Grundprojekt

M.Sc. Computer Science

Gerald Melles
Hochschule für Angewandte Wissenschaften Hamburg
Hamburg, Germany
contact@geraldmelles.com

ABSTRACT

This paper has two goals: Firstly, to propose a taxonomic basis for future works in the field of Virtual Reality technologies. Secondly, to provide insights into the design and implementation of three core technical aspects of immersive applications within the virtuality continuum: tracking, networking/streaming and sensory feedback channels. It is intended to aid researchers in making informed choices concerning these technical subsystems for their experiments.

KEYWORDS

VR, virtual reality, mixed reality, tracking, streaming, feedback

ACM Reference Format:
Gerald Melles. 2017. Grundprojekt: M.Sc. Computer Science. In *Proceedings of ACM Conference (Conference'17).* ACM, New York, NY, USA, 7 pages. https://doi.org/10.1145/nnnnnnn.nnnnnnn

1 INTRODUCTION

1.1 Related and foundational works

This paper discusses only a portion of the technologies relevant to AR and VR, and (almost) only from a technical point of view. For a broader perspective and more exhaustive discussions, therefore, the following works are recommended:
Alan B. Craig's "Understanding Augmented Reality" ([8]) is excellently suited as an introduction to the field of Augmented Reality. It maintains a very practical viewpoint and presents real-world examples for almost everything mentioned.
"The Engineering of Mixed Reality Systems" ([11]) provides a lot of information on the soft- and hardware engineering involved in the creation of Mixed Reality System.
"Human Factors in Augmented Reality Environments" ([18]) attempts to shed light on the H in HCI. It also suggests (and references) a large number of ways to evaluate aspects of Human-Computer-Interaction.

1.2 Taxonomy of the Virtuality Continuum

We will use the terms "Virtual Reality" or "Virtuality Continuum" to refer to the whole of all concepts and technologies in this field, all other relevant terminology is explained in this introduction or in the glossary at the end of this document.
Despite its fairly recent spike in popular recognition, virtual reality research has been conducted for quite a long time (at least by the standard of the relatively young field of Computer Science). As such, there exist taxonomic proposals dating back more than two decades.
Perhaps the most commonly cited taxonomy is that of Milgram and Kishino ([28], see also: "Ontological perspective', below).
Milgram and Kishino specifically focus on displays and visual technologies (note that the title of the paper is "A Taxonomy of Mixed Reality Visual Displays"). I do, however, agree with the assessment Thomas Erickson made yet another year previous (at the time he was working for Apple Inc.'s Advanced Technology Group): "[...] note that the word 'visualization" is really to narrow. "Perceptualization" is probably more apropos, although it doesn't roll readily off the tongue. Sound and touch, as well as visual appearance, may be profitably used to represent data." ([37], p. 8-9). Technologies like these, for the creation of auditory and haptic stimuli as part of virtual reality simulations, exist - in fact some will be discussed in this paper. Milgram and Kishino's taxonomy could be understood to include them. But if the concept of virtual reality is concerned with offering its users an *experience* - be it through augmentation or replacement of their natural perceptualization - then I would argue that its taxonomy should reflect that.
A taxonomy is always based on a perspective. I propose that there should be two main perspectives from which to propose taxonomies of virtual reality systems: *Phenomenologically* and *ontologically.*
I will only briefly elaborate on what these perspectives encompass, a more exhaustive discussion will be part of future works. Note, however, that while these perspectives are derived from the approaches in their synonymous fields of metaphysics, the fields in which they are applied here remain those of Human-Computer-Interaction and Computer Science. They simply serve as taxonomic bases for orthogonal viewpoints on Virtual Reality.
Additionally, a Virtual Reality system may be viewed from the technical perspective of its concrete implementation. This may perhaps be less useful for distinguishing between different types of VR on a theoretical level and providing the basis for interdisciplinary discourse. Nonetheless, this perspective requires its own terminology and is perhaps the most natural one to the engineers tasked with the system's creation.

1.3 Phenomenological perspective

"Phenomenology is the study of structures of consciousness as experienced from the first-person point of view. The central structure

---

Conference'17, July 2017, Washington, DC, USA

ACM ISBN 978-x-xxxx-xxxx-x/YY/MM.
https://doi.org/10.1145/nnnnnnn.nnnnnnn

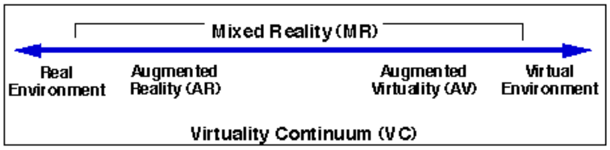

Figure 1: Simplified illustration of Milgram and Kishino’s Virtuality Continuum [28]

of an experience is its intentionality, its being directed toward something, as it is an experience of or about some object. An experience is directed toward an object by virtue of its content or meaning (which represents the object) together with appropriate enabling conditions." [13]

From a phenomenological perspective, therefore, virtual reality technologies should evaluated in terms of their *user experience (UX)*. This perspective especially includes classifications of its *immersiveness*. For example, a hand-held Augmented Reality application on a smartphone seems likely to be less immersive than one utilizing a Head-Mounted Display.

