Metaphor and Storytelling in Interface Design for Virtual Reality

Andreas Kratky(✉)

Interactive Media Division, School of Cinematic Arts,
University of Southern California, 3470 McClintock Ave., SCI 201Q,
Los Angeles, CA 90089-2211, USA
akratky@cinema.usc.edu

Abstract. Virtual Reality has – again – become the target of substantial interest, research and industry growth. In the current market situation it is aimed at a general audience rather than expert users and therefore requires a fundamental rethinking of how we conceive of human computer interaction in Virtual Reality. In comparison to the established methods of designing interfaces for the desktop environment and for mobile applications numerous changes need to be considered. Even though the use of metaphors has become looser and more abstract, it is still the common way of providing an easily graspable conceptual model of the functions and behaviors of an application, be it an operating system or a task-oriented application. How do metaphors work in Virtual Reality? How does the relationship between the metaphorical environment and the environment of operation change? What kind of cognitive support structures are necessary for the perceptual situation of Virtual Reality? Drawing from an interdisciplinary set of theories we will address these questions through high-level analysis and develop methodologies to recast the design principles for the creation of user interfaces for Virtual Reality.

Keywords: Virtual reality · User interface design · Storytelling · Metaphor · Urban planning · Worldbuilding

1 Introduction

1.1 Evolution of Interface Metaphors

For the last thirty years the desktop metaphor has helped "the rest of us [1]" to use computers more easily and more efficiently. Interface metaphors are the dominant approach to communicate the conceptual structure of software applications and support the user in forming cognitive models of what a program can do and how to operate it. They are the backbone of how we conceive of computer human interfaces. When Apple introduced the first Macintosh computer in 1984 it was marketed as a computer for all those who were not specifically trained in the use of computers and who did not want to go through an extensive training. The marketing highlighted the ease of use and quick learning process realized through several new inventions, such as the Graphical User Interface, direct object interaction with the mouse, and the desktop metaphor. The desktop metaphor was intended to provide a conceptual framework for users to more

M. Antona and C. Stephanidis (Eds.): UAHCI 2016, Part II, LNCS 9738, pp. 287–300, 2016.
DOI: 10.1007/978-3-319-40244-4_28

easily understand the environment of the computer, its actions, and the options it offered. As all users were expected to be familiar with the office environment and the handling of documents and folders they could use this existing knowledge and transfer it to the computer operations. The metaphor was intended to accommodate an intuitive understanding of the objects and operations that can be expected in the computer. The idea of direct interaction with documents, the possibility to move them in and out of folders, copy them, throw them into the trash bin by dragging them with the mouse, was a way of getting around learning the complicated commands that were used so far to carry out this type of operation. Since the early days of personal computing, the desktop has been the standard of human computer interface design. After Apple introduced the concept to the market in 1984, Microsoft adopted it a year later with the Windows 1.0 platform and since then numerous implementations of the same idea have been produced. Today the metaphor seems an almost natural image for the use of computers. A closer examination, though, complicates this image.

Many activities carried out with a personal computer are revolving around documents and many people use them for text processing and table calculations in an office environment, but the mainstay of Macintosh users was in the creative field and graphic design rather than in professional office work. The office environment was dominated by IBM and computers running Microsoft operating systems and Apple did not get a significant market share in this field. Along with the first Macintosh software applications like *MacDraw* and *MacPaint* were released, which founded the strong position of the Macintosh in the design field. This software also introduced a metaphor-mismatch: A paint brush, canvas and paint bucket are not normally used in an office, they are icons of the creative design or art studio, which stereotypically is perceived as the counter-concept to the stiff and rigid office environment. The office metaphor made perfect sense in its original setting on the Xerox Alto and Xerox Star computers as these machines were conceived specifically to prototype "the office of the future [2]." The concept was intended to serve users who were focused on producing and managing reports and neither knowledgeable nor interested in computers. For these users the decision was to make the computer as invisible as possible [3]. This was to be achieved by allowing the users to transfer their existing knowledge about the office procedures to the computer. Familiar objects like paper, folders and file cabinets translated into digital icons were supposed to make this transfer easy. The system was made purely for document processing and all software applications were started with the system, so users only interacted with the documents themselves and did not have to care about launching applications for specific tasks [4]. What may seem limiting from today's point of view made sense for Xerox, the companies business was document processing and after having made millions with copying machines the company wanted to make sure its business grew into the new era of digital document processing. For other computers and other purposes, the desktop metaphor was at best only a partial match and inevitably brought inconsistencies with it. While it helped to explain file management to users, it was out of tune with applications like Adobe's *Photoshop*, which employs the metaphor of the darkroom to explain its tools for photo retouching and many other applications that implement operations from other domains.

