IJCSE International Journal of Computer Sciences and Engineering Open Access
Research Paper Vol.-6, Issue-10, Oct. 2018 E-ISSN: 2347-2693

Visual Analytics: Need, Process, Scope, Tools & Techniques and Challenges

R. Shankar[1*] , S. Duraisamy[2]

[1,2]Dept. of Computer Science, Chikkanna Government Arts College, Tirupur, Bharathiar university, Coimbatore, Tamilnadu, India

[*]*Corresponding Author : shankarcgac@gmail.com, Mob:- +91 9443492044*

Available online at: www.ijcseonline.org

Accepted: 15/Oct/2018, Published: 31/Oct/2018

Abstract- Visual analytics employs interactive visualizations to integrate users' knowledge and inference capability into numerical/algorithmic data analysis processes. It is an active research field that has applications in many sectors, such as security, finance, and business.. This paper aims at providing an overview of visual analytics, need, scope, tools and techniques and the most important technical research challenges in the field.

Keywords- visual analytics, information visualization, data analysis, user interaction

I. INTRODUCTION

Big data is the term for data sets so bulky and complicated that it becomes not easy to process using conventional data management tools or processing applications. Visual analytics is a form of inquiry in which data that provides insight into solving a problem is displayed in an interactive, graphical manner. The approach uses data visualization technologies to help data scientists and other professionals identify trends, patterns and relationships in the data they are working with. Packaged visual analytics software tools make it easier for non-technical users to use by including drag-and-drop options for setting and modifying analytical parameters. Visual analytics contain various tools and techniques that help user to extract efficient knowledge from complex data and communicate effectively. Visual analytics is a multi-disciplinary area ranging from Bio Technology, Database research, Medicine, Business, Security and also Risk management [2].

Figure 1 Visual Analytics (VA)

Like other types of business intelligence (BI) and business analytics (BA) initiatives, visual analytics applications must incorporate effective data management strategies in order to integrate, unify and standardize data coming from different source systems. Visual analytics is especially helpful in applications involving large, complex data sets and analytical processes that require a high degree of monitoring and interaction -- for example, big data analytics and data mining uses. Figure 1 shows Visual Analytics of big data. Section I mainly deals with the overview of Visual Analytics (VA). This section gives brief introduction to VA, need for VA, VA process and Scope of VA. Section II contains VA tools and techniques and also its benefits. Section III contains research area of VA of Visualization, Interaction and Data Analysis. Section IV contains finding problems and directions of VA and also discuss about the challenges in VA in Big Data. Section V gives an overview on the ideas for future research of VA.

A. The Need for Visual Analytics (VA)

The mother of invention is the need. This new age digital era is full of data and which is exponentially increasing with websites, smart phones and smart technologies connected over the internet. Data is captured like never before. The amount of data generated every minute is mind-boggling [3]. With ever increasing competition and shrinking margins, using data to take better decision is serving a differentiator. This has created a need to look for visual based data analysis that can make sense out of the huge data.

In the field of Business intelligence or data analytics, the insights in visual form are tremendously helpful. It makes the consumption of insights easy to understand and take decisive action. Figure 2. shows that need for data

visualization tools to process and interpret the complex data sets and models. And also how complex analysis tools can be integrated with the coordinated multiple views framework to support exploratory visual analytics and present Meme Media, a framework that can support exploratory visual analytics and open science.

Figure 2 Need for VA Further growing need for Big Data.

Visual Analytics is especially indispensable while dealing with a large chunk of complex data, these days typically referred as Big data. Big data is data sets that are so voluminous and complex that traditional data processing application software is inadequate to deal with them [5]. The properties of Big Data that makes it difficult to handle using conventional techniques shows in table 1.

Table . Five V's of Big Data

Properties	**Meaning**
Volume	Massive Volume of Data
Velocity	Speed of data to be stored, retrieved and analyzed
Value	Ability to turn our data into value
Variety	Data comes from all streams structured, unstructured and semi-structured
Veracity	The data being stored and mined are meaningful to the problem being analyzed ie trustworthiness of data.

B. The Visual Analytics Process

Keim et al also introduced a seminal framework to depict the visual analytics process [4]. The Visual Analytics Process combines automatic and visual analysis methods with a tight coupling through human interaction in order to gain knowledge from data. Fig.1.3 illustrates the entire visual analysis process.

