OmniView: An Exploratory Study of 360 Degree Vision using Dynamic Distortion based on Direction of Interest

Feng Liang
Ignition Point, Inc.
Tokyo, Japan
feng.liang@ignitionpoint-inc.com

Stevanus Kevin
Holger Baldauf
Kai Kunze
Keio University, Tokyo, Japan
{stevanuskevin765,Holger.Baldauf}@gmail.com
kai@kmd.keio.ac.jp

Yun Suen Pai
The University of Auckland
Auckland, New Zealand
yspai1412@gmail.com

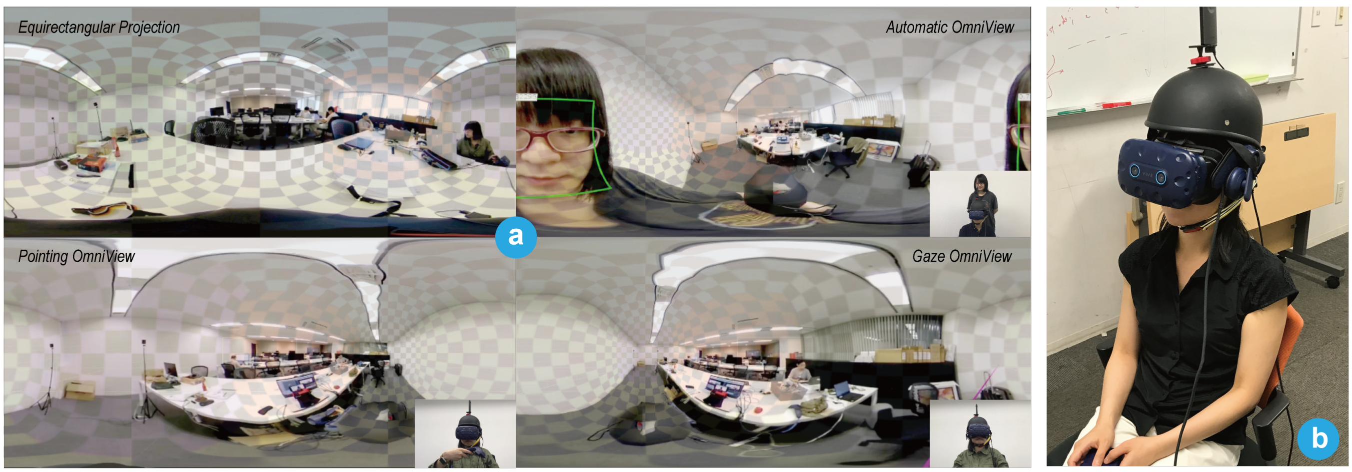

Figure 1: (a) The actual view seen by users using OmniView, with an equirectangular projection and 3 Variations of OmniView: Gaze OmniView, Pointing OmniView, Automatic OmniView (face marker in Automatic OmniView and line renderer in Gaze OmniView were turned off during user studies). (b) Participant wearing the OmniView prototype during user study.

ABSTRACT

The possibility of providing humans with a 360 field of view (FOV) is an area that has fascinated researchers for some time. We introduce OmniView, an exploratory study to determine an optimized 360 FOV vision using dynamic distortion methods for reducing distortion and enlarging the area of the direction of interest. We developed three variations of OmniView to trigger distortions: eye gaze selection, pointing selection, and automatic selection. Our first study (n=16) compares these methods with no dynamic distortion (equirectangular view) in terms of spatial perception change and reaction time. The second study (n=16) evaluated subjective simulator sickness while performing simple everyday tasks. There is no significant change between OmniView and a no distortion baseline regarding spatial perception and reaction time. However, there is a statistically significant rise in disorientation from before the study to after using Pointing OmniView and Automatic OmniView. Most users (n=13) have a strong preference towards Gaze OmniView. From the results, we provide design guidelines towards further optimizing 360 vision and preserving our spatial sense and reaction time while keeping simulator sickness to a minimum.

CCS CONCEPTS

• **Human-centered computing → Information visualization**; *Usability testing*.

KEYWORDS

360 degree vision; projections; visual augmentation

ACM Reference Format:
Feng Liang, Stevanus Kevin, Holger Baldauf, Kai Kunze, and Yun Suen Pai. 2020. OmniView: An Exploratory Study of 360 Degree Vision using Dynamic Distortion based on Direction of Interest. In *AHs '20: Augmented Humans International Conference (AHs '20), March 16–17, 2020, Kaiserslautern, Germany.* ACM, New York, NY, USA, 10 pages. https://doi.org/10.1145/3384657.3384796

AHs '20, March 16–17, 2020, Kaiserslautern, Germany

ACM ISBN 978-1-4503-7603-7/20/03…$15.00
https://doi.org/10.1145/3384657.3384796

1 INTRODUCTION

We humans have binocular vision. Our field of view (FOV) is limited to achieve better depth perception. In contrast, monocular vision can perceive the environment with one eye at a time and no overlap. This results in poorer depth estimation in exchange for superior FOV. Contrary to popular beliefs, binocularity is not solely linked to predators; the strains of hunting and being hunted leaves a clear mark in the development of visual streaks. Yet, binocularity depends highly on each animal's specific requirements derived from their environment [7]. As humans, the configuration of our eyes allows us to perform proper hand-eye coordination, but we cannot see behind or above us. Our FOV is limited to ca. 180 degrees. Therefore, it is not possible for us humans to perceive the world like an animal with monocular vision (e.g. a pigeon), simply because of the position of our eyes.

Such a point of view can be simulated using cameras, software stitching and rendering, pacing the entire view of the environment within the FOV of the user. The most common way to visualize a 360 environment onto a plane is by using equirectangular projection [1]. Yet, we could choose from over 250 available projection methods with a history of over 2,000 years [30], mainly developed for cartography. To our knowledge, nobody has evaluated different projection methods to extend human vision. In our research, we evaluate the effects of our dynamic projections with various activation methods.

Towards this end, we propose OmniView, a software-based approach to optimize the visualization of single plane projections in 360 environments. OmniView is a continuation of our previous work PanoFlex, which was accepted as a 2-page poster in VRST2019 (non-archival). Unlike conventional equirectangular projection, OmniView decreases the distortion and widens the direction of interest (DOI), making objects or people of interest more visible and easier to be interacted with. This allows the user to more easily adapt to 360 vision because content from other directions is still within the periphery while the DOI is less distorted. We propose three methods to activate OmniView: gaze-based, pointing-based, and automatic. For our first study, we compare these methods with conventional equirectangular projection to evaluate its effect on spatial perception. For the second study, we evaluate the subjective usability and simulator sickness after using the prototype device for a period of time. Increasing the capability of humans with an added view of their entire surroundings, opens the door for many potential use cases that can be quite beneficial. We believe that creating an effective 360 vision for daily use to be one of the first steps towards augmenting human vision capabilities beyond the ordinary. The contribution of this research is threefold:

(1) We present OmniView, an adaptable solution for optimizing 360 FOV vision. Our method dynamically changes the distortion for the DOI, making it easier to look at while maintaining the 360 degree panorama peripheral vision.
(2) We explore the effects of OmniView on spatial perception, reaction time and simulator sickness with an in-depth comparison using a baseline (equirectangular projection) and variations of activation method (gaze, pointing, and automatic distortion). Overall, gaze-based OmniView is preferred by most users.
(3) Finally, we discuss design solutions for future wide FOV-enabled devices based on the obtained results.