The limitations of interface metaphors and the constraints they put on designers have been criticized [5] and various concepts for improvement have been devised.

In the 1990s several attempts were made to make the metaphor more specific and to extend it to other fields and spaces for specific activities. Systems like Microsoft's *Bob*, Apple's *eWorld* or General Magic's *Magic Cap* introduced multi-room environments that allowed the user to switch rooms according to specific tasks. The login-screen, for example, was presented as the front door, there was a hallway to start different applications, which, within the metaphor, meant to enter different rooms. Magic Cap had a "game room" to play games, a "store room" to store data etc. [6]. The idea behind these systems was to provide more detail and specificity to remedy metaphor mismatches and the perceived problem of abstraction. They were intended as virtual environments inhabited by the user and additional characters such as the friendly assistant *Rover*, a retriever dog, in Microsoft's *Bob*. None of these systems had success with the users, they were perceived as overly cluttered, too cute and unusable [7–9]. Instead, the movement went away from interface metaphors altogether.

1.2 Non-metaphoric Interfaces

While the desktop metaphor has remained as the basis of most current operating systems with a graphical user interfaces, new application types, such as mobile operating systems, do not use strong metaphors anymore. A turning point was the introduction of the iPhone in 2007, a smartphone that was able to carry out several tasks that, until then, were a domain of the desktop computer. The iPhone came with a specifically designed mobile operating system, Apple's *iOS*. A year later the first version of the Android operating system was released by Google and then Microsoft's *Windows Phone.* All of these operating systems do not use an explicit comprehensive metaphor like the desktop anymore. There were local metaphors for certain applications, like the wooden bookcase in *iOS* for the iBooks application or the notepad with torn-off paper edges trying to make the digital representation of the action of note-taking more relatable and "friendly." These references to material reality disappeared in later releases. In 2007, comprehensive metaphors were not needed anymore to explain the concept of devices like smartphones or tablets, which are taking over more and more of higher-level computing tasks from desktop machines, such as real-time 3d rendering, complex design tasks, text processing of long and complex texts etc. Many of the targeted users are familiar with digital devices and the basics of computing from a user point of view, these concepts do not need to be explained anymore with the help of metaphors. The metaphors, as mentioned in Apple's *iOS* interface guidelines, are reduced to simple gestures like the operation of switches through tapping, sliding, flicking through pages etc. [10].

The maturing of digital technology and the fact that most current users have either grown up with these technologies, or at least had enough exposure to have familiarity with their basic functions, has changed the interface design methods and the cognitive models users can apply to become familiar with new technologies. The desktop metaphor still dominates the user interfaces of desktop computers mostly for questions of 'backwards compatibility:' users have become accustomed to the desktop interface as the way of operating a computer and the prior knowledge activated to use such a system does probably not come from the actual office environment anymore but rather from the digital representation of it. How people use computer operating systems has changed: familiar with the way Google's 'one-box search interface dominates internet search, users have become accustomed to using a search tool like *Spotlight* to launch applications rather than navigating through the directory structure or using a launch bar.

Now functions established in non-metaphoric operating systems like iOS migrate back into the desktop systems. An example is the split view-feature, introduced in iOS 9 and then brought to OS X 10.11, where normally this idea would have been in conflict with the metaphor of overlapping windows, simulating the overlapping documents on a desk. The uses and the prior knowledge to be applied to learning and operating new technology have shifted with a generational shift. In early mobile operating systems like *MagicCap*, which was made for devices such as Sony's *Magic Link* or Apple's *Newton*, the room and desktop metaphor was still used, trying to build on the success of the desktop operating system. Besides the overemphasized metaphor, it also seems that the office environment did not fit to a mobile device. What was appropriate for the desktop computer, which in many cases indeed stood on a desk in an office environment, was out of tune with mobile devices, which could be used in many different environments. Having an explicit setting like the office environment would be in potential dissonance with the actual operating environment and a more abstract design allows the device to merge more seamlessly with the environment of operation. The effectiveness of an interface metaphor lives from an understanding of the environment in which a task is carried out and a translation of the main components and actions of this environment. Because an office environment and the related tasks were not suitable for the interface of the *One Laptop Per Child* project, which was intended to be a learning tool for children in developing countries, the so called Zooming User Interface was created, which uses the metaphor of a community, seen in different zoom stages from *Neighborhood*, to *Groups*, to *Home*. The zoom stages also represent the network environment of the computer and are indicative of the experience using the computer [11].

