The Contribution of Stereopsis and Motion Parallax to Depth Segmentation

Teodora Neagu

A thesis submitted to the Faculty of Graduate Studies in Partial Fulfillment of the Requirements

for the Degree of Master of Science

Graduate Program in Biology

York University

Toronto, Ontario

August 2024

ABSTRACT

Studies comparing how stereopsis and motion parallax inform depth yield mixed results, with some showing poorer performance from motion parallax. We propose that this stems from use of depth magnitude tasks, which bias observers towards using stereopsis. In Experiment 1, we used a depth segmentation task to evaluate performance of stereopsis and self-generated motion parallax individually and in combination. Observers performed equally well with individual cues but appeared to rely on stereopsis when both were available. Follow-up experiments showed that increasing the range of self-motion did not improve performance, but the use of object motion did. Subsequently, we replicated Experiment 1 using object motion with experts and non-experts and found segmentation from object motion was always better than stereopsis, and non-experts needed more stereopsis than experts to perform the task. Thus, we suggest that the ability to use depth information depends on several factors including task demands and observers' experience.

TABLE OF CONTENTS

LIST OF FIGURES

INTRODUCTION

To interpret a 3-dimensional scene the human visual system must incorporate information from a number of sources. One of the most precise mechanisms contributing to depth perception is stereopsis which is the perception of depth from retinal disparity that is a consequence of having eyes that are horizontally separated in the head. The positional difference between the resulting retinal images is called binocular disparity and this information is used by the brain along with estimates of the distance to the object, to compute an estimate of the three-dimensional structure of the object or scene, relative to fixation (Figure 1a) (Wheatstone, 1838).

In addition to binocular cues, the visual system relies on monocular depth cues to support depth perception such as shading, perspective and size. Motion parallax is a monocular depth cue that is similar to stereopsis as depth information is computed via triangulation, and also relies on distance to scale depth (Ono, Rivest and Ono, 1986). It arises from the relative motion of objects at different distances, closer objects appear to move faster than objects that are further away. Self-generated motion parallax occurs when an individual translates laterally relative to the scene and object-generated motion parallax occurs when an object physically moves in space, causing the image to move along the retina. An object moving at a fixed speed will move a greater distance on the retina when it is close to the observer than when it is far away.

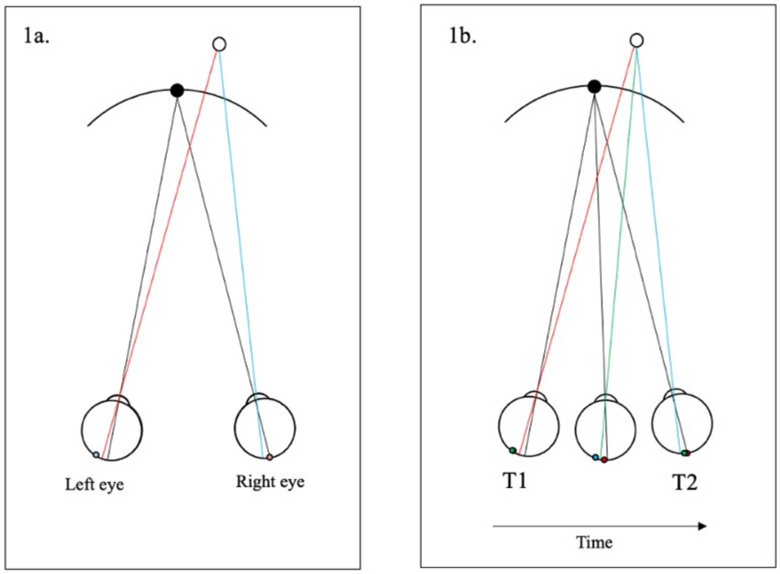

Figure 1. Geometric similarity between stereopsis (left) and self-generated motion parallax (right) as depth cues. In Figure 1a., the red and blue lines represent where the distant stimuli reach the retina in the left and right eyes respectively. The red and blue dots show the corresponding points from the opposite eye. In Figure 1b. the head translates to the right. T1 (red) and T2 (blue) represent the initial and final positions respectively, with the middle eye showing an intermediate position (green). The coloured dots represent the corresponding points. The position change of the open circle on the retina due to motion parallax is equivalent to the binocular disparity. The retinal disparity is shown as the offset between the lines and the coloured dots in each eye (for stereopsis) and at T1 and T2 (for motion parallax).

To achieve an accurate percept of depth from self-generated motion parallax, one must integrate information from the vestibular and proprioceptive systems alongside visual input (Rogers and Rogers, 1992). As the observer translates, objects behind and in front of the point of fixation will appear to move in the same or opposite directions as the head (Figure 2). In the more complex situation where an object in the scene is also moving, Hayashibe (1993) points out that head motion and apparent stimulus velocity must be compared to compute relative depth. It has also been suggested that the observer's eye rotation may play a role (Nawrot 2003, Nawrot

2006), but the extent to which eye movements inform motion parallax is debated. To form a depth percept from self-generated motion parallax, we also need to compute the position of the head in relation to the stimulus, which involves the vestibular and proprioceptive systems. In doing so, the observer needs to integrate more information to perceive three-dimensionality, which may introduce more noise, while stereopsis mainly relies on the fixed interpupillary distance to calculate depth.

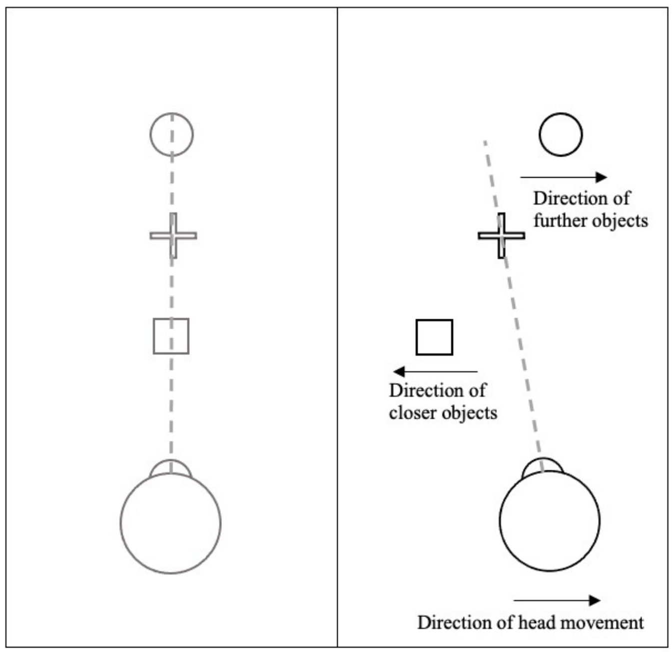

Figure 2. The change in head movement and perceived object motion. The first panel represents an observer fixating on the plus sign, with the square in front of the point of fixation, and the circle behind. The second panel shows that as the observer moves their head towards the right, objects in front of the point of fixation appear to move in the opposite direction of the head, and objects behind appear to move in the same direction as the head.

Studies have compared performance on depth-based tasks to determine whether stereopsis and motion parallax rely on common neural mechanisms, given that depth can be computed via triangulation for both cues. Seminal work by Rogers and Graham (1979) compared the two cues and found that motion parallax and stereopsis are both sufficient cues to depth when presented in isolation. In this study, observers were shown a corrugated surface which they viewed monocularly while moving their head laterally. Observers were asked to match the perceived depth on a stereoscopic display. They found that with smaller amounts of relative motion, observers were able to correctly match the stereoscopic depth to the depth specified by motion parallax, suggesting that the computation of depth from motion parallax is comparable to depth computation via stereopsis. However, in a subsequent study they reported that while the pattern of sensitivity functions for depth from motion parallax and stereopsis were similar, stereopsis produced lower thresholds (Rogers and Graham, 1982).

It is important to note that Rogers and Graham compared stereopsis and motion parallax using simple stimuli that lacked other depth cues. Subsequently, Rogers and Collett (1989) presented a corrugated surface where the two cues were present simultaneously, but in conflict; i.e., at a given location, disparity depicted a peak and relative motion showed a trough (and vice versa). When the binocular disparity modulation was larger than 1 arcmin from peak to trough, the motion parallax cue was perceived as relative motion along a row of dots, rather than relative motion specifying depth. They comment that with increasing reliability of the stereoscopic cue, observers rely less on motion parallax. This shows that when both cues are presented simultaneously but in conflict, depth percepts are primarily driven by stereopsis. Other such 'cue conflict' studies have reported similar biases towards the reliance on stereopsis when presented with structure from motion stimuli (among others see Ichikawa and Saida 2002).

