



Article

# Guidance in Cinematic Virtual Reality-Taxonomy, Research Status and Challenges

Sylvia Rothe <sup>\*</sup>, Daniel Buschek and Heinrich Hußmann

Media Informatics, LMU Munich, 80337 Munich, Germany; daniel.buschek@ifi.lmu.de (D.B.); hussmann@ifi.lmu.de (H.H.)

\* Correspondence: sylvia.rothe@ifi.lmu.de

Received: 31 January 2019; Accepted: 14 March 2019; Published: 19 March 2019



**Abstract:** In Cinematic Virtual Reality (CVR), the viewer of an omnidirectional movie can freely choose the viewing direction when watching a movie. Therefore, traditional techniques in filmmaking for guiding the viewers' attention cannot be adapted directly to CVR. Practices such as panning or changing the frame are no longer defined by the filmmaker; rather it is the viewer who decides where to look. In some stories, it is necessary to show certain details to the viewer, which should not be missed. At the same time, the freedom of the viewer to look around in the scene should not be destroyed. Therefore, techniques are needed which guide the attention of the spectator to visual information in the scene. Attention guiding also has the potential to improve the general viewing experience, since viewers will be less afraid to miss something when watching an omnidirectional movie where attention-guiding techniques have been applied. In recent years, there has been a lot of research about attention guiding in images, movies, virtual reality, augmented reality and also in CVR. We classify these methods and offer a taxonomy for attention-guiding methods. Discussing the different characteristics, we elaborate the advantages and disadvantages, give recommendations for use cases and apply the taxonomy to several examples of guiding methods.

**Keywords:** Cinematic Virtual Reality; 360° video; attention; gaze; guiding; subtle gaze direction; off-screen indicators; forced guiding; stylistic rendering

## 1. Introduction

Omnidirectional movies (360° movies) are attracting widespread interest and have many possible applications, e.g., telling stories about exciting experiences and locations in the world, or documenting places of historic interest. Even though the term 360° video is widespread, it does not accurately reflect this media. On the horizontal level, there are indeed 360° to explore, however, in full-surround videos, there are also vertical angular extents of  $\pm 90^\circ$  to be observed. Therefore, we use the term omnidirectional video, which is often used in scientific literature [1,2].

In Cinematic Virtual Reality (CVR), the viewer watches omnidirectional movies using head-mounted displays (HMD) or other Virtual Reality (VR) devices. Thus, the viewer can feel immersed within the scenes and can freely choose the viewing direction. It is possible that important details are outside the viewer's field of view. For some CVR experiences, this proves to be unproblematic—no additional guiding is necessary: The user discovers a storyline as constructed by the author [3]. In other story constructs, it can be important not to miss some elements. Guiding can prevent the user from becoming lost or confused. Often viewers express a fear of missing out (FOMO) because they do not know where to look [4,5]. In such cases, they wish to be guided in an unobtrusive way, for relaxing enjoyment without the fear of missing something.

The purposes of movies cover a wide range: entertainment, art, education, marketing or even instructions. How much guidance is needed depends to a large extent on the movie content. In some

cases, guiding is advisable for the continuity of the story, for interaction cues, subtitles, education information, and for social viewing applications. Cinematic VR is not a clear lean back media [6] and several key aspects motivate guiding the viewer: Choosing the frame by looking around is a very natural way of interaction which can be enhanced by other interaction possibilities, such as interactive scene changes. Drawing attention to interactive cues can be supported by guiding methods. Another motivation for guiding arises by subtitling. Since the viewer can freely choose the viewing direction, it is difficult to identify the speaker belonging to the subtitle. Here, the viewer can be guided towards the speaker [7]. Furthermore, VR can be used for education [8–11], for example in museums or classrooms. Since guiding techniques can increase the recall rate [12,13], such methods can support the learning process. Additionally, suitable guiding methods are needed if teachers, museum guides or students want to draw attention to something. Since viewers can feel isolated watching movies via HMD, techniques are needed to support social awareness and communication. For example, guiding techniques can visualise a region of interest or the own viewport to the co-watchers [14].

Even if viewers often do not notice it, filmmakers direct the gaze and attention of the viewer to relevant aspects in the movies. Cinematic tools such as sounds, lights and movements redirect the attention of the viewer. Studies have shown that the pattern of gaze fixations is often consistent across viewers of Hollywood-style movies [15,16]. In Hollywood-style movies, filmmakers use narrative and editing techniques to strongly guide the viewers to important aspects of a scene, often at the expense of more peripheral details [15]. Since the viewer has more freedom in CVR, most of these methods are less effective.

Images in traditional movies are framed, and in the frame the filmmaker arranges elements for the story. Investigations using eye trackers showed that viewers seldom explore the periphery of the movie image [17] if there are no subtitles. People are likely to look at the central area of a frame or screen. In reality, such a frame does not exist, in Cinematic VR, the position of the frame is determined by the viewer.

To understand how gaze and attention can be guided, we inspected several models from other research fields, such as psychology and biology. We explain the terms used in these models in Chapter 2. This knowledge is important for exploring guiding methods for the field of CVR.

In the last few years, several approaches for guiding in Cinematic VR have been published [18–22]. Since Cinematic VR is a relatively new field of research and since it is very close to virtual and augmented reality, we also looked into concepts of these fields, as well as methods for audio–visual content on flat screens (TV, monitor) and mobile devices. There are several techniques used in other areas which are adaptable to CVR. In Chapter 3, we give an overview of published work.

Each of the techniques is focused on one or two attributes of the guiding techniques. More research is expected in the next years, and this needs consistent terms for discussing these techniques. Clarifying the concepts is also helpful for finding new approaches. To discuss which attribute of a technique was relevant for the success or failure of a method, a single overarching terminology is required. With our taxonomy in Chapter 4, we contribute structure and clarification to work on guiding in CVR.

Applying this taxonomy on known guiding methods we distinguish between 2D and 3D media. Methods used in traditional filmmaking or for images can be applied in CVR to guide the viewer in the current field of view as described in Chapter 5. VR and Augmented Reality (AR), as well as CVR, have additional needs for guiding the viewer since objects can be outside of the screen. These guiding methods are described in Chapter 6.

At the end of this work, we discuss how the introduced taxonomy provides support for the design process of guidance in Cinematic VR. The taxonomy fosters understanding of the various attributes of guiding techniques, to find new methods and support filmmakers to select the right methods for their projects.

## 2. Terms and Insights from Various Research Fields

Researchers of different areas, such as psychology, biology and computer sciences, are working on topics about attention and gaze directing [23]. Basic knowledge of these areas is necessary to understand guiding methods.

### 2.1. Attention Theory

There are several factors responsible for where someone is looking. On the one side, there are bottom-up factors that characterize the scene. Bottom-up factors are stimuli which attract attention due to their properties such as color or shape and are normative. Methods are normative if they are working in the same way for all people, unless a person has a specific condition such as color blindness. On the other side, there are top-down factors such as task or goal. The performance of such factors can vary between individuals. Depending on the goal, the attention can be space-based (position of an object), feature-based (features of an object) or object-based [23,24].

Movie parts can be explored in a bottom-up (stimulus-driven) or top-down (task-driven) manner. Viewers can be guided by staging and compositional techniques by using lights, colors, and focal depth. Especially the bottom-up process is responsible for the fact that viewers often do not perceive the cuts in a movie. The task, following the story, causes edit blindness [25]. This effect could be also useful for some of the guiding methods.

Cues are able to direct the attention to a target. They can have various properties and positions. Posner [26] showed that viewers detect a target faster if the cue is a feature of the target (e.g., a colored border) than a cue positioned not on the target (e.g., an arrow). He introduced the terms exogenous and endogenous. Exogenous cues are stimulus-driven and work automatically, for example, a flash that attracts attention. They cause an unintentional orientation. Such cues are positioned on the target and can also be auditive or haptic [27]. They are working in a bottom-up manner. Since the reaction to such cues is reflexive, they act fast. However, if there is no interesting target cue, the attention is transient. Endogenous cues are goal-driven and work voluntarily [27]. Often, they are based on a sign that tells where to look or to listen and require first an interpretation, e.g., an arrow. Even if goal-driven attention works slower, it enhances the processing of the event [26] and can be sustained at a location for longer periods. Yarbus [28] showed that eye movements depend on the task. In his experiment, participants watched the same scene after having been asked different questions. The eye movements differed significantly.

One of the most influential models of human visual attention, the feature integration theory, was developed in 1980 by Treisman and Gelade [29]. It explains the role of visual attention for object recognition. Perceiving a stimulus, features are registered early, automatically and in parallel. Objects are identified later in a separate process. In the first step, called the pre-attentive stage, parts of the brain automatically gather information about features such as color and shape. During the second step, the focused attention state, the whole object is perceived by combining the individual features.

Exogenous and pre-attentive processes are mostly memory-free, subtle and caused by cues at the target (direct cues). In contrast, endogenous and attentive processes are memory-bound, overt and caused by indirect cues, which have to be interpreted (e.g. an arrow). Direct cues can be outstanding features of an object. Healey et al. [30] published a list of two-dimensional outstanding features complemented by literature which describes tasks using these features. Some examples relevant for CVR are: color, size, curvature, line orientation, intensity, flicker, direction of motion, lighting direction, and intersection. Wolfe and Horwitz [31,32] also described target attributes which can efficiently guide attention: color, orientation, size, depth, motion, and luminance.

The above-mentioned terms are all relevant for characterizing the process of drawing attention. They describe related but different aspects of attention directing and are not orthogonal to each other. Table 1 gives an overview without claim of completeness. For our taxonomy, we chose the cue property, since it best complements our other extracted dimensions considering their use in informing practical design choices.

**Table 1.** Attributes of attention-guiding techniques used in literature.

process	bottom-up	top-down
attention	exogenous	endogenous
impulse	stimulus-driven	goal-driven
automatism	automatically/reflexive	voluntary
attentiveness	pre-attentive	attentive
cognition	memory-free	memory-bound
awareness	subtle	overt
cues	direct cues	indirect cues (symbolic)

Knowing about work in psychology on explaining attention directing is fundamental to understanding the various guiding methods in HCI and filmmaking. Depending on the movie genre, other guiding methods can be suitable.

