Computer Science Review 36 (2020) 100245

Contents lists available at ScienceDirect

Computer Science Review

journal homepage: www.elsevier.com/locate/cosrev

Survey

Survey on Visualization and Visual Analytics pipeline-based models: Conceptual aspects, comparative studies and challenges

Hela Ltifi [a,b,*], Christophe Kolski [c], Mounir Ben Ayed [a,d]

[a] University of Sfax, National Engineering School of Sfax, REGIM-Lab, Research Groups in Intelligent Machines, LR11ES48, 3038, Sfax, Tunisia
[b] Computer Science and Mathematics Department, Faculty of Sciences and Techniques of Sidi Bouzid, University of Kairouan, Tunisia
[c] LAMIH-UMR CNRS 8201, Univ. Polytechnique Hauts-de-France, Valenciennes, France
[d] Computer Sciences and Communication Department, Faculty of Sciences of Sfax, Route Sokra Km 3.5, BP 1171 3000 Sfax, Tunisia

ARTICLE INFO

Article history:
Received 24 September 2019
Received in revised form 28 March 2020
Accepted 23 April 2020
Available online 4 May 2020

Keywords:
Information Visualization
Visual Analytics
Model
Pipeline
Survey

ABSTRACT

In the last decade, many studies have focused on visualization. The main key to make it practical for research and engineering applications is the suitable definition of a pipeline-based visualization model. It provides effective abstractions for designing visualization tools. In this article, we present a comprehensive survey on pipeline-based models for Information Visualization. The basic principles of visualization models and processes are reviewed in literature as proposed in the original versions and their different revisions. Their significant extensions are grouped into two main categories: Information Visualization process and Visual Analytics process. We propose three summaries: The first one focuses on the conceptual aspects; the second one compares existing studies, and the third one discusses the challenges and opportunities for future research in the field of visualization.

Contents

1. Introduction

Visualization is the study of interactive visual representations of abstract data to reinforce human cognition [1]. Abstract

* Corresponding author at: Computer Science and Mathematics Department, Faculty of Sciences and Techniques of Sidi Bouzid, University of Kairouan, Tunisia.
E-mail addresses: hela.ltifi@ieee.org (H. Ltifi), Christophe.Kolski@uphf.fr (C. Kolski), mounir.benayed@ieee.org (M. Ben Ayed).

https://doi.org/10.1016/j.cosrev.2020.100245
1574-0137/

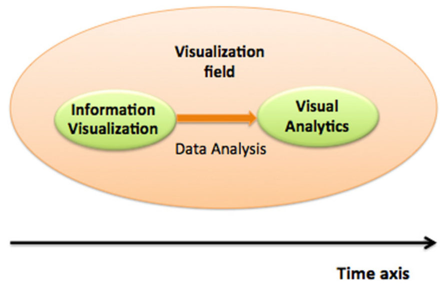

Fig. 1. Positioning in relation to the field of visualization.

data includes both numeric and non-numeric data. The field of visualization has emerged from research on Human–Computer Interaction, computer science graphics, visual design psychology, and business methods [1–4]. It is increasingly integrated as an essential component in scientific research. Visualization assumes that visual representations' forms and interaction methods profit from the human eyes broad bandwidth pathway into the mind to permit users seeing, exploring and comprehending vast amounts of information [1,5,6]. Visualization research focuses on the proposal of methods, techniques and frameworks for transmitting abstract information and analysing data. It is related to the cognitive abilities of human analysts allowing discovering unstructured exploitable insights, limited only by human imagination and creativity. Analyst does not need to learn any sophisticated methods to interpret visualizations [7]. Visualization is also considered as hypothesis generation scheme that can be tracked by more formal analysis [8].

The field of visualization has evolved into a prosperous research field in the recent 30 years [3]. It has progressed into three main branches: (1) Scientific Visualization that concerns scientific data, typically physically based [1], (2) Information Visualization that concerns nonnumeric, non-spatial, and high-dimensional data [1] [9], and (3) Visual Analytics [10] that supports users to make analytical reasoning using interactive visual human–computer interfaces.

As our study is not limited to a specific data type or application domain, we merge Scientific Visualization with Information Visualization. Data analysis makes the data useful by automatically identifying the most interesting aspects of its structure. The integration of this discipline in Information Visualization gave birth to Visual Analytics (cf. Fig. 1). This allowed generating graphical representations highlighting relationships difficult to capture by direct data analysis.

Several Information Visualization models and applications were proposed and published in literature. An outgrowth of this field is the Visual Analytics (cf. Fig. 1), which focuses on the analytical reasoning facilitated by interactive visual interfaces [10]. It is the key technology to deal with and understand massive amounts of complex data that are streaming into organizations, to discover relationships between pieces of data, to build knowledge and to make appropriate decisions [11,12]. Visual Analytics includes analysts who have long-term and strategic views [2,8]. Its primary target is the close association of human reasoning and automated methods.

1.1. Existing survey studies on Information Visualization and Visual Analytics

Numerous surveys on visualization models and processes exist. In 1998, Geisler performed a survey on Information Visualization applications and techniques based on data type [13]. In 2008, Zudilova and her colleagues [14] discussed interactive visualization by presenting a large range of topics such as data representation and user interface. In 2013, Moreland presented the most prevalent features of basic visualization pipelines [15]. In 2014, Liu and his colleagues [16] examined the research trends concerning the empirical methodologies, user interactions, visualization frameworks, and applications. Furthermore, some surveys mainly focused on the Visual Analytics techniques and applications and generally limited to a single data type or a specific application domain. For example, Andrienko and Andrienko [17] who reviewed visual-analytics techniques supporting the movement data analysis, and West et al. [18] who conducted a systematic literature review of Visual Analytics approaches dealing with complex clinical data. Other surveys emphasized on specific kinds of visualization, such as graph visualization [19], software visualization [20] and visualization construction tools [21]. In 2016, Wang et al. reviewed previous works presenting Visual Analytics pipelines from different perspectives (i.e. data, visualization, model and knowledge) [4]. In addition, more recent surveys, such as [22,23] and [24], partially deal with Visual Analytics pipelines.

1.2. Research objectives

Most of the previously presented literature surveys focused on reviewing the state-of-art in a certain direction in-depth. To the best of our knowledge, they do not show the scientific community's shift from Information Visualization to Visual Analytics. In this work, we aim to conduct a comprehensive survey that takes into account all of the most cited and latest Information Visualization and Visual Analytics literatures as a whole. The ultimate purpose of this survey is to draw a complete picture of the progression of the Information Visualization and the Visual Analytics pipeline-based models by investigative the related works. We underline in particular their versions and extensions in order to explore, take advantage and capitalize the evolution in this field, without being limited to a specific data domain or domain application.

In this article an organized comprehensive overview is introduced. It discusses the Information Visualization and Visual Analytics pipeline-based models to investigate their contribution for modelling decision support systems, the commonalities between them, and what conceptual aspects have been proposed/extended. This study describes their different evolution, comparing studies, and discussing challenges and research opportunities in the field of visualization.

The remainder of this paper is organized as follows: it first presents the used literature analysis methodology to explain how the visualization literature is analysed. Then, it describes the evolution of the Information Visualization models through their versions. Next, it reviews the various significant extensions related to the Visual Analytics. Finally, three summaries are provided at the end of the paper to present synthesis discussions and several challenges and research opportunities.

2. Literature analysis methodology

To have an overview of the research progress of Information Visualization and Visual Analytics until 2019, we examined and reviewed related literatures in the field of visualization. We started our survey from key literature sources in the field: (1) [25] and [5] for Information Visualization and (2) [26] and [27] for Visual Analytics. We covered premier conferences and journals on visualization, which are IEEE Information Visualization (IEEE InfoVis), IEEE Visual Analytics Science and Technology (IEEE VAST), Information Visualization (InfoVis) Journal, and IEEE transactions on Visualization and Computer Graphics Journal (IEEE TVCG). We reviewed 112 papers to identify proposed models, their evaluation and related perspectives. We used the survey methodology applied by [28] to conduct our literature review on existing relevant visualization pipeline-based models. The workflow of our literature analysis is composed of five main steps:

1. *Identify motivational requirements*: we found several survey papers (such as [20] and [29]) that provide us motivations for conducting the in-depth literature study in the area of visualization, which is considered as an important and timely topic [29]. We aim to describe in details the evolution of the visualization pipeline-based models and the various significant extensions. The purpose is to identify conceptual aspects and reasoning logics behind them for knowledge generation, to perform comparative studies and to present future challenges. Such knowledge generation acquiring has become a major goal for decision-makers.
2. *Conduct Literature review*: in this step, a set of specific descriptors is proposed to represent the main challenges of this work. Referred to a set of literatures about theoretical frameworks, pipeline-based models and processes on visualization and Visual Analytics, we identify and summarize our main findings based on four types of descriptors: (1) Fundamental information that includes the paper title, the affiliation and the conference or journal where the paper is published, (2) Domain type that can be Information Visualization or Visual Analytics, (3) Data type that concerns what type of data the paper works on (i.e. Temporal, Spatial, Multi-dimensional, Hierarchical, etc.) and shows if the paper proposal has specific requirement on the data type or not, and (4) Design category that depicts the paper proposal category (i.e. model, framework, technique, etc.).
3. *Report findings of the analysis*: in this step, authors code the literature using descriptors and resolve it with conflict descriptors (if exist). A hybrid quantitative and qualitative literature analysis approach is executed [28,29]. The quantitative analysis is gained from the statistics of descriptors where the qualitative one is based on the authors' experiences to derive the research topics of interest. This step is time consuming; where sometimes re-reading is required to re-adjust our observation. Nevertheless, this aided us to gain greater insights into our analysis.
4. *Outline and organize findings*: it consists of reporting, sorting and generalizing our findings that are of significance.
5. *Reveal further research avenues*: the findings from our survey will highlight and discuss a number of new visualization research axes that could be explored in future studies.

3. Survey on Visualization pipeline-based models

In this section, we present our survey on visualization research papers focusing on pipeline-based models. We classify them into two categories: visualization pipeline-based models and Visual Analytics pipeline-based models.

3.1. Survey on the Information Visualization pipeline-based models

A visualization pipeline based-model refers to the steps that data undergoes from their raw format to final visual representation. In literature, several models for the Information Visualization have been proposed. These models allowed interesting progress in the formalization and automation of visualization process, based on the input data. First, in 1983, Bertin explained fundamentals of Information Visualization [25]. He proposed a model, which allows creating visual representations using graphical symbols and marks to convey meaning.