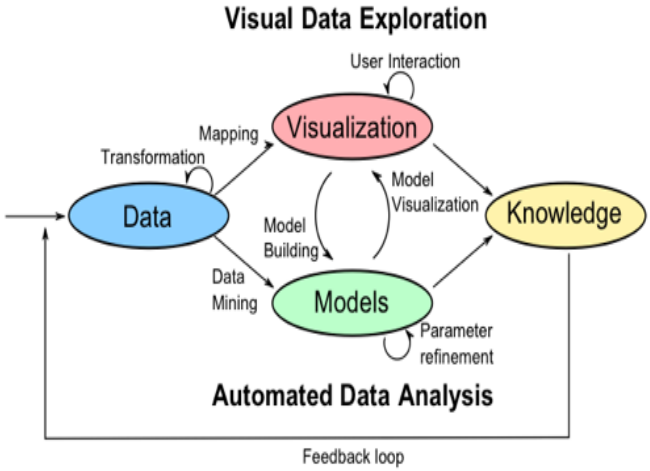

Figure 3 VA Process

In many application scenarios, heterogeneous data sources need to be integrated before visual or automatic analysis methods can be applied. Therefore, the first step is often to preprocess and transform the data to derive different representations for further exploration (as indicated by the Transformation arrow in the figure 3). Other typical preprocessing tasks include data cleaning, normalization, grouping, or integration of heterogeneous data sources [6].

After the transformation, the analyst may choose between applying visual or automatic analysis methods. If an automated analysis is used first, data mining methods are applied to generate models of the original data. Once a model is created the analyst has to evaluate and refine the models, which can best be done by interacting with the data.

Visualizations allow the analysts to interact with the automatic methods by modifying parameters or selecting other analysis algorithms.

Model visualization can then be used to evaluate the findings of the generated models. Alternating between visual and automatic methods is characteristic for the Visual Analytics process and leads to a continuous refinement and verification of preliminary results.

Misleading results in an intermediate step can thus be discovered at an early stage, leading to better results and a higher confidence. If a visual data exploration is performed first, the user has to confirm the generated hypotheses by an

automated analysis. User interaction with the visualization is needed to reveal insightful information, for instance by zooming in on different data areas or by considering different visual views on the data. Findings in the visualizations can be used to steer model building in the automatic analysis. Visual analytics has some overlapping goals and techniques with information visualization and scientific visualization.

There is currently no clear consensus on the boundaries between these fields, but broadly speaking the three areas can be distinguished as from table 2.

Table 2. Scientific Visualization Vs Information Visualization Vs VA

Scientific visualization	Information visualization	Visual analytics
Scientific visualization deals with data that has a natural geometric structure (e.g., MRI data, wind flows).	Information visualization handles abstract data structures such as trees or graphs.	Visual analytics is especially concerned with coupling interactive visual representations with underlying analytical processes (e.g., statistical procedures, data mining techniques) such that high-level, complex activities can be effectively performed (e.g., sense making, reasoning, decision making).

In the Visual Analytics Process knowledge can be gained from visualization, automatic analysis, as well as the preceding interactions between visualizations, models, and the human analysts.

C. Scope of Visual Analytics

Generally, large scale organizations have large amount of data and information to process. They need some strong procedures and techniques to collect, analyze, process and visualize the data in order to get required results as well as to take the right decision in order to get their long term goals and objectives. Several software and tools relating to big data analytics, visual analytics are being used by companies in order to manage their big data or large data sets. Are regards the big data analytics, it is the process of examining big data to uncover hidden patterns, unknown correlations and other useful information that can be used to make better decisions. Figure 4 illustrates the detailed scope of visual analytics. Concerning the field of visualization, visual analytics integrates methodology from information analytics, geospatial analytics, and scientific analytics[7].

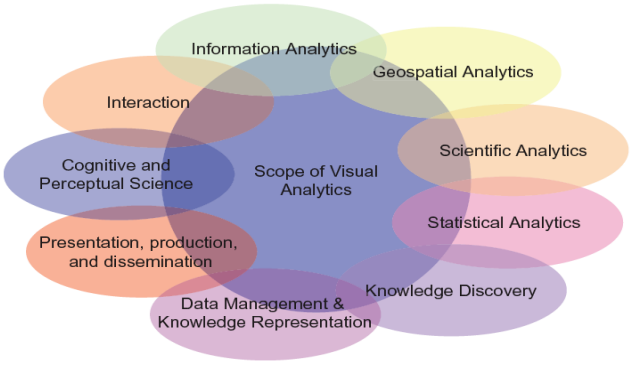

Figure 4 Scope of VA

Bid data tools and techniques are considered most effective for large scale organizations to get required results. Big data analytics can help organizations to better understand the information contained within the data and will also help identify the data that is most important to the business and future business decisions. Big data analysts basically need the knowledge that comes from analyzing the data. The ultimate purpose is to present the data in a perfect to get closer to the natural and ultimate results. While we come to the visual analytics, it is a form of analytics in which interactive graphical displays of data are used to generate analytical results and insights.