## 2.2. Basics about Physiology of the Eyes

For studying gaze directing, knowledge about eye physiology is essential. Signals such as colors or flickers are perceived differently depending on the eye region. In the periphery, rod cells are located, which are responsible for seeing in the darkness and are very sensitive for illumination and motion. This means guiding cues at the periphery could be flickering lights or moving elements. The cone cells in the fovea are needed for seeing colors during the day. They are less responsive to light. Colors could be used for drawing the attention to an object, over which the viewer lets wander the gaze.

The diverse characteristics of periphery and fovea are the reason for the differing perception of flickers depending on the viewing direction. The critical flicker fusion frequency (CFF) is the rate, at which the flickering fuses and is perceived as continuous. The CFF is about 22 to 25 Hz for low lights (rod cells). For higher light intensity (cone cells) the CFF increases to the logarithm of the light intensity (Ferry-Porter law) [33] and increases up to 80 Hz depending on the area of the light intensity (Granit-Harper law) [34]—the larger the flickering stimulus, the higher the CFF.

This is the reason for different critical flickering fusion rates (CFF) in different regions of the eye. Thus, a high-frequent flicker can be visible in the periphery but fused in the fovea. Additionally, the temporal resolution acuity increases for larger flickers ([35]). This property can be used for subtle gaze direction. However, the CFF is a very sensitive attribute and depends not only on the regions of the eye but also on [34,35]:

- the frequency of modulation
- the region on the retina perceiving the modulation
- the amplitude/depth of the modulation
- the luminance (Ferry-Porter law) [33]
- the size of the modulation (Granit-Harper law) [34]
- color
- contrast
- the person (age, fatigue)

CFF can be used for creating stimuli cues and developing guiding methods. Some research has been done which used the CFF for gaze directing in images [36]. However, such methods cannot be easily adapted to movies, since alteration in the movie image can make the flickering stimuli ineffective and the threshold has to be increased, making the method no longer subtle [13]. Even if it is difficult to consider all the mentioned parameters at once and to find CFF thresholds valid for each person, the knowledge about this behaviour has to be taken into account when analysing and developing guiding methods. Flickers can not only be designed as an artificial part of the method, but they can also be included in the movie (e.g., flickering lights).

### 3. Guiding Methods in Literature

There is a lot of research about gaze guiding in images, traditional movies, VR, AR, and CVR. Table 2 gives an overview of the methods and the environments in which they were tested. We inspected not only the methods evaluated in CVR environments since techniques of other fields could be adaptable to CVR, even if it needs closer inspection and adjustments. Some of them we have already evaluated in previous work, such as diegetic methods [22] and Subtle Gaze Directing (SGD) [13]. It depends on the type of movie, whether guiding methods are needed and how strong or subtle the guiding should be. More obvious methods perform better but decrease the experience [19]. Since gaze guiding can increase the recall rate of target objects [12,13] also, the aim of the movie is relevant for finding the most suitable technique. A guiding method can be important in an educational CVR application, but disturbing in a meditative movie.

**Table 2.** Overview of guiding methods from research projects and the environments for which they were tested. The last column indicates the name of the method in the paper. In some of the studies, eye-tracking (ET) was used additionally.

Project/Literature	Environment	Display	Method	Name
(Rothe and Hußmann, 2018)[22]	CVR	HMD	diegetic	
(Brown et al., 2016)[18]	CVR		diegetic	
(Nielsen et al., 2016)[21]	CVR	HMD	diegetic forced	
(Y.-C. Lin et al., 2017)[20]	CVR	HMD	forced sign	Autopilot Arrow
(Gugenheimer et al., 2016)[37]	CVR	HMD	forced	SwiVRChair
(Danieau et al., 2017)[19]	CVR	HMD	effects	Desaturation Fading
(Chang et al., 2018)[38]	CVR	HMD	haptic	FacePush
(Sassatelli et al., 2018)[39]	CVR	HMD	forced	Snap-Changes
(Gruenefeld et al., 2018b)[40]	CVR	HMD, LEDs	off - screen	RadialLight
(Y.-T. Lin et al., 2017)[41]	omnidirectional video	mobile device	off-screen	Outside-In
(Cole et al., 2006)[42]	3D models	monitor, ET	modulation	stylized rendering
(Tanaka et al., 2015)[43]	omnidirectional images	mobile device	angular shift	
(Mendez et al., 2010)[44]	images	monitor	SMT	
(Veas et al., 2011)[45]	video	monitor	SMT	SMT
(Hoffmann et al., 2008)[46]	desktop windows	large screen	on-screen	frame, beam, splash
(Renner and Pfeiffer, 2018)[47]	AR	HoloLens	arrow flicker	SWave. 3D-path
(Perea et al., 2017)[48]	AR	mobile device	off-screen	Halo3D
(Gruenefeld et al., 2018a)[49]	AR, VR	HMD	off -screen	HaloVR WedgeVR
(Gruenefeld et al., 2017b)[50]	AR	HMD	off -screen	EyeSee360
(Bork et al., 2018)[51]	AR	HMD	off - screen	Mirror Ball sidebARs u.a.
(Siu and Herskovic, 2013)[52]	AR	mobile device	off - screen	SidebARs

Table 2. Cont.

Project/Literature	Environment	Display	Method	Name
(Renner and Pfeiffer, 2017a)[53]	AR	HMD	screen-referenced word-referenced	sWave, arrow, flicker
(Burigat et al., 2006)[54]	maps	mobile device	off - screen	Halo, arrows
(Henze and Boll, 2010)[55]	AR	mobile device	off -screen	Magic Lens Peephole
(Schinke et al., 2010)[56]	AR	mobile device	off -screen	Mini-map 3d arrows
(Koskinen et al., 2017)[57]	VR	HMD, LEDs	off - screen	LEDs
(Zellweger et al., 2003)[58]	desktop windows	monitor	off -screen	CityLights
(Bailey et al., 2012)[12]	images	monitor	subtle	SGD
(Bailey et al., 2009)[59]	images	monitor	subtle	SGD
(McNamara et al., 2008a)[60]	images	monitor, ET	subtle	SGD
(Grogorick et al., 2017)[61]	VR	HMD, ET	subtle	SGD
(McNamara et al., 2012)[62]	images	monitor, ET	subtle	Art
(Wei-quan Lu et al., 2014)[63]	AR, video	monitor	subtle	subtle cues, visual clutter
(Waldin et al., 2017)[36]	images	High-frequent monitor	subtle	SGD
(Smith and Tadmor, 2013)[64]	images	monitor	blur	
(Hata et al., 2016)[65]	images	monitor	subtle, blur	
(Hagiwara et al., 2011)[66]	images	monitor	saliency editing	
(Kosek et al., 2017)[67]	lightfield video	HMD	visual, auditive, haptic	IRIDIUM+
(Kaul and Rohs, 2017)[68]	VR/AR	HMD	haptic	HapticHead
(Rantala et al., 2017)[69]	image	monitor	haptic	Headband
(Stratmann et al., 2018)[70]	cyber-physical systems	monitors	haptic	Vibrotactile Peripheral
(Knierim et al., 2017)[71]	VR	HMD	haptic	Tactile Drones
(Sridharan et al., 2015)[72]	VR	HMD, ET	subtle modulation	
(Kim and Varshney, 2006)[73]	image	monitor	saliency adjust	
(Jarodzka et al., 2013)[74]	video	monitor	blur	EMME
(Lintu and Carbonell, 2009)[75]	images	monitor	blur	
(Khan et al., 2005)[76]	images	large display	highlighting	spotlight
(Barth et al., 2006)[77]	video	GCD, ET	saliency adjust	
(Dorr et al., 2008)[78]	video	GCD, ET	saliency adjust	
(Sato et al., 2016)[79]	video	monitor	subtle diegetic	saliency robot gaze
(Vig et al., 2011)[80]	video	GCD	saliency adjust	
(Bocca et al., 2007)[81]	AR	AR-HMD	non-diegetic	funnel
(Sukan et al., 2014)[82]	AR	AR-HMD	non-diegetic	ParaFrustum
(Kosara et al., 2002)[83]	text, images	monitor, ET	blur, depth-of-field	
(Mateescu and Bajić, 2014)[84]	images	monitor	color manipulation	
(Delamare et al., 2017)[85]	AR	HMD	gaze gestures	

“GCD” = “Gaze-Contingent Display”; “ET” = “eye-tracking”.



Some of the most important guiding methods of different research fields we inspect deeper:

### 3.1. Diegetic Methods

Some research in recent years has focused on diegetic methods for guiding the viewer in CVR [18, 21, 22, 86]. Diegetic cues are part of the movie, for example moving characters, lights or sounds. The concept of diegesis in film theory was developed by Souriaux [87] and afterwards adapted to other fields, e.g., literary theory. Diegetic elements belong to the narrative world. The term diegetic is well-known in film theory and is mostly used for music and other sounds. Diegetic music in a movie is part of the story. It can be heard not only by the viewer (like film music) but also by the characters. Examples are: music from a radio in the movie or music played by musicians which are movie characters. For guiding the viewers' attention, the authors can use diegetic cues which are included in the story world: moving protagonists, lights or sounds.

On the other side, there are non-diegetic cues which are not part of the story, such as arrows, focus assistant tools or forced rotations methods [88, 89].

### 3.2. Saliency Modulation Technique (SMT)

Mendez et al. [44] described a Saliency Modulation Technique (SMT) for directing the viewer's attention to a target object. For analysing the original images, saliency maps were used and the material was modulated depending on the results of the analysis. A saliency map shows the saliency values on each region in the image [90, 91]. Thus, it was possible to apply minimal changes. This method works on video in real time. However, it can only be used if the target is in the viewer's field of view. Additionally, the method is less strong in environments with moving and blinking objects. Veas et al. [45] investigated SMT regarding modulation awareness, attention, and memory. They showed that SMT can shift the attention to selected targets without the viewer noticing the modulation. Moreover, SMT can increase the recall rate.

There are promising approaches in other fields of research. Gaze-Contingent Displays (GCD) reduce the resolution of peripheral areas to decrease the amount of data [92, 93]. The region where the user is looking is determined by eye tracking and shown in high-resolution quality. Since saliency modulation techniques can guide the gaze [1, 45, 66, 77, 80, 94], the techniques could be transferable to Cinematic VR. However, saliency modulation can only be effective if the modulated region is in the field of view (FoV) of the viewer.

