The populated web: Browsing, searching and inhabiting the WWW using collaborative virtual environments

Steve Benford, Dave Snowden, Chris Brown, Gail Reynard and Rob Ingram

Department of Computer Science
The University of Nottingham
Nottingham, NG7 2RD, UK
email: {sdb, dns, ccb, gtr, rji} @cs.nott.ac.uk

ABSTRACT. This paper presents new approaches for accessing the World Wide Web using multi-user three dimensional spaces, or collaborative virtual environments (CVEs.) The major sections of the paper deal with the aspects of browsing and navigating, searching, populating and displaying information within, and about the WWW. The methods described are illustrated with examples of working applications developed by the authors which together present an image of a possible future for Web access and collaboration based around the Web and its contents.

KEYWORDS. virtual reality, world wide web, visualisation, collaboration, virtual environments.

1. INTRODUCTION

The goal of this paper is to present new services for accessing, exploring and collaborating via the World Wide Web. Between them, these services demonstrate new approaches to browsing and searching and, perhaps more importantly, seek to transform the Web into an environment where users are naturally aware of one another's presence and are easily able to communicate with one another. Our approach utilises Collaborative Virtual Environment (CVE) technology (i.e. multi-user, distributed virtual reality) in order to provide navigable 3-D graphical, textual, audio and video representations of the structure of the Web and also of its users. Put another way, we aim to transform the Web from being a series of discrete locations which can be visited one at a time into being a landscape of many related locations which is visibly populated by its users.

This work has been motivated by a number of

problems with the Web in its current form.

Browsing and navigating - the problems of navigating hypermedia systems are well known (see below). A key aspect of these problems is that current Web browsers constrain the user to see only one page at a time instead of allowing them to get a sense of where they are currently located within a more general context.

Searching - searching is an intellectually demanding task which may involve iteratively comparing the results of many queries in order to home in on useful information. However, current Web search engines only allow users to enter a single query at a time and provide little opportunity for interactively comparing the results of many different queries.

Populating - although many users may be simultaneously accessing the same Web information, the current generation of Web browsers hides their presence from one another. Furthermore, although there have been

attempts to add real-time communication facilities to the Web (e.g. audio telephones), these have so far been limited in their capabilities and have not been well integrated with general Web browsing. In contrast, automatic awareness of the presence of other people and an ability to communicate with them is taken for granted in physical information environments (e.g. we are automatically aware of other people reading a notice board, browsing a newspaper stand or using a library without even having to think about it).

Displaying- the typical picture of Web access is of a single user browsing information via some kind of personal computer or workstation. Part of our work has involved considering alternative ways of displaying Web information within people's working environments (e.g. the use of larger scale projected displays) in order to enable new kinds of Web access and usage.

The remainder of this paper explores each of these issues in turn, presenting a series of Web-based applications and services which have emerged from our recent research. Between them, these applications offer a vision of future Web access which is radically different from that provided by the current generation of Web browsers.

2. BROWSING AND NAVIGATING - FROM PAGE TO LANDSCAPE

We begin our exploration with the issues of browsing and navigating. Jacob Nielsen addressed the problems of navigating large hypermedia structures, citing cases where users were unable to return to given locations [Nielsen90]. A number of solutions have been proposed to these kinds of navigation problems, including guides, tours and various backtracking mechanisms. Indeed, variants of these have been integrated in the current generation of World Wide Web browsers. These techniques, however, introduce their own problems; for example, the passive movement mode in guided tours may inhibit learning of the system and reduces the hypertext to a linear form. As a result, Nielsen has proposed the use of overview diagrams (effectively maps) and fish-eye techniques to show both local detail and global context. Some more recent browsers have made use of this idea to show a user's current location with respect the broader structure of the web (e.g. the Harmony browser for Hyper-G). With reference to our own work, this idea of accessing specific Web pages while still having a view of the surrounding context (i.e.

other nearby pages) has motivated the development of the WWW3D browser.