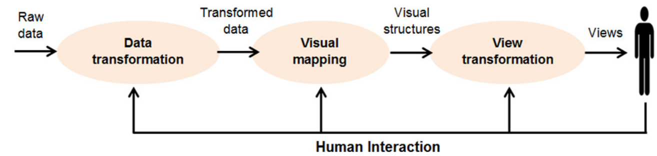

Fig. 2. Information Visualization reference model [1].

Shneiderman proposed a well-known information seeking mantra: "Overview first, zoom/filter, details on demand" [5]. He focused on: (1) the data types modelling by proposing taxonomy based on the dimension and data structure criteria, and (2) the control types modelling by proposing taxonomy of tasks in visualization. These control types are discussed within a survey performed by Roth who introduced three principal groups of controls [30]: (1) objective-based tasks and purposes of use, (2) operator-based Human–Computer Interfaces and controls, and (2) operand-based classifications of visual properties transformation. Such modelling can be used to specify the data that can be visualized, its visual form and user interactions with it.

Card and Mackinlay proposed a model that uses a table to characterize the visualization [31]. Its rows list the visualized information and its columns present the process of data filtering, transformation, projection, and possible interactions with it. In 1999, Card et al. proposed a visualization pipeline for data displaying, cf. Fig. 2 [1]. It is considered as the basic pipeline of Information Visualization. It explains the step-wise process of producing visual representations. The first step consists of transforming raw data into a well-organized data format (data set). The second step is to map the dataset into visual form containing visual structures corresponding to the dataset entities. The final step transforms this visual form into interactive views. The view is then displayed to the user across the human visual system. Users interpret this view and can interact with any of these steps to adjust the resulting visualization, and make additional interpretations.

A refinement of Card et al.'s [1] visualization pipeline is proposed by Ed Chi [6]. It allows supporting user interaction and data spaces (cf. Fig. 3). Three successive data transformation steps into visual variables describe the visualization process of Ed Chi [6]. This model differentiates between: (1) data called value, (2) metadata called analytical abstraction, (3) visual information called the visualization abstraction, and (4) visualization mapping where the user interprets the displayed representation, called view. In each of these four data spaces, the user can interact with information using controls. This model was applied to the Information Visualization and scientific visualization domains, where data do not always have predefined representations.

Numerous of today's applications are dynamic. Manipulated data are named temporal data. In this context, Daassi et al. extended Ed Chi's visualization pipeline to visualize temporal data [32]. The proposed extension consists of a combination of two dimensions: Temporal Dimension and Structural Dimension. For each dimension, they refined Ed Chi's process into a new structure composed of four phases, which is illustrated on Fig. 4: (1) data/Time, (2) point of view on Time/point of view on the structural dimension, (3) Time space/structural space and (4) point of view on the time space/ point of view on the structural space. The resulted visualization process is defined as an association between two visualization processes corresponding to the two dimensions where different visualizations configurations are possible.

Heer and Agrawala proposed adapting the data and visual space models to the application domain; this proposal is recognized as the *Reference Model pattern* [11] for Information Visualization. It is presented by Fig. 5. This reference model suggests

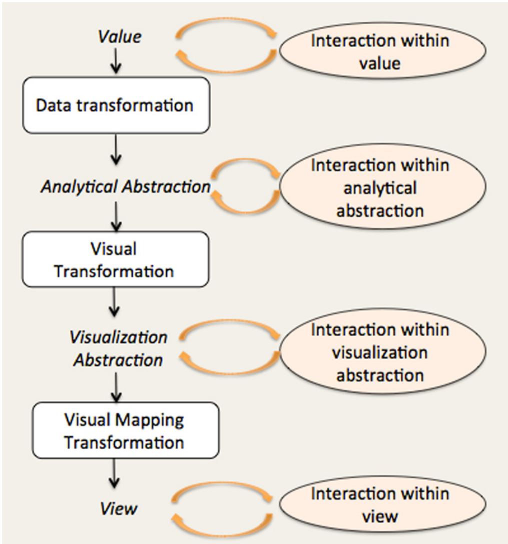

Fig. 3. Data State Reference Model [6].

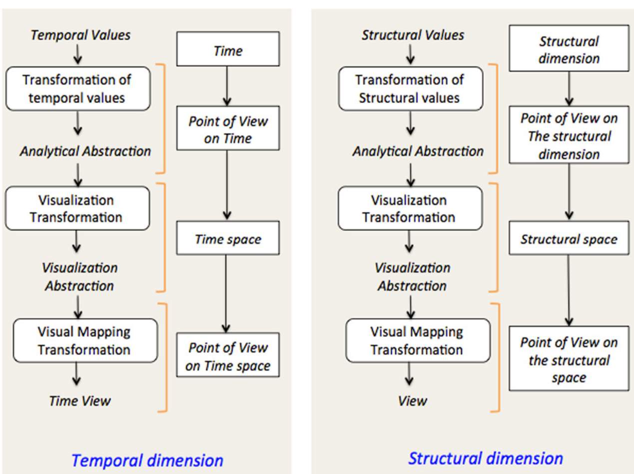

Fig. 4. Visualization process of temporal data [32].

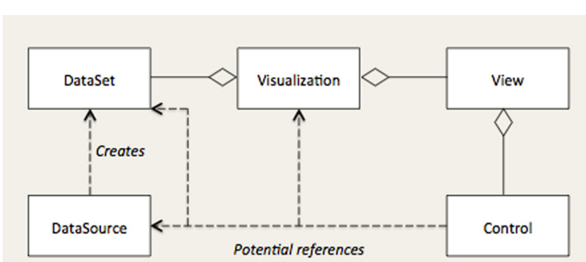

Fig. 5. Reference Model pattern [11].

five classes: the data source providing the data set, the view space and its controls, and the visualization representing a precise data set.

Based on these models and processes, software such as *Show Me* [33] or *AutoVis* [34] analysed data structure to generate automatic visualization. However, these visualization processes are only based on visualization theory, without taking into account the user's behaviour. To improve these models, a set of interesting research works has been proposed for intelligent visualization. We can cite, for instance, the work of Gotz and Wen [35]. They proposed a behaviour-driven visualization recommendation, instead of the data driven or task-driven models for Information Visualization, based on the current use of visualization and explored to infer a future analysis. Another work of interest has been introduced by Hipp and his colleagues who proposed the integration of a semantic representation of process, information and context for appropriate user visualizations [36]. Jansen et al. proposed an interaction model for beyond-desktop visualizations [37]; it consists of combining the visualization reference model of [6] with the instrumental interaction paradigm[1] of Wills and Wilkinson [34]. In this refined pipeline-based model, raw data is transformed into visualization and then rendered into the physical world. Users can modify and explore data by directly using visualizations or across instruments. Interactions may also be executed in the physical world.

In this section, we have listed the most-cited pipeline-based Information Visualization models since 1983. We have presented a representative survey that shows the fast-rising aspect of this visualization branch providing designers an understanding of well-known proposed models. This state-of-the-art cannot be sufficient without drawing comparative summaries in the field. It is the objective of Section 4. But before that and in order to complete our survey study, we investigate in the following section the branch of Visual Analytics.

3.2. Survey on the Visual Analytics pipeline-based models

While Information Visualization deals with representing data and providing interaction techniques, Visual Analytics focuses on leaning more towards computation and analytical reasoning using interactive visualizations. It is the key technology to make sense and understand large amount of data as well as to discover important relationships between data pieces [2]. We can classify Visual Analytics models into three categories: general Visual Analytics models, Visual Analytics models for knowledge generation and predictive Visual Analytics models.

3.2.1. General Visual Analytics pipeline-based models

Several Visual Analytics models are proposed in literature. They are useful to design tools for users (i.e. analysts, decision makers) in order to combine their human capacity with the huge processing capabilities of today's computers to better understand the complex issues. Keim et al. [27] introduced that only representing data using visual metaphor rarely gives any insight. These authors adapted the mantra of [5] for visual analysis in order to gain profound insights and highlight the combination of numerical data analysis with interactive visual interfaces: "Analyse first, show the important, zoom/filter, analyse further, details on demands".

Several researchers presented different means to improve the classic Information Visualization model of [1]. From these, Bertini et al. [39] introduced overlaying the Quality-Metrics-Driven Automation on the classic model of [1]. These metrics can be integrated into its different steps to automate the data analysis and support the visual analysis. The human cognition model introduced by Green et al. depicts the human hypotheses generation process [40]. They underline that computer and human collaborate to search and discover models from visualization. Hypotheses are made and analysed through the analytic process of competing hypotheses. This process is composed of the following steps: (1) generate hypotheses, (2) list evidences, (3)

[1] "It is an interaction paradigm based on the concept of tool, or instrument, that mediates interaction between the end-users and the on-line objects they want to manipulate" [38].

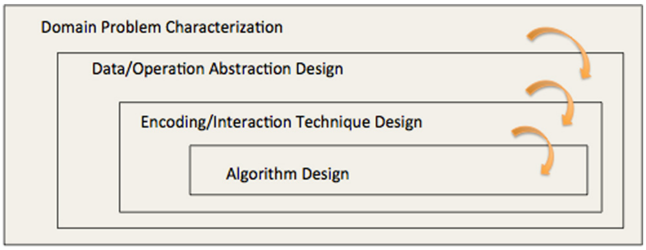

Fig. 6. Nested model of [51].

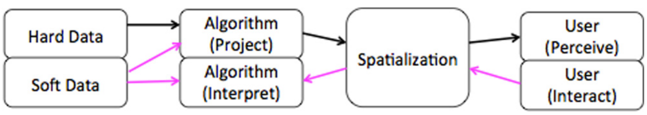

Fig. 7. Semantic interaction pipeline [54].

prove/disprove, (4) create the hypotheses and evidences matrix, (5) draw conclusions, and (6) re-analyse conclusion based on evidences.

Choo et al. [41] and Endert et al. [42] proposed theoretical Visual Analytics models to better understand how to integrate algorithmic and visual components. They focused on high-level and abstract views. Crouser et al. highlighted the prominence of the Human–Computer Interaction in the Visual Analytics process [43]. In addition, interesting Visual Analytics models and guidelines have been made to greatly increase the advancement of the field [44–50]. Munzner et al. [51] proposed the nested model that structured the visual analysis modelling into four layers, where each one is based on the output of the previous one (cf. Fig. 6). These layers are: domain problem characterization, data/operation abstraction design, encoding/interaction technique design, and algorithm design. Designers using nested model must identify and comprehend the system requirements (i.e. tasks and data).

