Timeline Visualization of Omnidirectional Videos

Shakeeb Shirazi

University of Tampere
Faculty of Communication Sciences
Human Technology Interaction
M.Sc. thesis
Supervisor: Jaakko Hakulinen
May 2018

University of Tampere
Faculty of Communication Sciences
Human-Technology Interaction
Shakeeb-ur-Rehman Shirazi: Timeline Visualization of Omni-Directional Videos
M.Sc. thesis, 38 pages, 13 index and appendix pages

May 2018

There has been an increasing need for timeline visualization of ODV (omnidirectional videos) to help users efficiently navigate to a specific moment or scene in videos. When ODV is viewed, especially when it is being edited, navigation in time is required. There exists massive amount of work over the decades on how to navigate in time in regular video. As the technology has enabled different solutions, traditional solutions like forward and rewind buttons can be replaced with freely draggable timelines. However, omnidirectional video adds some new challenges. Most importantly, in most cases, omnidirectional video is displayed with only one part of the video visible on player viewport. While navigating in time, the user often needs to be aware of what is happening all over in 360-degree domain in the video. This can be made possible in different ways: in addition to normal view, a smaller view called thumbnail displaying the entire video that may be shown, or the view may adjust so that the entire video is visible. In both cases multiple projections can be used. In the latter case, the transition between the normal mode and view mode based on time navigation can depend on projection. Both cases could be done smoothly and possibly in different ways.

A 360-degree video player was developed with two related features: thumbnail attached to timeline and a draggable line on the thumbnail for efficient panning of the main viewport. For evaluation of the system, four different test conditions with four different 360-degree videos were used. Two different shaped thumbnails were used in the evaluation, one is spherical in shape and the other rectangular. Out of the four test conditions the line enabling the panning was included in two cases. According to objective results, the rectangular thumbnail with the panning support is the most efficient one and has the shortest task completion time. According to subjective results, the implemented panning support on both types of thumbnails can be innovative, efficient, easy to use and practical.

Key words and terms: Omnidirectional videos, navigation, visualization, interaction, viewport, render

Acknowledgements

I would like to take the opportunity to thank all the people who have contributed and supported me in carrying out this research.

To begin with, I would like to express my sincere gratitude and deepest thanks to my supervisor, Jaakko Hakulinen who has been a great source of unrelenting guidance and moral support throughout the work. His constant support, assistance and collaboration has lifted up my moral and motivated me to complete the research enthusiastically. His meticulous and reflective comments helped me in maintaining inquisitiveness and motivation. I could complete my research in an industrious way owing to my supervisor's remarkable support.

I want to extend my gratitude to all the methodical, congenial and exuberant colleagues who have highly collaborated in the process of proposed system evaluations. Without their support and diligence, I would not have been able to carry out the effective evaluation.

My acknowledgement will never be complete without mentioning my family and friends. I express my heart-felt and profound gratefulness to them for their prayers, encouragement, affection, appreciation and valuable support at every stage of my study. Their affection was a continuous source of inspiration for me and was the driving force to face every challenge with perseverance during the research.

Table of Contents

1. Introduction

During recent years, the trend of online video streaming has been increasing with a great pace. Every social media platform has the option to upload and share video content. There are online streaming websites like YouTube, Netflix etc. based on video, movie and TV series content. These websites are gaining popularity with every passing day.

Video streaming has gained so much popularity globally that many of the videos are viewed by hundreds and millions of viewers. Following the emerging trends and peoples' interest in videos, online streaming websites try to offer innovative video navigation controls to enhance the UX of online streaming and to strive in the market competition. Recently, YouTube has introduced a new function i.e., with a double tap on the main view of video player, the user can now forward or rewind ten seconds. This new function has made video streaming more convenient for touch screen users as this technique has resolved the problem of navigation with seek bar. Traditional video is recorded with a camera with a limited field of view. Recently omnidirectional video (ODV) which covers the entire 360-degree horizontal area has become widely available.

Currently omnidirectional videos are gaining popularity in the field of online streaming websites and VR based video games. Online streaming website like YouTube and social media websites like Facebook have introduced ODV with navigation techniques to turn the view to any direction or angle as per the users' choice. This technique often features a small ring visualizing the current direction to the user. One problem with this technique is that, the user has to pan much to reach a specific direction. To address this problem, YouTube has recently implemented round shaped navigation key on top left side of viewport with four arrow keys for left, right, top and bottom directions. This improvement has resulted in decreased amount of panning interactions. However, this enhancement in video player still does not provide an improved idea about the angle of the viewport. Panning to a specific direction is still a difficult task to achieve.

The 360-degree viewing angle could be cylindrical or spherical in shape when a video is recorded. ODV is not as popular as it could be because of multiple reasons such as its requirement of very high resolutions and thus high bandwidth internet connection. Many people do not have cameras that can record ODVs also people are still unaware of the added value of 360-video point of view.

When ODV was introduced, it was supposed that it holds potential for providing viewers an enhanced feeling of presence and involvement compared to traditional video. In most solutions, ODV playback did not provide viewers with additional interaction except viewing direction manipulation. Advancements in the field of interactivity were not given considerable attention by the developers of ODV players which is why ODV has not received maximum popularity. [Maarten et al., 2016]

Videos are used now not only for entertainment purposes but also for learning. For example, there are plenty of video lectures available online. Surveillance for security purposes is also a popular use of video. Videos in health-care, product presentations and tutorials are gaining popularity. Browsing through videos is still a challenging task to perform. This type of browsing still requires complex interaction and is a time-consuming task.

In scientific literature, several videos browsing techniques have been proposed for regular video. However, there is need for improved navigation support for ODV playback. Before putting such navigation technique for practical application, they must comply with some set standards such as [Schoeffmann and Boeszoermenyi, 2009]:

- The newly designed tool should be user friendly. Approaches proposed till now generally require prior knowledge of the field which is difficult for untrained users to understand.
- Proposed browsing technique should be applicable in all application domains of the videos.
- A new browsing technique should take less time for content analysis and should acquire less space for meta-data storage.

This thesis presents two interconnected elements aiming to enable efficient navigation through omnidirectional videos. A 360-video player with these elements was developed and evaluated. Similarly to existing players, the interface consist of main viewport which display a selected part of the video. For time navigation there is a seekbar which the user can drag to move forward and back in the video. The seekbar was augmented with a thumbnail view. The thumbnail provides a small view to the entire 360-video and it moves together with the draggable seekbar selector element. As drags the seekbar selector, the thumbnail follows and its content updates in real time allowing the user to efficiently find the segment or scene of the video they are interested in. Finally, the thumbnail has a vertical line in red color which indicates the direction of the video which the main viewport is currently displaying. The user can pan the main viewport by dragging the line in horizontal direction. Compared to panning the view by dragging the main viewport, the line on the thumbnail can provide more efficient panning and it also communicates the current viewing direction. In the following, the components described above are called Main view (MV), Seek bar (SB) and Redline (RL).

The developed prototype system called “timeline visualization of omnidirectional videos using thumbnails” (TVODVT) has three different types of real time thumbnails available. The first thumbnail is spherical in shape while the second thumbnail is rectangular. Both of these display the entire 360-degree video frame. The third thumbnail is similar to the main view but in thumbnail size, i.e., it shows only part of the video matching the view in the main viewport. The third type of thumbnail provides less information with respect to complete 360 scene at a time on a timeline. Therefore, it was not included in the evaluation.

Before the introduction and proposal of any new feature or interaction technique a number of research questions arises. In this research work the following research questions will be answered to evaluate the feasibility of the introduced navigation techniques.

- What type of thumbnail view projections of ODVs best support efficient navigation in time?
- How panning through thumbnail and panning through main view of 360 are different in terms of efficiency?
- How thumbnails with and without panning support would affect the subjective preferences of the ODV users?
- How will the users be facilitated with the new features of navigation in ODVs?

To answer these questions, a user study with the developed prototype was conducted. Both objective and subjective data were collected and the results indicate that the introduced solution has potential to be used as a navigation tool and can provide better viewing experience of 360-degree videos.

This thesis has seven chapters. Literature review as chapter 2 discusses omnidirectional video and video navigation in detail. Description of the evaluated system in Chapter 3 describes the video player UI and its component. It also explains the files used in TVODVT prototype. Evaluation and experiment Chapter 4 describes procedure, participant and forms used in the evaluation. Chapter 5 describes the results of the evaluation. Chapter 6 contains discussion of the results in relation with present knowledge in literature. Chapter 7 recaps the work and discusses the future work.

2. Literature review

Omnidirectional video is an emerging trend from the last decade. ODV entails 360 degree view with user controllable navigation interactions. This research presents timeline visualization of omnidirectional videos through thumbnail-based interactions. Aim of this research is to develop a prototype that provides new navigation features at 360 degrees viewing angle and evaluate them. Designing and testing of such features requires prior knowledge of the basic working and functioning of omnidirectional videos and video navigation interfaces. Related work in these two areas is presented in this chapter.

2.1 Omnidirectional Video

Omnidirectional videos enable direct view of surrounding by combining images of a whole scene in correct perspective. [M. Geyer and Daniilidis, 2003]. There are three types of ODV's: Monoscopic, Stereoscopic and Lightfield. The most common type of ODV is Monoscopic. These are flat renderings of a 360-degree shot, which can be viewed on any headset or any type of screen. In Monoscopic type user can look around but has no real depth perception. Stereoscopic ODV is mostly captured with two lenses for every field of vision and is best viewed through VR headsets. It creates three-dimensional rendering of a 360-degree shot. Third one is Lightfield which is a capture of the light field that originates from a scene. [Unite Europe, 2017]

Due to developments in technology for recording content in 360-degrees, Omnidirectional videos are becoming more dominant in area of interactive media. Leading IT companies like YouTube, Facebook and Vimeo have dedicated players for the ODV content. ODVs are recorded with a specifically designed 360-degree camera or a set of cameras, which covers the whole 360 degree scene. These ODVs can be viewed on computers, hand-held devices and head mounted devices like VR headsets. In all viewing experience pan, tilt or roll the viewport are the basic interactions used in ODV navigation. [Kallioniemi et al., 2017]

There are potential domains where omnidirectional imaging and videos could be used. For example, industrial use such as remote operation [Saarinen et al., 2017] and telepresence applications [Onoe et al., 1998, De la Torre et al., 2005]. ODVs has also been used in consumer markets. For example, such work has been conducted in the context of museums [Saarinen et al., 2017] and theatre [Decock et al., 2011]. There are also uses of ODVs in education [Järvinen and Ekola., 2014], healthcare and therapy [Rizzo et al., 2003].

Omnidirectional video scenes are controllable through Pan-Tilt-Zoom (PTZ) principle that supports ODV players to provide a spatially restricted viewport to the users. This function involves two-dimensional control of the video with respect to the placement of view window. The users can also view spatial range of the viewport at different magnifications easily. As x axis is without any navigational restrictions, a viewer can carry out continuous panning at 360 degrees. However, to diminish cognitive load and to minimize possible motion sickness, tilt movement is restricted to 180 degrees. This indicates that viewport can be moved sideways but the tilt movement is confined. (see Fig. 2.1). [Maarten et al., 2016]

The ODV player [Maarten et al., 2016] seen in Figure 2.1 is designed to be entrenched in HTML page. The player does not require plug-ins as it is entirely attuned to the Web in a way that it powers the standardized HTML5 technologies. Input content of the ODV player is omnidirectional video in equirectangular projection.

