Stimulating the Human Visual System Beyond Real World Performance in Future Augmented Reality Displays

David Dunn
UNC Chapel Hill

Okan Tursun
MPI Informatik, Germany
Università della Svizzera
italiana, Switzerland

Hyeonseung Yu
MPI Informatik,
Germany

Piotr Didyk
Università della Svizzera
italiana, Switzerland

Karol Myszkowski
MPI Informatik,
Germany

Henry Fuchs
UNC Chapel Hill

(a) A conceptual example of an augmented reality scene with annotations at both near and far depths requiring a gaze change in both direction and depth illustrating the prevalence of depth switching in real world scenarios.
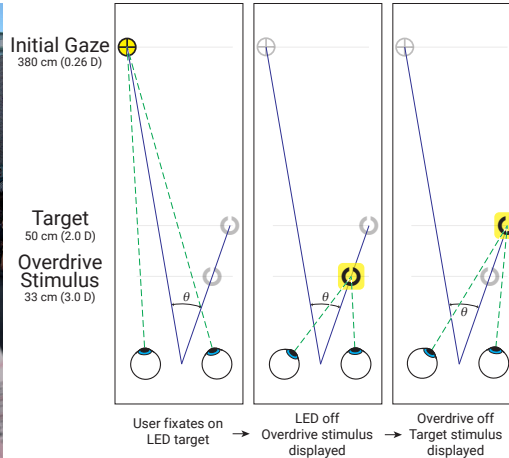
(b) Illustration of steps in the user study described in Section 5.2. First a user fixates on a distant LED target. Then the LED is turned off and an overdrive stimulus is displayed at a distance closer than the target for a specific duration. The overdrive stimulus is then turned off and the target stimulus is displayed for a specific duration before turning off.
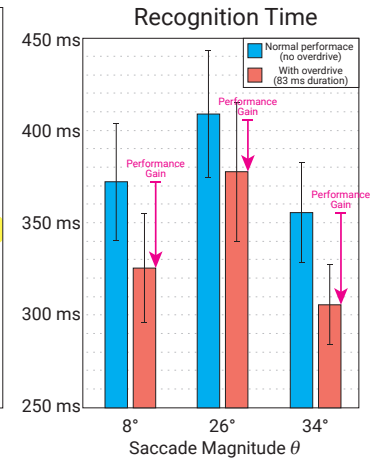
(c) Results of the user study described in Section 5.2 showing the decrease in total recognition time when an overdrive stimulus is added.

ABSTRACT

New augmented-reality near-eye displays provide capabilities for enriching real-world visual experiences with digital content. Most current research focuses on improving both hardware and software to provide digital content that seamlessly blends with the real world. This is believed to not only contribute to the visual experience but also increase human task performance. In this work, we take a step further and ask the question of whether the capabilities of current and future display designs combined with efficient perception-inspired content optimizations can be used to improve human task performance beyond the human capabilities in the natural world. Based on an in-depth analysis of previous literature, we hypothesize here that such enhancements can be achieved when the human visual system is provided with content that optimizes the oculomotor responses. To further investigate possible gains, we present a series of perceptual experiments that built upon this idea. More specifically, we focus on speeding up accommodation response, which significantly contributes to the eye-adaptation when a new stimulus is shown. Through our experiments, we demonstrate that such speedups can be achieved, and more importantly, they can lead to significant improvements in human task performance. While not all of our results give definite answers, we believe that they reveal plentiful opportunities for further enhancing the human experience and task performance when using new augmented-reality displays.

1 INTRODUCTION

Augmented reality (AR) displays offer novel methods of human-computer interaction, the interplay between complex technology and human subjects provides a rich field for many different avenues of research. In combining the real world with virtual elements, whether they be annotations of real objects (Figure 1a), spatially locked virtual objects, or even the diminishing of real objects, AR has the ability to replicate many visual cues, so very realistic stimuli may be generated. By definition, optical see-through (OST) AR is capable of injecting additional visual signal, which both mimics the real-world image and manipulates it. Therefore, the question arises: to what extent can we control and manipulate the visual signal beyond just showing virtual content, and what benefits can it have?

It has been shown that contextual, spatially aligned information made available in AR displays can improve performance on informationally-dense psychomotor tasks [23]. The most straightforward way of providing such an enhancement can be a system where important visual information, e.g., instructions, is made easily accessible or visible by a human observer. Such a goal can be achieved, for example, by enlarging a small portion of the real world - zoom capability, to enable a user to complete a task in a shorter amount of time. The above example relies on affecting the conscious visual perception, which controls, in this case, the gaze of the subject. We hypothesize that it is possible to also affect the unconscious visual

perception [33,54,65] by providing an additional visual signal which is not directly perceived by an observer, but enables performance enhancements beyond those that are based on the presentation of relevant data. Such enhancements may take many forms: improved visual acuity, decreased functional loss, and decreased reaction time. The mechanisms that may lead to such enhancements range from simple ones, such as early pupil dilation when transitioning from dark to bright environments, to more complex ones, such as increasing or suppressing specific spatial frequencies. These techniques should work by displaying a specific stimulus that leverages an existing mechanism of the human visual system (HVS) for improved performance, and ideally, they would operate to enhance the see-through image in both augmented and non-augmented environments.

This work aims to move beyond improvements caused by the better presentation of relevant data and improve the performance of the HVS, irrespective of content. To this end, we present the idea that the unique nature of OST AR displays creates opportunities for enhancement of the HVS beyond its capabilities in the natural world. While we believe that many improvements are worthy pursuits, this work targets explicitly improving human performance by decreasing the time the HVS needs to adapt to a new fixation location. More precisely, we focus on speeding up focal and fixation depth changes that are very common in free viewing, but relatively slow. They often require hundreds of milliseconds for the full transition, and as shown in this paper, these times correlate with human performance in simple tasks such as object recognition.

To achieve our goal, we make two key observations. First, it has been demonstrated that the main eye movements, i.e., saccade, vergence, and accommodation, can be significantly faster if they occur simultaneously [17, 19, 57]. This suggests that by providing content which sufficiently stimulates all the movements, the eye transition to a new target can be facilitated. Second, for each of these eye movements, the responses have a higher peak velocity when the motion is larger. This suggests that by initially presenting a stimulus beyond the intended target, and then backing off to the actual target, higher transition speeds may be achieved (Figure 1b). We call this mechanism *overdriving* the visual cues and hypothesize that it can further speed up eye adaptation to new depth targets (Figure 1c).

In this paper, we investigate the above ideas in a series of user experiments. While it is possible to measure eye response time directly, we argue that just because the physiological responses occur at an increased speed does not mean the overall perception, and in particular human performance, is also sped up. For example, it may take the HVS longer to program the motion, meaning a longer latent period at the onset, or a period of diminished perception similar to saccadic suppression may occur. Therefore, only by measuring task completion time, it is possible to know if the physiological speedups lead to perceptual speedups. Consequently, in our experiments, we measure the full time that is required for completing a given task. This allows us to draw conclusions that are relevant for practical applications in AR displays. The main contributions of this work include:

- an investigation if physiological response speedups lead to improved human performance,

- a review and analysis of previous oculomotor dynamics studies revealing potential paths for enabling faster depth transitions,

- task-based user studies combining saccade, vergence, and accommodation dynamics to measure the HVS response times,

- evidence of possible performance gains (12.5%) when overdriving visual cues, and

- an outline of future research on further improving the HVS response during refocusing.
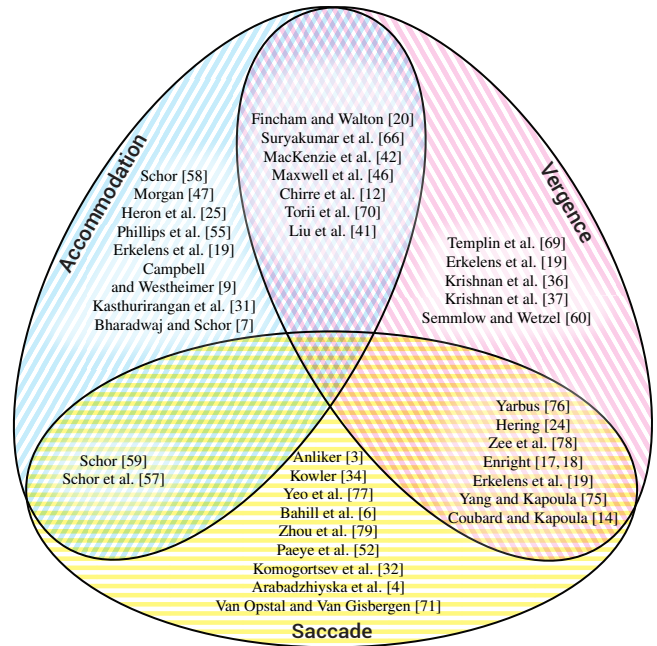

Figure 2: Taxonomy of oculomotor dynamics studies. Note the lack of studies at the intersection of all three responses.