Sedlmair et al. [52] suggested a design methodology of Visual Analytics systems composed of nine stages (learn, winnow, cast, discover, design, implement, deploy, reflect, and write). Numerous other models emphasized interaction in Visual Analytics, such as Pike et al. [53] and Endert et al. [54]. Pike et al. propose to model Visual Analytics system by underlying interaction and more exploring the relationship between interaction and cognition. Endert et al. [54] introduced a semantic interaction model to interact with high-dimensional data in a two-dimensional (2D) view (cf. Fig. 7). This model presents how the user interactions in a spatial visualization can be integrated into the computation of a Visual Analytics system.

3.2.2. Visual analytics pipeline-based models for knowledge generation

Knowledge can be gained from both computational models and visual models. Among the first Visual Analytics models for knowledge generation, we quote that of Van Wijk [26] (cf. Fig. 8). It is composed of three main modules: (1) an input data component to be filtered, pre-processed and analysed, (2) a visualization component where this data is mapped to a visual representation (image), and (3) a user component who can visually explore such image by using his/her visual perception and background knowledge to go back to visualization module for further analysis. Wijk [26] proposed this model with a precise description using merely mathematical notation about interactive knowledge build-up.

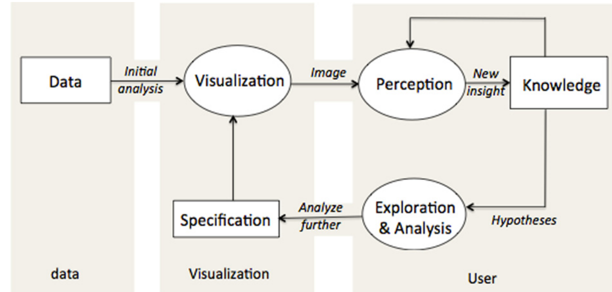

Fig. 8. Generic Visual Analytics model of [26].

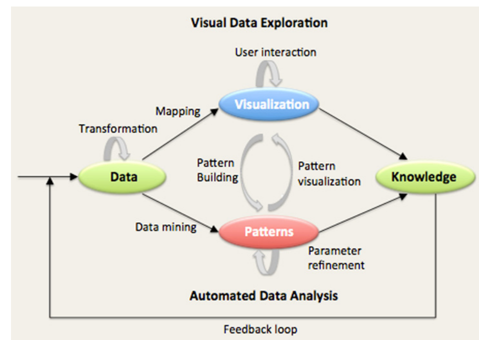

Fig. 9. Conventional Visual Analytics Model of Keim et al. [12].

Then Keim et al. [12] suggested the Visual Analytics process visible in Fig. 9. It is a conventional Visual Analytics pipeline-based model that is broadly adopted in the community. Fig. 9 depicts an abstract overview of the stages (in the form of ovals) and their transitions (in the form of arrows). The process starts with the data transformation (such as filtering and sampling) to generate graphical representations for further exploration purposes. Then, the analyst applies, separately, automatic analysis or visual methods. The automatic analysis includes, in general, data mining methods for discovering interesting patterns from original data. Once a pattern is generated, the human analyst has to interactively assess it and refine its parameters using visualization. Patterns visualization can be used for evaluating these generated data mining results. User interaction with the visual representation is essential to better understand these automatic results. Combining visual data exploration with automatic data analysis leads to a continuous refinement and verification of the automatic results in order to gain knowledge for decision-making.

Sacha et al. [7] proposed a refinement of the Keim et al.'s Visual Analytics model [12]. They introduced a knowledge generation model that involves computer and human parts (cf. Fig. 10). The left hand side of the model shows the Visual Analytics process of Keim et al. [12], while the right hand side shows the human knowledge generation process. The enriched part is in the reasoning process, composed of: (1) an exploration loop allowing selecting appropriate actions on the visualization based on the findings, (2) a verification loop to automatically formulate hypotheses about the data analysis from findings, and (3) a knowledge generation loop. These authors consider that automating the steps, from insights to knowledge and from knowledge to new hypotheses, is not possible.

There were two extensions of the Knowledge generation model for Visual Analytics of [7]: the first includes uncertainty spread and human trust building [55] and the second focuses specifically on interacting with Dimensionality Reduction methods for visualizing multidimensional data [56]. Another extension is of

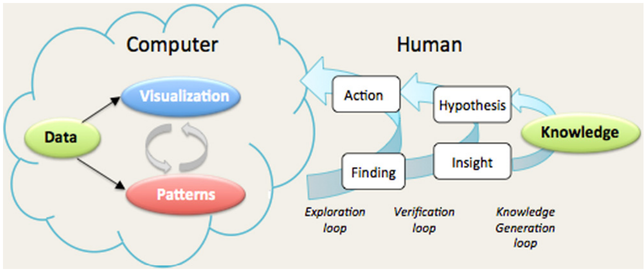

Fig. 10. Visual Analytics Model of Sacha et al. [7].

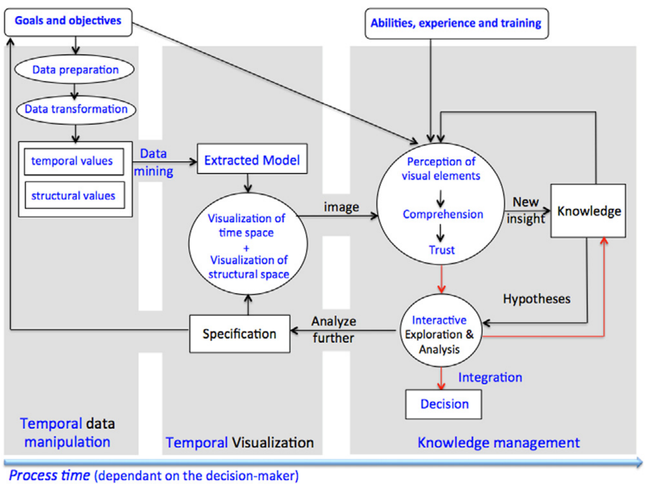

Fig. 11. Visual Analytics Model of Ltifi et al. [8].

Ltifi et al. [8] who proposed an adaptation of the Visual Analytics model of [26] considering specific data mining, temporal and cognitive considerations. They structured the adapted model into three stages: (1) temporal data manipulation, (2) temporal visualization, and (3) discovered knowledge management (cf. Fig. 11).

Considering the focus on knowledge, additional related work on knowledge-assisted Visual Analytics is by Federico et al. [57], and model building by Andrienko et al. [58]. The first one suggested an extension of the visualization model of van Wijk [26] by integrating the function and role of tacit and explicit knowledge in the analytical reasoning process. Such extension aims at designing a broad range of analytics systems [57]. The second one proposed a comprehensive conceptual framework considering the Visual Analytics process as a goal-oriented workflow and providing a model as a result [58].

3.2.3. Predictive Visual Analytics pipeline-based models

Predictive Visual Analytics models integrate Visual Analytics to assist in predictive analytics tasks (such as regression, classification, clustering, etc.). Such models cover the field of visualization methods and techniques that were used for supporting predictive analytics process steps. Understanding these models allows improving comprehensibility of the predictive modelling process and leading to desired results and trends (i.e. predictive patterns) including optimized predictions [59].

Lu et al. [22] structured a generic predictive Visual Analytics pipelines into five main steps:

(1) Data pre-processing that prepares data for analysis,
(2) Feature engineering that covers feature generation and feature selection techniques supported by visualization techniques (such as parallel coordinates [60], scatter plots [61], and matrix views [62]) to select appropriate features
(3) Modelling where machine learning and statistical techniques are applied to the data for predicting unknown data (such as tree view [63]), and
(4) Result exploration & model selection to explore the results and compare the performance among many patterns (if more than one pattern is discovered) where several visualization techniques can be used such as scatterplots [64], line charts [61], tree view [65], dimension reduction techniques [66], etc., and
(5) Validation to test pattern quality and where common visualization techniques can be used (examples are [67–69]) (cf. Fig. 12).

Several other predictive Visual Analytics models are available in literature. A specific model has been suggested by Lu et al. [70]. It highlights the important role played by visualization and adjustment for the knowledge generation from extracted data mining patterns (cf. Fig. 13(a)). This model proposed visualization during each knowledge discovery phases to improve comprehensibility and efficiency. The adjustment underlines the human knowledge involvement in all phases for predictive analysis. Sacha et al. [29] proposed interactive Visual Analytics/Machine Learning pipeline including several interaction options (cf. Fig. 14). It is an analyst' validation and refinement process, which allow him/her to interact with each step across a visual interface and to act as a mediator between the human and the machine learning components. Thereafter, they have improved such work by proposing ontology for Visual Analytics-assisted machine learning [71]. The work aims to describe and understand the Visual Analytics processes used in machine learning systems, as well as to identify the potential for introducing sophisticated Visual Analytics techniques into these systems.

A progressive learning model has been proposed by El-Assady et al. [72] (cf. Fig. 15). It includes an initial parameter space analysis and an iterative human-in-the-loop reinforcement learning process in which human annotators compare, evaluate, and optimize models using a Visual Analytics dashboard. Wagner et al. [73] proposed Knowledge-Assisted Visual Analytics Method for time-oriented data. They suggested taking advantage of explicit expert knowledge in the Visual Analytics process to make analytical reasoning more effective and efficient [73].

In this section, we have introduced the literature survey on Visual Analytics models. We have focused on articles and papers published between 2005 and 2019. Following, we will discuss these works and draw comparative summaries.

4. Summaries

We propose three complementary summaries that are described in succession. The first deals with the conceptual aspects of the Information Visualization and Visual Analytics pipeline-based models. The second presents comparative studies of the visualization and Visual Analytics pipeline-based models in literature. The third identifies some key themes for future visualization and Visual Analytics research.

