Panning and Zooming High-Resolution Panoramas in Virtual Reality Devices

Huiwen Chang
Facebook and Princeton Univeristy

Michael F. Cohen
Facebook

Figure 1: We explore four modes for panning high-resolution panoramas in virtual reality devices. Left to right: normal mode, circle mode, transparent circle mode, and zoom circle mode. The latter two provided equally successful interfaces.

ABSTRACT

Two recent innovations in immersive media include the ability to capture very high resolution panoramic imagery, and the rise of consumer level heads-up displays for virtual reality. Unfortunately, zooming to examine the high resolution in VR breaks the basic contract with the user, that the FOV of the visual field matches the FOV of the imagery. In this paper, we study methods to overcome this restriction to allow high resolution panoramic imagery to be able to be explored in VR.

We introduce and test new interface modalities for exploring high resolution panoramic imagery in VR. In particular, we demonstrate that limiting the visual FOV of the zoomed in imagery to the central portion of the visual field, and modulating the transparency or zoom level of the imagery during rapid panning, reduce simulator sickness and help with targeting tasks.

ACM Classification Keywords

H.5.m. Information Interfaces and Presentation (e.g. HCI): Miscellaneous

Author Keywords

virtual reality; gigapixel resolution.

INTRODUCTION

Over the past decade, we have seen new abilities to create and explore very high resolution panoramic imagery[17]. New robotic image acquisition technology have made high-resolution panoramic images easier to acquire. Support for online exploration of the results at sites such as Gigapan [32], allow anyone with a browser to zoom in to observe incredibly clear details of nature or cities. Artists are creating very large format panoramas combining high resolution imagery with AI-based Deep Dream to create new artistic and impressionistic worlds [1]. Scientists use them to examine high resolution microscopic scans and for environmental examination [11]. All interactive mapping applications are essentially pan/zoom interfaces on very large images. The viewer's ability to explore the panoramic imagery transforms the once static relationship between the viewer and the subject. The viewer is no longer a passive observer, but now determines which parts and what details of the image to view by panning and zooming the imagery.

More recently, the development of consumer level heads-up Virtual Reality (VR) devices has created interest in a new generation of user interaction and user experiences. Viewing full spherical is a natural medium to be explored in VR. Very high resolution imagery has the potential to open up many new applications ranging from embedding surprises in the details for treasure hunt style games, to embedding stories in cityscapes, to simply adding informational textual or auditory annotations at various levels of detail as in [18]. Video and photo sharing communities from social network such as Facebook [9] and Google+ [12] now support interfaces for viewing 360-degree content in various VR devices. Unfortunately, the paradigm of reproducing real-world viewing interaction by mapping head motion to panning in such a setting has precluded zooming, and thus one cannot examine high-resolution panoramas in VR. This paper presents a system to overcome this restriction.

Panning-and-zooming is a fundamental operation for image viewer applications on flat devices. This includes high resolution photographic imagery as well as mapping applications. Figure 2 shows an overview and two examples of zoomed-in details of the panorama. Research work [17, 18] improves such interaction and exploration for high-resolution panoramas on flat displays.

This raises the question of why not just do the same thing in VR. Simply allow zooming into the imagery, and leave panning as it is where a 10 degree head turn corresponds to

UIST 2017, October 22–25, 2017, QuÃl'bec City, QC, Canada.

ACM ISBN 978-1-4503-4981-9/17/10 ...$15.00.
http://dx.doi.org/10.1145/3126594.3126617

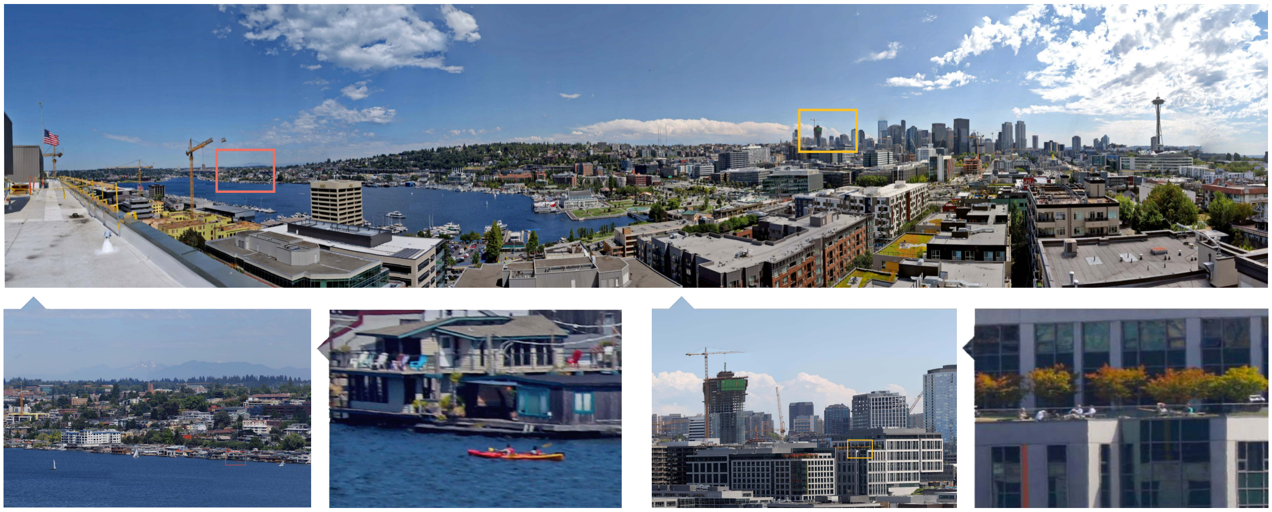

Figure 2: Two zoomed-in views for the high-resolution panorama.

a 10 degree pan on the imagery? The problem is that "visual motion", i.e., the image flow for a given head turn is much faster when zoomed in. Let's say the VR device displays imagery across a 90 degree field-of view, and the portion of the panorama displayed matches that 90 degrees. Panning 30 degrees would translate the imagery roughly one third of the way across the view. If one is zoomed into the imagery such that the field-of-view of the displayed panorama is only 10 degrees, that same 30 degree head rotation would induce a flow of three times the displayed imagery's width. This may lead users to severe simulation sickness which we will discuss in more details in Section 3. For this reason, interacting with high-resolution panoramas in VR devices is an unsolved problem.