### 3.3. Blurring

Another method is using blurred regions in the images. Smith et al. [64] investigated blurred and non-blurred regions for guiding the viewer's attention. They showed that the viewer tends towards regions with little or no spatial blur if the rest of the image is more blurred. This approach is very similar to methods used in traditional filmmaking. Hata et al. [65] extended this method for visual guidance with unnoticed blur effects for images. A threshold was found at which the viewer notices the blur and could be guided below it.

### 3.4. Stylistic Rendering

Guiding the viewer in a movie can also be done by stylistic elements: depth of field, colors, brightness, and sharpness. Cole et al. [42] investigated gaze direction in 3D models with stylized focus. They used local variations in shading effects (color saturation and contrast) and line qualities (texture and density) for drawing the viewer's gaze to the emphasized area. Additionally, they applied a dynamic technique: stylized focus pull. Focus pull is a creative camera technique in traditional filmmaking where the focus changes during the shot and so the attention switches from one area to another. In digital editing, focus-pulling can also be added in the post-production by an animated filter effect.

### 3.5. Subtle Gaze Direction (SGD) with Eye Tracking

Subtle gaze direction can guide the gaze without the viewer noticing it. The concept of subtle gaze direction (SGD) was first presented by Bailey [59]. The core of this concept is to modulate the target region if it is in the peripheral area for inducing the viewer to look there and to stop the modulation when the viewer is watching in this direction. In this way, the viewer's gaze can be guided without perceiving the modulation. In the research of Bailey et al. [59], two options were investigated: luminance modulation and warm-cool modulation, both with a rate of 10 Hz in a circular region of approximately 1-cm diameter. An eye tracker was used to observe the viewer's gaze and the modulation stopped when the viewer changes the view in the direction of the target region. It was shown that this technique can effectively guide the user's gaze in still images without the user noticing the modulation.

In the experiments of McNamara et al. [60] modulation of the luminance was more effective than warm-cool modulation. They showed that SGD improves the performance for search tasks in images without the participants noticing the modulation. The same method was investigated for guiding in narrative art, with static images [62]. Grogorick et al. extended this method to virtual environments [61]. A luminance modulation was used and a circle shape was dynamically adapted to ellipses for the wide FoV in VR. Additionally, the stimulus was dynamically positioned, so the method could be used also for targets, which are not in the FoV at the beginning of the stimulation. The experiments showed that results of search tasks can be improved for hidden objects.

### 3.6. Subtle Gaze Direction (SGD) with High Frequent Flickers

There is some research on how SGD can be used without an eye tracker. Waldin et al. [36] took advantage of the fact that the peripheral vision is more sensitive to highly frequent flickering than the foveal vision. Therefore, the critical fusion frequency (CFF) is different in these areas. A signal flickering in the periphery is no longer flickering when the viewer is looking at it and the signal is in the fovea. If the viewer is looking in the direction of the flicker, the flicker fuses to a stable image. In this way, no eye-tracking is necessary for stopping the modulation. The experiments used flickers of 60 Hz and 72 Hz, so a display of 120 Hz and 142 Hz was needed. In the first experiment, images with cycles were used, in the second experiment a highly complex image. The method worked effectively in both cases. It seems to be necessary to execute a personal calibration routine to find the size and luminance of the flicker modulation. At the moment, this method cannot be adapted to VR and CVR since the frequency of the HMD displays (90 Hz) are not high enough. Additionally, flickers are less effective in environments with dynamic changes. As mentioned in Chapter 2, the CFF does not only depend on the region of the eye region, and the thresholds are difficult to find.

### 3.7. Off-screen Indicators (Halo, Edge)

Since the viewer in CVR observes only an extract of the film image via HMD, the above methods are not always effective. Depending on the viewing direction, cues can be missed since they are not in the viewer's FoV. Therefore, methods are needed to indicate targets beyond the screen.

One way of visualizing off-screen objects on flat displays is the halo technique [95], where off-screen objects are surrounded by circles the size of which is sufficient to be visible at the edge of the display. From the curvature of the circle, the user can infer the position of the object. The halo method is not directly transferable to CVR because the CVR screen is a sphere. To ensure that a circle is still visible on the edge of the display, the center must not be more than 90° away. For points outside this region, for example on the opposite side of the gaze cursor, the circle cannot be made visible in the display. EdgeRadar [96] and Wedge [97] are modifications of this technique to avoid overloading and overlapping. These techniques were adapted to mobile AR [48] and for HMDs [49]. Gruenefeld et al. [98] compared several off-screen object visualization techniques (Arrow, Halo and Wedge) for out-of-view objects in Augmented Reality. In their experiments, the halo and wedge



technique performed best. However, the implemented methods were limited to a 90° area in front of the user and need further adaptation to 360°. EyeSee360 [50] is a visualization technique for out-of-view objects in Augmented Reality which could be adaptable for CVR.

### 3.8. Forced Rotation of the user (SwiVRChair)

Gugenheimer et al. [37] developed a chair which automatically rotates the viewer to look at predefined regions of interest. In their experiment, simulator sickness was very low. This may be caused by the fact that the viewer was turned around and so the rotation in the virtual world matched with the rotation in the real world. Additionally, the participants needed lower head movements for enjoying the VR experience in a more “lean back” way.

### 3.9. Forced Rotation of the VR world

Another possibility of forced guiding is to rotate the scene in a way that the region of interest (RoI) is in the field of view of the viewer. Nielsen et al. [21] compared forced rotation with diegetic guiding. In their experiments, the diegetic method was more helpful and caused higher presence. Lin et al. [20] compared forced rotation (called autopilot) with an arrow which points to the direction of the RoI (called visual guidance). The results depended on the type of movie, but no generally higher sickness was observed for the forced rotation. One reason for simulator sickness is the discrepancy between movements in the real and virtual world [99,100], and so rotating the VR world in front of the user often provokes sickness. There is no consistent opinion if rotating a scene causes simulator sickness or not [20].

### 3.10. Forced Rotation via Cutting

In traditional filmmaking, cutting can be used to show important details to the viewer. After the cut, the RoI is displayed. The same can be done in CVR: Independent of the viewing direction, the viewer will see the RoI after the cut [39,101]. However, it needs to be investigated if this can cause disorientation in case both scenes are in the same location and the viewing direction changes with the cut—similar to the crossing the line problem [102] in traditional movies.

### 3.11. Haptic Cues

Kaul et al. [68] developed HapticHead for guidance in virtual and augmented reality. Chang et al. [38] introduced FacePush, a system for haptic signals using HMDs. Their system generates forces on the face of the viewer and was tested for two VR experiences (boxing, diving) and for CVR guiding. In contrast to HapticHead or a vibrotactile headband [69], FacePush indicates the advised direction of rotation (left/right) and not the absolute direction of the RoI. For integrating haptic cues in a story as diegetic cues, it needs haptic stimulus on other parts of the body beyond the head. Drones can provide such haptic stimuli [103,104].

Summarizing all these methods, we found several properties of guiding techniques investigated in the literature: subtle, off-screen, forced, diegetic, haptic, and some others. Not all of them are comparable to each other, since they highlight different aspects of the guiding method. It is important to classify these attributes for finding the most relevant and qualified characteristics for guiding methods in CVR. We will do this in the next chapter.

## 4. Taxonomy

To find the appropriate techniques for guiding in CVR, we inspected methods for various media: images, movies, virtual and augmented reality (Chapter 3). Inspired by the large amount of papers about guiding methods and several taxonomies in Virtual Reality [21,105,106], we analysed these methods and classified their properties, also taking into account that they might be combined across papers: For example, even if one paper emphasised the subtleness of a method, that method might

also potentially address visual/auditive/haptic senses in future work, or might be investigated for on- or off-screen targets. Even if a paper emphasizes the subtleness of a method, the method can be additionally visual/auditive/haptic and on-screen/off-screen. In this process, we took into account if a dimension is needed in CVR.

With our classification, we found seven orthogonal dimensions. Nielsen et al. [21] described three dimensions for attention guiding. One of our dimensions (diegesis) is consistent, two others correspond to their taxonomy (directness and freedom).

Our taxonomy describes the most important attributes which we discovered in the literature and which are relevant in our own work without claim of completeness. It is conceivable that in the future new components should be added depending on the focus of research. Table 3 presents our taxonomy of important dimensions. They will be explained in the following subsections.

**Table 3.** The table shows the different dimensions of guiding methods and possible values.

Dimension/Property	Option 1	Option 2	Option 3
Diegesis	diegetic	non-diegetic	
Senses	visual	auditive	haptic
Target	on-screen	off-screen	
Reference	world-referenced	screen-referenced	
Directness	direct	indirect	
Awareness	subtle	overt	
Freedom	forced by system	forced by reflex	voluntary

#### 4.1. Diegetic and Non-Diegetic

Research results show that diegetic methods perform well in Cinematic VR [18,21,22]. For visual diegetic methods, the cue has to be in the field of view, e.g., movements, light, colors. In most cases, the location of the cue (e.g., the color of the target) will be identical to that of the target. One exception is: a protagonist looks or points into a certain direction.

However, one can imagine story parts where no suitable cues in the story world exist. If it is nevertheless necessary to guide the attention to a detail, non-diegetic methods can be applied. Depending on the use case, the method either has to be designed to be noticed easily or to avoid disturbance. Some advantages and disadvantages of diegetic and non-diegetic methods are listed in Table 4.

**Table 4.** Advantages and disadvantages of diegetic and non-diegetic methods. The third row indicates for which requirements the methods are suitable.

	Diegetic	Non-Diegetic
+	high presence and enjoyment [21,22]	easily usable and noticeable
-	depends on the story, not usable for all cases [18,22]	can disrupt the VR experience [21]
->	suitable for visual experiences and wide story structures	suitable for important information

#### 4.2. Visual, Auditive and Haptic

It is obvious to discuss visual methods for attention-guiding in CVR. Movements, lights and characters are well known for drawing attention in traditional movies [107]. However, these cues can only be used if they are in the field of view. If the viewer is looking in another direction, the cues will not be discovered. For motivating the user to change the viewing direction, sound coming from the direction of the Point of Interest (PoI) is a considerable method, since it can be used out of the field of view. Even if the source of a sound is not visible, it is possible to hear it—including the direction of the sound. In real life, a source of noise can get someone to change the viewing direction. The same is true for CVR [22]. Also, haptic cues can cause this behavior and it is worthwhile to discuss it as

a guiding method. Some advantages and disadvantages of visual, auditive and haptic methods are listed in Table 5.

**Table 5.** Advantages and disadvantages of visual, auditive and haptic methods. The third row indicates for which requirements the methods are suitable.

