Design and Evaluation of Guiding Methods for 360º videos

Tejal Mate

University of Tampere
Faculty of Communication Sciences
Human-Technology Interaction
M.Sc. thesis
Supervisor: Prof. Markku Turunen
December 2018

University of Tampere
Faculty of Communication Sciences
Human- Technology Interaction
Tejal Mate: Design and Evaluation of Guiding Methods for 360° videos
M.Sc. thesis, 68 pages, 2 appendix pages
December 2018

360° video is an emerging trend which is gaining popularity due to immersive nature. These videos can be watched on conventional displays as well as the Head Mounted displays (HMD). In contrast to the traditional 2D videos, in which entire video is displayed within the user's field of view, 360° videos always have content that is outside the user's field of view. Viewing 360° content on HMD gives the user an immersive feeling, the user feels like being in the recorded situation. This is a key feature of 360° videos. The content outside the current field of view can be viewed on HMD by turning head or body. However, there is always a possibility of getting disoriented as the content is present all around. Disorientation disrupts the user's immersion and some interesting event or object can be missed. This problem can be solved by giving guidance to the user to navigate in a 360° video.

Several studies have proposed solutions to the problem of missing the interesting content and getting disoriented. This thesis proposes a novel solution using a research through design approach to design a set of guiding methods. These guiding methods use visual and audio cues such as an arrow marker, a bull's eye mark, and speech synthesizer audio to guide the user while navigating in a 360° video. A software prototype has been developed which uses different combinations of the designed guiding methods with different videos from YouTube. A novel concept of using 3D audio cues with treble and bass elements is introduced as audio guidance for navigation in a 360° video.

The software prototype was tested with 22 users with a two-phase approach, where guiding methods were redesigned with changes suggested by users from the first phase and in the second phase the revised version was tested again with users. The thesis presents design recommendations for implementing guidance methods with focus on overall experience, usefulness and distraction. The choice of the guidance method depends on the content of the video. The guidance should be subtle, unobtrusive and non-repetitive. The audio guidance should be distinct from the content audio.

Keywords: 360° video, Guiding methods, Field of View (FOV), Hotspots, Interesting events or objects, User experience, Immersion

Acknowledgements

Writing this master’s thesis has been a very long journey. I have received a lot of help and support from many people during my journey whom I would like to thank.

First, I would like to thank my supervisor Prof. Markku Turunen for his guidance, encouragement and patience throughout the thesis process. I take this opportunity to thank Dr. Jaakko Hakulinen for his valuable suggestions in the prototype development and detailed review of the thesis. I would like to thank my colleague Ilkka Rönkä for the collaboration in this research. TAUCHI research group kindly helped me with the devices needed to develop and test the research prototype

I would like to thank my husband Dr. Sujeet Mate for his support. My master’s studies would have not been possible without him. I want to express a deep gratitude to my parents and my husband’s parents for their love and constant support. All my dear family members and wonderful friends thank you for always being there for me.

Finally, I want to say a BIG thank you to my dear daughters for their unconditional love and patience. This thesis is dedicated to them.

Contents

1. Introduction

We live in the world of 'Video on Demand'. Watching videos on YouTube, Vimeo, Facebook and many such other services has been common for several years. People are getting used to record the videos and share them on social networks regularly. Primarily these are 2D videos, referred to as traditional or conventional videos where all that one records can be seen on a display at once. Figure 1.1 is a pictorial representation of the conventional 2D video.

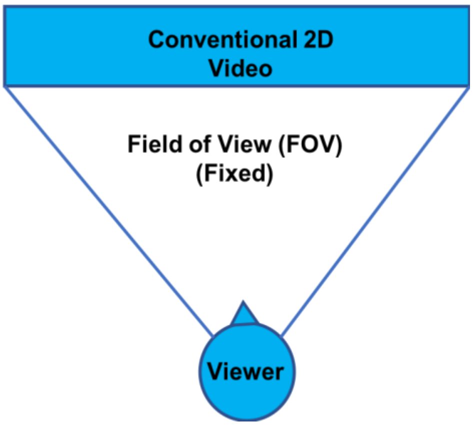

Figure 1.1. Pictorial representation of Conventional 2D video showing fixed field of view.

Video is a rich media type, having a lot of information that can change over time. In traditional videos, the field of view is restricted, which means the user is able to see the part of the scene where the camera was pointing during the recording [Neng and Chambel, 2012]. This is because the entire video is displayed within the human field of view. Nowadays, 360° videos are getting increasing prominence. A 360° video, also known as immersive video gives viewers a view in all directions and an immersive experience of the surrounding of the camera [Lin et al., 2017b]. In 360° videos, views in every direction are recorded at the same time.

1.1 Introduction to 360° video

A 360° video features more information than a conventional (limited field of view) video. There are additional challenges in viewing such videos since one cannot watch in every direction at the same time. Still, such videos provide the whole scene around the viewer, holding the potential to provide fully immersive user experiences. Hence, a 360° video can also be described as a video that has no field of view limitations. Figure 1.2 is a pictorial representation of a 360° video depicting the user's current field of view and entire 360° view.

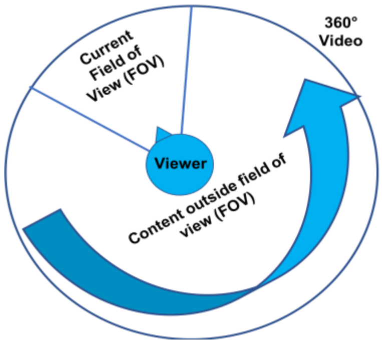

Figure 1.2. Pictorial representation of a 360° video.

When we say that a user can view the video in all the directions and video has no field of view limitations, we talk about the 3-degrees of freedom (3 DoF). In 3 DoF, there are roll, pitch, and yaw movements. Yaw is the sideways or horizontal movement of the head, pitch is the vertical movement and roll is turning the head around the roll axis. These movements can happen due to rotation of user's head or other body movements. For example, while sitting in a swivel chair wearing an HMD, a user may move the chair or just can rotate the head in the 3 directions. Another example is a user wearing an HMD standing in an open space where he/she can make 3DoF movements along with other body movements in the space. In the scope of this thesis, the 360° video content does not support 6-degrees of freedom (6 DoF). Consequently the 360° video content does not respond to user movements such as walking or leaning the body to the front or sides. Figure 1.3 demonstrates roll, pitch and yaw movements.

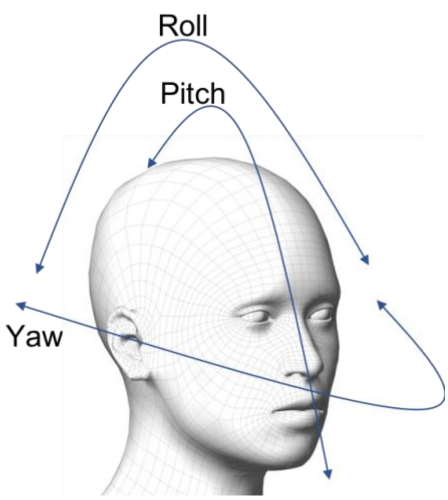

Figure 1.3. 360° viewing with 3DoF.

1.2 Introduction of HMD

A 360° video can be watched on various types of devices such as flat screens, personal computers, mobile phones or Head mounted Displays (HMDs) (Figure 1.4). If we watch a 360° video on a regular flat screen or a personal computer on a browser, controls which can be clicked to move the field of view left, right, up and down are provided on the video player interface. However, continuously clicking these controls to view the entire video content from each direction is very tedious and may not be an enjoyable experience to the user. Hence it becomes difficult to watch the entire content of the video at a glance on a flat screen on a browser.

Recent advances have made HMDs affordable and portable. Head Mounted Display (HMD) has made viewing 360° videos easier and accessible to end users. With the help of an HMD, one can turn the head in 360°, i.e., in all the directions (roll, pitch and yaw) and get an immersive experience of the 360° video. Users viewing such videos during user studies for the thesis, expressed that they actually felt to be present in the location where the video was shot watching the main attractions shown by the narrator in the video.

Figure 1.4. Gear VR (Head Mounted Display).

However, while watching 360° videos with HMD or on a flat screen, it is difficult to focus on all the things happening in the video. For example, a user is watching a view in front, but at the same time, something interesting is happening at the back side. Consequently, the user is missing the interesting events/ objects which are out of the current field of view. This thesis addresses the problem of missing out on important parts of the content which may adversely impacts the user's immersive viewing experience.

1.3 Scope and objective of thesis

360° videos are a very large topic of research. Considering the current popularity of these videos, they will come to our living rooms and will be casual soon. Amongst the different issues related to 360° videos like recording, stitching, recording devices i.e. cameras, audio and many more, this thesis will focus on user's smooth navigation in a 360° video environment and the immersive experience which is the key feature of 360° videos.

This thesis presents the research work performed to address a challenge of missing the important events or objects, endemic to 360° videos. The thesis will discuss in detail the development process of guidance method based on previous work and research done on this topic.

A detailed discussion about the design of the guiding methods represents the core of this thesis. A pictorial representation of the guiding methods is provided to give an idea of their working. The guiding methods were designed, prototyped and evaluated in order to validate their effectiveness and usefulness in achieving immersive experience without losing any important events or objects in the 360° video.

A working software prototype was developed as a part of this thesis. To test different types of guiding methods in 360° environment, variety of content such as sightseeing videos, nature videos, aerial videos of different places, synthetically made content, content with human narration, and many more is required. However, it is very difficult to generate such content by recording various videos with the given timeframe. Hence a variety of 360° content was downloaded from YouTube and combined with the designed guiding methods. An effort was made to make the best possible combinations of guiding methods and 360° content downloaded. These different videos were later combined together with the guiding methods using Unity software to form a working prototype which was used in a user test.

Hence, the main focus of this thesis is on improving an immersive experience of watching a 360° video with the help of different guiding methods to help users to identify and enjoy the important events or objects in the video. It also presents the user studies performed which helped to derive the conclusions which will support the design of good guidance methods in future.

1.4 Research motivation

Watching 360° content is becoming popular these days. For this emerging technique, already a vast variety of content is available on internet for viewing. HMDs are convenient devices to watch 360° videos and they provide an immersive experience to the user. As Petri and Huber [2015] state, 360° videos are composed of moving images that capture up to 360° of the recorded scenes. It is very difficult to watch all the content at the same time due to limited human field of view. Hence, there is a challenge of providing users an immersive experience without them missing the interesting events happening around the 360° video.

To address this two-fold challenge of not missing out on interesting events and achieving that within the diversity of available 360° video content, two research questions were formulated. The first one is *what will be the user experience of viewing 360° videos with HMDs with guidance which help avoid missing interesting events?* It was observed that while watching a 360° video there is a possibility that the user gets disoriented in the virtual environment. Hence there is a great chance of missing the important occurrences in the video which may be of great interest to the viewer. This will keep the user deprived of the immersive experience which is the main purpose of 360° video. If guidance is provided in the form of navigational cues, then there is a possibility that a user will be able to view all the important events or objects. Flawless navigation will improve user's immersive experience. The second question is *which type of guiding methods will be suitable for what type of 360° video content?* Through this research question, I would like to analyze what content and guidance methods match with each other, so that they fulfill the guiding methods requirements (Section 3.1.1) and provide a flawless navigation with immersive experience.

1.5 Thesis outline

This thesis considers different challenges associated with uninterrupted navigation in a 360° environment, in particular challenges with getting an immersive experience and of missing out on important events or object in a 360° video. To address these challenges, this thesis discusses different types of designs of guiding methods which make use of audio, visual and directive cues. Furthermore, the thesis describes the process of developing a software prototype using different types of 360° video content with the variety of guiding methods. In the later part, the thesis uncovers the user experience of using the prototype which indicates the most liked, moderately liked and completely disliked guiding methods. Lastly, this thesis concludes by proposing design implications for the design of guiding methods for 360° environment.

This thesis contains 6 chapters. In chapter 2, background of the thesis is explained along with some relevant concepts and their explanations. The chapter also presents a literature review which presents some solutions and considerations for the research problem of the thesis. The literature review mainly focuses on the 360° video viewing challenges and the different aspects to handle this challenge.

The third chapter describes the guiding methods design process and actual guiding methods. The two-phase design approach and the development cycles are explained in detail in this chapter. The key guidance design constraints, which helped to design the guiding methods, are discussed. This will help the reader to visualize the actual guiding methods implemented in the software prototype. Some visuals from the software prototype are also included.

The fourth chapter is dedicated to user evaluation design, implementation and execution. The details presented in the chapter will explain the user tests execution process and the data collection process.

In the fifth chapter, user evaluation results are presented in detail. These are a combination of quantitative and qualitative data. The results are analyzed in detail to derive the conclusions to identify the most appreciated guiding method and the usable guiding methods which can improve the user's overall and immersive experience. The design implications are presented to support future work in the area.

In chapter 6, conclusions for this thesis are presented. The chapter gives an overview of the problem statements and presents how the work was carried out to address the problem. It again gives an overview of the design implications which are the result of carrying out the detailed research work.

1.6 Author's contribution

My contribution in the thesis work was basic idea generation, developing the idea in detail to design a software prototype, designing the guiding methods, matching the designed guiding methods with the suitable 360° content which results in designing the software application, designing user tests and testing the prototype with users. My colleague Ilkka Rönkä developed the software prototype and assisted in designing the guiding methods, matching them with the suitable 360° content, user experience testing, user experience test environment setup. The software implementation description part is based on Ilkka's thesis description [Rönkä Ilkka, 2018].

2. Background and previous work

This chapter describes the terms related to 360° videos and other terms used in the thesis. Subsequently previous research related to viewing of and navigation in 360° videos is presented.

2.1 360° video and other terminology

Many different terms are used for 360° videos including Omnidirectional videos (ODV's) and 360° panoramic video. Kallioniemi et al. [2017a] specify term interactive omnidirectional video (iODV) as a media format that allows the user to explore and interact with a 360-degree view of the recorded scenery. 360° Hypervideo, iODVs are ODV applications with more interactive elements than just looking around the scene [Kallioniemi et al., 2017b].

Field of view (FOV) is the area or the part of a video which is being watched by the user at any given instance of time. This is a frequently used term in 360° videos. As mentioned in Chapter 1, in conventional 2D videos field of view is the content that is displayed on the screen.

Interesting events or interesting objects are the visuals present in a 360° video which exhibit something interesting to the user in the context of the video. For example, on Facebook360 website ["Facebook360°"] swimming bears are seen catching fishes. These interesting events or objects are also referred to as Regions of Interest or Hotspots.

Immersion can be stated as a feeling of being in the situation or a feeling of being present in the situation. According to Witmer and Singer [1998] immersion is a psychological state of oneself experiencing to be enveloped by, included in, and interacting with an environment.

User Experience (UX) describes the user's feeling about using a product. According to ISO 9241-110:2010 (clause 2.15), user experience is defined as a person's perceptions and responses that result from the use and/or anticipated use of a product, system or service [ISO DIS 9241-210, 2010]. Vermeeren et al. [2010] state that UX is not only something that is evaluated after interacting with an object or a product, but also before and during the interaction.

2.2 Background

360° videos or omnidirectional videos (ODVs) are an emerging new type of media. These are usually recorded with cameras that cover up to 360° of the recorded scenes. 360° videos cannot be viewed entirely at once due to limitation of human vision. Benjamin et al. [2015] emphasize that, users usually focus their attention on a particular region and relocate this region while watching. However, this oftentimes leads to missing relevant events of the video that happened in another location at the same time.

A large amount of 360° / omnidirectional video generated from different data sources such as panoramic imagery, astronomical data, street view data, and many more is readily available. However, the appropriate display options for consuming this content remain scarce due to their inherent immersive and borderless nature [Benko and Wilson, 2010].

Head mounted display is one such medium which offers a useful platform for watching 360°/ omnidirectional content. Lot of research work for viewing 360° videos has been carried out using head mounted displays (HMDs). Head mounted displays are immersive and offer an advantage of natural spatial navigation through head movement. HMDs provide a high degree of freedom for interaction [Benjamin and Huber, 2015].

Lin et al. [2017a] state that, while 360° videos can display full-field content, the limitations of viewing generally restrict views to only a subset of the full video, thus giving users a natural visual experience (conventional 2D video viewing experience).

The users navigate through a 360° video by orientating the screen or their head if they are using an HMD. As 360° videos cover a grand area, it becomes difficult for the users to search for the potential events. This degrades the user experience, leading the users to missing some important events while navigating through the entire 360° video.

Such instances of searching for the important events, objects or regions of interests become more problematic when the videos contain multiple important events or objects. Hence, users have to switch between multiple important events or objects to understand the whole story displayed in the video. This might lead to incomplete understanding about the video content which reduces the immersive experience, a key aspect of watching 360° videos. As a result, there are high chances of users feeling lost or disoriented in a 360° video environment.

2.3 Previous Work

There is an extensive ongoing research to find a concrete solution to address the above-mentioned problem. There are many different aspects to this problem and many different solutions have been proposed to address them. This section discusses the relevant research work.

In virtual environments efficient navigation is a major issue. In navigation and route guidance, landmarks play a very important role in communicating directions. According to Kallioniemi et al. [2013] availability of prominent landmarks affects how easy or hard it is to provide guidance in certain location.

Darken and Sibert [as cited in Kallioniemi et al., 2017] list important features when focusing on wayfinding in large virtual worlds

1) User needs adequate source of directional cues, for example landmarks.
2) A large world without explicit structure is difficult to search exhaustively.
3) Path following is a natural spatial behaviour.
4) A map allows for optimizations of search strategies.

Hence, based on the above study by Kallioniemi et al. [2013], it can be said that if navigational cues are provided in a virtual environment, users can navigate through a virtual environment achieving a sense of immersion without losing the important events or objects.

However, navigating in a 360° hypervideo is challenging. It has been observed that 360° hypervideo player must provide the users with appropriate affordances to understand the hypervideo structure and to navigate it effectively while allowing them to have an immersive experience. Main challenges in hypermedia are disorientation and cognitive load. Location awareness can become an issue due to the lack of boundaries [Neng and Chambel, 2012].

Neng and Chambel [2012] designed a system containing hotspot availability and location indicator to inform the user the location and availability of hotspots outside the current field of view: The bigger the size of a hotspot, the closer its availability will be. Hotspot markings are kept transparent to minimize their impact on the video content. Hotspots are also accompanied by textual information about them in a small text box.

Another study by Lin et al. [2017a] demonstrate designs of guidance techniques. In this study a PIP (picture in picture) technique was used to indicate hotspot availability. In this offscreen regions of interest are reintroduced in a spatial PIP previews, allowing the users to make sense of where the offscreen ROI’s are. In designing a prototype, Lin et al. [2017b] followed some design considerations such as ‘minimizing effect on the video content, assisting user to reach off screen content’.

Another type of work was done by Lin et al. [2017b] to assist the user in a 360° environment. They developed two focus assistance techniques which involved Auto Pilot (directly bringing viewers to the target), and visual guidance (indicating the direction of a target). An arrow was used to indicate the direction. This study has similarity with our prototype where our prototype used an arrow method to indicate the direction of the interesting event or object. Similarly, auto pilot assistance technique also shows a similarity with automatic transition guidance method used in our prototype. However, in this thesis other types of guidance methods involving 3D audio as well as a mix of different guidance techniques are also analysed.

Lin et al. [2017b] study was conducted for two different types of video content ‘Sport’ and ‘Tour’. There was an emphasis on matching guiding method with the appropriate video content by the author. This finding is corroborated in this thesis: a guidance method and video content should be complementary to each other. The results of this study also emphasize on slower auto-pilot method which is supported by our study feedback for the automatic transition method.

Saarinen et al. [2017] present a study where interaction with hotspots was studied in a 360° video environment. This study also present guidelines for designing and producing interactive content in an omnidirectional video. Following are some of these guidelines followed while designing the application. For example, guidelines regarding avoidance of objects very close to the camera, presenting details of 360° videos with embedded content, and using different visual cues were followed among others.

There has been research done in finding user experience evaluation approaches for virtual reality and 360° video watching. The research by Kauhanen et al. [2017] consider aspects such as Immersion, Presence, Disorientation, Sense of Control, Pleasantness, Exploration and Simulator Sickness. However, in this thesis Immersion, Presence, Disorientation, Sense of control, Pleasantness, and Exploration are evaluated as a composite overall experience with the guiding methods (Chapter 4). This was done to get an early feedback regarding the suitability of guiding method designs.

2.4 Research through design approach

Research through design (RtD) as a method for interaction design research in HCI is a new model proposed by Zimmerman et al. [2007]. This research approach employs methods and processes from design practice as a legitimate method of inquiry [Zimmerman et al.,2007; Zimmerman et al., 2010].

