User Experience of Panoramic Video in CAVE-like and Head Mounted Display Viewing Conditions

Adam Philpot
Middlesex University
London, UK
a.philpot@mdx.ac.uk

Maxine Glancy
BBC Research & Development
Salford, UK
maxine.glancy@bbc.co.uk

Peter J Passmore
Middlesex University
London, UK
p.passmore@mdx.ac.uk

Andrew Wood
BBC Research & Development
London, UK
andrew.wood1@bbc.co.uk

Bob Fields
Middlesex University
London, UK
b.fields@mdx.ac.uk

ABSTRACT
Panoramic 360 video is a rapidly growing part of interactive TV viewing experience due to the increase of both production by consumers and professionals and the availability of consumer headsets used to view it. Recent years have also seen proposals for the development of home systems that could ultimately approximate CAVE-like experiences. The question arises as to the nature of the user experience of viewing panoramic video in head mounted displays compared to CAVE-like systems. User preference seems hard to predict. Accordingly, this study took a qualitative approach to describing user experience of viewing a panoramic video on both platforms, using a thematic analysis. Sixteen users tried both viewing conditions and equal numbers expressed preferences for each display system. The differences in user experience by viewing condition are discussed in detail via themes emerging from the analysis.

ACM Classification Keywords
Human-centered computing User studies.

Author Keywords
panoramic video; CAVE; HMD; user experience

INTRODUCTION

Panoramic (also referred to as 360 degree) videos are produced using an array of cameras and video stitching software. The majority of panoramic videos currently available are monoscopic and typically viewed on computers, hand held phones or Head Mounted Displays (HMDs). The increasing availability of consumer and professional panoramic video cameras means that the amount of panoramic content is growing rapidly along with the sale of consumer HMDs used to view them. For example, Samsung [22] recently announced that over five million Gear VR headsets have been sold and over ten million hours of video has been watched. A large amount of these videos can be assumed to be panoramic (Samsung has not released a breakdown of this content, some of it may have been watching ordinary videos in 360 virtual cinemas). Compared to traditional TV, panoramic video viewed on a HMD is interactive, as the user has the ability to rotate their view to look anywhere around them. It is also reported to be more immersive than traditional TV, and to engender a sense of presence. The advent of easy to use consumer 360 cameras has lowered the threshold to producing content. At the same time, media companies are increasing production of 360 content; the New York Times is producing a new 360 video every day [27].

While HMDs are becoming cheap, CAVE-like systems are expensive and require dedicated space. However, in recent years, with developments in projection mapping and Augemented Reality (AR), a number of systems have been developed that could be used to approximate CAVE-like systems in the home, for example Microsofts Illumiroom [13] and RoomAlive projects [12], and Razer's Project Ariana [21]. Furthermore, inside out tracking AR systems, like Microsoft HoloLens and Intel's Project Alloy, can simulate projection mapping, and could (with wider field of view) also approximate CAVE-like systems. Consequently it is of interest to compare user experience in HMDs and CAVE-like systems.

HMD and CAVE-like Systems

While a HMD is an effective solution for viewing 360 video, indeed, 360 video experiences are seen as a way to entice consumers to buy such devices, there are a number of different ways of viewing panoramic video. At the top end of Virtual

 The definitive version was published in TVX '17, June 14-16, Hilversum, Netherlands 2017. http://dx.doi.org/10.1145/3077548.3077550

Figure 1. A (anonymised for review) panoramic video (left), viewed in a CAVE-like environment (right).

Reality (VR) applications, CAVE-like systems that use projection onto purpose built planar surfaces (essentially rooms), are often considered the most sophisticated viewing systems.

A HMD and a CAVE differ in a number of respects. Firstly, the resolution differs between the displays with a current HMD roughly one megapixel per eye and a CAVE typically more than double that per screen; however, it is the resolution of the panoramic video (typically 4K) that currently limits resolution respectively. The HMD is completely immersive as the viewer is entirely surrounded by video, whereas in a 4-sided CAVE, video is only projected on to the front, left and right walls, and the floor. The Field Of View (FOV) in the CAVE is that of human vision, roughly 180 degrees, whereas on a typical current HMD it is much less (e.g. the Gear VR has roughly 95 degrees FOV). In terms of embodiment in the HMD, concerning panoramic video, there is typically no representation of body; while in the CAVE, embodiment is natural, as the viewer can see their own body. In terms of comfort, the HMDs are bulky to wear, have to be strapped to the head, and may feel uncomfortable. In terms of viewing position, the HMD naturally places the viewer in the centre of the viewing sphere, which is harder to do in the CAVE and may lead to an unnatural viewpoint. Also, in terms of physiology, in the CAVE, vision is normal, whereas for the HMD, accommodation and vergence are decoupled, leading to possible discomfort [8]. For HMDs, the user is typically alone in the virtual world (although work in social VR is rapidly progressing), while in the CAVE, the space is big enough to naturally accommodate a number of viewers at the same time.

Therefore HMDs and CAVEs differ in many ways, and there are many reasons that a viewer may prefer one system over the other; making it hard to predict which systems users will prefer. The aim of this study is to discover, in detail, how the user experience of these viewing conditions differ, allowing viewers to express themselves in their own words.

RELATED WORK

Cinematic Virtual Reality refers to cinematic experiences in Virtual Reality which range from monoscopic 360 degree panoramic video, to various forms of stereoscopic 360 video, including those derived from computational photography (such as Google Jump), to emerging forms of navigable video. Content may be filmed, computer generated, or a mixture of both. The development of Light Field capture systems such as the Lytro Immerge [10] and playback systems such as Nozon Presenz [20] suggest that navigable six degrees of freedom video is on the horizon. However, monoscopic panoramic video currently predominates.