Like all of the applications described in this paper, WWW3D is realised as a Collaborative Virtual Environment - a multi-user distributed virtual reality system. This means that many users who are connected by a computer network can experience a shared virtual world constructed from interactive 3-D graphics, audio, video and text. Each user can freely navigate their own viewpoint within this world, can directly interact with the objects they find there, can see other graphically embodied users and can talk to them via audio and text channels. Users may access the world using a conventional workstation, a projected display or possibly a head mounted display (though in practice, few people have access to the latter and the quality of display provided is so low that they would be almost useless for accessing any textual information). The applications described in this paper have been built using the DIVE CVE platform from the Swedish Institute of Computer Science (SICS) [Fahlén93], a widely used CVE toolkit within the research community. More information on CVEs can be found in [Benford94].

For the time being, we will focus on the browsing functionality of WWW3D. We shall address multi-user issues in section 4 below. WWW3D provides a single 3-D graphical display which integrates 3-D representations of individual web pages (HTML), the structure of the part of the web that the user has explored and history information showing the links the user has followed in the recent past.

WWW3D uses the information contained in HTML tags to produce a representation of an HTML document in 3-D space. Each web document is represented as a sphere which is labelled on the outside with the document's title. The contents of the document are placed around the inside surface of the sphere and become visible when the user flies into the sphere. Textual information is represented by icons that can be unfolded to reveal an entire text. Images are displayed by texture mapping them onto polygons on the inside surface of the sphere. Finally, links to other documents are represented as icons, labelled with their destination. To reduce the visual complexity of the virtual environment WWW3D makes extensive use of Level of Detail (LOD) operations. When viewed from outside a document is represented as an opaque sphere and the actual document content is not displayed. When a user enters a document to view it, the sphere is drawn in wireframe so that the

rest of the virtual environment is still visible outside. Figure 1 shows the contents of an example web document as displayed by WWW3D.

When the user selects a link icon, WWW3D creates a new sphere representing the target document, locates it near the original document and draws an arrow between them. As more links are followed, so more spheres and arrows are added to the structure resulting in the gradual construction of a 3-D graph of the area of Web space that the user has explored. An incremental version of the Force Directed Placement (FDP) [Fruchterman 91] algorithm is used to dynamically regenerate the layout of this graph. This algorithm is based on a physical simulation of spring forces such that the spheres are treated as masses, the links as springs and the action of the springs on the masses is simulated until the whole structure settles into a stable state. As a result, the 3-D graph that is generated has the useful effect of pulling tightly interlinked Web pages into local clusters, thereby making the structure of the Web more legible.

Several other visual cues are provided to aid navigation. For example, the brightness of both the spheres and the link arrows is dependant on the time since the user last visited them, thereby providing the user with a visual representation of their browsing history. As WWW3D parses a newly retrieved document, it checks for links to documents that the user has already explored and also draws arrows to represent them. This means that at any given moment the complete set of links between document is displayed without the user having to follow every link. This is intended to aid the user by indicating links between documents of which the user might have been unaware. Figure 2 shows the display produced by WWW3D after a user has browsed a number of documents.

A user may start a WWW3D browsing session either by specifying one or more initial URLs from which to begin browsing or simply continuing where they left off, using the saved state from the last session.

In summary, WWW3D represents an innovative approach to browsing and navigating the Web in that it shows the detail of individual pages situated within the broader context of a 3-D map of the structure that a user has browsed so far and allows the user to easily move between these two levels of detail (e.g. by flying into a sphere to read a single page or pulling back their viewpoint to obtain an overview of structure).

Figure 1: A WWW3D representation of the contents of a web document.

Figure 2: A collection of web documents showing the links between them.

