

A Taskonomy of 3D Web Use

Jacek Jankowski*

Digital Enterprise Research Institute, NUI Galway, Ireland

Abstract

This paper aims to shed some light on the fundamental tasks users engage in while browsing the 3D Web. Firstly, we will provide an overview and discussion of several field studies that provide insights on how the current version of the Web is used focusing on the tasks that users carry out. Then we will present a comprehensive review of the state of the art of mouse-based 3D interaction techniques for Desktop 3D Virtual Environments (DVEs). Finally, we will introduce a simple taxonomy - a "3D Web Taskonomy" that was designed to assist designers and developers of interactive 3D Web applications to better evaluate their options when choosing and implementing an interaction technique.

CR Categories: H.5.2 [Information interfaces and presentation]: User Interfaces—Interaction Styles; H.5.4 [Information Interfaces and Presentation]: Hypertext and Hypermedia—User Issues; I.3.6 [Computer Graphics]: Methodology and Techniques—Interaction Techniques; I.3.7 [Computer Graphics]: Three-Dimensional Graphics and Realism—Virtual Reality;

Keywords: 3D Web, user tasks, interaction techniques, taxonomy

1 Introduction

There have been a number of approaches over the last years to integrate 3D technologies on the Web and most of these systems and standards disappeared or barely survived [Behr et al. 2009]. We argue that this is because of the fact that research was focused mostly on 3D graphics and 3D graphics alone. The 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 and interactive 3D graphics are at odds on the Web.

Human-computer interaction (HCI) is the study of interaction between people (users) and computers; it includes developing new interaction techniques for supporting user tasks, providing better access to information, and creating better forms of communication. Interaction between users and computers occurs at the human-computer interface (user interface (UI) or simply interface). A prerequisite to the effective UI design is an understanding of the users and their tasks - the tasks for which that interface will actually be used. Consequently, designers can translate this knowledge into usable systems and like in regular engineering tasks, previous experiences and theory on HCI can help them in the process.

This paper aims to shed some light on the fundamental tasks users engage in while browsing the 3D Web. Firstly, we will provide an overview and discussion of several field studies that provide insights on how the current version of the Web is used focusing on the tasks that users carry out. Then we will present a comprehensive

review of the state of the art of mouse-based 3D interaction techniques (Viewpoint Control, Selection and Manipulation, and System Control) for Desktop 3D Virtual Environments (DVEs), the environments that exploit common, general-purpose desktop hardware for interaction, i.e., mouse and keyboard. Finally, we will introduce a simple taxonomy - a "3D Web Taskonomy" that was designed to assist designers and developers of interactive 3D Web applications to better evaluate their options when choosing and implementing an interaction technique, and show how it can be used for the development of new designs.

2 UI Design - Hypertext and the Web

The World Wide Web is today's largest and most important hypertext-based online information infrastructure. A great deal of research interest has been devoted to the Web; however, most of this research effort has focused on the technical and application aspects. Our concern, however, is with how users interact with the Web. In order to design simple yet interactive 3D hypermedia interface for the Web users, it is necessary to first understand the tasks users engage in while browsing the Web. That is, in order to determine which HCI techniques and/or approaches are most likely to aid in 3D Web user interface design, it is first necessary to understand what users actually do with their time while using the Web.

2.1 Empirical studies

Empirical studies are valuable sources of knowledge on how users interact with computer systems, thus, in the next two sections we provide an overview and discussion of several field studies that provide insights on how the Web is used focusing on the tasks that users carry out. We describe field studies from two categories: observational short-term studies, which provide data on the users and their tasks as collected through screen captures, video recording, diaries or questionnaires, and click-through long-term studies, which provide quantitative data on the actual user navigation behavior. For a more detailed description of the topic at hand see [Herder 2006].

2.1.1 Observational studies

One of the most influential observational studies, with focus to gain a clearer understanding of the tasks users engage in while browsing the Web, was carried out by Byrne et al. [1999]. Participants were asked to browse the Web and provide verbal protocols describing what they were doing; a video camera was set up to record both the protocol and the user's screen as they browsed. Based on collected data, the authors presented a taxonomy of tasks ("taskonomy") undertaken on the WWW that consists of six general classes: (1) *Use Information*: using information from the Web, e.g., reading/viewing/downloading; (2) *Locate on Page*: locating something "interesting", a specific string or image; (3) *Go to Page*: any activity which causes the browser to display an URL (e.g., using a hyperlink, back button, bookmark); (4) *Provide Information*: providing authentication information or search string; (5) *Configure Browser*: changing the state of the browser, e.g., adding a bookmark or changing the cache size; (6) *React to Environment*: times when the browser demands something of the user, e.g., responding to a dialog. In terms of total time Use Information dominated users' browsing. The next most time-consuming activity was Locate on Page followed by Go to Page (due to the loading time).

*e-mail: jacek.jankowski@deri.org

Sellen et al. [2002] categorized various kinds of activities carried out on the Web. Their study was based on a combined diary and interview methodology. By examining a two-day set of Web activities from 24 people, the authors sorted the activities into six categories, each having common characteristics and goals: (1) *Finding* (24%): using the Web to find some specific, well-defined information; (2) *Information Gathering* (35%): researching a specific topic by gathering information. (3) *Browsing* (27%): visiting sites with no specific goal in mind, but rather to be informed or be entertained; (4) *Transacting* (5%): using the Web to execute transactions (e.g., on-line banking, ordering a product); (5) *Communicating* (4%): participating in chatrooms or discussion groups; (6) *Housekeeping* (5%): checking or maintaining the accuracy and functionality of Web resources. Information Gathering activities were the most frequent and were followed by Finding and Browsing categories.

