A Taxonomy of Tasks for Guiding the Evaluation of Multidimensional Visualizations

Eliane R. A. Valiati
Instituto de Informática
Universidade Federal do Rio Grande do Sul
Caixa Postal 15064 91.501-970
Porto Alegre, RS, Brazil
evaliati@inf.ufrgs.br

Marcelo S. Pimenta
Instituto de Informática
Universidade Federal do Rio Grande do Sul
Caixa Postal 15064 91.501-970
Porto Alegre, RS, Brazil
mpimenta@inf.ufrgs.br

Carla M.D.S. Freitas
Instituto de Informática
Universidade Federal do Rio Grande do Sul
Caixa Postal 15064 91.501-970
Porto Alegre, RS, Brazil
carla@inf.ufrgs.br

ABSTRACT

The design of multidimensional visualization techniques is based on the assumption that a graphical representation of a large dataset can give more insight to a user, by providing him/her a more intuitive support in the process of exploiting data. When developing a visualization technique, the analytic and exploratory tasks that a user might need or want to perform on the data should guide the choice of the visual and interaction metaphors implemented by the technique. Usability testing of visualization techniques also needs the definition of users' tasks. The identification and understanding of the nature of the users' tasks in the process of acquiring knowledge from visual representations of data is a recent branch in information visualization research. Some works have proposed taxonomies to organize tasks that a visualization technique should support. This paper proposes a taxonomy of visualization tasks, based on existing taxonomies as well as on the observation of users performing exploratory tasks in a multidimensional data set using two different visualization techniques, Parallel Coordinates and RadViz. Different scenarios involving low-level tasks were estimated for the completion of some high-level tasks, and they were compared to the scenarios observed during the users' experiments.

Categories and Subject Descriptors

H.5.2 [**Information Interfaces and Presentation**]: User Interfaces – *Graphical user interfaces (GUI).*

General Terms

Algorithms, Design, Experimentation.

Keywords

Information visualization, usability evaluation.

1. INTRODUCTION

The increasing volume of information provided by several applications, different instruments and mainly the Web has lead to the development of techniques for selecting among a large bulk of data the subset of information that is relevant for a particular goal or need. Research on scientific data visualization, visual query systems, data mining and interactive visualization techniques has resulted in a wide variety of visual presentation and interaction techniques that can be applied in different situations.

Although there is a great variety of models and techniques for information visualization (see, for example, Card et al. [1]), each application requires a particular study in order to determine if the selected technique is useful and usable. The type of data that should be represented as well as the user tasks or analysis process that the visualization should help or support usually guides these studies. Whereas the first information visualization techniques were presented without thorough evaluation studies, in the last years researchers have become aware of the importance of such usability studies [2,3]. Almost all the evaluations are accomplished through experiments with users, which leads to the problem of defining the set of user tasks that should be part of the experiments, as well as providing different datasets to be tested. Moreover, the evaluation of information visualization techniques should target both the visual representation and the interaction mechanisms.

The identification and understanding of the nature of the users' tasks in the process of acquiring knowledge from visual representations of data is a recent branch in information visualization research. Some works have proposed taxonomies to organize tasks that a visualization technique should support (see next section). This paper proposes a taxonomy of visualization users' tasks, based on existing taxonomies as well as on the observation of users performing exploratory tasks in a multidimensional data set. Two case studies were conducted using two different visualization techniques, Parallel Coordinates [4] and RadViz [5]. Different scenarios involving low-level tasks were estimated for the completion of some high-level tasks, and they were compared to the scenarios observed during the users' experiments.

The paper is organized as follows. Next section summarizes some relevant works related with taxonomies of tasks common in information visualization techniques. Section 3 presents our taxonomy of visualization tasks. Then, in section 4, we describe some experiments carried out for evaluating our taxonomy and the results obtained. Finally, a discussion and final considerations are presented in the last section.