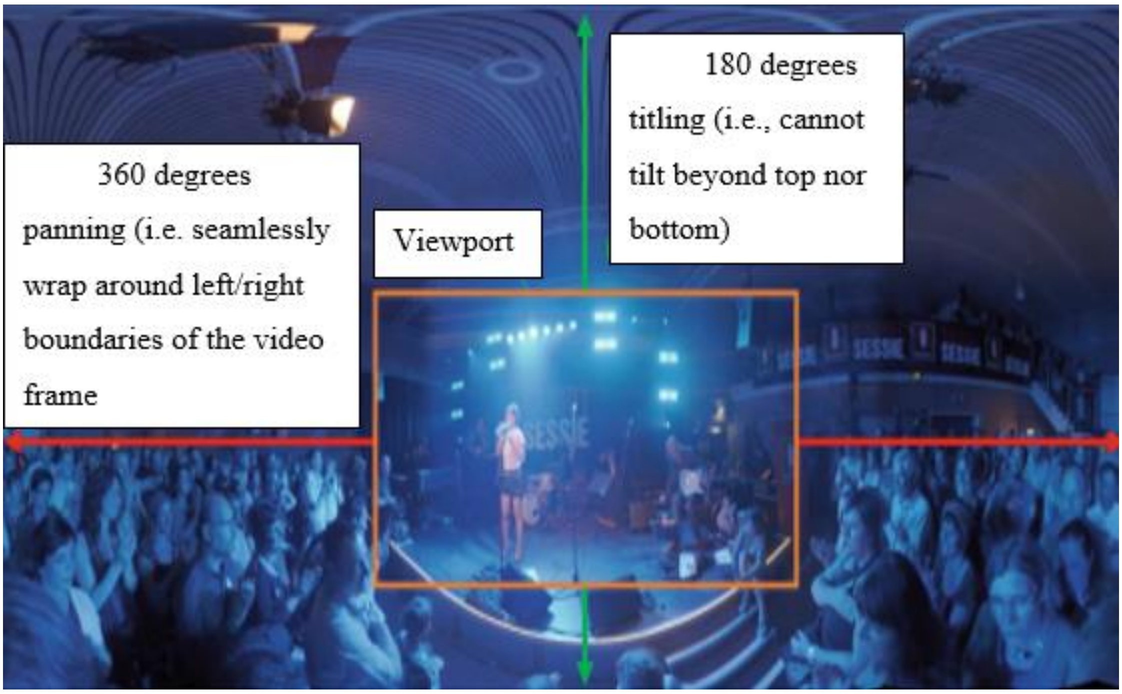


Figure 2.1 shows the pan and tilt restrictions enforced by equirectangular projection of the ODV content in single full frame. Adapter from [Maarten et al., 2016]

Omnidirectional video setups can be related to CAVE like environment but interactive curved and spherical displays seems to be challenging in ODV players. This involves the use of interaction techniques and development of walk-ups which intend to create a transparent environment where the user can interact with suitable devices for tasks. Such user interactions in ODV require designing suitable ODV interaction techniques.

There are plenty of applications of ODVs described in the literature and illustrating and demonstrating their usage possibilities. So far, most of those concepts and prototypes have been evaluated only on a small scale and mostly the evaluations have been laboratory based. After technological advancements there is a need to evaluate ODVs based application with real users and in the real environments. [Hakulinen et al., 2018]

2.2 Video Navigation

In old time when VCRs were used to play videos there were basic rewind, forward and step rewind/forward control functionalities available. Today video playback often takes place online with video streaming gaining popularity and to navigate through video the user has to use seek bar navigation instead of the traditional ways of time navigation controls. As an alternative, YouTube recently introduced 10 seconds forward or rewind feature by double tapping on video screen either on a touch screen or conventional display screens. Still, when a user tries to navigate to some specific frame or scene in a video, it is always a challenge to locate exact desired position.

Variety of research papers and articles can be found on video navigation. However, when usability evaluation testing on newly proposed navigation techniques are performed, very few satisfactory results could be attained and there is a room for improvement. Number of projects have provided enhancements to the efficiency of video navigation with added features of catering desired scene selection for a user.

Schoeffmann and Hudelist [2015] introduced a video navigation interface. The aim of this interface is to allow interactive random access to video content. These tools are meant to provide a sequence of thumbnails. In videos with better thumbnail visualization and provide features to enhance the experience of forwarding and rewinding. These features include better experience with seek bar in video navigation. It is possible to navigate through seek bar in an enlarged window. At the beginning of the enlarged window a small thumbnail is visualized to give user a hint about current position in the video.

A dynamic timeline approach was proposed for seek bar navigation in videos and it was based on an elastic timeline similar to a rubber band slider. It adjusts playback speed through a nonlinear function of the mouse providing smoother interaction. This technique provides a clear presentation of video content at high speed scrubbing. To be specific, it shows a portion of key clips from the video at high speed. The key clips are selected from frames of the video. [Schoeffmann and Hudelist, 2015]

Fluid interaction and visualization technique for navigating, segmenting, connecting and interpreting video using a pen based pressure sensitive interface was presented by Ramos Gonzalo, and Balakrishnan [2003]. In their solution, a timeline of thumbnails floats under the main viewport of the video and if the user press any of the thumbnails with a pen then that specific thumbnail is viewed in bigger size. The selected thumbnail floats with the previous and next thumbnails in the form of a sine wave. Different motions were introduced with the stylus. When the stylus is moved towards right, a frame is created on slight right similarly movement on left creates frame on left. A circular shaped motion with the stylus explains about the frame. There were plenty of different gestures introduced to navigate and control the video. The basic issue in their proposed technique was learning curve caused by the need to learn all the interactions and gestures to become an efficient user of that video player.

In the research paper by Suporn Pongnumkul, Jue Wang, Gonzalo Ramos and Michael Cohen [2010] a technique was proposed for video browsing in a manner which provides a precise movement of slider without loss of frames in a lengthy video. In a traditional video player scrolling through scroll bar is not accurate and efficient particularly in lengthy videos. The scrolling action causes sudden moves in the video so by skipping large part of the frames the viewer cannot reach the specific content to select a specific frame from a video. To address this problem, they proposed a content aware dynamic timeline control. Their technique separated video speed and play back speed and carried out video content analysis with controlled speed to present viewable shots. As compared to traditional video players, this proposed technique increases the navigation experience to select a specific frame both in short and lengthy videos.

With the VR headsets, during playback of 360-degree video, the user is engaged with the video. That is a fundamental aspect to consider when designing the control interactions for playback because the user with a VR headset on his/her head cannot see his/her hands or even be able to see the real environment. Researchers have compared three generally available interaction method used in VR interaction for playback control and navigation of 360-degree videos. Those three interaction methods were gestures performed by hands, directing with head orientation and a remote controller.

Gestures are preferred by the users for the interaction purposes [Bleumers et al., 2012]. Another study shows that the gestures are not reliable when used in user interfaces for interaction and navigation [Atienza et al. 2016]. Therefore, other methods are very common in VR based interactions, such as head orientation pointing. According to a study interactions performed with handheld devices like, a remote controller and tablets are more effective on interaction with smart TV compare to gestures [Bobeth et al., 2014]. The most common feature required in video navigation is pausing playback and skipping the content functions in video players [Darnell, 2007]. Another study shows that for panning the view and zooming the viewport handheld devices with touch screen function are the most needed features by users [Zoric et al., 2013].

These results show that these requirements by users are the same in the case of 360-degree video playback in a VR environment [Pakkanen et al, 2017]. To evaluate three different interaction techniques, they developed a 360-degree video player using JavaScript and utilizing WebGL and Web VR APIs through ThreeJS library In this thesis, a similar approach was used to develop a prototype for evaluation study but this 360-degree video player does not require Web VR API because it is only based on two dimensional display screens and mouse-based interaction, not head-mounted display type VR interactions.

Recently many navigation techniques have been implemented to navigate in videos in online video streaming services such as YouTube, Netflix and other similar websites. A group of Autodesk research from Toronto, Ontario, Canada **[**Matejka et al., 2013**]** present a very valuable approach to this field. They present a new navigation technique called swifter in which there pre-cached thumbnails used during scrubbing action. When a user holds mouse cursor at video seek bar, this is considered a scrubbing action. During the scrubbing the user can select his/her desired scene from a grids of images.

They also proposed three types of different scrolling techniques such as Page-at-a Time, Row-at-a-Time and Continuous Scrolling **[**Matejka et al., 2013**]**. In Page-at-a-Time, the set of visible thumbnails of video changes if play head (ring under cursor at seek bar in the case of YouTube) moves across the whole timeline and they divide those timelines into small timestamps. For example, a grid of thumbnails will not change when timestamp is t and t+1 and whole grid will update when timestamp equals to t+2. In row-at-a-time, scrolling grid is updated based on the horizontal movement of the cursor. For example, in 5x5 grid the view updates once for every 5 pixels moved horizontally. In continuous scrolling thumbnails, the grid is updated on the movement of play head either in horizontal or vertical direction and the entire grid updates slightly on every movement. The above mentioned concept is represented in Figure 2.2.

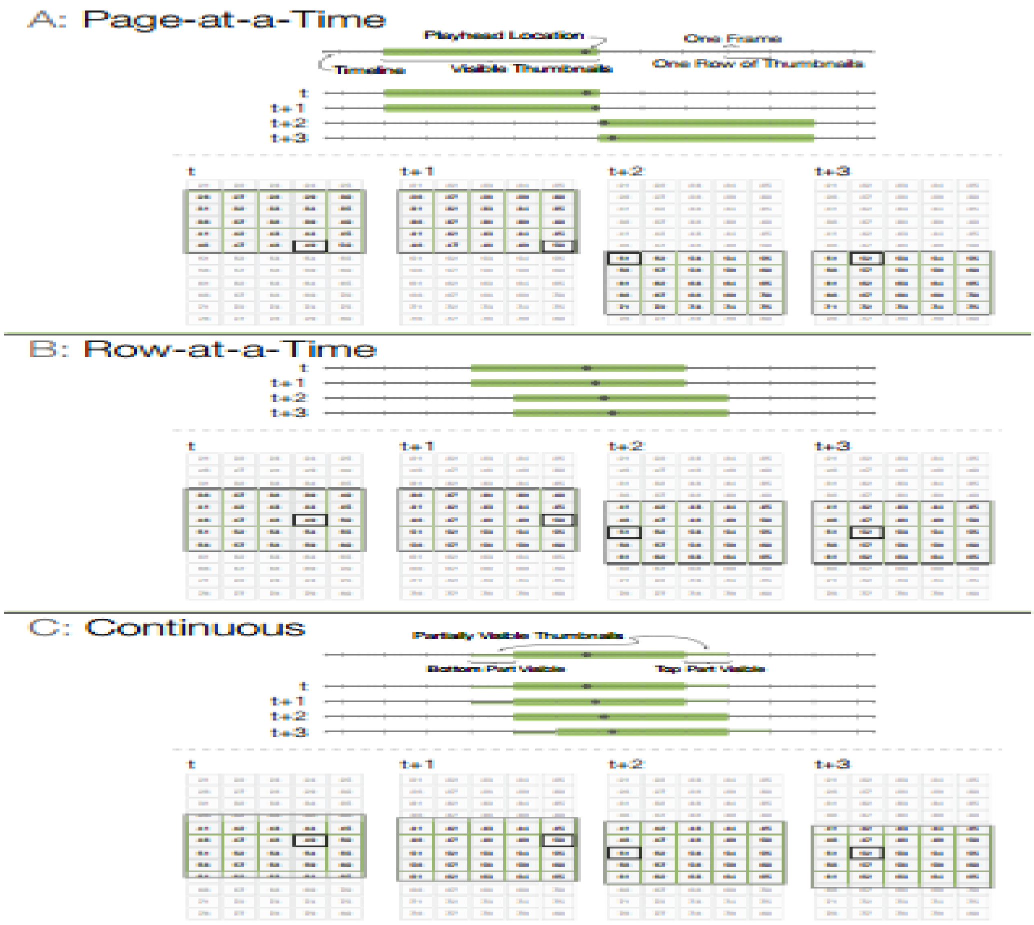


Figure 2.2: Position of visible thumbnails as the play head location updates in each of the scrolling techniques **[**Matejka et al., 2013**]**

The researchers also proposed two selection techniques for thumbnail view; indirect and direct selection (see Figure 2.3). In indirect selection, the technique is moreover similar to conventional ways of video navigation. The user highlights the desired thumbnail by moving play head left or right. However, this technique does not support small target areas and thus precise movement of play head is difficult to achieve.