This perspective is especially important for interdisciplinary research and discourse. For example, highly immersive virtual environments have been recognized as valuable research tools in both behaviorology and psychology for decades (cf. [2], [25]).

1.4 Ontological perspective

The Ontological perspective on the other hand concerns itself with what is *real* or *unreal (virtual)*. It allows for a distinction between Virtual Reality applications based on the *existence* versus the *simulation* of an entity perceived by the user. In other words: Any entity is considered virtual if it is only perceptible through the use of the Virtual Reality system.

The above-mentioned taxonomy proposed by Milgram and Kishino provides a decent basis for describing Virtual Reality applications from the Ontological perspective. It highlights the sliding scale between none and all of the user’s perception stemming from interaction with a virtual environment. Their taxonomy has been used in many forms over the years with various proposed changes and extensions.[1]

Since no technical implementation has - as of yet - been able to fool all of the user’s senses, the point could be made that this means all so-called ”virtual reality” systems would better be described as”mixed reality” systems. This may seem a purely academic distinction with no bearing on the market, but Microsoft Inc. would appear to disagree: Their newest systems for immersive (near?) virtual reality are called ”Windows Mixed Reality”.[2]

1.5 Technical aspects of immersive Virtuality

Immersive virtual reality setups are complex systems comprised of many of different hardware and software components. Often, rather than being concentrated on a single device, they are distributed across several devices, such as head-mounted displays, computers, trackers, input devices and cameras. These devices need to be networked and often have tight constraints regarding latency and QoS. In this paper, three types of subsystems - together encompassing the majority of AR/VR devices - are discussed with regards to their requirements and possible implementations: **Tracking**, **Networking/Streaming** and **Feedback Channels**.

The goal of this paper is to provide AR/VR researchers with enough insight into these subsystems to make informed architectural decisions when planning experiments concerning interactive AR/VR. Additionally, it serves as a basis for my upcoming Master’s Thesis, particularly where taxonomic disambiguation is concerned.

To achieve this goal, introductions to individual tracking, networking/streaming and feedback techniques will be given, followed by brief suggestions for their implementation. This may be in the form of references to publications where such implementations are discussed, or suggestions based on my own experiments. In the case of multiple suggestions, their most significant benefits and drawbacks are also briefly described.

2 TRACKING

For the purposes of this paper, we divide the multitude of existing and proposed tracking systems into categories, from the perspective of the tracked object.

2.1 Outside-In Tracking

Outside-In tracking systems are based on sensors which are capable of pin-pointing an object’s location. Most often, these systems use cameras and computer vision algorithms to simultaneously locate an object on a number of video feeds from different perspectives to extrapolate its position (and possibly rotation). These camera-based systems will be the focus of this chapter.

Some of these systems track only specific markers.[3] Others are capable of identifying objects even without them being specially marked.

While knowledge of the position and orientation of the sensors these systems rely on is generally essential for extrapolation, they are not necessarily static. Some systems such as the one proposed by Cortes et. al. in [7] are capable of moving (in their case: rotating) the sensors in order to increase the covered area. Another strategy involved fusioning sensor data from different systems to combine their coverage - an approach which is currently being researched in the CSTI.

Passive markers. Passive markers are affixed to an object whose position and/or orientation are to be tracked. They are either necessary to make an object trackable with a specific system or improve the chances of an object being tracked correctly. They do not contain any logic or active components and often work by reflecting specific radiation wavelengths, esp. infrared light. One proposed implementation can be found in [14].

The (proprietary) Advanced Realtime Tracking (ART) (cf. [1]) and OptiTrack (cf. [30]) systems for example predominantly use passive markers. These are typically based on cat’s eye type reflectors, which are designed to reflect light in the same direction as it arrived

---

[1] [23], for example, amusingly proposed its extension by their "climbing reality continuum" between real and virtual rock-climbing.

[2] However, note that in [9], Microsoft’s Director of Communications for Mixed Reality Greg Sullivan claimed that in their view, the point of view of the *device* was more important in making that distinction than that of the user. Clearly there seems to be a need for a taxonomy whose definition (at least) is universally agreed upon.