There is a simple solution to the visual motion problem that is common on flat devices. Why not just slow down the panning such that the visual motion matches the head turn? In other words, in the example above, a 30 degree head turn would always induce the same visual motion of approximately one third of the screen width. In fact, this works reasonably well, until you consider the fact that you would have to spin around nine times to circumnavigate a full 360 degree panorama. This may or may not be OK, but there is no such luxury for looking up and down. You cannot tilt you head up beyond 90 degrees. Thus you could only examine a narrow vertical swath around the horizon.

In this paper, we demonstrate and explore four different modes for displaying high-resolution panoramic imagery in VR (see Figure 1 and the accompanying video). The first mode is exactly what is described above, where the visual motion increase with zoom level. This is our control condition to try to improve upon. It has been shown that motion sickness is exacerbated by motion on the view periphery. The next three modes thus limit the zoomed in region to the center of the visual field. We discuss the variations between these modes in detail later and discuss reactions to these novel interfaces in a user test. Both qualitative measures and questionnaire data were gathered to assess how comfortable and intuitive the participants found the interfaces.

To our knowledge, this is the first paper to address the problem of interacting with high-resolution panoramas within VR platforms. The results of our experiments demonstrate quite clearly that our interfaces overcome the basic limitations of viewing high-resolution panoramic imagery in VR for most participants.

RELATED WORK

This work lies at the intersection of graphics and HCI, including panorama viewers, multi-scale image interfaces, as well as issues such as simulator sickness especially in virtual environments. We will discuss relevant work in each area.

Panorama Viewers. Panorama viewers have been around since early efforts such as QuickTime VR [6] and Plenoptic Modeling [20] which introduced the creation and viewing interface for standard-resolution panoramic images. Such imagery now resides in online panorama browsing systems such as Google Street View [30] and Bing Maps Streetside [16] that enable users to examine 360 panoramas from all around the world. Wide angle (up to 360x180 degree) panoramas are typically represented and stored in cylindrical or spherical coordinates. A subset of the imagery is then reprojected into a perspective projection at viewing time. Changing the direction of the view results in a different subset of the imagery being displayed. This motion across the panorama is typically called panning.

Multiscale Interface for High-Resolution Panoramas.

More recently, it has become possible to acquire very high-resolution panoramas typically through capturing and stitching multiple individual images. This led to zoomable viewers to allow users to pan and zoom over the high-resolution panoramas to appreciate all the detail.

Kopf et al. [17] propose a system to capture and view wide angle Gigapixel images that smoothly adapts the projection, particularly when zoomed out to depict very wide angle fields-of-view. Further assistance for exploring large imagery containing audio annotations is presented by Luan et al. [18]. Ip et al. [13] increases the efficiency of exploration by helping find the salient areas in such large images. These works all focus on panorama viewing on flat displays.

The idea of applying a "digital lens" to viewing 2D media can be traced back to the Magic Lens work of Bier et al[4]. This work has been extended in many ways for flat (monitor) displays, with many designs for interactive zooming interfaces. Previous work[5, 3, 7, 25, 2] proposes and investigates overview + detail and focus + context interaction techniques to navigate in multi-scale worlds. One work, Pietriga et al. [25] reach a similar conclusion that translucence could help achieve efficient transitions. Our work extends this work in the context of VR which raises significant new challenges such as motion sickness.

New touch sensitive devices plus orientation sensing such as gyros has taken the interaction beyond click-and-drag interfaces. The most common interaction for zooming on touch devices has become pinching. Malacria et al. [19] propose a zooming technique using a single finger called Cyclo which maps the circular movement to the zoom direction avoiding clutching when zooming. Nancel et al. [21] work on multi-scale interfaces specifically designed for wall-sized displays by mapping pan-and-zoom interfaces from mouse or touch input to mid-air gestures. Petkov et al. [24] introduce an interface allowing users to explore and interact with the huge canvas by walking along the display surface. In this work, we do not focus on the affordances for zooming itself, but rather on the display of the zoomed-in media while maintaining context to avoid motion sickness and allow easier targeting. It is important to note that all these methods are also limited to interactions with flat devices. Although much of this work can be adapted to VR to allow the user to guide the zoom level, our work here focuses on what to display, given that head motion will be used as the primary interface for panning. We believe our work is the first to offer solution to this particular problem.

Simulator Sickness. Interaction research has found that in virtual environments, a mismatch in visual context may lead to visual vertigo, or simulator sickness, and people will suffer from dizziness and nausea [15]. Any conflict between visual and body motion makes the signals sensed by user uncoupled. Particularly, if the visual input moves much faster than head movement, this mismatch leads to not only discomfort, but causes vestibular-ocular reflex (VOR) gain adaptation and invokes an uncomfortable re-calibration process [29].

Several works [28, 23, 27] show that such symptoms tend to be more severe with head-worn displays than with screen-based simulator. This discomfort hampers the VR user's experience and thus best practices have been encoded for developers to follow [22].