Observational studies provide us with qualitative insights on Web activities. However, it remains unclear to what extent the results - which may be biased by the test-environment setting and which rely heavily on the interpretation by the experimenter - can be observed in everyday Web usage. On the other hand, quantitative evidence can be provided by click-through long-term client-side studies.

2.1.2 Click-through studies

The second category of field studies, click-through long-term studies, focuses on gathering quantitative data on the actual user navigation behavior by logging user browsing activities. There are only a few studies of this kind - this may partly be explained by the fact that browsing is considered a private activity.

The first long-term click-through study was carried out in 1994, when the Web was relatively new, by Catledge and Pitkow [1995]. During a period of 21 days, the Web activity of 107 participants was collected using an instrumented version of XMosaic browser at university computers of Georgia Tech. Following hyperlinks (52%) and the browser's back button (40.6%) were by far the most frequent mechanisms for navigation, with regards to frequency of use. With regards to within-site navigation, users operated in a small area - the navigation structure resembled a hub and spoke distribution model due to the frequent use of backtracking. Tauscher and Greenberg [1997], who conducted their study on Web usage behavior in late 1995, confirmed Catledge and Pitkow's results that link following and back button usage were the most frequently used methods for Web navigation.

Similar study was carried out by Cockburn and McKenzie [2001] in 1999, who were motivated by the dramatic growth of the Web, the emergence of new browsers (Netscape and Internet Explorer), and the increased usage of search engines. The authors observed that browsing is a rapidly interactive activity; users often visit many pages within seconds successively. This implies that response time is an important issue (pages should be designed to load quickly); in addition, pages should be kept short and "scannable" (supporting evidence to Nielsen's web design guidelines [Nielsen 2000]).

From 1999 to 2004 the Web has dramatically changed. Web sites have evolved from static hypertext information resources to dynamic and interactive applications. To examine the consequences of these changes on user behavior, Weinreich et al. [2008] conducted another long-term click-through Web usage study in Winter 2004/2005. The authors' major findings were: link following continues to be the most common navigation action; the usage of multiple windows and tabs has partly replaced back button usage; Web browsing is a rapid activity even for pages with substantial content (again, this fact calls for designs that allow for cursory reading); click maps revealed that most activity happens on the top-left quarter of the page. Table 1 presents results of the study and compares them with previous long-term browser usage studies.

Table 1: Comparing web usage studies [Weinreich et al. 2008]

	Catledge et al.	Tauscher et al.	Weinreich et al.
Time of study	1994	1995 - 1996	2004 - 2005
Link following	45.7%	43.40%	43.50%
Direct access	12.60%	13.20%	9.40%
New window/tab	0.20%	0.80%	10.50%
Submit	-	4.40%	15.30%
Back	35.7%	31.70%	14.30%
Reload	4.30%	3.30%	1.70%
Forward	1.50%	0.80%	0.60%
Other	-	2.30%	4.80%

3 UI Design - 3D Graphics and DVEs

As we have already mentioned, and we cannot stress this enough, a prerequisite to the effective UI design is an understanding of the users and their tasks - the tasks for which that interface will actually be used. In the previous section we identified the hypertext tasks that a 3D Web user interface must support. In this section we will identify and describe the "3D tasks" that any 3D Web UI must implement.

The interaction in DVE systems can be characterized in terms of three universal interaction tasks [Hand 1997; Bowman et al. 2001]: (1) *Viewpoint Control* - refers to the task of moving user's viewpoint through an environment; (2) *Selection and Manipulation* - refers to techniques of choosing object and specifying its position and orientation; (3) *System Control* - refers to communication between user and system which is not part of the virtual environment.

3.1 Viewpoint Control

Virtual Environments usually represent more space than can be viewed from a single point. Users have to be able to get around within the environment in order to obtain different views of the scene. This process of getting around a virtual environment while keeping track of one's whereabouts and task objectives is the subject of a number of techniques often referred to by such names as wayfinding, locomotion, navigation, and camera control.

Developing an effective technique for mouse and keyboard based 3D Navigation is difficult. Viewpoint control involves six degrees of freedom (6DOF). The problem is the number of parameters to be controlled by the user - the major limitation inherent in using 2D devices is that at least one change of state is required to cover all translations and rotations. Another problem is the type of viewpoint control required by a given task. Mackinlay et al. [1990] distinguish four types of viewpoint movement for interactive 3D workspaces: General movement, Targeted movement, Specified coordinate movement, Specified trajectory movement. In addition to the difficulties of controlling the viewpoint, effective techniques for 3D Navigation should satisfy general interface requirements (e.g., is easy to use, etc.).

In recent years many types of viewpoint movement techniques for interactive 3D workspaces have been proposed (for detailed survey see [Christie et al. 2008]). In the following we will classify them into four categories proposed by Mackinlay et al. [1990].

3.1.1 General movement

In an exploratory movement (such as walking through a simulation of an architectural design), the user does not have any specific goal.

Its purpose is to rather gain knowledge of the environment. We will classify it into two groups:

Sweep-Pan-Zoom Sweeping, panning, and zooming are the primary camera movements used in almost every 3D modeling environments. They are natural ways to inspect objects, and work well with a pointing device such as a mouse since all of them are at most 2-dimensional operations. Sweep-Pan-Zoom technique requires the user to accomplish a movement by shifting back and forth among simple navigation modes. Such interaction model can be not optimal if the menu has to be used frequently.

