

An Overview of Interaction Techniques and 3D Representations for Data Mining

Ben Said Zohra¹, Guillet Fabrice¹, Richard Paul²,
Blanchard Julien¹ and Picarougne Fabien¹

¹*University of Nantes*

²*University of Angers*

France

1. Introduction

Since the emergence of databases in the 1960s, the volume of stored information has grown exponentially every year (Keim (2002)). This information accumulation in databases has motivated the development of a new research field: Knowledge Discovery in Databases (KDD) (Frawley et al. (1992)) which is commonly defined as the extraction of potentially useful knowledge from data. The KDD process is commonly defined in three stages: pre-processing, Data Mining (DM), and post-processing (Figure 1). At the output of the DM process (post-processing), the decision-maker must evaluate the results and select what is interesting. This task can be improved considerably with visual representations by taking advantage of human capabilities for 3D perception and spatial cognition. Visual representations can allow rapid information recognition and show complex ideas with clarity and efficacy (Card et al. (1999)). In everyday life, we interact with various information media which present us with facts and opinions based on knowledge extracted from data. It is common to communicate such facts and opinions in a virtual form, preferably interactive. For example, when watching weather forecast programs on TV, the icons of a landscape with clouds, rain and sun, allow us to quickly build a picture about the weather forecast. Such a picture is sufficient when we watch the weather forecast, but professional decision-making is a rather different situation. In professional situations, the decision-maker is overwhelmed by the DM algorithm results. Representing these results as static images limits the usefulness of their visualization. This explains why the decision-maker needs to be able to interact with the data representation in order to find relevant knowledge. Visual Data Mining (VDM), presented by Beilken & Spenke (1999) as an interactive visual methodology "to help a user to get a feeling for the data, to detect interesting knowledge, and to gain a deep visual understanding of the data set", can facilitate knowledge discovery in data.

In 2D space, VDM has been studied extensively and a number of visualization taxonomies have been proposed (Herman et al. (2000), Chi (2000)). More recently, hardware progress has led to the development of real-time interactive 3D data representation and immersive Virtual Reality (VR) techniques. Thus, aesthetically appealing element inclusion, such as 3D graphics and animation, increases the intuitiveness and memorability of visualization. Also, it eases the perception of the human visual system (Spence (1990), Brath et al. (2005)). Although there is still a debate concerning 2D vs 3D data visualization (Shneiderman (2003)), we believe that

3D and VR techniques have a better potential to assist the decision-maker in analytical tasks, and to deeply immerse the user's in the data sets. In many cases, the user needs to explore data and/or knowledge from the inside-out and not from the outside-in, like in 2D techniques (Nelson et al. (1999)). This is only possible in using VR and Virtual Environment (VEs). VEs allow users to navigate continuously to new positions inside the data sets, and thereby obtain more information about the data. Although the benefits offered by VR compared to desk-top 2D and 3D still need to be proven, more and more researchers are investigating its use with VDM (Cai et al. (2007)). In this context, we are trying to develop new 3D visual representations to overcome some limitations of 2D representations. VR has already been studied in different areas of VDM such as pre-processing (Nagel et al. (2008), Ogi et al. (2009)), classification (Einsfeld et al. (2006)), and clustering (Ahmed et al. (2006)).

In this context, we review some work that is relevant for researchers seeking or intending to use 3D representation and VR techniques for KDD. We propose a table that summarizes 14 VDM tools focusing on 3D - VR and interaction techniques based on 3 dimensions:

- Visual representations;
- Interaction techniques;
- Steps in the KDD process.

This paper is organized as follows: firstly, we introduce VDM. Then we define the terms related to this field of research. In Section 3, we explain our motivation for using 3D representation and VR techniques. In Section 4, we provide an overview of the current state of research concerning 3D visual representations. In Section 5, we present our motivation for interaction techniques in the context of KDD. In Section 6, we describe the related work about visualization taxonomy and interaction techniques. In Section 7, we propose a new classification for VDM based on both 3D representations and interaction techniques. In addition, we survey representative works on the use of 3D and VR interaction techniques in the context of KDD. Finally, we present possible directions for future research.

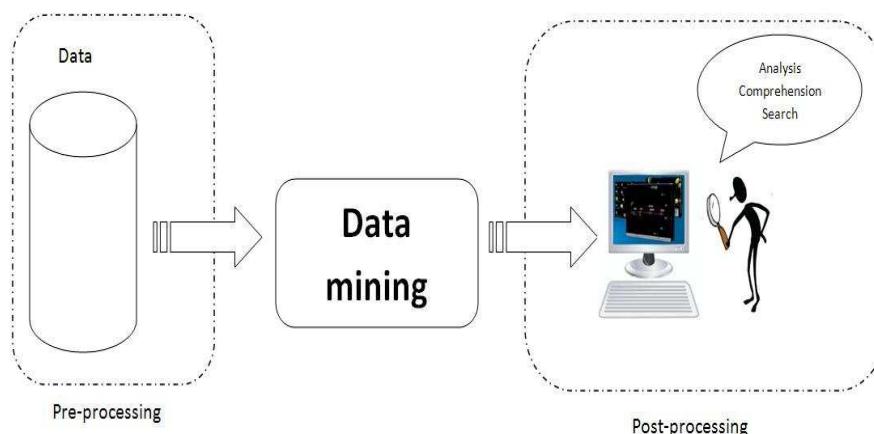


Fig. 1. The KDD process

2. Visual Data Mining

Historically, VDM has evolved from the fields of scientific visualization and information visualization. Both visualizations forms create visual representations from data that support user interaction with the aim to find useful information in the data. In scientific visualization, visual representations are typically constructed from measured or simulated data which represent objects or concepts of the physical world. Figure 2(a) shows an application that provides a VR interface to view the flow field around a space shuttle. In information visualization, graphic models present abstract concepts and relationships that do not necessarily have a counterpart in the physical world. For instance, figure 2(b) shows a 3D tree representation to visualize data clusters.



Fig. 2. Scientific visualization and information visualization examples: (a): visualization of the flow field around a space shuttle (Laviola (2000)) (b): GEOMIE (Ahmed et al. (2006)) information visualization framework

Beilken & Spenke (1999) presented the purpose of VDM as a way to "help a user to get a feeling for the data, to detect interesting knowledge, and to gain a deep visual understanding of the data set". Niggemann (2001) looked at VDM as a visual representation of the data close to the mental model. In this paper we focus on the interactive exploration of data and knowledge that is built on extensive visual computing(Gross (1994)).

As humans understand information by forming a mental model which captures only the main information, in the same way, data visualization, similar to the mental model, can reveal hidden information encoded in the data. In addition to the role of the visual data representation, Ankerst (2001) explored the relation between visualization and the KDD process. He defined VDM as "a step in the KDD process that utilizes visualization as a communication channel between the computer and the user to produce novel and interpreted patterns". He also explored three different approaches to VDM, two of which affect the final or intermediate visualization results. The third approach involves the interactive manipulation of the visual representation of the data rather than the results of the KDD methods. The three definitions recognize that VDM relies heavily on human perception capabilities and the use of interactivity to manipulate data representations. The three definitions also emphasize the key importance of the following three aspects of VDM: visual representations; interaction processes; and KDD tasks.

In most of the existing KDD tools, VDM is only used during two particular steps of the KDD process: in the first step (pre-processing) VDM can play an important role since analysts need tools to view and create hypotheses about complex (i.e. very large and / or high-dimensional) original data sets. VDM tools, with interactive data representation and query resources, allow domain experts to explore quickly the data set (de Oliveira & Levkowitz (2003)). In the last step (post-processing) VDM can be used to view and to validate the final results that are mostly multiple and complex. Between these two steps, an automatic algorithm is used to perform the DM task. Some new methods have recently appeared which aim at involving the user more significantly in the KDD process; they use visualization and interaction more intensively, with the ultimate goal of gaining insight into the KDD problem described by vast amounts of data or knowledge. In this context, VDM can turn the information overload into an opportunity by coupling the strengths of machines with that of humans. On the one hand, methods from KDD are the driving force of the automatic analysis side, while on the other hand, human capabilities to perceive, relate and make conclusions turn VDM into a very promising research field. Nowadays, fast computers and sophisticated output devices can create meaningful visualization and allow us not only to visualize data and concepts, but also to explore and interact with this data in real-time. Our goal is to look at VDM as an interactive process with the visual representation of data allowing KDD tasks to be performed. The transformation of data / knowledge into significant visualization is not a trivial task. Very often, there are many different ways to represent data and it is unclear which representations, perceptions and interaction techniques needs to be applied. This paper seeks to facilitate this task according to the data and the KDD goal to be achieved by reviewing representation and interaction techniques used in VDM. KDD tasks have different goals and diverse tasks need to be applied several times to achieve a desired result. Visual feedback has a role to play, since the decision-maker needs to analyze such intermediate results before making a decision. We can distinguish two types of cognitive process within which VDM assists users to make a decision:

- Exploration: the user does not know what he/she is looking for (discovery).
- Analysis: the user knows what he/she is looking for in the data and tries to verify it (visual analysis).

3. From 2D to 3D visualization and virtual reality

There is a controversial debate on the use of 2D versus 3D and VR for information visualization. In order to justify our choice of 3D and VR, we first review the difference between 3D visualizations and VR techniques:

- 3D visualization is a representation of an object in a 3D space by showing length, width and height coordinates on a 2D surface such as a computer monitor. 3D visual perception is achieved using visual depth cues such as lighting, shadows and perspective.
- VR techniques enable user immersion in a multi-sensorial VE and user interaction devices and stereoscopic images to increase depth perception and the relative 3D position of objects.