While CAVE systems have been around for some time [4], they are expensive and require dedicated room scale spaces to accommodate a permanent installation. There are a range of intermediate systems that could achieve some of the functionality of a CAVE and possibly be available to consumers to use in the home. The Ambilight system, an ambient lighting system [24], used low resolution projections around a TV to extend projection beyond a TV set into a room. The BBC's Surround Video system [18], projected a simultaneously filmed wide angle view onto the walls surrounding a television, to produce a more immersive experience. More recently, a higher resolution system, SurroundVideo+, featuring a TV with projected peripheral content in a three walled CAVE-like system [16], was used in user studies along with a TV and HMD conditions. Microsoft's IllumiRoom took the surround video concept a step further by using 3D scanning with the kinect and projection mapping to project images onto the walls around a games console [13] in a home environment. The Illumiroom was followed up by the RoomAlive project, which used multiple projectors and six kinects to produce higher definition room scale projection [12]. Razer has recently demonstrated Project Ariana, which also uses 3D scanning, a pair of 3D cameras - to scan the room geometry, detect lighting and the monitor, and use projection mapping to project graphics onto the surrounding environment [21].

Apart from projection based systems, recently developed, inside out, Augmented Reality (AR) systems like Microsoft's HoloLens [17], or Intel's Alloy system [11], could ultimately (with wider FOV) also be used to simulate CAVE systems in

real world environments. These systems build up 3D models of the environment so that graphics can be 'pinned' to the geometry in the environment.

Comparative Studies

There have been a range of comparative studies including HMDs or CAVEs over a range of scenarios, with varying results. In [23] collaboration between CAVE-like environments, was compared to collaboration with a CAVE and a desktop environment, on a Rubik's cube type puzzle. They found the best task performance with the CAVE-like system collaboration. Another study [1] investigated comparing human behaviour between HMD and a four-sided Spatially Immersive Display (SID), looking at natural turning (head movement) compared to manual turning (display rotation). They found that subjects have a significant preference for real turns in the HMD, and for virtual turns in the SID. In [7] performance on an oil well planning task was compared on a stereo enabled desktop display, and a CAVE-like system, with the latter system found to facilitate faster and more accurate performance. In [26] an experimental comparison of interaction in the real world and a CAVE virtual environment was carried out, varying interaction with and without virtual hands and comparing two manipulation tasks. They found users took longer and made more errors in the CAVE environment. In an analysis of Virtual Reality Induced Symptoms and Effects (VRISE) [25] results indicated that 60-70% of participants experience an increase in symptoms pre-post exposure for HMD, projection screen and reality theatre viewing conditions. Their most notable finding of inter- and intra-participant variability highlights the variability of individual susceptibility to VRISE. In a study on anxiety [14] showed that anxiety produced by fear of heights was higher in a CAVE compared to in a HMD. In [15] performance on a modified Stroop task under low and high stress conditions, was compared for desktop, HMD and CAVE-like viewing conditions. The CAVE-like system induced the greatest sense of presence. One study looked at how viewing of high or low emotional impact videos, viewed on a HMD compared to a tablet, affected the viewers pro-environmental attitude and behaviour [5]. They found that the increased immersion and higher emotional impact of the HMD appeared to increase self reported measures, but that higher immersion did not increase pro-environmental behaviour significantly. In [3] the higher resolution successor to the CAVE, the CAVE2, was compared to the HMD for collaborative sense making, in particular, the collaborative analysis of network connectivity. They found that participants using HMDs were faster than for CAVE2, and no different for accuracy or communication.

The current study aims to add to the body of comparative work by examining user experience of 360 video in HMDs and CAVEs.

METHODS

Aim and Experimental Conditions

The aim of the study was to compare the user experience of a monoscopic 360 video viewed either in a CAVE or HMD. An eight minute documentary profile of an artist was used. The video was a character led study of the artist's exploration of beekeeping and music making. The video was viewed either on a Samsung Gear VR using a Samsung Note 4 phone, hereafter referred to as the HMD condition, or on a CAVE-like system comprising three walls and a floor, hereafter referred to as the CAVE condition.

Subjects

The study was carried out with sixteen participants comprising twelve male and four female subjects in the age range from nineteen to fifty two. The subjects were mostly employees of a media company along with three university students. Each subject viewed the video once per condition, and the order of viewing for any two conditions was alternated to counter order effects (seven viewed the CAVE condition first, nine viewed the HMD condition first). All subjects had tried VR before but none had experience of a CAVE system previously. Subjects were volunteers that were recruited via email to all employees in the media company, and all members of a university based human computer interaction interest group.

Procedure

For the HMD condition, viewers sat in an office swivel chair. For the CAVE condition viewers sat on an inflatable white sofa, aligned with the (missing) back wall of the space, see Figure 1.

Subjects were first shown a short panoramic video clip about a Chinatown to orient them to the task and to familiarise them with the viewing scenario. Subjects then viewed the main video using their assigned viewing mode. Following this, a semi-structured interview was used to elicit conversation about the viewing experience. Guideline questions that could be used in the interview included: "How do/did you feel after/while watching the video? Did you find any differences to 'normal' videos that you watch on your phone, PC or TV? What's good about watching videos this way? What's bad about watching videos this way? Can you recall any moments that made you feel inclined to look around more? How much did the video hold your attention or focus?" A final question was: "Which viewing condition did you prefer and why?". These questions were the same for each participant and allowed expansion on replies of interest. The participants were video recorded and observed throughout the study.

Following the second viewing and interview, the participant was asked to complete six questions in the form of a Likert scale for each condition undertaken. The subjects were asked to circle the number: from 1 (not at all) to 5 (very much), that best matched their response to each statement. The questions and the median and mean responses to the questions are shown in the Results section in Table 2.

The general procedure was thus: users completed a pre-trial consent form, underwent orientation for the first platform, the user watched the documentary using the first platform, was interviewed about the first platform, underwent orientation for the second platform, the user watched the documentary using the second platform, was interviewed on the second platform, and finally, the user completed the Likert scale questionnaire for each viewing condition.