Walking/Driving/Flying The other way of allowing users to navigate in 3D environment is simply to let them walk along the ground, drive a vehicle or fly a plane [Bowman et al. 1997; Ware and Osborne 1990]. This technique requires the user to continuously modify the camera position using e.g., arrow keys or controls superimposed on the screen and controlled by the mouse. As the multi-touch technology is becoming more available, some research has been made on how to use this new input device. Kim et al. [2008] introduce a Finger Walking in Place (FWIP) interaction technique that allows a user to travel in a virtual world as her/his bare fingers (representing human legs) slide on a multi-touch surface.

3.1.2 Targeted movement

A targeted movement is a movement with respect to a specific target. Its purpose is to gain knowledge of a specific part of the environment chosen by the user.

Point of Interest (POI) Flight Mackinlay et al. [1990] describe Point of Interest (POI) Logarithmic Flight. It requires the user to simply choose a target "point of interest" on an object in the scene by using the mouse, and the virtual camera moves logarithmically towards this target. Tan et al. [2001] describe Speed-coupled Flying with Orbiting - technique that allows users to not only transition between local environment-views and global overviews of the world, but to also orbit around an object of interest in order to examine it. Navidget [Hachet et al. 2008] is another targeted movement approach where the endpoint of a trajectory can be selected.

Path Drawing Igarashi et al. [1998] propose a Path Drawing technique for walkthrough in 3D environments. The technique uses user-drawn strokes to control the virtual camera. A stroke is projected onto the walking surface and used as a path for the camera. Hagedorn and Dollner [2008] present another navigation method based on sketching navigation commands. It interprets the sketches according to their geometry, spatial and temporal context.

Bookmarked Views Some 3D viewing applications (e.g., VRML/X3D browsers) provide a viewpoint menu offering a choice of viewpoints, usually denoted by a brief textual description that helps make clear the intended purpose of each view. An interesting way to assist users in navigation in 3D environments was developed by Elvins et al. [1997]. They introduced a new technique that captures a 3D representation of a virtual environment landmark into a 3D thumbnail, which they call a worldlet.

Navigation by query Van Ballegooij and Eliens [2001] propose navigation by query, an interesting navigation technique based on Information Retrieval. This technique augments interface of virtual environment by allowing users to navigate a virtual world by means of querying its content.

3.1.3 Specified coordinate movement

Specified coordinate movement is a movement to a precise position and orientation, such as to a specific viewing position relative to a car model - the user has to supply the exact position and orientation of his destination. This type of camera positioning is used in CAD

and 3D editing software. The users can simply enter a pair of x, y, z coordinates (for position and orientation) using a keyboard.

3.1.4 Specified trajectory movement

Specified trajectory movement is a movement along a position and orientation trajectory, such as a cinematographic camera movement. Compared to previous techniques, where the users are free to roam and explore, the techniques of this category empower the author to bring structure to the experience. Such viewpoint control limits the user's freedom while traveling 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'.

Guided Navigation Guided Navigation was introduced by Galyean [1995], who proposed a new method for navigating virtual environments called "The River Analogy". He envisioned the navigation paths as the rivers flowing through a landscape, and the user as a metaphorical boat floating down the river with some control while also being pushed and pulled by the predefined current of the water. Elmqvist et al. [2008] have presented a similar method that uses motion constraints to guide the user on a pre-defined tour through the environment while still allowing users to control their speed as well as to deviate locally from the tour.

Constrained Navigation Hanson and Wernert [1997] propose a navigation technique that, rather than controlling an unconstrained vehicle in 3D space, uses the 2D control device to move the user on a constrained subspace, the '*guide manifold*', a kind of virtual 2D sidewalk. Burtnyk et al. [2002] describe StyleCam, a system for authoring 3D viewing experiences that seamlessly integrates spatial camera control with the temporal control of animation playback. Fitzmaurice et al. [2008] have developed a concept called Safe 3D Navigation where they augment navigation tools with properties to reduce the occurrence of confusing situations and improve the learning experience.

Path/Motion Planning Path/Motion Planning methods can be useful as a way to automatically create visits of monuments or engaging presentations of architectural visualization. They can also be useful in the context of interactive systems to create a first constrained path to help the user navigate an environment. Drucker and Zeltzer [1994] describe an intelligent camera control in a 3D virtual environment; the plan is computed using a well known graph searching technique called A*. Salomon et al. [2003] and Nieuwenhuisen and Overmars [2003] present similar approaches for interactive navigation in complex 3D environments. Their techniques are based on an algorithm that precomputes a global roadmap of the environment by using a variant of a randomized motion planning algorithm along with a reachability-based analysis. Oskam et al. [2009] extend the idea and present an approach for computing visibility-aware paths. Gran et al. [2004] present another automatic system for the generation of walkthroughs in 3D environments that can be segmented using a cell-and-portal approach.

Assisted navigation Guided tours technique, which is based on Guided Navigation, was developed to support (especially novice) users in navigating the virtual world, finding points of interests, and learning how to interact with them by exploiting an animated character that leads the user on a guided tour of the virtual environment. Wernert and Hanson [1999] propose a design that incorporates a personal "guide" that serves two important functions: keeping the user oriented in the navigation space, and "pointing" to interesting objects/places as they are approached. For the 3D Web context, Chittaro et al. [2003] propose the adoption of a humanoid animated character to lead the user on a guided tour of the world.

Object Inspection During inspection of an object, the navigation above its surface can become complex. Khan et al. [2005] describe HoverCam, a technique for navigating around 3D objects at close proximity. It allows users to move around the object while maintaining a fixed distance from the surface and while keeping the object centred in the field of view. Burnyk et al. [2006] describe ShowMotion, a system for 3D design review of CAD models. Their system replaces traditional "bookmarked" static views with moving "shots" interspersed with cinematic visual transitions.