2 RELATED WORK

In this section, we discuss literature related to spatial perception augmentation, FOV-based augmentation and the use of various projection methods that have been developed.

2.1 Non-visual Spatial Awareness Augmentation

Spatial awareness augmentation refers to the related work done on enabling users to gain sensory feedback from their surroundings without the help of vision. Most often, they are delivered by haptics. For example, SpiderSense [21] allows the user to sense the environment via proximity sensors and haptic actuators, which were placed around the body. Similarly, the haptic collar [26] conveys information to the user via vibrotactile motors, which were placed around the neck. ProximityHat [4] is another device that allows the user to sense nearby obstacles via ultrasonic sensors and pressure actuators. Heat-Nav [32] used thermal peltier elements, which were placed on the arm to successfully navigate users across a virtual maze. Sound is another method to deliver spatial information. Hindsight employs audio to help cyclists to stay aware of cars [27]. The system detects a nearby car and then delivers this information to the user via audio. These works use nonvisual sensory modalities to accommodate the limitations of our FOV and we find them relevant in understanding the effects of augmenting the user with different sensory perceptions.

2.2 Visual Augmentation

This subsection presents related works which instead focuses on visualizing cues within the FOV or on augmenting the human vision by increasing the FOV itself. For example, Gustafson et al. [15] and Gruenefeld et al.[13] both developed a method to view off-screen objects using a wedge and a head-mounted display (HMD) version of it, respectively. Jung et al. [17] also used a vehicle position estimation based on deep learning that allowed him to view approaching vehicles from the back within his FOV. For a hardware-based approach, Grunefeld et al. [14] used an LED ring attached around the HMD lens as an indicator for off-screen objects that is viewable on the periphery. Orlosky et al. [24] worked on an HMD that was able to provide an FOV of 238 using stereoscopic fisheye lenses. However, for 360 vision, one of the most related work to ours is FlyVIZ[1] developed by Ardouin et al., in which users are treated with an equirectangular projection that was able to show the full physical environment around them, since the camera renders the surroundings onto a flat plane in front of the user. This provided participants with one of the first experiences of actually having their entire periphery filled with content from all directions. Besides, there are also works like SpiderVision [11] and Unconstrained Neck, [29] which aimed to augment the human FOV. These works introduce several methods to view content beyond the typical FOV, yet only FlyVIZ actually allows for constant 360 viewing.

2.3 Projection Methods

One of the approaches to optimizing 360 vision is using different kinds of projection methods. Tran et al. [33] compared three projection formats with the forced-choice pairwise comparison method and found that cubemap projection and octahedron projection images performed better than equirectangular projection images. However as of this moment, equirectangular projection is still the most popular choice. Furthermore, Ardouin et al. [2] compared between perspective, equirectangular, hammer, Albers Equal Area Conic and azimuthal equidistant projections for virtual desktop environments and found that equirectangular, which is our baseline, performed better for object collection tasks. In the next study, the author then compared between perspectives with various forms of equirectangular and hammer projection that adapted to the geometry of the virtual environment [3]. Even though no user study was performed, the author showed that a 360 rendering did not significantly impact the frame rate. Debarba et al. [8] performed a comparative study between equirectangular, hammer, and perspective projection for virtual environments, and found that equirectangular and hammer projection outperformed the perspective projection in terms of search time. For projection optimization work, Rich360, proposed by Lee et al. [18], optimizes 360 video by changing the projection and shows that projection methods can be applied to improve the quality of 360 video. Finally, Boustila et al. [6] implemented a hybrid projection method that combines perspective (middle 71) with cylindrical projection (top and bottom 32) for a wider vertical viewing angle and increased perception of personal space for architectural viewing.

From the presented related work, we wish to emphasize that our research question is not "Can humans adapt to 360 vision?", but rather we look beyond, working on the question of "what is an appropriate 360 vision for humans?" Following this, OmniView is developed based on spherical-to-planar equirectangular projection and will be compared with equirectangular as a baseline in the study section. The user will always view the content as a 2D plane projected in front of them, as opposed to a VR spherical view for constant 360 viewing. Given our research question, all study conditions provide a specific type of 360 vision to the participants.

3 APPROACH AND IMPLEMENTATION

In terms of prototype appearance, our approach is similar to that of FlyVIZ [1]. The user wears an HMD with a helmet with a 360 camera mount. As mentioned earlier, to our knowledge this is the only known method that allows a user to view the surrounding environment in its entirety. The major differences to FlyVIZ are the projection distortion and the various ways to activate it.

3.1 Hardware

The computer used in our prototype is an Alienware Area-51m equipped with an Intel Core i7-9700K processor and an Nvidia RTX 2080 graphical processing unit. With these specifications, our system runs at around 20 frames per second. The computer is placed on a desk and connected to our prototype with long cables to ensure mobility. However, the mobility of the user is restricted by the tracking area of the VR headset (HTC Vive Eye Pro). To capture the panoramic image of the environment, we mount the Ricoh Theta V 360 camera on a helmet worn by the user.

Figure 2: Our prototype consists of HTC Vive Pro Eye and Ricoh Theta V mounted on helmet.

3.2 Software

The OmniView prototype was implemented using the Unity3D game engine. OmniView uses equirectangular projection. To create a custom equirectangular projection that is completely manipulatable, we mathematically create the projection by converting from the polar coordinate system to the Cartesian coordinate system. The virtual projection camera is placed in the middle of the 360 video from the Ricoh Theta V's video feed. From that camera the cubemap render is extracted to obtain renders from each of the 6 directions, each with a 90 FOV. Our render code then reads these textures with UV mapping, where the U and V components are normalized textures of the mapping starting from the lower left corner. θ refers to the azimuthal angle whereas ϕ refers to the inclination angles. The radius, r is equal to 1, because the sphere shares the same origin as the cubemap. r, θ and ϕ makes up the polar coordinates for the cubemap. In accordance to Unity's coordinate system (left hand coordinate system), we used the following formulas to convert these coordinates into a Cartesian coordinate system, where:

$$x = (\sin\phi * \sin\theta * r * -1) + x_1$$
$$y = (\cos\theta * r * -1) + y_1$$
$$z = (\cos\phi * \sin\theta * r * -1) + z_1$$

