Interaction in Virtual Worlds for Digital Storytelling

Gerald Melles

Hochschule für Angewandte Wissenschaften Hamburg (HAW), Berliner Tor 5, 20099 Hamburg, gerald.melles@haw-hamburg.de

Abstract. This paper provides an overview of Virtual Reality technologies as they currently being explored in the scientific community. It also examines key concepts and technologies which enable interaction in virtual worlds.

1 Motivation

Digital Storytelling is a term with widely varying definitions. Its widest interpretation may be the use of digital media in the conveyance of narrative. For the purposes of this paper, digital storytelling shall be defined as the resolution of linear narrative structures by means of user interaction in digital media. This definition highlights what may be its most significant difference from traditional narrative formats: Probably for the first time, technological advances have allowed the audience to interact with the narrative medium - and thereby influencing the narrative - without the need for a human storyteller to do so manually (or even be present at all). [1] This offers more freedom to storytellers and writers, who can now write complex, non-linear stories. While this ability to influence the narrative may make it easier for the audience to immerse themselves in the story, technical limitations of input and output devices can still make the interactions feel unintuitive, decreasing immersion and willing suspension of disbelief. New Virtual Reality (VR) technologies may mitigate this, enabling humans to experience stories in a natural, intuitive and above all realistic way - blurring the borders between imagination and reality.

2 Introduction and purpose of this paper

This paper will give a brief overview of the concepts and devices relating to virtual worlds in general and input and output in VR in particular. Its focus lies in the means which make (human) interaction with such worlds possible, as this is the enabling feature for non-linear digital storytelling within virtual worlds. Given the width of this field and the resulting briefness with which some of the

[1] Though arguably some very limited means for users to change the narrative themselves existed, such as pick-your-own-ending books

topics will have to be explained, the referenced literature may be considered recommended reading for further and in-depth study. [4] deserves special mention, as it offers more in-depth discussions of most of these abbreviated topics, in particular those relating to input and output in VR. The descriptions of the different concepts for sensory input and output channels in VR are primarily meant to provide the basis for the following and more verbose discussion of the interaction concepts which rely on them.

3 Virtual Reality

Virtual Reality (VR) can be described as the simulation of reality by means of generated and orchestrated stimuli. These stimuli are designed for humans to interpret them as part of a coherent reality which does not, in fact, exist.
To the present day, many concepts and devices have been developed for the generation of stimuli - though only a few of our senses have received the vast majority of attention. Primary focus of scientific research has been the generation of auditory, visual and haptic stimuli.

3.1 The subjective nature of reality

Reality often differs significantly from the human perception of it. Humans receive input from the world via their senses, then interpret that input based on their experience and instinct. The result may not correspond completely with reality. One example for this is the visual perception of color (wavelength of light in the spectrum visible to us). We can have difficulty distinguishing between colored light being reflected from a white object and white light being reflected from a colorful object. This allows for misinterpretations. Optical illusions, such as grid illusions (a.k.a. Hermann grids) also serve well to illustrate the subjectiveness of perceived and interpreted stimuli which shape our view on reality. Such discrepancies open possibilities for deceiving our senses, thereby allowing for the creation of virtual realities.

3.2 Senses

Human senses can be broadly categorized into interoceptive senses which receive stimuli from internal processes in our bodies and exteroceptive senses which receive stimuli from our surroundings. VR focuses on the generation of external stimuli, so the exteroceptive senses most relevant to VR shall be mentioned here briefly:

- sight
- hearing
- smell
- taste
- touch

– balance

The first four originate in Aristotles *De Anima* (cf. [8]), while balance (a.k.a. equilibrioreception) has more recently started being considered another sense (cf. [2]). The existence and relative importance of each sense is subject of discussion in different fields, such as philosophy, anthropology, biology and medicine. As such, an in-depth discussion is not within the purview of this paper but in that of specialized literature, such as [20]. In VR research, the most prominently featuring senses are sight, hearing and touch. Nevertheless, the others should be kept in mind when developing VR systems. This is especially true for balance, since a sensory mismatch between equilibrioreception and the other senses (most notably vision) has been found to be of considerable concern, potentially resulting in a loss of balance [2] or the so-called simulator sickness (cf. [3]).

