Information Visualization Approach In Marine Fisheries Landing Data

Nor Anita Fairos bt. Ismail
Department of Software Engineering
Faculty of Computing
University Technology of Malaysia,
Malaysia
e-mail: noranita@utm.my

Mohd Shafry Mohd Rahim
Department of Software Engineering
Faculty of Computing
University Technology of Malaysia,
Malaysia
e-mail: shafry@utm.my

Sheikh Nasir bin Kamarudin
Microsoft Certified Professional
Virtual Space Centre
CyberJaya, Malaysia
e-mail: sheikhnasir@virtualspaces.co

Farhan Mohamed
Department of Software Engineering
Faculty of Computing
University Technology of Malaysia,
Malaysia
e-mail: farhan@utm.my

***Abstract*—This paper studied the landings statistical data in marine fisheries by the state of Terengganu for the period of time 2000 until 2009 and to discuss some of the main features on how information visualization technique can be used as a keystone technology for represent these fisheries data. Information visualization (InfoVis) represents an abstract data in graphical representation concepts in such a way that is more natural or easier for human to comprehend. InfoVis is recognized as one of the important way to help users to study, explore, and present information in fisheries data. Today, this emerging technology is important in fisheries and plays a vital role in developing integrated approaches to fishery management and assessment. It helps to convey relatively complex technical information to scientists, managers and decision makers. Since visualization technology provide a high degree of functionality in sampling design, data assimilation, exploratory data analysis and model development, they will continue to play an increasing significant strategic role in fishery management and assessment.**

Keywords: Information visualization; graphical representation; interactive techniques.

I. INTRODUCTION

This development of fisheries sector has been identified as a catalyst to economic growth. The importance of the development of this sector continue to be emphasized by the government and this became apparent to the Era of the New Economic Policy (NEP) that emphasizes the fishing industry as a business that is regarded as an economic resource and the results of protein, as well as provide employment opportunities primarily through marine and fisheries aquaculture farming. In 2009, the fisheries sector accounted for 1.3% of KDNK compared to 1.2% in 2008 [1]. Increased production is in line with government incentives are through the Third National Agricultural Policy (NAP3) with proposals to make Malaysia as an exporter of food commodities in the world by 2010 [3].

Statistical distribution of the fishery landings statistics are important to reflect the performance of production and developments in the industry. Statistics from this figure can help the Department of Fisheries Malaysia to analyze. Furthermore it can identify the factors that influence the activity of fishing and the development of this industry from time to time. The Department of Fisheries Malaysia has published the Annual Statistical Report of Fisheries Malaysia that provides comprehensive and timely data to users on the performance of national fisheries production. This report is a reference to policy makers, researchers, private sector and civil society to obtain more comprehensive statistical information on trends and achievements of the fisheries sector [2]. As it is a printed material, it is difficult to analyze and review.

On top of that, Annual Statistical Report of Fisheries has more weaknesses in terms of performance and visual display issues if to be view in InfoVis. Several interviews have conducted to address these issues. The researchers has thrown up some problems, especially for the performance of landings data information is hard to see because the information only for text and numeric data.

The purposed of this paper is to seeks and highlight a technique that can be represented more efficiently by using an approach known as Information visualization techniques.

II. LITERATURE REVIEW

A. Information Visualization

Visualization is a use of graphical display techniques to representing of data that enhance the understanding. These visualization objective is to provide insight into the underlying

meaning of the data. While Information visualization (InfoVis) is transformation and manipulation of the data to produces information that "is accomplished by organizing it into a meaningful form, presenting it in meaningful and appropriate ways, and communicating the context around it" [6]. Card et al. [7] in the classic Readings on Information Visualization defined InfoVis as "the use of computer-supported, interactive, visual representations of data to amplify cognition".

The field of InfoVis is focuses on the use of visualization techniques to help people understand and analyze data [8]. It is relatively field of research that use of the visual representation approach for conveying abstracts data in intuitive ways. Visual representation and interaction techniques take advantage to allow people to see, use, explore and interact with large amounts of data, and finally help them to understand and interpret the essential facts of the data, thus analyze the data easier. These two components: representation and interaction allow the user to quickly detect patterns, trends, outliers, clusters, and to gain insights the data [9].

B. Non Multivariate

Non Multivariate mainly focus on attribute which are fix or changeable at fix rate over a period of time such year and month are consistence increasing. Type of net or species is fix based on certain period unless new species or new design of net found or created. These type of attribute are easily represented because of its consistence features.