In indirect selection technique, the user must hold and move the cursor vertically to select the desired thumbnail. On the chosen frame the user should release the button to select the corresponding video frame. For direct selection technique error prevention is handled by cursor movement steeper than 45 degrees so that no confusion on which selection techniques is used will arise. Another important design consideration in swifter technique is varying the size of the thumbnails. If the thumbnails are larger in size then greater amount of details could be seen and if the thumbnails are smaller in size then few details could be seen. On the other side if the size of thumbnails is small then several images could be displayed at once in the whole grid of thumbnails. Finally, to measure the pros and cons of each proposed aspect, they conducted three different studies with different key elements to identify the best configuration. **[**Matejka et al., 2013**]**

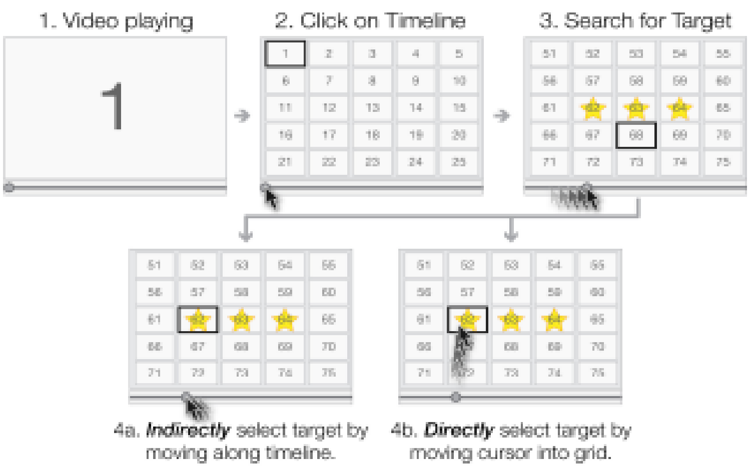


Figure 2.3: Mechanics of selecting a thumbnail using either the direct or indirect selection method. The cell with the thick outline indicates the currently selected thumbnail. **[**Matejka et al., 2013**]**

Justin Matejka, Tovi Grossman and George Fitzmaurice **[**2013**]** in their research paper have described in detail three implemented studies to understand the swifter technique. First, they have described in detail scene ordering, latency, video and timeline size and then video ordering within the context of video navigation, sequential and random order. Efficiency of any video navigation system is significantly dependent on latency. That is why swifter has low latency rate. They made a video player prototype for use in their study which is 854 pixels wide by 504 pixels tall and includes the timeline slider.

In their first study, the primary motive was to understand differences between the three scrolling techniques with respect to performance and subjective preferences. They recruited twelve paid volunteer participants for performing the evaluation study. The study started with the standard video player and by viewing some target scenes for every trial, the participants had to locate the same scene. However, the specific scene was randomized each time to avoid repetition. The participant took part in testing every scrolling technique including direct and indirect selection. After the completion of the study, results were gathered based on mean arithmetic approach and continuous technique had the fastest completion time of 8.9 seconds. The participants were also asked to tell about their liking and according to them, page-at-a-time technique was preferred.

In the second study by Autodesk research from Toronto, Ontario, Canada, the researchers focused to find out the effects of variation in grid dimensions and in detail two types of variations; the number and size of the thumbnails displayed. They hypothesize that those kinds of variations mostly depend on the complexity of some of the scenes in a video. To prove that, they conducted a testing with eight participants including three females aged 18 to 37. The test was performed within time duration of 1 hour. The test included low and high discernibility video material as one independent variable and eight different grid dimensions starting from 5x5 to 12x12. During the test, the size of the grid dimensions were randomized within the total length of a scene with 8 frames. The results were generated with respect to discernibility of frames.

In the case of low discernibility, lowest performance was in 7x7 and 8x8 grid size. On the other hand, in case of high discernibility, lowest performance was in 9x9 and 11x11 grid size. In short, the users have reported that visual searching is made easy in large dimensional size of thumbnails however, it could become worse if some scenes are placed within small size of grids. They concluded from this study that 8x8 grid is effective with respect to fast performance and it received high subjective ratings. They used the 8x8 grid for their third and final study. **[**Matejka et al., 2013**]**

In their final study, they tested swifter technique against existing online video navigation techniques to deal with the effects of latency. They tested four techniques with the same collection of thumbnails with one resolution of 134x72 pixels. Those four tested techniques were small thumbnail (Netflix), Swift, row of thumbnails and swifter. The results include the primary dependent variable which is completion time. Completion time for swifter techniques was 8.79s, for a row of thumbnails 14.03s, for swift 14.04s and for small frames it was 13.04s. **[**Matejka et al., 2013**]**

Axel Carlier, Vincent Charvillat, Wei Tsang Ooi [2015] write about video navigation in their research paper “A Video Timeline with Bookmarks and Prefetch State for Faster Video Browsing”. They present video prefetch technique to reduce seek latency. According to the results of their study, the proposed technique could be used to reduce seek latency by 40 % as compared to UI’s of Video players used today. Prefetch technique is basically a playback buffer in a non-continuous manner instead of conventional continuous buffer and for that entire seek bar is divided into parts called destinations. To define seek destination for prefetch segments of a video, the researchers proposed long video seeking and finding some specific scenes based on users' behavior. They also mention the increasing trend towards video content available online. Today, users often try to get a quick overview of the video content by scrolling the seek bar instead of watching the complete video.

Huang and Hsu proposed a data-mining approach to find out destination by seeking pattern of other users [Huang and Hsu, 2003]. In their work, they included such recommendations for video timeline and bookmarks. After conducting user study with their designed HTML5 based video player the results are remarkably great in the context of latency rate of proposed prefetched navigation technique as compare to conventional video player with continuous seek bar. From their study, they have concluded that a user could go outside the already prefetched segment to during seek. Users' behavior could become predictable if seek latency is low. This study improves on existing prediction techniques to enhance user experience during online video browsing. [Carlier et al., 2015]

Omnidirectional version of videos is an upgrade to traditional video providing the full 360 degrees of content. To utilize ODV different navigation interfaces which provide vertical and horizontal panning have been added to video players. By using these features, a user can, e.g., drag and take a look at a broader view. This navigation may also be supported be different thumbnail views. In addition to controlling the viewing direction, time navigation similar to regular video is often relevant in ODV playback. The current study explores the use of timeline visualization of omnidirectional videos through thumbnails (TVODVT) interactions. The proposed TVODVT prototype is described in the next chapter.

3. TVODVT Video Player System

To study the visualization of omnidirectional video at different time slots through the thumbnail navigation, a prototype named “timeline visualization of omnidirectional videos through thumbnails” (TVODVT) was developed. TVODVT is a web-based player for viewing omnidirectional videos on desktop and laptop computers. As seen in Figure 3.2, the interface is similar to common video players with some additional elements. Main view covering the entire interface area displays a 120 degrees wide area of the 360-degree video. The viewing direction can be panned by dragging the main view, similar to most existing ODV playback solutions. For time navigation, a seekbar is included on the bottom half of the interface. In addition to these regular elements, a thumbnail is included to help a user to have an overview of the complete 360 degrees view. This thumbnail is connected to the seekbar. In the thumbnail there is also a vertical line (Redline) which indicate the horizontal direction of the main view. The main view can also be panned by dragging Redline on the thumbnail. The TVODVT system was designed and developed to evaluate the thumbnails and the related Redline functionality. The implementation is a combination of markup language HTML5, CSS3 for styling and for behavior JavaScript code.

TVODVT system’s interface has three ways to navigate in ODV. The user can pan in the main view by holding it with mouse cursor and moving in either right, left, up or downward direction to change the view to that corresponding direction. Redline is introduced to horizontally change the main view and the user moves it in either left or right direction on a thumbnail and main view changes accordingly. Seek bar helps the user to navigate through a video by holding a round button and moving towards right which will forward a video or towards left, which will rewind the video. While dragging, the thumbnail follows the button and thumbnail content updates in real time to help the user find the desired content in the video. The following figure 3.1a and b shows the types of thumbnail used in TVODVT system.

Figure3.1a: Spherical thumbnail

Figure3.1b: Rectangular thumbnail

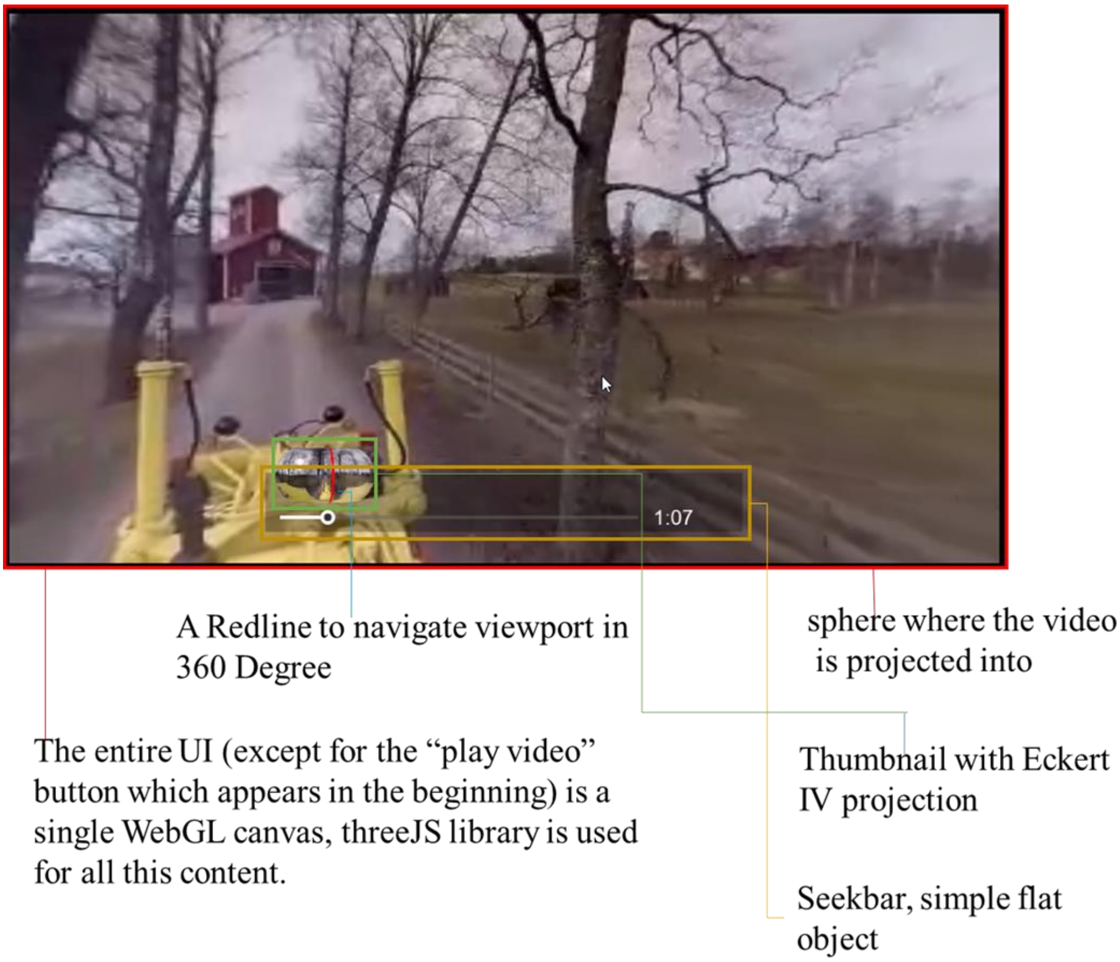


Figure 3.2: TVODVT System UI.

3.1 Public Interface

The following text explains the technical implementation of the TVODVT system. It elaborates the key elements in the system along with technical explanation. There are the following hidden elements in the Document Object Model of the application used to generate graphics displayed via the main 3D graphics based view.

- <video> element, playing the video file
- <canvas> where the seekbar is drawn
- <canvas> where the Redline is drawn

Viewports:

There are three different viewports, i.e., separate pictures drawn into the main canvas. Each has its own camera and scene (a collection of 3D models). Equi.js is the 3D sphere where the video is projected into to draw the main view and that the camera is inside the sphere.

1. The main view with equi.js and a perspective camera, this camera is controlled with mouse gestures. The control is implemented by RigControls.js.
2. Seek bar view with an orthographic camera
3. Video with thumbnail and Redline, the viewport moves as video time progresses the keep it next to the current video time in seekbar. There are two objects on top of each other. The topmost is transparent except for the red line. The lower one has the actual video. The 3D object is Eckert.js or plane2per1.json (plain rectangle), depending on the currently used thumbnail shape.

There is also a fourth viewport which replaces the 3rd when thumbnail matching the main view is used. Yet another 3D model is used. When one thumbnail type is used, other thumbnails are automatically deactivated by hiding the relevant 3D objects.

Used 3D models:

There are three JSON files used to create dimensions of a place, two for the thumbnails (one for spherical, another for rectangular) and one for main view video projection. One JavaScript file SeekbarModel.js is for simple flat object which display the seek bar.

3.2 Files

ODV implementation with thumbnails consist of the main player.html source file and several dependences on JavaScript libraries. Short descriptions of the libraries used in ODV thumbnails implementation are discussed below.

3.2.1 Player.html

Player.html contains most of the code custom to this software. It contains the web page, mouse and keyboard handlers except for the main view rotation, and handles the thumbnail updates.