2.4.1 RtD design thinking overview

RtD defines *design thinking* as the application of a design process that involves three steps. Firstly, grounding - investigation to gain multiple perspectives on a problem. Secondly, ideation - generation of many possible different solutions. Thirdly, iteration - cyclical process of refining concept with increasing fidelity and reflection [Zimmerman et al.,2007]. The research methodology in this thesis is inspired by this design thinking approach. In the following, the utility of the RtD approach in relation to this thesis is elaborated.

2.4.2 Motivation and application

“RtD allows researchers to rely on designerly activities as a way of approaching *messy situations with unclear or even conflicting agendas*; situations that are not well suited to other methods of inquiry. Additionally, RtD forces researchers to focus on research of the future, instead of on the present or the past. Finally, RtD provides an opportunity for the research community to *engage in discourse on what the preferred* state might be as an intentional outcome of the research, allowing us to consider the ethics of what we design.” is an excerpt from [Zimmerman et al., 2010] (italics not a part of the original text). This describes some key elements which were considered to be relevant for the research in this thesis.

The core part of the thesis work is designing guiding methods for a 360° video which will help the user to have an undisturbed immersive experience. Designing optimal guidance methods while watching 360° videos is a problem with many variables. For example, different content may require different types of guidance. Furthermore, the concept of guidance and unhindered immersive viewing of content are inherently contradictory. This experience is very individualistic because a certain guiding method can be useful for one user but disturbance for another user. Hence, achieving a guiding method that is probably useful for the majority of users is a difficult task.

In this thesis, a prototype is designed to first visualize and subsequently evaluate a particular guidance method. Thus, the objective of guiding users in a 360° video environment is approached with a two-step approach of design and user evaluation.

This thesis work has undergone the design thinking process as defined by Zimmerman et al. [Zimmerman et al., 2017]. The solutions to the problem presented in this thesis went through an iterative design process which explored multiple different options before narrowing down on the designs considered the best. These selected designs were used as the basis for the guidance methods which were subjected to user evaluation.

In the next chapter guiding methods design process as well as design details are presented. The design considerations for guiding methods in the software prototype corroborates some design considerations mentioned in the previous work.

3. Design process and implementation of guiding methods

"Location of a visual elements in UI has a huge impact on how user interprets information" Rick Oppedisan.

This chapter focuses on designing of guiding methods for 360° video and their implementation to make a working software prototype. The chapter discusses key design constraints and requirements taken into consideration before designing the guidance methods. It also discusses the prototype evolution process and actual design of the software prototype.

In the beginning, the chapter describes the developmental life cycle of the software prototype and evolution of designs of the guiding methods for 360° video. The software prototype was developed by undergoing two-phase design and development life cycle. Hence this chapter has a detailed discussion of phase one design and development of the guiding methods followed by phase two design and development of the guiding methods. Later, the chapter discusses guidance methods implementation process for 360° video in brief.

3.1 Overview of the design process of guidance methods

Designing guiding methods was the main task to solve the problem of subtle navigation in a 360° video. Main goal of this process is designing the guiding methods, applying these guiding methods to the suitable content extracted from YouTube and making a software prototype by combining these videos and the guiding methods. The designing process started from scratch where the only known aspect was the problem statement and we wanted to give a solution for it. Though there were some references (Section 2.3) to the previous work done to address this problem, we were aiming to have a novel component in our design. Hence, we started with brainstorming sessions where we reviewed current work done and related references, which comprised of prior publications as well as prototype demonstration setups involving ODVs (Section 2.3).

As the first step, some nascent ideas were proposed. Next these ideas were further developed to design the guiding methods for 360° videos. For example, an arrow-based navigation paradigm was one such idea which was developed further. We studied some 360° video work which was already been developed for HMDs in the university. Using these software prototypes gave us insights on working of 360° content on HMD and the challenges of designing guiding methods for 360° content on HMD (Section 2.2).

A basic understanding to design the guiding methods was developed from reviewing the prior work. For example, users should be provided with the visual cues to navigate in a 360° environment, the visual cues should be subtle and unobtrusive, and so on. This basic understanding from the previous work studied helped us work efficiently on the nascent ideas and design appropriate guiding methods for 360° videos. As a next step, the guiding methods were divided into three broad groups: Directive guiding methods, Audio guiding methods and Visual guiding methods. In addition, guiding methods which combined audio and directive methods to enhance the user experience were developed.

It was realized that in some scenarios, user's current field of view may be away from an interesting event which requires immediate attention. In such cases, it would make sense to automatically change the user's current field of view such that it covers the interesting event. For example, to ensure that a user watches an important occurrence which is out of current field of view. Consequently, in the third step, the guiding methods which included active guiding paradigm were included. In this method, the user is directed towards an important event or object by changing the user field of view automatically (Section 3.2.2, VII, VIII). This extended the guiding methods into active and passive guidance. See Figure 3.1 below.

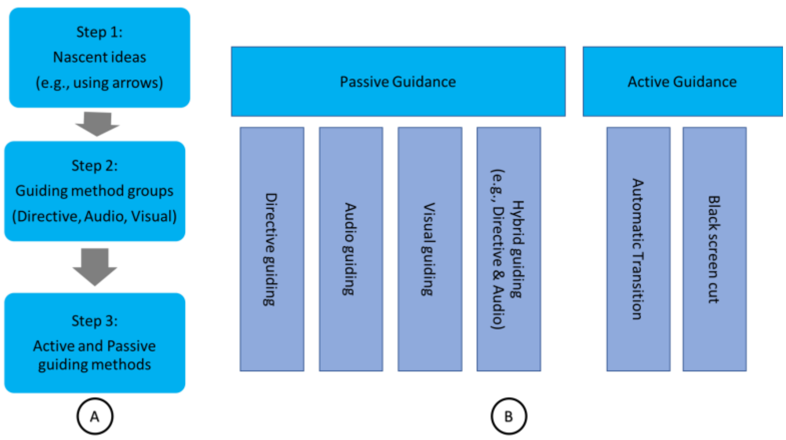

Figure 3.1. Active and Passive guiding paradigms.

3.1.1 **Key guidance design constraints and features**

Following are the main constraints from Lin et al. [2017a] study which were directly referred to while designing the guiding methods used in 360° videos to make a software prototype.

1) *Minimizing the impact to the main content of the video*
Adding visual cues to the main screen would degrade the immersive experience. Also, the assistance should not 'steal the show' from main screen. Hence guidance should be placed in the peripheral regions of the screen.

2) *Effectively guiding users to reach off-screen targets*
In addition to giving information about what is out-of-screen, a viewer should know how to find that. To guide the viewers spatial information which includes both directional and distance information should be displayed. The guidance should be natural and intuitive.

Main requirement was to provide guidance so that a user should be able to locate all the important occurrences happening in the 360°video. Guidance should be such that it guides the user in a subtle manner unobtrusively, Guidance should not become a distraction, guiding method should not pop up as a surprise [Argyriou et al., 2016], and while using an HMD, guiding method should be pleasing to eyes. This is important as the field of view is limited on an HMD and if the objects are very close to camera, they may obstruct useful information and become disturbing [Saarinen et al., 2017].

There is an interesting concept of audio guidance in our prototype. According to Kauhanen et al. [2017], if audio is used in a 360° video to guide the user it can, to some extent, help minimize the negative effects of poor visual quality of the video. Moreover, audio can be used to direct the attention of the user and it can also help facilitate the storytelling elements. In our prototype, there are three types of audio: a speech synthesizer audio, 3D treble sound (high pitch), and 3D bass sound (low pitch). *This is a novel concept which is used in guiding methods design. According to the current knowledge, we are the first ones to use different types of audios as guiding methods in an innovative way*.

3.2 Two-phase guiding methods development

This subsection presents the two-phase guiding methods design, refinement and evaluation process which was implemented in the thesis to exploit constructive feedback from such an iterative approach. Before going further, it is important to note the key constraints which guided the design process (Section 3.1.1).

3.2.1 Design and development cycle of Phase 1

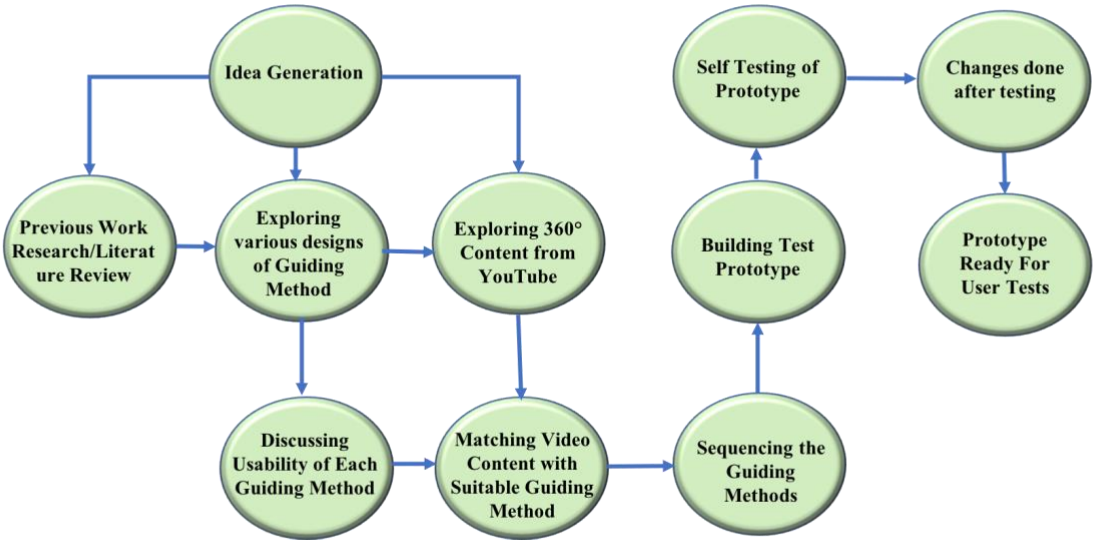

Figure 3.2. Phase 1 – Developmental life cycle.

In the first phase, the design process started with a very simple guiding technique of guiding the user to a specific direction with an arrow mark. This is '*Arrow method*', a small green arrow pointing to some important event or object in the video. Few important events in the video were selected and fixed for user guidance. The user would be guided to these pre-determined events with the help of the arrow to locate them.

Review of previous work related to guiding methods in a 360° content suggested that there were no '*Audio guiding methods*' used for exploring 360° videos. Hence, we are the early ones to design and evaluate an audio guiding method for locating a hotspot (important event or object) in 360° video. In audio guidance three types of audios were being used. The first type was *speech synthesizer audio* giving a direction such as right, left, behind etc. The second and third types were *Treble sound* (high pitch) and *Bass sound* (low pitch).

Some other markers (visual cues) were also used. One was '*a Hollow green rectangular marker'* which indicated a part of video which might be showing some interesting content. This rectangular marker also showed some '*textual information*' about the marked area. The addition of the text to the marked area was implemented to enhance the user experience by giving more information about the content of the video.

A green colored '*Bull's Eye*' mark was used as a mark to indicate an interesting event or object in the video as one of the guiding methods. Using different markers like a *Hollow Rectangle*, *Informative Text* and *Bull's eye* to indicate an interesting point of sight were used to find an appropriate marker (visual cue) which can guide the user.

During our discussions while designing the guiding methods, we came up with an interesting guiding method of '*Automatic transition*'. Automatic transition means user's field of view (FOV) is changed by blurring the content of the current field of view and another interesting content of the video is shown right in front of the user's current field of view. In this case the changed field of view does not have any markers which can point at some interesting event or object. Hence, another method was designed in which, after the automatic transition, a green colored '*eye marker'* is shown to indicate the interesting content of the video. This was designed to clearly indicate an interesting content after the user is automatically directed towards some part of the video.

A common thing available in all the videos from phase 1 and phase 2 was a green colored '*Guidance Available*' text. This text appears in a small font at the bottom of the 360° video. The text appeared only when the guidance was available in the video.

Final combinations of guidance methods and different 360° videos for Phase 1

The following table (Figure 3.3) shows the list of 360° videos and the guiding method attached to each and the guiding methods description. YouTube links of the videos used in the software prototypes are available in the reference section. Detailed description of each method is provided in Section 3.2.2

No.	Guiding Method name	Guiding method description	Video used
1.	Arrow	Arrow indicating the direction	London sightseeing
2.	Arrow with directional voice	Directional voice along with arrow indicating the direction (Left, Behind)	London sightseeing
3.	3D treble sound with eye marker	3D treble audio indicating the direction and eye marker indicating interesting event/object	Angel falls
4.	Speech synthesizer audio with hollow rectangle and eye marker	Speech synthesizer audio with a hollow rectangle and eye marker	Angel falls
5.	Bull's eye	Bull's eye	Air balloons
6.	Bull's Eye with speech synthesizer audio guidance	Speech synthesizer audio guidance and Bull's eye	Air balloon
7.	Automatic transition with blur effect	Automatic transition with blur effect	Venice carnival
8.	Automatic transition with blur effect and eye marker	Automatic transition with blur effect and eye marker	Venice carnival

Figure 3.3. List of videos and guiding methods combination phase 1.

3.2.2 Guidance methods in Phase 1

I. *Arrow method*: In the arrow method the user is guided with the help of a small green arrow. Arrow method is the simplest of all the methods and main intention to use this method is to guide the user with least distraction. A small green arrow is used which shows directions like 'Behind' and 'Up'. A London sightseeing ["A London City Guided Tour"] video was used to test this guiding method. It had many historical monuments and arrows could point the important monuments along with the human narration in the video itself.

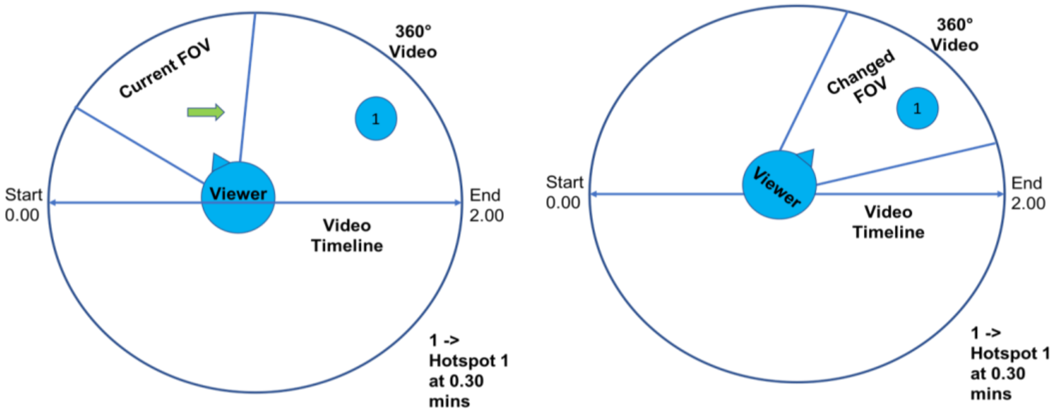

Figure 3.4. Arrow method.

As shown in the Figure 3.4, the user is watching a specific direction. When the software notices some interesting event or object at some other direction than the current FOV, a green arrow appears in the current FOV and shows the direction of the interesting event or object to the user at that particular time. This helps the user to know about the direction to watch to see the interesting event or object at that time. When the user turns in the direction shown by the arrow mark, it disappears from the screen.

II. *Arrow with a Directional Voice*: In this method the user is guided with the help of a small green arrow along with a speech synthesizer audio that indicates the direction. London sightseeing ["A London City Guided Tour"] video was used to test this method as well. Though the video itself had a person narrating about the historical monuments that were present in the video, the speech synthesizer audio is used only to tell the direction along with the arrow.

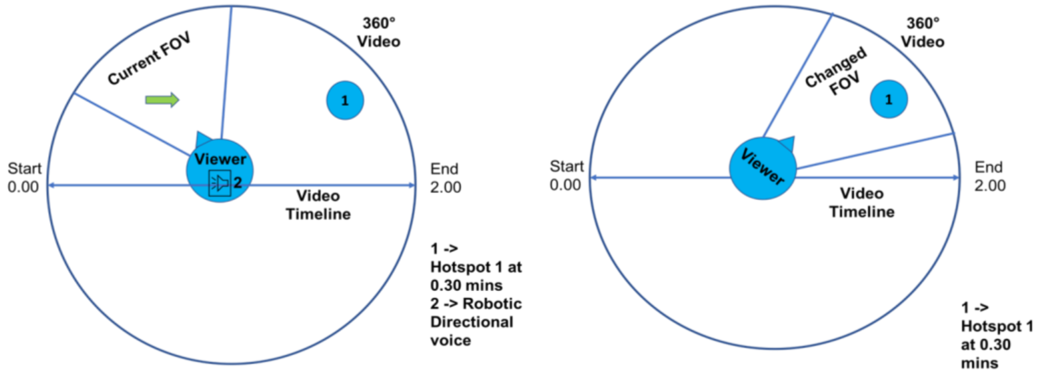

Figure 3.5. Arrow with directional voice.

As shown in the Figure 3.5, the user is watching a specific direction. When the software notices some interesting event or object at some other direction than the current FOV, a green arrow appears in the current FOV and audio is generated through speech synthesizer which gives direction of the interesting event or object to the user at that particular time. This helps the user to know about the direction to watch an interesting event or object at that particular time. When the user turns in the direction indicated by the arrow mark and the speech synthesizer audio, the arrow disappears from the screen and the audio stops. Audio keeps on repeating and the arrow mark remains on the screen till the user turns in the specified direction.

III. *3D Treble sound with Eye marker*: This method uses a 3D treble (high pitch) sound from right and left direction where the interesting event or object is located. The event or object is visually indicated with an eye marker. The idea behind this method is that the user should understand the voice direction and turn towards it to see the interesting event or object visible at that time. Angel falls [“Angel falls, Venezuela”] video was used to test this guiding method and the video itself had narration. Eye markers are used to clearly indicate the hotspots.

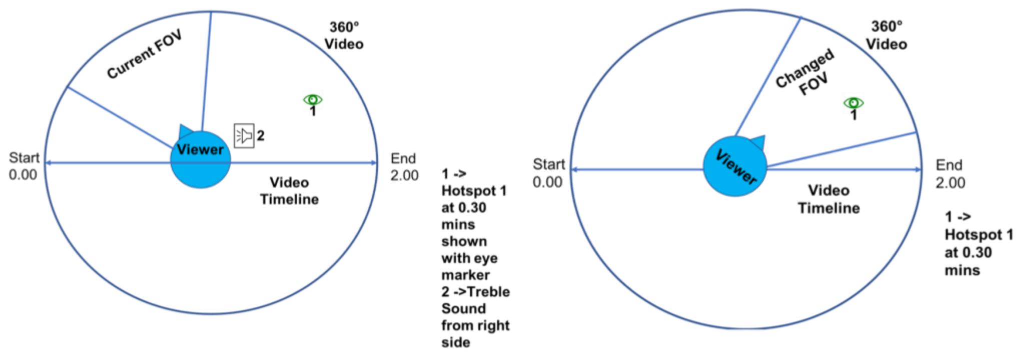

Figure 3.6. 3D treble sound with eye marker.

As shown in Figure 3.6, the user is watching a specific direction. When the software notices some interesting event or object at some other direction than the current FOV, a 3D treble sound is generated by the system from the direction where the interesting event or object is present. When the user notices the sound and turns to the direction of the sound, the interesting event or object is shown with a green eye marker. As soon as the user turns in the direction of the audio, the audio vanishes.

IV. *Speech synthesizer audio guidance with a Rectangle and eye Marker*: This method uses a speech synthesizer audio to direct the user towards a hotspot. The hotspots are marked with a hollow rectangle with green border. The rectangle is used to indicate not only a specific object but also the area around it as an interesting event or object. Angel Falls video [“Angel falls, Venezuela”] was used again with this method and a speech synthesizer audio giving directions to the user was included. The directions are ‘Behind’, ‘Left’, and ‘Right’. Hence depending on the gaze point, the user can turn in the particular direction to see the hotspots.

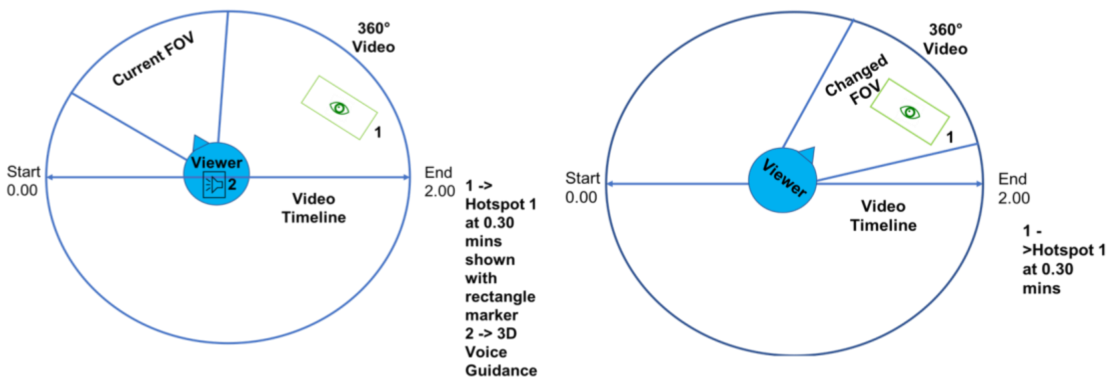

Figure 3.7. Speech synthesizer audio guidance with a rectangle and eye marker.