C. Multivariate information visualization

Multivariate data is a dataset that has many dependent variables and they might be correlated each to varying degree. Usually this type of dataset is associated with discrete data models [11]. Most of the data from areas such as engineering, science, business, statistical analysis, stock markets, etc. are multivariate, containing more than three attributes.

In many applications, the data is given in the form of a data table, where each column represents an attribute, and each row represents an observation of these attributes. The number of variates is typically quite large and the attributes may be categorical or numerical. In order to analyze such data, statistical methodologies can be used. However, such methodologies tend to summarize and compress the amount of information, so decisions made based on statistical results can be misleading due to the loss of the overall context of data.

The goal of multivariate visualization depends on the context of the problem but it usually involves the searching for patterns, structure (clusters), trends, behavior, or correlation among attributes [12], [13], [14], [15]. The resulting information is then fed into the exploratory stage of the knowledge-acquiring process to support the elaboration of a hypothesis about the phenomenon responsible for the targeted data. Hence visualization has been considered a helpful tool in augmenting analytic approaches of multivariate data, especially during the exploratory stages of the data analysis process.

D. Multivariate information visualization techniques

The use of multivariate information visualization techniques is intrinsically difficult because the multidimensional nature of data cannot be effectively presented and understood on real-world displays, which have limited dimensionalities. However, the necessity to use these techniques in daily life is increasing as the amount and complexity of data grows explosively in the information age. Thus, multivariate information visualization techniques that are easier to understand and more accessible are needed for the general population. In order to meet this need, there are a several of multivariate information visualization techniques that are usually used as a tool of communication for displaying a complex structure of data. The purpose is to give users a clear and intuitive data understanding.

Georges [22] has carried a study on numerous visualizations techniques that can be used to visualized multivariate data: 1D and 3D scatterplots, Matrix of scatterplots, Heat maps, Height maps, Table lens, Survey plots, Iconographic displays, Dimensional stacking (general logic diagrams), Parallel coordinates, Line graph, multiple line graph, Pixel techniques, circle segments, Multi-dimensional scaling and Sammon plots, Polar charts, RadViz, PolyViz, Principal component and principal curve analysis (PCA), Grand Tours, Kohonen self-organizing maps (SOM). [22]. However, when data volume is bigger with many dimensions, the visualization results of traditional method can be fuzzy. A good solution is dimension reduction [17].

E. Limitations of visual representation for multivariate data

Visualizations of multivariate data are often visually complex and interaction allows users to inspect and probe the presentation for better comprehension. Here are some of the visualization challenges in visualizing the multivariate data.

- Mapping. The importance of perceptual issues in visualization is mapping. The problem is how to be mapped the attributes of marine fisheries landing data into the visual variables representation to convey the information to the user. Finding a suitable mapping of multivariate data into 2D visual form is not a easy task. It usually depends on the nature of datasets to be visualized and is more related to human perception [16]. Bertin has defined seven visual variables such as color, sizes, value, shape, position, orientation, texture or motion [4].
- Dimensionality. Multivariate is datasets contain typically more than three attributes of data. Dimensions higher than three pose a challenge for our cognition to perceive the information [18].
- Visualization of its relationships of attributes. This is a formidable goal that many of the current visualization techniques fail to achieve. Since the number of possible variable relationships grows combinatorial with the number of variables, it is difficult to visualize them simultaneously [18].
- Can't perform exploration activities. Due the size of the data is larger, so it's becomes difficult to present such data in a single visual display and thus making it hard to navigate, relate, and compare data values in multivariate data set intuitively and interactively [19].

- Limitation of human eye-view. Multivariate data sets containing a large amount of variables or dimensions are often hard to analyze and represent visually, due to perceptual of human view and limitations in screen space [20]. The view is a difficult inherent problem. A quantity of data to be represented is too large for the available screen space. This is a problem that comes across rather frequently; given that very often real situation involves a very large amount of data. In these cases, when the display area is too small to visibly support all the element of a visual representation (visualization), certain interaction technique is needed. Elena [21] in carried out her studies on Parallel Coordinates as a 2D visual representation of Multidimensional data. The research mentions that the available screen limits the numbers of dimensions that can be visualized. Displaying a large number of dimensions can reduce the visualization's readability.

F. Interactive techniques and the role in Information Visualization

Information visualization (InfoVis) systems appear with two main components: representation and interaction. The representation component, which roots lie in the field of computer graphics, concerns the mapping from data to representation and how that representation is rendered on the display. The interaction component involves the dialog between the user and the system as the user explores the data set to uncover insights. The interaction component's roots lie in the area of human-computer interaction (HCI). Interaction is an important feature of Information visualization, allowing the user to encourage exploration and working with the data actively.