Other studies have examined the interaction between motion parallax and stereopsis under conditions that define the same 3D shape and are not in conflict. Bradshaw and Rogers (1996) used random dot stereograms that formed corrugated surfaces and measured detection thresholds using motion parallax, binocular disparity, and a combination of the two. In a pilot study, the observer's sensitivity to both cues was tested, and compound stimuli were designed for each participant so that the disparity and motion parallax information were proportionate to their sensitivity, which ensured both cues would reach thresholds at the same point. They reported that there was a facilitation effect, where detection thresholds were improved in the combined condition relative to when the individual cues were present in isolation (also shown by Tittle and Braunstein, 1993). Facilitation also implies that both cues rely on a common mechanism to compute depth, rather than treating the cues as independent, as is the case with the modified weak fusion model.

In a more recent study, Hartle and Wilcox (2021) found that when stereopsis and self-generated motion parallax were presented in isolation, performance in a depth magnitude task was more accurate and precise under the stereoscopic viewing condition compared to the motion parallax viewing condition. Moreover, when they had access to both cues, observers performed just as well as when stereopsis was presented alone. They used Bayesian modelling to understand how the two sources of information were combined. They implemented a modified weak fusion model initially proposed by Landy et al. (1995) which assumes that when combining multiple depth cues, the brain weights each cue by its individual reliability. A more reliable cue will be given more weight than a less reliable/noisier cue. This model would predict that the combined threshold will fall in between the two individual thresholds, but would be skewed towards the more reliable cue (as shown by Johnston, Cumming, Landy, 1994). Hartle and Wilcox (2021) found that a simple weighted cue integration approach did not adequately describe the data. Instead, they concluded that the results were better fit by a veto model as described by Bülthoff and Mallot (1988). According to this model, a combined percept threshold is expected to be the same as that for the more reliable cue, and significantly different than thresholds from the more variable cue(s).

The finding by Hartle and Wilcox (2021) that stereopsis alone drives a combined percept has been shown in several earlier studies (LeClair and Durgin, 2008; McKee and Taylor, 2010) which also measured depth magnitude. Due to the involvement of additional systems (vestibular and proprioceptive), the variability of depth estimation tends to be higher with motion parallax. We propose that one of the reasons for the bias towards stereopsis in these studies is the use of a task that requires computation of quantitative depth, and the goal of this work is to assess this possibility.

To this end, we have adopted a curve-tracing task initially used by Jolicoeur, Ullman and MacKay (1986) to assess the impact of different depth cues on the segmentation of curves in depth. In their study, observers were asked to report whether two indicator points were on the same curve, or different curves (Figure 3). Jolicoeur and colleagues found that observers mentally traced the curves to do this task. Subsequently, Sollenberger and Milgram (1993) adapted the task to assess the effects of stimulus rotation on depth perception by manipulating variables such as stimulus rotation speed or quality. They created three-dimensional branching, capillary-like structures and asked observers to report whether the box at the top belonged to branch A or Z at the bottom (Figure 4). They tested conditions in which the depth of the stimulus was defined by binocular disparity, relative motion, or both. They also evaluated the different types of motion, such as continually rotating surfaces, or sequentially presented images. They found that the presence of each cue alone helped disambiguate the stimulus, yet performance was optimal when both cues were present. Moreover, performance was enhanced when observers viewed a continually rotating surface. Milgram's work supports our intuition that this task could be useful for studying the interaction between stereopsis and motion parallax. However, it is important to note that they focused on segmentation information that was provided via stimulus motion (not self-motion), they did not vary the disparity information within the stimulus (it was either present or absent), and they did not evaluate how cues can contribute to a combined percept. Thus, while they showed that object motion is an effective cue for segmenting objects in depth, the relationship between binocular disparity and self-generated motion parallax remains unclear. Here, we build on the work of Jolicoeur, Ullman and MacKay and of Sollenberger and Milgram and further adapt the curve tracing paradigm to study cue integration.

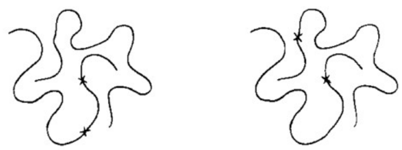

Figure 3. An example of the stimuli used by Jolicoeur, Ullman and Mackay (1986).

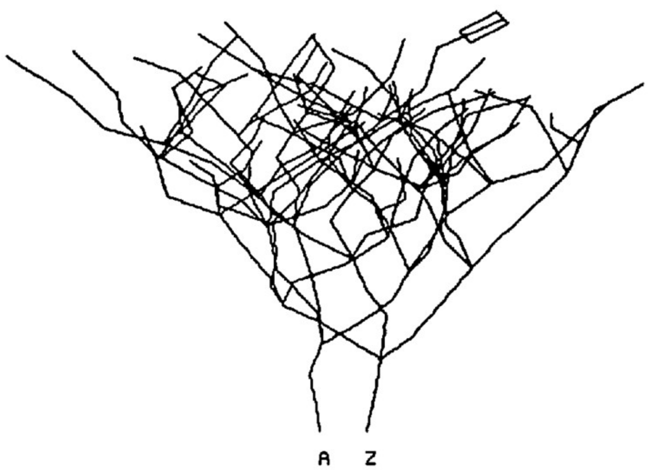

Figure 4. An example of the stimuli used by Sollenberger and Milgram (1993).

Objectives

In isolation, motion parallax is effective at providing the shape and depth of corrugated surfaces, yet when compared to stereopsis alone or in combination, stereopsis seems to consistently outperform motion parallax in depth magnitude tasks (LeClair and Durgin, 2008;

McKee and Taylor, 2010; Hartle and Wilcox, 2021). We propose that this is due to the unreliability inherent in the baseline information provided by motion that is needed to compute depth, while stereopsis relies on a fixed baseline (interocular separation). The modified curve tracing task developed here does not require that observers estimate depth magnitude. Instead, the task simply relies on the ability of motion parallax and stereopsis to promote depth segmentation. Thus, our goal is not to assess how much depth observers perceive, rather the purpose is to investigate how observers are using depth to segment the curves. Arguably this should be less impacted by the variability in the motion signal. In the first experiment, we compared the effects of (1) stereopsis, (2) self-generated motion parallax, and (3) a combination of stereopsis and motion parallax, on depth segmentation. In Experiment 1b, we replicated the self-generated motion parallax condition using a wider range of motion and in Experiment 1c we assessed object motion to evaluate the impact of the range and the type of motion, respectively. In Experiment 2a, we evaluate whether the interaction between stereopsis and motion-based segmentation seen in Experiment 1a is the same for object motion. We also investigate how experience levels impact the observer's ability to use depth cues to segment the curves in depth. In Experiment 2b, we evaluated the effect of stimulus motion speed on experienced observer performance.

EXPERIMENT 1a

INTRODUCTION

The objective of Experiment 1 was to test the new curve-tracing paradigm and to evaluate the ability to use information from stereopsis and motion parallax (in isolation or combination) to segment two curves in depth. Given the novelty of the task, this study will provide baseline data for comparison with subsequent experiments.

METHODS

Observers

A total of twelve observers were pre-screened using a RANDOT™ stereoacuity test. Observers were required to have a stereoacuity of at least 40 seconds of arc and normal or corrected-to-normal vision. Observer's inter-pupillary distance (IPD) was assessed before testing using a GR-4 digital pupillometer. Observers were recruited from surrounding labs, and most had experience with psychophysical tasks but were naïve to the experiment. The study was approved by the York University Research Ethics Board.

Stimuli and apparatus

Stimuli were created in Blender (version 2.93) using the Bezier curve function. After generating a curve, we added 0.01m of bevel depth to ensure that they were visible within the headset. A mesh was then created using the generated structure, and subsequently exported as an Filmbox (FBX) file. The experiment was designed using Unreal Game Engine version 4.27, where the FBX file was imported. The experiment was run using a Meta Quest 1 headset, which has a resolution of 1440x1600 pixels per eye, and a refresh rate of 72 Hz. A stimuli pair consisted of two overlaid white curves and two green probes (one on either side of the curves) displayed against a black background. To form a pair, one curve was duplicated, rotated 180° about the horizontal axis and overlaid. The complexity of the curves was varied by manipulating the number of peaks and valleys. Low complexity curves had two to four peaks, while higher complexity curves had five or more peaks (Figure 6).

Figure 5. The planar orientation of the stimuli in Experiment 1. The gray surfaces symbolize the planes which the curves occupy. The two planes are frontoparallel and superimposed. The disparity refers to the distance between the two planes, which ranges according to the predetermined step size.

