Research on the Application of Function-Technology-Aesthetics Framework in the Design Knowledge Modelling of Data Visualization

Chiung-Hui Chen

Abstract —**Data visualization is the combination of design aesthetics and rigorous science, and the selection on visual representations is ultimately determined by the integration of aesthetics and design information. For the design knowledge and experience on this aspect is rather deficient, designers often cannot keep the balance between design and functions. This study is aim at developing a data visualization design method targeted to design on design education, to enable students to have specific methods to follow when engaging in information design and to provide reference to designers for design. The content analysis method is employed to discuss relevant literature and to summarize the data visualization design cases, and the data visualization design patterns are established to facilitate learning and knowledge inheritance of data visualization design.**

Index Terms—**Information aesthetics, data visualization, information function, visualization exploration.**

I. Background and Purpose

Data calculation is one of the most important functions in computer science, as the descriptiveness and persuasiveness of complex data are enhanced by graphics, thus, presenting more effective graphical descriptions, which facilitate visual communication and interpretation. Models of mathematical functions and geometric graphics, and even economics, sociology, and meteorology, can be visualized, which extends the dimensions and meaning of design. Exploratory Data Analysis can integrate humans with data exploration, infer conclusions, and directly interact with data through the insight of humans on the visualized data, thus, effectively discovering hidden features and patterns, as well as their future changes, through the visual interface.

Data visualization combines design aesthetics and rigorous science. The selection of visual expressions depends on the integration of aesthetics and design information; in other words, data visualization must keep pace with aesthetic patterns and functions to achieve visualization needs. However, due to the lack of design knowledge and experience, designers often cannot grasp the balance between design and function, thus, creating boring or extremely colorful and complex data visualization patterns. Kalfoglou *et al.* [1] proposed that design knowledge is important for the design industry, design education, and related technologies. Design knowledge refers to exploring solutions in a wide range of possibilities in known and unknown fields, as its value is different from objective technical data. Therefore, the design education of data visualization cannot neglect the visual basis, and must return to communication; if it only focuses on colorful and complex data visualization patterns, it will produce invalid communication.

To sum up, how to effectively organize and use data design is a very important topic in the current mass production of information. This study aims to design a design-oriented data visualization method, which can guide students in specific design methods, and enable designers to incorporate data collection, analysis, and thinking in the design process. In this paper, content analysis is applied to explore the design knowledge in relevant data visualization literature, and the design principles of data visualization are established for the purpose of the learning and knowledge inheritance of data visualization design. The main purpose of this study is to summarize the design patterns, methods, and principles of data visualization.

II. Literature Review

This section discusses the design knowledge modeling of data visualization, and reviews related theoretical models and technical applications, including the functions of data visualization, the hierarchical and network data visualization technology, and data visualization aesthetics, which are described, as follows.

A. Functions of Data Visualization

Different from statistical analysis, Exploratory Data Analysis is not equal to the statistical graphic method based on statistical data visualization. Combining statistics and data analysis, exploratory data analysis is defined as a data analysis method based on data visualization. Its main purposes include: understanding the principles of data; discovering potential data structures; extracting important variables; detecting outliers; testing hypotheses; developing data reduction

Manuscript received August 12, 2020; revised December 3, 2020. This work was supported by the Foundation Name under MOST 108-2221-E-468-002.

Chiung-Hui Chen is with Asia University, Taichung, Taiwan (e-mail: 7451616@gmail.com).

doi: 10.18178/ijke.2020.6.2.131

models; determining optimization factors, etc.

From the perspective of data processing, exploratory data analysis and statistical analysis are also different. The process of statistical analysis is: problem-data-model-analysis-conclusion; while the process of exploratory data analysis is: problem-data-analysis-model-conclusion; for example, Pham *et al.* [2] proposed a web-based visualization analysis tool that allows ecologists to explore long-term ecological data with the aim of generating hypotheses from the data, rather than validating the hypotheses. The first part of this study includes a series of visual display and interaction techniques for data, which allows ecologists to communicate the important patterns and time series trends they discover from ecological data. The second part presents how ecologists can use visualization and exploration of interactive data to generate and test hypotheses.

