



# On the Design of a Dual-Mode User Interface for Accessing 3D Content on the World Wide Web

Jacek Jankowski, Stefan Decker

## ► To cite this version:

Jacek Jankowski, Stefan Decker. On the Design of a Dual-Mode User Interface for Accessing 3D Content on the World Wide Web. *International Journal of Human-Computer Studies*, 2013, 10.1016/j.ijhcs.2013.04.006 . hal-00822676

**HAL Id: hal-00822676**

**<https://inria.hal.science/hal-00822676>**

Submitted on 25 Nov 2013

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# On the Design of a Dual-Mode User Interface for Accessing 3D Content on the World Wide Web

Jacek Jankowski<sup>a,b,\*</sup>, Stefan Decker<sup>a</sup>

<sup>a</sup>Digital Enterprise Research Institute, National University of Ireland, Galway, Ireland

<sup>b</sup>Inria Bordeaux, France

---

## Abstract

The World Wide Web, today's largest and most important online information infrastructure, does not support 3D content and, although various approaches have been proposed, there is still no clear design methodology for user interfaces that tightly integrate hypertext and interactive 3D graphics. This paper presents a novel strategy for accessing information spaces, where hypertext and 3D graphics data are simultaneously available and interlinked. We introduce a Dual-Mode User Interface that has two modes between which a user can switch anytime: the driven by simple hypertext-based interactions "*hypertext mode*", where a 3D scene is embedded in hypertext and the more immersive "*3D mode*", which immerses the hypertextual annotations into the 3D scene. A user study is presented, which characterizes the interface in terms of its efficiency and usability.

**Keywords:** Hypertext, 3D Graphics, 3D Web, User Interface

---

## 1. Introduction

This article is an extended version of a paper: "*A Dual-Mode User Interface for Accessing 3D Content on the World Wide Web*" (Jankowski and Decker, 2012) presented at WWW'12.

### 1.1. Background and Motivation

Two different developments provide the foundation for this work. The first one is the evolution of hypertext and the spectacular success of the World Wide Web, which provides a publicly accessible, shared global information space - a phenomenon which has never before existed in the history of mankind. The second one is the growing popularity and importance of three-dimensional (3D) computer graphics for our economies.

*Hypertext and the Web.* Hypertext is usually seen as originating in the Vannevar Bush's historical article "*As We May Think*" (Bush, 1945), where he identified the information-overload problem and proposed cross-references within and across documents to better organize and manage the huge amount of available scientific information. His vision influenced and inspired two researchers generally credited with the invention of hypertext, Ted Nelson and Douglas Engelbart. Ted Nelson coined the term "*hypertext*" for non-sequential writing and reading in the 1965 while he was working on *Xanadu* (Nelson, 1965). During the 1960s, Douglas Engelbart worked on the "*oN-Line System*" (NLS), a revolutionary computer collaboration system that employed hypertext point-and-click features and mouse (Engelbart and English, 1968). By the mid-1980s many hypertext systems were developed.

Hypertext has become a mainstream interface paradigm with the emergence of the World Wide Web that opened the closed hypertext systems to the Internet and connected everyone's hypertexts into one hyperspace. Nowadays, billions of people use the Web to access information, take care of their daily chores, and also have a source for entertainment. They use Google to search; they read articles on Wikipedia to learn; they watch videos on YouTube to have fun; they hang out on Facebook to "*socialize*" or play online games. The Web has truly changed the way people spend their work and leisure time.

The Web evolved from a text-based system to the current rich and interactive medium that supports images, 2D graphics, audio and video. However, the vision was always bigger. The inventor of the World Wide Web, Tim Berners-Lee's dream behind the Web was to create "*a space in which anything could be linked to anything*" (Berners-Lee and Fischetti, 1999). The major media type that is still missing is 3D graphics. We argue that 3D graphics and virtual environments are equally viable and compelling "*documents*" of interest on the Web.

*3D Graphics and Virtual Environments.* Ivan Sutherland is acknowledged by many to be the "*grandfather*" of interactive Computer Graphics, Graphical User Interfaces, and Human-Computer Interaction. His Sketchpad (Sutherland, 1963) enabled users, for the very first time, to interactively create and directly manipulate images on a computer display. In his further work on the "*Ultimate Display*" (Sutherland, 1965), Sutherland demonstrated his vision of the immersive 3D experience: the ultimate display would act as a looking glass into a mathematical wonderland; it would display information in three-dimensional fashion, represented in utterly believable virtual world. In 1966, he used head-mounted display (HMD) to create the first Virtual Reality (VR).

---

\*Corresponding author. Tel:+33-535002605

Email address: jacek.jankowski@inria.fr (Jacek Jankowski)

Certainly, Ivan Sutherland helped to push the computer graphics discipline to where it is today, where expressing information through 3D visualization is bringing huge benefits for a variety of applications ranging from video games, through the latest special effects in movies, virtual environments to serious applications (e.g., engineering or medical visualizations, cultural heritage preservation). As computer graphics technology has reached the point where high quality 3D content can be rendered in real time on commodity desktop and mobile devices, it should be possible to use 3D models rather than 2D images to represent various objects and ultimately improve user experience on the Web. Moreover, as people spend all their lives in a 3D world and they develop skills for manipulating 3D objects, navigating around the 3D world, and interpreting 3D spatial relationships, we are convinced that 3D computer graphics and virtual environments have the potential to support web user interfaces.

## 1.2. Problem Statement

There have been a number of approaches over the years to integrate 3D technologies on the Web and most of these systems and standards disappeared or barely survived. We believe that this is partly because of the fact that research was focused mostly on 3D graphics and 3D graphics alone. The historical focus of research did not include the search for user interaction techniques that are optimal for the hypertext-based Web interface. However, what good is a realistic environment if one cannot interact with it? As a result, hypertext (the ultimate product in symbolic communication) and interactive 3D graphics (the ultimate achievement in visual media) are just now converging on the Web.

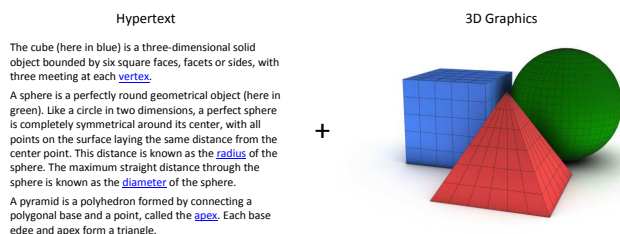


Figure 1: How to design an interface that combines hypertext and 3D graphics?

We believe that people can gain a lot from using integrated information spaces where hypertext and 3D graphics data are simultaneously available and linked, spaces we call Hypertextualized Virtual Environments (HiVEs). This paper focuses on user interface design that supports the integrated exploration of HiVEs; the design, where users can browse the text, look through general information and search for more specific information, and where they can also navigate freely through a 3D space, and examine and manipulate virtual 3D objects, to gain a better understanding of the data. The objective of our user interface is to pair interactive 3D graphics know-how with well established UI conventions of the Web to support all these user tasks. We believe that this issue is of a great importance, since there are upcoming new open (WebGL) and proprietary (Stage3D) proposals for 3D graphics APIs in the Web context.

## 1.3. Outline of the Paper

The paper at hand is structured into four parts. The first part (Section 2) introduces the foundations and lists related academic work. The second part (Section 3) develops a novel strategy for accessing Hypertextualized Virtual Environments (e.g., viewing 3D content on the Web). The third part (Section 4) describes a HiVE testbed environment. The fourth part (Section 5) presents the evaluations that assessed different aspects of our UI design. The paper ends with a summary and conclusions and by posing future directions for our work.

## 2. Foundations - Factors Driving the Design

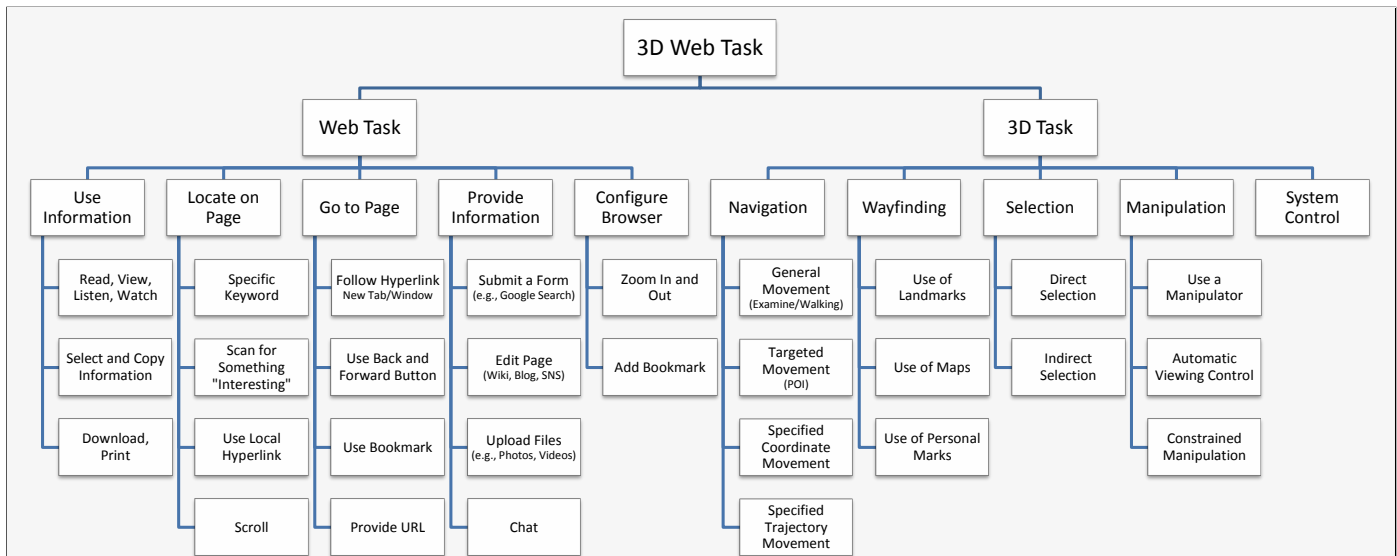
There were five main driving factors in designing our user interface: (1) User's primary tasks; (2) Meeting the needs of both novice and experienced users; (3) Principles of designing multimedia environments; (4) Existing body of UI design work for integrated information spaces where text and 3D graphics data are simultaneously available and linked; (5) Web usability principles and 3D design guidelines.

### 2.1. User's Primary Tasks

A prerequisite to the effective user interface design is an understanding of the users and their tasks - the tasks for which that interface will actually be used. An important component of this process is classification that can be done in the form of taxonomies. By classifying, categorizing, and decomposing the tasks, we gain a much deeper understanding of them and provide ourselves with a framework that can be further used to help design new interaction techniques. In (Jankowski, 2011), we focused on an understanding of the fundamental tasks users may engage in while exploring Web-based 3D virtual environments. We constructed a taxonomy of 3D Web user tasks (see Figure 2), where we divided these tasks into hypermedia tasks, such as following hyperlinks and placing bookmarks, and virtual environment tasks, such as 3D navigation and object manipulation. We also included a review of mouse-based 3D interaction techniques useful in the context of 3D Web. Our "taxonomy of 3D Web use" is based on: (1) the taxonomy of Web tasks created by Byrne et al. (1999) (the adaption was motivated by the changes in the capabilities of the Web and Web browsers) and (2) classification works of Hand (1997), Bowman et al. (2001) and Mackinlay et al. (1990).

### 2.2. Varying User Experience Levels

One of our major goals was to provide an interface that meets the needs of both novice and experienced users (Shneiderman and Plaisant, 2004). We assumed that most users would be new to 3D interactions. We therefore needed to design our UI in favor of novices. This meant adhering to well-established UI conventions and making 3D navigation as simple as possible. On the other hand, there are knowledgeable 3D users who can find the limitations and constraints of a novice interface frustrating. As we did not want to diminish the 3D interaction experience in any way, we needed to provide these expert users with much more freedom with regards to the 3D task support.



### Web User Tasks

The five subclasses of Web tasks include: Use Information, Locate on Page, Go To Page, Provide Information, and Configure Browser.

**Use Information** *Read, View, Listen, Watch*: Reading text is one of the most common activities while browsing the Web. The Web also allows users to view images, listen to music, watch videos, etc. *Select and Copy Information (Ctrl-C)*: This task refers to the popular, simple Copy-and-Paste method of reproducing text or other data from a source to a destination. Typically, it involves selecting the text by dragging over the text with the mouse and using of the key-combination Ctrl-C (or using a context menu). *Download, Print*: Information stored on the Web can be downloaded to a local disc or printed - it then can be used offline.

**Locate on Page** *Specific Keyword (Ctrl-F)*: Web users can find a keyword or phrase on a web page they are viewing by pressing Ctrl-F, entering the word they are looking for, and clicking on the "Find Next" button: their browsers will automatically locate and highlight the word. *Scan for something "Interesting"*: In this class of task a user is seeking no specific word or concept, but is simply scanning the page for something that might catch his or her interest. *Use Local Hyperlink*: A local hyperlink is a reference or navigation element in a Web document to another section of the same document. This type of link is often used to implement tables of contents. *Scroll*: If a Web page

is too long to be displayed on one screen, the user has to navigate on the page by scrolling it with the help of a scrollbar, arrow keys or using a mouse scroll wheel.

**Go to Page** *Follow Hyperlink*: Link following is the most common navigation action. It causes the target document to replace the document being displayed; users can also open the target document in a new tab or window using the browser's context menu. *Use Back and Forward Buttons*: These buttons allow users to return to previously visited webpages. They make the users more confident, since they know they can always backtrack. *Use Bookmark*: Bookmarks, also called favorites in Internet Explorer, can be accessed through a web browser' menu or toolbar. They allow users for direct jumps to their favorite places on the Web. *Provide URL*: Users can type URL, the "address" of a webpage, directly into the address bar of their browsers to open the webpage.

**Provide Information** *Submit a Form*: This task refers to providing information such as authentication information or search criteria in a webform. *Edit Page*: Web 2.0 websites, such as social-networking sites, blogs, and wikis allow their users to interact/collaborate with each other as content creators. *Upload Files*: Web 2.0 web sites also allow their users to upload and share photos (e.g., Flickr, Facebook), video clips (e.g., YouTube), etc. *Chat*: Some websites allow users to communicate in real time using easily accessible browser-based chat services (e.g., Facebook).

**Configure Browser** *Zoom In and Out*: All major web browsers allow users to increase or decrease the size of web pages to improve readability. *Add Bookmarks*: Web browsers let users add bookmarks as they work. They can also use web-based social bookmarking services that let them save bookmarks on a remote web server, accessible from anywhere.

### 3D User Tasks

The "3D task" class consists of five subclasses: Navigation, Wayfinding, Selection, Manipulation and System Control.

**Navigation** Navigation refers to the task of moving user's viewpoint through an environment. *General movement* - Exploratory movement, such as walking through a simulation of an architectural design; *Targeted movement* - Movement with respect to a specific target, such as moving in to examine a detail of an engineering model; *Specified coordinate movement* - Movement to a precise position and orientation, such as to a specific viewing position relative to a car model; *Specified trajectory movement* - Movement along a position and orientation trajectory, such as a cinematographic camera movement.

**Wayfinding** Wayfinding is related to how people build up an understanding (mental model) of a virtual environment. *Use of Landmarks* - Using landmarks (such as paths, fences, statues) support navigation in VEs; *Use of Maps* - Wayfinding can be a difficult also in the real world - people often have difficulties when dealing with unfamiliar physical environments. Including global maps

to the local views (a technique that became a common practice in games) can greatly improve navigation performance. *Use of Personal Marks* - Users should be able to use visual markers or "virtual breadcrumbs", which may be dropped as a means of assisting navigation.