2. RELATED WORK

Usually, information visualization techniques provide a set of operations or interactive (visual) techniques that allow a user to accomplish specific-domain tasks. Despite the considerable number of studies on visualization techniques, only recently, reports were published devoted to understanding and representing the tasks users perform to accomplish their goals for evaluation purposes.

Wehrend and Lewis [6] and Springmeyer [7] in the early 90's were among the first ones to explicitly address user operations and tasks characterizing the data analysis process in order to facilitate the selection of adequate visual representations. Both authors characterized domain-independent tasks, which allowed generalizing their classification.

With the goal of facilitating the choice of visual representations, Wehrend and Lewis [6] classified operations that a user might need to execute to analyze data as:

- *Locate:* the user knows a dataset entry and indicate it by pointing or describing it.
- *Identify*: similar to locate but the user describe the dataset entry without knowing it previously.
- *Distinguish*: different objects should be presented as distinct visual items.
- *Categorize:* objects may be different because they belong to different categories, which should be described by the user.
- *Cluster:* the system may find out categories and objects belonging to them are shown linked or grouped together.
- *Distribution:* the user specifies categories and objects belonging to them are distributed among them.
- *Rank:* the user is asked to indicate the order of the objects displayed.
- *Compare:* the user is asked to compare entities based on their attributes.
- *Compare within and between relations:* the user is asked to compare similar entities or different sets of objects.
- *Associate:* the user is asked to establish relations between objects displayed.
- *Correlate:* the user may observe shared attributes between objects.

In that work, through the classification of typical visual representations based on an integrated analysis of these categorization and the characteristics of datasets, one could find out which visual representation would lead to a better solution for an application problem.

Shneiderman [8] proposed a task by data type taxonomy for information visualizations combining seven data types (1-dimensional, 2-dimensional, 3-dimensional, temporal, multidimensional, tree and network) and seven users´ tasks the visualizations should support. The tasks specified by Shneiderman were: overview, zoom, filter, details-on-demand, relate, history and extract, and are on the basis of what is known as the visual information seeking mantra.

Later on, Zhou and Feiner [9] introduced another categorization of tasks. They separated *presentation intents* (goals a user has when using a visual representation) from low-level *visual techniques* (the exact operation performed on a given object presented in the display) by means of an intermediate level, the *visual tasks*. *Visual tasks* can be considered abstract visual techniques, since they indicate a desired visual effect in the representation while a visual technique is a *way* to achieve that desired effect, either by the user or the system.

Zhou and Feiner characterize visual tasks along two dimensions: *visual accomplishments* and *visual implications*. Visual accomplishments correspond to the presentation intents a visual task is supposed to support while visual implications specify the visual techniques that could be used to fulfill the visual task. Regarding visual accomplishments, two classes of visual tasks can be identified: *inform* and *enable*. *Inform tasks* can be further distinguished as *Elaborate* and *Summarize* tasks, while *enable* tasks can be divided in *explore* tasks and *compute* tasks. At the bottom level of this hierarchy of abstract visual tasks one still has tasks like *categorize*, *cluster*, *compare*, *correlate*, *identify*, etc., i.e., generic operations like those identified by Weherend and Lewis [6]. For example, one of the *explore* tasks is the abstract *search*, which can be represented by the visual tasks *categorize*, *cluster, compare, correlate, distinguish, emphasize, identify, locate, rank, and reveal* [9].

Based on the observation of principles for visual perception and cognition, the same authors [9] also establish a link from the visual tasks to the adequate visual techniques. For example, to identify a piece of information, one can give its name, point at it in the display, give a range of attributes as a profile, all of these implying certain concrete visual tasks like name input, mouse pointing and filtering.

Morse et al. [10] developed a procedure for mapping from this visual taxonomy to concrete tasks represented by 50 questions in the information retrieval domain. They used subsets of these questions and simple visual prototypes to test the role of visualizations in that domain. By defining tests based on this taxonomy, they tried to exhaustively evaluate capabilities of visualizations.

