Is a virtual environment feasible to support knowledge sharing?

Luis Manuel Borges Gouveia
lmbg@ufp.pt

Abstract: The paper discusses the design and use of a system that allows the creation of a virtual environment to assist knowledge sharing. Knowledge sharing is based on the manipulation of a concept space, which can be individually or collaborativelly built and refined. The system proposes a 3D interactive visualisation interface that has been evaluated using a prototype to support collaborative learning.

Index Terms-- visualisation, knowledge sharing, information discovery, collaborative learning

I. INTRODUCTION

There seems to be general agreement that 3D visualisation offers a more convenient and natural way for people to interact with information spaces (as distinct from environments that are naturally 3D) [Tufte, 1990] and [Benedikt, 1992]. Work of the type reported here is important in order to find out whether this is true and how [Gouveia, 1999]. We must consider that to date there is not much evidence to support it, other than in cases where the information has a natural spatial component [Hubbold et al., 1995]. A key problem for using a spatial layout which reflects potentially dynamically changing information is the user sense of position that can be lost if the layout changes [Ingram and Benford, 1995].

One application for testing the visualisation design is information discovery to support user efforts to find relevant information within a given knowledge domain [Gouveia, 1998] and [Li-Jen and Gaines, 1998]. In this case, providing a means for users setting up a context, a query generation tool and an Information Visualisation [Card et al., 1999] allows users to have context and information about a particular data source for analysis and comparison.

This application will serve as *proof-of-concept* that based on a given context, shared as a 3D interactive visualisation, users can be assisted to retrieve information and analyse it information discovery [Baeza-Yates and Ribeiro-Neto, 1990]. The integration between the shared visualisation and data source information is made possible by using an information visualisation integrated with the visualisation design.

A number of systems have already taken advantage of 3D facilities. Two examples are [Benford et al., 1995] who proposed a virtual environment to support the co-operative browsing and filtering of large document stores, and [Chevalier and Verlhac, 2000] who present a 3D graphical representation of search results. [Chen, 1999] provides a discussion on Information Visualisation and Virtual Environments, and [Hearst, 1999] discusses user interfaces for communication between human and information retrieval systems.

The evaluation was conducted using a prototype that implements [Gouveia, 1998]:

- a concept space as a 3D interactive visualisation;
- a visualisation design composed by two distinct visualisations: a concept space, representing the structure, and a criteria space that allows spatial positioning by specifying up to three criteria;
- data source integration by using an Information Visualisation within the criteria space visualisation;
- displaying of results using a search engine (the *AltaVista Search Personal eXtension 97*).

The rest of the paper is organised as follows: section 2 presents the goals of the system, section 3 describes the system, section 4 discusses the evaluation, section 5 analyses the results we obtained, and in section 6 we give our concluding remarks.

II. GOALS AND RATIONALE

The 3D interactive visualisation goal is to convey information about a structure for knowledge sharing. By knowledge sharing we mean activities that allow the dessimination of knowledge as the case of collaboration, repositories, training and context exploring [Clare and Detore, 2000]. The proposed application focus on how to represent a context considering a given knowledge theme. To test how this visualisation design could support knowledge sharing a system has been developed with one specific application in mind: give support to users in information discovery. The proposed system helps users to build their own queries by using a textual search engine based on information from the structure for knowledge sharing. It also allows the visualisation of data source

CEREM – Multimedia Resource Centre
University Fernando Pessoa
Pr 9 de Abril, 349 4249-004 Porto Portugal

information within the visualisation design and displaying of results using an HTML browser. The advantages of a tool like the one described are greater when data sources do not have an underlying structure and a query returns a vast amount of results as is the case of the World Wide Web. The tool is based on a shared interactive representation of a knowledge theme that can be used to construct queries and compare a data source with the domain representation, using a 3D interactive visualisation. To allow reuse of each user efforts in information retrieval, a basic support for collaboration is implemented within the system to share the knowledge domain representation and to enhance it.

III. DESCRIPTION OF THE SYSTEM

The system uses a 3D interactive visualisation based on a structured description of a domain based in concepts and weighted keywords the set of concepts is defined here as a *concept space*. In order to support sharing of these concept spaces, a user with proper rights can add new concepts and keywords, and modify the weights of the keywords. The tool implements a voting facility to support collaborative decision for each proposal.
Table 1 presents a partial structure for knowledge sharing with 5 concepts, defining the theme *Information Management*. Each concept is defined by a set of keywords that characterises it.