The representation is no more than a static image if not interaction exists. Static images clearly have analytic and expressive values [10] and their usefulness becomes more limited as the data set that they represent grows larger with more variables. By supporting further exploration of data items, interaction enables users to have multiple perspectives (example generating a different view) and gain insight on the data set. In order to achieve the same variety view of representation without interaction, a huge display space required to present multiple static images. Thus, as can be seen that these two components cannot be separate. Through interaction, some limits of a representation can be overcome, and the cognition of a user can be further amplified [9], [10].

A few frameworks and taxonomies of InfoVis interaction techniques exist, but they typically focus on low-level operations and do not address the variety of benefits interaction provided. Yi [10] have conducted an extensive review of InfoVis systems and his interactive capabilities proposing interaction techniques taxonomies. Based on the notion of user intent, they propose seven categories of interaction techniques widely used in InfoVis that are clear and comprehensive. The proposed categories are based on their own perspective on interaction InfoVis and thus inherently debatable.

III. MARINE FISHERIES LANDING DATA AND THE PROPOSED DATABASE DESIGN

A. Marine Fisheries Data

The data sources of this study derived from the Annual Fisheries Statistics Official Website of the Department of Fisheries Malaysia, Ministry of Agriculture & Agro-based Industry of Malaysia (http://dof.skali.my/web/guest/59). The data compile in pdf file and the report represent in tabular form. Fisheries data have been published in the Book of Fisheries Statistics Annual Statistics, Malaysia from 1954 to 2009. This sample of statistical data involving the number of fishermen, the amount of fishing, types of fishing equipment, types of fishing vessels, fish species, fishing trip activities, fishing zones, type of gear and the other as shown in Fig. 1.

The example of the sample fisheries dataset has been shown in Fig. 2(a) and Fig. 2(b). The Fig. 2(a) shown the sample of dataset in tabular form contains the total Landing of Marine Fish in Terengganu by Gear for the period of time 1996 until 2009. Fig. 2(b) visualize from tabular dataset to graph visualization show that graph has better impact rather than tabular.

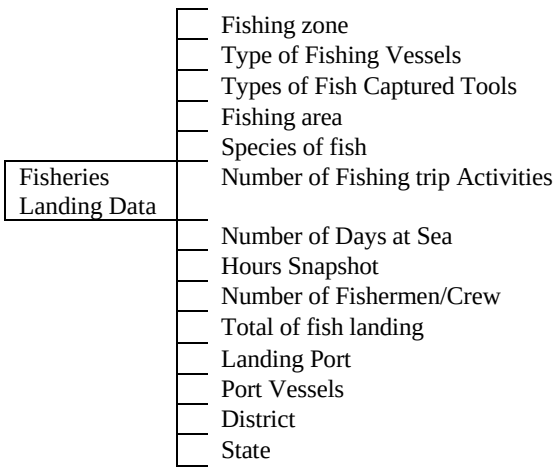

Figure 1: The data that influence the effectiveness of deep-sea fishing industry

Year	Trawl Nets	Fish Purse Seines	Anchovy Purse Seines	Drift/Gill Nets	Lift Nets	Portable Traps	Hook & Line	Total Tonnes
1996	26561	48879	2987	1865	2611	3345	5919	92493
1997	33538	73595	7974	1881	3027	4695	14871	139581
1998	50145	57910	2485	4239	3584	6992	18339	143582
1999	41368	80543	5322	7304	2101	7783	16479	160900
2000	15281	78810	2992	3772	1585	7232	10943	120615
2001	14801	67949	2406	3700	1523	8494	8404	107277
2002	13005	70306	1826	4146	939	7539	8483	105846
2003	15576	55339	846	5028	412	6050	7683	90911
2004	23431	54537	2360	11246	349	5401	10024	106412
2005	26023	43001	1664	6923	807	6729	7865	93011
2006	25090	60787	1496	7454	1754	5846	8966	111386
2007	16916	47275	1030	4204	344	4991	6248	81007
2008	21725	59479	2223	7022	240	6642	7368	104698
2009	14609	42494	1106	11477	503	5565	8566	84319

Figure 2(a): Landing of Marine Fish in Terengganu by Gear (1996-2009)