As shown in Figure 3.7, the user is watching a specific direction. When the software notices some interesting event or object at some direction other than the current FOV, the audio guidance is generated by the system via speech synthesizer telling the direction where the interesting event or object is. When the user turns to the direction indicated by the speech synthesizer, the interesting event or object can be seen with a green colored hollow rectangle mark and the eye marker at the event or object. As soon as the user turns in the direction indicated by the audio, the audio vanishes but markers stay on the screen. Audio keeps on repeating till the user turns to the specified direction.

V. *Bull's Eye Mark*: Bull's eye mark is used to indicate the interesting event or object in the video. An air balloon video ["The Golden Ring of Russia Air-Balloon Festival"] was used to test this guiding method, and few instances of different air balloons were marked as the interesting events or objects.

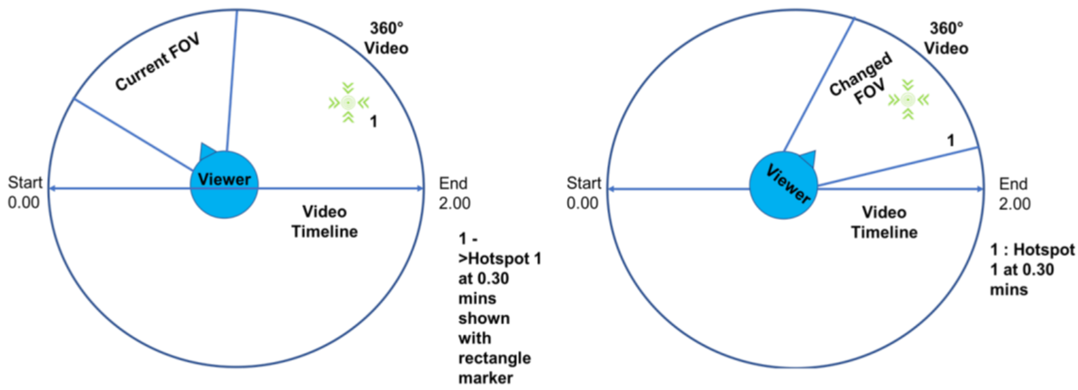

Figure 3.8. Bull's eye mark.

As shown in Figure 3.8, the user is watching a specific direction and bull's eye marks are already present at various places on interesting events or objects. With this guiding method, the interesting events or objects will be seen by the user only if the user explores around the video on his/her own accord.

VI. *Bull's Eye with a Voice Guidance*: The air balloon video ["The Golden Ring of Russia Air-Balloon Festival"] was used again with this method. In this method the interesting event or object is shown using a Bull's eye. If the user is watching somewhere else than the hotspot location, the user is guided to turn in the direction of the interesting event or object using voice guidance. Voice guidance is mainly giving the direction to turn to 'Behind', 'Left', and 'Right'.

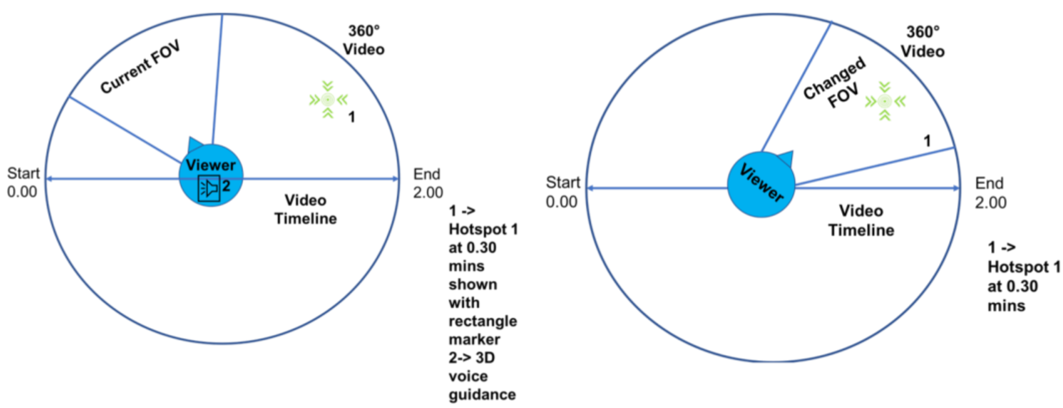

Figure 3.9. Bull's eye with voice guidance.

As shown in Figure 3.9, the user is watching a specific direction. When the software notices some interesting event or object at some place other than the current FOV, audio guidance is generated by the system via speech synthesizer telling the direction where the interesting event or object is. When the user turns in the direction indicated by the speech synthesizer, the interesting event or object can be seen with a bull's eye mark on the event or object. As soon as the user turns in the direction indicated by the audio, the audio vanishes but markers stay on the screen. Audio keeps on repeating till the user turns in the specified direction.

VII. *Automatic Transition with Blur Effect*: In this method the user's field of view is changed automatically and an interesting event or object in the video is brought in the user's FOV. While doing this transition, the content of the current FOV is blurred for a short duration (Figure 3.10, Step 2) and then the 360° video is rotated in yaw direction by the system to bring the interesting event or object in user's FOV (Figure 3.10, Step 3 and 4). A Venice carnival video ["Carnival of Venice, Italy"] was used with this guiding method. In the video a small boat takes the user through the waters of the Venice city. While the video itself is moving, few instances are marked as interesting event or objects and the blurring effect is applied to bring them in user's field of view.

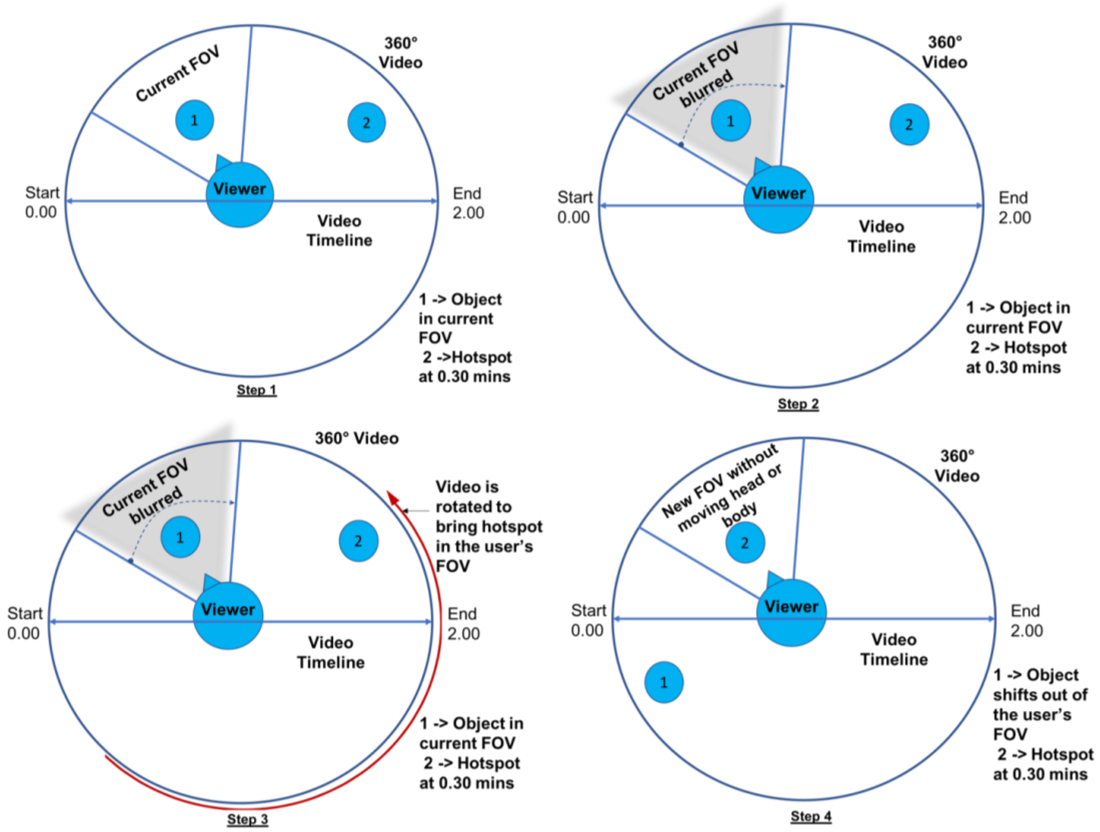

Figure 3.10. Automatic transition with blur effect.

VIII. *Automatic Transition with Blur effect and an Eye Marker*: The basic mechanism of this method is same as automatic transition with blur effect method

(Guidance method VII). The hotspot is marked with an eye marker (Figure 3.11) to clearly indicate the interesting event or object. Venice carnival video ["Carnival of Venice, Italy"] was used again, and the hotspots were marked with the green eye marker to simplify the hotspot search for the user.

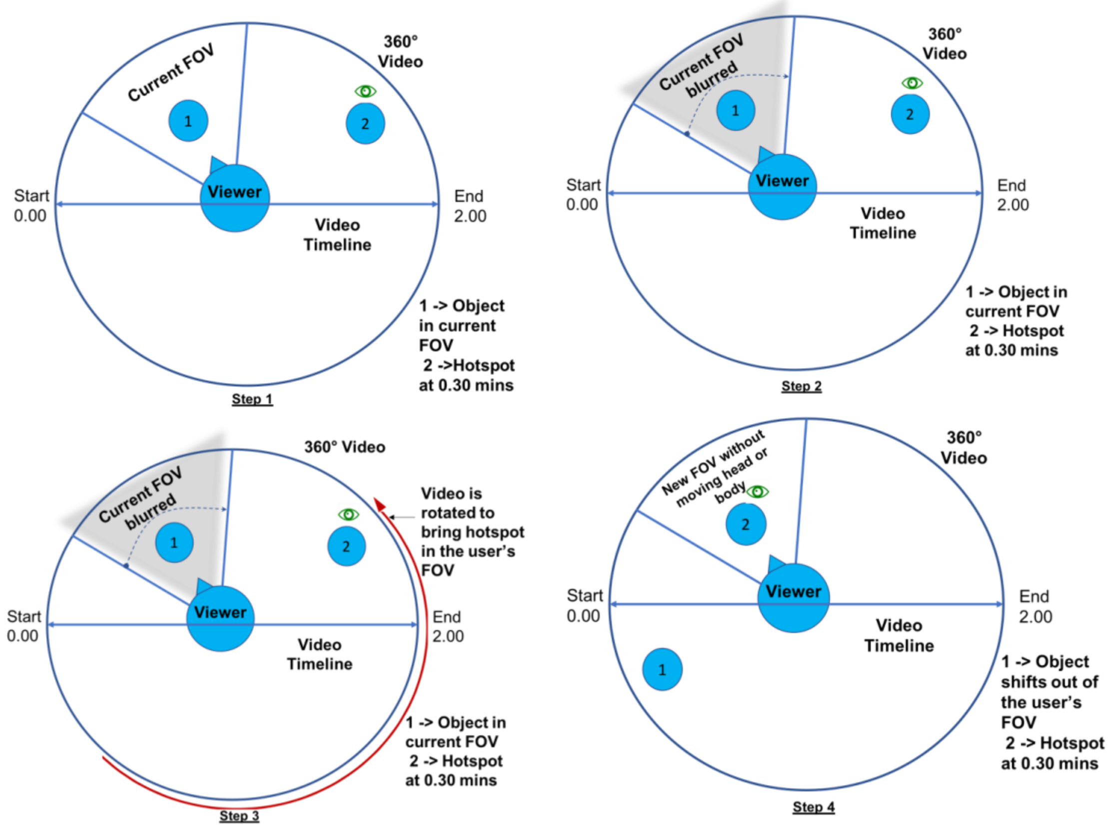

Figure 3.11. Automatic transition with blur effect and eye marker.

The following figure 3.12 shows actual visuals of the guiding methods from phase 1 software prototype.

Figure 3.12. Example of Actual Visuals of the videos in Phase 1.

Subsequent to designing and developing the prototype player with the guiding methods, we tested the prototype for a sanity check. After implementing minor changes actual user testing was done. This was the Phase 1 guidance method testing. After completing the user evaluation, the lessons learnt from the user feedback were incorporated in the guiding methods designs to improve our prototype and test it again (see Section 5.1 for user evaluation results for Phase 1). Hence, another round of brain storming session was done to redesign some guiding methods. A new revised prototype was developed as Phase 2 and evaluated again with another round of user evaluations. The next section will discuss Phase 2 guidance design development cycle in detail.

3.2.3 Design and development cycle of Phase 2

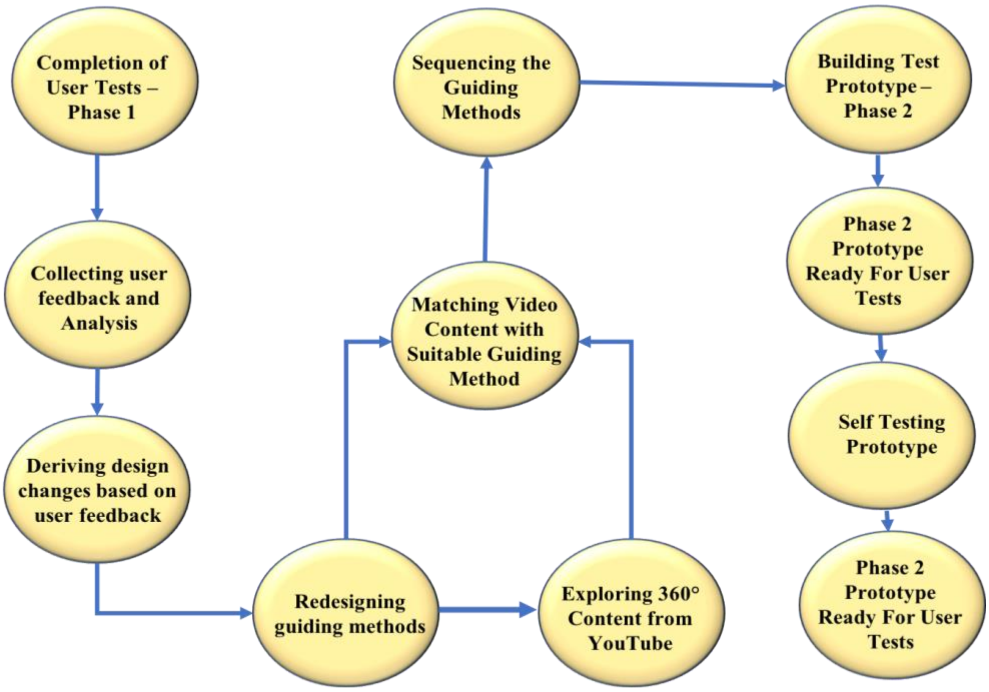

Figure 3.13. Phase 2 design development cycle.

After completing first round of user tests, based on the user feedback we redesigned some of the methods and added new methods to our evaluation prototype.

In the second phase of designing, new guiding methods were introduced based on the previous user test feedback (Section 5.1.1). Textual markers were introduced in many videos to give additional information about the video content. '*Textual markers*' also helped in building a story telling concept in some of the test videos. It was observed in Phase 1 that high pitch sound, i.e. treble sound, is hard to recognize by the users, so a bass sound, i.e. a low pitch sound, was introduced in Phase 2. We introduced a set of '*multiple arrows*' pointing the user towards a particular direction where interesting events or objects are present. In automatic transition method instead of blurring the content of the current field of view, a '*black screen cut*' was introduced along with the textual information about the video content in the current field of view.

As another development, we used some methods in combination with each other. A speech synthesizer audio guidance was accompanied with an arrow and some textual information about the content of the video. Such combinations of guiding methods were introduced to enhance the user experience.

The following table explains the guiding methods which were changed from Phase 1 to Phase 2.

No.	Phase 1 Guiding Methods	Changes for Phase 2 guiding Methods
1.	London Video: **Arrow indicating the direction(Arrow)**	London video: with **only Arrow** was kept **same**.
2.	London Video: **Directional voice along with arrow indicating the direction(Arrow with Directional voice)**	New Moscow video with **only music** with **Arrow**, **Textual Guidance** and **Voice Guidance** added.(Previous video changed as it had narrative which was distracting with the voice guidance)
3.	Angel Falls: **3D voice indicating the direction(3D Treble sound with Eye marker)**	Frozen Falls: New Frozen Falls video with **3D Bass Sound** and **Textual Guidance** added.(To check if people can recognize **3D Bass sound** over **3D High Note sound**)
4.	Angel Falls: **3D voice and a Rectangle Marker(3D Treble sound with Rectangle marker)**	Angel Falls: **High note 3D sound with Rectangle Marker(3D Treble sound with rectangle marker)** kept **same**
5.	Air Balloons: **Bull's eye(Bull's Eye)**	New Space Dream video added with **Directing Arrows without Bull's Eye.** (Bull's eye was distracting)
6.	Air Balloon: **Voice guidance and Bull's eye (Bull's Eye with Voice guidance)**	Air Balloon: **Bull's eye mark with directional voice (Bull's Eye with Voice guidance)** kept **same**
7.	Venice Carnival: **Automatic transition (Automatic Transition With Blur Effect)**	Kamchatka Volcano: New Kamchatka volcano video added with **Black Screen Cut** and **Changed field of view**.(Earlier blurring effect was removed)
8.	Venice Carnival: **Automatic Transition with blur effect (Automatic Transition with Blur effect and Eye marker)**	Venice Carnival: Venice Carnival video kept same but with **Slower Automatic Transition**

Figure 3.14. Changes done in guiding methods in phase 2 as compared to phase 1.

Final combinations of guidance methods and the different 360° videos for Phase 2

The following table shows the final list of 360° videos and the guiding method attached to each video and the guiding methods description. YouTube links of the videos used in the software prototypes are available in the reference section. Detailed description of each method is provided in section 3.2.4

No.	Guiding Method name	Guiding method description	Video used
1.	Arrow	Arrow indicating the direction	London sightseeing
2.	Arrow with directional voice and textual markers	Arrow with directional voice and Textual marker	Moscow kremlin
3.	3D Bass sound with arrow and Textual marker	3D Bass sound with arrow and Textual marker, story telling approach	Frozen falls
4.	Speech synthesizer audio with hollow rectangle and eye marker	Speech synthesizer audio with a hollow rectangle and eye marker	Angel falls
5.	Multiple directing arrows	Multiple directing arrows	Space dream
6.	Bull's Eye with speech synthesizer audio guidance	Speech synthesizer audio guidance and Bull's eye	Air balloon
7.	Black Screen cut with Textual marker	Black Screen cut with Textual marker	Kamchatka volcano
8.	Automatic transition with blur effect and eye marker	Automatic transition with blur effect and eye marker	Venice carnival

Figure 3.15. List of videos and guiding methods combination phase 2.

3.2.4 Updated guidance methods in Phase 2

I. *Arrow with 3D Voice Guidance and Textual Marker*: In this method (Figure 3.16), the user is guided with the help of a small green arrow. When the arrow is giving direction to the user a speech synthesizer audio is used to tell the direction. The direction given is ‘Behind’, ‘Left’ or ‘Right’. When the user turns and actually looks at the hotspot, textual information is displayed. We used a Moscow Kremlin video [“Moscow Kremlin”] with this guidance method. It contained many historical monuments from the city of Moscow. The textual information is displayed to make the user experience more informative.

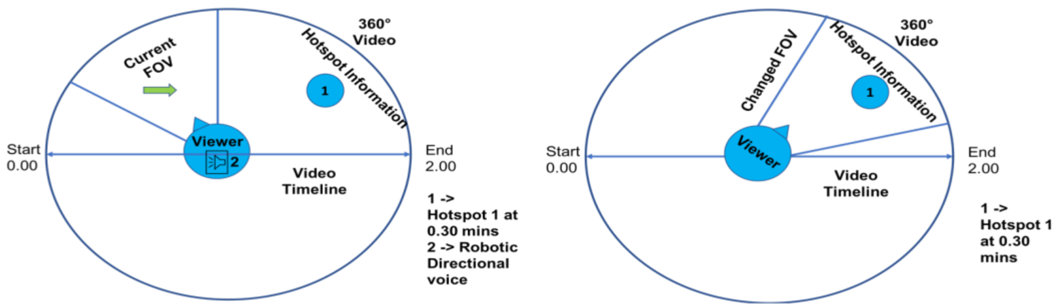

Figure 3.16. Arrow with 3D voice guidance and textual markers.