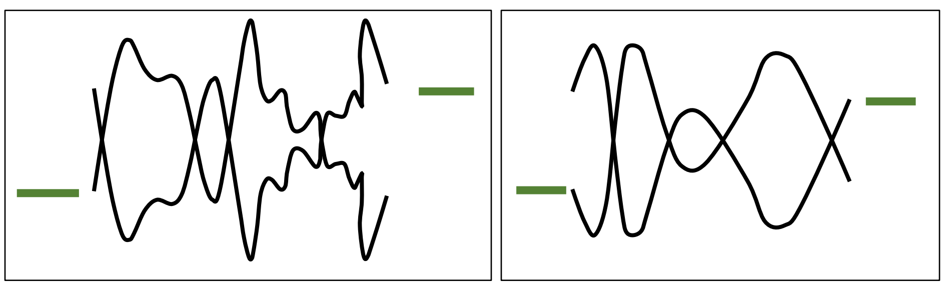

Figure 6. An example of the stimuli used in the study. The left image shows a high-complexity stimulus. The right image shows a low-complexity stimulus. In either case, a single curve is duplicated and rotated 180°. Two smaller green lines are displayed on either side of the curves to specify the endpoints of interest. The probe location is constant throughout trials, always indicating the bottom left and top right ends.

Procedure and Design

Observers were asked to report with a button press whether the probes aligned with the same curve or different curves. Before starting the experiment, observers participated in a truncated version to determine the appropriate depth offset range. Once the appropriate range was determined, observers were tested in three different conditions (1) static binocular condition,

(2) dynamic monocular condition, and (3) dynamic binocular condition. In the static binocular condition, observers' movement was limited by a chinrest. In the dynamic monocular condition, observers wore an eyepatch over their left eye to limit visual input. In the combined condition, observers performed the lateral movement described below while viewing the stimulus binocularly. In both dynamic conditions, observers were asked to sway side-to-side until red boundary walls appeared within the headset. Each observer was told to make at least one complete translation (move from one boundary to the other). The red boundary walls were used to limit observer motion to a total of 6 cm.

The depth offset between the two curves varied according to a method of constant stimuli with 6 test values bracketing zero disparity. We increased the distance between the curves with a step size of ± 0.2 cm, 0,4 cm, or 0.6 cm, which produced maximum separations between the curves of ± 1.2 cm, 2.4 cm, and 3.6 cm respectively. Each depth offset was tested in random order and the proportion correct data was fit with a psychometric function to compute the threshold (the 75% correct point). Conditions were randomized across observers to minimize order effects. Initially, each observer completed 140 trials per condition (2 complexities x 7 depth levels x 10 trials = 140 trials total), then they repeated this process in a subsequent session for a total of 280 trials per condition.

RESULTS AND DISCUSSION

Figure 7 shows the average proportion correct in the low complexity condition. Observers performed well above chance at zero depth offset, when there was no separation between the two curves. In general, we found that these data could not be meaningfully fit with a psychometric function so thresholds could not be computed.

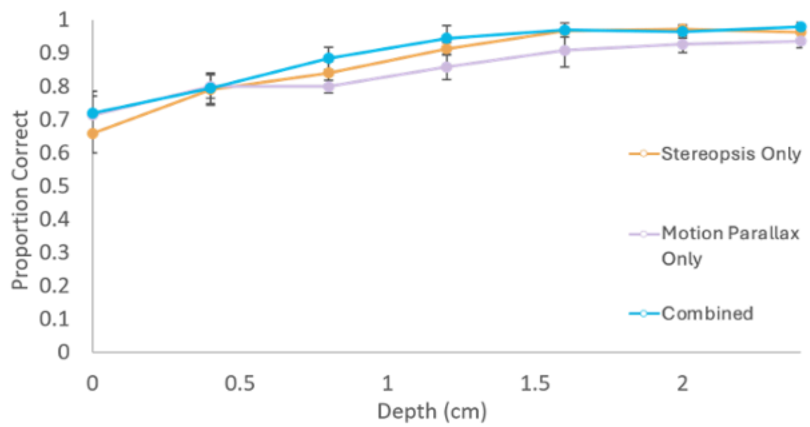

Figure 7. The average proportion correct of 12 observers across a range of depths. The data is obtained from the low complexity condition. The error bars represent SEM.

Even when there was no depth offset, observers performed well above chance in the low complexity condition. This is not the case for the high-complexity results. The psychometric fit for one representative observer in the high complexity condition is shown in Figure 8. The proportion correct data was fit by a cumulative normal distribution. If an observer's data could not be fit by a cumulative normal, or if the 75% threshold was outside the test range, they were considered outliers and were excluded from further analysis. Two observers were removed from average calculations since their data could not be fit with a psychometric function. Figure 8 shows performance from one observer who performed similarly for the stereopsis and combined conditions. In this case, at smaller depth offsets curve segmentation was worse in motion parallax only condition.

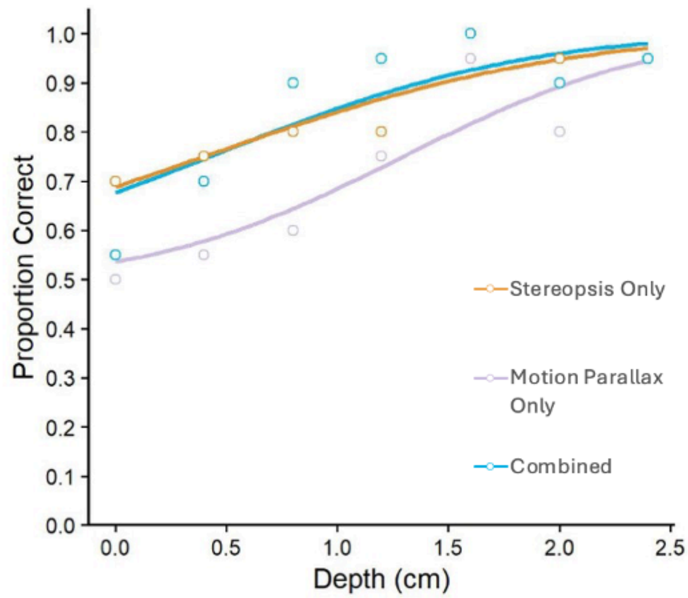

Figure 8: One observer's proportion correct data (shown by the open circles) fit with a cumulative normal function (solid lines) for the high complexity condition. Each data point represents the average of 20 responses.

Figure 9 shows the threshold values for the high-complexity data averaged across 10 observers. A lower threshold means that less depth is needed to accurately segment the two curves with 75% accuracy. The average threshold was highest for the motion parallax condition and lowest for the combined condition. It appears that performance in the combined condition is better than when motion parallax is presented in isolation. This is confirmed by a repeated measures ANOVA, which showed a statistically significant effect of cue ($F(2, 18) = 7.01$, $p<0.01$) and a pairwise t-test adjusted for false discovery rate using a Benjamini & Hochberg (B-H) adjustment. There were no significant differences between the motion parallax only condition and the stereopsis only condition ($p=0.21$), and the stereopsis only and combined condition ($p=0.10$). There was, however, a significant difference between the motion parallax only and the combined condition ($p=0.04$).

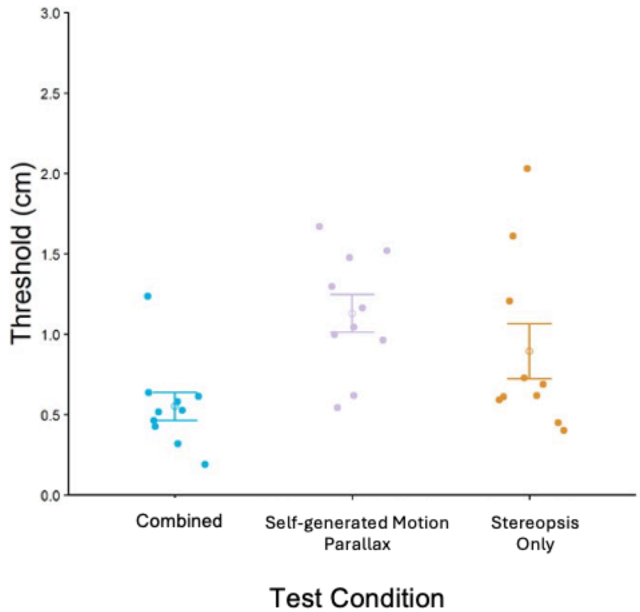

Figure 9. Average thresholds (75%) are plotted for the combined, self-generated motion parallax, and stereopsis conditions (n = 10). The open circle shows the average threshold, while the closed circles show individual thresholds. This graph shows the results for the high complexity condition only. Error bars show the standard error of the mean.

