Technische Hochschule Brandenburg
University of Applied Sciences

Fachbereich Informatik und Medien

Bachelorarbeit

Following Mia –
Immersion, Attention and Storytelling in Virtual Environments

Vorgelegt von: Felix Kroischke

am: 02.12.2019

zur Erlangung des akademischen Grades

Bachelor of Science

(B.Sc.)

Erstbetreuer: Prof. Alexander Urban

Zweitbetreuer: Prof. Eberhard Hasche

Abstract

This thesis provides an introduction to the field of Virtual Reality, with a focus on methods to grab and guide the attention of visitors inside a VR-experience. Moreover, a VR-experience that was developed in parallel is presented and explained. This experience – "Following Mia" – illustrates important aspects of VR and visualizes the role of attention in this context. Finally, ideas and approaches for future VR-projects are given.

Zusammenfassung

Diese Bachelorarbeit bietet eine Einführung in das Gebiet der Virtuellen Realität, wobei der Fokus auf Methoden liegt, um die Aufmerksamkeit von Besuchern innerhalb einer VR-Experience zu binden und zu steuern. Zudem wird eine VR-Experience vorgestellt und erläutert, die parallel zum Textteil entwickelt wurde. Diese Experience – "Following Mia" – illustriert wichtige Aspekte von VR und visualisiert die Rolle von Aufmerksamkeit in diesem Kontext. Im letzten Teil der Arbeit werden schließlich Ideen und Ansätze für zukünftige VR-Projekte präsentiert.

Content

Abstract 2

Zusammenfassung 2

1. Objective of this thesis 5

2. Introduction 5

   2.1. The definition of VR, AR, AV, MR and XR 5

   2.2. A word about the current core pieces of VR: headsets 8

   2.3. The potential of VR 8

   2.4. Content types 9

3. Cinematic VR 11

   3.1. New scopes and rules 11

   3.2. Potential stories 12

   3.3. Storyboarding in Cinematic VR 12

4. Following Mia 17

   4.1. The idea 17

   4.2. The environment 17

   4.3. The guide - “Mia” 19

   4.4. Cross-scene-animations and scripts 20

5. The Topics behind the scenes of Following Mia 21

   5.1. Fields of view (FOVs) 21

   5.2. Depth perception and visual cues 22

   5.3. Contrasts 24

   5.4. Lighting 24

   5.5. Motion 25

   5.6. Audio 25

   5.7. Social cues 26

   5.8. Narration 26

   5.9. Signals, arrows and text 26

6. Materials and software used 27

   6.1. Engine and rendering 27

   6.2. Head-mounted-display and graphics card 27

   6.3. Dialogue and Audio 27

7. Ideas for the future expansion of Following Mia 28

   7.1. Additional ways to grab and guide attention 28

   7.2. Additional ways for feedback and measurement of responses 28

7.3. Integration of interactivity and gaze tracking 29
7.4. Haptics as additional dimension 31
7.5. Completion of the theoretical background 31
8. Conclusion 32
9. Literature 33
10. Acknowledgements 37
11. Contact 37

1. Objective of this thesis

This bachelor's thesis primarily addresses creatives who are interested in working with virtual reality. Firstly, it wants to provide an overview about VR that includes definitions, potentials and references to interesting or important projects. Secondly, it wants to give an introduction into storytelling in VR with a focus on ways to guide attention in virtual environments.

Thirdly, it wants to illustrate these aspects with the help of a tutorial-like VR-experience, that was created as part of this thesis. This VR-experience – "Following Mia" – thereby acts as example, but it was developed in a way that makes it possible to use it independently from this thesis.

2. Introduction

2.1. The definition of VR, AR, AV, MR and XR

To date, there are still no uniform and generally accepted definitions for the terms Virtual Reality (VR), Augmented Reality (AR), Augmented Virtuality (AV), Mixed Reality (MR) and Extended Reality (XR). Moreover, there is usually no differentiation between hardware – usually headsets or smartphones – and contents.

Existing definitions mostly have a describing and blurred character and are usually based indirectly on a study by Milgram et al. (1994) [1]. In the study, a scale was used for classification, that ranged from a purely real environment to a purely virtual environment, the so-called Reality-Virtuality Continuum (RV Continuum).

For the purpose of this thesis, a modified version of the RV Continuum will be used for illustration (Fig. 1) and as a basis for the definitions that are proposed in this thesis (Tab. 1). Thereby, two general core statements can be formulated:

- The terms VR, AR, and AV describe to what extend an individual's perception of the current environment is altered. This alteration must not cause changes in the real environment.
- The term Mixed Reality (MR) is used when real and virtual objects can interact while an individual's perception of the current environment is altered in any way.

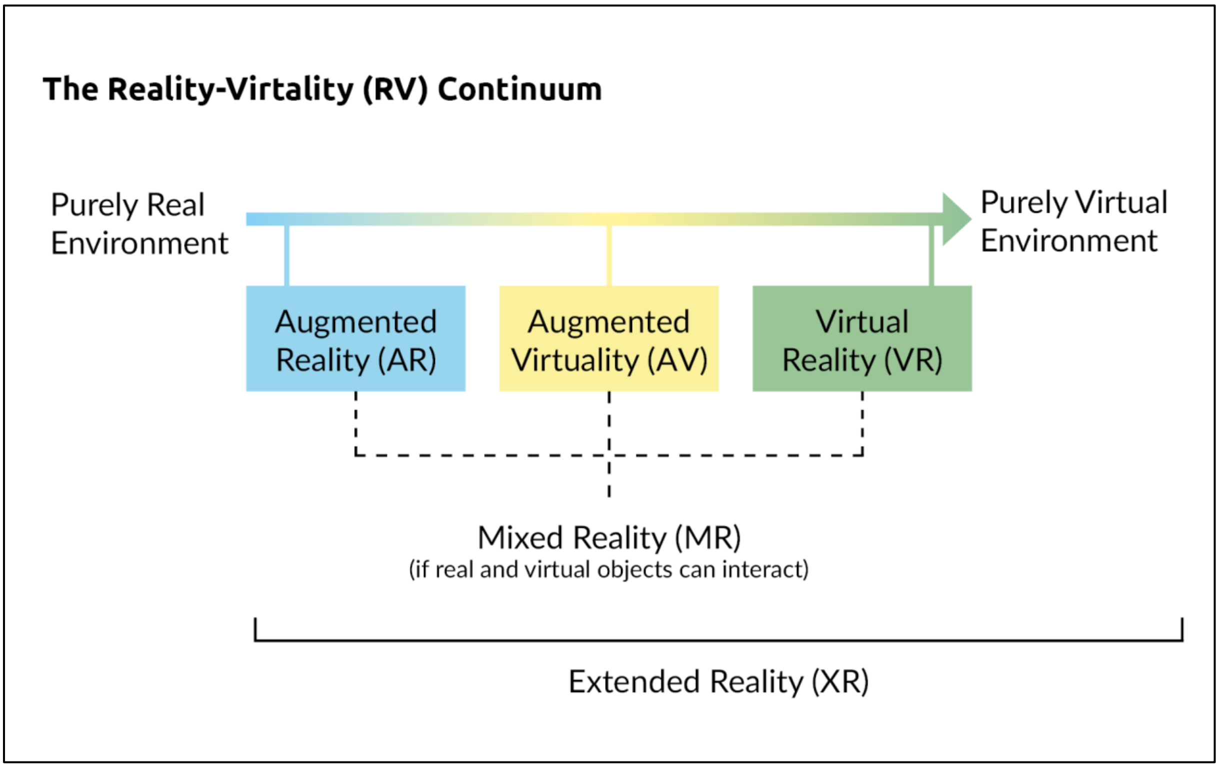

Fig. 1: Reality-Virtuality (RV) Continuum. The continuum spreads from a purely real environment to a purely virtual environment and comprises AR, AV and VR; terms, which describe to what extend the perception of an individuum is altered by generated information.

Tab. 1: Proposed definitions based on the RV Continuum

Term	Definition
Purely Real Environment	The perception of an individuum is not altered in any way.
Augmented Reality	The perception of an individuum is complemented by generated information that overlay parts of the real environment. Interactions between real and virtual objects are not possible.
Augmented Virtuality	The perception of an individuum is based on generated information that are complemented by single objects of the real environment. Interactions between real and virtual objects are not possible.
Virtual Reality	The perception of an individuum is based entirely on virtual, generated information. Interactions between real and virtual objects are not possible.
Purely Virtual Environment	= Virtual Reality
Mixed Reality	The perception of an individuum is altered like in AR, AV or VR, but simultaneously interactions between real and virtual objects are possible.
Extended Reality	Umbrella term that comprises all approaches to alter the perception of an individuum.

2.2. A word about the current core pieces of VR: headsets

The main component of modern VR- and AR-headsets - the so-called head-mounted displays – already look back on a long history of nearly 60 years [2]. However, it took a long time until Facebook launched the final version of the Oculus Rift Headset in 2016 [3], the first device that successfully introduced VR to a broad range of developers and customers. Over the last years, a rapid development set in, resulting in headsets with higher graphical quality, easier handling and a broader range of contents.

Inside-out tracking removed the need for external sensors [4], wireless technology allowed premium headsets to work without cables [5] and very recent headsets are even able to function without an external computer, offering an independence from location and separate hardware [6]. Simultaneously, approaches like Google Cardboard [7] made it possible to use smartphones as head-mounted displays, allowing people to experience VR without high investment costs.

2.3. The potential of VR

What makes VR stand out from other media is the extremely high degree of immersion that results from the complete replacement of the perceptible environment. As a result, visualizations, changes of perspective, journeys and trainings are possible like never before.