3.2.2 Tween.js

JavaScript library Tween.js is a tweeing JavaScript library use for easy, simple and smooth animation in web-based applications.

3.2.3 Three.js

Three.js is an open source JavaScript library that allows developers to create and render 3D scenes directly on browser. It has a large set of functions to support advanced visualization on modern browsers. Three.js use WebGL for rendering to browsers, almost all modern browsers support it and when there is no support, three.js will fall back to HTML5 canvas or SVG approach. WebGL is a cross-platform web standard for a low-end 3D graphics based on OpenGL and ECMAScript via the HTML5 canvas element [WebGL].

To display content in html canvas with three.js library, a scene, a camera and a renderer must be setup. For example, in ODV system Three.Texture(); method is used to create a texture to display thumbnails on the scrollbar.

```
var seekTexture = new THREE.Texture();
```

Here seekTexture is the texture where a thumbnail picture is copied into. Renderer used in ODV system to render a scene with a camera. WebGLRenderer is an object in three.js. library and WebGL is JavaScript API for rendering interactive 2D and 3D graphics inside an HTML canvas element.

```
this.renderer = new THREE.WebGLRenderer ();
```

Renderer is used to display the scene to users. To animate the scene, we used renderer to draw the scene several times per second.

Three.Scene() method is used to create a scene on the canvas. Three.PerspectiveCamera() method is used to create a camera. That method takes at maximum four attributes including field of view, aspect ratio and near and far clipping plane. Near and far means that objects further away from the camera than value of far or closer than near will not rendered.

ODV system contains one JavaScript file derived from a file in Three.js examples. Three.OrbitControls class allows a user to manipulate camera with mouse, keyboard or touchscreen. RigControls.js file in ODV source files is a modified form of that class and enables control of a multi camera rig instead of just a single camera. In this JavaScript file, there are some object methods used to define limit values such as how far a user can move in or out, how far an orbit vertically moves by applying upper and lower limits, and how far an orbit horizontally moves by applying minimum and maximum azimuth angle to infinity. In this JavaScript file, a generic 3D camera controller is setup. It rotates the camera which draws the main view and contains the relevant mouse listeners and implements to rotation interaction.

3.2.4 Queue.js

Queue.js specifies which 3D model and texture files to load and how the cameras, scenes and viewports are setup. For the player, this JavaScript file defines how scenes, cameras and 3D objects, including seek bar, seek Thumbnail including thumbnail in 2D and for thumbnail sphere, are setup. The implementation of thumbnails has one thumbnail for every 25 frames. This JavaScript file loads all pre-defined geometries in the form of JSON (JavaScript Object Notation) which is a modern alternative to XML data sets used in web development. In JSON files vertices are defined through JavaScript arrays and then combined to form 3D objects. The 3D model JSON files can be generated with 3D modeling tools and export using plugins included in ThreeJS distribution.

Queue.js generates queue entries which load all the thumbnail pictures into textures so that they are ready to use. This solution is not optimal due to two main reasons. First, it requires a lot of network traffic and second all thumbnails are stored in graphics card memory, which could be a performance issue on low-end system. In addition, the server must have a lot of thumbnails pre-generated or be able to generate them efficiently. Because of this, we cannot have a thumbnail for each frame. We decided to generate one thumbnail for every 25 frames which is approximately equal to one per second.

3.2.5 Seekbar.js

Seekbar.js is a JavaScript file to initialize and update seekbar graphics on player.html. It renders seekbar graphics into 2D canvas which is then copied into a ThreeJS texture. It also includes code to handle related mouse events. The seekbar is used for scrubbing the video. In this file, there is first definitions of constants which are used in update function. They are related to seekbar width, height and its range from left to right on the viewport.

3.2.6 SetUp3D.js

SetUp3D.js is a JavaScript file with loading methods. In setUp3D JavaScript is the code which loads the objects and registers the listeners specified in queue.js.

The following attributes can be found in SetUp3D object.

- DOM element where to render the graphics.
- A data structure, which will specify what to load and setup.
- A scene objects where each camera rig connects to a scene.
- A list which holds a set of cameras and related viewport information.
- Three separate attributes related to a skybox which is drawn before the main scenes.
- A Three.js renderer.
- A video reload threshold which is used to restart video playback once the video has reached the end.
- A Boolean attribute specifying whether to reload video at the end.
- Two attributes to statically define Width and Height of shadow map.
- A loading manager defined to reuse easily later in source code by assigning it only once to Three,js LoadingManager method.
  (this.loadingManager = new THREE.LoadingManager();)
- A texture loader method defined the same as above but with the previously defined loading manager given as a parameter.
  (this.textureLoader = new THREE.TextureLoader(this.loadingManager);)
- A scene loader method defined the same as above
  (this.sceneLoader = new THREE.ObjectLoader(this.loadingManager);)

- A JSON Loader method defined the same as above (this.jsonLoader = new THREE.JSONLoader(this.loadingManager);)

The setup method utilizes all previously defined data types and object methods. Before load call videos should be setup by calling setup video function which initializes video html element, creates a texture and returns the related values. Load method can then be used to load the defined queue and start the rendering. That function call receives the following values.

- Queue specifies what to load and how-to setup things.
- onProgress function is called to indicate progress on loading.
- onLoadComplete function is called once after the scene has been set up and is ready for display.
- viewMoveEndCallback is called every time controls defined in RigControls.js finish an interaction. For example, when mouse button is released, a call is made.

3.3 Logging

The player software has a logging system which registers the player setup and loading and all user interactions. The log files are stored in the web server serving the player and the videos. This logging allows quantitative analysis of user behavior.

TVODVT system has three main elements visible to the user; main view, seek bar and Redline. Main view is the sphere where video content is displayed. Seek bar is a simple floating line on main view which controls timely navigation of the video. Thumbnail floats on the seek bar in accordance with the video timing. Through navigation interactions, a user can view the display content across the 360 degrees video content. Using the timeline and thumbnail which updates in real time when the timeline is operated, the user can quickly seek for the wanted content in videos. The developed prototype was evaluated through user testing based evaluation procedure. Methods used for evaluation will be discussed in the next chapter.

4. Evaluation

To evaluate the developed navigation solutions, a user evaluation was conducted. The aim was to collect subjective and objective metrics to understand the usability of the proposed navigation techniques for ODV. The main purpose of this evaluation was to answer the following research questions.

- What type of thumbnail view projections of ODVs best support efficient navigation in time?
- How panning through thumbnail and panning through main view of 360 are different in terms of efficiency?
- How thumbnails with and without panning support would affect the subjective preferences of the ODV users?
- How will the users be facilitated with the new features of navigation in ODVs?

4.1. Participants

For the evaluation of TVODVT, we recruited 12 regular users of technology. All twelve participants were voluntary participated and most of them were my office colleagues. All the participants were familiar with the online video streaming players. The following table illustrates the demographics of the participants.

Table 4.1 Participant demographics

Participant	Gender	Age group	Familiarity with ODV	Familiarity with 360* video players
Pilot test 1	Male	>40	Yes	Yes
Pilot test 2	Male	18-28	Yes	No
P1	Male	18-28	No	No
P2	Male	18-28	Yes	Yes
P3	Female	18-28	No	No
P4	Male	>40	Yes	Yes
P5	Female	18-28	No	No
P6	Male	29-40	Yes	No
P7	Male	18-28	Yes	Yes
P8	Male	29-40	Yes	No
P9	Male	29-40	Yes	No
P10	Male	18-28	Yes	No
P11	Male	18-28	Yes	No
P12	Male	18-28	Yes	No

4.2. Method

We selected four different conditions for the evaluation. Two types of thumbnails were selected and both were used with and without the Redline functionality. The conditions were:

1. Spherical thumbnail, with Redline
2. Rectangular thumbnail, with Redline
3. Spherical thumbnail, without Redline
4. Rectangular thumbnail, without Redline

The evaluation was within-user, i.e., each participant tested all four configurations. We used the Latin squares algorithm for shuffling the order of the conditions. The ordering for participant 1 followed the conditions sequentially from condition 1 to condition 4. Table 4.2 shows the ordering for the first four participants. Numbers in the Table 4.2 are the condition numbers.

Participant five followed the same order as participant 1 and this technique is repeated for every group of four participants. Therefore, it is good to have N*4 number of participants to have equal number of participants with each order of conditions. We recruited 12 participants.

All participants performed the same tasks and to avoid learning effect of users on the evaluation we used four different videos, one per condition. The order of the four videos for each participant was same but the conditions were shuffled according to Latin squares algorithm. The following arrangement describes the implementation procedure.

Table 4.2 Evaluation strategy

Videos	Participant 1	Participant 2	Participant 3	Participant 4
Video 1	1	2	3	4
Video 2	2	3	4	1
Video 3	3	4	1	2
Video 4	4	1	2	3

All the participants followed the same evaluation procedure which is explained below.

1. Filling out user's background questionnaire (Appendix A).
2. The participants were introduced to the TVODVT system and the purpose of the evaluation.
3. The participants completed a set of tasks in each condition

- On completion of each condition, the participant had to fill out a form, which is based on Likert scale to evaluate things like ease of use and comfort of use (Appendix B). These forms were used for subjective evaluation. While the user filled out the forms, the moderator did setup the next condition with the TVODVT system.

4. After all the conditions were tested, the participant had to fill out a post-experiment questionnaire (Appendix C). In this, the participants compared the four different conditions. The questions in the questionnaire were:
   - With which of the techniques for navigation it was easiest to find asked information?
   - Which one of the techniques for navigation was the best overall?
   - Which type of thumbnail gave the best 360-degrees view on the scrollbar?
5. At the end, a short face-to-face interview was carried out with the participant. During the interview, the moderator was trying to learn about the participants personal thoughts about the TVODVT system.

Task were given to every participant in printed form. The tasks were written separately for each video used in the evaluation. The main purpose of these tasks was to enforce participants to interact with TVODVT prototype. Tasks were designed to consider the complete video playtime so that participants have to interact with whole video to find out the answers of the asked questions. Another reason to design such type of tasks was that participant have to view different parts of the video and possibly navigate in time as well. All four videos have almost same playtime i.e. approximately one minute. Same type of videos were selected for evaluation so that participant have to manipulate with each available navigation technique such as Redline, seek bar and main view. The following tasks were used in the four videos for evaluation.

Video 1 - Ship Outdoor

- How many male passengers are standing at the end of a balcony?
- On which floor are the male passengers standing?
- How many small boats are floating adjacent with the main ship?
- What is written in printed green box that is in-front of a cameraman?
- On which side of the ship is the sun visible on the horizon?

The following Figure 4.1 shows screen capture the ship outdoor video used in the evaluation study being viewed in TVODVT player.

Figure 4.1: Video 1, Ship outdoor with spherical thumbnail

Video 2 – Ship Indoor Corridor

- How many persons came from downstairs and move to upper floor?
- Is there a lady who came out from an elevator who is wearing a hijab?
- How many girls walked on the 11th floor who were wearing summer dresses (skirts) or with blonde hairs?
- How many old age women pass through 11th floor corridor?

Video 3 - Ship Control room (Bridge Center Rear)

- How many yellow colored walkie-talkies those are placed on the table just behind a landline telephone?
- How many people are working on the computer?
- Is there any red paper bucket placed on the floor of the control room beneath the table?
- How many screens are displaying security cameras view behind a man who is working on his computer?
- Is there a table lamp lit on the control room table?

The following Figure 4.2 shows a screen capture the ship control room video used in evaluation study in TVODVT player.

Figure 4.2: Video 3, Ship control room with rectangular thumbnail

Video 4 – Ship control room (Security monitor's room)

- How many bananas are placed on a computer table?
- How many people are standing in the control room?
- How many paper printers are placed on the table alongside walkie-talkies?
- How many big screens are displaying security cameras view?
- How many big screens are displaying ship blue prints?
- How many people are working on their laptop?

4.3 Parameters Investigated during Testing

The TVODVT system was tested with twelve participants to discover the most important design and interaction features of the system's UI. During the evaluation system logs were built to further evaluate our system.

4.4 Performance measured through objective data

We evaluated with each participant the four conditions with four different videos. User interactions during the evaluations were recorded in system logs. Based on those logs, total task completion times were extracted. The following quantitative analysis was derived from the logs.

- Completion time of the tasks on the four different conditions per each participant.
- Comparison of completion times between the participants to find out each condition median (overall or average) time.
- Overall completion time over all conditions.
- Total interaction time using MV (main view), RL (Redline) and SB (seek bar).

. The following information was collected into and extracted from the log files.