3.3 Immersion

The term immersion has no universally agreed-upon definition. For this paper, the used definition shall be the technical definition by Slater and Wilbur (1997): "Immersion is a description of a technology, and describes the extent to which the computer displays are capable of delivering an inclusive, extensive, surrounding and vivid illusion of reality to the senses of a human participant. Inclusive (I) indicates the extent to which physical reality is shut out. Extensive (E) indicates the range of sensory modalities accommodated. Surrounding (S) indicates the extent to which this virtual reality is panoramic rather than limited to a narrow field. Vivid (V) indicates the resolution, fidelity, and variety of energy simulated within a particular modality (for example, the visual and colour resolution). Vividness is concerned with the richness, information content, resolution and quality of the displays." [19]

3.4 AR

Augmented Reality (AR) is a variation of VR which differs in one key aspect: It does not endeavor to replace stimuli from reality entirely, but to "augment" them with generated stimuli instead. An example for this could be a heads up display (HUD) displayed in AR goggles, e.g. as semi-transparent numbers indicating wind speed and direction in a sailor's otherwise unaltered field of vision. Storytelling could similarly use AR to enrich (alter) reality by means of augmentation rather than replace it, such as by simulating artificial people onto a real environment. This "flavor" of VR use in storytelling is not in the focus of this paper.

3.5 VR input

In order to allow any form of interaction, input devices process user input, passing the result on to the simulation of a virtual world, which may then react in

[2] i.e. the ability to stand, not the sense

some way. User input can be any human action which may be detected by VR input devices. It is essentially the counterpart to the human perception of virtual worlds: the system's "perception" of a human. As such, most input devices could be assigned a sense they mimic. The devices' "perception" of human actions is most commonly called *tracking*, especially if they do so constantly (like a camera) rather than sporadically (such as a button). In recent years, the most prevalent tracking concept is that of optical tracking, which will be explained first and in more detail. Others include the use of acoustic, electromagnetic and inertial sensors (cf. [4], chapter 4). Other than by the sense they mimic, input devices may also be categorized by the kind of a human's observable actions they are intended to track, such as movements of the head and eyes, hands and fingers.
Traditional and less intuitive input methods such as the use of a mouse or keyboard are also still being employed. Their main advantage is that many users are practiced in their use already. There have been attempts at adapting them to make them more suitable in a three-dimensional, virtual world, such as the 3D mouse or the use of finger tracking (see below) with a virtual keyboard representation inside the VR.

Optical tracking

using markers Marker-based tracking uses two basic components: An optical sensor, such as a camera, and an optical marker, such as an LED (or combination of LEDs and reflectors). The markers are visible and clearly identifiable to the optical sensor, therefore their relative positioning (to each other and the sensor) is simple to detect. They may be attached to the user, allowing the system to extrapolate the user's state (such as position, orientation, movement of body parts, cf. [18]).

markerless Markerless tracking uses only an optical sensor. Algorithms endeavor to extrapolate relevant features from the optical input, such as which parts of the incoming feed belong to the user, other algorithms then interpret these features to determine the user's state (such as position, orientation, movement of body parts, cf. [12]). Markerless techniques tend to be less accurate and reliable than their marker-based counterparts, but the lack of markers makes for other benefits: The VR system is less obtrusive and does not require the user to have markers attached to them before use (also allowing for an easier switch from user to user).

Finger tracking Our fingers may be our most intuitive way to interact with objects. Among all creatures on this earth, humans have an unusually high degree of skill when it comes to manipulating objects with our extremities, most notably our hands. The word "manipulation" itself originates in part from the latin word for hand. It is logical, therefore, that tracking of our fingers may help facilitate natural interactions with virtual worlds. Optical tracking has proven to be particularly useful for this task (cf. [21]). A persistent problem remains that of our reliance on haptic feedback when handling objects: Our hands and fingers are one of the most sensitive areas when it comes to haptic (touch) sensations. As such, it is important to consider the connection of finger tracking as input and haptic feedback as output.

Eye tracking Our eyes are a powerful tool when perceiving our surroundings. For a number of interaction concepts, their ability to focus on an object with great speed and presicion is invaluable. A number of (markerless optical tracking) techniques exist for pinpointing positions we look at (e.g. cf. [6]) It should, however, not go without mention that terms such as "gaze pointing" or "gaze tracking" are often applied to tracking of the user's head, not their eyes.