3.1 2D versus 3D

Little research has been dedicated to the comparison of 2D and 3D representations. Concerning the non-interactive visualization of static graphs, 3D representations have

generally not been advised ever since the publications by Tufte (1983) and Cleveland & McGill (1984). Nevertheless, the experiments of Spence (1990) and Carswell et al. (1991) show that there is no significant difference of accuracy between 2D and 3D for the comparison of numerical values. In particular, Spence (1990) pointed out that it is not the apparent dimensionality of visual structures that counts but rather the actual number of parameters that show variability. Under some circumstances, information may be processed even faster when represented in 3D rather than in 2D. Concerning the perception of global trends in data, experimental results of Carswell et al. (1991) also show an improvement in answer times using 3D but to the detriment of accuracy. Other works compare 2D and 3D within the framework of interactive visualization. Ware & Franck (1994) indicated that displaying data in 3D instead of 2D can make it easier for users to understand the data. Finally, Tavanti & Lind (2001) pointed out that realistic 3D displays could support cognitive spatial abilities and memory tasks, namely remembering the place of an object, better than with 2D.

On the other hand, several problems arise such as intensive computation, more complex implementations than 2D interfaces, and user adaptation and disorientation. The first problem can be addressed by using powerful and specialized hardware. However, one of the main problems of 3D applications is user adaptation. In fact, most users just have experience with classical windows, icons, menu pointing devices (WIMP) and 2D-desktop metaphors. Therefore, interaction with 3D presentations and possibly the use of special devices demand considerable adaptation efforts to use this technology. There is still no commonly-accepted standard for interaction with 3D environments. Some research has shown that it takes users some time to understand what kind of interaction possibilities they actually have (Baumgärtner et al. (2007)). In particular, as a consequence of a richer set of interactions and a higher degree of freedom, users may be disoriented.

3.2 Toward virtual reality

To overcome limitations of interaction with 3D representations, VR interfaces and input devices have been proposed. These interfaces and devices offer simpler and more intuitive interaction techniques (selection, manipulation, navigation, etc.), and more compelling functionality (Shneiderman (2003)). In VR, the user can always access external information without leaving the environment and the context of the representation. Also, the user's immersion in the data allows him to take advantage of stereoscopic vision that enables him to disambiguate complex abstract representations (Maletic et al. (2001)). Ware & Franck (1996), compared the visualization of 2D and 3D graphs. Their work shows a significant improvement in *intelligibility* when using 3D. More precisely, they found that the ability to decide if two nodes are connected or not is improved by a factor 1.6 when adding stereo cues, by 2.2 when using motion parallax depth cues, and by a factor of 3 when using stereoscopic as well as motion parallax depth cues. Aitsisalmi & Holliman (2009), found that the participants obtained better scores if they were doing a mental rotation task on a stereoscopic screen instead of a 2D screen. This result demonstrates the *efficiency* of VR and shows that the extra depth information given by stereoscopic display makes it easier to move a shape mentally. It is generally considered that only stereoscopy allows one to fully exploit the characteristics of the 3D representations. It helps the viewer to judge the relative size of objects and the distances between them. It also helps him to mentally move a shape in the 3D visualization area. Finally, Cai et al. (2007), found that visualization increases *robustness* in object tracking and positive detection accuracy in object prediction. They also found that the interactive method enables

the user to process the image data 30 times faster than manually. As a result, they suggested that human interaction may significantly increase overall productivity.

We can therefore conclude that stereoscopy and interaction are the two most important components of VE and the most useful to users. Therefore, the equipment used should be taken into account from the very beginning of application design, and consequently be taken into account as a part of VDM techniques taxonomy.

4. Visual representations for Visual Data Mining

One of the problems that VDM must address is to find an effective representation of something that has no inherent form. In fact, it is crucial not only to determine which information to visualize but also to define an effective representation to convey the target information to the user. The design of a visualization representation must address a number of different issues: what information should be presented? How this should be done? What level of abstraction to support? etc. For example, a user tries to find out interesting relations between variables in large databases. This information may be visualized as a graph (Pryke & Beale (2005)) or as an abstract representation based on a sphere and cone (Blanchard et al. (2007)).

Many representations for VDM have been proposed. For instance, some visual representations are based on *abstract representations*, such as graphs (Ahmed et al. (2006)), trees (Einsfeld et al. (2007), Buntain (2008)), and geometrical shapes (Ogi et al. (2009), Nagel et al. (2008), Meiguins et al. (2006)) and others on *virtual worlds objects* (Baumgärtner et al. (2007)). The classification proposed in this chapter provides some initial insight into which techniques are oriented to certain data types, but does not assert that one visual representation is more suitable than others to explore a particular data set. Selecting a representation depends largely on the task being supported and is still a largely intuitive process.

4.1 Abstract visual representations

3D representations are still abstract and require the user to learn certain conventions, because they do not look like what they refer to or they do not have a counterpart in the real-world. There are 3 kinds of abstract representations: graphs, trees, and geometrical shapes.

1. Graphs

A graph (Figure.3) is a network of nodes and arcs, where the nodes represent entities while the arcs represent relationships between entities. For a review on the state of the art in graph visualization see Herman et al. (2000).

At the beginning, graph visualization was used in 2D space to represent components around simple boxes and lines. However, several authors think that larger graph structures can be viewed in 3D (Parker et al. (1998)). In the empirical study of Ware & Franck (1996), which measured path-tracing ability in 3D graphs, they suggested that the amount of information that can be displayed in 3D with stereoscopic and motion depth cues exceeds 2D representations by a factor of 3. Another experiment with new display technologies confirmed the previous experiment and showed much greater benefits than previous studies. Ware & Mitchell (2008) experiments showed that the use of stereoscopic display, kinetic depth and 3D tubes was much more beneficial than using lines to display the links as in previous studies.

A technique based on the hyper system (Hendley et al. (1999)) for force-based visualization can be used to create a graph representation. The visualization consists of nodes and links whose properties are given by the parameters of the data. Data elements affect parameters such as node size and color, link strength and elasticity. The dynamic graphs algorithm enables the self-organization of nodes in the visualization area by the use of a force system in order to find a steady state, and determine the position of the nodes. For example, Beale (2007) proposed a Haiku system (Figure.3(b)) which provides an abstract 3D perspective of clustering algorithm results based on the hyper system. One of the characteristics of this system is that the user can choose which parameters are used to create the distance metrics (distance between two nodes), and which ones affect the other characteristics of the visualization (node size, link elasticity, etc.). Using the hyper system allows related things (belonging to the same cluster) to be near to each other, and unrelated things to be far away.

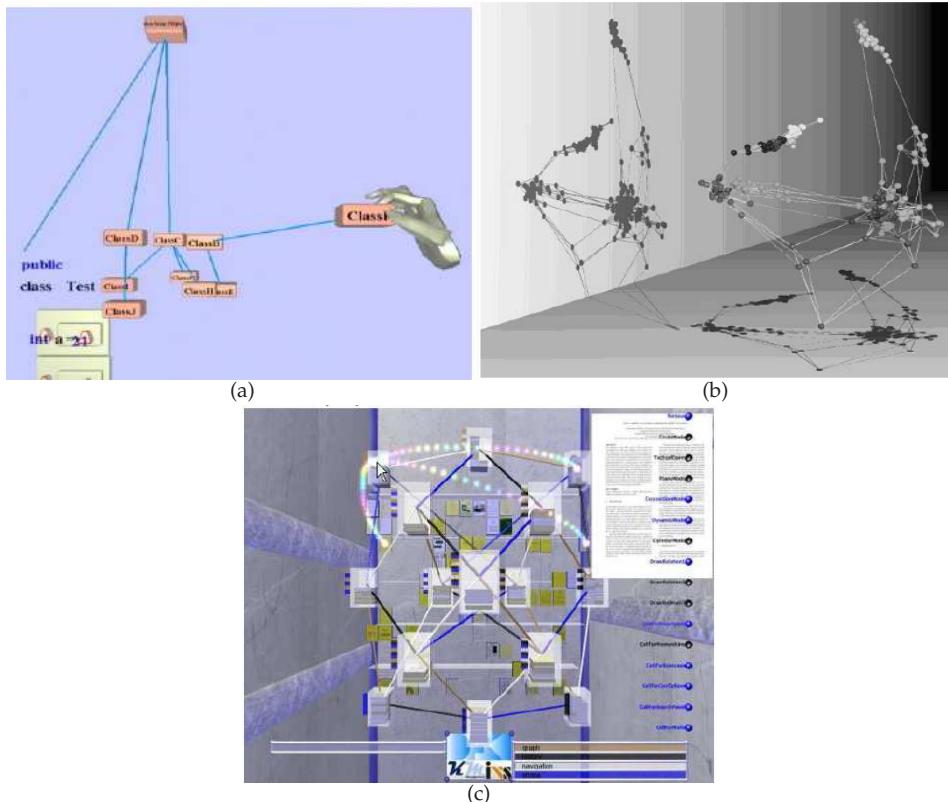


Fig. 3. An example of graph representations: (a) Source code Ougi (Osawa et al. (2002)), (b) Association rules: Haiku (Beale (2007)), (c) DocuWorld (Einsfeld et al. (2006))

2. 3D trees

3D trees (Figure.4) is a visualization technique based on the hierarchical organization of data. A tree can represent many entities and the relationships between them. In general, the

visualization of hierarchical information structures is an important topic in the information visualization community (Van Ham (2002)). Because trees are generally easy to layout and interpret (Card et al. (1999)), this approach finds many applications in classification visualization (Buntain (2008)). 3D trees were designed to display a larger number of entities than in 2D representations, in a comprehensible form (Wang et al. (2006)). Various methods have been developed for this purpose, among which, space-filling techniques and node-link techniques.