Until the user turns to the direction given by the speech synthesizer audio, the audio keeps on repeating and stops when the user sees the important event or object.

II. *3D Bass Sound with Arrow and Textual Marker*: In this guiding method (Figure 3.17), a low pitch 3D bass sound is used when an important event or object is near user’s field of view and an arrow marker is shown into the direction where the important event or object is. Idea behind using a 3D bass sound is that the user should recognize the direction from which the sound is coming. The user should make use of the sound to understand the direction. Textual markers are used to give more information about the important event or object. A story telling approach is used in the form of textual information. A frozen falls video [“Plitvice Lakes in Winter, Croatia”] was used with this guidance method.

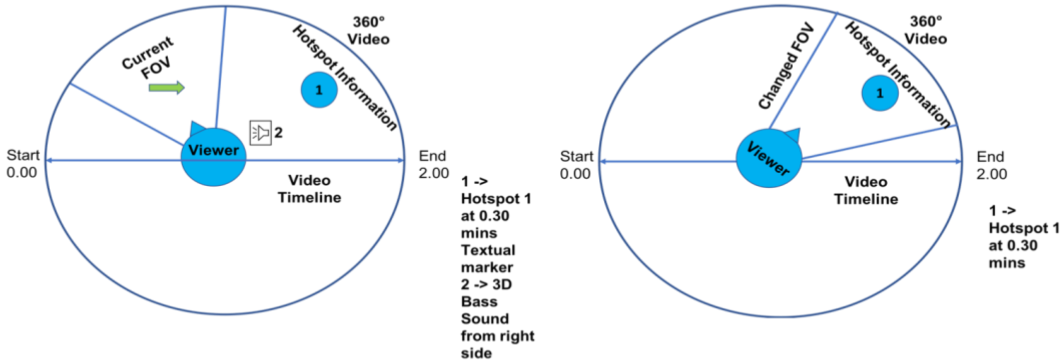

Figure 3.17. 3D Bass sound with arrow and textual markers.

Until the user turns to the direction indicated by the 3D bass sound, the audio keeps on repeating and stops when the user sees the important event or object.

III. *Multiple Directing Arrows*: In this guiding method (Figure 3.18), a series of arrows directs the user towards an important event or object. Main intention of this method is to direct the user towards a specific event or object without obstructing the field of view. This method was derived from the 'Bull's Eye' mark method based on Phase 1 user tests feedback. Hence a subtle use of multiple green '< 'arrows is applied. Once the user reaches the specific event or object the arrows vanish. A video called space dream consisting of synthetic content ["Space Dream 360°"] depicting space objects was used with this guiding method.

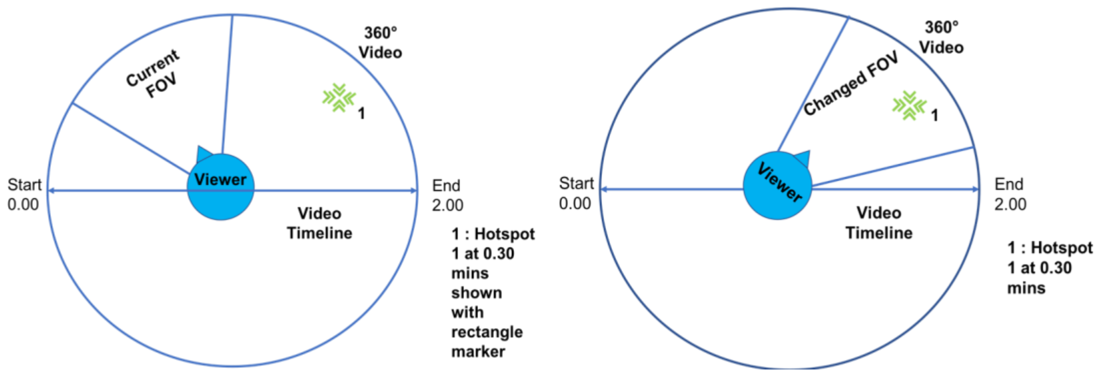

Figure 3.18. Multiple directing arrows.

IV. *Black Screen Cut with Textual Markers*: In this guiding method (Figure 3.19), while the video is playing and there is some important event or object in some other direction than the user field of view, a black screen appears in the current field of view (Figure 3.19, Step 2) for few seconds while the video is rotated to bring the changed field of view in front of the user (Figure 3.19, Step 3 and 4) (hence named it as Black screen cut). In this way the user is shown the important event or object with a textual marker that gives additional information about the event or object to the user. Kamchatka volcano from Russia ["Kamchatka Volcano Eruption"] video was used to test this guiding method as the content of the video was considered interesting and textual markers could give additional information about the video.

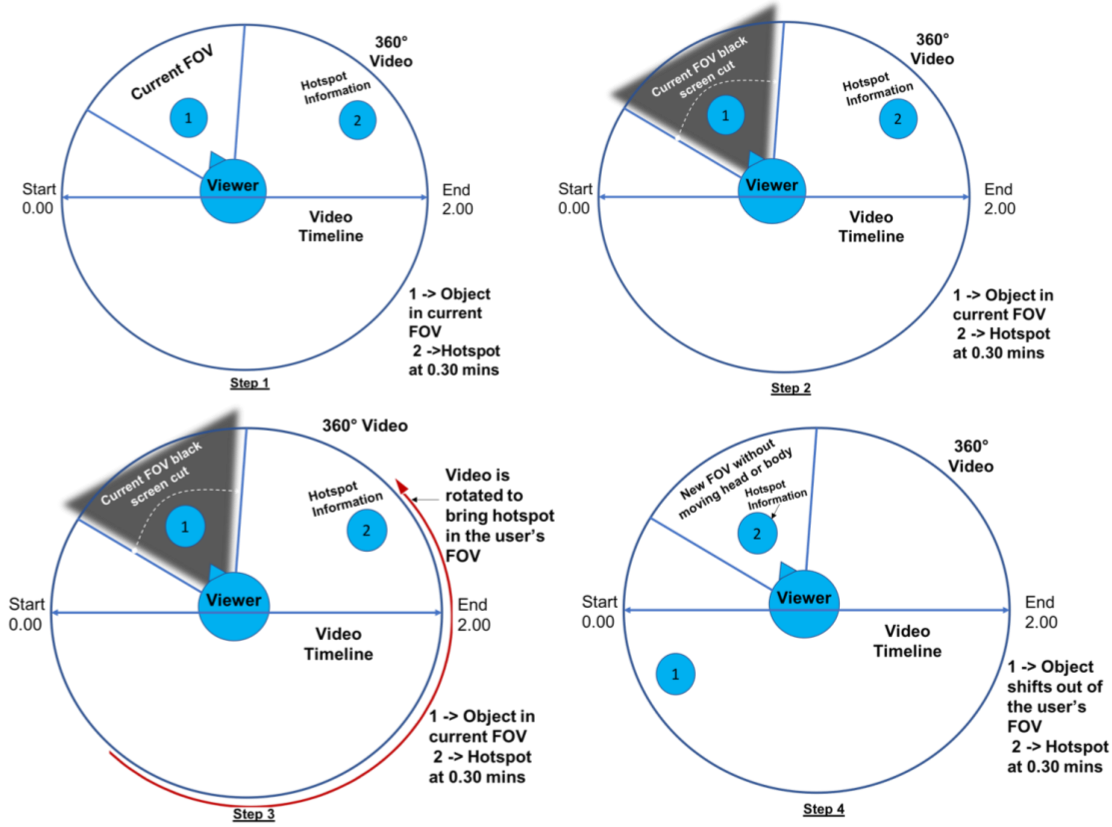

Figure 3.19. Black screen cut with textual markers.

The following are the screenshots of the actual guiding methods in Phase 2 from the test prototype.

Figure 3.20. Example of Actual Visuals of the videos in Phase 2.

This chapter explained the process of designing the guiding methods for a 360° video. The tabular demonstrations for Phase 1 (Figure 3.3) and Phase 2 (Figure 3.15) guiding methods explained the two-phase design approach and evolution of the prototype from Phase1 to Phase 2. Figure 3.14 described the difference between the two phases and explained the reasons for the changes in guiding methods. Actual visuals from Phase 1 and Phase 2 gave understanding of the guiding markers in the test prototype.

In the next chapter usability evaluation and the design and implementation of the related software prototype are discussed in detail.

4. Usability Evaluation

Usability refers to the quality of user's experience while interacting with products. Usability is not a single dimensional property of any product. It is a combination of Intuitive design, Ease of learning, Efficiency of use, Memorability and Subjective satisfaction ["Usability Evaluation Basics", Petrie and Bevan, 2009]. Though all these components should be assessed, Ease of learning – ability to quickly grasp how to use the system for the first time is most relevant in the context of our software prototype.

The ISO 9241 standard on Ergonomics of Human System Interaction (Part 11 1998) [ISO 9241-11, 1998] defines usability as: "The extent to which a product [service or environment] can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use." Usability evaluation focuses on how well the users can learn and use any software product to achieve their goal. It also depends on how satisfied the users are with the product.

This chapter is divided into two parts, User experience evaluation design and User experience evaluation execution. In the first part, objectives of the user experience evaluation based on the research questions (Section 1.4) are discussed. The second part of this chapter explains user test design and implementation of the evaluated software prototype. It also covers the process of usability evaluation planning and execution. The chapter explains the questionnaire included in the prototype, pre-test questionnaire and interview questions. It also includes the procedure of recruiting users, arrangements for the test venue and actual conduction of the user tests.

4.1 User experience evaluation for Phase 1 and Phase 2

The main objective of user experience evaluation was to assess how users perform in a 360° environment. Purpose of designing and implementing the software prototype was to find out answers for the research questions (Section 1.4) that were proposed.

Working on these research questions and going through the design and development cycles, we came up with the Phase 1 prototype (Section 3.2.2). After completion of the first phase user tests, results were obtained, and analysis of the results was done. Based on the analysis, the second phase (Section 3.2.4) was planned with the changes suggested by the users.

In order to plan the user experience evaluations, it was important to decide what output is expected through the user tests. The following results were expected from the user tests on a broad level:

1) Whether users feel it is useful and satisfactory to have a guidance to navigate in a 360° video?
2) Whether users are able to view all the important events or objects in a 360° video with the help of the guidance methods provided.
3) How do different types of guiding methods affect the users' immersive experience of the 360° video?
4) How useful are the different combinations of content types and guidance methods?
5) Can we confirm that some guidance methods with a specific type of 360° video is always useful?

Keeping the above questions in mind we designed the user evaluation questionnaires. The user test included a questionnaire before the user started the test, a questionnaire inside the prototype after each video, and a set of questions after the user finished using the prototype.

A brief questionnaire asking about the participant's background, name, age and familiarity with the technology was prepared (Appendix 1). The purpose of this questionnaire is to know the background of the participant. In the software prototype, each video had a set of questions at the end. These questions helped a participant to analyze the last video watched. The participants answered the questions on a scaled questionnaire having a range (1-7). In this '1' denotes a poor output and '7' denotes an excellent output.

A scaled question answering method was chosen for three reasons: it is *easy to decide*, it *can be done quickly*, the *total time needed to report about the guiding method experience can be very short.* This leads to acceptable total time taken by each participant to complete the entire test. Main benefit of using test questionnaire inside the prototype (Figure 4.2) was to reduce the need to remember the different guidance methods. It would not be possible for the participant to remember each video, the guidance method used attached to it and the resulting experience of it at the end of the entire test.

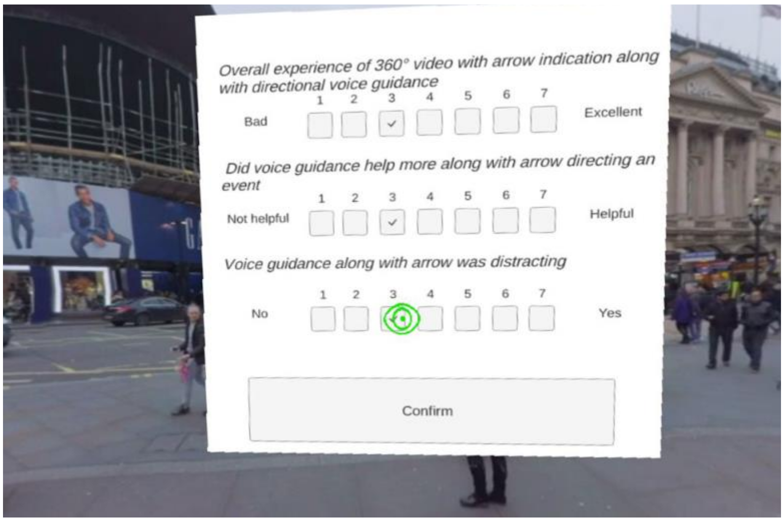

Figure 4.2. In-prototype questionnaire.

The questions presented inside the prototype (Figure 4.2) were formed to know about the overall experience of the video and the guidance method used in it, if the guidance method was useful, did it affect the overall user experience, and lastly how did the participant manage to explore the important events or objects in the video with the help of guidance method.

After testing the prototype, the participants were interviewed. To evaluate any type of a system, it is very important to receive first-hand information from the test participants. Interviews help in getting the deep answers or what the user actually thinks about the system which cannot be otherwise stated in the 'check-box' questions or the scaled questions. Interview can also open up some new dimensions which can contribute to future research work. Hence an interview session was conducted after the prototype testing.

The questionnaire was designed to ask user about the software prototype in detail. The prototype questionnaire gave us quantitative data about the overall experience of the prototype and the guiding methods. Detailed discussion about the prototype and the guiding methods was done through interviews (Appendix 2) to get the qualitative data.

4.2 Prototype implementation

The following section will discuss the prototype implementation considerations, how the prototype videos were arranged in a sequence to make it a continuous test and software implementation in detail.

4.2.1 Prototype implementation considerations

Designing the guiding methods was one part of the work and arranging them in the form of a working test prototype was another challenge. Initial challenge was to select a video and selecting a suitable guidance method based on the content of the video. We searched many videos on YouTube and downloaded them to use as example content in our prototype. The main purpose of using YouTube material was the availability of the 360° content. Another challenge was arranging them in a sequence so that a usable software prototype (Section 1.3) which feels like a continuous test and not a random collection of videos could be created.

We started with a simple method i.e. the arrow method (Section 3.2.2, I) which can guide the user to a direction to look for an interesting event or object. We increased the complexity of guiding methods by combining more than one method in a single video depending on the content of the video. After the combinations of guiding methods and the suitable 360° videos were finalized, the prototype was reviewed multiple times by us and analysed from the user's perspective to check the usability of the prototype for the testing purpose. When the prototype was reviewed some lessons were learnt which will be discussed based on the Neilson's usability heuristics guidelines [Nielsen, 1994]. These findings relate to our software prototype design which used HMD to view 360° videos. Some of them can be generalized to design a system using HMD to watch 360° videos.

It has been observed that though 360° videos are becoming common these days, *watching them on HMD is not very common* and using *HMD is not very common*. As this is a *head worn device*, it can *create pressure* on *head as well as neck* and *can be uncomfortable to wear for longer durations*. Wearing HMD and watching 360° videos on it for a longer duration can make user somewhat *dizzy and disoriented*. This is known as *simulator sickness* [Kauhanen et al., 2017].

Further adding to the findings, it can be stated that it is very important that the *user is well informed about the test and its details*. The user should know everything about the tasks that needs to be performed during the test as the *technology and devices can be new to the user*. Hence, a *set of instructions was provided* before actually starting to use the prototype. These instructions were provided in the prototype via a female speech synthesizer audio (Neilson heuristic principle: speak users language) [Nielsen, 1994]. *All these factors should be considered to simplify the usability of a prototype with efficiency of use* [Nielsen, 1994].

According to Nielson’s heuristic principles [Nielsen, 1994], *status information must be provided to the user along with the feedback to keep user informed about what is going on*. This principle is applicable to our prototype because, before starting each video with a different guiding method, the user will be informed about what is going to happen in the next video. Hence, there are instructions provided before starting each video and what user needs to do in the respective video.

There are buttons provided in the prototype to progress through the test. These buttons can be pressed by focussing one’s gaze on them and then tapping on the tap-pad of GearVR headset located on the right side. *This hand-eye coordination should be taken into consideration and should be clearly explained to the user*. The test prototype contains eight videos to test different guiding methods which is somewhat long duration. Hence, a break needs to be provided. *This break is important to relieve the pressure from the user both physically and mentally* (Neilson heuristic principle: Clearly marked exits) [Nielsen, 1994].

An immediate feedback for each guidance method must be received from the user before the test moves on to the next video. *If a user is asked for feedback about all the guidance methods at the end of the entire test, it is impossible for the user to remember everything about each method. Hence, a questionnaire was created for each guiding method which appears in the software prototype at the end of each video* (Neilson heuristic principle: Consistency) [Nielsen, 1994]. *The user needs to fill in this questionnaire at the end of each video. This will ease the user’s task of remembering everything from the watched videos reducing the cognitive load*. Hence the user need not remember the content and guidance from the first video till the last video (Neilson heuristic principle: Recognition rather than recall,) [Nielsen, 1994].

These design considerations from the iterative testing of the prototype helped to design a software prototype which was used by the users with ease and efficiency.

4.2.2 Final prototype

To test the effect of guiding methods while watching 360° videos, we designed a software prototype which contained a series of 360° videos with different types of guiding methods in them. We designed a set of guiding methods and applied them to eight different videos. In the first phase we took four different 360° videos and used one video for two different guiding methods. Hence the first phase, prototype had eight videos with different guiding methods.

The test started with a 360° video of 'Aurora Borealis' ["Aurora Borealis"] where the user is introduced to the test by giving instructions on what to do during the test. Instructions were given through a female speech synthesizer audio. Then the user was shown a 'Dali'["Dali"] video. Main purpose of this video was to make the user comfortable in a 360° environment. As 360° videos are not commonly watched on HMD, the user might forget to explore around a 360° video, i.e., watch in all the directions by rolling, pitching and possibly yawing the field of view. Hence this introductory video was shown to make the user familiar with the 360° video and allow the user to explore the possibilities to move around in the 360° immersive environment. It also helped to orient them to the task and get familiar with the viewing scenario. After this video that user starts the actual test. In the beginning of each video there is an introductory information given by the speech synthesizer audio, which tells the user about what can be expected in the coming video.

At the end of each video, the user fills in the questionnaire which is embedded in the software prototype. A scaled questionnaire ranging from 1-7 is used to give the answers to the questions. Selecting the answers from the questionnaire and moving on to the next video by pressing button was done by pointing gaze and single clicking the tap pad on the right side of HMD.

4.2.3 Software implementation

To implement the design ideas many things were taken into consideration. One is the choice of a suitable software development tool. Choosing correct technology can be crucial to actually implement the design ideas. Hence a combination of a new design idea with the correct choice of technology is very important to successfully develop a good working prototype. Our prototype was being developed for Samsung Gear VR which uses Samsung mobile phone, so Unity software development tool was used.

The application prototype was created to evaluate the effectiveness and usability of the guiding methods. The application prototype was made with version 5.6.1 of Unity3D Game Development Tool. The prototype uses part of the Unity3D base of the University of Tampere's Amaze360 project. The C# programming language was used to create interactions and events in the application. To display 360° videos, the application used the Unity3D Video player component that allows viewing 360° videos in Android applications. UI elements were made using Unity3D UI library components. The audio playback application uses the Unity3D Audio source component. The Oculus SDK library was used to create a virtual reality glass view. The application was made for

Samsung Gear virtual reality glasses and the application uses the touchpad of the device to allow the user activate user interface elements.

The application progresses step by step, and in addition to the start-up view, there are nine different steps. In the first video, the participant gets acquainted with the 360° video. In eight other sections, the participant uses the guiding methods. After each video there is a checkbox-based questionnaire (Figure 4.2), in which the participant can select an answer by pointing the gaze on the desired choice and pressing the touch pad of the Samsung GearVR. Between the videos, the buttons appear in front of the user's field of vision. Following Figure 4.3 is an example of the type of button which appears in the actual test prototype.

Figure 4.3. In-prototype button to navigate between the videos

The text of the button changes according to the videos as the test progresses forward. There are check boxes below the question and the user can point his/her gaze on a particular box and can select the option by tapping on the touch pad on right side of GearVR. At the end of the questionnaire, a button for the next video appears in the application and the speech synthesizer audio is repeated. This audio tells the participant about the next action to be performed.

The application control of the 360° videos and guiding methods has been implemented with control text files, where the event name, description, event time and event duration are stored. Each video in the application has its own text file from which control events are read. The application compares the event time of the control text file to the time of the 360°video. If the 360° video time is the same as the control text file, then the application activates the selected guiding method. The application will remove the previous guiding elements if the guiding has been active. The guiding event will be deleted after the event duration has expired.

