualizations. A method is presented for automatically evaluating and optimizing data visualizations for an analytical task using a computational model of human vision. The method relies on a neural network simulation of early perceptual processing in the retina and visual cortex. The neural activity resulting from viewing information visualization is simulated and evaluated to produce metrics of visualization effectiveness for analytical tasks. Visualization optimization is achieved by applying these effectiveness metrics as the utility function in a hill-climbing algorithm. This method is applied to the evaluation and optimization of two visualization types 2D flow visualizations and node-link graph visualizations. The computational perceptual model is applied to various visual representations of flow fields evaluated using the advection task of Laidlaw et al. [1]. The predictive power of the model is examined by comparing its performance to that of human subjects on the advection task using four flow visualization types. The results show the same overall pattern for humans and the model. In both cases, the best performance was obtained from visualizations containing aligned visual edges. Flow visualization optimization is done using both streaklet-based and pixel-based visualization parameterizations. An emergent property of the streaklet-based optimization is head-to-tail streaklet alignment, the pixel-based parameterization results in a LIC-like result. Here we propose M3 Visualization for mathematical approach for improve HCI performance in statistical representation.

Keywords – Model Data, Data visualization, Data Mining, HCI performance.

I. INTRODUCTION

Perception is the means through which all data visualizations are processed, and the characteristics of successful data visualizations can often be traced back to perceptual mechanisms of the human visual system. It is becoming increasingly recognized that the properties of human perception play a vital role in determining the effectiveness of data visualizations [2]. For example, the most successful flow visualizations contain contours tangential to the flow field. These visualizations take advantage of the edge detection mechanisms of human vision, which respond strongly to continuous contours. As a second example, visualizations that must convey fine spatial details are found to be more effective when they are rendered in greyscale than in a red-green scale or yellow-blue scale. This is because luminance contrast processing mechanisms of the visual system can convey far more detail than the chromatic contrast processing mechanisms.

Perception is the means through which all data visualizations are processed, and the characteristics of successful data visualizations can often be traced back to perceptual mechanisms of the human visual system. It is becoming increasingly recognized that the properties of human [1] perception play a vital role in determining the effectiveness of data visualizations. For example, the most successful flow visualizations contain contours tangential to the flow field. These visualizations take advantage of the edge detection mechanisms of human vision, which respond strongly to continuous contours. As a second example, visualizations that must convey fine spatial details are found to be more effective when they are rendered in grey scale than in a red-green scale or yellow-blue scale. This is because luminance contrast processing mechanisms of the visual system can convey far more detail than the chromatic contrast processing mechanisms.

Data visualization is the presentation of data in a pictorial or graphical format. It enables decision makers to see analytics presented visually, so they can grasp difficult concepts or identify new patterns. With interactive visualization, you can take the concept a step further by using technology to drill down into charts and graphs for more detail, interactively changing what data you see and how it's processed [3]. The concept of using pictures to understand data has been around for centuries, from maps and graphs in the 17th century to the invention of the pie chart in the early 1800s. Several decades later, one of the most cited examples of statistical graphics occurred when Charles Minard mapped Napoleon's invasion of Russia. The map depicted the size of the army as well as the path of Napoleon's retreat from Moscow – and tied that information to temperature and time scales for a more in-depth understanding of the event. It's technology, however, that truly lit the fire under data visualization. Computers made it possible to process large amounts of data at lightning-fast speeds [4]. Today, data visualization has become a rapidly evolving blend of science and art that is certain to change the corporate landscape over the next few years.

Application of data visualization

Big data is set to offer companies tremendous insight. But with terabytes and petabytes of data pouring in to organizations today, traditional architectures and infrastructures are not up to the challenge. IT teams are burdened with ever-growing requests for data, ad hoc analyses and one-off reports. Decision makers become frustrated because it takes hours or days to get answers to questions, if at all. More users are expecting self-service access to information in a form they can easily understand and share with others [4]. This begs the question: How do you present big data in a way that business leaders can quickly understand and use? This is not a minor consideration. Mining millions of rows of data creates a big headache for analysts tasked with sorting and presenting data. Organizations often approach the problem in one of two ways: Build "samples" so that it is easier to both analyze and present the data, or create template charts and graphs that can accept certain types of information. Both approaches miss the potential for big data. Instead, consider pairing big data with visual analytics so that you use all the data and receive automated help in selecting the best ways to present the data. This frees staff to deploy insights from data. Think of your data as a great, but messy, story [5]. Visual analytics is the master filmmaker and the gifted editor who bring the story to life.

Challenges faced by visualization

Meeting the need for speed: In today's hypercompetitive business environment, companies not only have to find and analyze the relevant data they need, they must find it quickly. Visualization helps organizations perform analyses and make decisions much more rapidly, but the challenge is going through the sheer volumes of data and accessing the level of detail needed, all at a high speed [6, 7]. The challenge only grows as the degree of granularity increases. One possible solution is hardware. Some vendors are using increased memory and powerful parallel processing to crunch large volumes of data extremely quickly. Another method is putting data in-memory but using a grid computing approach, where many machines are used to solve a problem. Both approaches allow organizations to explore huge data volumes and gain business insights in near-real time.