Space-filling techniques (Van Ham (2002), Wang et al. (2006)) based upon 2D tree-maps visualization proposed by Johnson & Shneiderman (1991) have been successful for visualizing trees that have attributes values at the node level. Space-filling techniques are particularly useful when users care mostly about nodes and their attributes but do not need to focus on the topology of the tree, or consider that the topology of the tree is trivial (e.g 2 or 3 levels). The users of space-filling techniques also require training because of the unfamiliar layout (Plaisant et al. (2002)).

Node-link techniques, on the other hand, have long been frowned upon in the information visualization community because they typically make inefficient use of screen space. Even trees of a hundred nodes often need multiple screens to be completely displayed, or require scrolling since only part of the tree is visible at a given time. A well-known node-link representation in cone trees was introduced by Robertson et al. (1991) for visualizing large hierarchical structures in a more intuitive way. 3D trees may be displayed vertically (Cone Tree) or horizontally (Cam Tree).

Buntain (2008) used 3D trees for ontology classification visualization (Figure.4(a)). Each leaf represents a unique concept in the ontology, and the transparency and size of each leaf is governed by the number of documents associated with the given concept. A molecule is constructed by clustering together spheres that share common documents, and surrounds the leaves with a semi transparent shell (Figure.4(b)).

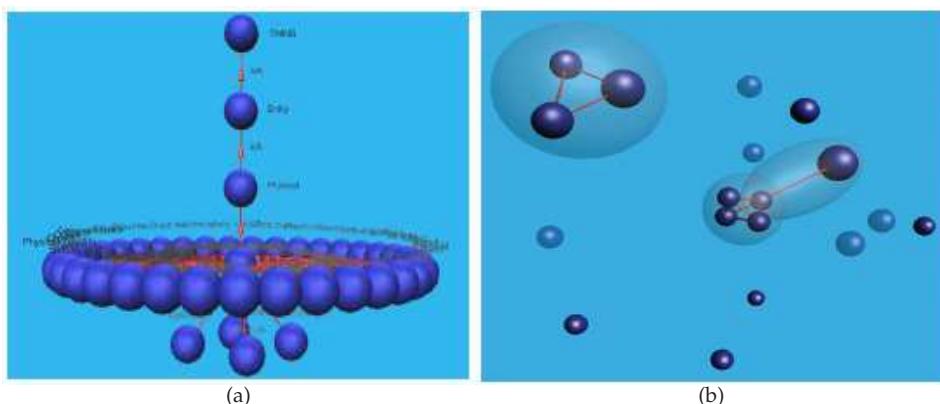


Fig. 4. An example of trees representing ontology classification: SUMO (Buntain (2008))

3. Geometric shapes

In this technique, 3D objects with certain attributes are used to represent data and knowledge. The 3D scatter-plot visualization technique (Nagel et al. (2001)) is one of the most common

representations based on geometric shapes (Figure 5). The main innovation compared to 2D visualization techniques is the use of volume rendering which is a conventional technique used in scientific visualization. 3D rendering techniques use voxels (instead of pixels in 2D) to present a certain density of the data. 3D scatter-plot has been adapted by Becker (1997), making the opacity of each voxel a function of points density. Using scatter-plots is intuitive since each data is faithfully displayed. Scatter-plots have been used successfully for detecting relationships in two dimensions (Bukauskas & Böhnen (2001), Eidenberger (2004)). This technique hit limitations if the dataset is large, noisy, or if it contains multiple structures. With large amounts of data, the amount of displayed objects makes it difficult to detect any structure at all.

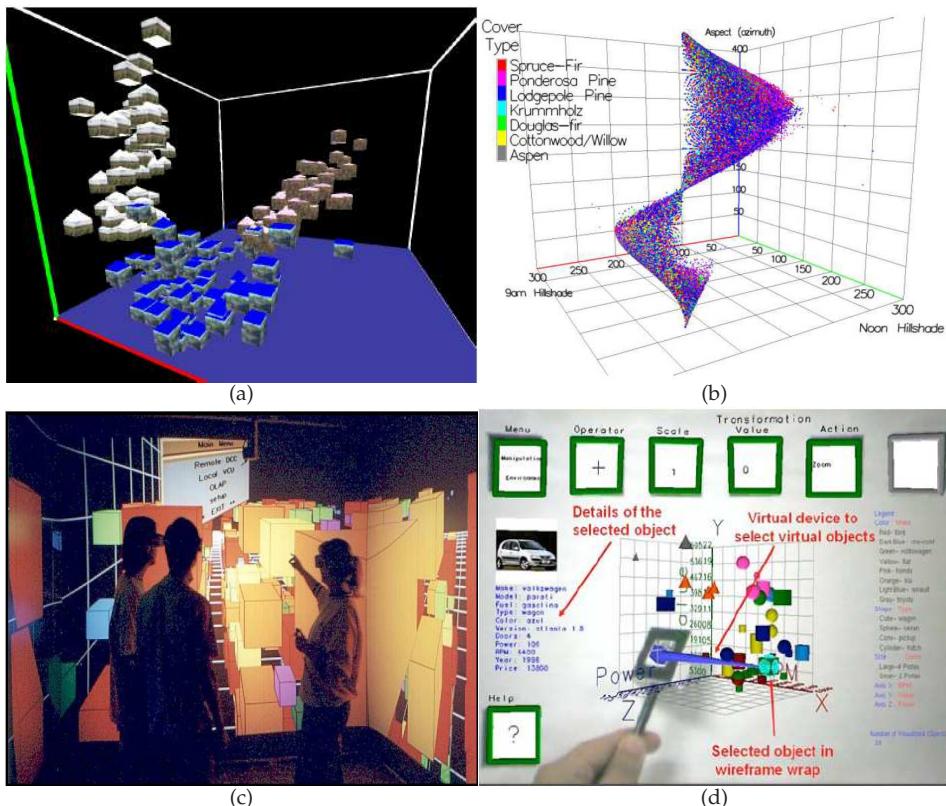


Fig. 5. Different 3D scatter plot representations: (a) VRMiner (Azzag et al. (2005)), (b) 3DVDM (Nagel et al. (2008)), (c) DIVE-ON (Ammoura et al. (2001)), (d) Visualization with augmented reality (Meiguins et al. (2006))

4.2 Virtual worlds

Trying to find easily-understandable data representations, several researchers proposed the use of real-world metaphors. This technique uses elements of the real-world to provide insights about data. For example, some of these techniques are based on a city abstraction

(Figure.6). The virtual worlds (sometimes called cyber-spaces) for VDM are generally based either on the information galaxy (Krohn (1996)) or the information landscape metaphor (Robertson et al. (1998)). The difference between the two metaphors is that in the information landscape, the elevation of objects is not used to represent information (objects are placed on a horizontal floor). The specificity of virtual worlds is that they provide the user with some real world representations.

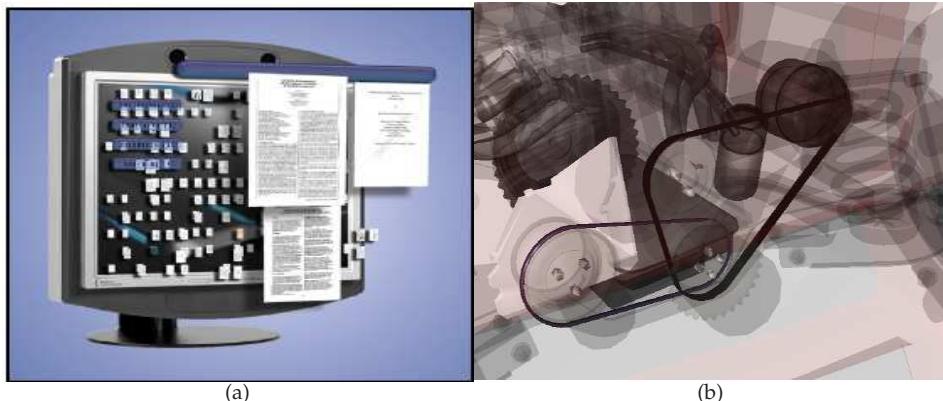


Fig. 6. Example of virtual world representation (a) faults projected onto a car model in (Götzelmann et al. (2007)) (b) documents classification in @VISOR Baumgärtner et al. (2007)

5. Interaction techniques for Visual Data Mining

Interaction techniques can empower the user's perception of information when visually exploring a data set (Hibbard et al. (1995)). The ability to interact with visual representations can greatly reduce the drawbacks of visualization techniques, particularly those related to visual clutter and object overlap, providing the user with mechanisms for handling complexity in large data sets. Pike et al. (2009) explored the relationship between interaction and cognition. They consider that the central percept of VDM is that the development of human insight is aided by interaction with a visual interface. As VDM is concerned with the relationship between visual displays and human cognition, merely developing only novel visual metaphors is rarely sufficient to make new discoveries provide or confirmation or negation of a prior belief.

