An Alternative Definition and Model for Knowledge Visualization

Dong Hyun Jeong*
UNC Charlotte
Viscenter

Remco Chang†
UNC Charlotte
Viscenter

William Ribarsky‡
UNC Charlotte
Viscenter

ABSTRACT

The term *knowledge visualization* has been used in many different fields with many different definitions. In this paper, we propose a new definition of knowledge visualization specifically in the context of visual analysis and reasoning. Our definition begins with the differentiation of knowledge as either explicit or tacit knowledge. We then present a model for the relationship between the two through the use visualization. Instead of directly representing data in a visualization, we first determine the value of the explicit knowledge associated with the data based on a cost/benefit analysis and display the knowledge in accordance to its importance. We propose that the displayed explicit knowledge leads us to create our own tacit knowledge through visual analytical reasoning and discovery.

1 INTRODUCTION

The research agenda by Thomas and Cook [9] outlines the fundamental research areas for the field of visual analytics based on the considerations of data representations, interactive techniques, perceptual effects, etc. Although understanding these areas are essential in creating visual analytical tools, we believe that there exists a different type of visualization approach that focuses on the knowledge products that are closer to the reasoning artifacts needed in complex analytical processes. Our hypothesis is that since the goal of visual analysis is in acquiring more knowledge, an approach that emphasizes the knowledge products could be more direct and beneficial. We call this approach "knowledge visualization." In this paper, we present the initial path and focus on identifying what knowledge visualization is and what is needed to making it possible.

It is relevant to note that the term "knowledge visualization" has already been defined by other researchers such as Burkhard as "the use of visual representations to improve the transfer of knowledge between at least two persons or group of persons" [2]. Our definition of the term knowledge visualization is different. We believe that knowledge is a deterministic process that encapsulates a collection of information [8]. Throughout the rest of the paper, we will be using the later definition when using the term "knowledge."

Our approach of knowledge visualization deals with representing knowledge rather than data or information by determining the values of knowledge through understanding the data or information. Although finding and understanding important knowledge are extremely difficult, we begin with understanding the definition of knowledge and the user's analytical procedures. Specifically, we differentiate knowledge into tacit knowledge (personal, context-specific, hard to formalize and communicate) and explicit knowledge (transmittable in a formal, systematic language) [7]. With this differentiation, we can formulate a model on how knowledge that inherently exist in data (explicit knowledge) interacts with knowledge within an analyst (tacit knowledge) through a visual analytical tool. A more detailed definition of explicit and tacit knowledge is presented in section 2, while the model showing the relationships between the two types of knowledge is introduced in section 3.

*e-mail: dhjeong@uncc.edu
†e-mail:rchang@uncc.edu
‡e-mail:ribarsky@uncc.edu

2 DEFINITION OF KNOWLEDGE

To develop a new approach for visualizing knowledge, we must first know what knowledge is. In knowledge management literature, it has been established that distinguishing between data, information, and knowledge is important to designing knowledge management programs [4]. Work by Syed and Shah [8] reviews various definitions and explanations of the DIKW (data, information, knowledge, wisdom) hierarchy and focuses on presenting a model that explicates the relationship between data, information, and knowledge. In Syed and Shah's model, knowledge is defined as the range of one's information. However, Davenport and Prusak [3] state that "knowledge derives from information as information derives from data" and further define knowledge as "a fluid mix of framed experience, contextual information, values and expert insight that provides a framework for evaluating and incorporating new experiences and information." In Davenport and Prusak's perspective, knowledge is the refined information, in which human cognition has added value. In other words, information becomes knowledge through cognitive effort. Based on this definition, we think that knowledge can only result from human cognitive process that includes perceiving, recognizing, conceiving judging, reasoning, and imagining [1]. It also shows that knowledge fundamentally involves relationships either among ideas or other pieces of knowledge.

Nonaka and Takeuchi [7] differentiates knowledge as tacit knowledge and explicit knowledge to understand how knowledge is shaped and how knowledge can be applied. In their definition, explicit knowledge can be processed by a computer, transmitted electronically, or stored in database. On the other hand, tacit knowledge is personal and specialized and can only be extracted by human. We extend Nonaka and Takeuchi's concept and apply it to our concept of knowledge visualization. We believe that through the use of interactive visual analytical tools, analysts can understand the explicit knowledge within data by transforming it into tacit (internal) knowledge.

To demonstrate our extension of Nonaka and Takeuchi's definition of explicit and tacit knowledge and their relationships in the context of visual analytics, we present a knowledge model based on a model by van Wijk [10] on the value of visualization in the following section.

3 KNOWLEDGE VISUALIZATION MODEL

In *The Value of Visualization* van Wijk [10] presents a model that demonstrates how the value of visualization can be quantified and calculated in a deterministic fashion. Supplementing van Wijk's model, we add the two types of knowledge described above (explicit and tacit). While van Wijk believes that "visualization is subjective" and extracting knowledge from data is an objective process that is often not considered in visualization, we propose that it is possible to incorporate general properties of knowledge (explicit) and personal properties of knowledge (tacit) in visualization. We base our modifications to van Wijk's model on the beliefs that:

- Explicit knowledge is different from information.

- Tacit knowledge can only result from human cognitive processing (reasoning).

- Explicit knowledge exists in data, and is independent from the user or his tacit knowledge.

- Explicit and tacit knowledge are related and can be connected through the use of interactive visualization.