Understanding the data: It takes a lot of understanding to get data in the right shape so that you can use visualization as part of data analysis [8]. For example, if the data comes from social media content, you need to know who the user is in a general sense – such as a customer using a particular set of products – and understand what it is you're trying to visualize out of the data. Without some sort of context, visualization tools are likely to be of less value to the user. One solution to this challenge is to have the proper domain expertise in place. Make sure the people analyzing the data have a deep understanding of where the data comes from, what audience will be consuming the data and how that audience will interpret the information.

Addressing data quality: Even if you can find and analyze data quickly and put it in the proper context for the audience that will be consuming the information, the value of data for decision-making purposes will be jeopardized if the data is not accurate or timely. This is a challenge with any data analysis, but when considering the volumes of information involved in big data projects, it becomes even more pronounced. Again, data visualization will only

prove to be a valuable tool if the data quality is assured [9]. To address this issue, companies need to have a data governance or information management process in place to ensure the data is clean. It's always best to have a proactive method to address data quality issues so problems won't arise later.

Displaying meaningful results: Plotting points on a graph for analysis becomes difficult when dealing with extremely large amounts of information or a variety of categories of information. For example, imagine you have 10 billion rows of retail SKU data that you're trying to compare [10]. The user trying to view 10 billion plots on the screen will have a hard time seeing so many data points. One way to resolve this is to cluster data into a higher-level view where smaller groups of data become visible. By grouping the data together, or "binning," you can more effectively visualize the data.

Dealing with outliers: The graphical representations of data made possible by visualization can communicate trends and outliers much faster than tables containing numbers and text [11]. Users can easily spot issues that need attention simply by glancing at a chart. Outliers typically represent about 1 to 5 percent of data, but when you're working with massive amounts of data, viewing 1 to 5 percent of the data is rather difficult. How do you represent those points without getting into plotting issues? Possible solutions are to remove the outliers from the data (and therefore from the chart) or to create a separate chart for the outliers. You can also bin the results to both view the distribution of data and see the outliers [12]. While outliers may not be representative of the data, they may also reveal previously unseen and potentially valuable insights.

II. EXISTING SYSTEM IMPACT

An alternative method of quantifying the effectiveness of data visualization, one that addresses the shortcomings of descriptive theories and psychophysical tests, is highly desirable. What is needed is not just a perceptual theory, but a perceptual computational model that allows data visualization designers to explore in detail how the perceptual mechanisms of the human visual system respond to data visualization. Computational Modeling-Computational Modeling has become indispensable addition to the field of fluid dynamics. While closed form theories of fluid dynamics are capable of describing idealized systems, in even simple real world systems the behavior can quickly become too complex for them to be directly applied in practice. For example, the problem of designing of supersonic blunt nose aircraft confounded the fluid dynamics community throughout the 1950's and early 1960's. The complexity of the mathematical analysis of this seemingly simple problem proved to be essentially intractable. Experimental approaches to fluid dynamics via wind tunnels have the limitations of being slow and expensive, and only yield a few variables such as lift and drag. These limitations have a striking resemblance to the current limitations of perceptual theory and psychophysical testing.

The problem of modelling perception at a higher more cognitive level is even more challenging. It is likely that humans adopt different perceptual strategies depending on the visualization. For example, with a regular arrow grid people are likely to rely on the contour information available in the display, which is degraded due to perceptual aliasing, and more on higher-level cognitive strategies, such as looking ahead from a particular arrow, mentally interpolating the flow direction at the look ahead point and repeating to the edge of the field. Conversely, in the case of aligned streaklets, a simpler contour following finding strategy will be more successful. Layered graph diagrams are a notable subclass of node-link diagrams. The nodes of these diagrams are arranged such that they lie in distinct layers. This layout facilitates the understanding of structural attributes of the graph, such hierarchical and sibling

III. Proposed Methodology

Recent theories are of restricted use to knowledge visualisation designers as a result of they're usually expressed in general and descriptive forms that don't seem to be promptly applicable. For instance, the idea of contour perception states that 2 close visual objects that exhibit edges destined near the linear axis between them are perceived as a continual contour. Whereas this statement suggests that we should always limit our flow visualisation styles to those with aligned arrows, it still leaves and large variety of degrees of freedom unspecified.

So current theory has suffers to provide these areas Simplicity/Adaptation, visibility and Extendibility of representation, finally we should say it's has boundary of visualizing the ideas. But there are many potential advantages to applying a computationalM3 approach to data visualization:

Simplicity - Computational models express complex behaviour in terms of a few fundamental behavioural rules, enabling the behaviour of the system to be more easily comprehended. By applying a computational model of visual perception to data visualization, a visualization practitioner could observe the effects that visualization design have on the neural dynamics of the visual system, thus gaining insight into the reason for its effectiveness.