**Selection** Selection refers to techniques of choosing an object. *Direct Selection* - The most common way to select objects in 3D space is positioning a mouse cursor over given object and clicking a mouse button (it is based on ray casting). *Indirect Selection* - Selection of objects can also be done indirectly from a collection of objects (similarly to selecting viewpoints that are defined for a 3D scene).

**Manipulation** Manipulation refers to techniques of specifying object's position, orientation and scale. *Use of Manipulators* - Clicking and dragging on the handles of the manipulator displayed together with the object. *Automatic Viewing Control* - Manipulation can be augmented by automatically finding an effective placement for the virtual camera. *Constrained Manipulation* - Manipulation can also be simplified by exploiting knowledge about real world (e.g., gravity) and natural behaviour of objects (e.g., solidity).

**System Control** System Control refers to communication between a user and a system which is not part of the virtual environment. It refers to a task in which a command is applied to change either the state of the system or the mode of interaction.

Figure 2: Taxonomy of 3D Web user tasks (Jankowski, 2011).

### 2.3. Principles of Designing Multimedia Environments

Multimedia presentation was studied extensively within psychology, aiming at extracting principles that guarantee an effective design and facilitate learning; the central claim of multimedia is that providing information in more than one medium of communication will lead to better understanding (Mayer, 2009). The theory based on *Dual channel* (Paivio, 1986), *Limited capacity* (Chandler and Sweller, 1991), and *Active process-*

*ing* (Wittrock, 1989) assumptions suggests that if active processing takes place in both visual and verbal cognitive subsystems, learning is improved; dual coding of information is more effective than single coding; it is also critical that both visual and verbal representation are actively constructed, together with the connections between them. Supporting multimedia theory, studies have shown that verbal and visual representations in combination are often more effective than either

in isolation (Mousavi et al., 1995; Faraday and Sutcliffe, 1997, 1999; Mayer and Moreno, 1998). On the other hand, Nielsen (2000) warns that unconstrained use of multimedia can result in user interfaces that confuse users and make it harder for them to understand the information.

Mayer listed the basic principles for how to design multimedia learning environments (Mayer, 2005): *Multimedia principle*: People learn better from words and pictures than from words alone; *Spatial contiguity principle*: People learn better when corresponding words and pictures are presented near to each other rather than far from each other on the page or screen; *Temporal contiguity principle*: People learn better when corresponding words and pictures are presented simultaneously rather than successively; *Coherence and signalling principles*: People learn better when extraneous material is excluded rather than included and when cues are added that highlight the organization of the essential material; *Segmenting principles*: People learn better when a multimedia message is presented in learned-paced segments rather than as a continuous unit;

We believe that understanding and adhering to these principles is crucial to integrating hypertext and 3D graphics. Consequently, they will guide our design decisions throughout the user interface development process.

## 2.4. Existing Body of UI Design Work

In this section we survey the work that has been done in the area of UI design for information spaces where hypertext and 3D graphics data are simultaneously available and linked: we start from describing two early 3D hypermedia systems - Intermedia and Hyper-G, we then describe current approaches for integrating HTML and 3D graphics on the Web, we talk about Information-Rich Virtual Environments, and finally we present the body of research related to labeling and annotating 3D illustrations.

### 2.4.1. Early 3D Hypermedia Systems

Brown's Intermedia was one of the first hypermedia systems that integrated 3D documents (Yankelovich et al., 1988). It allowed authors to create a linked body of material that included not only text and static images but also animated graphics, timelines, video, and sound. With the InterSpect viewer, users could manipulate three-dimensional images by rotating them, zooming in or out, or hiding parts of the model. Another remarkable 3D hypermedia system was developed by researchers from Graz University of Technology (Andrews, 1992, 1994, 1996; Andrews and Pichler, 1994; Pichler, 1993). Their Hyper-G system was the first to combine 3D graphics and hyperlinking across the Internet. Harmony, a graphical hypermedia client for Hyper-G, would later become the VRweb/VRwave browser for the VRML. These hypermedia systems were window-oriented, which means that they used document clusters to form groups of related documents and all document types stayed separated in their own windows. As a result, they were not bound to a particular UI metaphor. They also had separate link databases that enabled linking from and to different kinds of media; this allowed links to be bidirectional.

### 2.4.2. Integrating HTML and 3D Graphics on the Web

In contrast to the presented early window-oriented hypermedia systems, the Web is a multimedia document based hypermedia system. Its user interface is based on single documents (HTML web pages) consisting of several parts of information of different types. The documents are designed and composed by a web designer, who is responsible for placement of texts and media elements, and the overall aesthetics of the site.

Supporting 3D content on the Web has been a focus of the research community since the 1990s with the early introduction of the Virtual Reality Markup/Modelling Language (VRML), a simple, text-based format for describing virtual worlds that has been later superseded by X3D, which adds XML capabilities to better integrate with other WWW technologies.

In order to integrate VRML/X3D with HTML, a special browser plug-in is required that allows the 3D scene and UI controls to be displayed within the web browser window. Following VRML/X3D, a number of other plug-in based approaches to integrate 3D technologies on the Web has been proposed including Flash, Unity3D, WPF, Java3D, and MPEG4 (a broader overview of the current state of 3D on the Web can be found in (Behr et al., 2009)).

There are generally two techniques that are used to integrate HTML with VRML/X3D (some discussion on this topic can be found in (Balaguer and Gobbetti, 1995; Campbell, 1997)).

In *Immersive Approach*, a VRML/X3D scene appears in the viewing window that spans the entire available screen space of the browser window (Mohageg et al., 1996). When the user activates a hyperlink in the 3D scene, the related HTML file is opened and given full presentation by the browser in the browser window. The user needs to close the file to return to the scene. A big disadvantage of such approach is that it is not in line with the temporal contiguity principle (see Section 2.3). In this approach, annotations can also be attached to objects in the 3D scene and to do that, VRML/X3D developers typically use Billboard node, a grouping node that automatically turns to face the world visitor as he or she moves in the scene. HTML can be inserted on a Billboard by converting it to a bitmap representation. The problems with this approach include lack of control over HTML and poor readability.

In *Document-based Approach*, HTML and VRML/X3D documents are presented separately side by side in a browser. Creating such multimedia document involves using frames (not recommended: problems with linking, indexing, and security) or the <object> HTML tag (the replacement for the <embed> element). The <object> tag allows objects, such as audio, videos, Java applets, ActiveX, PDF, and Flash, as well as VRML/X3D, to be incorporated within a HTML page and displayed by the appropriate plug-in application. In this approach, Scene Authoring Interface (SAI) can be employed to control embedded in HTML VRML/X3D scenes via JavaScript.

Today, WebGL (Khronos, 2010), a successor of Canvas3D (Vukicevic, 2007), a JavaScript API specification analogous to the OpenGL API for C++, aims to bring plugin-free 3D to the web, implemented right into the browser. At the time of this writing (March, 2013), WebGL is supported in most major web

browsers including Google's Chrome, Mozilla's Firefox and Apple's Safari.

Both plug-in-based and WebGL techniques for developing 3D Web sites do not really allow for a tight integration of hypertext with 3D graphics. The two parts, textual and visual, are typically developed separately and there is no real connection between them. Therefore, in the Declarative 3D for the Web Architecture W3C Community Group we explore new declarative ways (such as X3DOM (Behr et al., 2009) and XML3D (Sons et al., 2010)) of incorporating 3D graphics natively in the document object model (DOM).

With regards to viewpoint control, both plug-in-based and WebGL techniques usually provide users with the means for navigation through a 3D scene: on the one hand, they implement only one 3D navigation technique - the technique that is best suited for a given task (e.g., car configurators offer a sweep/tubmle navigation mode or a guided tour); on the other hand, VRML/X3D browsers offer multiple methods of interaction based on examine, fly, walk and fly-to. The first approach limits interaction for the sake of simplicity. The second offers more freedom in terms of viewpoint control. Our goal is to combine the most useful features of these approaches.

#### 2.4.3. Information-Rich Virtual Environments

Another closely related work is the research on integrating perceptual and symbolic information in VEs, and the further work on Information-Rich Virtual Environments (IRVEs) done by Bolter et al. (1995); Bowman et al. (1998, 1999, 2003); Polys and Bowman (2004); Polys et al. (2005). IRVEs combine the power of VEs and information visualization, augmenting VEs with additional abstract information such as text, numbers, or graphs. IRVE applications show promise in increasing the utility of the VE experience. Usability studies conducted in the Virtual Venue (Bowman et al., 1998) proved that the experience was enjoyable and useful. Work on Virtual Habitat (Bowman et al., 1999) shown the potential of IRVEs for education.

In his PhD thesis, Polys (2006) focused on information design for IRVEs; he investigated rendering and layout issues including managing the display space for visibility, legibility, association, and occlusion of the various data. In one of his experiments (reported in (Polys et al., 2005, 2011)), he evaluated the depth and association cues between objects and their labels using two categories of annotation techniques:

- *Object-space techniques*: annotations are embedded into a 3D scene relative to their referent; this technique provides strong association cues; on the other hand text displayed in 3D undergoes pitch, yaw, and roll transformations.
- *Screen-space techniques*: annotations are placed on a 2D plane that overlays a 3D scene; such techniques ensure better legibility of text.

Their experiment showed that screen-space interfaces outperformed object-space layouts on nearly all counts of accuracy, time, and ratings of satisfaction and difficulty across tasks. Therefore, we decided to use solely screen-space techniques for displaying annotations in 3D. Polys et al. (2005) also suggested

the development and evaluation of a richer set of viewport-space layout capabilities. In (Jankowski et al., 2010) we explored the effects of different screen-space techniques of integrating text with animated 3D graphics on text readability.

#### 2.4.4. Labeling and Annotating 3D Illustrations

Another closely related work is the small but growing body of research related to labeling and annotating 3D illustrations.

*Labeling 3D Illustrations.* Illustrations in scientific or technical documents typically employ labels to effectively convey co-referential relations between textual and visual elements. For centuries labels have proven their worth as powerful tools to support learning tasks. With regards to 3D interactive illustrations, most notable work was carried out by a group of researchers under the direction of Thomas Strothotte: Ali et al. (2005) proposed a variety of real-time label layout algorithms; the authors also state the requirements for dynamic labeling, classify layout styles; they also proposed the evaluation criteria to measure the quality of layouts. In their further work, Hartmann et al. (2005) and Gotzelmann et al. (2005) focused on functional and aesthetic attributes of label layouts. Another algorithms for labeling were presented Stein and Decoret (2008); Cmolik and Bittner (2010). Real-time automatic labeling algorithms were also used in Augmented Reality environments (e.g., (Bell et al., 2001; Azuma and Furmanski, 2003)).

*Annotating 3D Illustrations.* One of the first approaches for correlating text with illustrations of 3D objects was presented by Preim et al. (1997). Their system, Zoom Illustrator, maintains a close relationship between images and associated text, with interaction on the textual part influencing the appearance of the graphical part and vice versa; it reserves part of a view plane for textual descriptions connected by lines to the referenced objects. In a study that strongly affected our design, Sonnet et al. (2005) compared methods of associating text with its 3D model; they evaluated the effects of text positioning, connectivity, and visual hints on comprehension under three conditions: (a) annotations are attached to objects using translucent shapes; (b) annotations are located within the objects' shadows; (c) area showing the 3D model and text area are separated. The authors suggest that *setting a* works well for short labels, while for extensive texts, *setting c* seems to be applicable because a user can explore a scene without any occlusions from the text.

Gotzelmann et al. (2006) present a real-time agent-based algorithm to integrate labels of arbitrary size into 3D visualizations. In their approach comprehensive dynamic content can be displayed in annotation boxes. Their layout system displays annotation boxes as large as possible while preventing overlaps with scene objects; the distance to their associated visual elements is minimized. Users can pin annotation boxes on their current position, enlarge or minimize them (they then become labels). In his further work, Gotzelmann et al. (2007) present a tool where students can interactively explore tutoring material that is enhanced with 3D illustrations. By selecting text segments or by adjusting the view on 3D models, the users can request additional contextual information.



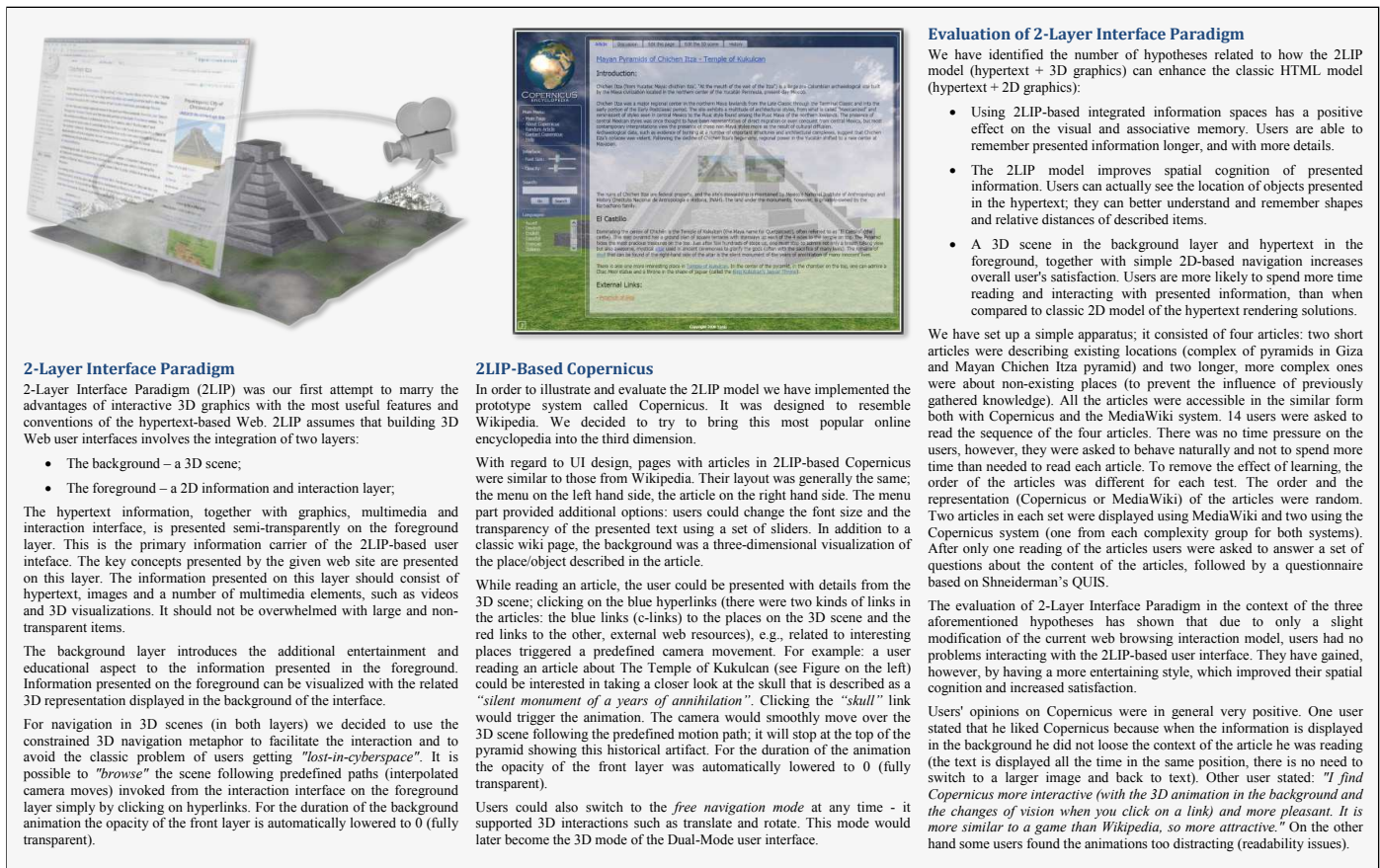


Figure 3: 2-Layer Interface Paradigm (2LIP) (Jankowski and Kruk, 2008; Jankowski and Decker, 2009)

## 2.5. Web Usability Principles and 3D Design Guidelines

A review of recent HCI literature indicates that there is a debate as to the most effective methodology or philosophy for the development of good user interfaces. A growing number of researchers and developers is considering usability as a major focus of interactive system development (e.g., Nielsen (2000); Nielsen and Loranger (2006); Shneiderman and Plaisant (2004)). Therefore, usability was another major driving factor for our design.