[1] **Techniques** List of keywords (name, weight) Indexing 0.6 Pull 0.4 Push 0.4 Reading 0.6 Scanning 0.4 Textmining 0.8	[4] **People** List of keywords (name, weight) interpretation 0.7 knowledge 0.7 people 1.0 skills 0.6
[2] **Information** List of keywords (name, weight) Architecture 0.6 Cost 0.6 Information 1.0 Management 0.8 Meaning 0.8 Overload 0.6 Quantity 0.6 Resources 0.6 System 0.8 Value 0.8	[5] **Transformation** List of keywords (name, weight) cost 0.5 optimisation 0.6 people 0.6 planning 0.6 strategy 0.6
[3] **Management** List of keywords (name, weight) Information 0.8 Management 1.0 Organisation 0.8 People 0.8 Planning 0.6 Procedures 0.7 Strategy 0.7	

Table 1: A partial example of a structure for knowledge sharing about *Information Management*

The exact number of keywords varies, as each concept can need more or less keywords to be described. Later, more keywords can be added to the concept. Each keyword consists of a word that can be used for searching, and an associated weight. The weight can also be modified later by users. The keyword weight is a value between 0.0 and 1.0 and can be understood as a membership value of the relation between the keyword and the concept. The sum of all weights for a given concept does not have to equal 1, since weights are not probabilities.
In this example, shown in Table 1, five concepts were defined, each one with a different number of keywords. Some of the keywords exist in more than one concept. For example, the *strategy* keyword exists in the concepts *Management* and *Transformation* with weights of 0.7 and 0.6 respectively. This may be used to detect a relationship between the two concepts based on the number of common keywords in each concept and on their weights.
The components of this structure (concepts, keywords and weights) were used to build a 3D interactive visualisation which also allows to visualise the relationships, if any, between the concepts. Users can more easily analyse the structure and its content by using the 3D interactive visualisation. The visualisation uses colour and virtual world style user navigation to take advantage of the visualisation design. Each concept is represented as a sphere with size and colour computed based on the concept keywords. The relationship between two concepts is visualised as colour coded lines. The colour code results from computing the keyword degree of similarity taking into account the common keywords and their weights in the two concepts. This visualisation is named a concept space. Figure 1 shows one possible concept space perspective.
The spatial position of each sphere is given by the user who proposes the new concept. After that, the spatial position remains constant. The constant spatial position for each concept (sphere) allows the construction of spatial references that can be recalled for the virtual world navigation. It also allows the creation of a visual image for the structure being shared and the opportunity to externalise the knowledge context for discussion and enhancement by a group of users.
Each user interacts with this shared visualisation the concept space and can produce a second visualisation from it for personal use, not to be shared. The second visualisation supports each user organisation of search and browse tactics allowing the creation of a criteria space using the existing concepts.

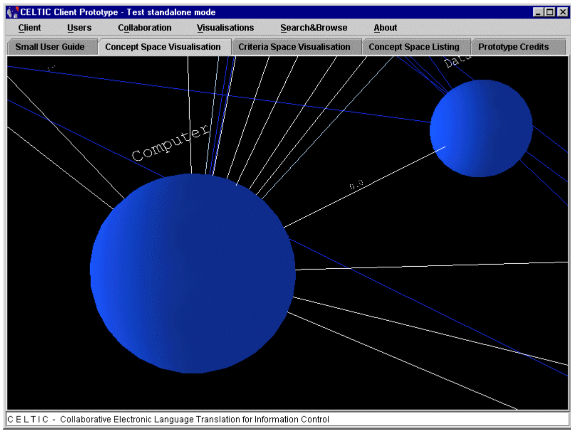

Figure 1: A perspective of the concept space visualisation

Figure 2 presents an example of a criteria space created by a user. As the criteria space is a three-dimensional space, the user can enter three criteria that are used to compute the spatial position of each concept based on its keyword weights.

