MULTI-MODAL WEB-BROWSING
An Empirical Approach to Improve the Browsing Process of Internet Retrieved Results

Dimitrios Rigas and Antonio Ciuffreda
Department of Computing, University of Bradford, Bradford, UK

Keywords: Internet, multimedia, graphics, browsing, earcons, musical stimuli.

Abstract: This paper describes a survey and an experiment which were carried out to measure some usability aspects of a multi-modal interface to browse documents retrieved from the Internet. An experimental platform, called AVBRO, was developed in order to be used as basis for the experiments. This study investigates the use of audio-visual stimuli as part of a multi-modal interface to communicate the results retrieved from Internet. The experiments were based on a set of Internet queries performed by an experimental and a control group of users. The experimental group of users performed Internet-based search operations using the AVBRO platform, and the control group using the Google search engine. On overall the users in the experimental group performed better than the ones in the control group. This was particular evident when users had to perform complex search queries with a large number of keywords (e.g. 4 to 5). The results of the experiments demonstrated that the experimental group, aided by the AVBRO platform, provided additional feedback about documents retrieved and helped users to access the desired information by visiting fewer web pages and in effect improved usability of browsing documents. A number of conclusions regarding issues of presentation and combination of different modalities were identified.

1 INTRODUCTION

Today most of the popular web search engines display information of retrieved documents in a textual format. The documents are displayed in a relevance-based order and consist of a title, a brief description and a URL. Although this approach used by most of the search engines is fairly simple and efficient, different problems still exist. For example, many users may struggle to read, analyse and visit the retrieved documents. This becomes particularly a problem when there is a large amount of textual entries displayed in the interface. A browsing process of this nature may frustrate users if it is carried out on a frequent basis. Common user activities involve scrolling up and down to locate desired textual entries and repeated visiting of links.

As the Internet becomes bigger and users search queries becomes complex, the results obtained increasingly become difficult to browse. Therefore the way the results are communicated to users becomes critical and additional user feedback might help users to locate the desired document easier. Many search engines are trying to improve the way the retrieved documents and links are presented and often additional information for these documents is provided.

However the approach is mainly textual and using a single sensory channel. The use of a single channel limits the scope for improvements.

In an ideal scenario, users should be able to browse as many results as possible using as fewer keystrokes or more movements as possible. Some of the experiments described in this paper address some of these issues by incorporating multimodal graphs that use of visual (e.g., graphics and colour) and auditory stimuli (e.g., earcons).

2 PREVIOUS WORK

A different number of research studies have been done in the past years for providing alternative approaches to text-based interfaces of Internet search engines. Most of these works are based on the development of two-dimensional or three-dimensional graphical objects aimed at displaying retrieved documents from search queries. Other

Rigas D. and Ciuffreda A. (2006).
MULTI-MODAL WEB-BROWSING - An Empirical Approach to Improve the Browsing Process of Internet Retrieved Results.
In *Proceedings of the International Conference on Signal Processing and Multimedia Applications*, pages 269-276
DOI: 10.5220/0001568402690276

studies have also investigated the use of auditory stimuli (e.g., speech and non-speech) to communicate different types of information in user interfaces. The sections below briefly review some of those studies.

2.1 The Use of Graphics

Several research studies have investigated the development of three-dimensional or two-dimensional graphs for browsing Internet and databases search results. The use of three-dimensional graphs on interfaces has been considered in different experiments carried out in recent years. Periscope (Wiza, 2004), based on AVE technology (Wiza, 2003) makes use of a series of three-dimensional Interface Models of holistic, analytical or hybrid nature. Periscope allows documents to be displayed at different levels of abstractions and therefore improving the visualization of documents. A similar approach has been applied on other systems, such as Cat-a-Cone and NIRVE. Cat-a-Cone (Hearst, 1997). These systems use a single three-dimensional graph based on a tree shape in order to organize the collection of retrieved information according to their categories. These techniques enable users to easily scroll the tree branches in order to browse and locate the desired documents. NIRVE (Sebrechts, 1999) is an original platform that uses a set of advanced three-dimensional graphs to display the entire set of retrieved documents. The interesting feature of NIRVE is on the organised arrangement of documents according to clusters (Cugini, 2000). Three-dimensional graphs are excellent tools for data visualization but they are often very complex. The lack of simplicity in these graphs is one of the common problems in these platforms. This is often the result of the overexploitation of the visual channel and in some cases to inappropriate usage of visual metaphors.