- Tasks Completion Time

To find out the usability and efficiency of the four available condition the time is the most important factor. By collecting the system logs of the evaluation session, we were able to calculate time differences. It is calculated as the start of every condition till the end of it.

- Panning through main view pane vs dragging through thumbnail. The number of pan operations was calculated to see which method of panning the participants used and how often.

There were two options available for navigation in our TVODVT system. Out of all four test conditions, in conditions 3 and 4 the users could control the main view only by dragging the main view while in the other two conditions they could also use Redline in a thumbnail to pan the view. We measured task completion time in these cases. Out of the total time, how much time was spent panning and how much spent on dragging through thumbnails to perform a task was also calculated.

5. Results

This section covers the results of the evaluations, which were conducted to evaluate the effectiveness of TVODVT system. Data collected with the 12 participant is reported. System evaluation was divided into two main parts. First, one is the objective evaluation where a setup was built for the recording of participant interaction in the form of system log files. Those log files were then analyzed with the help of a Python script to extract interaction times from each case. Those results are discussed in detail under the objective evaluation section of this chapter. The second section reports the subjective evaluation results which originate from participants' answers consisting of the participants' task answers, condition evaluation questionnaires and post experiment questionnaire forms. The third section summarizes interview results and other miscellaneous findings.

5.1 Objective Metrics

Objective evaluation is based on data that was fetched from log files. Those log files were collected during the evaluation session and record every interaction on TVODVT system. We extracted each interaction with their type, the place of interaction, completion time and the number of their occurrences.

Figure 5.1 shows the tasks' completion time of all tasks in minutes for the four conditions. Condition 1 (C1, Spherical thumbnail with Redline) took 17.3 minutes, condition 2 (C2, Rectangular thumbnail with Redline) took 11.227 minutes, condition 3 (C3, Spherical thumbnail without Redline) took 22.167 minutes and condition 4 (C4, Rectangular thumbnail without Redline) took 13.92 minutes in average for all 12 participants to complete. C2 required least time to complete while C3 was the most time consuming. C4 is the second fastest and C1 is the third fastest condition with respect to tasks completion time.

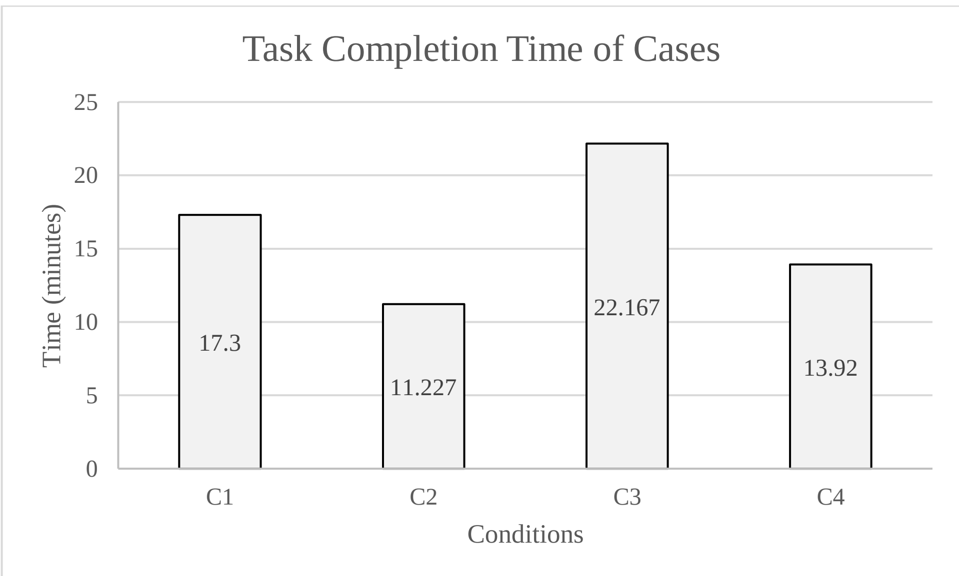


Figure 5.1: Task Completion Time of Conditions

Figure 5.2 shows the task completion time including the details of mouse interaction time on MV (main view), RL (Redline) and SB (seek bar). In the following figure condition 1's interaction time is visualized with dotted, condition 2 with vertical line, condition 3 with horizontal line and condition 4 is solid filled pattern. In condition 1 participants dragged MV more to complete the given tasks as compared to RL. 556 seconds were spent on panning the view by dragging the MV and the least interaction was carried out with RL because their completion time was 138 seconds. In condition 2 participants were dragged MV the most as compared to RL. 528 seconds were spent on MV then 114 seconds spent on RL. Condition 3 and condition 4 did not involve RL interaction. In condition 3 almost equal time was spent on MV and SB. Maximum time was spent on MV in condition 4 however. Overall, MV is the most used interaction in all four conditions. After MV, RL is the most used interaction technique in condition 2.

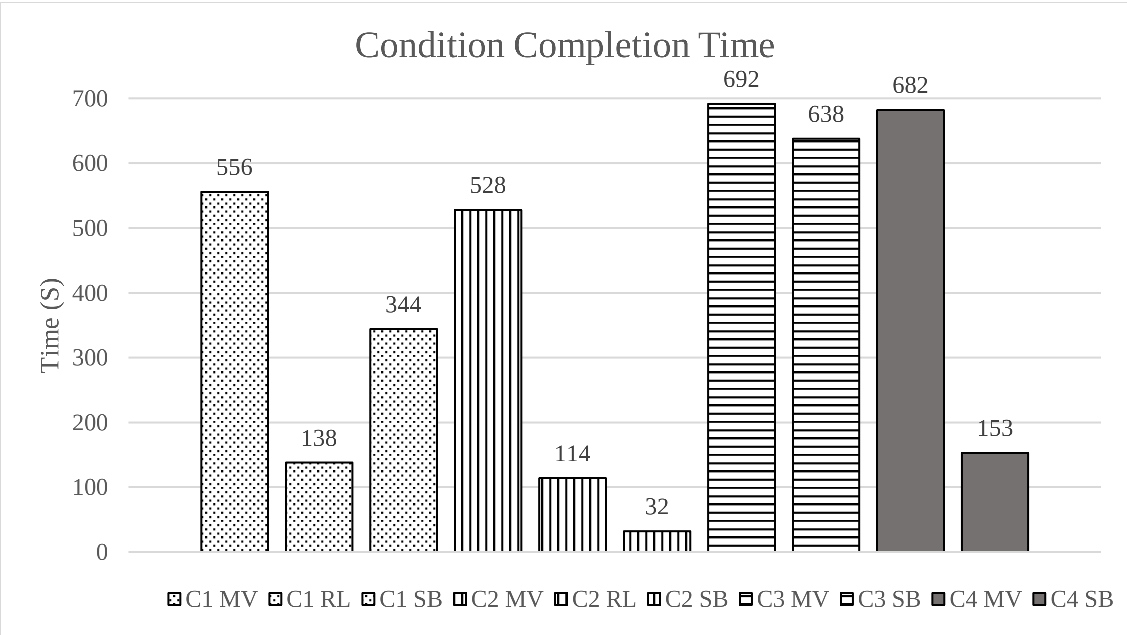
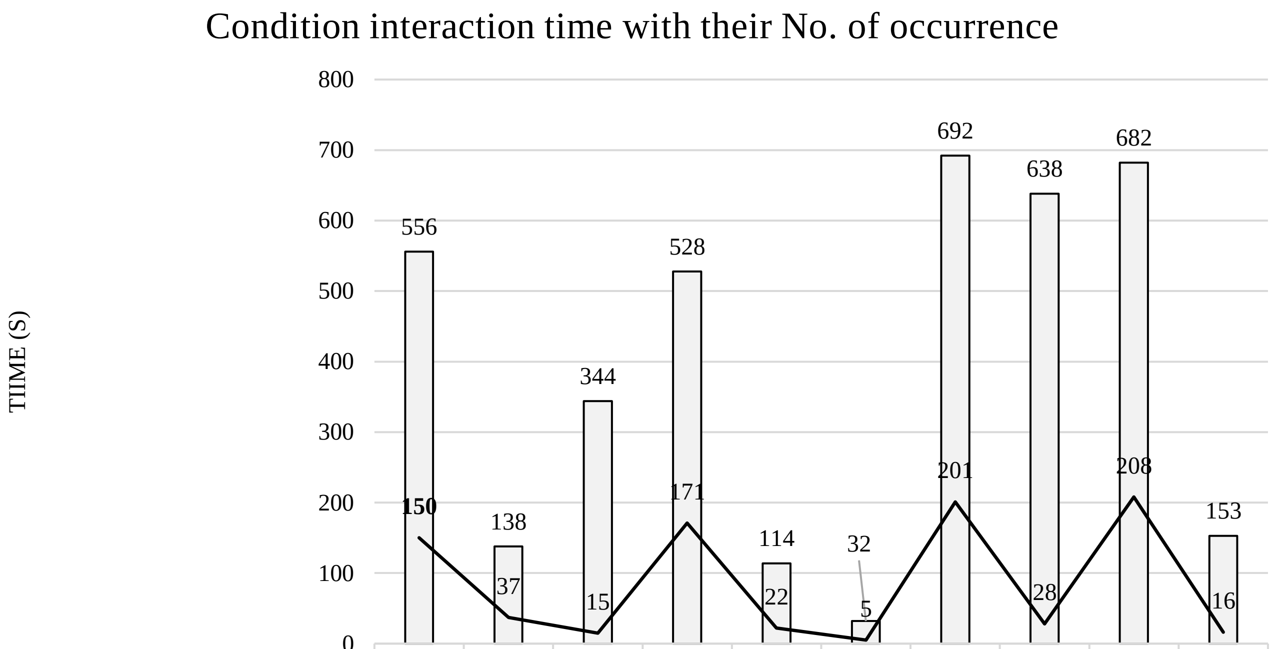


Figure 5.2: Condition completion time with interaction type (MV, RL, SB)

Figure 5.3 shows the completion time per case with their number of occurrences of different types of interactions. In the following figure, bars in solid pattern show completion time in seconds and line graph shows the interaction occurrence count per type and condition. In MV, 208 is the maximum number of events recorded in condition 4, 201 in C3 because C3 and C4 did not include RL interaction. The cases in which RL is active, i.e., C1 and C2, 172 MV events were logged for C2 and 150 for C1. In RL the largest number of events recorded is 37 in C1 and in C2 and 22 RL events logged. In case of SB 28 events were logged in C3 and 16 in C4, 15 in C1 and just 5 in C2.

In MV 208 events were logged with interaction time of 682 seconds. In RL maximum 37 events were logged with 138 seconds completion time. In SB maximum 28 events were logged with 682 seconds completion time.



CASES --->	C1 MV	C1 RL	C1 SB	C2 MV	C2 RL	C2 SB	C3 MV	C3 SB	C4 MV	C4 SB
Completion time (Secs)	556	138	344	528	114	32	692	638	682	153
frequency (No. of Occurance)	150	37	15	171	22	5	201	28	208	16

Figure 5.3 Condition interaction time with their No. of occurrence

5.2 Subjective Evaluation

In the subjective evaluation for ease of use, inventiveness, efficiency, practicality and pleasance of the technique (Appendix B), C1 was rated the most efficient, practical, and inventive out of all four evaluated conditions. C2 was rated slightly easier to use and more pleasing than C1. C3 and C4 were rated almost the same in as can be seen in a boxplot of ratings for each condition. Figure 5.4, 5.5, 5.6, 5.7 and 5.8 shows the boxplots of the ratings for each condition for the questions "*Was the technique inventive?*", "*Was the technique easy to use*", "*Was the technique efficient*", "*Was the technique practical*" and "*Was the technique pleasing*" in the condition evaluation questionnaire. These questions were selected to subjectively evaluate TVODVT prototype.

Below, there is a figure of the subjective evaluation of ease of use. The easiest technique to use is the one in C1 then C3 and C4 on third and C3 was the least easy technique to use.