3.1.5 Wayfinding

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]: (1) users may wander without direction when attempting to find a place for the first time; (2) they may then have difficulty relocating recently visited places; (3) they are often unable to grasp the overall topological structure of the space. This problem is known as a problem of users getting "lost-in-cyberspace".

Efficient wayfinding is based on the navigator's ability to conceptualize the space. This type of knowledge, as defined by Thorndyke [1982], who studied the differences in spatial knowledge acquired from maps and exploration, is based on: (1) *survey knowledge* - knowledge about object locations, inter-object distances, and spatial relations; (2) *procedural knowledge* - the sequence of actions required to follow a particular route.

Based on the role of spatial knowledge in wayfinding tasks, designers have concerned themselves with developing design methodologies that aid navigation. Darken and Sibert [1996] present a toolkit of techniques based on principles of navigation derived from real-world analogs. Their evaluation shows that subjects' wayfinding strategies and behaviours were strongly influenced by the environmental cues. Vinson [1999] proposes design guidelines for landmarks to support navigation in Virtual Environments; he focuses on the design and placement of such navigation aids. Another technique for supporting interaction and navigation is augmenting virtual environments with interactive personal marks [Grammenos et al. 2006]. Darken and Sibert [1993] were first to present the concept: users can drop breadcrumbs, which are markers in the form of simple cubes floating just above the ground plane. Edwards and Hand [1997] describe similar technique, trailblazing, and the use of maps as examples of planning tools for navigation. While researching the relation between wayfinding and motion constraints, Elmqvist et al. [2008] found out that navigation guidance allows users to focus less on the mechanics of navigation, helping them in building a more accurate cognitive map of the 3D world.

3.2 Selection and Manipulation

Another typical task performed in a 3D virtual environment is object selection and its direct manipulation.

3.2.1 Object Selection

The most common way to select objects in desktop VEs is positioning a mouse cursor over given object and clicking a mouse button. The technique is based on ray casting; it uses the ray from the eye point through the pixel currently selected by the mouse pointer to find the first intersection point with the scene (the surface of the target or to its approximated surface, e.g., bounding box). If the virtual ray intersects an object, it can be selected for further manipulation.

Elmqvist and Fekete [2008] adapt the concept of semantic pointing to a 3D context for improving object picking performance. The

adaptation works by shrinking empty space and expanding target sizes in motor space. Looser et al. [2005] examined pan-based target acquisition in 3D first-person computer games. Another technique used in CAD and 3D editing software is indirect selection of objects from a list.

3.2.2 Object Manipulation

Manipulating objects is another basic task performed in a 3D virtual environment, especially during scene construction. When designing a scene with multiple objects, users have to repeatedly realign and adjust different parts. It is important to provide an efficient object manipulation technique.

The mouse, an excellent input device for 2D user interfaces, is commonly used to position objects in 3D environments. This causes the problem of how to map a fundamentally 2D input device into different operations with different degrees of freedom.

One solution to the problem is to provide users with a set of tools, dialog boxes, menu items, or other kinds of 2D UI elements for changing between manipulation modes, where mode is a mapping of an input to a possible output. Much simpler and more efficient solution is to provide users with manipulators for explicit and more direct manipulation [Nielson and Olsen 1987; Strauss and Carey 1992]. Strauss et al. [2002] define manipulators as visible graphic representations of an operation on, or state of, an object, that is displayed together with that object. This state or operation can be controlled by clicking and dragging on the graphic elements (handles) of the manipulator.

Manipulators, which currently are adopted by most CAD and 3D editing software applications, evolved initially from ideas associated with interactive manipulation using 2D control devices. Positioning manipulators originate from the work carried out by Bier [1987] on skitters and jacks tools for interactive 3D positioning. Bier also developed a very useful technique for snap-dragging in 3D [Bier 1990]. Rotation manipulators are built upon Chen et al.'s work on Virtual Sphere [1988] and Shoemaker's ArcBall [1992]. Both techniques are based on a concept of a virtual ball that contains the object to manipulate. Comparison of mouse-based interaction techniques for 3D rotation can be found in [Bade et al. 2005].

Some research has been done to exploit knowledge about real world (e.g., gravity) and natural behaviour of objects (e.g., solidity) to provide simple and intuitive interaction techniques for object manipulation. Bukowski and Sequin [1995] employ a combination of pseudo-physical behaviour and goal-oriented properties called 'Object Associations' to position objects in a 3D scene. In their approach objects selected for relocation actively look for nearby objects to associate and align themselves with. Goesele and Stuerzlinger [1999] introduce Semantic Constraints. They noticed that in addition to the physical properties people expect objects in a natural environment to behave according to their experience: "People know that a chair will be standing on a floor and a painting will be hung on a wall. Hanging the chair on the wall and placing the painting on the floor is against common sense". Smith et al. [2001] presented the evaluation of Bukowski and Sequin's 'Object Associations' [1995] and Goesele and Stuerzlinger's 'Semantic Constraints' [1999]. They showed that in most cases constraints provide more than a factor of two speed-up combined with a significant increase in accuracy. Reisman et al. [2009] describe a screen-space method for multi-touch 3D interaction. Just like 2D multi-touch interfaces allow users to directly manipulate 2D contexts with two or more points, their method allow the user to directly manipulate 3D objects with three or more points. Phillips et al. [1992] describe a technique for augmenting the process of 3D manipulation by automatically finding an effective placement for the virtual camera.