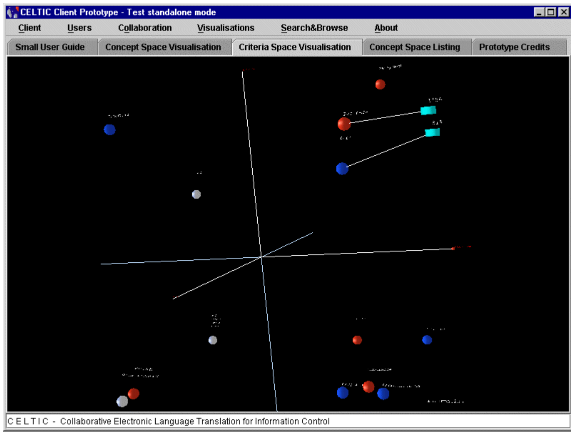

Figure 2: A perspective of a criteria space visualisation

The spatial position for each concept is calculated from comparing the criteria with existing keywords on the concept and using the keyword weight as a co-ordinate value for the criteria. If the criterion does not exist for a particular concept, a co-ordinate value of –1 is given to the concept for the corresponding criteria dimension. The resulting criteria space produces a visualisation of eight possible quadrants resulting from the three criteria combination of three dimensions.
In the criteria space visualisation, the placing of each concept depends on the weight of concept keywords. Each one of the eight quadrants represents a combination of the criteria no criteria, c1, c2, c3, c1 and c2, c2 and c3, c3 and c1, c1 and c2 and c3 being c1, c2 and c3, the first, second and third criteria.

The criteria space can be integrated with a data source. Using an Information Visualisation to compare concepts of the criteria space from the first quadrant the one that contains all the concepts where all three criteria exist. The Information Visualisation symbols are green cylinders linked to the related concepts by lines see figure 2. Figure 3 presents the results displayed in an HTML window after using the tool to generate the search query.

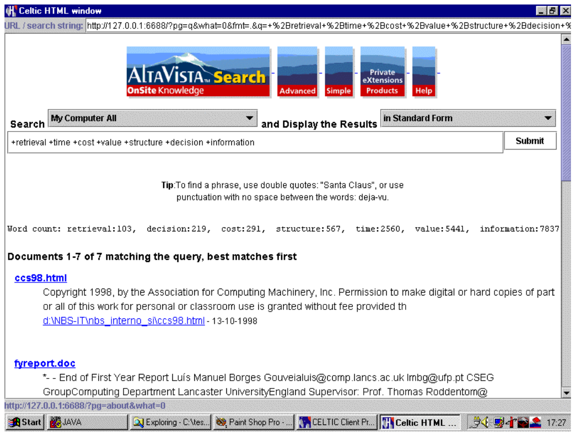

Figure 3: The HTML windows with query results

IV. EVALUATION

System variables are set by scripting commands or are determined by the information your enter when you set up a Dial-Up Networking connection. System variables are read-only, which means they cannot be changed within the script.
The evaluation was conducted using the approach defended by [Shneiderman, 1998] for virtual environments. Work from [Cohen et al., 2000] concerning education issues was also considered for the evaluation design. [Sebrechts et al., 1999] claim in their study that there were high interface costs for the visualisation of search results, although those costs decrease substantially with user experience. An evaluation study conducted to research Web search behaviour of Internet experts and newbies uses as performance evaluation factors web expertise and knowledge domain expertise [Holscher and Strube, 2000]. This study allows to verify that both factors were important although deficits in one or the two factors led to compensatory behaviour been knowledge expertise the most difficult to work since internet skills can with easy be trained for learning how to use search engines [Holscher and Strube, 2000].
To evaluate the proposed system we selected eleven undergraduate students from Fernando Pessoa University. The subjects were volunteers and no payment has been made for their participation. The knowledge domain was *Information Management*. The subjects were asked to use the prototype in the following six activities:

1. use the concept space and describe its meaning, by filling with words empty spaces in a set of sentences;

2. use the criteria space to relate the existing concepts giving the three criteria: *information*, *management* and *cost*;
3. choose the concept *Computer* and analyse its relation with other concepts;
4. try to create a criteria space where the *Knowledge* and *Enterprise* concepts would be related;
5. perform a search based on the *Information* concept;
6. perform a search based on the *management* criteria.

A. Evaluation script

The evaluation test was designed to have a one-hour and half and was composed of the following activities:

- subject fills a pre-experiment questionnaire (5 minutes);
- subject is given a general overview of the tool functionality (10 minutes demo);
- subject undertakes a lab training period (10 minutes);
- break (5 minutes);
- continuous session for performing the described six activities (50 minutes);
- subject fills a post-experiment questionnaire (10 minutes).

The test was repeated for each subject.

