Evaluating the Effects of Image Persistence on Dynamic Target Acquisition in Low Frame Rate Virtual Environments

David J. Zielinski*
Duke Immersive Virtual Environment
Duke University

Hrishikesh M. Rao †
Department of Biomedical Engineering
Duke University

Nicholas D. Potter‡
Athletic Department
Duke University

Marc A. Sommer §
Department of Biomedical Engineering
Department of Neurobiology
Duke University

Lawrence G. Appelbaum¶
Department of Psychiatry and
Behavioral Sciences
Duke University

Regis Kopper‖
Department of Mechanical Engineering
and Materials Science
Duke University

Abstract

User performance in virtual environments with degraded visual conditions due to low frame rates is an interesting area of inquiry. Visual content shown in a low frame rate simulation has the quality of the original image, but persists for an extended period until the next frame is displayed (so-called high persistence - HP). An alternative, called low persistence (LP), involves displaying the rendered frame for a single display frame and blanking the screen while waiting for the next frame to be generated. Previous research has evaluated the usefulness of the LP technique in low frame rate simulations during a static target acquisition task. To gain greater knowledge about the LP technique, we have conducted a user study to evaluate user performance and learning during a dynamic target acquisition task. The acquisition task was evaluated under a high frame rate, (60 fps) condition, a traditional low frame rate HP condition (10 fps), and the experimental low frame rate LP technique. The task involved the acquisition of targets moving along several different trajectories, modeled after a shotgun trap shooting task. The results of our study indicate the LP condition approaches high frame rate performance within certain classes of target trajectories. Interestingly we also see that learning is consistent across conditions, indicating that it may not always be necessary to train under a visually high frame rate system to learn a particular task. We discuss implications of using the LP technique to mitigate low frame rate issues as well as its potential usefulness for training in low frame rate virtual environments.

Index Terms: H.5.1 [Information Interfaces and Presentation]: Multimedia Information Systems—Artificial, augmented, and virtual realities

1 Introduction

Computer graphics systems have made significant improvements over the decades, and today it is possible to render complex graphics at high frame rates in real-time [14]. However, as the quality of rendering improves, so does the demand for graphics, which pushes the ability of a computing system to render frames at the maximum display refresh rate. GPU load is just one case which can force virtual environments (VEs) to run at a low frame rate. When running a VE at low frame rates the system presents visual artifacts and also suffers input latency and input update rate issues. We are interested in understanding how providing visually low frame rate feedback affects user performance on a VE task, particularly in dynamic tasks, where artifacts are constantly present.

In an ideal world, all virtual reality (VR) applications would run at the maximum frame rate of our display system (often 50hz and up). However, there are several conditions that may make it necessary to display simulations at low frame rates. For example, the low update rate of a tracking system (e.g. GPS), situations where heavy load is placed on the CPU (e.g. physics calculations), heavy GPU load (e.g. too many polygons or shaders), and network transport issues (e.g. real-time streaming of data or images) may all force VEs to be shown at lower-than-maximum display frame rates.

In most VR paradigms, the simulation gets input device data and then starts preparing a frame to render. However, in a low frame rate system, it may take a long time to render the frame. This means that by the time the image is finally displayed to the user, the input device (e.g. hand controller) may have moved significantly. This total time from physical movement to user viewing the rendered results is referred to as the end-to-end latency of the system. We will refer to the time difference between the user's physical actions and the system's processing of of the actions (but not yet rendering and displaying) as input latency. Previous work has shown that increasing input latency lowers task performance [21, 9].

Another related factor is input update rate, which is the rate at which input device data is sampled by the application. Input update rate becomes a problem in low frame rate VEs because in most VR paradigms, the system preforms all calculations and updates once per visual frame.

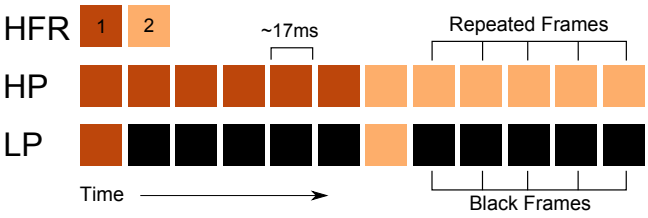

Figure 1: Comparison of rendering two frames (dark orange and then light orange) of content for the different experimental conditions. HFR displayed content at 60 fps, HP and LP displayed content at 10 fps.

1.1 Low Persistence Technique

The term low persistence has become familiar in the VR community after the Oculus Rift DK2 utilized the term to describe the shortening of the frame duty cycle to produce crisper images while running at a high frame rate (75 hz) [3]. In contrast, previous research explored the effects of low persistence on low frame rate

*e-mail:djzielin@duke.edu
†e-mail:hrishikesh.rao@duke.edu
‡e-mail:npotter@duaa.duke.edu
§e-mail:marc.sommer@duke.edu
¶e-mail:lawrence.appelbaum@duke.edu
‖e-mail:regis.kopper@duke.edu

IEEE Symposium on 3D User Interfaces 2016
19–23 March, Greenville, SC, USA
978-1-5090-0842-1/16/$31.00 ©2016 IEEE

situations [22], which was called a low frame rate low persistence (LP) technique. To understand the LP technique, lets look into the high persistence case first. In a low frame rate situation, we get one frame of "fresh" content, and repeat that frame some number of times until we get a new frame. This condition was referred to as low frame rate high persistence (HP).

The HP technique works well in most cases, but it introduces interesting visual artifacts, such as a phenomenon called phantom objects [10] or visually perceiving multiple copies of the original object. This breakdown stems from the limited capacity of the brain to interpolate over images that are slower than the cortical flicker fusion threshold [15]. With the LP technique, we get the original fresh frame of content, and then instead of showing it multiple times, we show black frames while waiting for the next frame to be generated. While we no longer observe artifacts present in the HP technique, a visual strobing effect and decreased brightness are introduced.

