Design of video player and user interfaces for branched 360 degree video

Design av videospelare och användargränssnitt för förgrenade 360 graders video

Martin Christensson
Mimmi Cromsjö

Supervisor : Niklas Carlsson
Examiner : Marcus Bendtsen
Upphovsrätt

Detta dokument hålls tillgängligt på Internet - eller dess framtidiga ersättare - under 25 år från publiceringsdatum under förutsättning att inga extraordinära omständigheter uppstår.

Tillgång till dokumentet innebär tillstånd för var och en att läsa, ladda ner, skriva ut enstaka kopior för enskilt bruk och att använda det oförändrat för ickekommersiell forskning och för undervisning. Överföring av upphovsrätten vid en senare tidpunkt kan inte upphäva detta tillstånd. All annan användning av dokumentet kräver upphovsmannens medgivande. För att garantera äktheten, säkerheten och tillgängligheten finns lösningar av teknisk och administrativ art.

Upphovsmannens ideella rätt innefattar rätt att bli nämnt som upphovsman i den omfattning som god sed kräver vid användning av dokumentet på ovan beskrivna sätt samt skydd mot att dokumentet ändras eller presenteras i sådan form eller i sådant sammanhang som är kränkande för upphovsmannens litterära eller konstnärliga anseende eller egenart.

För ytterligare information om Linköping University Electronic Press se förlagets hemsida http://www.ep.liu.se/.

Copyright

The publishers will keep this document online on the Internet - or its possible replacement - for a period of 25 years starting from the date of publication barring exceptional circumstances.

The online availability of the document implies permanent permission for anyone to read, to download, or to print out single copies for his/hers own use and to use it unchanged for non-commercial research and educational purpose. Subsequent transfers of copyright cannot revoke this permission. All other uses of the document are conditional upon the consent of the copyright owner. The publisher has taken technical and administrative measures to assure authenticity, security and accessibility.

According to intellectual property law the author has the right to be mentioned when his/her work is accessed as described above and to be protected against infringement.

For additional information about the Linköping University Electronic Press and its procedures for publication and for assurance of document integrity, please refer to its www home page: http://www.ep.liu.se/.

Martin Christensson
© Mimmi Cromsjö
Students in the 5 year Information Technology program complete a semester-long software development project during their sixth semester (third year). The project is completed in mid-sized groups, and the students implement a mobile application intended to be used in a multi-actor setting, currently a search and rescue scenario. In parallel they study several topics relevant to the technical and ethical considerations in the project. The project culminates by demonstrating a working product and a written report documenting the results of the practical development process including requirements elicitation. During the final stage of the semester, students create small groups and specialise in one topic, resulting in a bachelor thesis. The current report represents the results obtained during this specialisation work. Hence, the thesis should be viewed as part of a larger body of work required to pass the semester, including the conditions and requirements for a bachelor thesis.
Abstract

Branched video enables users to interact with the video content and to choose a unique path through that content. With the further development of Virtual Reality (VR) technology the integration of branched video in these systems becomes increasingly convincing. This thesis develops a 360 degree branched video player in Unity Editor and three different user interfaces (UI) tailored for the needs of branched 360 degree video content, both the design and implementation focus on the aspects of usability. To expedite development of the UIs, a preliminary user study was conducted to identify a promising design direction. From the study it was established that plain buttons with descriptive text attached, anchored in front of the field of view at the time of appearing, was to be preferred for the selection of the subsequent branch. This design was preferred both in a short motion picture and exploratory film setting. However, world-space anchored symbols as depicted to appear transporting the user to that world-space location showed promise in an exploratory video setting. Based on the result of the study and subjective feedback from study participants, additional features was implemented to the user interfaces. Lastly, further development of both features for the user interface, tools regarding the video player, and a follow-up user study are suggested.
Acknowledgments

We would like to thank our supervisor Niklas Carlsson for his guidance and support during the project. We would also like to thank Linn Hallonqvist and Madeleine Bäckström for their feedback and support.
# Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abstract</td>
<td>iv</td>
</tr>
<tr>
<td>Acknowledgments</td>
<td>v</td>
</tr>
<tr>
<td>Contents</td>
<td>vi</td>
</tr>
<tr>
<td>List of Figures</td>
<td>viii</td>
</tr>
<tr>
<td>1 Introduction</td>
<td>1</td>
</tr>
<tr>
<td>1.1 Motivation</td>
<td>1</td>
</tr>
<tr>
<td>1.2 Aims</td>
<td>2</td>
</tr>
<tr>
<td>1.3 Approach</td>
<td>2</td>
</tr>
<tr>
<td>1.4 Limitations</td>
<td>2</td>
</tr>
<tr>
<td>1.5 Report outline</td>
<td>2</td>
</tr>
<tr>
<td>2 Background</td>
<td>4</td>
</tr>
<tr>
<td>2.1 Interactive video</td>
<td>4</td>
</tr>
<tr>
<td>2.2 Branched video</td>
<td>4</td>
</tr>
<tr>
<td>2.3 360 degree video</td>
<td>5</td>
</tr>
<tr>
<td>2.4 Unity editor</td>
<td>6</td>
</tr>
<tr>
<td>2.5 Oculus</td>
<td>6</td>
</tr>
<tr>
<td>2.6 Immersion and presence</td>
<td>6</td>
</tr>
<tr>
<td>2.7 Usability testing: Thinking aloud &amp; System Usability Scale</td>
<td>7</td>
</tr>
<tr>
<td>2.8 Related work</td>
<td>8</td>
</tr>
<tr>
<td>3 Preliminary System Design and Implementation</td>
<td>10</td>
</tr>
<tr>
<td>3.1 Environmental setup</td>
<td>10</td>
</tr>
<tr>
<td>3.2 Branched video player</td>
<td>10</td>
</tr>
<tr>
<td>3.3 Shared UI elements</td>
<td>11</td>
</tr>
<tr>
<td>3.4 Standard button approach</td>
<td>12</td>
</tr>
<tr>
<td>3.5 Button thumbnail approach</td>
<td>13</td>
</tr>
<tr>
<td>3.6 Hotspot approach</td>
<td>14</td>
</tr>
<tr>
<td>4 Preliminary User Study</td>
<td>16</td>
</tr>
<tr>
<td>4.1 Methodology of pre-user study</td>
<td>16</td>
</tr>
<tr>
<td>4.2 Results of pre-user study</td>
<td>17</td>
</tr>
<tr>
<td>5 Improved Implementations</td>
<td>21</td>
</tr>
<tr>
<td>5.1 Position and number of UI</td>
<td>21</td>
</tr>
<tr>
<td>5.2 Visibility of reticle</td>
<td>21</td>
</tr>
<tr>
<td>5.3 Auto-generated selection</td>
<td>22</td>
</tr>
<tr>
<td>5.4 Timer bar</td>
<td>22</td>
</tr>
<tr>
<td>5.5 Hotspot approach with additional functions</td>
<td>23</td>
</tr>
<tr>
<td>Section</td>
<td>Title</td>
</tr>
<tr>
<td>---------</td>
<td>--------------------------------------------</td>
</tr>
<tr>
<td></td>
<td>5.6  Fade</td>
</tr>
<tr>
<td>6</td>
<td>Discussion</td>
</tr>
<tr>
<td>6.1</td>
<td>Pre-study result</td>
</tr>
<tr>
<td>6.2</td>
<td>Pre-study methodology</td>
</tr>
<tr>
<td>6.3</td>
<td>Future ideas and implementations</td>
</tr>
<tr>
<td>6.4</td>
<td>Wider context</td>
</tr>
<tr>
<td>7</td>
<td>Conclusions</td>
</tr>
<tr>
<td></td>
<td>Bibliography</td>
</tr>
</tbody>
</table>
List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Tree structure showing how a branched video is constructed. The arrows represent a video sequence playing and the circles the vertex, also called branch points.</td>
</tr>
<tr>
<td>2.2</td>
<td>Caption for LOF</td>
</tr>
<tr>
<td>3.1</td>
<td>Modelling video segments out of video clips and formation of branched tree structure. At the bottom of the figure displays an example of a segment, including the information forming the segment.</td>
</tr>
<tr>
<td>3.2</td>
<td>Standard button approach including two selection buttons, each with text associated with the content of the path.</td>
</tr>
<tr>
<td>3.3</td>
<td>Overview of the environment containing standard buttons.</td>
</tr>
<tr>
<td>3.4</td>
<td>Button Thumbnail approach including two buttons, each with an image of a preview from the video content of the path.</td>
</tr>
<tr>
<td>3.5</td>
<td>Overview of the environment containing four set of thumbnail.</td>
</tr>
<tr>
<td>3.6</td>
<td>Hotspot approach with an interactive symbol of two feet.</td>
</tr>
<tr>
<td>3.7</td>
<td>Overview of the environment containing two hotspots.</td>
</tr>
<tr>
<td>4.1</td>
<td>Score received from System Usability Scale, SUS, for the demo videos.</td>
</tr>
<tr>
<td>5.1</td>
<td>Visible reticle when hovering over panel with play and pause buttons</td>
</tr>
<tr>
<td>5.2</td>
<td>Standard button approach with added timer bar showing for how long the buttons are available</td>
</tr>
<tr>
<td>5.3</td>
<td>Counter in Hotspot approach tracking the number of found Hotspots in the current segment</td>
</tr>
</tbody>
</table>
The ever growing application use cases for Virtual Reality (VR) and the ingenuity that this technology enables are expected to take computer-generated experience to new heights. An increasing number of fields take notice of different VR applications and investigates the possibilities of integration in their respective fields. For video consumption in particular, the integration is apparent. By moving into three dimensions, the user can experience and interact with the video content in a completely new way than have been possible in the past. In particular, the three dimensional space of VR results in a set of unique options for novel user interface (UI) implementations. By investigating and exposing users to interfaces of different character, this thesis sets out to obtain knowledge regarding which UI elements works well in this new space.