Interaction also allows the integration of the user in the KDD process. KDD is not a completely human-guided process, since DM algorithms analyze a data set searching for useful information and statistically valid knowledge. The degree of automation of the KDD process actually varies considerably since different levels of humans guidance and interaction are usually required. But it is still the algorithm, and not the user, that is looking for knowledge. In this context, de Oliveira & Levkowitz (2003) suggested that VDM should have a greater role than a traditional application of visualization techniques to support the non-analytic stages of a KDD process. It is through the interactive manipulation of a visual interface that knowledge is constructed, tested, refined and shared.

We can distinguish 3 different interaction categories: exploration, manipulation and human-centered approaches.

5.1 Visual exploration

Visual exploration techniques are designed to take advantage of the considerable visual capabilities of human beings, especially when users try to analyze tens or even hundreds of graphic variables in a particular investigation. Visual exploration allows the discovery of data trends, correlations and clusters, to take place quickly, and can support users in formulating hypotheses about the data. It is essential in some situations to allow the user to simply look at the visual representation in a passive sense. This may mean moving around the view point in order to reveal structure in the data that may be otherwise masked and overlooked. In this way, exploration provides the means to view information from different perspectives to avoid occlusion and to see object details. It can be very useful to have the ability to move the image to resolve any perceptual ambiguities that exist in a static representation when a large amount of information is displayed at once. The absence of certain visual cues (when viewing a static image) can mask important results (Kalawsky & Simpkin (2006)).

Navigation is often the primary task in 3D worlds and refers to the activity of moving through the scene. The task of navigation presents challenges such as supporting spatial awareness and providing efficient and comfortable movements between distant locations. Some systems enable users to navigate without constraint through the information space (Nagel et al. (2008), Einsfeld et al. (2006), Azzag et al. (2005)). Other systems restrict movement in order to reduce possible user disorientation (Ahmed et al. (2006)). As an illustration, in VRMiner (Azzag et al. (2005)) a six-degree freedom sensor is fixed to the user's hand (Figure.7) allowing him/her to easily define a virtual camera in 3D space. For example, when the user moves his hand forward in the direction of the object, he/she may zoom in or out. The 3DVDM system (Nagel et al. (2008)) allows the user to fly around and within the visualized scatter-plot. The navigation is controlled by the direction of a "wanda" device tracked with 6 degrees of freedom. Dissimilarly, in GEOMI (Ahmed et al. (2006)), the user can only rotate the representation along the X and Y axes but not along the Z axis.

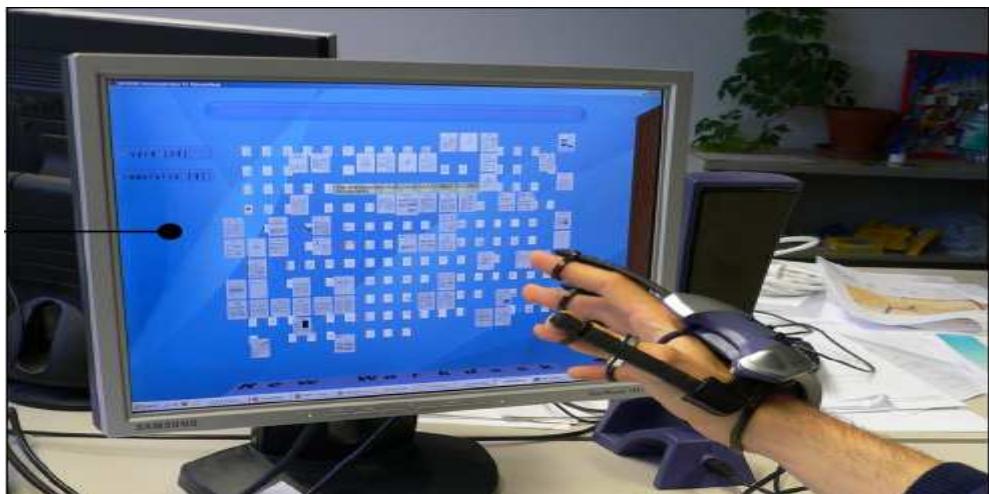


Fig. 7. Illustration of navigation through a virtual environment with data-glove (Baumgärtner et al. (2007))

In visual exploration, the user can also manipulate the objects in the scene. In order to do this, interaction techniques provide means to select and zoom-in and zoom-out to change the scale of the representation. Beale (2007) has demonstrated that using a system which supports the free exploration and manipulation of information delivers increased knowledge even from a well known dataset. Many systems provide a virtual hand or a virtual pointer (Einsfeld et al. (2007)), a typical approach used in VE, which is considered as being intuitive as it simulates real-world interaction (Bowman et al. (2001)).

- Select: this technique provides users with the ability to mark interesting data items in order to keep track of them when too many data items are visible, or when the perspective is changed. In these two cases, it is difficult for users to follow interesting items. By making items visually distinctive, users can easily keep track of them even in large data sets and/or with changed perspectives.
- Zoom: by zooming, users can simply change the scale of a representation so that they can see an overview (context) of a larger data set (using zoom-out) or the detailed view (focus) of a smaller data set (using zoom-in). The essential purpose is to allow hidden characteristics of data to be seen. A key point here is that the representation is not fundamentally altered during zooming. Details simply come into focus more clearly or disappear into context.

Visual exploration (as we can see in Section.7) can be used in the pre-processing of the KDD process to identify interesting data (Nagel et al. (2008)), and in post-processing to validate DM algorithm results (Azzag et al. (2005)). For example, in VRMiner (Azzag et al. (2005)) and in ArVis (Blanchard et al. (2007)), the user can point to an object to select it and then obtain informations about it.

5.2 Visual manipulation

In KDD, the user is essentially faced with a mass of data that he/she is trying to make sense of. He/she should look for something *interesting*. However, *interest* is an essentially human construct, a perspective of relationships among data that is influenced by tasks, personal preferences, and past experience. For this reason, the search for knowledge should not only be left to computers; the user has to guide it depending upon what he/she is looking for, and hence which area to focus computing power on. Manipulation techniques provide users with different perspectives of the visualized data by changing the representation. One of this techniques is the capability of changing the attributes presented in the representation. For example, in the system shown by Ogi et al. (2009), the user can change the combination of presented data. Other systems have interaction techniques that allow users to move data items more freely in order to make the arrangement more suitable for their particular mental model (Einsfeld et al. (2006)). Filter interaction techniques enable users to change the set of data items being presented on some specific conditions. In this type of interaction, the user specifies a range or condition, so that only data meeting those criteria are presented. Data outside the range or not satisfying the conditions are hidden from the display or shown differently; even so, the actual data usually remain unchanged so that whenever users reset the criteria, the hidden or differently-illustrated data can be recovered. The user is not changing data perspectives, just specifying conditions within which data are shown. ArVis (Blanchard et al. (2007)), allows the user to look for a rule with a particular item in it. To do this, the user can search for it in a menu which lists all the rule items and allows the wanted object to be shown.

5.3 Human-centered approach

In most existing KDD tools, interaction can be used in two different ways: exploration and manipulation. Some new methods have recently appeared (Baumgärtner et al. (2007), Poulet & Do (2008)), trying to involve the user in the DM process more significantly and using visualization and interaction more intensively. In this task, the user manipulates the DM algorithm and not only the graphical representation. The user sends commands to the algorithm in order to manipulate the data to be extracted. We speak here about local knowledge discovery. This technique allows the user to focus on interesting knowledge from user's point of view, in order to make the DM tool more generically useful to the user. It is also necessary for the user to either change the view point or manipulate a given parameter of the knowledge discovery algorithm and observe its effect. There must therefore be some way in which the user can indicate what it is considered interesting and what is not, and to do this the KDD tool needs to be dynamic and versatile (Ceglar et al. (2003)). The human-centered process should be iterative since it is repeated until the desired results are obtained. From a human interaction perspective, a human-centered approach closes the loop between the user and the DM algorithm in a way that allows them to respond to results as they occur by interactively manipulating the input parameters (Figure.8).

With the purpose of involving the user more intensively in the KDD process, this new kind of approach has the following advantages (Poulet & Do (2008))

- The quality of the results is improved by the use of human-knowledge recognition capabilities;
- Using the domain knowledge during the whole process (and not only in the interpretation of the results) allows guided searching for knowledge.
- The confidence in the results is improved as the DM process gives more comprehensible results.

In Arvis (Blanchard et al. (2007)), the user can navigate among the subsets of rules via a menu providing neighborhood relations. By applying a neighborhood relation to a rule, the mining algorithm extracts a new subset of rules. The previous subset is replaced by the new subset in the visualization area.