4.1. Summary of the conceptual aspects

58 out of 112 publications propose Information Visualization or visual analytic pipeline-based models. The other 54 publications depict more about techniques, evaluation, performance

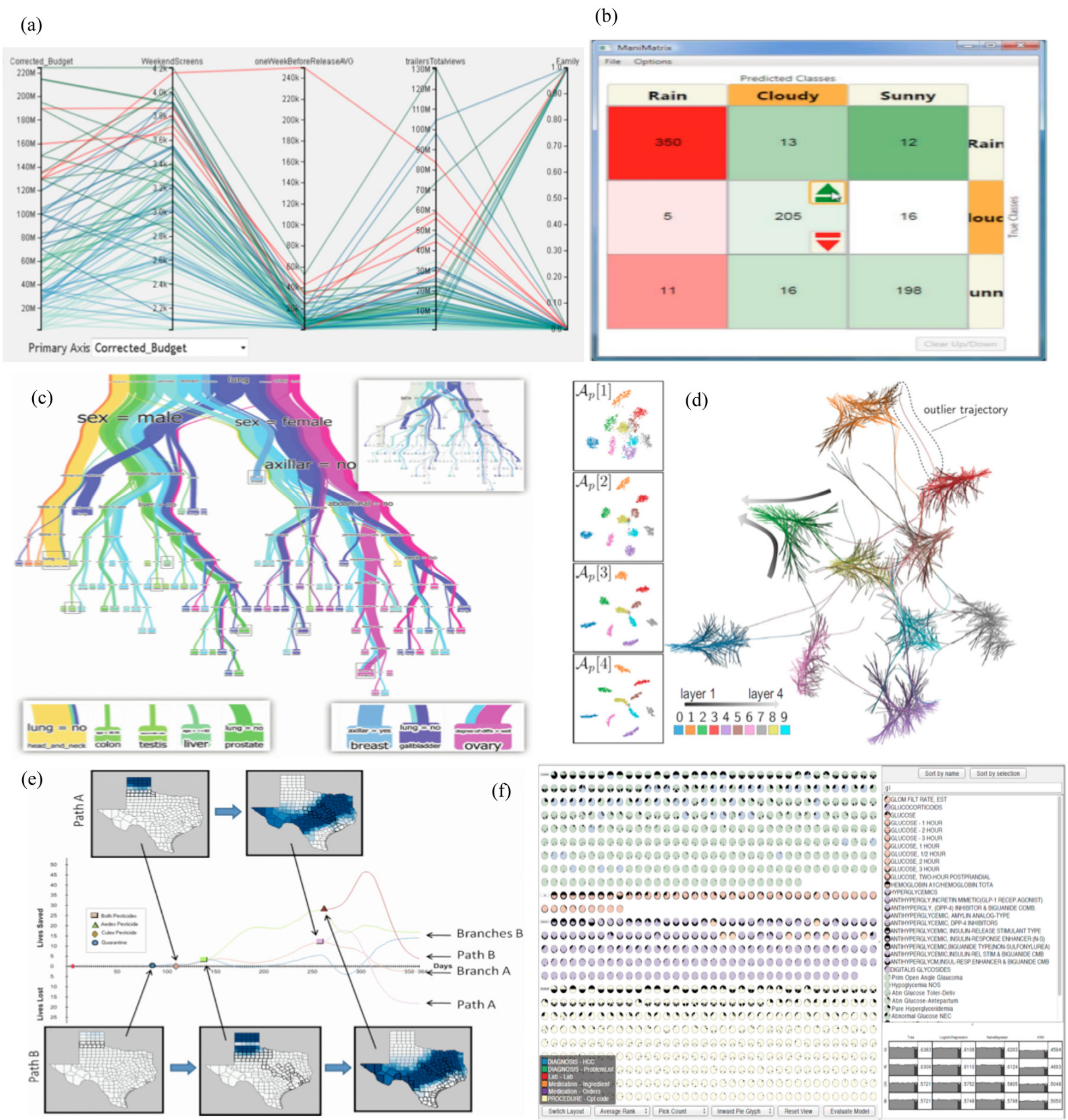

Fig. 12. (a) Example of Parallel Coordinates Plot for feature engineering. It presents the 5 most similar objects coloured in red. Users can update the display by acting on the drop-down menu [60], (b) Interactive Confusion Matrix, where users can directly manipulate it [62], (c) Tree-like interactive view of [63] for supporting a decision tree construction process, (d) Rauber et al.' tool [66] for visualizing the similarities between artificial neurons and show the inter-layer evolution of hidden layers after training, (e) Decision history tree view to depict prediction patterns' results in a branching time path form [65], and (f) A glyph that allows feature visualization and selection [69]. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

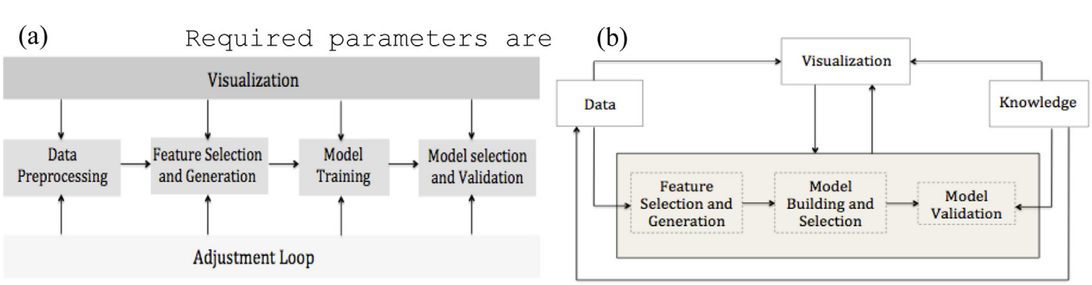

Fig. 13. (a) The specific predictive model of [70] and (b) The generalized version of [22].

improvement or challenges. Table 1 presents the publication distribution over the different examined conferences and journals.

To keep the study focused and manageable, we had to limit our literature analysis to a representative set of examples as visible

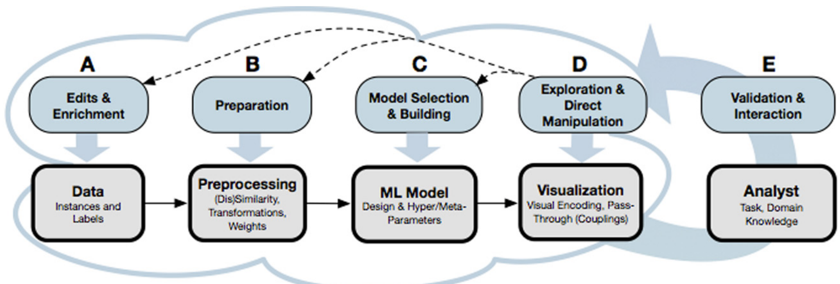

Fig. 14. Interactive Visual Analytics/Machine Learning of [29].

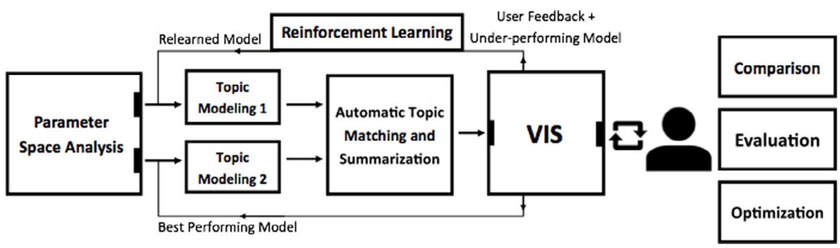

Fig. 15. The progressive learning model of [72].

Table 1
Publication distribution.

Conference/Journal references	IEEE TVCG	InfoVis	IEEE VAST	IEEE InfoVis	Others	Total
Information Visualization	3	1	0	2	8	14
Visual Analytics	21	3	9	2	9	44
Total	24	4	9	4	17	58

in Table 1. Reviewing papers from our four key literature sources (i.e. IEEE TVCG, InfoVis journal, IEEE VAST, IEEE InfoVis) has brought us to study other different interesting papers published by other sources (named "Others" in Table 1). We notice that the number of studied papers introducing Information Visualization models is lower than the number of papers introducing visual analysis models. This is because recent research in this area is moving more towards visual analysis that integrates Information Visualization by adding analysis.

In this survey, we focused on the visualization models literature. We aim to identify papers that include Information Visualization and Visual Analytics models. We find these essentially in the visualization community. We mainly aimed at presenting and discussing extensions and versions of visualization models. Yet, we are convinced that we have analysed a representative subset of the literature. Fig. 16 shows the number of surveyed papers per publication year.

The cited papers (cf. Fig. 16) are about general theory model, without limitation to a specific data type or application domain. This kind of general research work consolidates the foundation of the field of visualization. Due to its success, practitioners and researchers interested in these models tried to improve them taking into account specific theoretical or conceptual aspects. Based on the provided review in Sections 3 and 4, all visualization and Visual Analytics models transform data abstraction to visual abstraction. Visual abstractions will be then combined into views. Humans can thus explore these views to respond to their needs. It is a common conceptual process structured on the following steps:

(1) *Visual mapping*: converts data abstraction into visual abstraction, which refers to displayed elements on the screen and transmits information to human through the sense of sight. We talk about common visual channels used for encoding information (including position, size, shape, direction, etc.). A visual abstraction can have multiple visual channels simultaneously to represent important data attributes allowing users to get information easily and clearly. Several literatures give evaluations [74–77] and improvements [78–81] on existing visual mapping methods.

(2) *Views' Generation and Combination:* views allow users to obtain and explore relevant information and knowledge. Combining visual abstractions with specifications such as menus, captions, and legends help users to not feel confused and to explore more information on views. Generating a single complex view can be stressful for users to understand and recognize the information. Thus, multiple distinct views can be generated using "divide and conquer" method[2] [82]. However, these views can distract users' attention. For this, designers must coordinate views by considering and highlighting their relationships [83] [84]. Examples of variant of views are the overview + detail [84] and the small multiples [85] [86]. Users can update generated views based on their analytic tasks by applying replacements, replications or overlays techniques [87].

(3) *Interaction:* it is a Human–Computer Interaction aspect essential for visual exploration process. It brings vitality to visualization system. Without it, visual representations are simply static images [88]. There are two main classifications of interaction in visualization. We summarize them in Table 2.

[2] It consists of separating a big group of views into several smaller ones. It aids the user memory by reducing the amount of data they need to consider at one time.

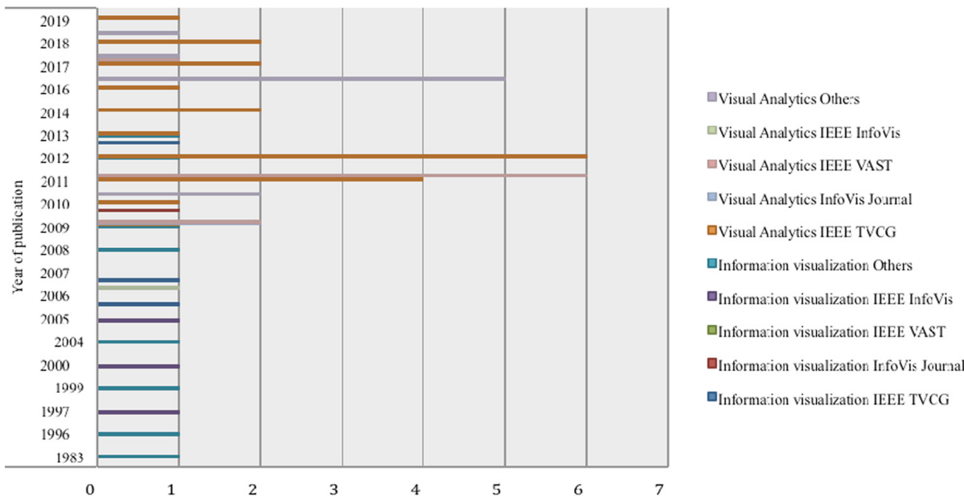

Fig. 16. Evolution of surveyed papers per publication year.