With the help of artificial intelligence and big data, we have more information models, such as intelligent geometry, building information systems, and intelligent buildings. Based on the pre-processing and post-processing information diagrams of the design of operation systems and interfaces, designers can judge the appropriateness of the diagrams according to the design purposes, and then, consider how to apply it to the design plan, space capacity, use evaluation, and urban space. Therefore, design data visualization can assist in the data exploration and knowledge mining of building design. Then, the data can be classified and extracted through comparison and statistical analysis, in order to facilitate subsequent design decisions [3].

B. Hierarchical and Network Data Visualization Technology

Trees are one of the earliest and most widely used visualization metaphors. Through the tree structure, we can understand the evolution of human consciousness, culture, and society. In the real world, data may contain an internal hierarchy, which can be abstracted into a tree structure, meaning a nonlinear structure defined by branch relationships. Hierarchical data visualization is a long-term research topic, and with the emergence of new requirements, the innovations of hierarchical data visualization are thriving.

Many relevant cases and researches at home and abroad have discussed the effectiveness of visualization technology and the performance of visual graphics. Two typical hierarchical and network visualization technologies, meaning the Space-Filling method and the Node-Link method, are discussed in this study.

1) Space-filling

Three most common space-filling methods are Treemaps, Sunburst, and Hyperbolic Viewer. Treemaps is a visualization technology proposed by Shneiderman [4], which presents the hierarchical data structure in a nested rectangle through clever spatial configuration. The size of the rectangle shows the relationship of certain attributes of the information object.

The second space-filling method adopts the radial distribution of trees in space allocation and division. This method was first proposed by Johnson [5] with the expansion of the tree diagram, called the Polar Treemap. Stasko *et al.* proposed a more complete design in 2000 called Sunburst [6].

Derived from the Polar Treemap, the Hyperbolic Viewer is a visualization technology proposed by Rao and Pirolli [7], which is used for large hierarchical structures. It presents data as a tree structure, and places the tree structure on the Hyperbolic Plane. Through the interactive operations of users, important information will be moved to the center and magnified with the Fisheye View, which clearly shows the hierarchical relationship between objects.

2) Node-link

The Node-Link method mainly includes the Force-Directed Layout, as first proposed by Eades in 1984 [8], which reduces the exchange of edges in the layout and attempts to maintain consistent edges. This method refers to a spring model to simulate the layout process, where the spring is used to simulate the relationship between two points; under the elastic force, the closer points will be bounced away, while the farther points will be drawn closer. Through continuous iterations, the entire layout reaches a dynamic equilibrium and tends to be stable. Thereafter, the concept of “Force-Directed” is proposed, which evolves the layout into a force-directed distribution algorithm.

Through the application of a sub-group partitioning algorithm, the classification categories are calculated. The force-directed layout of nodes is easy to understand and implement, can be used in most network datasets, and produces better symmetry and local aggregation effects; therefore, it is more aesthetically appealing and interactive. Users can see the entire gradual dynamic equilibrium process in the interface, which renders the entire layout result more acceptable.

C. Data Visualization Aesthetics

As the network technology and the global information network become more sophisticated, the volume of data accumulated by humans on the Internet is increasing at an astonishing rate. Artists are interested in the social significance represented by massive data, and have begun to explore network technologies, such as network software and hardware, which logically analyze and operate the data, and reproduce the content structure of the data in visualization graphics, thus, data aesthetics is in network art [9]. The focus of its aesthetic significance is to explore subjective experience constructed by network data, and present the subjective experience in a visual abstraction and reflection form. Artistic data visualization works do not address a certain need, but re-question and construct the topics under discussion in different ways from media cultural products [10]. In other words, it is more important for viewers to understand the idea of the works, rather than reading the information about the works [11].