1.1 Motivation

Different types of interactive video has been established in a two dimensional video setting for some time. Interactive video allow the user in various ways to decide on the sequence of events. This ability is presented to the user in a relatively simple and intuitive user interface that the majority of users will be familiar with. With interactive video in 360 degrees video, on the other hand, the possibilities for interaction are greatly deepen and can lead to significant changes in the experience. In a 360 degree environment, interactive elements can be used in infinitely many ways and for different purposes. One possible usage is branched video; thus, allowing the user to determine the path in a tree structure through the video.

Prior works have investigated the problems of how to best deliver either 360 degree video or branched video to the user without playback interruptions. Due to significantly larger file sizes, this types of videos typically are more difficult to successfully deliver when there is limited bandwidth available (see section 2.8). In this work we therefore assume a high-bandwidth scenario, and focus on the user experience in the case that the bandwidth is enough to deliver each path of the video content in the playback.

A useful case for future development in the field of interactive 360 degree video is to examine how user interface can be designed to provide the best possible experience for the user when watching a branched video. In general, there has been limited research developing complete
user interfaces and determine desirable as well as undesirable elements for interaction in this environment.

1.2 Aims

The purpose of this thesis is to create a video player for branched video in 360 degree environment, then to examine how different user interfaces can be implemented to contribute making branched video as pleasant experience as possible for the user. To evaluate the user experience, we perform a user study on interfaces to examine what users think of both the interfaces in its entirety as well as individual elements. This thesis aim is to:

- Create a 360 degree video player for branched video.
- Create different types of user interfaces and implement them in the video player. The user interface enable the user in ways to choose between different video paths in the branched video structure.
- Perform a preliminary user study to collect data of users feedback of different interfaces when observing demo 360 degree videos displayed inside a VR headset.
- Design and implement improvements to the user interfaces based on the data collected during the preliminary user study.
- Discuss and establish a general direction for user interfaces for branched 360 degree video.

1.3 Approach

To achieve the aims of this thesis, a branched 360 degree video player is designed in Unity (see Section 2.4). Three different types of user interfaces are designed and integrated into the player, allowing us to identify elements of the UIs as either positive or negative. To succeed in developing user interfaces that can be used for branched 360 degree video, we focus on the usability aspects of the UIs. Throughout the thesis we define usability by the same definition made by Nielsen [14]. This implies defining usability as multi-dimensional property with multiple components and attributes. The attributes associated with usability are: Learnability, Efficiency, Memorability, Errors and Satisfaction. To gain insight into the value of different design choices, a preliminary user study was conducted. This study highlights a potential general design directions and brings forward some less promising UI elements that can be ruled out of further development.

1.4 Limitations

In this work, we have developed a limited selection of user interfaces. Due to the limited time available, it led us to delineate us to design three different types of user interfaces implemented in two types of videos: short motion picture and exploratory video. These interfaces were presented to the participators of a user study and data regarding the general opinion and any annoyances the participant experiences with each UI were collected. The user study was performed as a preliminary user study considering that the scope of the study is limited to under 10 participants.

1.5 Report outline

This thesis is structured as follows. First, we present some background of relevant areas which is necessary to understand the following work. Then, in Chapter 3 we present a preliminary system design and implementation, how the branched video player and the user
interfaces where designed. In Chapter 4, we describe how the preliminary user study was conducted and present the results. In Chapter 5, improved implementations based on the user study is presented. In Chapter 6 we discuss the result and methodology used in the thesis. As well, suggestions of improvements and ideas for future work is stated. Finally, conclusions are presented in Chapter 7.
Background research has been made to assign information relevant to the work. The areas explained are interactive video, branched video, 360 degree video, Unity editor and Oculus. In additions the terms immersion and presence are explained, as well as two methods for usability testing. Lastly, related work is presented.

2.1 Interactive video

In linear media, such as TV, series and movies, the user is usually limited when it comes to influencing how the video is played. The user view the video in a passive manner where only a few and minor interactions occur. Such a video player usually only assists the user in being able to play, pause, fast forward, rewind and stop the video. With an interactive video, this property changes and the user becomes more involved in the experience. An interactive video enable to a greater extent input from the user, for example by the user interacting with objects in the image. Alternatively to watching video in a static fashion, you can use an interactive video which navigates through the video with options such as “next appearance”, “skip scene”, “previous scene” and “last event”. Hammoud [9] draws the following definition of an interactive video: “Interactive video is a digitally enriched form of the original raw video sequence, allowing viewers attractive and powerful interactivity forms and navigational possibilities”.