Table 2
Classifications of interaction in Information Visualization and Visual Analytics.

References	Brief description
[89]	They classified basic visualization interaction by output states. The categories include: (1) graphical operations affecting graphical representations, (2) data operations affecting data state, and (3) set operations affecting control state. They mentioned that graphical representations change when the two others are changed.
[88]	They classified interaction from the perspective of the users' demands. The classes include selecting, exploring, reconfiguring, encoding, abstracting, filtering and connecting. Lam [90] affirms that diverse interaction selections allow users acting freely, but can lead to confusion and need more time to spend on understanding the entire interactive visualization.

Concerning Visual Analytics models for knowledge generation, we can understand that the reasoning logics behind are the same: new knowledge is transformed from either inductive reasoning process or deductive reasoning process [91]. We describe them in Table 3.

4.2. Summary of the comparative studies

A variety of different visualization pipeline-based models including a large set of modelling guidelines and options do exist. They reflect a high-level system comprehension of human concepts. However, each model has its own strengths and weaknesses. Several comparative studies on Information Visualization and Visual Analytics models have been discussed in the literature. Table 4 presents a set of these comparative studies.

Based on the comparative results presented in Table 4, we can classify the studied models by the underlined key factors (cf. Table 5) (cf. Fig. 17).

Discussion of the Comparative Studies: Fig. 17 illustrates the analysis of the key factors highlighted by the surveyed pipeline-based models visible in Table 4. It clearly indicates the imbalanced distribution of Visual Analytics research in different key factors across pipeline-based models. In fact, the majority of models aims at improving the human interaction measures and mechanisms, which shows that Human–Computer Interaction still plays an important role in the field of visualization. This proves that the interaction and user interfaces will continue to be a key theme for future visualization research (cf. Section 4.3.1). Then, we find that Knowledge generation and machine learning (including prediction) factors take more attention than the others. Such observation emphasizes the partnership between humans and machines. Machines have the task of analysing and transforming data into information, which allows humans to turn that information into knowledge based on visual tasks.

Design and Cognition factors come in the third position to highlight the growing interest of the research community in cognitive modelling of visualization systems. The others take less attention by recent pipeline-based visualization models. This can be explained by the fact that the research community can turn to new future challenges in which the proposal of visualization pipelines can take place. These challenges can be related to emerging research fields such as big data and Internet of things. In fact, during our literature review, we found several recent visual analysis works that propose tools, techniques and methods but not pipelines in the fields of big data [95–100], Internet of Things [101,102], provenance [103–105] and intelligent evaluation [106–110]. It is in this sense that we believe that these areas represent challenges for future work.

From our survey analysis of literature, we have found that most recent works proposed pipeline-based Visual Analytics models (cf. Table 4). Such an evolution is due to the fact that recent computational advances have made it possible to perform effective automatic analysis (using data mining and machine learning algorithms) of data (in particular complex data) in order to identify valuable insights and make informed decisions. Although such analysis has proven its usefulness in several practical applications, it still faces significant challenges such as increasing data dimensions, data heterogeneity and real-time aspect. To cope these challenges, Visual Analytics has been developed in recent years through a right combination of automatic analysis with interactive visualizations (c.f. Section 4). For this, future challenges are related to Visual Analytics more than Information Visualization.

4.3. Summary of challenges in Visual Analytics field

From our survey analysis, we have identified a set of key themes for future Visual Analytics research. These challenges and

Table 3
Reasoning in Visual Analytics knowledge generation models.

Reasoning process	Brief description
Inductive Reasoning	The user (i.e. analyst) builds ideas from observations. Nevertheless, users with diverse prior knowledge would have different ideas based on the same observations [92]. Practices of inductive reasoning are shown in the exploration to knowledge process in the Generic Visual Analytics model of [26], the hypotheses' generation in the human cognition model of [40], and the exploration-verification-knowledge path in the knowledge generation model of [7].
Deductive Reasoning	The user searches for evidences that confirm or negate the initialized hypotheses. Beginning with the same hypotheses allows users reaching the same conclusion. Practices of deductive reasoning are shown in the knowledge to exploration process in the Generic Visual Analytics model of [26], the hypotheses' analysis process in the human cognition model of [40] and the knowledge-verification-exploration path in the knowledge generation model of [7].
Inductive and Deductive Reasoning	Inductive and deductive reasoning processes can be performed alternately, where generated knowledge in the inductive reasoning is applied to the deductive reasoning. Alternatively, the knowledge verification can lead to new hypotheses, which explains the knowledge accumulation [26] [40] [7].

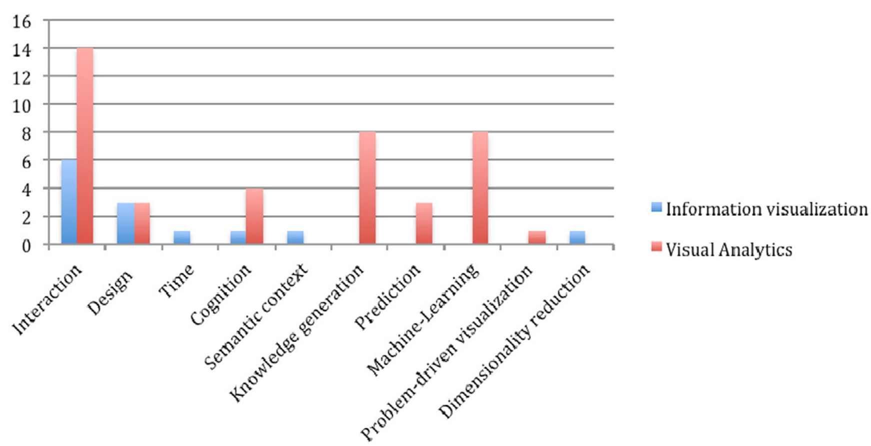

Fig. 17. Number of surveyed pipeline-based models per key factors.

recommendations for future research aim to give an interesting set of ideas for the research community concerning basic and emerging research fields in developing Visual Analytics pipelines. We summarize them in six main key themes categorized into two types: (1) basic fields and (2) emergent fields:

4.3.1. Visual Analytics challenges related to basic fields

These challenges are related to interaction, knowledge generation, machine learning and cognitive learning. We have identified them based on our survey analysis:

Interaction and user Interface: as presented in our survey, interaction and user interfaces play an increasingly prominent role in Information Visualization and Visual Analytics [56,111] (cf. Fig. 17). While data volume is continuously and rapidly growing, human cognition needs to be more and more taken into consideration in Visual Analytics pipelines for the design of appropriate human–computer interfaces. Research in this area opens continuously interesting challenges such as user-centred analysis, perception-driven design, multimodal interaction techniques for Visual Analytics and adaptive user interfaces [112].

Knowledge generation and machine learning for prediction: intelligent Machine Learning algorithms, allowing systems to both learn and reason, intended for Visual Analytics to generate knowledge, have been the subject of interesting research in recent years (cf. Fig. 17), both in theory and in practice and will continue to be a challenge for the research community [29,71]. By integrating machine learning, the Visual Analytics solutions will see a strengthening of their innovation and prediction capabilities through more and more reliable knowledge models improving decision-making.

Cognitive modelling: the modelling of cognitive phenomena raises a series of challenges, some of which are general, but others are specific to Visual Analytics systems. Recent work has taken impressive steps towards dealing with the uncertainty and complexity generated by the Visual Analytics process [8,40,55]. However, cognitive modelling is still challenged in such field by introducing cognitive design for specifying decision-maker behaviours more naturally and intuitively. Future work can be planed to reconstruct cognitive entities from an experimental approach producing a set of laws involving directly observable visual phenomena (situations and) behaviours).

4.3.2. Visual Analytics challenges related to emergent fields

These challenges are related to big data, Internet of Things, provenance and intelligent evaluation.

Big data: Visual Analytics for big data is challenged in visualization literature. It is identified as a promising method for big data sensemaking [95–97]. Qin et al. [98] claimed that visualization is a basic function for analysts because of its ability to make data to speak to the user at an intuitive level. However, the intersection of both fields (Visual Analytics and big data) is not without difficulties. Permitting sensemaking of big data is a main challenge of Visual Analytics from a human and technical perspective [99]. Future challenges related to the big data volume, velocity and variety are discussed by [100]. These challenges are related to system scalability, available Information Space, visual representations, interpretability, multidimensional data, network or relational data, workflow and real-time analysis. We think that these challenges continue to gain interest in the Visual Analytics pipelines research.

Table 4
Comparative Studies on surveyed pipeline-based models.