### 2.5.1. Web Usability

After publishing countless articles and several books on Web Usability, Jakob Nielsen is rightly seen as one of the pioneers of making the web more usable and useful to everyday life. In his works the author shares his insightful thoughts on the subject, thoughts that are backed up with empirical data. Based on his "Designing Web Usability" (2000) and Steve Krug's "Don't Make Me Think" (2005), we summarize what we consider to be the main principles of Web usability:

- **Websites should explain themselves:** Web pages should be self-evident. Obvious. Self-explanatory. Users should be able to learn what they are and how to use them without expending any effort thinking about it.
- **Do not waste people's time:** Much of the web use is motivated by the desire to save time.

- **People do not read pages; they scan them:** Most people spend little time reading web pages, they simply scan (or skim) them for the first thing that looks good.
- **People are creatures of habit:** If they find something that works, they stick to it; they tend not to look for a better way. Designers should use existing Web conventions.
- **People tend to get "lost in space", so:** Make it easy to go home (so users can start over); Highlight current location in navigational UI components; Allow local search;
- **Typography - Readability and Legibility:** The primary goal of Web design is to convey information. Designers should choose typography that communicates.
- **Allow going back:** People are more comfortable when they know that they can undo or alter their actions.

### 2.5.2. 3D Design Guidelines

Currently, computer and video game industry leads the development of 3D interactive graphics and it is where many cutting edge interface ideas arise. Based on the observation of interfaces of popular 3D games, several works dedicated to game design (e.g., Rouse (2000)) and design guidelines for virtual environments (Kaur, 1998; Shneiderman, 2003), we summarize what we consider to be the main 3D design guidelines:

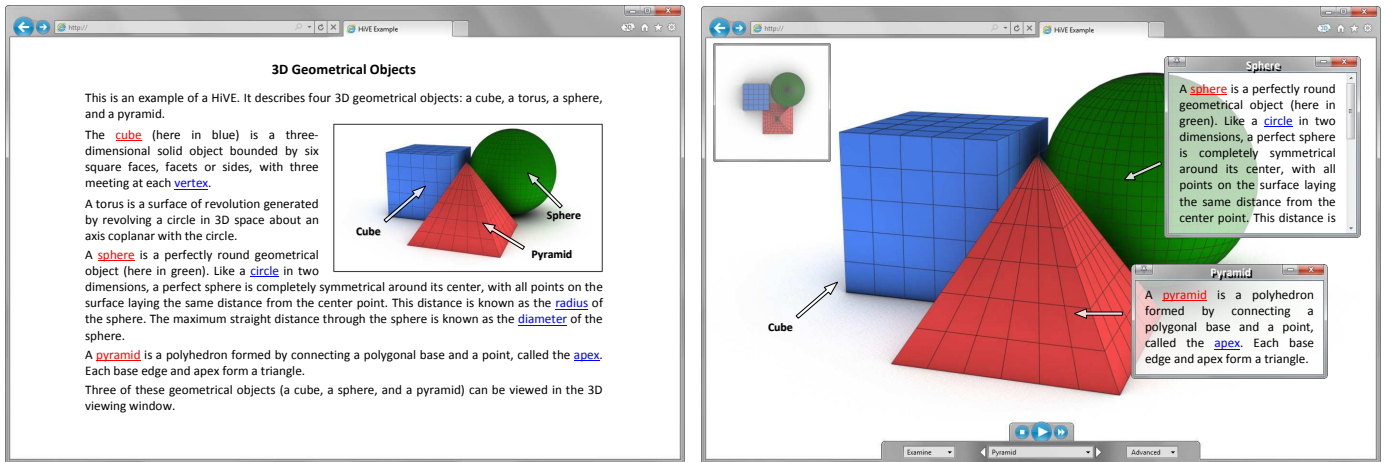


Figure 4: The same information presented in the hypertext and the 3D mode of the dual-mode user interface.

- **Text:** Keep text readable, give users control over explanatory text and let users select for details on demand;
- **Help:** Help should be embedded throughout the interface. It should allow users to click on any new UI component for an explanation;
- **Navigation:** Minimize the number of navigation steps for users to accomplish their tasks, simplify user movement (keep movements planar, use collision detection), allow teleportation (selecting destination in an overview);
- **Wayfinding:** Provide overviews so users can see the big picture (e.g., maps), provide history keeping (undo/redo);
- **Manipulation:** Provide users with manipulators, exploit knowledge about real world (e.g., gravity) and natural behaviour of objects (e.g., solidity);

The presented Web usability principles and 3D design guidelines will guide our design decisions throughout the our user interface development process.

### 3. Proposed User Interface Design

2-Layer Interface Paradigm (2LIP) (Jankowski and Kruk, 2008; Jankowski and Decker, 2009) was our first design iteration of a user interface that combines the advantages of interactive 3D graphics with the most useful features and conventions of the hypertext-based Web. The paradigm assumes that building 3D Web user interfaces involves the integration of two layers: (1) the background layer is a 3D scene; (2) the foreground layer, above the 3D view is the hypertextual content. Hyperlinks are used for navigation in the 3D scene (see the more detailed description of 2LIP in Figure 3).

This evaluation of our 2LIP paradigm (Jankowski and Decker, 2009), its further informal user studies, observations of people using 2LIP-based Copernicus, and expert reviews confirmed that 2LIP's very simple hyperlink-based interaction model can fit into the existing Web infrastructure and is suitable

for Web users - the users that typically have little experience with 3D graphics. On the other hand, we detected some usability problems, for instance: hypertextual content could only be displayed in a transparent form (readability issues); one could not reverse the index in Copernicus and click on an element of the 3D scene and be taken to the relevant part of the article. Moreover, HCI experts pointed out that having a 3D graphics underlay invites interaction and having rich, immersive, and enjoyable experiences.

Therefore, the concept of 2LIP has been substantially re-elaborated - we decided to develop a new user interface for information spaces integrating hypertext with 3D graphics called Dual-Mode User Interface (DMUI) that has two modes between which the user can switch anytime (see Figure 4):

1. **Hypertext "don't-make-me-think" Mode** driven by simple hypertext-based interactions, where a 3D scene is embedded in hypertext;
2. **3D "take-me-to-the-Wonderland" Mode**, which immerses the hypertextual annotations into the 3D scene.

The aforementioned 2-Layer Interface Paradigm usability problems were corrected (e.g., readability issues) and further functionality was added (e.g., introduction of annotations and the Mini-Map in the "free navigation mode", now called the "3D mode"). In the following we will describe in detail the two modes of the dual-mode user interface.

#### 3.1. Hypertext Mode

3D Web was (and still is) believed to have potential to be the next step in the Web's evolution, since it could benefit from graphics hardware and provide users with new and exciting experiences. Nevertheless, while various approaches have been proposed (most notably VRML/X3D), they have never seen much serious widespread use. One reason for the limited acceptance is the lack of 3D interaction techniques that are tailored for the hypertext-based Web interface.



Our approach to this problem is the hypertext mode of our interface (see Figure 4 (left)). This mode introduces a level of 3D-based interactivity and couples it with well adapted hypertext-based interactions. Our intention was to create, based on the Nielsen's (2000) and Krug's (2005) work on Web usability, a "don't-make-me-think" type of user interface. In the following we will describe the components of the hypertext mode of DMUI: hypertextual information and the embedded viewing window, where the 3D content appears.

### 3.1.1. Hypertextual Information

We define hypertextual information as an information set that can contain: textual information, non-textual information (e.g., static and animated images, audio, video), interactive information (e.g., flash interactive illustrations), navigational means (e.g., hyperlinks). In the hypertext mode, hypertextual information is the primary information carrier. It is possible to read it without any interaction with a 3D scene - the information is not embedded into the 3D scene, but rather presented in a concise form familiar to the Internet users. Compared with standard hypertextual information that can be found e.g., on the Web, the hypertext mode of DMUI introduces two new user interface components/mechanisms: *3D-hyperlinks* and *hypersections*.

**3D-Hyperlinks.** Ben Shneiderman and his colleagues first proposed the idea of having the text itself be the link marker (a concept also known as embedded menus or illuminated links) (Koved and Shneiderman, 1986). Since then, people using Web applications have become intimately familiar with this navigational user interface component.

In our user interface, hyperlinks constitute not only a mechanism for navigation between hypertext documents, but also for navigation within 3D scenes. If a 3D scene contains a viewpoint node named *viewpoint1*, selecting a hyperlink connected to this viewpoint should smoothly animate the camera from its current position to the selected vantage point. By introducing 3D links, we aim to provide users with the ability to view 3D content from different points of view with a single mouse click.

Visual cues are given as to where source anchors are located in a document. We use light blue highlighting as the default color for "hypertext links" and light red highlighting as the default color for "3D links" (see Figure 4 (left)). Both types of links can be embedded in hypersections.

**Hypersections.** Hypersections define sections of hypertextual information in a document; they are analogous to divisions in an HTML document defined by the `<div>` tag that is often used to group block-elements to format them with styles. The difference is that hypersections are not only designed to contain the hypertextually structured information. They are also designed to be:

- *Annotations of related 3D objects:* Hypersections become annotations when the user switches to the 3D mode;
- *Links between hypertextual and 3D information:* When a mouse cursor passes over a hypersection, the hypersection and the corresponding object in the 3D viewing window are automatically highlighted and the mouse cursor

changes its shape; the user is given visual cues as to what information is related to what object and where the related object is on the scene.

- *Navigational UI components:* Pressing the left mouse button over the hypersection animates the camera from its current position to the corresponding object and optionally triggers the camera movement that may be defined for the given object by the author of the scene (e.g., orbiting) (Jankowski, 2012a).

We believe that 3D-hyperlinks and hypersections can greatly facilitate the interaction. We wanted to make possible for users with little knowledge of 3D interaction techniques to browse a 3D scene simply by making a single mouse click.

### 3.1.2. 3D Viewing Window

The hypertext mode of DMUI builds upon well established principles for including graphical resources in hypertext. For example, in the HTML code, we use the `<img>` tag to link to the image; the browser then displays the image where the tag occurs with the text flowing around the image (it is default behavior; CSS is usually used to define the appearance and layout of the page). Similarly, in the hypertext mode, the 3D viewing window is displayed by the browser "inline" with presented hypertextual information. This window renders 3D scenes through which the users can navigate and in which they can interact with objects in real time.

There is also a practical, as well as aesthetic, reason for such design decision. As the size of the 3D viewing window is a factor in the time it takes to render the 3D scene, making the window smaller may make animations and movement through the scene noticeably smoother, especially for more complex virtual environments.

**Selectable Objects.** Hypertext browsers often draw hyperlink objects in a different color so that users can easily find them. The Dual-Mode UI (in both hypertext and 3D mode) provides two mechanisms to make certain that users could find selectable objects (objects in the scene that are designated as anchors for links to related hypertextual information or external resources) in a virtual environment:

- They can be identified using labels, short textual descriptions connected to their referent objects with lines extending into the virtual scene (these labels are also links to more comprehensive explanations displayed in the 3D mode);
- Cyclic temporary highlighting of all selectable objects allows users to view all objects of a possible interest in the visible area of the 3D scene. Some alternatives for highlighting selectable objects are described by Pichler (1993);
- When a mouse cursor passes over a selectable object, the object, its label and the related hypersection are automatically highlighted; information at both ends of the link is made more visible/prominent to indicate a particular scope of information to the reader.

*Easy Modeless Navigation.* While designing the hypertext mode of the dual-mode UI we have tried to accommodate as broad audience as possible by offering multiple ways to control the viewpoint, either by clicking selectable objects (easy for everybody), dragging the mouse across the 3D viewing window (more immersive, but also requiring some practice), or scrolling that gives people the option to see the whole scene in a guided tour (for more information on navigation refer to (Jankowski and Hachet, 2013)).

- We reserved a single left mouse button click in the 3D viewing window while the cursor is over a selectable object for *targeted movement navigation* (Mackinlay et al., 1990; Tan et al., 2001). The click on an object of interest smoothly animates the camera from its current position to the selected object and optionally triggers a predefined for the given object camera movement (e.g., orbiting). Such approach was used to preserve the primary interpretation of clicking in the web browser window as following a hyperlink. The technique is easy to use (users simply have to select the point to which they want to fly to), fast (it operates within controlled completion times), and cognitive-friendly (enhances the perception of the virtual environment); it can also be easily integrated with other techniques (Mackinlay et al., 1990; Hachet et al., 2008). On the other hand it has a major drawback: the target is always a selectable object.
- The second possible interaction in the 3D viewing window is mouse dragging (moving the mouse while holding the left or right mouse button down) and is reserved for *general movement navigation*. This approach should, if possible, emulate real world behaviors and take into account information about the scene and the task at hand. For example, geographic VEs often employ a walking metaphor of camera motion where user positions are restricted to the 2D plane of the terrain; examine metaphor is often used to view different sides of objects and it is suitable for tasks where the user's goal is to view an object as though he or she were holding it. If the user's goal can be reliably determined, the moding between the navigation techniques should be automated.
- There are some problems inherent in using general movement techniques. As they are designed to allow for unconstrained movement to any part of the virtual environment, the user may move to unknown and unwanted locations, look at things from awkward angles or miss seeing important locations/features (Fitzmaurice et al., 2008). As a result, one cannot ensure that the user receives the intended message. Like Galyean (1995) and Burtnyk et al. (2002), we believe that empowering the author to bring some structure to the interaction experience can make VEs more suitable for the *new-to-3D* users. Therefore, in addition to the above mentioned navigation techniques, our design balances the exploration methods with an ability to guide the user, while at the same time maintaining a sense of pacing or flow through the experience. We reserved

scrolling for *specified trajectory movement navigation*. As users can navigate on a page by scrolling it, when the cursor hovers over a 3D scene, the mouse scroll wheel can also be used to navigate between the viewpoints that are defined for this scene. Such guided or constrained 3D navigation limits the user's freedom while travelling through a virtual world. It constrains the audience's movement to interesting and compelling places, and thus avoids the classic problem of users getting 'lost-in-cyberspace'.

If the user wants to have more freedom in terms of viewpoint control, he or she can switch to the 3D mode using a button located on a browser's tool bar (see Figure 4). \*In the test application, this button is in the UI's bottom left corner.

### 3.2. 3D Mode

Having a 3D graphics underlay invites interaction and having rich, immersive, and enjoyable experiences. Yet, for sake of simplicity, the hypertext mode design limits interaction with that layer. This can lead to a problem with supporting immersion and the feeling of presence. However, what is immersion and presence and why do we need them?

In VR context, *immersion* refers to the objective level of sensory fidelity a VR system provides, while *presence* refers to the user's subjective psychological response to the VR system, related to the experience of "*being there*" (Witmer and Singer, 1998; Slater, 1999, 2003; Bowman and McMahan, 2007). A level of immersion depends only on the systems rendering software and display technology. It is objective and measurable - one system can have a higher level of immersion than another. It can best be attained by visually immersing the user with an HMD (Sutherland, 1968), CAVE (Cruz-Neira et al., 1993), or Fish Tank VR (Ware et al., 1993), by using stereo displays and head tracking. However, some level of immersion is also possible in the low-immersion Desktop Virtual Environments (Robertson et al., 1997).