To alleviate the sickness, Whittinghill et al. [31] proposed providing a stable visual reference such as a virtual nose to reconcile differences between the vestibular and visual senses. Other work [26, 8] showed that reducing the display's field of view (FOV) could help combat the nausea. This indicates that our peripheral vision is the most sensitive to mismatches between visual and head motion. Recently, Fernandes and Feiner [10] propose optimizing the tradeoff between presence and FOV and thus add dynamic FOV modification according to a user's motion.

We take many of the insights in these works as motivation in what we demonstrate to reduce motion sickness when viewing multi-scale panoramas in a VR headset.

PROBLEM

On desktop computers or laptops, affordances such as rolling the scroll wheel, or double-clicking is provided to zoom in and out, and dragging the mouse to pan across the panorama. For mobile devices and tablets, finger motions like clicking, dragging and pinching allows for panning-and-zooming. Gesture interaction has also been used for large wall sized displays to manipulate large scale imagery as proposed by Nancel et al. [21]. In all of these cases, panning distance is modulated by the zoom level to keep the visual motion closely related to the scale of the motion in the particular interaction modality. As we discuss below, such modulation by zoom level does not translate well to VR. None of these interactions use head motion to control panning as is typical in VR. However, some mobile interfaces do leverage the gyro to induce panning across wide FOV imagery, but this interaction also runs into the same problems above when zoomed into high resolution imagery.

We refer to the zoom level z as $\frac{FOV_0}{FOV_i}$, where FOV_0 is the heads-up display's field-of view, and FOV_i is the field-of-view of the imagery. Thus, the zoom level is identically 1.0 if not zoomed in at all. A zoom level of 1 represents the basic contract with the user in most VR applications and leads to a sense of immersion as if one is embodied in the virtual scene.

Initially the image is fully zoomed out and in this case $z = 1$. As the viewer zooms in, the portion of the scene displayed corresponds to a smaller region of the panorama, thus the field of view is smaller and the zoom level increases.

In traditional VR interfaces, the pan position changes with head rotation and thus a head rotation of FOV_0 would indicate that all imagery in the scene has changed completely. In other words a pixel at one edge of the view will have traversed the view completely to the other side. As the zoom level increases, smaller and smaller head motions will induce the equivalent visual flow. In other words, the same head rotation of FOV_0 at zoom level z will cause the imagery to traverse the user's entire view from side to side z times. This rapid visual motion is what leads to difficulty in orientation and discomfort such as simulator sickness.

Our goal is to develop a user interface for heads-up VR viewing of high resolution wide angle imagery that allows for both panning and zooming, while avoiding issues of simulator sickness. Ideally such a system will preserve as many aspects of the immersive quality of VR. The interaction should be intuitive and engaging to users in that it does not require specific training to utilize the high level zoom functionality. It should combat simulator sickness as much as possible. In the next section, we describe several designs we explored and their pros and cons.

DESIGN CHOICES

This section describes several designs for the interface of panning and zooming high-resolution panoramas in VR devices. This paper does not address the specific affordance for zooming, thus in all cases we simply use a joystick for setting the zoom, while head turning effects the pan location.

We conducted a formal experiment to compare three modes with what we will term *Normal Mode*. This default interface is to have the pan position match the head position and simply fill the FOV with the zoomed in imagery. Shown in Figure 1 left, as the zoom level becomes higher, the perceived visual motion becomes faster. As mentioned in previous experiments [22], the mismatch between the slow head motion and fast display motion often leads to strong feeling of simulator sickness. Besides the modes included in the experiments, we will also describe two modes we thought as candidates but rejected due to certain problems.

Slow Mode

We design Slow Mode which adapts the visual motion velocity to the zoom level to match head rotation velocity. More specially we set the view motion speed to be proportional to wearer's head motion but be inversely proportional to zoom level. This is essentially analogous to the standard interface on flat displays where panning is afforded through mouse and/or touch interaction.

This interface provides a simple stable viewing experience. Unfortunately, the motion speed adjustment unavoidably brings along two unacceptable downsides: physical challenges and a proprioceptive problem. First, users may need to move very far to navigate to a desired location, since at different zoom levels, the same global change in direction requires different amounts of head motion. For example, to turn 90-degrees to the right, the user may have to spin multiple times in a circle. This problem is even more challenging when moving vertically. So it dramatically limits the range users could pan at high zoom level, as shown in Figure 3. A more severe problem is a proprioceptive mismatch between what the user sees and what they feel. Since the view direction change does not match user's head rotation, the horizon may change when user's head moves vertically at different zoom levels. These problems are illustrated in Figure 3. Because the proprioceptive problem results in severe sickness and discomfort, it makes zoom-and-pan operation impossible along vertical direction in Slow Mode.

Having experienced these problems ourselves using our early prototype, we dismissed this mode in our final experiment.

Location Mode

Another idea we explored is to treat zooming as a virtual location change within a sphere texture mapped with the high resolution panorama. In other words, instead of narrowing the FOV, keep the FOV of the virtual camera constant but rather move the camera forward towards the virtual sphere. This leverages the natural ambiguity between zooming and forward motion.

Once near the sphere (zoomed in) a turn of the head would then reveal an uneven distribution of zoom level over the visual field, but would provide a paradigm to explore the full resolution of the imagery.

Although intriguing as an experiment, the distortion of the imagery when not facing directly outward from the center made the overall effect unintuitive. We thus also dropped this interface from final consideration.

Circle Mode

The attempts above help us identify some requirements for a pan-and-zoom interface. First, the panning speed should be consistent with the direction of user's head motion; Second, it's better to stay at the center to avoid distortion. Third, a good interface should enable users to recognize their location when zoomed in, to make it easy for overall navigation. Finally, visual motion should be limited, especially on the view periphery to avoid simulator sickness.