Thematic Analysis

The thematic analysis methodology of Braun and Clark [2] was used for the analysis of the data collected during the study. The overall aim of the analysis is to capture, as a collection of 'themes', an understanding of what is really going on in the mass of data captured in the full set of interview transcripts. The data consisted of sixteen transcribed interviews from the video recordings. The transcriptions were loaded into NVivo 10, a qualitative data analysis computer software package. The software was used to code relevant sentences in the transcripts and in the analysis phase for querying the data when coded, (e.g. example query: display all sentences coded with both the presence and the HMD themes). Transcripts were coded using an open coding procedure, during which the coding scheme was inductively defined and refined as the coding proceeded, very much in the spirit of Grounded Theory's *constant comparative* method [9, 6]. A starting point for the analytic process was a set of seven codes found to be relevant in a previous study [19]. These codes are shown in the left hand column of Table 1 and are referred to as the Established Themes in subsequent discussion.

Items of the interview data were considered in turn, and compared to the emerging coding scheme, to find existing codes that apply, to refine the definition of previously generated codes, or to produce new codes as appropriate. Thus, while reading the text, 'nodes' (as they are termed in NVivo) or themes, were created as necessary and sentences assigned to them. The nodes were created according to the judgement of one of the authors and refined and modified during the coding process. Subsequent words or sentences were allocated to the newly formed nodes, with new nodes created as and when the author felt they were needed. Five of the sixteen transcripts were also coded by a second author, and discrepancies between the two sets of nodes were resolved through discussion. The analysis resulted in the initial set of seven codes being extended by the addition of a further seven new codes, and the combined collection provided a structure for the discussion in the following sections.

Established Themes	Discriminating Themes
Presence.	The size of images.
Certainty (about what should be attended to).	Embodiment.
Comfort.	Peripheral Vision.
Attention.	Projection of image on self.
Concentration on story.	Physical surfaces particularly the floor.
Engagement.	Cube effects.
Social ease.	Confined or Trapped.

Table 1. Themes

RESULTS - ESTABLISHED THEMES

As a starting point to guide the thematic analysis, we used the Established Themes described in the previous section. In this section, we concentrate on themes for which user experience is similar for both conditions, under these themes, the two experiences are only subtly different. In the next section we will discuss themes that discriminate between user experience of the two viewing conditions, which are referred to as the Discriminating Themes, and which are displayed in the right hand column of Table 1.

The following text contains quotes from users and are labelled thus [User ID, the condition they are talking about, H or C - short for HMD or CAVE] e.g. [P99,H].

Presence

Both CAVE and HMD viewers referenced presence in various ways. Some were more explicit than others, with some users suggesting presence through their use of words by describing parts of the video as when they were 'in' the bee hive, studio or park.

A number of users referred to the environment and/or video as being immersive, or feeling immersed in the video or environment. It was quite similar for both conditions with neither appearing to be more immersive than the other. Attributes which contributed to the users immersion included feeling as though they were a part of the environment: *"You are in a more real environment and that made you feel you are actually in there; in the display."*[P12,H], and *"Definitely, the actual sense of being sat down watching something happening and being sat in that moment."*[P11,C]. Many users' comments referred to being 'in' particular scenes of the video which contributed to their increased sense of immersion and presence.

Some users felt that the HMD was more immersive than the CAVE: *"It's probably more immersive because you're actually sitting within and obviously there's the 360 element."*[P6,H], and *"Here [CAVE] I had a good overview, but it just wasn't as immersive."*[P15,C].

Some users mentioned the height of the camera in specific scenes of the video. When the camera position was high up, some users mentioned feeling tall while another felt a sense of vertigo, and when the camera position was lower some users felt small: *"You're in the bee keepers hut and you're very tall, and then you're outside and you're at the vent where bees are coming out, and you're really small again."*[P8,H].

One particular scene, where the camera is positioned between the Bee Man and a table, and the Bee Man is working with some electronics, provoked many comments, mainly from HMD viewers, both positive and negative: *"Then I was thinking, it feels like I'm in his lap or something, that feels weird."*[P7,H], and *"One really good bit was, it was kind of a first-person perspective where he was fiddling with the controls so I could see his hands and see all the different electronic bits."*[P15,H].

Some users referred to feeling like a 'fly on the wall': *"I was just looking and it felt like I was a fly on the wall up in the ceiling."*[P7,H]. Some users remarked on the size that they felt within the scene also relating that their reduced size made them feel more a part of the scene, for example, feeling small in the scene when viewing in the CAVE: *"There's a marked difference in this environment, because you feel like you're shrunk down and you're much more in the scene."*[P14,C].

Certainty (about what should be attended to)
Users in both conditions expressed concerns about whether they were looking in the correct direction or where the focus of the action was and whether they were missing anything by looking somewhere else. CAVE users did not express a fear of missing out directly, but hinted at the fact that there may be other things to look at while watching one area of the video: *"It wasn't like I was focused on one point, I could kind of just glance left and right and still see a bit of what was going on."*[P8,C]. Whereas HMD users were slightly more specific with mentioning that they could miss information somewhere else; while looking in a particular direction, or when looking around the environment: *"I was wondering if anything else is going on at the time, whether I was looking at the right thing."*[P11,H].

Many users sited a scene change as a good time to look around. This was to familiarise themselves with their surroundings and to explore the environment: *"I think when the scene changed, I'd look around to get to understand where I am, so to speak."* [P13,C]. Once the scene remained the same for a period of time, some users said that they felt that it was a good time to look around, also when they found themselves not so interested by a particular part of the video: *"I think essentially, whenever a shot was kind of in the same place, not moving for 10 or 15 seconds, that's when I start looking around."*[P16,C]. Some users said that while the character/Bee Man was talking, they felt that they could listen and look around at the same time without fear of missing anything: *"I was listening to what he was saying, and looking at what he was doing, just looking around the environment as he was talking."*[P9,C]. Additionally, the fixed voice to camera affords users to look around confidently.