2.2 Branched video

Branched video is a form of interactive video and intend that the video you are watching is a non-linear, multi-path video. This implies that different paths can be taken in the video, making a variety of possible endings. This facilitates the user to interact with the video and consciously choose the path the users themselves desire. Branched video enables the user to be part of the creation of the experience and make it unique. Krishnamoorthi et al. [11] defines a multi-path, nonlinear form of video in which "non-contiguous fragments of video can be stitched together to form what we term a nonlinear video segment, and the video can include branch points at which there are multiple choices of which segment to play back".
2.3 360 degree video

In a 360 degree video, the camera is placed in the centre of the video; thus, giving the user the opportunity to view in all surrounding directions of the video. The user therefore access full control over which part of the video is displayed. The current section of space observed covered by the viewers field of view is called the view frustum [19]. The viewing frustum allows only a portion of the video to be rendered to the user. The ability to render a video in 360 degrees is given by the video clip being recorded with an omnidirectional camera or multiple cameras that simultaneously record in every direction, which is then put together into one single 360 degree video [20].

There are two types of 360 degree videos, monoscopic and stereoscopic. Monoscopic is the type of 360 degree video commonly used in, for example, Google Street View, Facebook and Youtube 360[1]. These are flat 360 degree renderings that can be seen on all types of screens or with a Head mounted display (HMD). In monoscopic videos, the camera can rotate around but the video lacks depth [18]. Stereoscopic and the other side are filmed with two lenses and can be viewed in 360 degree with a VR set. By creating 3D rendering of a 360 degree video by using a dual camera setup to compensate for the offset for each eye, the feeling of a virtual reality is enhanced when each eye will see pictures that differ slightly and add depth [18].

---

1http://www.immersionvr.co.uk/mono-vs-stereo
2.4 Unity editor

For ease of implementation of both 360 degree video, displaying a Graphical User Interface (GUI) as well as creating a virtual three dimensional (3D) environment, the Unity Editor can be used. It is a widely used editor which feature a wide array of tools, both for playing video and implementing functionality as well as customization of the GUI. The editor enable the ability to animate a visual 3D environment with the flurry of tools available. Unity also comes with a Software Development Kit (SDK) for Oculus (see Section 2.5), which makes Unity an appropriate choice for the purpose.

2.5 Oculus

To access a virtual 3D environment and watching 360 degree video, an Oculus Virtual Reality headset can be used. The rig consists of goggles with dual 1080x1200 pixel OLED displays for each eye and integrated headphones which deliver real-time 3D audio effect. The display is refresh rate of 90 Hz and a 110 degree field of view. The Oculus uses sensors called Constellation sensors which are tracking the position of the headset which is fitted with embedded infrared LEDs. This is then tracked by the Constellation sensor and fed to a workstation. One model is called Oculus Rift Consumer version 1 (CV1) and can be seen in Figure 2.2.

2.6 Immersion and presence

The general aim with VR is for the user to truly feel as if they were physically at the location where the clip was filmed; to create a sense of realism of the VR-experience to the user. To achieve this, the implementation has to reliably convince the user that what they are experiencing are truly real. To appeal to this idea, the VR-experience must stimulate the senses of the user the same way real-world stimuli would affect the user. Using the Oculus HMD and the workstation, it can affect the users visual as well as their auditory sensory experience.

---

3 https://unity3d.com/unity/editor
4 https://www.oculus.com/blog/powering-the-rift/
4 The picture is taken from https://www.tomsguide.com/us/oculus-rift-user-guide,review-3496.html
Slater [17] describes this fidelity of users senses of the VR-experience as *immersion* and *presence*. It is defined as following by Slater:

- "Immersion refers to the objective level of sensory fidelity a VR system provides."
- "Presence refers to a user’s subjective psychological response to a VR system."

The VR-experiences level of immersion depends heavily on the performance of the workstation and to what extent the rendering software can deliver a seamless, close to reality experience. Immersion is objective in the sense that it only has measurable parameters. Immersion is a spectrum of many factors, such as field of view, field of regard, display size and resolution for example. If the level of immersion is high, the *display fidelity* is high. The highest level of of display fidelity would be indistinguishable from the real world.

Presence on the contrary is the subjective physiological response. It is the sense of conviction by the user that what they are experiencing is closely related to reality. It is a feeling of being engulfed by the system. It is highly individual what response a unique user will experience subjected to the same VR event. The same user can also experience different level of presence at different space in time to the same VR event. This is due to the users state of mind, recent history as well as other factors [17].

### 2.7 Usability testing: Thinking aloud & System Usability Scale

Nielsen [15] explain *Thinking aloud* as a method that involves participants in a usability test to speak out loud while using the system to be evaluated. The method is considered to be a valuable usability engineering method and is increasingly used for the practical evaluation of human-computer interfaces.

The main advantage of the method is that, by the test participant putting words on their thoughts while actively using the system, it can provide a better understanding of the participant’s experience. The Thinking aloud method captures how users interpret each individual interface item, which gives an idea of what the participant experience as difficult or annoying. During the test, information regarding small irritants that would not show up in other forms of testing can be noted. The method can therefore provide qualitative data from a relatively small group of users [15].

The disadvantage of Thinking aloud is that it is not suitable for the purpose of receiving quantifiable results. The method can also be considered unnatural to apply as users may find it difficult to express themselves and their opinions throughout the test. This can affect the test by impairing the user’s ability to work with the system at its actual pace as the user is occupied having to think aloud. Even the user’s problem solving ability can be lowered when the user has several things to focus on simultaneously [15].

There are plenty of different methods for measuring usability for a system, one of which is the *System Usability Scale* (SUS). SUS was developed by Brooke [4] as a survey scale in order to easily examine the usability of a system or product. The SUS instrument is composed of 10 statements that are scored on a 5-point scale from strongly agree to strongly disagree. The 10 statements are then counted according to a certain method to finally become a score range from 0 to 100, where higher scores indicate better usability [2].

SUS has several advantages. First of all, it is flexible and can therefore be used in a wide range of interface technologies including computer graphical interfaces and web sites. SUS
Additionally has the advantage of being quick and easy to use for both participants and administrators of the study. Finally, the result is easy to interpret [2].

2.8 Related work

When watching a 360 degree video, the user has great flexibility when it comes to which direction you want to look in. However, this flexibility requires significant bandwidth consumption for content providers delivering these services over the internet. In previous work, Almquist et al. [1] made an empirical characterisation based on data from viewing sessions of head movements while watching different categories of 360° video. From the study, Almquist et al. found that the predicted head movement of the study participants varied greatly depending on the type of video shown to the participant. From the data collected, an optimisation model was implemented to explore the prefetching aggressiveness trade off more precisely. This provided solid ground to build discussion of measurement-based biases to increase the user’s quality of experience as well as further system optimizations.