The control text file was created using a separate Unity3D view, where 360° video is viewed by moving the mouse and pressing the right mouse button to mark the event to the file. The event mark is added to the indicated position in the 360° video. The second press of the mouse button marks the end of the event, storing the duration of the event, and an event form will open. The form is used to enter the name and description of the event. 360° video playback stops for the duration of filling in the form. The text file is updated with the event information. The control text file is exported to the application data folder on the Android phone memory from which the app installed on the phone reads the control text file.

The application has been built so that its content can be easily changed. For Test Menu and Questionnaire C# scripts are used in the application. Unity3D Inspector allows one to set the section guides, button texts, event list names and Questionnaires used in the application. One can change the number of events by changing the size of the list size.

This section introduced the application designed to evaluate the guidance methods. Firstly, what techniques were applied to the application and how the application progresses step by step was described. Next, we discussed how guidance control was implemented using control text files and how the control text files are created. Subsequently, we introduced how to easily customize an application content.

4.3 User evaluation execution

This section will describe how the user evaluations were executed in Phase 1 and Phase 2.

4.3.1 User test arrangements

After the development of the software prototype is complete, it is important to test the usability of the idea which was generated as some innovation to solve a particular problem. After the iterative testing the software prototype by us, a decision was taken to go ahead with the user tests. Planning the user tests was a big task for the successful result generation.

Planning a user test goes through many important steps such as designing the questionnaires by forecasting the expected results, designing interviews, fixing the number of participants, searching for the participants who are willing to participate in the test as the HMD technology is not very common to use, searching for the venue which can fulfil the requirements and making the participants comfortable.

I. Selecting the participants

Considering all these practicalities some professors were contacted who sent emails to various student groups from different courses. Few students showed willingness to participate in the user test. Some friends were contacted too. As there was not any particular requirement about the background of the participants, we got the users from varied background. In the first phase, 10 participants took the test. In the second phase, there were total 12 participants in the user test.

II. User demographics

The following tables show the demographics of the participants in Phase 1 (Figure 4.4) and Phase 2 (Figure 4.5).

User	Age	Gender	Use of Computer and other applications on computer	Previous experience of watching 360° video	Previous experience of using a GearVR
1	29	M	Daily	Yes	Yes
2	26	F	Daily	Yes	Yes for Games
3	33	F	Very Often	Some Promotional	No
4	23	F	Almost Everyday	Yes	Yes 2 times
5	40	F	Everyday	Yes	Yes
6	37	M	Daily	No	No
7	30	M	Always	Yes	Yes
8	40	M	Regularly	No	No
9	36	F	Regularly	No	No
10	25	M	Daily	Yes	Yes

Figure 4.4. Phase 1 user demographics.

User	Age	Gender	Use of Computer and other applications on computer	Previous experience of watching 360° video	Previous experience of using a GearVR
1	23	F	Always	Yes (YouTube/Laptop)	Yes
2	22	M	Everyday	Yes, Computer	Yes/ Familiar
3	18	M	Everyday	Once, on Phone	No
4	23	M	Everyday	Yes, Computer	Yes 2 times
5	25	F	Everyday	Yes, GearVR	5-6 times
6	21	M	Daily	No	Once
7	35	F	Daily	Once	No
8	21	M	Everyday	No	No
9	25	F	Daily	Yes, Computer	Yes
10	22	M	Everyday	Yes	Yes
11	22	F	Daily	Once	Once
12	24	M	Everyday	Not Sure	No

Figure 4.5. Phase 2 user demographics.

III. Test venue

All the tests were conducted in the University of Tampere campus and University Library Linna.

IV. Test environment Set Up

Test environment set up was the same for both phases.

We used

1) Samsung GearVR.
2) Samsung mobile S7.
3) 3D Noise cancelling Headphones.
4) Rotating Chair to sit in. (office swivel chair).
5) A mobile phone (To record the user interviews).

We performed two rounds of User Tests

1) Round 1: 10 Users
2) Round 2: 12 Users
3) Participants filled in a questionnaire inside the prototype related to each video.

The following Figure 4.6 will show the Test environment Set Up

Figure 4.6. Test environment set up.

4.3.2 User test procedure

The user tests were taken in a silent room in University of Tampere and University library. The participant was warmly welcomed and made comfortable. Later the participant was explained about the devices that will be used during the test, how to navigate throughout the test wearing the HMD and head phones etc.

The participant signed a consent form ensuring that their personal information and the data gathered from the test will not be used anywhere other than the theses. There was a short pre-test interview for each participant. The participant was allowed to take a break and take off the device during the test as the test was long. After the participant had finished the test and relaxed a bit, an interview took place to understand more about the experiences during the test and overall experience and opinions about the prototype and the new idea behind the prototype. A small reward of participation was given at the end of the interview. Post-test interview was audio recorded and some video and photographs were taken during the user test with the participant's consent.

I. Timeline of the User Tests

Both rounds were completed in a span of 10-15 days during December 2017 – February 2018.

From the self-testing it had been estimated that the entire test would need around 30-45 minutes based on the individual user's performance. Hence a slot of 1 hour was reserved for each user test. This was done by anticipating some delays during the test.

Each participant was given a small gift as a mark of thanks for participating in the user tests. This gift was sponsored by University of Tampere, Faculty of Communication Sciences.

II. Data Collection

Data collection was done in 3 steps. Pre-test questionnaire to collect basic information about the participant, in-prototype questionnaire and interview after using the prototype. On an average each participant took around 25-30 minutes to complete the test irrespective of the previous experience of using HMD to watch a 360° video. Total data collected through interviews in Phase 1 (10 participants) was 96.43 minutes which was on an average 9.6 minutes per participant. In Phase 2 (12 participants) it was 106.96 minutes which was on an average 8.9 minutes per participant. Later, interview transcripts were created for both phases which helped to analyze the data.

As mentioned before, the participant answered a questionnaire which was available in the prototype after each video. A log file got created in the software which contains the answers for each video. As shown above, scaled questionnaire data was collected in the log file. Data from these log files was then entered into an Excel file to view entire data at the same time for analysis. In addition to the log file, data was collected in the form of pre-test and post-test interviews. These interviews were recorded and then transcripts were created to extract data from the interview. This data was then entered into Excel files for further analysis.

In the two phases of evaluation significant amount of data was collected from the user evaluation tests. This was used for subsequent user evaluation findings and analysis.

5. Results and Analysis

This chapter will focus on the results from the user tests as well as the conclusions derived from the data collected. The chapter will show the findings from Phase 1 and Phase 2 separately which will help to understand which research question responses are consistent across both the phases and which are not. Data from the log files along with interview transcripts were studied and analyzed for Phase 1 and Phase 2.

5.1 Phase results

This section discusses in detail the results of user experience evaluation of Phase 1 in detail. The section describes the user experience evaluation with graphical representations for the properties of guiding methods that were evaluated with the in-prototype questionnaire.

5.1.1 360° video without any guidance

In phase 1, 10 users participated to test the prototype (Section 4.3.1). The data was retrieved through the in-prototype questionnaires presented after each video. The questionnaire contained scaled questions on a range from 1 = Poor output to 7 = Excellent output.

In the beginning of the interview, the participant was asked about the overall experience of watching a 360° video on a GearVR. This question was asked to know the user's comfort level in the test environment, as they were from varied backgrounds. Results from this question helped us know the effect of familiarity with the technology. Some responses from the interview for the overall experience question were *'Overall it was enlightening to see the places this way'* – a participant with some experience of 360° videos using HMD; *'It was nice, I got to see more than just the ordinary videos'* – a participant with no experience of watching 360° videos with HMD. Some participants did feel that it is very easy to get disoriented in the 360° environment, hence guidance is very important. Some of the remarks were ***'****Guidance is the key, as you can easily get lost in the 360° environment****'***– a participant with no experience of watching 360° videos and HMD.

In the first phase there were eight different guidance methods (Section 3.2.2). The participants were asked the questions about the overall experience, usefulness of the guiding methods, whether the method was distracting, or the method helped the user to locate all the important events happening in the video via in- prototype questionnaire (Figure 4.2).

Different Guidance Methods (Phase 1)	360° Video Mean (SD)	M1 Mean (SD)	M2 Mean (SD)	M3 Mean (SD)	M4 Mean (SD)	M5 Mean (SD)	M6 Mean (SD)	M7 Mean (SD)	M8 Mean (SD)
Overall experience	5.1 (1.10)	5.4 (0.96)	4.6 (0.69)	4.8 (1.31)	4.8 (1.03)	5.9 (0.99)	6.8 (0.63)	2.8 (0.78)	3.8 (1.31)
Guidance Helpful/ Easy/Useful	**(360° video Vs 2D Video)** 4.7 (1.25)	5.5 (0.70)	4.2 (1.13)	3.5 (1.26)	5.4 (1.71)	5.2 (1.31)	5 (1.76)	4.8 (1.54)	4.4 (1.42)
Could Locate All Important Events/ *Guiding method combination is distracting	**(Previous Experience)** 3.7 (2.31)	6.4 (1.26)	3.4 * (1.50)	5.6 (1.26)	5.4 (1.71)	4.2 * (2.04)	6.1 (0.99)	3.5 * (1.08)	1.9 (1.59)

Figure 5.1. Phase 1 results for all the guiding methods, Mean and SD values.

M1: Arrow guidance, M2: Arrow with directional Voice, M3: 3D treble sound with eye marker, M4: Speech synthesizer audio with hollow rectangle and eye marker, M5: Bull's Eye, M6: Bull's Eye with speech synthesizer audio guidance, M7: Automatic Transition with blur effect, M8: Automatic Transition with blur effect and eye marker.

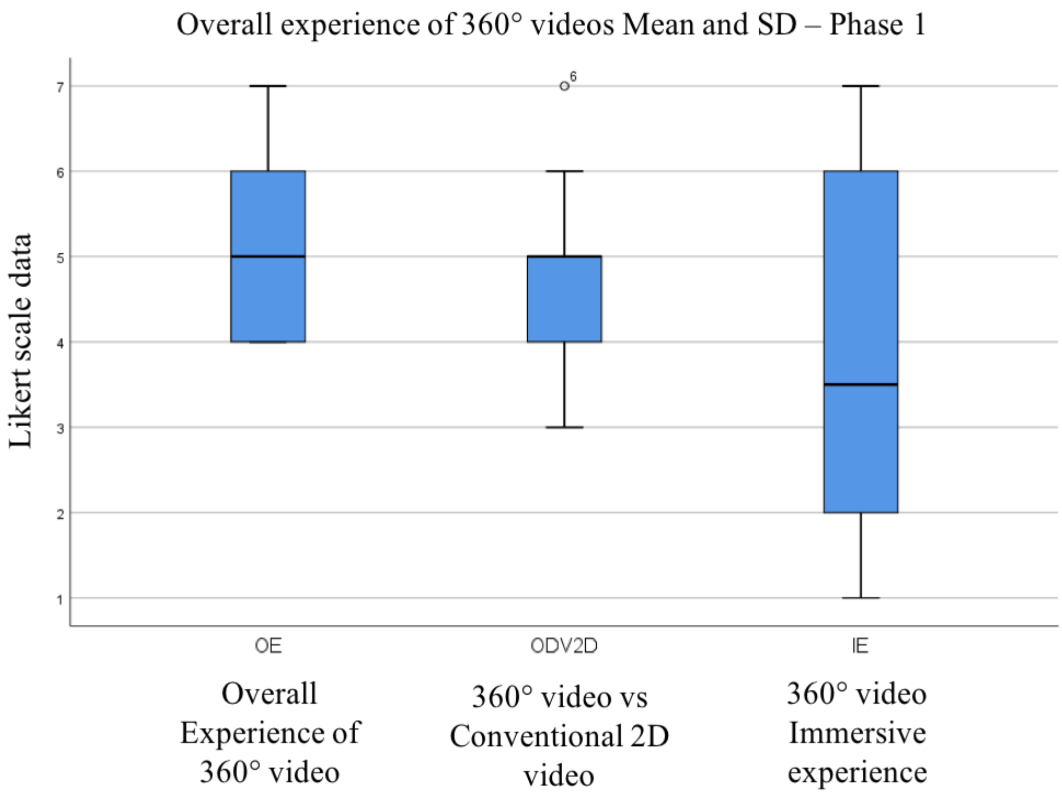

Figure 5.2 Phase 1- Overall experience of a 360° video.

Figure 5.1 and 5.2 explains the data which was gathered through the in-prototype questionnaire. Figure 5.1 presents the results for each guidance method in terms of overall experience and usefulness. Figure 5.2 is the graphical representation of user's overall experience of a 360° video, comparison of 360° video and a conventional 2D video and 360° video immersive experience.

In Phase 1, the test started with a 360° video without guiding methods. This video was included in the test to give the participant a feel of a 360° video. All the participants expressed that the overall experience of watching a 360° video on HMD was very nice irrespective of the previous experience of using technology. The graph (Figure 5.2) also shows the user experience of comparing 360° videos and 2D conventional videos. Overall the participants were satisfied with the immersive experience of a 360° video over a conventional 2D video. The following sections will address three properties which were evaluated inside the software prototype.

5.1.2 Overall experience of all guiding methods

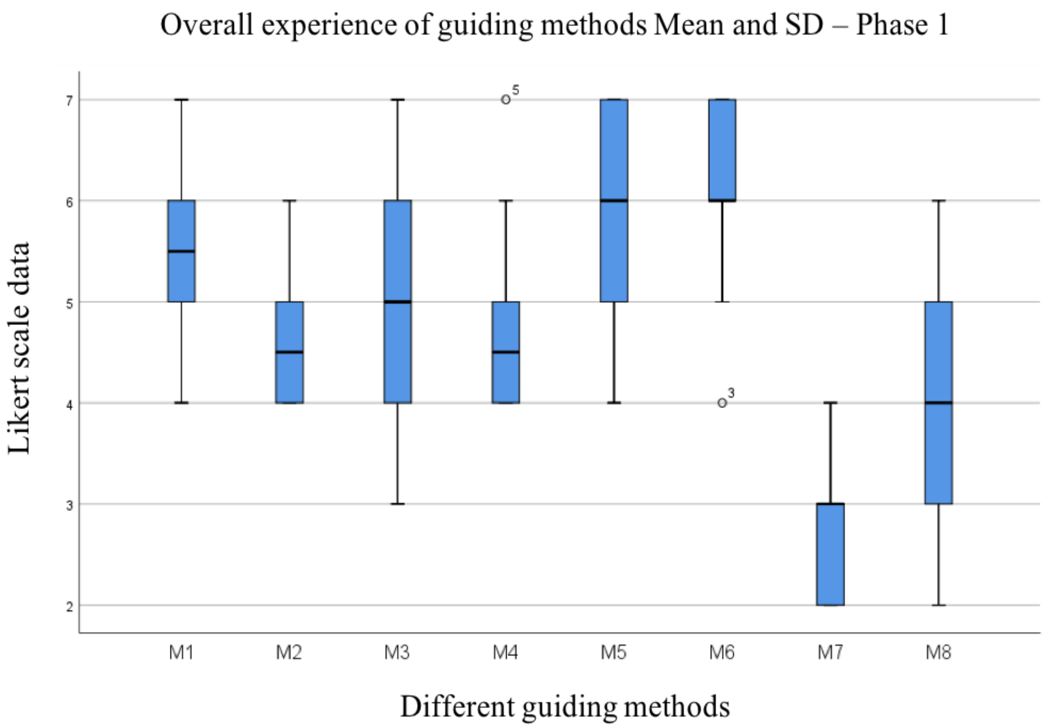

Figure 5.3. Phase 1 - Overall experience of all the methods.

Figure 5.3 is a graphical representation of overall representation of all the guiding methods from Phase 1. The first guiding method in the prototype was an *Arrow method (M1). The arrow method was the simplest method in the prototype.* This method was liked by majority of the users. This method was non-intrusive, so the participant got an immersive experience as if they were present in London. Few responses from the interview were *'Arrow was OK, it was not very Intrusive' and 'Arrows were probably the Best'*. Next method was *Arrow with directional voice (M2)* method. In this method a speech synthesizer audio is added to give directions to the user where a probable interesting event or object can be seen. This method used in the London sightseeing ["A London City Guided Tour"] video. This method was also liked by many participants except few who felt that the video narration and the speech synthesizer audio got mixed together.

The next method was *3D Treble sound with eye marker (M3)*. In this method a rectangular marker with the information text was shown to enhance user experience so that the user will be able to locate all the important visuals in the video. This method was moderately liked by the participants because they did not notice the 3D treble audio and the direction it was coming from. Hence, they could not make use of it. However, they used the speech synthesizer audio guidance (mean = 4.8). Overall the experience with this method was 'OK' with the speech synthesizer audio. Few responses from the interview were *'I did not notice the 3D sound at all' and 'Later on I realized that it was a 3D sound'*.

3D Treble sound with speech synthesizer audio guidance with rectangular marker (M4) was the next guidance method. In this method a rectangular marker with the information text was shown to enhance user experience so that the user will be able to locate all the important visuals in the video. This method was moderately liked by many users. Many participants could locate the rectangles, and few said that it was good to have rectangle markers to indicate an interesting field of view. One response was *'I really liked the rectangle markers especially because they also show the textual information about a particular field of view'.* The results indicate a good overall experience.

Moving further to the next method *M5,* a *Bull's eye mark* was provided as a type of guidance. From the results (Figure 5.3) it can be clearly seen that the *participants liked this method a lot in the overall experience though some participants suggested that the size of bull's eye mark was too big, so it was distracting user's view of the video content.* One participant response was *'The method is useful, but the size of bull's eye mark is very big'*. It was observed that the participants did not like the bull's eye mark, but the overall experience numbers are quite high (Figure 5.3). This can be interpreted so that the method is good but due to the size, it is not working optimally.

In the next video, bull's eye guidance was combined with speech synthesizer audio giving directions to the important events or objects (M6). The numbers from the table (Figure 5.1) and graph representation (Figure 5.3) show that the overall experience mean value for this method is 6.8 which says that this combination of guidance methods was liked by almost all participants. This guidance helped participants to find maximum number of interesting points which was air balloons in this case and voice guidance giving the direction assisted them to find the objects successfully. One participant response was *'If I completely loose the track to locate balloons then 'turn behind' voice guidance was very helpful'.*

The last two guidance methods M7 and M8 were almost rejected by the participants (see Figure 5.3). The feedback for these methods was given instantaneously after finishing the test due to their last position in the software prototype. The *Automatic transition with blur effect (M7)* method made the participants feel confused and dizzy. Many participants did not understand the purpose of blurring at the particular moments and the method created a load on user's eyes because of the close proximity of the blurred visuals.

In the *M8* guidance method, blurring effect was accompanied with an eye marker to indicate the important event or object. Due to the blurring effect the user could not concentrate on the eye marker. Few responses were *'Blurring effect removes the immersive effect of 360° video and makes you realize that you are watching something and not into something'* and *'It was unpleasant'.* Interpretation of these responses can be that blurring effect has removed the immersive effect resulting in a traditional video watching experience. The visual effects which may suit traditional videos best may not work with 360° videos.

5.1.3 Helpfulness, ease and usefulness of the guiding Methods

In this section, results for the usefulness of the guiding methods are presented. Some guiding methods are interpreted as easy to use and helpful in finding the interesting events or objects.

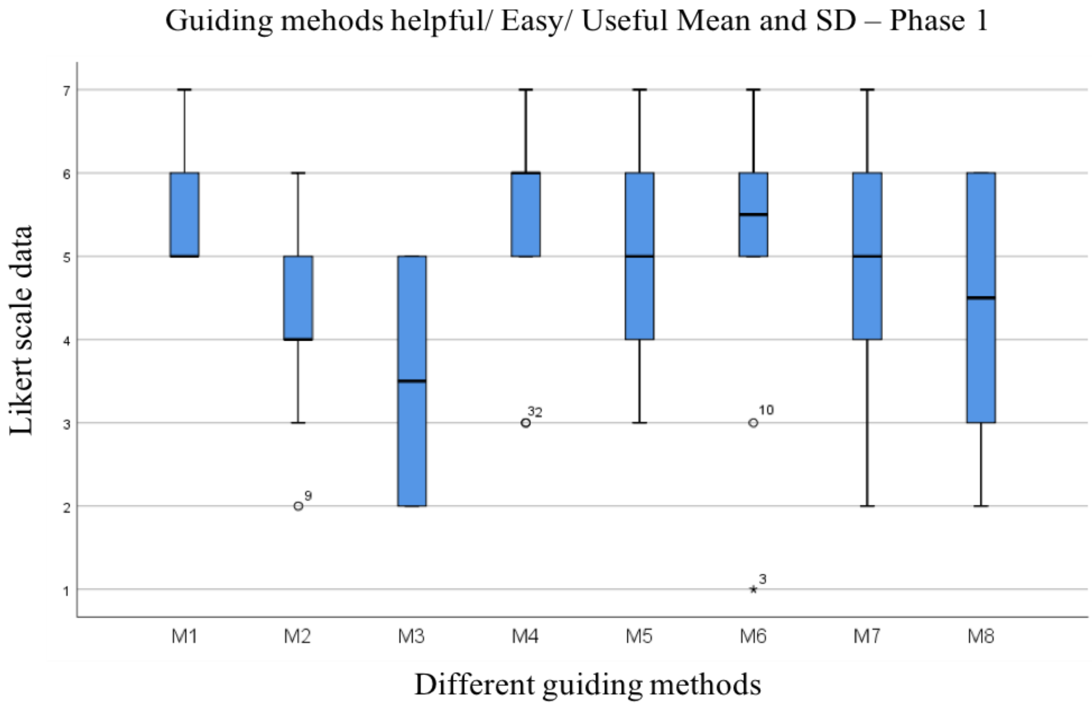

Figure 5.4. Phase 1 – Guiding methods helpful/ easy/ useful.