[3] Though a distinction is made in this paper, the term "tracker" is often used by others synonymously with "marker".

in. The combination of these markers and LED-Rings around its cameras allows the ART system to reliably track the position of objects with a minimum of unintended refraction.

Active markers. Active trackers work largely like passive trackers, but contain some form of active component. An example for this is the active tracking marker concept for the ART system, which works much like the aforementioned IR reflectors but uses infrared LEDs on the objects rather than around the cameras. This way, no reflectors are needed, drastically increasing the marker's luminosity and simultaneously reducing its size. Active marker systems are more complex, requiring electronic components and their own power supply.

Markerless. Markerless camera-based tracking systems are capable of identifying features in their field of vision without those having to be specially marked. This tends to require more complex identification techniques: While for infrared marker based systems such as the ART system it could be sufficient to find out which pixels in a still video frame are brightest, the identification of an object's outline without markers requires advanced computer vision algorithms. That said, many computer vision algorithms have been implemented on an open source basis in libraries such as OpenCV (Open Computer Vision).
An example of markerless camera-based tracking systems can be found in the LeapMotion system: It uses a stereoscopic setup of two infrared cameras in combination with a custom software driver to accurately track human hands in an area in front of it. (cf. [34])

2.2 Inside-Out Tracking

Inside-out tracking systems also contain active components, but rather than using actors providing an external system with additional information to become trackable, it tracks its own position using sensors. More precisely: It uses sensors to gather information about its surroundings, analyzing these observations to estimate its own rotation and/or position. Often, these sensors are cameras: The camera feeds can then be used to identify static features (such as furniture) in the user's surroundings. This in turn allows for an estimation of its own position and rotation relative to them.
One example for this is the HTC Vive Tracker which identifies its own position in three-dimensional space by receiving infrared pulses from "Lighthouse" base stations. A key benefit of this design is scalability: So long as they have an unimpeded line of sight to the base stations, the number of objects in a system may be unlimited. This may, however, be undercut by doing the necessary calculations on the same server the simulation itself runs on, rather than on the microcontrollers of the objects.
To distinguish between those variants which require base stations (such as the HTC Vive) and those who do not, HTC has begun calling its new inside-out headset - the Vive Focus - a "world-scale" (as opposed to "room-scale") VR technology (cf. [12]).

3 NETWORKING AND STREAMING

Augmented and virtual reality systems rely heavily on a variety of hardware components. Examples for this include the aforementioned tracking systems, gyroscopic and acceleration sensors and cameras. Networking these components can be a challenge, especially due to two factors:

Platform variety. Hardware components may not all utilize the same operating system, limiting the interoperability of their software.

Time constraints. The data may need to be transmitted in (near) real-time[4]. AR and VR systems are especially sensitive to latencies causing a temporal mismatch between the user's senses. For example, turning one's head while wearing an HMD is likely to cause extreme disorientation and motion sickness if the turning motion is is not translated accurately and in real-time.

In this section, the most important technical challenges will be mentioned. A more in-depth analysis of problems and possible solutions is planned to follow in another paper.

Most sensors used in AR/VR systems create and transmit data at a rate low enough not to present any particular problems in terms of bandwidth and processing throughput. Gyroscopic and acceleration sensors for example mostly return little more than a few bytes of data representing angles, speeds and forces every time they are polled by a microcontroller. While that still leaves potential problems related to the facilitation of reliable real-time networking and QoS constraints, the greatest challenges of these sensors tend to be "cleaning" and processing the values to make them usable. For example, this can include heuristics to remove unlikely values and outliers and algorithms for detecting trends and generating values based upon them when the sensor itself provides false values or none at all.
But not all sensors transmit small portions of data. Cameras especially can create vast amounts of data, and should the AR/VR system rely on this data for simulation, it may be subject to the same near real-time constraints. Furthermore, media streaming in general is a large part of many interactive applications, such as games, telepresence and playback software. Research has shown that users are more sensitive to latency in AR and VR systems than on traditional desktop systems.
In the following sections, I will discuss the technologies involved and propose the creation of a streaming framework designed explicitly for use in AR/VR applications.

Mixed Reality. Even more than most VR systems, mixed reality (MR) environments often consist of many networked and integrated subsystems. For them, guaranteeing the required QoS can be even more challenging. It may even be sensible to consider the implementation of a dedicated messaging system designed specially for this purpose.