Pio et al. [16] presents another way to tackle the same problem. They took well-known ideas and techniques that have been powerful tools in computer graphics and image processing and extended them to meet this new problems. This where done by remap equirectangular layouts to cube maps and let the original sphere be divided into 30 smaller pieces. Then by using view-dependent streaming, out of the full 360 degree frame, they only stream what can be seen in the current field of vision.

One problem that follows the concept of branched video is the freedom to choose their own path through the video easily results in playback interruptions. Krishnamoorthi et al. [12] have worked on developing a solution to the problem by creating an implementation of an interactive branched video player using HTTP-based Adaptive Streaming (HAS). Using a simple optimization framework, they design optimized prefetching policies that maximise the playback quality while ensuring sufficient workahead to avoid stall events.

In a paper written by Carlsson et al. [5] they introduce the concept of “multi-video stream bundling”, meaning multiple “parallel” video streams that are synchronised in time, where each stream provides a video from a camera filming the same event. Each stream is then delivered individually using HAS. Since each current stream is divided into different qualities using HAS, bandwidth can thus be allocated to the currently played stream and alternative streams that the user may choose to switch to next.

Within the area of multimedia Maugey et al. [13] have examined the concept of interactive multiview navigation applications which enable the users to freely navigate and choose their viewpoint. This means that depending on the users interaction they will be provided with different data. Maugey et al.‘s paper investigates a solution for the challenges consider compression performance and viewing quality trade-off, while also considering low delay cost and bandwidth to smooth out the transitions between different viewpoints for users. By developing a optimized model for the trade-off, Maugey et al. proposes an interesting model which presents promise in providing lower resource consumption and higher viewing quality. Hamza et al. [10] also proposes a model for improving the quality of the viewing experience in multiview video streaming. The model was focused on quality-aware rate adaptation on a virtual view distortion model. Their proposed model presents solid improvements in regards of user quality-of-experience and the quality of the rendered virtual view.

Early work was conducted on the topic of branched video, including on-demand streaming
protocols for non-linear media and the scalabillity of such streaming [21]. For example, by applying a tight lower bound to the server bandwidth, Zhao et al. [21] showed significant potential for reducing the required bandwidth for implementations using multicast delivery. Insights from the bound analysis showed that precise branch choice predictions can greatly reduce the required bandwidth. In the absence of precise model, a simple policy can be implemented which only listens to broadcasts from the current and immediately subsequent branch paths. Lastly, Zhao et al. develop different scalable protocols for delivering the non-linear media for both broadcast and stream merging.
Preliminary System Design and Implementation

To enable the comparison between different user interface elements, a system platform must first be established in order to base our examination of these elements on. To achieve this, both a physical setup with a HMD and workstation is required, and a software platform on which to implement a varied selection of UI elements. The UIs implemented took inspiration from different sources and were primarily focused on presenting a wide spectrum of different UI elements.

3.1 Environmental setup

For this implementation we used the HMD Oculus Rift CV1, running on a workstation with the following specifications:

- Windows 10
- Intel Xeon CPU E5-1620 v4 (@3.50 GHz)
- NVIDIA GeForce GTX 1080
- 32GB
- 2x USB 3.0 ports, 1x USB 2.0 ports

3.2 Branched video player

To be able to evaluate different user interfaces, we have developed a video player for branched 360 degree video, in which the interfaces will be applied on in order to be tested. In a 360 degree video the user is positioned in the centre of a hollow sphere on which the video is projected to its inside. The videos we have chosen to use are stereoscopic 360 degree videos. To get an interactive branched video, one or more linear video clips have been used by dividing them into segments. A segment is created by selecting a video clip to be used and determining the start and end time in the given clip which should represent the particular segment (see Figure 3.1). Then, a complete branched video can be created by using multiple segments and assemble them as a tree structure.
The video player we developed for branched videos was used by implementing the already integrated video player extension\(^1\) in Unity. We approached this by projecting the media files onto a render texture\(^2\), which is then applied to the Skybox material\(^3\) in the environment and then played by the Unity video player. To achieve the branched video structure a C# script was implemented to handle the transition between segments in a single or multiple linear videos. The segments are defined as objects in the script and has variable entries for each specific UI, but commonly entry such as video path, video clip, segments start- and end time as well as the IDs of the branching video segments.

![Diagram of video segments]

\[ A = \{\text{video clip: 1, start time: 15, end time: 75, branch 1: B, branch 2: C}\} \]

Figure 3.1: Modelling video segments out of video clips and formation of branched tree structure. At the bottom of the figure displays an example of a segment, including the information forming the segment.

### 3.3 Shared UI elements

All the user interfaces produced shared some elements which set a standard for basic features. This was a necessity for the user interfaces to function. These include C# scripts och graphical implementations to allow the user to access basic features of the 360 degree video player.

#### Gaze and click

Branched video in any dimension requires interaction from the user. The method of achieving this in a 360 degree environment without access to a conventional cursor was by using the gaze of the user. The gaze is a Unity Ray-cast\(^4\) projection which has its origin in the centre of the HMDs field of view. The ray cast is projected against collidable objects in the environment and detected as hovering over the object. This ray-cast projection is graphically represented to the user as a small dot in the centre of their field of view.

---

\(^1\)https://docs.unity3d.com/ScriptReference/Video.VideoPlayer.html
\(^2\)https://docs.unity3d.com/Manual/class-RenderTexture.html
\(^3\)https://docs.unity3d.com/ScriptReference/Skybox-material.html
\(^4\)Unity Ray-casting documentation: https://docs.unity3d.com/ScriptReference/Physics.Raycast.html
To perform a selection, a "click", the user has to focus its gaze on one of the collidable objects continuously for a set period of time. The period is variable, but for the purpose of standardising, the period was set to a constant two seconds throughout all UIs. To indicate this to the user, a radial progression bar is visible as soon as a selection is initiated. This radial is representative of the continuous gaze of the user, and as of executing a successful selection the progress of the radial is completed and removed from the view of the user.

**Play and pause**

In all implementations of our UIs, an element that remain constant throughout are the Play- and Pause buttons. These are located at the bottom of the field of view of the initial rotation of the player camera and remain anchored to this position in the environment throughout the entirety of all video segments. The buttons are white and smaller in size and mostly hidden out of view, both for ease of access to the user with the consideration to not be distracting and taking away from the user feeling a sense of presence in the environment. These buttons are highlighted in a bright red colour while hovered over.

### 3.4 Standard button approach

This user interface consists of two square, white *selection buttons* which are located next to each other. The buttons are attached to a static, invisible canvas\(^5\) in the environment. These buttons are used to choose between the two selectable paths and determines which segment should be the next to be played. Attached to the buttons are a descriptive text of the content of the next segment to aid the user in choosing between them. How the standard button approach looks from the user’s perspective can be seen in Figure 3.2 and an overview in Figure 3.3. When the buttons are hovered over by the gaze of the user, they highlight in bright red to indicate initiation of the selection process.