II. VISUAL ANALYTICAL TOOLS AND TECHNIQUES

Visual analytics tools and techniques create an interactive view of data that reveals the patterns within it, enabling everyone to become researchers and analysts. It brings together computer science, information visualization, cognitive and perceptual sciences, interactive design, graphic design, and social sciences. Unstructured and semi-structured data types typically don't fit well in traditional data warehouses that are based on relational databases oriented to structured data sets. Further, data warehouses may not be able to handle the processing demands posed by sets of big data that need to be updated frequently -- or even continually, as in the case of real-time data on stock trading, the online activities of website visitors or the performance of mobile applications [8]. Visual analytics enable the human-information, communication by putting together new computational and theory-based tools with innovative interactive techniques and visual representations. Figure 5. shows the design of the visual analytics tools and techniques is based on cognitive, design, and perceptual principles.

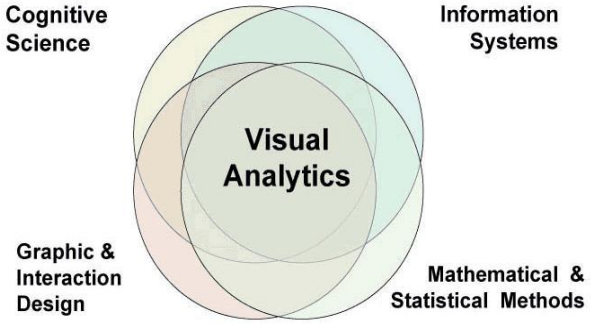

Figure 5 VA Tools and Techniques

To get more precise views, we can say it combines automated analysis techniques with interactive visualizations for an effective understanding, reasoning and decision making on the basis of very large and complex data sets. The ultimate goal is to create the tools and techniques to enable concerned people to identify the expected and discover the unexpected, synthesize information and get insight from massive, dynamic, ambiguous, and often contradictory data. Provide opportune, defensible, and comprehensible assessments and communicate assessment successfully for action to get results [9].

Visual analytics mainly focuses on analytical reasoning techniques that enable users to get deeper insights that directly support assessment, planning, and decision making. Furthermore, it focuses on the techniques to support production, presentation, and dissemination of the results of an analysis to communicate information in the appropriate context to a variety of audiences. It also gets close to the data representations and transformations that convert all types of conflicting and dynamic data in ways that support visualization and analysis and well as visual representations and interaction techniques that take advantage of the human eye's broad bandwidth pathway into the mind to allow users to see, explore, and understand large amounts of organization's data and information at a time.

Some of the most popular tools of visual analytics are given below [6]:

Tableau is an integrated business intelligence (BI) and analytics solution that helps to analyze key business data and generate meaningful insights. The solution helps businesses to collect data from multiple source points such as SQL databases, spreadsheets, cloud apps like Google Analytics and Sales force to create a collective dataset.

QlikView is a Business Intelligence (BI) data discovery product for creating guided analytics applications and dashboards tailor-made for business challenges. The software enables user to uncover data insights and relationships across various sources with QlikView's Associative Data Indexing Engine.

Sisense is an agile business intelligence (BI) solution that provides advanced tools to manage and support business data with analytics, visuals and reporting. The solution allows businesses to analyze big and disparate datasets and generate relevant business trends for them.

Dundas BI, from Dundas Data Visualization, is a browser-based business intelligence and data visualization platform that includes integrated dashboards, reporting tools, and data analytics. It provides end users the ability to create interactive, customizable dashboards, build their own reports, run ad-hoc queries and analyze and drill-down into their data and performance metrics [10].

Domo is a cloud-based business management suite that integrates with multiple data sources, including spreadsheets, databases, social media and any existing cloud-based or on-premise software solution. It is suitable for company sizes ranging from small businesses to large enterprises, and is compatible with Windows or Mac platforms, iPad tablets, and also works on mobile devices.

BOARD Created to combine business intelligence, corporate performance management, and business analytics, BOARD is a full-featured business intelligence system that serves midsize and enterprise-level companies in a variety of different industry segments.

Cognos Analytics is an upgrade to Cognos Business Intelligence (Cognos BI). By adding cognitive guidance, a web-based interface and new data visualization features, Cognos Analytics provides self-service analytics to large and midsize organizations across all industries. At the same time, it provides security features, data governance and management features.

Yellowfin, a business intelligence platform, is a single integrated solution developed for companies across varying industries and scaling sizes.