The obtained values of x, y and z are the coordinates of each pixel present within the cubemap, which was previously expressed as polar coordinates. We then added the modifiers x_1, y_1 and z_1 with a value range of -1 to 1 to tweak each distortion direction. Since each of these values represents an axis, a positive modifier would enlarge the distortion of one direction, and a negative modifier would enlarge the opposite direction. For example, the modifier $z1$ with a value of -1 enlarges the distortion for the back view, whereas a value of 1 increases the distortion for the front view. When one direction is enlarged, the distortion for that direction decreases (closer to regular perspective projection), making it easier to view that direction, and its distortion is compensated by other directions. The resulting distortion becomes a form of hybrid distortion similar to the work proposed by [6], which marries perspective with cylindrical projection in a non-VR environment. It also resembles

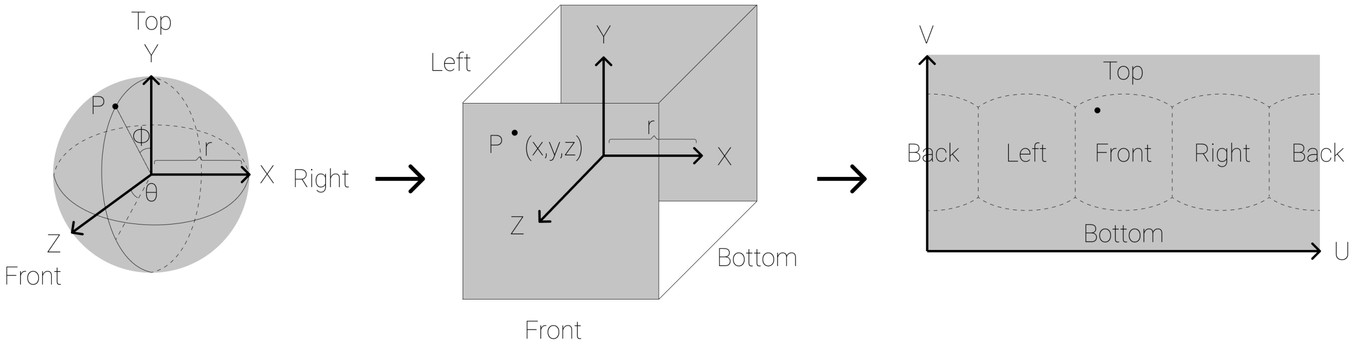

Figure 3: Conversion from polar coordinates to Cartesian coordinates for example point P.

the work by Sarkar et al. [25], who proposed fisheye viewing for graphs, in which the distortion changes depending on the graphical information while keeping both the local detail (the direction of interest) and the global context (surrounding periphery). The key difference is that our solution uses eye gaze, pointing and face detection to dynamically reduce distortion and increase viewing space of the selected point using an HMD and a 360 camera. Also, we apply it to 360 vision not only to camera images.

The three variations of OmniView that we developed are Gaze OmniView, Pointing OmniView and Automatic OmniView. Each of these versions mainly differs in the method to activate the distortion. They were implemented in the Unity 3D game engine.

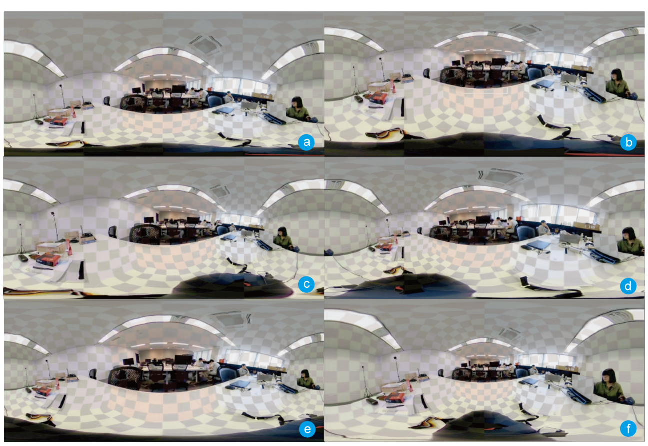

Figure 4: User's view while using OmniView for each direction. In the user studies, we set the distortion value of each direction to 0.3/-0.3 while leaving other parameters at 0 to enlarge (a)top (b)bottom (c)left (d)right (e)front (f)back.

3.2.1 Gaze OmniView. Gaze OmniView leverages the users' gaze point to initialize distortion. For example, activating the back distortion requires the user to look towards the far right or left in accordance with the distortion. Each user is required to first perform a short calibration to ensure accurate eye tracking.

From our pilot test, we found that making the distortion too sensitive to the user's saccades actually causes confusion and disorientation. Therefore, we implemented a dwell-based selection method. The user has to look at a particular direction for a fixed amount of time before distortion is initiated. Over informal tests with users, we selected a dwell period of 1 second and this dwell timing is used for the other two variations of OmniView as well in order to maintain consistency in activation. The benefit of Gaze OmniView is that it remains hands-free because it uses gaze-based activation.

3.2.2 Pointing OmniView. Pointing OmniView differs from Gaze OmniView in a key factor: the input method. The input method uses a more conventional VR controller for pointing towards the direction, in this case a Vive controller. Furthermore, unlike Gaze Omniview, in which the user looks at the projection directions to activate, Pointing OmniView allows the user to point towards the direction to activate the distortion for that particular direction. For example, activating the back distortion requires the user to point towards his/her back.

Even though this method loses the hands-free nature, we hypothesize that it might improve spatial perception because the user has to physically point the controller.

3.2.3 Automatic OmniView. For Automatic OmniView, we remove the control of the distortion from the user. Instead, we utilize OpenCV's DNN (Deep Neural Network) face detector module, which is based on the SSD (Single Shot MultiBox Detector) framework [19] with ResNet-10 model, to automatically detect a person that is within close proximity to the user. When the camera detects a person's face at a certain direction, that direction becomes the DOI. This triggers the change in distortion to prioritize the face direction.

This method removes control from the user and is designed to allow automatic distortion for conversation. However, at this point of time, it only supports one person at a time and only activates for a person's face instead of an object. Even though we can easily replace face detection with object detection, we choose to prioritize communication functions for now. Furthermore, object detection would highly depend on context, and objects to be detected need to be specified accordingly, otherwise false positive triggers could easily occur.

4 STUDY 1: EFFECT ON SPATIAL PERCEPTION AND REACTION TIME

The goal of this study is to evaluate the spatial perception and reaction time of the participant using our system compared with the conventional equirectangular view for 360 vision. We also gather feedback for induced simulator sickness using SSQ. We choose equirectangular vision as our baseline because 1) it is, to our knowledge, the only other projection method that has been used for full 360 vision [1], 2) it is the gold standard used for VR cinematic content by many companies like Google (Photos[1] and Maps[2]), Facebook[3] and also 360 cameras like Ricoh Theta[4] and Insta360[5] because it projects accurately to the inside of a sphere for VR viewing, and 3) it is a very popular choice for cartography and even used by World Wind Explorer developed by NASA which was based on the WorldWindJS package [28]. Other popular projection methods like Mercator [22], Winkel-Tripel [16], etc. are mainly well-known for cartography and equirectangular projection is still the most widely used method for VR-based applications [9, 10, 20, 23, 31].