Many experiments have also been performed using two-dimensional graphs. Kartoo (Kartoo) is an innovative search engine in which retrieved results are communicated to users as icons. The retrieved documents are scattered in an interactive map together with suggested words. The correlations of the retrieved results and keywords are displayed using bonds between the icons. Another application called Insyder (Reiterer, 2000), (Mann, 2000) used synchronized multiple views. These allowed users to browse retrieved documents using a variety of two-dimensional graphs. Another application, called Envision (Heath, 1995), (Wang, 2002), allowed users to browse documents using a specific two-dimensional graph. The two axes of the graph communicated different attributes of the documents. Documents were communicated as icons and grouped in the graph in ellipses that formed clusters. A different application called VQuery (Jones, 1998a), (Jones, 1998b) was developed for libraries and utilised the Venn diagram methodology to display documents according to the query terms they contained using a Boolean approach.

Most of the visual metaphors used in several two-dimensional graph-based applications appear to be effective as they incorporate intuitive and easy to read and understand graphs. These graphs communicate small volumes of data and so they do not visually overload the user. These simple and well-understood displays, however, limit considerably the amount of data that can be presented to the user at any given time during the interaction process.

2.2 The Use of Sound

A series of experimental studies have demonstrated the successful application of non-speech and speech sound as a means to communicate information. Non-speech sound can be broadly divided into earcons, auditory icons and special sound effects. Auditory icons (Gaver, 1986) are short sounds that we hear in our everyday life. They are also often referred to as ‘environmental sounds’. The use of auditory icons (Gaver, 1993) was implemented in an application called SonicFinder.

Earcons are another form of auditory stimuli. They are short musical messages that have attracted research attention in recent years. Earcons have been used in graphical interfaces for communicating information (Brewster, 1993), (Brewster, 1994), (Rigas, 1996). Earcons have also demonstrated to be useful in interfaces for visually impaired users (Alty, 2005), (Edwards, 1987), (Rigas, 1997) (Rigas, 2005).

Text-To-Speech technology (Duggan, 2003) has been widely adopted in a wide range of applications. The success of this type of synthetic speech can be explained by the technical improvements made into the naturalness and fluency of the sound. The use of recorded speech can communicate a valid tool for delivering information, thanks to the level of naturalness and intelligibility that can be obtained. The main drawback of this type of speech sound can be found to the manual recording needed for each single instruction. For this reason this technology is mainly used in applications where the number of instructions is limited.