The first method in the prototype was an *arrow method (M1).* Arrow method helped participant to locate all the interesting points in the London sightseeing [“A London City Guided Tour”] video. Figure 5.4 shows that arrows were very helpful and easy to use and were pointing in the right direction where the interesting objects or events were present. Hence, they helped the user to find all the important events or objects in the video. One response from the interview was *‘In London video it was very important to point at the objects of interest with an arrow which the narrator is talking about’*.

In *Arrow with directional voice (M2)* method, a speech synthesizer audio is added to give directions was combined with the arrow method in the London sightseeing [“A London City Guided Tour”] video. Though some participants felt that video narration and speech synthesizer audio got mixed together, there was a positive feedback that only the ‘behind’ direction is useful as participants might fail to look behind (Figure 5.1, mean = 4.2, SD = 1.13). One response from the interview was *‘For me left and right became an intrusion but voice guidance for ‘Behind’ worked best’*.

A *3D Treble sound with eye marker (M3)* was the guidance method with the next video Angel Falls [“Angel falls, Venezuela”]. Many participants did not feel that 3D treble sound indicating the direction of some important event or object was helpful because they did not notice the 3D treble audio and the direction it was coming from. Hence, they could not make use of it. However, they used the speech synthesizer audio guidance (mean = 4.8). One response in the interview was *'I did not notice the 3D sound at all', 'Later on I realized that it was a 3D sound'.*

3D Treble sound with speech synthesizer audio guidance with rectangular marker (M4) was the next guidance method. In this method a rectangular marker with the information text was shown to enhance user experience. As seen from the above table (Figure 5.4) this method was moderately liked by many participants, many participants could locate the rectangles, and few said that it was good to have rectangle markers to indicate an interesting field of view. Overall, we got a mixed response for the usefulness of this method. Few responses were *'To locate all the rectangles some arrow should be provided because 3D treble sound is hard to notice many times'* and *'I didn't like the rectangles that much, I prefer to see the bigger picture'.* Hence the participant suggested guidance method M4 with some changes so that it can be helpful to locate the important events or objects.

Moving further in the next method *M5* a *Bull's eye mark* was provided as a type of guidance. From the results (Figure 5.4) it can be clearly seen that this method helped the participants to find the air balloons easily, though many felt that the size of bull's eye mark is very big. Hence there was a suggestion that if multiple arrows are provided then the user experience can be more immersive. One response was *'The method is useful but as soon as user starts to respond to the guidance it should disappear else it can be distracting due to the size'.*

In the next video, *bull's eye guidance was combined with speech synthesizer audio* giving directions to the important events or objects (*M6*). This guidance helped the participants to find maximum number of interesting events or objects which was air balloons in this case and voice guidance giving the direction assisted to find them successfully. A user response was *'If I completely loose the track to locate balloons then 'turn behind' voice guidance was very helpful'.*

The last two guidance methods M7 and M8 received mostly negative remarks from the participants. The *Automatic transition with blur effect (M7)* method made the participants feel dizzy. Many participants did not understand the utility of blurring at a particular moment and the method created a lot of confusion for some participants. Though the responses for this method in the interviews were negative, the scaled question numbers were opposite hence it is very difficult to draw any conclusion.

In the guidance method *M8*, the blurring effect was accompanied with an eye marker to indicate the important event or object. Due to the blurring effect the participant could not concentrate on the eye marker. However, some participants said that at eye marker helped them understand the important event or object though they could not concentrate on it well. Hence the graph shows that the method was moderately useful, but it can be because of the eye marker.

5.1.4 Locating all important events or objects

In this section, results for each guiding method are presented from the in-prototype questionnaire answers to indicate if each guiding method helped locating all the important events or objects or if the guiding method was distracting.

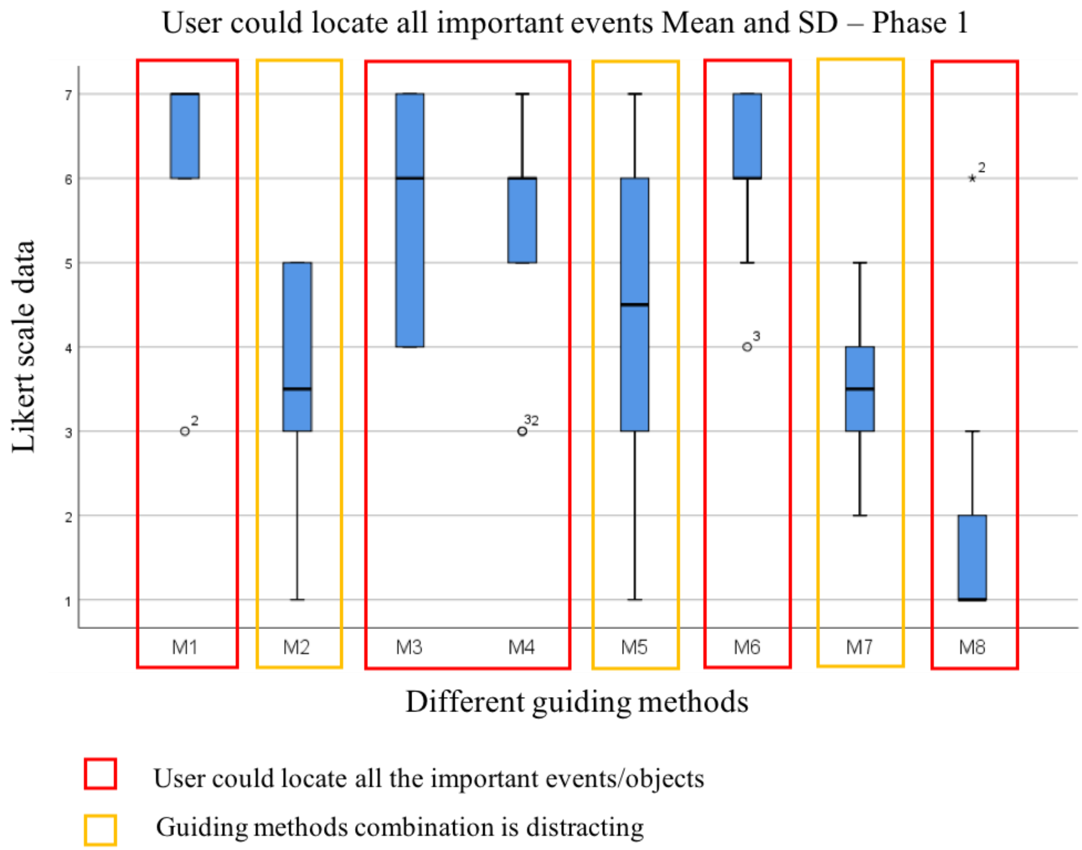

Figure 5.5. Guiding method could locate important events/ distracting guiding methods.

The first *arrow method (M1)* was the simplest of all methods and as seen from the Figure 5.5. It helped user locate all the important events or objects. This method was liked by many participants and it was non-intrusive. *Arrow with speech synthesizer audio giving directions (M2)* method was moderately liked by the participants. This method helped participants to find the important events or objects but at the same time the speech synthesizer audio and the video narration got mixed with each other. Due to this combination of video narration and speech synthesizer audio, some participants found this method distracting. However, the response values inside the prototype questionnaire were good for 'Behind' direction given by the speech synthesizer audio, as these events are always outside the field of view. In *3D Treble sound with eye marker (M3) and 3D Treble sound with speech synthesizer audio guidance with rectangular marker (M4)* methods, participants could locate most of the important events or objects because of the speech synthesizer audio and the rectangle marker. The *Bull's eye mark (M5)* method helped the participant to find all the important events or objects but received a slight negative feedback due to its big size. However, *bull's eye with speech synthesizer audio guidance (M6)* helped in a positive way due to the 'behind' direction given by the speech synthesizer audio. Hence this method received good response from the participants. As stated earlier in Section 5.1.3, guiding methods M7 and M8 received a lot of negative feedback but the quantitative data does not match the interview responses.

From the first round of the user tests, we found that there are some popular methods, some moderately liked methods and some unpopular methods based on the participants' responses. Hence, we decided to make few changes in the guiding methods and redesigned some guiding methods based in the Phase 1 user feedback, explained here (Section 5.1.1). By making these changes we modified the user evaluation prototype and the guidance methods before conducting Phase 2 user tests.

5.2 Phase 2 results

In Phase 2, 12 users participated to test the prototype (Section 4.3.1, I and II). Main data was retrieved through the in- prototype questionnaires presented after each video.

In Phase 2, while some methods from Phase 1 were kept the same, some methods were replaced with the newly redesigned methods (Section 3.2.3, Figure 3.14). *The main reason behind keeping some of the methods the same in Phase 1 and 2 was that the participants were different in the two phases hence, if the responses coming for the methods are somewhat similar then the findings for the common methods can be reassured.* As new participants were recruited for Phase 2, questions about the overall experience of watching a 360° video on GearVR were kept same. The main purpose of this question was to know the participants' comfort level in the test environment because they were from varied backgrounds. Results from this question helped us know the effect of familiarity with the technology. Some responses for the same question from the interview were *'I am feeling excited and It feels like I am in another world', 'It was an immersive experience and I felt that I was in waterfall'* and *'I have not used GearVR before still it was easy to use it and watch 360° video'.* Some users also felt that guidance in a 360° environment is very useful because there are many things which are happening at the same time, for instance beyond the user's field of view. The participants were familiar with watching conventional 2D videos in which everything that is happening in the video can be seen. In case of 360° videos, a user might forget to turn around in all the directions and watch everything from the 360° video. Some responses telling about this experience were *'Guidance in 360° video is useful as I might forget to turn around and explore the complete video'* and *'Guidance is helpful so that I can find the right place in the video'.* The following graph (Figure 5.6) shows the evaluations of overall experience of a 360° video as compared to a 2D video and the immersive property of 360° videos without any guidance.

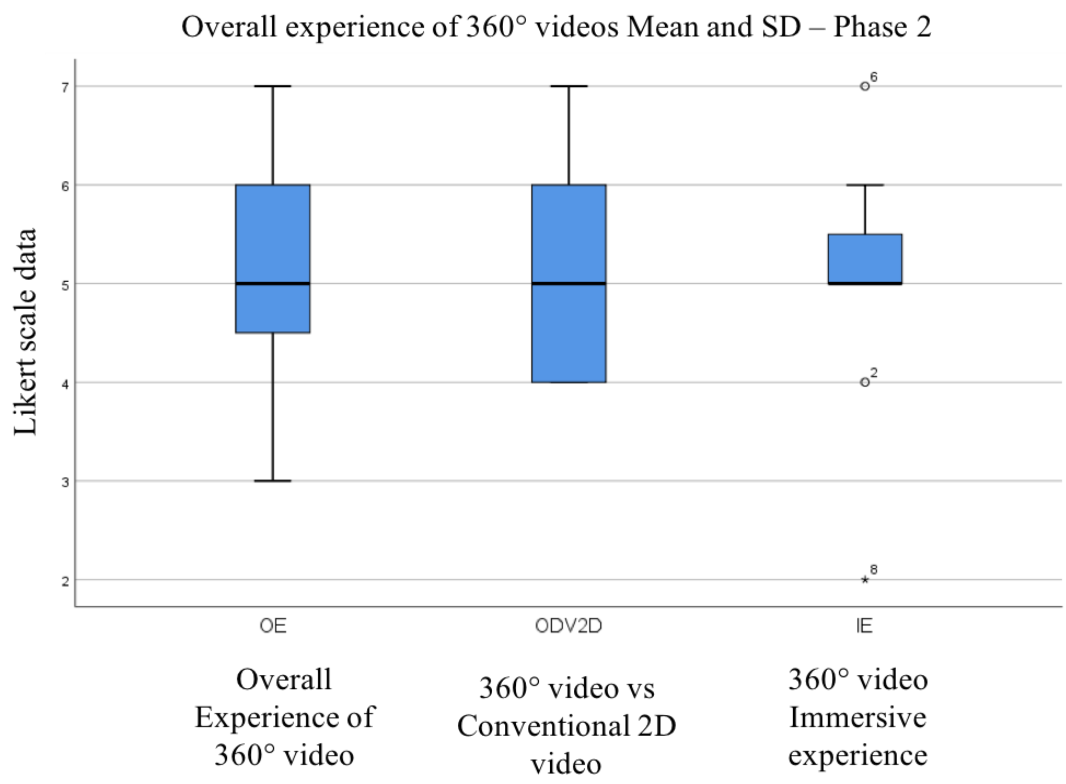

Figure 5.6 Phase 2 - Overall experience of a 360° video.

Phase 2 prototype contained eight videos (Section 3.2.3). In Phase 2, four new methods were introduced which were designed based on Phase 1 user feedback and four methods from the Phase 1 were kept unchanged. The following table (Figure 5.7) shows the changed methods in Phase 2 from Phase 1.

Phase 1 Guidance method		Phase 2 changed Guidance method
M2: Arrow guidance with speech synthesizer audio	Changed to	**M2: Arrow Guidance with speech synthesizer audio and textual markers**
M3: 3D treble sound with speech synthesizer audio guidance	Changed to	**M3: 3D bass sound with speech synthesizer audio guidance**
M5: Bull's Eye marker	Changed to	**M5: Directing arrows**
M7: Automatic transition	Changed to	**M7: Automatic transition with black screen cut**

Figure 5.7. Changed guidance methods from Phase 1 to Phase 2.

The following table (Figure 5.8) shows the results from round 2 of user tests. The results show the overall experience for all the methods and their usefulness.

Different Guidance Methods (Phase 2)	360° Video Mean (SD)	M1 Mean (SD)	M2 Mean (SD)	M3 Mean (SD)	M4 Mean (SD)	M5 Mean (SD)	M6 Mean (SD)	M7 Mean (SD)	M8 Mean (SD)
Overall experience	5 (1.33)	5.1 (1.37)	4.7 (1.49)	4 (1.63)	5.2 (0.91)	5.2 (1.31)	6.7 (0.67)	4.5 (1.35)	2.7 (1.25)
Guidance Helpful/ Easy/Useful	**(360° video Vs 2D Video)** 5.2 (1.22)	5.5 (1.08)	4.5 (2.12)	4.5 (1.43)	6.1 (0.87)	5.5 (1.58)	4.6 (1.89)	4.5 (1.35)	3.2 (0.91)
Could Locate All Important Events/ *Guiding method combination can be co-ordinated at the same time/# High Transition Speed	**(Immersive Experience)** 5 (1.33)	6.7 (0.94)	6.6 (0.96)	5.9 (1.37)	5.8 (0.91)	5.5 (1.58)	6.3 * (1.05)	4.4 # (0.51)	3.8 # (1.54)

Figure 5.8. Phase 2 results for all the guiding methods, Mean and SD values.

M1: Arrow Guidance, M2: Arrow Guidance with Directional Voice and Textual Markers, M3: 3D Bass Sound with Speech synthesizer audio guidance, M4: 3D Treble sound with Event Rectangle Marker, M5: Directing Arrows, M6: Bull's Eye with Direction indicating Voice Guidance, M7: Automatic Transition with Black Screen Cut, M8: Slower Automatic Transition with Eye Marker for Hotspot Indication.

As can be seen in the above table (Figure 5.6 and 5.8), the values for overall experience of watching 360° videos are very high which indicate that users were very much satisfied with the overall experience of watching a 360° video with an HMD. Regarding the immersive experience, we got many responses indicating that in many videos, the participants felt like they were in the location of the video, such as the heights of the waterfalls and volcanos. These responses were very interesting and satisfying. When we asked the participants about the 360° video 'without guidance' and 'with guidance', many said that 'with guidance' is a better option as the content is also available beyond the current field of view and there are high chances of missing content which might be very interesting. *However, almost all the participants insisted on subtle and non-intrusive guidance.* The following sections will address three properties which were evaluated inside the software prototype.

5.2.1 Overall experience of all the methods

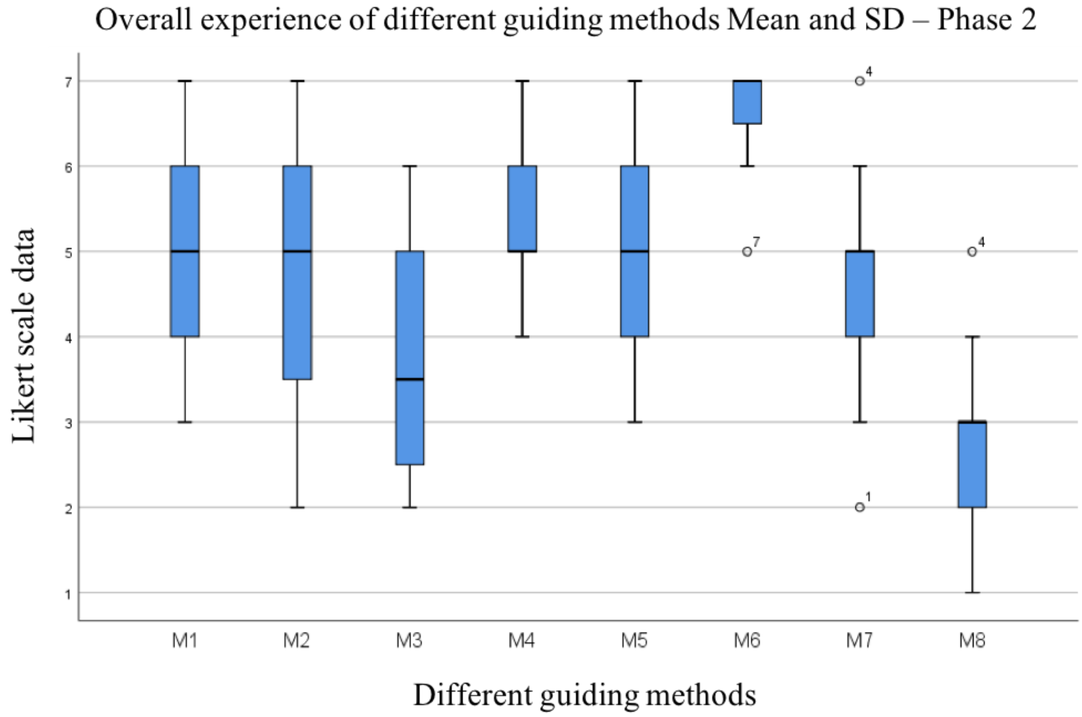

Figure 5.9. Phase 2 - Overall experience of all the methods.

This section will report the 'overall experience' of all the methods. Feedbacks for the first method *M1*, i.e. *arrow method* which was common in both the phases were almost the same. If we look at the values (Figure 5.9, Figure 5.3) for M1 in both the phases, they almost match each other. Hence from the values it can be interpreted that the *arrow method was the most liked method by all the participants*.

The next method M2 was different in Phase 2 from Phase 1 (Figure 5.7). In M2 during Phase 2, there were arrows and directional voice accompanied with textual markers giving information about the interesting object or event. In addition to the guidance method, the video was also changed for this method. Phase 1 video for M2 had its own narration ["A London City Guided Tour"] but in Phase 2 M2 method there was only music and no narration as the soundtrack ["Moscow Kremlin"]. The numbers from the table (Figure 5.8 and Figure 5.9) suggest that this method was also liked by many participants. Some responses for this method were *'I like the arrows with voice, which made me look and find what it is telling me to watch'* and *'It was a bit easier with voice and arrow with textual markers'*.