The successful sensual immersion of the user in an imaginary 3D space is a very important part of the experience while interacting in a virtual environment. Achieving a close to real-life experience in a virtual world, creating a feeling of being there is crucial to give a potential virtual tourist the sensation of what the site is really like. Tan et al. (2001) assert that the level of immersion that the user experiences greatly affect the navigation task and performance. The more immersed the user is, and the more easily the user can mentally integrate information acquired, the greater the chances of efficient navigation (Robertson et al., 1997). Other benefits of immersion relate to the positive effect on spatial understanding and to a decrease in information clutter (Bowman and McMahan, 2007). Certainly, much of a presence has to do with a quality of the presented material and the manner in which the user experiences it. High-presence experience can take place while we are playing a well designed video game, watching a quality movie, or even while reading of good novel, in spite of the lack of visual or perceptual immersion.

Comparing to the hypertext mode, the 3D mode of the dual-mode UI was designed to make users feel more present in an

environment by increasing the level of immersion through increased *3D Viewing Window* size (Ni et al., 2006; Polys et al., 2005) that spans the entire available screen space. In this mode 3D graphics is the main information carrier. The 3D mode provides users with much more freedom with regard to the 3D task support - it was designed to support unconstrained interactive navigation through a 3D scene. Furthermore, in this mode hypertextual data relating to an environment is embedded into that environment. The design of this, what we call "*take-me-to-the-Wonderland*" mode, was inspired by the work on IRVEs Bowman et al. (2003) and the work on annotating 3D illustrations Sonnet et al. (2005).

In the following we will describe the components of the 3D mode of our interface: the viewing window, where the 3D content appears, hypertextual annotations, a dashboard designed to manage interactive browsing and navigation, and a mini-map designed to aid wayfinding.

### 3.2.1. 3D Viewing Window

In the 3D mode of DMUI, 3D scenes appear in the viewing window that spans the entire available screen space. A dashboard, a map and hypertextual information about an environment are rendered on top of the viewing window in *screen-space*, a 2D plane called HUD (head-up display) that overlays the 3D scene. Such technique ensures better legibility of the hypertextual information.

In general, with regard to functionality, the viewing window in the 3D mode is similar to the one in the hypertext mode. By default, it offers the same navigation metaphors (a viewpoint control technique that is best suited for a given task), the same behavior of selectable objects, etc. However, this mode provides a user with more than one "default" way of interacting with the 3D scene - the user can choose to use other methods of navigation (e.g., change from *walk* to *fly* or *examine*).

For example, a visitor of a virtual museum may be provided with the *walk* navigation metaphor that is the only way of interaction in the hypertext mode and the default one in the 3D mode; as the user may want to see the artifacts from different angles, he/she can change the metaphor to *examine* while being in the 3D mode of the Dual-Mode UI.

### 3.2.2. Hypertextual Information

One of the underlying premises of this research is that communicating information about 3D environments can be significantly improved by attaching annotations to the environment's objects. The 3D mode allows users to interactively recall and view the attached hypertextual information by clicking on labels connected to objects of interest during navigation. In response, the hypertextual information is presented in *screen-space* adjacent to associated objects in scrollable annotation boxes (as we have already mentioned, Polys (2006) examined existing approaches to information layout in IRVEs for *details-on-demand* interaction and suggested using *screen-space* techniques). In fact, these hypertextual explanations are hypersections from the hypertext mode of our UI. Consequently, the annotation boxes may contain information from hypertext, through images and movies, to multimedia animations. They

may also contain 3D hyperlinks for navigation within the 3D scene. Users may move and resize the annotation boxes, and toggle their visibility; they can also minimize them into labels (Gotzelmann et al., 2006).

In some sense, this idea of presenting additional information about an environment within that environment derives from the augmented reality (AR) concept (Feiner, 2002). Instead of creating a layer over a view of the real world, we put the information hypertext-based layers on the top of the computer generated three-dimensional scene.

Overlapping opaque annotations can obscure parts of a 3D scene users may need to see. Moreover, having two sources of information, visual in the background and (mainly) verbal in the foreground, users must make choices about what to attend to and when. On the one hand, they may need to focus their attention exclusively on e.g., textual description of an object without interference from the virtual scene. On the other hand they may need to time share or divide their attention between the object and its referent description.

To better support human attention, better maintain the fluency of work, and to improve workspace visibility without harming interaction performance, in a 3D mode of the Dual-Mode user interface, annotations are rendered as semi-transparent user interface objects (Harrison et al., 1995; Harrison and Vicente, 1996; Ishak and Feiner, 2004). However, as the contents of the virtual scene and annotations will always interfere with each other to some extent, the advantages of transparent layered displays must be weighed against the perceptual interference between the 3D scene in the background and the annotations in the foreground. To minimize visual interference, in our previous work we explored the effects of different screen space techniques of integrating text with animated 3D graphics on text readability (Jankowski et al., 2010) and we applied *positive billboard style* as a default technique for rendering annotations in the 3D mode of the Dual-Mode UI.

### 3.2.3. Dashboard

As we have already mentioned, for a given scene type and a task at hand, a designer should decide on the most intuitive mapping between input and interaction technique. However, very often there is insufficient input DOF for the task and user input must be moded. Therefore, the user has to be given explicit control of the different modes of navigation.

A graphical dashboard (in Figure 4 presented at the bottom of the viewing window) provides ready access to the most important 3D interaction tools. By default, this control panel should provide the methods of navigation based on examine, walk, fly and fly-to (POI); it should also allow the user to switch between the viewpoints that are defined for the 3D scene. If a hypertextualized virtual environment is designed in a way that allows users to add objects to the scene, manipulate and modify them, the users should be able to switch to the more advanced dashboard that provides such editing functionality. A user should be able to hide the dashboard to create more space in a 3D viewing window.

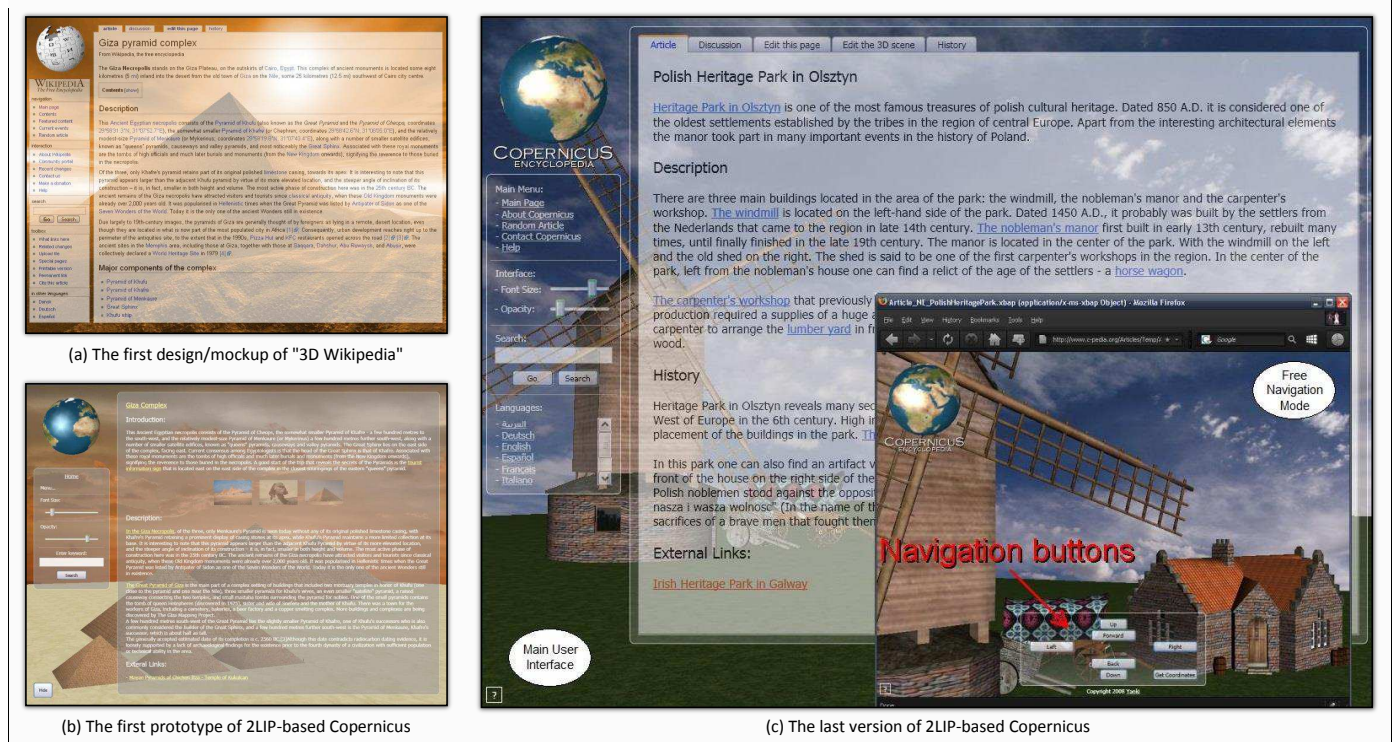


Figure 5: The evolution of 2LIP-based Copernicus: (a) the first design/mockup of "3D Wikipedia" created using GIMP for one of the weekly meetings of Corrib Clan (April 2007); (b) the first working prototype of 2LIP-based Copernicus (September 2007); (c) the last version of 2LIP-based Copernicus - it had some authoring features, free navigation mode (b), and implemented physics engine (July 2008).

### 3.2.4. Mini-Map

In addition to the difficulties of controlling the viewpoint, there is a problem of wayfinding, especially in large virtual worlds. It is related to how people build up an understanding (mental model) of a virtual environment. This problem may manifest itself in a number of ways Darken and Sibert (1996):

- Users may wander without direction when attempting to find a place for the first time.
- They may then have difficulty relocating recently visited places.
- They are also often unable to grasp the overall topological structure of the space.

Efficient wayfinding is based on the navigator's ability to conceptualize the space. Maps proved to be an invaluable tool for acquiring and maintaining orientation and position in a real environment. According to Darken and Sibert study (1996), this is also the case in a virtual environment.

Therefore, also influenced by computer games, we decided to include a mini-map to the 3D mode of the Dual-Mode User Interface (see the miniature map, placed in a top left corner of the 3D mode of the UI - Figure 4(right)). It displays terrain, important locations and objects. It dynamically updates the current position of the user with respect to the surrounding environment.

## 4. Copernicus

In the previous section we described in detail the dual-mode user interface design for information spaces combining hypertext with interactive 3D graphics, the interface largely based on the concept of 2LIP. To put our new design into practice and evaluate it, we needed a platform for the experimentation and for the assessment of both hypertext and 3D modes of our interface, a system that would allow to find a balance between 3D interaction techniques and well established hypertext interactions. Therefore, the implementation of Copernicus underwent many enhancements. From the 2LIP-based prototype application that mimicked Wikipedia's look and feel, it grew to a full-fledged wiki engine on top of which we started to implement our Dual-Mode user interface. Copernicus became a true wiki - a simple to use, asynchronous, Web-based collaborative HiVE authoring system, where users are allowed to create and later modify hypertextualized virtual environments.

### 4.1. User Interface

The new version of Copernicus implements the Dual-Mode UI, so it has two modes between which a user can switch any-time: the driven by simple hypertext-based interactions "hypertext mode", where a 3D scene is embedded in hypertext and the more immersive "3D mode", which immerses the hypertextual annotations into the 3D scene.

Figure 6 shows several examples of the hypertextualized virtual environments displayed both in the hypertext and the 3D





Figure 6: Examples of hypertextualized virtual environments displayed both in the hypertext and the 3D mode of the Copernicus's user interface. The button for switching between the two modes is located in the bottom left corner of the UI.

mode of the interface. One of the examples is a virtual environment representing a museum of World of Warcraft. It allows users, while in the hypertext mode, to read about the collection and easily navigate through the rooms of the virtual museum using 3D hyperlinks and hypersections. The same hive also allows users to walk through the rooms of the museum and view the exhibition; the user can click on object's labels to read more comprehensive explanations displayed in annotation boxes.

#### 4.2. Author Interface

Copernicus is a wiki - it implements an editor, i.e. a tool that can be used to create and update HiVE documents directly on the Web. Our testbed offers a workflow that unites 3D design with wiki-way editing approach. It ensures that the hypertextual information, 3D graphics, and underlying interaction model can be simultaneously evolved, resulting in the creation of rich interactive hypertextualized virtual environments.

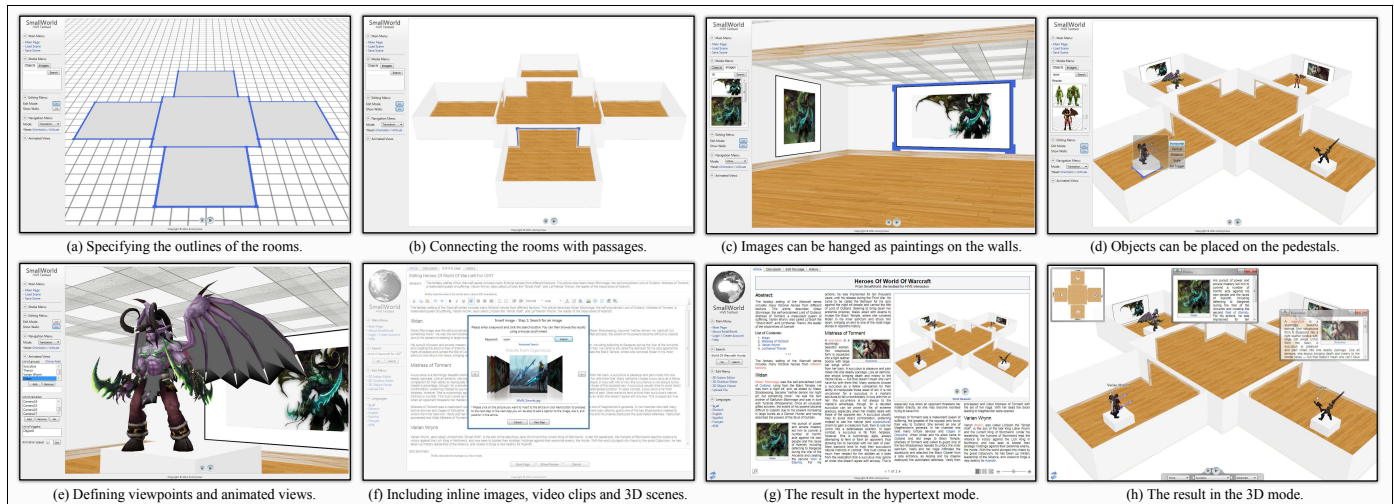
In Copernicus, different types of information, from text, through images and video clips, to 3D graphics, can be easily collected, linked, and later made available as integrated information spaces in the hypertext based environment (the hypertext mode) or within the context of a virtual environment (the

3D mode). Our goal was to engineer an authoring environment that will allow to create HiVEs without any specialized training. In Figure 7 we describe, step-by-step, the process of authoring an article for Copernicus: from preparing multimedia content, through creating a 3D scene, creating 3D viewpoints, creating and writing a new article, including images, video clips and 3D scenes in the article, creating 3D and hypertext links in the article, to saving the article in the database. It is important to note that Copernicus was used to create the virtual museum of Warcraft for the user study described in this article.

#### 4.3. Implementation

Copernicus is implemented using Windows Presentation Foundation (WPF). In WPF, the underlying graphics technology is DirectX: from displaying plain text to rendering complex 3D graphics, all the drawing work is processed through the DirectX pipeline. As a result, the application can benefit from hardware acceleration. WPF uses Extensible Application Markup Language (XAML) as the user interface markup language. XAML allows the programmer to separate the user interface definition from the underlying business logic; it allows





### Preparing Multimedia Content

The author can upload images (bmp, png, jpg), video clips (wmv, avi) and 3D objects (at the moment it is possible to import zipped Collada 3D objects from Google 3D Warehouse).