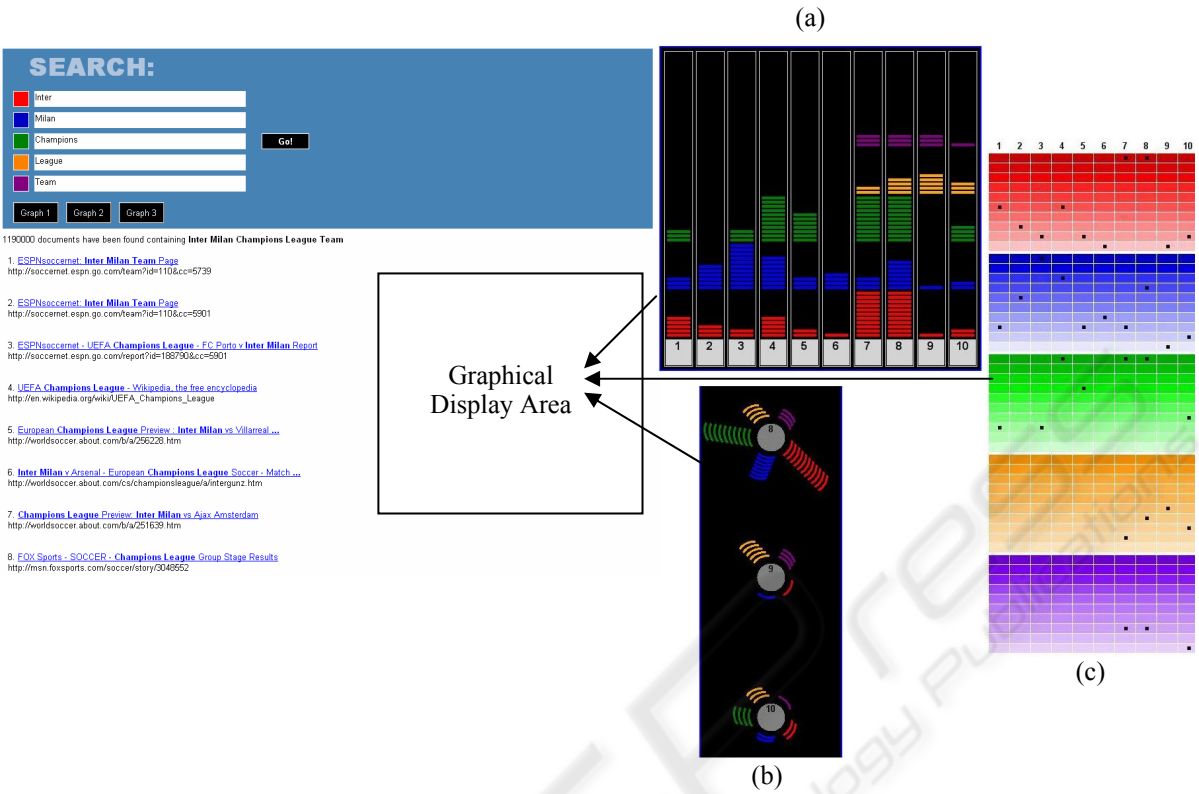

Figure 1: A visual illustration of the AVBRO experimental platform with three examples of the multi-modal graphs used.

3 AVBRO: AN EXPERIMENTAL AUDIO-VISUAL PLATFORM

An experimental browsing platform, called AVBRO (Audio Visual BROwser), has been developed in order to serve as a basis for this experimental programme. The functionality of this platform is mainly focused on typical search engine browsing activities. The aim is to investigate usability aspects of a multi-modal approach in search engines for delivering a larger but usable amount of information related to the retrieved results from Internet queries. The term usable also refers to the meaningfulness of the information provided to the users in order to make their decisions regarding the suitable document or link to follow.

AVBRO uses Google API technology (Google Web APIs) and the Goggle search engine to perform query operations over the Internet and obtain the required documents or links. For each retrieved result, AVBRO counts the occurrences of each user provided keyword within the document. The results are communicated in a multi-modal context consisting of graphs and earcons. These metaphors have been designed within the AVBRO platform to help to communicate the information (i.e., documents or links obtained as a result of an internet search) and they are evaluated in terms of their usability and usefulness for user browsing.

Figure 1 shows the visual interfaces of the AVBRO experimental platform. The initial page of the platform provides an interface with five different text fields. Each text field is associated with a specific colour and a specific musical instrument. Users need to enter one keyword for each of the five text fields. The results page is divided into two areas. In the left area of the interface, the retrieved results are displayed textually like in any other search engine. Each result entry consists of a title, a description and the URL. In the right part of the interface, a multi-modal graph is presented. The graphs also have auditory output. By clicking with the mouse to a specific area of the graph, a sequence of notes using different musical instrument is played in a rising pitch order. For example, if a keyword was encountered once within a document, a single note using a specific timbre would be played.

3.1 Graphical Modalities Used

The interface of the AVBRO platform allows users to choose three different two-dimensional multi-

modal graphs. The purpose of these graphs was to communicate the frequency of occurrences of each query word for each retrieved document which is displayed in the result page.

Figure 1 show the three graphs used. The first two graphs (a) (b) communicate each document in a particular shaped object, where lines (communicating occurrences) of specific colours (communicating the keyword used) are displayed within it. Each object displayed is numbered according to the ranked order number of the document it communicates.

The third graph (c) is shaped similar to a table where the numbered columns communicate documents. Frequency of query words are communicated by the location of black spherical object within coloured cells inside the columns, where the colour of the cell communicates the query word used.

3.2 Sound Used

As previously stated, each textfield included in the search page is associated with a different musical instrument. The musical instruments chosen for this platform are the piano, the organ, the saxophone, the drum and the guitar. The graphs allow a set of rising pitch musical notes from these instruments to be played when the user clicked over the area of the displayed graph communicating a specific document. The number of played notes communicates the number of occurrences (up to 10) of the corresponding query word.

4 INITIAL SURVEY

The starting point of this investigation was an initial survey. A short pre-experimental questionnaire was handed to users in order to gather information regarding their qualifications, views, experience and their computer background.

More specifically, the questionnaire aimed to identify the level of their browsing expertise and the number of keywords often used to search the internet, their internet searching habits and their opinions on multimodality for browsing activities.