When compared to perceptual operators [11, 12], which indicate the perceptual tasks a user accomplishes in a visual environment, visual tasks are significantly different. In order to achieve the presentation intents, perceptual operators emphasize what a user must do (*search, determine, verify, compare, look up, add, subtract*) while visual tasks describe the support a visual representation must provide for the accomplishment of user tasks.

Byrne et al. [13] created a taxonomy of user tasks for the web, based on the analyses of most frequent tasks performed by users while using web applications. This study describes patterns for user tasks in terms of six classes of sub-tasks, *use information, locate on page, go to page, provide information, configure browser* and *react to environment,* which can be employed to provide a comprehensive vocabulary for user activity in such application domain.

In a recent work, Amar and Stasko [14] discuss the notion of analytic gap, representing the obstacles presented by visualization systems in facilitating high-level analytical tasks, such as domain learning and decision making under uncertainty, which are not covered usually by the existing works in design and evaluation of information visualization systems.

The authors claim that, although Wehrend & Lewis´s and Zhou & Feiner´s low-level tasks are essential, they do not offer a consistent basis to fill the analytic gaps. Thus, they have proposed a new taxonomy [14], with higher level tasks that can provide a better support to visualization systems designers and evaluators. Limitations of the existing visualization systems were grouped into two major categories: the Rationale Gap and the Worldview Gap. The first one is defined as the gap between perceiving a relationship and actually being able to explain confidence in that relationship, and the usefulness of that relationship. In fact, users need to be able to relate data sets to the realms in which decisions are being made. The second one is defined as the gap between what is being shown and what actually needs to be shown to draw a straightforward representational conclusion for making a decision. In fact, users need to be able to formulate a strategy for browsing a visualization and for creating, acquiring and transferring knowledge or metadata about important domain parameters within a data set.

Then, each gap motivates the proposition of three high level tasks that visualization systems should support (although overlapping is possible):

a) rationale-based tasks: *expose uncertainty, concretize relationships, formulate cause and effect*; and

b) worldview-based tasks: *determination of domain parameters, multivariate explanation,* and *confirm hypotheses.*

In a very recent work, Amar et al. [15] proposed a taxonomy of 10 low level tasks based on 196 analytic questions found by students when analyzing data with commercial visualization systems.

The understanding and representation of tasks that a user performs while analyzing data analysis are essential for an effective evaluation of information visualization systems.

The relationships between tasks and goals are clear as described in Norman's theory of action [16]. The user's behavior during interaction with a system corresponds to a 7-stages cycle: the user has goals; formulates intents; verifies possible actions and selects the most appropriate one according to their intentions; executes the chosen action; perceives, interprets and evaluates system's results until completion of the task. Sometimes, a user goal can be mapped to only one task, but often it may require the coordinated execution of more than one task. Clearly, users' needs and goals must ideally be taken into account throughout design and development. Evaluation is the process responsible for validating: a) how a system cover efficiently users' goals and b) how users' tasks using (tasks of a) system meet the exact user goals in an effective, efficient, safe and satisfying way [17]. However, few authors explicitly explore the set of user tasks for evaluation purposes. In particular, we are interested in discussing how tasks understanding and representation (for example, the well-known HCI's task model) can be used for evaluation purposes.

In a previous work [18], we investigated how to model these tasks using a formal method and its corresponding environment to take advantage of the possibilities of automatically generating different scenarios that cover all the accomplishments and implications for each visual task.

Ideally, the evaluation of visualization techniques must be able to:

- Identify the user goals and verify if the user can reach them with an application which implements an information visualization technique;
- Identify which interaction mechanisms made available to the user by the visualization techniques are useful to accomplish the user task;
- Identify the graphical rendering functions that have been employed by the visualization techniques to show information;
- Relate user goals, interaction mechanisms and graphical rendering.

The requirements stated above need more precise users' tasks descriptions and our taxonomy is a step towards this goal.