Both HP and LP are in contrast to the systems high frame rate (usually at the limit of the display hardware) which will be referred to as the HFR condition. In our experiment the high frame rate (HFR) condition will be at 60 fps (frames per second), and the HP and LP conditions will be at 10 fps. See Figure 1 for a comparison of the different techniques.

In this work, we present the evaluation of a trap shooting simulation on performance and learning using three visual conditions – HFR, LP and HP. We were interested in evaluating if training under the degraded visual conditions (HP and LP) transferred during the higher fidelity (HFR) assessment. Specifically, learning was evaluated by a pre- and post-assessment of shooting performance under the best visual condition (HFR), after training was conducted with one of the three visual conditions. We evaluated the performance of the visual conditions by comparing the training blocks performed at each different visual condition.

2 Related Work

2.1 Previous Low Persistence Experiment

The work presented here follows up on work by Zielinski et al. 2015 [22]. Regarding the results of the original low frame rate low persistence study, the authors conducted a within-subjects design and tested two tasks, selection and navigation. They found no significant differences in performance between the HP and LP, but did find that HFR was significantly better.

Thinking about their selection task, we think the static targets may have minimized the HP motion artifacts, as the only thing moving during the selection task was the user's wand. Also there was no randomization in the task, meaning that the user may have been able to easily predict the next target position. Finally, in their work, they utilized a simulator which updated the simulation (e.g. processing button clicks and collision updates) at 11fps for HP and LP (update every 91ms), and 55fps for HFR (update every 18ms). We are interested to see if this difference in time for registering clicks (max 73ms) could explain some of the results in the original experiment which showed HFR selection being faster.

In order to extend the exploration of the LP condition, and better understand the difference between HFR, HP, and LP, we decided to evaluate a target acquisition task with moving targets to explore situations where HP visual artifacts are continuously present. We also took care to make our simulator software have equivalent simulation update rates and input latency across all conditions to hopefully more isolate the visual components of the conditions. However, even though we are controlling for input latency, since the different visual conditions have different display output rates, we will still have different end-to-end latencies for the different display conditions.

2.2 Interpolation

One way to address the low frame rate issue is to consider interpolating between the frames of rendered content. Most modern TV's must deal with this issue, as their native display rates can be quite high (120+ hz) but the previously recorded content can be quite low (24hz/60hz). Complex algorithms have been developed to achieve this interpolation [20]. However, these systems rely on introducing delay, so that one or more frames in the future can help guide the interpolation process. This works fine for a TV, but for real-time interactive systems we need to minimize any additional delays.

VR researchers have investigated predictive methods for frame interpolation [17]. However, such techniques have some issues with display artifacts and specific issues with occlusion. Also, while we are good at predicting head motions (since the head does not start or stop motions instantaneously), we may have difficulty predicting motions for other objects in the scene, which can start and stop instantaneously.

A related technique is the Oculus Rift so-called "time warp technique" [2]. In this technique the recent head tracking data is sampled at the last moment before final rendering, and the image is warped to better fit a more recent head position/orientation. However, this type of technique works on high frame rate systems where head position/orientation is close between frames. On a low frame rate system, with larger differences in head position/orientation, the reprojection process would show artifacts.

2.3 Degraded Visual Conditions and Learning

While visual perception seems to be continuous, the human visual system actually receives an intermittent flow of visual information due to interruptions from blinks, sacades, and other sources of transient occlusion from interfering objects. In recent years, it has been realized that under structured conditions, such intermittent vision may actually offer an opportunity to enhance learning. For example, stroboscopic training techniques are now used in athletic training with companies such as Nike Inc. developing eyewear to simulate intermittent visual conditions. These stroboscopic training glasses are in an open condition for a period of 100ms, and then close for 100ms to 900ms. Research with such tools has shown evidence that practice in such degraded visual conditions does in fact lead to some types of learning. For example, comparison of pre- versus post-training has shown that aspects of perceptual skill such as foveal motion coherence and attention [5], short-term visual working memory [4], and coincidence-anticipation timing [18] are all improved relative to control participants who did not undergo such training. Such findings suggest that practice in conditions where vision is transiently interrupted might lead to selective enhancements in perceptual and motor performance.

2.4 Learning in VR

Much work has been done to establish that users can achieve learning in VR (and also that the learning can transfer back to the real world). Previous work has evaluated if VR can be used to train surgical residents for the task of gallbladder surgery [16]. Results indicated that the VR-trained group performed the real task faster, and with fewer errors than the control group. Abstract task learning has also been evaluated, which involved having users remember moving particular objects from source locations to destination locations [7]. In that work it was found that higher levels of immersion (e.g. increasing the field of view) increased memorization recall of a task sequence (placing colored shapes in specific grid locations). Other projects have experimented with virtual embodiment, where a user experiences the world as someone else. Specifically researchers have placed students in the role and appearance of a gorilla inside a virtual zoo, so the students can better gain knowledge about gorillas [6]. Finally, the Navy has shown that VR can be effective for learning how to navigate inside a ship for success

in shipboard fire fighting [19]. This body of research shows that VR can be useful for learning a wide range of tasks, and that these learned skills can often transfer over to the real world.

3 User Study

In order to evaluate the effects of LP and HP on a dynamic target acquisition task we conducted a user study. Our goal was to understand the differences among the conditions in terms of user performance, learning, and simulator sickness.

3.1 Apparatus

We used a six-sided CAVE-type [8] system to perform the experiment. Given that the experiment involved the user always looking towards the front screen, the door (which is the back wall) was left open and the rear projector was turned off, effectively making the system a 5-sided CAVE-type system. This allowed the experimenter to monitor the participant for compliance with the protocol. Head tracking with 6 degrees of freedom (DoF) was provided by an Intersense IS-900 LT tracker. Hand tracking was provided by an Intersense 6DoF wand controller.

Each wall of the CAVE-type system had two Christie Digital WU7K-M projectors running at 120hz. These projectors were overlapped and blended, and gave a total resolution of 1920x1920 pixels per wall. Active stereo was achieved via Volfoni Edge RF liquid crystal shutter glasses. A cluster of computers (1 computer per 2 projectors) plus one computer as a master computer powered the system. The computers utilized high end NVIDIA K6000 graphics cards which feature 12GB of video RAM, 2880 GPU cores, and a performance of 2000 million tri/sec. Our simulation was visually simple enough that we were able to run at our full system performance, which was 120hz active stereo - effectively 60 fps.