Comfort
Some CAVE users remarked about how relaxing and comfortable it was to view the video in the CAVE. This could be attributed to the fact that they were sitting on a soft sofa (as opposed to a swivel chair in the HMD condition) and they can relate to the situation as the way that they usually watch videos: *"[I preferred] The CAVE, because I generally felt more relaxed. I felt more comfortable, it was nice to be surrounded and it was atmospheric. This one [HMD] just felt not as comfortable."*[P1,H]. Some users directly compared the relaxation of the CAVE to the HMD: *"Sitting on a sofa without a head thing on is a bit more comfortable."*[P13,H], and *"Compared to the headset, it was more relaxed because it's not so... There's nothing pressed up against your face."*[P14,C].

None of the users exhibited any signs of cyber sickness, though two users mentioned that they felt slightly disoriented when removing the HMD. However, they soon adjusted to the actual world, with no lasting implications: *"Once I'd taken the headset off, I feel a bit sort of um, my mind feels a little bit kind of lost for a moment, just trying to re-orientate myself as you come out of it."*[P8,H].

Some users found the HMD quite uncomfortable due to its weight pressing on the face, some eye strain due to the closeness of the screen to the eye and feeling slightly claustrophobic: *"Comfortable, though the weight of the actual headgear at times felt a bit uncomfortable."*[P13,H], and *"My eyes were beginning to strain a bit. There was slight discomfort."*[P3,H].

Attention
Both CAVE and HMD viewers appeared to be aware that most of the 'action' in the video took place in the forward viewing direction of the camera (in front), as in traditional documentary. This provided a reference point of where to return one's view following looking around or exploring the environment: *"Especially in a piece like this where most of the action takes place in front of you, so you know to go back to looking forward."*[P14,H], and *"The bulk of the activity still happened in front of you."*[P13,C]. In the CAVE condition, the use of a fixed seat (as opposed to a swivel chair in the HMD condition) encouraged this, and also discouraged viewers from turning and looking behind them; it was not essential to view behind, considering that the majority of the action was taking place in front of the camera: *"When I was in the CAVE, probably because my chair was fixed, I was looking in front of me."*[P9,H.

Some users noticed the absence of the ceiling and the back wall when viewing in the CAVE and others liked the fact that they could look all the way round and up when viewing using the HMD: *"Well you missed the ceiling and the back, but they don't matter at all."*[P10,C], *"I preferred the head set, because I could look all the way around."*[P5,H] and *"In the CAVE, I knew there wasn't a back wall, so I didn't look right around."*[P1,H].

Concentration on Story
Some users were able to recall specific information about the video at will when speaking about their experience. This occurred both following the first and second viewings regardless of condition: *"No I can't remember, but I remember a whole load of stuff about him being allergic to animals and stuff like that."*[P15,H], and *"The fact the guy made music from bees, and then there were all these alternative ways that he was finding to make the music from the bees, yeah, I found that quite interesting."*[P8,C].

Two users mentioned that they were sometimes following the visuals as opposed to what the Bee Man was saying: *"When I had the headset on, I wasn't listening as much to what he said because it's visually more immersive, so you don't listen, you're too busy looking round."*[P6,C]. Whereas others were able to concentrate on the narrative as well as look around: *"I was listening to what he was saying and looking at what he was doing, just looking around the environment as he was talking."*[P9,C].

Engagement
Twelve of the sixteen participants mentioned that they either had some interest in the video or found elements of the video interesting; if a user has no interest in the subject however, they do not engage with the content of the video: *"Nice to be surrounded by things and the story is interesting."*[P1,C], and *"You'd never put bees with music making, so that alone is interesting."*[P3,H]. Conversely, some users were not interested in the video. The video was described as many things,

ranging from unique, quirky and interesting to dark, sinister and weird. This eclectic collection of descriptors contributed to the interest and range of user responses exhibited towards the content of the video: *"I thought it was very silence of the lambs stroke Aphex Twin a little bit sinister."*[P6,C], and *"I felt that it was a little bit quirky. Towards creepy because they never got their hats off."*[P7,H].

Some users commented that sound was good to help create atmosphere and also blocked out the surrounding noise allowing enhanced engagement and focus: *"And especially with the sound it goes a long way because it tends to block out what you are hearing in the room you're in or where you are."*[P14,H]. The video production included a lot of work on the sound to attempt to achieve an enhanced viewer experience.

Social Ease

There was not much said by users; however it appears that users would not watch videos with other people while wearing HMDs as they found it quite a solitary, isolated experience: *"But all wearing your headsets in the living room in isolation, so it's an immersive experience, but you're quite isolated."*[P3,H]. The CAVE was also described as *"less for social watching."*[P1,C], but two users believed that one could watch a video in a CAVE with others and have an immersive experience: *"And being able to sit with someone else and watch a movie and still have the immersive experience, that would be really cool."*[P11,C]. It is not unthinkable that a sofa can be shared and, ultimately, the viewing experience in a CAVE. Users were specifically asked if they could imagine viewing with other people in the CAVE, surprisingly not many considered that they would.

RESULTS - DISTINGUISHING THEMES

The Size of Images

Some users liked the apparently larger scale of image in the CAVE (though the image size on the retina should have been the same for both conditions): *"I enjoyed the size of it, the ability to be able to look at really small details whilst still, because the main image is so big you can't miss it."*[P16,C].

Embodiment

Some users mentioned that they felt disembodied, more so when using the HMD. The fact that when they looked down and did not see their legs or any kind of body representation was quite strange for some users: *"It's a disembodied experience; I wasn't a person sitting in space, I was an undefined body voyeur."*[P4,H], and *"When I looked down I realised I wasn't standing there."*[P5,H]. One user quite liked the experience: *"That's what was interesting, I felt disembodied a load of the time. I really liked instances, where, again going back to disembodied, so put in places that weren't to scale with my body."*[P4,H].

One user reported feeling disoriented due to feeling like they were 'Hovering in mid-air': *"Initially it was disorientation, when you look down it's like you are hovering in mid-air."*[P5,H].