	Reference	Summary	Underlined factor(s)
Information Visualization models	[25] and [31]	These models focus on the perceptual properties and visual features.	Interaction
	[5]	The guidelines of [5] can be useful for designing and developing visualization tools. Many practitioners find Mantra useful in different design scenarios. However, [93] argue that guidelines alone are not sufficient. Based on this idea, others such as Chi [6] and Daassi and his colleagues [32] have studied conceptual models.	Design
	[1]	It emphasizes on human interactions in the visualization process.	Interaction
	[6]	It emphasizes the purposes of the visualization and the role of human. It focuses on the visualization stages rather than on early data treatment stages [37].	Interaction
	[11]	It highlights the conceptual distinction between partially developed visualizations and ready-to-render visualizations [37]. It presents two main components of the visual interface: (1) the visualization that shows the information, and (2) the Graphical User Interface that consists of graphical controls or widgets (sliders, knobs, etc.) [23]. According to [23], this model is an effective user interaction framework for Information Visualization.	Design Interaction
	[32]	It considers the temporal and structural dimensions of data to be visualized.	Time
	[35]	It stresses the importance of user' behaviour for recommending appropriate visualizations. However, it does not consider the other aspects of the activity context.	Cognition
	[36]	It underlines the semantic context for the visualization design.	Design Semantic context
	[37]	It combines conceptual interaction and visual notation to model visualization systems. It is specific for beyond-desktop visualization systems.	Interaction
Visual Analytics models	Generic Visual Analytics model of [26] and its adapted version [8]	It presents a basic knowledge generation model. It mathematically expresses gained knowledge by quantifying its amount basing on the perceived image, the current user knowledge, and the user perception. However, it does not take into account the idea of the cognitive reasoning. [8] adapted it by considering the comprehension and trust processes in order to make more effective and human-perceptive dynamic visualizations.	Knowledge generation Cognition
	Conventional Visual Analytics Model of [12]	The analysis process is characterized by interactions between data, visualizations, patterns, and users for knowledge discovery. It takes into account Machine Learning interaction. It aimed at model/pattern building and parameter refinement.	Interaction Machine Learning
	[43]	It highlights the importance of human interaction mechanisms and measures in the Visual Analytics process.	Interaction
	The nested model of [51]	It suggests a design order. The temporal sequence is not always strictly executed and the improvements should not be limited in the current stage.	Design
	[40], and [55]	These models underline human cognitive aspects (comprehension, trust, sub-processes of human thinking, etc.) in order to facilitate the flow of human reasoning in Visual Analytics process.	Cognition
	[41] and [42]	They underline the importance of providing fitting mechanisms to analysts in order to reflect their knowledge concerning data across interactions that directly change computational results [23].	Interaction
	[39]	It provides quality metrics based automation. It obtains information from the stages of the visualization pipeline [1] and influences its processes through the metrics it calculates. This model underlines that user is always in control. It is an interesting work that helps in the visual exploration of meaningful patterns [7].	Interaction
	[53] and [54]	They give an interesting description of the human data analysis process. They depicted the interaction process between a human analyst and automated analysis techniques. The work of [54] shows how the user interactions in a spatial visualization can be incorporated into the computation of a visual analytic system [23].	Interaction
	[30]	It is a comprehensive modelling that distinguishes between goals, objectives, operators, and operands of Visual Analytics tasks [94].	Design
	[52]	It is a reflection on Visual design considerations. It highlights the importance of Human–Computer interaction, user-centred design and problem-driven visualization for designing Visual Analytics systems.	Design Interaction Problem-driven visualization
	[12], [7], [8], [55], [70], [57] and [58]	These works propose models that: – Encompass the process of human knowledge generation. – Clarify the role of humans in knowledge generation, and highlight the importance of supporting tighter integration of human and machine. – Depict the interplay between Machine learning and visualization approaches. However, they were mostly designed from an interactive visualization perspective characterizing the role of the "human in the loop".	Knowledge generation Machine Learning Interaction
	[55]	It integrates uncertainty propagation and trust building within the knowledge generation process. It aims to produce a more complete knowledge generation [23].	Knowledge generation Cognition

(*continued on next page*)

Table 4 (*continued*).

Reference	Summary	Underlined factor(s)
[29] and [71]	It insists on a tighter connection between machine learning and visualization researches for more effective data analysis.	Machine Learning
[56]	It emphases the specific area of dimensionality reduction and how its techniques integrate with interactive visualization in Visual Analytics systems [23].	Dimensionality reduction Interaction
[22]	It presents a standard pipeline that generalize the structure of predictive Visual Analytics pipeline-based models	Prediction
[72]	It focuses on the importance on parameter analysis and iterative human-in-the-loop reinforcement learning process.	Knowledge generation Prediction Machine Learning
[73]	It underlines the importance of explicit expert knowledge for making analytical reasoning.	Prediction Knowledge generation

Table 5
Distribution of surveyed pipeline-based models per key factors.

	Interaction	Design	*Time*	Cognition	Semantic context	Knowledge generation	Prediction	Machine Learning	Problem-driven visualization	Dimensionality reduction
Information Visualization	[25], [31], [1], [6], [11], [37]	[5], [11], [36]	[32]	[35]	[36]					
Visual Analytics	[12], [43], [41], [42], [39], [53], [54], [52], [7], [8], [55], [70], [56]	[51], [30], [52]		[8], [40], [54]		[26], [12], [7], [8], [55], [70], [72], [73], [57], [58]	[22], [72], [73]	[12], [7], [8], [55], [70], [29], [71] [72]	[52]	[56]

Internet of Things (IoT): Nowadays, practitioners are interested by smart technologies and connected sensors devices. As these devices are connected, they become part of the Internet of Things [101]. Networked together, they generate an unprecedented amount of data. The challenge is to understand what that data means and to draw insight from it. Visual Analytics can be considered as a promising solution to do that. It is the visual analysis made of data from connected objects. Such solution offers data acquisition, telecommunication, visualization, analysis and decision-making services (machine learning, data science, visualization, etc.), temporal reactivity (flow analysis) and transmission of actions [102]. This is what we propose to call Visual Analytics of Things (VAoT).

Provenance: it consists of keeping track of Visual Analytics process to provide the 'context' of data, analysis, reasoning and decision. Information provenance allows tracing the data source for automatic update, redo/undo user interactions, or avoiding repeated analysis processes. In addition, it can facilitate the evaluation of the extracted knowledge. Visualization research community agrees on the prominence of supporting provenance, where many tools and systems have been developed to assist analysts to record data, computational workflows and reasoning processes [103–105]. It is important that research into provenance topic continue to propose a set of pipeline-based models to support provenance through a wide range of fields and for different application purposes.

Intelligent evaluation: it is essential to evaluate the effectiveness of Visual Analytics systems [106,107]. Visualization practitioners apply different approaches (such as case studies, expert review, questionnaires, etc.) to assess the systems' usability [108]. Each approach has its own strengths and weaknesses. Using an interesting user study can produce valuable user feedback to identify problems with the systems. Nevertheless, such study might be insufficient to provide high-level insights [109]. Because of the multiple data analysis and visualization components, Visual Analytics systems are, in general, complex, which needs more effective and intelligent evaluation approaches. Intelligent evaluation must integrate intelligent techniques (such as deep learning techniques) to analyse and learn from evaluation results provided by formal user studies, which poses a challenge expected to gain more interest in the field of visualization [110].

5. Related work: Positioning compared with other surveys

At the end of this paper, we investigate what this study brings compared with other recent surveys examining visualization pipeline-based models.

Table 6 classifies these surveys based on the objectives set a priori. In fact, the title of this paper sums up the objectives we set ourselves: performed a survey on Information Visualization and Visual Analytics pipeline-based models, their conceptual aspects, comparative studies and challenges.

On the basis of the classification shown in Table 6, we can identify that our comprehensive survey deals with all the objectives set a priori; which is not the case for the other related surveys (i.e. [4,15,22–24]). It sheds a light on specific interesting conceptual aspects for guiding the designing process of intelligent Decision Support Systems. It aims to be valuable to guide future visualization research studies in various application domains.

6. Conclusion

Visualization is considered as a powerful tool to assist users (decision-makers) in understanding data and recognizing patterns, especially in situations where human reasoning is indispensable. This field is used in many research and engineering areas. It allows transforming a set of raw and often complex data into one or more visual representation(s) in order to facilitate the understanding of what data means.

The field of visualization started with Information Visualization and evolved into visual analysis. It proposed techniques, evaluation performance improvement and pipeline-based models. In this survey we have presented a comprehensive summarization of Information Visualization and Visual Analytics pipeline-based models. We reviewed them while highlighting their versions and extensions. These models can be used as a guideline for structuring and developing visualization systems in real life. Additionally, we have presented and discussed their conceptual

Table 6
Positioning compared with other recent surveys.

Reference	Studying pipelines	Information Visualization	Visual analytics	Conceptual aspects	Comparative studies	Future directions
(Moreland 2013) [15]	Yes	Yes	No	Partially (related to the visualization process)	No	No
(Wang et al. 2016) [4]	Yes	No	Yes	Yes	No	Yes
(Lu et al. 2017) [22]	Yes	No	Yes	Yes	No	Yes
(Endert et al. 2018) [23]	Partially	No	Yes	Partially (related to machine learning)	Yes	Yes
(Cui 2019) [24]	Partially	No	Yes	Partially (related to the Visual Analytics process)	Yes	Yes
Our survey	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**

aspects and comparative studies. We have identified common conceptual process structured on three steps (i.e. visual mapping, views' generation and combination and Interaction). We have also presented the reasoning logics behind Visual Analytics pipeline-based models for knowledge generation (i.e. inductive reasoning process or deductive reasoning process). We have performed comparative studies on information and Visual Analytics models in order to enumerate the underlined factors (such as interaction, knowledge generation, design, etc.).

Based on this survey, we have presented a set of main future challenges that can offer a rich set of ideas for the research community. Finally, we have investigated what this study brings compared with other recent surveys presenting visualization pipeline-based models. We have concluded that our comprehensive survey deals with all the objectives set a priori. We are attentive to new proposals and developments of pipeline-based models that could be introduced in the coming years, both generic and specific to areas of current application (including: Big data, IoT, provenance and intelligent evaluation).

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

The research leading to these results has received funding from the Ministry of Higher Education and Scientific Research of Tunisia under the grant agreement number LR11ES48.