VR-experiences can provide a realistic but safe trainings environment for prospective surgeons [8]. Children with autism can improve their communication skills unstrained through simulated conversations [9]. VR can act as a supportive tool when treating post traumatic stress disorders [10] or during rehabilitation phases of brain damage patients [11]. VR makes it possible to visit restricted or dangerous places first hand, for example the White House in Washington [12] or the city of Raqqa, the former "capital" of the Islamic State group [13].

Additionally, there are more and more applications and contents that break up the single-person-nature that was a core characteristic of VR until now. Games like Star Trek Bridge Crew [14] or the different experiences of The Void [15] require intense interactions between players to be successful. According to studies, it could even be possible that VR could even replace face to face communication one day [16], a perspective that should not be underestimated in a time where extensive travelling becomes questionable against the background of climate change.

However, aside of training, documentation, games and communication, VR offers another potential: the potential to tell stories like never before. The term “viewer”, that was used in the cinematic context for decades, is no longer appropriate. By putting recipient directly in the center of the action they become “visitors”, who feel like a part of the virtual world that surrounds them.

2.4. Content types

It is difficult to classify different contents for VR because they differ when it comes to origin, intention and degree of possible interaction. However, it is possible to give a rough overview when contemplating these aspects separately (Tab. 2).

Tab. 2: Classification of VR-contents

Classification by		Description
Origin of material	Reality	Reality-based contents usually persist of spherical photos or videos of real environments or live-action scenes that are stitched together from flat photos or recordings.
	Engine	Engine-based contents persist of digitally generated and designed environments, objects etc. *
Possible depth perception	Monoscopic	Monoscopic contents appear flat to the visitor.
	Stereoscopic	Stereoscopic contents allow depth perception.
Interactivity	Photos or non-interactive videos	Photos or non-interactive videos can be viewed, but their process cannot be altered, and the position of the visitor is stationary.
	Explorable content	Explorable content cannot be altered through interaction, but the visitor is able to change his position and / or perspective.
	Fully interactive content	Fully interactive content reacts to a visitor's actions.

** Sometimes, engine-based contents are also referred as "Game Engine Based" [17] due to the fact that the engines used for content creation are primarily used in game development [18], [19].*

3. Cinematic VR

3.1. New scopes and rules

To properly understand the idea and the background of Following Mia, it is necessary to know more about a specific field of VR: the so-called Cinematic VR.

Cinematography - “the art or process of making films/movies” [17] – was a field which always lend itself to VR. The new medium promised new scopes for design and a way higher immersion.

Over the last years, many experiments and ideas were realized, reaching from simple 360°-documentaries by National Geographic [18], [19] to extensive movies like Crow: The Legend [20], Invasion [21], or Gloomy Eyes [22]. Specialized studios like the Oculus Story Studio [23] or Google Spotlight Stories [24] were founded to test the limits of VR-stories. And scientific studies confirmed that the subjective experience and the physiological reaction showed significantly stronger emotional effects in VR compared to 2D films [25].

Nevertheless, it is still unclear how to exactly define and distinguish this new field, for example when it comes to partially interactive movies. Even the term “Cinematic VR” is controversial and can be seen besides other expressions like “virtual reality cinema” or “VR film” [26]. It can be assumed that this uncertainty mainly results from the fact that many established rules for filmmaking do not apply well to the new medium VR. To mention one example: the “three-point lighting” – the usage of three static but well positioned lights to model a subject with illumination – can lose its effect in VR due to the possible movement of the visitors.

At the moment, the number of structured, developable approaches to find or create rules or general principles for Cinematic VR is still relatively low. So far, the probably most promising approach was proposed by the director, writer and theorist Jessica Brillhart, wo introduced the idea of “potential stories” [27].

3.2. Potential stories

When it comes to traditional movies, it is possible to identify some commonalities. Usually, the viewer has the role of an observer and the visible part of the world is predefined through camera angles, transitions and cuts. The story of these movies is also predefined, linear and unalterable.

In contrast, visitors in a VR-experience can decide themselves where they want to look at a given moment. They are not limited by a frame or predefined camera angle and can therefore examine a scene as they wish. As a result, every visitor experiences a different story, depending on the details he pays attention to and the perspectives he chooses. It is possible to grab and guide attention in a way that most visitors will have a similar experience, but there will always be a gradient from a main plot to an infinite number of other, more or less different experiences.

For content creators, the existence of those "potential stories" leads to new challenges, but also to new possibilities. On the one hand it requires decent planning and an intelligent management of resources to create scenes and worlds that extend behind a single, predefined frame. On the other hand, the new, central position in the scene and the freedom of choice allow an unmatched degree of immersion for the visitors. They are no longer limited to the role of an observer.

In VR, the storytelling shifts away from a main plot towards the creation of vivid worlds that act as stage where potential stories can occur. To use the words of director and theorist Jessica Brillhart [27]: "VR is more about the universe, […] and my job as a creator is to help pull you along through that universe."

3.3. Storyboarding in Cinematic VR

Due to the three-dimensional, possibly interactive nature of VR and the lack of a determined frame, new methods for pre-production have to be found. Especially the storyboarding must be approached differently, complementing traditional storyboards (Fig. 2) with additional information about the environment around the visitor.

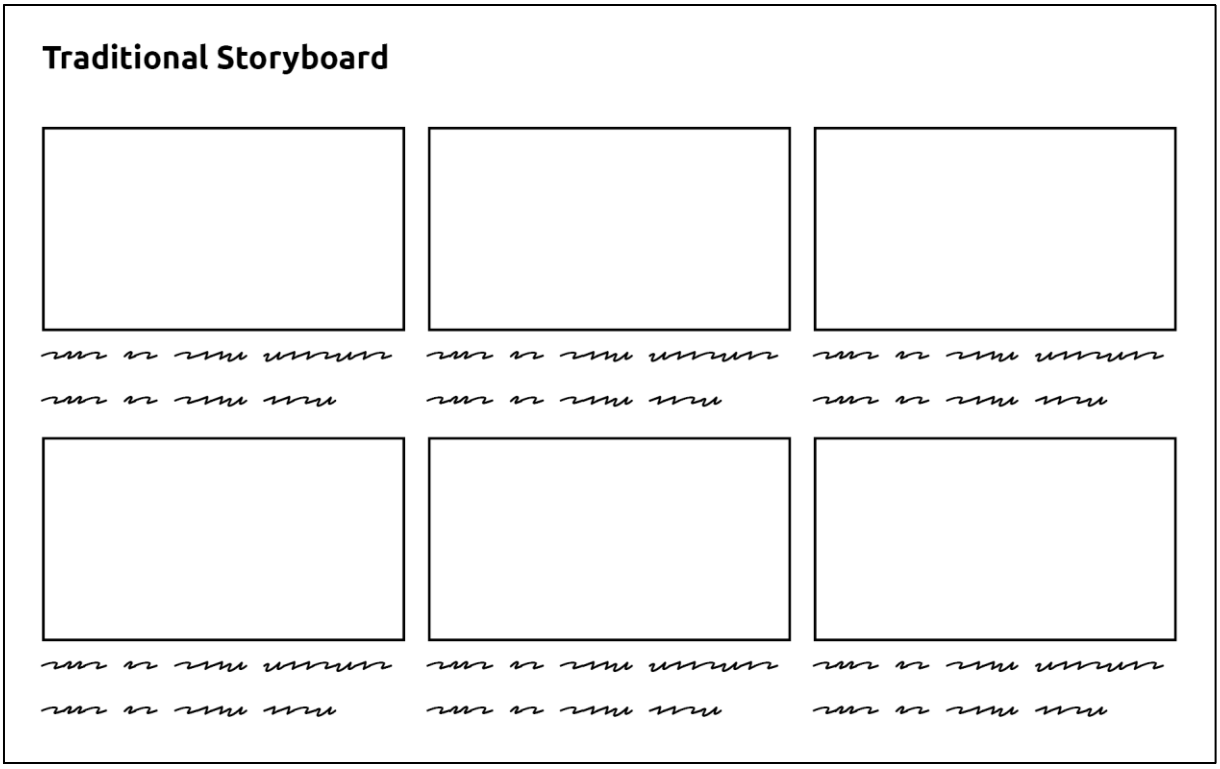

Fig. 2: Traditional storyboard used for filmmaking. Although such traditional storyboards differ in the information provided (descriptions, dialogues, shots, location, sound, etc.), they all follow a similar approach: A scene is visualized by a single or a few frames that act as illustration. However, for VR, this two-dimensional approach is not practicable.

New approaches try to provide this information in different ways. Some simply expand the frame of traditional storyboards by new areas that contain what a visitor could see when turning his head slightly (Fig. 3). However, this is only practicable for situations where it is unlikely that the visitor examines his environment extensively. Nevertheless, it is an easy way to concept games like Beat Saber [28] or applications for visualization like Nanome [29].