3. SEARCHING - GRAPHICAL REPRESENTATION AND MANIPULATION OF QUERIES AND RESULTS

Next we focus on the problem of searching. The typical operation of a current Web search engine is that the user enters one or more keywords and in return is

presented with a list of matching pages, rank ordered according to how well they matched. The problem with this approach is that it is difficult for the user to compare the results returned by different queries in order to home in on pages that match successive or multiple combinations of keywords. We have therefore developed the VR-VIBE application in order to allow users to dynamically compare the results of multiple search queries [Benford95]. VR-VIBE is based on an earlier 2-D system by Olsen et al. called VIBE [Olsen93] and works as follows:

1. The user enters a number of queries, each consisting of several keywords.
2. Each query is assigned a position in 3-D space, either directly by the user (i.e. by picking it up and dragging it) or automatically by the system if the user wishes.
3. The application matches each page in its local catalogue against each query, resulting in a score that represents the *relative attraction* of that page to that query. From this, the *overall relevance* of a specific page to all of the queries can be calculated (the sum of all scores for this page), as can the *overall significance* of each query (the sum of all scores across for this query).
4. A 3-D representation of the search space is then constructed in the following way. An icon representing each matched page is located at its relative centre of attraction to all of the currently active queries. Thus, a page which is attracted to only one query will be located next to that query, whereas a page which is attracted to several queries will be located somewhere between them, with its exact position depending on the relative strengths of the attractions. In other words, it is as if the queries exert a gravitational pull on the pages. The overall relevance of each page is mapped onto the size and shade of its icon.; thus, the bigger and brighter it is, the greater its overall relevance to all of the queries. Similarly, the overall significance of each query is mapped onto its brightness; thus, the brighter it is the greater its "pull" on the pages in the space. The resulting 3-D visualisation gives the user a sense of which pages are generally interesting (big and bright ones) and also which pages are most relevant to which specific keywords (spatial position).

However, perhaps the most important aspect of VR-VIBE is the way in which users may dynamically manipulate the visualisation. Current functionality includes:

- Flying through the 3-D scene in order to home in on specific areas or zoom out for an overview.

- Selecting the icons of pages in order to retrieve them from the Web and display them.
- Selecting queries in order to activate or de-activate them.
- Dynamically adding new queries to the visualisation.
- Repositioning queries either by dragging them or asking the system to "tidy" them into a sensible configuration.
- Running an additional temporary search and highlighting pages which match it.
- Using a relevance filter widget to remove all pages whose overall relevance falls below a set threshold.

We anticipate that it is this ability to dynamically manipulate the visualisation, adding and removing queries, moving them about and so forth, that is the real power of this approach over the current generation of "one query at a time" interfaces to Web search engines.

Figure 3 shows VR-VIBE being used to browse a catalogue of 1500 documents on the subject of CSCW. We can see the documents arranged into the shape of a pyramid, and simple embodiments of other users.

4. POPULATING - EMBODIMENT AND COMMUNICATION

So far we have considered how to visualise the underlying structure of the Web and also how to visualise the results of searching the Web. In this section we turn our attention to the issue of populating the Web - making its users directly visible to one another and enabling them to communicate. We separate this discussion into two parts: embodiment and communication for 3-D users (i.e. the users of future multi-user VR based Web browsers) and tracking and representation of 2-D users (i.e. the users of the current generation of 2-D Web browsers).

4.1. Embodiment and communication for 3-D users

Both of the above visualisations are inherently multi-user. This means that several people may share the same virtual world at the same time, freely navigating their own viewpoints. Each person sees the visualisation from their own autonomous perspective in 3-D space in much the same way as we perceive shared physical space. Each person may also be represented by an embodiment, a graphical object which is attached to their viewpoint and which shows that they are present, where they are located

FIGURE 3: Multiple users browsing a document space using VR-VIBE

and where they are looking in the virtual world. Embodiments (sometimes called avatars or clones), may range from simple graphical objects which essentially act as 3-D tele-pointers through to complex animated graphical forms which track the physical movements of their owners (through VR tracking devices), support gestures and even facial expressions (e.g. through the use of texture mapped video as discussed below). Embodiments may also support some other more subtle aspects of communication in virtual worlds including:

- **Degree of presence** - our experience with VR tele-conferencing has shown that the presence of an embodiment in a virtual world does not necessarily mean the active presence of its owner "inside" of it (particularly where simple, relatively static, embodiments are involved) [Greenhalgh95]. Users are often distracted by events in their physical space and sometimes even leave their terminal equipment unattended. This can cause some social embarrassment for other 3-D users (e.g. talking to an uninhabited body). As a result, our embodiments can show various degrees of presence (e.g. people may put them into a "sleeping" position if they have to absent themselves from the virtual world).
- **Points of presence** - an extension of the above idea involves adding some simple agent functionality to an embodiment so that it may carry out rudimentary communication functions on behalf of its owner while they are absent. For example, we have extended some of our

embodiments to record messages, to give out messages and to try to alert their owner when directly "prodded".