With these objectives in mind, we propose the Circle Mode: In Circle Mode, only a circular region in the center is zoomable, and outside the circle the zoom level remains 1. This interface is similar to a magnifying glass, which integrates two fields-of-view at the same time. There are effectively two layers in the screen, the one in the circle called the *zoom* layer, and the other one is called *background* layer whose center part is occluded by the zoom layer. The zoom level of the background layer is always 1.

Maintaining the background layer greatly reduces the motion sickness. By limiting the zoomed in motion to only the center of the users visual field reduces the degree of dizziness [26, 8]. Moreover, the context of the background stabilizes the peripheral motion and also benefits navigation. When panning slowly the background layer serves as a stable reference to alleviate the sickness as confirmed in [31].

However, since the central portion of the view is zoomed in, it offers little navigational assistance during rapid panning as the imagery moves so quickly. Since the circular region obscures the background, it is hard to locate objects. This leads us to two new interfaces described below that both maintain the benefits of Circle Mode while alleviating the downsides.

Alpha Circle Mode

The next mode we explored is to begin with the Circle Mode above, but to modify the transparency of the central circular zoomed-in region based on the panning velocity (i.e., the users' head motion). This allows the slower moving background to be visible through zoomed-in imagery during rapid head motion. The relationship between panning velocity, v, and transparency, α, is designed such that the circle is fully opaque for small velocities then becomes transparent smoothly as the velocity increases until fully transparent and returns to opaque as the user comes back to rest in one location. We also add some hysteresis (damping) to the transparency changes to smooth out the temporal variation in transparency. The hysteresis

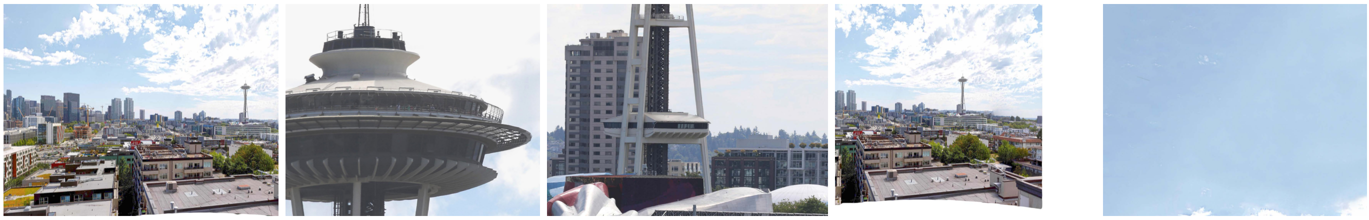
(a) Starting view (b) Zoom in (c) Pan down (d) Zoom out (e) Face forward

Figure 3: Problems of Slow Mode: In this example, a user first (a) looks at the top of the space needle and then (b) zooms in. Next the user pans down to the bottom of space needle. To achieve this pan in Slow Mode, the user might tilt down to face their feet, or perhaps even put their head down between their knees facing backward, which are the physical challenges. Next consider when they zoom out – they are still seeing the horizon, but they feel like they are looking down at their feet. And when they turn their head facing forward, their view in VR display is however sky. Those are the proprioceptive problems. Note that any pan in vertical direction at different zoom level will lead to proprioceptive problems.

values differ between whether the transparency is increasing (faster response), or decreasing (slower response) based on empirical experiments. Finally, to maintain the context of the circle and also to aid in targeting during rapid panning, we include a black circle outline as well as targeting cross-hairs that lie behind the zoomed-in circle, but in front of the non-zoomed-in background. Thus these become visible as the circle become transparent.

The specific mapping from panning velocity, v, to target transparency, α^*, independent of hysteresis is given by

$$\alpha^*(v) = \begin{cases} 1 & \text{if } \mathrm{v} < \mathrm{v}_1 \\ (v_2 - v)/(v_2 - v_1) & \text{if } \mathrm{v}_1 < \mathrm{v} < \mathrm{v}_2 \\ 0 & \text{if } \mathrm{v} > \mathrm{v}_2 \end{cases}$$

where, v_1 = 10 degrees/second, and v_2 = 40 degrees/second, and with hysteresis, the final transparency, α is given by

$$\alpha_{t+1} = \gamma * \alpha_t + (1 - \gamma) * (\alpha^*_{t+1} - \alpha_t)$$

where, $\gamma = 0.9$ if transparency is increasing, and 0.95 is transparency is decreasing. This difference leads to rapid increase of transparency as the head begins to turn, but eases back to opaque more slowly when the head comes to a rest. These values were selected through informal experimentation.

Zoom Circle Mode

Another similar mode inspired by the same issues that led to the Alpha Circle Mode automatically varies the zoom level within the circle based on head motion, instead of the transparency (again, best viewed in the video).

The user always sets the target zoom, typically while not moving. By lowering the actual zoom during rapid head motion, the visual motion across the field-of-view is lowered, and more context is provided within the circle for targeting to reach a goal location. When the user stops moving, the circle slowly zooms back in to the target zoom level giving the user time to make subtle adjustments to head direction to zoom in on a target location.

The implementation follows the same pattern as with the circle transparency. In particular, given a target zoom level, Z^*, the actual zoom level, z, is computed in two steps, first setting a target, z^*, given the head velocity, v:

$$z^*(v) = \begin{cases} Z^* & \text{if } \mathrm{v} < \mathrm{v}_1 \\ 1 + (Z^* - 1)(v - v_1)/(v_2 - v_1) & \text{if } \mathrm{v}_1 < \mathrm{v} < \mathrm{v}_2 \\ 1 & \text{if } \mathrm{v} > \mathrm{v}_2 \end{cases}$$

where v_1 and v_2 are as above. The final zoom level, z, after hysteresis is given by

$$z_{t+1} = \gamma * z_t + (1 - \gamma) * (z^*_{t+1} - z_t)$$

where, $\gamma = 0.95$ if zoom is decreasing, and 0.98 is increasing is decreasing. This difference leads to rapid zoom out as the head begins to turn, but eases back to the target zoom level more slowly when the head comes to a rest.