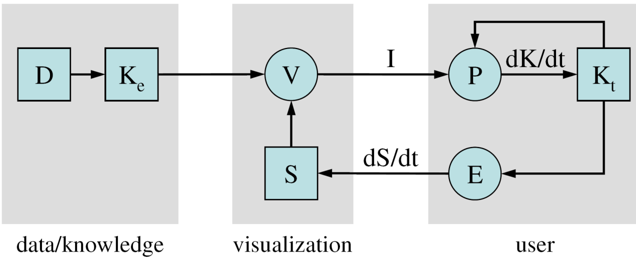

Figure 1: A model of knowledge-based visualization that integrates van Wijk's model [10] with explicit and tacit knowledge

Figure 1 shows an operation model of knowledge visualization that includes our consideration of knowledge (explicit K_e and tacit K_t). Explicit knowledge (K_e), extracted from data (D) is represented as a visualization (V), which is received both perceptually and cognitively (P) by the user via an image (I). The cognitive processing, leading to understanding and an increase of user tacit knowledge (K_t) which recursively affects subsequent perception and cognition. Tacit knowledge guides the user's interaction and exploration (E), so that the specifications (S) that control the visualization change over time.

The relationship between these variables can also be expressed as an set of equations:

$$K_e = f(D); I(t) = V(K_e, S, t) \tag{1}$$

where explicit knowledge K_e is extracted from data D based on some function f. Based on the explicit knowledge K_e, some specification S, an image at time t, $I(t)$, can be created through the visualization V. This image is then perceived and understood by the user, resulting in an increase of the user's tacit knowledge K_t.

While we believe this model is accurate conceptually, in practice, we have come to realize that the amount of explicit knowledge K_e that exists in a complex dataset D could be nearly infinite. The process of converting explicit knowledge into visualization is thus limited by the screen resolution and computer hardware and often cannot be displayed in its entirety. To this end, we refine the model in Figure 1 and subdivide explicit knowledge K_e into smaller collections of knowledge K_{e_1} to K_{e_n} (Figure 2). The challenge of creating a visualization therefore becomes the determination of the most relevant and valuable collection of knowledge to display to the user.

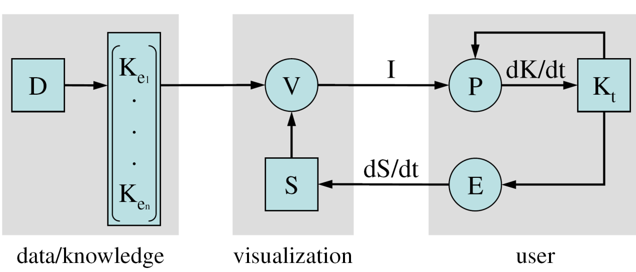

Figure 2: The refined model of the operational model in Figure 1 having the collections of explicit knowledge K_{e_1} to K_{e_n}.

We consider the value of a piece of explicit knowledge to be a function of the relationship between the data elements within the knowledge [6]. However, this value is subject to the cost of displaying the knowledge in terms of rendering cost, clutter, hysteresis, etc. Maximizing the value of a visualization hence becomes an optimization problem of maximizing the overall value of the displayed knowledge while minimizing their cost. This relationship can be written as [5]:

$$\text{maximize} \sum_{i=1}^{n} b_i x_i \text{ subject to } \sum_{i=1}^{n} c_i x_i \leq t_c, \qquad x_i \in \{0,1\} \tag{2}$$

where n is the set of knowledge, b is the cognitive benefit, c is the cost, and x_i is the decision of whether or not a piece of knowledge K_{e_i} is to be displayed. This equation is subject to a target cost t_c, which could be the maximum resolution of a screen, the amount of memory a computer has, or whatever the constraint is that limits all pieces of knowledge to be displayed.

4 Conclusion and Future Work

In this paper, we propose a new approach of visualization that emphasizes the visualization of knowledge instead of data or information. We begin by defining knowledge as either explicit or tacit knowledge, and then we insert this definition into our model of knowledge visualization based on the work by van Wijk. In our model, although explicit and tacit knowledge are different, we show that they can be related and connected through the use of interactive visualization. Finally, we demonstrate that the value of every piece of explicit knowledge can be quantitatively measured as a sum of its benefit and cost. Given limited resources, this measurement can be used to determine the maximum value of visualization by maximizing the total benefits of the displayed knowledge while minimizing their costs.

For future work, we look to apply the framework into our existing visualization applications. We believe that by adopting this model, these visualizations will be able to directly build on the user's existing tacit knowledge through reasoning and cognitive understanding of the explicit knowledge in the data.

References

[1] The American Heritage Medical Dictionary 2007.

[2] R. A. Burkhard. Learning from architects: the difference between knowledge visualization and information visualization. In *Eighth International Conference on Information Visualization*, pages 519–524, 2004.

[3] T. H. Davenport and L. Prusak. *Working Knowledge: How Organizations Manage What They Know*. Harvard Business School Press, 1998.

[4] I. Jurisica and D. Wigle. *Knowledge Discovery in Proteomics*. CRC Press, second edition, 2005.

[5] H. Kellerer, U. Pferschy, and D. Pisinger. *Knapsack Problems*. Springer, 1 edition, 2004.

[6] A. Koike and T. Takagi. Knowledge discovery based on an implicit and explicit conceptual network. *Journal of the American Society for Information Science and Technology*, 58:51–65, 2006.

[7] I. Nonaka and H. Takeuchi. *The Knowledge Creating Company*. Oxford University Press, 1995.

[8] A. Syed and A. Shah. Data, information, knowledge, wisdom: A doubly linked chain? In *International Conference on Information and Knowledge Engineering*, pages 270–278, 2006.

[9] J. Thomas and K. Cook. *Illuminating the Path: The Research and Development Agenda for Visual Analytics*. National Visualization and Analytics Center, 2005.

[10] J. J. van Wijk. The value of visualization. In *Proc. IEEE Visualization 2005*, pages 79–86, 2005.