- **Capabilities** - As we shall see below, we anticipate that the Web will continue to be a heterogeneous environment in that people will access it using a wide range of equipment, network services and so forth. For example, some people may be audio capable and some only text; some may have 3-D browsers and some 2-D. A key aspect of embodiments is that they should naturally convey the underlying communication capabilities of their owners. Thus, if one of our bodies has ears then it is audio capable; a letter "T" on its head indicates text capability and a clear distinction is made between representations of 2-D and 3-D users.

In addition to embodiment, our applications support direct real-time communication between users through a combination of audio, text, graphics and video. Thus, on encountering another's embodiment one may directly address them using whatever combination of media is appropriate as is indicated by their embodiment. In the case of video, one or more real-time video streams can be texture mapped onto polygonal surfaces in the virtual world. For example, a video image of a face may be attached to the front of an embodiment, so that the video window moves and turns in the virtual world as the user does, or a video view into a remote physical scene may be realised as a "video window" pasted on a wall in the virtual world (see the Internet Foyer application below). Figure 4 shows an example of a three person virtual meeting from the point of view of one of the participants. The users are embodied with simple graphical objects but greater detail and identity is provided by the real time video faces.

Our integration of real-time media into virtual environments utilises an underlying mechanism called the spatial model of interaction which is intended to support scalability in terms of number of simultaneous participants whilst maintaining flexibility of communication [Greenhalgh95]. The model allows users to scope their interaction in terms of maintaining only a limited network connectivity to others as they move about the virtual world and also to flexibly control the quality of service provided across these different connections. To quickly summarise, connectivity is enabled through a concept called *aura* - a personal subspace which scopes ones presence. Quality of service is controlled through a concept called *awareness*, a measure of how much each object is aware of each other connected object. In turn, awareness is controlled by two

further concepts called *focus* and *nimbus*. A focus represents the spatial allocation of a users attention and a nimbus the spatial projection of their information. Thus, A's awareness of B is some function of A's focus on B and B's nimbus on A. Finally, various *adapter* objects change a users aura, focus and nimbus allowing them to reconfigure their communication capabilities (e.g. standing at the virtual podium expands one's aura and nimbus allowing one to address a greater number of people than usual). The net effect of aura, awareness, focus and nimbus is that users move around a virtual space and use various objects in order to manage connectivity and quality of service. More specifically, audio, textual, graphical and video are sensitive to relative positions and orientations (e.g. transmitted audio may become quieter or the frame rate of video images may be reduced as one moves or turns away from them).

In summary, we have discussed a number of issues concerned with representing the presence of other VR-based users of Web visualisations and of enabling and managing communication between them over different media. Next, we consider the issues involved in tracking and representing the presence of the users of the current generation of Web browsers.

4.2. Tracking and representation for 2-D users

Our browsers also support users in being aware of the presence of so called 2-D users (i.e. the users of current graphical multi-media browsers such as Netscape and Mosaic). This involves three key issues: tracking, representation and mutuality.

- **Tracking** - we have developed a module that tracks the presence of users as they access our Web server and move from page to page and notifies the various visualisation tools that were described above. This module works with both live data and also from Web logs.
- **Representation** - we represent the presence of these tracked users through embodiments. 2-D users' embodiments are placed close to and facing the page that they are currently browsing. Animated movements then show their progress as they move from page to page across the visualisation. Thus, one can directly see their progress as they wander about.
- **Mutuality** - tracking and representation raise the further issue of privacy. Our approach is based on a general principle of mutuality, whereby anyone who enters a server that supports tracking is warned of this in ad-

vance and is given the opportunity to decline being tracked (our Web pages carry a standard form allowing people to sign in and identify themselves). On electing to be tracked, a 2-D user gets to see a textual list of other visitors (3-D and 2-D) appended to each Web page. If they decline, then they do not see this list. Thus, either users are mutually aware or mutually invisible.

