Natural language Processing based Visualization - A Survey

C.Vijayalakshmi *(vijayalakshmi@bsauniv.ac.in)*

Dept of Information Technology, B S Abdur Rahman University, Chennai.

Dr. R. Shriram *(shriram@bsauniv.ac.in)*

Dept of Computer Science and Engg, B S Abdur Rahman University, Chennai.

Abstract

Visualization technologies help to categorize patterns and to mine insights from huge amounts of information. Visualization technology shows significant promise from raising the value of large-scales collections of information. Visualization has been used to communicate information, to monitor trends embedded in data, and to explore huge volumes of data from hypothesis generation. This study gives the review about visualization, its types, search engines, tools and its usage in Natural Language Processing.

Keywords- Natural language processing, Visualization and Search Engine.

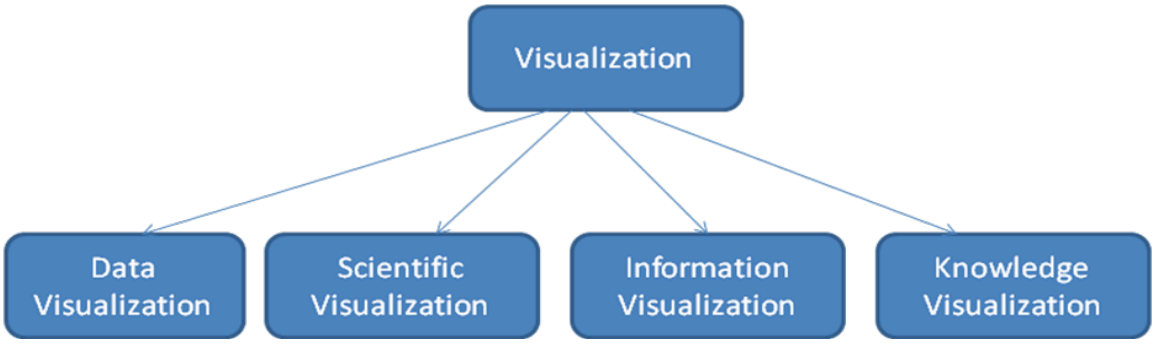
Introduction

Visualization is the graphical presentation of information, with the target of providing the viewer Visualization with a qualitative understanding of the information contents. It is also the process of transforming concepts, objects, and numbers into a structure that is perceptible to the human eyes.[1]

Visualization in Natural language Processing

The process of analyzing the input provided in a human language and transformation of this input into a helpful form of representation. The field of Natural language Processing is mainly concerned with getting computers to perform useful and interesting tasks with human languages. Natural language Processing (NLP) data is wide ranging in scope but has specific characteristics that mean the problems with visualizing large amounts of data are less significant. This is because either the information is visualized as coloured markup on the text, or the information is grouped over small segments of text, such as paragraphs or sentences.[2]

Types of Visualization



Data Visualization

Data visualization is a graphical representation of numerical data. It understands the patterns, trends and relationships that exist in groups of numbers. It involves detection, measurement, and comparison, and is improved by means of interactive techniques and providing the information from various views and with multiple techniques. The precise data visualization tool can present a difficult data set in a way that is simple to understand. [1]

Scientific Visualization

Scientific visualization is concerned with use of computer graphics for the analysis and presentation of computed or measured scientific data. It exemplifies the skill of using computer graphics techniques to explore results from numerical analysis and extract meaning from complex multi-dimensional data sets. Few areas of scientific visualization are flow visualization, chemical visualization, astrophysical visualization and medical visualization. There are a number of different techniques to visualize scientific data, with reconstruction of isosurface and direct volume rendering being the more common.

Information visualization

Information visualization is a computer-aided process that aims to reveal insights into an abstract phenomenon by transforming conceptual data into visual-spatial forms. The information visualization optimizes the use of our perceptual and visual-thinking ability in dealing with phenomena that might not readily lead themselves to visual-spatial representations [3].

Knowledge Visualization

Knowledge visualization represents the data results of a computer-simulated cognitive process, such as learning, perception, reasoning and association of some knowledge acquired by human beings [4]

Visual Search Engines

Interactive graphical and visualization techniques are recommended to increase the ability of the display to handle huge numbers of results while simultaneously presenting several attributes for each Web page. Query reformulation and reconstruction is usually controlled by the search engine leading to redundancy. Integrating the user in the progression of query reformulation is done by visualizing the process itself and it benefits the overall search relevance. An interactive Visual Search Engine visualizes both processes of query reformulation and results presentation [5].

