Published in Proceedings of the 2019 ACM International Conference on Interactive Experiences for TV and Online Video (TVX '19). Salford (Manchester), United Kingdom — June 05 - 07, 2019. New York, NY, USA: ACM, 2019. ISBN:978-1-4503-6017-3.
https://doi.org/10.1145/3317697.3323351

What Are Others Looking at? Exploring 360° Videos on HMDs with Visual Cues about Other Viewers

Ville Mäkelä*
LMU Munich
ville.maekelae@ifi.lmu.de

Tuuli Keskinen
Tampere University
tuuli.keskinen@tuni.fi

John Mäkelä
Tampere University
john.makela@tuni.fi

Pekka Kallioniemi
Tampere University
pekka.kallioniemi@tuni.fi

Jussi Karhu
Tampere University
jussi.karhu@tuni.fi

Kimmo Ronkainen
Tampere University
kimmo.ronkainen@tuni.fi

Alisa Burova
Tampere University
alissa.burova@tuni.fi

Jaakko Hakulinen
Tampere University
jaakko.hakulinen@tuni.fi

Markku Turunen
Tampere University
markku.turunen@tuni.fi

ABSTRACT

Viewing 360° videos on a head-mounted display (HMD) can be an immersive experience. However, viewers must often be guided, as the freedom to rotate the view may make them miss things. We explore a unique, automatic approach to this problem with dynamic guidance methods called **social indicators**. They use the viewers' gaze data to recognize popular areas in 360° videos, which are then visualized to subsequent viewers. We developed and evaluated two different social indicators in a 30-participant user study. Although the indicators show great potential in subtly guiding users and improving the experience, finding the balance between guidance and self-exploration is vital. Also, users had varying interest towards indicators that represented a larger audience but reported a clear desire to use the indicators with their friends. We also present guidelines for providing dynamic guidance for 360° videos.

CCS CONCEPTS

• **Human-centered computing → Virtual reality**;

KEYWORDS

360 Videos; Omnidirectional Videos; Head-Mounted Displays; Guidance; Attention; Gaze; Virtual Environments; Virtual Reality

Figure 1: *Social indicators* provide visual cues about where other viewers have looked at in 360° videos. *Top:* the Shadow Indicator darkens the edge of the screen when a popular area of focus is detected in the opposite direction. *Bottom:* the Bubble Indicator displays a bubble that follows the most popular area of focus. The size of the bubble changes depending on the spread of the viewers' focus.

1 INTRODUCTION

Virtual reality (VR) has received much attention from the research community in the recent years. A popular application area in VR is the viewing of 360° videos, also known as omnidirectional videos (ODVs in short) [13, 17, 28]. These videos are often viewed using a head-mounted display (HMD), which enables looking around naturally by turning one's head. This form of viewing can provide an immersive and enjoyable experience, and lead to a high sense of presence [1, 3, 7, 17, 21, 24, 35, 36].

Current research commonly identifies the need to guide 360° video viewers, as enabling viewers to choose their own viewpoint may result in them feeling lost or missing things [8, 11, 23, 26, 27, 32, 35, 39]. While several methods to guide viewers have been presented, they have focused on limited scenarios and remained small in scale.

In this paper, we contribute to this research problem by introducing **social indicators**: dynamic guidance methods that automatically visualize areas in the 360° environment that other viewers commonly looked at (Figure 1). Social indicators collect gaze data from viewers, which are used to recognize popular areas in the content. These areas are then visualized to other viewers according to certain thresholds. We therefore differentiate our work from existing research by 1) basing our guidance on authentic viewers, and 2) providing a guidance method that can be automatically applied to any 360° content. This approach has the potential to solve several issues and presents several advantages.

First, existing research has largely worked with the assumption that there is prior knowledge of *where* the viewer should be guided, as is the case with narrative- and task-driven applications (and therefore, guidance can be designed and tailored for that specific experience). However, there are massive amounts of 360° content that have no inherent tasks, that viewers explore freely and that may contain multiple points of interest simultaneously. With such content, it is not necessarily clear what should be highlighted to the viewer, and to what extent do viewers want to be guided.

Second, our approach offers flexibility and scalability: at its core, it can be applied automatically to any kind of 360° content, used for both recorded and real-time viewing, and on a large scale. We predict that dynamic guidance will be particularly important in the near future, as 360° videos are becoming more popular and live streaming services for 360° content are on the way.

Third, using viewers' attention to produce guidance for subsequent viewers creates a variety of interesting implications. It is generally unclear how users perceive guidance that is based on other viewers, that is, what is popular among the larger audience. For example, are viewers curious about what was interesting to others? To what extent do viewers want to follow what was popular?

We developed two different social indicators, the *Shadow Indicator* and the *Bubble Indicator*. The Shadow Indicator darkens the edge of the viewport to nudge users towards the popular area (Figure 1, left). The Bubble Indicator displays the popular area more specifically with a bubble (Figure 1, right). In our study, the indicators were based on gaze data collected from 14 people.

We conducted a 30-participant user study where participants watched 360° videos with and without the social indicators. Our research questions were:

- **RQ1:** What are the general characteristics, strengths, and weaknesses of exploring 360° content on HMDs?
- **RQ2:** How do social indicators in 360° videos affect the viewing behavior and experience?
- **RQ3:** What characteristics must be accounted for when designing dynamic guidance methods?

The results suggest that the social indicators are a promising method for guiding users that can be applied to any 360° content. Of our two tested social indicators, the Bubble Indicator successfully guided users towards popular areas and helped them focus without dramatically affecting their viewing behavior. However, users also wanted to be able to explore the content on their own without any visual cues. Therefore, striking a balance between guidance and self-exploration is a critical goal in future virtual experiences. Moreover, participants had varying amounts of interest in what unknown viewers had viewed but expressed a clear desire to know what their friends would be looking at.

Our contribution in this paper is twofold. First, our work gives insight into the general viewing behavior and preferences of users when viewing 360° content. Second, we report the implementation and evaluation of two dynamic guidance methods: we recognize their strengths and weaknesses and assess how they change the viewers' behavior, and derive recommendations for providing successful dynamic guidance in the future. Our results are valuable for building future immersive systems utilizing 360° content.

2 RELATED WORK

Many methods for guiding viewers in virtual environments (VEs) have been presented. A common context for this is cinematic virtual reality (CVR), wherein users view immersive 360° stories using head-mounted displays [8, 23, 26, 27]. Understandably, in filmmaking (and storytelling) it is imperative that viewers are looking in the right direction during key parts of the story. Prior research has also found that viewers may experience fear of missing out (FOMO) when presented with the freedom to choose their own viewpoint [1, 3, 32, 35, 39], which is another argument for providing guidance to viewers.