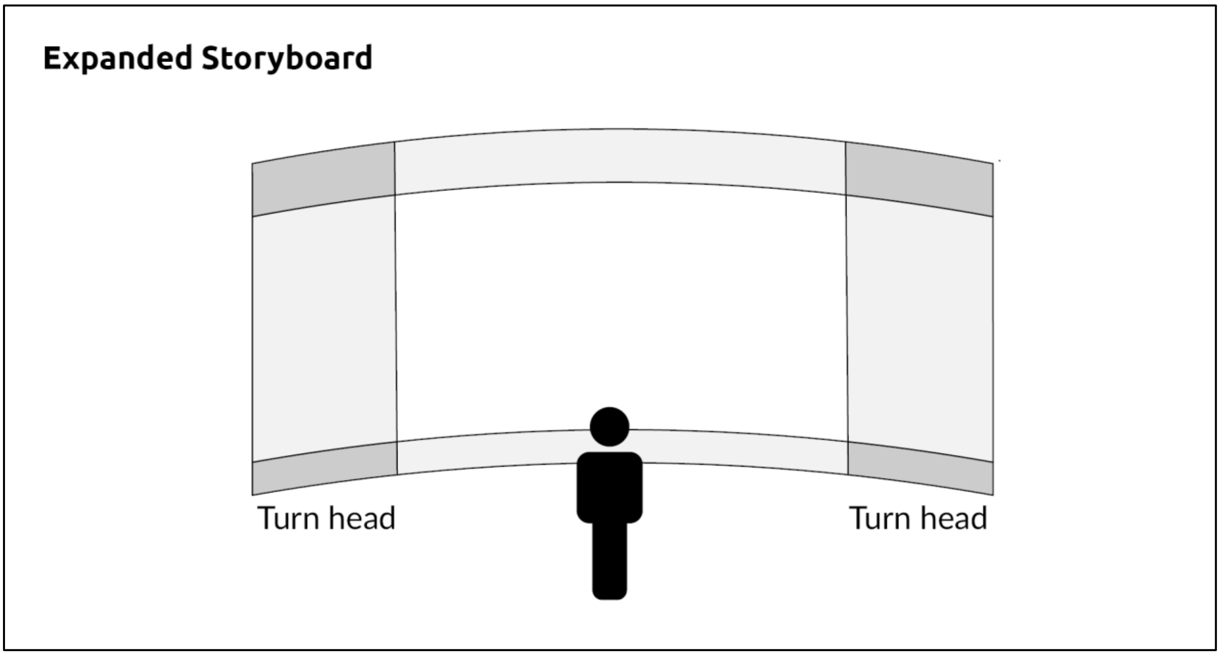

Fig. 3: Expanded storyboard. The traditional two-dimensional frame is replaced by a curved one, which is more convenient to represent a visitor's perspective. By separating the frame into different areas, it is possible to not only sketch what a visitor can see in a specific moment, but also what he could see if he moves his head slightly in one direction (light grey areas) or if he moves his head slightly in two directions (dark grey areas).

A more complex approach are "Unfolded Cube storyboards" (Fig. 4), where the surroundings of a visitor are represented by six coherent frames (front, left, right, top, bottom, back), as if the visitor would stand inside a big cube. In this way, complex environments can be illustrated.

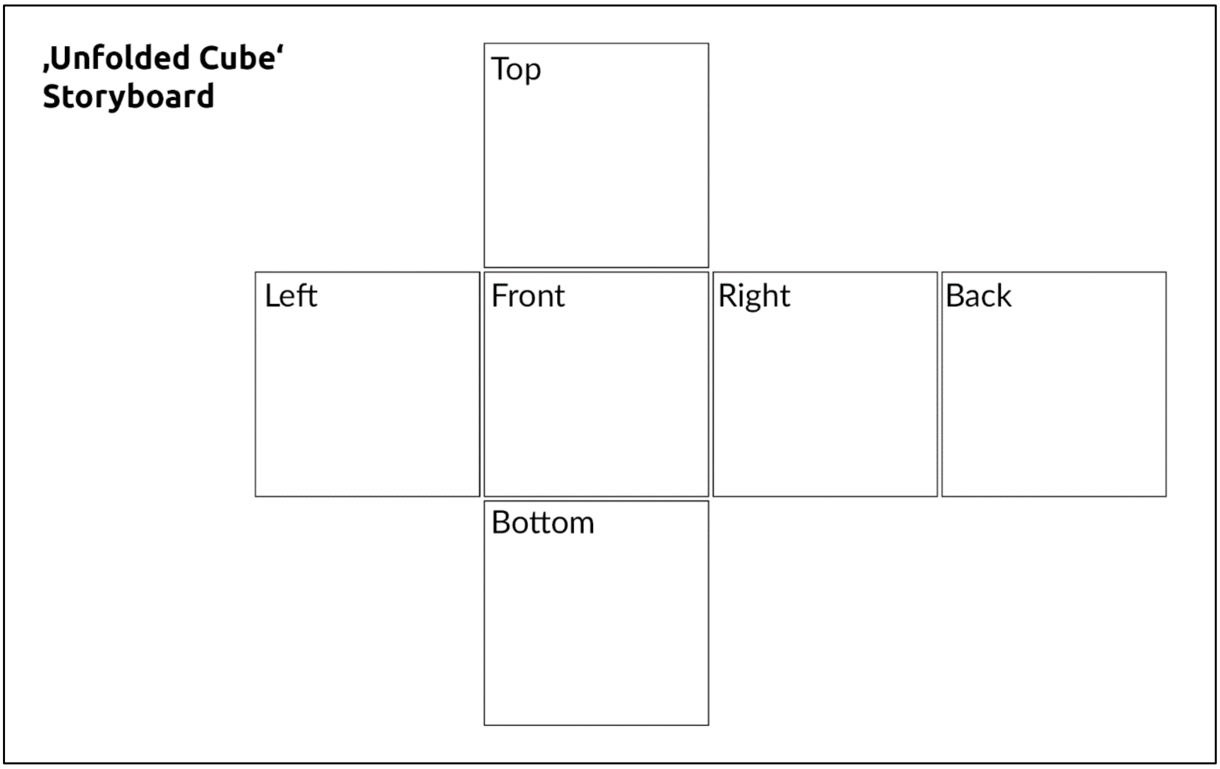

Fig. 4: "Unfolded Cube"-Layout for VR-storyboarding. This storyboard-type is used to sketch the complete surroundings of a visitor at a specific moment. The different directions (in front of the viewer, on his left, on his right, …) are represented by coherent frames, as if a cube around the visitor was fold out.

A third storyboard-type, that could probably be called “Isometric storyboards” (Fig. 5), gets by without frames or frame-like adaptions completely. Here, the visitor and his surroundings are displayed as isometric projection. This allows a better estimation and planning of the distances and the spatial atmosphere of the environment. At the same time, a shading of the circle still allows the classification into different viewing areas. The only problem with this approach is the difficulty to display height differences between objects. However, it is possible to use vertical lines of different length to illustrate such differences.

For Following Mia, a slightly adjusted isometric layout was used (Fig. 6). Instead of a circle a hexagon was used as central plane, representing the hexagonal base area of the room that acts as environment (p. 17).

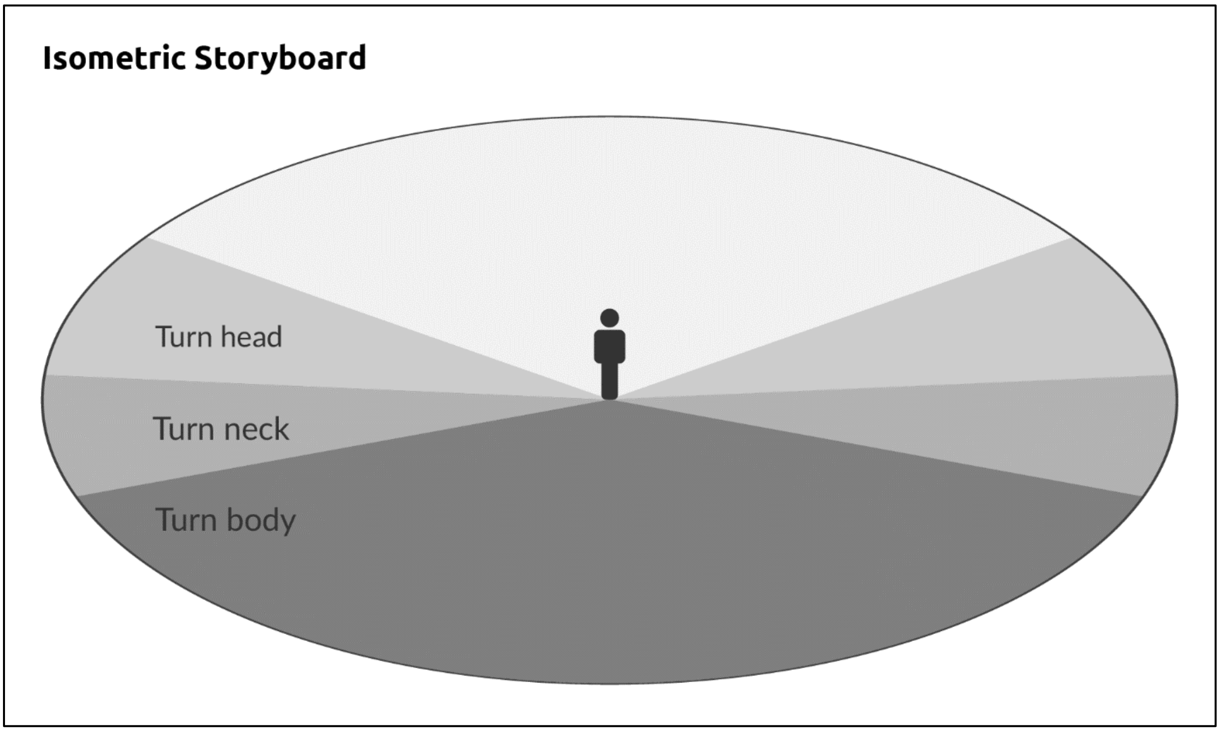

Fig. 5: Isometric storyboard. The visitor and his surroundings are sketched as isometric projection. This allows it to better estimate the spatial atmosphere of a scene as well as distances. The different shades on the circle thereby represent the areas that are only visible to the visitor if he turns his head, neck or body.