3.2 Simulation Software

The simulation software was written in C++ and OpenGL, and utilized the virtual reality library Syzygy [13]. As previously discussed, in an actual low frame situation, we increase input latency (since we often sample the input devices, before beginning to render the frame) as well as reduce our input update rate (as we often only sample the position from the tracker and compute new object collisions/positions once per rendered frame). In our experiment we would like to isolate the visual components from the input components, and we have thus simulated the low frame rate conditions. Specifically, the initial image is rendered to a texture. When the render callback function is called, if it is a repeated frame, the already-rendered image is displayed again. This allows us to repeat the image as long as necessary, and also allows us to obtain low latency and a full speed (60 fps) simulation update rate for all display conditions.

3.3 Participants

Participants were screened for family history of photosensitive epileptic seizure and asked to have normal or corrected-to-normal vision. We recruited 30 volunteers (10 in each training condition HFR, HP, LP). The ages ranged from 18-57, (M=27.63, SD=10.10) years. Eight participants (26.67%) were female and two (6.67%) held the wand with a left hand dominant grip. In terms of previous VR experience, 9 participants (30.00%) had previously experienced a CAVE-type system, and 9 participants (30.00%) had previous experience with a head-mounted display (HMD, e.g. Oculus) system. Overall, 14 participants (46.67%) had previously experienced some sort of immersive VR.

3.4 Study Design

We utilized a between-subjects design. Each participant was assigned one training method (HFR, HP, or LP). The independent variables were training method (HFR,HP,LP), and block number.

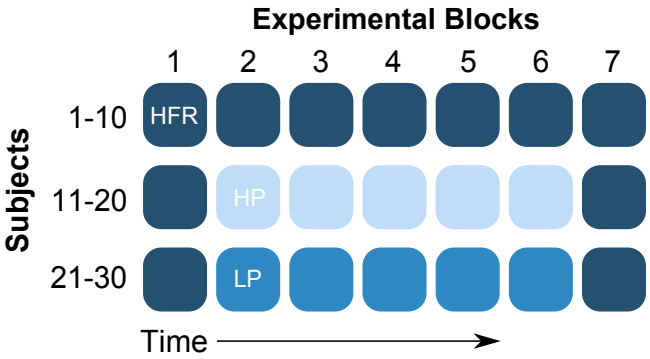

Figure 2: Layout of training blocks for the 30 participants. 7 Blocks total with a pre and post HFR block. Blocks 2-6 were considered training blocks and were one of the experimental conditions (HFR, HP, LP).

Each participant completed a total of 7 blocks (see figure 2). Each block consisted of 50 trials, for a total of 350 trials per participant. In order to give the participants adequate breaks between blocks, the participants did the experiment in pairs. While one participant was completing a block of trials, the other participant took a break. They alternated back in forth until all blocks had been completed. This design helped prevent fatigue in the participants (by forcing breaks to be taken), and also allowed us to run the experiment in a more timely manner. Each pair of participants did the same experimental condition (ie, HFR, HP, LP) and the pair took no longer than two hours of time to complete the whole session.

3.5 Procedure

Upon arrival, participants signed an informed consent form and completed a brief background survey containing demographic items. A tutorial followed where participants were shown how to complete the task, learned what constituted successful and error trials, and understood the different paths targets would take. This tutorial phase was conducted with easier (larger) targets than the actual study. The study itself had 7 blocks of 50 trials each. The block structure was HFR, 5 training blocks, then a final HFR (see Figure 2).

Each block began with a calibration phase, where the participant adjusted their position based on feedback from on screen display. By adjusting where they stood, it was verified that the hand controller was within 5 cm of the center of the system. Once close enough participants could click a button to proceed to the task. This ensured all participants were near the center of the system before each block of trials began. Participants were asked to remain with their feet in a fixed position, and to keep their elbows next to the torso, without extending the arm. Each participant was also instructed to use his or her non-dominant hand to steady the grip for the target selection.

After the completion of all 7 blocks, participants completed a video game experience survey and the Simulator Sickness Questionnaire (SSQ) [11]

3.6 Virtual Environment

The virtual environment was a large dark open field (see figure 3). A darker environment was used to reduce visual fatigue caused by contrast with the black frame on the LP condition. Grid lines were added to the ground to give the user a sense of perspective. The target was an orange sphere (radius=0.30 meters). A red ray extended from the user's hand controller (i.e., the wand) out into the distance. This gave the user real-time feedback about where they were currently pointing.

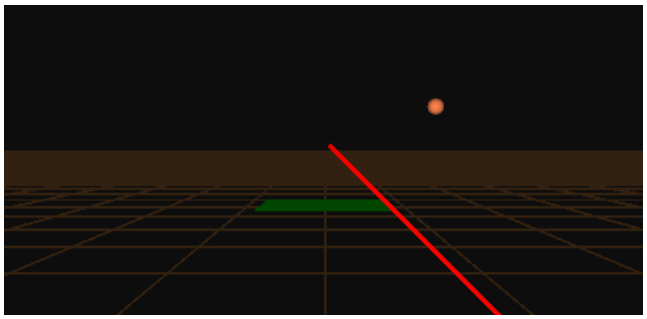

Figure 3: Virtual Environment utilized in the task. Note the red ray casting wand, the green trap house (used to request another target to be launched), and the orange target sphere traveling away from the user.