Visual cues can be roughly divided into diegetic cues and non-diegetic cues [23, 27]. *Diegetic cues* refer to cues that are within the virtual environment, as a natural part of it. An example of this would be a character in the virtual environment pointing the viewer in the right direction. *Non-diegetic cues*, on the other hand, are separate from the environment, such as an arrow appearing on the heads-up display to point to the direction of the action.

Nielsen et al. [23] studied two guidance methods in a virtual environment: forced rotation, a non-diegetic cue where the user's virtual body would always face the action; and firefly, a diegetic cue where an animated 3D firefly would fly around the virtual environment directing the user's attention to the action. Although the differences between the two visualizations and a baseline condition (no guidance) were minimal, participants felt that the firefly helped them focus their attention better.

Similar to the forced rotation [23], other works have also investigated automatic panning [15, 19]. While this approach removes the need for the viewer to rotate the view, it also serves fundamentally different use cases, and is not suitable for free viewing of 360° content on HMDs. In addition to automatic panning, Lin et al. [19] evaluated an arrow that is displayed on the edge of the viewport to point in the right direction. Both guidance methods led to increased focus and higher engagement as opposed to a baseline without guidance.

An interesting method for informing the viewer of events outside their view was proposed by Lin et al. [20]. They overlaid screenshots of other parts of the 360° environment on the user's viewport, allowing them to see what was going on elsewhere in the scene.

Rothe et al. [26, 27] studied the use of lights, movement, and sounds to guide viewers. In their study, the three types of cues performed relatively comparably, but when put into pairs, moving objects with sound were the most effective in drawing the viewers' attention.

Grogorick et al. [9] optimized and evaluated subtle gaze guidance by smoothly altering the luminance of the region of interest. This guidance is removed once the gaze tracking indicates the user looks towards the target. If the indicated area is outside of the HMD viewport, their solution moves the pulsating area within the visible area to indicate the direction where to look.

In summary, current research has largely focused on predetermined and carefully planned visual cues, usually set by the designers. While such cues may work well in specific contexts, they must be designed and implemented for each experience individually. In this paper, we look into guidance methods that can be applied to any 360° content, even in real-time, by basing the guidance to what is (or has been) popular to other viewers.

We are motivated to explore viewer-based guidance because overall users are highly interested in shared VR experiences [12], and current research has worked extensively on various shared and collaborative contexts. Examples include viewing 360° videos together on separate tablets [33], sharing one's surroundings through a 360° camera [29, 34], allowing co-located HMD users and non-HMD users to interact with the virtual environment together [10], shared, embodied VR experiences [31], shared outdoor activities through a telepresence robot [14], and remotely located collaborative wayfinding activities [16]. Still, only a few existing works present visualizations of other viewers' attention in virtual environments. CollaVR [22] allowed two viewers to see the edges of each other's viewport in their HMD, and Sharedsphere [18] allowed a host user to share their view through a head-mounted camera, which was then embedded in the view of a guest VR user.

Existing research has proposed various technical solutions for detecting or predicting viewer attention in 360° images or videos [2, 5, 30, 38]. We use a similar approach by gathering gaze data from viewers and analyzing it to find popular areas in the videos. However, in this paper, we do not focus on the technical implementation (i.e., saliency detection), but on the user experience and implications of viewer-based guidance.

Because our aim is to explore dynamic guidance that is not dependent on the specific 360° content it is used with, certain considerations apply. Diegetic cues that blend into the environment may be problematic and not necessarily desirable, and hence we focus on non-diegetic cues. Another consideration is that visual cues representing other viewers' focus are moving constantly at various speeds. Hence, briefly appearing visual cues are unlikely to work, as the direction may change in an instant.

We further separate ourselves from existing work by focusing on free exploration scenarios, where users do not follow a predetermined narrative. Therefore, our aim is to provide guidance that is subtle and helps the user discover interesting things and make decisions on what to view. In contexts presented by existing work, such as storytelling, there is an underlying motivation for the viewer to be guided, but it is unclear how this works in free-form exploration scenarios that contain a lot of movement, such as people and traffic, and whether it has a positive impact on the experience.

Figure 2: The *Shadow Indicator*. When a popular area is detected outside the view, a shadow appears on the side of the screen opposite from the popular area. In this case, the shadow appears on the left, hinting that many viewers are focusing on something to the right. The Shadow Indicator is somewhat less obtrusive on the HMD than it appears in the figure.

In light of these considerations, we designed our dynamic guidance methods, which we will describe in detail in the following chapter.

3 SOCIAL INDICATORS

For our study, we designed and developed two different social indicators, the *Shadow Indicator* and the *Bubble Indicator*. In the following, we provide a rationale for their design, and describe both techniques as well as the details of their implementation.

Our common design goal for the social indicators was to provide users with a subtle guidance system that supports users in the exploration of 360° content, by offering them a conscious choice between following and ignoring the indicators. To strike this balance, we aimed to make them subtle, i.e., non-intrusive, but also clear enough so that the guidance is separate from the environment (non-diegetic). Therefore, conventional or otherwise highly visible interface elements, such as arrows or pictures on the display [19, 20], were not considered, and also, guidance methods with a more subconscious design [9] were not considered.

3.1 Shadow Indicator

The Shadow Indicator darkens the edge of the screen opposite from the direction in which the point of interest is (Figure 2), that is, the object(s) other users are looking at. In this manner, the Shadow Indicator aims to nudge users towards the other direction in a subtle manner, by affecting the viewer's peripheral vision. The indicator only considers horizontal movement, i.e., it appears either on the left or right side of the viewport, to not be too obtrusive and to maintain a clear function. The shadow's strength, i.e., breadth and transparency, is dependent on how far the popular area is. The shadow shrinks and goes transparent as the user nears the area. Fade animations are used to smoothen the appearance and disappearance of the indicator.

Figure 3: The *Bubble Indicator*. *Left*: a bubble appearing on the edge of the viewport hints that a large number of other viewers are focusing on something in the bubble's direction. *Right:* The bubble appearing in the viewport visualizes the area that other viewers focused on. The size of the bubble indicates the spread of focus points. In this case, viewers focused on a firetruck that was passing by. The Bubble Indicator is somewhat more visible on the HMD than it appears on a still image (due to its movement and changing size).

3.2 Bubble Indicator

The Bubble Indicator visualizes the area that others are looking at with a bubble, i.e., a white circle with no fill. When the target is outside the viewport, the bubble is displayed on the edge of the screen in the target's direction (Figure 3, left). The size of the bubble indicates the distance to the target and shrinks as the user gets closer. When the target comes into view, the bubble follows the target (Figure 3, right). In this phase, the bubble's size changes based on how spread the focus points of other viewers are. When a popular area is no longer detected, the bubble fades out.

The Bubble Indicator was partly based on existing work. Bubbles have been used in mobile environments to visualize the locations of off-screen objects [4], and they were found to be helpful in estimating the distance to the target. The Bubble Indicator works similarly in this respect when the object of interest is off-screen.