2 BACKGROUND

In real-world viewing conditions, vergence, accommodation and saccade are closely coordinated [48, 76, 78]. Due to the complex nature of the interaction of various cues, dynamics of each process are mostly studied independently and we present a summary of those works in Sections 2.1, 2.2, and 2.3. A depth switch mandates both a change of vergence and accommodation, and the interaction between the two have been well-studied, especially with the recent wide distribution of virtual reality (VR) and AR head-mounted displays [26, 35, 62]. However, measuring their dynamics in concert is limited by instrumentation as presented in Section 2.4. Saccades do not occur with all depth switches, but are usually present, and interestingly have quite a significant impact on the other two responses as discussed in Sections 2.5 and 2.6. We further present in Section 2.7 the apparent lack of research performed when all three responses are combined, and discuss other factors related to this work. In general, the responses can be characterized by an initial latent period, velocity, and acceleration. Figure 2 summarizes the taxonomy of oculomotor dynamics studies that we discuss in this work. In contrast to studies that directly measure the dynamics, studies that aim to measure task-based performance are examined in Section 2.8. Compared to those studies, our work is dedicated to leveraging the physiological speedups due to combined eye movements with the aim of improving task performance. While the concept of using AR displays for improving the HVS performance is new, there have been previous display prototypes which enhance real-world perception and we discuss them in Section 2.9.

2.1 Accommodation Dynamics

The initial reaction time or latency for accommodation has been observed in the range of 300 ms to 500 ms [7,9,25,47,55,58]. While Phillips et al. [55] have observed latencies as short as 200 ms for accommodation, the probability of their occurrence is very low.

Once the accommodation is initiated, the lens accommodation dynamics can be mainly characterized by the velocity and total duration. Bharadwaj and Schor [7] observed a smooth increase in velocity to its peak value and then its slightly slower reduction to

a steady state. As accommodation magnitude increased, so did the peak velocity – with a maximum value of around $10\,\mathrm{D(diopter)/s}$. Kasthurirangan et al. [31] observed a similar average peak velocity for the lens accommodation, but a high variance can be observed in their data. Also, for near-to-far accommodation, peak velocities over $20\,\mathrm{D/s}$ have been measured for the large accommodation magnitudes of $4\,\mathrm{D}$ to $5\,\mathrm{D}$. The duration of lens accommodation is in the range of $500\,\mathrm{ms}$ to $800\,\mathrm{ms}$ [7, 9, 19, 25, 47, 55, 58].

The interesting feature of the accommodation dynamics is that near-to-far accommodation is independent of the destination depth [8], while the response of the far-to-near accommodation depends on the target depth [7]. In particular for the latter case, the total duration of the acceleration increases with the dioptric difference between the current focal state and target depth. Therefore when over-driven stimulus is provided, it is expected that the peak velocity increases during the initial phase of the accommodation.

2.2 Vergence Dynamics

The latency for vergence has been reported in the range of $150\,\mathrm{ms}$ to $200\,\mathrm{ms}$ [36, 37]. Erkelens et al. [19] report a maximum vergence velocity of convergence as $120\,^\circ/\mathrm{s}$ to $190\,^\circ/\mathrm{s}$ and divergence as $160\,^\circ/\mathrm{s}$ to $180\,^\circ/\mathrm{s}$ for 32° vergence magnitude. Mean vergence velocities for the same magnitudes were reported as $38\,^\circ/\mathrm{s}$ to $59\,^\circ/\mathrm{s}$ for convergence and $45\,^\circ/\mathrm{s}$ to $53\,^\circ/\mathrm{s}$ for divergence.

Response time is dependent on the distance traveled and the direction of the fixation change, either far-to-near accommodation and convergence, or near-to-far accommodation and divergence. The duration of convergence and divergence are typically within $200\,\mathrm{ms}$ to $800\,\mathrm{ms}$ [60, 69].

2.3 Saccade Dynamics

Saccades are the fastest type of eye movements, which take place in approximately $20\text{-}80\,\mathrm{ms}$ depending on their magnitude. They are preprogrammed by oculomotor mechanisms of the brain and their duration and velocity cannot be voluntarily controlled [34]. This property of saccades makes them easy to model and predict. Bahill et al. [6] introduced one of the earliest models which explain the relationship between saccade magnitude, duration and its peak velocity (a.k.a. *the main sequence*). Their study shows a linear relationship between the saccade magnitude and the peak velocity, which reaches a saturation limit around $650\,^\circ/\mathrm{s}$ for saccades longer than $15\text{-}20^\circ$. Later, Anliker [3], Paeye et al. [52] and Yeo et al. [77] showed that the change in the velocity during a saccade can be expressed as a bell-shaped curve, which is symmetric around the peak velocity. However, the symmetry assumption holds only for small saccade magnitudes, because velocity profiles get more skewed for longer saccades, such that the peak velocity is reached before the midpoint of a saccade [71]. Arabadzhiyska et al. [4] recently showed that displacement profiles explain the dynamics of longer saccades accurately enough to predict the landing position during a saccade. There is also a growing body of research which introduced biologically inspired models of saccade dynamics [32, 79]. Compared to aforementioned statistical methods, these models are relatively more complex and have a large number of parameters but they are capable of explaining the role of underlying anatomical and neurological mechanisms in the execution of saccades.

2.4 Interaction of Accommodation and Vergence

Although isolated measurement of dynamics of accommodation and vergence provide useful insights into the behavior of the human visual system, more complex interactions are observed in real-world viewing conditions. The vergence and accommodation of human vision are neurally coupled [20], however it is challenging to study the complete picture of interaction between vergence and accommodation due to the technical difficulty of measuring all those cues simultaneously. A possible solution is to keep the stimuli always in front on one eye for which its accommodative state has been sampled by an autorefractor, while for another eye the stimuli have been drawn consistently at the position consistent with the target vergence distance [42, 46]. Recently, an advanced experimental setup to measure the accommodation in the presence of the gaze change have been proposed [12, 41, 66], but little investigation has been done regarding the interaction between vergence and accommodation. One important insight from these studies is that the accommodation response becomes faster under binocular viewing conditions compared to monocular case [12]. This can be attributed to the dynamics of the vergence-accommodation cross-links that cause vergence and accommodation to respond when either one is simulated [70].

2.5 Interaction of Vergence and Saccade

One of the earliest hypotheses regarding the dynamics of saccade and vergence movements assumes that there is no interaction between their neurological mechanisms in the human brain, and the combined eye movements can be explained by a simple addition [24, 76]. However, Enright [17, 18] showed that for a saccade between $0.9\text{-}4^\circ$ and a change in vergence between $0.7\text{-}2.3^\circ$, $40\text{-}90\,\%$ of vergence is executed during the saccade. This observation is several folds larger than what is predicted by the additivity hypothesis, and it clearly shows that saccades and vergence are not summed, but they interact. Later, Erkelens et al. [19] also confirmed the facilitation of vergence by saccades. In their experiments, they observed that $95\,\%$ of divergence and $75\,\%$ of convergence were performed during the saccade when a 45° saccade and 11° vergence were combined. That facilitation reduced the duration of divergence from $500\,\mathrm{ms}$ to $220\text{-}280\,\mathrm{ms}$. A similar speedup was also observed for convergence, although it was much less pronounced. Unfortunately, the number of participants in both Enright and Erkelens et al. are limited to 2–3.

One interesting observation in the study of Erkelens et al. is that even when the targets aligned to elicit pure vergence movements, saccades still occur frequently (e.g., when fixation points are aligned along the median plane). In their study of saccade-vergence interactions, Zee et al. [78] also made this observation; however, they decided to discard the majority of such vergence movements from their analysis due to "contamination" by saccades. Coubard and Kapoula [14] quantified those saccades in detail, and they identified six different patterns of saccade trajectories that naturally occur in approximately $84\text{-}97.8\,\%$ of symmetric vergence movements.

Yang and Kapoula [75] conducted a detailed analysis of vergence facilitation by measuring durations, peak, and mean velocities of eye movements for combined saccades (20° magnitude) and vergence (15° magnitude). Their measurements show that vergence movement speeds up by $7\text{-}22\,\%$, and saccade reciprocally slows down by $50\,\%$, which results in $20\text{-}110\,\mathrm{ms}$ reduction in the total duration of the eye movement. However, their measurements show a high variability both between participants and across different trials for the same individual.

2.6 Interaction of Saccade and Accommodation

When accommodation is performed in conjunction with saccades, both the reaction latency and the response duration of accommodation decrease. Schor [59] reports that the period of latency for accommodation is reduced by $13\,\%$ and the velocity of accommodation response increased by $27\,\%$ when a simultaneous saccade is performed. In absolute terms this amounted to accommodative latencies $100\,\mathrm{ms}$ to $300\,\mathrm{ms}$ shorter and accommodation velocities $1.5\,\mathrm{D/s}$ to $2\,\mathrm{D/s}$ faster for the saccade condition than the no-saccade condition.