As mentioned above, the play and pause buttons are always visible throughout the entire video while the selection buttons are only visible at a set time before the branch point. If the segment reaches its end without the user having made a choice the video will pause and wait for a choice to be made.

---

\(^5\)https://docs.unity3d.com/Manual/UICanvas.html
3.5 Button thumbnail approach

The third interface is based on the same concept as the first interface but has two significant differences. Equally as the Standard button approach, the interface has a play and pause button that is always available to the user. The first main difference is that the two selection buttons here are replaced by two thumbnails that the user can interact with, can be seen in Figure 3.4. These thumbnails are image previews of what the next two selectable segments contains. The user selects the path by interacting with the thumbnail they invoke to be displayed next. The second aspect is that the interface with the two thumbnails is shown in four cardinal directions, thus a set of the thumbnails is shown at four positions around the user, which can be seen as an overview in Figure 3.5.

![Figure 3.3: Overview of the environment containing standard buttons.](image)

![Figure 3.4: Button Thumbnail approach including two buttons, each with an image of a preview from the video content of the path.](image)
3.6 Hotspot approach

This interface is based on symbols, here called Hotspots, representing the different choices for the user. In each segment, one or more hotspots will be anchored in the environment in a way that it is perceived to the user as being part of the video. If the user chooses to interact with any of the hotspots, it transfers the user to a new segment of the video. These hotspots will be anchored at unique coordinates, individually customised for each specific segment. The location should be that it fits into the environment, for example on a door or a road. In this interface there is, as with Standard button approach, no requirement for a choice to have been made before the segment ends, the video will be paused until the user has completed a choice of a segment. In Figure 3.6 the users view of the Hotspot approach is shown and in Figure 3.7 an overview.
3.6. Hotspot approach

Figure 3.6: Hotspot approach with an interactive symbol of two feet.

Figure 3.7: Overview of the environment containing two hotspots.
To evaluate the preliminary system design and implementation that have been created a preliminary user study was performed. In this section the methodology and results are presented. The purpose of this study is to gain knowledge of how users experience the user interface implementations that have been made and to identify the most promising directions for further development.

4.1 Methodology of pre-user study

The experiment involves participants watching several branched videos that have different interfaces to interact within the video. Prior to the study, the participants were informed of the basic concepts of branched video and 360 degree video. The purpose was that the participant should be well understood in the technology to be used. Later a test video was shown to the participant to familiarises with our setup. During the entire user study, a thinking aloud protocol (see Section 2.7) was used and the participants were encouraged to comment at any time during the test. The thinking aloud approach was used to get information about what the user was experiencing as they watched the demonstrations. This is likely to provide a good understanding of the user’s cognitive process and the users experience while interacting with the system[8].

The participant conducted a total of four tests where for each test the participant was instructed to watch a demo video and explore how the user interface works. Each demo consists of one of the three UI implemented in one type of video. The types of videos used in the pre-study is short motion picture (story-driven videos where the attention is moving around the 360 degree sphere) and exploratory video (the user is expected to explore the scene in a still environment). The four demonstrations where as following:


2. Thumbnail in SMP: Button Thumbnail approach implemented in a short motion picture.

4.2 Results of pre-user study

After each demo video, the system’s usability scale (SUS) described in Section 2.7 was used to collect data to later be able to examine the participant’s experience. As a reminder; SUS is the most well-known questionnaire used in UX research (see Section 2.7). It consists of a 10 questionnaire with five response options for the participant; from strongly agree to be strongly disagree [4]. The highest SUS score of a test is 100 and lowest 0. These scores are then converted to the scale; excellent, good, okay, poor or awful.

In addition to Thinking aloud and SUS, supplementary questions were asked to answer the concerns that existed regarding differences in the interface. The participants were asked to motivate and elaborate on the overall experience; comment on positive, as well as negative, experiences being exposed to the different user interfaces. The aim of these post-test questions were to get the participant confirming and motivating the result of the SUS result. If general consensus around an aspect or function of a certain UI was uniform, a conclusion for a specific SUS score can be attested or disregarded. This approach of open questions suited our purpose of receiving early initially feedback and could aid further implementations by identifying desirable elements and functions of the interfaces. The questions could likewise indicate undesirable elements which narrows the spectrum of possible implementations.

Participants

Eight participants were recruited to be subjected to the study. All the participants were Swedish university students with an average age of 23 years. The gender demographic of the eight participants were seven identifying as men and one as women. Four out of the eight participants had previously been in contact with VR and AR systems before. Two of the participants had experience with the Oculus HMD in particular. All eight participants had tried or understood the concepts and tasks required with interacting with branched video.

4.2 Results of pre-user study

Figure 4.1 shows the SUS score obtained for each of the four demo videos.

Standard in SMP

Demo video Standard in SMP received a mean SUS score of 84. Out of the eight participants six thought it was excellent, one poor and one good (see Figure 4.1a). According to the observations, about half of the participant had difficulty finding the selection buttons when they appeared. It could take several seconds after the video segment ended before they started to rotate and look around to locate the position of the buttons. Some participants noticed that the selection buttons always appeared where the play and pause buttons were placed which made it easier for these participants to keep track of whether the selection buttons were visible. During the test, the system experienced a severe drop in frame rate due to unknown causes at a certain point in a single segment directly following a vertex.

Two participants expressed during the test that they were confused regarding if the selection buttons where visible and when they would appear the next time.

Thumbnail in SMP

Demo video Thumbnail in SMP received a mean SUS score of 76.5. Out of the eight participants three thought it was excellent, three good and two poor (see Figure 4.1b). According to observations, some of the participants had difficulty detecting how the sets of thumbnails appeared in every cardinal directions. Other participants noticed the displaying of the four sets
4.2. Results of pre-user study

![SUS score graphs for different conditions](image)

Figure 4.1: Score received from System Usability Scale, SUS, for the demo videos.

of thumbnails quickly but needed some time to understand that all sets where the same. One of the participant asked if all thumbnails with the same appearance would generate the same option. As with demo Standard in SMP, the system experienced severe drop in frame rate at the same point in a single segment right after a vertex. The cause of this issue is unknown.

**Standard in exploratory**

Demo video *Standard in exploratory* received a mean SUS score of 84. Out of the eight participants five thought it was excellent, two good and one poor (see Figure 4.1c). From observation during the test, the participants were comfortable with the selection buttons from the exposure to the buttons in *Standard in SMP*. All of the participants were successful in completing the task provided. Altogether the participants had similar feedback of the selection buttons as in demo *Standard in SMP* regarding its position.

**Hotspot in exploratory**

Demo video *Hotspot in exploratory* received a mean SUS score of 79.4. Out of the eight participants five thought it was excellent, one good and two awful (see Figure 4.1d). According to observations, all participants interacted with a hotspot rapidly before the video segment was over. This left some of them unaware of the second hotspot in the scene. One participant verbally expressed difficulties seeing the hotspots when they were placed far away from the centre of the video, making the hotspot perceived smaller. Another participant disclosed it was unclear what the selection of a hotspot would yield. A third expressed overwhelming positively and excitement watching the demo.
Summary of pre-study results

The result of SUS reviled that the two demo videos; Standard in SMP and Standard in exploratory received the highest score, both gained a average value of 84. Second comes the demo video Hotspot in exploratory with SUS score 79.4 and last demo Thumbnail in SMP with score 76.5. This suggests that the participants preferred using Standard in SMP and exploratory in their entirety.

