Author preprint of 2020 IEEE ISMAR, 73-79, 2020, doi:10.1109/ISMAR50242.2020.00027

Optical distortions in VR bias the perceived slant of moving surfaces

Jonathan Tong*
Centre for Vision Research
Department of Psychology
York University
Toronto, Ontario, Canada

Robert S. Allison†
Centre for Vision Research
Department of Electrical Engineering and Computer Science
York University
Toronto, Ontario, Canada

Laurie M. Wilcox‡
Centre for Vision Research
Department of Psychology
York University
Toronto, Ontario, Canada

ABSTRACT

The magnifying optics of virtual reality (VR) head-mounted displays (HMD) often cause undesirable pincushion distortion in the displayed imagery. Eccentrically increasing magnification radially displaces image-points away from the optical axis, causing straight lines to curve outwards. This, in turn, should affect the 3D perception of surface shape by warping binocular and monocular depth cues. Previous research has shown that distortion-induced biases in perceived slant do occur in static images. However, most use cases in VR involve moving images. Here we evaluate the impact of motion on biases in perceived slant. An HMD was used to present flat, textured surfaces that varied in slant and were either stationary, or translated laterally by the observer. In separate studies we varied the degree of distortion and evaluated the impact on perceived slant at several locations along the surface. We found that, irrespective of whether the surface was moving or stationary, distortion introduced significant bias into local slant estimates. The pattern of results is consistent with the surface appearing to be concave (as if viewing the inside surface of a bowl), as predicted from the warping of binocular and monocular cues. Importantly, the intermediate distortion level produced the same, but weaker, pattern of biases seen in the fully-distorted condition. When an appropriate level of pre-warping was applied, slant perception was veridical. Overall, our results highlight the importance of sufficiently correcting for optical distortions in VR HMDs to enable veridical perception of surface attitude.

Index Terms:

Human-centered computing—Virtual reality; Human-centered computing—Empirical studies in HCI; Computing methodologies—Perception; Computing methodologies—Virtual reality

1 INTRODUCTION

With increasing demand and interest in virtual reality (VR), human perception and performance while using a head-mounted display (HMD) is of growing concern [1, 7, 10, 11, 15, 16, 21, 24, 27]. One of the less frequently studied aspects of HMDs is the impact of image distortion due to display optics on perception and performance. Modern VR HMDs contain powerful lenses to focus images and make virtual objects appear at a distance from the user. However, the magnification power is not uniform, as it typically increases with the square of the radial distance from the optical axis, making images appear distorted [7]. This pattern of distortion is known as "pincushion distortion" as it induces a pincushion-like shape in rectangular images. This distortion can be modelled as a remapping of image points radially outwards, according to a polynomial function (usually 3rd or 5th order; see Equation 1), such that points farther from the optical axis are displaced more than points near the optical axis. The inverse of this type of distortion, known as "barrel distortion", is applied to images during the rendering stage to make them appear undistorted when viewed through the lenses (a process known as pre-distortion) [5, 20, 23]. Therefore, the accuracy of the lens distortion model, and its associated coefficients, determine the effectiveness of this pre-distortion.

*e-mail: tongj86@yorku.ca
†e-mail:allison@eecs.yorku.ca
‡e-mail:lwilcox@yorku.ca

2 PREVIOUS WORK AND MOTIVATION

Although the technical aspects of lens distortion and its correction have received much attention [4–6, 8, 20, 23], less consideration has been paid to how such distortions affect a viewer's perception of space. There is some evidence that distortion correction has little impact on distance estimation, assessed via blind walking [17]. It is possible that the null result in [17] is due to the placement of the distance estimation target near the center of the field of view, where distortion is minimal. Given the non-linear nature of the distortion we would expect to see much greater impact on perceived shape for surfaces that extend across the field of view. As described above, little to no warping occurs in the center of the display but warping increases non-linearly toward the periphery [8]. Thus, pincushion distortion causes flat surfaces to appear curved, much like the inside surface of a bowl. Consistent with this pattern of curvature, recent experiments have shown that there are significant biases in local slant estimation of surfaces viewed in a modern HMD with uncorrected lens distortion [25], with larger biases in the top and bottom portions of the display relative to the center.

The observed slant biases in static scenes reflect the combined effects of the distortion on monocular (texture gradient) and binocular (binocular disparity) depth information. Pincushion lens distortion increasingly magnifies texture elements from the optical axis outwards, systematically causing them to project images that are larger and less densely spaced toward the periphery (Fig. 1, top-middle). This pattern of texture gradient is consistent with that created by a uniformly textured concave surface [2, 17]. Similarly, for crossed disparities typically encountered in VR HMDs pincushion distortion will give rise to disparity gradients consistent with concave curvature, that is with disparity increasing in magnitude toward the periphery (Fig. 1, top-left).

To this point we have focused on how monocular and binocular cues are affected by distortion in static scenarios. However, in typical use-cases content is rarely motionless; the addition of motion to the virtual scene should interact with the effects of optical distortion. For instance, the amount of distortion at given surface location depends on its radial distance from the center of the optical axis. Therefore, when a surface moves relative to the viewer, the amount of distortion at any given point on the surface changes. If observers can integrate local slant information from monocular texture over time in the presence of this changing distortion they might be less susceptible to bias.