3.3 Base Data

For the social indicators to work, we collected a set of base data prior to the actual study. This base data would then be analyzed by the social indicators during the study to identify and visualize popular areas of focus.

To gather the base data, we recruited 14 participants (8 males, 6 females) to watch the video material that we would use in the user study. We used the HTC Vive headset with Tobii eye tracker. For each viewer, we calibrated the eye tracker, after which we advised them to watch the videos and look around as they pleased. During their viewing, we recorded their gaze data and head orientation. The equipment and setup were identical to the study, described in the following sections.

3.4 Implementation

3.4.1 Gaze Data Clustering. To derive the most popular areas of focus, the same methodology was used for both indicators. For performance reasons, to make our approach suitable for real-time use, the original data from the eye tracker (Rs=120Hz) was downsampled (without low-pass filtering) to 60Hz. For study purposes, however, we performed clustering in a pre-processing step to guarantee that the indicators performed exactly the same for all participants. To compensate for the slower response time caused by the pre-processing step, all base data was advanced 0.3 seconds for both indicators, so users would not perceive the indicators as lagging behind the events in the video. Downsampling was not performed when analyzing the final gaze data from the study participants.

We performed k-means clustering on the gaze direction vectors using a $1/10^{th}$ second time window, and with the number of clusters set to k=3 and cosine similarity as the distance metric. Using the value of $1/10^{th}$ seconds resulted in a sufficiently fast clustering for real-time use with the number of users in our base data, while still providing enough data to lessen the effects of individual sample errors. The number of clusters was chosen to increase stability, and this was coupled with a threshold value: the most popular cluster was required to have a size of 1.35 times the second most popular cluster for the indicator to be shown. We added this threshold to lessen the amount of spurious movement that would otherwise result from having multiple similarly sized clusters. A value of 1.35 was found to significantly reduce spurious movement while still showing the indicator for enough time in each video.

With the clustering and base data described above, a popular area of focus was detected for 68% of the time for the used video material. We hypothesized that this was a suitable value for our evaluation, as the indicators would be there for most of the time and therefore constantly offer viewers the choice between following the indicator and focusing on individual exploration, presumably resulting in better insight about their behavior and preferences.

3.4.2 Implementation of the Shadow Indicator. The gradient function for the Shadow Indicator was chosen to produce a sharp falloff to dark when fully in view ("curtain") and a smooth gradient or no visible gradient at all when the view direction is approaching the cluster centroid. The input to the gradient function was a function of the angle between the HMD forward direction and the target vector, and was smoothed using a low-pass filter.

3.4.3 Implementation of the Bubble Indicator. The Bubble Indicator was positioned at a z-distance of 2 meters from the viewer (as opposed to the video, which was 9 virtual meters from the camera). When the target was in view, the bubble was positioned at the target and size of the bubble was scaled linearly based on the number of gaze points in the target cluster. The size was clamped to a minimum of 0.54m and maximum of 1.8m in diameter. When the target was not in view, the Bubble Indicator was positioned on the left or right side of the HMD and the size was scaled linearly, depending on the angle between the target and the HMD forward vector. The position, size, and opacity of the bubble were further filtered using a low-pass filter. The filter for the position had a faster response time than the filters for the size and opacity.

4 STUDY

To study the effects of the social indicators and gain an in-depth understanding of user preferences during the viewing of 360° videos, we conducted a 30-participant user study. We followed a within-subjects design, where all participants experienced all three conditions: No Indicator, Shadow Indicator, and Bubble Indicator.

4.1 Video Material

We used three 3-minute 360° video clips in 6K resolution. All three videos consisted of three scenes, which were changed every minute. Hence, participants viewed a total of nine scenes lasting nine minutes altogether. The viewer's angle on the horizontal plane was reset at each scene change, so that every participant started exploring each scene from the same direction. The reset in direction was covered with a short fade-in effect during scene change to avoid disorienting the user.

All scenes were filmed in the streets of San Francisco and were similar in terms of content (examples in Figures 1–3). The scenes aimed to capture a steady flow of people and vehicles and the "buzz" of a city. This way, we aimed to ensure that there were multiple objects and events to focus on, without introducing any clear focus points that would take priority over everything else. The camera height was roughly 170 cm for each scene.

4.2 Participants

We recruited 30 participants (15 males, 15 females) through various channels. From the university, we recruited not only students but also administrative staff, and we also sent invitations to nearby companies, and asked them to spread the invitation further. On average, the participants were 32 years old (SD = 9.4), the minimum being 19, and the maximum 58. 11 of the participants were students from varying disciplines, such as computer science, medicine, and social sciences. The remaining 19 participants varied from construction workers to professionals in marketing and administration. All participants were rewarded with a movie ticket.

Eight participants were wearing glasses, of which two removed the glasses for the study. In addition, two participants reported wearing contact lenses. Participants were rather inexperienced regarding 360° videos and HMDs. 16 participants had no prior experience with 360° videos, 12 had watched 360° videos a few times at most, while only two reported having watched such videos several times. 11 participants had no prior experience with any VR technology, e.g., HMDs, and the remaining 19 participants had experienced them a few times at most.

4.3 Procedure

Participants first signed an informed consent and filled out a background questionnaire. Next, participants were handed brief instructions about the study procedure in their native language. Most importantly, participants were instructed that they will be watching a set of 360° videos with an HMD, but they do not have to complete any specific tasks during the viewing of the videos. Instead, they were encouraged to imagine that they were standing at the location waiting for a friend, and that they could look around to pass the time, for instance, by looking at interesting events or people around them. They were also instructed that prior to the actual videos, there will be a quick calibration and a practice video during which they can get accustomed to the HMD.

Participants then put on the HMD. Throughout the study, participants stood in the middle of the room when using the HMD. We used the HTC Vive with Tobii eye tracking. In the beginning, the software showed the participant's eyes and informed them how well the eyes were visible, displaying them in green when they were tracked reliably, and red when the tracking was unstable. Participants could then adjust the HMD so that it was placed firmly and comfortably, and the eyes were visible. Then, a quick calibration was run, wherein participants followed red dots on the screen with their eyes without moving their head. Calibration was done only once and lasted only around 5–10 seconds. However, every time participants put the HMD back on, it was checked with the calibration tool that the eyes were being tracked without problems.

Next, the practice video was shown, during which participants could look around and get a feel for viewing 360° videos with an HMD. The practice scene was from a park during winter with little to no movement in it. The video lasted for one minute.

Participants then watched three 3-minute videos described earlier. Each video was viewed once. One video was viewed using the Shadow Indicator, one using the Bubble Indicator, and one with no indicator. The order of the three conditions was counter-balanced. Before viewing a video with either indicator, corresponding instructions were handed to the participant on paper, to inform them of what the indicators do and how they work. It was made clear that participants are in no way required to follow the indicators, but rather, they can behave like they want, and make use of the indicators as little or as much as they like. The number of viewers that the indicators were based on (14) was not disclosed to the participants.