Some notes on damping the head motion

As anyone trying to point a hand held telescope or binoculars at a target can attest to, it is very hard to keep it stable, and thus the enlarged image appears to shake. In the VR, small head motions contribute to similar instability. In addition, even a perfectly still headset incurs some positional read noise. We over come both of these with some gentle damping. This does not fully eliminate the natural shake as we found that too much damping leads to the well known phenomena of "over steering" which occurs due to the latency required for the damping.

We implement the damping effect in much the same way as with the hysteresis term in the transparent and Zoom Circle Modes. More specifically, the quaternion representing the currently reported head position, $Q^*(t)$, is *slerped* (linearly interpolated in quaternion space) with the most recent quaternion used for rendering, Q_{t-1}. The interpolation value we found to work best was 20%, i.e., Q_t = slerp(Q_{t-1}, $Q^*(t)$, 0.2).

USER STUDY

We conducted a user study to assess the comfort and efficiency of four of the modes we discussed above: Normal Mode(N), Circle Mode(C), Alpha Circle Mode(AC), and Zoom Circle Mode(ZC). We ruled out both the slow and Location Modes based on the observations described above. In addition to simply exploring a wide angle high resolution panorama, we also asked the users to find specific elements in the scene and timed them as we describe below. This provided us with both qualitative and quantitative feedback which we report below. We used a single 3 gigapixel panorama covering about 150 degrees horizontally which appears in our figures and the accompanying video. (Note there is a portion of the sky missing due to the way our image stitcher processed the input imagery from a gigapan capture.)

Participants

We recruited thirty three participants (9 female and 24 males, aged from 22 to 60 years) all employees at a technology firm. Among the participants, seven subjects did not have any previous experience in VR, nine had some minimal experience, seventeen had more extensive experience with VR devices.

Apparatus

Participants were immersed in the virtual environment using an Oculus Rift (DK2), in which head tracking was provided by the Oculus Rift. Participant could control the zoom level with an Oculus Remote during free form exploration. We locked the zoom level as 32 during the specific target finding. The application was developed using Unity and driven by a standard graphical workstation. During the experiment, participants were seated on a swivel desk chair to afford easy turning around for navigation.

Protocol

At the beginning of the experiment, participants were briefed about the equipment and the experimental tasks. Once the experimenter set up the VR equipment, we asked the participants to first get familiar with the four modes by asking them to follow a short tutorial on how to zoom, pan, and switch modes. The participants then began the tasks, which included 13 assignments. After finishing the tasks, the experimenter removed the head mounted display and participants were asked to fill a subjective questionnaire related to their experience. Participants could request a rest anytime during the tasks or even quit midway through each task. On average the experiment lasted 25 minutes.

Tasks

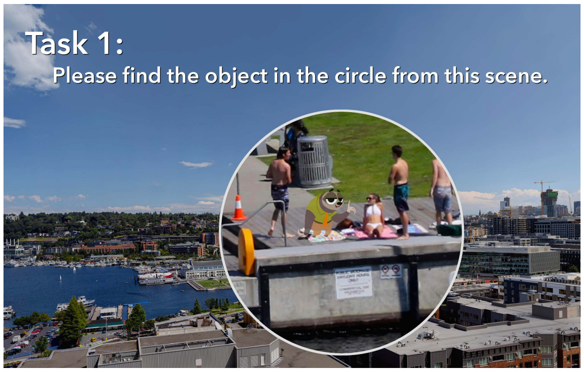

Figure 4: An example of our tasks in the experiment. The subjects were asked to find the character in the circle, three-toed sloth in Zootopia in this case, hidden at the panorama. A hint is that the character is located within the range of the background zoom-out scene.

The tasks consisted of 13 assignments. An example of the task can be seen in Figure 4. Each task asked a participant to search for a specific target character in a high-resolution panorama. We selected the characters from a set of emoticons and placed them deep into the scene.

In a pilot study, we found it to be difficult to find such a tiny character randomly placed in the very large panorama. Therefore, we provide a hint of where to search, by showing the target in the circle roughly in front of the correct portion of the background scene. Users are told this is the general location to search. They can flip back and forth between the static goal/hint and active searching.

An example is shown in Figure 4. For each task, we fix the mode and the zoom level to limit the number of free parameters. Since we use the same panorama during the experiment, it is likely that users gain familiarity with the scene as the experiments proceed. Thus we randomly shuffle both the mode and the target order across participants respectively. Each participant is given 12 targets to find, thus experiencing each mode 3 times. In the last assignment, the task remains the same but participants have control over which mode to use. We want to observe which mode they choose to use.

For each task, we record the time needed for participants to find the target. As we can observe what they see, we mark the task as complete when they come essentially to rest with the target centered in their field of view. Considering some people experience different levels of motion sickness with VR devices, we allow participants to quit a task whenever they feel fatigue or discomfort. We also call a task as having failed if they cannot find the target after 180 seconds. We also record the time as 180 seconds (for visualization purposes) if they quit the task.

Questionnaires

Qualitative data was collected for each mode using a questionnaire after the experiment. Three questions, arranged on 5-level Likert scales, were administered to gather data on the aspects listed below.

Motion sickness *"I felt dizzy when navigating the panorama".* It measures the compatibility between the visual scene change and the motion in the physical world. Although the SSQ [14] is the standard instrument to measure motion sickness, our experiment used a simpler metric asking the users for their overall perception of motion sickness. This score is more limited than SSQ and does not provide a breakdown of severity of each symptom.