Moere [12] argued that data aesthetics in network art places too much emphasis on the connotation of art, while neglecting the functionality that should be considered in visualization, which renders the works too abstract and incomprehensible. From another perspective, information visualization is mainly aimed at functions and efficiency; if it is too biased towards practical applications, the potential impact of the aesthetic elements in the works may be ignored. As an independent medium, aesthetics can effectively fill the gap between the functionality and art of the works, which increases the value and functionality of information.

III. Research Content

As constructing a complete design knowledge base is a long-term task that includes many aspects, this study only makes some pioneering attempts in this regard, which aim at the exploration and development of methods. Based on the above literature analysis and research purposes, this study was conducted in three stages. The following describes the implementation method and content in detail:

A. *Function-Technology-Aesthetics (FTA)*

The DIKW model (Data, Information, Knowledge, Wisdom), as proposed by Rowley [13], shows that, through quality cases and literature, data can be converted into experience, which becomes knowledge through analysis and induction, and finally, forms wisdom through application testing. Based on this process, this study starts with motivation, and discusses the meaning of the data visualization model, the functions of data visualization are (abbreviated as F), the hierarchy and network data visualization technology (abbreviated as T), and the data visualization aesthetics (abbreviated as A).

In the field of design education, such as graphic business design, architecture and interior design, industrial product design, and digital media design, the design teaching process must be based on training, as well as surveying or collecting relevant user behavior records, in order to achieve its ultimate goal according to the different characteristics and needs of each field. Then, through the changes in the data, the key issues were identified to help designers make choices between different versions with similar functions. However, due to the lack of design knowledge and experience in data visualization, this study proposed a design-oriented data visualization design framework (FTA), which allows data visualization design to participate in analysis and guide students in a specific design method; moreover, this framework will enable designers to incorporate data collection, analysis, and thinking in the design process, as shown in Fig. 1.

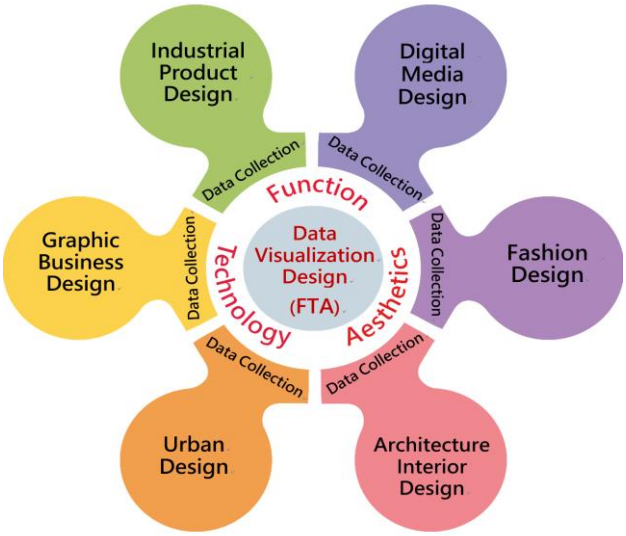

Fig. 1. Design-oriented data visualization design framework (FTA).

B. *Content Analysis*

The case data sources in this study are mainly from the case platform website created by Manuel Lima, VisualComplexity.com [14], which is one of the most influential data visualization case libraries at present. The main goal of this website is to collect and utilize different visualization design methods, such as shapes, colors, sizes, and distances, and provide critical understanding of disciplines and social contexts, thus, inspiring people to engage in research in architecture, urban planning, art, design, fashion, technology, religion, maps, biology, astronomy, and physics. Regarding book references, this study referred to three important visualization books written by Manuel Lima, namely, The Book of Circles-Visualizing Spheres of Knowledge [15], The Book of Trees- Visualizing Branches of Knowledge [16], and Visual Complexity-Mapping Pattern of Information [17], which provided a number of classic cases on information design.