The questionnaire also included questions about the use of multimedia tools for searching and browsing the internet in order to collect users' opinions and previous experience with multimedia platforms.

The results of this survey demonstrated that most of the users regularly used more than one keyword and that the number of pages accessed for each search was most of the times greater than one.

The results of this pre-experimental questionnaire also revealed that no one has ever had any extensive experience with the use of multimodality for web searching purpose. This finding demonstrates that multi-modal browsing is not widely used for browsing internet based retrieved results. Most of users agreed that sound and graphics could actually communicate information of retrieved search results.

5 EXPERIMENTS

On completion of the initial survey, a set of experiments have been carried out to test the functionality and usability of these multi-modal graphs in terms of successful communication of information to the users, the attitude of users and the overall validity of the approach taken.

Users had to perform eight tasks. All tasks involved search operations over the Internet with a number of keywords ranging from one to five. The eight tasks were divided in three different levels of difficulty. These included two *simple tasks* with one to two keywords, three *intermediate tasks* with three to four keywords and three *complex tasks* with five keywords.

A total of sixty users took part in these experiments. The users were all students from the University of Bradford and had different levels of Internet browsing and computer skills.

The simple and intermediate tasks were performed by a total of thirty users. These users were divided into two groups of fifteen each. An experimental group that used the AVBRO platform and a control group that used the Google search engine. The complex tasks were performed by another thirty users who were again divided in the same way into two groups of fifteen users for each group.

In both groups, users were requested to use the graphical objects displayed and the musical stimuli played. After performing each of the search operations users were required to enter the number of pages accessed in order to complete the task and obtain the required information.

In addition to the number of pages accessed, users in the experimental group were also required to note for each task performed the level of their perceived difficulty and the perceived usefulness of the multimodal graphs used. Users from the control group were required to note the level of usefulness of the text used for displaying information in addition to the level of their perceived difficulty of the task performed.