	Visual	Auditive	Haptic
+	can be easily integrated	works also for off-screen RoIs	novel experience
-	not always visible, depending on the viewing direction	difficult to distinguish between diegetic and non-diegetic	difficult to realize
->	suitable for visual experiences and wide story structures	suitable for changing the viewing direction	needs additional devices suitable for public application

#### 4.3. On- and Off-screen

Depending on the viewing direction, a PoI can be in the FoV of the viewer or outside of it. To guide the attention to an object on the screen, methods can be discussed which are already investigated for images or traditional movies. We call this on-screen guiding. However, in CVR it can happen that the viewer first has to change the viewing direction for seeing the PoI on the screen. For this case, off-screen methods are needed.

Which of both methods is used does not depend on the author, rather on the viewing direction. The author has to decide if both are needed. Visual methods such as saliency modulation of the ROI can only work if the region is in the FoV. If the viewer should not miss it, an off-screen technique has to be added.

#### 4.4. World- and Screen-referenced

Cues in VR can be differentiated between screen-referenced and world-referenced indicators [108,109]. Screen-referenced items are connected to the display and move along with it in case the viewer is turning the head. World-referenced items are connected to the virtual world, in our case to the movie. They stay fixed at their place in the movie world, even if the viewer turns the head. The term “screen-referenced” corresponds to the notion “in-view” used in augmented reality and “world-referenced” matches “in-situ” (e.g., Reference [110]).

Even if diegetic cues are world-referenced, the opposite is not true. A cue added on top of the movie for guiding the viewer, which cannot be seen by the movie characters, is non-diegetic. Screen-referenced cues are always non-diegetic since they cannot be part of the story world (movie). They are well suited for menus. Some advantages and disadvantages of world-referenced and screen-referenced methods are listed in Table 6.

**Table 6.** Advantages and disadvantages of world-referenced and screen-referenced methods. The third row indicates for which requirements the methods are suitable.

	World-Referenced	Screen-Referenced
+	integrated in VR world, higher presence	always visible
-	not always visible	disrupt the VR experience
->	suitable for experiences where presence is important	suitable for off-screen guiding and important information

#### 4.5. Direct and Indirect Cues

There are two main types of cues: direct and indirect cues [111]. Direct cues are at the target, e.g., outlines, colors or lights. Indirect cues are based on symbolic information and have first to be interpreted, for example an arrow. The cues do not have to be visual. For example, a sudden bang can work as an auditive direct cue and a voice, that says what can be seen, as an indirect cue.

Direct cues work mostly stimulus-driven and are based on the characteristics of the scene (exogenous, memory-free), e.g., an abrupt light or sudden movements and working in a bottom-up manner. For that, regions of interest have to be sufficiently different from the surroundings. Direct cues act fast, transient and spontaneous [24,112].

Indirect cues involve a conscious effort (endogenous, memory-bound), e.g., interpreting a sign. They work in a top-down manner by cognitive properties such as knowledge, expectations and tasks. Indirect methods are slow, sustained and voluntary [24,112,113]. Some advantages and disadvantages of direct and indirect cues are listed in Table 7.

**Table 7.** Advantages and disadvantages of world-referenced and screen-referenced methods. The third row indicates for which requirements the methods are suitable.

	Direct Cues	Indirect Cues
+	fast	sustainable
-	transient, not always visible	must be interpreted
->	suitable for on-screen guiding	suitable for recallable RoIs

#### 4.6. Subtle and Overt

In case the user is not aware of a method, the method is called subtle. In contrast to this, overt techniques will be noticed by the user [106]. There are several subtle guiding techniques, which are based on the physiology of the eye, and the term Subtle Gaze Guiding (SGD) is already established for these methods. However, the term subtle is not used consistently in the literature. The term subliminal is also common for stimuli, below the threshold for conscious perception [65]. As already mentioned, such thresholds (e.g., CFF) depend usually on several factors and vary between people. Thus, a cue can be subliminal for one person, but supraliminal for another.

Subtleness can also be achieved otherwise. Examples are diegetic methods, where elements of the movie guide the gaze. The user notices the cue but is normally not aware of the guiding property. Even if subtlety of techniques can be defined as a continuum, we agree with Suma et al. [106] in choosing a dichotomical categorization (subtle vs overt), whereby subliminal is included in subtle.

Some advantages and disadvantages of subtle and overt methods are listed in Table 8.

**Table 8.** Advantages and disadvantages of subtle and overt methods. The third row indicates for which requirements the methods are suitable.

	Subtle	Overt
+	no disruption [59]	easily noticeable [106] can increase recall rates
-	not always effective [13]	can be disrupting
->	suitable for wide story structures	suitable for learning task [12,13]

#### 4.7. Forced by System, Forced by Reflex and Voluntary

Most of the discussed methods are voluntary: Viewers can freely decide if they follow any guiding cues or if they explore the scene on their own. However, also forced methods can be applied [20,21,37,70]. There are different ways of forced guiding. On the one hand, the viewer can be rotated, as in SwiVRChair [37]. This has the advantage that the viewer can feel the rotary motion. On the other hand, the VR-world/movie can be rotated [20]. These methods force the user to change the viewing direction in a technical way. This can also be done by using the methods based on the physiological models described in Chapter 2. Stimuli can provoke the viewer to change the viewing direction reflexively in a fast way. Some advantages and disadvantages of forced and voluntary methods are listed in Table 9.

**Table 9.** Advantages and disadvantages of voluntary and forced methods. The third row indicates for which requirements the methods are suitable.

	Forced by System	Forced by Reflex	Voluntary
+	RoI always shown	fast, can be integrated in the story	remains the freedom of viewing direction
-	can disrupt the VR experience [21]	not always usable	RoIs can be missed
->	suitable for important and fast direction changes	suitable for visual experiences and fast reactions	suitable for visual experiences and wide story structures (world narratives)

#### 4.8. Usage of the Taxonomy for CVR Guiding Methods from Literature

The previous section described our identified dimensions. Now, Table 10 shows methods from literature for guiding in CVR and their attributes in the introduced taxonomy. It shows that less subtle methods have been studied so far. We could find only one haptic guiding method for CVR likely due to the state of technology. Most literature about guiding in CVR is concentrated on off-screen guiding since this is one of the challenges of this medium. It is expected that methods from traditional movies work if the RoI is in the field of view. However, it needs some effort to find the best way to implement them in CVR.

**Table 10.** Guiding methods evaluated for CVR and the attributes of the taxonomy.

Literature	Diegesis	Senses	Target	Reference	Directness	Awareness	Freedom	Name
(Rothe and Hußmann, 2018)[22]	diegetic	visual	on-scr.	world	dir.	subtle	voluntary	
(Brown et al., 2016)[18]	diegetic	audio	off-scr.	world	dir.	subtle	voluntary	
(S. Rothe et al., 2018)[13]	non-dieg.	visual	on-scr.	world	indir.	subtle	voluntary	SGD
(Nielsen et al., 2016)[21]	diegetic	visual	off-scr.	world	dir.	subtle	voluntary	
	non-dieg.	/	off-scr.	screen	dir.	overt	forced sys	
(Y.-C. Lin et al., 2017)[20]	non-dieg.	/	off-scr.	screen	dir.	overt	forced sys	Autopilot
(Gugenheimer et al., 2016)[37]	non-dieg.	visual	off-scr.	world	indir.	overt	voluntary	Arrow
(Danieau et al., 2017)[19]	non-dieg.	/	off-scr.	world	dir.	overt	forced sys	SwiVRChair
(Chang et al., 2018)[38]	non-dieg.	visual	on-scr.	world	dir.	overt	voluntary	Fading, Desaturation
(Sassatelli et al., 2018)[39]	non-dieg.	haptic	off-scr.	screen	indir.	overt	voluntary	FacePush
(Gruenefeld et al., 2018b)[40]	non-dieg.	/	off-scr.	screen	dir.	overt	forced sys	Snap-Changes
	non-dieg.	visual	off-scr.	screen	indir.	overt	voluntary	RadialLight

“dieg.” = “diegetic”; “scr.” = “screen”; “forced sys” = “forced by system”; For “forced sys” methods, no sense is assigned (/) since this method cannot be influenced by the users’ sense.

## 5. Methods for CVR Adapted from Guiding in Traditional Movies and Images (2D)

In this chapter, we present well-known guiding methods used in traditional movies or images and classify them according to the taxonomy of chapter 4. These methods can be used as *on-screen* methods in CVR. Using the taxonomy, differences and similarities of methods could be found and unique characteristics identified.

### 5.1. Diegetic Methods Diegetic, Visual/Auditive, on-Screen/off-Screen, World-Referenced, Subtle, Voluntary

In traditional filmmaking, movements, sounds or lights included in the story can guide the viewers' attention [107]. Such diegetic techniques can guide the gaze also in Cinematic VR if they are in the field of view (on-screen). Even if most diegetic methods are on-screen techniques, there are some exceptions, where such methods can be applied to off-screen guiding:

- Diegetic visual cues: If a person is looking in a direction out of the screen, the viewer will mostly follow it [18]. The same is true for moving objects [22].
- Diegetic auditive cues: Sounds motivate the user to search for the source of the sound and therefore to change the viewing direction [22].

Diegetic cues are subtle and the viewer is free to follow them. Due to the nature of diegetic methods, they are always world-referenced. Non-diegetic methods can be both world-referenced or screen-referenced. Diegetic cues are mostly direct, however, exceptions are conceivable, for example a person speaking about an object in the room.

### 5.2. Image Modulation non-Diegetic, Visual, on-Screen, world-Referenced, Subtle/overt, Voluntary

Image modulations, such as changing color, saliency or saturation, are mostly non-diegetic, visual, on-screen, world-referenced, and voluntary. If the modulation is subtle or overt depends on the degree of modification. Saliency modulation, as well as blurring, are effects which are used in traditional movies for guiding viewers' attention. Danieau et al. [19] applied them to CVR and compared four video effects for CVR: (1) fading-to-black for the area out of interest, (2) desaturation (like SMT), (3) blurring, and (4) deformation by displaying a wavelike effect on the side of the viewer's field of view. In an informal user study, blur and deformation were not successful in guiding. Comparing fading-to-black, desaturation, no guiding, and forced rotation in the main study, they found a trade-off between the efficiency and noticeability of the effects. They were either disturbing (fading-to-black) or ineffective (desaturation). In some of our user studies, we made similar experiences [13]. Deformations were either not subtle or not working. We think that for blurring methods, the resolution of movies and displays are not high enough for noticing a relevant difference between the blurred and non-blurred area.