### Creating a 3D Scene

The task of creating a 3D scene from scratch is difficult. To simplify the problem, we focused on the creation of editing environment, that would allow to build 3D scenes based on a library of existing objects. We have built two distinct 3D editors for indoor and outdoor scenes. Indoors, the virtual environment is created as a set of interconnected rooms. Outdoors, the virtual environment is a landscape of plains and hills upon which the author can place buildings, trees, cars, etc.

Our 3D indoor editor allows a user to create simple single-floor museum-like environments. The procedure is straightforward and involves: (1) changing the mode to the edit mode through a toggle on/off button; (2) specifying the outlines of the rooms by drawing and dragging squares on the surface that represents the floor; (3) making the walls visible by using the Show Walls toggle button; (4) connecting the rooms with doors or passages – clicking with the right mouse button on a wall brings up a context menu that allows to create such connections; (5) once the layout is finished, the author can decorate the rooms: (5a) the user can search for images in the Media menu – the found images can be hung as paintings on the walls by simply selecting the image in the results list and clicking on some space on the wall; the created paintings can be then repositioned and scaled; (5b) the user can also search for 3D objects – the objects can be placed on the pedestals which the user has to create beforehand (by using the floor context menu); the objects can be then positioned and scaled (the user has to click twice on the object while being in the edit mode); (6) finally, when the author is happy with the design, the 3D scene can be saved in the database.

### Creating 3D Destinations

Once the author has created the 3D scene, the interesting places in that scene can be selected and marked as 3D destinations. Our 3D editor supports defining viewpoints and animated views for such points of interest in order to let the author later connect them to the links in the text and ultimately allow easy link-based navigation for the end-users.

**Viewpoints** A viewpoint is simply a specific camera position and orientation defined by a pair of x, y, z coordinates, that is, a specific view of the 3D scene. Viewpoints can be defined by navigating to the desired viewing position and then defining the current position as a destination.

**Animated Views** Static viewpoints very often do not show "3D-ness" of virtual objects – as Andy van Dam mentioned: "If it ain't moving, it ain't 3D". Therefore, the editor allows to define a destination anchor that is an animated view of a model. To produce such anchor, the author has to create two or more viewpoints as discussed earlier and use them as a set of key frames in an animation of the camera. The spline interpolation is used as an animation's interpolation method since it produces smooth camera motions.

**Selectable Objects** An object on a scene can also be made a source anchor, that is, a selectable object. In the World of Warcraft museum shown here, a selectable object is being created from the Illidan object. Such object can be then connected to a viewpoint or an animated view. As a result, a click on the object will smoothly animate the camera from its current position to the selected object and optionally trigger an animated view.

### Creating and Writing a New Article

To create a new page for SmallWorld the author needs to type the title of the new article in the search box and then click Go. If the Search Results page reports "You can create a new article about" followed by the article name in red, then the user can click the red article name to start editing the article. We developed a WYSIWIG-type rich-text content editor that allows

for the rich formatting of text, including common structural treatments like lists and headings, and formatting treatments like bold and italic text.

### Including Images, Video Clips and 3D Scenes in an Article

SmallWorld supports including inline images, video clips and 3D scenes in text documents in a similar way. The procedure is simple: the user needs to select a cursor position in a document, click on one of the Add-Media type buttons located in the toolbar, search for a resource, and finally insert the resource to the document.

### Creating Links and Hypersections in an Article

Links, both internal links to other articles and 3D links to specific views of a 3D model, can also be defined interactively in the 3dWiki's article editor. The procedure is similar for both types of links: (1) First, the user has to select a region of text to define the region as a source anchor; (2) Then the user needs to click on Add a Link or Add a 3D Link button located in the toolbar – this brings up the Create New Link dialog box; (2a) For an internal link, the user has to provide the title of the page he or she wants to link to; (2b) For a 3D link, the user has to choose a "3D destination" from the list of viewpoints and animated views defined for the given (used in the article) 3D scene; (3) Finally, when the user clicks on the Add Link button, the page appears featuring a new hyperlink. If the link points to another wiki page, it will be highlighted in light blue; if the link's purpose is to trigger an animation on a 3D scene, it will be shown in light red.

### Saving an Article

When the article is done, the author can press "Show Preview" to take a look at how the page will appear in the default hypertext interface mode. If there are no errors, the author should write a short edit summary in the text field below the edit-box and press the "Save Page" button. The article should immediately be visible to all users of the wiki. To re-edit the page, all the user needs to do is to click on the "Edit" tab at the top of that page.

Figure 7: Step-by-step process of authoring an article for Copernicus.

to simply construct user interfaces using well-defined UI elements: from simple 2D controls (such as panels or buttons), through more sophisticated controls for displaying documents, to interactive 3D drawings.

WPF applications can be deployed on the Web. XAML Browser Applications (XBAPs) are programs that can be hosted inside a web browser. If the .NET framework is installed, users can use an XBAP just like a Flash-enhanced web page. The only but very significant shortcoming of using XBAPs is that they can only be rendered using Internet Explorer or Firefox on Windows. As WPF/XAML features allow to easily combine hypertextual information and 3D graphics, we have decided to use it to implement our user interface designs. The project's source code is available under GPL license.

We should note here that the choice to use the WPF/XAML approach was made before WebGL became more widely available. We are currently working on the new WebGL and standard HTML/CSS/JavaScript implementation that is based on *Declarative 3D* principles (see the Future Work section).

### 4.4. Copernicus in Use

Like many other researchers, we recognize the importance of observing real users in real-world contexts in addition to the more controlled laboratory contexts. We had several opportunities to observe novices interacting with the dual-mode user interface using Copernicus. Most observations were made in primary and secondary schools (one of the objectives of DERI is to popularize the knowledge of science and promote engineering among young students) as well as at local science fairs. We also observed individual users at their personal work spaces (mostly at DERI and NUIG). These users were free to access and navigate any content they preferred; they then provide us with feedback on the positives and negatives of the system.

The comments from the participants of this initial evaluation were extremely positive. The study has shown that due to only a slight modification of hypertext-based interface paradigm, the users had no problems interacting with Copernicus. The simplicity of the interaction techniques, essentially a single click in the hypertext mode and a drag action in the 3D mode, were immediately understood and usable by all our users.

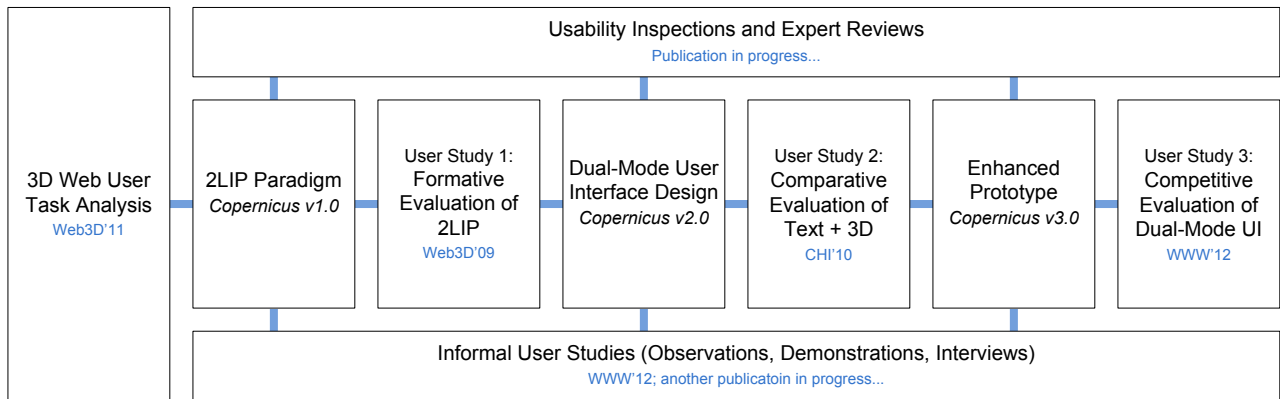


Figure 8: Our methodology for the design and evaluation of the Dual-Mode user interface.

## 5. Evaluation

A user interface for hypertextualized virtual environments should support interaction styles that are quite different from the ones that can be found in common traditional user interfaces such as GUIs. Therefore, to design and evaluate an interface for HiVEs, we decided to use a systematic and methodical approach based on the sequential evaluation developed by Gabbard et al. (1999) - structured, iterative method for user-centered design and evaluation of VE user interaction (Figure 8 illustrates the approach).

As a prerequisite to the effective user interface design is an understanding of the users and their tasks (the tasks for which that interface will actually be used), we first focused on an understanding of the fundamental tasks users may engage in while exploring Web-based 3D virtual environments. We introduced a "*taskonomy of 3D Web use*", a simple taxonomy we used later to identify the tasks that the user interface for HiVEs must support (see Figure 2 and (Jankowski, 2011)).

This research led to the development of 2-Layer Interface Paradigm (2LIP) (Jankowski and Decker, 2009), our first attempt to build an interface that tightly combines hypertext with 3D graphics. The formative evaluation of our 2LIP paradigm (see Figure 3), its further informal user studies, observations of people using 2LIP-based Copernicus, and expert reviews confirmed that 2LIP's very simple hyperlink-based interaction model can fit into the existing Web infrastructure and is suitable for Web users - the users that typically have little experience with 3D graphics. On the other hand, we detected some usability problems, for instance: hypertextual content could only be displayed in a transparent form (readability issues); one could not reverse the index in Copernicus and click on an element of the 3D scene and be taken to the relevant part of the article. Moreover, HCI experts pointed out that having a 3D graphics underlay invites interaction and having rich, immersive, and enjoyable experiences.

Therefore, we decided to develop a new user interface for information spaces integrating hypertext with 3D graphics. We designed the Dual-Mode user interface and developed the second version of Copernicus. The detected usability problems were corrected (e.g., the problem of readability of hypertextual content) and further functionality was added (e.g., introduc-

tion of annotations and the Mini-Map in the "*free navigation mode*", now called the "*3D mode*").

Further usability inspections and user studies brought us back to the problem of readability of text annotations rendered on top of an animated 3D background in the 3D mode of DMUI (the problem manifested itself already in the main user interface mode in 2LIP). As there are no clear guidelines on how best to display readable text on top of an animated 3D background, we decided to design an experiment, in which we investigate the effects of varying: (a) text drawing style, (b) image polarity, and (c) background style on readability of text (Jankowski et al., 2010). The outcomes of the evaluation were used to improve the usability of an evolving user interface design: we decided to use the *positive billboard style* as a default technique for rendering annotations in the 3D mode of the Dual-Mode UI.

Finally, to characterize the Dual-Mode user interface in terms of its efficiency and usability, we conducted a user study, where we compared our interface to two other, currently most common user interfaces for integrated information spaces, where text and 3D graphics data are simultaneously available and linked (Jankowski and Decker, 2012). In the following we will describe this study in more detail. First we discuss the evaluation setup used in the experiment. This discussion is followed by the description of the evaluated user interfaces and the procedure used in this evaluation. Finally we present and discuss the results of the experiment.

### 5.1. Participants

20 students, researchers and members of staff with normal or corrected-to-normal vision participated in the experiment. 5 of the participants were native English speakers. 5 of our participants were female. The participants ranged in age from 27 to 42, with 12 participants in the 27-30 range and 8 participants in 30-42 range. 8 participants were PhD students, 9 had higher education, and 3 were post-doctoral researchers. All of the participants were familiar with hypertext navigation; 3 of them had no or very little experience with 3D interaction, 13 had some experience navigating 3D scenes (playing 3D games sporadically), and 4 had considerable experience in 3D navigation from playing 3D computer games. Subjects were given gifts worth 15€/20\$ for their participation. Additionally, an iPod was awarded to one of the 5 best performing participants.

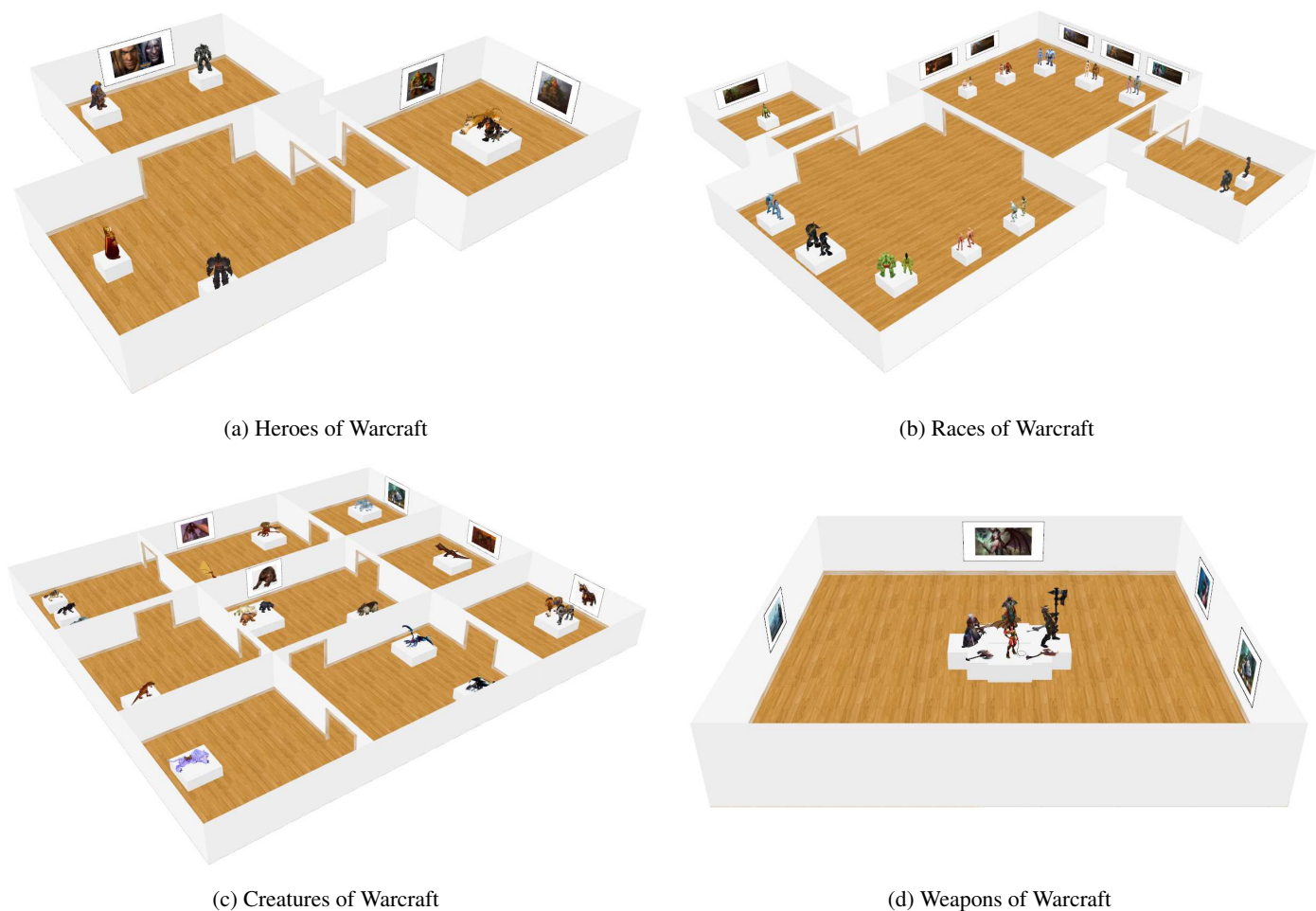


Figure 9: The 3D scenes used in the evaluation.

## 5.2. Apparatus

The experiment was conducted on the Intel Core 2 Extreme laptop computer equipped with 4GB of memory, GeForce 9800M GTX graphics card, connected to a 24-inch widescreen LCD display running at 1920x1200 resolution. Input devices were a standard 3-button optical mouse and a keyboard. The computer operating system used was Microsoft's Windows 7. The test application used for the evaluation was developed based on the Copernicus source code; the content for the evaluation (see next section) was authored using our 3D wiki as well. The participants sat in a lab which was illuminated by overhead fluorescent lights.