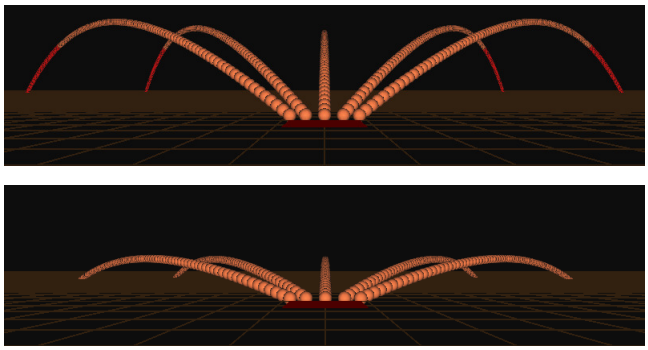

Figure 4: Tracings of the 10 different paths the target object would take during the study. Top picture illustrates the 5 high elevation trajectories. The bottom image illustrates the 5 low elevation trajectories.

3.7 Dynamic Target Acquisition Task

In an effort to create a quantifiable task that relied heavily on visual ability and provided existing logistical parameters we modeled our dynamic target acquisition task on the International Sport Shooting Trap game [1] as well as previous virtual reality distal pointing tasks [12].

The targets were launched from the ground plane behind the trap house. This was represented as a green square on the ground 16.46m directly in front of the user. Users would launch the next target by hovering the ray over the trap house until it turned green. There were 10 possible paths the targets could take (see Figure 4). These paths represented 5 horizontal directions of -45°, -30°, 0°, 30°, 45° relative to the center of the trap house. Targets were launched at two elevations high (25.17°) and low (12.95°) relative to the ground plane. All targets had a speed of 29.06 m/sec and proceeded along trajectories that incorporated gravitational pull, air resistance and lift force that were set to mimic typical flight times and distances observed in actual trap events in the International Trap game. In the real life Trap shooting event the variety of trajectories of the targets are much greater. We included the extreme and average target flight paths [1], as well as the average target flight speed in our dynamic target acquisition task.

After the target was launched, the user would position the hand controller to intercept the red ray with the launched target sphere. Then the under button on the IS-900 wand (resembling a trigger), was pressed with the index finger. If the user was contacting the ray with the target at the time of the button press, a shattering sound was heard, and the target visually broke into pieces. This was counted as a hit. For each launch of the target the user had 2 attempts (button

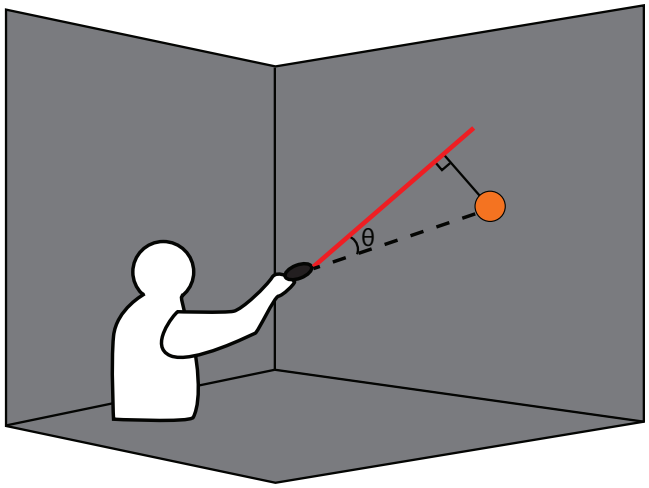

Figure 5: Diagram showing how angular closeness is determined. Note the target sphere in orange and the wand controller emanating the red ray casting ray.

presses) to hit the target. At 59.44 meters the target sphere turned red, to indicate it could no longer be hit. This design choice was made to mimic real trap shooting in which the physical force of the shot is only sufficient to break the clay pigeon at closer distances. If the user did not manage to hit the target it would continue traveling until hitting the ground plane, at which point it would vanish. After each trial the remaining trials left in the block (each block being 50 total trials) was displayed to the user in the lower right corner of the screen. To initiate the next trial, the participant would again hover the ray over the trap house to launch the next target. Each block of 50 trials was randomized such that 5 trials of each launch path (10) were shuffled.

For each block we produced a summary file, which indicated the time the user clicked, the closeness of the ray to the center of the target at the click, and the overall result (did the user get the target). We additionally created a larger detail log file. This contained data about the target, the users head, and the users hand location for each frame of the simulation.

4 Results

4.1 Data Analysis

A combination of Matlab 2015b, IBM SPSS 23, and SAS 9.4 were utilized to analyze the data. Data was tested for sphericity, using the Maulchly's test, and when the assumption was violated, we used the Greenhouse-Geisser correction on the degrees of freedom. For post-hoc tests of the between-subjects factor, we used the Tukey's test, and for pairwise multiple comparisons, the Bonferroni correction was used.

A 4-way mixed-design factorial analysis of variance (ANOVA) with three within-subjects factors and one between-subjects factor was used. The within-subjects factors were block number, elevation and horizontal direction of the target trajectory. The task contained 10 target trajectories and for each trajectory, the participant performed 5 trials. In the ANOVA, we computed the mean closeness of those 5 trials. This provides a smoother mean per participant per condition [12]. There were three situations in which there were fewer than 5 data points to average. On rare occasions, there would be a signal loss between the wand and the tracker thus freezing the wand. A single trial would be omitted when this happens (18 trials, 0.18%). The second situation is if the button press was an outlier (162 trials, 1.57%). Finally, if the participant did not make a single press during the trial (229 trials, 2.23%).

Table 1: Main effects for training blocks (2-6) with means and standard deviations of angular closeness.

Visual Condition (p<.001 ***)	HFR	M=0.332	SD=0.096
	LP	M=0.426	SD=0.127
	HP	M=0.488	SD=0.208
Block (p<.001 ***)	2	M=0.485	SD=0.200
	3	M=0.449	SD=0.139
	4	M=0.439	SD=0.148
	5	M=0.428	SD=0.148
	6	M=0.399	SD=0.159
Target Direction (p<.005 **)	Center	M=0.348	SD=0.159
	Near	M=0.416	SD=0.130
	Far	M=0.483	SD=0.173
Target Elevation (p<.005 **)	Low	M=0.467	SD=0.160
	High	M=0.384	SD=0.133

Angular Closeness Metric During the experiment, participants received visual feedback when a target was successfully acquired. Traditionally, the only metric of performance would be the hit rate (number of acquisitions in a block) and this parallels the real world game of trap shooting where points are only awarded when a target explodes. However, in a virtual reality setting, we have the advantage of knowing just how close to the center of target the participant was aiming when the button was pressed regardless of whether the attempt was a hit or a miss. Instead of just analyzing whether or not the target was acquired on a trial (a binary variable based on an arbitrary design parameter) we analyzed the distance from the intersection of the ray to the target center when the trigger button was pressed, which gave a continuous performance variable. To properly account for the fact that the target is moving away from the participant, we used the angular distance between the ray and the target center [12]. Henceforth, we will simply refer to this metric of performance as "closeness" (see figure 5).