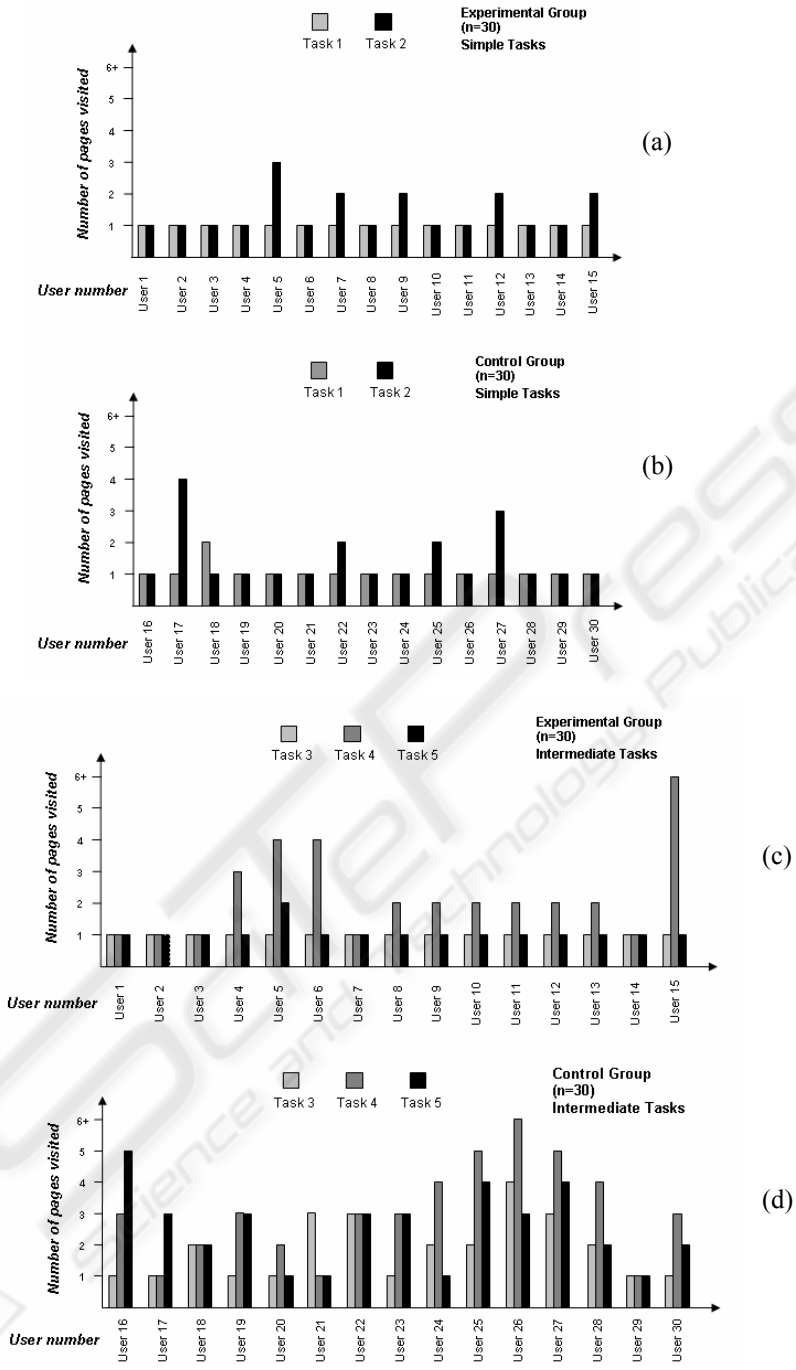

Figure 2: The results of the easy (a, b) and intermediate tasks (c, d) for the control and experimental group.

Figure 2 shows the results of the simple and intermediate tasks for the control and experimental groups. It can be seen that there is a small difference for the simple tasks between control and experimental groups. However, there is a noticeable difference for intermediate and complex tasks. The results of complex tasks for the control and experimental groups can be seen in Figure 3. It can be seen that the number of pages visited by users is typically double in the control group when compared to the experimental group.

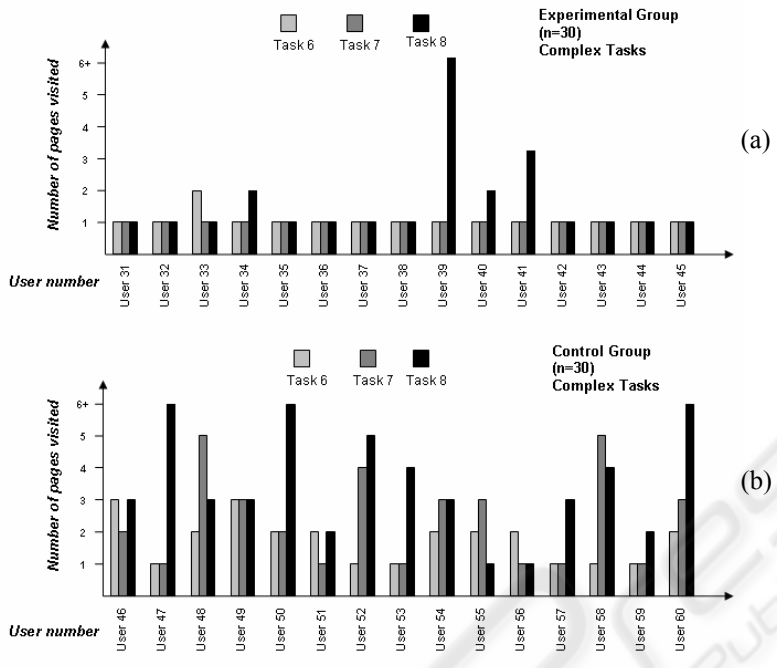

Figure 3: The results of the complex tasks for the experimental (a) and control (b) groups.

6 DISCUSSION

In these experiments the functionality of the multimedia-based interface was evaluated and compared with the functionality of a normal interface for search engines.

The mean number of the pages accessed was 1.2 for the experimental and 1.26 for the control groups in order to complete the two simple tasks, 1.44 for the experimental and 2.48 for the three intermediate tasks and 1.22 for the experimental and 2.53 for the control groups for the three complex tasks.

The mean numbers and standard deviation of all the tasks performed for the control and experimental groups can be seen in Figure 4. Statistical calculations of the T-test demonstrated that the two simple tasks and one intermediate level task were not significant but the remaining five tasks (intermediate and complex) were statistically significant. By using a critical value of 2.763, the T-Tests were for Task 1: T = 1, for Task 2: T = 0.22, for Task 3: T = 3.38, for Task 4: T = 1.48, for Task 5: T = 4.46, for Task 6: T = 3.56, for Task 7: T = 3.72 and for Task 8: T = 3.33.

The analysis of the results gave good insights about the continuation of this research programme. Examples are the use of additional visual and graphical metaphors that communicate different types of information in a non redundant way and the inclusion of speech and non-speech sound.