Technologies and Protocols. When choosing the technological basis for the electronic exchange of data, the primary tradeoff is between the complexity of the network and the roundtrip speed. This is due to increases in latency based on routing procedures.
Since AR/VR systems have a need for low latencies, the tendency here is toward the use of direct point-to-point wiring and relatively

---

[4] Real-time in the scope of this work may be understood to be a time constraint which makes delays imperceptible to humans. Note that this ideally includes even subconscious perception.

simple architectures, such as USB. However, integrations of devices in a (mixed reality) environment may benefit from the potentials offered by more complex network topologies. Lastly, if the VR system is not to exist in isolation but is to be loosely coupled with environmental and web services, there is no way around the certain complexity imposed by networking technologies such as the Internet Protocol Stack.

Another concern may be portability[5]. Some personal systems (see "Feedback Channels" below) such as the HTC Vive HMD require wired connections with the computer the simulation runs on. Wireless systems allow for a greater freedom of movement for the user and may avoid unintended haptic sensations caused by the user dragging or touching cables (or even falling over them).

4 FEEDBACK CHANNELS

Thus far we have primarily discussed ways in which a VR system can receive and process a variety of *inputs*. This section will elaborate on 'Feedback Channels' as part of the system's *output*.

Feedback here refers to ways in which the system responds to user interaction in ways the user may perceive. The most obvious way may be the HMD's visual display itself, as it tends to be the most prominent feature of VR systems. When a user interacts with the VR system, such as by moving their head, the feedback provided by the system (and expected by the user) is a matching translation of that movement into virtual space. But aside from this visual channel, there are also other senses which may be used to provide feedback to the user. In this paper, we will discuss the three most prominent of these sensory output channels: visual, auditory and mechanical/Haptic.

Aside from distinctions based on which of our senses they address, feedback systems may be classified by whether they are *personal* or *environmental*.

Personal feedback systems in the context of this paper shall be understood as being devices which are attached to the user's body. Most often that means wearable devices such as smartwatches, HMDs or gloves.

Environmental feedback systems are not commonly attached to the user. Examples include screens/displays, sound systems and actors in IoT devices (such as LEDs, speakers or motors).

Devices designed to be *wielded* by the user can exist in a gray area between these categories. As such, we shall define them to be personal whenever they are being wielded (or are attached to the user, for example by means of wrist-bands like the HTC Vive's controllers) and environmental whenever they are put down.

4.1 Visual

personal. Their personal displays is perhaps the place where the differences between VR and other devices in the Mixed Reality Continuum are most obvious. While VR HMDs focus on replacing *all* of the user's evironment's visual stimuli, AR devices most often allow some of them to still be visible. That includes the most well-known AR device - the (cancelled) Google Glass project - as well as Vufine's "Vufine+", currently the most easily available and probably most often bought AR device. Most of the devices in development, such as those by Vuzix or Lumus (currently partnered with HTC), also follow this principle. Nevertheless, it should be mentioned that designs which replace all natural visuals but use front-mounted cameras to replicate natural vision could still be considered AR HMDs. This highlights the difficulty of clearly distinguishing between VR and AR. Lumus' Website even proudly quotes HTC's COO David Chen, claiming to be helping HTC with a "natural extension into augmented reality"[26].

environmental. Environmental visuals could be interpreted as including literally everything we can see that is not attached to ourselves. Their significance is more obvious in AR systems, where their visibility to the user is the very definition of AR. But VR systems may also rely heavily on environmental visual systems, since not all VR systems are based on HMDs: CAVE-Systems, for example, project the visuals of a virtual reality onto surfaces around the user. movement, light, illusions, screens

4.2 Auditory

personal. Personal auditory systems are primarily ear- and headphones. These may be combined with a tracking system to allow for the modulation of sounds to reflect the user's stance, creating the illusion of three-dimensional audio.

environmental. Environmental audio in general may again include all sources of auditory stimuli which the user does not wield or wear. Of particular interest, however, are *surround sound* and *spatial sound* systems.

Common surround sound systems create the impression of a three-dimensional soundscape by playing back sounds through a variety of speakers positioned around the user. These systems are generally unable to maintain this illusion for more than one user located in a very specific place.

Spatial sound systems (Wave field synthesis systems in particular) endeavor to approximate the exact propagation of sound waves not just in one place but in a larger area (surrounded by a large number of individual speakers). This way, they maintain the illusion even for multiple and moving users. (cf. [15]) [6]

4.3 Mechanical/Haptic

Mechanical perception is, in many ways, more challenging to generate than visual or auditory feedback. This is mainly because of the fact that, while our visual and auditory senses are each located in only up to two places on our body, our mechanical senses instead cover and fill our entire body. Furthermore, mechanical (such as haptic) sensation is a complex amalgamation of neural impulses not just related to our skin, but also our muscles, joints and internal organs. A worn or wielded object's inertia and weight, for example, are felt in many parts of out bodies simultaneously. As such, human-computer-interfaces are often limited to *tactile* sensation. This means they only endeavor to replicate the mechanical aspect of touching a surface. In contrast, *haptic* interfaces attempt to also generate enough stimuli to allow the user to explore an object using

---

[5] here in the sense of not inhibiting the user's movement rather than hardware independence/abstraction

[6] Note that a method utilizing Wave-Field Synthesis for haptic (rather than auditory) feedback has recently been proposed by a research team within Facebook and Oculus Inc. in [27]

mechanical sensations.
While these two types represent the majority of mechanical feedback systems, there have been attempts at also replicating other mechanical sensations. An example for this can be found in [17], where users had flywheels attached to their heads to simulate inertia.