C. *Design Patterns of Data Visualization*

The ancient Roman architect, Marcus Vitruvius Pollio, stated that buildings should include three elements: Firmitas, Utilitas, and Venustas, and discussed how to achieve these standards [18]. This study referred to 20 well-known cases to compile data on data visualization design, and analyzed the theories and various data visualization technologies in these cases. For each case, the framework, as proposed by this study, was used to analyze the performance and content of the constituent elements, and then, the possible and conditional restrictions of the research framework were discussed. Next, the visualization of the design data collection was supported by the proposed Data Visualization Design Framework (FTA) (Fig. 1), the 20 well-known cases were analyzed and summarized, and their dependencies were further constructed, as shown in Table I. Finally, the patterns of each case are described, as follow.

1. The function element (abbreviated as F), is summed up into four items:

(1) Integrate multi-domain experts

(2) Viewed content is based on a non-disposable template

(3) Views are generated by computer operations

(4) View analysis is dominant instead of an application edition

2. The technology element (abbreviated as T), is summed up into five items:

(1) The data itself adopts a content structure

(2) A dedicated data system

(3) Data presents semantic structure

(4) Process can be applied to other data modules

(5) Structure is presented in multiple dimensions, and data objects can be classified, sorted, and combined

3. The aesthetics element (abbreviated as A), is summed up into four items:

(1) Explanatory

(2) Narrative

(3) Ideographic

(4) Story

(5) Social

TABLE I: DESIGN CASE PATTERNS OF DATA VISUALIZATION (EIGENFACTOR)

Case name	Eigenfactor.org (Visualizing Information Flow in Science), 2009
Case author	Maritz Stefaner
Case drawing	
Timing	Bibliographies are often found in academic literature and are an important indicator of the prestige and credibility of scholarly papers. With a bibliography, we can understand the connection between a paper (or a book) and other papers (or books). If 2 papers (or 2 books) are cited by a third party, then, there is an indirect connection between the two papers (or books). In this way, most books and research papers can be used to understand the relationship between these books and papers. If specific data can be obtained, a huge connection graphic can be created to identify similarities between different fields.
How	The data in the view is derived from Thomson Reuter's 2007 Journal Citation Reports, and includes 634926 book references in 6,218 scientific literature. The segmentation in the circle in the 1997-2007 Citation Report of the Thomson Reuters Journal represents different disciplines, such as molecular and cellular biology, medicinal chemistry, earth science, etc., while connection represents the references between disciplinary papers. The small graphics show the details.
F	(4) View analysis is dominant instead of an application edition
T	(2) A dedicated data system (3) Data presents semantic structure
A	(1) Explanatory
Node-Link	The Node-Link methods draw single individuals into nodes, and the connection between nodes represents the hierarchical relationship between individuals. These methods are intuitive and clear, and particularly suitable for representing hierarchical relationships. However, when there are too many individuals, especially when the breadth and depth are hugely different, the readability is poor, and massive data points are concentrated in the local area, making it difficult to efficiently utilize limited screen space.

IV. CONCLUSIONS AND FUTURE RESEARCH

Design knowledge is of great significance to the design industry, design education, and related technologies, and refers to exploring solutions in a wide range of possibilities in known and unknown fields, thus, its value is different from objective technical data. To understand the learnability and knowledge of the data visualization design framework (FTA), as proposed in this study, the results of this study are introduced into formal teaching to verify the learnability of the FTA framework model. The case patterns summarized above are implemented in design teaching, and targets the students of an "Information Visualization Design Research" course. Each case pattern includes a thematic information design application exercise for at least 2 students, in order to test validity (2 Students* 20 well-known cases = 40 sample verifications), and then, constructs the data visualization design knowledge base.