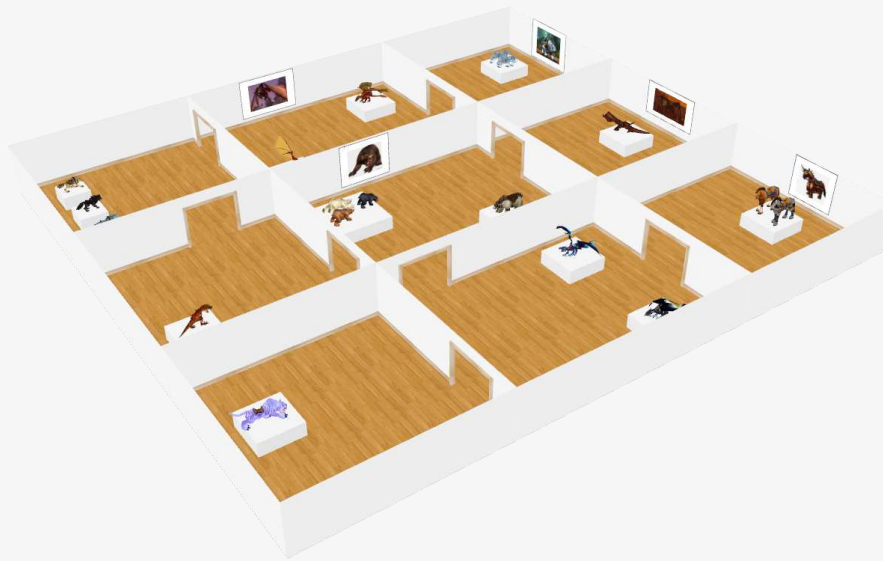
## 5.3. Stimuli

For the purpose of this experiment we prepared a virtual museum featuring heroes, races, creatures, and weapons from the fantasy setting of the World of Warcraft (WoW) game; this choice was made to prevent the influence of previously gathered knowledge. The museum was divided into four exhibitions, one for the training session (heroes) and three for the main experiment (races, creatures and weapons). Each exhibition conveyed integrated hypertextual and 3D visual information. The virtual worlds created for the study were simple single-floor museum-like environments, populated with 3D objects collected from

Google 3D Warehouse and images/painting (see Figure 9). Below follows a description of each exhibition:

- *Heroes of Warcraft* - an exhibition of four notable characters from the Warcraft game. The article consisted of about 480 words and 4 images. The 3D scene consisted of three rooms populated by 5 objects and 5 paintings (see Figure 9a).
- *Races of Warcraft* - an exhibition of the twelve playable races of the Alliance (i.e. Humans, Night Elves, Dwarves, Gnomes, Draenei, and Worgen) and the Horde (Orcs, Trolls, Tauren, Forsaken, Blood Elves, and Goblins) factions from WoW. The article consisted of about 1350 words and 12 images representing each race. The 3D scene consisted of four rooms populated by 12 race objects and 12 paintings (see Figure 9b).
- *Creatures of Warcraft* - an exhibition of common creatures that can be found in the World of Warcraft, such as bears, saber cats, drakes, and wolves. The article consisted of about 920 words and 3 images; each creature was characterized by strength, agility, and intellect values. The 3D scene consisted of nine rooms populated by 14 objects and 13 paintings (see Figure 9c).





## Article

### Bears

Bear is a very common, bulky, furry, four-legged beast that lopes along on all fours, but may occasionally rise up on hind legs. Most bears are hostile when approached, despite their sometimes cuddly appearance. They tend to be larger than human sized at their smallest.

#### Grizzly Bear

Grizzly bears are massive carnivores who weigh more than 800 kg and stand nearly 2.7 meters tall when they rear up onto their hind legs. They are bad-tempered and territorial. A grizzly bear attacks mainly by tearing at opponents with its claws. They are found in cold forests, and are found either solitary or in pairs.

— Strength 21, Agility 9, Intellect 10

#### Polar Bear

The polar bear (aka arctic bear) is a species of bear with a white fur. These long, lean carnivores are slightly taller than grizzly bears. Polar bears fight just as grizzly bears do. A polar bear's white coat bestows camouflage in snowy areas. They are found on cold planes, either solitary or in pairs.

— Strength 25, Agility 10, Intellect 14

### Saber Cats

Saber cats are a special type of large fanged cat, similar to the (extinct) sabertooth cats of our world. They come in a variety of colors and have either spotted or striped coats, the striped type being the more common. Night elves carry a special affinity for sabers and use them as mounts.

#### Frostsaber

Frostsaber cats are similar in appearance to a real-world white tiger. They are intelligent and loyal when domesticated, but feral and aggressive when left in their natural environments. The frostsaber cat is a fearless opponent, its thick fur having the resilience to shrug off minor damage. It uses its fangs and claws against those foolish enough to engage it in melee combat.

— Strength 15, Agility 26, Intellect 21

#### Winterspring Frostsaber

A purple frostsaber native to Winterspring is bigger and stronger than frostsaber. In combat, these frostsaber cats are deadly foes, leaping from the shadows to rend their foes with its fangs and claws.

— Strength 21, Agility 25, Intellect 25

### Drakes

Drakes are dragons that are still relatively young in comparison to their ancient brethren. Drakes are almost always flying while moving, while nether drakes tend to walk.

### Red Drake

Red drakes are the adults of the red dragonflight, and most are fully grown at this stage. Once friends to nearly every living creature, red drakes are now much more suspicious, tending to seek only the company of their own kind. They are also friendly to humans, elves, and their other traditional friends and allies. Red drakes have lithe frames; though large, the crimson drake is graceful and noble in form. The fiery horns atop the dragon's skull resemble a crown upon the head of a worthy king. Golden eyes analyze every motion, obviously judging whether one might be friend or foe.

— Strength 42, Agility 22, Intellect 51

### Nether Drake

Nether drakes are the immature members of the netherwing dragonflight that are believed to be the spawn of Deathwing. Without Deathwing's guidance, they now seek to find their identity beyond the shadows of their father's destructive heritage. Living in Draenor's ruins, now known as Outland, the nether drakes are available as flying mounts to swiftly transport players around the shattered world.

— Strength 51, Agility 21, Intellect 63

### Proto Drake

The proto-drakens were sort of the early version before dragons really evolved into the sapient, beautiful majestic creatures that they are known today.

— Strength 58, Agility 15, Intellect 43

### Horses

A horse is a large four-legged, four-hoofed animal with a long face, substantial body, and fast, narrow legs. This is the primary mount for humans. Warhorses are trained and bred for strength and aggression. Human knights ride such beasts into battle.

— Strength 10, Agility 15, Intellect 15

### Raptors

Raptors are large, aggressive theropods (this means, they stand upright on two large hind legs in a bird-like fashion). As they seem to be very intelligent for beasts, Trolls use them as a mount.

— Strength 21, Agility 18, Intellect 17

### Wolves

Wolves are carnivorous, furry, canine beasts that often run in packs. Wolves encountered in the wilderness are typically hostile and attack characters that approach them too closely. These cunning and ferocious creatures have found a welcome home among the orcs, who admire their endurance and keen survival instincts.

### White Wolf

White wolves can be found in cold forests. Unlike other wolves, they like to hunt during a day.

— Strength 18, Agility 14, Intellect 18

### Black Wolf

These wolves like to live and hunt in dark forests, where their fur makes them hard to detect.

— Strength 16, Agility 16, Intellect 18

### Brown Wolf

Brown wolves are the smallest of the wolf kind.

— Strength 14, Agility 18, Intellect 10

## Questions

### S:H-3D

(1) Find the room with the creature with the strength value of 10. How many rooms are connected to this room?

- 1:1
- 0:2
- 0:3
- 0:4

(2) Find the creature with the agility value of 9. What is the price of the painting depicting this creature? (The price is displayed in the right bottom corner of the painting.)

- 0:160
- 0:200
- 0:540
- 0:There is no price on this painting

(3) Find the creature with the intellect value of 15. This race is afraid of some other race. What is the price of the painting depicting this other race? (The price should be displayed in the right bottom corner of the painting.)

- 0:160
- 0:200
- 0:540
- 1:There is no price on this painting

### S:3D-H

(1) There are three types of drakes in the museum. Find the one with the front legs much shorter than the back legs. What is the agility of this type?

- 0:Below 14
- 1:Between 14-18
- 0:Between 18-21
- 0:Above 22

(2) Find the room with two paintings connected to the room with three pedestals/stands. What is the strength value of the creature in the room with two paintings?

- 0:10
- 0:21
- 1:42
- 0:51

(3) Find the room connected to the room with horses. What is the intellect of the presented in this room species?

- 0:25
- 0:43
- 1:63
- 0:47

### C:H-3D

(1) There are two saber cat species in the museum. What is the eye color of the stronger kind? Find the paintings behind the models of saber cats. What are the hair colors of the hunters depicted on those paintings?

- 0:Yellow; hair colors: red and green
- 1:Yellow; hair colors: green and blue
- 0:Green; hair colors: red and green
- 0:Green; hair colors: blue and green

(2) Find the strongest creature in the museum. What is the color of the creature's eyes?

- 0:Yellow
- 0:Blue
- 0:Red
- 1:Green

(3) There are three types of creatures with the strength value of 21. Find the one that has the biggest jaws. This race dislikes some other race. What is the eye color of this other race?

- 0:Yellow
- 1:Blue
- 0:Red
- 0:Green

### C:3D-H

(1) In the museum, there are five paintings with a price displayed in the right bottom corner of a painting. Find the most expensive one. What is the agility of the depicted on the painting creature?

- 0:10
- 1:15
- 0:18
- 0:26

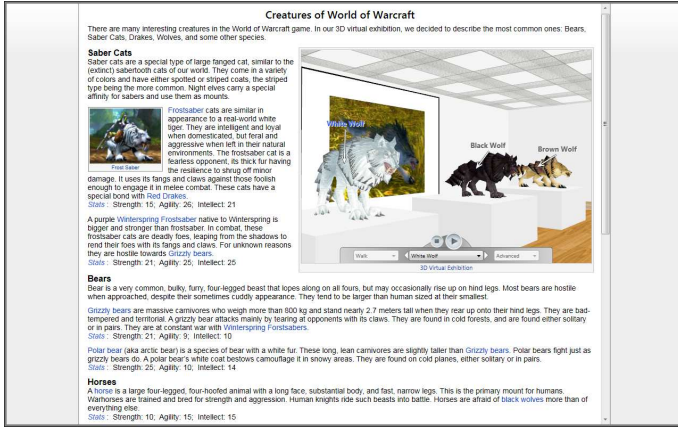
(2) Assuming that the room with Red Drakes is to the north, and the room with Nether Drakes is to the south, find the strongest race among all the races on the west side of the museum. What is the intellect of this creature?

- 0:18
- 0:21
- 0:25
- 1:There are two creatures with the same strength!

(3) Assuming that the room with Red Drakes is to the north, and the room with Nether Drakes is to the south, find the race that is furthest north-west. What is the intellect of this creature?

- 1:Below 12
- 0:Between 12-15
- 0:Between 16-19
- 0:Above 20

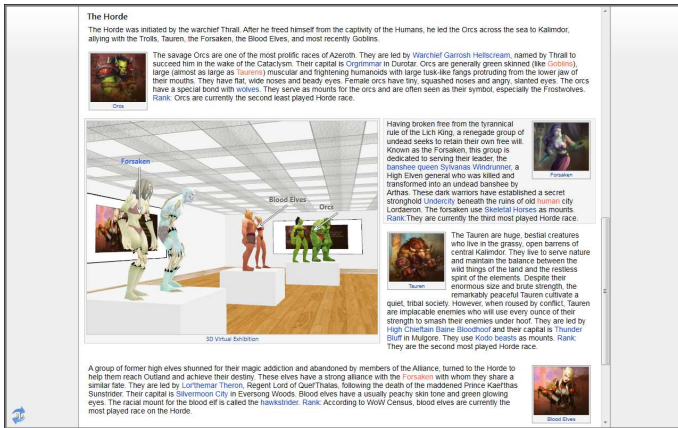
Figure 10: Article and tasks for the "Creatures of Warcraft" HiVE.



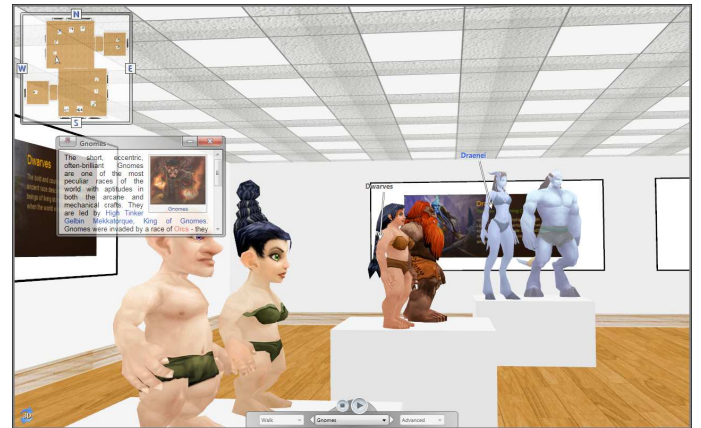
(a) Creatures of Warcraft in the hypertext UI



(b) Weapons of Warcraft in the 3D UI



(c) Races of Warcraft in the hypertext mode of the dual-mode UI



(d) Races of Warcraft in the 3D mode of the dual-mode UI

Figure 11: The exhibitions presented in the evaluated user interfaces.

- **Weapons of Warcraft** - an exhibition of weapons (such as swords and axes) from the Warcraft universe. The article consisted of about 1060 words; each of the 9 weapons was characterized by damage, bonuses (e.g. to strength, agility, etc.), and a price. The 3D scene consisted of one room; all 9 objects were positioned in the center of the room (see Figure 9d).

According to the classification of virtual worlds Darken and Sibert (1993), all our environments are dense (relatively large number of objects and cues in the space) and static (the positions and values of the objects do not change over time). Moreover, the exhibition of weapons is a small world (a world in which all or most of the world can be seen from a single viewpoint), while all other environments are large (there is no van-tage point from which the entire world can be seen in detail).

#### 5.4. User Interfaces

As we have already mentioned, the hypertext *"don't-make-me-think"* mode of the dual-mode user interface was inspired by the state-of-the-art practice of embedding 3D scenes as part of an HTML page. The design of what we call 3D *"take-me-to-the-Wonderland"* mode was inspired by the work on IRVEs (Bowman et al., 2003) and the work on annotating 3D illustrations (Sonnet et al., 2005). To characterize the dual-mode

user interface in terms of its efficiency and usability, we decided to compare it to these two inspirations that are currently user interfaces of choice for integrated information spaces, where text and 3D graphics data were simultaneously available and linked.

- **Hypertext User Interface** - this interface was created by modifying the hypertext mode of the dual-mode UI. Features like an ability to switch to 3D mode, hypersections, and 3D hyperlinks were disabled. On the other hand, the dashboard UI component was added to the 3D viewing window (see Figure 11a).
- **3D User Interface** - this interface was created by modifying the 3D mode of the dual-mode user interface. Features like an ability to switch to the hypertext mode and 3D hyperlinks were disabled (see Figure 11b).
- **Dual-Mode User Interface** - this interface integrates Hypertext UI and 3D UI into one modal interface. It allows users, while in the hypertext mode, to read about the collections and easily navigate through the rooms of the virtual museum using 3D hyperlinks and hypersections (see Figure 11c). The same UI, while in the 3D mode, also allows users to experience the 3D scenes that span the entire available screen space of the browser window; just like in the 3D UI, the user can walk through the rooms of the



museum and click on object's labels to read more comprehensive explanations displayed in annotation boxes (see Figure 11d).