3. TAXONOMY OF TASKS

This section presents the taxonomy of specific users' tasks we used to guide the selection of tasks of our experiment. The taxonomy was proposed to support the design of different scenarios for the evaluation of multidimensional visualization techniques. It integrates, at different levels, analytic tasks, cognitive tasks and operational tasks, that a user might need to accomplish when using a visualization technique, either for an exploratory analysis or for supporting or preceding a more conventional statistical analysis.

Figure 1 presents an overview of our taxonomy, which comprehends seven tasks: identify, determine, visualize, compare, infer, configure and locate. Five of these tasks can be considered as goals a user might have when using a visualization technique for either visually exploring or analyzing the data set through some statistics (identify, determine, compare, infer and locate). The other two tasks (visualize and configure) are typical intermediate level tasks that support the analytical ones. Events can be identified as marking the beginning and end of each task. The analytical tasks are very high level and, although one can not establish a hierarchy among them, their accomplishment might involve the execution of other ones. In the following paragraphs we explain each of these tasks.

Identify. Corresponds to any action of finding, discovering or estimating visually: clusters, through proximity, similarity, continuity or closed shapes; correlation; properties like values, dispersion, symmetrical or asymmetrical distribution; patterns; thresholds, similarities or differences, data dependency or independency, uncertainty and/or data variation. The identify task begins each time the user begins a new activity with the goal of finding, discovering or estimating some new information regarding the data. The task is considered completed when the user finds the information s/he is interested in or explicitly changes the current goal.

Determine. This task refers to the actions of calculating, defining or precisely indicating values like: mean; median (for asymmetrical distribution); variance, standard deviation and amplitude as measure of dispersion; percentile; sum; proportions; differences; correlation coefficients, probabilities and other statistics as for example, hypotheses test. This task begins each time a user needs to calculate a specific value, and consequently ends up when calculation is completed, or the user changes his/her goal.

Visualize. Each one of the previous tasks often requires the visualization of the data space as well as navigation through the data set. This implies to represent graphically all (or the desired) dimensions or data items. Also, sometimes the user needs to access metadata, which corresponds to information about the domain space. The Visualize task begins each time a user has established parameters for the graphical representation, proceeds with changes in those parameters as well as navigation and browsing. The task ends when the user changes his/her goal, passing from simple visualization to other analytic or support task.

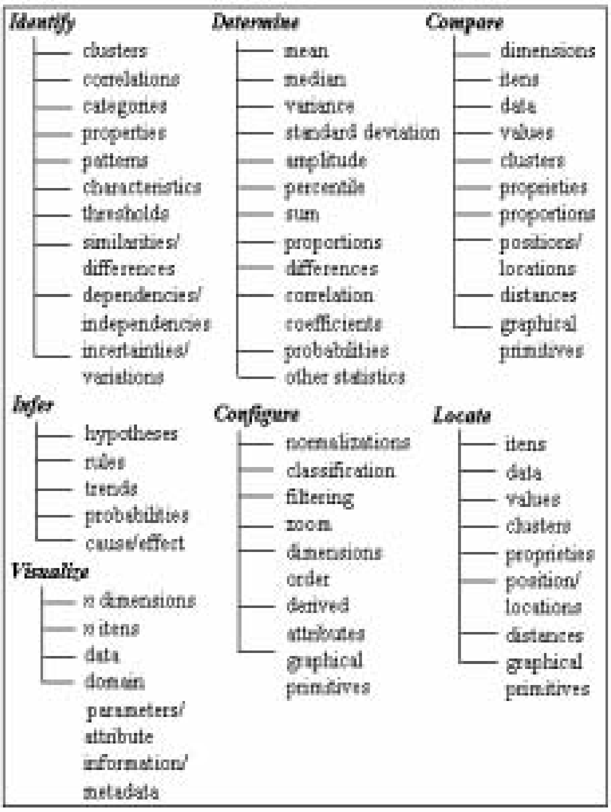

Figure 1: Taxonomy of user's tasks for interaction with multidimensional visualization systems

Compare. Once data have been identified, visualized, determined or located, the user can compare them by analyzing dimensions, data items, values, clusters, properties, proportions, locations and distances, as well as visual characteristics. The Compare task is an analytic task performed specifically if the user has to compare data items presented in the graphical representation.