Ease of use *"It was easy to move and find objects in the scene."* It measures if the interface could help user to investigate the details of the panorama.

Further exploration *"I would use this mode over the others to explore new panoramas."* This measures the satisfaction of the mode.

Analysis

Hypotheses

In section 3 we discuss the problems associated with Normal Mode that led us to explore several modes of interaction. Based on our explorations we developed two general hypotheses:

Question	N	C	AC	ZC
Motion sickness	$(1.0, 1.3)$	$(1.8, 1.1)$	$(3.1, 1.0)$	$(2.5, 1.2)$
Ease of use	$(2.1, 1.0)$	$(2.5, 1.2)$	$(3.8, 0.9)$	$(3.7, 0.8)$
Exploration	$(1.9, 1.1)$	$(2.5, 1.1)$	$(4.1, 0.9)$	$(3.9, 0.9)$

Table 1: Mean and standard deviations for the questionnaire responses (5 point Likert scale). For Motion Sickness we report 5 - score, thus in each case a higher score is better.

Comparison	Motion sickness		Ease of use		Exploration	
	t	p	t	p	t	p
N vs C	2.36	0.01	1.33	0.095	2.29	0.013
C vs AC	4.61	$7.6x10^{-6}$	4.8	$3.3x10^{-8}$	-5.96	$2.7x10^{-8}$
C vs ZC	2.5	0.008	4.27	$2.1x10^{-5}$	5.05	$1.2x10^{-6}$
ZC vs AC	2.4	0.01	0.68	0.25	1.08	0.14

Table 2: T-test

H1 Limiting the zoomed in area to the central portion of the visual field and providing non-zoomed context would be helpful, particularly for reducing simulator sickness. Thus, the various Circle Modes will be preferred.

H2 Transparent and Zoom Circle Modes will perform better than the simple Circle Mode, particularly for targeting/finding tasks. We are not sure a priori which of these two will perform better.

In the following, we test our hypotheses with respect to user experience and the task performance.

User experience

The Table 1 provides the summary of the questionnaires results. A Friedman test on each of the 3 questions produces a value greater than 20 indicating high levels of agreement across participants on the review of the four modes. A t-test allows us validate our hypothesis. See the table in Table 2.

Normal Mode vs Circle Mode On motion-sickness and exploration, Circle Mode shows significant advantages over Normal Mode. However, for ease-of-use, the advantage there is not as clear (t = 1.33, p near 0.1). A possible reason is that our task requires participants to pan a far distance through the panorama. In such cases, the high visual velocity in both normal and Circle Modes may induce some nausea.

Alpha Circle and Zoom Circle wins By comparing the Circle Mode with transparent circle and zoom circle, we see both of these variants are assessed significantly better than Circle Mode on all three questions.

Alpha Circle vs Zoom Circle From the p-value, we don't see a clear preference between these two modes. Some of the participants have a strong preference for one mode based while some preferred the other. Several participants commented that the automatic zooming is slower or faster than they expected, which can be tuned by the γ values.

We can thus conclude from the user experience questions, that our hypothesis is supported by the user feedback.

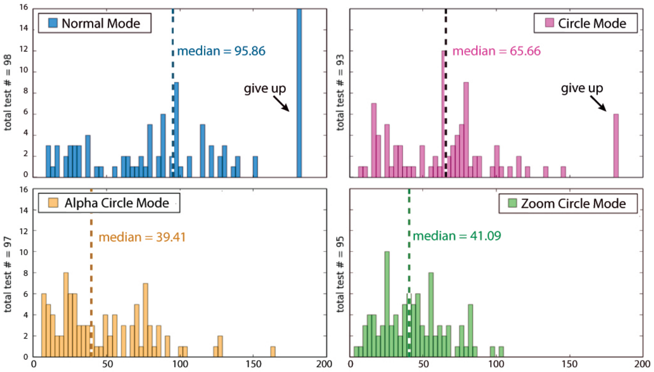

Figure 5: The time users spent to find the target in each mode. We mark the test time as 180 seconds when the users gave up the task.

Mode	Mean	STD	Median	Min	Max
N	77.0	39.4	84.9	9.1	150.0*
C	61.1	30.7	65.7	7.9	146.5*
AC	47.7	32.0	39.4	6.7	164.7
ZC	43.5	22.1	41.3	5.2	104.5

Table 3: The statistic summary for the completion times in each mode. The table shows time (seconds) to find an object in a scene, for each of the mode in our user study. Here we only include the tasks completed by users. Note that the user study task were capped at max seconds when users gave up and possibly spent more time than the time shown in the table if there was no option to quit.

Task performance

Completion time Figure 5 shows completion times of all subjects, for each of the four modes. If the participant chose to "give up" on a task (either because of complaints of dizziness or sheer frustration), we mark their search time as 180 seconds. In 16 out of 98 assignments participants chose to give up in Normal Mode, which is the most. There are 6 quits in Circle Mode, and no quits in zoom circle or Alpha Circle Modes. This, in itself, suggests that zoom circle or alpha Circle Mode are the most suitable way to investigate the deep detail of a panorama. Circle Mode performs worse than Normal Mode which has the longest median search

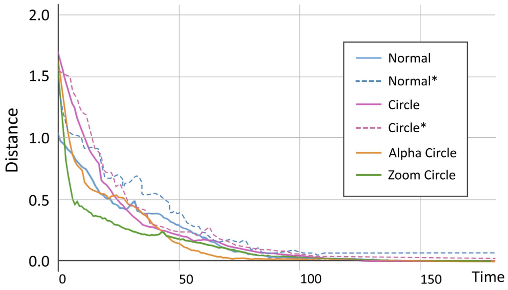

Figure 6: Users get closer to the target in four modes along the time. The solid line only includes subjects completing the task, whereas the dashed line includes people who gave up during the task. Here the distance is the great circle distance normalized by the radius, in $[0, \pi]$.

time. This observation is consistent with the questionnaires about the motion-sickness.