6. Related work on taxonomies of visual representations and interaction techniques

Many researchers have attempted to construct a taxonomy for visualization. Chi (2000) used the Data State Model (Chi & Riedl (1998)) to classify information visualization techniques. This model is composed of 3 dimensions with categorical values: data stages (value, analytical abstraction, visualization abstraction, and view), transformation operators (data transformation, visualization transformation, and visual mapping transformation), and within-stage operators (value stage, analytical stage, visualization stage, and view stage). This model shows how data change from one stage to another requiring one of the three types of data transformation operators. This state model helps implementers understand how to apply and implement information visualization techniques. Tory & Moller (2004), present a high-level taxonomy for visualization which classifies visualization algorithms rather than data. Algorithms are categorized according to the assumption that they make about the data being visualized. Their taxonomy is based on 2 dimensions:

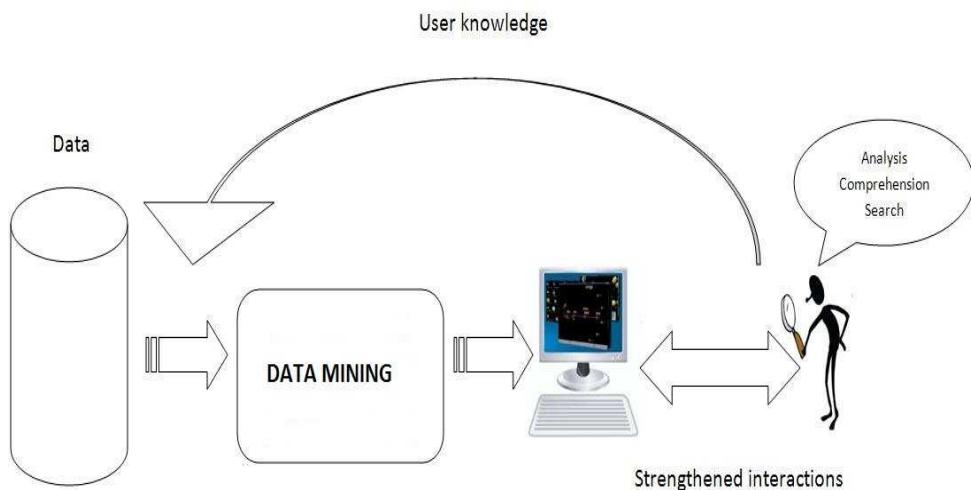


Fig. 8. The human-centered approach

- Data values: discrete or continuous
- How the algorithm designer chooses to display attributes: specialization, timing, color, and transparency.

Another area of related research is interaction and user interfaces. In this area, (Bowman et al. (2001)) present an overview of 3D interaction and user interfaces (3DUI). This paper also discusses the effect of common VE hardware devices on user interaction, as well as interaction techniques for generic 3D tasks and the use of traditional WIMP styles in 3D environments. They divide most user interaction tasks into three categories: navigation, selection/manipulation and system control. Arns (2002) thinks that Bowman's taxonomy is general and can encompass too many parts of a VR system. For that reason, she created a classification for virtual locomotion (travel) methods. This classification includes information on display devices, interaction devices, travel tasks, and the two primary elements of virtual travel: translation and rotation. Dachselt & Hinz (2005) have proposed a classification of 3D-widget solutions by interaction purpose/intention of use, e.g. direct 3D object interaction, 3D scene manipulation, exploration and visualization. Finally, Teyseyre & Campo (2009) presented an overview of 3D representations for visualizing software, describing several major aspects such as visual representations, interaction issues, evaluation methods, and development tools.

7. A new classification of Visual Data Mining based on visual representations and interaction techniques

In this section, we present a new classification of VDM tools composed of 3 dimensions: visual representations, interaction techniques, and KDD tasks. Table.1 presents the different modalities of each of the three dimensions. The proposed taxonomy takes into account both the representation and the interaction technique. In addition, many visualization

Dimension	Modalities
Visual representation	Graphs, 3D trees, geometrical shapes, virtual worlds
Interaction techniques	Visual exploration, visual manipulation, human-centered
KDD tasks	Pre-processing, classification, clustering, association rules

Table 1. Dimension modalities

design taxonomies include only a small subset of techniques (e.g., locomotion Arns (2002)). Currently, visualization tools have to provide not only effective visual representations but also effective interaction metaphors to facilitate the exploration and help users achieve insight. Having a good 3D representation without a good interaction technique does not mean having a good tool. This classification looks at some representative tools for doing different KDD tasks, e.g., pre-processing and post-processing (classification, clustering and association rules). Different tables summarize the main characteristics of the reported VDM tools with regard to visual representations and interaction techniques. Other relevant information such as interaction actions (navigation, selection and manipulation, and system control), input-output devices (CAVE, mouse, hand tracker, etc.) presentation (3D representation or VR representation) and year of creation is also reported.

7.1 Pre-processing

Pre-processing (in VDM) is the task of data visualization before the DM algorithm is used. It is generally required as a starting point of KDD projects so that analysts may identify interesting and previously unknown data by the interactive exploration of graphical representations of a data set without heavy dependence on preconceived assumptions and models. The basic visualization technique used for data pre-processing is the 3D scatter-plots method, where 3D objects with attributes are used as markers. The main principle behind the design of traditional VDM techniques, such as The Grand Tour (Asimov (1985)), the parallel coordinate (Inselberg & Dimsdale (1990)), etc., is that they are viewed from the outside-in. In contrast to this, VR lets users explore the data from inside-out by allowing users to navigate continuously to new positions inside the VE in order to obtain more information about the data. Nelson et al. (1999) demonstrated through comparisons between 2D and VR versions of the VDM tool XGobi that the VR version of XGobi performed better.

In the Ogi et al. (2009) system, the user can see several data set representations integrated in the same space. The user can switch the visible condition of each data set. This system could be used to represent the relationships among several data sets in 3D space, but it does not allow the user to navigate through the data set and interact with it. The user can only change the visual mapping of the data set. However, the main advantage of this system is that the data can be presented with a high degree of accuracy using high-definition stereo-images that can be beneficial especially when visualizing a large amount of data. This system has been applied to the visualization and analysis of earthquake data. Using the 3rd dimension has allowed the visualization of both the overall distribution of the hypocenter data and the individual location on any earthquake, which is not possible with the conventional 2D display. Figure 9 shows hypocenter data recorded over 3 years. The system allows the visualization of several databases at the same time e.g. map data, terrain data, basement depth, etc and the user can switch the visible condition of each data in the VE. For example, the user can change the visualization data from the combination of hypocenter data and basement depth

System	Visual Representation	Interaction techniques	Interaction actions			Input-Output devices	3D/ VR	year
			Navigation	Selection and Manipulation	System control			
Ogi et al. (2009)	Geometric shape	Visual manipulation	-	-	Graphical menu	CAVE	VR	2009
3DVDM Nagel et al. (2008) (Fig.5(c))	Geometric shape	Visual exploration	Manual view point manipulation + thought wizard metaphor	-	-	Three-button 'wand' + CAVE	VR	2008
Nested circles Wang et al. (2006)	Visual exploration and visual manipulation	Tree	-	Object selection + virtual pointer	Graphic menus	Mouse + 2D screen	3D	2006
Visualization with augmented reality Meiguins et al. (2006) (Fig.5(e))	Geometric shape	Visual manipulation	-	Object selection + virtual hand	Graphical menus	Hand tracker	VR	2006
Dive-On Ammoura et al. (2001) (Fig.5(d))	Geometric shape	Visual exploration and manipulation	Physical movement + steering + target-based travel	Object selection	Graphical menus	Hand + head tracker + CAVE	VR	2001

Table 2. 3D VDM tool summary for pre-processing KDD task

data to the combination of hypocenter data and terrain data. Thus, the system can shows the relationships between only any two data sets among the others.

As a result of using VR, the 3DVDM system (Nagel et al. (2008)) is capable of providing real-time user response and navigation as well as showing dynamic visualization of large amounts of data. Nagel et al. (2008) demonstrated that the 3DVDM visualization system allows faster detection of non-linear relationships and substructures in data than traditional methods of data analysis. An alternative proposal is available with DIVE-ON (Data mining in an Immersed Visual Environment Over a Network) system, proposed by Ammoura et al. (2001). The main idea of DIVE-ON is visualizing and interacting with data from distributed data warehouses in an immersed VE. The user can interact with such sources by walking or flying toward's them. He/she also can pop up a menu, scroll through it and execute all environment, remote, and local functions. Thereby, DIVE-ON makes intelligent use of the natural human capability of interacting with spatial objects and offers considerable navigation possibilities e.g. walking, flying, transporting and climbing.



Fig. 9. Visualization of earthquakes data using a 4K stereo projection system (Ogi et al. (2009))

Inspired by treemaps Wang et al. (2006) presented a novel space-filling approach for tree visualization of file systems (Figure.10). This system provides a good overview for a large hierarchical data set and uses nested circles to make it easier to see groupings and structural relationships. By clicking on an item (a circle), the user can see the associated sub-items represented by the nested circles in a new view. The system provides the user with a control panel allowing him/her to filter files by types; by clicking on one file type, the other files types are filtered out. A zoom-in/zoom-out function allows the user to see folder or file characteristics such as name, size, and date. A user-feedback system means that user interaction techniques are friendly and easy to use.