3.6 VR output

Output devices as part of a VR system are meant to provide artificial stimuli to humans as part of a coherent virtual reality. Humans rely on their senses to different extents, which means that the artificial creation of some stimuli is more important to a coherent virtual reality than that of others. For example, most humans rely mainly on visual and auditory perception of their surroundings, making the correct simulation of corresponding stimuli paramount for the simulation of a virtual environment. Nevertheless, stimuli for other senses may be just as important in different contexts. Haptic sensations, for example, are important as a feedback channel for other interactions (cf. [17]).

Visual The human reliance on visual perception of their surroundings makes correct visual simulation of a virtual world the most important aspect of a VR system in most cases. The most important categorization of visual output devices is that of whether the device is physically connected (mounted) to the user. So-called "Head-Mounted Displays" or HMDs have recently and rapidly grown in popularity and use (see section "Future Trends"). Other visual output devices project the simulation onto walls, ceiling and/or floor (such as the "CAVE" system) or use a number of displays to "tile" an area in much the same way.

Acoustic Especially for our perception of things outside of our optical field of view, we rely on acoustic senses. Humans are capable of interpreting minute differences such as latencies between the sound reaching either of their ears as sound coming from a direction. This is also known as three-dimensional hearing (cf. [13]). To facilitate a convincing three-dimensional simulation, a VR system needs to emulate sound sources in the virtual world. Much like with displays, the two main strategies are to either mount the speakers directly on or in the user's ears, such as using headphones, or to distribute a number of sound sources around the user (such as speakers for surround sound or wavefield synthesis systems). Given the user's ability to turn their head, head-mounted speaker based VR-Audio systems need to be equipped with the ability to react to this very quickly, change the corresponding output of the speakers and keep intact the illusion of the sound sources being stationary.

Haptic Haptic sensations on our skin allow us to make assumptions about an object's characteristics, such as surface structure, weight, mass and acceleration. As such, replicating haptic sensations is another focus of VR research (cf. [14], [15], [11], [17])

Other Output devices exist even for other than the aforementioned senses, but tend to be in early stages of development and of limited use. Olfactory output devices for example tend to be cumbersome (due to the difficult delivery of molecules to the nose), prohibitively expensive and can only capable of providing limited sets of stimuli ([7]).

4 VR interaction

Given the input from the aforementioned input devices, this input then needs to be interpreted. In other words: A user's action can only be used to facilitate interaction if a meaning, an intention, can be assigned to it. Following this, the system then needs to respond accordingly. This often includes some sort of output as feedback, be it to show the progress of an interaction or its outcome. Most interaction with computers outside of VR is still achieved via metaphors such as WIMP (windows, icons, menus, pointers). In virtual worlds, traditional metaphors used in 2D graphical user interfaces may be problematic or impossible to apply (cf. [4], p. 160).
For example, the user may want to select a position to indicate wanting to interact with something (such as a file) whose representation (such as an icon) is located there. On a traditional two-dimensional interface they would move the mouse to move the cursor to that position. In three-dimensional space, the first need which arises is that for an input device which permits selection in three dimensions, which could be satisfied by devices such as a 3D mouse or optical tracking system of the user's hand. However, not all points in that space can be shown to the user with the same density - they may be closer to or further away from the origin of the user's perspective. This is problematic for a number of reasons, such as the points to being either out of the user's reach (such as being farther away than they could reach with their hand) or difficult to select due to the lack of precision at far distances. These problems lead (among other things) to the need for navigational controls, which allow the user to change their perspective on the world they interact in.
This section will categorize and describe the most commonly used ways in which a user interacts with virtual worlds. To do so, some ways of categorization will have to be established. The first and most common distinction is based on the user's intention:

- Selection (of a position or part of the virtual world)

- Manipulation (of selected objects' parameters)
- Navigation (of the user's representation and/or point of view in the virtual world)
- System Interaction (with VR system attributes not directly represented in the virtual world)