The experiments clearly demonstrated that the use of multimodal graphs helped users to browse retrieved results. This was particularly the case for intermediate to complex queries. The use of non-speech sound was not tested on its own in the experiments. It was observed that users during the browsing process assigned a secondary role to the non-speech sound and their browsing decisions were based on the graphs. This user behavior is attributed to the fact that both graphs and non-speech sound communicated the same information redundantly. It is believed that utilisation of auditory and visual stimuli in a non redundant way will enable to communicate a larger volume of information at any given time during the browsing process. Auditory stimuli could also involve the use of auditory icons and speech.

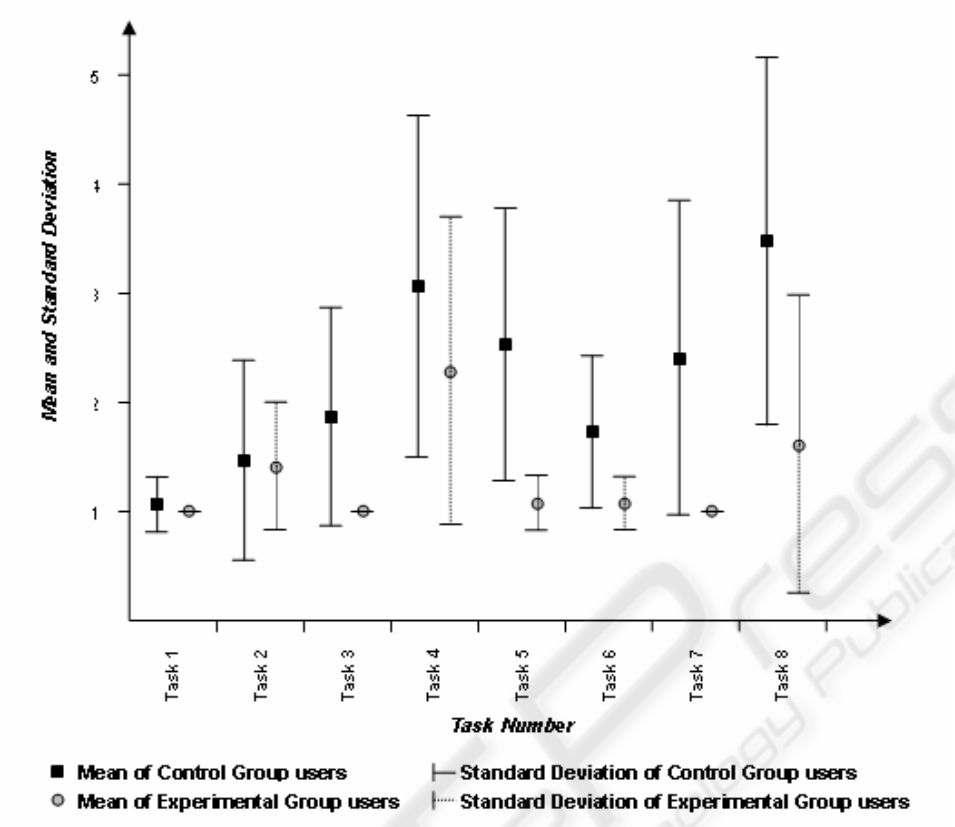

Figure 4: The mean and standard deviation values of pages visited for the entire set of tasks performed by users in the experimental and control group during the experiments.

7 CONCLUSION

In this paper, the AVBRO platform and experiments performed under this platform were described. The experimental platform utilised multimodal metaphors in addition to text to communicate internet based retrieved results. The experiments performed involved two experimental and two control groups in order to measure any difference in user performance. The experimental results demonstrated that the users were particularly helped when they had to perform tasks ranging from intermediate to complex. The level of complexity refers to the queries performed. No significant improvement was observed when users had to perform simple tasks.

The results of these experiments demonstrated that the use of audio and visual metaphors for providing additional feedback about documents retrieved can in most cases significantly help users to access the desired information. Additionally the experiments demonstrated a positive attitude by users towards some of the multimedia features offered in the experimental platform, proving the validity of the multimodal graphs approach and the need of further research in this direction. Currently, a series of further experiments are performed in the light of the results of these experiments.

REFERENCES

Alty, J. L., Rigas, D., 2005. Exploring the Use of Structured Musical Stimuli to Communicate Simple Diagrams: The Role of Context. In International Journal of Human-Computer Studies, Elsevier.