After experiencing all three conditions, participants filled out a questionnaire containing questions and statements about the indicators. The order of the question sets was dependent on the balancing of the conditions: participants first answered questions about whichever indicator they experienced first. Finally, participants were interviewed. Each session lasted approximately 45 minutes.

5 RESULTS

In the following, we first analyze the gaze data collected during the experiment. Then, we report the results from the questionnaires and interviews.

5.1 Gaze Data

For the gaze data, we had three primary interests. First, we assessed how much the HMD users' gaze differs from their head orientation in different conditions. Second, we evaluated whether the indicators made viewers look into the popular visualized areas, i.e., whether viewers followed the cues. Third, we measured the overall distance that the participants' head and gaze moved in each condition, to assess whether the social indicators affected this aspect of viewing behavior.

5.1.1 Gaze Versus Head Orientation. On average, the viewers' gaze was 14.8 degrees away from the head's orientation, i.e., the HMD's viewport center. This difference is significant and argues for the use of gaze data instead of only head orientation. Previous research has similar results: head movement may serve as an initial indicator of attention, but misses much of what can be uncovered with gaze data [6, 25].

For the conditions, the average values were 14.8 degrees for No Indicator, 14.5 degrees for Shadow Indicator, and 15.1 degrees for Bubble Indicator, and no statistically significant differences were found. Hence, the indicators did not result in increased discrepancy between gaze and head orientation. We find this somewhat surprising considering that the indicators often appeared on the edge of the viewport, which we assumed would lead to users' gaze targeting the edges more often, but this did not seem to be the case.

5.1.2 Attention Towards the Popular Areas. To assess whether the indicators affected viewing behavior, we measured the average distance (in degrees) of the viewers' gaze to the center of the popular focus area across all conditions. For this measurement, we only accounted for the time periods when a popular focus area was recognized, from the first moment when the indicator starts fading in and has not yet reached full opacity, until the moment when the indicator first starts to fade out. The spread of the indicator (the popular area) did not affect this metric. We also note that while the popular areas were not visualized in any way in the No Indicator condition, the popular areas were still known, and hence we included the No Indicator as a baseline for this measurement.

When no indicator was active, participants looked 81.7 degrees away from the popular areas on average. With the Shadow Indicator, the average distance was slightly shorter at 79.1 degrees, and with the Bubble Indicator noticeably shorter at 74.5 degrees. A repeated measures ANOVA revealed a significant main effect ($F(2, 58) = 4.134$, $p = 0.021$), and pairwise comparison with Bonferroni correction showed a statistically significant difference between the Bubble Indicator and the No Indicator conditions ($p = 0.016$). Hence, the Bubble Indicator was successful in guiding the viewers' attention towards the popular areas.

5.1.3 Viewing Activity. We measured how much the users' gaze and head moved in total (in degrees) in each condition (Table 1). Since gaze data can be jittery, we ran the analysis using both median-filtered and unfiltered data to be safe. Nonetheless, the amount of head and gaze movement in each condition was similar, as no statistical differences were found.

Since our previous analysis shows that participants successfully followed the Bubble Indicator, the lack of differences in viewing activity means that the indicators did not result in more "hectic" behavior. Rather, participants shifted their attention more towards the indicator instead of, e.g., going rapidly back and forth between the indicator and their own interests.

Table 1: Average distance moved in degrees per participant in each 3-minute 360° video (distances per minute in parentheses). Gaze[1] is unfiltered, Gaze[2] is median-filtered (n = 5).

	Head	Gaze[1]	Gaze[2]
No Indicator	5 341° (1 780°)	37 541° (12 514°)	15 974° (5 325°)
Shadow Indicator	5 219° (1 740°)	33 299° (11 100°)	14 828° (4 943°)
Bubble Indicator	5 184° (1 728°)	38 525° (12 842°)	15 951° (5 317°)

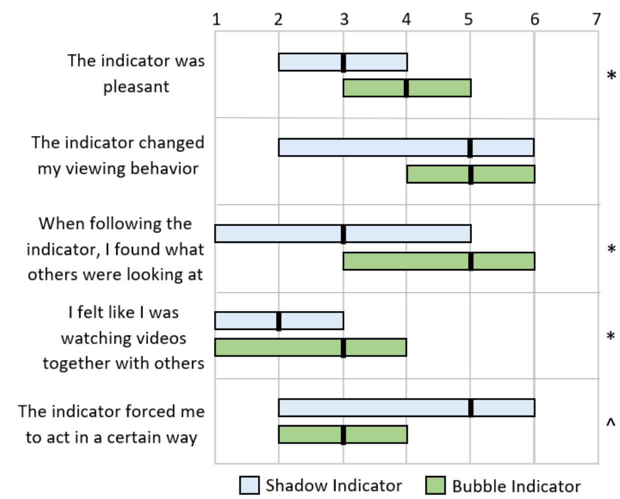

Figure 4: Statements about the indicators, where 1 = totally disagree, 4 = neither agree or disagree, and 7 = totally agree. The boxes represent inner quartiles and the thick lines represent medians. Statements marked with ∗ had a statistically significant difference between the indicators ($p < 0.05$), and the statement marked with ∧ had a near-significant difference ($p = 0.053$).

5.2 Subjective Results

Participants answered statements about the indicators on a 7-point Likert scale (Figure 4). In three of the five statements, a Wilcoxon Signed Rank test revealed a statistically significant difference between the two indicators ($p < 0.05$), and in one statement, a near-significant difference was found ($p = 0.053$).

First, participants found the Bubble Indicator (Mdn = 4) more pleasant than the Shadow Indicator (Mdn = 3), although neither was rated very favorably. Second, users somewhat agreed that both indicators had an effect on their viewing behavior (Mdn = 5). Third, the Bubble Indicator was expectedly more helpful in pointing out what others were looking at (Mdn = 5) compared to the Shadow Indicator (Mdn = 3) that only visualized the general direction. Fourth, neither indicator made users feel like they were watching videos together with other viewers. Although the difference between the indicators was significant, the more interesting result is that both were rated negatively in this regard (Mdn = 2–3). Finally, the Shadow Indicator was rated more forceful (Mdn = 5) than the Bubble Indicator (Mdn = 3) with a near-significant difference.

In addition, participants were asked to estimate how much they utilized the indicators from the total time that they were visible on an 11-point scale (from 0% to 100%). The Shadow Indicator and the Bubble Indicator were estimated to have been used 30% and 40% of the available time on average, respectively. No statistical differences were found.

5.3 Interview Results

In the interview, we asked questions about the participants' experiences and preferences related to viewing 360° videos in general, their general feelings towards the visualization of other viewers, as well as specific questions about the indicators. We conducted a lightweight inductive content analysis [37] with two researchers where applicable, to identify common themes among participants.