2 Problems of Interface Design for VR

From this evolution of the interface metaphor we can deduct a set of questions relevant to user interface design for Virtual Reality applications. These questions are becoming important as VR is moving into general usage. In contrast to the first wave of Virtual Reality in the late 1980s and early 1990s, which was producing research and products for task-oriented applications such as simulation, vehicular control and telepresence, the current wave of VR is targeted on entertainment. In this sector we are used to specific interface devices and control schemes from computer games, using game controllers or other custom tools, and do not expect general purpose interaction strategies as the WIMP (Window, Icon, Menu, Pointer) interfaces. But as the boundary gets blurred between entertainment and task-oriented applications carried out in a VR environment, the need for effective interface structures and efficient support for users to build cognitive models for what they are doing becomes important. We begin to see a variety of task-oriented applications emerging in VR. Specifically spatial tasks such as three-dimensional drawing in Virtual Reality promise a good match between task and the affordance of the VR environment. The application *Tilt Brush* provides a new and seemingly intuitive interface for spatial drawing [12]. The option to view immersive content in a web browser such as MozVR [13] or YouTube merges the experience of VR with that of the conventional web browser. Finally VR input device-maker *Control VR* [14] imagines users at work sitting on a normal desktop interface inside the Virtual Reality space.

A direct transfer of the desktop metaphor into Virtual Reality is probably not the best solution. The role and function of metaphors is different in Virtual Reality. The most obvious difference is in the environment structure. The screen-view is modeled on the framed window into a space, which is not shared by the user. Like in a painting the user remains outside of the depicted reality and is looking at it or into it. Despite 3-dimensional elements in traditional interfaces the visual paradigm is flat perspective projection in the style of Renaissance perspective [15]. The framed view has an inherent organization that structures areas like center and periphery and the stacking order of depth layering. Most traditional interfaces make use of these principles. They place the menu bar in the top-peripheral zone of the screen, other tools and information panels are generally also organized in peripheral zones to free the view to the center where the object of attention, the document that is being written or the drawing that is being made, resides. The center periphery structure is essential to most interfaces in both categories entertainment as well as task oriented applications. A related visual rhetoric is used for the stacking windows, which, simulating the overlapping documents on a desktop, can easily be navigated by the user based on the notion of a flat surface and a stacking order along one depth axis directly perpendicular to the screen surface.

The framed, flat surface of the screen is also used to negotiate the different kinds of information that needs to be communicated to users when navigating an entertainment experience: There is the information pertaining to the content of the experience and information necessary to operate it. Both are very different levels of address, the operating information is meant solely for the user and not accessible for characters in the experience. Integrating these two different layers in a seamless and compelling way has been traditionally a challenge in interactive entertainment. The metaphor of the fourth wall, which originally comes from theater, indicating the separation between the stage and the audience, has been used to achieve this integration [16]. For example heads-up displays (HUDs) are often used in computer games, simulating a helmet display or the information a pilot sees through instruments in his cockpit or in a projected HUD on the windshield of his plane [17]. The notion of the fourth wall, the canopy separating the pilot from the exterior, implies generally a degree of separation between the user and the experience as the framing of the view provides a means of orientation and structure of the information display. In a virtual environment this separation would go against the idea of full immersion.

In VR a perspective that is similarly “restricted” as the perspective window, exists, which is the field of view in which we see in focus. The normal field of view of human beings spans ca. 200º in horizontal direction and 150º in the vertical [18]. Current VR headsets have a span of ca. 100º, which could be seen as limiting the field of view in a similar way as the framed window does. But the peripheral areas of the field of view function differently – we could say in opposition – than the frame boundaries. While the frame boundaries are limiting the view and directing the attention to the center, peripheral vision scans the extreme edges of the field of view and directs attention outwards [19]. Peripheral vision is particularly suited to detect fast motion and visual cues while the observer is in motion [20]. This kind of scanning function is an important counterpart to the high-quality vision in the small area of foveal vision [21]. In contrast to the traditional screen-based interface it is the characteristic of VR experiences that the view can indeed be attracted outward and – similar to real vision – by turning the head and directing the foveal vision to phenomena that are not in front of the user a much wider range of space can be used for interaction. The immersive space extends far beyond the bounded space of the desktop – which has tremendous potential, while at the same time bringing new questions and problems to the interface design.