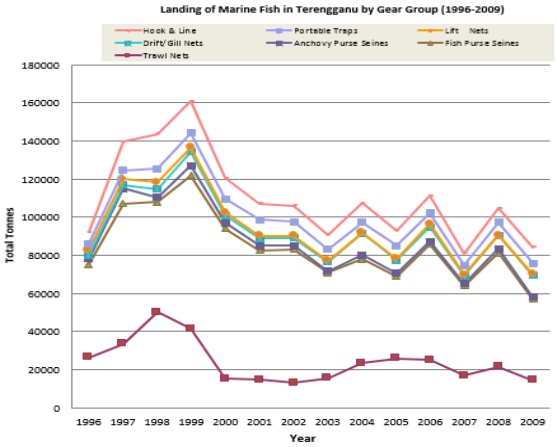

Figure 2(b): From tabular dataset to graph visualization.

B. *Problem Identified in Marine Fisheries Landing Data*

The statistical report of fisheries marine landing data are presented in traditional visualization that is tabular form (column and row). Although the table report is the simplest and most straightforward way to present the tabular data, but it becomes difficult or not quickly interpreted. These problems has been addressed and identified from the statistical report representation of Fishery Annual Statistics after conducting the discussion with marine fisheries researcher at Universiti Malaysia Terengganu (UMT) are:

- Tabular data and tabular report is a format for storing information and represent the data in fisheries. However, as the data grows in size, it becomes increasingly difficult to discover its intrinsic structure or see the comparison among cells from the traditional column-row presentation.
- The information displayed in the form of text and numeric values. They claim that the information display by textual or numeric in context representation is hard to recognize and slow to recall or retrieve the data that are required.
- The fishery landing data is displayed in the traditional visualization technique that is view in tabular form (table) and a graph analysis is presented in a static graphic representation (info- graphics) such as charts, graphs, tables, diagrams, etc. and the presentation is not dependent to technology. Infographics have to maintain a regular presence in the media to display quantitative data, but their value and use as information channels for the general public has not been clearly established as a result of a popular perception of statistical graphics were characterizes as dull or personally irrelevant [9].
- Most of the data in the form of tabular report, and they find it challenging to quickly and effectively absorb the information, identified trends and patterns. The identification, aberrations and see hidden relationships in the data that are often not apparent in traditional tabular representations.
- Data and information analysis that presented in the Annual Fisheries Statistics report is limited. The report only displays information for a certain data analysis. This situation is not giving user the freedom and choice to decide what is the information and data analysis that they is required to see.

C. *Database Design*

In this study, the researchers have developed a proposed model for database design of marine fisheries landings data for the state of Terengganu (See fig. 3 and fig. 4). This proposed model is using a relational database model or better well known as the Relational Database. In the concept of the Relational database model, the relationship between the entities is specified by the key and the type of relationship with other entities. While the physical design database model is developed using SQL Server and ASP.net programming.

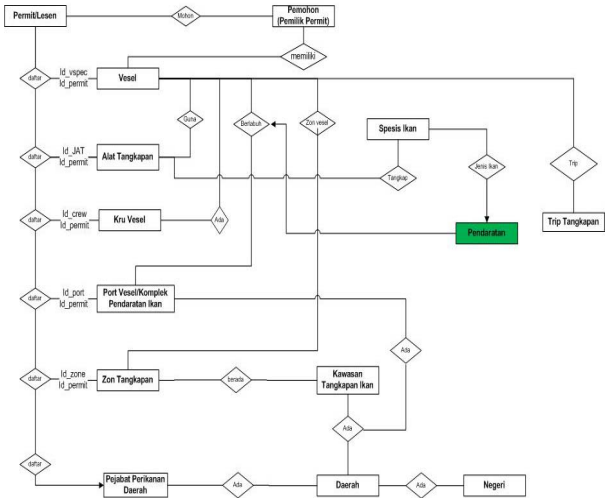

Figure 3: The Proposed model for database design (Logical Design).