Most participants (19, 63%) did not know the city in which the videos were filmed and were unable to provide a reasonable guess. Six participants (20%) recognized the city, and five participants (17%) guessed it correctly. This general unfamiliarity with the viewed surroundings likely somewhat encouraged exploration. Eight participants reported that they tried to focus on things that would hint where they were. This goal-setting behavior was visible during the study as well: many participants were thinking out loud, reading street names and other signage and wondering where they are.

When asked whether participants are generally interested in what other viewers have looked at in 360° videos, 10 participants (33%) replied *yes*, nine replied *no* (30%), and the remaining 11 (37%) replied *depends* or *somewhat*. From the latter group, as many as nine participants provided an extremely similar reasoning about the balance between self-exploration and following the indicators. For instance, two participants explained: *"I'd like to explore by myself first, and then see what others are looking at."* (P27, Female/25). *"I was curious about others, but I also wanted to give room for myself for exploration."* (P16, Male/29).

When asked whether it would make a difference if the other viewers were people they knew, the majority of participants (24, 80%) strongly agreed, while three somewhat agreed, and three were unsure. Those who agreed provided various reasons for this, the most common being that they are generally interested in what their friends do and would therefore be interested in what they are looking at in a video. Many participants added that they often share the same interests with their friends, and therefore they would assume that an indicator generated by their friends will point to something interesting. Many also reported that the indicators could serve as a point of reference for later discussions about the videos among friends.

5.3.1 Shadow Indicator Feedback.

The Shadow Indicator received generally negative feedback. The most common reported issue was that it was too *forceful* (19), not allowing viewers to explore freely. 12 participants further explained that the shadow blocked parts of the screen, sometimes preventing them from looking where they wanted to. Four participants explicitly mentioned that upon following the Shadow Indicator, it was sometimes difficult to tell what exactly it was guiding the viewers to.

However, some participants also noted the potential of the Shadow Indicator in certain contexts: *"In situations where it's important to guide the viewer the shadow works."* (P8, Male/35). *"Works if there's a need to strongly guide the viewer."* (P18, Male/34).

5.3.2 Bubble Indicator Feedback.

The Bubble Indicator was generally well received by the participants. It was most commonly described as *precise* (10), *subtle* (10), and *clear* (9), and nine participants said that it helped them *focus* and *locate things* to look at. Three participants said that it was *fun*, and another three mentioned that it was *easy to follow*. Another set of three participants explained that it was easy to quickly glance at the bubble and then return to self-exploration. Two participants summarized their experiences with the Bubble Indicator: *"It gave me a choice because it was so subtle. It was easy to notice, but also easy to ignore if I wanted to explore on my own."* (P18, Male/34). *"It was subtle and guided me a little bit without being distracting. I also kind of liked following it."* (P10, Female/24).

However, similar to the Shadow Indicator, some participants (12) felt that the Bubble Indicator was also forceful, although this criticism was reported with a noticeably lighter tone. In the case of the Bubble Indicator, the complaint did not seem to come from the visualization itself, but rather, from the frequency of the indicator: *"If used less, the bubble would be nice. Now there was a bit too much of it."* (P9, Male/25).

Another reported issue (10) about the Bubble Indicator was that it sometimes jumped too quickly from one place to another. This was indeed the case in the occasional situation where the base data viewers' attention shifted quickly. It should be noted, however, that this happened equally with both indicators, but it seems that it was more apparent to viewers with the Bubble Indicator.

5.3.3 Individual Viewing Feedback.

Users appreciated the *freedom* (15), *lack of distractions* (11), and *peacefulness* (9) provided by individual viewing, i.e., the No Indicator condition: *"It was a more intimate experience because I was truly alone, and no one was pressuring me where to look."* (P17, Female/30). Four participants also mentioned that the experience was more immersive without the indicators, and three reported that without the indicators they were more motivated to explore the environment on their own.

We identified two common negative traits in individual viewing. First, participants reported a *lack of engagement* (12) without the indicators, mentioning things such as feeling lonely or bored, or lacking focus: *"I felt lonelier and more lost, even though I generally like being alone."* (P4, Female/47). The second negative trait was that without any guidance it is *easy to miss things* (10), as also identified by previous work [1, 32, 35, 39].

6 DISCUSSION

Our results provide insight into the experience and preferences of users viewing 360° videos on HMDs, and provide understanding on the use of social indicators as a form of dynamic guidance. In this discussion, we go through the three research questions we set in the beginning of the study and address them based on our results.

6.1 Characteristics of Standard Viewing of 360° Content on HMDs (RQ1)

Participants seemingly enjoyed the experience and immersed themselves when watching the 360° videos, which is in line with existing research [3, 7, 17, 21, 24, 35]. In standard viewing of 360° content (without indicators), a major positive aspect as reported by the participants is that it truly allows freedom to explore as one pleases, without distractions or pressure.

However, we identified two primary negative aspects. First, our study supports the existing notion that there is a need to guide viewers in situations where they can choose their own viewpoint [1, 32, 35, 39], as our participants also reported the fear of missing out (FOMO), that is, missing something interesting due to looking at something else at the time. Second, participants stated being generally less engaged, by reporting feelings of being lonely or lost, or having difficulty deciding which parts of the environments to explore. These results suggest that there is indeed demand for coming up with ways to improve the viewing experience of virtual environments.

We also provide measures for head and gaze movement during exploration of 360° videos. The viewers' head moved around 1 700–1 800 degrees per minute, whereas their gaze moved around 11 100–12 800 degrees per minute (4 900–5 300 degrees with median-filtering). Moreover, the viewers' gaze differed an average of 14.5–15 degrees from their head orientation. We also found that the visual cues did not significantly influence these measures. The findings are useful for future studies regarding the exploration of virtual environments, as they offer a set of measures for comparison and help in forming further hypotheses about viewer behavior. Similar measures have been provided by previous research [6, 25]; however they focused purely on viewing behavior without external cues.

6.2 Social Indicators in 360° Videos (RQ2)

Our goal with the social indicators was to not force users to behave in a certain way, but rather, to provide *support* for the exploration of 360° videos – help them discover interesting things, make decisions, and provide a more compelling experience overall. With this goal in mind, we discuss our results regarding the effect of the social indicators on the viewing behavior and experience.

One of our interesting findings is that the social indicators did not significantly affect the amount of head or gaze movement, nor did they lead to increased discrepancy between head and gaze orientation. Considering our goals about subtle guidance, we argue this is a positive finding, as users did not dramatically change their physical behavior. For instance, users did not resort to more "chaotic" behavior by, e.g., going rapidly back and forth between the indicators and their own interests, or by looking more towards the edges of the viewport during indicator use.