4.3.1 thermal. Thermal feedback is generally seen as a form of haptic and therefore mechanical perception (though some, such as Löchtefeld et. al. in [24], highlight contrasts between mechanical haptic and thermal perception).
Recently, quite a number of papers have been published proposing the use of thermal feedback modules. The predominant argument here seems to be that multi-sensory feedback increases immersion (presence) and memory of the experience (e.g.: thermal combined with visual feedback as proposed in [5] and [32]). It has also been theorized that thermal perception is closely linked with a feeling of personal danger - a theory which prompted Wilson, Maxwell and Just to experiment with a thermal feedback system to signal to users that they were browsing an unsafe (not SSL-secured) website ([38])
In terms of the technology involved, the generation of heat is not particularly challenging. In fact, much of the hardware design in electronic circuitry is dedicated to *avoiding* unnecessary loss of energy through heat. (Re)Moving heat in order to create the perception of coldness, on the other hand, is somewhat more involved. Thermoelectric devices using the Peltier effect represent the most cost and space effective means of doing both, while also not being as sensitive to movement as compression-based heat exchangers. This makes them ideal for creating thermal feedback wearables and hand-held interaction devices.
It should be noted that they are also quite energy-inefficient, however, as well as relatively slow to react. Our thermoreceptive senses are also slower to react than some of our other senses: "One challenge is that temperature changes are perceived over the course of seconds." [36] As such, thermoelectric systems themselves are not normally suitable for a feedback system on real-time constraints. This could be circumvented by preheating (or precooling) the systems and using mechanical systems to only move them close enough to the user to be felt when thermal feedback is required - which may of course also cause potentially undesirable tactile sensations.

personal. Personal systems for the creation of mechanical feedback tend to be based on one of the following principles: Vibration, mechanical pressure, mechanical inhibition or direct neural or muscular stimulation.
Of these, vibration is the most commonly used concept. This is likely due to its low complexity, space requirements and costs. Vibration as a means for creating tactile sensations is also used commonly in a number of other non-VR applicances, such as watches and smartphones. Another example can be found in [20], where a head-mounted ’vibrotactile grid’ (consisting of 20 small vibration motors as may be found in phones attached to what looks to be a swimmer’s cap) is used in an attempt to provide ’haptic guidance’ to users. Since vibration is distinctly *not* the same as the feeling of touching an immobile surface, the exact sensations that can be produced using vibration should be considered and the most appropriate ones chosen for each implementation. Strohmeier and Hornbæk offer an excellent starting point for this in [35].
Exerting specifically modulated mechanical pressure on a user is somewhat more difficult to achieve than vibration (which is basically an oscillation of mechanical pressure). Mechanical pressure systems tend to require exoskeletal structures and be much more bulky than vibration systems. There are a great many means by which mechanical pressure may be created, such as pneumatics, hydraulics, servo-motors and linear actuators. An example for a servo-driven exoskeletal haptic feedback system in the form of a glove is provided by Dexta Robotics’ ”Dexmo” in [16]. In [3], a pen-like instrument is outfitted with electromagnets to provide noticeable attractive and repulsive forces toward or away from specific points on a surface in order to guide its user.
A specialized form of mechanical pressure system uses forces generated by the user, such as their limbs’ muscle power and inertia, to generate pressure on the user by inhibiting their movement. Wearable pants-like systems may, for example, inhibit the movement of their legs to create the illusion of moving through difficult terrain, such as water or sand. An example for this can be found in [33]. Here, a compression feedback system (inflatable straps around the user’s extremities) is used to not just generate pressure on its own but to purposefully inhibit the user’s movements.
Finally, there remains the option of directly stimulating a user’s body, e.g. by use of EMS (electro-magnetic stimulation). This permits the use of the user’s own muscles as mechanical actuators. An example may be found in Kaul et. al.’s [19].

environmental. Environmental haptic feedback systems include all those systems designed to create haptic sensations as part of a virtual reality setup that are not wieldeed or worn. They are highly heterogenous in terms of the technology involved: There exist pneumatic, hydraulic and electrically actuated systems to mention some of the most common. [22], Knierim et. al. suggest using highly mobile flying drones of a Quadcopter design to provide tactile feedback in mid-air. Aerial haptic systems have been proposed, using air jets and vortices or ultrasound. In [31], Ochiai et. al. propose the combination of femtosecond laser light fields and acoustic (ultrasound) waves to allow for non-contact haptics (meaning the user would not physically have to touch the actor systems) of astounding precision.