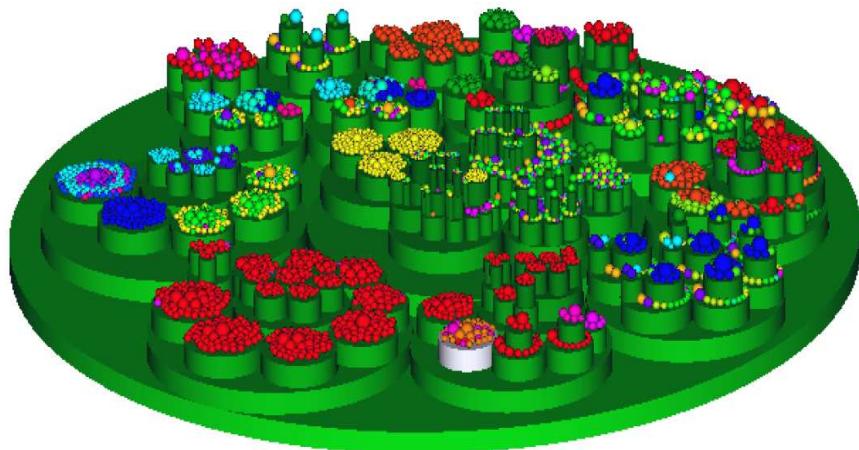


Fig. 10. Representation of a file system with 3D-nested cylinders and spheresWang et al. (2006)

Meiguins et al. (2006) presented a tool for multidimensional VDM visualization in an augmented-reality environment where the user may visualize and manipulate information in real time VE without the use of devices such as a keyboard or mouse and interact simultaneously with other users in order to make a decision related to the task being analyzed. This tool uses a 3D scatter-plot to visualize the objects. Each visualized object has specific characteristics of position (x, y and z axes), color, shape, and size that directly represent data item values. The main advantages of this tool is that provide users with a dynamic menu which is displayed in an empty area when the user wants to execute certain actions. The tool also allows users to perform many manipulation interactions tasks such as real-time filter attributes, semantic zoom, rotation and translation of objects in the visualization area. A detailed comparison of these techniques is presented in Table.2.

7.2 Post-processing

Post-processing is the final step of the KDD process. Upon receiving the output of the DM algorithm, the decision-maker must evaluate and select the interesting part of the results.

7.2.1 Clustering

Clustering is used for finding groups of items that are similar. Given a set of data items, this set can be partitioned into a set of classes, so that items with similar characteristics are grouped together.

The GEOMI system proposed by Ahmed et al. (2006) is a visual analysis tool for the visualization of clustered graphs or trees. The system implements block model methods to associate each group of nodes to corresponding cluster. Two nodes are in the same cluster if they have the same neighbor set. This tool allows immersive navigation in the data using 3D head gestures instead of the classical mouse input. The system only allows the user visual exploration. Users can walk into the network, move closer to nodes or clusters by simply aiming in their direction. Nodding or tilting the head rotates the entire graph along the X and Y axes respectively, which provides users with intuitive interaction.

The objective of @VSIOR (Baumgärtner et al. (2007)), which is a human-centered approach, is to create a system for interaction with document, meta-data, and semantic relations. Human capabilities in this context are spatial memory and the fast visual processing of attributes and patterns. Artificial intelligence techniques assist the user, e.g. in searching for documents and calculating document similarities.

Otherwise, VRMiner (Azzag et al. (2005)) uses stereoscopic and intuitive navigation; these allow the user to easily select the interesting view point. VRMiner users have found that using this tool helps them solve 3 major problems: detecting correlation between data dimensions, checking the quality of discovered clusters, and presenting the data to a panel of experts. In this context, the stereoscopic display plays a crucial role in addition to the intuitive navigation which allows the user to easily select the interesting view point.

A detailed comparison of these techniques is presented in Table.3.

7.2.2 Classification

Given a set of pre-defined categorical classes, determine which of these classes a specific data item belongs to.

System	Visual Representation	Interaction techniques	Interaction actions			Input-Output devices	3D/ VR	year
			Navigation	Selection and Manipulation	System control			
@VISOR Baumgärtner et al. (2007) (fig.6(b))	Graph	Human-centered	-	Object selection + object positioning + virtual hand	Graphical menu gestural interaction	Tablet PC(2D) + Data glove + stereoscopic	VR	2007
GEOMI Ahmed et al. (2006)	Graph tree	+ Visual exploration	Steering	-	Gestural interaction	Head tracker + stereoscopic	VR	2006
VRMiner Azzag et al. (2005) (Fig.5(b))	Abstract geometrical shape	Visual exploration	Manual view point manipulation	Object selection	Gestural interaction	Data glove + stereoscopic	VR	2005

Table 3. 3D VDM tool summary for clustering KDD task

In SUMO (Figure.4), a tool for document-class visualization is proposed (Buntain (2008)). The structure classes and relations among those classes can be presented to the user in a graphic form to facilitate understanding of the knowledge domain. This view can then be mapped onto the document space where shapes, sizes, and locations are governed by the sizes, overlaps, and other properties of the document classes. This view provides a clear picture of the relations between the resulting documents. Additionally, the user can manipulate the view to show only those documents that appear in a list of results from a query. Furthermore, if the results view includes details about subclasses of results and "near miss" elements in conjunction with positive results, the user can refine the query to find more appropriate results or widen the query to include more results if insufficient information is forthcoming. The third dimension allows the user a more expressive space, complete with navigation methods such as rotation and translation. In 3D, overlapping lines or labels can be avoided by rotating the layout to a better point of view.

DocuWorld (Einsfeld et al. (2006)), is a prototype for a dynamic semantic information system. This tool allows computed structures as well as documents to be organized by users. Compared to the web Forager (Card et al. (1996)), a workspace to organize documents with different degrees of interest at different distances to the user, DocuWorld provides the user with more flexible possibilities to store documents at locations defined by the user and visually indicates cluster-document relations (different semantics of connecting clusters to each other).

A detailed comparison of these techniques is presented in Table.4.

7.2.3 Association rules

On account of the enormous quantities of rules that can be produced by DM algorithms, association rule post-processing is a difficult stage in an association rule discovery process.

System	Visual Representation	Interaction techniques	Interaction actions			Input-Output devices	3D/ VR	year
			Navigation	Selection and Manipulation	System control			
SUMO Buntain (2008) (Fig.4)	Tree	Visual exploration	Manual view point manipulation	-	-	2D Mouse + 2D screen	3D	2008
DocuWorld Einsfeld et al. (2006) (Fig.3(c))	Graph	human-centred	Thought wizard metaphor	Object selection + object positioning + virtual pointer	Gestural interaction + voice commands	Mouse stereoscopic	+ VR	2006

Table 4. 3D VDM tool summary for classification KDD task

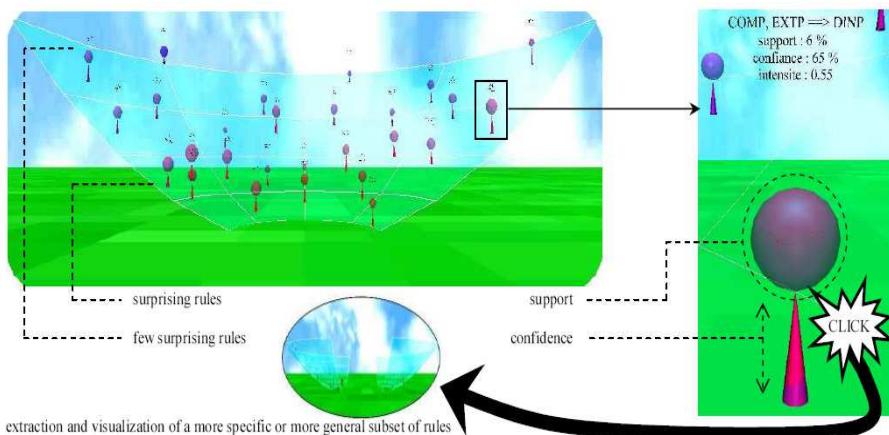


Fig. 11. ArVis a tool for association rules visualization Blanchard et al. (2007)

In order to find relevant knowledge for decision-making, the user needs to rummage through the rules.

ArVis proposed by Blanchard et al. (2007) is a human-centred approach. This approach consists of letting the user navigate freely inside the large set of rules by focusing on successive limited subsets via a visual representation of the rules (Figure.11). In other words, the user gradually drives a series of visual local explorations according to his/her interest for the rules. This approach is original compared to other rule visualization methods (Couturier et al. (2007), Gordal & Demiriz (2006), Zhao & Liu (2005)). Moreover, ARVis generates the rules dynamically during exploration by the user. Thus, the user's guidance during association rule

System	Visual Representation	Interaction techniques	Interaction actions			Input-Output devices	3D/ VR	year
			Navigation	Selection and Manipulation	System control			
ARVis Blanchard et al. (2007) (Fig.11)	Geometric shape	Human-centered	Manual view point manipulation	Object selecting + virtual pointer	Graphical menus	Mouse + 2D screen	3D	2007
3D spatial data mining on document sets (Fig.6(a)) Götzelmann et al. (2007)	Virtual world	Visual navigation and visual manipulation	-	Object selection	Graphical menus	-	3D	2006

Table 5. 3D VDM tool summary for association rules KDD task

post-processing is also exploited during association rule mining to reduce the search space and avoid generating huge amounts of rules.

Götzelmann et al. (2007) proposed a VDM system to analyze error sources of complex technical devices. The aims of the proposed approach is to extract association rules from a set of documents that describe malfunctions and errors for complex technical devices, followed by a projection of the results on a corresponding 3D model. Domain experts can evaluate the results gained by the DM algorithm by exploring a 3D model interactively in order to find spatial relationships between different components of the product. 3D enables a flexible spatial mapping of the results of statistical analysis. The visualization of statistical data on their spatial reference object by modifying visual properties to encode data (Figure.6(a)) can reveal apriori unknown facts, which where hidden in the database. By interactively exploring the 3D model, unknown sources and correlations of failures can be discovered that rely on the spatial configuration of several components and the shape of complex geometric objects.