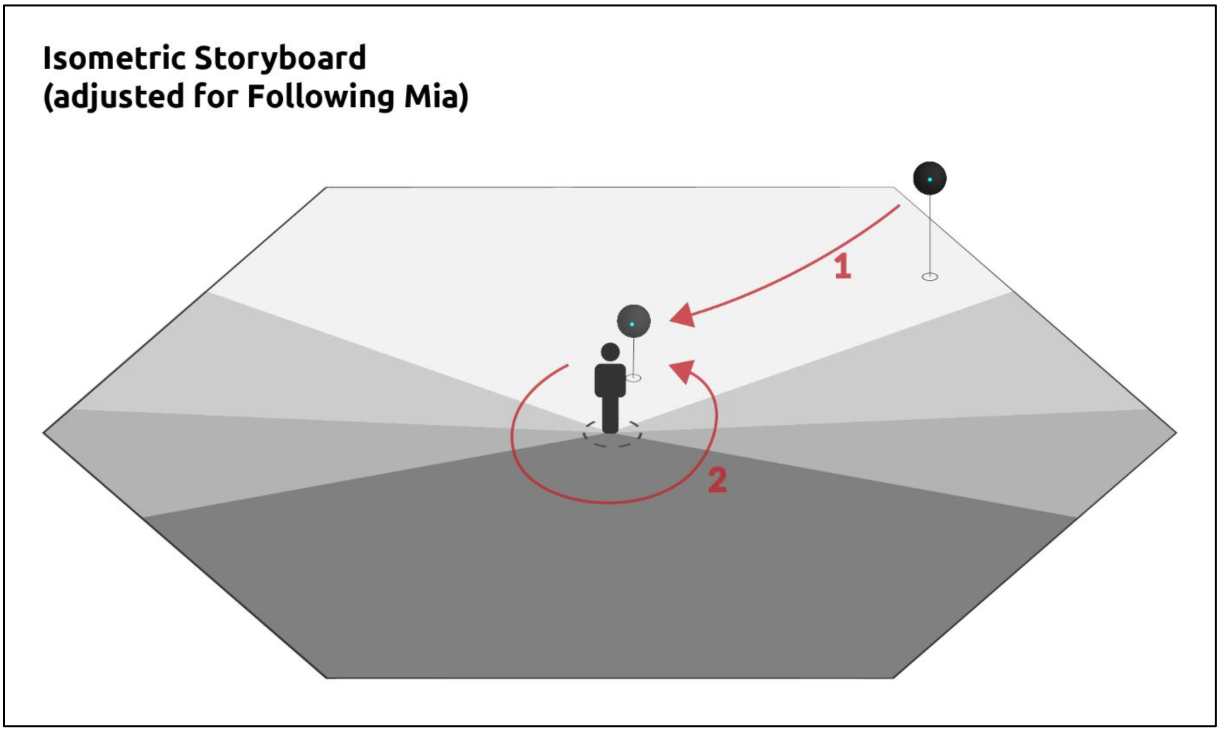

Fig. 6: Isometric storyboard, adjusted for Following Mia. The layout used for Following Mia was altered slightly to directly represent the six-sided nature of the room that acted as environment. This example also illustrates how it is possible to display height differences in an isometric projection: lines of different length, which stand vertical on the plane represent different heights.

4. Following Mia

4.1. The idea

The idea of the VR-experience “Following Mia” was to create a starting point for people who want to know more about storytelling and attention in VR. The experience was intended to be enjoyable and informative at the same time, condensing and illustrating the current state of knowledge and offer inspiration and approaches for new VR-projects. Although demonstrations of the concepts could have been based on actual examples (see chapter 2.3 / “The potential of VR”), this approach was discarded because it would have been possible that the examples shift the focus away from the actual explanations. Therefore, Following Mia was designed from scratch and without foreign material.

After a brief conception phase, the experience was designed like a film, presenting the content in consecutive but self-contained scenes that illustrate explanations visually.

4.2. The environment

Following Mia takes place in a single, spacious room without furniture or other decorations. The design of the room is simplistic but avoids dullness through the unusual layout and the contrasts between white and grey (Fig. 7). While most elements from early concepts (Fig. 8) persisted, the shape of the room changed during the development process. The originally rectangular base was found to be boring by different testers and the round edges and corners were hard to identify due to the reduced depth perception in VR (see chapter 5.2 / “Depth perception and visual cues”). The final concept addresses these issues and also uses a striped texture to support a depth perception of the now angularly walls.

The blue light bars act a futuristic element and loosen up the otherwise rather sterile atmosphere.

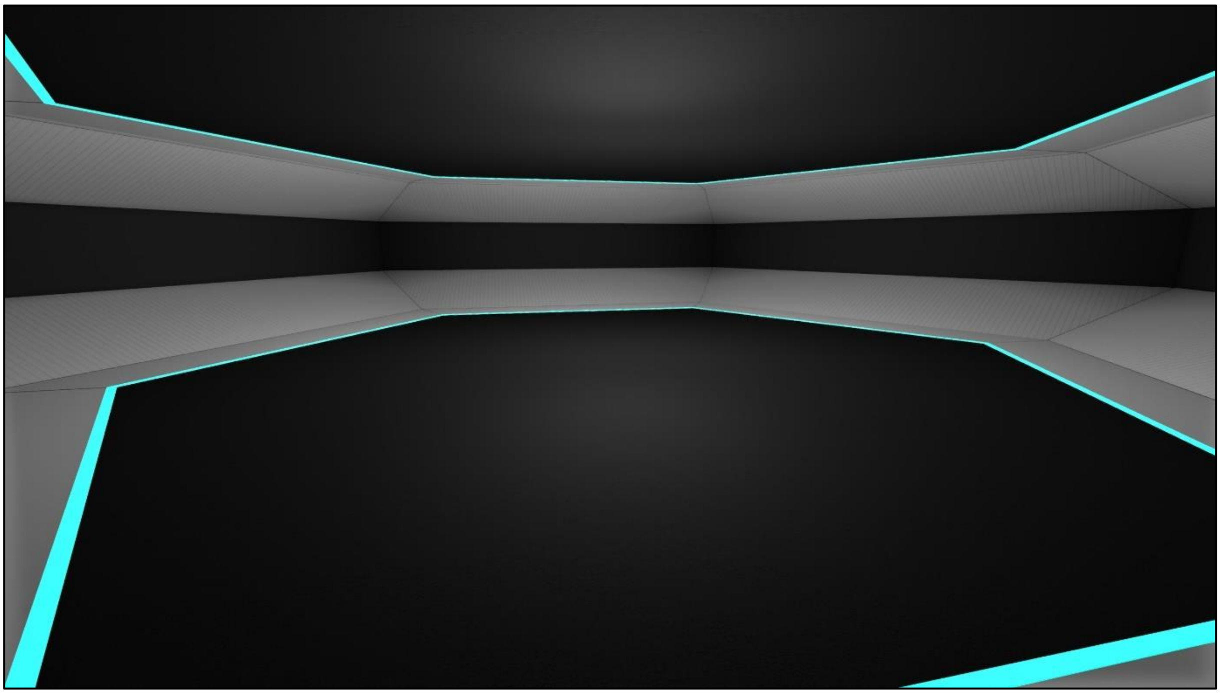

Fig. 7: The environment where Following Mia takes place. The room tries to achieve a combination of simplicity, general neutrality and interesting and loosening details.

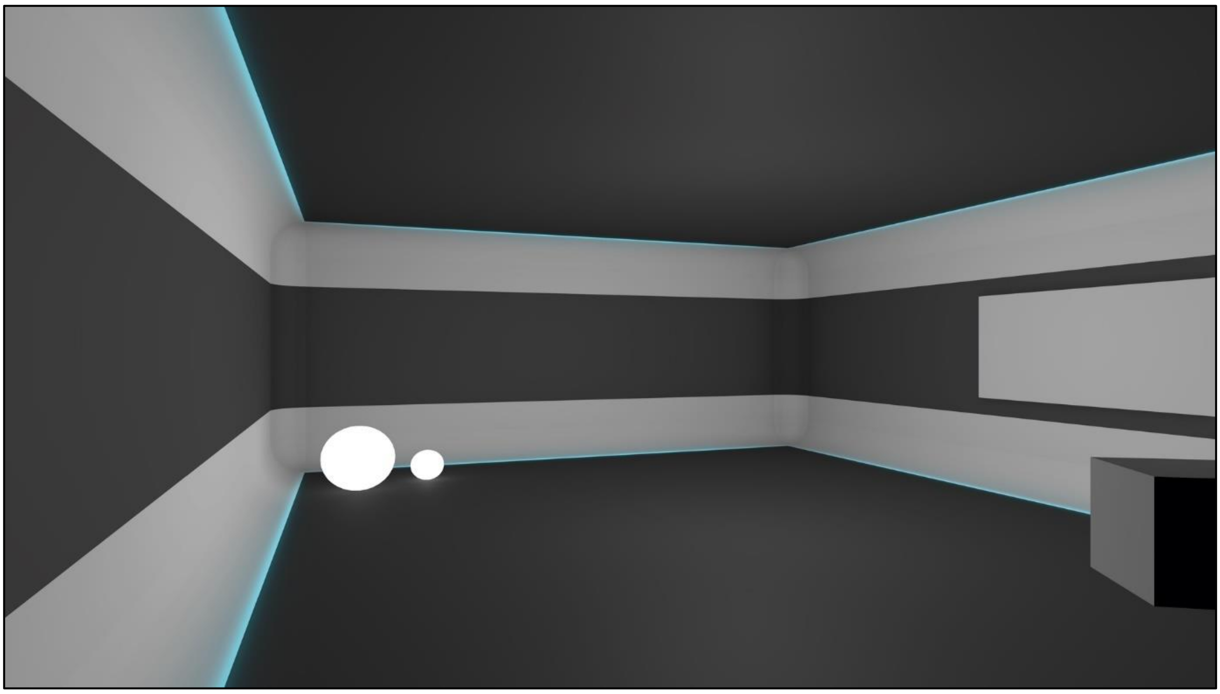

Fig. 8: An early concept of the room that acts as environment in Following Mia. Many of the elements, like the colors or the light bars, were taken over to the final concept, but the shape of the room changed extensively.