The aforementioned types provide the structure of this section.
Another categorization could be made using the aspect of *natural* vs. *magical* interaction: Natural interaction concepts try to replicate the way we may interact with reality as much as possible, whereas magical interaction permits unnatural or unrealistic but more powerful and versatile interaction. For example: It is likely to feel most natural to the user to only be able to "grasp" objects in one's own immediate vicinity. Drastically increasing the reach of the user, such as by enabling interactions with all objects in the user's vision independent of their distance, is less natural, but significantly more powerful. The most appropriate interaction design generally falls somewhere between these extremes. In the following sections, the words "natural" and "magical" will be used to refer to these concepts.
A third distinction can be the user's point of view during the interaction. Egocentric interactions allow the user to interact directly with objects from their perceived position in the virtual world (i.e. that of the camera and their avatar). Exocentric interactions on the other hand permit the user a point of view on the target of the interaction not dependent on their own location and orientation. For example, a user may use a smaller, scaled down and possibly simplified representation of a part of the virtual world (like a map) to interact with the objects represented therein. This is also known as World in Miniature (or WIM).

4.1 Importance of Feedback

Feedback is another noteworthy aspect of interaction design. In reality, many interactions result in immediate feedback via sensory stimuli. For example: When we touch an object we feel its surface texture, temperature and pressure against our skin. Depending on the object and interaction, we may also hear the collision, or see the object move. Instinct and experience will lead us to expect a certain feedback from objects, a lack may result in reduced immersion. We also rely on this feedback during an interaction, for example: We need the sensations caused by an objects weight and mass to accurately throw it. Intended interactions may even be canceled as a result: When intending to pick up a cup of hot coffee, we may stop doing so and even drop it once we perceive the heat it is emitting. Feedback is not limited to touch, either: when we look at or talk to a person or animal, they may react to this, resulting in feedback. Designing appropriate feedback mechanisms may therefore be essential for some forms of natural interaction.

4.2 Selection Interactions

Selection as a VR interaction category may be defined as the set of user input which the VR system interprets as indicating a position, point, surface or object in the space of the virtual world. The most important one could be called "pointing": The user indicates a ray originating from at or near their point of view, toward the target of the selection. Its widespread use could be attributed to our reliance on sight, especially when initiating an interaction with a real-world object. Pointing and related selection concepts are discussed in a dedicated section. Many other ways to indicate a position, point or surface exist, however. For example: The user may use a designated name or coordinates of an object, by typing them into a keyboard or talking to a voice interface. While less intuitive in some cases (especially when unnatural, such as typing into a keyboard rather than speaking), these concepts are nonetheless useful in certain circumstances, such as the object being out of the user's line of sight. Especially selection by means of verbal communication may feel natural to the user while providing a potentially very powerful means of selecting objects outside of the user's field of vision.

Pointing Pointing interaction at its simplest can be defined using two vectors: a positional and a directional vector. The first specifies an origin, typically at or near the user's own position in the virtual world and therefore implicitly given. The second indicates a direction. The directional vector may be specified implicitly, such as by the orientation of the user's view frustum, or explicitly using a pointing device. This could for example be the user's tracked hand or finger, or a 3D mouse. The simplest way of using such a two-vector system to indicate a position or object is to determine the first intersection of a ray (from origin along the direction) with a surface. Most often, this is done when the user performs another simple action, such as pressing a button or pulling a trigger, though it is also possible select without such an explicit trigger. One problem with systems which allow for selection by means of *only* pointing (no trigger) is that users may have a hard time*not* selecting objects. This is also known as the "Midas Touch Problem" (cf. [4], p. 163). An alternative for trigger mechanisms, which only attempt one selection along the pointing ray, is the use of *modes*: One mode for "selection active" (where selection is constantly attempted) and one for "selection inactive" (where no selection is attempted). The user can then switch between these modes using other (system) interactions). In some VR systems, such as ones utilizing Smartphone-based HMDs (see section below), it is also not unusual to trigger a selection by specifying both vectors (by turning one's head) and keeping the resulting ray on an object for a few seconds. Here, the "Midas Touch Problem" can be avoided by not looking at any other object for long enough to trigger the selection.
The pointing range is another factor in selection interaction design. Typical limits include the approximate length of the user's arm or the boundaries of a room the user is in - most often ranges at which the user is considered to still be able to accurately select all relevant objects. It has been shown that longer ranges make

selection difficult due to a lack of accuracy, but techniques exist for leveraging this problem (see below).
A problem with pointing using rays can be the interpretation of the extent intended for selection: A user pointing at another user's avatar's eye may mean the eye, head, avatar or corresponding other user. Contextual information is needed to interpret the selection correctly. Simple pointing such as the one explained above can be enhanced in any number of ways. Using contextual targeting aids at longer distances means that even pointing near an object causes the object to be selected, provided no other object is closer to the targeted point. Alternatively, instead of using a ray along which selection may be performed a cone could be used. This is known as a "flashlight technique", since such a cone may resemble the cone of light from a flashlight. Depending on its implementation, objects which are located within the cone and perhaps also those intersecting with it would be selected.
Pointing techniques can be exocentrical as well. World in Miniature (WIM) implementations for example may allow the user to use pointing techniques on small-scale representations.