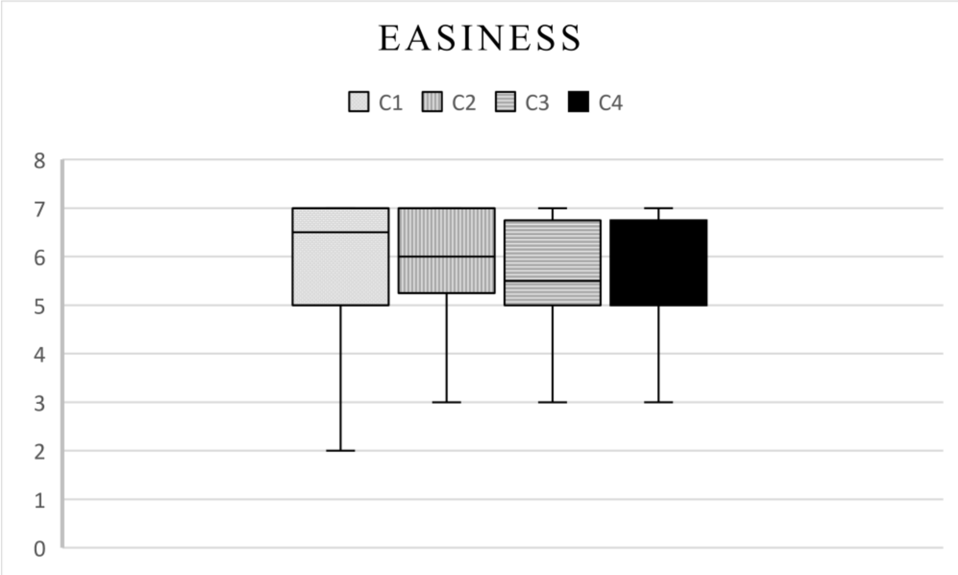


Figure 5.4: Subjective evaluation of ease of use

The following figure shows subjective evaluation of inventiveness. The most innovative technique is C1 then C2 on second, C4 on third and C3 is the least innovative technique used in TVODVT system.

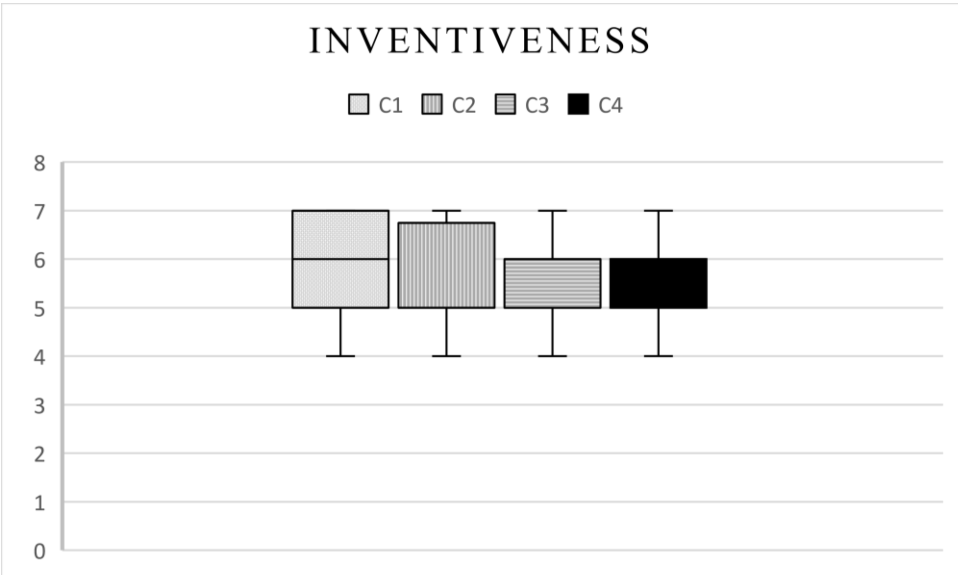


Figure 5.5: Subjective evaluation of inventiveness

Figure 5.6 displays subjective evaluation of efficiency. The technique used in condition 1 turned out to be most efficient while that used in C3 was recorded to be least effective in TVODVT system. However techniques used in C2 and C4 were rated the same on the basis of the recorded mean square values.

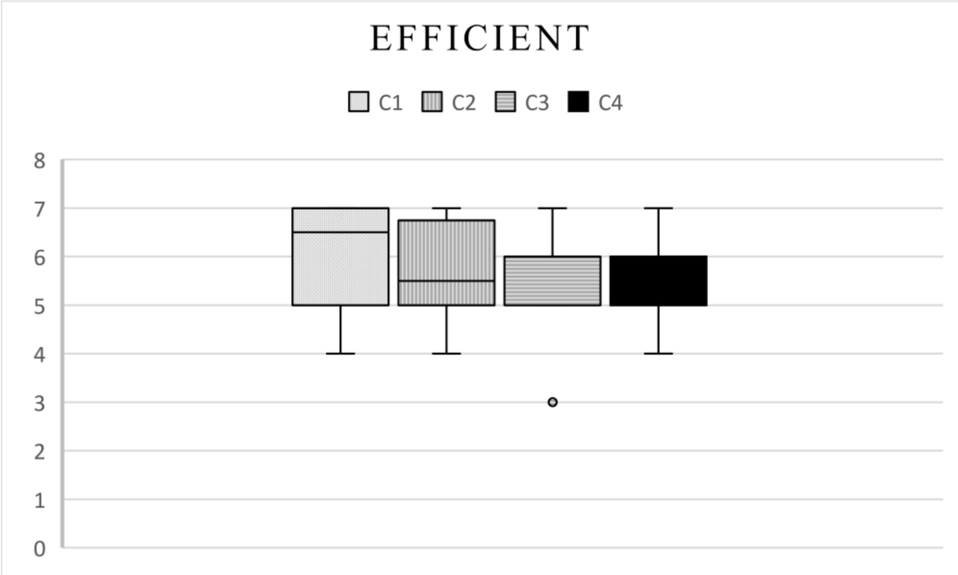


Figure 5.6: Subjective evaluation of efficiency

Subjective evaluation of practicality is shown in the figure 5.7. The technique used in C1 was recorded to be the most practical. C2 turned out to be slightly less practical however, similar to evaluation of efficiency, practicality of techniques for C3 and C4 were rated the same.

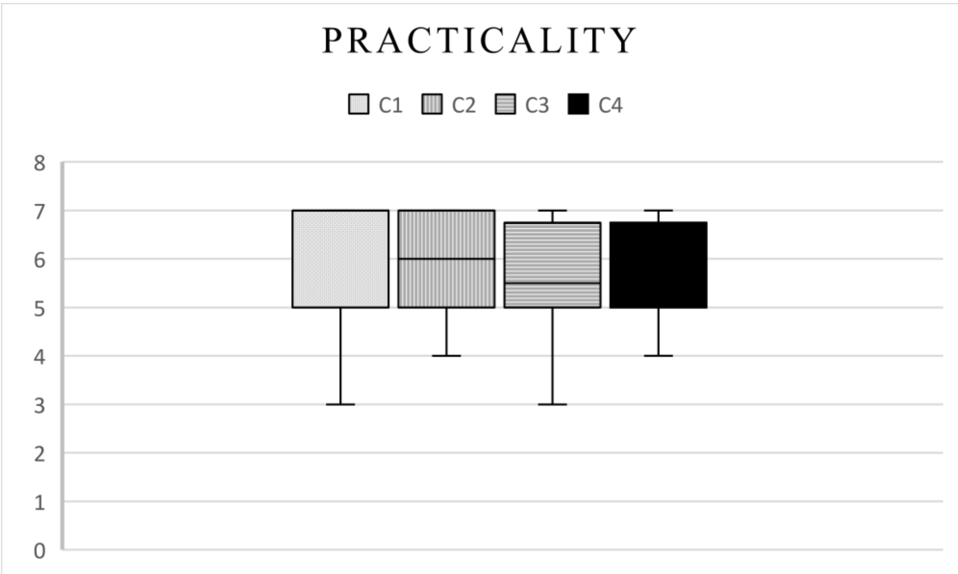


Figure 5.7: Subjective evaluation of practicality

Figure 5.8 illustrates the variations in terms of pleasantness. The technique used in C2 was ranked the most attractive one. C1 involved slightly less pleasing technique as compared to that in C2. The least pleasing techniques were observed in C3 and C4.

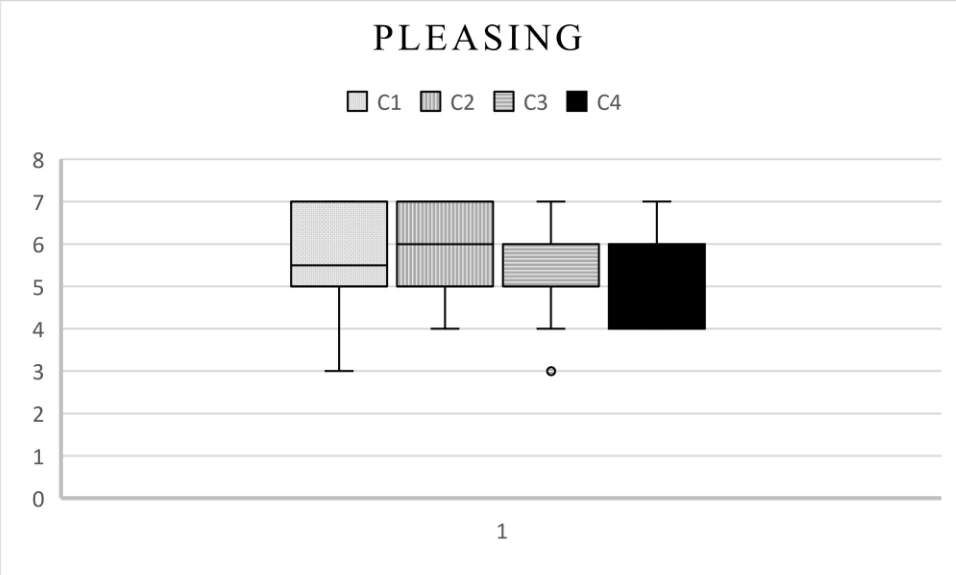


Figure 5.8: Subjective evaluation of pleasantness

5.3 Other Results

In the post experiment questionnaire (appendix C), 10 out of the 12 participants preferred the Redline navigation technique. As the answers to the question "*Which type of thumbnail gave the best 360 degree view on the seek bar*?" 7 participants think that spherical shape thumbnail gives the best view, 4 participants think that rectangular shape thumbnail gives the best view and one participant thinks that both types of thumbnail give equally good 360 degree viewing experience on the seek bar.

Out of the 12 participants, half of them consider that the spherical thumbnail with the use of Redline for navigation is the best navigation technique introduced in the TVODVT system. Four out of the 12 participants consider that the rectangular thumbnail with Redline is the best navigation technique and in one participant's opinion both types of thumbnails with the use of Redline is the best overall technique. Only one participant considered spherical thumbnail without the use of Redline as the best navigation technique.

Overall, 7 out of 12 participants considered the best ODV navigation technique of those in the TVODVT system to be the spherical thumbnail with the use of Redline. 8 out of the 12 participants think that no introduced navigation technique was a reason of confusion or difficulty in use. 4 out of the 12 participants felt that the introduced navigation technique was confusing. Two participants gave comments that Redline movement should be reversed horizontally as they are used to such mapping. One of the participants suggested that navigation techniques themselves are not confusing, however the mouse direction was counter intuitive and it takes time to getting used to. Another participant commented on the speed of movement of the main view, which was quite fast as he was not able to view the changes of the main view during interaction on a thumbnail.

In the answer to question, "Would you like to use this type of 360 degree video player in your smart electronic devices?" all 12 participants said yes because they found

TVODVT efficient, innovative, to provide a 360° view, ease in navigation, ease and faster scrolling.

Overall the participants gave very useful and meaningful comments on the TVODVT system. Two participants suggested that Redline should also support vertical interaction alongside horizontal interaction. Three participants gave positive feedback about the whole TVODVT system and the idea to view objects in ODV. Three participants liked Redline interaction on thumbnails as they considered it as the best navigation technique in 360° videos.

6. Discussion

ODV requires interactive elements to be integrated on ODV player. To present the content of ODV, different mechanisms are required which might have some advantages and disadvantages in regards to the UX of the users. Almost all the proposed mechanisms are built on established web standards like HTLM5, CSS, JavaScript and WebGL. This factor reduces the learning curve that is mostly related to the adoption of a new technology to content producers, engineers and end users. [Maarten et al., 2016]

Our TVODVT prototype system is also based on well-known web standards. Due to that, it requires less time to introduce a new interaction technique on this ODV player. We introduced thumbnails displaying the entire 360-content of the video attached to seekbar and an element called Redline on the thumbnails enabling efficient panning of the main view to the 360 video. Two different shapes of thumbnails were selected for the user evaluation of the proposed system. With the combination of the two different thumbnails and Redline interaction on it we came up with four different test conditions. To avoid user learnability factor we used four different ODVs for all four cases in every evaluation test per participant. By doing that, every participant had to interact with four different videos for completing tasks on four conditions.