B. Evaluation factors

The evaluation was conducted based on a reduced number of variables. For organising data gathering a number of evaluation factors were considered:

- Asking each student:
  - what they have *learned* (as measured by a multiple-choice questionnaire);
  - how they think the system *helped* them (with a like/dislike rating);
  - what is their *opinion* about using the system(with a like/dislike rating);
- taking the *time to complete* of the six activities;
- performance is examined taking into account students opinion for rating as low or high their own expertise considering:
  - the Web expertise
  - the Knowledge domain expertise (*Information Management*)

C. Gathered data

Table 2 summarises the data collected during the evaluation test. The eleven subjects were considered and the values for each of the evaluation factors were collected into Table 2. At the bottom of each column, the value range of each of the evaluation factors is shown.

Subject	**Learn**	**Help**	**Opinion**	**Complete**	**time to complete**	**Web expertise**	**knowledge expertise**
1	7	4	6	6	34	5	5
2	5	5	5	6	41	4	4
3	6	2	4	6	44	7	4
4	6	6	3	6	32	7	5
5	9	6	6	6	39	6	6
6	8	7	5	6	41	5	6
7	7	4	6	4	50	4	5
8	3	4	2	5	45	3	3
9	6	5	5	6	30	6	3
10	5	6	4	5	50	2	2
11	4	1	3	4	50	2	2
	0 - 10	**1 - 7**	**1 – 7**	**1 - 6**	**0 - 50 mm**	**1 - 7**	**1 - 7**

Learn	ten test questions
Help	value 1-7 low/high
Opinion	value 1-7 low/high
Complete	1 to 6 situations
time to complete	minutes
web expertise	value 1-7 low/high
knowledge expertise	value 1-7 low/high

Table 2: Evaluation data

V. ANALYSING THE RESULTS

A. Statistical analysis

The first step is to consider *learn, help*, *opinion* and *time to complete* as the independent variables. The dependent variables are Web expertise (*webexp*) and knowledge expertise (*knowexp*). The normalised data table for statistical treatment is presented in Table 3.

The variable *learn* keeps its values between 1 and 10. The variables *help* and *opinion* were dichotomised. These value were dichotomised by grouping the values 1, 2, 3, and 4 converted to 0, and the values 5, 6, and 7 converted to 1. The *complete* variable was also dichotomised converting the values 4 and 5 to 0, and the value 6 to 1. In this case, the variable reports subjects that accomplished all the proposed tasks.

Dependent variables *web* and *knowledge* expertise where also dichotomised converting the values 1, 2, 3, and 4 to 1, and values 5, 6, and 7 to 2.

We used the Binomial model for *learn*, *help* and *opinion* variables. For the *time to complete* variable (taking into account the *complete* variable), the Cox proportional hazards model was selected.

The statistical software package was *Glim 4*, version 8 from the Royal Statistical Society, running in a Sun SPARCstation. The statistical tests were conducted with the help of the Centre for Applied Statistics at Lancaster University.

Subject	Learn	Help	Opinion	Complete	time to complete	Web expertise	knowledge expertise
1	7	0	1	1	34	2	2
2	5	1	1	1	41	1	1
3	6	0	0	1	44	2	1
4	6	1	0	1	32	2	2
5	9	1	1	1	39	2	2
6	8	1	1	1	41	2	2
7	7	0	1	0	50	1	2
8	3	0	0	0	45	1	1
9	6	1	1	1	30	2	1
10	5	1	0	0	50	1	1
11	4	0	0	0	50	1	1

Table 3: Normalised data

B. *General observations*

Using Table 3 for the statistical tests, the following observations can be made.

About the *learn* variable (test questionnaire):

- *web* expertise has significance at a 5% level;
- *knowledge* expertise has significance at a 1% level;
- both *web* and *knowledge* expertise are significant but with *knowledge* subject being more significance. No important interaction between both variables has been reported.

About the relation between *web* and *knowledge* experience:

- in the presence of *knowledge* expertise, the *web* expertise is no more significant at a 5% level;
- in the presence of web expertise, the *knowledge* subject expertise is approximately significant at a 5% level.

About the *help* variable (low/high help for the users):

- there is no evidence of meaningful effects with *web* and *knowledge* expertise;
- with both *web* and *knowledge* expertise together there is also no effects.

About the *opinion* variable (low/high help for the users):

- *web* expertise is not significant;
- *knowledge* expertise is approximately significant at a 10% level;
- with both *web* and *knowledge* expertise there are no effects.

About the *time to complete* variable (taking into account the subjects that have completed all the tasks):

- *web* expertise is significant at 1% level;
- *knowledge* expertise is significant at 5% level;
- both *web* and *knowledge* expertise do not have any relation (there's no reason to change the above statement).

