# VISUALIZATION – A REVIEW

K.Tamilselvi[1], K.Ramesh Kumar[2]

*[1]Research Scholar, Research and Development, Bharathiar University, Coimbatore - India*
*[2] Research Guide, Bharathiar University, Coimbatore - India*

---

Abstract-*Visualization is one of the most important components of research presentation and communication due to its ability to synthesize large amount of data into effective graphics. The area of visualization has not yet received the attention in data mining. The importance of visualization is expected to grow and there should be more tools and research areas where the application of visualization techniques provides more insight for the user. The insightness regarding the data for the user can be further obtained by the research area Interactive visualization. The dimensionality of the data and the need of visualization techniques for different types of data will be discussed in this paper.*

KEYWORDS: Visualization, techniques, interactive visualization

## INTRODUCTION

Large Volume of data processing and extracting meaningful data from it has become a challenge in the Bigdata Scenario. Presentation of these explored data requires proper and efficient visualization techniques.

Visualization techniques can support three categories of user tasks [1]:

Explore data: The user does not necessarily have a priori knowledge about the data, nor precise exploration goals. The user looks for meaningful structure, patterns or trends, and hence for formulating a relevant hypothesis.

Confirm a hypothesis: the user looks for certain patterns or structure in data (the user's goal is to verify a hypothesis). Analytic tools may be needed for confirming or refuting the hypothesis.

Produce presentation: the user has a validated hypothesis and his/her goal is to communicate the knowledge to other parties. The focus is on refining the visualization to optimize the presentation.

## 1. Visualization Strategy

Visualization is one strategy for making sense of data. Whether visualization could help to work more efficiently and effectively, it should be focused on understanding data, especially when the understanding relates to a task that must be performed.
Shneiderman [2] proposed the Task by Data Type Taxonomy (TTT) for information visualizations, dividing the visualization techniques into seven data types (one-, two-, and three-dimensional data, temporal and multi-dimensional data, tree and network data) and seven tasks (overview, zoom, filter, details-on-demand, relate, history and extracts) The table visualizations of multidimensional data need to be concerned. The tasks for which the visualizations are employed and evaluated are of exploratory nature.

Keim[3] identifies five categories of techniques:
Standard 2D/3D displays, geometrically transformed displays, icon-based displays, dense pixel displays, and stacked displays. The techniques differ with respect to the ways in which they graphically represent the data dimensions and arrange the data on the screen [4].
Table 1 : Techniques

| Type of technique | Name of technique |
|---|---|
| Variations of standard 2D Displays | Multiple Line Graphs, Permutation Matrix, Survey Plot, Bar graph |
| Geometrically transformed Displays | SOM, Scatter Plot Matrix, Parallel Coordinates, Sammon's Mapping, PCA, Radviz, Star Coordinates |
| Stacked Displays | Tree Map |

 

Standard 2D/3D displays: This 2D/3D displays consists of the effective way for presenting one-, two- and three dimensional data on a standard 2D or 3D display. Examples of techniques are line graphs, histograms, pie charts, doughnut charts, box plots, x-y(-z) plots(or scatter plots), bar and column charts, radar charts, area graphs, stackedbar and columns graphs[5] [6].Variations of the standard 2D/3D displays can be employed for representing multidimensional data, for example, Multiple Line Graphs [5],Survey Plots [7] and Permutation Matrix [5].

Geometrically transformed displays: These techniques helps at finding different transformations and projections of multidimensional datasets [3]. They use geometric transformations and projections to produce useful visualizations [4]. Included are techniques from exploratory statistics (such as Scatter Plot Matrix [8], Principal Components Analysis [9], Dendrograms [10] ).Other techniques are Sammon's Mapping [11], Parallel Coordinates [12], Radial Coordinate Visualization [13], Self-Organizing Map [14] , Star Coordinates [15], etc.

Stacked displays: Stacked displays are representations of data that are partitioned in a hierarchical fashion. When the data are multidimensional, the data dimensions to be used in building the hierarchy have to be selected carefully[16] [3] [4]. An example of technique in this category is Treemap [17]; [18].

To increase the popularity of information visualization techniques among users, Information visualization techniques are to be evaluated. Techniques belonging to variations of standard 2D displays, geometrically transformed displays and stacked displays need to be evaluated so that it can be useful for the user to have understanding and insight about the data.

## 2. Human Factor in Data Mining

Data mining is the process of extracting knowledge from very large amounts of data. The discovered knowledge takes the form of patterns found within the data, patterns that must be interesting to the user (valid, novel, potentially useful and understandable) [19] [20]. This process is also known as the knowledge discovery in databases (KDD). In the context of KDD, data mining (DM) is a step in the process, which is responsible for automatically extracting patterns from data in order to be effective, DM has to have a human in the data exploration process. In this way, the human abilities (flexibility, creativity and general knowledge) are combined with computer performance (storage capacity and computational power). The solution to involve the human directly in the process of data exploration is called by Keim visual data mining (VDM) or visual data exploration.