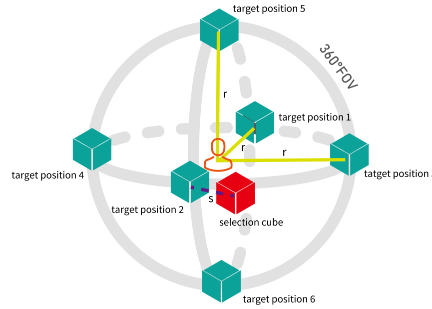

Figure 6: The positions of the target cubes in spatial perception study. r is the fixed distance between each cube and participant. s, which is the difference between the center points of what they see (target cube) and what they select (selection cube), is the distance we measure to evaluate the effect on spatial perception by using OmniView.

4.1 Apparatus

For this study, the VR HMD device used is the HTC Vive Pro Eye powered by an Alienware Area-51m mentioned previously. Since the study is simulated in VR, a 360 camera is not necessary for now.

4.2 Study Design

Based on the goal of this study, we wish to determine if there is any change in 1) spatial perception, 2) reaction time, and 3) induced simulator sickness between our OmniView variations with equirectangular projection. It is a modified version of the study employed by [5] but with a full 360 view around the user and it is essentially a search task.

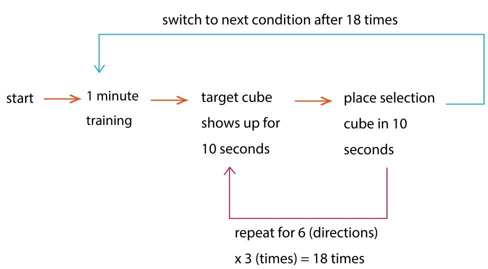

Figure 5: Flow Chart for Spatial Perception Study

When the participant puts on the HMD, they first see the projection in front of them as a plane (all conditions are the same). There is also a small avatar in front of them with a semi-transparent sphere around it representing the space around the participant. When a target object appears behind the participant for example, they will be able to see it on the planar information. In this example, the target object will be rendered either at the far left or right of the projection, which shows the view of the participant's back. Since this entire study runs in a virtual environment, the automatic condition becomes not viable because it depends on face detection. Therefore, for the purpose of this study, we modified it to distort the DOI depending on the direction of where the target cube appears (other conditions were not modified). After a period of time, the target object will disappear. A selection object will appear. This object can only move on the surface of the semi-transparent sphere. The participant is required to move the selection object to where they think the target object's position was. In this example, they need to place the selection object to the back of the avatar by the end of the given time. The task is repeated for each direction. Throughout the study, we measure the difference in distance between the center point of the target object and the projected center point of the selection object and compute that as the perceived error. We also log the reaction time, which is the time required by the participant to find where they think the selection cube should be. After each condition, they answer the SSQ questionnaire.

[1] https://www.google.com/photos/about/
[2] https://www.google.com/maps
[3] https://www.facebook.com/
[4] https://theta360.com/uk/
[5] https://www.insta360.com/

4.3 Participants

We recruited a total of 16 participants (9 male, 7 female, mean:30.56, SD:7.61). The participants first answer the SSQ questionnaire so we can establish the baseline values. The participant views the environment using 4 versions of the system depending on the condition (The order is based on Latin Square to eliminate possible ordering effects).

4.4 Procedure

Participants first fill a consent form that states that the collected data will be preserved and stored securely, and that they can freely discontinue the experiment at any time. They then proceed to put on the HMD, which is when the experimenter explains the full procedure. As described earlier, the participant sees the full 360 projection of their virtual position on a plane in the form of either an equirectangular projection or one of the OmniView versions. An avatar with a semi-transparent sphere around it is visible in front of the user. For better understanding of the procedure, we divide this study into the training phase, target phase, and selection phase.

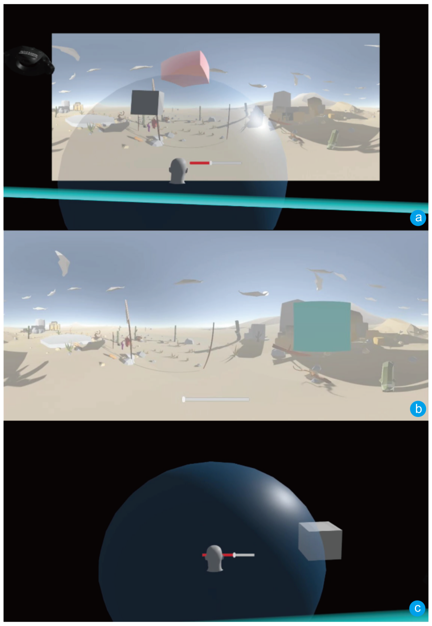

Figure 7: Procedure of spatial perception study is divided into 3 phases:(a)training phase, (b)target phase and (c)selection phase.

During the training phase, which is shown in Figure 7(a), we make the selection object's transformation equal to that of the target object, where both are always visible. The participant can freely manipulate the selection object which is constrained to the surface of the semi-transparent sphere. When they move it, the target cube around them also moves and is visible in the planar projection view. This allows them to understand how the space around the avatar works in relation to what they see on the planar projection. The training phase lasts for 1 minute.

During the target phase that is shown in Figure 7(b), the target object will appear around the user at a fixed distance from them and projected on the planar projection (depending on the condition). There are 6 possible directions for the object to spawn (front, back, left, right, up and bottom). Only one object appears at a time. Once the participant completes a round (finishes the selection phase), another will spawn at a different position, with each position repeating 3 times. The target cube will only be visible for 10 seconds. During this period, the participant is simply required to memorize the position of the target object and figure out where they think that object is relative to them.

After the target object disappears, the selection phase that is shown in Figure 7(c) begins. The participant is required to place the selection object, which is constrained on the semi-transparent sphere's surface, in order to indicate where they think the target object is relative to them. Within 10 seconds, the user may freely adjust the position of the cube until they are satisfied. When the selection cube's position is finalized, they hit a button on the controller to inform the system that they have completed the selection, allowing us to log the reaction time. After that, the target phase begins again for the next position, until the task is completed.

When a condition is completed, the participant is required to answer the SSQ and NASA TLX questionnaire. He/She will then rest for 2 minutes before redoing the experiment for the next condition. Each participant will undergo a total of 4 conditions x 6 positions x 3 repeats = 72 runs. The experiments lasts for about 30 minutes.

4.5 Results and Discussion

Since this is an exploratory study on optimising 360 vision (as opposed to introducing one like FlyViz [1]), our primary concern is to find if OmniView does or does not have any significant effect on spatial perception and reaction time.