Schor et al. [57] show that transient vergence responses associated with saccades did not affect accommodation when not stimulated by defocus. This observation suggests that the saccadic enhancement of accommodation did not result from vergence driven accommodation, therefore one cannot benefit from such interactions with

traditional head-mounted displays (HMDs) having a fixed accommodation depth. In our study, we run the task-driven measurements on multi-focal plane setups in which the multiple physical accommodation depths are supported.

2.7 Accommodation, Vergence, and Saccade Interactions

To our best knowledge, there has been lack of studies directly measuring the combined dynamics of accommodation, vergence, and saccade. As discussed in Section 2.4, this is mainly due to technical difficulties of measuring accommodation changes in the presence of vergence and saccade. In particular when saccade is present, the optical path of accommodation measurement system should follow the gaze change governed by the fast dynamics of saccade, which requires high-speed tracking and imaging systems. This technical challenge motivates us to study the interactions of those dynamics through task-based measurements that implicitly consider the effect of accommodation, vergence and saccade simultaneously. Therefore, our task-driven measurement allow us to study those mechanisms all together without building sophisticated measurement devices.

All of the studies mentioned so far measure the physiological speeds of the responses through use of instrumentation. In the past, there has been some interest in investigating the perceptual aspect of this problem for a very specific case of abrupt changes of disparity in videos [49]. But it is still unclear that the physiological speed correlates perfectly with perceptual speed. Thus it is essential, with our application, to run task-based experiments to verify that the timings which are reported are not just physiological but are actually perceptual.

2.8 Task-based Experiments in AR

Kruijff et al. [40] have previously surveyed the perceptual issues in AR and their potential implications on task performance such as poor depth judgement and slow object recognition. They identified the causes of poor task performance such as the hardware limitations in resolution, field of view, color and contrast reproduction, the effects of poor augmentation due to registration errors and increased visual clutter as well as individual differences in users' perception.

More recent works focused on specific tasks for investigating the impacts on performance. Smith et al. [64] compared the visual search performance on Head-Up (HUD) and Head-Down (HDD) displays while driving. They observed significantly faster response times when reading unstructured bodies of text from HUDs. Swan et al. [67] studied perceptual depth matching performance because it is a crucial task in many medical and industrial AR applications that require simultaneous interaction with real and virtual objects. They observed systematical overestimation of depth in the presence of collimating optics. Later, Singh et al. [63] used a custom built AR haploscope as their testbed and reproduced previous findings of Swan et al. [67]. They also investigated the effect of age-related loss of accommodative ability and found no degradation in depth matching performance due to age. Recently, Gabbard et al. [22] studied the effect of frequent context switching and refocusing between real world and a monocular optical see-through display with adjustable focal distance. The majority of the participants reported eye fatigue and had poor task completion performance due to repetitive context and focal distance switching. In a later study, Arefin et al. [5] were able to replicate those findings on a custom built AR haploscope including both monocular and binocular viewing conditions. These studies measure AR task performance for specific applications, while the focus of our task-based experiments is to quantify the amount of speedup in physiological and neurological processes of combined eye movements.

2.9 AR Displays for Real World Enhancement

OST AR displays occupy a unique space: they act as display de-

Figure 3: View of physical display hardware at Location A for running experiments on binocular accommodation and vergence overdriving as described in Section 3.

vices by presenting virtual content to the wearer, but they also act as passive glasses for viewing the real world. While most displays seek to modify the real-world view as little as possible, some displays manipulate the view purposefully. Chakravarthula et al. [11] leverage the real world view by providing an additional dynamic focal power lens capable of enhancing the vision of myopic, hyperopic, and presbyoipic users. Wetzstein et al. [73] employ a transparent spatial light modulator to attenuate light for enhancing perception in a variety of ways including contrast enhancement and reduction, attention calling, and various forms of color alteration. Our experiments look for additional methods OST AR displays can employ in improving perception of the real world.

3 EXPERIMENTAL HARDWARE

In order to test various strategies for stimulating the human visual systems, our strategy sought to determine the period of time required for recognizing visual stimuli near the visual acuity threshold for different cases of depth switching. Thus we required displays with multiple depth support and angular resolution greater than the human eye. There were several candidate display technologies capable of providing our depth needs.

Lightfield displays can provide full-depth volumetric images [28, 29, 43], but the spatial resolution is significantly below the human visual acuity. Holographic displays are capable of presenting 3D images with accurate focus cues [44, 61]; however, the narrow eyebox makes perceptual experiments extremely challenging. Varifocal displays support multiple depths dynamically [2, 16], but the latency required to switch between different depths prevents testing response times of the human visual systems shorter than the depth transient time. Multiplane displays [1, 56] were best suited for our needs; however, the maximum resolution reported in the literature is approximately 15 cpd [27], which is far below human visual acuity.

As some of the motivations of our work are to find strategies for stimulating visual systems in future display devices and to serve as guidance to new display research, we decided not to limit ourselves to existing display prototypes. Therefore, we built our own display prototypes using multiplane optics to achieve both multiple focal-depth support and near-retina resolution. Photographs of one of the custom-built multiplane displays are shown in Figure 3.
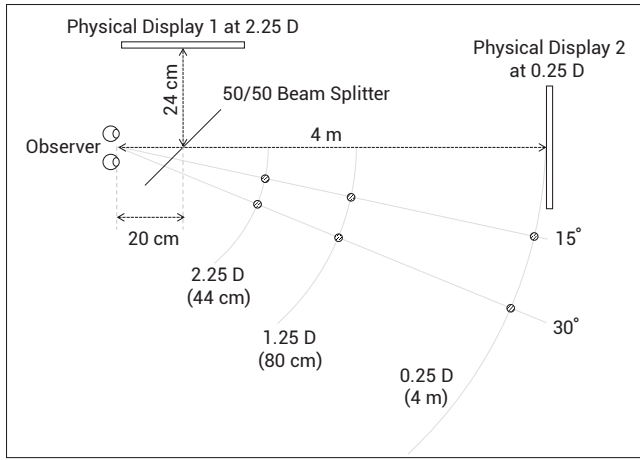
Figure 4: Top view of the experimental configuration used to study the recognition time during combined accommodation, vergence, and saccade.
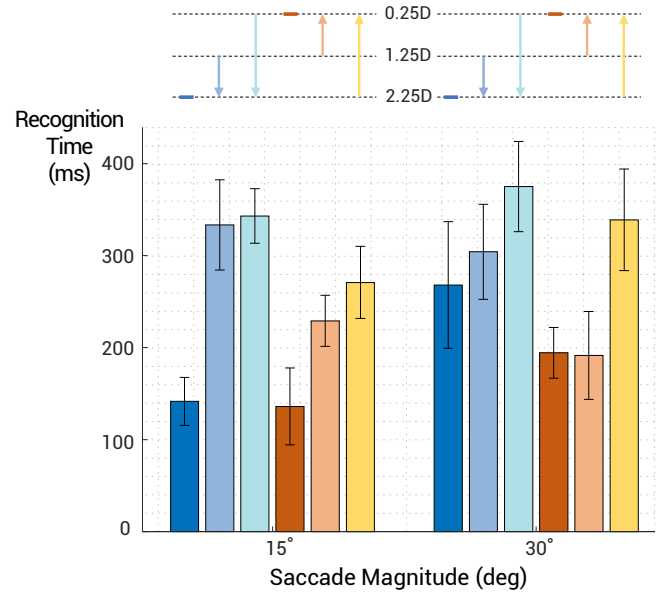

Figure 5: Mean recognition times for various saccade magnitudes and accommodation changes for experiment described in Section 4. The upper diagram shows the accommodation changes for the corresponding bars in the bottom. The error bars represent the Standard Error of the Mean (SEM).

4 MEASURING THE RECOGNITION TIME DURING COMBINED ACCOMMODATION, VERGENCE, AND SACCADE

We first conducted a task-based perceptual experiment[1] to measure the latency of pattern recognition when accommodation, vergence and saccade all work in concert. In the results subsection, we compare the measured latency to the values reported from previous physiological measurements.