The user interface designs evaluated in this study differed in the method used to integrate the textual information with the objects in the 3D scene. On the other hand, the interfaces allowed for the same interactive exploration of 3D scenes; there were no differences in the techniques that enabled a user to navigate the 3D scenes. Movement was confined to 'walk' mode; guided tour navigation (scrolling over 3D scene) was disabled; collision detection was used to prevent users moving through objects and walls.

### 5.5. Tasks

Different possible measures could be used to determine the effectiveness and usability of the evaluated interfaces. In choosing tasks for the study, we looked for ones that are both valid (resemble a 'real' act of browsing 3D content on the Web) and that are recognized for being able to detect significant differences. We decided to adopt tasks that were introduced by Chen et al. (2004) and were later successfully used by Polys et al. (2005, 2011) and Ni et al. (2006) to evaluate IRVEs. Thus, the participants performed 4 types of tasks, representing various conditions a user is likely to experience on a 3D Web site:

1. Search for textual information and then search for visual information (S:H-3D). Task 1 requires the users to first search for text information, and then to find the corresponding visual information in the 3D scene. An example task is: *Find the Horde race that uses Hawkstriders for mounts. What other races are to the left and right of this race?*
2. Search for visual information followed by textual information (S:3D-H). Task 2 is conceptually reversed, in that the users are required to find the visual information on the 3D scene first, and then to answer questions about the related text information. An example task is: *Find the sword which hilt/handle has a yellow dragon eye and ends with dragon claws. What is the price of this weapon?*
3. Compare text information and derive visual information (C:H-3D) (find visual attributes of items with a given text criteria). An example task is: *Find the strongest creature in the museum. What is the color of the creature's eyes?*
4. Compare visual information and derive textual information (C:3D-H) (search for textual attributes of items with a given visual criteria). An example task is: *There are two races with tails. What are their mounts?*

### 5.6. Measurements

Like in the works of Chen et al. (2004) and Polys et al. (2011), the study measured relative effectiveness of our user interfaces by both examining time taken to answer each question and correctness of answers. In addition, we developed a questionnaire to measure participants' subjective impressions of the user interfaces. The questionnaire contained continuous Likert scales regarding ease of use, learnability, efficiency, aesthetics,

presentation and access to text, 3D navigation, wayfinding, immersion, and overall preference. Subjects were also welcome to add any comments relevant to their experiences. The test application also recorded the usage of different UI components (e.g., a number of selections in a viewpoint menu, the use of hypersections, 3D hyperlinks, etc.)

### 5.7. Procedure

Each test session started with an introduction to the test application. Its interface was divided into two parts: the window, where the user was presented with tasks and the browser window (1280x800), where the user could interact with the prepared exhibitions through the user interfaces evaluated in this study. The introduction was followed by a training session (4 practice tasks for each interface) to allow the subject to get familiarized with the test application, the interfaces, and the test procedure. The users were educated and guided on how to use the walk and go-to navigation metaphors, and the viewpoint menu for control in a virtual world; they were also introduced to the concepts of hypersection and 3D hyperlink. After the subjects indicated that they were satisfied, we proceeded with the actual trials.

The tasks in the main part of the evaluation were similar to the ones from the training session: for each exhibition-UI combination we asked 4 questions related to the content of the exhibitions (presentation of variables was counterbalanced by means of Latin square design). For each question there was a choice of 4 answers from which the user had to choose 1 and only 1 answer. The subjects were asked to complete the assigned tasks "as accurately and as fast as possible". They were also told that it was more important to solve the tasks correctly rather than to be quick. They were allowed to take a break between each set of questions. The participants were video recorded during the tasks and notes were taken about their actions and comments. After being presented with all 36 tasks (3 UI modes \* 3 exhibitions \* 4 tasks), the users were given the questionnaire and asked to directly compare the evaluated user interfaces.

Each evaluation session lasted approximately 120 minutes - here it is important to stress the fact that for most of the participants the experiment was not tiring and seemed much shorter (actually, some participants expected more questions).

### 5.8. Results

In this section we present the results of the experiment. We collected a total of 720 time and accuracy measurements (20 subjects \* 3 UI modes \* 3 exhibitions \* 4 tasks), and 660 measurements of subjective impressions (20 subjects \* 11 questionnaire parts \* 3 UI Modes). We analyzed our results with analysis of variance (ANOVA); the collected data was checked to ascertain whether it met the assumptions of this test. With ANOVA we modeled our experiment as a repeated-measures 3x3x4 design (UI x Environment x Task). Bonferroni correction procedure was used for evaluating the pairwise comparisons.

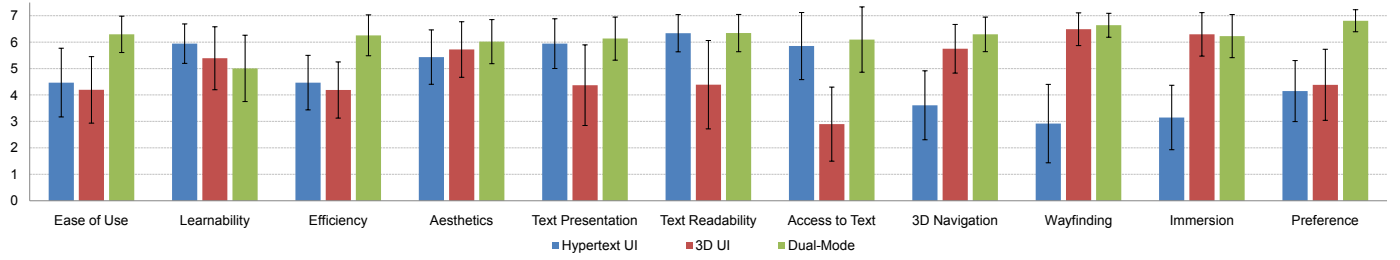


Figure 12: Subjective results from the questionnaire.

### 5.8.1. Objective Results

Times for completion of each task were normalized on the overall average completion time. Normalization was used to remove any effects of base reading speed and 3D navigation experience among participants. As overall accuracy was very high (0.985%), we decided to simply double the times of wrong answers.

Analysis of the task completion time revealed significant main effects of all variables and their interactions ( $p < 0.003$ ). Most importantly, it found significant main effects of UI ( $F(2, 38) = 44.32$ ,  $p < 0.00001$ ), interaction between UI and environment type ( $F(4, 76) = 4.49$ ,  $p = .0026$ ), and interaction between UI and task type ( $F(6, 114) = 25.66$ ,  $p < 0.00001$ ).

Post-hoc comparisons of means revealed that the dual-mode UI condition resulted in the best overall task performance ( $p < 0.0001$ ), while the hypertext UI condition was marginally worse than the 3D UI ( $p < 0.041$ ). To be more precise, executing tasks using the dual-mode UI was about 43% faster than using the hypertext UI (99s vs. 141s), and about 31% better than using the 3D UI (99s vs. 129s), while executing tasks using the 3D UI was about 9% faster than using the hypertext UI (129s vs. 141s).

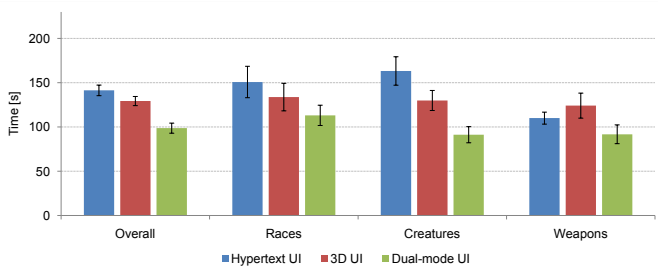


Figure 13: Overall results.

Comparisons of means for each exhibition revealed that the dual-mode UI was significantly better than the hypertext UI ( $p < 0.015$ ) for Races and Creatures. It was also better than the 3D UI ( $p < 0.011$ ) for Creatures. Figure 13 illustrates the overall results of our experiment and the results for each exhibition with respect to task completion time (error bars denote 0.95 confidence intervals). An interesting finding, visible in Figure 13, is that the hypertext UI was worse than the 3D UI in large environments, and it was better in a small one (Weapons). We believe this is because the small environment did not require much 3D navigation and users could not get lost in 3D space.

As we have already mentioned, we also found a signifi-

cant main effect of interaction between UI and task type ( $F(6, 114) = 25.663$ ,  $p < 0.0001$ ) on task completion time. Not surprisingly, since the task types differed significantly (see Section 5.5). Post-hoc comparisons of means revealed that executing tasks using the dual-mode and hypertext UIs was significantly faster than using the 3D UI for S:H-3D and C:H-3D tasks ( $p < 0.0001$ ); using the dual-mode and 3D UIs was significantly faster than using the hypertext UI for S:3D-H and C:3D-H tasks ( $p < 0.00006$ ) - see Figure 14.

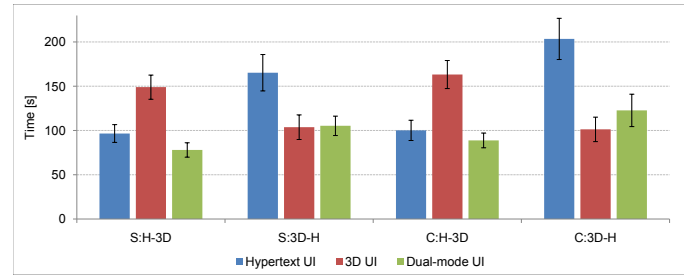


Figure 14: Interaction between UI and task type.

### 5.8.2. Subjective Results

The average subject ratings with respect to ease of use (difficult/easy), learnability (difficult/easy), efficiency (completing the tasks fast and accurately was: difficult/easy), aesthetics (non-aesthetic/aesthetic), presentation of text (confusing/clear), readability of text (hard to read/easy to read), search and access to text (difficult/easy), 3D navigation (complicated/simple), wayfinding (complicated/simple), immersion (not involved/really immersed), and overall preference are illustrated in Figure 12, together with standard deviations.

Analysis of the ratings revealed significant main effects of UIs on all scores. The dual-mode UI was perceived easier to use and more efficient than the hypertext and 3D UIs ( $p < 0.0001$ ). It was also perceived as more aesthetic than the hypertext UI ( $p < 0.003$ ). On the other hand, the hypertext UI was perceived as easier to learn than the 3D UI ( $p < 0.045$ ) and the dual-mode UI ( $p < 0.0003$ ). Subjects perceived presentation, readability and access to text in the 3D UI as worse than in the dual-mode and hypertext UIs ( $p < 0.0001$ ). In contrast, 3D navigation, wayfinding, and immersion in the hypertext UI were ranked lower than in the dual-mode and 3D UIs ( $p < 0.0001$ ). Finally, the dual-mode UI was evidently preferred over the alternatives ( $p < 0.0001$ ). These findings clearly support our analysis of task performance.

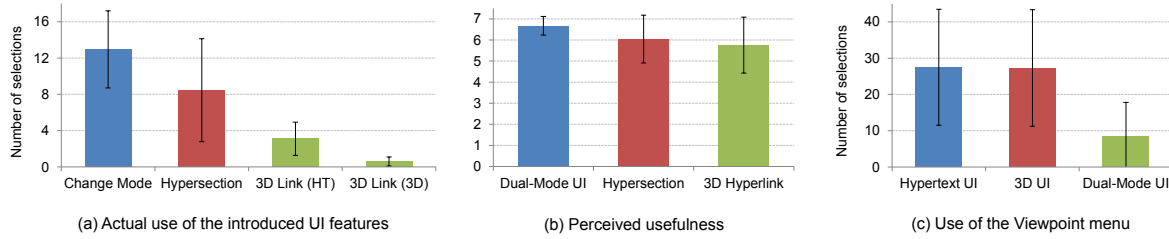


Figure 15: (a) Actual use and (b) perceived usefulness of the introduced UI features; (c) Use of the Viewpoint menu

As we have already mentioned, our test application also recorded the usage of hypersections and 3D hyperlinks as well as how many times users switched between the UI modes when using the Dual-Mode user interface. Figure 15 (a) illustrates the results: the option to change the mode in DMUI was used the most, followed by the use of hypersections and 3D hyperlinks in the hypertext and the 3D mode respectively. These results are in line with the the average subject ratings with respect to usefulness of these new UI features (the results are illustrated in Figure 15 (b), together with standard deviations).

### 5.8.3. Discussion

The results from this comparative user study suggest users performed better with the dual-mode user interface over alternatives, i.e. the hypertext and 3D UIs, on tasks, which we believe are representative of a variety of 3D Web applications. The subjective comments also showed a preference for the dual-mode interface. Moreover, the evaluation results can help to understand better the relationship between the hypertext and 3D UIs. Like most controlled user-based studies, this one had some limitations that restrict the generality of our findings: although we tested three different virtual environments, we still managed to test only a small sample of possible uses of Text+3D content.

A viewpoint menu turned out to be very important navigation tool, especially in the hypertext and 3D UIs (the use of the menu is illustrated in Figure 15 (c)). Interestingly, some users preferred alphabetic order of viewpoints, some preferred order based on the distance between the viewpoints; one subject noted that he would like to have categories in the menu. On the other hand, it is not clear whether it would be useful for environments with a large number of points of interest. A map also proved to be very important wayfinding aid to the majority of users. We noticed that few users clicked on the map - they expected instant teleportation to the selected rooms.

3 of the 4 participants with self-rated prior considerable experience in 3D navigation stated that they at times would have liked keyboard-based control of the camera, in addition to the mouse-only control we provided. Some participants asked about search support (Ctrl-F), both for text and 3D. One user did not like the grey highlighting of 3D objects in the 3D scene: *"such highlighting makes colors and details less visible"*. With regard to the dual-mode UI, one suggestion was to move the button for switching modes and position it in the corner of the 3D viewer (in the hypertext mode).

A few users criticized the label and annotation layout used in the experiment as we implemented a very simple layout scheme that places the label/annotation box on the left corner of a box that encloses the geometry's bounds (bounding box). As an effective label and annotation layout must fulfill a number of requirements for readability (labels must not overlap), unambiguity (labels clearly refer to their objects), pleasing (prevent visual clutter), real-time (compute layouts at interactive rates), spatial-coherency (prevent visual discontinuities), temporal-contiguity, compaction (reduce the layout area) (Fekete and Plaisant, 1999; Ali et al., 2005; Mayer, 2005; Polys, 2006), we plan to implement a more sophisticated view management component (see e.g., Bell et al. (2001); Gotzelmann et al. (2006)).

## 6. Conclusions

During the First Hypertext Conference in 1987, Andries van Dam gave a keynote speech and listed user-controlled 3D graphics as one of the key issues we should be looking at while researching and developing hypertext systems:

*"If a picture is worth a thousand words, a dynamic picture of time-varying objects is worth a thousand static ones. We need dynamics at the nodes, not just static pictures and text."*

Andries van Dam van Dam (1988)

24 years have passed and Andy's vision is still just a vision. The Web, today's largest and most important hypertext-based online information infrastructure, does not support 3D content and, although various approaches have been proposed (most notably VRML/X3D and now WebGL), there is still no clear design methodology for user interfaces that integrate hypertext and interactive 3D graphics.

### 6.1. Summary

This paper began with the introduction of hypertext and the Web as well as 3D computer graphics and virtual environments. We have presented arguments for combining these mediums and outlined our research and design approach to this topic. Further, we have clarified the foundations of this paper - we gave a summary of the research on UI design that underlies our work. We then have presented a novel strategy for accessing 3D content on the Web, where the objective was to pair interactive 3D graphics know-how with well established UI conventions of

the Web. We have introduced a Dual-Mode user interface, an interface that has two modes between which a user can switch anytime: the driven by simple hypertext-based interactions "*hypertext mode*", where a 3D scene is embedded in hypertext and the more immersive "*3D mode*", which immerses the hyper-textual annotations into the 3D scene. Furthermore, we have described Copernicus, a wiki-type authoring environment allowing the rapid development of HiVEs, a testbed allowing for the experimentation and for the assessment of both hypertext and 3D modes of the Dual-Mode user interface. Our Dual-Mode user interface, implemented in Copernicus, was evaluated iteratively using several methods for usability studies. The evaluations illustrated that people like to use Copernicus and appreciate the interactive possibilities offered by its user interface. The results from the latest comparative user study suggest users performed better with the Dual-Mode user interface over alternatives.