4.4 Combined and multi-sensory stimuli

Some studies suggest that the combination of different sensory feedback channels can cause the user experience to differ significantly from what might be expected. Particularly sensory mismatches have been discovered to contribute strongly to the phenomenon of VR sickness (cf. [21]).
Another example is the combination if visual and vibrotactile feedback discussed in [4]: The visual perception of a rectangular object combined with a well-timed vibrotactile feedback made users believe they had run their finger over an edge when it was actually a round surface.

5 CONCLUSION

We have introduced a measure of taxonomic disambiguation to the field of Virtual Reality research by introducing the orthogonal

Phenomenological and Ontological perspectives. Then we elaborated on three of the most important technical aspects of VR and AR systems as part of a third, technical perspective. This was then used to discuss options for implementation, with examples given from current institutional research, industrial developments and our own experimentation.

Once the suggested implementations for tracking, networking / streaming and feedback techniques mentioned above are available as either detailed implementation manuals or hardware and software modules, that should significantly aid researchers in the implementation of experiments researching immersion, efficiency and - ultimately - suitability of VR/AR interaction designs. That is especially true in the case of researchers lacking either resources (such as time and personnel) or experience with the technical aspects AR/VR technologies, such as may be the case in interdisciplinary research.

6 FUTURE WORKS

Most VR research so far has been focused on individual subsystems. Insofar as comparisons exist, they are limited to those between similar subsystems.

Rather than follow that trend, we would like to evaluate promising combinations of subsystems more holistically. To this end, we are planning the development of a number of prototypes for most of the abovementioned categories.

This fits well into the CSTI's research focus and promises synergies with its other teams' research efforts. Multiple projects in the CSTI have, for example, already proposed designs for personal haptic feedback devices.

7 GLOSSARY

Installation. VR "Installation" is used as a holistic term to describe the combination of a technical VR system (see below) and its users. The term is intended to put the *purpose* of the system in the foreground.

Setup. A VR "Setup" describes the hard- and software components and subsystems of a system designed to create the illusion of augmented and/or virtual reality. As opposed to a VR installation (see above) this does not include the users. The term is supposed to highlight the concrete *implementation* of the system.

Sensor. Any entity capable of reacting to a physical input.

Actor. Any entity capable of providing physical output.

HMD. Head Mounted Display. Not limited to devices providing visual stimuli but encompassing all head-mounted devices for enabling virtual perceptualization.

QoS. Quality of Service. This refers to agreed upon traits of a rendered service, such as guaranteed delivery or maximum latency.

REFERENCES

[1] ART. 2017. Advanced Realtime Tracking. (2017). https://ar-tracking.com/

[2] Jeremy N. Bailenson, Jim Blascovich, Andrew C. Beall, and Jack M. Loomis. 2003. Interpersonal Distance in Immersive Virtual Environments. *Personality and Social Psychology Bulletin* 29, 7 (2003), 819–833. https://doi.org/10.1177/0146167203029007002 arXiv:https://doi.org/10.1177/0146167203029007002 PMID: 15018671.

[3] James Burnside, Ben Elgar, Sam Healer, Alexander Hill, Zac Ioannidis, Luke Mitchell, Paul Worgan, and Anne Roudaut. 2016. Force Attraction Pen: A Haptic Pen with Variable Attraction Force. In *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems (CHI EA '16).* ACM, New York, NY, USA, 2655–2660. https://doi.org/10.1145/2851581.2892441

[4] Juan Pablo Carrascal and Roel Vertegaal. 2017. Effects of Tactile Feedback on the Perception of Virtual Shapes on Non-Planar DisplayObjects. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '17).* ACM, New York, NY, USA, 4417–4423. https://doi.org/10.1145/3025453.3025488

[5] Zikun Chen, Roshan Lalintha Peiris, and Kouta Minamizawa. 2017. A Thermally Enhanced Weather Checking System in VR. In *Adjunct Publication of the 30th Annual ACM Symposium on User Interface Software and Technology (UIST '17).* ACM, New York, NY, USA, 123–125. https://doi.org/10.1145/3131785.3131825

[6] M. K. D. Coomans and H. J. P. Timmermans. 1997. Towards a taxonomy of virtual reality user interfaces. In *Proceedings. 1997 IEEE Conference on Information Visualization (Cat. No.97TB100165).* 279–284. https://doi.org/10.1109/IV.1997.626531

[7] G. Cortes, E. Marchand, J. Ardouinz, and A. Lécuyer. 2017. Increasing optical tracking workspace of VR applications using controlled cameras. In *2017 IEEE Symposium on 3D User Interfaces (3DUI).* 22–25. https://doi.org/10.1109/3DUI.2017.7893313

[8] Alan B. Craig. 2013. *Understanding Augmented Reality: Concepts And Applications.* Morgan Kaufmann, Waltham, MA, USA.