We first look at the SSQ results, where there was a statistically significant difference in induced disorientation, $x^2(2) = 10.7, p < 0.05$. However, post-hoc analysis with Wilcoxon signed-rank tests with Bonferroni correction for adjusted significance showed no statistical significance between the conditions. This leads us to believe that, for a controlled study like this where only virtual content is used, participants do not perceive any noticeable increase in nausea, oculomotor problems or disorientation. The next study will look into SSQ results for when OmniView is used in a physical environment instead.

Secondly, to analyze the results of distance between target cube and selection cube, as well as the reaction time of the participants, we ran Shapiro-Wilk test and Kolmogorov-Smirnov test for normality for each direction. As a result, all directions showed normality ($p > 0.05$). Therefore, we performed repeated measures ANOVA for all directions on distance and reaction time. After running ANOVA, p-value of all directions on distance and reaction time are > 0.05, which means there is no statistical significance between equirectangular projection and OmniView.

Looking at the NASA TLX score for each method, there was no statistical significance, suggesting that each of the implementations did not have any higher perceived task load over another. The highest score on average would be the pointing method, with a TLX score of 3.24. Even so, the score is quite low, leading us to believe that the human vision has no issues in perceiving any additional load when it comes to both 360 FOV and dynamic distortion. This could be due to how the human vision system functions based on optical flow [12]. Our current implementation distorts the DOI based on interpolation over time; if it changes to a new distortion instantly, this could potentially lead to higher perceived load, though this is currently not within the scope of this study.

For the next analysis, we look into the correlation between reaction time and spatial distance error across all conditions. We computed the Pearson's rank order correlation coefficient, with the tabled results shown in Figure 11. In this figure, we are mainly interested in the relationship between the dependent variables. There was a strong positive correlation between each of the conditions' reaction time with its spatial distance error. The automatic condition has a value of $r_s = .254, p < 0.05$, the pointing condition has a value of $r_s = .218, p < 0.05$, the equirectangular condition has a value of $r_s = .384, p < 0.01$, and the gaze condition has a value of $r_s = .261, p < 0.05$. Only the p value for the equirectangular condition was below 0.01, suggesting that the longer it took to react to an object, the bigger the spatial distance error, and this is most

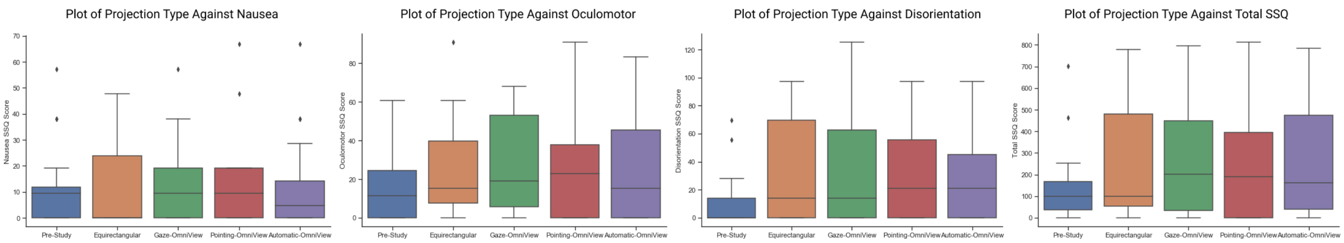

Figure 8: SSQ Results for Study 1

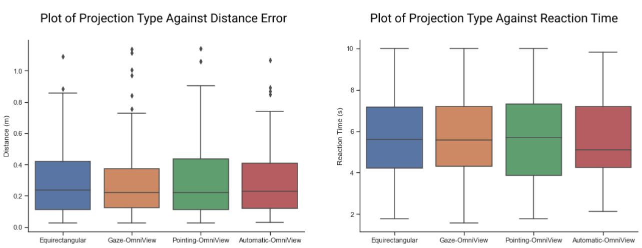

Figure 9: Results for Spatial Perception Error and Reaction Time

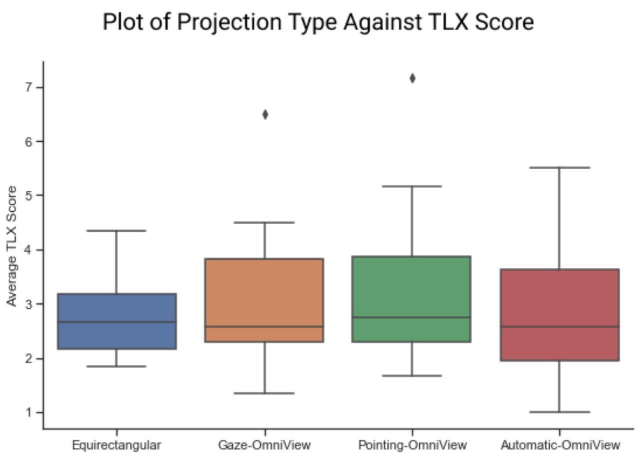

Figure 10: Results for NASA TLX

	auto-dist	auto-react	point-dist	point-react	equirect-dist	equirect-react	gaze-dist	gaze-react
auto-dist	1							
auto-react	.254*	1						
point-dist	.670**	0.032	1					
point-react	.244*	.633**	.218*	1				
equirect-dist	.639**	0.18	.596**	.370**	1			
equirect-react	.282**	.658**	0.109	.670**	.384**	1		
gaze-dist	.666**	0.132	.757**	.297**	.631**	0.162	1	
gaze-react	0.145	.589**	0.157	.757**	.329**	.673**	.261*	1
* Correlation is significant at the 0.05 level (2-tailed).								
** Correlation is significant at the 0.01 level (2-tailed).								

Figure 11: Correlation Matrix for Reaction Time and Spatial Perception Error

prevalent using equirectangular. This could be because equirectangular was the only condition that did not adjust its view to suit the DOI. So when participants took longer time to locate an object, they found it more difficult to place it correctly, possibly because the DOI is either small or too distorted.

	tlx_auto	ssq_auto	tlx_equi	ssq_equi	tlx_gaze	ssq_gaze	tlx_point	ssq_point
tlx_auto	1							
ssq_auto	.676**	1						
tlx_equi	.732**	0.465	1					
ssq_equi	0.483	.808**	.548*	1				
tlx_gaze	.693**	0.464	.880**	.618*	1			
ssq_gaze	0.352	.744**	.512*	.899**	.535*	1		
tlx_point	.637**	.533*	.831**	.726**	.892**	.652**	1	
ssq_point	0.378	.669**	0.492	.868**	.621*	.880**	.759**	1
** Correlation is significant at the 0.01 level (2-tailed).								
* Correlation is significant at the 0.05 level (2-tailed).								