In Table 3, we calculate summary statistics for the time needed for task completion of each mode. Comparing with the Normal Mode, people using Circle Mode need approximately 70 percent of the time, and other two required less than half the time, which shows the clear advantage using circle related modes. Also notice that the worst performance using Zoom Circle Mode is just slightly worse than mean or median performance using normal circle. This validates our hypotheses that Normal Mode is not an ideal interface.

Additionally, Figure 6 shows in each mode, the curve of how the distance between the user's pan position and target changes with time for a task. From the figure, we can see that when using transparent and Zoom Circle Modes, in the early stage the distance drops significantly faster than in other modes. Notice that panning is crucial at the beginning to approximate target's location, and the figure shows that these two modes help the fast initial localization, suggesting a better panning experience. In the figure, we show both the dash line which includes people who gave up during the task, and the solid line that only includes those completing the task. The dashed curve becomes more tortuous than the one for people completing tasks, since some of the people pan crazily to find where the target is and the distance curve changes dramatically. These cases also demonstrate it's hard to navigate so people don't know where to find the target.

Free mode As mentioned before, during the user study, we force the participant use specific modes in the first 12 tasks, and lock the zoom level as well. In the last task, they have the freedom to choose the mode and can control the zoom. There are 32 out of 33 participants choosing to use transparent or Zoom Circle Mode to find the character. We also observe that most of the users quickly shift to the Circle Modes for better navigation when they have control.

Overall, an inspection of the pool of results leads us to believe that Alpha Circle and Zoom Circle Modes are more user-friendly than the other two modes for zoom-and-pan in VR devices.

CONCLUSION AND FUTURE WORK

We bring together two recent innovations in immersive media. First, the ability to capture very high resolution imagery than can be explored only through panning and zooming. Second, the rise of consumer level heads-up displays for virtual reality. Unfortunately, zooming breaks the basic contract with the user in VR, that the FOV of the visual field matches the FOV of the imagery. In this paper, we study methods to overcome this restriction to allow high resolution panoramic imagery to be explorable in VR.

We demonstrate that limiting the visual FOV containing zoomed in imagery to the central portion of the visual field and adding non-zoomed media outside this area limits simulator sickness. We also show how modulating the transparency or zoom level during rapid panning help, particularly in targeting tasks. We report results from a user-study to exercise the proposed interfaces empirically and it shows that our solutions successfully alleviate motion sickness and improves the experience for user to investigate the details in high-resolution panorama.

We do not pretend that there are no other possible interfaces that may work well. We would invite comparisons to other techniques. We also hope to explore whether some combination of transparency and zoom at the same time may be better than either on its own. We are also interested to explore other task spaces as well as types of imagery such as high resolution medical imagery and maps to see if the same modes will work well in those contexts.

As VR devices become more popular, it is interesting to rethink how we can solve the traditional tasks which are well-studied on flat displays. Sometimes it's challenging as the new control/interaction system leads to different requirements for user experience. Panning and zooming high-resolution panoramas, which is discussed in this paper, is just one example. We hope this research will act a role model to motivating more research on improving user experiences in VR devices.

One more result: High Resolution Deep Dream Imagery

In the accompanying video we show this work applied to a high resolution stitched panorama overlooking a lake and urban scene. We also show a high resolution HDR panorama by Daniel Ambrosi (see Figure 7) that has been passed through a variant of Deep Dream (`https://research.googleblog.com/2015/06/inceptionism-going-deeper-into-neural.html`). You can see other similar works by Ambrosi at `http://www.danielambrosi.com/`. This type of imagery could not be appreciated without either a zooming interface or extremely large format prints (which he has also created).

REFERENCES

1. Daniel Ambrosi. 2015. Dreamscapes. `http://www.danielambrosi.com/Dreamscapes/`. (2015).
2. Caroline Appert, Olivier Chapuis, and Emmanuel Pietriga. 2010. High-precision magnification lenses. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 273–282.
3. Patrick Baudisch, Nathaniel Good, Victoria Bellotti, and Pamela Schraedley. 2002. Keeping things in context: a comparative evaluation of focus plus context screens, overviews, and zooming. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. ACM, 259–266.
4. Eric A. Bier, Maureen C. Stone, Ken Pier, William Buxton, and Tony D. DeRose. 1993. Toolglass and Magic Lenses: The See-through Interface. In *Proceedings of the 20th Annual Conference on Computer Graphics and Interactive Techniques (SIGGRAPH '93)*. ACM, New York, NY, USA, 73–80. DOI: `http://dx.doi.org/10.1145/166117.166126`
5. Marianne Sheelagh Therese Carpendale. 1999. *A framework for elastic presentation space*. Ph.D. Dissertation. Simon Fraser University.

Figure 7: A high resolution Deep Dream photograph and a zoomed-in detail. Courtesy of Daniel Ambrosi.