3.3 Application Control

Application Control describes communication between user and system which is not part of the virtual environment [Hand 1997]. It refers to a task in which a command is applied to change either the state of the system or the mode of interaction [Bowman et al. 2001]. Hand and Bowman et al. point out that although viewpoint movement and selection/manipulation have been studied extensively, very little research has been done on system control tasks. However, application control techniques have been studied intensively over the past 40 years in 2D "point-and-click" WIMP graphical user interfaces (interfaces based on windows, icons, menus, and a pointing device, typically a mouse).

3.3.1 Screen-Space WIMP Application Control

The HCI and UI communities agree that a good application control technique should be easy to learn for novice users and efficient to use for experts; it has to also provide the means for the novice users to gradually learn new ways of using the interface. WIMP interfaces have proven to support all these characteristics. People using computer applications have become intimately familiar with a particular set of WIMP UI components: windows, icons, menus, and a mouse. Theoretical models, experimental tools and real applications have resulted in a generally very good understanding of traditional WIMP interfaces for 2D applications. However, what if the information is represented in three-dimensional space?

Currently, computer and video game industry leads the development of 3D interactive graphics and it is where many cutting edge interface ideas arise. From what we can observe, the most of game interfaces are designed in a way, where application control interface components are placed in a screen-space - on a 2D plane called HUD (head-up display) that is displayed side-by-side with a 3D scene or overlays a 3D scene.

World of Warcraft is one of the most successful games of all time. From its release in 2004, its interface evolved only a little and we believe that it is a good example of how to design a game interface properly. Warcraft's interface provides players with a great amount of useful information, while allowing them to easily and intuitively make actions they want. This balance was achieved through an emphasis on the following: (1) The controls are clear - their physical appearance tells how they work; (2) At the beginning of the adventure in the Warcraft universe there are only few available UI components; as the players gain more experience and more abilities, they can gradually enhance their interfaces and add more functionality to their toolbars; (3) Help is embedded throughout the interface.

We found several works that are dedicated to game design, where the authors include some discussion around creating 3D game's interfaces [Rouse 2000; Salen and Zimmerman 2003]. Richard Rouse, in his book *Game Design: Theory & Practice* [2000] pointed out that the best interface is the one that is difficult to notice. Greg Wilson, in his "Off With Their HUDs!" article in *Gamasutra* [2006], describe the shift of the game user interfaces away from a reliance on traditional heads-up displays (e.g., in *Far Cry 2* HUD appears only in certain situations, and then fades out when it is not needed).

Screen-space application control splits the user interface into two parts with very different UI metaphors. While the navigation and manipulation functionality is accessible through a 3D user interface, the rest of system functionality can only be controlled through a conventional 2D GUI. This can lead to a problem with supporting the feeling of immersion and directness. Hand points out that this is a general problem with application control, since it may require the user to change from talking directly to the interface objects, to talking about them, thereby stepping outside the frame of reference used when manipulating objects or the viewpoint. If this shift is too great then the engagement of the user may be broken.

3.3.2 Object-Space Post-WIMP System Control

WIMP GUIs, with their 2D interface components, were designed for 2D applications (such as word processing and spreadsheets) controlled by a 2D pointing device (such as a mouse). However, when the information is represented in a 3D space, the mapping between 3D tasks and 2D control widgets can be much less natural and introduce significant indirection and "cognitive distance" [van Dam 1997]. Andy van Dam argues that this new forms of computing necessitate new thinking about user interfaces. He calls for the development of new "post-WIMP" UIs that rely on, e.g., gesture and speech recognition, eye/head/body tracking, etc.

In this paper we are concerned with a design of user interfaces and interaction techniques that exploit common mouse and keyboard as input devices. Object-space post-WIMP application control involves integration of UI components into a 3D scene. Typically, user interface designers transform system control techniques used in 2D interfaces, such as buttons and menus, and implement them in 3D environments. However, such adaptation is not straightforward - 2D objects that are placed in 3D space may undergo disproportion, disparity, occlusion. Moreover, selection (manipulation of a cursor) in 3D can be difficult to perform; users, for example, can have problems with selecting an item in a 3D menu floating in space.

3D Information Workspaces In the seminal work on information visualization using 3D interactive animation, Robertson, Card, and Mackinlay [1993] use Interactive Objects as a basic input/output mechanisms for the Information Visualizer's user interface. They are used to form the basis for coupling user interaction with application behaviour and build 3D widgets that represent information or information structure. The Web Forager [Card et al. 1996] is an information workspace for the Web built upon the Information Visualizer. Pierce et al. [1999] present toolspaces and glances to provide a mechanism for storing, accessing, and retrieving 3D widgets and other objects in interactive 3D worlds. Robertson et al., during their further work on information workspaces, developed Task Gallery, a window manager that uses interactive 3D graphics to provide direct support for task management and document comparison [2000].

Three-dimensional Widgets Dachsel and Hinz [2005] propose a classification of existing 3D widgets for system control. The classification is based on interaction purpose/intention of use and it was created for DVEs. Dachsel, in his further work on the CONTIGRA project, focused on menu selection subgroup. Together with Hubner [2007], they provide a comprehensive survey of graphical 3D menu solutions for all areas of the mixed reality continuum.

In Section 2 we have described how users interact with the Web. Then we have surveyed a multitude of 3D interaction techniques from the area of DVEs. As we aim to assist researchers and developers in creating new interaction techniques and improving existing ones for the context of the 3D Web, we constructed a taxonomy of user tasks of the most common activities while exploring 3D Web.