4.3. The guide - “Mia”

Mia – a little, robotic sphere with a big glowing eye – is the main character of Following Mia and acts as guide trough the experience (Fig. 9). Nearly all of the examples and scenes included in the experience need additional explanations that are provided as audio. Mia gives these explanations a face, so that visitors do not get the impression that they are only guinea pigs in a laboratory environment.

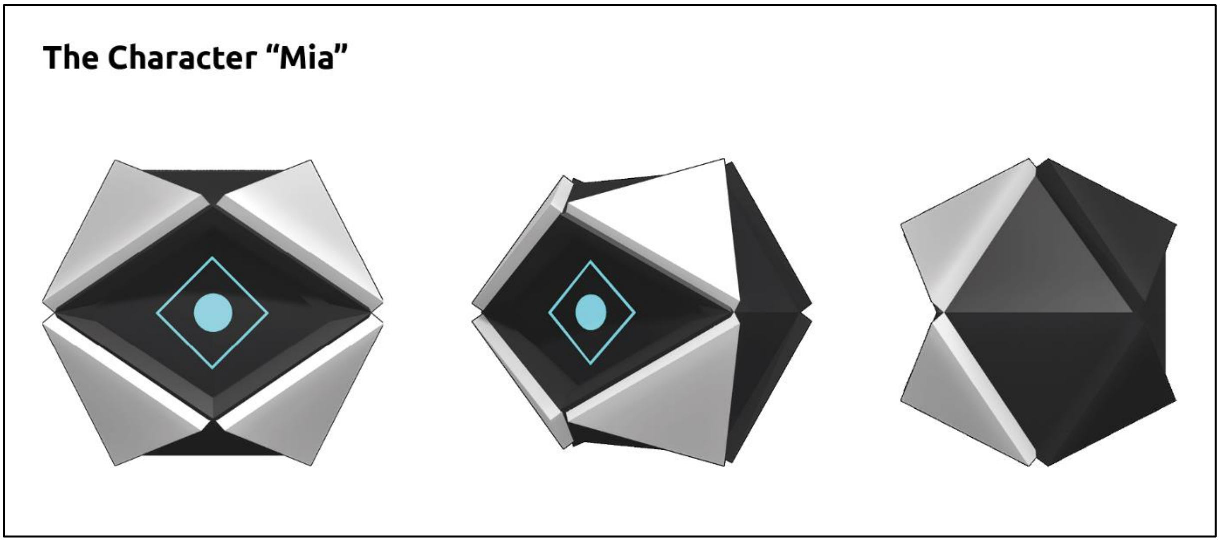

Fig. 9: "Mia", a floating robot that acts as main character during the experience. Her name was originally just a working title but was kept after positive feedback by testers.

Mia’s robotic form was chosen for two reasons. Firstly, it intensifies the impression of a futuristic environment that is provided by the room. Secondly, it allows uncomplicated and easy horizontal and vertical movement in the room without the need for walking animations, what freed resources for other aspects of the thesis during the development.

Moreover, the concept of a floating sphere was already found to be successfully in other games, movies and VR-experiences, e.g. in the Destiny-franchise (Fig. 10) [30] or the Steam VR tutorial [31].

Fig. 10: Render of a "Ghost", a little robot that acts as a companion and assistant in the Game Destiny 2 [30]. The Ghosts where an inspiration for Mia, which guides the visitors through the VR-experience that was created along this thesis.

4.4. Cross-scene-animations and scripts

While the different scenes of Following Mia have their individual approaches and animations, there is a general and cross-scene difficulty: because of the high amount of explanations required for the different aspects of VR, attention and storytelling, there are phases of low visual activity. To prevent that these phases lead to restlessness and boredom, a continuous but unintrusive visual stimulus was implemented in the form of Mia’s eye, which changes its brightness depending on her current audio output. By linking her voice-audio-track to the emission level of the material that forms her eye it is possible to achieve an effect similar to the iconic glowing eye of HAL 9000, the machine-antagonist from the movie 2001: A Space Odyssey [32].

The second cross-scene-script links the rotation of Mia’s body to the current position of the visitor (Fig. 11). In this way it is possible to make Mia look towards the visitor, even if he moves during the experience. Additionally, this automation removes the need to adjust Mia’s body rotation during her own movements during the experience and it can also be adjusted target-wise to direct Mia’s gaze towards other objects.

```
using UnityEngine;

public class FollowVisitor : MonoBehaviour
{
    [SerializeField]
    // The Target to look at (= the visitor) and the speed of the rotation towards
    the target
    private Transform target;
    private float speed = 5f;

    private void Update()
    {
        // Comparation of the position of the object (= Mia) compared to the
        position of the target
        Vector3 direction = transform.position - target.position;
        // Rotation of the object (= Mia)
        Quaternion rotation = Quaternion.LookRotation(direction);
        transform.rotation = Quaternion.Lerp(transform.rotation, rotation, speed *
        Time.deltaTime);
    }
}
```

Fig. 11.: Script to make it possible for Mia to track the visitor or another object of choice. Because of Mia's fixed eye, a simple rotation of her spherical body can create the impression that she is looking in a specific direction. Because it is possible to calculate this rotation at runtime, Mia can also "look at the visitor", even if he is moving.

5. The topics behind the scenes of Following Mia

Following Mia is a cinematic VR-experience with a predefined sequence. However, it consists of scenes that are consecutive but also self-contained to a high degree. Each of these scenes deals with one aspect of VR respectively one way to guide attention in VR.

Within the experience, the scenes are meant to illustrate problems or possible approaches to guide attention in a demonstrative, enjoyable way without overtaxing the visitor through too much information. However, within this chapter of the thesis, the backgrounds of the different topics are explained in a more detailed way. The background of the first scene – potential stories – was already explained earlier in this thesis (see p. 12).

5.1. Fields of view (FOVs)

According to the Oxford Advanced Learner's Dictionary [33], the field of view (FOV) is defined as "the total amount of space that you can see from a particular point without moving your head”. Humans have a natural FOV of about 210° horizontally and 150° vertically.

However, current VR-headsets still narrow this FOV significantly, mostly due to the sheer amount of computing power that would be required for wider screens and therefore higher resolutions. The average horizontal FOV of current VR-headsets is about 110°, only a few professional or experimental models like to Pixmax Vision 8K [34] or the StarVR One [35] reach 210°. As a result, content creators have to pay attention to the scene design in VR to prevent that objects or characters that belong together cannot be seen at once due to a too small FOV.

5.2. Depth perception and visual cues

In his article for the Oxford Research Encyclopedia [36], Neuroscientist Andrew J. Parker summed up depth perception as follows: “Humans and some animals can use their two eyes in cooperation to detect and discriminate parts of the visual scene based on depth. Owing to the horizontal separation of the eyes, each eye obtains a slightly different view of the scene in front of the head. These small differences are processed by the nervous system to generate a sense of binocular depth. As humans, we experience an impression of solidity that is fully three-dimensional; this impression is called stereopsis and is what we appreciate when we watch a 3D movie or look into a stereoscopic viewer.”

This summary also applies to VR. However, the limited resolution of VR-headsets leads to the problem that characteristics of objects like shading, texture, relative motion or contours become harder to identify the farther away an object is. Additionally, the less pixels are used to display an object, the lower are the perceptible differences between the images shown to the left and the right eye of a visitor.

It is possible to calculate the maximum perceivable distance for a head mounted display by using horizontal FOV and resolution of the corresponding headset as well as the so called interpupillary distance (IPD) of the user, which describes the distance between the centers of the pupils of the eyes (Fig. 1) [37]. As a rule of thumb, it can be assumed that UI elements should be placed in a distance of circa one meter [38], while the general comfortable viewing range with strong depth perception lies between 0,75 m and

10 m. Behind that distance, the sense of stereo depth fades of until objects appear flat and cannot be separated anymore without additional visual cues (Fig. 12) [37].

Visual cues can give additional information about objects and are mostly monocular, meaning they can be perceived with only one eye although they support stereopsis. Examples for visual cues that can be used in VR-experiences are:

› Motion parallax - Objects that are closer to the visitor appear to move faster than object which are farther away.

› Relative scale - Objects get smaller with increasing distance from the visitor.

› Occlusion - Closer objects block the view of objects that are farther away / behind the closer objects.

› Texture gradients - Repeating patterns appear denser as they recede into the distance.

› Lighting - Shadows or highlights make it easier to perceive the shape and position of objects.