VI. CONCLUDING REMARKS

Based on the statistical analysis it is possible to make several observations concerning the experiment.

People *learn* more when they had already some expertise in the *knowledge* area (in this case, *Information Management*). The importance of using the *web* before was moderate although not so important as the *knowledge* expertise to explain the questionnaire results (*learn*).

The users feeling about how the system *helps* them has not any impact from their *web* or *knowledge* expertise. When considering the user *opinion* about the system, then *knowledge* expertise seems to have some importance, regardless of the web expertise.

Operation of the system seems to be influenced by the users *web* expertise in a very important way. The *knowledge* expertise also assists users in system operation.

Overall, the system tends to be better supporting people with some *knowledge* expertise and little *web* expertise. This seems to show some potential as an interface to access information for people that have already some *knowledge* expertise. However more evaluation is needed in more knowledge domains and with more people using the system.

The use of visualisation techniques can improve the interface by supporting familiar cues to the user perception and thus convey information for knowledge sharing.

People were able to use the visualisation design which allowed them to take advantage of context information about a given knowledge theme (*Information Management*).

The proposed system shows that a user can take advantage of sharing knowledge to support information retrieval by representing it as a 3D interactive visualisation context.

REFERENCES

[1] Baeza-Yates, R. and Ribeiro-Neto, B. (1999). *Modern information retrieval.* Addison Wesley.

[2] Benedikt, M. (1992). Cyberspace: Some proposals. In Benedikt, Michael. Cyberspace, first steps. Third printing. MIT Press, pp 119-224.

[3] Benford S. and Snowdon D. and Greenhalgh, C. and Ingram, R. and Knox, I. and Brown, C. (1995). VR-Vibe: A virtual environment for co-operative information retrieval. Computer Graphics Forum, 14(3), pp C349-C360.

[4] Card, S. and Mackinlay, J. and Shneiderman, B. (1999). Readings in Information Visualization. Using Vision to Think. Morgan Kaufmann.

[5] Chen, C. (1999). Information Visualisation and Virtual Environments. Springer.

[6] Chevalier, M. and Verlhac, M. (2000). ISIDOR: A Visualisation Interface for Advanced Information Retrieval. Proceeding of ICEIS 2000, 2nd International Conference On Enterprise Information Systems, pp 414 - 418, 4-7 june, Stafford, UK.

[7] Clare, M. and Detore, A. (2000). Knowledge Assets. Harcout Brace.

[8] Cohen, L. and Manion, L. and Morrison, K. (2000). Research Methods in Education. 5^{th} Edition. Routledge Falmer.

[9] Gouveia, L. (1998). End of first year report. PhD first year assessment. CSEG Group, Computing Department. Lancaster University. Internal Report, September.

[10] Gouveia, L. (1999). Shared Visualisation and Virtual Environments for Co-operative Learning. ECSCW'99 conference. Doctoral Colloquium. 12-16 September. Conference proceedings supplement, pp. 70-72.

[11] Hearst, M. (1999) User Interfaces and Visualization. In (Baeza-Yates and Ribeiro-Neto, 1999). Modern Information Retrieval. Addison Wesley.

[12] Holscher, C. and Strube, G. (2000). Web search behavior of Internet Experts and newbies. 9th International World wide Web conference WWW9, Amsterdam.

[13] Hubbold, R. and Murta, A. and West, A. and Howard, T. (1995). Design Issues for Virtual Reality Systems. In Gobel, M (ed.) Virtual Environments'95. Springer, pp 224-236.
[14] Ingram, R. and Benford, S. (1995). Improving the Legibility of Virtual Environments. In Gobel, M (ed.) Virtual Environments'95. Springer, pp 211-223.
[15] Li-Jen, L. and Gaines, B. (1998). Modelling and Supporting Virtual Cooperative Interaction Through the World Wide Web. In Sudweeks, F. et al. (eds) Network & Netplay. MIT Press.
[16] Sebrechts, M. and Vasilakis, J. and Miller, S. and Cugini, J. and Laskowski, S. (1999). Visualisation of search results: a comparative evaluation of text, 2D and 3D interfaces. 22nd Conference on Research and Development in Information Retrieval ACM SIGIR, Berkley, pp 3-10.
[17] Shneiderman, B. (1998). Designing the User Interface. Strategies for Effective Human-Computer Interaction. Third Edition. Addison Wesley.
[18] Tufte, E. (1990). Envisioning Information. Graphics Press.