Out of the related questions we added, significant insights was gained regarding the different implementations that had been made. The majority of the participants expressed that one set of selections buttons displayed at the same time is the optimum number, because it is most easily comprehended by the user. Although positively was conveyed by some participants, having several duplicates of the UI element displayed at the same time, it became increasingly complicated to comprehend for first-time users. The participant expressed that the advantage of several UIs was that you discover it quickly without the need to rotate unnecessarily.

When we asked if the participant preferred Standard button approach or Button Thumbnail approach in a short motion picture, every participant unanimously agrees that Standard button is the better option. The participants described that the text attached to the selection buttons in the Standard approach clarifies what the outcome of choice will generate.

In all of the implementations, the video paused when a segment reached the branch point, waiting for the user to actively select the next segment. The selection buttons in the Standard button approach had then been visible to the user for five seconds while the video still played. Three participants replied that they prefer a choice to be made for one, randomly, when the segment reaches its end. Otherwise it removes the “flow” of the video if it is often paused. At the same time, three participants answered that they prefer the video to be paused so that you are unable to miss the opportunity to make the choice. The other participants were unsure. All participants agreed collectively that a time frame of five seconds prior to a segment ending, was an appropriate window for the selection buttons to be visible to the user.

The overall comments from the participants suggests that hotspots suited the exploratory video settings. The participants perceived that it was a quick and easy to use way to be transported between different connected places. Participants mentioned that it is important that the hotspots are large enough in size to be noticeable to the user. Some of the participants mentioned that to take full advantage of the variable location of the hotspots they have to indicate to the user that the hotspots have a importance of location, that you are “teleporting” or “walking” over to that specific location in the environment. Some participants wanted more hints about what the hotspot led to.

When asked if the participant preferred the Hotspot approach or the Standard button approach in exploratory videos, the majority of participants answered that Standard button approach made it easy to find the selection buttons in the scene. On the contrary, hotspots encouraged the user to explore further and imply where the selection transfers the user.

Regarding if the videos should contain visual or/and auditory feedback linked to buttons and interactions with the UI the participants were quite satisfied without it. One participant mentioned it would be beneficial if a timer was displayed to indicate that the selection buttons are visible and for how long it will remain in view. The user could possibly benefit from indications, e.g, arrows, directing the user to the location of hotspots in the scene.

Overall the participants were satisfied with the way they interacted in the system using
the reticle along with the radial progression bar to indicated that a selection was performed. Several participants expressed that the reticle did not always needed to be visible in the users field of view, it is only necessary to be visible when selection of a new segment is possible. Only a single participant preferred a physical controller rather than using the gaze to interact. The majority thought the reticle was of proper size but the colour might have been beneficial to change to an alternative colour.

The participants stated that they would prefer the selection buttons to be displayed in the direction the user is watching at the time when the selection buttons first appears in the picture, instead of being locked in one place or being in four cardinal directions. They suggest that it would be helpful discover the buttons and be aware of its position.
5 Improved Implementations

From the user study we received valuable information about which of the demo videos the users favour. Information was collected regarding the participants thoughts of the UIs and the specific elements that define them. With the feedback, we are now informed of the direction in which the development of the video player and UI may best proceed. Therefore, the development has continued and new implementations have been carried out which is presented below. From the result of the pre-study we have decided to continue developing Standard button approach in SMP and Hotspot approach in exploratory video.

5.1 Position and number of UI

The most significant remark the participants had regarding Standard button approach and Button thumbnail approach culminated in the desire to only have one set of buttons or thumbnails visible. Furthermore, the participants requested those buttons to be visible in the direction the user is looking, making them immediately detectable. This has been improved and implemented such that one set becomes anchored in the direction the user is looking at, using the forward vector of the camera and a Quaternion look rotation at the moment of the first frame the selection buttons appear. The distance from the camera of the selection button is a fixed distance, variable in setup. The rotation and direction of the camera at that frame decides the anchoring point of the selection buttons which remain static until the playback of another segment.

5.2 Visibility of reticle

From the pre-study it emerged that several participants expressed that it was unnecessary having the reticle visible throughout the whole video when using Standard button approach. Given this information, the visibility of the reticle has now been altered when Standard button approach is used. The reticle has been implemented such that it is only visible at the branch point, when the buttons are available and a selection can be performed. When selection buttons are not visible, the play and pause buttons are still available for the user. To enable

1https://docs.unity3d.com/ScriptReference/Transform-forward.html
2https://docs.unity3d.com/ScriptReference/Quaternion.LookRotation.html
5.3 Auto-generated selection

Since the participants had different opinions whether an automatic selection should be performed after a set time, such a function has been implemented. In the video player it is possible to choose whether this function should be enabled or not. If the function is active, at the end when reaching the branch point, a selection is generated for the user out of the two choices that were visible in the interface. Therefore the video continues to play without any interruptions if actions is absent from the user.

5.4 Timer bar

One participant mentioned during the user study that the addition of further indication of how long the selection buttons will remain visible would be convenient. Therefore, a timer bar has been implemented in the Standard button approach. This timer bar is progressively filled up, from left to right, during the time that the buttons are visible and indicates how much time the user has left to perform the selection. The time for how far in advance the selection buttons appear ahead of the segments end is variable within the video player. The additional implementation can be seen in Figure 5.2.
5.5 Hotspot approach with additional functions

From the pre-user study, it became apparent that the participants appreciate the concept of hotspots but expressed some dissatisfaction with the implementation as a result of the lack of information transparency the hotspots gave. The satisfaction of the users would increase with additional information or indication attached to the hotspot to hint at where the user would “transfer” subsequent to selection. To solve this issue, an implementation has been made which adds a brief descriptive text to the hotspot. To maintain the property of being a UI that encourages the user to explore, this text becomes visible beside the hotspot only when the user has hovered over the hotspot with the reticle. The text then disappear when the user hover over the hotspot again. The participants also asked for tools to help find every hotspots in the current segment. The solution implemented to the UI to combat this concern, was the use of a hotspot counter. The counter is placed in front of the user, above eye level, and indicates how many hotspots in the current scene that the user has hovered over. The visual implementation of the counter can be seen in Figure 5.3.

5.6 Fade

From observations made during the test, certain confusion among some of the participants of the preliminary study arose when selecting a new path from a branch point and the video player was loading a new path into playback. In the implementation used in the preliminary study, the change of path was practically instantaneous to the user and caused some confusion whether the interaction from the user was successful. The implementation of a fading of the display during these path changes could serve to clarify to the user that they were successful in their interaction at the branch point and a new path has been loaded into playback. The fade implemented currently dims the screen to a complete black then back to full brightness at a linear pace which can be varied. During this time, the playback of the video is paused and then resumed at the time of the screen reaching full brightness once again.
6 Discussion

From the preliminary user study and methodology surrounding the execution there exist a lot of uncontrolled variables due to the limitations of this thesis. We discuss potential pitfalls of the result of our study and what implications might have been produced. Secondly, discussions of further continuation of the revised UI implementations for the 360 degree branched video player is being presented. Moreover, we discuss additional tools to enhance both the viewing- and the authoring experience of the 360 degree branched video player.