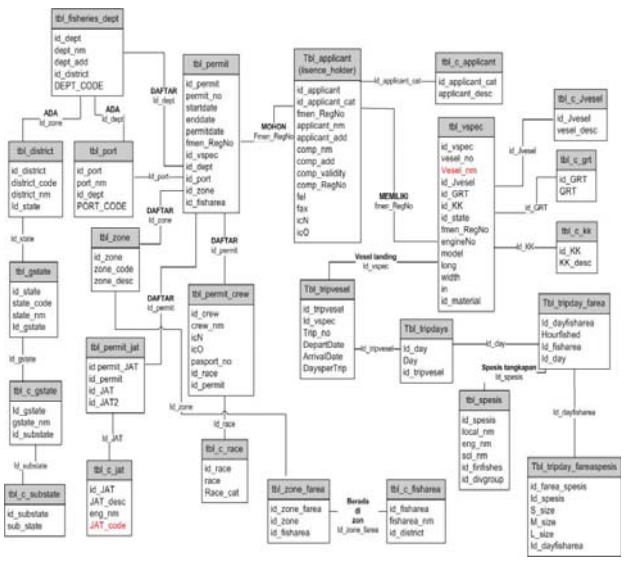

Figure 4: The Proposed model for database design (Physical Design).

IV. Discussions and Conclusion

We have discussed the field of information visualization and the two important components that is representation and interaction techniques that are available to visualization designers. Interaction is most important and the need for it's in marine fisheries domain is required to make an effective visual representation.

Interaction allows the user to work with much larger data sets. It help user to navigate, selection and manipulate the visualization to meet their specific data analysis needs. The user can explore the data better by interaction capabilities such as animation, changing axes, or adjusting the data model and eliminate data by filtering, zooming & panning, and querying. Others interaction mechanisms used to improve visual representation are rotation, linking & brushing, interactive selection, interactive mapping, focus+content & distortion and detail-on-demand. These interactive capabilities providing the user to directly interact with the visual representation (e.g. chart or image) to study the underlying data and discover new information.

There are several potential benefits applying visualization techniques in marine fisheries studies. First, the user may need to spend less memory and processing resources when interacting with the interface. The pattern detection abilities of the user may be enhanced, and the time required to search for information is potentially reduced. By allowing a user to interact with the graphical representation of the data, the structure may be rotated and transformed to allow a better view of previously occluded elements and potentially allows the user to examine relationships, which are not as apparent in a static view.

In marine Fisheries Product profitability analysis can be improved by using visualization to identify the key variable driving profitability and factors influence the activities of fishing (fish landing). Planning and forecasting with visualization providing a quick to understand changes in the pattern and trends of fish landing and their impact on the business plan in this industry for the future. With directly analysis through visualization enables management to dynamically evaluate alternative scenarios. Forecast potential fishing areas. Information on potential fishing area will be extended to fishermen. Indirectly, the fishermen in the future will go down to the sea with a bundle of information was obtained and then help the fishermen to save time finding the fish.

The purpose of this research is referring to the issue of looking for suitable visual representation for marine fisheries landing data and how visualization technique can is used to explore and improve the user comprehension on the fisheries data. The second purpose is to find an intuitive way and efficient interaction methods that capable of enabling the user to interactively investigate and navigate through the visual representation in marine fisheries data. We believe that data visualization provide more sophisticated graphical representation to visually review the output data through graphical indicators such as color-coding, graphics icon with interaction capabilities, which is beneficial to individual, organizations and the society in helping them to gain insight into the marine fisheries data.

References

[1] Jabatan Perikanan Malaysia, Laporan Perikanan Tahunan, pelbagai keluaran. (http://dof.skali.my/web/guest/59).

[2] Perangkaan Pertanian, Kementerian Pertanian dan Industri Asas Tani Malaysia, Oktober 2008.

[3] B.A Talib and A.H Jaafar, "Penangkapan Ikan oleh Nelayan Negeri Sembilan: Satu Kajian Empirika", IJMS 14 (1), 219-230 (2007).

[4] C.Gorg et al, "Chapter 4:Visual Representation, Book Chapter A. Kerren et al. (Eds.): Human-Centered Visualization Environments 2006", LNCS 4417, pp. 163–230, 2007.

[5] Andrews and Keith, Information Visualisation, http://courses.iicm.tugraz.at/ivis/ in: Helmut Mader, "Visualizing Multidimensional Metadata Development of the Visualization Framework MD2VS". Master's Thesis. Institute for Information Systems and Computer Media (IICM), Graz University of Technology, Austria. March 12th, 2007

[6] R. Mazza, "Chapter 1: Introduction to Visual Representation, Book Chapter: Introduction to Information Visualization", Springer-Verlag London Limited, pp. 9,2009.

[7] Card, S., Mackinlay, J., Shneiderman, B, "Readings in information visualization", Morgan Stanley Publishers, San Francisco, 1999.

[8] A.Kerren et. al., "Abstracts Collection: Information Visualization-Human-Centered Issues in Visual Representation, Interaction, and Evaluation", Dagstuhl Seminar Proceedings 07221, 2007.