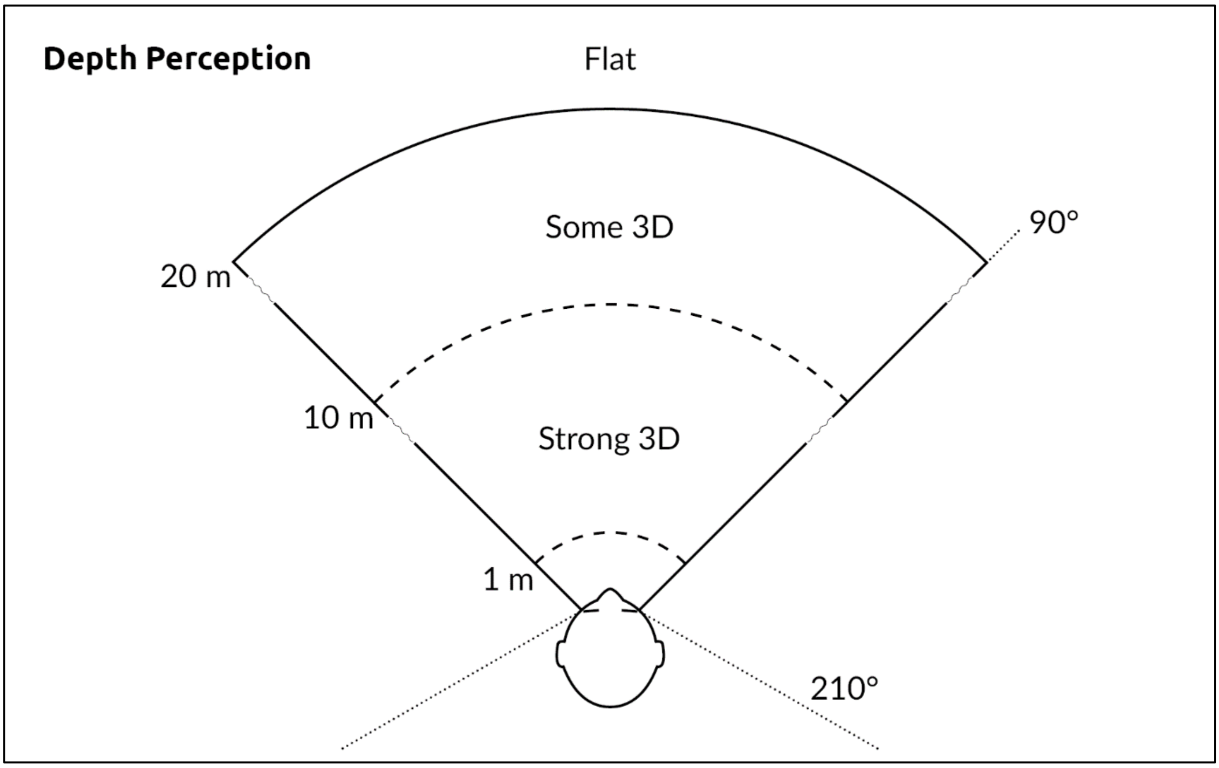

Fig. 12: Depth perception in VR. While the depth perception is strong between 0,75 m and 10 m, it fades behind that distance until objects appear to be flat at roughly 20 m. The figure is based on a graphic by Mike Alger [37].

5.3. Contrasts

The importance of contrast is widely accepted and well documented when it comes to the design of websites or print products [39]. Contrasts help to distinguish objects, form a visual hierarchy and are also useful to improve visual processing as well as fixations in natural scenes [40].

When it comes to VR tough, the impact and effects of contrasts were researched only to a minor degree so far. However, there are currently no signs indicating that contrast would work different in the new medium.

5.4. Lighting

In traditional filmmaking, lighting is the key to set up a scene and to create desired moods and to intensify or mellow the emotional responses that naturally come from the narrative [41].

In VR, this still applies. However, here light can also be a way to highlight characters, objects or areas of interest, like it was described in detail by Oriyomi A. Adenuga [42]. Possible approaches are light contrasts, colored lights, moving lights, staged lighting, fading lights, pulsing lights and follow spotlights.

5.5. Motion

Our history as hunters and gatherers hard-wired our eyes and brains to motion, especially if this motion includes animate objects [43]. As soon as something in our environment moves unexpectedly or the direction or speed of a present motion changes suddenly, we react heavily to it [44] – an impulse that surely saved many of the lives of our ancestors. As a result, motion can be a strong tool to grab the attention of visitors in a VR-environment and it can moreover be used to increase the effectiveness of other approaches like lighting [45].

5.6. Audio

Audio is a very important, yet often overlooked topic when it comes to filmmaking and game development [46]. Voices, music and sound effects make a scene feel realistic, increase emotional responses or are even able to trigger automatic, involuntary experiences in a second sensory or cognitive pathway through a neurological effect called Synesthesia [47].

In a three-dimensional environment like a VR-experience, audio can also act as a tool for location and orientation, provided that the experience uses proper, binaural audio (a three-dimensional stereo sound sensation for the listener) [48]. This possibility is especially important when it comes to attention, because auditive cues do not rely on a line of sight like motion, lighting or text. Instead, auditive cues work permanently and are therefore able to grab a visitor's attention independently from his currently chosen FOV.

Lastly, it has also to be mentioned that three-dimensional environments are able to significantly increase emotional responses produced by audio, as the example of Senua's Sacrifice VR demonstrates impressively [49].

5.7. Social cues

As highly social beings, humans have developed automatisms to rapidly analyze facial expressions, gestures and body language. By observing others, humans gather so-called social cues that significantly influence attention and behavior. Social cues are often preferred to other stimuli when it comes to orientation tasks [50] and especially the direction of another person's gaze can trigger a reflexive shift of an observer's visual attention [51].

This reactivity can also be utilized for VR-experiences. By using other characters present in the scene, visitors can be manipulated without the need for possibly immersion-breaking changes to the environment. Gestures, gazes, interactions and expressions of joy, fear curiosity etc. offer many ways to direct a story.

5.8. Narration

Narration – here in the broader sense the commentary of a narrator – is usually an element to give the audience additional information about the story or the thoughts and feelings of characters. This commentary usually has no effect on the events in the movie because the acting characters are not aware of it. However, if the characters are aware of it, it can influence the events significantly, as shown in the short film "The Gunslinger" [52].

In VR, where a visitor can decide where to look, narration can be used to influence this decision by directly or indirectly pointing out details and interesting aspects or by giving additional information about things that could happen next.

5.9. Signals, arrows and text

In traditional filmmaking, text is usually used rarely, apart from advertisements, trailers or the visualization of chats and thoughts like in the BBC series "Sherlock" [53]. The usage of arrows is even more extraordinary.

In traditional movies, "Exit" or road signs as well as labels on objects are usually part of the background and only attract notice when the characters in the scene pay attention to them or they are shown in a close-up. In VR however, where the focus is defined by the visitor, they can give cues and direct attention without breaking the immersion.

6. Materials and software used

6.1. Engine and rendering

Following Mia was developed in Unity, a game engine that supports different types of head-mounted displays and is usable free of charge for non-commercial projects [54].

Even tough Unity offers a rendering pipeline tailored specifically to mobile devices and VR-applications [55], the standard pipeline of the engine was used to achieve a higher graphical quality, e.g. through real-time shadows or ambient occlusion. Because of the abstract and simplified approach of the experience this was possible without a high loss in frame rates that would have been a problem otherwise.

6.2. Head-mounted-display and graphics card

For the development of Following Mia, an Oculus Rift S headset [56] was used. The graphic card used was a GeForce RTX 2060.

6.3. Dialogue and audio

Mia's voice was generated via the cloud text-to-speech service of Google [57]. The other audio files used where either royalty-free or self-recorded. To enable high-quality binaural audio, the Oculus Native Spatializer for Unity [58] was used instead of the less efficient standard audio spatializer of the Unity engine.

7. Ideas for the future expansion of Following Mia

This thesis and the corresponding VR-experience were produced over a period of less than six weeks. As a result, many ideas could not be tested or implemented. However, they could be useful for a future expansion of Following Mia or the creation of similar projects. Therefore, some of them are presented shortly in this part of the thesis.

7.1. Additional ways to grab and guide attention

Although many different ways to grab and guide attention were presented in this thesis, there are still several more possible approaches that could be researched, evaluated and integrated. Examples for such approaches include the influence of size, color, alignment, repetition, proximity, space and texture – characteristics that are all part of visual hierarchies [59].

Another concrete starting point for future expansions could be the article “Guidance in Cinematic Virtual Reality-Taxonomy, Research Status and Challenges” by Rothe et al. [60], which reviewed and classified several more attention-guiding methods that also exceed the design of the virtual environment, e.g. blurring, forced rotation or on-screen radars to indicate the direction of POIs.

7.2. Additional ways for feedback and measurement of responses

Although the feedback of several people contributed to this thesis, reactions and opinions were not measured in a standardized way. Similar or future projects could probably change that through the use of 360°-heatmaps (Fig. 13). While Following Mia was designed completely within one Unity scene to prevent loading times, a future approach could divide the experience into several parts, eventually selectable via a menu. This division would simplify evaluation.

Heatmaps however could be implemented into Unity with relative ease [61] and would allow a more scientific feedback approach because they do not rely on possibly imprecise or unconsciously compromised, subjective verbal or written feedback.