6. Shenchang Eric Chen. 1995. QuickTime VR: An Image-based Approach to Virtual Environment Navigation. In *Proceedings of the 22Nd Annual Conference on Computer Graphics and Interactive Techniques (SIGGRAPH '95)*. ACM, New York, NY, USA, 29–38. DOI: http://dx.doi.org/10.1145/218380.218395
7. Andy Cockburn, Amy Karlson, and Benjamin B. Bederson. 2008. A Review of Overview+Detail, Zooming, and Focus+Context Interfaces, In ACM Surveys. *Comput. Surveys* (January 2008).
8. Mark H Draper, Erik S Viirre, Thomas A Furness, and Valerie J Gawron. 2001. Effects of image scale and system time delay on simulator sickness within head-coupled virtual environments. *Human Factors: The Journal of the Human Factors and Ergonomics Society* 43, 1 (2001), 129–146.
9. Facebook. 2016. 360. https://facebook360.fb.com/. (05 2016).
10. Ajoy S Fernandes and Steven K Feiner. 2016. Combating VR sickness through subtle dynamic field-of-view modification. In *2016 IEEE Symposium on 3D User Interfaces (3DUI)*. IEEE, 201–210.
11. Karen A. Frenkel. 2010. Panning for Science. *Science* 330, 6005 (2010), 748–749. http://science.sciencemag.org/content/330/6005/748
12. Google+. 2015. 360 Panoramas. https://plus.google.com/s/%23Panorama_360. (2015).
13. C.Y. Ip and A. Varshney. 2011. Saliency-assisted navigation of very large landscape images. *IEEE Transactions on Visualization and Computer Graphics* 17, 12 (2011), 1737–1746.
14. Robert S Kennedy, Norman E Lane, Kevin S Berbaum, and Michael G Lilienthal. 1993. Simulator sickness questionnaire: An enhanced method for quantifying simulator sickness. (1993).
15. Eugenia M Kolasinski. 1995. *Simulator Sickness in Virtual Environments.* Technical Report. DTIC Document.
16. Johannes Kopf, Billy Chen, Richard Szeliski, and Michael Cohen. 2010. Street Slide: Browsing Street Level Imagery. In *ACM SIGGRAPH 2010 Papers (SIGGRAPH '10)*. ACM, New York, NY, USA, Article 96, 8 pages. DOI: http://dx.doi.org/10.1145/1833349.1778833
17. Johannes Kopf, Matt Uyttendaele, Oliver Deussen, and Michael F. Cohen. 2007. Capturing and Viewing Gigapixel Images. In *ACM SIGGRAPH 2007 Papers (SIGGRAPH '07)*. ACM, New York, NY, USA, Article 93. DOI: http://dx.doi.org/10.1145/1275808.1276494
18. Qing Luan, Steven M Drucker, Johannes Kopf, Ying-Qing Xu, and Michael F Cohen. 2008. Annotating gigapixel images. In *Proceedings of the 21st annual ACM symposium on User interface software and technology*. ACM, 33–36.
19. Sylvain Malacria, Eric Lecolinet, and Yves Guiard. 2010. Clutch-free Panning and Integrated Pan-zoom Control on Touch-sensitive Surfaces: The Cyclostar Approach. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '10)*. ACM, New York, NY, USA, 2615–2624. DOI: http://dx.doi.org/10.1145/1753326.1753724
20. Leonard McMillan and Gary Bishop. 1995. Plenoptic modeling: An image-based rendering system. In *Proceedings of the 22nd annual conference on Computer graphics and interactive techniques*. ACM, 39–46.
21. Mathieu Nancel, Julie Wagner, Emmanuel Pietriga, Olivier Chapuis, and Wendy Mackay. 2011. Mid-air Pan-and-zoom on Wall-sized Displays. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '11)*. ACM, New York, NY, USA, 177–186. DOI: http://dx.doi.org/10.1145/1978942.1978969
22. Oculus. 2016. Simulator Sickness in Oculus Best Practice. https://developer3.oculus.com/documentation/intro-vr/latest/concepts/bp_app_simulator_sickness/. (2016).

23. Emilee Patrick, Dennis Cosgrove, Aleksandra Slavkovic, Jennifer A Rode, Thom Verratti, and Greg Chiselko. 2000. Using a large projection screen as an alternative to head-mounted displays for virtual environments. In *Proceedings of the SIGCHI conference on Human Factors in Computing Systems*. ACM, 478–485.

24. Kaloian Petkov, Charilaos Papadopoulos, and Arie E Kaufman. 2013. Visual exploration of the infinite canvas. In *Virtual Reality (VR), 2013 IEEE*. IEEE, 11–14.

25. Emmanuel Pietriga and Caroline Appert. 2008. Sigma lenses: focus-context transitions combining space, time and translucence. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 1343–1352.

26. JD Prothero and HG Hoffman. 1995. Widening the field-of-view increases the sense of presence in immersive virtual environments. *Human Interface Technology Laboratory Technical Report TR-95* 2 (1995).

27. Sarah Sharples, Sue Cobb, Amanda Moody, and John R Wilson. 2008. Virtual reality induced symptoms and effects (VRISE): Comparison of head mounted display (HMD), desktop and projection display systems. *Displays* 29, 2 (2008), 58–69.

28. Kay M Stanney, Robert S Kennedy, and Julie M Drexler. 1997. Cybersickness is not simulator sickness. In *Proceedings of the Human Factors and Ergonomics Society annual meeting*, Vol. 41. SAGE Publications, 1138–1142.

29. Thomas A Stoffregen, Mark H Draper, Robert S Kennedy, and Daniel Compton. 2002. Vestibular adaptation and aftereffects. *Handbook of virtual environments: Design, implementation, and applications* (2002), 773–790.

30. Luc Vincent. 2007. Taking online maps down to street level. *Computer* 40, 12 (2007), 118–120.

31. DM Whittinghill, Bradley Ziegler, T Case, and B Moore. 2015. Nasum Virtualis: A Simple Technique for Reducing Simulator Sickness. In *Games Developers Conference (GDC)*.

32. Wikipedia. 2010. Gigapan. `http://en.wikipedia.org/wiki/Gigapan`. (2010).