Stimulus   We used a Landolt C, which is a ring with a gap at one of the four different positions (left, right, top or bottom) as the target stimulus. The stroke width and the gap size are $1/5$ of the external diameter. It is a standardized symbol commonly used in visual acuity tests, where observers are asked to look at the shape from a distance and identify the random position of the gap. The minimum gap size that can be correctly resolved by the observer in acuity tests is used to measure their visual acuity level. It is a common practice to use a gap size of $1'$ (arc minute) to test for normal vision under ideal viewing conditions. In our experiments, our goal was to use a Landolt C as an indicator for the restoration of full visual sensitivity at the end of eye movements and prevent an early detection of the gap position in the presence of defocus blur (ongoing accommodation), diplopia (ongoing vergence), retinal motion blur (during saccade), reduced visual sensitivity (due to saccadic suppression) and during ongoing cognitive processes which take place in the brain after completing the eye movement. During a pilot study, we noticed that the participants could not resolve a gap size of $1'$ even during a direct fixation on the stimulus. We relate this to the absence of standard viewing conditions suggested by the International Council of Ophthalmology, such as conducting the experiment in a dark room instead of ambient luminance levels between 85-300 cd/m^2, loss of stimulus contrast due to beam-splitter transmission and having a foreground-background color combination different from the recommendation (bright on dark instead of dark on bright) [51]. Therefore, we opted for a slightly larger stimulus with a gap size of $2'$ in our experiment.

Experimental Configuration   We used a two-plane display equipped with light-emitting diodes (LEDs) as shown in Figure 4 to conduct our experiment. The two physical displays were located at 2.25 D and 0.25 D, and they were optically combined by a large beam splitter (30"×24") so that displays at each depth could be

observed by left and right eyes simultaneously. The displays were gamma corrected and their peak luminance was adjusted to the same level (20 cd/m^2 including the attenuation effects of beam splitters). Six LEDs were located at two different eccentricities (15° and 30°) and three different depths (2.25 D, 1.25 D and 0.25 D). A light diffuser with a cross was placed in front of the LEDs to aid the observers focusing on the targets. The luminance of the LEDs was adjusted to the same level as the peak luminance of the displays (20 cd/m^2). The experiment was conducted in a dark room and a chin-rest was used for fixing the head position.

Participants   Twelve paid undergraduate and graduate students participated in this experiment. Their age ranged between 20 and 31 years, and four of them were female and eight of them were male. All had normal or corrected-to-normal vision.

Procedure   One of the the LEDs was turned on and the participant was directed to fixate upon it. The light remained on for 1–2 sec to allow users to accommodate to the depth of the light, then a Landolt C stimulus was shown on one of the displays for a duration of T. We used the vPEST [21] procedure to adaptively select T in each trial and determine the duration required for the observer to identify the position of the gap in the stimulus. Our experiment was based on four alternative forced choice (4AFC) paradigm where the gap position of the stimulus was randomized for each trial. The target detection threshold was set to one just noticeable difference (1 JND), which corresponds to 62.5 % detection rate. For the twelve different configurations of initial LED positions and target display planes, we run twelve vPEST procedures in parallel with randomized order. The total number of trials was variable for each participant as defined by the adaptive staircase procedure. In order to avoid visual fatigue, the experiment was conducted in two sessions and the duration of each session was limited to 30-40 mins. Participants were allowed to rest for at least 2 hours between sessions.

Results   The average recognition times measured in this experiment are shown in Figure 5. The complete set of measurements can
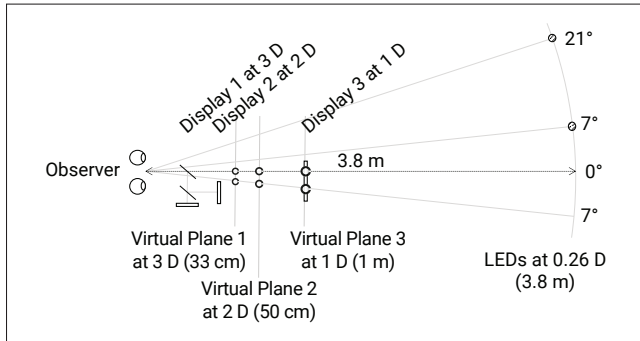
---

[1]All described experiments were approved by Saarland University ERB No. 18-9-5 or by UNC-Chapel Hill IRB #16-0368.
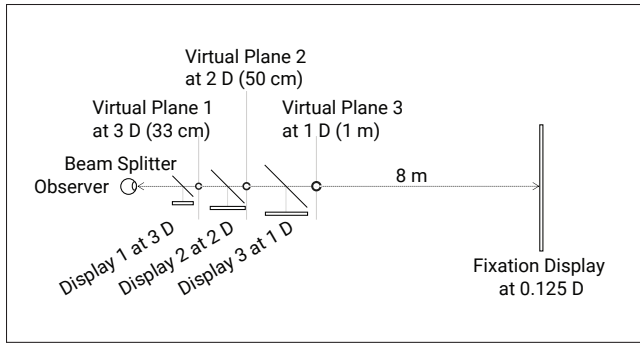
(a) Side view of multiplane display for pilot experiment setup (Location A)
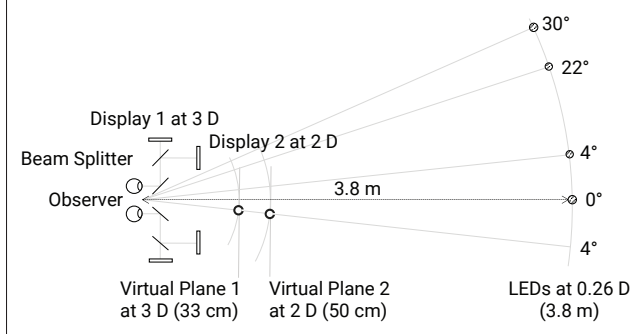
(b) Top view of multiplane display for experiment setup 1 (Location B)

(c) Top view of multiplane display for experiment setup 2 (Location B)

Figure 6: Experimental configurations used in the accommodation and convergence overdriving experiments with saccadic facilitation. The experiment is conducted at multiple locations, which are indicated in the captions above.

be found in the supplementary material. We have several observations from the experimental results. The recognition time increases as the accommodation depth increases. Without accommodation change, the recognition time can be as short as 141 ms, which is comparable to the fast characteristics of saccade dynamics. However, the recognition time in the absence of accommodation change reaches up to 268 ms for a 30° saccade. This indicates that it takes longer to recognize the orientation of Landolt C pattern in the far periphery. In addition, the overall recognition time is less than 400 ms for every trial. This duration is shorter than the duration of pure accommodation responses from previous studies, which is reported as 500 ms to 800 ms [7, 9, 19, 25, 47, 55, 58]. This can be explained by the effect of saccadic facilitation, which provides a reduction in latency of 100 ms to 300 ms and an increase in accommodation velocities of 1.5 D/s to 2 D/s as reported by Schor [59].
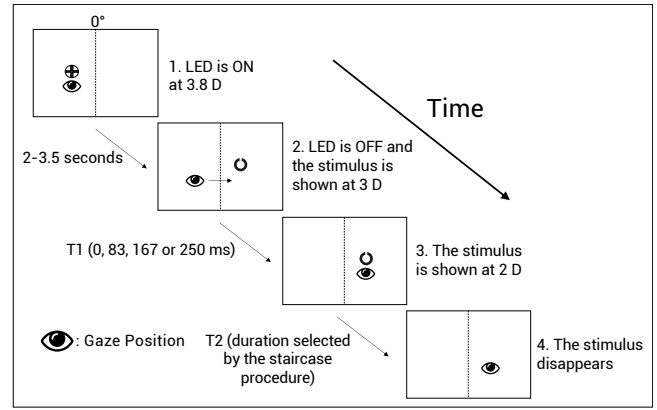

Figure 7: The schematic of the accommodation and convergence overdriving experiment with saccadic facilitation.

5 OVERDRIVING THE VISUAL SYSTEM

With the goal of achieving faster times for depth switches, we could either attempt to decrease the length of the latent period, or increase the maximum velocity of the responses. As stated before, by overdriving the visual cues, we can induce higher transition speeds and hypothesize we could decrease the time required for physiological response. To verify that the physiological speed correlated with perceptual speed, task-based experiments would be ideal. In this section, we describe a series of experiments which attempt to induce this over-driven behavior in subjects by presenting overdrive visual stimuli and measuring results based on performance in a time-constrained identification task.

5.1 Early Investigations

Our overdrive experimentation began with a monocular accommodation-only experiment simplified from the experiment described in Section 4, but with an overdriven stimulus displayed for 0 ms, 83 ms, 167 ms, or 250 ms before the target stimulus was displayed. This limited study, described in more detail in the supplemental material, showed that the overdriving principle had promise and should be extended to include vergence and saccade.

The next experiment we performed was a small pilot study in overdriving the accommodation and vergence responses with and without a saccade with 4 and 6 users each. Figure 6(a) shows the configuration used in this study, and a full description can be found in the supplemental material, but as seen in Figure 8 it too showed promising results when the saccade was present. With these promising results, we elected to proceed with larger user studies to investigate the effects of the saccade magnitude and overdrive duration.

5.2 Overdriving Accommodation and Convergence During Saccade

The goal of this experiment was to check if there are task performance speedups in correctly recognizing visual stimulus in the presence of saccadic facilitation of convergence as well as target-depth overdrive.

Stimulus   We used the same stimulus as described in Section 4.