Peripheral Vision

Some users liked the wider field of view and that the image was 'everywhere': *"The wider view, so you could see more at the same time from a better perspective."*[P12,C], and *"What I like, more in theory than in actual, is the idea that nothing ends in my field of vision, which is great."*[P4,C]. HMD viewers did not mention peripheral vision

Some users really liked seeing things in their peripheral vision: *"I like the use of my peripheral vision. The use of my ability to use my peripheral vision."*[P9,C], and *"You can see, I suppose, the interesting difference in the headset; you seem to be restricted to your line of vision you have to look around to see it, but here you can catch it more easily out of your peripheral vision."*[P14,C].

In the CAVE some users were distracted by things happening in the background, such as a car moving past or other people walking by: *"Times he might be talking about stuff that's interesting but I'd be distracted by a car."*[P3,C]. This is related to having a wide field of view as in real life. One does not have to contend with this in usual TV viewing. One user pointed out that it was more visually stimulating and therefore found that they gave more attention to looking than listening: *"When I had the headset on, I wasn't listening as much to what he said because it's visually more immersive, so you don't listen, you're too busy looking round."*[P6,C]. HMD users were distracted by the ability to look around the environment, and while doing so, to not pay attention to what was happening in the video: *"Sometimes you were so consumed by the fact you were in this world, and you could look around, that it's easy to be distracted."*[P3,H]. Some users mentioned that they were not distracted by external influences due to wearing the HMD: *"It's not like you can be distracted by anything else going on outside of the screen, the screen is all you can see."*[P14,H].

Projection of Image on Self

The projection of the image onto the user in the CAVE provoked various comments. Some users found it quite pleasant and interesting; one user did not like it. Some users found that it increased the immersion, made them feel physically involved and added to the experience: *"What really did it was the projection on the floor and on you as well, it was noticeable and interesting."*[P3,C].

The projection of the image upon the user raised some interest with one or two users commenting to this effect. It did not increase engagement, but did add 'something' to the experience, one user commenting that: *"You notice it's on your legs therefore you sort of, there's an element of feeling physically involved."*[P6,C]. One user particularly did not like being projected upon stating: *"And I didn't feel like I liked it projected upon me, like the flowers or the bee hive upon me, it was like, 'eh?'."*[P7,C].

Physical Surfaces Particularly the Floor

Some users noticed that the image on the floor made a difference to their viewing experience and liked it; in one part of the video the electronics were projected on to the floor and they looked down to see that: *"Oh yeah yeah the floor especially that bit with the table when he was doing the, building some*

electronics stuff. So yeah that was based on the bottom, so definitely, I looked at the bottom part."[P12,C]. However, one user found it strange that they had to look down to the floor to see what was happening and one user did not like the image projected onto the floor: *"No, I think it was a little strange, for instance, looking down to the floor to see, to focus on the activity when he was pulling out the honeycomb."*[P13,C]. The user observations could be attributed to the novelty of the CAVE, as it is something different to the 'normal' way of viewing, on a TV, phone etc.

Cube Effects.

Something that CAVE viewers remarked upon was the visibility of the angles in the room where each wall met each other and the floor, at 90 degree angles: *"Things that let it down were you could see the joins between the wall and the floor, and that made it slightly less than the headset."*[P5,C], and *"With the CAVE, you've got these clear lines between it, which kind of throw you off a little bit with it."*[P8,C].

The video had scenes inside a shed, and other rooms, and some users felt that the fact that these scenes did not map to the walls and floor in the room via the projection was problematic: *"So the nature of the environment means for example the walls of the beehive didn't look straight because of where it was dissected by the lines of the room."*[P13,C], and *"Sometimes the wall was on the floor."*[P9,C]. Some users suggested a solution could be to project onto a dome or curved walls: *"I guess if it was a dome, it would be fantastic."*[P7,C].

Confined or Trapped

One user found that being cut off from the actual world when viewing the video using the HMD was a good thing: *"Specifically 360 video, you are so much more part of the experience because you, there's no getting away from it the same way."*[P14,H]. Whereas others found it unpleasant and in one instance claustrophobic: *"It's a bit claustrophobic at times compared to a normal usual screen."*[P16,H], and *"And the bee hive, it was so close that I wanted to get some distance."*[P7,H]. One user liked the fact that the HMD does not have any 'sides' like a TV screen: *"Just the fact that you can explore the environment more instead of being confined to the sides of a screen."*[P10,H].

RESULTS - PREFERENCE

Users were asked which condition (C or H) they preferred and why, at the end of their second viewing. Seven users preferred the HMD and seven users preferred the CAVE; two users could not decide and have been assigned the preference of 'BOTH'. In the following section users are denoted by their participant number with their preference following in parentheses.

Although users P5(H), P6(H) and P13(H) preferred viewing the video using the HMD, they all remarked that the CAVE was more comfortable. They found the HMD uncomfortable, simply due to the fact that they were wearing it, and in the CAVE they were not: *"The CAVE experience was more comfortable, because I didn't have this thing on my head."*[P5], *"But the headset's quite uncomfortable, so this [CAVE] is comfier."*[P6], and *"sitting on a sofa without a head thing on is a bit more comfortable."*[P13].

Users P1(C), P3(C) and P6(H) specifically mentioned the discomfort of the HMD: *"This one (H) just felt not as comfortable."*[P1], *"It's more immersive, you felt you were there. But with that comes discomfort."*[P3], and *"But the headset's quite uncomfortable."*[P6]. Also, pertaining to discomfort when using the HMD, two users, P3(C) and P16(C), both mentioned the feeling of claustrophobia when viewing using the HMD: *"It's less claustrophobic in the CAVE."*[P3], and *"The full immersion of the headset I found actually quite overbearing and claustrophobic."*[P16].