The participants performed tasks most precisely and in the shortest time in rectangular shape thumbnail with Redline (condition 2). Condition 2 took only 11.227 minutes to complete the given tasks as per tasks completion time. Subjectively the participants considered the interface featured in condition 1 (Spherical shape thumbnail with Redline) as the most efficient, practical and easy to use approach for interaction through ODV.

Overall, there was a positive feedback from the 12 participants on the TVODVT system. Some of them suggested little improvements to the system. Two of them suggested that interaction through Redline should be reversed so that if they move Redline to the right then MV moves to the other direction. This is the logic they are used to it in multimedia games. One participant suggested vertical interaction with Redline because in the proposed system Redline moves only in horizontal direction. The participants also suggested possible uses of our TVODVT system in applications like security surveillance system, multimedia, and automobile infotainment system.

Further, our system is not quite as efficient as commercially available ODV players because our system is still in prototype phase. The architecture for thumbnail loading which loads pre-rendered thumbnails is not efficient and most of the time during evaluation it took more time to load the player that it would to load commonly available players with use of thumbnails. However, the use of thumbnails in video players is commercially utilized on 2D video players and in ODV. Thumbnails introduced in ODVs have the same functionality as in regular 2D videos but not optimized for 360 degree content. In our opinion, the mechanism of thumbnail loading could be revised to make ODV player more efficient and effective on low-end hardware.

Our evaluation study was task-based and the participants had to perform the given tasks with the use of three different interaction elements in the ODV player to navigate through videos. Prior to the first task the moderator introduced the system to the participants in order to avoid confusion in the case out of all four conditions to be performed by each participant. Real time interaction with TVODVT system with pilot video was also provided for better understanding of the system. In our opinion, this evaluation method best suits to evaluate this type of ODV system. During all 12 evaluations, no complaints regarding any stress while interacting with mouse and any type of viewing and interaction point's difficulty like on SB, MV and RL were mentioned.

The findings from the evaluation of the proposed TVODVT system are next discussed with respect to the research questions.

- What type of thumbnail view projections of ODVs best support efficient navigation in time?

Our objective evaluation results suggest that rectangular shape thumbnail with Redline is the most efficient of the proposed techniques with respect to task completion time. According to subjective evaluation in which five questions were asked from the participants, out of those questions four results were in the favor of spherical thumbnail with Redline.

- Is navigation through a thumbnail better than panning by using the main view on 360-degree video?

Thumbnail gives better idea of the whole 360-degree view. Using Redline with a thumbnail indicates the exact viewing angle in the main viewport. According to the subjective evaluation, thumbnail with Redline is very useful tool for navigation in ODVs and has better usability than the traditional panning on 360-degree video.

- Does the user have any subjective preferences over the type of thumbnails with or without Redline?

According to subjective evaluation, 11 out of 12 participants preferred thumbnails with Redline. Four participants suggested rectangular shape thumbnail with the use of Redline. Six participants suggested spherical shape thumbnail with the use of Redline. One participant liked both shapes of thumbnails with the use of Redline. One participant only liked the spherical thumbnail without the use of Redline.

- Lastly, how does the user find navigation and viewing experience in ODVs?

Overall, the participants liked the TVODVT system with respect to viewing experience in 360-degree with the help of thumbnail and introduction of Redline on thumbnail for navigation purposes. Some of them suggested improvements in the way of movement of

MV after interacting with Redline. One participant suggested vertical interaction with Redline.

All the feedback recorded after the interaction with the proposed TVODVT system suggest that interaction with the system is effective and feasible in the light of all research questions.

7. Conclusion and Future work

ODV is gaining popularity in mainstream media services like online video streaming, security surveillance, multimedia games and in other channels. Under the subject line of ODV and VR there is plenty of research going on. In all of that research, navigation and interaction through 360-degree video are one of the main topics under discussion. This thesis also ideates a concept related to ODV of which a prototype was made to evaluate that interaction.

The current research established TVODVT prototype with thumbnail interaction. Thumbnails of two different shapes were used for the evaluation of the proposed system. Evaluation was carried out with four different conditions with 12 participants in total. Each participant interacted with different ODVs in all four conditions. Completion time for each case was recorded. Overall the participants gave positive response to the TVODVT system. The findings of the research were discussed with respect to the research questions.

In short, the use of thumbnail with Redline on seek bar of ODVs has potential to be used as a navigation tool and can provide better viewing experience of 360-degree videos. This type of interaction can answer some of the problems with existing ODV players which are mostly used in online streaming websites or in other multimedia channels.

Further, after improvement in TVODVT system, another evaluation test could be conducted to assess the usability factor on handheld devices. The use of hand held equipment for viewing video content is increasing with the emerging trend of smart devices. More research is required in other scenarios regarding its usage and to refine the interaction and navigation techniques used in system. I am anticipative that the results of the proposed TVODVT system evaluation and future implementation will be a source of encouragement for researchers in this field.

References

[Atienza et al. 2016] R. Atienza, R. Blonna, M. I. Saludares, J. Casimiro, and V. Fuentes, "Interaction techniques using head gaze for virtual reality", 2016 IEEE Region 10 Symposium (TENSYMP), Bali, 2016, pp. 110-114. https://doi.org/10.1109/TENCONSpring.2016.7519387

[Bleumers et al., 2012] L. Bleumers, W. van den Broeck, B. Lievens, and J. Pierson, “Seeing the bigger picture: a user perspective on 360° TV.” In Proceedings of the 10th European conference on Interactive tv and video (EuroiTV '12). ACM, New York, NY, USA, 115-124. https://doi.org/10.1145/2325616.2325640

[Bobeth et al., 2014] J. Bobeth, J. Schrammel, S. Deutsch, M. Klein, M. Drobics, C. Hochleitner, and M. Tscheligi, “Tablet, gestures, remote control?: influence of age on performance and user experience with iTV applications.” In Proceedings of the 2014 ACM international conference on Interactive experiences for TV and online video (TVX '14). ACM, New York, NY, USA, 139-146. https://doi.org/10.1145/2602299.2602315

[Carlier et al., 2015] Axel Carlier, Vincent Charvillat, Wei Tsang Ooi, “A Video Timeline with Bookmarks and Prefetch State for Faster Video Browsing” MM '15 Proceedings of the 23rd ACM international conference on Multimedia Pages 967-970 Brisbane, Australia 26-30th October 2015. https://doi.org/10.1145/2733373.2806376

[Darnell, 2007] M. Darnell, “How do people really interact with TV?: naturalistic observations of digital tv and digital video recorder users.” Comput. Entertain. 5, 2, Article 10. 2007.

[Decock et al., 2011] Jan Decock, Jan Van Looy, Lizzy Bleumers and Philippe Bekaert, “The Pleasure of Being (There?). An Explorative Study into the Effects of Presence and Identification on the Enjoyment of an Interactive Theatrical Performance using Omnidirectional Video.” In Proceedings of the International Society for Presence Research Annual Conference. ISPR 2011. Edinburgh Napier University, Edinburgh, Scotland, UK. https://doi.org/10.1007/s00146-013-0487-6

[De la Torre et al., 2005] Fernando De la Torre, Carlos Vallespi-Gonzales, Paul Rybski, Manuela Veloso, and Takeo Kanade, “Omnidirectional video capturing, multiple people tracking and identification for meeting monitoring. Technical report.” http://repository.cmu.edu/robotics/128/

[Gonzalo and Balakrishnan, 2003] Gonzalo Ramos and Ravin Balakrishnan, “Fluid Interaction Techniques for the control and Annotation of Digital Video” In Proceedings of the 16th annual ACM symposium on User interface software and technology (UIST '03). ACM, New York, NY, USA, 105-114. DOI: https://doi.org/10.1145/964696.964708

[Hakulinen et al., 2018] Jaakko Hakulinen, Keskinen Tuuli, Makela Ville, Saarinen Santeri and Turunen Markku, “Omnidirectional Video in Museums- Authentic, Immersive and Entertaining.” A. D. Cheok et al. (Eds.): ACE 2017, LNCS 10714, pp. 567–587, 2018. Springer International Publishing AG. https://doi.org/10.1007/978-3-319-76270-8_39

[Huang and Hsu, 2003] C.-M. Huang and T.-H. Hsu, “A user-aware prefetching mechanism for video streaming”, World Wide Web, 6(4):353–374, 2003. https://doi.org/10.1023/A:1025661921237

[Järvinen and Ekola., 2014] Anu-Maria Järvinen and Leena Ekola, “Turning point – A practical assessment of using 360 video in sign language interpreting studies.” Diaconia University of Applied Sciences. Degree Programme in Sign Language Interpretation. Bachelor Thesis. http://urn.fi/URN:NBN:fi:amk-201404295266

[Kallioniemi et al., 2017] Pekka Kallioniemi, 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, 299-318. https://doi.org/10.1007/978-3-319-68059-0_20

[Maarten et al., 2016] Wijnants Maarten, Erum Kris Van, Quax Peter, Lamotte Wim, “Augmented ODV: Wed-Driven Annotation and Interactivity Enhancement of 360 Degree Video in Both 2D and 3D” Web Information Systems and Technologies. WEBIST 2015. Lecture Notes in Business Information Processing, vol 246. Springer, Cham. https://doi.org/10.1007/978-3-319-30996-5_3

[Matejka et al., 2013] Justin Matejka, Tovi Grossman, George Fitzmaurice, “Swifter: Improved Online Video Scrubbing”, In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '13). ACM, New York, NY, USA, 1159-1168. DOI: https://doi.org/10.1145/2470654.2466149

[M. Geyer and Daniilidis, 2003] Christopher M. Geyer, Kostas Daniilidis, “Omnidirectional Video”, The Visual Computer, Volume 19, Number 6, October 2003, pages 405-416. http://repository.upenn.edu/cis_papers/8

[Onoe et al., 1998] Yoshio Onoe, Kazumasa Yamazawa, Haruo Takemura, and Naokazu Yokoya, “Telepresence by Real-Time View-Dependent Image Generation from Omnidirectional Video Streams.” Comput. Vis. Image Underst. 71, 2 (August 1998), 154-165. http://dx.doi.org/10.1006/cviu.1998.0705

[Pakkanen et al, 2017] Toni Pakkanen, Jaakko Hakulinen, Tero Jokela, Ismo Rakkolainen, Jari Kangas, Petrim Piippo, Roope Raisamo and Marja Salmimma, “Interaction with WebVR 360-degree Video Player: Comparing Three Interaction Paradigms” 2017 IEEE Virtual Reality (VR), 279-280. https://ieeexplore.ieee.org/abstract/document/7892285/

[Rizzo et al., 2003] Albert A. Rizzo, Kambiz Ghahremani, Larry Pryor, and Susanna Gardner, “Immersive 360-degree panoramic video environments: research on creating useful and usable applications.” In Human-computer interaction: theory and practice, Vol 1, J. Jacko and C. Stephanidis, Eds. Lawrence Erlbaum Associates, Mahwah, NJ, USA, pp. 1233–1237.

[Saarinen et al., 2017] Santeri Saarinen, Ville Mäkelä, Pekka Kallioniemi, Jaakko Hakulinen, and Markku Turunen, “Guidelines for Designing Interactive Omnidirectional Video Applications.” 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, 263-272. https://doi.org/10.1007/978-3-319-68059-0_17

[Schoeffmann and Boeszoermenyi, 2009] Klaus Schoeffmann and Laszlo Boeszoermenyi, “Video Browsing Using Interactive Navigation” 7th International Workshop on Content-Based Multimedia Indexing, CBMI 2009, June 2009. https://doi.org/10.1109/ICME.2009.5202582

[Schoeffmann and Hudelist, 2015] Klaus Schoeffmann and Marco A. Hudelist, “Video Interaction Tools A survey of Recent Work” ACM Comput. Surv. 48, 1, Article 14 (September 2015), 34 pages. DOI: https://doi.org/10.1145/2808796

[Suporn et al., 2010] Suporn Pongnumkul, Jue Wang, Gonzalo Ramos and Michael Cohen, “Content-Aware Dynamics Timeline for Video Browsing” In Proceedings of the 23nd annual ACM symposium on User interface software and technology (UIST '10). ACM, New York, NY, USA, 139-142. DOI: https://doi.org/10.1145/1866029.1866053

[Unite Europe, 2017] https://unite.unity.com/

[webGL] https://www.khronos.org/webgl/
https://developer.mozilla.org/en-US/docs/Web/API/WebGL_API

[Zoric et al., 2013] G. Zoric, L. Barkhuus, A. Engström, and E. Önnevall. “Panoramic video: design challenges and implications for content interaction.” In Proceedings of the 11th european conference on Interactive TV and video (EuroITV '13). ACM, New York, NY, USA, 153-162.