A detailed comparison of these techniques is presented in Table.5.

7.2.4 Combining several methods

The Haiku tool (Figure.3(b)) combines several DM methods: clustering, classification and association rules (Beale (2007)). In this tool, the use of 3D graphs allows the visualization of high-dimensional data in a comprehensible and compact representation. The interface provides a large set of 3D manipulation feature of the structure, such as zooming in and out, moving through the representation (flying), rotating, jumping to specific location, viewing data details, and defining an area of interest . The only downside is that the control is done using a mouse. A detailed presentation is shown in Table.6.

System	Visual Representation	Interaction techniques	Interaction actions			Input-Output devices	3D/ VR	year
			Navigation	Selection and Manipulation	System control			
Heiku Pryke & Beale (2005) (Fig.3(b))	Graph	Human-centred	Manual view point manipulation + target based	Object selection	-	Mouse + 2D screen	3D	2005

Table 6. 3D VDM tool combining several methods

8. Conclusion

A new classification of VDM tools composed of 3 dimensions: visual representations; interaction techniques; and DM tasks, has been presented along with a survey of visual representations and interaction techniques in VDM. We can see that most of the recent VDM tools still rely on interaction metaphors developed more than a decade ago, and do not take into account the new interaction metaphors and techniques offered by VR technology. It is questionable whether these classical visualization/interaction techniques are able to meet the demands of the ever-increasing mass of information, or whether we are losing ground because we still lack the possibilities to properly interact with the databases to extract relevant knowledge. Devising intuitive visual interactive representations for DM and providing real-time interaction and mapping techniques that are scalable to the huge size of many current databases, are some of the research challenges that need to be addressed. In answer to this challenge, Mackinlay (1986) proposes two essential criterias to evaluate data mapping by visual representation: expressiveness and effectiveness. Firstly, expressiveness criteria determine whether a visual representation can express the desired information. Secondly, effectiveness criteria determine whether a visual representation exploits the capabilities of the output medium and the human visual system. Although the criteria were discussed in a 2D-graphic context, they can be extended to 3D and VR visualization. Finally, VDM is inherently cooperative requiring many experts to coordinate their activities to make decisions. Thus, collaborative research visualization may help to improve VDM processes. For example, current technology provided by 3D collaborative virtual worlds for gaming and social interaction, may support new methods of KDD.

9. References

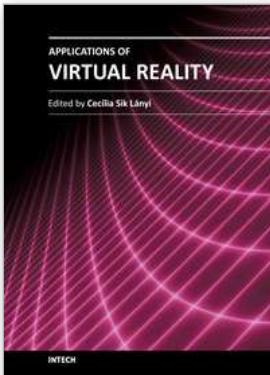
- Ahmed, A., Dwyer, T., Forster, M., Fu, X., Ho, J., Hong, S.-H., Koschutzki, D., Murray, C., Nikolov, N. S., Tarassov, R. T. A. & Xu, K. (2006). Geomi: Geometry for maximum insight, *Graph Drawing* 3843: 468–479.
- Aitsiselmi, Y. & Holliman, N. S. (2009). Using mental rotation to evaluate the benefits of stereoscopic displays, *Proceedings of SPIE, the International Society for Optical Engineering*, pp. 1–12.
- Ammoura, A., Zaâane, O. R. & Ji, Y. (2001). Immersed visual data mining: Walking the walk, *Proceedings of the 18th British National Conference on Databases*, pp. 202–218.

- Ankerst, M. (2001). *Visual Data Mining*, PhD thesis, Institute for Computer Science Database and Information Systems, University of Munich.
- Arns, L. L. (2002). *A new taxonomy for locomotion in virtual environments*, PhD thesis, Iowa State University, USA.
- Asimov, D. (1985). The grand tour: a tool for viewing multidimensional data, *SIAM Journal on Scientific and Statistical Computing* 6(1): 128–143.
- Azzag, H., Picarougne, F., Guinot, C. & Venturini, G. (2005). Vrminer: a tool for multimedia databases mining with virtual reality, in J. Darmont & O. Boussaid (eds), *Processing and Managing Complex Data for Decision Support*, pp. 318–339.
- Baumgärtner, S., Ebert, A., Deller, M. & Agne, S. (2007). 2d meets 3d: a human-centered interface for visual data exploration, *Extended abstracts on Human factors in computing systems*, pp. 2273–2278.
- Beale, R. (2007). Supporting serendipity: Using ambient intelligence to augment user exploration for data mining and web browsing, *International Journal of Human-Computer Studies* 65(5): 421–433.
- Becker, B. (1997). Volume rendering for relational data, *Proceedings of the IEEE Symposium on Information Visualization*, pp. 87–91.
- Beilken, C. & Spenke, M. (1999). Interactive data mining with infozoom : the medical data set, *Workshop Notes on Discovery Challenge, at the 3rd European Conference on Principles and Practice of Knowledge Discovery in Databases*, pp. 49–54.
- Blanchard, J., Guillet, F. & Briand, H. (2007). Interactive visual exploration of association rules with rule-focusing methodology, *Knowledge and Information Systems* 13(1): 43–75.
- Bowman, D. A., Kruijff, E., LaViola, J. J. & Poupyrev, I. (2001). An introduction to 3-d user interface design, *Presence: Teleoper. Virtual Environ.* 10(1): 96–108.
- Brath, R., Peters, M. & Senior, R. (2005). Visualization for communication: The importance of aesthetic sizzle, *Proceedings of the 9th International Conference on Information Visualisation*, pp. 724–729.
- Bukauskas, L. & Böhlen, M. (2001). Observer relative data extraction, *In Proceedings of the International Workshop on Visual Data Mining*, pp. 1–2.
- Buntain, C. (2008). 3d ontology visualization in semantic search, *Proceedings of the 46th Annual Southeast Regional Conference*, pp. 204–208.
- Cai, Y., Stumpf, R., Wynne, T., Tomlinson, M., Chung, D. S. H., Boutonnier, X., Ihmig, M., Franco, R. & Bauernfeind, N. (2007). Visual transformation for interactive spatiotemporal data mining, *Knowledge and Information Systems* 13(2): 119–142.
- Card, S. K., Mackinlay, J. D. & Schneiderman, B. (1999). *Readings in information visualization : using vision to think*, Morgan Kaufmann publishers.
- Card, S. K., Robertson, G. G. & York, W. (1996). The webbook and the web forager: an information workspace for the world-wide web, *Proceedings of the SIGCHI conference on Human factors in computing systems*, pp. 416–417.
- Carswell, C. M., Frankenberger, S. & Bernhard, D. (1991). Graphing in depth: Perspectives on the use of three-dimensional graphs to represent lower-dimensional data., *Behaviour & Information Technology*. 10(6): 459–474.
- Ceglar, A., Roddick, J. F. & Calder, P. (2003). *Managing data mining technologies in organizations*, IGI Publishing, chapter Guiding knowledge discovery through interactive data mining, pp. 45–87.
- Chi, E. (2000). A taxonomy of visualization techniques using the data statereference model, *Proceedings of IEEE Symposium on Information Visualization*, pp. 69–75.

- Chi, E. H. & Riedl, J. T. (1998). An operator interaction framework for visualization systems, *Proceedings of IEEE Symposium on Information Visualization*, pp. 63–70.
- Cleveland, W. S. & McGill, R. (1984). Graphical perception: Theory, experimentation, and application to the development of graphical methods., *Journal of the American Statistical Association* 79(387): 531–554.
- Couturier, O., Rouillard, J. J. L. & Chevrin, V. (2007). An interactive approach to display large sets of association rules, *Proceedings of the 2007 conference on Human interface*, pp. 258–267.
- Dachselt, R. & Hinz, M. (2005). Three-dimensional widgets revisited: towards future standardization, in K. Y. Bowman D, Froehlich B & S. W (eds), *New directions in 3D user interfaces*, pp. 89 – 92.
- de Oliveira, M. C. F. & Levkowitz, H. (2003). From visual data exploration to visual data mining: A survey, *Visualization and Computer Graphic* 9(3): 378–394.
- Eidenberger, H. (2004). Visual data mining, *SPIE Information Technology and Communication Symposium*, pp. 121–132.
- Einsfeld, K., Agne, S., Deller, M., Ebert, A., Klein, B. & Reuschling, C. (2006). Dynamic visualization and navigation of semantic virtual environments, *Proceedings of the conference on Information Visualization*, pp. 569–574.
- Einsfeld, K., Ebert, A. & Wolle, J. (2007). Hannah: A vivid and flexible 3d information visualization framework, *Proceedings of the 11th International Conference on Information Visualization*, pp. 720–725.
- Frawley, W. J., Piatetsky-Shapiro, G. & Matheus, C. J. (1992). Knowledge discovery in databases: An overview, *AI Magazine* 13(3): 57–70.
- Gordal & Demiriz, A. (2006). A framework for visualizing association mining results, *Lecture Notes in Computer Science* 4263/2006: 593–602.
- Götzelmann, T., Hartmann, K., Nürnberger, A. & Strothotte, T. (2007). 3d spatial data mining on document sets for the discovery of failure causes in complex technical devices, *Proceedings of the I2nd Int. Conf. on Computer Graphics Theory and Applications*, pp. 137–145.
- Gross, M. (1994). *Visual computing : the integration of computer graphics, visual perception and imaging*, Springer-Verlag.
- Hendley, R. J., Drew, N. S., Wood, A. M. & Beale, R. (1999). Narcissus: visualising information, *Proceedings of the IEEE Symposium on Information Visualization*, pp. 90 – 96.
- Herman, I., Melançon, G. & Marshall, M. S. (2000). Graph visualization and navigation in information visualization: A survey, *IEEE Transactions on Visualization and Computer Graphics* 6(1): 24–43.
- Hibbard, W., Levkowitz, H., Haswell, J., Rheingans, P. & Schroeder, F. (1995). Interaction in perceptually-based visualization, *Perceptual Issues in Visualization IFIP Series on Computer Graphics*: 23–32.
- Inselberg, A. & Dimsdale, B. (1990). Parallel coordinates: a tool for visualizing multi-dimensional geometry, *Proceedings of the 1st conference on Visualization*, pp. 361–378.
- Johnson, B. & Shneiderman, B. (1991). Tree-maps: a space-filling approach to the visualization of hierarchical information structures, *Proceedings of the 2nd conference on Visualization*, pp. 284–291.
- Kalawsky, R. & Simpkin, G. (2006). Automating the display of third person/stealth views of virtual environments, *Presence: Teleoper. Virtual Environ.* 15(6): 717–739.