4.3 Manipulation Interactions

Manipulation interactions may be defined as the set of user input which the VR system interprets as indication of a user's desire to change an object's parameters (such as position, orientation, size or shape). This kind of interaction has a close relationship with feedback mechanisms: Having selected an object and now executing a manipulation interaction such as wanting to move then object, the user may expect to be able to see (or feel) the object move. Manipulation most often directly follows selection (as in the example above), which leads to a necessity to design both selection and manipulation interactions together, so the user may use them coherently. Exceptions exist, however, especially when a VR system is designed to mimic interactions in reality precisely: In [5], a VR system for training personnel in the manipulation of nanotubes is described. Due to the extremely small scale of these nanotubes, the same probe which allows their manipulation is also used for scanning their current position and orientation. In this case, the users needed to rely on the generated and purely haptic feedback of the VR. They had to sense and manipulate objects in the same step, without a way to separate the tasks of selection and manipulation, because one necessitated the other.
Manipulation in virtual worlds does not always need to resemble reality either, this depends on the purpose of the VR system. In some cases, it may be beneficial to give the user a way to manipulate aspects of the VR in powerful, magical ways not possible in reality, such as moving objects with gestures. This is especially important when considering interaction ranges: The range of physical interaction (esp. grasping) available to us in reality is very limited, especially when compared to our much greater visual range. Especially exocentrical techniques provide powerful alternatives. WIM implementations for example can allow the user to keep track of large-scale manipulations of entire sections of the virtual

world. For manipulation of a single object, it can also be beneficial to (upon its selection) provide the user with a simplified version of only the object, without its surroundings. This is especially the case if available input devices have significant limitations, such as allowing only two-dimensional interactions: The "Virtual Sphere" and "Arcball" techniques for example have been developed for allowing the use of essentially two-dimensional input for rotating a single three-dimensional object (cf. [9]).

4.4 Navigation Interactions

The larger the virtual world and the less knowledge of its topology the user has, the greater the need for well-designed navigation interaction. To go into available techniques, the complex task shall be split into two basic concepts: *Way finding* (or path finding) to find a path from origin to destination and *traveling* to change position. It should be noted, however, that the term navigation is often used to describe only what is defined here as "traveling": Interactions allowing changes in orientation and position of an object such as the user's avatar (e.g. in [1]).
The task of navigation in VR can be accomplished in one of several basic ways: A user may indicate their intended destination (i.e. selecting it), expecting the VR to guide them to it. This requires path finding algorithms as well as a means for guiding the user, such as visual cues or written/spoken instructions, and traveling interactions to actually change position. Alternatively, users may attempt to find the path themselves, memorizing it (forming a mental map of sorts) and traveling accordingly. For this, they need detailed information of the virtual world, either through prior knowledge or sources of information such as a WIM representation or a map. This is also known as way finding (see below). Lastly, the user may forgo traveling altogether and only select the desired destination, resulting in instantaneous relocation. This is arguably the easiest and least natural way, reducing way finding to destination selection.

Way finding To get to a destination, we first need to know the path to it. As many ways as we have to find a path in reality (route planners, asking other people, maps, signs), way finding in VR can be quite different. First of all, virtual worlds do not necessarily much resemble reality. For example, they may not have other people we could ask, or streets we could follow. Secondly, they may allow for traveling modalities impossible or improbably difficult in reality (see below), such as flying, extreme velocity changes or instantaneous relocation from one point to the next (a.k.a. teleportation).
VR simulations designed to represent reality as naturally as possible often resemble it far enough that traditional way finding techniques can be used with little or no effort for adaption. For example: In [16], Ropinski et al. propose a technique for navigating virtual representations of cities using road maps.
Not so natural and realistic virtual worlds may offer more powerful navigation concepts. The downside is that they will also require that the users know how to use these concepts. Through careful design (for which game design may provide

inspiration), this can still be intuitive and be learned efficiently, but it may also necessitate the users' explicit instruction (e.g. through tutorials, manuals).