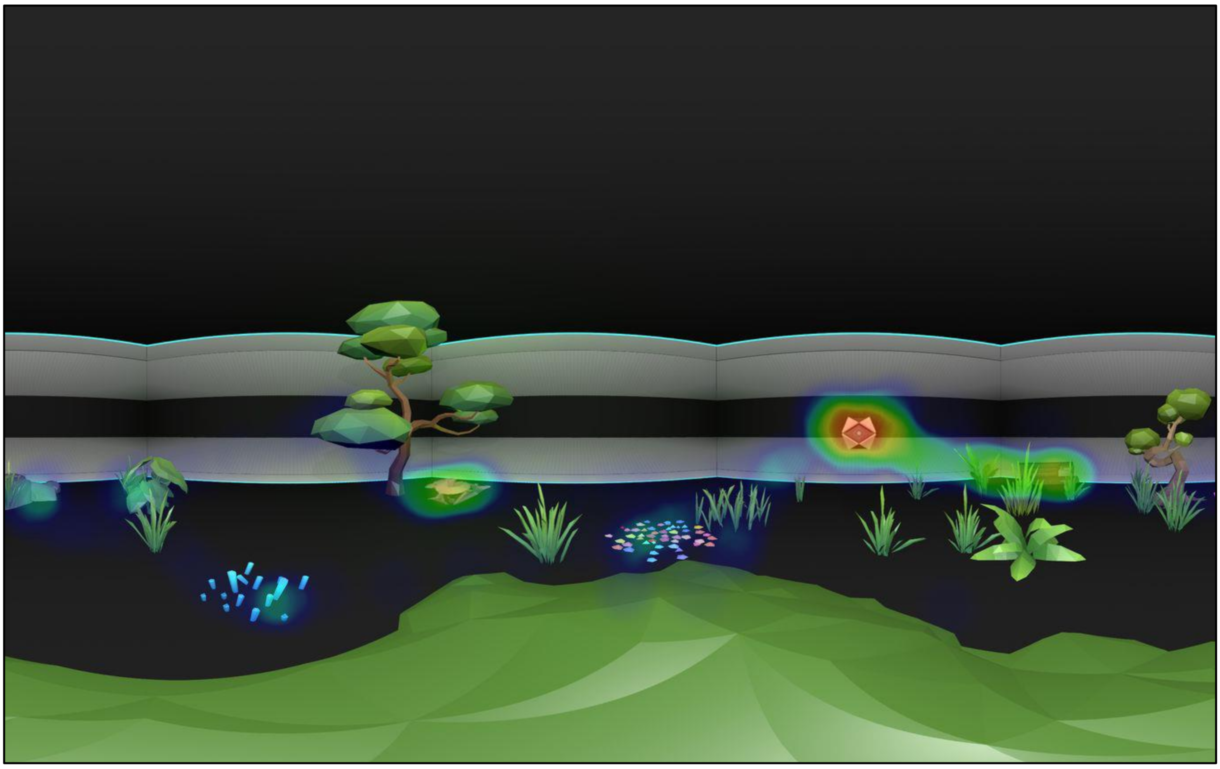

Fig. 13: Mockup of a heatmap. The longer a visitor looks at a specific area of the scene, the more it gets highlighted. After the experience, such results can be evaluated to gather reliable feedback.

7.3. Integration of interactivity and gaze tracking

By now, Following Mia is a passive experience. However, it could make sense to add some basic interactivity to also cover this possible part of storytelling. As an example, there could be a scene that demonstrates what ways on interaction between visitor and VR-environment are possible, e.g. controller movement / input, voice input, movement, etc.

Thereby, a focal point should be interactions through head tracking respectively tracking of a visitor's gaze, because this approach offers an easy and simultaneously immersive way of interaction. So, it requires only a few lines of code to check if an object is currently in front of the visitor (Fig. 14). By implementing a counter, this could be used to trigger events by looking to a specific direction for a fixed amount of time, enhancing immersion and removing the need for controllers.

```
using UnityEngine;
public class VisibilityCheck : MonoBehaviour
{
    [SerializeField]
    // Target of the check, the ganze time required until an effect is triggered and the
    current gaze time
    private Transform target;
    private float requiredGazeTimeInSeconds = 3;
    private float currentGazeTimeInSeconds = 0;
    Camera cam;

    void Start()
    {
        // Identify the current camera (= the FOV of the visitor)
        cam = GetComponent<Camera>();
    }

    void Update()
    {
        // Check if the target is in front of the Camera (= in front of the visitor)
        Vector3 screenPoint = cam.WorldToViewportPoint(target.position);
        if (screenPoint.z > 0 && screenPoint.x > 0 && screenPoint.x < 1 && screenPoint.y
        > 0 && screenPoint.y < 1)
        {
            // As long as the target is in front of the camera, increase the current gaze
            time
            currentGazeTimeInSeconds += Time.deltaTime;
            if (currentGazeTimeInSeconds >= requiredGazeTimeInSeconds)
            {
                // When the gaze time reaches the predefined value, trigger an event
                currentGazeTimeInSeconds = 0;
                // Triggered event (code)
            }
        }
        else
        {
            // As soon as the object is not in front of the visitor anymore, reset the
            gaze timer
            currentGazeTimeInSeconds = 0;
 }  }   }
```

Fig. 14.: Simple script to trigger events by looking at them. In Unity, this script could be attached to the camera to check if a target object is in front of it (and thus in the current FOV). When looking at the object for a certain amount of time, the visitor could trigger an event in this way.

7.4. Haptics as additional dimension

At the moment, VR can replace our visual and auditive perception, but haptics is still limited to plastic controllers which can only give feedback via vibration. However, several companies are already working on universally applicable solutions that could expand future VR-experiences, e.g. gloves (Fig. 15) [62] or body suits [63]. In the long term, even such features could be part of Following Mia because the ability to touch objects could also be used for attention and storytelling, e.g. when it comes to interactions with other characters in the scene.

Fig. 15: Glove which allows haptic feedback in VR-experiences [64]. Feedback points make shapes, textures and motions of virtual objects tangible while an exoskeleton provides resistive force. At the same time, magnetic motion tracking allows interactions with the environment.

7.5. Completion of the theoretical background

Besides the practical additions, many additional theoretical topics could and should be covered in future expansions of Following Mia. Terms like motion sickness, presence, or locomotion (the ways of movement that are possible in VR) should be defined, explained and integrated to complete the theoretical background of VR.

8. Conclusion

This thesis and the VR-experience that was developed in parallel offer a solid starting point for people that are interested beginners when it comes to VR. However, two important insights were gained during the development respectively writing process of this thesis:

› Although many aspects of VR are already covered in this thesis and although there are suggestions for expansions, it will still be difficult to give a complete overview of VR in the future. The field of VR is still developing rapidly.

› Most of the “findings” about VR are currently based on assumptions or opinions of filmmakers who already produced some experiences by trial and error. Therefore, it is difficult to find reliable, science-based information that could be used to formulate clear, reliable rules and definitions that could be used in the future.

9. Literature

[1] H. T. A. U. F. K. Paul Milgram, “Augmented Reality: A class of displays on the reality-virtuality continuum,” *Proceedings of SPIE - The International Society for Optical Engineering,* pp. 282-292, January 1994.

[2] I. E. Sutherland, “A head-mounted three dimensional display,” in *Fall Joint Computer Conference*, San Francisco, California, 1968.

[3] Facebook Technologies, LLC, “Oculus Rift,” [Online]. Available: https://www.oculus.com/rift/. [Accessed 1 November 2019].

[4] A. P. A. L. F. Miguel Ribo, “A new Optical Tracking System for Virtual and Augmented Reality Applications,” *Instrumentation and Measurement Technology Conference,* pp. 1932 - 1936, February 2001.

[5] HTC Corporation, “HTC Vive Wireless Adapter,” [Online]. Available: https://www.vive.com/eu/wireless-adapter/. [Accessed 1 November 2019].

[6] Facebook Technologies, LLC, “Oculus Quest,” [Online]. Available: https://www.oculus.com/quest. [Accessed 1 November 2019].

[7] Google LLC, “Google Cardboard,” [Online]. Available: https://arvr.google.com/intl/en_uk/cardboard/. [Accessed 1 November 2019].

[8] L. E. A. G. G. L. H. C. H. D. A. M. C. S. R. D. S. D. A. Gunnar Ahlberg, “Proficiency-based virtual reality training significantly reduces the error rate for residents during their first 10 laparoscopic cholecystectomies,” *The American Journal of Surgery,* pp. 797-804, June 2007.

[9] K. L. J. R. K. D. L. B. E. L. Natalia Stewart Rosenfield, “A Virtual Reality System for Practicing Conversation Skills for Children with Autism,” *Multimodal Technologies Interact,* 20 April 2019.

[10] A. G. J. X. K. R. L. V. E. P. L. P. A. F. Oswald D. Kothgassner, “Virtual reality exposure therapy for posttraumatic stress disorder (PTSD): a meta-analysis,” *European Journal of Psychotraumatology,* 19 August 2019.

[11] M. F. Levin, “Can virtual reality offer enriched environments for rehabilitation?,” *Expert Review of Neurotherapeutics,* pp. 153-155, 2011.

[12] Felix & Paul Studios, “The People's House - Inside the White House with Barack and Michelle Obama,” [Online]. Available: https://www.youtube.com/watch?v=bqW2qm02jwI. [Accessed 1 November 2019].

[13] BBC, “The city fit for no-one - Inside the ruined 'capital' of the Islamic State group,” [Online]. Available: https://www.bbc.co.uk/news/resources/idt-sh/the_city_fit_for_no_one_raqqa_syria_islamic_state_group. [Accessed 1 November 2019].

[14] Ubisoft Entertainment SA, “Star Treck Bridge Crew,” [Online]. Available: https://www.ubisoft.com/en-us/game/star-trek-bridge-crew/. [Accessed 1 November 2019].

[15] The Void, “The Void,” [Online]. Available: https://www.thevoid.com/. [Accessed 1 November 2019].

[16] M. N. Harrison Jesse Smith, “Communication Behavior in Embodied Virtual Reality,” *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems,* 21 April 2018.

[17] Oxford University Press, “Oxford Advanced Learner's Dictionary - "Cinematography",” [Online]. Available: https://www.oxfordlearnersdictionaries.com/definition/english/cinematography. [Accessed 11 November 2019].

[18] National Geographic Partners, LLC, “How We Made the First 3-D Virtual Reality Video in Space,” [Online]. Available: https://www.nationalgeographic.com/news/2018/04/making-first-360-video-space-nespoli-one-strange-rock-science/. [Accessed 1 November 2019].

[19] National Geographic Partners, LLC, “Venture into the African wetlands,” [Online]. Available: https://www.nationalgeographic.com/animals/2018/12/okavango-wilderness-360-video-launch/. [Accessed 1 November 2019].

[20] Baobab Studios Inc., “Crow: The Legend,” [Online]. Available: https://www.baobabstudios.com/crow-the-legend. [Accessed 1 November 2019].

[21] Baobab Studios Inc., “Invasion,” [Online]. Available: https://www.baobabstudios.com/invasion. [Accessed 1 Novemebr 2019].