Halo is an end-to-end supply chain management and business intelligence platform that helps in business planning and forecasting inventory for supply chain management. The system uses data from all sources - big, small, and in-between - to form a cumulative view of all business information.

Looker is a cloud-based business intelligence (BI) platform designed to explore and analyze data. The solution helps businesses to capture and analyze data from multiple sources and make data-driven decisions.

Birst, an Infor Company, is a web-based networked BI and analytics solution that connects insights from various teams and helps in making informed decisions. The tool enables decentralized users to augment the enterprise data model

virtually without compromising data governance. Birst also offers a unified semantic layer that maintains common definitions and key metrics.

InetSoft Style Intelligence is a business intelligence software platform that allows users to create dashboards, visual analyses and reports via a data mashup engine—a tool that integrates data in real time from multiple sources. These sources can include OLAP servers, ERP apps Web services, relational databases and spreadsheets.

Corporater is a cloud-based OSHA- and ISO-compliant business intelligence (BI) solution that provides midsize and large organizations across various industries with applications to plan for and execute their business outcomes. Features include budgeting and forecasting, predictive analytics, strategic planning, project portfolio management and more.

Exago BI is a web-based solution that's designed to be embedded in web-based applications. Embedding Exago BI allows SaaS companies of all sizes to provide their customers with self-service ad hoc, operational reporting, and interactive dashboard capabilities.

Chartio is a cloud-based business intelligence solution that provides founders, business teams, data analysts and product teams in an organization tools to manage day-to-day business operations.

Rapid Insight is an on-premise Business Intelligence solution for higher education institutions and fundraising, healthcare and data science corporations. The suite of applications includes dashboards and scorecards, data mining and predictive analytics, ETL and query and report writing.

icCube Data Analysis & Reporting is a cloud-based business intelligence platform that offers real-time data analysis and visualizations through configurable dashboards, charts and widgets. The solution is accessible on tablets, smart phones and desktop devices

A. Benefits of Visual Analytics

The visual analytics has become an essential part of business. Companies and enterprises are using data visualization technologies in order to speed up their business performance and improve their business decisions making process. So, some key benefits of visual analytics for business:

- Visual analytics software improves the data exploration, minimizes the overall cost and improves the data analysis.
- Visual analytics (VA) make easier the bulk of complex information for better decisions.
- VA enables enterprises to understand data much more quickly and to make faster, better decisions.
- With sharp improvements in computing and data storage, it helps to businesses to solve relevant issues.
- Having the capabilities of solving large and complex issues it offers more accurate results for more profitable decisions for business.
- It offers the different trends of visualization, so the understandable data presentation modes are guaranteed.

III. RELATED RESEARH AREA

Visual analytics is a multidisciplinary field that includes the following focus areas:

- Analytical reasoning techniques that enable users to obtain deep insights that directly support assessment, planning, and decision making
- Data representations and transformations that convert all types of conflicting and dynamic data in ways that support visualization and analysis.
- Techniques to support production, presentation, and dissemination of the results of an analysis to communicate information in the appropriate context to a variety of audiences.
- Visual representations and interaction techniques that take advantage of the human eye's broad bandwidth pathway into the mind to allow users to see, explore, and understand large amounts of information at once[11].

Visual Analytics can be seen as an integral approach combining visualization, human factors, and data analysis.

The figure 6 illustrates the research areas related to Visual Analytics[12]. Besides visualization and data analysis, especially human factors, including the areas of cognition and perception, play an important role in the communication between the human and the computer as well as in the decision-making process. With respect to visualization[8], Visual Analytics relates to the areas of Information Visualization and Computer Graphics, and with respect to data analysis, it profits from methodologies developed in the fields of information retrieval, data management and knowledge representation as well as data mining.

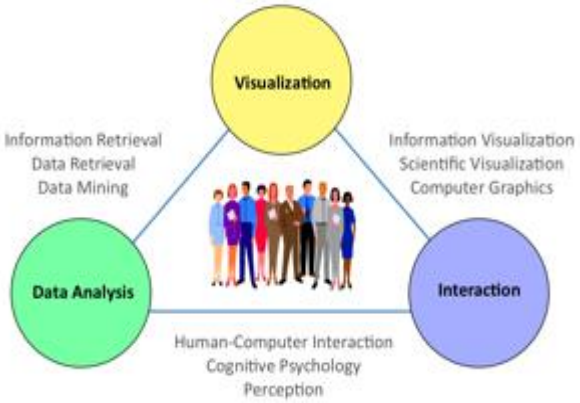

Figure 6 Research Area Related to VA

A. Visual Analytics Challenges

Major challenges of Visual Analytics are given below:

Big data is different, and the three V's and other sense making challenges presented earlier also apply to visual analytics[13][14][15].