One of these questions is how users resolve inconsistencies in interface metaphors. The virtual environment, in order to support orientation in the wide immersive space, has to rely on a higher degree of consistency of how this space is represented and constituted. A user working in a desktop environment is perfectly aware of the two worlds separated by the 'fourth wall' of the screen surface, his own reality and the world of the task or experience. Inconsistent metaphors, such as switching between a virtual office environment for some parts of a task and a darkroom or painter's studio for other parts of the same task, can easily be resolved. In the immersive space of virtual reality the user has a harder time to reconcile this type of mismatch because he is cut off from his physical reality, which is therefore missing as a corrective. Disposing of two transparently coupled environments, the task environment and the environment of operation, allows to more flexibly negotiate between target and source environment of the metaphor [22]. To provide the necessary cues for operation and navigation of the immersive space the conceptual models provided for the user have to be tighter and more comprehensive than the loose metaphors we are using in desktop interface design. Given this difference between the traditional desktop environment and VR we might even question if metaphors can work at all in VR.

It is clear, though, that support for the user to build a mental model of the system he is operating in and the task he is aiming to solve needs to be provided. This support, since the success of the desktop, is coming largely from interface metaphors. As Lakoff and Johnson argue, "the human conceptual system is metaphorically structured and defined [23]. A set of metaphors that is to effectively support a user in the operation of a virtual environment needs to be structured differently and has to include different considerations than the traditional design approaches. The design for such a metaphor will have to draw on multiple fields of research that have so far been considered independently and for different types of applications. In the following we will develop a heuristic for addressing the problems we elaborated here and how new methodologies for user interface design with specific consideration of the affordances of Virtual Reality environments can be formulated.

3 Interface Design Heuristics for Virtual Reality Environments

3.1 Physiological Aspects

The main characteristic of a Virtual Reality environment is that it provides a complete space which is shared between the user and the experience or the work area in which a task is solved. In contrast to the desktop interface the user's gaze in not fixed but can move around freely. In many of the current VR systems the user has to remain statically in one place and can only change the perspective by moving the head. Since the user cannot see the real space in which the interaction takes place it is very difficult to move around freely if the dimensions and potential obstacles of the virtual space and the real space do not coincide [24]. Remaining static has the downside that the user is less inclined to change the view. If seated in one position looking straight is the most comfortable direction, and the more the object of interest moves away from this central axis the less comfortable it is to look at it. Liberating the physical movement by not being seated in a chair makes the user significantly more mobile and use of the full environment more likely. Even though it is difficult to realize in the normal setting of a casual user, the first consumer VR systems with user tracking in space are approaching the market [25].

With a mobile gaze and eventually a mobile viewer the main organizational principle of the Virtual Reality environment is spatial and the design of it has to ensure that the user is able to navigate it with ease. Rather than stacking documents on a table and organizing them within a fixed frame information and interaction affordances are distributed in space. Designs that aim to support the user in constructing cognitive models of the VR application will have to focus on spatial orientation. The desktop environment mostly excluded the physical reality of the user and reduced orientation in the environment the user interacts with to rational and visual procedures. In the VR environment the physiological reality plays an important role and is part of the processes used for orientation and the construction of a mental model of the environment. We have to distinguish two categories of orientation in the case of VR, the first considering a static user, seated in a chair and only navigating by moving the head and upper body to a limited degree; and the second, considering a user freely moving within the bounds of a room in which user's movements are tracked and translated into the VR environment. In the first case the physiological senses involved into determining the body's position in space are visual cues and the vestibular system, and, to a comparatively smaller degree, proprioceptive cues stemming from the neck and upper body movement, and auditory cues [26]. In the second case the proprioceptive component as well as the tactile component play important roles. The cues from these different sensory systems are integrated into an internal model of the physical position in space [27] and they have to be considered in the design and orientational structure of the VR application. Disregarding these cues and their consistent integration can lead to feelings of motion- or 'cybersickness' and disorientation [28, 29].

The spatial organization of information and interactive affordances enables users to interact through gestures and body movement. In the case of a static user this will mostly be deictic gestures of pointing or looking in certain directions and at certain objects in the virtual environment. A Virtual Reality display in contrast to the traditional bounded-screen display is perceived as an extremely large display that exetends beyond the user's field of view. Large displays incite users to use more deictic gestures as part of their conceptual reference system regarding the displayed environment [30]. Gestures like these can be used for selecting, activating etc. functions in the environment [31]. In the second case this interaction vocabulary is extended to include movement through space, such as walking up to certain objects, revealing more information through a change of position.