Appendix A: Background Questionnaire

Gender: Male Female

Age group: <18years 18-28 years 29-40 years >40 years

Are you familiar with online video streaming players like YouTube, Vimeo?

Yes No

Are you familiar with 360* videos?

Yes No

Are you familiar with 360* video players?

Yes No

Appendix B.1: Condition Evaluation Questionnaire

Please evaluate the interaction used to pan the view.

Was the technique easy to use?

1	2	3	4	5	6	7
Very difficult			neutral			Very easy

Was the technique inventive?

1	2	3	4	5	6	7
Conventional						inventive

Was the technique efficient?

1	2	3	4	5	6	7
Inefficient						efficient

Was the technique practical?

1	2	3	4	5	6	7
Impractical						practical

Was the technique pleasing?

1	2	3	4	5	6	7
Unlikeable						pleasing

Overall comments?

Appendix B.2: Condition Evaluation Questionnaire Answers

Participant	C1					C2					C3					C4				
Pilot test 1	4	5	6	6	5	4	5	6	6	3	-	-	-	-	-	-	-	-	-	-
Pilot test 2	4	3	4	5	4	4	4	4	4	4	2	2	2	2	2	-	-	-	-	-
P1	7	7	7	7	7	7	5	5	7	7	7	7	7	7	7	7	7	7	7	7
P2	7	6	6	6	7	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5
P3	7	6	7	7	6	7	6	7	7	6	6	6	7	7	5	6	6	7	7	5
P4	3	5	4	4	4	2	4	4	3	3	4	4	4	4	4	4	4	3	3	3
P5	6	6	7	7	7	6	7	7	7	7	5	6	6	6	5	5	5	5	5	4
P6	7	7	7	7	7	6	6	5	6	6	6	6	6	6	6	7	7	7	7	7
P7	6	5	5	5	5	6	5	5	5	5	7	6	6	4	5	6	5	5	5	5
P8	5	4	6	6	5	5	4	5	6	5	5	5	5	5	5	5	5	5	5	5
P9	7	5	6	7	6	7	5	6	6	6	5	5	5	6	5	6	6	6	5	6
P10	6	7	6	6	7	6	6	6	7	6	7	7	6	7	7	7	7	7	7	7
P11	7	5	5	5	6	5	6	6	5	4	7	7	7	7	6	6	5	4	6	6
P12	3	5	5	5	4	3	5	5	5	4	3	5	5	5	4	3	5	5	4	4

Appendix C.1: Post experiment Questionnaire

Which of the navigation technique do you prefer?

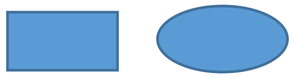
With Redline/without Redline

Which type of thumbnail gave the best 360 view on the seek bar?

Rectangular shape Thumbnail/Spherical shape thumbnail

Which of the navigation technique was the best overall?

A. Spherical thumbnail, with Redline
B. Rectangular thumbnail, with Redline
C. Rectangular thumbnail, without Redline
D. Spherical thumbnail, without Redline



Rectangular Spherical

Did you find any of the techniques confusing? Yes No

If Yes, why?

Would you like to use this type of 360* video player in your smart electronic devices? Yes No

If no, why?

If yes, why?

Overall comments?

Appendix C.2: Post experiment Questionnaire

Participant	Preferred Technique	Thumbnail gave best 360* view on the seek bar	The best navigation technique	Any of the technique confusing	360* video player for your smart devices	Overall comments
Pilot test 1	with Redline	Spherical shape	Spherical thumbnail with Redline	No	Yes	
Pilot test 2	with Redline	Spherical shape	Spherical thumbnail with Redline	Yes, B Difficult to navigate to require side.	Yes, Sometimes it can be useful to point out the exact location on screen in just one click.	
P1	without Redline	Spherical shape	Spherical thumbnail without Redline	No	Yes, If I have a company then for security purpose	Both technique are good but spherical thumbnail without Redline is well and up to date.
P2	without Redline	Rectangular shape	Rectangular thumbnail with Redline	No	Yes, Efficient.	
P3	with Redline	Rectangular shape	Rectangular thumbnail with Redline	No	Yes, Interesting to have all visions around me.	
P4	with Redline	Spherical shape	Spherical thumbnail with Redline	Yes, Moving R/L was reversed to what I am used to.	Yes, Easier when works correctly.	1)start with Redline in the middle of thumbnail 2)Add vertical interaction to the Redline
P5	with Redline	Rectangular shape	Rectangular thumbnail with Redline	Yes, I could not able to navigate properly	Yes, It is very handy for 360* overview.	Great Job!
P6	with Redline	Both	Both thumbnails with Redline	No	Yes, Those were easy to use and navigate as compared to other known tools.	

P7	with Redline	Spherical shape	Spherical thumbnail with Redline	No, But may be a bit. The one with inverted control.	Yes, with some adjustment I see it useful.	Anything with 360* videos is interesting to me and I believe this is a good use of a navigation system for 360* videos.
P8	with Redline	Spherical shape	Spherical thumbnail with Redline	No	Yes, Fun and reminds me FPS games	Nice Job!
P9	with Redline	Spherical shape	Spherical thumbnail with Redline	No, Not the techniques themselves, But the mouse direction was counter intuitive. (Takes time getting used to)	Yes, New technologies are always important and something I always look forward with AR. A player like this would be very interesting but should work 360* from top to bottom and side to side.	The more it is used, the more practical and easy it gets. The thumbnails didn't make much of a difference outside of scrolling to video frames but mouse navigation could be inverted.
P10	with Redline	Rectangular shape	Rectangular thumbnail with Redline	Yes	Yes	Redline is helpful. Otherwise, it is not easy to look things in 360*.
P11	with Redline	Spherical shape	Spherical thumbnail with Redline	No	Yes, Easy to scroll faster	Scrolling with Redline was very good and scrolling without Redline is not smooth.
P12	with Redline	Spherical shape	Spherical thumbnail with Redline	Yes, Because view moves-changes quite quickly to notice the things in the view.	Yes, Because it helps to have a better view and to understand what is happening in video more effectively.	Good idea and inventive approach to provide better view of a video.

Appendix D: Objective Evaluation test results

Video Type	Participant 1				Participant 2				Participant 3			
Type 1	C1	MV	RL	SB	C2	MV	RL	SB	C3	MV	RL	SB
	No.	12	2	0	No.	35	0	2	No.	18	/	2
	T	74664	3145	0	T	77488	0	9193	T	129147	/	63609
Type 2	C2	MV	RL	SB	C3	MV	RL	SB	C4	MV	RL	SB
	No.	11	0	0	No.	21	/	0	No.	16	/	1
	T	93217	0	0	T	57027	/	0	T	113825	/	57751
Type 3	C3	MV	RL	SB	C4	MV	RL	SB	C1	MV	RL	SB
	No.	6	/	10	No.	4	/	0	No.	15	0	9
	T	33225	/	74248	T	25130	/	0	T	58648	0	243168
Type 4	C4	MV	RL	SB	C1	MV	RL	SB	C2	MV	RL	SB
	No.	3	/	0	No.	16	/	0	No.	13	0	0
	T	37696	/	0	T	20952	/	0	T	53456	0	0
Video Type	Participant 4				Participant 5				Participant 6			
Type 1	C4	MV	RL	SB	C1	MV	RL	SB	C2	MV	RL	SB
	No.	9	/	0	No.	16	4	0	No.	18	4	1
	T	27888	/	0	T	74099	7744	0	T	50156	23907	2212
Type 2	C1	MV	RL	SB	C2	MV	RL	SB	C3	MV	RL	SB
	No.	32	5	0	No.	8	10	0	No.	9	/	0
	T	103684	26586	0	T	18712	37954	0	T	32081	/	0

Type 3	C2	MV	RL	SB	C3	MV	RL	SB	C4	MV	RL	SB
	No.	14	1	0	No.	15	/	5	No.	42	/	0
	T	53573	2656	0	T	31168	/	46051	T	85291	/	0
Type 4	C3	MV	RL	SB	C4	MV	RL	SB	C1	MV	RL	SB
	No.	11	/	0	No.	25	/	0	No.	3	7	0
	T	54414	/	0	T	63376	/	0	T	12681	35224	0

Video Type	Participant 7				Participant 8				Participant 9			
Type 1	C3	MV	RL	SB	C4	MV	RL	SB	C1	MV	RL	SB
	No.	38	/	1	No.	15	/	0	No.	12	1	0
	T	112348	/	3478	T	27626	/	0	T	76981	2671	0
Type 2	C4	MV	RL	SB	C1	MV	RL	SB	C2	MV	RL	SB
	No.	17	/	0	No.	13	6	0	No.	8	0	0
	T	77185	/	0	T	20532	16617	0	T	32151	0	0
Type 3	C1	MV	RL	SB	C2	MV	RL	SB	C3	MV	RL	SB
	No.	6	4	2	No.	10	1	2	No.	8	/	10
	T	10885	20854	55535	T	20767	5910	20572	T	42281	/	450806
Type 4	C2	MV	RL	SB	C3	MV	RL	SB	C4	MV	RL	SB
	No.	15	0	0	No.	10	/	0	No.	11	/	0
	T	17795	0	0	T	23505	/	0	T	34120	/	0

Video Type	Participant 10				Participant 11				Participant 12			
Type 1	C2	MV	RL	SB	C3	MV	RL	SB	C4	MV	RL	SB
	No.	20	3	0	No.	39	/	0	No.	10	/	0
	T	57383	24499	0	T	57414	/	0	T	62958	/	0

Type 2	C3 / MV / RL / SB	C4 / MV / RL / SB	C1 / MV / RL / SB

Type	Cond.	Row	MV	RL	SB
Type 2	C3	No.	17	/	0
		T	66859	/	0
	C4	No.	51	/	15
		T	115228	/	95298
	C1	No.	10	1	0
		T	54876	1073	0
Type 3	C4	No.	5	/	0
		T	11650	/	0
	C1	No.	11	1	4
		T	38470	7368	45134
	C2	No.	3	1	0
		T	12167	15782	0
Type 4	C1	No.	4	6	0
		T	9825	16724	0
	C2	No.	16	2	0
		T	40971	3097	0
	C3	No.	9	/	0
		T	52336	/	0

Note: Time in millisecond.

C1: Condition 1

C2: Condition 2

C3: Condition 3

C4: Condition 4

MV: Main View

RL: Redline

SB: Seek bar

T: Time

Appendix E: Tasks Answers

Participant	C1					C2						C3				C4				
Pilot test 1	2	6	1	A1	left	3	2	3	1	4	1	2	yes	2	1	4	1	yes	1	yes
Pilot test 2	2	5th last	1	A1	left	3	2	3	1	1	1	2	No	2	1	4	1	yes	1	yes
P1	2	1st floor	1	A1	left	3	1	2	1	1	1	2	yes	2	1	4	1	yes	1	yes
P2	2	6	1	A1	West	3	3	3	1	1	1	2	yes	2	1	4	1	yes	1	yes
P3	2	6	1	A1	left	3	2	3	3	1	1	2	yes	2	1	4	1	yes	1	yes
P4	2	6th	1	A1	back-right	3	2	3	2	2	1	2	yes	3	1	4	1	yes	1	yes
P5	2	6th	1	A1	right	3	3	1	1	1	1	2	yes	2	1	4	1	yes	1	yes
P6	2	6	1	A1	back	3	2	3	1	13	1	2	1	2	1	4	1	1	3	yes
P7	2	6th	1	A1	right	3	2	3	1	1	1	2	yes	2	1	sorry	1	yes	1	yes
P8	2	1st	1	A1	right	3	2	3	1	1	1	2	yes	2	2	4	1	yes	1	yes
P9	2	3	1	A1	left	3	2	3	2	5	1	2	1	2	1	4	1	yes	1	yes
P10	2	6	1	A1	right	3	2	3	1	1	1	2	yes	2	1	4	1	yes	1	yes
P11	2	3rd	1	A1	right	3	3	3	1	2	1	2	yes	2	0	4	1	yes	1	yes
P12	2	6	1	A1	back-side	3	3	3	1	1	1	2	yes	2	1	4	1	yes	1	yes