4 Taxonomy of User Tasks

As we have already mentioned, to design a good user interface we need to possess a good understanding of the UI's users and their tasks. 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.

The interaction tasks that a web-based 3D user interface needs to support can be divided into hypermedia tasks, such as following hyperlinks and placing bookmarks, and virtual environment tasks, such as navigation and object manipulation. We constructed a tax-

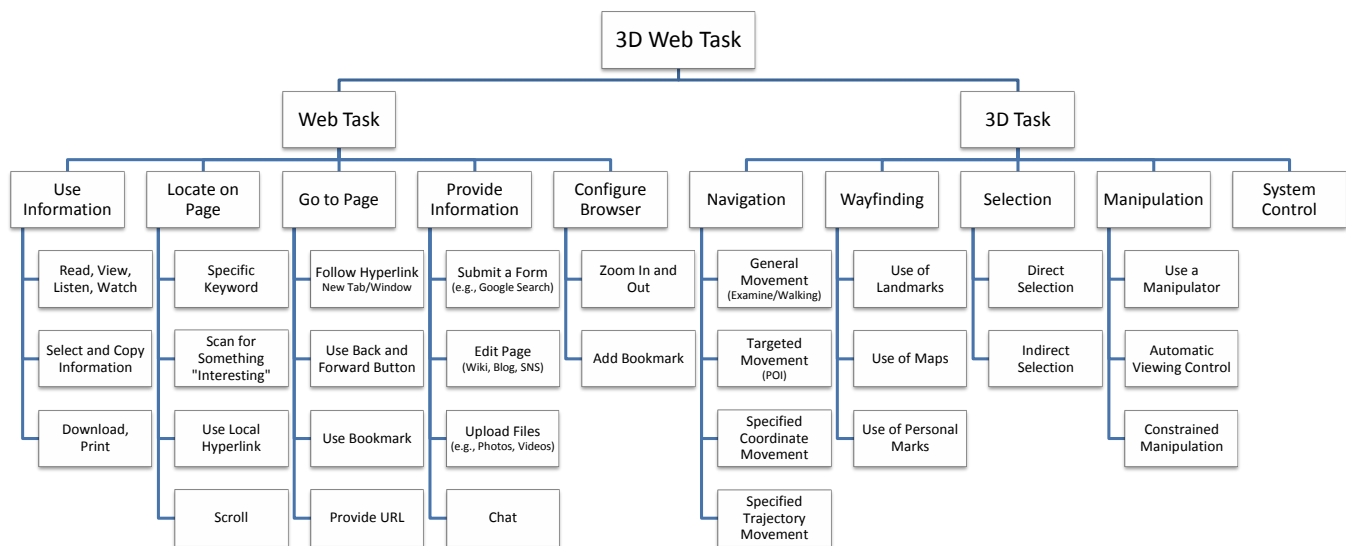


Figure 1: Taxonomy of 3D Web user tasks.

onomy of 3D Web user tasks (see Figure 1) that illustrates these two categories (both Web and "3D") of tasks, the taxonomy that integrates the most common activities while browsing the Web and exploring a 3D virtual environment. 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].

4.1 Web User Tasks

The taxonomy consists of two general classes: Web tasks and 3D 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.

4.2 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. *Using a Manipulator* - 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.

5 Conclusions

We believe that people can gain a lot from using 3D Web applications, integrated information spaces where textual (hypertext) and spatial (3D graphics) data are simultaneously available and linked. However, viewing and interacting with this kind of spaces can be complicated: users need to simultaneously deal with hypertext and three-dimensional graphics, two very different mediums that should both be appreciated equally. Users need to be able to browse the text - look through general information and search for more specific information. They also need to navigate freely through a three-dimensional space, and examine and manipulate virtual 3D objects, to gain a better understanding of the data. An interface for the exploration of this kind of integrated information spaces must support all these user tasks.

This paper aimed to clarify some of the foundations of 3D Web user interface design. We focused on an understanding of the fundamental tasks users engage in while exploring Web-based 3D virtual environments. We provided an overview and discussion of several field studies that provide insights on how the current version of the Web is used. We then presented a comprehensive review of the state of the art of mouse-based 3D interaction techniques (Viewpoint Control, Selection and Manipulation, and System Control) for Desktop 3D Virtual Environments. Finally, we introduced a simple taxonomy - a "3D Web Taskonomy" that was designed to assist designers and developers of interactive 3D Web applications to better evaluate their options when choosing and implementing an interaction technique. We intended to develop the classification at a level that allows researchers of HCI, the Web and Computer Graphics domains to easily understand the topic at hand.

In [Jankowski 2011] we develop a theory for combining hypertext with interactive 3D graphics - we introduce Hypertextualized Virtual Environments (HiVEs), integrated information spaces where

symbolic (hypertext) and visual (3D graphics) data are simultaneously available and linked. The taxonomy developed in this paper is used to identify the tasks that the HIVE user interface must support.

Acknowledgements This work has been funded in part by Science Foundation Ireland under Grant No. SFI/08/CE/I1380 (Lion2) and Enterprise Ireland under Grant No. PC/2008/353 (Copernicus).