As outlined above, the low complexity condition was not sufficiently challenging and performance was consistently at or above 75% correct. Given this, from this point on we will consider the high complexity condition alone.

We predicted that by using a segmentation task, the uncertainty regarding the parallax baseline from self-generated motion would have less impact on performance making it more comparable to the stereopsis-only condition. Our results do not contradict this prediction, as there was no significant difference between the thresholds in the two isolated cue conditions. It is also evident from Figure 9 that when both cues were available, performance was significantly better than when observers only had access to the motion parallax cue, suggesting that the addition of stereopsis enhances performance. There was a significant difference between our combined and self-generated motion thresholds. This suggests that the addition of relative motion to a static, binocular viewing condition does not significantly improve performance. One explanation for the latter result could be that the range of motion was too small and a larger range might increase the impact of the motion parallax cue. To evaluate this, we conducted a follow-up study where we replicated the motion parallax condition with twice the total range of motion.

EXPERIMENT 1b

INTRODUCTION

In the previous study, we saw that there was no significant difference between performance for the isolated cues, and performance in the combined condition was not significantly different from stereopsis alone but it was significantly different from self-generated motion parallax. The motion alone condition had higher thresholds. This pattern suggests that while motion parallax alone can be effectively used to segment curves in depth, it may not be as effective as when both cues are available.

In Experiment 1b, we doubled the range of motion to 12 cm to investigate if the added range of motion improves thresholds by providing a stronger relative motion signal. Additionally, we also tracked head position to confirm that observers were translating the full amount specified by the motion boundary walls. We then compared this data to the self-motion condition from Experiment 1a.

METHODS

Observers

A subset of five observers from Experiment 1a participated in this follow-up experiment.

Stimuli and apparatus

The stimuli, apparatus and range of depth offsets were the same as in Experiment 1a. Viewing time was also limited to 4 seconds, yet observers were able to terminate the trial by responding. The key difference is the change in translation distance specified by the red boundary walls. In Experiment 1b, the red boundary walls specify 12 cm of total side-to-side translation, or approximately double the average observer IPD. The observer's head motion was tracked using the Quest 1 head tracking system.

Procedure and Design

The design of Experiment 1b was the same as Experiment 1a, with the key difference being that observers were instructed to move a total of 12 cm. Observers completed this block once, for a total of 10 trials per point.

RESULTS AND DISCUSSION

Figure 10 shows the thresholds for the two motion conditions for the same 5 observers. This figure shows that thresholds are very similar in the two conditions, which is confirmed by a pairwise t-test with a BH adjustment ($t(4)= -0.22$, $p=0.84$).

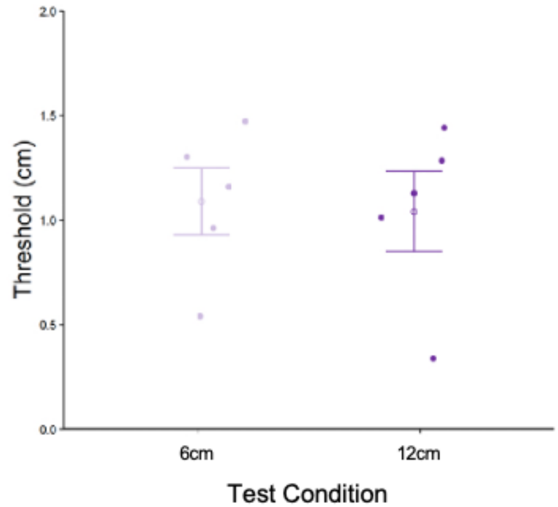

Figure 10. The average 75% thresholds for 6cm and 12 cm of observer motion. The threshold was obtained from psychometric fits. The dark purple represents the data collected from 5 observers in Experiment 2. The light purple shows the 75% threshold for the same 5 observers from Experiment 1a. The error bars represent the standard error of the mean.

To verify that our motion extent manipulation was effective, we analyzed the movement paths to see if observers constrained their motion as instructed. Figure 11 shows the average translation distance for all observers for the 6 cm and 12 cm motion conditions. It is clear that regardless of the range of motion specified by the motion boundary walls, observers moved a total of 20 cm per translation, or well beyond the specified range. It is possible that this range of motion is the most physically comfortable for a seated observer when asked to move their upper body from side to side.

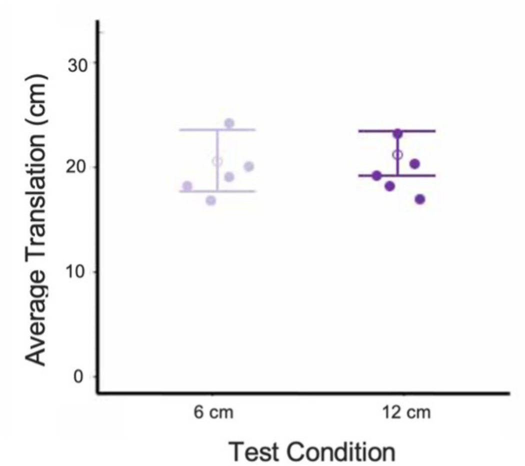

Figure 11. The average head translation of 5 observers. The open dark purple represents the average translation across all observers when 12 cm was specified. The open light purple shows the results for the same observers in Experiment 1a when the translation was 6 cm. The filled circles represent the average distance moved from each observer. The error bars show the standard error of the mean.

Our results show no difference in performance with increasing the motion baseline, although this is most likely due to the fact that observers were moving 20cm in both conditions. Given that, we can be confident that the limited range of motion was not the cause of the relatively poor motion segmentation in Experiment 1a. Given that the extent of motion does not explain the pattern of results in Experiment 1a, we then asked about the role of self-motion. While vestibular and proprioceptive input is not required to compute relative depth magnitude in this segmentation task, demand is placed on the visuomotor and cognitive systems as they try to maintain fixation, segment the curves in depth and make their response. To evaluate the effects of maintaining self-motion during this task, in Experiment 1c we use object motion in place of self-motion. If the requirement to move is disrupting the task, we expect to see improved performance in the object motion condition when compared with the self-generated motion parallax condition.

EXPERIMENT 1c

INTRODUCTION

In the previous experiment, we saw even when observers were moving a total of 20 cm, it appeared that they did not use the information provided by motion when they had access to stereopsis and motion parallax. This confirms that the strength of the motion parallax signal is not limiting observers' ability to segregate the two depth planes. In a second follow-up experiment, we assess whether the act of moving side to side was responsible for elevating thresholds. In this case, while the vestibular and proprioceptive input are not needed to compute depth, the requirement to move may simply have made the task more difficult. To assess this we created versions of the stimuli with the same motion profiles that were viewed by stationary observers.

METHODS

Observers

A total of six observers from Experiment 1a participated in this experiment.

Stimuli and Apparatus

The stimuli and apparatus used here were the same as those used in Experiment 1a. The only difference is that the observer was stationary, with their head in a chin rest, and the object moved. The same stimuli were used from the previous experiment, but they were moved to replicate the change in angle that an observer moving 6 cm would experience if they were moving their heads laterally. This calculation was done for each surface pair at all step sizes.

Procedure and Design

Each individual participated in a single session with ten trials per depth offset for each condition. The results were then fit with a psychometric function to obtain a 75% threshold, and these thresholds were compared to those from the monocular condition in Experiment 1a.

RESULTS AND DISCUSSION

Figure 12 shows the average threshold of the object motion condition collected from six observers in green. The light purple point plots the average threshold from the same six observers in Experiment 1a. A pairwise t-test (with a Benjamini-Hochberg correction) does not show a significant difference between the two thresholds ($t(5)=-2.36$, $p=0.12$). While the thresholds are not significantly different, the smaller SEM suggests that observers are more accurate when they have access to object motion versus when they rely on self-motion.