Last Shot In the task, participants were instructed that, on a given trial, only the first two button presses would be counted towards the target acquisition task. In 97.8% of the trials, at least one button press was recorded and in 39.9% of the trials, a second button press was recorded. For our analysis, we chose to include just the last press on a given trial. Meaning, if the button was pressed only once, that closeness would be utilized in further analysis. If the button was pressed a second time, only the second press was utilized in the analysis and not the first press.

Outlier exclusion The mean closeness was calculated individually for each visual condition using all participants. We considered as an outlier any closeness that was more than 3 standard deviations away from the mean. For the HFR, LP and HP conditions, we excluded 50 (1.44%), 51 (1.48%) and 61 (1.81%) trials respectively.

Data transformation The distribution of closenesses was right-skewed (skewness = 1.42, kurtosis = 2.46). We transformed the data using a natural log to approximate a Gaussian distribution (skewness = -0.92, kurtosis = 2.43).

Collapsing directions across the vertical midline Two participants were left-handed and to account for that, we considered the horizontal directions as ipsilateral/contralateral rather than leftward/rightward. Pairwise comparisons of the 5 horizontal directions in the task showed that no significant differences were found between the two most eccentric trajectories ($p = 1.0$) or between the two near trajectories ($p = 1.0$). Thus we collapsed the directions across the midline to yield 3 directions which refer to as Center, Near, and Far. Henceforth, the results will be reported using the 3 collapsed horizontal directions.

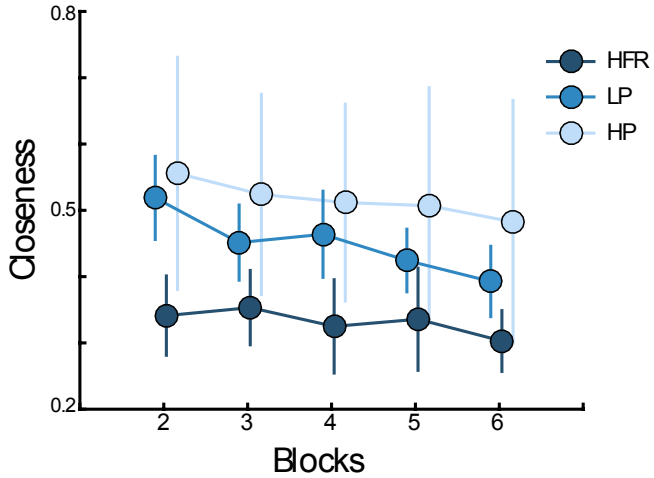

Figure 6: Angular closeness during the training blocks. Plotted are the means and standard deviations for the 10 participants in each of the three visual conditions. The data points are separated horizontally for viewing clarity. Lower closeness scores are better.

4.2 Training Blocks

Looking at the training blocks (2-6) in which the participants performed under the three visual conditions (See table 1) there was a significant main effect of visual condition ($f_{2,27} = 8.87, p < .001$). A post-hoc Tukey HSD test showed that participants performed better with HFR as compared to HP ($p < .001$) and LP ($p < .05$). There was no significant difference between HP and LP ($p = .93$).

There was a main effect of block ($f_{4,24} = 10.56, p < .001$) and no interaction with condition ($f_{8,50} = 1.41, p = .202$). Across the visual conditions, participants steadily improved their performance. A pairwise comparison showed that the performance in block 6 was significantly better than the performance in block 2-5. Similarly, the performance in blocks 5 and 4, were better than that in block 2 ($p < .05$ and $p < .05$ respectively). This indicates that learning occured within the training blocks (see figure 6) . Taken together, the results indicate that participants improved their performance through the training blocks and this rate of training was not different for the three visual conditions.

Our results show that there was a main effect of target direction ($f_{2,26} = 83.89, p < .005$) and further, there was an interaction between direction and visual condition ($p < .05$). Across the visual conditions, participants using HFR had the best performance followed by LP and lastly, HP (see figure 7). Statistically, in the Far direction, HFR was significantly better than both HP ($p < .005$) and LP ($p < .005$). In the Near direction, the same was true for HFR when compared to HP ($p < .005$) and to LP ($p < .05$). While the trend is similar in the Center condition, HFR is not significantly different than LP ($p = .384$) but it is better than HP ($p < .05$) and there is no statistical difference between LP and HP ($p = .748$).

There was a main effect of elevation ($f_{1,27} = 7.60, p < .005$). Participants had significantly higher performance in the upper elevation (see figure 8) as compared to the lower elevation ($p < .005$). Additionally, there was an interaction between elevation and visual condition ($p < .05$). In the lower elevation, participants using HFR performed significantly better than those using HP ($p < .005$) and LP ($p < .005$) and there was no difference better those using HP and LP ($p = .934$). However, in the upper elevation, there was no significant difference between the performance in HFR as compared to LP ($p = .104$) but the performance in HFR was better than that in HP ($p < .05$). There was also no significant difference between

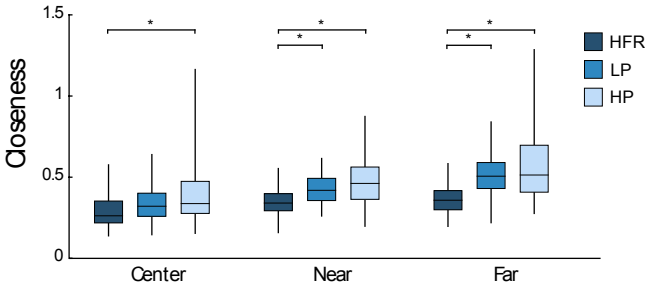

Figure 7: Angular closeness separated by visual condition for the three target horizontal directions. In the Near and Far target trajectories, performance with the HFR condition was significantly better than that with LP and HP. In the Center condition, performance with HFR is not significantly different than with LP.