Experimental configuration   This experiment was run at a location different from the pilot study using two different prototype displays, which are shown in Figure 6(b) and (c). Experiment Setup 1 consisted of a stereoscopic multi-plane display which had two virtual planes at 2 D (0.5 m) and 3 D (0.33 m). This prototype display closely resembles the arrangement of screens widely used in HMDs. To address any variability in interpupillary distances of the
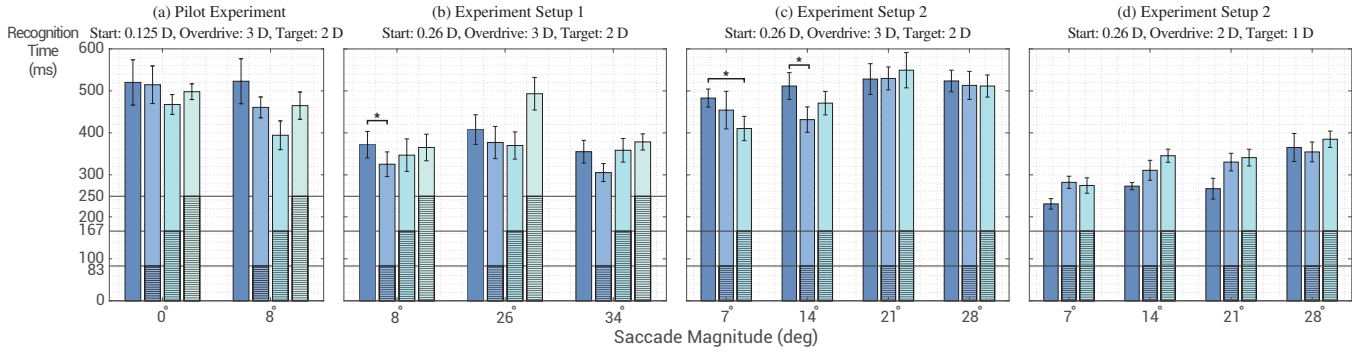
Figure 8: Mean recognition times measured in the accommodation and convergence overdriving experiments with saccadic facilitation. Hatched patterns represent the overdrive duration (T1) within total recognition time (T1+T2). The error bars represent the Standard Error of the Mean (SEM). The speedups marked with asterisks (*) are statistically significant ($p \leq 0.05$, Wilcoxon signed-rank test).

participants, the multi-plane display was calibrated for each participant at the beginning of the experiment using the reference LED at 3.8 m along the medial axis (at 0° eccentricity). Three LEDs were positioned at 4°, 22°, and 30° eccentricities relative to the medial axis as fixation targets. Landolt C stimulus was shown on the virtual planes at 3 D (overdrive) and at 2 D (final depth). The stimuli were shown at 4° of eccentricity to the right of the medial axis, which allowed us to test saccades with three different magnitudes of 8°, 26°, and 34° with this display.

Experiment Setup 2 consisted of three desktop displays, which were visible through beam splitters. Their arrangement provided us with one extra virtual plane, which totals up to three virtual planes at 1 D (1 m), 2 D (0.5 m) and 3 D (0.33 m) (Figure 6(c)). It was possible to show the stimuli along the medial axis (0°) and 7° to the right of the medial axis with this setup. Two LEDs, which were used as fixation targets, were positioned at 7° and 14° to the left of the medial axis. This setup allowed us to test four different saccade magnitudes (7°, 14°, 21°, and 28°) and two target depth levels at 1 D (overdriven at 2 D) and 2 D (overdriven at 3 D). The aim of this setup was to allow an additional depth and more evenly distributed saccade magnitudes when compared to experiment setup 1.

In both setups, a light diffuser with a cross pattern was placed in front of the LEDs to trigger proper accommodation response on the targets. The display gamma correction and peak luminance adjustment were done following the same protocol as described in Section 4. This experiment was also conducted in the same dark room conditions using the chin rest for additional stability.

Participants   Two different groups of paid undergraduate and graduate students participated in this experiment. The first group consisted of 10 participants, who did the experiment on Experiment Setup 1, each completing between 1 and 3 of the 3 different saccade magnitude trials. The second group consisted of 13 participants (5 female, 9 male), who did the experiment on Experiment Setup 2. Two participants in this group were treated as outliers and their data is discarded from the analyses because the timings were significantly longer than the rest of the group indicating inattentiveness (please refer to the supplemental materials for more details and complete raw data). The ages of the participants ranged from 20 to 31 years. All participants had normal or corrected-to-normal vision.

Procedure   The experiment started with the calibration of multi-plane display for each participant to achieve proper stereopsis. During this phase, the participant was asked to align virtual crosses shown on the displays with the LED positioned at the center of their visual field (0°).

Once calibration was complete, we ran interleaved adaptive staircase procedures in parallel for different values of the *overdrive duration* (T1), which is the period of time for holding the stimulus

at the overdrive depth level before moving it to the target depth (Figure 7). Each trial started with the participant's fixation on one of the target LEDs. After a random duration of time between 2-3.5 s, the target LED was turned off and the stimulus appeared at the overdrive depth level, where it was displayed for the overdrive duration (T1). Then the stimulus was moved to the target depth, where it was displayed for *target duration* (T2) selected by the adaptive staircase procedure depending on the responses of the participant during each trial. The stimulus disappeared at the end of T2 and the participant was asked to identify the gap position. We used vPEST [21] as our adaptive staircase procedure similarly to Section 4 to estimate the actual value of T2. Based on 62.5 % detection rate, we fit a logistic psychometric function to the collected responses to improve the efficiency of threshold estimation [53]. The experiment was conducted in multiple sessions as described in Section 4 except that the experiment was divided into four sessions with a maximum of 2 sessions per day.

Results   Total recognition times (T1+T2) observed on Experiment Setups 1 and 2 are shown in Figure 8. When we compare the measurements obtained by overdriving convergence with those from pure saccadic facilitation, we observe a mean speedup between 30.95-80 ms when T1=83 ms. With a saccade magnitude of 8° and T1=83 ms, the total duration for recognition is reduced from 373 ms to 327 ms with a speedup of 46 ms on Experiment Setup 1 for 10 participants. This improvement is statistically significant ($p = 0.02$, Wilcoxon signed-rank test). On that setup, the maximum speedup is observed for T1=83 ms with a saccade magnitude of 34°, where the recognition time is reduced from 356 ms to 306 ms. However, that improvement is not statistically significant ($p = 0.22$). One potential factor contributing to the lack of statistical significance is relatively smaller number of participants who did the experiment with 26° and 34° saccades on this experiment setup (7 and 5 participants, respectively).

The maximum speedup on Experiment Setup 2 is measured as 80 ms with an overdrive duration T1=83 ms, when the convergence is combined with a 14° saccade ($p = 0.05$). On that setup, when the convergence is combined with a 7° saccade, a statistically significant ($p = 0.03$) amount of speedup is observed for a relatively longer overdrive duration (T1=167 ms). Both speedups are observed at the target depth of 2 D (overdrive depth: 3 D). In both experiment setups, other combinations of T1, saccade magnitude and target depth levels do not show a significant improvement in the recognition time ($p > 0.25$).

5.3   Discussion

Overall we show a trend indicating that for three different experiments in two locations, overdriving the depth of the stimulus leads to
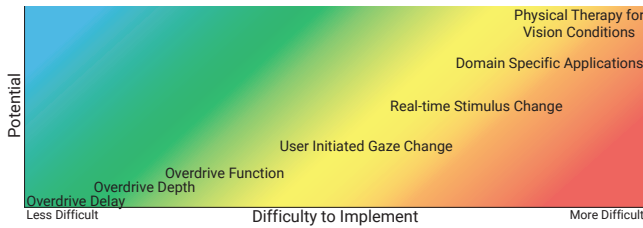
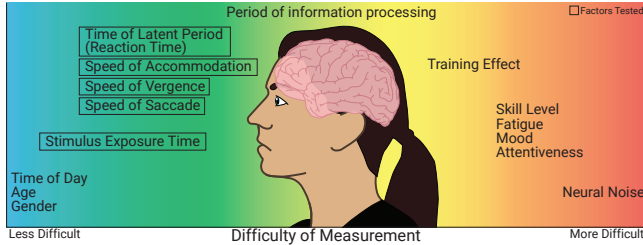Figure 9: Methods and applications of possible future research.

Figure 10: Possible user factors affecting task-based performance time organized by difficulty of measurement. Factors with box are those tested in our experiments, but many other factors may contribute to the results of our work and should be tested in future work.

faster depth switches. Both by combining the accommodation, vergence, and saccade, and by briefly presenting a stimulus beyond the intended target, we are able to speed up the recognition of stimuli.