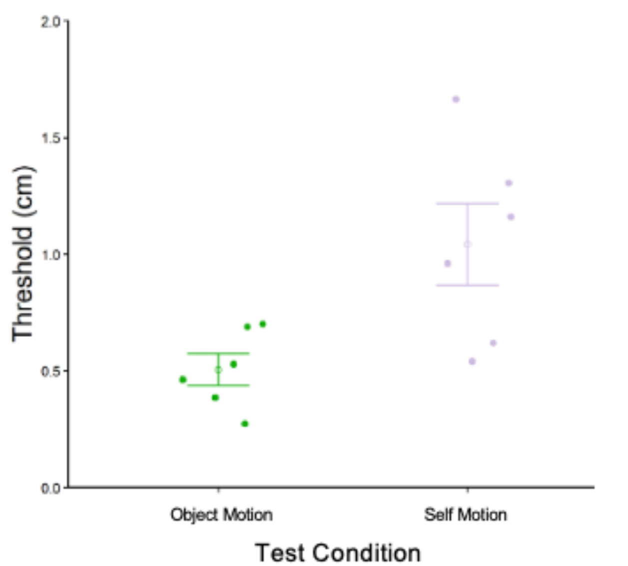

Figure 12. Average thresholds plotted for object motion and self-motion. Threshold values were computed from individual psychometric functions are shown by the open circles. The green shows the thresholds collected from six observers in Experiment 1c, and the light purple

represents the same observer's data from the self-generated motion parallax condition in Experiment 1a. The filled circles show individual thresholds.

The error bars show that observers are less variable in their judgments when rely on object motion to do the task. This may support the argument that the self-motion signal introduces additional noise that is not present in the object motion condition.

Given the increased reliability seen in the object motion condition, we hypothesized that when presented along with consistent binocular disparity information, combined performance would likely differ from that seen in Experiment 1a. To assess this, in a subsequent study we replicated Experiment 1a but used object motion rather than self-motion. If the reduction in variability promotes cue integration we might expect to see thresholds in the combined condition that fall between those of the isolated cue conditions. Given that the same group of observers participated in Experiments 1a-c, in Experiment 2a we tested both experienced and novice observers to evaluate the role of experience.

EXPERIMENT 2a

INTRODUCTION

The results of our previous study show that thresholds for curve segmentation based on object motion were less variable than when segmentation is based on self-motion. Yet it is unclear how this more reliable source of information will compare to stereopsis, and how it will be integrated to form a combined depth percept. Therefore, in Experiment 2a, we replicate the conditions from Experiment 1a using object motion instead of self-motion. We also consider that the reduction in variability could be the result of practiced observers since we used the same observers in Experiments 1a-c. Therefore, in this experiment, we will also evaluate the role that observer experience plays in segmenting curves in depth.

METHODS

Observers

Twenty inexperienced observers and eleven experienced observers were pre-screened using a RANDOT™ stereoacuity test and were required to have a stereoacuity of at least 40 seconds of arc and normal or corrected-to-normal vision. An experienced observer was considered to be an individual who had participated in stereoscopic psychophysical tasks for a year or more before participating in the present study.

Stimuli and Apparatus

The stimuli used here were the same as in previous experiments. The moving stimuli were the same as those from Experiment 1c, in which observers viewed with an eyepatch over their left eye for the isolated object motion condition, and binocularly for the combined condition. Static stimuli were the same as the high-complexity stimuli from Experiment 1. Observers used a Meta Quest 2 headset which has a resolution of 1832x1920 pixels per eye and a refresh rate of 120Hz. Observer IPD was also measured using a GR-4 digital pupillometer, and the headset was adjusted to match IPD.

Procedure and Design

As in Experiment 1 we used a method of constant stimuli with six test offsets (defined via motion and/or disparity) as outlined in Experiment 1. The same step sizes from Experiment 1a were used, resulting in depth offsets ranging from ± 1.2 to 3.6 cm, which resulted in a maximum lateral stimulus displacement of 3.002 cm to 3.1 cm for the front curve, and 2.96 cm to 2.89 cm for the back curve. Participants were first tested with a brief version of the isolated cue conditions to determine the appropriate step size and to ensure that they understood the task. Then participants were then tested in three conditions; (1) stereopsis only, (2) object motion only, and (3) the combination of the two. The three conditions were counterbalanced across observers to control for order effects. Each block consisted of seven depth levels (determined by the pretest) and each depth level was tested 20 times, for a total of 140 trials. The proportion correct data was then fit using a cumulative normal distribution to extract the 75% threshold.

RESULTS AND DISCUSSION

Figures 13a and b show the average 75% thresholds for inexperienced and experienced observers respectively. Of the 20 inexperienced observers, five participants were excluded because their data could not be fit with a psychometric function, and another five participants were excluded because their 75% threshold was outside the range of depth values that was tested. In contrast, all 11 of the experienced observers fulfilled the inclusion criteria and are included in Figure 13b.

From Figure 13 it is clear that there is a difference between how the novice and experienced observers perform the depth segmentation task. A mixed ANOVA confirms these results, showing that there is a significant effect of experience, the results of which are shown in Table 1. Inexperienced observers seem to be most effective at segmenting curves when they had access to a combined viewing of the stimulus, and they were better at segmenting the two curves when they had access to object motion instead of stereopsis (Figure 13a). These findings were supported by a repeated measures ANOVA ($F(2, 18) = 18.99$, $p < 0.001$) and a pairwise t-test (with an Benjamini-Hochberg adjustment) which shows that stereopsis thresholds are significantly higher than object motion thresholds ($p = 0.04$) and combined thresholds ($p = 0.01$), and that object motion thresholds are significantly higher than combined thresholds ($p < 0.01$). In Figure 13b., we can see that experienced observers were better at segmenting curves when they were able to use information from object motion than when they relied on stereopsis (much like the inexperienced observers), yet there was no difference between object motion and combined thresholds. Again, these results are supported by an ANOVA ($F(2, 20), = 12.71$, $p < 0.01$) and a pairwise t-test, which shows that there is a significant difference between object motion and stereopsis thresholds ($p < 0.01$) and combined and stereopsis thresholds ($p < 0.01$), but no significant difference between the object motion condition and the combined condition ($p = 0.13$). With respect to observer experience levels, all three conditions show higher thresholds for the inexperienced group. There is also a notable difference in variability across the two groups; inexperienced observers exhibit consistently larger SEM. Previous exposure to VR headsets or 3D displays may be an important difference between our inexperienced and experienced groups. Head-mounted displays, like the Quest 2 used in the present study, are known to cause cue conflicts, like the accommodation-vergence conflict. It may take practice to become accustomed to a VR headset, which could also explain why inexperienced observers are more prone to error.

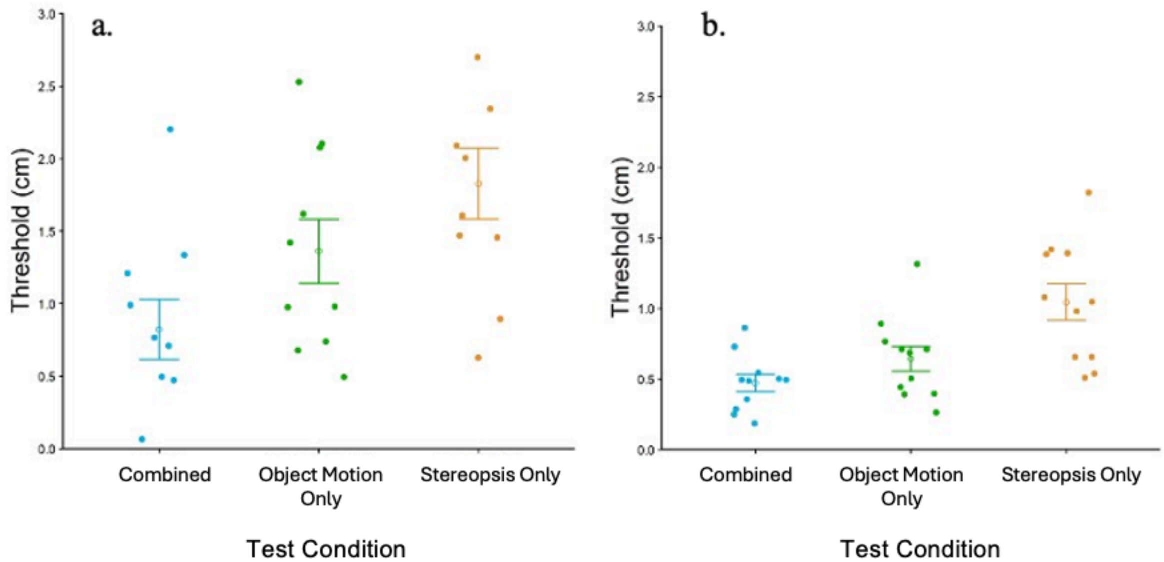

Figure 13. Average thresholds plotted for combined, object-motion, and stereopsis conditions. Figure 13a. shows the 75% thresholds for inexperienced observers and b. plots the 75% thresholds for experienced observers. The open points represent the average thresholds, and the filled points represent individual observer thresholds. The error bars represent SEM.