Traveling Once a path to a destination has been identified, most often movement is required - in every case aside from instantaneous travel. Movement interaction in a three-dimensional environment is, in some ways, not unlike pointing and may (for every instant/frame) be defined using the following information:

- an object to be moved, with an origin and orientation
- a directional vector pointing in the direction in which travel is desired

Velocity may be predefined or based on user input. The object to be moved is normally the camera object which represents the user's location and view frustum in the virtual world. Often, an avatar representation is attached to it. As such, the origin and orientation is known implicitly; the user needs to indicate a direction. The following approaches are among the most common.
Walking as the standard mode for human travel in reality is basis for the most natural traveling interactions concepts. The user's movements can be tracked using the aforementioned tracking techniques. It does impose some drastic limitations, however: The user changes position in reality, too, which requires a) the VR devices to cover a free area large enough to accommodate such movement as the user will exhibit or b) the VR devices to move with them. Given the finite space in reality available in most use cases, there are two basic ways to allow the user to move freely: Having the user walk without moving at all, or redirecting the user with VR output, making them think they are walking freely but actually causing them to only walk within the available space.
The simplest option is the user purposely walking on the spot. This feels less natural, but can deliver enough input to be interpreted as walking in a direction and with a certain speed. Other methods include the use of treadmills. Traditional treadmills allow for movement only in one direction, limiting the options where indication of direction is concerned. Omnidirectional treadmills solve that problem by allowing the user to walk in any direction, just like they are used to. In addition to *permitting* omnidirectional movement, they may also be used for the *input* of that direction into the VR.
Redirected walking has been the focus of a number of researchers in the recent years, who have demonstrated that VR-based stimuli can cause the user's real movements to be redirected in an unintended direction without their notice (cf. [10]).
It is also not uncommon to use a traditional combination of keyboard and mouse for traveling interactions of the user's own avatar. At perhaps its most basic, the mouse is used for yaw and pitch axes, with one to four additional key interaction for movement forward as well as left, right, up and down (relative to orientation). A flight stick is an alternative, adding a roll axis. These combination are commonly found in 3D games such as first-person shooters and flight simulators and therefore easily accessible for many people. Nevertheless, they are far from intuitive and require the user to use one or both of their hands. Other concepts

track the user's head, torso or legs to allow for defining an orientation and other methods (such as buttons) for movement, freeing at least one of the user's hands as a mouse or flight stick are not required.
The simulation of physics is another factor in designing traveling interaction. In reality, all aspects of movement are defined by many factors, such as the mass of the moving object, gravity and friction. The more realistic a simulation is meant to be, the more it will need to take physics into account. This may also mean that the user's intention of moving a certain way (such as into the air or through a wall) may have to be overruled.

4.5 System Interactions

VR systems generally need to permit some interactions which have less to do with the virtual world being represented and more with the VR simulation as a whole. Examples for this could be: turning simulations on or off, selecting a virtual world, saving progress, switching modes or changing settings. These interactions are inherently unnatural. This makes their intuitive use difficult, but not impossible: An example could be the interpretation of voluntary (and even involuntary) indications of a user's distress as interaction, such as shutting one's eyes, screaming, a high heart rate or spasms, resulting in the simulation being turned off immediately.

5 Future trends

It seems likely that the market for VR application and therefore the push for technological advances will grow quickly in the near future, a trend which is fueled by companies like Google and Samsung. They recently introduced a two varieties of VR devices to the market: a lower-budget technique using smartphones with their screens and sensors as the center of portable VR goggles (Google Cardboard, Samsung Gear VR) and high-budget stationary VR goggles with higher resolutions and different features. Game developers, especially small independent studios, are already intensifying efforts to come up with ways to use these devices and market the results. In light of these trends, research into VR technologies will likely grow in recognition and popularity.
It may be worth asking why these new products have sparked so much interest. This seems to be due to two separate factors: The inclusion of game developers and the ubiquity of high-quality graphics capable smartphones.