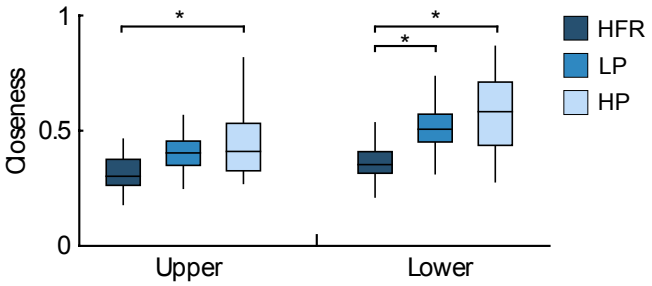

Figure 8: Angular closeness for the three visual conditions separated by the two target elevations. In the Lower elevation, participants using HFR perform significantly better than those using LP or HP. But in the Upper elevation, performance in the HFR condition and that in the LP condition are not significantly different.

LP and HP ($p = 1.0$).

4.3 Learning

In both the first and last blocks, participants performed under HFR conditions. The feature to be examined is the possible effect of training under the different visual conditions in blocks 2-6. Results show that was a main effect of block ($f_{1,27} = 30.05, p < .001$) and not for condition ($f_{2,27} = 2.04, p = .149$).

Just as we saw in the training blocks, there is a main effect of direction ($f_{2,26} = 32.27, p < .001$) and of elevation ($f_{1,27} = 14.74, p < .001$). For both, there was no interaction with visual condition. With regard to direction, participants performed best for the Center as compared to the Near ($p < .001$) and the Far ($p < .001$). Also, participants performed better in the Near as compared to the Far ($p < .005$). Likewise, as it was in the training blocks, participants had better performance in the upper elevation as compared to the lower elevation ($p < .005$).

4.4 Simulator Sickness

Analysis was performed with a non-parametric Kruskal-Wallis test and found no significant effects for total SSQ score or any of the subscores (nausea, disorientation, oculomotor). There was high intersubject variability, and in all conditions there were participants who scored 0 (no sickness). While not significant, the means do show a trend in total sickness and all subscores $HFR < HP < LP$ (see figure 10).

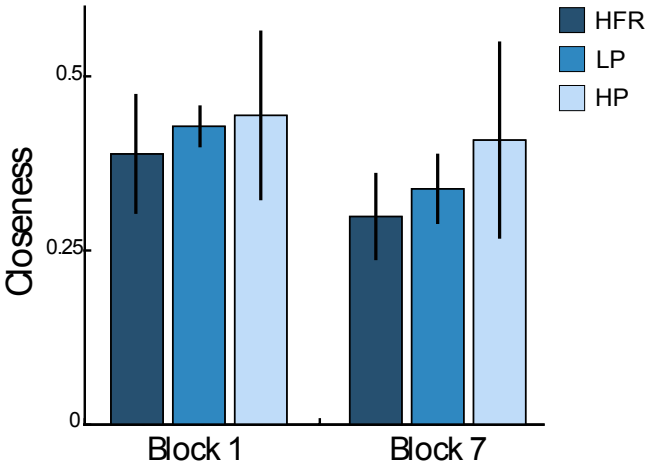

Figure 9: Plotted are the means and standard deviations of angular closeness from the pre-test (block 1) and post-test (block 7). Irrespective of the visual condition in Blocks 2-6, participants performed significantly better in Block 7 as compared to Block 1. There is no significant difference in performance in Block 7 across the three groups.

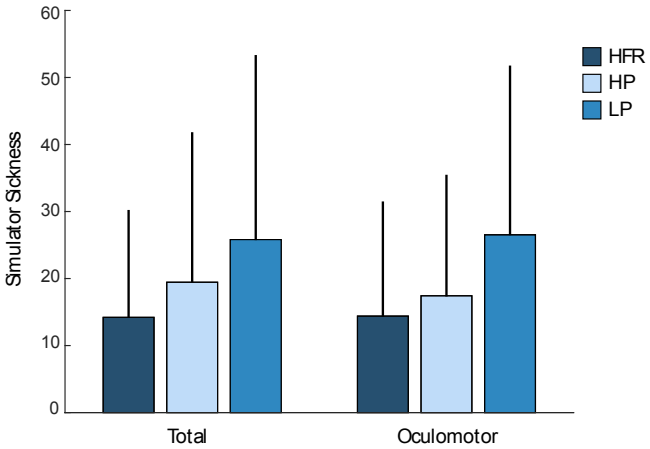

Figure 10: Simulator sickness total and oculomotor subscore for the three different visual conditions. Lower sickness scores are better.

5 Discussion

5.1 Performance on Different Target Trajectories

The interactions of both direction and elevation with condition illuminate where the differences between visual conditions are large and where they are not. The targets that follow the center trajectory are the easiest to acquire judging by it having the highest performance with smallest closeness values. Similarly, the upper elevation appears to be easier than the lower elevation. In the cases where the task is easier to perform, the LP condition recovers the detriments that were introduced due to low frame rates. But even with the easiest conditions, HP does significantly worse than HFR. As the task gets tougher, with peripheral targets and lower elevations, the high fidelity visual feedback is essential in acquiring the target. In the hard cases, participants in the LP condition perform no different than those in HP.

We believe that the reasons for high elevations to be easier in general may be because the user has more time to react due to longer flight time of target. Directional results may be explained by the fact that the center condition involves the manipulation of only one axis, and thus requires fewer motor and cognitive loads. Per-

haps this extra time or reduced cognitive load may have left room for a better refinement step [12].