### 5.3. Overlays non-Diegetic, Visual, on-Screen/off-Screen, World/Screen-Referenced, Overt, Voluntary

Overlays, such as arrows, are indirect indicators. It requires interpretation to find the right direction. Such methods are well-known on flat-screens, but also available in VR environments. Lin et al. [20] compared an arrow with a forced rotation (autopilot). Both methods are very obvious and were evaluated for a sports video and a city tour. Forced rotation is suitable in cases where the viewer needs to see a detail in time whereas an arrow indicates something or gives hints.

### 5.4. Subtle Gaze Direction non-Diegetic, Visual, on-Screen/off-Screen, World-Referenced, Subtle, Voluntary

Subtle methods for gaze direction (SGD) were investigated for static images on flat displays [12,60]. Such methods can improve the success in search tasks [60] and reduce the error rate in remembering regions and their locations [12]. To extend these methods to CVR or VR, there are several issues to consider:

- A method developed for images works in a static environment. The remaining part of the picture does not change. This is not the case for videos.
- A method developed for clear test environments sometimes might not work for complex images or videos with a lot of objects competing for attention.
- A method developed for a monitor has to be extended for the case where the target object is not in the FoV.
- A method using flickering must take into account the frame rate of the movie and the HMD.



We tested subtle gaze directing for CVR [13] and achieved similar results as Danieau et al. for video effects [19]: searching the right parameters for the method resulted in a technique which either was not subtle or did not work well. SGD which is working well for still images is difficult to adapt to CVR. That may be because of the available hardware. Depending on the used type of SGD, high display frequencies or a wide field of view are necessary. For using the different sensory perception of fovea and periphery of the eye, the FoV of an HMD does not seem to be large enough. To adapt high frequency subtle flickering methods, the frequency of the HMD display needs to be higher. On the other side, movements in the movie might render subtle cues ineffective. However, we could find a higher recall rate with (non-subtle) flickering.

## 6. Methods for CVR adapted from VR and AR (3D)

Following the above taxonomy, we present several known methods from VR and AR and classify them according to the taxonomy. For each method, we discuss if and how it can be adapted to CVR.

### 6.1. Arrows and Similar Signs non-Diegetic, Visual, on-Screen/off-Screen, World-Referenced/Screen-Referenced, Overt, Voluntary

Arrows are well known (Figure 1a) and often used for showing directions in real life. In several papers, they were compared with other guiding methods [20,51,53,54]. They work well but can be disturbing. Augmented Reality methods such as attention funnel [81,114] or ParaFrustum [82] could be suitable for instruction or education application. Both are realized by drawing augmented elements, which start at the viewer's eyes and lead to the region of interest. The methods are overt and usable for only one PoI. Since the overlay partially covers the RoI, it is less suitable for CVR movie experiences.

### 6.2. Stylistic Rendering non-Diegetic, Visual, on-Screen, World-Referenced, subtle/Overt, Voluntary

Stylistic rendering methods [42] for 3D models are similar to image modulation methods described in 5.2. and can be adapted to CVR. To find the perfect rendering style for each target can be a creative part of CVR filmmaking. However, for noticing such an effect, it has to be in the field of view (on-screen).

### 6.3. Picture-in-Picture Displays non-Diegetic, Visual, off-Screen, Screen-Referenced, Overt, Voluntary

All methods described so far indicate the direction of the RoI. In contrast, showing the RoI in a small inline-window (Figure 1b, example from our work) at the screen offers the advantage that the viewer knows what to expect and thus can decide if the viewing direction should be changed. One disadvantage of this method is that the window covers a part of the content. The other drawback, the missing information about the position of the RoI, can be solved by placing the display on the side near the RoI.

Lin et al. [20] evaluated this method for omnidirectional movies on mobile phones. The method outperformed arrow-based guidance for most aspects, even if it occupied more space.

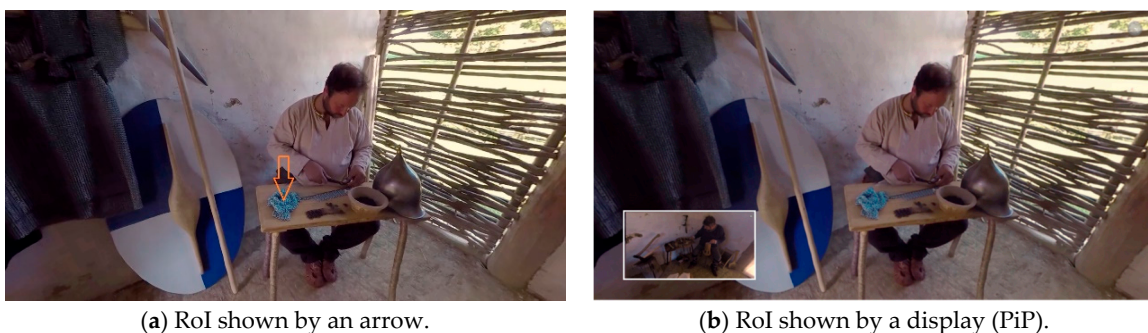


Figure 1. Two overlay methods for CVR (Images taken from our work).

#### 6.4. Radar non-Diegetic, Visual, off-Screen, Screen-Referenced, Overt, Voluntary

Methods used by sailplanes for collision avoidance systems could be used to show the RoI. Such systems show from which direction another sailplane comes. We implemented a method to indicate the direction of the PoI (Figure 2a). The bar at the bottom shows if the PoI is on the right or on the left side. The bar on the right shows if the PoI is higher or lower than the own viewing direction. Another example can be seen in Figure 2b, where the direction is shown by a circle and the height by a bar.



(a) The bar at the bottom shows if the PoI is to the right or left. The bar on the right shows if the PoI is higher or lower than the own viewing direction.



(b) The direction is shown by a circle and the height by a bar (radar circle method).

**Figure 2.** Collision avoidance methods of sailplanes transferred to indicate the direction of the PoI. In both cases, the direction is on the left side behind the viewer, a bit below the own viewing direction.

### 7. Practical Considerations when Applying the Taxonomy

The introduced taxonomy supports researchers and practitioners in designing guiding methods for Cinematic VR. This chapter connects the dimensions with the design questions of the filmmaker. When developing a CVR experience, filmmakers know their material and can think about the desired effect of a guiding method and about its attributes. To make decisions about the most appropriate guiding technique, various aspects are relevant and are captured by answering the following questions:

- How fast should the method work (tempo)?
- Is presence more important than the effectiveness of the method (effectiveness)?
- Is it important that the viewer remembers the target (recall rate)?
- Are there more than one RoI simultaneously (number of RoIs)?
- Is there a problem if indicators (arrows, halo) cover movie content (covering)?
- How complex is the content (visual or auditive clutter)?
- Should the guiding method be part of the CVR movie (experience)?

Answering these questions is the first step in finding the right technique.

*Tempo:* In traditional movies, the filmmaker can determine the pace by cutting and showing image sections for a short or long time. In CVR, the user explores the scene by changing the viewing direction in its own tempo. The filmmaker can influence this process by choosing the right guiding method. To affect the pace of a movie, one can choose between slow- and fast-acting methods. For example, forced methods work very fast, but they can destroy the experience. However, there are CVR movies imaginable, where a forced rotation is part of the experience.

*Effectiveness:* It depends on the purpose of the movie if effectiveness is more important than presence. More obvious techniques often are more effective, but they can destroy the movie experience. For a relaxing movie event, it might be less important to ensure that viewers always follow the guiding. In that case, diegetic, voluntary methods should be preferred. In contrast, for a sport event, it can be substantial to see the details at the right moment. Here, forced or stimulus-driven methods could be a good choice. For instructional movies, an arrow fits perfectly. It is an overt technique using an indirect cue, hard to overlook, but viewers remain free to follow it or not.

*Recall Rate:* In case the CVR movie is used for learning applications, the recall rate can be important, and more obvious methods can be applied, such as overt, non-diegetic ones. Also, for other genres, it can be relevant that the viewer remembers details of the story, but the indicator should not be so obvious. Based on attention theory, voluntary processes are memory-bound, keep the attention longer, and thus increase the recall rate. Also, modulation techniques (e.g., stylistic rendering, SMT) are applicable since they influence memory [45].

*Number of RoIs:* The above methods are mostly evaluated for a single RoI. Especially for nonlinear storytelling, more than one RoI may be required at the same time. Not all of the mentioned methods are able to handle this case. Although some methods are able to manage more than one RoI, this can lead to overcrowding the display and overtaxing the viewer. It needs more research for finding methods and adjustments to handle more than one RoI simultaneously.

*Covering:* Diegetic or modulation methods do not cover movie content. When using indicators, such as arrows or halos, parts of the movie are not visible. This can be disturbing. However, purposes are conceivable where the indicator is more important than complete visibility, e.g., for instruction videos. Choosing the right parameters (size or color) can make the cue more obvious/effective or subtler. World-referenced overlays (e.g., arrows) stay at the same place in the movie and cover an area permanently. Screen-referenced overlays (such as signs at the display edge) change the covered area if the head is moving.

*Clutter:* To be able to assess whether a method will be suitable, the complexity of an image must be considered. Cluttered images with a lot of details require clear and obvious techniques. If the image is clear, more subtle methods can work. The same is true for audio: If there are a lot of sounds in the movie, it can be difficult to follow a spatial audio signal which should guide to a RoI.

*Experience:* It is not always necessary to make guiding imperceptible, it can be also part of the experience, in the same way, as scene transitions influence the movie. The technique can affect the style, the pace and the atmosphere of a movie.

Overall, the taxonomy above provides support in the process of finding the most suitable technique to address these practical questions.

## 8. Conclusions

Since in Cinematic VR the viewer can freely choose the viewing direction, the selection of the visible image section is no longer defined by the filmmaker, but by the viewer. This can cause problems if the viewer misses an important detail of the story. Also, the viewing experience can suffer because the viewer is afraid to miss something. Additionally, an important aspect of influencing style and pace rests no longer exclusively in the hands of the filmmaker. On the other hand, CVR provides a lot of new opportunities. With the added space component, non-linear, interactive stories, intuitive for the viewer, can be realized. Guiding the viewer in such experiences is not only a requirement, but it is also a chance for the filmmaker to influence the style and pace in novel ways. It may be used like transitions in traditional movies—the filmmaker chooses the best fitting techniques for each Region of Interest. Also, changing between the methods within a movie could be useful, e.g., for changing the pace of certain movie sections.