It should be noted that not all cases show positive results, particularly the smaller depth switch case of Experiment Setup 2 with target depth of 1 D (overdrive depth: 2 D). In this case the initial fixation to final target depth difference is only 0.74 D. The depth of focus for humans depends on many factors, but Wang and Ciuffreda [72] report values larger than 1 D in some studies, meaning the target may already be in focus before any depth change has occurred. Thus for small depth changes only the saccade and vergence are necessary, while very little accommodation needs to occur, if any at all. In this condition, displaying the additional overdrive stimulus extends the period of adaptation leading to the longer overdrive times seen in Figure 8(d). This means there is a lower bound on the distance of depth change for which providing overdriven stimuli has any benefit.

6 FUTURE RESEARCH

While we have some hypotheses, unfortunately we have been unable get definite answers as to why the results are seen in some cases and not others. The diversity in these studies shows that some cases demonstrate larger impact of overdriving the stimulus during a depth switch. Some of the first work to be done would seek to answer this question. One path would be to recreate the pilot experiment setup and determine any underlying reasons for these apparent differences with a larger set of users that would be able to provide significant positive or negative results in every case. In particular, we believe that looking at stimulus selection, understanding user factors, leveraging instrumentation capability, and a clearer picture of the underlying processes can lead to more consistent results.

Stimulus Selection   The effect of the visual qualities of the stimulus and the dynamics of the overdrive need to be further teased out. Our task being dependent on visual acuity necessitated a small stimulus, but what effect would a larger stimulus have? Or a stimulus of different shape, frequency, color, etc? Beyond just the visual qualities of the stimulus, the dynamics of the overdrive present fertile ground for future research. We investigated 3 different timings for
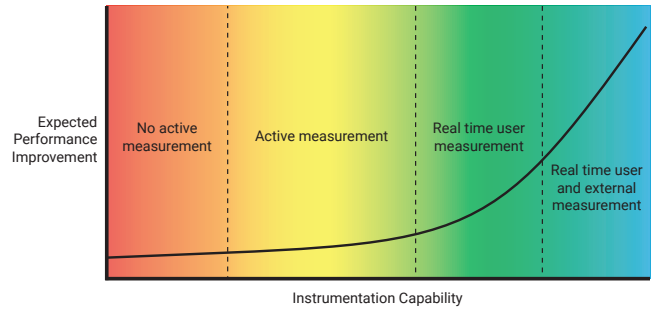

Figure 11: Effect of instrumentation capability on expected performance improvement. With the development and employment of new instruments capable of measuring the accommodation of the eye during vergence and saccade, it is expected performance will further improve.

the display of the overdrive stimulus, but were limited by the 60 Hz of our experimental equipment. Higher frame rate displays could much more finely tune the amount of temporal overdrive, while high quality accommodation-providing stereoscopic displays could provide more precision for physical distances of overdrive. Different strategies of presenting fixation targets in temporal dimension may be explored to improve the recognition time by reducing the latency to initiate the eye movements as reported by Takagi et al. [68]. Along with the time and location of the stimuli, the overdrive function itself could be further investigated. Our simple step function may not be the optimal presentation of the stimulus, continuous or even non-monotonic functions may further decrease the depth switching time. Stimulus selection process should also take into account potential effects of saccadic suppression and other phenomena related to reduced perceptual sensitivity during eye movements [30, 45]. Figure 9 illustrates the difficulty and potential of these and some other potential applications.

Understanding User Factors   Humans are notoriously difficult to quantify. As we seek to drive portions of the HVS, there are many factors which may or may not affect the success of our experiments. Figure 10 displays some of these user factors on a scale of how difficult they are to measure. Each user reacts at a different speed, which can be affected by a set of factors which can be controllable or uncontrollable as well as measurable and immeasurable. For example, saccades tend to undershoot the target, their accuracy is reduced and corrective saccades are needed more frequently when they are combined with vergence [75]. The interaction between saccade and vergence is not fully predictable in advance because various types of gaze displacement patterns are observed during vergence with different probabilities [14]. In addition, although a speedup is observed in mean velocities, the latency of eye movements increases when they are performed in combination [75]. Each of these factors increase the variability of the measurements and require a careful analysis of results. Any future work in attempting to enhance human performance should attempt to classify each relevant factor and rate it according to its impact. By performing a threshold staircase study, confounding events may have a masking effect on the benefits of the response in our experiments, so ideally future studies will isolate these factors. Additionally a larger number of participants would be able to possibly overcome any confounding effects.

Instrumentation Capability   None of the experiments described in this work used any active eye measurement techniques; however, as discussed in Section 2.4, if instrumentation were employed which is able to measure the focal state of the eye during saccade and vergence, further insight into the overdriving dynamics could lead to more optimized performance gains. With better accommodation

measurement and more precise eye tracking, better tests could be designed which would lead to better insight into what is happening and could be used to generate ideas on how to change the stimulus to further improve performance. A drastic increase in benefit will come when these measurements are able to be achieved and stimuli are able to be updated in real-time reacting as the user does. By adding full environmental scanning, depth processing, segmentation, and classification, the potential improvements are expanded by offering application and domain specific capabilities. This progression of instrument capability leading to improvements in performance is illustrated in Figure 11.

Near-to-far Accommodation and Divergence   All overdrive experiments in this work are strictly far-to-near accommodation and convergence. Yang and Kapoula [75] reports higher speedups for divergence compared to convergence when they are performed in combination with saccades (more than $2\times$ reduction in the total duration). Similarly, convergence and divergence have different latencies and mean velocities [74, 75]. Therefore, it would be worthwhile to replicate an experiment with divergence and near-to-far accommodation as the responses have different dynamics, to see if the overdrive works similarly.

Near Response Triad   In normal viewing conditions as gaze moves from a distant point toward the viewer, three responses occur: convergence, far-to-near accommodation, and pupillary dilation. This is referred to as the near response triad. The connections between these three responses have been well established and have even been mapped neurologically from the Edinger-Westphal nucleus along the ocularmotor nerve [15, 50]. In this work, we did not investigate pupillary dilation due to lack of instrumentation, no clear connection to the speed of the other responses, and questionable perceptual impact. It may provide an interesting line of study in future work.

Stimulus vs User Directed   All expressed values for the dynamics of vergence and accommodation are for the case when a subject is presented with a stimulus. However, the case when voluntary depth change occurs, as with the user-driven fixation in varifocal optical see-through (OST) AR displays, has not fully been studied. Some discussion on the topic by Ciuffreda and Kruger [13] indicates that under the user-driven condition an estimated amount of accommodation, which could be used to resolve details, would occur before any retinal-blur information is received. This preprogramming would make the visual scanning process more efficient. Kruger and Pola [38] suggest that the natural mode of the accommodation system is anticipation, while [39] show that accommodative prediction is a key component in tracking regularly approaching and receding objects. These discussions indicate that for voluntary depth changes, the latent period would be drastically reduced, if not eliminated entirely, and pre-accommodation may occur, such that the response time would happen partially before the fixation change is initiated. Unfortunately, until further studies are made on user-driven changes to fixation depth, we must base our requirements on reported values, and therefore use the stimulus-driven results described above.

Physiological and perceptual speedups   With the complex mechanisms and processes in the HVS, it is unclear if a faster physiological response as measured by instrumentation would necessarily lead to faster perception. We have shown one case where the physiological speedup will lead to a perceptual speedup, but this question must continue to be asked for each and every such potential case.

Physical Therapy   Recent work has examined ways VR may be used to treat ophthalmologic disorders such as strabismus and amblyopia [10, 80]. As AR HMDs become more commonplace, it will become possible to use the displays for prescribed physical therapy. Overdriving the depth switches could provide a workout for the eye muscles to strengthen and relax them. Unexpected benefits such as delaying the onset of myopia or alleviating presbyopia may become possible with the right treatment. We look forward to seeing how the medical industry takes advantage of these new devices.

7 CONCLUSION

Due to the nature of the investigated new phenomena, we only scratched the surface of possible opportunities for both speeding up perception during depth switches and the overarching goal of enhancing the HVS beyond its natural capabilities. Our results demonstrate that significant perceptual speedups are possible in the correct circumstances.

We recommend for AR practitioners to be aware of the time required for depth switching, and that they should design experiences with these delays in mind. Avoiding augmentations that are too near or too far would be good practice, as well as avoiding rapid or large depth switches. With current displays, they should consider methods of minimizing the total amount of refocusing by the user, and as new displays with greater capabilities become available, they should work toward making depth switches faster, more efficient, and less burdensome.

Several recent OST AR displays have included technology supporting accommodation while others embrace eye tracking. We add one more reason for display designers to continue pursuits in these areas as both of these technologies are required in seeking to overdrive user depth switching. If these recent trends in display technology continue, in the near future, displays will be capable of enabling accommodation overdrive as investigated for the first time in this work.

While we find the topic engaging and our initial results are promising, we believe the greater goal is worth a much broader investigation. Thus, we present our work intending to inspire others to pursue a similar research direction. It is our hope to inspire additional AR researchers to join in the quest for better understanding the interactions between saccades, vergence, and accommodation, and to spark further work in employing OST AR displays to enhance perception in other ways.   The space that has to be explored is large, but we believe that the potential benefit will be rewarding. Therefore, we welcome co-investigators to consider the promise of using AR displays to stimulate the HVS beyond what it is capable of in the natural world.