In both cases self-perception of the user within the virtual environment is required. Already in the first wave of Virtual Reality research the data-glove was developed to provide a means to the user to be present and have manifest agency in the environment [32]. While this form of presence required complicated and cumbersome technology back then, today multiple tracking methods are available to track the user's hand movement and translate it into the environment [33]. Features of structuring an environment for predominantly visual navigation rely on visibility, salience and other visual cues allowing to estimate distance, order, occlusion etc.

In case of a mobile user the dominant ordering principles include, besides the described visual cues, multiple aspects of body movement. Gross body movement, such as movement through a tracked space, can be used to support the construction of mental models and support orientation of users in a virtual environment. Structuring this kind of movement through space in a meaningful way is part of the design aspects that need to be considered to support users in carrying out tasks in a virtual environment. Research in the relationship between gross body movement and cognitive processes has been done in several separate fields. The research threads that inform our set of questions most fruitfully are haptics and grounded cognition [34] on one side and choreography [35] on the other. Both fields reveal specific processes of building mental models and constituting memory distinct from other sensory modalities like vision, which can be used to structure the movement sequences and rhythms of traversing and interacting with objects in the virtual environment. The physiological involvement and its targeted design can leverage additional potential for meaning-making and retention of actions necessary to accomplish a task. In this respect a VR environment may be superior to the traditional desktop environment for certain tasks and entertainment purposes. In designing movements and gestures the designer has to be aware of physiological constraints of the human body and its biomechanics. Large gestures can be straining and causing fatigue and hyperextension of joints has to be avoided [36].

3.2 Conceptual Modeling Through Story and Metaphor

In order to successfully benefit from the physiological aspect of operating in a virtual environment and to benefit from the large space it can provide, the user has to be able to orient herself with ease within the environment. Head-mounted displays have a restricted field of view, requiring the user to be able to accommodate to the restrictions [37] and to develop a memory for the parts of the environment that are out of sight as user is moving and changing perspective. This locational memory is part of the mental model the user constructs when solving a task or going through an entertainment experience in Virtual Reality. It is in part achieved through the physiological memory acquired through body movement and in part through metaphors or other forms of conceptual modeling. As elaborated earlier, metaphors work well for the predominantly rationally structured operations of desktop user interfaces. To support the conceptual modeling of a virtual environment we need potentially more complex and comprehensive models.

Metaphorically structured models normally employ several different metaphors that correspond to separate aspects of the model. As Lakoff and Johnson write, metaphorical structuring is partial and relies on a form of lexicon rooted in language to resolve the individual metaphors and construct a higher-level model from them [23]. To provide the comprehensive structuring that is desirable in a virtual environment instead of using partial metaphors we can use elements of storytelling, which have a more specific and tighter structure.

Storytelling can work as a structure to organize both space and processes that take place in the space. Classic narrative theory distinguishes between two parts of a narrative text, the story, which refers to the actions that make up the story, and the discourse, the order of how the actions are communicated. The actions exist independently from the way they are relayed through words, images or other media forms [38]. They can, for example, be communicated in a different temporal structure than they actually happened, which enables different ways of organizing the same narrative content through effects like flash-backs and time inversions. Both story and discourse have a temporal structure as well as a spatial structure. The spatial structure of the story organizes all elements that are part of the story, lists them, describes and places them inside the space in which the story unfolds [38]. The existence and location of these elements is again independent from how relationships between them unfold over time. The discourse space, in contrast, is the order of how the story elements (also called "existents" [38]) are arranged in the flow of communicating them, i.e. which elements are visible, in which perspective they are visible etc. This fundamental structure of narrative indicates how it could be used as a method of structuring and communicating operations in a virtual environment. It can serve as an ontological structure providing all elements of an environment and within this space different temporal processes of task execution or entertainment can be devised. The discourse layer of the narrative structure can provide perspectives and selections for how a user perceives the environment and in response plans her interactions. Storytelling provides the foundational layers for an efficient structuring of the interaction flow in a virtual environment.