Source	DFn	DFd	F	p
Experience	1	19	8.86	$p<0.01$
Condition	2	38	31.19	$p<0.001$
Experience: Condition	2	38	2.82	$p=0.07$

Table 1. A mixed ANOVA table with results for Experiment 2a.

Despite the effects of experience discussed above, the data show that both groups of observers may be more effective at using motion information to segment the curves, unlike Experiment 1, where stereoscopic and self-motion thresholds were the same. The improvement in the object motion condition shown in Figure 13 supports our hypothesis that the observers' tendency to rely on the information from stereopsis alone during the combined viewing is due to the added effort of moving while making a judgment. Inexperienced observers show a significant improvement in the combined condition when compared to each of the isolated cues. These results echo the facilitatory effects described by Bradshaw and Rogers (1996) and suggests that for these observers, the visual system integrates the information from the two depth cues non-linearly.

In contrast, experienced observers do not show this facilitation effect. Performance in the combined cue condition is not statistically different from that of object motion alone; it appears that they disregard the information from stereopsis. This is the reverse of the results of Experiment 1a, where the combined percept seemed to be driven by stereopsis alone. The difference in integration across experience levels, again, could be due to more naïve observers needing more information to complete a task within an impoverished viewing condition like that present in our study. Inexperienced observers may simply need more information than experienced observers to complete a psychophysical task, and are therefore more likely to use all the depth cues available to them. It appears that practiced observers can complete the same task with less information and may even ignore depth information from certain cues.

EXPERIMENT 2b

INTRODUCTION

In Experiment 1a, where we found that when observers were moving 20 cm, self generated motion parallax was as effective as stereopsis in isolation, yet it did not contribute to a combined depth percept. In Experiment 2a, we saw that regardless of experience level, observers needed less depth to segment two curves when they experienced 6 cm of object motion alone compared to stereopsis in isolation (as shown by the lower thresholds). Given that the trial duration is fixed and observers are asked to move their heads one full cycle, this means that the stimulus moved faster on the retina than expected. In Experiment 2b, we compare segmentation performance for object motion alone at 6 cm and 20 cm (their true movements) to understand what effect this might have had on thresholds. That is, we replicate the object motion condition seen in Experiment 2a with a faster stimulus to verify that the improved thresholds seen in the object motion condition were not the result of the slower motion.

METHODS

Observers

Six experienced observers who participated in Experiment 2a also participated in Experiment 2b.

Stimuli and Apparatus

The stimulus and apparatus in Experiment 2b is the same as described for the object motion condition in Experiment 2a. Observers are kept stationary with a chinrest, and are asked to use an eyepatch to limit visual input to one eye. The key difference in Experiment 2b is the speed at which the curve pair is completing a translation. In the analysis of the motion data collected from Experiment 1a observers, we have found that observers are moving 20 cm, and are taking an average of 2.5 seconds to complete a full translation (from right to left), meaning that the observer is moving at a rate of 8 cm per second. Contrarily, in Experiment 2a, the object was moving at a rate of 6 cm per second. Therefore, in Experiment 2b, we have increased the speed with which the object is moving to match the speed at which the observer is moving at when they complete a 20 cm translation.

Procedure and Design

The procedure is similar to that used in Experiment 2a but the participants only completed one object-motion condition which simulated 20 cm of observer motion. The psychometric functions were fit with a cumulative normal distribution to compute the 75% thresholds.

RESULTS AND DISCUSSION

Figure 14 shows the average thresholds obtained from the individual psychometric fits of the 6 observers in the slow object motion condition from Experiment 2a (light green) and the faster motion (dark green). A pairwise t-test with a BH correction shows that there is a significant difference between the two conditions ($t(5)=-2.67$, $p=0.04$), with thresholds in the fast-motion condition slightly lower on average than the original motion condition.

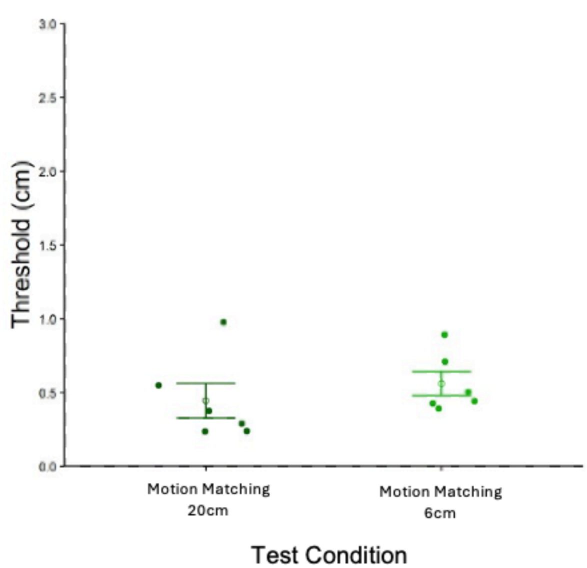

Figure 14. Average thresholds plotted for motion matching 20 cm and motion matching 6 cm. Threshold values were computed from individual psychometric functions. The light green shows the thresholds collected from 6 observers in Experiment 2a, and the dark

green represents the same observer's data from the object speed matching 20 cm of motion collected in Experiment 2b.

Observers exhibited lower thresholds on our segmentation task when they were viewing motion consistent with a larger head movement over the same period of time, resulting in higher movement speeds. If the stimulus speed was responsible for the reduction in sensitivity, then thresholds in the Motion Matching 6 cm condition would be significantly lower than in the Motion Matching 20 cm condition. Our results do not support this prediction so we conclude that the self-motion distance was not responsible for poor performance in Experiment 1a.

GENERAL DISCUSSION

SUMMARY

The goal of this thesis was to explore the role of task demands on how depth from stereopsis and motion parallax are combined. Motion parallax in isolation has been shown to be a sufficient cue for informing depth (Gibson et al. 1959; Steinbach, Ono, Wolf, 1991; Ono and Wade, 2005). However, work comparing motion parallax to stereopsis yields inconsistent results, with some arguing that motion parallax is comparable to stereopsis (Rogers and Graham, 1979), while others suggest it is worse than stereopsis at informing depth percepts (Rogers and Graham, 1982; Rogers and Collett, 1989; Durgin et al. 1995). In Experiment 1a we introduce a depth segmentation task that relies on curvature tracing; we see that there was no significant difference between thresholds from the stereopsis-only condition and the self-generated motion parallax only condition. Additionally, we find that there is no difference in thresholds between the combined condition and stereopsis only condition, however, combined thresholds are significantly better than when motion parallax is present in isolation. Taken together, the results of this experiment can suggest that there is no difference in performance when observers are segmenting two curves when presented with either cue (stereopsis and motion parallax) in isolation. However, when observers have simultaneous access to both cues, they appear to rely more heavily on information from stereopsis, and the addition of depth information from self-generated motion parallax does not improve performance. In Experiment 1b we assess if this was due to the relatively noisy motion parallax information by increasing the range of motion.

In Experiment 1b, we doubled the range of motion from 6cm to 12cm and found that there was no difference in performance between the two conditions. Importantly, when analyzing the motion tracking data, we found that regardless of condition, observers were in fact moving an average of 20cm. This explains why there was no significant change in Experiment 1b thresholds, since the amount the observer moved in both conditions was the same. However, it is interesting that even when observers have access to a larger range of motion (20cm), thresholds do not improve. It seems clear that monitoring self-motion through proprioception and the vestibular system degrades the relative depth information. If so, then object-based motion should improve performance.

To test this, in Experiment 1c, we replicated the motion parallax condition using object motion. We found that there was no significant difference in thresholds, yet variability was lower in the object motion condition. In a subsequent study (Experiment 2b) we confirmed that the speed of object motion was not a factor. To assess the impact of object motion on cue integration we replicated the conditions in Experiment 1a with object motion instead of self-motion and tested separate groups of experienced and inexperienced observers.

The results show that experienced observers have significantly lower thresholds in the object motion condition than in the stereopsis-only condition, suggesting that these observers are better at using the information from object motion than they are at using the information from stereopsis. Interestingly, there is no significant difference between the combined and object motion thresholds, suggesting that stereopsis may not advantage performance when observers have access to object motion. We found a similar pattern when looking at our inexperienced observers. Here, the stereopsis-only condition still showed a larger threshold than object motion, yet we did find that a combined percept is driven by both cues since combined thresholds were significantly lower than each of the isolated conditions. We also see that inexperienced observers are less consistent, as there is large variability both within and between observers. Generally, regardless of experience level, observers perform better when they have access to object motion than when they have access to stereopsis alone. Experienced observers appear to be more biased to using the information provided by object motion, yet naïve observers depend on both cues to perform depth segmentation.