- System scalability
- Available Information Space
- Visual Representations
- Interpretability
- Multidimensional data
- Network or relational data
- Workflow and real time analysis

IV. PROBLEMS AND DIRECTIONS

A. Foster Interdisciplinary Research

The fastest path to new solutions to highly complex problems, as targeted by visual analytics, is the promotion of interdisciplinary research projects where heterogeneous teams address real-world problems and develop visual analytics methods and tools that help in investigating and solving these problems.

B. Large-Data Visualization

This challenge focuses primarily on data presentation in VA, which includes visualization techniques and the visual display of information. Recent R&D in abstract visualization, highly scalable data projection, dimension reduction, high-resolution displays, and power wall displays has helped overcome aspects of this challenge.

V. CONCLUSION

Visual Analytics is a new and up growing technology. VA is a form of analytics in which interactive graphical displays of data are used to generate analytical results and insights. In this paper, We believe that in future researchers will pay more attention to VA tools and techniques to solve problems of big data effectively and efficiently.

REFERENCES

[1] K. Parimala, G. Rajkumar, A. Ruba, S.Vijayalakshmi," Challenges and Opportunities with Big Data" International Journal of Scientific Research in Computer Sciences and Engineering, Vol.5, Issue.5, pp.16-20, October (2017)

[2] Shilpa M Patil, Dr. H S Guruprasad," A Study on Visual Analytics" International Journal of Computer Science and Information Technology Research, Vol. 3, Issue 1, pp: (303-308), Month: January - March 2015.

[3] Yuzuru Tanaka, Jonas Sjobergh, Keisuke Takahashi," A Need for Exploratory Visual Analytics in Big Data Research and for Open Science, 2016.20th international Conference information Visualization (IV), ISSN: 2375-0138, **DOI:** 10.1109/IV.2016.42

[4] Keim D A, Kohlhammer J, Ellis G, Mansmann F. Mastering the Information Age: Solving Problems with Visual Analytics. Florian Mansmann, 2010.

[5] https://www.proschoolonline.com/blog/visual-analytics-need-model-tools-applications-2/[4].

[6] http://www.vismaster.eu/faq/the-visual-analytics-process/[5].

[7] Daniel A. Keim, Florian Mansmann, J¨orn Schneidewind, Jim Thomas, and Hartmut Ziegler" Visual Analytics: Scope and Challenges",First Publ in: Lecture notes in Computer Science No: 4404(2008), pp 76-79[6].

[8] https://searchbusinessanalytics.techtarget.com/definition/big-data-analytics.

[9] http://infinitylimited.co.uk/visual-analytics-key-attributes-scope-and-advantages/

[10] https://www.softwareadvice.com/bi/visual-analytics-tools-comparison/

[11] https://x.smu.edu.sg/node/68.

[12] http://www.visual-analytics.eu/faq/

[13] Boyd, D. and Crawford, K.,Critical Questions for Big Data, Information, Communication & Society, May 2012.

[14] Keim, D.A., Kohlhammer, J., Pohl, M., Santucci, G., Solving Problems with Visual Analytics, Procedia Computer Science, 2011.

[15] Czerwinski,M., DeLine, R., Drucker, S., Fisher, D.,Interactions with Big Data Analytics, 2012.

Authors Profile

Dr.R.Shankar born in 1973 completed PG degree in the year 1998 and Ph.D from Bharathiar University, Coimbatore, India in 2018. He is guiding students towards M.Phil programme in the areas like Network Security and Cryptography and Routing and Data Mining. He has published 12_papers in International Journals and more than 7_ conferences National and International level. He has over 20 years 6 months in teaching experience in the field of higher education. Presently he is working as Assistant Professor in computer science Department at Chikkanna Government Arts College, Tirupur. His research interests specialized in Speech Quality in VoIP, Video Streaming and Big Data.

Dr. S. Duraisamy born in 1973 completed PG degree in the year 1997 and Ph.D from Alagappa University, Karaikudi, India in 2008. He is guiding students towards Ph.D. programme in the areas like Software Engineering, Software Quality Assessment, Software Quality Management and Big Data. He has published 70 papers in International Journals and more than 20 conferences National and International level. He is also organized Seminars/ Conferences/Workshops from funding agency like AICTE, DRDO and TCS. He is a life time member of ISTE. He has 21 years of teaching experience and 18 years of research experience. Presently he is working as Assistant Professor in computer science Department at Chikkanna Government Arts College, Tirupur.