As well as the comfort and discomfort of the two conditions, three users, P1(C), P8(C) and P16(C) felt relaxed in the CAVE setting. This was not attributed to anything in particular; however the inflatable sofa could have been an influential factor: *"The CAVE because I generally felt more relaxed."*[P1], *"I was more relaxed in it. I felt I could just chill out and I didn't have this thing stuck to my face and that was more relaxing."*[P8].

Presence and Immersion was quite prominent in the reasons for some users' preference. Users P2(C), P3(C), P5(H), P6(H), P10(C), P12(H), P14(BOTH) and P15(H) all either specifically mentioned, or alluded to, presence and/or immersion as a key factor in their preference of condition: *"Gives a greater sense that you are there."*[P2], *"The headset felt far more as if you were 'in', it's more immersive, you felt you were there."*[P3], *"The headset, just because it's more immersive, more interesting."*[P6], *"Again it feels a lot more immersive."*[P10], *"You are in a more real environment and that made you feel you are actually in there; in the display."*[P12], *"I certainly felt more immersed in the headset."*[P14], and *"It's an immersive experience."*[P15]. Clearly, both conditions provide presence and immersion to the user which ultimately contribute to a pleasurable experience.

There were two users P11(BOTH) and P14(BOTH), who could not decide which they preferred, citing the dependency on what type of content they viewed in each condition as the crux of their indecision: *"For this particular bit of content the headset. But I would say, with different content, like a movie, probably this [CAVE]."*[P11], and *"Depends on content."*[P14].

Overall, the fact that the users were split equally in their preference of condition, in this instance, did not highlight whether one is preferred over the other. Although just two users mentioned content when speaking about their preference, it is likely to be quite influential in a viewer's experience when watching panoramic video using various conditions. The comfort and relaxation associated with the CAVE contributed to users' preference, but was not a distinguishing factor. The outstanding attribute relating to preference of condition in this study appears to have been whether the viewer experiences presence and immersion when watching panoramic video in a CAVE or using a HMD.

RESULTS - QUESTIONNAIRE

After completing the two viewing conditions, subjects were asked to rate their responses to six questions on an Likert scale of five points ranging from agree 'not at all' to agree 'very much'. The responses to these questions are summarized in

	Statement	CAVE Median	HMD Median	CAVE Mean	HMD Mean	P Value
S1	I felt like I was there, in the scenes of the video	3	4	3.25	3.94	0.044
S2	I felt I could interact with the displayed environment.	2	2.5	2.31	2.56	0.43
S3	I paid more attention to the displayed environment than I did to my own thoughts (e.g., personal preoccupations, daydreams etc).	3	4	3.31	3.88	0.169
S4	I felt as though I was in the same space as the character and/or objects.	3	4	3.25	3.81	0.147
S5	How much did you enjoy the content of the clip	4	4	3.69	3.63	0.705
S6	How much did you enjoy the way you viewed the clip?	4	4	4.06	3.88	0.417

Table 2. Median and mean values for the responses to the instructions "Please circle the number; from 1 (not at all) to 5 (very much), that best matches your response to each statement". The last column is the P value from Related Samples Wilcoxon Signed Rank Tests (n=16) on responses to the six statements. Only S1 was significant at P=0.05 level.

Table 2. The results per viewing condition were aggregated ignoring order and the medians and means were calculated as measures of central tendency. The median and mean results are generally similar for the two conditions.

The Wilcoxon Signed Rank test was chosen to analyse the results, as it makes no assumption of normality or equal variance in the distributions. Related Samples Wilcoxon Signed Rank Tests were thus performed on responses to the six statements. A significant difference between the two viewing conditions was only found for the first statement concerning presence. HMD viewers rated higher agreement with this statement than CAVE viewers, suggesting the sense of presence was higher for the HMD condition. This may be due to the fact that in the CAVE condition the video was not completely 360 degrees around the viewer and due to the anomalies affecting the illusion shown in Figure 2.

S4 also related to presence, and while the P value was lower than for the other statements apart from S1, it was not significant.

S2 queried how much viewers felt they could interact with the environment, as interaction is limited to changing viewpoint, it is not surprising that both conditions scored low. S3 queried how much viewers paid attention to the display, again this was a little higher for the HMD than for the CAVE condition but not a significant difference.

S5 asked how much viewers enjoyed the video, and for this group it was generally rated positively with means and medians above three. In order to look at the format of the presentation, S6 asked whether viewers enjoyed the way they had viewed the content, and again, they gave generally positive responses with means and medians values above three for both CAVE and HMD conditions.

DISCUSSION

Drawing Comparisons

Clearly, a comparison of CAVE and HMD is only relevant for content that may be shown in a CAVE missing the ceiling and back wall. However, casual sampling of 360 content shows that many videos do not fully use 360 most of the time. Furthermore, it is a consideration that making people look around a lot could be tiring. Given the position of the sofa, and the fact that the CAVE system had no projection on the ceiling or back wall, viewers could view the remaining walls and floor by looking around, without the need to rotate the seating. However the restriction in the freedom to move around and to sit comfortably possibly caused biases which are not part of the systems compared.

That presence is an overriding feature of panoramic video is evident in user discussion around that concept, and it is also largely supported by responses to the questionnaire. It is interesting that subjects tended to cite presence and immersive factors as reasons for their preference of viewing condition, regardless of which it was.

There are some obvious differences between the two viewing conditions, such as in peripheral vision and sense of embodiment, as noted in the introduction. Thus, it was not surprising when these terms came up during the interviews, and that they eventually emerged as themes. However, even though obvious differences came up, it was still difficult to predict which of the two viewing conditions the users would prefer.

The discriminating themes found in this study have identified issues which could be addressed to improve the user experience of watching monoscopic panoramic video within a CAVE-like environment. Careful consideration should be paid to the video environment mapping to the physical space i.e. the walls of the CAVE, and care in technical post-production i.e. stitching.