[9] D.Brodbeck, R.Mazza, D.Lalanne. Interactive Visualization - A Survey. In Collection of Human Machine Interaction, 2009: 27-46

[10] J. S.Yi, Y.a.Kang, J.Stasko, J.A.Jacko, "Toward a deeper understanding of the role of interaction in information visualization". IEEE Transactions on Visualization and Computer Graphics, 13, 6, 1224–1231 (2007).

[11] Selan, "A Framework for the visualization of Multidimensional and Multivariate Data", 2004.

[12] W. Basalaj, "Proximity visualization of abstract data", Technical Report 509, University of Cambridge Computer Laboratory, January 2001. url:http://www.pavis.org/essay/index.html.

[13] A.Inselberg and B. Dimsdale. "Parallel coordinates: A tool for visualizing multi- dimensional geometry". In A. Kaufman, editor, Proceedings of the First IEEE Con- ference on Visualization (VIS '90), pages 361–370. IEEE Computer Society Press, 1990.

[14] V. Interrante. "Harnessing natural textures for multi-variate visualization", IEEE Computer Graphics and Applications, 20(6):6–11, 2000.

[15] E. J. Wegman and Q. Luo, "High dimensional clustering using parallel coordinates and the grand tour", Technical Report 124, Center for Computational Statistics, George Mason University, 1996.

[16] W.Wing, "A Survey on Multivariate Data Visualization",2006.

[17] C.Hui and S.Yichuan, "A Multi-perspective and Interactive Method for Multidimensional Data Visualization", 2nd Conference on Environmental Science and Information Application Technology. IEEE, ESIAT 2010.

[18] H.Siirtola. "Interactive Visualization of Multidimensional Data". Dissertations in Interactive Technology, Number 7. University of Tampere, 2007.

[19] D. A. Keim, "Visual exploration of large data sets". Communications of the ACM, 44(8),39-44, 2001.

[20] S.Johansson. "High Dimensional Data in information Visualization", 2009.

[21] E.Fanea, “Establishing Graphical and Formal Relationships between Visualization of Multi-Dimensional Data” Master’s Thesis, University of Calgary, Calgary, Alberta, 2006.

[22] G.Grinstein et al, “High-Dimensional Visualizations”, Proceedings of the 7th Data Mining, 2001.

[23] A. Buja, D. Cook, and D. F. Swayne, "Interactive High-Dimensional Data Visualization," Journal of Computational and Graphical Statistics,vol. 5, pp. 78-99, 1996.

[24] M. C. Chuah and S. F. Roth, "On the Semantics of Interactive Visualizations," presented at IEEE Symposium on Information Visualization (InfoVis '96), San Francisco, CA, USA, pp. 29-36, 1996.

[25] A. Dix and G. Ellis, "Starting simple: adding value to static visualisation through simple interaction," presented at the working conference on Advanced visual interfaces (AVI '98), L'Aquila, Italy, pp. 124-134, 1998.

[26] D. A. Keim, "Information Visualization and Visual Data Mining," IEEE Transactions on Visualization and Computer Graphics, vol. 8, pp. 1-8, 2002.

[27] L. Wilkinson, *The Grammar of Graphics*, 2nd ed. New York, NY, USA: Springer, 2005.

[28] B. Shneiderman, "The eyes have it: a task by data type taxonomy for information visualizations," presented at IEEE Symposium on Visual Languages, 1996, Boulder, CO, USA, pp. 336-343, 1996.

[29] R. Spence, Information Visualization: Design for Interaction, 2nd ed: Prentice Hall, 2007.

[30] M. O. Ward and J. Yang, "Interaction Spaces in Data and Information Visualization," presented at Joint Eurographics/IEEE TCVG Symposium on Visualization, Konstanz, Germany, pp. 137-145, 2004.

[31] M. X. Zhou and S. K. Feiner, "Visual task characterization for automated visual discourse synthesis," presented at Conference on Human Factors in Computing Systems (CHI '98), Los Angeles, CA, USA, pp. 392-399, 1998.

[32] R. Amar, J. Eagan, and J. T. Stasko, "Low-Level Components of Analytic Activity in Information Visualization," presented at IEEE Symposium on Information Visualization (InfoVis '05), pp. 111-117, 2005.

[33] L.Tweedie, "Characterizing Interactive Externalizations," presented at Conference on Human Factors in Computing Systems (CHI '97), Atlanta, GA, pp. 375 - 382, 1997.