Figure 12: Correlation Matrix for total SSQ and NASA TLX

Finally, we look into the correlation between the SSQ score and average perceived task load. We separated this table because there were no significant correlations between the reaction time and spatial distance error. We ran the Spearman's rank order correlation coefficient with the tabled results shown in Figure 12. Again, we are only interested in the relationship between the dependent variables. Overall, there was a statistically significant positive correlation between the TLX and SSQ for all conditions. The pointing method ($r_s = .759, p < 0.01$) and automatic method ($r_s = .676, p < 0.01$) have the highest correlation coefficient, followed by the equirectangular ($r_s = .548, p < 0.05$) and gaze method ($r_s = .535, p < 0.05$). Despite each method exhibiting correlation between SSQ and TLX, only the pointing and automatic methods had a p value below 0.01. The Pointing OmniView differs in that it is not hands-free and its pointing position depends on the physical space as opposed to the planar projection. This involves more physical movement from the participant and possibly more cognitive load, which in turn could lead to strong correlation between simulator sickness and perceived task load. The automatic method, despite requiring less movement, relies completely on the system to decide the DOI. The lack of control given to the participant may effect both the perceived load and sickness simultaneously.

Overall, this means that when compared with equirectangular projection as baseline, there was no significant effect towards the spatial perception and reaction time when performing a search task. Unlike equirectangular projection, OmniView relies on a constantly shifting distortion that, while it was meant to make the DOI easier, could potentially also negatively effect spatial perception, which would then defeat the purpose of having a wider FOV. Nevertheless, we interpret these findings as a way for OmniView to further push the boundaries of dynamic distortion methods.

5 STUDY 2: EFFECTS WHEN PERFORMING EVERYDAY TASKS

For this next study, we look into the possible effects of OmniView on simulator sickness when used in physical space and performing simple everyday tasks. This is also of interest in order to understand humans' compatibility with an enhanced vision system. Debarba et al. [8], who conducted a simulator sickness comparative study between hammer, equirectangular and perspective projection for a non-VR environment, found that there was no significant differences between them. Motivated by these results, we wish to ascertain this for OmniView as well when comparing it to something more conventional like equirectangular. This study is not a long-term exposure but rather an initial exploratory study towards understanding OmniView's effects when using the system in the physical environment, instead of just being simulated in VR.

5.1 Apparatus

The apparatus for this study is similar to the previous study but with the addition of the 360 helmet. The helmet is an ordinary bicycle helmet, on top of which we mount the Ricoh Theta V for streaming real-time 360 video to Unity. The participant then sees the projected content of the camera through the Vive Pro, where we manipulate the experimental conditions.

5.2 Study Design

Participants are required to perform simple everyday tasks while using the full prototype for a short period of time, after which we evaluate simulator sickness based on SSQ. The tasks are performed back-to-back and the order of the task is based on Latin Square. Each task is just 2 minutes long as opposed to long term exposure because 360 vision is relatively new to most people and the purpose of this study is to see if there is any onset of the aforementioned simulator sickness. The results from this experiment will let us know if long term exposure in the real world is a viable future study for any of the tested conditions.

5.3 Participants

For this study, we used the same 16 participants (9 males, 7 females, mean:30.56, SD:7.61) as the spatial perception study. The study was conducted about 30 minutes after the previous study.

5.4 Procedure

The experiment is conducted in a room with minimal distraction. The participant is first required to wear the prototype and perform three activities (2 minutes each) for a period of 6 minutes. Each of these activities are selected specifically to not only reflect simple daily tasks, but also allow us to ascertain if a specific task favours specific implementations. The first activity is talking while sitting. For this activity, the participant is simply required to sit down and have a conversation with one of the experimenters. During the conversation, the experimenter will move around the participant, while the participant remains static and seated. This requires the participant to change the DOI activation while conversing. The second activity is reading a book on a tablet device while sitting.

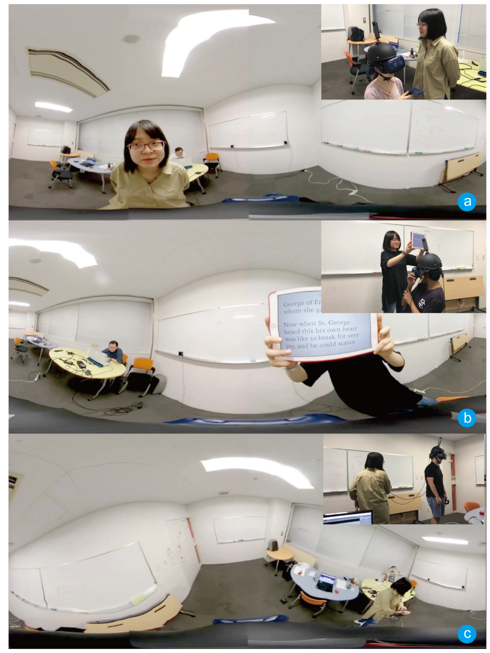

Figure 13: 3 simple everyday tasks in study 2: (a)talking while seated, (b)reading and (c)walking to and fro.

This task tests the ability to focus on small texts within a 360 projection with minimal motion. We wish to ascertain if a small focus point within the projection also causes any discomfort. Finally, the last activity is walking to and fro. This final task introduces body motion when they must walk 2 meters to the front and to the back within the range of the HTC Vive's tracking zone. This task introduces optical flow in the projection for all directions, including the DOI. Each participant performs these activities back-to-back without pause. They also perform these activities for each of the conditions (equirectangular, Gaze OmniView, Pointing OmniView and Automatic OmniView) with a rest period of 2 minutes in between. The SSQ questionnaire is answered before any activity is performed and after each condition.

5.5 Results and Discussion

Looking at the SSQ results, there was a statistically significant difference in induced disorientation, $x^2(2) = 13.03, p < 0.05$. Post hoc analysis with Wilcoxon signed-rank tests was conducted with a Bonferroni correction applied, resulting in a significance level set at p < 0.01. Median (IQR) disorientation score for before the study, after using the automatic, equirectangular, gaze and pointing method was 0 (0 to 13.92), 55.68 (0 to 93.96), 48.72 (0 to 97.44), 48.72 (0 to 97.44) and 48.72 (3.48 to 83.52), respectively. There was a statistically significant rise in disorientation from the pre study to after the use of the automatic ($Z = -2.671 p < 0.01$) and pointing ($Z = -2.944, p < 0.01$) methods. We believe that this is due to how OmniView works, during which a spherical-to-planar projection is shown to the participant at all times, allowing them to actually see all directions projected onto a plane. On the other hand, gaze-based methods allow direct activation of the plane and looking at any

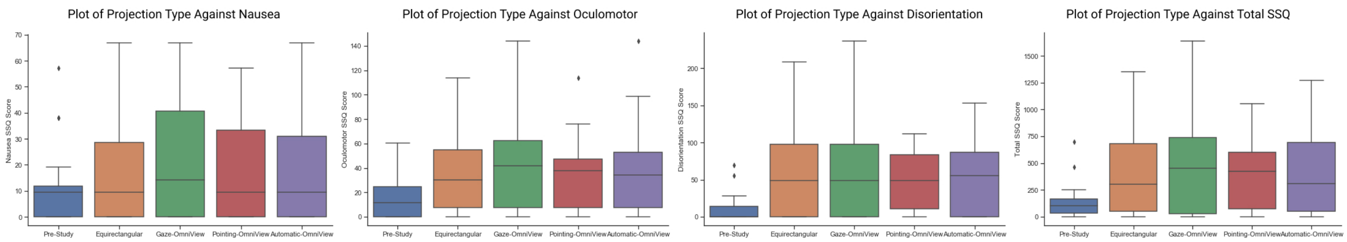

Figure 14: SSQ Results for Study 2

direction triggers that DOI directly. Pointing does not work based on the planar information, but rather the actual space of the user, which may possibly lead to confusion. Automatic on the other hand, does not give any control to the user, making spatial information the hardest to obtain (most significant increase in disorientation).