REFERENCES

[1]   K. Akeley, S. J. Watt, A. R. Girshick, and M. S. Banks. A stereo display prototype with multiple focal distances. In *ACM Transactions on Graphics (TOG)*, vol. 23, pp. 804–813. ACM, 2004.

[2]   K. Akşit, W. Lopes, J. Kim, P. Shirley, and D. Luebke. Near-Eye Varifocal Augmented Reality Display using See-Through Screens. *ACM Trans. Graph. (SIGGRAPH)*, (6), 2017.

[3]   J. Anliker. Eye movement: On-line measurement, analysis, and control. *R. A. Monty & J. W. Senders (Eds.), Eye movements and psychological processes*, pp. 185–202, 1976.

[4]   E. Arabadzhiyska, O. T. Tursun, K. Myszkowski, H.-P. Seidel, and P. Didyk. Saccade landing position prediction for gaze-contingent rendering. *ACM Transactions on Graphics (Proc. SIGGRAPH)*, 36(4), 2017. doi: 10.1145/3072959.3073642

[5] M. S. Arefin, N. Phillips, A. Plopski, J. L. Gabbard, and J. E. Swan II. Impact of ar display context switching and focal distance switching on human performance: Replication on an ar haploscope. In *2020 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*, pp. 571–572. IEEE, 2020.

[6] A. T. Bahill, M. R. Clark, and L. Stark. The main sequence, a tool for studying human eye movements. *Mathematical Biosciences*, 24(3-4):191–204, 1975.

[7] S. R. Bharadwaj and C. M. Schor. Acceleration characteristics of human ocular accommodation. *Vision Research*, 45(1):17–28, 2005.

[8] S. R. Bharadwaj and C. M. Schor. Dynamic control of ocular dis-accommodation: First and second-order dynamics. *Vision Research*, 46(6):1019–1037, Mar. 2006. doi: 10.1016/j.visres.2005.06.005

[9] F. W. Campbell and G. Westheimer. Dynamics of accommodation responses of the human eye. *J. Physiol.*, 151(2):285–295, 1960.

[10] L. K. Cepeda-Zapata, F. O. Romero-Soto, V. A. D. de León, J. L. Roa-Huertas, N. E. Naal-Ruiz, D. Ibarra-Zarate, and L. M. Alonso-Valerdi. Implementation of a virtual reality rendered in portable devices for strabismus treatment based on conventional visual therapy. In *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pp. 7189–7192. IEEE, 2019.

[11] P. Chakravarthula, D. Dunn, K. Akşit, and H. Fuchs. FocusAR: Auto-focus Augmented Reality Eyeglasses for both Real World and Virtual Imagery. *IEEE Transactions on Visualization and Computer Graphics*, 24(11):2906–2916, Nov. 2018. doi: 10.1109/TVCG.2018.2868532

[12] E. Chirre, P. Prieto, and P. Artal. Dynamics of the near response under natural viewing conditions with an open-view sensor. *Biomedical Optics Express*, 6(10):4200–4211, Oct. 2015. doi: 10.1364/BOE.6.004200

[13] K. J. Ciuffreda and P. B. Kruger. Dynamics of human voluntary accommodation. *American Journal of Optometry and Physiological Optics*, 65(5):365–370, May 1988.

[14] O. A. Coubard and Z. Kapoula. Saccades during symmetrical vergence. *Graefe's Archive for Clinical and Experimental Ophthalmology*, 246(4):521–536, 2008.

[15] H. Davson. The Near Response. In H. Davson, ed., *Physiology of the Eye*, pp. 783–785. Macmilan Education UK, London, 1990. doi: 10.1007/978-1-349-09997-9_28

[16] D. Dunn, C. Tippets, K. Torell, P. Kellnhofer, K. Akşit, P. Didyk, K. Myszkowski, D. Luebke, and H. Fuchs. Wide Field Of View Varifocal Near-Eye Display Using See-Through Deformable Membrane Mirrors. *IEEE Transactions on Visualization and Computer Graphics*, 23(4):1322–1331, Apr. 2017. doi: 10.1109/TVCG.2017.2657058

[17] J. Enright. Facilitation of vergence changes by saccades: Influences of misfocused images and of disparity stimuli in man. *The Journal of physiology*, 371(1):69–87, 1986.

[18] J. T. Enright. Changes in vergence mediated by saccades. *The Journal of Physiology*, 350(1):9–31, 1984. doi: 10.1113/jphysiol.1984.sp015186

[19] C. J. Erkelens, R. M. Steinman, and H. Collewijn. Ocular Vergence Under Natural Conditions. II. Gaze Shifts Between Real Targets Differing in Distance and Direction. *Proceedings of the Royal Society of London. Series B, Biological Sciences*, 236(1285):441–465, 1989.

[20] E. F. Fincham and J. Walton. The reciprocal actions of accommodation and convergence. *The Journal of Physiology*, 137(3):488–508, 1957. doi: 10.1113/jphysiol.1957.sp005829

[21] J. M. Findlay. Estimates on probability functions: A more virulent PEST. *Perception & Psychophysics*, 23(2):181–185, Mar. 1978. doi: 10.3758/BF03208300

[22] J. L. Gabbard, D. G. Mehra, and J. E. Swan. Effects of ar display context switching and focal distance switching on human performance. *IEEE transactions on visualization and computer graphics*, 25(6):2228–2241, 2018.

[23] S. J. Henderson and S. K. Feiner. Augmented reality in the psychomotor phase of a procedural task. In *Mixed and Augmented Reality (ISMAR), 2011 10th IEEE International Symposium On*, pp. 191–200. IEEE, 2011.

[24] E. Hering. *Die Lehre Vom Binocularen Sehen*. Leipzig: Engelmann, 1868.

[25] G. Heron, W. N. Charman, and C. Schor. Dynamics of the accommo-dation response to abrupt changes in target vergence as a function of age. *Vision Research*, 41(4):507–519, 2001. doi: 10.1016/S0042-6989(00)00282-0

[26] D. M. Hoffman, A. R. Girshick, K. Akeley, and M. S. Banks. Vergence–accommodation conflicts hinder visual performance and cause visual fatigue. *Journal of vision*, 8(3):33–33, 2008.

[27] X. Hu and H. Hua. High-resolution optical see-through multi-focal-plane head-mounted display using freeform optics. *Optics express*, 22(11):13896–13903, 2014.

[28] H. Hua and B. Javidi. A 3d integral imaging optical see-through head-mounted display. *Opt. Express*, 22(11):13484–13491, Jun 2014. doi: 10.1364/OE.22.013484

[29] F.-C. Huang, K. Chen, and G. Wetzstein. The Light Field Stereoscope: Immersive Computer Graphics via Factored Near-eye Light Field Displays with Focus Cues. *ACM Transactions on Graphics (TOG)*, 34(4):60:1–60:12, July 2015. doi: 10.1145/2766922

[30] G. K. Hung, L. Sun, J. L. Semmlow, and K. J. Ciuffreda. Suppression of sensitivity to change in target disparity during vergence eye movements. *Experimental neurology*, 110(3):291–297, 1990.

[31] S. Kasthurirangan, A. S. Vilupuru, and A. Glasser. Amplitude dependent accommodative dynamics in humans. *Vision Research*, 43(27):2945–2956, 2003. doi: 10.1016/j.visres.2003.08.004

[32] O. V. Komogortsev and J. I. Khan. Eye movement prediction by oculomotor plant Kalman filter with brainstem control. *Journal of Control Theory and Applications*, 7(1):14–22, 2009.

[33] S. Kouider and S. Dehaene. Levels of processing during non-conscious perception: A critical review of visual masking. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 362(1481):857–875, 2007.

[34] E. Kowler. Eye movements: The past 25 years. *Vision Research*, 51(13):1457–1483, 2011.

[35] G. Kramida. Resolving the Vergence-Accommodation Conflict in Head-Mounted Displays. *IEEE Transactions on Visualization and Computer Graphics*, 22(7):1912–1931, 2016.

[36] V. Krishnan, F. Farazian, and L. Stark. An analysis of latencies and prediction in the fusional vergence system. *American Journal of Optometry and Archives of American Academy of Optometry*, 50(12):933—939, 1973.

[37] V. V. Krishnan, D. Shirachi, and L. Stark. Dynamic measures of vergence accommodation. *American Journal of Optometry and Physiological Optics*, 54(7):470–473, July 1977.