6.1 Pre-study result

After the preliminary user study, we found the greatest success with the Standard button approach in short motion picture and Hotspot approach in exploratory video. This is based on that demo Standard in SMP received the highest SUS score which suggests that participants considered this demo having the highest level of usability. This conclusion is also supported by the majority of the participants expressing that it had the optimal number of visible sets of buttons and they preferred to have selection buttons with text in front of thumbnails. The Hotspot in exploratory on the other hand received a lower SUS score compared with Standard in exploratory, which both are demos with exploratory video. Despite this, we see the potential of using hotspots in exploratory videos compared to standard buttons. This because we received a lot of positive feedback from some of the participants who thought hotspots helped them making the experience more vivid, which we value highly in this type of video because we consider it improving the experience. We believe that the users that gave "Awful" as their SUS score for the hotspots felt that the implementation was not complete enough. We believe these users saw shortcomings in the implementation and not in the concept of using hotspots.

Insights of cross implementation of user interfaces in additional video types

Due to the limited scope of our work, we could not examine our three user interface in more than two different types of videos, short motion picture and exploratory video. Based on the results we obtained from the pre-study, we can come to insights about whether the user interfaces would be suitable for additional types of 360 degree video. The video types we have in mind are static focus video (main focus at the same location in the video) and rides (a
virtual ride in which the camera is moving forward).

From the result for *Standard button approach*, the user interface seems to work well regardless of the video type since the interface was perceived well by the study participants in both short motion picture och exploratory video. We believe that the interface is conceptually complete and the UI fits well into most video settings. Therefore, the user interface could be used in static focus video as well as video rides because of its simple but clear design. Hotspots, on the other hand, is harder to draw general conclusions regarding because it is more dependent on the individual video. Use cases for hotspots suit *static focus video* when used for choosing doors or other objects of interest in the video. The initial conceptual design characteristic of hotspots encourages the user to explore the scene around them. This causes a conflict with static video where the user’s focus should be directed to one place. The same reasoning applies to the video type *rides*, where the moving nature of the video makes it difficult for position of hotspots. However, we do not exclude the implementation of hotspots in all videos of the type *rides*, but we can not see apparent potential use cases.

### 6.2 Pre-study methodology

**Thinking aloud and SUS**

In the preliminary user study the *Thinking aloud* method was used. The purpose of using Thinking aloud was to get immediate feedback from the user whilst the system was under use. The advantages of Thinking aloud was clear during the test, and we could benefit positively from the unstructured data which was produced by the participants using this method. We received important constructive feedback on multiple details of the interfaces. The method prompted users to comment on both minor and major elements in the interface that they themselves experienced as positive or negative. By us listening to what the participants communicated and observing their actions, the method also contributed to a better understanding of what the participants experienced in regards of non-usable aspects of the UIs.

Besides Thinking aloud the System Usability Score, SUS, was used during the pre-study. SUS presented useful and gave a result that was easy to interpret. The result from SUS could help perceive the participants thoughts about the systems in their entirety after they had some time for consideration. However, the weakness of the SUS scale is that it does not declare anything about what the participants thought was good or bad. Nor does it provide any basis for us to analyse what should be change in the interface.

When conducting the preliminary study there is also a responsibility on the people who are the acting as moderator and secretary of the study. To get as reliable result as possible, these should try not to have an affect on the participants response. During the Thinking aloud portion of the study, the moderator asked the user to verbally communicate their thoughts during the test. Furthermore, no questions or requests from the moderator were made during the test to avoid any possible influence. The added questions that we asked the user were formulated without being suggestive. There were open questions as well as comparative questions where the user got the answer which option it preferred, this to ensure minimal impact. The secretary’s role was to record the users’ comments from Thinking aloud as well as answers to questions that we asked. When some users commented a lot during the test, it was difficult for the secretary to be able to write down all the quotes from the user, but instead had to write down a summary. This may have contributed to the answers being interpreted incorrectly. Similarly, the secretary could interpret the content of what the participant was expressing or the participant themselves could experience difficulties communicating the exact message they were trying to utter.
Another area which might have caused influence of the result was by adding multiple features which differentiated standard button approach and button thumbnail approach. With similarities in design and both being implemented and presented in a short motion picture to the participants of the preliminary user study, multiple features differentiated the UIs. The approach of first showing the user a single set of selection buttons, then enabling the feature to receive four replicas of the set of selection buttons in all four cardinal directions; for both the standard button approach and the button thumbnail approach could be beneficial. This could have clearly separated the exact feature of having a descriptive picture from the feature of having four replicas surrounding the user. However, in the supplementary questions the participants were asked to specifically describe any particular feature of any UI they disliked. Additionally, the specific question were asked to the participant whether they preferred a single UI or multiple replicas. This convincingly gave a distinction of which features was desirable of each of the UIs.

The effects of immersion and presence on the result

As presented in Section 2.6, the overall aim of a VR system is to deliver a experience to the user that convinces their senses that what they are experiencing are truly real. However, the preliminary study did not primarily focus on delivering this but instead to get initial data on the subjective opinion on UI elements. Nonetheless, we are aware that immersion and presence can have major implications on the result of the study. As far of aspects that may have caused a disruption in the users level of presence, we have identified two major potential causes from the preliminary user study: the absence of sound and the branching structure of the videos. Firstly, the absence of sound related to the content of the video might detach or cause disinterest from the user in the video content and ultimately result in harming the overall impression of the VR system. Similarly, we also observed disinterest from the participants in careful selection of the paths. This might be due to the nature of the video, which originally was not intended for use in an interactive setting, but instead was a originally a linear motion picture. To achieve a non-linear interactive video, the overall feeling of the video experience was disjoint. We did not compare how different branching structures or videos could affect the result or if it would contribute to an increase in the level of presence the user experienced. We also experienced a severe drop in frame rate in the Short motion picture, for both the Standard button approach and Button thumbnail approach at a single certain point in the transition between two specific segments. This have caused a negative influence in the users level of immersion, but it most likely did not have any significant effect on the result of the preliminary user study, since the error was brief and did not impact the function och visual appearance of the UIs.

6.3 Future ideas and implementations

Considering the lack of previous research, further work is required for determining the most suitable implementations for the user interface. It would be beneficial to do subsequent studies on a wider test group and with additional implementations. Below are presented suggestions on how a user study could be performed based on results and experience gained in the preliminary study.

User study

First of all, a larger number of participants in the study would give a greater basis to base further development of the UI. It would provide a more quantitative study which generates
6.3. Future ideas and implementations

a result closer to reality. If more participants are used, demo videos can also be examined in
different schemes to get a more fair assessment. Otherwise it could be that the participants
thinks that the first UI is more difficult than the last because it collects experience during
the test. Furthermore, the study would have to include a larger breadth of participants.
Participants in a wider range of ages can provide a result that would better reflect future
users as branched 360 degree videos may be used by a broad audience.