There were some references to both embodiment and disembodiment when viewing the video in both conditions. The fact that the viewer has no reference to themselves within the video environment, when watching using the HMD, was pleasant to some users and off-putting to others. In respect to viewing panoramic videos using the HMD, there may be advantages to representing the viewer as part of the video in order to ease the negative sense of disembodiment.

A range of differences between viewing conditions are thus evident in the reports of participants. In the remainder of this section, a framework is presented that aims to make sense of the differences, organising themes into clusters, based on how they appear to influence the users' experience of panoramic video.

Towards a Framework

The thematic analysis research method adopted here sought, through the careful examination and coding of interview data, to arrive at an understanding of the key concepts at work when people draw comparisons between the two viewing platforms. In common with methods like Grounded Theory, the current research aims to make connections between themes produced, looking for common patterns, aiming to identify which, of all the issues raised, are the key ones – analogous to the Grounded Theory notion of a *core category* or *core variable* that can help to explain the majority of the variability observed in the data.

Looking at the discussions of themes, the concepts that appear most central are those to do with presence and immersion – the words and the theme crop up in much of what people say (both in the analysis of themes and in the participants explanations of their preference for one system over another). Many of the other themes described above are not explicitly about presence, but are still connected (for example, the comments relating to projection of image on self refer to the significance of such projections in helping to create or diminish peoples sense of being there, being involved, and so on.) Identification of Presence as a central theme is supported by the qualitative analysis, in which the Likert item S1, relating to peoples' sense of 'being there' – closely connected to presence, produced a significant difference between viewing platforms.

Having identified the central theme of Presence, the relationships between the remaining themes and Presence can be considered, producing clusters of themes that stand in a similar relation to the central theme of Presence. The emergent framework, illustrated in Figure 2, identifies clusters of themes that capture the ways that presence can be influenced, positively or negatively. Three clusters have been identified: Anomalies affecting the illusion, Affect and feeling, and Cognitive and perceptual effects.

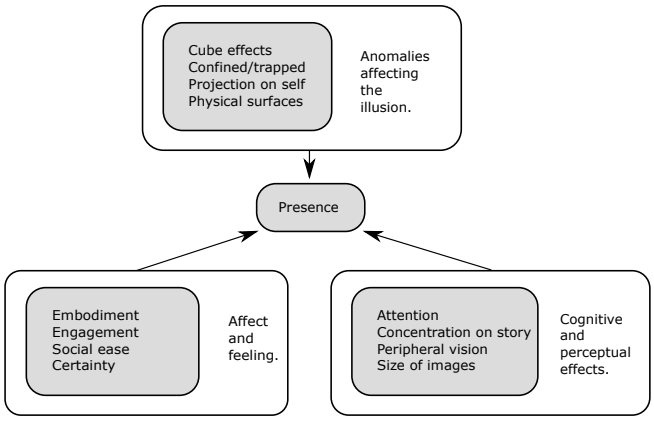

Figure 2. A framework for understanding user experience of panoramic video in a HMD and CAVE.

CONCLUSIONS

This study has added to the body of knowledge concerning the user experience of panoramic video by considering the differences between viewing in a headset and viewing in a CAVE-like environment. A set of themes that were previously applied to comparing HMD and phone and screen viewing were used to compare viewing in the HMD and CAVE conditions. However, consideration of these themes did not discriminate between viewing experiences for these two viewing conditions, and the user responses in relation to these themes was very similar. Consequently, a set of themes which do discriminate between the two was identified and elucidated. We also asked which viewing condition users preferred, and their opinion was split equally between the two options. Presence is a major feature of 360 video, compared to traditional film, and it appears that the users reported sense of presence was the deciding factor in choosing a preference. The emergent themes that summarise what participants reported when asked to describe various aspects of their experience of panoramic video on the two platforms, have been organised into a framework that puts Presence centre-stage, and aims to identify the clusters of sub-themes that influence the experience of being present in the panoramic video scene. Future work will develop this research in three directions:

- Development of the theoretical framework further, for instance, to better understand how the minor themes influence Presence.
- Exploration of how cinematic conventions are used in the production of panoramic video, and how viewers understand and orient towards directorial elements. Such an exploration will allow us to develop an understanding of how a filmic literacy is emerging for this novel medium.
- Identification of consequences of the findings for the production of panoramic video content. This could take the form of 'design advice' or guidelines, or could lead to an empirically informed evaluative critique of emerging guidance.

REFERENCES

1. Doug A Bowman, Ameya Datey, Young Sam Ryu, Umer Farooq, and Omar Vasnaik. 2002. Empirical Comparison of Human Behavior and Performance with Different Display Devices for Virtual Environments. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*. 2134–2138.
2. Virginia Braun and Victoria Clarke. 2006. Using thematic analysis in psychology. *Qualitative Research in Psychology* 3, 2 (jan 2006), 77–101. DOI: http://dx.doi.org/10.1191/1478088706qp063oa
3. Maxime Cordeil, Tim Dwyer, Karsten Klein, Bireswar Laha, Kim Marriott, and Bruce H. Thomas. 2017. Immersive Collaborative Analysis of Network Connectivity: CAVE-style or Head-Mounted Display? *IEEE Transactions on Visualization and Computer*

Graphics 23, 1 (jan 2017), 441–450. DOI: http://dx.doi.org/10.1109/TVCG.2016.2599107