References

[1] S.K. Card, J.D. Mackinlay, B. Shneiderman, Readings in Information Visualization: Using Vision To Think, Morgan Kaufmann Publishers Inc, 1999, pp. 1–34, (Chapter 1).
[2] D.A. Keim, J. Kohlhammer, G.P. Ellis, F. Mansmann, Mastering the Information Age - Solving Problems with Visual Analytics, Eurographics Association, 2010.
[3] M. Lu, S. Chen, C. Lai, L. Lin, X. Yuan, Frontier of information visualization and visual analytics in 2016, J. Vis. 20 (4) (2017) 667–686.
[4] X.M. Wang, T.Y. Zhang, Y.X. Ma, et al., A survey of visual analytic pipelines, J. Comput. Sci. Tech. 31 (4) (2016) 787–804.
[5] B. Shneiderman, The eyes have it: a task by data type taxonomy for information visualizations, in: Proceedings of the IEEE Symposium on Visual Languages, Boulder, USA, 1996, pp. 336–343.
[6] Ed-H. Chi, A taxonomy of visualization techniques using the data state reference model, in: IEEE Symposium on Information Visualization, InfoVis 2000, IEEE Computer Society, 2000, pp. 69–75.
[7] D. Sacha, A. Stoffel, F. Stoffel, B.-C. Kwon, G. Ellis, D.-A. Keim, Knowledge generation model for visual analytics, IEEE Trans. Vis. Comput. Graphics 20 (12) (2014) 1604–1613.
[8] H. Ltifi, E. Ben Mohamed, M. Ben Ayed, Interactive visual KDD based temporal decision support system, Inf. Vis. 15 (1) (2016) 31–50.
[9] C. Chen, Top 10 unsolved information visualization problems, IEEE Comput. Graph. Appl. 25 (4) (2005) 12–16.
[10] J.J. Thomas, Cook K.A. (Eds.), Illuminating the Path: The R & D Agenda for Visual Analytics, National Visualization and Analytics Center, 2005.
[11] J. Heer, M. Agrawala, Software design patterns for information visualization, IEEE Trans. Vis. Comput. Graphics 12 (5) (2006) 853–860.
[12] D. Keim, G. Andrienko, J.D. Fekete, C. Görg, J. Kohlhammer, G. Melançon, Visual analytics: Definition, process, and challenges, in: Andreas Kerren, John T. Stasko, Jean-Daniel Fekete, Chris North (Eds.), Information Visualization - Human-Centered Issues and Perspectives, in: LNCS, Springer, 2008, pp. pp.154–175.
[13] G. Geisler, Making information more accessible: a survey of information visualization applications and techniques, 1998, URL: https://www.cs.nmt.edu/viz/papers/GeislerMarrero.pdf. (Accessed July 2018).
[14] E. Zudilova-Seinstra, T. Adriaansen, R.v. Liere, Trends in Interactive Visualization: State-of-the-Art Survey, first ed., Springer Publishing Company, 2008, Incorporated.
[15] K. Morel, A survey of visualization pipelines, IEEE Trans. Vis. Comput. Graphics 19 (3) (2013) 367–378.
[16] S. Liu, W. Cui, Y. Wu, M. Liu, A survey on information visualization: recent advances and challenges, Vis. Comput. 30 (12) (2014) 1373–1393.
[17] N. Andrienko, G. Andrienko, Visual analytics of movement: An overview of methods, tools and procedures, Inf. Vis. 12 (1) (2012) 3–24.
[18] V.L. West, D. Borland, W.E. Hammond, Innovative information visualization of electronic health record data: A systematic review, J. Amer. Med. Inform. Assoc. 22 (2) (2014) 330–339.
[19] T. Von Landesberger, A. Kuijper, T. Schreck, J. Kohlhammer, J.J. van Wijk, J.-D. Fekete, D.W. Fellner, Visual analysis of large graphs: state-of-the-art and future research challenges, Comput. Graph. Forum 30 (2011) 1719–1749.
[20] P. Caserta, O. Zendra, Visualization of the static aspects of software: a survey, IEEE Trans. Vis. Comput. Graphics 17 (7) (2011) 913–933.
[21] L. Grammel, C. Bennett, M. Tory, M.-A. Storey, A survey of visualization construction user interfaces, in: M. Hlawitschka, T. Weinkauf (Eds.), EuroVis - Short Papers, The Eurographics Association, 2013.
[22] Y. Lu, R. Garcia, B. Hansen, M. Gleicher, R. Maciejewski, The state-of-the-art in predictive visual analytics, in: EuroVis 2017, vol. 36, no. 3, 2017.
[23] A. Endert, W. Ribarsky, C. Turkay, B.L.W. Wong, I. Nabney, I. Díaz Blanco, F. Rossi, The state of the art in integrating machine learning into visual analytics, Comput. Graph. Forum 36 (8) (2018) 458–486.
[24] W. Cui, Visual analytics: A comprehensive overview, IEEE Access 7 (2019) 81555–81573.
[25] J. Bertin, Semiology of Graphics, University of Wisconsin Press, 1983.
[26] V.J.J. Wijk, The value of visualization, in: Proceedings of IEEE Visualization, vol. 7, 2005, pp. 9–86.
[27] D.A. Keim, F. Mansmann, J. Schneidewind, H. Ziegler, Challenges in visual data analysis, in: Proc. the IEEE Conference on Information Visualization, 2006, pp. 9–16.
[28] S.J. Miah, G.M. McGrath, D. Kerr, Design science research for decision support systems development: recent publication trends in the premier IS journals, Aust. J. Inf. Syst. 20 (2016) 1–14.
[29] D. Sacha, M. Sedlmair, L. Zhang, J.A. Lee, D. Weiskopf, S.C. North, D.A. Keim, Human- centered machine learning through interactive visualization: Review and open challenges, in: 24th European Symposium on Artificial Neural Networks, ESANN, 2016, pp. 641–646.
[30] R.E. Roth, Cartographic interaction primitives: framework and synthesis, Cartogr. J. 49 (4) (2012) 376–395.
[31] S.K. Card, J.D. Mackinlay, The structure of the information visualization design space, in: Proceedings of the 1997 IEEE Symposium on Information Visualization, Washington, DC, USA, 1997, pp. 92–99.

[32] C. Daassi, L. Nigay, M.C. Fauvet, Visualization process of temporal data, in: Proceedings of DEXA Conference, Zaragoza, Spain, 30 August– 3 September 2004, in: LNCS, vol. 3180, Springer, Amsterdam, 2004, pp. 914–924.

[33] J.D. Mackinlay, P. Hanrahan, C. Stolte, Show me: automatic presentation for visual analysis, IEEE Trans. Vis. Comput. Graphics 13 (6) (2007) 1137–1144.

[34] G. Wills, L. Wilkinson, AutoVis: Automatic visualization, Inf. Vis. 9 (1) (2010) 47–69.

[35] D. Gotz, Z. Wen, Behavior-driven visualization recommendation, in: Proceedings of UI'09, New York, USA, 2009, pp. 315–324.

[36] M. Hipp, B. Michelberger, B. Mutschler, M. Reichert, A framework for the intelligent delivery and user-adequate visualization of process information, in: Proceedings of the 28th Symposium On Applied Computing, SAC'13, Coimbra, Portugal, 2013, pp. 1383–1390.

[37] Y. Jansen, P. Dragicevic, An interaction model for visualizations beyond the desktop, IEEE Trans. Vis. Comput. Graphics 19 (12) (2013) 2396–2405.

[38] M. Beaudouin-Lafon, Instrumental interaction: an interaction model for designing post-wimp user interfaces, in: Proc. CHI 2000, ACM, 2000, pp. 446–453.

[39] E. Bertini, A. Tatu, D.A. Keim, Quality metrics in high-dimensional data visualization: An overview and systematization, IEEE Trans. Vis. Comput. Graphics 17 (12) (2011) 2203–2212.

[40] T.M. Green, W. Ribarsky, B. Fisher, Building and applying a human cognition model for visual analytics, Inf. Vis. 8 (1) (2009) 1–13.

[41] J. Choo, S. Bohn, H. Park, Two-stage framework for visualization of clustered high dimensional data, in: IEEE Conf. on Visual Analytics in Science and Technology, VAST, 2009, pp. 67–74.

[42] A. Endert, C. Han, D. Maiti, L. House, S. Leman, C. North, Observation-level interaction with statistical models for visual analytics, in: IEEE Conf. on Visual Analytics in Science and Technology, VAST, 2011, pp. 121–130.

[43] R.J. Crouser, R. Chang, An affordance-based framework for human computation and human–computer collaboration, IEEE Trans. Vis. Comput. Graphics 18 (12) (2012) 2859–2868.

[44] A. Dasgupta, R. Kosara, Adaptive privacy-preserving visualization using parallel coordinates, IEEE Trans. Vis. Comput. Graphics 17 (12) (2011) 2241–2248.

[45] J. Hullman, E. Adar, P. Shah, Benefitting infovis with visual difficulties, IEEE Trans. Vis. Comput. Graphics 17 (12) (2011) 2213–2222.

[46] H. Wickham, H. Hofmann, Product plots, IEEE Trans. Vis. Comput. Graphics 17 (12) (2011) 2223–2230.

[47] S. Kandel, A. Paepcke, J.M. Hellerstein, J. Heer, Enterprise data analysis and visualization: An interview study, IEEE Trans. Vis. Comput. Graphics 18 (12) (2012) 2917–2926.

[48] Y. Wu, F. Wei, S. Liu, et al., Opinionseer: Interactive visualization of hotel customer feedback, IEEE Trans. Vis. Comput. Graphics 16 (6) (2010) 1109–1118.

[49] Y. Wu, G. Yuan, K.L. Ma, Visualizing flow of uncertainty through analytical processes, IEEE Trans. Vis. Comput. Graphics 18 (12) (2012) 2526–2535.

[50] Y. Wu, X. Liu, S. Liu, K.L. Ma, Visizer: A visualization resizing framework, IEEE Trans. Vis. Comput. Graphics 19 (2) (2013) 278–290.

[51] T. Munzner, A nested model for visualization design and validation, IEEE Trans. Vis. Comput. Graphics 15 (6) (2009) 921–928.

[52] M. Sedlmair, M. Meyer, T. Munzner, Design study methodology: Reflections from the trenches and the stacks, IEEE Trans. Vis. Comput. Graphics 18 (12) (2012) 2431–2440.

[53] W.A. Pike, J.T. Stasko, R. Chang, T.A. O'Connell, The science of interaction, Inf. Vis. 8 (4) (2009) 263–27.

[54] A. Endert, P. Fiaux, C. North, Semantic interaction for sense-making: Inferring analytical reasoning for model steering, IEEE Trans. Vis. Comput. Graphics 18 (12) (2012) 2879–2888.

[55] D. Sacha, H. Senaratne, B.C. Kwon, G. Ellis, D.A. Keim, The role of uncertainty, awareness, and trust in visual analytics, IEEE Trans. Vis. Comput. Graphics 22 (1) (2016) 240–249.

[56] D. Sacha, L. Zhang, M. Sedlmair, J.A. Lee, J. Peltonen, D. Weiskopf, S.C. North, D.A. Keim, Visual interaction with dimensionality reduction: A structured literature analysis, IEEE Trans. Vis. Comput. Graphics 23 (1) (2017) 241–250.

[57] P. Federico, M. Wagner, A. Rind, A. Amor-Amorós, S. Miksch, W. Aigner, The role of explicit knowledge: A conceptual model of knowledge-assisted visual analytics, in: 2017 IEEE Conference on Visual Analytics Science and Technology, VAST, IEEE, 2017.

[58] N. Andrienko, T. Lammarsch, G. Andrienko, G. Fuchs, D. Keim, S. Miksch, A. Rind, Viewing visual analytics as model building, Comput. Graph. Forum (2018) Published online before print, https://doi.org/10.1111/cgf.13324.

[59] M. Gleicher, A framework for considering comprehensibility in modeling, Big Data 4 (2) (2016) 75–88.

[60] Y. Lu, R. Kruger, D. Thom, F. Wang, S. Koch, T. Ertl, R. Maciejewski, Integrating predictive analytics and social media, in: IEEE Symposium on Visual Analytics Science and Technology, IEEE, 2014, pp. 193–202.

[61] S. Bremm, T. von Landesberger, M. Hess, T. Schreck, P. Weil, K. Hamacher, Interactive visual comparison of multiple trees, in: IEEE Conference on Visual Analytics Science and Technology, IEEE, 2011, pp. 31–40.

[62] A. Kapoor, B. Lee, D. Tan, E. Horvitz, Interactive optimization for steering machine classification, in: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, ACM, 2010, pp. 1343–1352.

[63] S. Van Den Elzen, J.J. Wijk van, Baobabview: Interactive construction and analysis of decision trees, in: IEEE Conference on Visual Analytics Science and Technology, VAST, pp. 151–160.