These results were aligned with the significant correlation of SSQ with TLX that we found from study 1. Even though study 1 did not induce any significant rise in perceived sickness and task load, this could be largely due to the environment being in VR and the task not actually requiring much physical movement. Since the study is streamed from a 360 camera with limited resolution, which increases the disparity with their actual vision, this could also overall contribute to stronger sickness.

6 DESIGN IMPLICATION

For a fair comparison, we did not compare them with normal FOV such as VR spherical view, but rather with equirectangular projection and found that OmniView did not effect the user's spatial perception and reaction time significantly. However, the pointing and automatic OmniView induced significant disorientation, therefore we recommend gaze-based OmniView as the preferred mode for distortion change. These results allow us to push the boundaries of human vision even further, to not only accommodate wide FOV, but also use different interactive methods to manipulate the vision.

Even though we did not find any significant positive effects of OmniView over conventional equirectangular, the exploratory studies provided us with several useful insights on pushing the boundaries of 360 vision augmentation. It is important that we established that OmniView also had no significant negative impact. The nature of OmniView projection is that content from every direction is directly viewable by the user at all times. Thus, viewing the back would mean looking at the far right or left. Therefore, we find it interesting that even though pointing towards a direction to activate the DOI should be less confusing and disorienting, our results showed the opposite, in that using gaze on a mapped plane was instead less disorienting. We also think activation by gaze is simply deemed more natural because the human foveal vision naturally has the highest acuity at the gaze point, with the periphery being blurrier. Gaze OmniView acts the same way as human foveal vision, in that only the gaze point with highest acuity has the least distortion. The periphery with lower acuity is however sensitive to motion, and OmniView nevertheless shows information from the remaining directions in the periphery.

Below are the following design guidelines for future 360 vision optimization:

(1) Dynamic distortion solutions like OmniView did not present any negative effects concerning spatial perception and reaction time. More extreme distortion parameters can be further investigated.
(2) Pointing in OmniView causes significant disorientation and is the least preferable, whereas gaze-based tracking was on average the superior choice, though further studies are needed to validate this.

7 CONCLUSION AND FUTURE WORKS

In this work, we propose an exploratory study on optimizing 360 FOV towards human vision augmentation. OmniView is an approach to alter distortion parameters to prioritize the direction of interet (DOI). We introduced three variations of OmniView: activating distortion based on eye gaze, on pointing by hand, or automatic activation based on face detection.

For future works, we will next look at long-term studies and evaluation using OmniView. Furthermore, we also plan to integrate OmniView with other popular spherical-to-planar projections like Mercator, Winkel Triple, etc., allowing us to compare between wider options for OmniView.

8 ACKNOWLEDGMENTS

This project was a collaboration between Keio Graduate School of Media Design and Ignition Point, Inc.

REFERENCES

[1] Jérôme Ardouin, Anatole Lécuyer, Maud Marchal, Clément Riant, and Eric Marchand. 2012. FlyVIZ: A Novel Display Device to Provide Humans with 360 Vision by Coupling Catadioptric Camera with Hmd. In *Proceedings of the 18th ACM Symposium on Virtual Reality Software and Technology (VRST '12)*. ACM, New York, NY, USA, 41–44. https://doi.org/10.1145/2407336.2407344
[2] J. Ardouin, A. Lécuyer, M. Marchal, and E. Marchand. 2013. Navigating in virtual environments with 360° omnidirectional rendering. In *2013 IEEE Symposium on 3D User Interfaces (3DUI)*. 95–98. https://doi.org/10.1109/3DUI.2013.6550203
[3] J. Ardouin, A. Lécuyer, M. Marchal, and E. Marchand. 2014. Stereoscopic rendering of virtual environments with wide Field-of-Views up to 360°. In *2014 IEEE Virtual Reality (VR)*. 3–8. https://doi.org/10.1109/VR.2014.6802042
[4] Matthias Berning, Florian Braun, Till Riedel, and Michael Beigl. 2015. ProximityHat: A Head-worn System for Subtle Sensory Augmentation with Tactile Stimulation. In *Proceedings of the 2015 ACM International Symposium on Wearable Computers (ISWC '15)*. ACM, New York, NY, USA, 31–38. https://doi.org/10.1145/2802083.2802088
[5] Wutthigrai Boonsuk, Stephen Gilbert, and Jonathan Kelly. 2012. The Impact of Three Interfaces for 360-degree Video on Spatial Cognition. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '12)*. ACM, New York, NY, USA, 2579–2588. https://doi.org/10.1145/2207676.2208647
[6] S. Boustila, A. Capobianco, O. Génevaux, and D. Bechmann. 2016. New hybrid projection to widen the vertical field of view with large screen to improve the perception of personal space in architectural project review. In *2016 IEEE Virtual Reality (VR)*. 159–160. https://doi.org/10.1109/VR.2016.7504702

[7] Thomas W Cronin. 2005. The visual ecology of predator-prey interactions. *Ecology of predator-prey interactions* (2005), 105–138.

[8] Henrique G. Debarba, Sami Perrin, Bruno Herbelin, and Ronan Boulic. 2015. Embodied Interaction Using Non-planar Projections in Immersive Virtual Reality. In *Proceedings of the 21st ACM Symposium on Virtual Reality Software and Technology (VRST '15)*. ACM, New York, NY, USA, 125–128. https://doi.org/10.1145/2821592.2821603

[9] Tarek El-Ganainy and Mohamed Hefeeda. 2016. Streaming virtual reality content. *arXiv preprint arXiv:1612.08350* (2016).

[10] Ching-Ling Fan, Jean Lee, Wen-Chih Lo, Chun-Ying Huang, Kuan-Ta Chen, and Cheng-Hsin Hsu. 2017. Fixation prediction for 360 video streaming in head-mounted virtual reality. In *Proceedings of the 27th Workshop on Network and Operating Systems Support for Digital Audio and Video*. ACM, 67–72.