Of the two tested social indicators, the Bubble Indicator was successful in guiding users, as they focused on the popular areas more than in the other two conditions. On average the popular area with the Bubble Indicator was around 7.2 degrees closer than with no indicator. We argue that this number is significant when considering the situation: users did not follow the indicator all the time. Rather, they were shifting between following the indicator and exploring the environment on their own.

The Bubble Indicator also received positive feedback from the participants, being described as clear, precise, and subtle. Moreover, many participants reported that the Bubble indicator helped them focus and locate things, a finding also reported by Nielsen et al. [23] regarding their diegetic firefly cue. The Shadow Indicator, on the other hand, was unsuccessful. Based on the feedback, users did follow it (although they did not want to) but failed to ascertain which area the indicator was pointing to.

Both indicators were somewhat criticized, however, for being occasionally too forceful. With the Shadow Indicator this was much more evident, as users complained that it blocked the edge of the screen, partly preventing users from looking at what they wanted. The Bubble Indicator received some of the same criticism, albeit more subtly: they reported that the frequent appearance of the indicator made them feel like it was expecting them to follow it. This suggests that the design of the Bubble Indicator per se was successful – it simply appeared too frequently, which can be adjusted as discussed in the next subsection.

6.3 Considerations for Dynamic Guidance Methods (RQ3)

Based on our results and the discussion above, we formulate three recommendations for providing dynamically guided viewing experiences for 360° videos.

6.3.1 Balance Self-Exploration and Guidance. Users highly valued the freedom and calmness of self-exploration, but at the same time reported a lack of engagement and fear of missing things without the indicators. Both indicators were criticized for the fact that they were occasionally difficult to ignore, but the Bubble Indicator received mostly favorable feedback, and helped users to focus on popular areas. A significant number of users explicitly stated that while they were interested in the indicators and what others were looking at, they also wanted room for peaceful exploration of the environment without any additional distractions.

This mix of results strongly suggests that social indicators have great potential in improving the viewing experience, but enough time should be given to viewers to explore on their own without distractions. In our study, the indicators were visible for 68% of the time. We hypothesized that the indicators should appear frequently to offer users a choice between following the indicator and exploring on their own; however, their frequent presence made users feel somewhat pressured. Participants estimated having followed the indicators around 30–40% of the time they were visible.

Based on these results, we estimate that social indicators should not be visible for more than around 40–50% of the total runtime of the experience. This could be achieved, for instance, by introducing another threshold based on which the indicator would pause, e.g., after a long visualization. Another consideration is to give users time in the beginning to familiarize themselves with their surroundings before enabling the indicators.

6.3.2 Pre-Process the Base Data when Possible. To maintain an implementation that could equally be used for real-time visualization of other, simultaneous viewers, we did not pre-analyze the base data. We refrained from doing pre-analysis to achieve an all-purpose implementation that could also be used for, e.g., shared real-time experiences in the future. As a result, however, the indicators sometimes jumped quickly from one area to another or vanished very soon after appearing. This problem was also pointed out by the participants.

When real-time visualization is not needed, i.e., when the social indicators are used purely as a form of dynamic, supportive guidance, we recommend pre-processing the base data, through which many improvements can be achieved. Most importantly, the jumping issue can be solved completely by analyzing the data a couple of seconds in advance, as then very briefly appearing popular areas can be ignored. Moreover, pre-processing allows the visualization of popular areas slightly ahead of time, giving users more time to react. Other thresholds could also be added to ensure stability with more action-focused content by, e.g., discarding focus points that move too fast to be comfortably followed.

6.3.3 Consider Social Scenarios. Participants generally did not feel that they were watching 360° videos together with other people when using the indicators. This is not surprising considering that the videos were not watched simultaneously with others. Interestingly, though, some participants still reported feeling lonelier without the indicators.

Users had varying amounts of interest in the viewing behavior of other people, as 33% stated being interested, 30% stated being uninterested, and 37% stated being somewhat interested, or being interested with certain conditions, primarily regarding the balance between self-exploration and guidance as already discussed.

However, participants reported a considerably stronger desire to view 360° videos with social indicators if the indicators were based on their friends. Therefore, social indicators should be considered for social scenarios where, e.g., remotely located friends want to explore virtual environments together.

7 CONCLUSIONS AND FUTURE WORK

We explored **social indicators**: dynamic viewer-based guidance methods for viewing 360° videos on HMDs. Social indicators gather gaze data from viewers to identify popular areas in the videos, which are then visualized to subsequent viewers. Social indicators can be automatically applied to any 360° content. We designed two social indicators, the *Shadow Indicator* and the *Bubble Indicator*, and evaluated them in a 30-participant user study.

Our results can be summarized in three key findings. First, standard viewing of 360° content on HMDs is appreciated for its calm and distraction-free nature but criticized for lack of engagement and the potential for missing things, which calls for an investigation of subtle guidance methods. Second, social indicators can subtly guide users towards popular areas and help them focus without dramatically affecting the viewing behavior, but they can also be distracting if they are too frequent. Therefore, careful balancing of guidance and self-exploration is required to provide an optimal experience that brings out the strengths of both viewing modes. This balancing can be achieved by introducing a variety of thresholds that control the frequency of the indicators. Third, users exhibit varying amounts of interest towards indicators that are generated by a large, unknown audience, but have a clear desire to utilize social indicators with people they know.

Our work focused on free exploration scenarios, where the 360° content does not present a clear narrative with specific objects or events to focus on. For this, we chose to use relatively neutral content from the streets of a busy city. We argue this type of content is a good overall representative of content fitting for dynamic guidance, containing many points of interest and elements fundamental to a lot of 360° content, such as interesting scenery, traffic, and people moving all around the scene. Still, it would be interesting to study dynamic guidance with content that presents more unique experiences that are difficult to access otherwise, such as extreme sports or scenes with rare animals. In addition, in this study we focused on scenarios where other people had already watched the videos, and those viewers were not known to participants. Hence, dynamic guidance methods should be evaluated in social, real-time settings, such as with groups of remotely located friends. Finally, our study was conducted using head-mounted displays, which allow the viewport to be turned quickly and naturally. It is therefore unlikely that the presented designs for social indicators would perform similarly with other platforms, such as standard mobile devices. As such, dynamic guidance techniques could be a research topic for a variety of different platforms.

ACKNOWLEDGMENTS

This work was done as a part of *Immersive Media Disruption* (IMD) project, financed by Tekes — the Finnish Funding Agency for Innovation (later Business Finland). We thank all the study participants.