In sum, while self-motion may provide sufficient information to do depth segmentation at levels equivalent to stereopsis, observers seem to rely on stereopsis when both sources of information are present. When observers perform the same task using object motion, we find that segmentation performance is superior to stereopsis-only condition, and the motion cue appears to be relied upon when both are present. Importantly, this latter result critically depends on the level of experience with psychophysical experiments using 3D displays.

THE EFFECT OF CURVE COMPLEXITY

In Experiment 1a, we evaluated the role of curve complexity by introducing two types of curves; a low complexity curve, with 4 or fewer modulations, and a high complexity curve, with 5 or more modulations. Even when there was no depth separation, observers were able to distinguish between the two low-complexity curves correctly around 70% of the time. This is most likely because observers were able to mentally trace the stimuli (Jolicoeur et al., 1986), suggesting that they were using non-depth related cues to separate the two curves. Given this finding, we removed the lower complexity curves from subsequent experiments.

THE ROLE OF EXPERIENCE

Observer experience levels can substantially impact performance in psychophysical tasks and VR tasks. In Experiment 2a, we found that inexperienced observers had elevated thresholds and were less consistent than experienced observers. All of our experienced participants had been involved as observers in numerous psychophysical experiments. Moreover, all of the experienced observers had extensive practice with VR headsets in experimental settings. In contrast, our inexperienced observers had minimal or no practice with psychophysical tasks in VR.

Studies with impoverished viewing conditions, such as this curve segmentation task, are different from the natural world that we normally interact with. In our study, the stimulus is limited to two white curves and two green probes against a black background. Other than the depth cue(s) of interest to us, there is little visual information that is available to the observer to disambiguate the two curves.

The level of experience an observer has with virtual reality headsets, like the head-mounted displays (HMDs) used in this study, could affect an observer's performance in many ways. One known limitation of head-mounted displays is the vergence-accommodation conflict. Vergence refers to the rotation of the eyes that is necessary to fixate on a point, while accommodation refers to the change in the lens required to bring an object into focus. Under normal viewing conditions, vergence and accommodation covary. In a head-mounted display, observers can change their vergence substantially, but accommodation is fixed to the screen plane. In other words, there is an uncoupling of the two cues (Hoffman et al. 2008). The Meta Quest 2 headset, used in Experiment 2a, has a focal plane distance of 1.3 m, whereas our curves were placed at or around 1 m away from the observer. With such a small difference between the two, it is unlikely that the vergence-accommodation conflict affected the performance of inexperienced observers.

Another limitation of head-mounted displays is the consistent misestimation of absolute distance (Li, Phillips, Durgin, 2011; Messing & Durgin, 2005; Naceri et al., 2010). In an impoverished viewing environment, there are few depth cues that inform the absolute distance of an object, so depth estimation is thought to obtained from a combination of vergence eye movements and vertical disparities, although the strength of each cue seems to be dependent on the distance or size of the object or display (Foley, 1980; Rogers and Bradshaw, 1995). Given the importance of vergence cues, it is perhaps not surprising that this phenomenon of absolute distance underestimation can be worse for inexperienced observers since it has been shown that naïve participants are consistently less accurate in their vergence responses. (Horwood and Riddell, 2010). In contrast, our experienced observers who have used VR headsets are likely more accurate in their vergence responses under these conditions, which could improve their ability to estimate absolute distance, thus enhancing their depth segmentation performance.

Under normal viewing conditions in the real world, we rely on a multitude of consistent depth cues in physical environments; the limited access to additional depth cues in the simulated environment could degrade performance by introducing additional variability. Experienced observers, on the other hand, may be accustomed to the reduced information available to them and are better at attending to the visual information that is relevant to the task (Gibson, 1953, Fendick and Westheimer, 1983; Stransky, Wilcox, and Allison, 2014; Chopin, Levi and Bavelier, 2017), which could contribute to their improved performance. With respect to our study, the experienced observers participated in all experiments, so by the time they completed Experiment 2a, they also had extensive experience with this particular curve segmentation task.

TASK DEPENDENCE OF CUE INTEGRATION

As outlined in the Introduction, studies that have compared depth processing from motion parallax and stereopsis have found similarities, but stereopsis is often used preferentially over motion parallax. We have pointed out that this is true particularly when magnitude estimation is involved. We have argued that such tasks bias observers towards using stereopsis because it potentially has a less noisy estimate due to the fixed baseline (IOD). Depth from motion parallax, however, relies on encoding head position, based on vestibular and proprioceptive signals, which adds variability to the baseline used for triangulation. Thus, it is possible that the poor performance from motion parallax can be attributed to the nature of the task used to investigate psychophysical performance, as suggested by Bradshaw, Parton, Eagle (1998) and Bradshaw, Patron and Glennerster (2000). In this series of experiments, we have introduced a curve tracing task that requires that observers simply segregate two curves in depth. As outlined above, we have suggested that self-generated motion parallax can provide sufficient depth information to support depth segmentation, and observers seem to perform similarly using self-motion as they do with stereopsis. Further, we also found that motion-based depth cues aid in segmentation, with or without the addition of stereopsis, which we discuss in the following section. In sum, we have successfully identified a task that reduces or eliminates the bias towards using stereopsis that is otherwise present in depth estimation tasks. While Bayesian modelling was not possible here, the data from Experiment 1a resembles cue-vetoing. In this case, while performance on the segmentation task appears to be similar when observers have access to self-motion and stereopsis in isolation, they seem to ignore the information from self-motion, when both cues are available. Our expert observers in Experiment 2a also show a pattern of results that resembles cue vetoing (Bülthoff and Mallot, 1988), however here it seems that stereopsis is ignored, and object motion appears to drive the combined percept. Unlike our experts, our naïve observers seem to integrate the two cues to achieve better performance than when each cue is presented alone, much like a facilitation model would predict. It is clear from this work that when multiple cues are available there is a strong interaction between the task demands and observer experience. Modelling such integration processes requires careful consideration of both of these factors.

SELF-MOTION VERSUS OBJECT MOTION

Our results show that observers are better at using the information from object motion than they are at using information from self-motion to segregate stimuli in depth. There are reports that observers are better at using information from self-motion than object motion for depth judgements. Rogers and Graham (1979) found that both a moving observer and a moving oscilloscope can provide sufficient information to serve as an independent cue to depth, yet they found that depth was more pronounced in the self-motion condition. Likewise, Ono and Steinbach (1990) found that observers were more likely to perceive depth in random-dot stereograms using head motion instead of object motion. As noted previously these studies rely on quantitative depth perception, and not just segmentation, which may explain the different pattern of effects. In addition, in these studies the researchers introduce a fixation point that lies at a midpoint within the stimulus, resulting in a shearing retinal motion whereas in our study, observers fixate on the frontal surface/curve so all of the retinal motion is in the same direction. Likely, the additional information gained from the patterns of relative retinal motion aided depth segmentation in those studies, which could also explain why we saw poor performance when relying on self-motion.

CONCLUSION

In this series of experiments, we have evaluated a new depth segmentation task to give insight into how stereopsis and motion-based depth information are combined. The results suggest that observers can use stereopsis and self-generated motion parallax to do this task in a similar way. However, when both are presented in combination, there remains a bias towards using stereopsis. Our subsequent studies show that the use of these depth cues depends critically, not only on the variability within cues (e.g. self-motion vs. object motion) but also on the amount of experience observers have with the tasks and the viewing environment. The role of experience is rarely considered in studies of cue combination, and this issue requires further study to answer key questions such as 1.) what type of experience is necessary to see these differences, 2.) how much experience is needed to alter how this information is used and 3.) by what means does training affect the use of different depth cues?