[11] Kevin Fan, Jochen Huber, Suranga Nanayakkara, and Masahiko Inami. 2014. SpiderVision: Extending the Human Field of View for Augmented Awareness. In *Proceedings of the 5th Augmented Human International Conference (AH '14)*. ACM, New York, NY, USA, Article 49, 8 pages. https://doi.org/10.1145/2582051.2582100

[12] John V. Forrester, Andrew D. Dick, Paul G. McMenamin, Fiona Roberts, and Eric Pearlman. 2016. Chapter 5 - Physiology of vision and the visual system. In *The Eye (Fourth Edition)* (fourth edition ed.), John V. Forrester, Andrew D. Dick, Paul G. McMenamin, Fiona Roberts, and Eric Pearlman (Eds.). W.B. Saunders, 269 – 337.e2. https://doi.org/10.1016/B978-0-7020-5554-6.00005-8

[13] Uwe Gruenefeld, Abdallah El Ali, Susanne Boll, and Wilko Heuten. 2018. Beyond Halo and Wedge: Visualizing Out-of-view Objects on Head-mounted Virtual and Augmented Reality Devices. In *Proceedings of the 20th International Conference on Human-Computer Interaction with Mobile Devices and Services (MobileHCI '18)*. ACM, New York, NY, USA, Article 40, 11 pages. https://doi.org/10.1145/3229434.3229438

[14] Uwe Gruenefeld, Tim Claudius Stratmann, Lars Prädel, and Wilko Heuten. 2018. MonoculAR: A Radial Light Display to Point Towards Out-of-view Objects on Augmented Reality Devices. In *Proceedings of the 20th International Conference on Human-Computer Interaction with Mobile Devices and Services Adjunct (MobileHCI '18)*. ACM, New York, NY, USA, 16–22. https://doi.org/10.1145/3236112.3236115

[15] Sean Gustafson, Patrick Baudisch, Carl Gutwin, and Pourang Irani. 2008. Wedge: Clutter-free Visualization of Off-screen Locations. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '08)*. ACM, New York, NY, USA, 787–796. https://doi.org/10.1145/1357054.1357179

[16] Bernhard Jenny, Tom Patterson, and Lorenz Hurni. 2008. Flex Projector–interactive software for designing world map projections. *Cartographic Perspectives* 59 (2008), 12–27.

[17] Jinki Jung, Hyeopwoo Lee, Jeehye Choi, Abhilasha Nanda, Uwe Gruenefeld, Tim Stratmann, and Wilko Heuten. 2018. Ensuring Safety in Augmented Reality from Trade-off Between Immersion and Situation Awareness. In *2018 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. IEEE, 70–79.

[18] Jungjin Lee, Bumki Kim, Kyehyun Kim, Younghui Kim, and Junyong Noh. 2016. Rich360: Optimized Spherical Representation from Structured Panoramic Camera Arrays. *ACM Trans. Graph.* 35, 4, Article 63 (July 2016), 11 pages. https://doi.org/10.1145/2897824.2925983

[19] Wei Liu, Dragomir Anguelov, Dumitru Erhan, Christian Szegedy, Scott E. Reed, Cheng-Yang Fu, and Alexander C. Berg. 2015. SSD: Single Shot MultiBox Detector. *CoRR* abs/1512.02325 (2015). arXiv:1512.02325 http://arxiv.org/abs/1512.02325

[20] Wen-Chih Lo, Ching-Ling Fan, Jean Lee, Chun-Ying Huang, Kuan-Ta Chen, and Cheng-Hsin Hsu. 2017. 360 video viewing dataset in head-mounted virtual reality. In *Proceedings of the 8th ACM on Multimedia Systems Conference*. ACM, 211–216.

[21] Victor Mateevitsi, Brad Haggadone, Jason Leigh, Brian Kunzer, and Robert V Kenyon. 2013. Sensing the environment through SpiderSense. In *Proceedings of the 4th augmented human international conference*. ACM, 51–57.

[22] Mark Monmonier. 2010. *Rhumb lines and map wars: A social history of the Mercator projection*. University of Chicago Press.

[23] Akira Ohashi, Yuki Tanaka, Gakuto Masuyama, Kazunori Umeda, Daisuke Fukuda, Takehito Ogata, Tatsurou Narita, Shuzo Kaneko, Yoshitaka Uchida, and Kota Irie. 2016. Fisheye stereo camera using equirectangular images. In *2016 11th France-Japan & 9th Europe-Asia Congress on Mechatronics (MECATRONICS)/17th International Conference on Research and Education in Mechatronics (REM)*. IEEE, 284–289.

[24] Jason Orlosky, Qifan Wu, Kiyoshi Kiyokawa, Haruo Takemura, and Christian Nitschke. 2014. Fisheye vision: peripheral spatial compression for improved field of view in head mounted displays. In *Proceedings of the 2nd ACM symposium on Spatial user interaction*. ACM, 54–61.

[25] Manojit Sarkar and Marc H Brown. 1992. Graphical fisheye views of graphs. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. ACM, 83–91.

[26] Stefanie Schaack, George Chernyshov, Kirill Ragozin, Benjamin Tag, Roshan Peiris, and Kai Kunze. 2019. Haptic Collar: Vibrotactile Feedback Around the Neck for Guidance Applications. In *Proceedings of the 10th Augmented Human International Conference 2019 (AH2019)*. ACM, New York, NY, USA, Article 12, 4 pages. https://doi.org/10.1145/3311823.3311840

[27] Eldon Schoop, James Smith, and Bjoern Hartmann. 2018. HindSight: Enhancing Spatial Awareness by Sonifying Detected Objects in Real-Time 360-Degree Video. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18)*. ACM, New York, NY, USA, Article 143, 12 pages. https://doi.org/10.1145/3173574.3173717

[28] Bruce Schubert. [n. d.]. https://worldwind.earth/explorer/

[29] Lichao Shen, Mhd Yamen Saraji, Kai Kunze, and Kouta Minamizawa. 2018. Unconstrained Neck: Omnidirectional Observation from an Extra Robotic Neck. In *Proceedings of the 9th Augmented Human International Conference (AH '18)*. ACM, New York, NY, USA, Article 38, 2 pages. https://doi.org/10.1145/3174910.3174955

[30] John P Snyder. 1997. *Flattening the earth: two thousand years of map projections*. University of Chicago Press.

[31] Kashyap Kammachi Sreedhar, Alireza Aminlou, Miska M Hannuksela, and Moncef Gabbouj. 2016. Viewport-adaptive encoding and streaming of 360-degree video for virtual reality applications. In *2016 IEEE International Symposium on Multimedia (ISM)*. IEEE, 583–586.

[32] Jordan Tewell, Jon Bird, and George R. Buchanan. 2017. Heat-Nav: Using Temperature Changes As Navigation Cues. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '17)*. ACM, New York, NY, USA, 1131–1135. https://doi.org/10.1145/3025453.3025965

[33] Tran Thi Hai Uyen, Seungcheol Choi, and Oh-Jin Kwon. 2019. Subjective Evaluation of the 360-Degree Projection Formats Using Absolute Category Rating. In *Proceedings of the 2019 4th International Conference on Intelligent Information Technology (ICIIT '19)*. ACM, New York, NY, USA, 10–15. https://doi.org/10.1145/3321454.3321471