- Keim, D. A. (2002). Information visualization and visual data mining, *IEEE Transactions on Visualization and Computer Graphics* 8(1): 1–8.
- Krohn, U. (1996). Vineta: navigation through virtual information spaces, *Proceedings of the workshop on Advanced visual interfaces*, pp. 49–58.
- Laviola, J. J. (2000). Msvt: A multimodal scientific visualization tool, *the 3ed IASTED International Conference on Computer Graphics and Imaging*, pp. 1–17.
- Mackinlay, J. (1986). Automating the design of graphical presentations of relational information, *ACM Transactions on Graphics* 5(2): 110–141.
- Maletic, J. I., Marcus, A., Dunlap, G. & Leigh, J. (2001). Visualizing object-oriented software in virtual reality, *Proceedings of the 9th International Workshop on Program Comprehension (IWPC'01)*, pp. 26–38.
- Meiguins, B. S., Melo, R., do Carmo, C., Almeida, L., Gonçalves, A. S., Pinheiro, S. C. V. & de Brito Garcia, M. (2006). Multidimensional information visualization using augmented reality, *Proceedings of ACM international conference on Virtual reality continuum and its applications*, pp. 391 – 394.
- Nagel, H. R., Granum, E., Bovbjerg, S. & Vittrup, M. (2008). *Visual Data Mining*, Springer-Verlag, chapter Immersive Visual Data Mining: The 3DVDM Approach, pp. 281–311.
- Nagel, H. R., Granum, E. & Musaeus, P. (2001). Methods for visual mining of data in virtual reality, *Proceedings of the International Workshop on Visual Data Mining in conjunction with 2nd European Conference on Machine Learning and 5th European Conference on Principles and Practice of Knowledge Discovery in Databases*, pp. 13–27.
- Nelson, L., Cook, D. & Cruz-Neira, C. (1999). Xgobi vs the c2: Results of an experiment comparing data visualization in a 3-d immersive virtual reality environment with a 2-d workstation display, *computational statistics* 14: 39–51.
- Niggemann, O. (2001). *Visual Data Mining of Graph-Based Data*, PhD thesis, Department of Mathematics and Computer Science of the University of Paderborn, Germany.
- Ogi, T., Tateyama, Y. & Sato, S. (2009). Visual data mining in immersive virtual environment based on 4k stereo images, *Proceedings of the 3rd International Conference on Virtual and Mixed Reality*, pp. 472 – 481.
- Osawa, K., Asai, N., Suzuki, M., Sugimoto, Y. & Saito, F. (2002). An immersive programming system: Ougi, *Proceedings of the 12th International Conference on Artificial Reality and Telexistence*, pp. 36 – 43.
- Parker, G., Franck, G. & Ware, C. (1998). Visualization of large nested graphs in 3d: Navigation and interaction, *Journal of Visual Languages & Computing* 9(3): 299–317.
- Pike, W. A., Staskob, J., Changc, R. & O'Connellld, T. A. (2009). The science of interaction, *Information Visualization* 8(4): 263–274.
- Plaisant, C., Grosjean, J. & Bederson, B. B. (2002). Spacetree: Supporting exploration in large node link tree, design evolution and empirical evaluation, *Proceedings of the IEEE Symposium on Information Visualization*, pp. 57–64.
- Poulet, F. & Do, T. N. (2008). *Visual Data Mining*, Springer-Verlag, chapter Interactive Decision Tree Construction for Interval and Taxonomical Data, pp. 123–135.
- Pryke, A. & Beale, R. (2005). Interactive comprehensible data mining, *Ambient Intelligence for Scientific Discovery* 3345/2005: 48–65.
- Robertson, G., Czerwinski, M., Larson, K., Robbins, D. C., Thiel, D. & van Dantzich, M. (1998). Data mountain: using spatial memory for document management, *Proceedings of the 11th annual ACM symposium on User interface software and technology*, pp. 153–162.

- Robertson, G. G., Mackinlay, J. D. & Card, S. K. (1991). Cone trees: animated 3d visualizations of hierarchical information, *Proceedings of the SIGCHI conference on Human factors in computing systems: Reaching through technology*, pp. 189 – 194.
- Shneiderman, B. (2003). Why not make interfaces better than 3d reality?, *IEEE Computer Graphics and Applications* 23(6): 12–15.
- Spence, I. (1990). Visual psychophysics of simple graphical elements., *Journal of Experimental Psychology: Human Perception and Performance*. 16(4): 683–692.
- Tavanti, M. & Lind, M. (2001). 2d vs 3d, implications on spatial memory, *Proceedings of the IEEE Symposium on Information Visualization*, p. 139.
- Teyseyre, A. R. & Campo, M. R. (2009). An overview of 3d software visualization, *IEEE Transactions on Visualization and Computer Graphics* 15(1): 87–105.
- Tory, M. & Moller, T. (2004). Rethinking visualization: A high-level taxonomy, *Proceedings of IEEE Symposium on Information Visualization*, pp. 151–158.
- Tufte, E. R. (1983). *The Visual Display of Quantitative Information*, Graphics Press.
- Van Ham, F. van Wijk, J. (2002). Beamtrees: compact visualization of large hierarchies, *Proceedings of IEEE Symposium on Information Visualization* pp. 93– 100.
- Wang, W., Wang, H., Dai, G. & Wang, H. (2006). Visualization of large hierarchical data by circle packing, *Proceedings of SIGCHI conference on Human Factors in computing systems*, pp. 517–520.
- Ware, C. & Franck, G. (1994). Viewing a graph in a virtual reality display is three times as good as a 2d diagram, *Proceedings of IEEE Visual Languages*, pp. 182–183.
- Ware, C. & Franck, G. (1996). Evaluating stereo and motion cues for visualizing information nets in three dimensions, *ACM Transactions on Graphics* 15(2): 121–140.
- Ware, C. & Mitchell, P. (2008). Visualizing graphs in three dimensions, *ACM Transactions on Applied Perception* 5(1): 1–15.
- Zhao, K. & Liu, B. (2005). *Opportunity Map: A Visualization Framework for Fast Identification of Actionable Knowledge*, PhD thesis, University of Illinois at Chicago.



Applications of Virtual Reality

Edited by Dr. Cecília Sík Lányi

ISBN 978-953-51-0583-1

Hard cover, 210 pages

Publisher InTech

Published online 02, May, 2012

Published in print edition May, 2012

Information Technology is growing rapidly. With the birth of high-resolution graphics, high-speed computing and user interaction devices Virtual Reality has emerged as a major new technology in the mid 90es, last century. Virtual Reality technology is currently used in a broad range of applications. The best known are games, movies, simulations, therapy. From a manufacturing standpoint, there are some attractive applications including training, education, collaborative work and learning. This book provides an up-to-date discussion of the current research in Virtual Reality and its applications. It describes the current Virtual Reality state-of-the-art and points out many areas where there is still work to be done. We have chosen certain areas to cover in this book, which we believe will have potential significant impact on Virtual Reality and its applications. This book provides a definitive resource for wide variety of people including academicians, designers, developers, educators, engineers, practitioners, researchers, and graduate students.

How to reference

In order to correctly reference this scholarly work, feel free to copy and paste the following:

Ben Said Zohra, Guillet Fabrice, Richard Paul, Blanchard Julien and Picarouge Fabien (2012). An Overview of Interaction Techniques and 3D Representations for Data Mining, Applications of Virtual Reality, Dr. Cecília Sík Lányi (Ed.), ISBN: 978-953-51-0583-1, InTech, Available from:
<http://www.intechopen.com/books/applications-of-virtual-reality/an-overview-of-interaction-techniques-and-virtual-reality-representations-for-data-mining>



InTech Europe

University Campus STeP Ri
Slavka Krautzeka 83/A
51000 Rijeka, Croatia
Phone: +385 (51) 770 447
Fax: +385 (51) 686 166
www.intechopen.com

InTech China

Unit 405, Office Block, Hotel Equatorial Shanghai
No.65, Yan An Road (West), Shanghai, 200040, China
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元
Phone: +86-21-62489820
Fax: +86-21-62489821

© 2012 The Author(s). Licensee IntechOpen. This is an open access article distributed under the terms of the [Creative Commons Attribution 3.0 License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.