References

- BADE, R., RITTER, F., AND PREIM, B. 2005. Usability comparison of mouse-based interaction techniques for predictable 3d rotation. In *Smart Graphics'05*.
- BEHR, J., ESCHLER, P., JUNG, Y., AND ZÖLLNER, M. 2009. X3dom: a dom-based html5/x3d integration model. In *Web3D'09*, ACM.
- BIER, E. A. 1987. Skitters and jacks: interactive 3d positioning tools. In *SI3D'87*, ACM.
- BIER, E. A. 1990. Snap-dragging in three dimensions. In *SI3D'90*, ACM.
- BOWMAN, D. A., KOLLER, D., AND HODGES, L. F. 1997. Travel in immersive virtual environments: An evaluation of viewpoint motion control techniques. In *VRAIS'97*.
- BOWMAN, D. A., KRUIJFF, E., LAVIOLA, J. J., AND POUPLYREV, I. 2001. An introduction to 3-d user interface design. *Presence: Teleoper. Virtual Environ.*
- BUKOWSKI, R. W., AND SÉQUIN, C. H. 1995. Object associations: a simple and practical approach to virtual 3d manipulation. In *SI3D'95*, ACM.
- BURTNYK, N., KHAN, A., FITZMAURICE, G., BALAKRISHNAN, R., AND KURTENBACH, G. 2002. Stylecam: Interactive stylized 3d navigation using integrated spatial and temporal controls. In *UIST'02*, ACM.
- BURTNYK, N., KHAN, A., FITZMAURICE, G., AND KURTENBACH, G. 2006. Showmotion: camera motion based 3d design review. In *I3D'06*, ACM.
- BYRNE, M. D., JOHN, B. E., WEHRLE, N. S., AND CROW, D. C. 1999. The tangled web we wove: a taskonomy of www use. In *CHI'99*, ACM.
- CARD, S. K., ROBERTSON, G. G., AND YORK, W. 1996. The webbook and the web forager: an information workspace for the world-wide web. In *CHI'96*, ACM.
- CATLEDGE, L. D., AND PITKOW, J. E. 1995. Characterizing browsing strategies in the world-wide web. *Comput. Netw. ISDN Syst.*
- CHEN, M., MOUNTFORD, S. J., AND SELLEN, A. 1988. A study in interactive 3-d rotation using 2-d control devices. In *SIGGRAPH'88*, ACM.
- CHITTARO, L., RANON, R., AND IERONUTTI, L. 2003. Guiding visitors of web3d worlds through automatically generated tours. In *Web3D'03*, ACM.
- CHRISTIE, M., OLIVIER, P., AND NORMAND, J.-M. 2008. Camera control in computer graphics. *Computer Graphics Forum*.
- COCKBURN, AND MCKENZIE, B. 2001. What do web users do? an empirical analysis of web use. *Int. J. Hum.-Comput. Stud.*
- DACHSELT, R., AND HINZ, M. 2005. Three-dimensional widgets revisited - towards future standardization. In *IEEE VR 2005 Workshop New directions in 3D user interfaces*.
- DACHSELT, R., AND HÜBNER, A. 2007. Three-dimensional menus: A survey and taxonomy. *Computers & Graphics*.
- DARKEN, R. P., AND SIBERT, J. L. 1993. A toolset for navigation in virtual environments. In *UIST'93*, ACM.
- DARKEN, R. P., AND SIBERT, J. L. 1996. Wayfinding strategies and behaviors in large virtual worlds. In *CHI'96*, ACM.