5.2 Favorability of LP vs HP

For dynamic target acquisition tasks, it seems clear that performing the task at a high frame rate is preferred. However, in easier tasks, a low frame rate system may benefit from a low persistence adjustment. This may be due to the fact that, by adding black frames during the next frame generation time, a "masking" effect occurs, and the user may perceive a smoother motion of the target. The harder difficulties of tasks where the target elevation was low and thus active for a short a period, or too eccentric and thus became visually too small too fast, exposed the actual benefit of HFR over LP. Nevertheless, it seems clear that LP will perform at least as well as HP, but may, in some situations, perform as well as HFR.

5.3 Task Difficulty

We can analyze the task difficulty in light of the distal pointing model proposed by Kopper at al. [12]. At the beginning of the task, targets were rather large, with a diameter of approximately 2°. When targets reached their farthest point, the diameter ranged from approximately 0.59°on high elevations ($\sim 1.8s$ after launch) to approximately 0.69°on low elevations ($\sim 1.3s$ after launch). According to Kopper's distal pointing model, the index of difficulty (ID_{DP}) of the task at the moment of launch was rather small (~ 0.5), and the task got progressively more difficult as the target flew away, reaching, at the farthest distance, $ID_{DP} \sim 43$ for low elevations and $ID_{DP} \sim 52$ for high elevations. If we look at the selection time, the predicted movement time (MT) was approximately $1.1s$ at target launch and increased over time, up to $MT \sim 2.1s$ for low elevations and $MT \sim 2.5s$ for high elevations.

This presented a paradox to the users. The predicted time for successful target selection was always higher than the time the target was actually in flight (1.32s low, 1.80s high elevation). This, coupled with the fact that targets were not only changing in visual size, but also moving in different directions, caused the users to have to adapt to a time critical task, where they had to favor time over accuracy in order to have a chance of hitting the target. Basically, the nature of the trap shooting simulation made the task too difficult to accomplish and most targets were missed. For this reason, we decided to analyze task performance based on the closeness of the target, rather than on success of the trial.

5.4 Angular Closeness Metric

In the real world task of trap shooting, competitors are not awarded any points unless the targets are hit. Similarly, in our task, visual and auditory feedback were only provided when the participants successfully hit the target, which had an arbitrary size. The nature of the trap shooting simulation, however, made the task too difficult and most targets were missed. This happened because of an interesting trade-off. Time was necessary for users to lock on a target before shooting, but that very time caused targets to become increasingly smaller in visual size, increasing the tasks' index of difficulty exponentially [12]. Ultimately, using "hit" as a metric of performance turned out to have too low sensitivity, as performance differences were washed out due to the difficulty of the hit task.

To address this issue, we decided to use the angular distance from the intersection of the ray with the target plane to the target center, which we referred to as closeness, as the main analysis metric. Essentially, whether the ray hit the target or not became irrelevant to the analysis. One limitation of this approach is the feedback that participants received when a target was hit, which may have confounded the analysis of closeness due to a possible motivation factor. This assumption, however, is diminished by the fact that closeness got consistently smaller over training blocks for all visual conditions. Also, as participants trained, both hit rate increased and closeness values decreased marking the benefits of practice.

As the goal of this study was to understand the effect of target acquisition performance under different visual conditions, we also make the point that closeness is a better metric than hit rate, as the results on distance to the target center can be generalized to targets of any size.

5.5 Learning Effects

While users display lower performance while training in degraded conditions, when they are assessed in HFR for the final training block, the performance is comparable to those who trained in HFR. All conditions showed learning equally well. This was evidenced when looking at the assessment trials (pre and post) where there was appreciable training occurring through the experiment. The extent of such training did not depend on the visual condition the participants performed. This indicates that the task of dynamic target acquisition using a high fidelity frame rate can be trained using low fidelity conditions and still get comparable levels of benefit.

One possible explanation is that users are able to learn the target flight paths and timing, even under degraded conditions. Another explanation may come from the psychophysics literature, where it has been shown that in many situations, training under degraded visual conditions causes an actual improvement of task performance in normal conditions [4, 5, 18].

5.6 Simulator Sickness

Contradicting the results of the study presented here, previous low persistence research found significant differences on the oculomotor sub-scores of the SSQ for LP [22]. Although we found trends showing LP with higher oculomotor subscore, no significant differences were found. This may be due to the fact that previously, a within-subjects study was performed, such that participants had their own ratings to compare to when rating their level of simulator sickness. Another possibility is that the previous experiment contained a navigation task, whose high optical flow may have caused the increased oculomotor subscores.

5.7 Advantages and Limitations of Low Persistence

When the option of HFR is unavailable, LP seems to provide an environment that allows comparable performance to HFR in high elevation and center horizons. In those conditions HP performs significantly worse than HFR. Also we see overall trends in the training blocks (figure 6) and the pre-post analysis (figure 9) about LP approaching HFR in performance when compared to HP. Finally we have an unexpected variability that can be seen in the standard deviations in participant performance who utilized HP, while LP was more consistent across participants. It seems that there may be an advantage to utilizing LP in a dynamic target acquisition task. Future work should be done to gain more definitive answers about the usefulness of LP.

The low persistence technique poses some clear limitations over traditional low frame rate VEs. Although there is a perceived smoothness of motion introduced by the blank frames masking effect, it introduces a stroboscopic effect, which is visually fatiguing. Moreover, by replacing content frames with black frames, there is an overall decrease in visual brightness for the user. We addressed some of these limitations by taking steps to carefully design the VE with lower contrast between the blanking frames and the world. For realistic task scenarios, future work should consider exploring other image degradation techniques, such as blurring and noise insertion to reduce the high contrast induced by LP.

5.8 Study Limitations

The study was limited by the small number of participants tested (10 per condition). A larger number of participants would have

provided increased power of the analyses. The study could have also benefited and gained greater statistical power from a within-subjects design and a reduced number of independent variables being tested. Finally, the task may have been too difficult, and future work should consider larger targets (or perhaps targets of varying size).