The methodology of 'worldbuilding' has become a popular tool to develop environment descriptions that focus on the story-space aspect. Inspired by the philosophy of possible worlds, which regards a given state of a world as just one possible state among many other alternative world descriptions contained in one larger, maximally inclusive world. The notion of possible worlds is an epistemic tool that considers situations as "simply structured collections of physical objects [39] " and aims to formulate the set of possible logical statements that are part of the maximally inclusive world. In storytelling the idea of possible worlds has been used as a speculative tool to design alternative worlds that could have been possible if certain historic events did not happen, or happened differently. Possible worlds are consistent within themselves and even if they do not correspond to reality as it is known they provide compelling descriptions of alternative realities. Worldbuilding is the act of constructing possible fictional worlds. It is a creative method used in writing as well as filmmaking and game design to create environments for possible stories. This approach of building narrative from the space-component is in contrast with traditional models of storytelling, which tend to construct the story from the chain of events rather than from the world in which the events happen. As Mark J. P. Wolf writes, storytelling and worldbuilding are often in conflict in that the latter produces digressions and descriptions that slow down the narrative flow [40]. In game design, for example, worldbuilding is a method of creating environments in which multiple stories can take place and different characters can be accommodated, in this sense allowing for a bandwidth of variations in gameplay. For our current set of questions worldbuilding offers a method to formulate consistent models of worlds that can help users to conceptualize the virtual environment they are operating in. Those descriptions can be in conflict with reality, as long as they provide a compelling and consistent description of the environment and enable the user to form expectations and concepts about the environment. An example for such a world-description could be Edwin A. Abbott's novella "Flatland [41]," which describes a world that exists in only two dimensions.

3.3 Patterns of Visual Organization

A method of investigation that merges the aspects of space, spatial perception and choreography in a unique way is Kevin Lynch's analysis of the modern cityscape. This method can serve as a model for the creation of environmental cues to facilitate orientation and mental mapping of users in a virtual environment and it provides a suitable complement to the conceptual modeling through metaphors and storytelling. The urban planner Kevin Lynch conceives of the city as a spatio-temporal construct, a large built environment that is experienced by its users over long time spans. His approach examines the city as not only consisting of its architectural units, but equally of the inhabitants moving across it and pursuing multiple activities in it. Each single inhabitant has a different experience and perception of the city depending on her interests, experiences, associations and memories, which makes for a double structure of the image of a city: the public image, shared by many inhabitants, and the private image, shaped by the specific perception of an individual. In this sense Lynch formulates a method that directly fuses with the narrative structure we discussed above as the basis for conceptual modeling techniques and delivers a suitable aesthetic and experiential heuristic to address the design of environment-based user interfaces for Virtual Reality. Through observations and interviews Lynch studied how people create a mental image of the city, elaborating principles for what he called the "legibility of the city [42]." The term legibility refers to organizational principles of the urban structure that facilitate recognition and memorization. He identifies five main properties that support imageability: paths, edges, districts, nodes and landmarks. These properties are the salient features of a city inhabitants use to structure their mental representation of their surroundings. Paths are the streets, transit lines and walkways people use to navigate the city. They structure the sequence in which places are perceived and form the orientational grid of the users. The edges are boundaries that delimit sections in the cityscape. These can be walls, canals, streets etc., not in their function as a pathway, but in their function as delimiters akin to the barrier a freeway poses to pedestrians. Districts are areas that are perceived as larger geometric units. These may coincide with neighborhoods but they can also be individually construed subsections. Nodes are places in which main pathways intersect, meeting points with high traffic frequency or different kinds of discernible concentrations. And finally landmarks are points of reference in a city that stand out such as high buildings, signs, mountains or other features of singular quality and visibility [42].

The legibility properties elaborated by Lynch can serve as both a vocabulary to construct virtual environments and deliberately integrate certain features, as well as a methodology of designing and evaluating interface designs. Lynch's approach to urban planning, even though it originated in a completely different field, can function as an extension of the methods of design ethnography and ethnomethodology that emerged in the late 1980s and 1990s, when the personal computer entered the workplace and the home at a large scale. Ethnography as a tool for user interface design developed at Xerox at the time when the company was working on the Alto and Star systems. Triggered by customer complaints about usability issues with copiers observational studies were used to understand the reasons for those issues. This approach was generalized and extended as a tool to systematically study the underlying conceptions about the nature of a task that inform the design of every tools made to support the task [43]. Design ethnography has since become an established component of systems design and human computer interaction, focusing on "studying work in the wild" to analyze potential mismatches between systems functionality and the methods employed by users to organize their interactions and do their work [44]. The notions of Lynch's studies are apt to extend the methods of design ethnography in a direction specifically relevant to the design of user interfaces for virtual environments.