Guidance method M3 from Phase 1 was modified with a new video and introduced a 3D bass sound in combination with textual guidance (Figure 3.14). This textual guidance was information about the places in the video which portrayed as telling a story about the content of the video. As seen from the numbers (Figure 5.8 and Figure 5.9) this method was moderately liked by the participantsand 3D bass sound was better recognized than 3D treble sound. Some feedback comments were *'It was easy to locate the direction and sound'* and *'3D bass sound was a solid effect and also noticed the direction it was coming from'.*

Guidance method M4 was kept unchanged (Section 3.2.3, Figure 3.14) in Phase 2. As seen from the values (Figure 5.8 and Figure 5.9) the responses for the overall experience were almost same for both the phases. An example response is *'The base sound I didn't realize but high pitch sound was prominent'.*

In Phase 1, we got the feedback that bull's eye was distracting, hence we changed bull's eye guidance method to *directing arrows* in *M5* (Figure 3.14). The video for this method was changed. We used a synthetic content video ["Space Dream 360°"]. There is significant difference in numbers for overall experience (Figure 5.9), the participants liked the directing arrows more. Few related responses were *'Yes, directing arrows were helpful and I saw all the objects'* and *'It was helpful, but it was blocking my way somewhat'.*

M6 method having air balloon video ["The Golden Ring of Russia Air-Balloon Festival"] with bull's eye and directional voice guidance method was kept same in Phase 2 to check if we get the same responses. As expected, it can be seen from numbers (Figure 5.3 and Figure 5.9) in the graph feedback was almost the same. The participants liked the overall experience of the video and the guiding method and the directive voice helped to enhance the user experience.

Based on the Phase 1 user feedback for automatic transition guidance method M7 (Section 5.1), changes were made to achieve a better user experience. Blurring effect was replaced by a black screen cut to change the user's field of view (Figure 5.2). With a more interesting content of Kamchatka Volcano video ["Kamchatka Volcano Eruption"] black screen cut guidance was provided with some textual information about the content of the video. As seen from the numbers and graph (Figure 5.3 and Figure 5.9) there was a significant difference in the overall experience of the changed method. Blurring effect was giving a dizziness to the users. Black screen cut method minimized the dizzy feeling, but the users felt uneasy due to high transition speed of the black screen cut. Hence the responses included *'Black screen cut method is OK but did not enhance user experience as it should'* and *'It is somehow hindering my freedom of watching the video'.*

The last video in the prototype was kept the same having automatic transition with blurring effect and an eye marker to point at the interesting objects (Section 3.2.3, Figure 3.14). The transition speed was reduced to see the effect on the user experience. In this phase, the participants did not like the guidance method either. As seen from the table and graph (Figure 5.8 and Figure 5.9) the numbers are not very high. Hence it can be concluded that M8 guidance method was not liked by the participants. Some responses were *'It's a sudden change that makes me dizzy'* and *'Automatic transition is OK, but it breaks my flow and immersive experience'.*

5.2.2 Helpfulness, ease and usefulness of the guiding Methods

In this section, results for usefulness of the guiding methods are presented. Some guiding methods were considered as easy to use and helpful in finding the interesting events or objects.

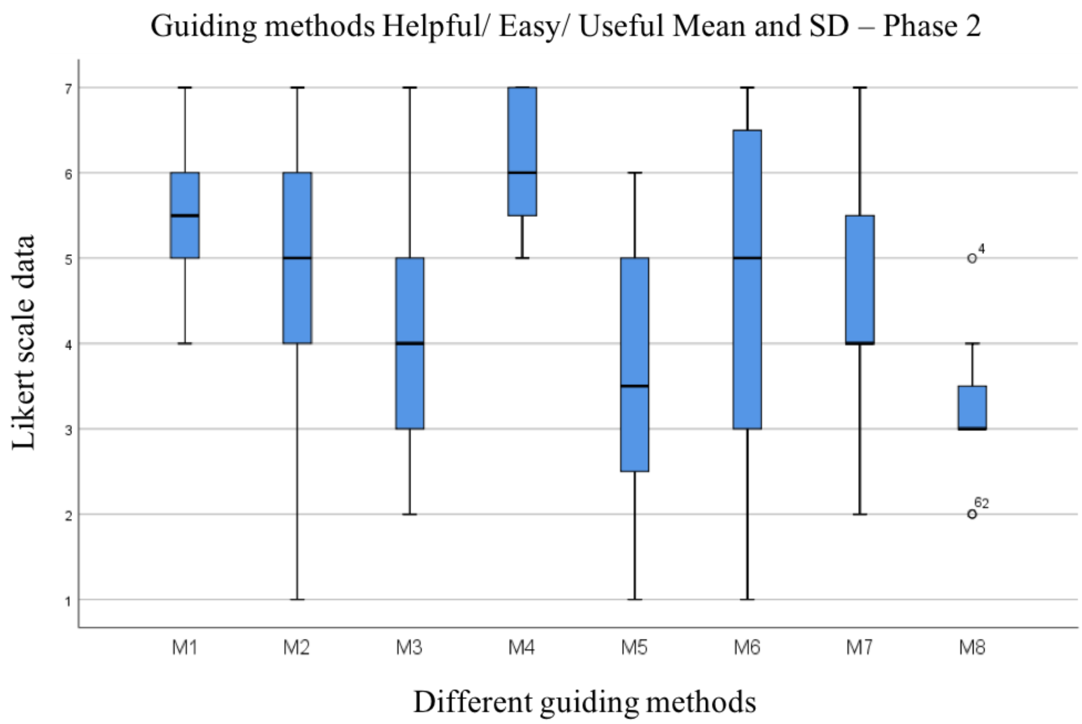

Figure 5.10. Phase 2 - Guiding Methods helpful/easy/useful.

The graph shows that *arrow method (M1)* was very useful. Both the phases show almost similar graphs (Figure 5.4 and Figure 5.10) which confirms that the arrow method is a simple and useful method which helped the participants find the important events or objects. The next method *M2* having arrows and directional voice accompanied with textual markers giving information about the interesting object or event in a [“Moscow Kremlin”] video helped the participants in finding the important events or objects. The graph shows that this method got good response (mean = 4.5). The next method was changed from 3D treble sound to 3D bass sound with a new video of [“Plitvice Lakes in Winter, Croatia”]. The story telling approach and 3D bass sound helped the participants positively to navigate in the 360° video (Figure 5.10). The participants recognized 3D bass sound better than 3D treble sound. In the case of *M4* method, which was same as Phase 1, the participants found the rectangles more useful to locate the important events or objects as seen in the figure 5.10. Method *M5 with directional arrows* helped participants to locate important events or objects and received a good response for the usefulness property of the guiding method. *Bull’s eye method with speech synthesizer (M6)* method was kept same in Phase 2. In this phase too, the participants used the guiding method efficiently to locate the important events or objects and it received good responses from the participants as seen in the Figure 5.10. In Phase 2, *blur effect of automatic transition* was replaced by *black screen cut (M7).* This change brought more positive responses from the participants and showed a positive change in the numbers (Figure 5.8 and Figure 5.10). The last method *automatic transition with blur effect and slower transition speed (M8),* received the same negative feedback from the participants which is reflected in the Figure 5.10.

5.2.3 Locating all important events or objects

In this section, results for each guiding method are presented which indicate if each guiding method helped locating all the important events or objects or if the guiding method combination helped the participant to coordinate between the guiding method and finding important event or object.

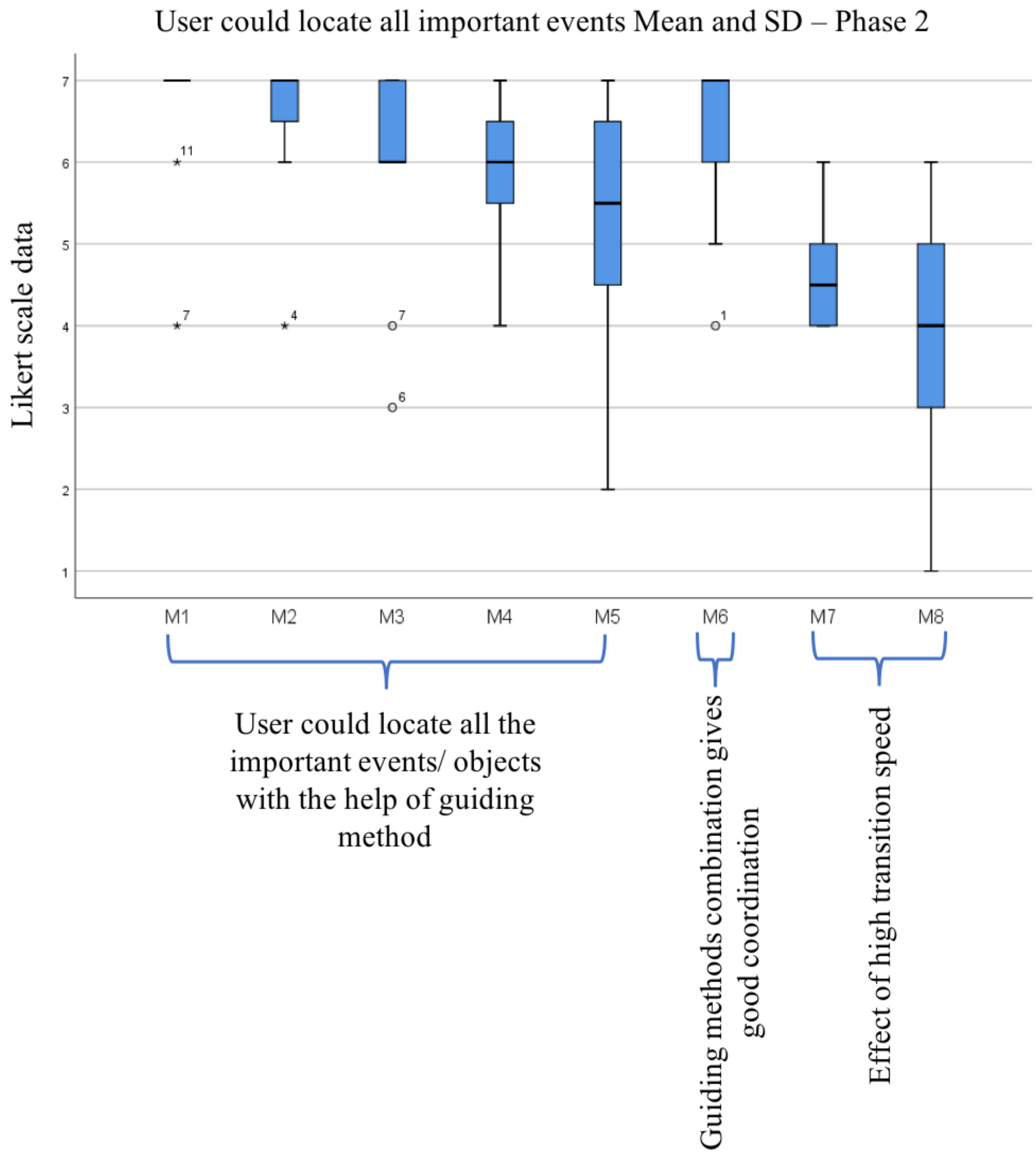

Figure 5.11. Phase 2 – Guiding method helps locate all important events or objects.

As seen from the Figure 5.11, guiding methods combinations in methods *M1, M2, M3, M4 and M5* helped the participants to find all the important events or objects resulting in a good response. *Bull's eye with speech synthesizer audio method (M6)* got the user response that the guidance method combination was good and helped user achieve good coordination in following the guiding method to locate all the important events or objects. Hence the participants could locate most of the important events from the video but found the bull's eye mark large which obstructed the user's field of view. Methods having *automatic transition with black screen cut (M7)* and *blur effect with slower automatic transition (M8)* show somewhat similar response about the transition speed that it must be slow. (though the speed was slower than in Phase 1). Hence the responses were good for M7 with a suggestion of slower transition speed and M8 method still received negative responses for the blur effect and transition speed as seen in the Figure 5.11.

5.2.4 Comparison between methods changed between Phase 1 and Phase 2

The following table shows the comparison between the methods which were changed from Phase 1 to Phase 2.

Different Guidance Methods (Phase 1)	M2 Mean (SD)	M3 Mean (SD)	M5 Mean (SD)	M7 Mean (SD)
Overall experience	4.6 (0.69)	4.8 (1.31)	5.9 (0.99)	2.8 (0.78)
Guidance Helpful/ Easy/Useful	4.2 (1.13)	3.5 (1.26)	5.2 (1.31)	4.8 (1.54)
Could Locate All Important Events/ *Guiding method combination is distracting	3.4 * (1.50)	5.6 (1.26)	4.2 * (2.04)	3.5 * (1.08)

M2: Arrow Guidance with speech synthesizer audio
M3: 3D Treble Sound with speech synthesizer audio
M5: Bull's Eye Marker
M7: Automatic Transition

Different Guidance Methods (Phase 2)	M2 Mean (SD)	M3 Mean (SD)	M5 Mean (SD)	M7 Mean (SD)
Overall experience	4.7 (1.49)	4 (1.63)	5.2 (1.31)	4.5 (1.35)
Guidance Helpful/ Easy/Useful	4.5 (2.12)	4.5 (1.43)	5.5 (1.58)	4.5 (1.35)
Could Locate All Important Events/# High Transition Speed	6.6 (0.96)	5.9 (1.37)	5.5 (1.58)	4.4 # (0.51)

M2: Arrow with speech synthesizer audio and Textual Markers
M3: 3D Bass Sound with speech synthesizer audio guidance
M5: Directing Arrows
M7: Automatic Transition with Black Screen Cut

Figure 5.12. Comparison between guidance methods changed between Phase 1 and Phase 2.

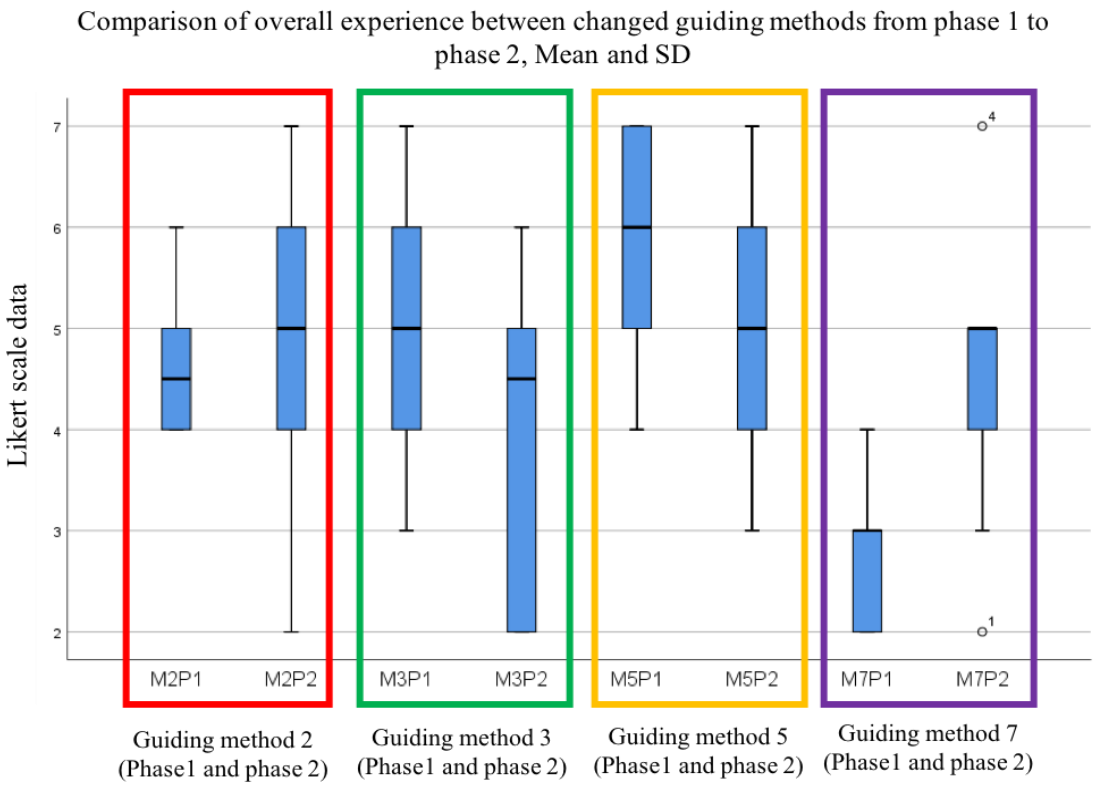

Figure 5.13. Comparison between the changed methods from Phase 1 to Phase 2 for overall experience.

As seen here (Figure 5.12 and Figure 5.13) a significant change in the user experiences can be found with the changed methods from Phase 1 to Phase 2 (Figure 5.7). Method M2 where the video was changed having only music from the video having narration in video, the overall experience improved (Section 3.2.3, Figure 3.14). The participants felt the guidance was more effective because video narration and speech synthesizer audio giving directions in the video did not mix with each other. This made the guidance clear and helped the participants to follow the directions. Hence significant improvement in the guidance method performance was achieved.

Method M3 was changed from a 3D high pitch (treble) direction giving sound to 3D Bass sound (Section 3.2.3, Figure 3.14). Change in the pitch of the guidance sound helped the participants to locate the sound and make use of it. As this video ["Plitvice Lakes in Winter, Croatia"] had narration about the places in the video, the contrasting low pitch bass sound became more noticeable. Hence this change made a significant difference in the user performance. This change is not reflected in the graphs because the rectangle marker in M3P1 gave more positive responses than the 3D treble sound whereas 3D bass sound was more liked by the participants over 3D treble sound.

Method M5, in which Bull's eye mark was obstructing the user's field of view was replaced by directing arrows that lead the user to a particular object in the video (Section 3.2.3, Figure 3.14). Though from the user responses, there were some changes in user performance in Phase 2 as compared to Phase 1, the results from the table and graph (Figure 5.12 and Figure 5.13) above show that there was no significant difference in the values for overall performance and usefulness of the guiding method. Therefore, bull's eye method and directing arrows method cannot be compared.

The last method changed was method M7 in which automatic transition with blur effect was replaced with a black screen cut to change the user's field of view (Section 3.2.3, Figure 3.14). As seen from the table above (Figure 5.12 and Figure 5.13) there was a significant improvement in user performance. The main reason for the performance improvement was that the participants felt less dizzy during the field of view change. Though there was some negative feedback and the participants did not like M7 in Phase 2 (Figure 5.2), the performance was better than M7 method in Phase 1 (Figure 5.13).

The changes done in few methods in Phase 2 indicate positive impact on user experience for some methods while some of the changes did not gain much improvement in the user performance (Figure 5.13).

5.3 User evaluation results analysis

In this section analysis of the user test results will be presented for both the phases. The process of analysis takes into consideration the user demographics and in-demo questionnaire feedback. The values from questionnaire have been processed to calculate Mean values for each method and standard deviation. Lastly, the feedback from post-test interviews is included. *Some general findings after analyzing both the phases are as follows:*

After successfully completing the user tests for Phase 1 and Phase 2 and finding out the results (Section 5.1 and Section 5.2), it can be said that all the participants liked our test prototype. It was clear from the user responses that guidance is very important in watching a 360° video on HMD as in 360° environment there are many things happenings around the user which are out of the current field of view. To capture these important events a proper guidance method should be used. Guidance method should be subtle and improve the immersive user experience.

Amongst the participants not everyone had a previous experience of watching a 360° video or using an HMD. While some had watched a 360° video before, not every one of them have watched it on an HMD. Hence familiarity with the technology was different for different participants. Considering this fact, all the participants successfully used the devices and worked with the prototype fluently once they were explained the technique. For all the participants the *overall experience* was exciting, and they were thrilled to watch some videos like angel falls videos where camera goes to a high altitude and deep jungle is seen from there.

Some participants complained about the pressure which GearVR creates on the neck. All the participants took breaks which were provided in the prototype software during the test. There was feedback that it is somewhat difficult to continuously wear the device.

'Audio guidance' is an innovative guidance type which was introduced in our prototype (Section 3.1.1). 3D voice guidance was used as a tool of guidance with some videos. 3D treble sound or high pitch sound was provided as a guidance to indicate the direction where some important event will be present in the video. This concept was totally new and went unnoticed for the few participants. Only some participants recognized the sound and could interpret it as a way of giving guidance to indicate the direction of probable important event or object. Some participants also missed this 3D audio

guidance on the first occurrence, but they realised the guidance sound on the second occurrence and used it to explore the video content.

In Phase 1 and Phase 2 textual guidance was given in all the methods. ‘Guidance Available’ text appeared at the bottom of the screen whenever any type of guidance was given to the user. In both the phases when the participants were asked about it, it was observed that many of them did not notice it. However, some noticed and used it. Some participants said that they actually made use of the text to realize that some guidance will be given to them as the text appeared on the screen. Hence there were mixed responses from the participants about understanding the purpose of the text appearing on the screen. Some said it was useful and they actually used it to realize there is guidance and some said that it was not at all needed because on HMD already the screen is small and there is no time to read.

The *overall experience* of watching 360° videos was very interesting and enlightening for all the participants. The participants felt as if they are in a different world. Some of the guidance methods helped the participants to have an immersive and undisturbed pleasant experience.