5. DISPLAYING - BOUNDARIES BETWEEN PHYSICAL AND VIRTUAL

Our final issue in this exploration of future Web interfaces concerns the nature of physical displays. The current picture of Web access is of users sitting down at individual machines and browsing through the Web as a focused activity. Although we would not expect this kind of access to disappear, there are other interesting possibilities that might be considered. In particular, in considering the Web as a navigable space in its own right, we have become interested in its relation to people's existing physical spaces. This has led us to experiment with new forms of Web access based on the use of shared projected displays which enable more peripheral and casual awareness of the Web and its use. We shall illustrate this idea through an example application - the Internet Foyer - which integrates many of the techniques described in previous section [Benford96].

5.1. Goal of the Internet Foyer

We begin with the goal of the Internet Foyer. Foyers are important areas of physical buildings for a variety of reasons. First, they present the public face of a building and the organisation(s) which it houses. The significance of this function alone should not be underestimated; large sums of money are spent on making foyers interesting and impressive places. Second, foyers provide a context for locating useful information for visitors such as maps, directories and displays. Third, they may be home to various people whose job it is to help these visitors (e.g. receptionists). Fourth, they enhance security by providing a single point of entry into the organisation within which incoming and outgoing people are made publicly visible and hence accountable. Fifth, they provide public meeting places, either for arranged rendezvous or for chance encounters.

Turning to the Web, it is clear that, in parallel to the physical manifestation of buildings, many organisations are increasingly acquiring an electronic manifestation

through the World Wide Web. In such cases, the public "home" pages of an organisation on the WWW could be considered to be a kind of foyer: a public entry point into the organisation's Web space. There is certainly evidence that organisations are investing considerable effort into turning their WWW pages into interesting and impressive places. However, when compared to physical foyers, WWW pages leave much to be desired. For example, the people who pass through them are not generally visible to one another and there are no opportunities for rendezvous and social encounters.

There are therefore two main goals to the Internet Foyer. First, by using a combination of the techniques described above, to construct a virtual foyer based on a populated visualisation of an organisation's WWW space which might come closer to the functionality of a real foyer. Second, to merge this virtual foyer with a real physical foyer so as to provide an integrated shared space which spans both the physical and virtual worlds. Thus, visitors entering an organisation's WWW space would be able to interact with those entering its physical space and vice versa.

5.2. Functionality of the Internet Foyer

The overall concept of the Internet Foyer is summarised below. Visitors in a traditional physical foyer see a graphical representation of a virtual foyer projected onto the wall. This virtual foyer is a 3-D graphical visualisation of an organisation's home pages whose electronic visitors are mutually embodied (i.e. visible to one another) and are able to communicate with one another through open audio channels. A video window in the virtual foyer affords its occupants a corresponding view back into the real foyer and an open audio channel allows communication between the two spaces. There are therefore three ways to experience the Internet Foyer:

As a visitor to the physical foyer one users enter a normal physical space. Projected onto a wall of that space is the graphical representation of the virtual foyer. This representation is based on the approach described in section 2. Both 3-D and 2-D visitors to this visualisation and represented using the techniques described in section 4 and an open audio link allows communication with the virtual reality users in the virtual foyer.

As the user of a VR-based Web browser, one sees the same basic visualisation as those in the physical foyer. However, one is able to freely navigate around this

visualisation, homing in on specific details or backing off in order to obtain a perspective view. A real-time video window which is texture mapped onto a wall of the virtual foyer allows CVE users to look out into the physical foyer and to see its occupants looking back at them.

As the user of a traditional (2-D) WWW browser one sees the Internet foyer as a series of WWW pages. On entering the foyer one is asked to register. Beyond this, additional pages provide information about current and recent visitors, displayed as simple textual lists and may display captured images from both the real and virtual foyers.