[22] Atlas V, “Gloomy Eyes,” [Online]. Available: https://atlasv.io/gloomy-eyes. [Accessed 1 November 2019].

[23] Facebook Technologies, LLC., “Oculus Story Studio,” [Online]. Available: https://www.oculus.com/story-studio/. [Accessed 1 November 2019].

[24] Google LLC, “Google Spotlight Stories,” [Online]. Available: https://atap.google.com/spotlight-stories/. [Accessed 1 November 2019].

[25] N. Ding, W. Zhou and A. Y. H. Fung, “Emotional effect of cinematic VR compared with traditional 2D film,” *Telematics and Informatics,* pp. 1572-1579, September 2018.

[26] J. M. Pérez, “Cinematic Virtual Reality: A Definition and Classification Proposal,” *Departament de Publicitat, Relacions Públiques i Comunicació Audiovisual / Universitat Autònoma de Barcelona,* 2 September 2016.

[27] The Conference / Media Evolution, “Jessica Brillhart – Getting Real with Virtual Reality | The Conference 2016,” [Online]. Available: https://www.youtube.com/watch?v=tQ0clqdPADA. [Accessed 13 November 2019].

[28] BEAT GAMES, “Beat Saber,” [Online]. Available: https://beatsaber.com/. [Accessed 9 November 2019].

[29] Nanome, Inc., “Nanome,” [Online]. Available: https://nanome.ai/nanome/. [Accessed 9 November 2019].

[30] Bungie, Inc., “Destiny 2,” [Online]. Available: https://www.bungie.net/7/en/Destiny. [Accessed 8 November 2019].

[31] B. Lynch, “Steam VR Tutorial,” [Online]. Available: https://www.youtube.com/watch?v=Kg7gPiz8-SU. [Accessed 11 November 2019].

[32] S. Kubrick, Director, *2001: A Space Odyssey.* [Film]. United Kingdom / United States: Stanley Kubrick Productions, 1968.

[33] Oxford University Press, “Oxford Advanced Learner's Dictionary - "Field of vision",” [Online]. Available: https://www.oxfordlearnersdictionaries.com/definition/english/field-of-vision. [Accessed 14 November 2019].

[34] Pixmax, “Pixmax,” [Online]. Available: https://www.pimax.com/. [Accessed 14 November 2019].

[35] StarVR Corp., “StarVR One,” [Online]. Available: https://www.starvr.com/products/. [Accessed 14 November 2019].

[36] A. J. Parker, “OXFORD RESEARCH ENCYCLOPEDIA, NEUROSCIENCE,” May 2019. [Online]. Available: https://oxfordre.com/neuroscience/neuroscience/view/10.1093/acrefore/9780190264086.001.0001/acrefore-9780190264086-e-77. [Accessed 12 November 2019].

[37] M. Alger, “Visual Design Methods for Virtual Reality,” 17 October 2015. [Online]. Available: http://aperturesciencellc.com/vr/VisualDesignMethodsforVR_MikeAlger.pdf. [Accessed 14 November 2019].

[38] Facebook Technologies, LLC., “Oculus - VR Best Practices,” [Online]. Available: https://developer.oculus.com/design/latest/concepts/bp-vision/. [Accessed 15 November 2019].

[39] A. Alarakha, “Contrast Principle Of Design,” [Online]. Available: https://254-online.com/contrast-principle-of-design/. [Accessed 20 November 2019].

[40] B. M. '. Hart, H. C. E. F. Schmidt, I. Klein-Harmeyer and W. Einhäuser, “Attention in natural scenes: contrast affects rapid visual processing and fixations alike,” *Philosophical Transactions of the Royal Society B: Biological Sciences,* 19 Oct 2013.

[41] J. L. Poland, *Lights, Camera, Emotion!: an Examination on Film Lighting and Its Impact on Audiences' Emotional Response,* 2015.

[42] O. A. Adenuga, “Adapting Cinematic and Theatrical Lighting to Virtual Reality,” December 2016.

[43] J. Pratt, V. Petre, R. M. Guo and R. A. Abrams, “It's Alive!: Animate Motion Captures Visual Attention,” *Psychological Science,* October 2010.

[44] C. J. Howard and A. O. Holcombe, “Unexpected changes in direction of motion attract attention,” *Attention, Perception, & Psychophysics,* November 2010.

[45] S. Rothe, H. Hußmann and M. Allary, “Diegetic Cues for Guiding the Viewer in Cinematic Virtual,” *Proceedings of Virtual Reality Software and Technology,* 2017.

[46] The Los Angeles Film School, “THE IMPORTANCE OF SOUND - IS SOUND REALLY JUST AS IMPORTANT AS VIDEO QUALITY WHEN PRODUCING FILMS AND GAMES?,” 9 January 2017. [Online]. Available: https://www.lafilm.edu/blog/the-importance-of-sound/. [Accessed 20 November 2019].

[47] B. Nanay, “Synesthesia,” [Online]. Available: https://www.psychologytoday.com/intl/basics/synesthesia. [Accessed 20 November 2019].

[48] M. Lalwani, “Surrounded by sound: how 3D audio hacks your brain,” 12 February 2015. [Online]. Available: https://www.theverge.com/2015/2/12/8021733/3d-audio-3dio-binaural-immersive-vr-sound-times-square-new-york. [Accessed 20 November 2019].

[49] Sony Interactive Entertainment LLC., “Hellblade: Senua's Sacrifice - 360° Trailer,” [Online]. Available: https://www.youtube.com/watch?v=pBEtUnr38do&t=0s&list=PLU8wpH_LfhmuVhyt9XZU3DKXVlN4Pq2zh. [Accessed 20 November 2019].

[50] L. Paletta, N. J. Neuhuber, M. Schwarz, V. Wagner and K. W. Kallus, “Social Cueing in Human Attention within the Context of Orientation Tasks,” *Procedia Manufacturing,* pp. 3775-3781, 2015.

[51] S. R. H. Langton, R. J. Watt and V. Bruce, “Do the eyes have it? Cues to the direction of social attention,” *Trends in Cognitive Sciences,* pp. 50-59, 1 February 2000.

[52] E. Kissack, Director, *The Gunfighter.* [Film]. 2014.

[53] T. Zhou, “A Brief Look at Texting and the Internet in Film,” [Online]. Available: https://www.youtube.com/watch?v=uFfq2zblGXw. [Accessed 20 November 2019].

[54] Unity Technologies, “Unity.com,” [Online]. Available: https://unity.com/. [Accessed 30 November 2019].

[55] Unity Technologies, “Unity - Universal Render Pipeline,” [Online]. Available: https://unity.com/srp/universal-render-pipeline. [Accessed 25 November 2019].

[56] Facebook Technologies, LLC, “Oculus Rift S,” [Online]. Available: https://www.oculus.com/rift-s/. [Accessed 30 November 2019].

[57] Google LLC, “Google Text-to-speech,” [Online]. Available: https://cloud.google.com/text-to-speech. [Accessed 30 November 2019].

[58] Facebook Technologies, LLC., “Oculus Native Spatializer for Unity,” [Online]. Available: https://developer.oculus.com/documentation/audiosdk/latest/concepts/book-ospnative-unity/. [Accessed 25 November 2019].

[59] Interaction Design Foundation, “Visual Hierarchy - Your constantly-updated definition of Visual Hierarchy and collection of topical content and literature,” [Online]. Available:

https://www.interaction-design.org/literature/topics/visual-hierarchy. [Accessed 26 November 2019].

[60] S. Rothe, D. Buschek and H. Hußmann, “Guidance in Cinematic Virtual Reality-Taxonomy, Research Status and Challenges,” *Multimodal Technologies and Interaction,* 19 March 2019.

[61] (. n. u. rambodk, “Documentation to 360° Heatmaps,” [Online]. Available: https://github.com/Unity-Technologies/360-Video-Heatmaps. [Accessed 8 November 2019].

[62] HaptX Inc., “HaptX,” [Online]. Available: https://haptx.com/. [Accessed 12 November 2019].

[63] VR Electronics Ltd., “TeslaSuit,” [Online]. Available: https://teslasuit.io/. [Accessed 12 November 2019].

[64] HaptX Inc., “HaptX Press Images,” [Online]. Available: https://haptx.com/press-kit/. [Accessed 11 November 2019].

[65] O. Fong-Surdenas, “olandafs.artstation.com,” 30 July 2019. [Online]. Available: https://olandafs.artstation.com/projects/m4oD8.

[66] B. D. A. F. C. K. S. R. B. Rorik Henrikson, “Multi-Device Storyboards for Cinematic Narratives in VR,” *Proceedings of the 29th Annual Symposium on User Interface Software and Technology,* pp. 787-796, 10 2016.

[67] A. Coppens, “Merging real and virtual worlds: An analysis of the stateof the art and practical evaluation of Microsoft Hololens,” *Faculty of Sciences - University of Mons,* June 2017.

[68] Jaunt, Inc., “The Cinematic VR Field Guide - A Guide to Best Practices for Shooting 360°,” p. 7, 2017.

[69] Epic Games, Inc., “Unreal Engine,” [Online]. Available: https://www.unrealengine.com/. [Accessed 1 November 2019].

[70] Bungie, Inc., “Artist Reference Portal,” [Online]. Available: https://www.bungie.net/en/News/Article/11987/7_Artists-Reference-Portal. [Accessed 8 November 2019].

[71] Q. Féret, “Binaural Audio : How 3D audio hacks your brain,” [Online]. Available: https://arvrjourney.com/binaural-audio-how-3d-audio-hacks-your-brain-a3de0ceb4196. [Accessed 20 November 2019].

10. Acknowledgements

Thanks to K, K, V, V, C, T and T. ;)

11. Contact

Felix Kroischke – www.projectcv.de – mail@projectcv.de