[64] S.K. Badam, J. Zhao, S. Sen, N. Elmqvist, D. Ebert, TimeFork: Interactive prediction of time series, in: Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems, ACM, 2016, pp. 5409–5420.

[65] S. Afzal, R. Maciejewski, D.S. Ebert, Visual analytics decision support environment for epidemic modeling and response evaluation, in: IEEE Conference on Visual Analytics Science and Technology, IEEE, 2011, pp. 191–200.

[66] P.E. Rauber, S.G. Fadel, A.X. Falcao, A.C. Telea, Visualizing the hidden activity of artificial neural networks, IEEE Trans. Vis. Comput. Graphics 23 (1) (2017) 101–110.

[67] G. Andrienko, N. Andrienko, S. Rinzivillo, M. Nanni, D. Pedreschi, F. Giannotti, Interactive visual clustering of large collections of trajectories, in: IEEE Symposium on Visual Analytics Science and Technology, 2009, pp. 3–10.

[68] F. Heimerl, S. Koch, H. Bosch, T. Ertl, Visual classifier training for text document retrieval, IEEE Trans. Vis. Comput. Graphics 18 (12) (2012) 2839–2848.

[69] J. Krause, A. Perer, E. Bertini, INFUSE: Interactive feature selection for predictive modeling of high dimensional data, IEEE Trans. Vis. Comput. Graphics 20 (12) (2014) 1614–1623.

[70] J. Lu, Y. Ma, W. Chen, et al., Recent progress and trends in predictive visual analytics, Front. Comput. Sci. (2016) 1–16.

[71] D. Sacha, M. Kraus, D. Keim, M. Chen, VIS4ML: An ontology for visual analytics assisted machine learning, IEEE Trans. Vis. Comput. Graphics 25 (1) (2019) 385–395.

[72] M. El-Assady, R. Sevastjanova, F. Sperrle, D.A. Keim, C. Collins, Progressive learning of topic modeling parameters: A visual analytics framework, IEEE Trans. Vis. Comput. Graphics 24 (1) (2018) 382–391.

[73] M. Wagner, A. Slijepcevic, M. Rind, D. Zeppelzauer, W. Aigner, KAVAGait: Knowledge-assisted visual analytics for clinical gait analysis, IEEE Trans. Vis. Comput. Graph. (TVCG) 25 (3) (2018) 1528–1542, http://dx.doi.org/10.1109/TVCG.2017.2785271.

[74] R. Jianu, A. Rusu, Y. Hu, D. Taggart, How to display group information on node-link diagrams: An evaluation, IEEE Trans. Vis. Comput. Graphics 20 (11) (2014) 1530–1541.

[75] Y. Albo, J. Lanir, P. Bak, S. Rafaeli, Off the radar: Comparative evaluation of radial visualization solutions for composite indicators, IEEE Trans. Vis. Comput. Graphics 22 (1) (2016) 569–578.

[76] T. Gschwandtner, M. Bogl, P. Federico, S. Miksch, Visual encodings of temporal uncertainty: A comparative user study, IEEE Trans. Vis. Comput. Graphics 22 (1) (2016) 539–548.

[77] J. Johansson, C. Forsell, Evaluation of parallel coordinates: Overview, categorization and guidelines for future research, IEEE Trans. Vis. Comput. Graphics 22 (1) (2016) 579–588.

[78] J.H. Lee, K.T. McDonnell, A. Zelenyuk, D. Imre, K. Mueller, A structure-based distance metric for high-dimensional space exploration with multidimensional scaling, IEEE Trans. Vis. Comput. Graphics 20 (3) (2014) 351–364.

[79] S. Kieffer, T. Dwyer, K. Marriott, M. Hola Wybrow, Human-like orthogonal network layout, IEEE Trans. Vis. Comput. Graphics 22 (1) (2016) 349–358.

[80] R.G. Raidou, M. Eisemann, M. Breeuwer, E. Eisemann, A. Vilanova, Orientation-enhanced parallel coordinate plots, IEEE Trans. Vis. Comput. Graphics 22 (1) (2016) 589–598.

[81] V. Yoghourdjian, T. Dwyer, G. Gange, et al., High-quality ultra-compact grid layout of grouped networks, IEEE Trans. Vis. Comput. Graphics 22 (1) (2016) 339–348.

[82] M.Q. Wang Baldonado, A. Woodruff, A. Kuchinsky, Guidelines for using multiple views in information visualization, in: Proc. of the Working Conference on Advanced Visual Interfaces, 2000, pp. 110–119.

[83] J.C. Roberts, State of the art: Coordinated & multiple views in exploratory visualization, in: Proc. the 5th International Conference on Coordinated and Multiple Views in Exploratory Visualization, 2007, pp. 61–71.

[84] C. Papadopoulos, I. Gutenko, A. Kaufman, VEEVVIE: Visual explorer for empirical visualization, VR and interaction experiments, IEEE Trans. Vis. Comput. Graphics 22 (1) (2016) 111–120.

[85] M. Glueck, P. Hamilton, F. Chevalier, et al., PhenoBlocks: Phenotype comparison visualizations, IEEE Trans. Vis. Comput. Graphics 22 (1) (2016) 101–110.

[86] S. Goodwin, J. Dykes, A. Slingsby, C. Turkay, Visualizing multiple variables across scale and geography, IEEE Trans. Vis. Comput. Graphics 22 (1) (2016) 599–608.

[87] J.C. Roberts, Display models: Ways to classify visual representations, in: Proc. IEEE Conference on Information Visualization, 1999.
[88] J.S. Yi, Y. Kang, J.T. Stasko, J.A. Jacko, Toward a deeper understanding of the role of interaction in information visualization, IEEE Trans. Vis. Comput. Graphics 13 (6) (2007) 1224–1231.
[89] M.C. Chuah, S.F. Roth, On the semantics of interactive visualizations, in: Proceedings of the IEEE Symposium on Information Visualization, 1996, pp. 29–36.
[90] H. Lam, A framework of interaction costs in information visualization, IEEE Trans. Vis. Comput. Graphics 14 (6) (2008) 1149–1156.
[91] P. Legrenzi, V. Girotto, P.N. Johnson-Laird, Focussing in reasoning and decision making, Cognition 49 (1/2) (1993) 37–66.
[92] P. Pirolli, S. Card, The sensemaking process and leverage points for analyst technology as identified through cognitive task analysis, in: Proc. of the International Conference on Intelligence Analysis, 2005, pp. 2–4.
[93] B. Craft, P. Caims, Beyond guidelines: What can we learn from the visual information seeking mantra? in: Proceedings of the Ninth International Conference on Information Visualisation, IEEE Computer Society Washington, DC, USA, 2005, pp. 110–118.
[94] M. Brehmer, T. Munzner, A multi-level typology of abstract visualization tasks, IEEE Trans. Vis. Comput. Graphics 19 (12) (2013) 2376–2385.
[95] S. Kaisler, F. Armour, J.A. Espinosa, W. Money, Big data: Issues and challenges moving forward, in: 46th Hawaii International Conference on System Sciences, 2013, pp. 995–1004.
[96] V.L. Lemieux, B. Gormly, L. Rowledge, Meeting big data challenges with visual analytics: The role of records management, Records Manag. J. 24 (2) (2014) 122–141.
[97] S. Pouyanfar, Y. Yang, S.-C. Chen, M.-L. Shyu, S.S. Iyengar, Multimedia big data analytics: A survey, ACM Comput. Surv. 51 (1) (2018) 10:1–10:34.
[98] Qin. H.F., Li. Z.H., Research on the method of big data analysis, Inf. Technol. J. 72 (10) (2013) 1–7.
[99] J. Fekete, Visual analytics infrastructures: From data management to exploration, Computer 46 (2013) 22–29.
[100] M. Blytt, Big challenges for visual analytics. assisting sensemaking of big data with visual analytics, 2014, Retrieved on 26 April 2018 from, https://www.ntnu.no/documents/10401/1264433962/MikkelArtikkel.pdf/.
[101] G. Oguntala, R. Abd-Alhameed, S. Jones, J. Noras, M. Patwary, J. Rodriguez, Indoor location identification technologies for real-time IoT-based applications: An inclusive survey, Comp. Sci. Rev. 30 (2018) 55–79.
[102] M. Mohammadi, A. Al-Fuqaha, Enabling cognitive smart cities using big data and machine learning: Approaches and challenges, IEEE Commun. Mag. 56 (2) (2018) 94–101.
[103] W. Dou, D.H. Jeong, F. Stukes, W. Ribarsky, H.R. Lipford, R. Chang, Recovering reasoning process from user interactions, IEEE Comput. Graph. Appl. 29 (3) (2009) 52–61.
[104] D. Gotz, M.X. Zhou, Characterizing users' visual analytic activity for insight provenance, Inf. Vis. 8 (1) (2009) 42–55.
[105] E.D. Ragan, A. Endert, J. Sanyal, J. Chen, Characterizing provenance in visualization and data analysis: An organizational framework of provenance types and purposes, IEEE Trans. Vis. Comput. Graphics 22 (1) (2016) 31–40.
[106] H. Lam, E. Bertini, P. Isenberg, C. Plaisant, S. Carpendale, Empirical studies in information visualization: Seven scenarios, IEEE Trans. Vis. Comput. Graphics 18 (9) (2012) 1520–1536.
[107] S. Amri, H. Ltifi, M. Ben Ayed, Comprehensive evaluation method of visual analytics tools based on fuzzy theory and artificial neural network, in: The 12th International Conference on Intelligent Systems and Knowledge and Engineering (ISKE), 24–26 November, IEEE, Nanjing, China, 2017.
[108] M. Tory, T. Moeller, Evaluating visualizations: Do expert reviews work, IEEE Comput. Graph. Appl. 25 (5) (2005) 8–11.
[109] G.D. Sun, Y.C. Wu, R.H. Liang, et al., A survey of visual analytics techniques and applications: State-of-the-art research and future challenges, J. Comput. Sci. Technol. 28 (5) (2013) 852–867.
[110] H. Ltifi, S. Amri, M. Ben Ayed, Fuzzy logic–based evaluation of visualizations generated by intelligent decision support systems, Inf. Vis. 17 (1) (2018) 3–21.
[111] P.C. Wong, H.-W. Shen, C.R. Johnson, C. Chen, R.B. Ross, The top 10 challenges in extreme-scale visual analytics, IEEE Comput. Graph. Appl. 32 (4) (2012) 63–67.
[112] M. Sturdee, J. Alexander, Analysis and classification of shape-changing interfaces for design and application-based research, ACM Comput. Surv. 51 (1) (2018) 2:1–2:32.