In the pre-study deliberately omitted the measurements of the time it took for each par-
ticipant to perform a task and how high the failure rate was. In a future study it would
be interesting to investigate these parameters, quantify how many of the participants who
manage to fulfil the task as it is assigned, e.g., find all the hotspots in the scene, and the
elapsed time. It could reveal information regarding how user-friendly the UI is, how natural
it is to use and how well integrated it is.

Just as in our preliminary user study, the new implementations can be examined in a
similar way, e.g., timer bar, text to hotspots, hotspot counter, reticle and the position of the
UI. In order to gain a deeper knowledge of the participants views regarding individual parts
of the UI you would have to do more demo videos where in each demo, only one detail at a
time changes. To see what participants think about the time when the buttons is visible, you
would have several demos where you, for example, show the buttons for 5, 7 and 10 seconds
and then let the participants take a position in which it prefers. Regarding if an automatic
selection of path is to be performed and if interactions with the UI should be done using
the gaze or an traditional hand controller. All of these options should be implemented in a
separate demo video. The question was presented in the preliminary user study, but exposed
great differences of opinion so it poses interesting aspects to investigate further.

System visual and auditory feedback

Although all the users in the preliminary study had previously been exposed to VR, the ex-
tent of their experience using a VR system was not accounted for in the preliminary study.
We believe that the level of usability will increase with the level of confidence and certainty
of the user. Part of the confidence comes from the habit of using the system, but a major factor is
the feedback that the user receives in using the system. The branched video player gave the
user limited visual feedback with only a filling radial indicating that a gaze was to be main-
tained over an interactive item in the scene (see “Gaze and Click” Section 3.3), whilst giving
nothing in regards of auditory feedback. With the absence of physical controls, only visual
and auditory feedback is enabled in the VR system and has to make for a satisfactory expe-
rience alone. Our belief is that visual and auditory indicators of interactions in the system
will benefit the users confidence in using the system. Although none of the participants was
unable or had any difficulties in quickly learning the procedure of “clicking”, no comparison
could be performed and measured. Another element of the UI which was implemented after
the first iteration was the Timer bar (see Section 5.4). To more convincingly imply the meaning
and function of this element to the user, some auditory feedback could be complementary.

Authoring tool

The approach currently used to create the branched video in our 360 degree environment is
considerably complicated and not suited for the majority of users, demanding prior experi-
ence in both C# and Unity. Therefore it would be advantageous to design and implement a
video player for 360 degree video that allow interactive video to be stitched together in ways
that allow users to interactively select different non-linear media paths through the media.
The tool should allow the creator to author a segment, without knowledge in programming,
by selecting the desired video clip as well as the start and end time that will result in a
complete segment. Subsequently, the creator could, in a graphic representation, place the segments in the desired order, creating branched tree structures.

Since information about segments are stored in an C# object, the parameters required for each UI is structured and can easily be assigned with e.g. a well structured metafile. However, challenges lies within the individual parameters needed to correctly describe each UI. For example, the Hotspot button approach needs information about the location and amount of hotspots in a particular segment, while this is not in case needed for any other UI we have created.

6.4 Wider context

The use cases for VR, and 360 degree video in particular, are almost endless. By integrating a non-linear structure to these videos a wide variety of decision making training in a simulated real world environment can be achieved. For example, the technology could be well suited to use in some respects in education. By letting users experience an event in something that is similar to a real situation, it can contribute to better learning, e.g. a historical event or how to manage an emergency situation. VR and interactive video could as well be an appropriate tool for learning how to drive a car, as it removes a risks considering being a new driver is a danger to himself and others.

Although we can see several advantages of the technology, there can be negative consequences in the future. As the technology can be used for education, it would be possible to use it for the training of unwanted and unethical areas such as warfare. We can also speculate that VR can have a negative impact on people’s social skills as the user does not have to interact with others and they can be immersed by an imaginary world. Research has also shown that VR can cause anxiety and paranoia in users with both pre-existing mental illness and not [7].

With additional research and improvements to VR in general but also 360 degree video streaming, the demand will likely increase from users. With the future seemingly offering online streaming services for 360 video degree video the challenge to combat the significantly larger size of these videos than regular 2D videos. Although the user only can view a small portion of the video through its viewing frustum, the entire 360 degree video must be streamed which can be a huge constraint on the limited bandwidth. To enable streaming over the internet, an optimized prefetching algorithm can be implemented to efficiently use the limited bandwidth. This, we believe, will be the most considerable hurdle when bringing the technology of 360 degree video to a wider audience and more readily available.

From our work, we present a first step in introducing the means to accomplish a easy-to-learn, usable user interface for this structure. We believe that the context of this work can lay ground to further develop a general consensus around user interface standards in this form of 360 degree video. Due to large differing in the UIs, the scope - and the few participants of the user preliminary user study we recognise the limitations of our work. However, this work should suffice as a qualified initial stepping stone to further development.
Conclusions

In this thesis we propose a method for developing a 360 degree branched video player in Unity. The player utilised already existing functionality for 360 degree video playback integrated in the Unity editor, along with C# scripts produced to enable a branching tree structure. The aim of this thesis was to investigate different user interfaces to enable the user to select a path they desired in a 360 degree branched video environment. With the 360 degree branched video player as base, different UIs were produced with identical task-related functionality but presenting diverse and particular graphical elements. Subsequently, a preliminary user study was carried out utilising System Usability Scale (SUS) and Thinking aloud methods as suggested by Nielsen [15]. As a result of the study, the four main findings were:

- Selection buttons in the form of plain buttons with descriptive text attached was most successful, in the context of a short motion picture. When applied to an exploratory 360 degree video, the result showed it achieved higher SUS score than the Hotspot approach in the same type of video, but subjective feedback exhibited more promise from the Hotspot approach.

- All participants preferred the buttons anchoring at the centre of their field of view at the time of first becoming visible. Overall feedback on a fixed anchoring position in the environment was negatively perceived.

- The Hotspot approach received a considerable amount of positive subjective evaluations. This UI shows a lot of promise from the comments made by participants. However, the implementation of this UI requires thorough thought and the anchoring location in the environment must remain purposeful to appeal and entice the user. The user should be provided information about the number of hotspots in the current segment since these are not necessarily immediately visible to the user.

- Selection buttons with attached descriptive images (thumbnails) was not received well by the participants in the study, where it was not experienced sufficiently descriptive for its purpose. Selection buttons with descriptive text showed similarly traits, but was according to the study a more suited option for the implementation.

From the result of the study, implementations of a timer at a set time before the branch point, a graphical counter for number of hotspots found in the scene, a visual fading effect for the
transfer between segments and some options about the visibility of the reticle was implemented. From this further development we suggest some future work that can be based on the results. We suggest further iterations of user studies and subsequent implementations and improvements can be made to the 360 degree branched video player. We also suggest an authoring tool for the creation of the branched video that can be attached to the player for ease of use.