[9] Heise c't. 2017. Warum Microsoft seine VR-Brillen "Mixed Reality" nennt. (2017). https://www.heise.de/ct/artikel/Warum-Microsoft-seine-VR-Brillen-Mixed-Reality-nennt-3820657.html

[10] Alex Davies. 2010. Adrift in the Virtuality Continuum. *Comput. Entertain.* 8, 1, Article 2 (Nov. 2010), 14 pages. https://doi.org/10.1145/1857940.1857942

[11] Emmanuel Dubois and Philip Gray (Eds.). 2010. *The Engineering of Mixed Reality Systems.* Springer, London, UK.

[12] Engadget. 2017. HTC Vive Focus is a standalone VR headset with 'world-scale' tracking. (2017). https://www.engadget.com/2017/11/13/htc-vive-focus-standalone-vr-headset-daydream/

[13] Center for the Study of Language and Stanford University Information (CSLI). 2017. The Stanford Encyclopedia of Phylosophy: Phenomenology. (2017). https://plato.stanford.edu/entries/phenomenology/

[14] M. Foursa. 2004. Real-time Infrared Tracking System for Virtual Environments. In *Proceedings of the 2004 ACM SIGGRAPH International Conference on Virtual Reality Continuum and Its Applications in Industry (VRCAI '04).* ACM, New York, NY, USA, 427–430. https://doi.org/10.1145/1044588.1044681

[15] Francesco Grani, Dan Overholt, Cumhur Erkut, Steven Gelineck, Georgios Triantafyllidis, Rolf Nordahl, and Stefania Serafin. 2015. Spatial Sound and Multimodal Interaction in Immersive Environments. In *Proceedings of the Audio Mostly 2015 on Interaction With Sound (AM '15).* ACM, New York, NY, USA, Article 17, 5 pages. https://doi.org/10.1145/2814895.2814919

[16] Xiaochi Gu, Yifei Zhang, Weize Sun, Yuanzhe Bian, Dao Zhou, and Per Ola Kristensson. 2016. Dexmo: An Inexpensive and Lightweight Mechanical Exoskeleton for Motion Capture and Force Feedback in VR. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (CHI '16).* ACM, New York, NY, USA, 1991–1995. https://doi.org/10.1145/2858036.2858487

[17] Jan Gugenheimer, Dennis Wolf, Eythor R. Eiriksson, Pattie Maes, and Enrico Rukzio. 2016. GyroVR: Simulating Inertia in Virtual Reality Using Head Worn Flywheels. In *Proceedings of the 29th Annual Symposium on User Interface Software and Technology (UIST '16).* ACM, New York, NY, USA, 227–232. https://doi.org/10.1145/2984511.2984535

[18] Weidong Huang, Leila Alem, and Mark A. Livingston (Eds.). 2013. *Human Factors in Augmented Reality Environments.* Springer Science+Business, New York, USA.

[19] Oliver Beren Kaul, Max Pfeiffer, and Michael Rohs. 2016. Follow the Force: Steering the Index Finger Towards Targets Using EMS. In *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems (CHI EA '16).* ACM, New York, NY, USA, 2526–2532. https://doi.org/10.1145/2851581.2892352

[20] Oliver Beren Kaul and Michael Rohs. 2016. HapticHead: 3D Guidance and Target Acquisition Through a Vibrotactile Grid. In *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems (CHI EA '16).* ACM, New York, NY, USA, 2533–2539. https://doi.org/10.1145/2851581.2892355

[21] Michiteru Kitazaki, Tomoaki Nakano, Naoyuki Matsuzaki, and Hiroaki Shigemasu. 2006. Control of Eye-movement to Decrease VE-sickness. In *Proceedings of the ACM Symposium on Virtual Reality Software and Technology (VRST '06).* ACM, New York, NY, USA, 350–355. https://doi.org/10.1145/1180495.1180567