5.1 Smartphones (Or: did you know you already own a VR device?)

The concept of using smartphones as the basis for VR devices has recently crossed the boundary from research and development to large-scale consumer market offerings. The two main offers here are the Oculus Gear VR and Google Cardboard, though a number of less well known firms have produced similar products (often even simpler and cheaper). Both are little more than straps and holders allowing the user to mount their smartphone right in front of their eyes, while dividing the phone's screen into two sections, one for each eye. This allows for a simple stereoscopic 3D simulation. While they may indeed serve well to demonstrate the basic idea of VR, they are far from state-of-the-art VR systems. Aside from lacking the high resolution and wide view frustums of other VR systems, they also do not offer many ways of delivering input into the simulation. Only some of them offer anything other than tracking of head orientation, such as a hole in the phone holder to allow the user to press the phone's buttons. Nevertheless, their mobility offers new possibilities as well, and the increased popularity and presence of VR in the public consciousness may give rise to new ideas and interest in research.

5.2 Inclusion of game developers

Computer games represent the most widespread use of three-dimensional virtual environments. Game developers and programmers in that industry have a great deal of experience creating convincing and immersive 3D environments. This experience is of great value in the development of VR applications, as VR-Worlds have significantly more in common with 3D Games than with traditional 2D desktop systems and metaphors such as WIMP. It would appear that the firms developing market-ready VR solutions have noticed this: In recent years, they have started cooperating closely with game developers and game engine makers.
Another result of this is that a number of games on the market already support VR devices before those devices themselves are being offered to consumers. This means that there is a large number of consumers willing to buy the VR devices right from the start, when their prices (and related profits for the producers) are still very high.
Three-dimensional games are still a trove of interaction concepts. Research over the next few years will show which of them are applicable in VR, as well as the resulting benefits and drawbacks.

6 Conclusion

The design of interaction concepts and corresponding techniques and devices is a key component in the creation of VR systems. I would like to propose the development of a framework for testing such VR interactions. The framework will be based on a simple user story. This could, for example, be a VR consisting of a virtual room containing a number of objects, implementations of interaction concepts for these objects, as well as tasks for the testers. Evaluation of each combination of interaction concepts shall be based upon:

- observation and measurement of the testers' performance in combination with appropriate metrics

– evaluation of the user experience by means of questionnaires

The primary goal of this framework will be the evaluation of the usability of VR interaction concepts.

References

1. Anthes, C., Heinzlreiter, P., Kurka, G., Volkert, J.: Navigation models for a flexible, multi-mode vr navigation framework. In: Proceedings of the 2004 ACM SIGGRAPH International Conference on Virtual Reality Continuum and Its Applications in Industry. pp. 476–479. VRCAI '04, ACM, New York, NY, USA (2004), http://doi.acm.org/10.1145/1044588.1044693
2. Damann, N., Voets, T., Nilius, B.: Trps in our senses. Current Biology 18(18), R880 – R889 (2008), http://www.sciencedirect.com/science/article/pii/S096098220800972X
3. Davis, S., Nesbitt, K., Nalivaiko, E.: A systematic review of cybersickness. In: Proceedings of the 2014 Conference on Interactive Entertainment. pp. 8:1–8:9. IE2014, ACM, New York, NY, USA (2014), http://doi.acm.org/10.1145/2677758.2677780
4. Dörner, R., Broll, W., Grimm, P., Jung, B.: Virtual und Augmented Reality (VR / AR). Springer Vieweg (2013)
5. Gao, Z., Lécuyer, A.: A vr simulator for training and prototyping of telemanipulation of nanotubes. In: Proceedings of the 2008 ACM Symposium on Virtual Reality Software and Technology. pp. 101–104. VRST '08, ACM, New York, NY, USA (2008), http://doi.acm.org/10.1145/1450579.1450601
6. Haffegee, A., Alexandrov, V., Barrow, R.: Eye tracking and gaze vector calculation within immersive virtual environments. In: Proceedings of the 2007 ACM Symposium on Virtual Reality Software and Technology. pp. 225–226. VRST '07, ACM, New York, NY, USA (2007), http://doi.acm.org/10.1145/1315184.1315232
7. Herrera, N.S., McMahan, R.P.: Development of a simple and low-cost olfactory display for immersive media experiences. In: Proceedings of the 2Nd ACM International Workshop on Immersive Media Experiences. pp. 1–6. ImmersiveMe '14, ACM, New York, NY, USA (2014), http://doi.acm.org/10.1145/2660579.2660584
8. Hicks, R.D.: Aristotle De Anima. Cambridge: At The University Press (1907), https://archive.org/details/aristotledeanima005947mbp
9. Hinckley, K., Tullio, J., Pausch, R., Proffitt, D., Kassell, N.: Usability analysis of 3d rotation techniques. In: Proceedings of the 10th Annual ACM Symposium on User Interface Software and Technology. pp. 1–10. UIST '97, ACM, New York, NY, USA (1997), http://doi.acm.org/10.1145/263407.263408
10. Hodgson, E., Bachmann, E.: Comparing four approaches to generalized redirected walking: Simulation and live user data. IEEE Transactions on Visualization and Computer Graphics 19(4), 634–643 (April 2013)
11. Ishibashi, K., Da Luz, T., Eynard, R., Kita, N., Jiang, N., Segi, H., Terada, K., Fujita, K., Miyata, K.: Spider hero: A vr application using pulling force feedback system. In: Proceedings of the 8th International Conference on Virtual Reality Continuum and Its Applications in Industry. pp. 197–202. VRCAI '09, ACM, New York, NY, USA (2009), http://doi.acm.org/10.1145/1670252.1670294
12. Jenkins, O.C., González, G., Loper, M.M.: Tracking human motion and actions for interactive robots. In: Proceedings of the ACM/IEEE International Conference on Human-robot Interaction. pp. 365–372. HRI '07, ACM, New York, NY, USA (2007), http://doi.acm.org/10.1145/1228716.1228765