4. Carolina Cruz-Neira, Daniel J. Sandin, Thomas A. DeFanti, Robert V. Kenyon, and John C. Hart. 1992. The CAVE: audio visual experience automatic virtual environment. *Commun. ACM* 35, 6 (jun 1992), 64–72. DOI:http://dx.doi.org/10.1145/129888.129892
5. Diana Fonseca and Martin Kraus. 2016. A comparison of head-mounted and hand-held displays for 360 degree videos with focus on attitude and behavior change. In *Proceedings of the 20th International Academic Mindtrek Conference on - AcademicMindtrek '16*. ACM Press, New York, New York, USA, 287–296. DOI: http://dx.doi.org/10.1145/2994310.2994334
6. Barney G. Glaser. 1965. The Constant Comparative Method of Qualitative Analysis. *Social Problems* 12, 4 (apr 1965), 436–445. DOI: http://dx.doi.org/10.2307/798843
7. K. Gruchalla. Immersive well-path editing: investigating the added value of immersion. In *IEEE Virtual Reality 2004*. IEEE, 157–164. DOI: http://dx.doi.org/10.1109/VR.2004.1310069
8. David M Hoffman, Ahna R Girshick, Kurt Akeley, and Martin S Banks. 2008. Vergence-accommodation conflicts hinder visual performance and cause visual fatigue. *Journal of vision* 8, 3 (2008), 33.1–30. DOI: http://dx.doi.org/10.1167/8.3.33
9. Judith A. Holton. The Coding Process and Its Challenges. In *The SAGE Handbook of Grounded Theory*. SAGE Publications Ltd, 1 Oliver's Yard, 55 City Road, London England EC1Y 1SP United Kingdom, 265–289. DOI: http://dx.doi.org/10.4135/9781848607941.n13
10. Lytro Immerge. 2017. Lytro Immerge. (Jan 2017). Retrieved Janurary 9, 2017 from https://www.lytro.com/press/releases/
11. Intel. 2017. Intel Alloy. (Jan. 2017). Retrieved Janurary 9, 2017 from https://newsroom.intel.com/press-kits/project-alloy/#news-coverage
12. Brett Jones, Lior Shapira, Rajinder Sodhi, Michael Murdock, Ravish Mehra, Hrvoje Benko, Andrew Wilson, Eyal Ofek, Blair MacIntyre, and Nikunj Raghuvanshi. 2014. RoomAlive. In *Proceedings of the 27th annual ACM symposium on User interface software and technology - UIST '14*. ACM Press, New York, New York, USA, 637–644. DOI: http://dx.doi.org/10.1145/2642918.2647383
13. Brett R. Jones, Hrvoje Benko, Eyal Ofek, and Andrew D. Wilson. 2013. IllumiRoom. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI '13*. ACM Press, New York, New York, USA, 869–878. DOI: http://dx.doi.org/10.1145/2470654.2466112
14. M. Carmen Juan and David Pérez. 2009. Comparison of the Levels of Presence and Anxiety in an Acrophobic Environment Viewed via HMD or CAVE. *Presence: Teleoperators and Virtual Environments* 18, 3 (jun 2009), 232–248. DOI:http://dx.doi.org/10.1162/pres.18.3.232
15. Kwanguk Kim, M. Zachary Rosenthal, David Zielinski, and Rachel Brady. 2012. Comparison of desktop, head mounted display, and six wall fully immersive systems using a stressful task. In *2012 IEEE Virtual Reality (VR)*. IEEE, 143–144. DOI: http://dx.doi.org/10.1109/VR.2012.6180922
16. Andrew Macquarrie and Andrew Steed. 2017. Cinematic Virtual Reality Evaluating The Effect Of Display Type On The Viewing Experience For Panoramic Video. In *IEEE VR 2017: the 24th IEEE Virtual Reality Conference.*
17. Microsoft. 2017. Microsoft Hololens. (Jan 2017). Retrieved Janurary 12, 2017 from https://www.microsoft.com/microsoft-hololens//
18. Peter Mills, Alia Sheikh, Graham Thomas, and Paul Debenham. 2011. Surround video.. In *Proceedings of the 2011 Networked and Electronic Media Summit*. 55–63.
19. Peter J Passmore, Maxine Glancy, Adam Philpot, Amelia Roscoe, Andrew Wood, and Bob Fields. 2016. Effects of Viewing Condition on User Experience of Panoramic Video. In *ICAT-EGVE 2016 - International Conference on Artificial Reality and Telexistence and Eurographics Symposium on Virtual Environments*, Dirk Reiners, Daisuke Iwai, and Frank Steinicke (Eds.). The Eurographics Association. DOI: http://dx.doi.org/10.2312/egve.20161428
20. Nozon Presenz. 2017. Nozon Presenze. (Jan 2017). Retrieved Janurary 9, 2017 from http://www.nozon.com/presenz
21. Razer. 2017. Project Ariana. (Jan 2017). Retrieved Janurary 9, 2017 from http://www.razerzone.com/project-ariana
22. Samsung. 2017. Samsung sold 5 million Gear VR headsets. (Jan 2017). Retrieved Janurary 9, 2017 from http://www.roadtovr.com/samsung-sold-5-million-gear-vr-headsets/
23. Ralph Schroeder, Anthony Steed, Ann-Sofie Axelsson, Ilona Heldal, Åsa Abelin, Josef Wideström, Alexander Nilsson, and Mel Slater. 2001. Collaborating in networked immersive spaces: as good as being there together? *Computers & Graphics* 25, 5 (oct 2001), 781–788. DOI: http://dx.doi.org/10.1016/S0097-8493(01)00120-0
24. P Seuntiens, I Vogels, and A Van Keersop. 2007. Visual experience of 3D-TV with pixelated ambilight.. In *Proceedings of PRESENCE 2007*.
25. Sarah Sharples, Sue Cobb, Amanda Moody, and John R. Wilson. 2008. Virtual reality induced symptoms and effects (VRISE): Comparison of head mounted display (HMD), desktop and projection display systems. *Displays* 29, 2 (mar 2008), 58–69. DOI: http://dx.doi.org/10.1016/j.displa.2007.09.005

26. Alistair Sutcliffe, Brian Gault, Terence Fernando, and Kevin Tan. 2006. Investigating interaction in CAVE virtual environments. *ACM Transactions on Computer-Human Interaction* 13, 2 (jun 2006), 235–267. DOI:http://dx.doi.org/10.1145/1165734.1165738

27. New York Times. 2017. The Daily 360. (Jan 2017). Retrieved Janurary 14, 2017 from https://www.nytimes.com/video/the-daily-360