REFERENCES

Bradshaw, M. F., Parton, A. D., & Eagle, R. A. (1998). The Interaction of Binocular Disparity and Motion Parallax in Determining Perceived Depth and Perceived Size. *Perception*, *27*(11), 1317–1331. https://doi.org/10.1068/p271317

Bradshaw, M. F., Parton, A. D., & Glennerster, A. (2000). The task-dependent use of binocular disparity and motion parallax information. *Vision Research*, *40*(27), 3725–3734. https://doi.org/10.1016/S0042-6989(00)00214-5

Bradshaw, M. F., & Rogers, B. J. (1996). The Interaction of Binocular Disparity and Motion Parallax in the Computation of Depth. *Vision Research*, *36*(21), 3457–3468. https://doi.org/10.1016/0042-6989(96)00072-7

Bülthoff, H. H., & Mallot, H. A. (1988). Integration of depth modules: stereo and shading. *Journal of the Optical Society of America A*, *5*(10), 1749–1758. https://doi.org/10.1364/JOSAA.5.001749

Chopin, A., Levi, D. M., & Bavelier, D. (2017). Dressmakers show enhanced stereoscopic vision. *Scientific Reports*, *7*(1), 3435. https://doi.org/10.1038/s41598-017-03425-1

Durgin, F. H., Proffitt, D. R., Olson, T. J., & Reinke, K. S. (1995). Comparing depth from motion with depth from binocular disparity. *Journal of Experimental Psychology: Human Perception and Performance*, *21*(3), 679–699. https://doi.org/10.1037/0096-1523.21.3.679

Fendick, M., & Westheimer, G. (1983). Effects of practice and the separation of test targets on foveal and peripheral stereoacuity. *Vision Research*, *23*(2), 145–150. https://doi.org/10.1016/0042-6989(83)90137-2

Foley, J. M. (1980). Binocular distance perception. *Psychological Review*, *87*(5), 411–434. https://doi.org/10.1037/0033-295X.87.5.411

Gibson, E. J. (1953). Improvement in perceptual judgments as a function of controlled practice or training. *Psychological Bulletin*, *50*(6), 401–431. https://doi.org/10.1037/h0055517

Gibson, E. J., Gibson, J. J., Smith, O. W., & Flock, H. (1959). Motion parallax as a determinant of perceived depth. *Journal of Experimental Psychology*, *58*(1), 40–51. https://doi.org/10.1037/h0043883

Hartle, B., & Wilcox, L. M. (2021). Cue vetoing in depth estimation: Physical and virtual stimuli. *Vision Research*, *188*, 51–64. https://doi.org/10.1016/j.visres.2021.07.003

Hayashibe, K. (1993). Head Movement Changes Apparent Depth Order in a Motion-Parallax Display. *Perception*, *22*(6), 643–652. https://doi.org/10.1068/p220643

Hoffman, D. M., Girshick, A. R., Akeley, K., & Banks, M. S. (2008). Vergence–accommodation conflicts hinder visual performance and cause visual fatigue. *Journal of Vision*, *8*(3), 33. https://doi.org/10.1167/8.3.33

Horwood, A. M., & Riddell, P. M. (2010). Differences between naïve and expert observers' vergence and accommodative responses to a range of targets. *Ophthalmic and Physiological Optics*, *30*(2), 152–159. https://doi.org/10.1111/j.1475-1313.2009.00706.x

Ichikawa, M., & Saida, S. (2002). Integration of motion parallax with binocular disparity specifying different surface shapes. *Japanese Psychological Research*, *44*(1), 34–44. https://doi.org/10.1111/1468-5884.00005

Johnston, E. B., Cumming, B. G., & Landy, M. S. (1994). Integration of stereopsis and motion shape cues. *Vision Research*, *34*(17), 2259–2275. https://doi.org/10.1016/0042-6989(94)90106-6

Jolicoeur, P., Ullman, S., & Mackay, M. (1986). Curve tracing: A possible basic operation in the perception of spatial relations. *Memory & Cognition*, *14*(2), 129–140. https://doi.org/10.3758/BF03198373

Landy, M., & Brenner, E. (2001). Motion-disparity interaction and the scaling of stereoscopic disparity. In L. R. Harris, & M. Jenkin (Eds.), *Vision and attention* (pp. 129-151). Springer-Verlag.

Landy, M. S., Maloney, L. T., Johnston, E. B., & Young, M. (1995). Measurement and modeling of depth cue combination: in defense of weak fusion. *Vision Research*, *35*(3), 389–412. https://doi.org/10.1016/0042-6989(94)00176-M

LeClair, A., & Durgin, F. H. (2008). Depth interval perception: Comparing binocular stereopsis with motion parallax in "action space." *Journal of Vision*, *8*(6), 857–857. https://doi.org/10.1167/8.6.857

Li, Z., Phillips, J., & Durgin, F. H. (2011). The underestimation of egocentric distance: evidence from frontal matching tasks. *Attention, Perception, & Psychophysics*, *73*(7), 2205–2217.

https://doi.org/10.3758/s13414-011-0170-2

McKee, S. P., & Taylor, D. G. (2010). The precision of binocular and monocular depth judgments in natural settings. *Journal of Vision*, *10*(10), 5–5. https://doi.org/10.1167/10.10.5

Messing, R., & Durgin, F. H. (2005). Distance Perception and the Visual Horizon in Head-Mounted Displays. *ACM Transactions on Applied Perception*, *2*(3), 234–250. https://doi.org/10.1145/1077399.1077403

Naceri, A., Chellali, R., Dionnet, F., & Toma, S. (2009). Depth Perception within Virtual Environments: A Comparative Study Between Wide Screen Stereoscopic Displays and Head Mounted Devices. *2009 Computation World: Future Computing, Service Computation, Cognitive, Adaptive, Content, Patterns*, 460–466. https://doi.org/10.1109/ComputationWorld.2009.91

Nawrot, M. (2003). Eye movements provide the extra-retinal signal required for the perception of depth from motion parallax. *Vision Research*, *43*(14), 1553–1562. https://doi.org/10.1016/S0042-6989(03)00144-5

Nawrot, M., & Joyce, L. (2006). The pursuit theory of motion parallax. *Vision Research*, *46*(28), 4709–4725. https://doi.org/10.1016/j.visres.2006.07.006

Ono, H., & Steinbach, M. J. (1990). Monocular without stereopsis with and head movement. *Perception & Psychophysics*, *48*(2), 179–187. https://doi.org/10.3758/BF03207085

Ono, H., & Wade, N. J. (2005). Depth and Motion in Historical Descriptions of Motion Parallax. *Perception*, *34*(10), 1263–1273. https://doi.org/10.1068/p5232

Ono, M. E., Rivest, J., & Ono, H. (1986). Depth perception as a function of motion parallax and absolute-distance information. *Journal of Experimental Psychology: Human Perception and Performance*, *12*(3), 331–337. https://doi.org/10.1037/0096-1523.12.3.331

Rogers, B., & Graham, M. (1979). Motion Parallax as an Independent Cue for Depth Perception. *Perception*, *8*(2), 125–134. https://doi.org/10.1068/p080125

Rogers, B., & Graham, M. (1982). Similarities between motion parallax and stereopsis in human depth perception. *Vision Research*, *22*(2), 261–270. https://doi.org/10.1016/0042-6989(82)90126-2

Rogers, B. J., & Bradshaw, M. F. (1995). Disparity Scaling and the Perception of Frontoparallel Surfaces. *Perception*, *24*(2), 155–179. https://doi.org/10.1068/p240155

Rogers, B. J., & Collett, T. S. (1989). The Appearance of Surfaces Specified by Motion Parallax and Binocular Disparity. *The Quarterly Journal of Experimental Psychology Section A*, *41*(4), 697–717. https://doi.org/10.1080/14640748908402390

Rogers, S., & Rogers, B. J. (1992). Visual and nonvisual information disambiguate surfaces specified by motion parallax. *Perception & Psychophysics*, *52*(4), 446–452. https://doi.org/10.3758/BF03206704

Sollenberger, R. L., & Milgram, P. (1993). Effects of Stereoscopic and Rotational Displays in a Three-Dimensional Path- Tracing Task. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *35*(3), 483–499. https://doi.org/10.1177/001872089303500306

Steinbach, M. J., Ono, H., & Wolf, M. E. (1991). Motion parallax judgements of depth as a function of the direction and type of head movement. *Canadian Journal of Psychology / Revue Canadienne de Psychologie*, *45*(1), 92–98. https://doi.org/10.1037/h0084272

Stransky, D., Wilcox, L. M., & Allison, R. S. (2014). Effects of Long-Term Exposure on Sensitivity and Comfort with Stereoscopic Displays. *ACM Transactions on Applied Perception*, *11*(1), 1–13. https://doi.org/10.1145/2536810

Tittle, J. S., & Braunstein, M. L. (1993). Recovery of 3-D shape from binocular disparity and structure from motion. *Perception & Psychophysics*, *54*(2), 157–169. https://doi.org/10.3758/BF03211751

Wheatstone, C. (1838). XVIII. Contributions to the physiology of vision. —Part the first. On some remarkable, and hitherto unobserved, phenomena of binocular vision. *Philosophical Transactions of the Royal Society of London*, *128*, 371–394. https://doi.org/10.1098/rstl.1838.0019