Through this teaching operation, it can be observed whether the framework pattern can be changed according to the design constraints when the design conditions are adjusted, thus, exploring the various possibilities of design. Whether this method helps learners in the data visualization design process, including data collection, analysis and thinking, and whether it complements the interpretation of the qualitative dimension, is one of the future development trends of data visualization technology and intelligent service. Furthermore, suggestions and conclusions are presented regarding the educational direction and thinking of "data visualization design" from the angles of technical application and significance.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ACKNOWLEDGMENT

This study is supported by the Taiwan Ministry of Science and Technology, grant-MOST 108-2221-E-468-002.

REFERENCES

[1] Y. Kalfoglou, T. Menzies, K. D. Althoff, and E. Motta, "Meta-knowledge in systems design: Panacea... or undelivered promise?" *The Knowledge Engineering Review*, vol. 15, no. 4, pp. 381-404, 2000.

[2] T. Pham, J. A. Jones, and R. Metoyer, "Interactive visual analysis promotes exploration of long-term ecological data," *Ecosphere*, vol. 4, pp.1-22, 2013.

[3] N. Cawthon and A. Vande Moere, "The effect of aesthetic on the usability of data visualization," in *Proc. IEEE International Conference on Information Visualisation (IV'07)*, Zurich, Switzerland, 2007, pp. 637-648.

[4] B. Shneiderman, "Tree visualization with Treemaps: A 2-D space-filling approach," *ACM Transactions on Computer Graphics*, vol. 11, pp. 92-99, 1992.

[5] B. Johnson and B. Shneiderman, "Treemaps: A space-filling approach to the visualization of hierarchical information structures," in *Proc. the IEEE Information Visualization '91,* 1991, pp. 275-282.

[6] J. Stasko, "An evaluation of space-filling information visualizations for depicting hierarchical structures," *International Journal of Human-Computer Studies*, pp. 663-694, 2000.

[7] J. Lamping, R. Rao, and P. Pirolli, "A focus+context technique based on hyperbolic geometry for visualizing large hierarchies, proceedings of CHI'95," in *Proc. ACM Conference on Human Factors in Computing Systems*, 1995, pp. 401-408, New York.

[8] P. Eades, "A heuristic for graph drawing," *Congressus Numerantium*, vol. 42, no. 11, pp. 149-160, 1984.

[9] R. Greene, *Internet Art*, Thames & Hudson, London, UK, 2004.

[10] A. Ramirez Gaviria, "When is information visualization art?" *Determining the Critical Criteria. Leonardo*, vol. 41, no. 5, pp. 479-482, 2008.

[11] R. Kosara, "Visualization criticism — The missing link between information visualization and art," in *Proc. the 11th International Conference Information Visualization*, Washington, DC, USA, 2007, pp. 631-636.

[12] A. Vande Moere, "Form follows data: The symbiosis between design & information visualization," in *Proc. International Conference on Computer-Aided Architectural Design (CAADfutures'05)*, OKK Verlag, Vienna, Austria, 2005, pp. 31-40.

[13] J. Rowley, "The wisdom hierarchy: Representations of the DIKW hierarchy," *Journal of Information Science*, vol. 33, pp.163-180, 2007.

[14] M. Lima. (2005). Visual Complexity.com. [Online]. Available: http://www.visualcomplexity.com/vc/

[15] M. Lima, *The Book of Circles: Visualizing Spheres of Knowledge*, New York: Princeton Architectural Press, 2017.

[16] M. Lima, *The Book of Trees: Visualizing Branches of Knowledge*, New York: Princeton Architectural Press, 2014.

[17] M. Lima, *Visual Complexity: Mapping Patterns of Information*, New York: Princeton Architectural Press, 2011.

[18] D. Rowland, Ingrid and H. Thomas Noble, *Vitruvius: Ten Books on Architecture*, New York: Cambridge University Press, 1999.

Chiung-Hui Chen is an associate professor in the Department of Visual Communication Design, Asia University, Taiwan. Her research focuses on intelligent environment design, data visualization, design computing, and information aesthetics theory.