REFERENCES

[1] Tanja Aitamurto, Shuo Zhou, Sukolsak Sakshuwong, Jorge Saldivar, Yasamin Sadeghi, and Amy Tran. 2018. Sense of Presence, Attitude Change, Perspective-Taking and Usability in First-Person Split-Sphere 360° Video. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18)*. ACM, New York, NY, USA, Article 545, 12 pages. https://doi.org/10.1145/3173574.3174119

[2] M. Assens, X. Giro-i-Nieto, K. McGuinness, and N. E. O'Connor. 2017. SaltiNet: Scan-Path Prediction on 360 Degree Images Using Saliency Volumes. In *2017 IEEE International Conference on Computer Vision Workshops (ICCVW)*. 2331–2338. https://doi.org/10.1109/ICCVW.2017.275

[3] Marc Van den Broeck, Fahim Kawsar, and Johannes Schöning. 2017. It's All Around You: Exploring 360° Video Viewing Experiences on Mobile Devices. In *Proceedings of the 2017 ACM on Multimedia Conference (MM '17)*. ACM, New York, NY, USA, 762–768. https://doi.org/10.1145/3123266.3123347

[4] Stefano Burigat, Luca Chittaro, and Silvia Gabrielli. 2006. Visualizing Locations of Off-screen Objects on Mobile Devices: A Comparative Evaluation of Three Approaches. In *Proceedings of the 8th Conference on Human-computer Interaction with Mobile Devices and Services (MobileHCI '06)*. ACM, New York, NY, USA, 239–246. https://doi.org/10.1145/1152215.1152266

[5] Hsien-Tzu Cheng, Chun-Hung Chao, Jin-Dong Dong, Hao-Kai Wen, Tyng-Luh Liu, and Min Sun. 2018. Cube Padding for Weakly-Supervised Saliency Prediction in 360° Videos. *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition* (2018), 1420–1429.

[6] Erwan J. David, Jesús Gutiérrez, Antoine Coutrot, Matthieu Perreira Da Silva, and Patrick Le Callet. 2018. A Dataset of Head and Eye Movements for 360&Deg; Videos. In *Proceedings of the 9th ACM Multimedia Systems Conference (MMSys '18)*. ACM, New York, NY, USA, 432–437. https://doi.org/10.1145/3204949.3208139

[7] Diana Fonseca and Martin Kraus. 2016. A Comparison of Head-mounted and Hand-held Displays for 360° Videos with Focus on Attitude and Behavior Change. In *Proceedings of the 20th International Academic Mindtrek Conference (AcademicMindtrek '16)*. ACM, New York, NY, USA, 287–296. https://doi.org/10.1145/2994310.2994334

[8] Michael Gödde, Frank Gabler, Dirk Siegmund, and Andreas Braun. 2018. Cinematic Narration in VR–Rethinking Film Conventions for 360 Degrees. In *International Conference on Virtual, Augmented and Mixed Reality*. Springer, 184–201.

[9] Steve Grogorick, Michael Stengel, Elmar Eisemann, and Marcus Magnor. 2017. Subtle Gaze Guidance for Immersive Environments. In *Proceedings of the ACM Symposium on Applied Perception (SAP '17)*. ACM, New York, NY, USA, Article 4, 7 pages. https://doi.org/10.1145/3119881.3119890

[10] Jan Gugenheimer, Evgeny Stemasov, Julian Frommel, and Enrico Rukzio. 2017. ShareVR: Enabling Co-Located Experiences for Virtual Reality Between HMD and Non-HMD Users. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '17)*. ACM, New York, NY, USA, 4021–4033. https://doi.org/10.1145/3025453.3025683

[11] Jan Gugenheimer, Dennis Wolf, Gabriel Haas, Sebastian Krebs, and Enrico Rukzio. 2016. SwiVRChair: A Motorized Swivel Chair to Nudge Users' Orientation for 360 Degree Storytelling in Virtual Reality. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (CHI '16)*. ACM, New York, NY, USA, 1996–2000. https://doi.org/10.1145/2858036.2858040

[12] Simon Gunkel, Hans Stokking, Martin Prins, Omar Niamut, Ernestasia Siahaan, and Pablo Cesar. 2018. Experiencing Virtual Reality Together: Social VR Use Case Study. In *Proceedings of the 2018 ACM International Conference on Interactive Experiences for TV and Online Video (TVX '18)*. ACM, New York, NY, USA, 233–238. https://doi.org/10.1145/3210825.3213566

[13] Jaakko Hakulinen, Tuuli Keskinen, Ville Mäkelä, Santeri Saarinen, and Markku Turunen. 2018. Omnidirectional Video in Museums – Authentic, Immersive and Entertaining. In *Advances in Computer Entertainment Technology*, Adrian David Cheok, Masahiko Inami, and Teresa Romão (Eds.). Springer International Publishing, Cham, 567–587.

[14] Yasamin Heshmat, Brennan Jones, Xiaoxuan Xiong, Carman Neustaedter, Anthony Tang, Bernhard E. Riecke, and Lillian Yang. 2018. Geocaching with a Beam: Shared Outdoor Activities Through a Telepresence Robot with 360 Degree Viewing. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18)*. ACM, New York, NY, USA, Article 359, 13 pages. https://doi.org/10.1145/3173574.3173933

[15] H. Hu, Y. Lin, M. Liu, H. Cheng, Y. Chang, and M. Sun. 2017. Deep 360 Pilot: Learning a Deep Agent for Piloting through 360° Sports Videos. In *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 1396–1405. https://doi.org/10.1109/CVPR.2017.153

[16] Pekka Kallioniemi, Tuuli Keskinen, Jaakko Hakulinen, Markku Turunen, Jussi Karhu, and Kimmo Ronkainen. 2017. Effect of Gender on Immersion in Collaborative iODV Applications. In *Proceedings of the 16th International Conference on Mobile and Ubiquitous Multimedia (MUM '17)*. ACM, New York, NY, USA, 199–207. https://doi.org/10.1145/3152832.3152869

[17] Pekka Kallioniemi, Ville Mäkelä, Santeri Saarinen, Markku Turunen, York Winter, and Andrei Istudor. 2017. User Experience and Immersion of Interactive Omnidirectional Videos in CAVE Systems and Head-Mounted Displays. In *Human-Computer Interaction – INTERACT 2017*, Regina Bernhaupt, Girish Dalvi, Anirudha Joshi, Devanuj K. Balkrishan, Jacki O'Neill, and Marco Winckler (Eds.). Springer International Publishing, Cham, 299–318.