Based on previous literature, we described a taxonomy for guiding methods. Our focus was on CVR, yet most dimensions are transferable to augmented and virtual reality. Classifying these methods corresponding to the taxonomy assists researchers and practitioners in finding the right technique for different requirements. We listed the advantages and disadvantages of attributes along an identified set of key dimensions. We illustrated each such dimension with concrete examples of guiding methods.

This taxonomy can help to understand the various characteristics of guiding techniques, to find new methods which have not yet been analysed and support filmmakers to find the right methods for their projects.

**Author Contributions:** Conceptualization, S.R.; methodology, S.R. and H.H.; investigation, S.R.; writing—original draft preparation, S.R.; writing—review and editing, S.R., D.B. and H.H.; supervision, H.H.

**Funding:** This research received no external funding.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. De Abreu, A.; Ozcinar, C.; Smolic, A. Look around you: Saliency maps for omnidirectional images in VR applications. In Proceedings of the 2017 Ninth International Conference on Quality of Multimedia Experience (QoMEX), Erfurt, Germany, 31 May–2 June 2017.
2. Petry, B.; Huber, J. Towards effective interaction with omnidirectional videos using immersive virtual reality headsets. In Proceedings of the 6th Augmented Human International Conference on—AH '15, Singapore, 9–11 March 2015; ACM Press: New York, NY, USA, 2015; pp. 217–218.
3. 5 Lessons Learned While Making Lost | Oculus. Available online: <https://www.oculus.com/story-studio/blog/5-lessons-learned-while-making-lost/> (accessed on 10 December 2018).
4. Tse, A.; Jennett, C.; Moore, J.; Watson, Z.; Rigby, J.; Cox, A.L. Was I There? In Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems—CHI EA '17, Denver, CO, USA, 6–11 May 2017; ACM Press: New York, NY, USA, 2017; pp. 2967–2974.
5. MacQuarrie, A.; Steed, A. Cinematic virtual reality: Evaluating the effect of display type on the viewing experience for panoramic video. In Proceedings of the 2017 IEEE Virtual Reality (VR), Los Angeles, CA, USA, 18–22 March 2017; pp. 45–54.
6. Vosmeer, M.; Schouten, B. Interactive cinema: Engagement and interaction. In Proceedings of the International Conference on Interactive Digital Storytelling, Singapore, 3–6 November 2014; pp. 140–147.
7. Rothe, S.; Tran, K.; Hußmann, H. Dynamic Subtitles in Cinematic Virtual Reality. In *Proceedings of the 2018 ACM International Conference on Interactive Experiences for TV and Online Video—TVX '18*; ACM Press: New York, NY, USA, 2018; pp. 209–214.
8. Liu, D.; Bhagat, K.; Gao, Y.; Chang, T.-W.; Huang, R. The Potentials and Trends of VR in Education: A Bibliometric Analysis on Top Research Studies in the last Two decades. In *Augmented and Mixed Realities in Education*; Springer: Singapore, 2017; pp. 105–113.
9. Howard, S.; Serpanchy, K.; Lewin, K. Virtual reality content for higher education curriculum. In Proceedings of the VALA, Melbourne, Australia, 13–15 February 2018.
10. Stojšić, I.; Ivkov-Džigurski, A.; Maričić, O. Virtual Reality as a Learning Tool: How and Where to Start with Immersive Teaching. In *Didactics of Smart Pedagogy*; Springer International Publishing: Cham, Switzerland, 2019; pp. 353–369.
11. Merchant, Z.; Goetz, E.T.; Cifuentes, L.; Keeney-Kennicutt, W.; Davis, T.J. Effectiveness of virtual reality-based instruction on students' learning outcomes in K-12 and higher education: A meta-analysis. *Comput. Educ.* **2014**, *70*, 29–40. [[CrossRef](#)]
12. Bailey, R.; McNamara, A.; Costello, A.; Sridharan, S.; Grimm, C. Impact of subtle gaze direction on short-term spatial information recall. *Proc. Symp. Eye Track. Res. Appl.* **2012**, 67–74. [[CrossRef](#)]
13. Rothe, S.; Althammer, F.; Khamis, M. GazeRecall: Using Gaze Direction to Increase Recall of Details in Cinematic Virtual Reality. In Proceedings of the 11th International Conference on Mobile and Ubiquitous Multimedia (MUM'18), Cairo, Egypt, 25–28 November 2018; ACM: Kairo, Egypt, 2018.
14. Rothe, S.; Montagud, M.; Mai, C.; Buschek, D.; Hußmann, H. Social Viewing in Cinematic Virtual Reality: Challenges and Opportunities. In *Interactive Storytelling. ICIDS 2018*; Lecture Notes in Computer Science; Springer: Cham, Switzerland, 2018; Volume 11318, Presented at the 5 December 2018.
15. Subramanian, R.; Shankar, D.; Sebe, N.; Melcher, D. Emotion modulates eye movement patterns and subsequent memory for the gist and details of movie scenes. *J. Vis.* **2014**, *14*, 31. [[CrossRef](#)] [[PubMed](#)]
16. Dorr, M.; Vig, E.; Barth, E. Eye movement prediction and variability on natural video data sets. *Vis. Cogn.* **2012**, *20*, 495–514. [[CrossRef](#)]
17. Smith, T.J. Watching You Watch Movies: Using Eye Tracking to Inform Cognitive Film Theory. In *Psychocinematics*; Oxford University Press: New York, NY, USA, 2013; pp. 165–192.
18. Brown, A.; Sheikh, A.; Evans, M.; Watson, Z. Directing attention in 360-degree video. In *Proceedings of the IBC 2016 Conference*; Institution of Engineering and Technology: Amsterdam, The Netherlands, 2016.
19. Danieau, F.; Guillo, A.; Dore, R. Attention guidance for immersive video content in head-mounted displays. In Proceedings of the IEEE Virtual Reality, Los Angeles, CA, USA, 18–22 March 2017.



20. Lin, Y.-C.; Chang, Y.-J.; Hu, H.-N.; Cheng, H.-T.; Huang, C.-W.; Sun, M. 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 Press: New York, NY, USA, 2017; pp. 2535–2545.
21. Nielsen, L.T.; Møller, M.B.; Hartmeyer, S.D.; Ljung, T.C.M.; Nilsson, N.C.; Nordahl, R.; Serafin, S. 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 Press: New York, NY, USA, 2016; pp. 229–232.
22. Rothe, S.; Hußmann, H. Guiding the Viewer in Cinematic Virtual Reality by Diegetic Cues. In *Proceedings of the International Conference on Augmented Reality, Virtual Reality and Computer Graphics*; Springer: Cham, Switzerland, 2018; pp. 101–117.
23. Frintrop, S.; Rome, E.; Christensen, H.I. Computational visual attention systems and their cognitive foundations. *ACM Trans. Appl. Percept.* **2010**, *7*, 1–39. [[CrossRef](#)]
24. Borji, A.; Itti, L. State-of-the-art in visual attention modeling. *IEEE Trans. Pattern Anal. Mach. Intell.* **2013**, *35*, 185–207. [[CrossRef](#)] [[PubMed](#)]
25. Smith, T. Edit blindness: The relationship between attention and global change blindness in dynamic scenes. *J. Eye Mov. Res.* **2008**, *2*, 1–17.
26. Posner, M.I. Orienting of attention. *Q. J. Exp. Psychol.* **1980**, *32*, 3–25. [[CrossRef](#)] [[PubMed](#)]
27. Ward, L. Scholarpedia: Attention. *Scholarpedia* **2008**, *3*, 1538. [[CrossRef](#)]
28. Yarbus, A.L. *Eye Movements and Vision*; Springer: Boston, MA, USA, 1967.
29. Treisman, A.M.; Gelade, G. A feature-integration theory of attention. *Cogn. Psychol.* **1980**, *12*, 97–136. [[CrossRef](#)]
30. Healey, C.G.; Booth, K.S.; Enns, J.T. High-speed visual estimation using preattentive processing. *ACM Trans. Comput. Interact.* **1996**, *3*, 107–135. [[CrossRef](#)]
31. Wolfe, J.M.; Horowitz, T.S. Five factors that guide attention in visual search. *Nat. Hum. Behav.* **2017**, *1*, 0058. [[CrossRef](#)]
32. Wolfe, J.M.; Horowitz, T.S. What attributes guide the deployment of visual attention and how do they do it? *Nat. Rev. Neurosci.* **2004**, *5*, 495–501. [[CrossRef](#)]
33. Tyler, C.W.; Hamer, R.D. Eccentricity and the Ferry–Porter law. *J. Opt. Soc. Am. A* **1993**, *10*, 2084. [[CrossRef](#)]
34. Rovamo, J.; Raninen, A. Critical flicker frequency as a function of stimulus area and luminance at various eccentricities in human cone vision: A revision of granit-harper and ferry-porter laws. *Vis. Res.* **1988**, *28*, 785–790. [[CrossRef](#)]
35. Grimes, J.D. Effects of Patterning on Flicker Frequency. In *Proceedings of the Human Factors Society Annual Meeting*; SAGE Publications: Los Angeles, CA, USA, 1983; pp. 46–50.
36. Waldin, N.; Waldner, M.; Viola, I. Flicker Observer Effect: Guiding Attention Through High Frequency Flicker in Images. *Comput. Graph. Forum* **2017**, *36*, 467–476. [[CrossRef](#)]
37. Gugenheimer, J.; Wolf, D.; Haas, G.; Krebs, S.; Rukzio, E. 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 Press: New York, NY, USA, 2016; pp. 1996–2000.
38. Chang, H.-Y.; Tseng, W.-J.; Tsai, C.-E.; Chen, H.-Y.; Peiris, R.L.; Chan, L. FacePush: Introducing Normal Force on Face with Head-Mounted Displays. In *Proceedings of the 31st Annual ACM Symposium on User Interface Software and Technology—UIST '18*; ACM Press: New York, NY, USA, 2018; pp. 927–935.
39. Sassatelli, L.; Pinna-Déry, A.-M.; Winckler, M.; Dambra, S.; Samela, G.; Pighetti, R.; Aparicio-Pardo, R. Snap-changes: A Dynamic Editing Strategy for Directing Viewer's Attention in Streaming Virtual Reality Videos. In *Proceedings of the 2018 International Conference on Advanced Visual Interfaces—AVI '18*; ACM Press: New York, NY, USA, 2018; pp. 1–5.
40. Gruenefeld, U.; Stratmann, T.C.; El Ali, A.; Boll, S.; Heuten, W. RadialLight. In *Proceedings of the 20th International Conference on Human-Computer Interaction with Mobile Devices and Services—MobileHCI '18*, Barcelona, Spain, 3–6 September 2018; ACM Press: New York, NY, USA, 2018; pp. 1–6.
41. Lin, Y.-T.; Liao, Y.-C.; Teng, S.-Y.; Chung, Y.-J.; Chan, L.; Chen, B.-Y. 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 Press: New York, NY, USA, 2017; pp. 255–265.