Figures 4(a) and 4(b) present images of the Internet Foyer as it appears to different users. Figure 4(a) shows an overview of the Internet Foyer as it appears to a VR-based user, with a 3D graph style visualisation of a set of WWW pages, as described earlier, and a real time video window onto the physical world. Figure 4(b) shows how the Internet Foyer appears to visitors in the physical foyer.

In summary, the Internet Foyer represents an interesting Web application from several points of view. First it combines all of the ideas of visualisation, embodiment, tracking and communication as described above. Second, the use of combined projection and video texturing techniques has enabled us to create a new kind of display where the Web is seen more as an extension to a normal physical space (i.e. where access is shared and casual), rather than as a specific application which is launched on an individual computer.

6. Conclusion

Many people talk about the Web as if it were a space. Indeed, the populist term "Cyberspace" has become synonymous with the Web. In this paper, our aim has been to consider how the utilisation of Collaborative Virtual Environment technology might actually help transform the Web into being a truer space than it is today - specifically, into being a space that is both navigable and populated. Our paper has presented a series of applications which, between them, offer a vision of future Web browsers and services. The paper has also explored a number of key issues which underlay this vision, including:

- Allowing users to perceive the inherent structure of a region of the Web through the use of 3-D visualisation techniques, thus aiding navigation.

- Supporting more dynamic and interactive forms of searching.
- Embodying both 3-D and 2-D Web users so that they are mutually aware.
- Supporting integrated and scaleable real-time communication (including audio and video) between users through the concepts of aura, focus and nimbus which provide new ways of managing network connectivity and quality of service based on spatial proximity.
- Supporting the tracking of users as they progress across the Web, coupled with the notion of mutuality so as to maintain personal privacy.
- Considering the use of projected displays to enable new forms of Web access.

We hope that kinds of techniques may soon become an integral part of Web browsers. The necessary infrastructure is certainly maturing rapidly in the form of low-cost, high performance 3-D graphical platforms and the recent emergence of 3-D standards for the Web (i.e. VRML). Given this emerging infrastructure, we believe that it may soon be possible to develop a new generation of more usable, dynamic and populated Web services and tools.

REFERENCES

[Benford94] Benford, S. D., Bowers, J. M., Fahlén, L. E., Mariani, J. and Rodden, T. R., Supporting Co-operative Work in Virtual Environments, *The Computer Journal*, 37 (8), Oxford University Press, 1994.

[Benford95] Benford, S., Snowdon, D., Greenhalgh, C., Ingram, R., and Knox, I., VR-VIBE: A Virtual Environment for Co-operative Information Retrieval, *Computer Graphics Forum*, 14, 3, 1995, NCC Blackwell.

[Benford96] Benford, S. D., Brown, C. C., Reynard, G. T. and Greenhalgh, C. M., Shared Spaces: Transportation, Artificiality and Spatiality, *Proc. ACM Conference on Computer Supported Cooperative Work (CSCW'96)*, Boston, ACM Press (in press).

[Fahlén93] Fahlén, L. E., Brown C. G., Stahl, O., Carlsson, C., A Space Based Model for User Interaction in Shared Synthetic Environments, in *Proc. InterCHI'93*, Amsterdam, 1993, ACM Press.

[Fruchterman91] Fruchterman, T. M. J. and Reingold, E. M., Graph Drawing by Force Directed Placement, *Software Practice and Experience*, Vol. 21 (11), 1129-1164.

[Greenhalgh95] Greenhalgh, C. M., and Benford, S. D., MASSIVE: A Virtual Reality System for Tele-

FIGURE 4: The Internet Foyer as viewed (a) from within the virtual world and (b) from the physical world using a projection display.

conferencing, *ACM Transactions on Computer Human Interfaces (TOCHI)*, 2 (3), pp. 239-261, ISSN 1073-0516, ACM Press, September 1995.

[Nielsen90] Nielsen, J., The Art of Navigating Hypertext, *Communications of the ACM*, 33, 3, pp. 297-310, ACM Press.

[Olsen93] Olsen, K. A., Korfhage, R. R., Sochats, K. M., Spring, M. B. and Williams, J. G., Visualisation of a Document Collection: The VIBE System, *Information Processing and Management* 29 (1), pp. 69-81.