- DRUCKER, S. M., AND ZELTZER, D. 1994. Intelligent camera control in a virtual environment. In *Graphics Interface'94*.
- EDWARDS, J. D. M., AND HAND, C. 1997. Maps: Movement and planning support for navigation in an immersive vrml browser. In *VRML'97*, ACM.
- ELMQVIST, N., AND FEKETE, J.-D. 2008. Semantic pointing for object picking in complex 3d environments. In *Graphics Interface'08*.
- ELMQVIST, N., TUDOREANU, M. E., AND TSIGAS, P. 2008. Evaluating motion constraints for 3d wayfinding in immersive and desktop virtual environments. In *CHI'08*, ACM.
- ELVINS, T. T., NADEAU, D. R., AND KIRSH, D. 1997. Worldlets—3d thumbnails for wayfinding in virtual environments. In *UIST'97*, ACM.
- FITZMAURICE, G., MATEJKA, J., MORDATCH, I., KHAN, A., AND KURTENBACH, G. 2008. Safe 3d navigation. In *I3D'08*, ACM.
- GALYEAN, T. A. 1995. Guided navigation of virtual environments. In *SI3D'95*, ACM.
- GOESELE, M., AND STUERZLINGER, W. 1999. Semantic constraints for scene manipulation. In *Spring Conference in Computer Graphics'99*.
- GRAMMENOS, D., MOUROUZIS, A., AND STEPHANIDIS, C. 2006. Virtual prints: Augmenting virtual environments with interactive personal marks. *Int. J. Hum.-Comput. Stud.*
- GRAN, C. A., ALCOCER, P. P. V., AND GONZALEZ, M. F. 2004. Way-finder: Guided tours through complex walkthrough models. *Comput. Graph. Forum*.
- HACHET, M., DECLE, F., KNODEL, S., AND GUITTON, P. 2008. Navidget for easy 3d camera positioning from 2d inputs. In *3DUI'08*, IEEE.
- HAGEDORN, B., AND DÖLLNER, J. 2008. Sketch-based navigation in 3d virtual environments. In *Smart Graphics'08*, Springer.
- HAND, C. 1997. A survey of 3d interaction techniques. *Comput. Graph. Forum*.
- HANSON, A. J., AND WERNERT, E. A. 1997. Constrained 3d navigation with 2d controllers. In *Visualization'97*, IEEE.
- HERDER, E. 2006. *Forward, back and home again: analyzing user behavior on the web*. PhD thesis, University of Twente, Twente, The Netherlands.
- IGARASHI, T., KADOBAYASHI, R., MASE, K., AND TANAKA, H. 1998. Path drawing for 3d walkthrough. In *UIST'98*, ACM.
- JANKOWSKI, J. 2011. *Hypertextualized Virtual Environments*. PhD thesis, DERI NUIG, Galway, Ireland.
- KHAN, A., KOMALO, B., STAM, J., FITZMAURICE, G., AND KURTENBACH, G. 2005. Hovercam: interactive 3d navigation for proximal object inspection. In *I3D'05*, ACM.
- KIM, J.-S., GRAČANIN, D., MATKOVIĆ, K., AND QUEK, F. 2008. Finger walking in place (fwip): A traveling technique in virtual environments. In *Smart Graphics'08*, Springer.
- LOOSER, J., COCKBURN, A., AND SAVAGE, J. 2005. On the validity of using first-person shooters for fitts' law studies. In *People and Computers XIX (Volume 2): British Computer Society Conference on Human Computer Interaction*.
- MACKINLAY, J. D., CARD, S. K., AND ROBERTSON, G. G. 1990. Rapid controlled movement through a virtual 3d workspace. *SIGGRAPH*.
- NIELSEN, J. 2000. *Designing Web Usability: The Practice of Simplicity*. New Riders Publishing.
- NIELSON, G. M., AND OLSEN, JR., D. R. 1987. Direct manipulation techniques for 3d objects using 2d locator devices. In *SI3D'87*, ACM.
- NIEUWENHUISEN, D., AND OVERMARS, M. H. 2003. Motion planning for camera movements in virtual environments. Tech. rep., Utrecht University.
- OSKAM, T., SUMNER, R. W., THUEREY, N., AND GROSS, M. 2009. Visibility transition planning for dynamic camera control. In *Symposium on Computer Animation'09*, ACM/Eurographics.
- PHILLIPS, C. B., BADLER, N. I., AND GRANIERI, J. P. 1992. Automatic viewing control for 3d direct manipulation. In *SI3D'92*.
- PIERCE, J. S., CONWAY, M., VAN DANTZICH, M., AND ROBERTSON, G. 1999. Toolspaces and glances: storing, accessing, and retrieving objects in 3d desktop applications. In *I3D'99*, ACM.
- REISMAN, J. L., DAVIDSON, P. L., AND HAN, J. Y. 2009. A screen-space formulation for 2d and 3d direct manipulation. In *UIST'09*, ACM.
- ROBERTSON, G. G., CARD, S. K., AND MACKINLAY, J. D. 1993. Information visualization using 3d interactive animation. *Commun. ACM*.
- ROBERTSON, G., VAN DANTZICH, M., ROBBINS, D., CZERWINSKI, M., HINCKLEY, K., RISDEN, K., THIEL, D., AND GOROKHOVSKY, V. 2000. The task gallery: a 3d window manager. In *CHI'00*, ACM.
- ROUSE, R. 2000. *Game Design - Theory and Practice*. Wordware Publishing Inc.
- SALÉN, K., AND ZIMMERMAN, E. 2003. *Rules of Play: Game Design Fundamentals*. The MIT Press.
- SALOMON, B., GARBER, M., LIN, M. C., AND MANOCHA, D. 2003. Interactive navigation in complex environments using path planning. In *I3D'03*, ACM.
- SELLEN, A. J., MURPHY, R., AND SHAW, K. L. 2002. How knowledge workers use the web. In *CHI'02*, ACM.
- SHOEMAKE, K. 1992. Arcball: a user interface for specifying three-dimensional orientation using a mouse. In *Graphics Interface'92*.
- SMITH, G., SALZMAN, T., AND STUERZLINGER, W. 2001. 3d scene manipulation with 2d devices and constraints. In *Graphics Interface'01*.
- STRAUSS, P. S., AND CAREY, R. 1992. An object-oriented 3d graphics toolkit. In *SIGGRAPH'92*, ACM.
- STRAUSS, P. S., ISSACS, P., AND SHRAG, J. 2002. The design and implementation of direct manipulation in 3d. In *SIGGRAPH'02 Course Notes*, ACM.
- TAN, D. S., ROBERTSON, G. G., AND CZERWINSKI, M. 2001. Exploring 3d navigation: combining speed-coupled flying with orbiting. In *CHI'01*, ACM.
- TAUSCHER, L., AND GREENBERG, S. 1997. How people revisit web pages: empirical findings and implications for the design of history systems. *Int. J. Hum.-Comput. Stud.*
- THORNDYKE, P. 1982. Differences in spatial knowledge acquired from maps and navigation. *Cognitive Psychology*.
- VAN BALLEGOOIJ, A., AND ELIENS, A. 2001. Navigation by query in virtual worlds. In *Web3D'01*, ACM.
- VAN DAM, A. 1997. Post-wimp user interfaces. *Communications*.
- VINSON, N. G. 1999. Design guidelines for landmarks to support navigation in virtual environments. In *CHI'99*, ACM.
- WARE, C., AND OSBORNE, S. 1990. Exploration and virtual camera control in virtual three dimensional environments. In *SI3D'90*, ACM.
- WEINREICH, H., OBENDORF, H., HERDER, E., AND MAYER, M. 2008. Not quite the average: An empirical study of web use. *ACM Trans. Web*.
- WERNERT, E. A., AND HANSON, A. J. 1999. A framework for assisted exploration with collaboration. In *Visualization'99*, IEEE.
- WILSON, G., 2006. Off with their huds!: Rethinking the heads-up display in console game design. <http://www.gamasutra.com/>.