A common analysis drawn from the interviews from Phase 1 and Phase 2 was that all the participants stated that *the type of guidance method depends on the content of the video*. Even the speed or pace of the video can have an impact on the usefulness of a guidance method. Hence which type of guidance method is suitable for what type of video content is a topic that needs to be further studied in detail.

5.3.1 Results analysis for Phase 1

Considering all the results of guidance methods in Phase 1 (Section 5.1), it was found that the *arrow only* method was the most popular method which helped the user in a subtle way unobtrusively in locating all the important events happening around a 360° video. It was observed that once the arrow showed a direction to the user, it did not disappear from the screen. Hence the user felt that, the guidance should be provided once and when the user follows the guidance, then the arrow mark should disappear.

The same analysis is applicable for *Arrows with voice guidance* method. This method was appreciated by some users because the speech synthesizer audio feels different than the normal voice and it is easy to recognize. However, some felt that the directional voice guidance is disturbing or influencing user freedom. It felt distracting because the video had narration in itself, so both voices interfered with each other hampering the user experience. There was a suggestion that *Arrow with voice* guidance should be applied to the video which has only music. This change was further applied to one of the videos in the Phase 2 (Figure 3.14). User feedback (Section 5.1) also suggested that voice guidance was repeating till the user actually turns into the guided direction. This feels annoying and needs to be changed.

3D sound with rectangle marker was moderately liked by the participants. Some participants felt that they were being forced to watch a limited content and there is no freedom. Only a few participants found it useful. The rectangle markers were appreciated by some of the participants, but some did not understand its purpose.

Bull's eye marker method was the second most popular method (after arrows) amongst the participants but majority of them complained about the size of the marker. User feedback also suggested that a smaller icon and the voice guidance will work best.

Automatic transition with blur effect to change the user's field of view method received most negative comments as the forceful field of view change made user dizzy and hampered the overall immersive experience of watching the 360° content. This method confused the participants and felt that their *freedom of watching the content according to their wish is overridden*. The user suggested to *lower the speed of transition*.

Analysis of all the methods from Phase 1 lead to redesigning of some methods and formation of Phase 2 prototype. It was realized that specific type of guidance method should be based on the content of the video. Hence in Phase 2 different content was used for some methods to see the change in the user experience.

5.3.2 Results analysis for Phase 2

From the experience of Phase 1 user tests, Phase 2 had some base information to analyse the results from the user tests (Section 3.2.3). In Phase 2 the main concerns of analysis were the results for the methods which were modified or changed in Phase 2 from Phase 1 and the overall user experience for all the methods. The technology was very much new for most of the users. However, all the users performed well irrespective of the previous knowledge about a 360° video and HMD. Still, it was felt that it affected the depth of answers received in from the interview.

Arrows method was the most popular method in this phase too. In Phase 2 prototype, we changed the video content of *arrow with voice guidance* method based on the feedback from Phase1. It was important to see the feedback in Phase 2 for this guidance methods combination. Though the visuals from the graph (Figure.5.3 and Figure 5.9) appear similar, verbal responses were more significant and detailed. Many participants liked the arrow guidance with voice. In fact, some participants responded that *voice guidance was more useful along with the textual information which enhanced the user experience.*

3D bass sound guidance with a story telling approach created some difference on identification of the guidance in the video. *Participants liked the overall experience and along with the bass sound directive guidance textual markers added more to the video making viewing an immersive experience.* It was observed that the *participants noticed low pitch sound better than the high pitch sound*. Hence results for M3 method in Phase 2 were more positive than in Phase 1.

In the next method M4 the rectangle showing the field of view was appreciated by many participants, hence numbers from the table show some improvement in the overall experience and the participants could locate most of the interesting events or objects in the video. *Though some participants did not notice the treble sound, the rectangles helped to locate the important locations*. Hence in Phase 2 this method was well received by the users.

Based on the feedback for M5 method a new method having only directive arrows with a new video was introduced in Phase 2. *This method replaced bull's eye method to check if the size of guidance marker can improve the overall experience and unobtrusiveness*. It can be seen from the Figure.5.13 that though the overall experience values were almost same for both the phases, the values for distraction caused by the guidance showed some improvement compared to *only directing arrows* (Figure 5.5 and Figure 5.11). The directing arrows had a fading effect while approaching the interesting object or event in the prototype. This helped the participants to locate the interesting object or event without distractions.

M6 method which was same for both the phases. It received almost same results. The interview responses indicated that the bull's eye mark size was too big and there were too many arrows. Some participants also liked the voice guidance.

Based on the user feedback in Phase 1, M7 method was changed to automatic transition with black screen cut from automatic transition with blur effect to change user's field of view. Values from the Figure 5.13 show that black screen cut received more positive feedback than the blur effect. Blur effect with automatic transition made user dizzy. This might happen due to close proximity of HMD lenses. Therefore, black screen cut was more appreciated by the participants in Phase 2. Textual markers giving information about the important events or objects added more usefulness to the method.

The last method M8 having slower automatic transition with blur effect received the same feedback as in Phase1 (Figure.5.3 and Figure 5.9). This method was not received well by the participants.

5.4 Design implications

The following are the guidance method design implications that were formed based on the results and analysis of the two-phase evaluation study. These design implications should be considered in reference with viewing a 360° video on HMD:

Subtle and non-distracting guiding methods: 360° videos provide a great all-around view to the user. Hence the user is always interested in watching the interesting content shown in the 360° video. If there are too many visual cues provided in the video to guide the user, then user can easily get distracted and this affects the immersive experience of the 360° video. For example, Arrow method came up as the most subtle, unobtrusive and non-distracting method in the user experience analysis (Section 5.1 and Section 5.2). This also means that too much of guidance should be avoided. For example, in the [“The Golden Ring of Russia Air-Balloon Festival”] video size of bull’s eye mark was very big, which blocked the field of view; users got guided in a useful way but method itself received some negative feedback.

Type of guidance method depends on the content of the video: This is a very important design implication. It is essential that guiding method and content of the video should be compatible. For example, [“A London City Guided Tour”] video had the human narration in the video. When the video was combined with the speech synthesizer audio guidance, human narration in the video and the speech synthesizer audio got mixed which obstructed the immersive experience of the video. At the same time when the speech synthesizer audio guidance was combined with the [“Moscow Kremlin”] video which contained only music in the video, it helped the user to find the off-screen targets. A detailed study to find out which type of guiding methods are suitable for what type of content must be done to create more concrete design implications.

Guidance should be limited to reach the off-screen targets: When deciding the guiding methods for a 360° video, it is very important that the guidance method should be visible for a time duration that will not interfere with the immersive experience. This means that any type of visual cue or audio cue should be present in the video for a limited time. For example, in the directing arrows method the arrows were present on the screen till the target was reached. Later on, arrows vanished from the screen. This will maintain the immersive experience.

Audio guidance should be distinct from the content audio: It is very important that audio guidance must be distinctly audible from the audio of the video content. Otherwise it becomes difficult to recognize the difference between them. If such situation arises, the user is unable to hear the difference. For example, in the [“Angel Falls”] video, 3D treble sound was attached to show the direction of the off-screen targets. However, this sound got mixed with the audio of the video content. Hence the users could not notice it and the purpose of providing the guidance was wasted.

Repetitive guidance is not recommended: This is a very important design implication. If the guidance is repeated multiple times, then the user feels a compulsion for doing some action. Some reaction time should be given to the user to make use of the guidance. For example, in arrow with speech synthesizer method the user received directional audio guidance till the user turns and views in the specified direction like ‘Behind’. This continuous audio guidance annoyed the users.

Transitions in the field of view must be slow: If there are automatic transitions offered as a method of guidance in a 360° video, it should be taken into consideration that the speed of the transition is slow while changing the field of view. When the user is watching a 360° video on HMD, HMD lenses are very close to eyes. Hence, user might feel dizzy if the speed is very fast. For example, consider the automatic transition with blur effect and automatic transition with black screen cut from the software prototype (Chapter 3).

User must be informed about the guidance in advance: It is very important that the user is informed well in advance if some change is going to happen automatically in the current field of view. Otherwise the user feels confused and disoriented. For example, consider automatic transition with blur effect and automatic transition with black screen cut from the software prototype (Chapter 3). Burring effect made the user confused and the black screen cut came as a surprise to the user.

6. Conclusion and Summary

A 360° video watched on HMD can give a full immersive experience and make the user feel like they are in the place or situation which the video content is about. This fully immersive experience is enhanced if the user is able to enjoy all the important and interesting events or objects in the video. If some of them are missed by the user, then the purpose of the 360° video is not fully accomplished. In this thesis, this problem was addressed, and the solutions were proposed to achieve good results.

Conventional 2D videos have a limited field of view. Whatever the user records can be viewed in the current field of view. Hence conventional videos have a limited focus area and content. However, in a 360° video this limitation is overcome by giving a 3-dimention freedom where the user is able to enjoy the content in multiple directions i.e. roll, pitch and yaw (Chapter 1). In 360° videos the content is always available beyond the current field of view. Hence, this content is missed if the user does not know the direction or the field of view of it. To assist in enhancing the viewing experience, the user can be provided with guidance methods while navigating in a 360° video.

This thesis started with the basic concepts related to 360° videos, 3 DoF, HMD and viewing a 360° video with HMD. The problem handled in the thesis was explained in detail with references to previous work done in this area handling similar problems related to 360° videos. The idea of this thesis is to provide some novel solutions which will guide the user in an unobtrusive way and will help the user explore the content around the 360° video. To solve this problem different guidance methods were proposed, and they were tested with a number of users from varied backgrounds.

The guidance methods were designed with minute detail and matched with a variety of video content to make video and guidance method combinations useful. In many videos different guidance methods were combined together to positively enhance the user experience. It is not sufficient to design the prototype on paper but with constructive design approach, the ideas were implemented to make a working software prototype. Before testing this prototype with the users, self-testing was done iteratively to make sure the prototype can be presented to the users.

Completion of the first round of user tests gave many useful insights about the guiding methods and their combinations with each other for different types of 360° video content. Analyzing these results, suggested redesigning of some of the methods to improve the viewing experience. Hence another round of user test was performed with the revised guidance methods. It was found that these revisions in the guidance methods enhanced the user experience which lead to confirmation of certain design approaches.

Usability evaluations were carried out using a scaled questionnaire, a Pre-test questionnaire and post-test interviews. This generated a large amount data which was analyzed later. The analysis results emphasized the key design considerations.

The novel ideas in the guidance methods were the use of 3D audio (Treble and Bass sounds). Previous research work done did not show any references of use of 3D sounds in the same way which they were used in our prototype. Hence it can be claimed as a novel technique to guide the user to indicate the directions. We used many directional cues such as arrows, visual cues like hollow rectangles, bull's eye mark, etc., and a speech synthesizer female voice to give directions as a guidance technique. Even the combinations of these different types of guiding cues were used to enhance the user experience.

The main contribution of this thesis is the in-depth study of different guiding methods for 360° videos of different content types. This thesis puts forward some design implications which can be used in future work (Section 5.4).

This approach of guiding users in a 360° environment to get an immersive and enjoyable experience should be further studied in more detail to provide concrete guidelines for designing guidance methods for navigation in a 360° video.

References

[Argyriou et al., 2016] Argyriou, Lemonia, Daphne Economou, Vassiliki Bouki, and Ioannis Doumanis, "Engaging immersive video consumers: Challenges regarding 360-degree gamified video applications." In Ubiquitous Computing and Communications and 2016 International Symposium on Cyberspace and Security (IUCC-CSS), International Conference on, pp. 145-152. IEEE.

[“A London City Guided Tour”]. [Online]. Available: https://www.youtube.com/watch?v=m9EClKA1VeQ [Accessed: 2017]

[“Angel falls, Venezuela”]. [Online]. Available: https://www.youtube.com/watch?v=L_tqK4eqelA [Accessed: 2017]

[“Aurora borealis”]. [Online]. Available: https://www.youtube.com/watch?v=ehJg_OlcjpE [Accessed: 2017]

[Benjamin and Huber, 2015] Petry, Benjamin, and Jochen Huber, "Towards effective interaction with omnidirectional videos using immersive virtual reality headsets." Proceedings of the 6th Augmented Human International Conference. ACM, 217-218.

[Benko and Wilson, 2010] Benko, Hrvoje, and Andrew D. Wilson, "Multi-point interactions with immersive omnidirectional visualizations in a dome." ACM International Conference on Interactive Tabletops and Surfaces. ACM, 19-28.

[Brown et al., 2017] Brown, Andy, Jayson Turner, Jake Patterson, Anastasia Schmitz, Mike Armstrong, and Maxine Glancy, "Subtitles in 360-degree Video." Adjunct Publication of the 2017 ACM International Conference on Interactive Experiences for TV and Online Video. ACM,3-8.

[“Carnival of Venice, Italy”]. [Online]. Available: https://www.youtube.com/watch?v=pwivE6bvD8w [Accessed: 2017]

[“Dali”]. [Online]. Available: https://www.youtube.com/watch?v=F1eLeIocAcU [Accessed: 2017]

[Darken and Sibert, 1996] Darken, Rudolph P., and John L. Sibert. "Wayfinding strategies and behaviors in large virtual worlds." In Proceedings of the SIGCHI conference on Human factors in computing systems, ACM, 142-149.

[“Facebook360”]. [Online]. Available: https://facebook360.fb.com [Accessed: Aug-2017]
[“Interaction Design Foundation”]. [Online]. Available: www.interaction-design.org [Accessed: July-2018]

[ISO 9241-11, 1998] International Standards Organization. ISO 9241-11: Ergonomic requirements for office work with visual display terminals (VDTs). Part 11: Guidance on usability. Geneva: International Standards Organization.

[ISO DIS 9241-210, 2010] Standardization, I. O. F. "Ergonomics of Human-system Interaction: Part 210: Human-centered Design for Interactive Systems.".

[Kallioniemi et al., 2013] Kallioniemi, Pekka, Jaakko Hakulinen, Tuuli Keskinen, Markku Turunen, Tomi Heimonen, Laura Pihkala-Posti, Mikael Uusi-Mäkelä, Pentti Hietala, Jussi Okkonen, and Roope Raisamo, "Evaluating landmark attraction model in collaborative wayfinding in virtual learning environments." Proceedings of the 12th International Conference on Mobile and Ubiquitous Multimedia. ACM, 33.

[Kallioniemi et al., 2017a] Kallioniemi, Pekka, Tuuli Keskinen, Jaakko Hakulinen, Markku Turunen, Jussi Karhu, and Kimmo Ronkainen, "Effect of gender on immersion in collaborative iODV applications." Proceedings of the 16th International Conference on Mobile and Ubiquitous Multimedia. ACM, 199-207.

[Kallioniemi et al., 2017b] Kallioniemi, Pekka, Ville Mäkelä, Santeri Saarinen, Markku Turunen, York Winter, and Andrei Istudor, “User Experience and Immersion of Interactive Omnidirectional Videos in CAVE Systems and Head-Mounted Displays.” In: Bernhaupt R., Dalvi G., Joshi A., K. Balkrishan D., O’Neill J., Winckler M. (eds) Human- Computer Interaction – INTERACT 2017. INTERACT 2017. Lecture Notes in Computer Science, vol 10516. Springer, Cham, 299-31

[“Kamchatka Volcano Eruption”]. [Online]. Available: https://www.youtube.com/watch?v=o3a1fkLsNS4 [Accessed: 2017]

[Kauhanen et al. 2017] Kauhanen, Otto, Heli Väätäjä, Markku Turunen, Tuuli Keskinen, Esa Sirkkunen, Turo Uskali, Vesa Lindqvist, Chelsea Kelling, and Jussi Karhu, "Assisting immersive virtual reality development with user experience design approach." Proceedings of the 21st International Academic Mindtrek Conference. ACM, 127-136.

[Lin et al., 2017a] Lin, Yung-Ta, Yi-Chi Liao, Shan-Yuan Teng, Yi-Ju Chung, Liwei Chan, and Bing-Yu Chen, "Outside-In: Visualizing Out-of-Sight Regions-of-Interest in a 360 Video Using Spatial Picture-in-Picture Previews." Proceedings of the 30th Annual ACM Symposium on User Interface Software and Technology. ACM, 255-265.

[Lin et al., 2017b] Lin, Yen-Chen, Yung-Ju Chang, Hou-Ning Hu, Hsien-Tzu Cheng, Chi-Wen Huang, and Min Sun, "Tell me where to look: Investigating ways for assisting focus in 360 video." Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems. ACM, 2535-2545.

[“Moscow Kremlin”]. [Online]. Available: https://www.youtube.com/watch?v=5fyVCocJoks [Accessed: 2017]

[Neng and Chambel, 2012] Neng, Luís AR, and Teresa Chambel, "Get Around 360º Hypervideo Its Design and Evaluation." International Journal of Ambient Computing and Intelligence (IJACI) 4.4, 40-57.

[Nielsen, 1994] Jakob Nielsen, Enhancing the explanatory power of usability heuristics. In: Proceedings of the SIGCHI conference on Human factors in computing systems celebrating interdependence - CHI ’94, ACM, 152–158.

[Petrie and Bevan, 2009] Petrie, Helen, and Nigel Bevan. "The Evaluation of Accessibility, Usability, and User Experience.", 1-16.

[Philpot et al., 2017] Philpot, Adam, Maxine Glancy, Peter J. Passmore, Andrew Wood, and Bob Fields, "User experience of panoramic video in CAVE-like and head mounted display Viewing Conditions." Proceedings of the 2017 ACM International Conference on Interactive Experiences for TV and Online Video. ACM, 65-75.

[“Plitvice Lakes in Winter, Croatia”]. [Online]. Available: https://www.youtube.com/watch?v=A5ubcGt8Gfo [Accessed: 2017]

[Rovelo et al., 2014] Rovelo Ruiz, Gustavo Alberto, Davy Vanacken, Kris Luyten, Francisco Abad, and Emilio Camahort, "Multi-viewer gesture-based interaction for omni-directional video." Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. ACM, 4077-4086.

[Rönkä Ilkka, 2018]. [Online]. Available: http://urn.fi/URN:NBN:fi:uta-201807022203 [Accessed: September-2018]

[Saarinen et al., 2017] Saarinen, Santeri, Ville Mäkelä, Pekka Kallioniemi, Jaakko Hakulinen, and Markku Turunen, "Guidelines for designing interactive omnidirectional video applications." IFIP Conference on Human-Computer Interaction. Springer, Cham, 263-272.

[“Samsung GearVR”]. [Online]. Available: https://www.samsung.com/global/galaxy/gear-vr/images/ [Accessed: September-2018]

[“Space Dream 360°”]. [Online]. Available: https://www.youtube.com/watch?v=cPkj2C-rl24 [Accessed: 2017]

[“The Golden Ring of Russia Air-Balloon Festival”]. [Online]. Available: https://www.youtube.com/watch?v=zVOow7SVJbM [Accessed: 2017]

[“Usability Evaluation Basics”]. [Online]. Available: https://www.usability.gov/what-and-why/usability-evaluation.html [Accessed: June-2018]

[Vermeeren et al., 2010] Vermeeren, Arnold POS, Effie Lai-Chong Law, Virpi Roto, Marianna Obrist, Jettie Hoonhout, and Kaisa Väänänen-Vainio-Mattila, "User experience evaluation methods: current state and development needs." Proceedings of the 6th Nordic Conference on Human-Computer Interaction: Extending Boundaries. ACM, 2010.

[Witmer and Singer, 1998] Witmer, Bob G., and Michael J. Singer. "Measuring presence in virtual environments: A presence questionnaire." Presence7, no. 3, 225-240.

[Zimmerman et al., 2007] Zimmerman, John, Jodi Forlizzi, and Shelley Evenson, "Research through design as a method for interaction design research in HCI." Proceedings of the SIGCHI conference on Human factors in computing systems. ACM, 493-502.

[Zimmerman et al., 2010] Zimmerman, John, Erik Stolterman, and Jodi Forlizzi, "An analysis and critique of Research through Design: towards a formalization of a research approach." Proceedings of the 8th ACM Conference on Designing Interactive Systems. ACM, 310-319.

Appendix 1: User background questionnaire

Date:

User Evaluation Interview: Evaluation of Guiding Methods for 360° Videos

User Name:

Questionnaire before the demo:

A. Age:
B. Gender:
C. How often do you use computers & different applications on computers?

A. Have you watched a 360° video before?

A. Have you used GearVR / Head Mounted Display (HMD) device before?

Appendix 2: User interview Questions

Interview Questions:

1. How are you feeling after watching the demo?

2. How was your experience of watching a 360° video?

3. Was it an immersive experience?

4. Do you feel watching a 360° video with guidance is more useful than watching them without guidance?

5. What did you like about the guidance for 360 videos?
a. Anything specific about the different methods

b. Anything to change

6. What did you dislike about the guidance methods?
a. About specific methods

b. Anything to change for that

7. If you are given an option of guidance in a 360° video which method will you choose?