4 Conclusions and Future Work

This paper brings together several threads of theory and design research from a heterogeneous set of fields that are normally considered completely separate from each other to formulate a heuristic for the design of user interfaces for Virtual Reality applications. The challenges that emerge in the task of creating interfaces and accommodating work and entertainment in Virtual Reality have distinctively different properties than the design principles in classic desktop interface design. As prominent aspects stand out the involvement of the physiological reality of the user as part of her interaction with a computer and the particular focus on spatial navigation and action. Drawing from linguistics, narrative theory, urban planning and ethnography we are giving a high-level overview over the specific problems of user interface design for virtual environments and develop a heuristic to address these problems. The heuristic is formulated with consideration of the requirements of task-oriented applications that start emerging in the realm of Virtual Reality, but the foundational principles promise to be applicable also for entertainment oriented experiences.

The perspective of this paper focuses on aspects of visuality and how it influences the construction of mental representations. As future areas of inquiry it will be necessary to investigate the role of haptics in this process. Given the importance of the physiological involvement, haptics play an important role and one of the drawbacks of current VR systems is that haptics are only addressed in a rudimentary way. It is imaginable that specifically designed interaction objects that can be imbued with varying functions and roles are suitable avenues to involve the haptic senses in a more controlled and developed way into the design process.

The perspective shifts and interdisciplinary inspirations described in this paper are apt to also provide useful correctives in the consideration of traditional and non-traditional user interface design. In particular in mobile and distributed user interfaces aspects of environmental properties and orientation come into play that can be approached with similar methods as developed here in the context of VR.

References

1. Apple advertisement campaign (1984)
2. Perry, T.S., Wallich, P.: Inside the PARC The information architects: Insiders who became outsiders describe the trials and successes of the Xerox Palo Alto Research Center as it sought to create the electronic office of the future. IEEE Spectr. **22**(10), 62–76 (1985). IEEE
3. Johnson, J., Roberts, T.L., Verplank, W., Smith, D.C., Irby, C.H., Beard, M., Mackey, K.: The xerox star: a retrospective. Computer **22**(9), 11–26 (1989)
4. Smith, D.C., Irby, C., Kimball, R., Harslem, E.: The star user interface: an overview. In: Proceedings of the National Computer Conference, AFIPS 1982, 7–10 June 1982, pp. 515–528. ACM (1982)
5. Gentner, D., Nielsen, J.: The anti-mac interface. Commun. ACM **39**(8), 70–82 (1996)
6. General Magic Inc.: Design and Magic Cap. Icras, Sunnyvale, CA, 24 April 2000
7. Hassard, S.T., Blandford, A., Cox, A.L.: Analogies in design decision-making. In: Proceedings of the 23rd British HCI Group Annual Conference on People and Computers: Celebrating People and Technology, BCS-HCI 2009, pp. 140–148. British Computer Society (2009)
8. Dvorak, J.: The bottom 10: Worst Software Disasters. PC Magazine (2004)
9. Newman, M.: Bob is dead; long live Bob. Pittsburg Post-Gazette (1999)
10. Apple Inc.: IOS Human Interface Guidelines: UI Design Basics. Apple Inc. (2015)
11. Sugar Labs: Human interface guidelines/the laptop experience/zoom metaphor. Sugar Labs (2009)
12. Cass, S.: Tiltbrush: The killer app for VR. IEEE Spectrum. IEEE (2016)
13. MozVR. http://mozvr.com/
14. Control VR- the future of virtual reality, animation & more by the control VR team – kickstarter. https://www.kickstarter.com/projects/controlvr/control-vr-motion-capture-for-vr-animation-and-mor/description
15. Friedberg, A.: The Virtual Window: From Alberti to Microsoft. MIT Press, Cambridge (2009)
16. Conway, S.: A circular wall? Reformulating the fourth wall for video games (2009). www.gamasutra.com, http://www.gamasutra.com/view/feature/4086/a_circular_wall_reformulating_the_.php