[18] Gun A. Lee, Theophilus Teo, Seungwon Kim, and Mark Billinghurst. 2017. Sharedsphere: MR Collaboration Through Shared Live Panorama. In *SIGGRAPH Asia 2017 Emerging Technologies (SA '17)*. ACM, New York, NY, USA, Article 12, 2 pages. https://doi.org/10.1145/3132818.3132827

[19] Yen-Chen Lin, Yung-Ju Chang, Hou-Ning Hu, Hsien-Tzu Cheng, Chi-Wen Huang, and Min Sun. 2017. Tell Me Where to Look: Investigating Ways for Assisting Focus in 360° Video. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '17)*. ACM, New York, NY, USA, 2535–2545. https://doi.org/10.1145/3025453.3025757

[20] Yung-Ta Lin, Yi-Chi Liao, Shan-Yuan Teng, Yi-Ju Chung, Liwei Chan, and Bing-Yu Chen. 2017. Outside-In: Visualizing Out-of-Sight Regions-of-Interest in a 360° Video Using Spatial Picture-in-Picture Previews. In *Proceedings of the 30th Annual ACM Symposium on User Interface Software and Technology (UIST '17)*. ACM, New York, NY, USA, 255–265. https://doi.org/10.1145/3126594.3126656

[21] A. MacQuarrie and A. Steed. 2017. Cinematic virtual reality: Evaluating the effect of display type on the viewing experience for panoramic video. In *2017 IEEE Virtual Reality (VR)*. 45–54. https://doi.org/10.1109/VR.2017.7892230

[22] Cuong Nguyen, Stephen DiVerdi, Aaron Hertzmann, and Feng Liu. 2017. CollaVR: Collaborative In-Headset Review for VR Video. In *Proceedings of the 30th Annual ACM Symposium on User Interface Software and Technology (UIST '17)*. ACM, New York, NY, USA, 267–277. https://doi.org/10.1145/3126594.3126659

[23] Lasse T. Nielsen, Matias B. Møller, Sune D. Hartmeyer, Troels C. M. Ljung, Niels C. Nilsson, Rolf Nordahl, and Stefania Serafin. 2016. Missing the Point: An Exploration of How to Guide Users' Attention During Cinematic Virtual Reality. In *Proceedings of the 22Nd ACM Conference on Virtual Reality Software and Technology (VRST '16)*. ACM, New York, NY, USA, 229–232. https://doi.org/10.1145/2993369.2993405

[24] P. J. Passmore, M. Glancy, A. Philpot, A. Roscoe, A. Wood, and B. Fields. 2016. Effects of Viewing Condition on User Experience of Panoramic Video. In *Proceedings of the 26th International Conference on Artificial Reality and Telexistence and the 21st Eurographics Symposium on Virtual Environments (ICAT-EGVE '16)*. Eurographics Association, Goslar Germany, Germany, 9–16. https://doi.org/10.2312/egve.20161428

[25] Yashas Rai, Jesús Gutiérrez, and Patrick Le Callet. 2017. A Dataset of Head and Eye Movements for 360 Degree Images. In *Proceedings of the 8th ACM on Multimedia Systems Conference (MMSys'17)*. ACM, New York, NY, USA, 205–210. https://doi.org/10.1145/3083187.3083218

[26] Sylvia Rothe and Heinrich Hußmann. 2018. Guiding the Viewer in Cinematic Virtual Reality by Diegetic Cues. In *Augmented Reality, Virtual Reality, and Computer Graphics*, Lucio Tommaso De Paolis and Patrick Bourdot (Eds.). Springer International Publishing, Cham, 101–117.

[27] Sylvia Rothe, Heinrich Hußmann, and Mathias Allary. 2017. Diegetic Cues for Guiding the Viewer in Cinematic Virtual Reality. In *Proceedings of the 23rd ACM Symposium on Virtual Reality Software and Technology (VRST '17)*. ACM, New York, NY, USA, Article 54, 2 pages. https://doi.org/10.1145/3139131.3143421

[28] Santeri Saarinen, Ville Mäkelä, Pekka Kallioniemi, Jaakko Hakulinen, and Markku Turunen. 2017. Guidelines for Designing Interactive Omnidirectional Video Applications. In *Human-Computer Interaction – INTERACT 2017*, Regina Bernhaupt, Girish Dalvi, Anirudha Joshi, Devanuj K. Balkrishan, Jacki O'Neill, and Marco Winckler (Eds.). Springer International Publishing, Cham, 263–272.

[29] Samarth Singhal and Carman Neustaedter. 2017. BeWithMe: An Immersive Telepresence System for Distance Separated Couples. In *Companion of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing (CSCW '17 Companion)*. ACM, New York, NY, USA, 307–310. https://doi.org/10.1145/3022198.3026310

[30] V. Sitzmann, A. Serrano, A. Pavel, M. Agrawala, D. Gutierrez, B. Masia, and G. Wetzstein. 2018. Saliency in VR: How Do People Explore Virtual Environments? *IEEE Transactions on Visualization and Computer Graphics* 24, 4 (April 2018), 1633–1642. https://doi.org/10.1109/TVCG.2018.2793599

[31] Misha Sra, Aske Mottelson, and Pattie Maes. 2018. Your Place and Mine: Designing a Shared VR Experience for Remotely Located Users. In *Proceedings of the 2018 Designing Interactive Systems Conference (DIS '18)*. ACM, New York, NY, USA, 85–97. https://doi.org/10.1145/3196709.3196788

[32] Hannah Syrett, Licia Calvi, and Marnix van Gisbergen. 2017. The Oculus Rift Film Experience: A Case Study on Understanding Films in a Head Mounted Display. In *Intelligent Technologies for Interactive Entertainment*, Ronald Poppe, John-Jules Meyer, Remco Veltkamp, and Mehdi Dastani (Eds.). Springer International Publishing, Cham, 197–208.

[33] Anthony Tang and Omid Fakourfar. 2017. Watching 360° Videos Together. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '17)*. ACM, New York, NY, USA, 4501–4506. https://doi.org/10.1145/3025453.3025519

[34] Anthony Tang, Omid Fakourfar, Carman Neustaedter, and Scott Bateman. 2017. Collaboration with 360° Videochat: Challenges and Opportunities. In *Proceedings of the 2017 Conference on Designing Interactive Systems (DIS '17)*. ACM, New York, NY, USA, 1327–1339. https://doi.org/10.1145/3064663.3064707

[35] Audrey Tse, Charlene Jennett, Joanne Moore, Zillah Watson, Jacob Rigby, and Anna L. Cox. 2017. Was I There?: Impact of Platform and Headphones on 360 Video Immersion. In *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems (CHI EA '17)*. ACM, New York, NY, USA, 2967–2974. https://doi.org/10.1145/3027063.3053225

[36] Mirjam Vosmeer and Ben Schouten. 2017. Project Orpheus A Research Study into 360° Cinematic VR. In *Proceedings of the 2017 ACM International Conference on Interactive Experiences for TV and Online Video (TVX '17)*. ACM, New York, NY, USA, 85–90. https://doi.org/10.1145/3077548.3077559

[37] Yan Zhang and Barbara M Wildemuth. 2016. Qualitative analysis of content. *Applications of social research methods to questions in information and library science* 318 (2016).

[38] Ziheng Zhang, Yanyu Xu, Jingyi Yu, and Shenghua Gao. 2018. Saliency Detection in 360° Videos. In *The European Conference on Computer Vision (ECCV)*.

[39] Goranka Zoric, Louise Barkhuus, Arvid Engström, and Elin Önnevall. 2013. 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. https://doi.org/10.1145/2465958.2465959