6 Conclusion and Future Work

We have presented a user study evaluating the effects of image persistence on a dynamic target acquisition task. Results found that learning occurs consistently regardless of the visual condition used for training. This indicates that training of dynamic target acquisition under degraded visual conditions transfers well to performance under a baseline high frame rate task.

While the low persistence display technique was found to perform no different than the traditional low frame rate high persistence technique, under some circumstances low persistence approached the baseline high-frame rate condition. Specifically, we observed that when tasks are not too difficult (high elevation, center horizon), it may be beneficial to use a low persistence adjustment of the display frames in order to improve dynamic target acquisition in a low frame rate virtual environment.

Ultimately, the results from this study indicate that there are some situations in which low persistence should be considered as a candidate to potentially increase performance in low frame rate virtual environments, but that needs to be measured with increased levels of simulator sickness, particularly if the application contains high optical flow. In those cases, and where target's movements are reduced, using the traditional high persistence technique may be preferred.

We identify two areas of valuable future work moving forward. First, more research needs to be done better understand VR training under degraded visual conditions, and whether and how it can be used to improve training experience. Second, there is an opportunity for investigating what VE tasks may benefit the most from low persistence adjustments to low frame rate virtual environments.

References

[1] International shooting sport federation rule book. `http://www.issf-sports.org/documents/rules/2013/ISSFRuleBook2013-3rdPrint-ENG.pdf`. Accessed: 2015-09-21.

[2] John Carmack's Time Warp Technique, howpublished = `http://www.roadtovr.com/oculus-rift-sdk-v0-3-1-preview-distortion-john-carmack-timewarp/`, note = Accessed: 2015-09-21.

[3] Oculus Rift DK2. `https://www.oculus.com/en-us/dk2/`. Accessed: 2015-09-21.

[4] L. G. Appelbaum, M. S. Cain, J. E. Schroeder, E. F. Darling, and S. R. Mitroff. Stroboscopic visual training improves information encoding in short-term memory. *Attention, Perception, & Psychophysics*, 74(8):1681–1691, 2012.

[5] L. G. Appelbaum, J. E. Schroeder, M. S. Cain, and S. R. Mitroff. Improved visual cognition through stroboscopic training. *Frontiers in psychology*, 2, 2011.

[6] D. A. Bowman, L. F. Hodges, D. Allison, and J. Wineman. The educational value of an information-rich virtual environment. *Presence: Teleoperators and Virtual Environments*, 8(3):317–331, 1999.

[7] D. A. Bowman, A. Sowndararajan, E. D. Ragan, and R. Kopper. Higher levels of immersion improve procedure memorization performance. *Proceedings of the 15th Joint Virtual Reality Eurographics Conference on Virtual Environments*, pages 121–128, 2009.

[8] C. Cruz-Neira, D. J. Sandin, and T. A. DeFanti. Surround-screen projection-based virtual reality: the design and implementation of the CAVE. In *Proceedings of the 20th annual conference on Computer graphics and interactive techniques*, pages 135–142. ACM, 1993.

[9] S. R. Ellis, F. Bréant, B. Manges, R. Jacoby, and B. D. Adelstein. Factors influencing operator interaction with virtual objects viewed via head-mounted see-through displays: viewing conditions and rendering latency. In *Virtual Reality Annual International Symposium*, pages 138–145. IEEE, 1997.

[10] W. A. Hershberger and J. S. Jordan. The phantom array: a perisaccadic illusion of visual direction. *The Psychological Record*, 48(1):2, 2012.

[11] R. S. Kennedy, N. E. Lane, K. S. Berbaum, and M. G. Lilienthal. Simulator sickness questionnaire: An enhanced method for quantifying simulator sickness. *The international journal of aviation psychology*, 3(3):203–220, 1993.

[12] R. Kopper, D. A. Bowman, M. G. Silva, and R. P. McMahan. A human motor behavior model for distal pointing tasks. *International journal of human-computer studies*, 68(10):603–615, 2010.

[13] B. Schaeffer and C. Goudeseune. Syzygy: native PC cluster VR. In *Virtual Reality Conference Proceedings*, pages 15–22. IEEE, 2003.

[14] H. Schäfer, M. Niessner, B. Keinert, M. Stamminger, and C. Loop. State of the art report on real-time rendering with hardware tessellation. *Eurographics State of the Art Reports*, pages 93–117, 2014.

[15] A. R. Seitz, J. E. Nanez, S. R. Holloway, and T. Watanabe. Visual experience can substantially alter critical flicker fusion thresholds. *Human Psychopharmacology: Clinical and Experimental*, 20(1):55–60, 2005.

[16] N. E. Seymour, A. G. Gallagher, S. A. Roman, M. K. OBrien, V. K. Bansal, D. K. Andersen, and R. M. Satava. Virtual reality training improves operating room performance: results of a randomized, double-blinded study. *Annals of surgery*, 236(4):458, 2002.

[17] F. A. Smit, R. Van Liere, and B. Froehlich. A programmable display layer for virtual reality system architectures. *Visualization and Computer Graphics, IEEE Transactions on*, 16(1):28–42, 2010.

[18] T. Q. Smith and S. R. Mitroff. Stroboscopic training enhances anticipatory timing. *International Journal of Exercise Science*, 5(4):4, 2012.

[19] D. L. Tate, L. Sibert, and T. King. Using virtual environments to train firefighters. *Computer Graphics and Applications, IEEE*, 17(6):23–29, 1997.

[20] D. Wang, A. Vincent, P. Blanchfield, and R. Klepko. Motion-compensated frame rate up-conversionpart ii: New algorithms for frame interpolation. *Broadcasting, IEEE Transactions on*, 56(2):142–149, 2010.

[21] C. Ware and R. Balakrishnan. Reaching for objects in vr displays: lag and frame rate. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 1(4):331–356, 1994.

[22] D. Zielinski, H. Rao, M. Sommer, and R. Kopper. Exploring the effects of image persistence in low frame rate virtual environments. In *Virtual Reality Conference Proceedings*, pages 19–26. IEEE, 2015.