17. Iacovides, I., Cox, A., Kennedy, R., Cairns, P., Jennett, C.: Removing the HUD: The impact of non-diegetic game elements and expertise on player involvement. In: Proceedings of the 2015 Annual Symposium on Computer-Human Interaction in Play, CHI PLAY 2015, pp. 13–22. ACM (2015)
18. Steinicke, F., Bruder, G., Hinrichs, K., Kuhl, S., Lappe, M., Willemsen, P.: Judgment of natural perspective projections in head-mounted display environments. In: Proceedings of the 16th ACM Symposium on Virtual Reality Software and Technology, VRST 2009, pp. 35–42. ACM (2009)
19. Johnson, J.: Designing with the Mind in Mind: Simple Guide to Understanding User Interface Design Rules. Morgan Kaufmann Publishers, Elsevier, Amsterdam, Boston (2010)
20. McKee, S.P., Nakayama, K.: The detection of motion in the peripheral visual field. Vis. Res. **24**(1), 25–32 (1984)
21. Lachenmayr, B.: Visual field and road traffic. How does peripheral vision function? Ophthalmologe **103**(5), 373–381 (2006)
22. Khoury, G.R., Simoff, S.J.: Elastic metaphors: Expanding the philosophy of interface design. In: Selected Papers from Conference on Computers and Philosophy, CRPIT 2003, vol. 37, pp. 65–71. Australian Computer Society, Inc., (2003)
23. Lakoff, G., Johnson, M.: Metaphors We Live By. University of Chicago Press, Chicago (1980)
24. Greuter, S., Roberts, D.J.: SpaceWalk: movement and interaction in virtual space with commodity hardware. In: Proceedings of the 2014 Conference on Interactive Entertainment, IE 2014, pp. 1:1–1:7. ACM (2014)
25. Vive | home. http://www.htcvive.com/us/
26. Bortolami, S.B., Rocca, S., Daros, S., DiZio, P., Lackner, J.R.: Mechanisms of human static spatial orientation. Exp. Brain Res. **173**(3), 374–388 (2006)
27. Merfeld, D.M., Zupan, L., Peterka, R.J.: Humans use internal models to estimate gravity and linear acceleration. Nature **398**(6728), 615–618 (1999)
28. Davis, S., Nesbitt, K., Nalivaiko, E.: A systematic review of cybersickness. In: Proceedings of the 2014 Conference on Interactive Entertainment, IE 2014, pp. 8:1–8:9. ACM (2014)
29. Kitazaki, M., Nakano, T., Matsuzaki, N., Shigemasu, H.: Control of eye-movement to decrease ve-sickness. In: Proceedings of the ACM Symposium on Virtual Reality Software and Technology, VRST 2006, pp. 350–355. ACM (2006)
30. Bao, P., Gergle, D.: What's "this" you say?: The use of local references on distant displays. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI 2009, pp. 1029–1032. ACM (2009)
31. Haffegee, A., Alexandrov, V., Barrow, R.: Eye tracking and gaze vector calculation within immersive virtual environments. In: 2007 Proceedings of the ACM Symposium on Virtual Reality Software and Technology, VRST 2007, pp. 225–226. ACM (2007)
32. Sturman, D.J., Zeltzer, D.: A survey of glove-based input. IEEE Comput. Graph. Appl. **14**(1), 30–39 (1994)
33. Leap motion | mac & PC motion controller for games, design, virtual reality & more. https://www.leapmotion.com/
34. Barsalou, L.W.: Grounded cognition. Annu. Rev. Psychol. **59**, 617–645 (2008)
35. McCarthy, R., Blackwell, A., de Lahunta, S., Wing, A., Hollands, K., Barnard, P., Marcel, A.: Bodies meet minds: Choreography and cognition. Leonardo **39**(5), 475–477 (2006)
36. Lin, J., Wu, Y., Huang, T.S.: Modeling the constraints of human hand motion. In: Proceedings of the Workshop on Human Motion (HUMO 2000), p. 121. IEEE Computer Society (2000)

37. Creem-Regehr, S.H., Willemsen, P., Gooch, A.A., Thompson, W.B.: The influence of restricted viewing conditions on egocentric distance perception: Implications for real and virtual indoor environments. Perception **34**, 191–204 (2005). SAGE Publications
38. Chatman, S.B.: Story and Discourse: Narrative Structure in Fiction and Film. Cornell University Press, Ithaca (1978)
39. Menzel, C.: Possible worlds. In: The Stanford Encyclopedia of Philosophy (2016)
40. Wolf, M.J.P.: Building Imaginary Worlds: The Theory and History of Subcreation. Routledge, New York (2013)
41. Abbott, E.: Flatland: A Romance of Many Dimensions. Penguin Books, New York (1998)
42. Lynch, K.: The Image of the City. MIT Press, Cambridge (1960)
43. Suchman, L.A.: Human-Machine Reconfigurations: Plans and Situated Actions. Cambridge University Press, Cambridge (2007)
44. Crabtree, A., Rouncefield, M., Tolmie, P.: Doing Design Ethnography. Springer, New York (2012)