Infer. Usually, after identifying, determining or comparing information, the user is able to infer knowledge from the information, defining hypotheses, rules, probabilities or trends, characteristics of cause and effect. This task is performed as part of the process of data analysis and may not be completed at once, requiring successive tasks of visualize, identify, determine, compare, etc. It ends up as soon as the user gets his/her insight or changes goal.

Configure. This is a support task in the sense that it is needed for the execution of the analytic ones. For visualizing the data space, the user usually has to configure the visual representation. This task is related to the possible actions available to normalize data, filter, classify, derive attributes, re-order dimensions, and change visual characteristics used to represent data items or attributes. This task begins with the user interacting with the system selecting options for visualization and ends when the user obtains the desired visualization.

Locate. This task is related to the actions of searching and finding (precisely in the display) information already visualized, identified or determined. These can be data items, values, clusters, distances, properties and visual characteristics. The task begins with the user examining the visual representation and finishes when he/she spots the precise position, situation or desired information (or changes his/her goal).

4. EVALUATION

With the goal of evaluating the proposed taxonomy, two case studies were conducted: the first study had Computer Science students in our laboratory as subjects, and served as a preliminary case study; in the second one, a biologist was the subject. We chose to report here the preliminary case study.

We chose two multidimensional visualization techniques [19] implemented within the same framework, the InfoVis toolkit [20], proposed a set of questions to be answered about a specific data set, and observed the sequence of subtasks performed by the users. We then compared scenarios we estimated for the solution of each question with the observed ones (performed by the subjects).

Users. Five graduate volunteered students (3 men and 2 women) participated as subjects in the evaluation experiments. Ages varied from 23 to 28 years, all them being M.Sc. candidates and having just finished an Information Visualization course at the Graduate Program in Computer Science, Federal University of Rio Grande do Sul.

Set of data. The classical data set containing information about American, Japanese and European cars[1] manufactured between 1970 and 1982 was used. This data set was selected due to the familiarity all the students would have with the domain, facilitating the understanding of questions as well as their accomplishment. Moreover, it has been used in many data mining and visualization systems for evaluation purposes (see [21] e [22]), making easier further comparison of results.

Tasks, scenarios and procedure. We examined the questions/tasks listed by other authors (see [21] e [22]) in the evaluation of visualization techniques using the same data set, and defined a minimum set of four questions trying to encompass different user actions. For each question we estimated scenarios, i.e., the sequence of possible tasks and subtasks needed for solving each question. We then recorded all the actions the users performed during the tests. Our goal was to observe if during the experiments different tasks (not included in our taxonomy) would arise.

The questions were: (1) When was the greater production of Japanese cars manufactured? (2) Analyze the data and describe the main characteristics of the American cars.(3) Are the Japanese cars with 4 cylinders generally heavier than American with 6 cylinders? (4) Which is the tendency shown by European cars during the last years in relation to their characteristics?

Results. The real (recorded) scenarios showed the tasks performed by each user in the resolution of each question using each technique.

[1] http://www.ics.uci.edu/AI/ML/MLDBRepository.html

It was verified that independently of the technique used, during the solution of each question, all users followed the same real scenarios with few variations. Moreover, comparing these real scenarios with the respective estimated ones (see Question 3's scenarios in Figure 2 as an example), one might observe that all estimated subtasks occurred although in a slightly different order, generally in relation to the first subtasks estimated in each scenario.