Brewster, S. A., Wright, P. C., Edwards, A. D. N., 1993. An Evaluation Of Earcons For Use In Auditory Human-Computer Interfaces. In *Proceedings of InterCHI'93*. ACM Press.

Brewster, S. A., 1994. Providing A Structured Method For Integrating Non-Speech Audio Into Human-Computer Interfaces. *PhD Thesis*, University of York, UK.

Cugini, J., Laskowski, S. J., Sebrechts, M. M., 2000. Design of 3-D Visualization of Search Results - Evolution And Evaluation.
http://www.itl.nist.gov/iaui/vvrg/cugini/uicd/nirve-paper.ps.gz

Duggan, B., Deegan, M., 2003. Consideration In The Usage Of Text To Speech (TTS) In The creation Of Natural Sounding Voice Enabled Web Systems. In *Proceedings of the 1st International Symposium on Information and Communication Technologies ISICT '03*. Trinity College Dublin.

Edwards, A.D.N., 1987. Adapting user interfaces for visually disabled users. *PhD Thesis*, The Open UniversityYork, UK.

Gaver, W. W., 1993. Synthesizing Auditory Icons. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems.* ACM Press.

Gaver, W. W., 1986. Auditory Icons: Using Sound In Computer Interfaces. HumanComputer Interaction. 2, 167-177.

Google Web APIs*. http://www.google.com/apis.*

Heath, L. S., Hix, D., Nowell, L. T., Wake, W. C., Averboch, G. A., Labow, E., Guyer, S. A., Brueni, D. J., France, R. K., Dalal, K., Fox, E. A., 1995. Envision: A User-Centered Database Of Computer Science Literature. In *Communications of the ACM,* ACM Press.

Hearst, M. A., Karadi, C., 1997. Cat-a-Cone: An Interactive Interface For Specifying Searches And Viewing Retrieval Results Using A Large Category Hierarchy. In *Proceedings of the 20th International ACM SIGIR Conference.* ACM Press.

Jones, S., 1998. Graphical Query Specification and Dynamic Result Previews For A Digital Library. In *Proceedings of the 11th Annual ACM Symposium on User Interface Software and Technology.* ACM Press.

Jones, S., 1998. Dynamic Query Result Previews For A Digital Library. In *Proceedings of the Third ACM Conference on Digital libraries.* ACM Press.

Kartoo, *http://www.kartoo.com.*

Mann, T. M., Reiterer, H., 2000. Evaluation Of Different Visualizations Of Web Search Results. In *Proceedings of the 11th International Workshop on Database and Expert Systems Applications.* IEEE Computer Society.

Reiterer, H., Mußler, G., Mann, T. M., Handschuh, S., 2000. INSYDER — An Information Assistant For Business Intelligence. In *Proceedings of the 23rd Annual International ACM SIGIR Conference on Research and Development in Information Retrieval.* ACM Press.

Rigas, D. I., 1996. Guidelines For Auditory Interface Design: An Empirical Investigation. *PhD Thesis*, University of Loughborough, UK.

Rigas, D. I., Alty, J. L., 1997. The Use Of Music In A Graphical Interface For The Visually Impaired. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM Press.

Rigas, D., Alty, J. L., 2005. The Rising Pitch Metaphor: An Empirical Study. In International Journal of Human Computer Studies. Elsevier.

Sebrechts, M. M., Cugini, J., Laskowski, S. J., Vasilakis, J., Miller, M. S., 1999. Visualization Of Search Results: A Comparative Evaluation Of Text, 2D, And 3D Interfaces. In *Proceedings of the 22nd Annual International ACM SIGIR Conference on Research and Development in Information Retrieval.* ACM Press.

Wang, J., Agrawal, A., Bazaza, A., Angle, S., Fox, E. A., North, C., 2002. Enhancing The ENVISION Interface For Digital Libraries. In *Proceedings of the 2nd ACM/IEEE-CS Joint Conference on Digital Libraries.* ACM Press.

Wiza, W., Walczak, W., Cellary, W., 2003. AVE-Method For 3D Visualization Of Search Results. In *Lecture Notes in Computer Science.* Springer-Verlag.

Wiza, W., Walczak, W., Cellary, W., 2004**.** Periscope: A System For Adaptive 3D Visualization Of Search Results. In *Proceedings of the Ninth International Conference on 3D Web Technology***.** ACM Press.