[38] P. B. Kruger and J. Pola. Stimuli for accommodation: Blur, chromatic aberration and size. *Vision Research*, 26(6):957–971, Jan. 1986. doi: 10.1016/0042-6989(86)90153-7

[39] P. B. Kruger and J. Pola. Dioptric and non-dioptric stimuli for accommodation: Target size alone and with blur and chromatic aberration. *Vision Research*, 27(4):555–567, Jan. 1987. doi: 10.1016/0042-6989(87)90042-3

[40] E. Kruijff, J. E. Swan, and S. Feiner. Perceptual issues in augmented reality revisited. In *2010 IEEE International Symposium on Mixed and Augmented Reality*, pp. 3–12. IEEE, 2010.

[41] T. Liu, V. Sreenivasan, and L. N. Thibos. Uniformity of accommodation across the visual field. *Journal of Vision*, 16(3):6–6, Feb. 2016. doi: 10.1167/16.3.6

[42] K. J. MacKenzie, D. M. Hoffman, and S. J. Watt. Accommodation to multiple-focal-plane displays: Implications for improving stereoscopic displays and for accommodation control. *Journal of Vision*, 10(8):22, 2010. doi: 10.1167/10.8.22

[43] A. Maimone and H. Fuchs. Computational augmented reality eyeglasses. In *2013 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pp. 29–38, 2013.

[44] A. Maimone, A. Georgiou, and J. S. Kollin. Holographic near-eye displays for virtual and augmented reality. *ACM Transactions on Graphics (TOG)*, 36(4):85, 2017.

[45] E. Matin. Saccadic suppression: A review and an analysis. *Psychological bulletin*, 81(12):899, 1974.

[46] J. Maxwell, J. Tong, and C. M. Schor. The first and second order dynamics of accommodative convergence and disparity convergence. *Vision Research*, 50(17):1728–1739, Aug. 2010. doi: 10.1016/j.visres.2010.05.029

[47] M. W. Morgan. Accommodation and Vergence. *Optometry and Vision Science*, 45(7):417, July 1968.

[48] M. W. J. Morgan. Accommodation and its relationship to convergence. *Optometry and Vision Science*, 21(5):183–195, 1944.

[49] T.-J. Mu, J.-J. Sun, R. R. Martin, and S.-M. Hu. A response time model for abrupt changes in binocular disparity. *The Visual Computer*, 31(5):675–687, 2015.

[50] G. A. Myers and L. Stark. Topology of the near response triad. *Ophthalmic and Physiological Optics*, 10(2):175–181, 1990. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1475-1313.1990.tb00972.x. doi: 10.1111/j.1475-1313.1990.tb00972.x

[51] I. C. of Ophthalmology. Visual acuity measurement standard. 1984.

[52] C. Paeye, A. C. Schütz, and K. R. Gegenfurtner. Visual reinforcement shapes eye movements in visual search. *Journal of Vision*, 16(10):15–15, 2016.

[53] A. Pentland. Maximum likelihood estimation: The best PEST. *Perception & psychophysics*, 1980.

[54] M. A. Peters, R. W. Kentridge, I. Phillips, and N. Block. Does unconscious perception really exist? Continuing the ASSC20 debate. *Neuroscience of consciousness*, 3(1), 2017.

[55] S. Phillips, D. Shirachi, and L. Stark. Analysis of accommodative response times using histogram information. *American Journal of Optometry & Archives of American Academy of Optometry*, 49(5):389–400, 1972.

[56] K. Rathinavel, H. Wang, A. Blate, and H. Fuchs. An extended depth-at-field volumetric near-eye augmented reality display. *IEEE Transactions on Visualization and Computer Graphics*, 24(11):2857–2866, 2018.

[57] C. Schor. The influence of interactions between accommodation and convergence on the lag of accommodation. *Ophthalmic and Physiological Optics*, 19(2):134–150, 1999. doi: 10.1046/j.1475-1313.1999.00409.x

[58] C. M. Schor. A Dynamic Model of Cross-coupling Between Accommodation and Convergence: Simulations of Step and Frequency Responses. *Optometry and Vision Science*, 69(4):258–269, Apr. 1992.

[59] C. M. Schor, L. A. Lott, D. Pope, and A. D. Graham. Saccades reduce latency and increase velocity of ocular accommodation. *Vision Research*, 39(22):3769–3795, Nov. 1999. doi: 10.1016/S0042-6989(99)00094-2

[60] J. Semmlow and P. Wetzel. Dynamic contributions of the components of binocular vergence. *Journal of the Optical Society of America*, 69(5):639–645, 1979.

[61] L. Shi, F.-C. Huang, W. Lopes, W. Matusik, and D. Luebke. Near-eye light field holographic rendering with spherical waves for wide field of view interactive 3d computer graphics. *ACM Trans. Graph.*, 36(6), Nov. 2017. doi: 10.1145/3130800.3130832

[62] T. Shibata, J. Kim, D. M. Hoffman, and M. S. Banks. Visual discomfort with stereo displays: Effects of viewing distance and direction of vergence-accommodation conflict. In *IS&T/SPIE Electronic Imaging*, pp. 78630P–78630P. International Society for Optics and Photonics, 2011.

[63] G. Singh, S. R. Ellis, and J. E. Swan II. The effect of focal distance, age, and brightness on near-field augmented reality depth matching. *IEEE transactions on visualization and computer graphics*, 2018.

[64] M. Smith, J. Streeter, G. Burnett, and J. L. Gabbard. Visual search tasks: the effects of head-up displays on driving and task performance. In *Proceedings of the 7th international conference on Automotive User Interfaces and Interactive Vehicular Applications*, pp. 80–87, 2015.

[65] P. Sterzer, T. Stein, K. Ludwig, M. Rothkirch, and G. Hesselmann. Neural processing of visual information under interocular suppression: A critical review. *Frontiers in psychology*, 5:453, 2014.

[66] R. Suryakumar, J. P. Meyers, E. L. Irving, and W. R. Bobier. Application of video-based technology for the simultaneous measurement of accommodation and vergence. *Vision Research*, 47(2):260–268, Jan. 2007. doi: 10.1016/j.visres.2006.10.003

[67] J. E. Swan, G. Singh, and S. R. Ellis. Matching and reaching depth judgments with real and augmented reality targets. *IEEE transactions on visualization and computer graphics*, 21(11):1289–1298, 2015.

[68] M. Takagi, E. M. Frohman, and D. S. Zee. Gap-overlap effects on latencies of saccades, vergence and combined vergence-saccades in humans. *Vision Research*, 35(23-24):3373–3388, 1995.

[69] K. Templin, P. Didyk, K. Myszkowski, M. M. Hefeeda, H.-P. Seidel, and W. Matusik. Modeling and optimizing eye vergence response to stereoscopic cuts. *ACM Transactions on Graphics (Proc. SIGGRAPH)*, 33(4), 2014.

[70] M. Torii, Y. Okada, K. Ukai, J. S. Wolffsohn, and B. Gilmartin. Dynamic measurement of accommodative responses while viewing stereoscopic images. *Journal of Modern Optics*, 55(4-5):557–567, 2008.

[71] A. Van Opstal and J. Van Gisbergen. Skewness of saccadic velocity profiles: A unifying parameter for normal and slow saccades. *Vision Research*, 27(5):731–745, 1987.

[72] B. Wang and K. J. Ciuffreda. Depth-of-Focus of the Human Eye: Theory and Clinical Implications. *Survey of Ophthalmology*, 51(1):75–85, 2006. doi: 10.1016/j.survophthal.2005.11.003

[73] G. Wetzstein, W. Heidrich, and D. Luebke. Optical image processing using light modulation displays. In *Computer Graphics Forum*, vol. 29, pp. 1934–1944, 2010.

[74] Q. Yang, M. P. Bucci, and Z. Kapoula. The Latency of Saccades, Vergence, and Combined Eye Movements in Children and in Adults. *Investigative Ophthalmology & Visual Science*, 43(9):2939–2949, Sept. 2002.

[75] Q. Yang and Z. Kapoula. Saccade–vergence dynamics and interaction in children and in adults. *Experimental Brain Research*, 156(2):212–223, 2004.

[76] A. L. Yarbus. *Eye Movements and Vision*. New York: Plenum Press, 1967.

[77] S. H. Yeo, M. Lesmana, D. R. Neog, and D. K. Pai. Eyecatch: Simulating visuomotor coordination for object interception. *ACM Transactions on Graphics (TOG)*, 31(4):42, 2012.

[78] D. S. Zee, E. J. Fitzgibbon, and L. M. Optican. Saccade-vergence interactions in humans. *Journal of Neurophysiology*, 68(5):1624–1641, 1992.

[79] W. Zhou, X. Chen, and J. Enderle. An updated time-optimal 3rd-order linear saccadic eye plant model. *International Journal of Neural Systems*, 19(05):309–330, 2009.

[80] P. Žiak, A. Holm, J. Halička, P. Mojžiš, and D. P. Piñero. Amblyopia treatment of adults with dichoptic training using the virtual reality oculus rift head mounted display: Preliminary results. *BMC ophthalmology*, 17(1):1–8, 2017.