When observing the real scenarios, it was also noticed the occurrence of some subtasks in a third level, as well as a tendency of users in using different resources to configure the data presentation according to their preferences. As an example, in question 3 at the moment of locating the Japanese and American cars with different number of cylinders, the task Configure (that was not estimated) was performed as a subtask of Locate. This reflects the preference of each user in the use of different resources to configure data presentation (indicated in the real scenarios by square brackets, in the following way: [filters, colors, shapes, size, etc.]). The occurrence of a third level reinforces the fact that many tasks can be composed of subtasks, and there is no a priori hierarchy that can be imposed by the taxonomy.

More important in our observations was that no new task was identified besides those already defined in the taxonomy and estimated in the scenarios. We also observed that in our second case study, performed with a biologist analyzing his own data.

5. DISCUSSION AND FINAL COMMENTS

Multidimensional visualization techniques, as other interactive systems, are designed to help users to perform their tasks. However, to design information visualization techniques with better usability levels it is necessary to comprehend the tasks they intend to support. This work presented a taxonomy of specific tasks to guide the evaluation and design of multidimensional visualizations. We report the evaluation of two techniques (Parallel Coordinates and Radviz) through interaction tests, involving tasks and subtasks of the proposed taxonomy, comparing the observed scenarios with estimated ones in the solution of each question.

The taxonomy proposed, although is strongly based on other taxonomies, intends to be used specifically for multidimensional visualizations, taking into account the generic objectives that a user has when using such techniques to perform exploratory analyses as a previous step of statistical analysis. Other significant contribution of our work is the fact that the taxonomy was evaluated through two case studies with users.

However, although the results obtained in this work were significant, mainly in relation to the evaluation of taxonomy, it is necessary to consider that they were carried out involving few users and not all the tasks of the taxonomy were addressed (observed and evaluated) in the solution of the proposed tasks/questions. Thus, future work will address:

- studies involving a larger number of users, using the real scenarios obtained herein as estimated ones;
- other experiments using other questions that imply different tasks;
- a conformity inspection, guided by the tasks/questions proposed, with specific criteria defined for visualization techniques;
- analysis of the proposed taxonomy, carrying out interaction tests with the same techniques evaluated herein but involving different implementations;
- using the proposed taxonomy to guide other usability experiments with different multidimensional visualization techniques.

Question 3: Are the Japanese cars with 4 cylinders generally heavier than American with 6 cylinders?

Estimated scenario	Real scenario
Infer (tendency in relation to kind of car, weight and n° of cylinders)	**Infer (tendency in relation to kind of cars, weigh and n° of cylinders)**
Visualize (cylinders dimension)	Locate (Japanese cars with 4 cylinders) Configure (Japanese cars visualization), using [filters, colors, shape, size, etc.]
Visualize (weight dimension)	Locate (American cars with 6 cylinders) Configurate (American cars visualization), using [filters, colors, shape, size, etc.]
Locate (Japanese cars with 4 cylinders)	Visualize (cylinders dimension) Configurate (cylinders dimension visualization), using [filters, colors, shape, size, etc.]
Locate (American cars with 6 cylinders)	Visualize (weight dimension)
Identify (average weight of Japanese cars with 4 cylinders)	Identify (average weight of the Japanese cars with 4 cylinders)
Identify (average weight of the American cars with 6 cylinders)	Identify (average weight of the American cars with 6 cylinders)
Compare (average)	Compare (average)
Identify (correlation between weight and nr. of cylinders)	Identify (correlation between the weight and nr. of cylinders)

Figure 2. Comparison of estimated and real scenarios for Question 3

6. ACKNOWLEDGMENTS

This project has been sponsored by CNPq (National Agency for Scientific and Technological Development), SPIDER/WEB (CAPES/COFECUB project nr. 399/02/04) and EDGE (CNPQ/INRIA project nr. 490087/2005-1).

7. REFERENCES

[1] Card, S.K., Mackinlay, J.D. and Shneiderman, B., *Readings in Information Visualization - Using Visualization to Think*, Morgan Kaufmann, San Francisco, (1999).

[2] Chen, C., Czerwinski, M., Empirical studies of information visualization: A meta-analysis, *International Journal of Human-Computer Studies*, 53, 5 (2000), 851-866.