13. Keyrouz, F.: Humanoid hearing: A novel three-dimensional approach. In: Robotic and Sensors Environments (ROSE), 2011 IEEE International Symposium on. pp. 214–219 (Sept 2011)
14. Kurita, Y., Ikeda, A., Tamaki, T., Ogasawara, T., Nagata, K.: Haptic augmented reality interface using the real force response of an object. In: Proceedings of the 16th ACM Symposium on Virtual Reality Software and Technology. pp. 83–86. VRST '09, ACM, New York, NY, USA (2009), http://doi.acm.org/10.1145/1643928.1643948
15. Qi, W.: Geometry based haptic interaction with scientific data. In: Proceedings of the 2006 ACM International Conference on Virtual Reality Continuum and Its Applications. pp. 401–404. VRCIA '06, ACM, New York, NY, USA (2006), http://doi.acm.org/10.1145/1128923.1128999
16. Ropinski, T., Steinicke, F., Hinrichs, K.: A constrained road-based vr navigation technique for travelling in 3d city models. In: Proceedings of the 2005 International Conference on Augmented Tele-existence. pp. 228–235. ICAT '05, ACM, New York, NY, USA (2005), http://doi.acm.org/10.1145/1152399.1152441
17. Rosa, N., Hürst, W., Vos, W., Werkhoven, P.: The influence of visual cues on passive tactile sensations in a multimodal immersive virtual environment. In: Proceedings of the 2015 ACM on International Conference on Multimodal Interaction. pp. 327–334. ICMI '15, ACM, New York, NY, USA (2015), http://doi.acm.org/10.1145/2818346.2820744
18. Sementille, A.C., Lourenço, L.E., Brega, J.R.F., Rodello, I.: A motion capture system using passive markers. In: Proceedings of the 2004 ACM SIGGRAPH International Conference on Virtual Reality Continuum and Its Applications in Industry. pp. 440–447. VRCAI '04, ACM, New York, NY, USA (2004), http://doi.acm.org/10.1145/1044588.1044684
19. Slater, M., Wilbur, S.: A framework for immersive virtual environments (five): Speculations on the role of presence in virtual environments. Presence Teleoperators and Virtual Environments 6(6), 603–616 (1997)
20. Stoller, P.: The Taste of Ethnographic Things - The Senses in Anthropology. University of Pennsylvania Press (1989)
21. Wu, A., Shah, M., Lobo, N.D.V.: A virtual 3d blackboard: 3d finger tracking using a single camera. In: Automatic Face and Gesture Recognition, 2000. Proceedings. Fourth IEEE International Conference on. pp. 536–543 (2000)