42. Cole, F.; DeCarlo, D.; Finkelstein, A.; Kin, K.; Morley, K.; Santella, A. Directing gaze in 3D models with stylized focus. *Proc. 17th Eurographics Conf. Render. Tech.* **2006**, 377–387. [[CrossRef](#)]
43. Tanaka, R.; Narumi, T.; Tanikawa, T.; Hirose, M. Attracting User's Attention in Spherical Image by Angular Shift of Virtual Camera Direction. In *Proceedings of the 3rd ACM Symposium on Spatial User Interaction—SUI '15*; ACM Press: New York, NY, USA, 2015; pp. 61–64.
44. Mendez, E.; Feiner, S.; Schmalstieg, D. Focus and Context in Mixed Reality by Modulating First Order Salient Features. In *Proceedings of the International Symposium on Smart GraphicsL*; Springer: Berlin/Heidelberg, Germany, 2010; pp. 232–243.
45. Veas, E.E.; Mendez, E.; Feiner, S.K.; Schmalstieg, D. Directing attention and influencing memory with visual saliency modulation. In *Proceedings of the 2011 Annual Conference on Human Factors in Computing Systems—CHI '11*; ACM Press: New York, NY, USA, 2011; p. 1471.
46. Hoffmann, R.; Baudisch, P.; Weld, D.S. Evaluating visual cues for window switching on large screens. In *Proceeding of the Twenty-Sixth Annual CHI Conference on Human Factors in Computing Systems—CHI '08*; ACM Press: New York, NY, USA, 2008; p. 929.
47. Renner, P.; Pfeiffer, T. Attention Guiding Using Augmented Reality in Complex Environments. In *Proceeding of the 2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*; IEEE: Reutlingen, Germany, 2018; pp. 771–772.
48. Perea, P.; Morand, D.; Nigay, L. [POSTER] Halo3D: A Technique for Visualizing Off-Screen Points of Interest in Mobile Augmented Reality. In *Proceeding of the 2017 IEEE International Symposium on Mixed and Augmented Reality (ISMAR-Adjunct)*, Nantes, France, 9–13 October 2017; pp. 170–175.
49. Gruenefeld, U.; El Ali, A.; Boll, S.; Heuten, W. Beyond Halo and Wedge. In *Proceedings of the 20th International Conference on Human-Computer Interaction with Mobile Devices and Services—MobileHCI '18*; ACM Press: New York, NY, USA, 2018; pp. 1–11.
50. Gruenefeld, U.; Ennenga, D.; El Ali, A.; Heuten, W.; Boll, S. EyeSee360. In *Proceedings of the 5th Symposium on Spatial User Interaction—SUI '17*; ACM Press: New York, NY, USA, 2017; pp. 109–118.
51. Bork, F.; Schnelzer, C.; Eck, U.; Navab, N. Towards Efficient Visual Guidance in Limited Field-of-View Head-Mounted Displays. *IEEE Trans. Vis. Comput. Graph.* **2018**, *24*, 2983–2992. [[CrossRef](#)] [[PubMed](#)]
52. Siu, T.; Herskovic, V. SideBARs: Improving awareness of off-screen elements in mobile augmented reality. In *Proceedings of the 2013 Chilean Conference on Human-Computer Interaction—ChileCHI '13*; ACM Press: New York, NY, USA, 2013; pp. 36–41.
53. Renner, P.; Pfeiffer, T. Attention guiding techniques using peripheral vision and eye tracking for feedback in augmented-reality-based assistance systems. In *Proceedings of the 2017 IEEE Symposium on 3D User Interfaces (3DUI)*; IEEE: Los Angeles, CA, USA, 2017; pp. 186–194.
54. Burigat, S.; Chittaro, L.; Gabrielli, S. Visualizing locations of off-screen objects on mobile devices. In *Proceedings of the 8th Conference on Human-Computer Interaction with Mobile Devices and Services—MobileHCI '06*; ACM Press: New York, NY, USA, 2006; p. 239.
55. Henze, N.; Boll, S. Evaluation of an off-screen visualization for magic lens and dynamic peephole interfaces. In *Proceedings of the 12th International Conference on Human Computer Interaction with Mobile Devices and Services—MobileHCI '10*; ACM Press: New York, NY, USA, 2010; p. 191.
56. Schinke, T.; Henze, N.; Boll, S. Visualization of off-screen objects in mobile augmented reality. In *Proceedings of the 12th International Conference on Human Computer Interaction with Mobile Devices and Services—MobileHCI '10*; ACM Press: New York, NY, USA, 2010; p. 313.
57. Koskinen, E.; Rakkolainen, I.; Raisamo, R. Direct retinal signals for virtual environments. In *Proceedings of the 23rd ACM Symposium on Virtual Reality Software and Technology—VRST '17*; ACM Press: New York, NY, USA, 2017; pp. 1–2.
58. Zellweger, P.T.; Mackinlay, J.D.; Good, L.; Stefik, M.; Baudisch, P. City lights. In *Proceedings of the CHI '03 Extended Abstracts on Human Factors in Computing Systems—CHI '03*; ACM Press: New York, NY, USA, 2003; p. 838.
59. Bailey, R.; McNamara, A.; Sudarsanam, N.; Grimm, C. Subtle gaze direction. *ACM Trans. Graph.* **2009**, *28*, 1–14. [[CrossRef](#)]
60. McNamara, A.; Bailey, R.; Grimm, C. Improving search task performance using subtle gaze direction. In *Proceedings of the 5th Symposium on Applied Perception in Graphics and Visualization—APGV '08*; ACM Press: New York, NY, USA, 2008; p. 51.



61. Grogorick, S.; Stengel, M.; Eisemann, E.; Magnor, M. Subtle gaze guidance for immersive environments. In *Proceedings of the ACM Symposium on Applied Perception—SAP '17*; ACM Press: New York, NY, USA, 2017; pp. 1–7.
62. McNamara, A.; Booth, T.; Sridharan, S.; Caffey, S.; Grimm, C.; Bailey, R. Directing gaze in narrative art. In *Proceedings of the ACM Symposium on Applied Perception—SAP '12*; ACM Press: New York, NY, USA, 2012; p. 63.
63. Lu, W.; Duh, H.B.-L.; Feiner, S.; Zhao, Q. Attributes of Subtle Cues for Facilitating Visual Search in Augmented Reality. *IEEE Trans. Vis. Comput. Graph.* **2014**, *20*, 404–412. [[CrossRef](#)] [[PubMed](#)]
64. Smith, W.S.; Tadmor, Y. Nonblurred regions show priority for gaze direction over spatial blur. *Q. J. Exp. Psychol.* **2013**, *66*, 927–945. [[CrossRef](#)] [[PubMed](#)]
65. Hata, H.; Koike, H.; Sato, Y. Visual Guidance with Unnoticed Blur Effect. In *Proceedings of the International Working Conference on Advanced Visual Interfaces—AVI '16*; ACM Press: New York, NY, USA, 2016; pp. 28–35.
66. Hagiwara, A.; Sugimoto, A.; Kawamoto, K. Saliency-based image editing for guiding visual attention. In *Proceedings of the 1st International Workshop on Pervasive Eye Tracking & Mobile Eye-Based Interaction—PETMEI '11*; ACM Press: New York, NY, USA, 2011; p. 43.
67. Kosek, M.; Koniaris, B.; Sinclair, D.; Markova, D.; Rothnie, F.; Smoot, L.; Mitchell, K. IRIDIUM+: Deep Media Storytelling with Non-linear Light Field Video. In *Proceedings of the ACM SIGGRAPH 2017 VR Village on—SIGGRAPH '17*; ACM Press: New York, NY, USA, 2017; pp. 1–2.
68. Kaul, O.B.; Rohs, M. HapticHead: A Spherical Vibrotactile Grid around the Head for 3D Guidance in Virtual and Augmented Reality. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems—CHI '17*; ACM Press: New York, NY, USA, 2017; pp. 3729–3740.
69. Rantala, J.; Kangas, J.; Raisamo, R. Directional cueing of gaze with a vibrotactile headband. In *Proceedings of the 8th Augmented Human International Conference on—AH '17*; ACM Press: New York, NY, USA, 2017; pp. 1–7.
70. Stratmann, T.C.; Löcken, A.; Gruenefeld, U.; Heuten, W.; Boll, S. Exploring Vibrotactile and Peripheral Cues for Spatial Attention Guidance. In *Proceedings of the 7th ACM International Symposium on Pervasive Displays—PerDis '18*; ACM Press: New York, NY, USA, 2018; pp. 1–8.
71. Knierim, P.; Kosch, T.; Schwind, V.; Funk, M.; Kiss, F.; Schneegass, S.; Henze, N. Tactile Drones - Providing Immersive Tactile Feedback in Virtual Reality through Quadcopters. In *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems—CHI EA '17*; ACM Press: New York, NY, USA, 2017; pp. 433–436.
72. Sridharan, S.; Pieszala, J.; Bailey, R. Depth-based subtle gaze guidance in virtual reality environments. In *Proceedings of the ACM SIGGRAPH Symposium on Applied Perception—SAP '15*; ACM Press: New York, NY, USA, 2015; p. 132.
73. Kim, Y.; Varshney, A. Saliency-guided Enhancement for Volume Visualization. *IEEE Trans. Vis. Comput. Graph.* **2006**, *12*, 925–932. [[CrossRef](#)]
74. Jarodzka, H.; van Gog, T.; Dorr, M.; Scheiter, K.; Gerjets, P. Learning to see: Guiding students' attention via a Model's eye movements fosters learning. *Learn. Instr.* **2013**, *25*, 62–70. [[CrossRef](#)]
75. Lintu, A.; Carbonell, N. Gaze Guidance through Peripheral Stimuli. 2009. Available online: <https://hal.inria.fr/inria-00421151/> (accessed on 31 January 2019).
76. Khan, A.; Matejka, J.; Fitzmaurice, G.; Kurtenbach, G. Spotlight. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems—CHI '05*; ACM Press: New York, NY, USA, 2005; p. 791.
77. Barth, E.; Dorr, M.; Böhme, M.; Gegenfurtner, K.; Martinetz, T. Guiding the mind's eye: Improving communication and vision by external control of the scanpath. In *Human Vision and Electronic Imaging XI*; International Society for Optics and Photonics: Washington, DC, USA, 2006; Volume 6057, p. 60570D.
78. Dorr, M.; Dorr, M.; Vig, E.; Gegenfurtner, K.R.; Martinetz, T.; Barth, E. Eye movement modelling and gaze guidance. In *Proceedings of the Fourth International Workshop on Human-Computer Conversation*, Bellagio, Italy, 6–7 October 2008.
79. Sato, Y.; Sugano, Y.; Sugimoto, A.; Kuno, Y.; Koike, H. Sensing and Controlling Human Gaze in Daily Living Space for Human-Harmonized Information Environments. In *Human-Harmonized Information Technology*; Springer: Tokyo, Japan, 2016; Volume 1, pp. 199–237.
80. Vig, E.; Dorr, M.; Barth, E. Learned saliency transformations for gaze guidance. In *Human Vision and Electronic Imaging XVI*; International Society for Optics and Photonics: Bellingham, WA, USA, 2011; p. 78650W.