The information visualization techniques used for VDM can support the user needs in the following ways

1. The user can be involved in direct data exploration, the user can view the data in various graphical representation and gain insight of the data.
2. The user without the knowledge of the data can perform data visualization. The user can also interact with the data.
3. When direct manipulation and visualization of the data is required the user can control the process of obtaining and visualizations generated by a data mining technique.
4. The generated models are represented in a visual form and the user has the possibility to modify the model's or data mining technique's parameters and see the effects of his/her modifications directly on the visualization

## 3. Future Directions in Research

The Interactive visualization and visualization techniques can be useful in many areas for the data insight for the user

1.Nowadays lot of shopping is done on the online. A typical scenario is customers selecting and deselecting a wide range of configuration options of a product in the website. While selecting a product the customer can give his specification for the product. This requires lot of data and visualization techniques and interaction from the user. This area can be expanded for the interactive visualization research

2. Teaching is another field where visualization technique is required. A scenario where virtual reality is used in the educational scheme and the laboratory exercises are incorporated with the student asking queries changing the data and incorporating their ideas even if the data selection (ideas) is wrong. Building an application with right set of data is easier. The application with interactivity as well as virtual reality will pose serious challenges.

 

3. A scenario where medical diagnosis is required for a person with details provided by the patient interactively. This area when developed can be useful when the patient can be involved with the Doctor to identify his ailments where immediate attention has to be paid for diseases like Dengue.

4.A condition where a Spatial Data Mining can be made interactive. For example, to identify nearest ambulance service or a hospital location in case of an emergency, interacting with the system can be helpful to the user community.

## 4. Conclusion

The dimensionality of the data and the various visualization techniques for different types of data are discussed. Visualization and the interactive part of it can provide solutions to the problems in various fields like Medical field, crime field, education field etc. Selecting the proper visualization techniques and interactive intensity will pose serious research challenges.

## REFERENCES

[1] Grinstein, G. G., & Ward, M. O. (2002). Introduction to data visualization. *Information visualization in data mining and knowledge discovery*, *1*, 21-45.

[2] Shneiderman, B. (1996, September). The eyes have it: A task by data type taxonomy for information. In *Visual Languages, 1996. Proceedings., IEEE Symposium on* (pp. 336-343). IEEE.

[3] Keim, D. A. (2002). Information visualization and visual data mining. *Visualization and Computer Graphics, IEEE Transactions on*, *8*(1), 1-8.

[4] Keim, D. A. (2001). Visual exploration of large data sets. *Communications of the ACM*, *44*(8), 38-44

[5] Bertin, J. (*Graphics and graphic information processing*. Waltier de Gruyter.1981).

[6] Davidson, I., & Soukup, T. (2002). Visual data mining: techniques and tools for data

[7] Demšar, J. (2006). Statistical comparisons of classifiers over multiple data sets. *The Journal of Machine Learning Research*, *7*, 1-30.

[8] Cleveland, W. S. (1993). *Visualizing data*. Hobart Press.

[9] Duda, R. O., Hart, P. E., and Stork, D. G. (2000) *Pattern Classification* (2nd Edition).Wiley-Interscience.

[10] Sharma, S. (1995). *Applied multivariate techniques*. John Wiley & Sons, Inc..

[11] Sammon, J. W. (1969). A nonlinear mapping for data structure analysis. *IEEE Transactions on computers*, *18*(5), 401-409.

[12] Inselberg, A. (1985). The plane with parallel coordinates. *The Visual Computer*, *1*(2), 69-91.

[13] HOFFMAN, P. *Table Visualization: A formal Model and Its Applications. 1999* (Doctoral dissertation, Doctoral dissertation, Computer Science Dept., University of Massachusetts at Lowell).

[14] Kohonen, T. (2001). Self-Organizing Maps, ser. *Information Sciences. Berlin: Springer*, *30*.

[15] Kandogan, E. (2000). Star coordinates: A multi-dimensional visualization technique with uniform treatment of dimensions. In *Proceedings of the IEEE Information Visualization Symposium* (Vol. 650, p. 22).

[16] Marghescu, D. (2006). Evaluating the effectiveness of projection techniques in visual data mining. In *Proceedings of the Sixth IASTED International Conference on Visualization, Imaging, And Image Processing* (pp. 186-193).

[17] Johnson, B., & Shneiderman, B. (1991, October). Tree-maps: A space-filling approach to the visualization of hierarchical information structures. In *Visualization, 1991. Visualization'91, Proceedings., IEEE Conference on* (pp. 284-291). IEEE.

[18] Shneiderman, B. (1992). Tree visualization with tree-maps: 2-d space-filling approach. *ACM Transactions on graphics (TOG)*, *11*(1), 92-99.

 