WhatsOnWeb is a search engine which provides a new efficient clustering algorithm, which integrates new graph visualization algorithms and introduces a novel visual strategy for exploring clustered graphs. Experimental analysis is performed to evaluate the effectiveness of WhatsOnWeb+ as a search engine and performances is compared. It features three new illustrative interfaces, based on radial, layered, and treemap and measures the Efficiency of the clustering technique [6].

Yale Image Finder (YIF), a search engine that indexes text found inside biomedical images. YIF provides more comprehensive research results by searching over text that may not be present in the image description, and offers the ability to find related images and associated papers by directly comparing image content [7].

Types of Graph

Graphs represent structured and connected data. Graph visualization has several areas of application. A file hierarchy in a computer system is represented as a tree. It is necessary to navigate through the file hierarchy in order to find a particular file. The problem in viewing graph visualization is the size of the graph. Large graphs cause several complex problems. If the number of elements is huge, it can compromise performance or reach the limits of the viewing platform. This study presents a tree layout, H-tree layout., Balloon view, a cone tree, Hyperbolic view of a tree in 3D, the Klein model for the hyperbolic plane, Fisheye distortion and different schematic views of a tree namely ghosting, hiding, and grouping [8].

Tools

Visualization modules define graphical interfaces for viewing and manipulating data structures, and graphical tools for experimenting with Natural language Processing (NLP)

tasks. This paper presents various modules for interfacing, building and displaying tree structures. It also provides a graphical tool for displaying and simulating finite state automata and interactive graphical tool for experimenting with chart parsers. [9]

Conclusion

Complete human-level natural language understanding is still a distant goal, but there are practical and usable partial NLP systems applicable to many problems NLP methods have opened up new possibilities for high performance text understanding systems and visualization. This work presented a general survey of text visualization, natural language processing, visualization types, visual search engine, graphs and about visualization tools.

References

[1] Zhao Kaidi, Data Visualization, Matrix Number: HT00-6177E.

[2] Rodgers. P, Gaizauskas.R, Humphreys.K and Cunningham.H,(1997). Visual Execution and Data Visualisation in Natural Language Processing. *Proceedings of IEEE symposium on Visual Languages.*

[3] Alhenshiri.A, Brooks.S, Watters.C, Shepherd.M,(2010). Augmenting the Visual Presentation of Web Search Results. *Fifth International Conference on Digital Information Management (ICDIM).*

[4] Chen.M, Ebert.D, Hagen.H, S.Laramee.R, van Liere.R, Kwan-Liu Ma, Ribarsky.W, Scheuermann.G, Silver.D,(2009) Data, Information, and Knowledge in Visualization, Published by the IEEE Computer Society.

[5] Pao-Nan Chou & Ziyan Ma ( ).Trends in Information Visualization Research: A Content Analysis in a Referred Journal

[6] Di Giacomo .E, Didimo.W, Grilli.L, Liotta.G, Palladino.P,(2008). WhatsOnWeb: An Enhanced Visual Search Clustering Engine, *PacificVIS '08. IEEE Pacific Visualization Symposium*

[7] Songhua Xu, James McCusker and Michael Krauthammer,(2008). Yale Image Finder (YIF): A new search engine for retrieving biomedical images. *Bioinformatics Applications Note* Vol. 24 no. 17 pages 1968–1970

[6] Ivan Herman, Guy MelancËon, and M. Scott Marshall,(2000). Graph Visualization and Navigation in Information Visualization: A Survey. *IEEE Transactions on Visualization and Computer Graphics, vol. 6, no. 1.*

[8 Edward Loper and Steven Bird,(2004). NLTK: The Natural Language Toolkit,

[9] Okajima.S and Okada.Y,(2007). Treecube+3D-ViSOM: Combinational Visualization Tool for Browsing 3D Multimedia Data, *11th International Conference on Information Visualization*. IV '07.

[9] Morris.K.R and Mathew R. Schwaller,(2010). Data Visualization and Analysis Tools for the Global Precipitation Measurement (Gpm) Validation Network*, IEEE International Conference on Geoscience and Remote Sensing Symposium (IGARSS),*

[10] Ningning Dang, Fei Yang, Bingjia Xiao, and Yingfei Zhu,(2009). WebScope: A New Tool for Fusion Data Analysis and Visualization. *RT '09. 16th IEEE-NPSS Real Time Conference.*