Utility - Computational models produce numerical values that are suitable for use within a larger system. We explore this possibility in this thesis work by encompassing a mutational model of perception within an optimization loop.

Visibility - The numerical values with which computational models are expressed afford the creation of visualizations that show the effects that the input design has on the model.

Extensibility - Computational models can be easily integrated with other computational models. For example, a computational model of the early visual system might be integrated with a computational model of working memory, allowing designers to explore the cognitive load associated with visualization.

The contribution of this work is the application of the computational Modelling paradigm to the problem of data visualization, and the demonstration that such an approach can yield tangible results. Given the tremendous impact that computational science has had on other scientific disciplines, it seems clear that exploring the viability of its application to data visualization is a worthwhile endeavour.

My contribution has five components:

1. Development of a computational model of human perception that simulates the visual mechanisms critical to the perception of data visualizations.
2. Comparison of the model evaluations with human performance on psychophysical tasks
3. Automatic optimization of data visualizations based on the effectiveness predicted by the model.
4. Demonstration of the generality of the approach by the application to two distinct classes of visualizations: 2D vector fields and node-link diagrams.
5. Establishment of computational Modelling perception as a “third pillar of science” in data visualization by developing its relationship with perceptual theory and psychophysical experimentation.

IV.CONCLUSION

In this level of research, we investigated the application of computational Modelling data visualization and explored some of the opportunities that such an approach presents. Reviewing the neural architecture of the visual system, we recognized the suitability of a computational Modelling approach to the problem of data visualization. The neural activity resulting from viewing information visualization is simulated and evaluated to produce metrics of visualization effectiveness for analytical tasks. Visualization optimization is achieved by applying these effectiveness metrics as the utility function in a hill-climbing algorithm

REFERENCE

- [1] Pineo, D.; Ware, C., "Data Visualization Optimization via Computational Modelling of Perception," in Visualization and Computer Graphics, IEEE Transactions on , vol.18, no.2, pp.309-320, Feb. 2012doi: 10.1109/TVCG.2011.52
- [2] Jay Wang, Computational Modelling and Visualization of Physical Systems with Python.ISBN: 978-1-119-17918-4492 pagesJanuary 2016, ©2016
- [3] Bryson, S. Virtual reality takes on real physics applications. Comput. Phys. 6, 4 (July/Aug. 1992).
- [4] Sheridan, T.B., and Ferrill, W.R. Man-Machine Systems. MIT Press, Cambridge, Mass., 1974.
- [5] Donalek, Ciro, et al. "Immersive and collaborative data visualization using virtual reality platforms." *Big Data (Big Data), 2014 IEEE International Conference on*. IEEE, 2014.
- [6] Laha, Bireswar, Doug A. Bowman, and John J. Socha. "Effects of VR system fidelity on analyzing isosurface visualization of volume datasets." *Visualization and Computer Graphics, IEEE Transactions on* 20.4 (2014): 513-522.
- [7] Barfield, Woodrow, Claudia Hendrix, and Karl Bystrom. "Visualizing the structure of virtual objects using head tracked stereoscopic displays." Virtual Reality Annual International Symposium, 1997., IEEE 1997. IEEE, 1997.Kundur D., Hatzinakos D., Towards robust logo watermarking using multiresolution image fusion, IEEE Transactions on Multimedia 6 (2004) 185-197.
- [8] Zhang, Song, et al. "An immersive virtual environment for DT-MRI volume visualization applications: a case study." *Visualization, 2001. VIS'01. Proceedings.* IEEE, 2001.

- [9] Ragan, Eric D., et al. "Studying the effects of stereo, head tracking, and field of regard on a small-scale spatial judgment task." *Visualization and Computer Graphics, IEEE Transactions on* 19.5 (2013): 886-896.
- [10] Chen, Jian, et al. "Effects of stereo and screen size on the legibility of three-dimensional streamtube visualization." *Visualization and Computer Graphics, IEEE Transactions on* 18.12 (2012): 2130-2139.
- [11] Garlandini, Simone, and Sara Irina Fabrikant. "Evaluating the effectiveness and efficiency of visual variables for geographic information visualization." *Spatial information theory*. Springer Berlin Heidelberg, 2009. 195-211..
- [12] 12. Bowman, Doug A., et al. "Evaluating effectiveness in virtual environments with MR simulation." Interservice/Industry Training, Simulation, and Education Conference. 2012. C. Hsu and J. Wu, "Multi-resolution Watermarking for Digital Images", *IEEE Transactions on Circuits and Systems- II*, Vol. 45, No. 8, pp. 1097-1101, August 1998.
- [13] Drake, A. J., et al. "First results from the Catalina real-time transient survey." *The Astrophysical Journal* 696.1 (2009): 870.