[22] Pascal Knierim, Thomas Kosch, Valentin Schwind, Markus Funk, Francisco Kiss, Stefan Schneegass, and Niels Henze. 2017. Tactile Drones - Providing Immersive Tactile Feedback in Virtual Reality Through Quadcopters. In *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems (CHI EA '17).* ACM, New York, NY, USA, 433–436. https://doi.org/10.1145/3027063.3050426

[23] Felix Kosmalla, André Zenner, Marco Speicher, Florian Daiber, Nico Herbig, and Antonio Krüger. 2017. Exploring Rock Climbing in Mixed Reality Environments. In *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems (CHI EA '17)*. ACM, New York, NY, USA, 1787–1793. https://doi.org/10.1145/3027063.3053110

[24] Markus Löchtefeld, Tuomas Lappalainen, Jani Väyrynen, Ashley Colley, and Jonna Häkkilä. 2017. Comparing Thermal and Haptic Feedback Mechanisms for Game Controllers. In *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems (CHI EA '17)*. ACM, New York, NY, USA, 1829–1836. https://doi.org/10.1145/3027063.3053172

[25] Jack M. Loomis, James J. Blascovich, and Andrew C. Beall. 1999. Immersive virtual environment technology as a basic research tool in psychology. *Behavior Research Methods, Instruments, & Computers* 31, 4 (01 Dec 1999), 557–564. https://doi.org/10.3758/BF03200735

[26] Lumus. 2017. Lumus Partners. (2017). lumusvision.com/partners/

[27] Ravish Mehra, Christoph Hohnerlein, David Perek, Elia Gatti, Riccardo DeSalvo, and Sean Keller. 2016. HapticWave: Directional Surface Vibrations Using Wave-field Synthesis. In *ACM SIGGRAPH 2016 Emerging Technologies (SIGGRAPH '16)*. ACM, New York, NY, USA, Article 11, 2 pages. https://doi.org/10.1145/2929464.2929469

[28] Paul Milgram and Fumio Kishino. 1994. A Taxonomy of Mixed Reality Visual Displays. vol. E77-D, no. 12 (12 1994), 1321–1329.

[29] Mohammad Mehdi Moniri, Fabio Andres Espinosa Valcarcel, Dieter Merkel, Winfried Schuffert, and Tim Schwartz. 2016. Hybrid Team Interaction in the Mixed Reality Continuum. In *Proceedings of the 22Nd ACM Conference on Virtual Reality Software and Technology (VRST '16)*. ACM, New York, NY, USA, 335–336. https://doi.org/10.1145/2993369.2996318

[30] Inc NaturalPoint. 2017. OptiTrack. (2017). http://www.optitrack.com

[31] Yoichi Ochiai, Kota Kumagai, Takayuki Hoshi, Satoshi Hasegawa, and Yoshio Hayasaki. 2016. Cross-Field Aerial Haptics: Rendering Haptic Feedback in Air with Light and Acoustic Fields. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (CHI '16)*. ACM, New York, NY, USA, 3238–3247. https://doi.org/10.1145/2858036.2858489

[32] Roshan Lalintha Peiris, Wei Peng, Zikun Chen, Liwei Chan, and Kouta Minamizawa. 2017. ThermoVR: Exploring Integrated Thermal Haptic Feedback with Head Mounted Displays. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '17)*. ACM, New York, NY, USA, 5452–5456. https://doi.org/10.1145/3025453.3025824

[33] Henning Pohl, Franziska Hoheisel, and Michael Rohs. 2017. Inhibiting Freedom of Movement with Compression Feedback. In *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems (CHI EA '17)*. ACM, New York, NY, USA, 1962–1969. https://doi.org/10.1145/3027063.3053081

[34] SFUptownMaker. 2017. Leap Motion Teardown. (2017). https://learn.sparkfun.com/tutorials/leap-motion-teardown

[35] Paul Strohmeier and Kasper Hornbæk. 2017. Generating Haptic Textures with a Vibrotactile Actuator. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '17)*. ACM, New York, NY, USA, 4994–5005. https://doi.org/10.1145/3025453.3025812

[36] Jordan Tewell, Jon Bird, and George R. Buchanan. 2017. Heat-Nav: Using Temperature Changes As Navigation Cues. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '17)*. ACM, New York, NY, USA, 1131–1135. https://doi.org/10.1145/3025453.3025965

[37] Alan Wexelblat (Ed.). 1993. *Virtual Reality Applications and Explorations.* ACADEMIC PRESS PROFESSIONAL, Cambridge, MA, USA.

[38] Graham Wilson, Harry Maxwell, and Mike Just. 2017. Everything's Cool: Extending Security Warnings with Thermal Feedback. In *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems (CHI EA '17)*. ACM, New York, NY, USA, 2232–2239. https://doi.org/10.1145/3027063.3053127