## 6.2. Future Work

The work on user interface design for applications that tightly integrate hypertext and interactive 3D graphics opens a new field, and as such leads to many new questions; hence, there is a huge amount of work to be done. We believe that we have merely scratched the surface in this area of research, and that many interesting and exciting directions still lie ahead. Therefore, we plan to continue our efforts in exploring how the experiences accumulated in two different communities, namely the Web Development and Computer Graphics communities, can be capitalized upon to achieve the long term goal of using 3D on the Web the same way as we do with 2D graphics and video today.

In cooperation with researchers primarily from Fraunhofer IGD (X3DOM), DFKI (XML3D), and Web3D Consortium (X3D), we will continue our current efforts in the Declarative 3D for the Web Architecture W3C Community Group on an integration of interactive 3D graphics capabilities into the Web technology stack in a declarative way. We see our main role as helping to determine the requirements and use cases for the declarative integration of interactive 3D to the HTML-DOM, which will provide a foundation for future standardization (Behr et al., 2012; Jankowski, 2012b). Moreover, as the Dual-Mode user interface design is not yet grounded in any formal standard (we have focused on the HCI perspective of the problem), we want to explore the possibilities of applying our design methodology in the HTML5 specification, since it does not yet define how the integration of 3D imagery into HTML actually should look like.

Another future avenue that we intend to explore is collaborative web browsing. Currently, browsing the Web is mostly an individual experience. People visiting the same web page at the same time are not aware of each other. The dual-mode user interface supports this personal nature of browsing. On the other hand, research on Collaborative Virtual Environments Benford et al. (2001) has shown that their entertainment and social nature can provide virtual communities with richer content and greater interactivity and greatly support collaborative work. We envision the Web (or rather part of it) evolving into a virtual

space in which people, while being in the 3D mode of the dual-mode UI, can interact and communicate with each other through their avatars (graphical representations of themselves).

## Acknowledgements

This work has been supported in part by Science Foundation Ireland under Grant No. SFI/08/CE/I1380 (Lion2) and Enterprise Ireland under Grant No. PC/2008/353 (Copernicus). Please visit <http://copernicus.deri.ie> to learn more about the Copernicus Project.

## References

- Ali, K., Hartmann, K., Strothotte, T., 2005. Label layout for interactive 3d illustrations. In: WSCG'05.
- Andrews, K., 1992. Using 3d scenes as hypermedia nodes (poster). In: ECHT'92.
- Andrews, K., 1994. Soaring through hyperspace: A snapshot of hyper-g and its harmony client. In: Eurographics Symposium on Multimedia/Hypermedia.
- Andrews, K., 1996. Browsing, building, and beholding cyberspace: New approaches to the navigation, construction, and visualisation of hypermedia on the internet. Ph.D. thesis, Graz University of Technology, Austria.
- Andrews, K., Pichler, M., 1994. Hooking up 3-space: Threedimensional models as fully-fledged hypermedia documents. In: MHVR'94.
- Azuma, R., Furmanski, C., 2003. Evaluating label placement for augmented reality view management. In: ISMAR'03.
- Balaguer, J.-F., Gobbetti, E., 1995. i3d: a high-speed 3d web browser. In: VRML '95.
- Behr, J., Brutzman, D. P., Herman, I., Jankowski, J., Sons, K. (Eds.), 2012. Proceedings of the 1st International Workshop on Declarative 3D for the Web Architecture (Dec3D2012 at WWW2012), Lyon, France, April 17, 2012. Vol. 869 of CEUR Workshop Proceedings. CEUR-WS.org.
- Behr, J., Eschler, P., Jung, Y., Zöllner, M., 2009. X3dom: a dom-based html5/x3d integration model. In: Web3D'09. ACM.
- Bell, B., Feiner, S., Hollerer, T., 2001. View management for virtual and augmented reality. In: UIST'01. ACM.
- Benford, S., Greenhalgh, C., Rodden, T., Pycok, J., 2001. Collaborative virtual environments. CACM.
- Berners-Lee, T., Fischetti, M., 1999. Weaving the Web: The Past, Present and Future of the World Wide Web by its inventor. Orion Business.
- Bolter, J. D., Hodges, L. F., Meyer, T., Nichols, A., 1995. Integrating perceptual and symbolic information in vr. IEEE Comput. Graph. Appl.
- Bowman, D. A., Hodges, L. F., Allison, D., Wineman, J., 1999. The educational value of an information-rich virtual environment. Presence.
- Bowman, D. A., Hodges, L. F., Bolter, J., 1998. The virtual venue: User-computer interaction in information-rich virtual environments. Presence.
- Bowman, D. A., Kruijff, E., LaViola, J. J., Poupyrev, I., 2001. An introduction to 3-d user interface design. Presence: Teleoper. Virtual Environ.
- Bowman, D. A., McMahan, R. P., 2007. Virtual reality: How much immersion is enough? Computer.
- Bowman, D. A., North, C., Chen, J., Polys, N. F., Pyla, P. S., Yilmaz, U., 2003. Information-rich virtual environments: theory, tools, and research agenda. In: VRST'03. ACM.
- Burtnyk, N., Khan, A., Fitzmaurice, G., Balakrishnan, R., Kurtenbach, G., 2002. Stylecam: Interactive stylized 3d navigation using integrated spatial and temporal controls. In: UIST'02. ACM.
- Bush, V., 1945. As we may think. Interactions (July 1996).
- Byrne, M. D., John, B. E., Wehrle, N. S., Crow, D. C., 1999. The tangled web we wove: a taskonomy of www use. In: CHI'99. ACM.
- Campbell, B., 1997. Vml as a superset of html - an approach to consolidation. In: Annual Computer Science Conference.
- Chandler, P., Sweller, J., 1991. Cognitive load theory and the format of instruction. Cognit. and Instruct.
- Chen, J., Pyla, P. S., Bowman, D. A., 2004. Testbed evaluation of navigation and text display techniques in an information-rich virtual environment. In: Virtual Reality'04. IEEE.

- Cmolik, L., Bittner, J., 2010. Layout-aware optimization for interactive labeling of 3d models. *Computers & Graphics*.
- Cruz-Neira, C., Sandin, D. J., DeFanti, T. A., 1993. Surround-screen projection-based virtual reality: the design and implementation of the cave. In: *SIGGRAPH'93*. ACM.
- Darken, R. P., Sibert, J. L., 1993. A toolset for navigation in virtual environments. In: *UIST'93*. ACM.
- Darken, R. P., Sibert, J. L., 1996. Wayfinding strategies and behaviors in large virtual worlds. In: *CHI'96*. ACM.
- Engelbart, D. C., English, W. K., 1968. A research center for augmenting human intellect. In: *Proceedings of the Fall Joint Computer Conference*. ACM.
- Faraday, P., Sutcliffe, A., 1997. Designing effective multimedia presentations. In: *CHI'97*. ACM.
- Faraday, P., Sutcliffe, A., 1999. Authoring animated web pages using "contact points". In: *CHI'99*. ACM.
- Feiner, S., 2002. Augmented reality: A new way of seeing. *Scientific American*.
- Fekete, J.-D., Plaisant, C., 1999. Excentric labeling: dynamic neighborhood labeling for data visualization. In: *CHI'99*. ACM.
- Fitzmaurice, G., Matejka, J., Mordatch, I., Khan, A., Kurtenbach, G., 2008. Safe 3d navigation. In: *I3D'08*. ACM.
- Gabbard, J. L., Hix, D., Swan, J. E., 1999. User-centered design and evaluation of virtual environments. *IEEE Comput. Graph. Appl.*
- Galyean, T. A., 1995. Guided navigation of virtual environments. In: *SI3D'95*. ACM.
- Gotzelmann, T., Ali, K., Hartmann, K., Strothotte, T., 2005. Form follows function: Aesthetic interactive labels. In: *Computational Aesthetics'05*.
- Gotzelmann, T., Hartmann, K., Strothotte, T., 2006. Agent-based annotation of interactive 3d visualizations. In: *Smart Graphics'06*.
- Gotzelmann, T., Vazquez, P.-P., Hartmann, K., Nurnberger, A., Strothotte, T., 2007. Correlating text and images: Concept and evaluation. In: *Smart Graphics'07*.
- Hachet, M., Decle, F., Knodel, S., Guitton, P., 2008. Navidget for easy 3d camera positioning from 2d inputs. In: *3DUT'08*. IEEE.
- Hand, C., 1997. A survey of 3d interaction techniques. *Comput. Graph. Forum*.
- Harrison, B. L., Kurtenbach, G., Vicente, K. J., 1995. An experimental evaluation of transparent user interface tools and information content. In: *UIST'95*. ACM.
- Harrison, B. L., Vicente, K. J., 1996. An experimental evaluation of transparent menu usage. In: *CHI'96*.
- Hartmann, K., Gotzelmann, T., Ali, K., Strothotte, T., 2005. Metrics for functional and aesthetic label layouts. In: *Smart Graphics'05*.
- Ishak, E. W., Feiner, S. K., 2004. Interacting with hidden content using content-aware free-space transparency. In: *UIST'04*. ACM.
- Jankowski, J., 2011. A taskonomy of 3d web use. In: *Web3D'11*. ACM.
- Jankowski, J., 2012a. Evaluation of static vs. animated views in 3d web user interfaces. In: *Web3D'12*. ACM.
- Jankowski, J., 2012b. Writing effective use cases for the declarative 3d for the web architecture. In: *Dec3D'12*. CEUR-WS.
- Jankowski, J., Decker, S., 2009. 2lip: Filling the gap between the current and the three-dimensional web. In: *Web3D'09*. ACM.
- Jankowski, J., Decker, S., 2012. A dual-mode user interface for accessing 3d content on the world wide web. In: *WWW'12*. ACM.
- Jankowski, J., Hachet, M., 2013. A survey of interaction techniques for interactive 3d environments. In: *Eurographics'13*.
- Jankowski, J., Kruk, S. R., 2008. 2lip: The step towards the web3d (poster). In: *WWW'08*. ACM.
- Jankowski, J., Samp, K., Irzynska, I., Jozwicz, M., Decker, S., 2010. Integrating text with video and 3d graphics: The effects of text drawing styles on text readability. In: *CHI'10*. ACM.
- Kaur, K., 1998. Designing virtual environments for usability. Ph.D. thesis, City University London.
- Khronos, 2010. WebGL. <http://www.khronos.org/webgl/>.
- Koved, L., Shneiderman, B., 1986. Embedded menus: selecting items in context. *Communications*.
- Krug, S., 2005. *Don't Make Me Think: A Common Sense Approach to the Web* (2nd Edition).
- Mackinlay, J. D., Card, S. K., Robertson, G. G., 1990. Rapid controlled movement through a virtual 3d workspace. *SIGGRAPH*.
- Mayer, R. E., 2005. *Cambridge handbook of multimedia learning*. Cambridge University Press.
- Mayer, R. E., 2009. *Multimedia Learning* (2nd ed.).
- Mayer, R. E., Moreno, R., 1998. A split-attention effect in multimedia learning: Evidence for dual processing systems in working memory. *J. Educ. Psychol.*
- Mohageg, M., Myers, R., Marrin, C., Kent, J., Mott, D., Isaacs, P., 1996. A user interface for accessing 3d content on the world wide web. In: *CHI'96*. ACM.
- Mousavi, S. Y., Low, R., Sweller, J., 1995. Reducing cognitive load by mixing auditory and visual presentation modes. *J. Educ. Psychol.*
- Nelson, T. H., 1965. Complex information processing: a file structure for the complex, the changing and the indeterminate. In: *ACM '65: Proceedings of the 1965 20th national conference*. ACM.
- Ni, T., Bowman, D. A., Chen, J., 2006. Increased display size and resolution improve task performance in information-rich virtual environments. In: *GI'06*.
- Nielsen, J., 2000. *Designing Web Usability: The Practice of Simplicity*. New Riders Publishing.
- Nielsen, J., Loranger, H., 2006. *Prioritizing Web Usability*. New Riders Publishing.
- Paivio, A., 1986. *Mental representations. A dual coding approach*. Oxford University Press.
- Pichler, M., 1993. Interactive browsing of 3d scenes in hypermedia: The hyper-g 3d viewer. Master's thesis, Graz University of Technology, Austria.
- Polys, N. F., 2006. Display techniques in information-rich virtual environments. Ph.D. thesis, Blacksburg, VA, USA.
- Polys, N. F., Bowman, D. A., 2004. Design and display of enhancing information in desktop information-rich virtual environments: challenges and techniques. *Virtual Reality*.
- Polys, N. F., Bowman, D. A., North, C., 2011. The role of depth and gestalt cues in information-rich virtual environments. *IJHCS*.
- Polys, N. F., Kim, S., Bowman, D. A., 2005. Effects of information layout, screen size, and field of view on user performance in information-rich virtual environments. In: *VRST'05*. ACM.
- Preim, B., Raab, A., Strothotte, T., 1997. Coherent zooming of illustrations with 3d-graphics and text. In: *Graphics Interface'97*. Canadian Information Processing Society.
- Robertson, G., Czerwinski, M., van Dantzich, M., 1997. Immersion in desktop virtual reality. In: *UIST'97*. ACM.
- Rouse, R., 2000. *Game Design - Theory and Practice*.
- Shneiderman, B., 2003. Why not make interfaces better than 3d reality? *IEEE Comput. Graph. Appl.*
- Shneiderman, B., Plaisant, C., 2004. *Designing the user interface: strategies for effective Human-Computer Interaction*. Addison Wesley.
- Slater, M., 1999. Measuring presence: A response to the witmer and singer presence questionnaire. *Presence: Teleoper. Virtual Environ.*
- Slater, M., 2003. A note on presence terminology. In: *Presence-Connect*.
- Sonnet, H., Carpendale, M. S. T., Strothotte, T., 2005. Integration of 3d data and text: The effects of text positioning, connectivity, and visual hints on comprehension. In: *INTERACT'05*.
- Sons, K., Klein, F., Rubinstein, D., Byelozorov, S., Slusallek, P., 2010. Xml3d: interactive 3d graphics for the web. In: *Web3D '10*. ACM.
- Stein, T., Decoret, X., 2008. Dynamic label placement for improved interactive exploration. In: *NPAR'08*. ACM.
- Sutherland, I., 1963. *Sketchpad: A Man-Machine Graphical Communications System*. MIT PhD Thesis.
- Sutherland, I., 1965. The ultimate display. In: *IFIPS Congress*.
- Sutherland, I. E., 1968. A head-mounted three dimensional display. In: *Fall Joint Computer Conference*. ACM.
- Tan, D. S., Robertson, G. G., Czerwinski, M., 2001. Exploring 3d navigation: combining speed-coupled flying with orbiting. In: *CHI'01*. ACM.
- van Dam, A., 1988. Hypertext '87: keynote. *CACM*.
- Vukicevic, V., 2007. Canvas 3d: Gl power, webstyle (retrived 11.2010). <http://blog.vlad1.com/2007/11/26/canvas-3d-gl-power-web-style/>.
- Ware, C., Arthur, K., Booth, K. S., 1993. Fish tank virtual reality. In: *INTERACT'93 and CHI'93*. ACM.
- Witmer, B. G., Singer, M. J., 1998. Measuring presence in virtual environments: A presence questionnaire. *Presence: Teleoper. Virtual Environ.*
- Witrock, M. C., 1989. Generative processes of comprehension. *Educational Psychologist*.
- Yankelovich, N., Haan, B. J., Meyrowitz, N. K., Drucker, S. M., 1988. *Intermedia: The concept and the construction of a seamless information environment*. Computer.