81. Biocca, F.; Owen, C.; Tang, A.; Bohil, C. Attention Issues in Spatial Information Systems: Directing Mobile Users' Visual Attention Using Augmented Reality. *J. Manag. Inf. Syst.* **2007**, *23*, 163–184. [[CrossRef](#)]
82. Sukan, M.; Elvezio, C.; Oda, O.; Feiner, S.; Tversky, B. ParaFrustum. In *Proceedings of the 27th Annual ACM Symposium on User Interface Software and Technology—UIST '14*; ACM Press: New York, NY, USA, 2014; pp. 331–340.
83. Kosara, R.; Miksch, S.; Hauser, H. Focus+context taken literally. *IEEE Comput. Graph. Appl.* **2002**, *22*, 22–29. [[CrossRef](#)]
84. Mateescu, V.A.; Bajić, I.V. Attention Retargeting by Color Manipulation in Images. In *Proceedings of the 1st International Workshop on Perception Inspired Video Processing—PIVP '14*; ACM Press: New York, NY, USA, 2014; pp. 15–20.
85. Delamare, W.; Han, T.; Irani, P. Designing a gaze gesture guiding system. In *Proceedings of the 19th International Conference on Human-Computer Interaction with Mobile Devices and Services—MobileHCI '17*; ACM Press: New York, NY, USA, 2017; pp. 1–13.
86. Pausch, R.; Snoddy, J.; Taylor, R.; Watson, S.; Haseltine, E. Disney's Aladdin. In *Proceedings of the 23rd Annual Conference on Computer Graphics and Interactive Techniques—SIGGRAPH '96*; ACM Press: New York, NY, USA, 1996; pp. 193–203.
87. Souriau, E. La structure de l'univers filmique et le vocabulaire de la filmologie. | Interdisciplinary Center for Narratology. *Rev. Int. de Filmol.* **1951**, 7–8, 231–240.
88. Silva, A.; Raimundo, G.; Paiva, A. Tell me that bit again . . . bringing interactivity to a virtual storyteller. In *Proceedings of the International Conference on Virtual Storytelling*, Toulouse, France, 20–21 November 2003; pp. 146–154.
89. Brown, C.; Bhutra, G.; Suhail, M.; Xu, Q.; Ragan, E.D. Coordinating attention and cooperation in multi-user virtual reality narratives. In *Proceedings of the 2017 IEEE Virtual Reality (VR)*, Los Angeles, CA, USA, 18–22 March 2017; pp. 377–378.
90. Niebur, E. Saliency map. *Scholarpedia* **2007**, *2*, 2675. [[CrossRef](#)]
91. Itti, L.; Koch, C.; Niebur, E. A model of saliency-based visual attention for rapid scene analysis. *IEEE Trans. Pattern Anal. Mach. Intell.* **1998**, *20*, 1254–1259. [[CrossRef](#)]
92. Baudisch, P.; DeCarlo, D.; Duchowski, A.T.; Geisler, W.S. Focusing on the essential. *Commun. ACM* **2003**, *46*, 60. [[CrossRef](#)]
93. Duchowski, A.T.; Cournia, N.; Murphy, H. Gaze-Contingent Displays: A Review. *CyberPsychol. Behav.* **2004**, *7*, 621–634. [[CrossRef](#)] [[PubMed](#)]
94. Sitzmann, V.; Serrano, A.; Pavel, A.; Agrawala, M.; Gutierrez, D.; Masia, B.; Wetzstein, G. Saliency in VR: How Do People Explore Virtual Environments? *IEEE Trans. Vis. Comput. Graph.* **2018**, *24*, 1633–1642. [[CrossRef](#)] [[PubMed](#)]
95. Baudisch, P.; Rosenholtz, R. Halo: A technique for visualizing offscreen location. In *Proceedings of the Conference on Human Factors in Computing Systems CHI'03*, Ft. Lauderdale, FL, USA, 5–10 April 1993.
96. Gustafson, S.G.; Irani, P.P. Comparing visualizations for tracking off-screen moving targets. In *Proceedings of the CHI'07 Extended Abstracts on Human Factors in Computing Systems*, San Jose, CA, USA, 28 April–3 May 2007; pp. 2399–2404.
97. Gustafson, S.; Baudisch, P.; Gutwin, C.; Irani, P. Wedge: Clutter-free visualization of off-screen locations. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, Florence, Italy, 5–10 April 2008; pp. 787–796.
98. Gruenefeld, U.; El Ali, A.; Heuten, W.; Boll, S. Visualizing out-of-view objects in head-mounted augmented reality. In *Proceedings of the 19th International Conference on Human-Computer Interaction with Mobile Devices and Services—MobileHCI '17*; ACM Press: New York, NY, USA, 2017; pp. 1–7.
99. Kolasinski, E.M. Simulator Sickness in Virtual Environments. Available online: <https://apps.dtic.mil/docs/citations/ADA295861> (accessed on 17 March 2019).
100. Davis, S.; Nesbitt, K.; Nalivaiko, E. A Systematic Review of Cybersickness. In *Proceedings of the 2014 Conference on Interactive Entertainment—IE2014*; ACM Press: New York, NY, USA, 2014; pp. 1–9.
101. Pavel, A.; Hartmann, B.; Agrawala, M. Shot Orientation Controls for Interactive Cinematography with 360 Video. In *Proceedings of the 30th Annual ACM Symposium on User Interface Software and Technology—UIST '17*, Québec City, QC, Canada, 22–25 October 2017; ACM Press: New York, NY, USA, 2017; pp. 289–297.

102. Crossing the Line. Available online: <https://www.mediacollege.com/video/editing/transition/reverse-cut.html> (accessed on 10 December 2018).
103. Knierim, P.; Kosch, T.; Achberger, A.; Funk, M. Flyables: Exploring 3D Interaction Spaces for Levitating Tangibles. In *Proceedings of the Twelfth International Conference on Tangible, Embedded, and Embodied Interaction—TEI '18*; ACM Press: New York, NY, USA, 2018; pp. 329–336.
104. Hoppe, M.; Knierim, P.; Kosch, T.; Funk, M.; Futami, L.; Schneegass, S.; Henze, N.; Schmidt, A.; Machulla, T. VRHapticDrones: Providing Haptics in Virtual Reality through Quadcopters. In *Proceedings of the 17th International Conference on Mobile and Ubiquitous Multimedia—MUM 2018*; ACM Press: New York, NY, USA, 2018; pp. 7–18.
105. Nilsson, N.C.; Serafin, S.; Nordahl, R. Walking in Place Through Virtual Worlds. In *Human-Computer Interaction. Interaction Platforms and Techniques. HCI 2016*; Lecture Notes in Computer Science; Springer: Cham, Switzerland, 2016; Volume 9732.
106. Suma, E.A.; Bruder, G.; Steinicke, F.; Krum, D.M.; Bolas, M. A taxonomy for deploying redirection techniques in immersive virtual environments. In *Proceedings of the 2012 IEEE Virtual Reality (VR)*; IEEE: Costa Mesa, CA, USA, 2012; pp. 43–46.
107. Bordwell, D.; Thompson, K. *Film Art: An Introduction*; McGraw-Hill: New York, NY, USA, 2013.
108. Shah, P.; Miyake, A. *The Cambridge Handbook of Visuospatial Thinking*; Cambridge University Press: Cambridge, UK, 2005.
109. Yeh, M.; Wickens, C.D.; Seagull, F.J. Target Cuing in Visual Search: The Effects of Conformality and Display Location on the Allocation of Visual Attention. *Hum. Factors J. Hum. Factors Ergon. Soc.* **1999**, *41*, 524–542. [[CrossRef](#)] [[PubMed](#)]
110. Renner, P.; Pfeiffer, T. Evaluation of Attention Guiding Techniques for Augmented Reality-based Assistance in Picking and Assembly Tasks. In *Proceedings of the 22nd International Conference on Intelligent User Interfaces Companion—IUI '17 Companion*; ACM Press: New York, NY, USA, 2017; pp. 89–92.
111. Wright, R.D. *Visual Attention*; Oxford University Press: New York, NY, USA, 1998.
112. Carrasco, M. Visual attention: The past 25 years. *Vis. Res.* **2011**, *51*, 1484–1525. [[CrossRef](#)] [[PubMed](#)]
113. Itti, L.; Koch, C. Computational modelling of visual attention. *Nat. Rev. Neurosci.* **2001**, *2*, 194–203. [[CrossRef](#)] [[PubMed](#)]
114. Biocca, F.; Tang, A.; Owen, C.; Xiao, F. Attention Funnel: Omnidirectional 3D Cursor for Mobile Augmented Reality Platforms. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems—CHI '06*; ACM Press: New York, NY, USA, 2006; p. 1115.



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).