[3] Plaisant, C., The Challenge of Information Visualization Evaluation, *Proc. AVI´04 - ACM Conference on Advanced Visual Interfaces*, ACM Press (2004), 109-116.

[4] Inselberg, A., The Plane with Parallel Coordinates, Special Issue on Computational Geometry, *The Visual Computer*, 1, (1995), 69-91.

[5] Hoffman, P. E. et al, DNA Visual and Analytic Data Mining, *Proc. IEEE Visualization '97*, Phoenix, AZ, (1997), 437-441.

[6] Wehrend, S. and Lewis, C., A Problem-oriented Classification of Visualization Technique,. *Proc. IEEE Visualization'90*, IEEE Computer Society, (1990), 139-143.

[7] Springmeyer, R., Blattner, M. and Max, N.L., A characterization of the scientific data analysis process, *Proc. IEEE Visualization'92*, IEEE Computer Society, (1992), 235 – 242.

[8] Shneiderman, B., The Eyes Have It: A Task by Data Type Taxonomy for Information Visualization, *IEEE Symposium on Visual Languages*, 12, (1996), 326-343.

[9] Zhou, M. and Feiner, S. K., Visual Task Characterization for Automated Visual Discourse Synthesis, *Proc. CHI'98 Conference*, ACM Press, (1998), 392-399.

[10] Morse, E., Lewis, M. and Olsen, K., Evaluating visualizations: using a taxonomic guide. Int. J. Human-Computer Studies, 53 (2000), 637-662.

[11] Casner, S., A task-analytic approach to the automated design of graphic presentations, *ACM Trans. on Graphics*, 10, 2, (1991), 111-151.

[12] Ignatius, E., Senay, H. and Favre J., An intelligent system for task-specific visualization assistance, *Journal of Visual Language and Computing*, 5, (1994), 321-338.

[13] Byrne, M., John, B. E., Wehrle, N. S. and Crow, D. C., The Tangled Web We Wove: A Tasknomy of WWW Use, *Proc. CHI'99 Conference*, ACM Press, (1999), 544-551.

[14] Amar, R., Stasko, J., Knowledge Task-Based Framework for Design and Evaluation of Information Visualizations, *Proc. IEEE Symposium on Information Visualization (2004)*, 143-149.

[15] Amar, R., Eagan, J. and Stasko, J., Low-Level Components of Analytic Activity in Information Visualization. *Proc. IEEE Symposium on Information Visualization (2005)*, 111-147.

[16] Norman, D., Cognitive Engineering, *User Centred System Design: New Perspectives on Human-Computer Interaction*, Lawrence Erlbaum Associates, (1986), 31-61.

[17] Preece, J., Rogers, Y. and Sharp, H., *Interaction Design: Beyond Human-Computer Interaction*, John Wiley & Sons, New York, (2002).

[18] Winckler, M., Freitas, C. M. D. S. and Palanque, P., Tasks and Scenario-based Evaluation of Information Visualization Techniques, *Proc. 3rd International Workshop on Task Models and Diagrams for User Interface Design - TAMODIA'2004*, Prague, Czech Republic, (2004).

[19] Pillat, R.M. and Freitas, C.M.D.S., Coordinating Views in the InfoVis Toolkit, *Proc. of Advanced Visual Interfaces (AVI 2006)*.

[20] Fekete, J. D., The InfoVis Toolkit, *Proc. IEEE Symposium on Information Visualization (2004)*, 167-174.

[21] Kobsa, A., An Empirical Comparison of Three Commercial Information Visualization Systems. *Proc. IEEE Symposium on Information Visualization (2001).*

[22] Grinstein, G.G. et al., Benchmark Development for the Evaluation of Visualization for Data Mining, in *Information Visualization in Data Mininig and Knowledge Discovery*, Fayyad, U., Grinstein, G.G. and Wierse, A.(Eds.), Academic Press, San Francisco, (2002).