It is possible that the presence of additional depth information will help observers discount the effects of the distortion field on rigid objects. One potentially useful source of information regarding 3D structure is that of relative motion. When a scene, or a slanted surface, translates relative to an observer, the direction of near points

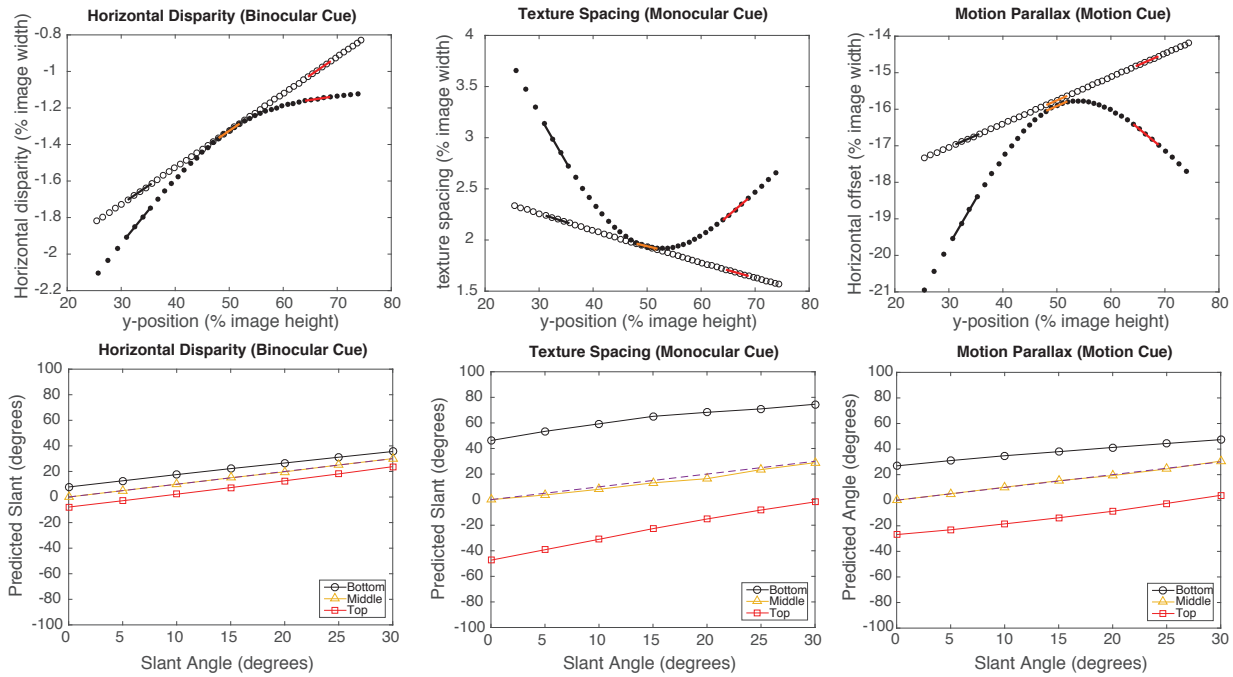

Figure 1: Top: Visual cues to surface slant derived from perspective projection of a surface mid-line, from left to right: Binocular Disparity, Monocular Texture Spacing and Horizontal offset as a function of vertical position in the image of a line slanted by 10°. Open points represent the undistorted image, filled points represent the pincushion distorted image. The colored lines represent the best fit slopes in the bottom (black), middle (orange) and top (red) positions along the mid-line. Bottom: The model's predicted local slant, at different positions along the mid-line (bottom, middle and top), from the Binocular cue, Monocular cue and Motion Parallax (left to right). The dashed line represents the veridical prediction.

on the surface changes faster than that of far points. This motion parallax cue can be used to infer the relative depth of objects in a scene or, for our purposes, points along a surface [9, 22, 26].

In the first of two experiments described here, we assess whether surface motion influences distortion-induced biases in local slant perception. Our results show that the errors in perceived slant resulting from pincushion distortion is the same for moving and static surfaces. Given that manufacturers do not typically generate individual pre-distortion correction equations for each HMD, and developers may apply incorrect pre-distortion, there is ample opportunity for insufficient distortion correction in these devices. To evaluate whether such partial distortion correction would result in similar or reduced slant biases, in our second experiment we tested observers on the same slant estimation task with moving surfaces and an intermediate level of distortion-correction condition. Our results confirm that errors in perceived slant vary directly with the degree of residual distortion highlighting the importance of accurate correction of lens distortion in VR HMDs.

3 Modelling the Consequences of Distortion

We used a computational model to predict the direction of biases in local slant perception resulting from the warping of several visual cues to slant: texture spacing, horizontal disparity and motion parallax (in moving surfaces). We computed the perspective projection of uniformly spaced discrete points along a vertical line (representing the surface mid-line), placed in front of the focal plane and slanted top-away from the observer. This was done from the perspectives of both the left and right eyes to generate stereoscopic image pairs. To model the effects of distortion, the images were then distorted with a level of pincushion distortion equivalent to a 1% change in the diagonal image size (eqn. 1, $k_1 = 1 \times 10^{-9}, k_2 = 0$). To model surface motion, the mid-line was initially placed 10 cm from the right of the mid-line then translated 10 cm to the left of the mid-line, and the perspective projection was computed for each line position. The image-coordinates of the points comprising the mid-line (either before or after distortion) were used to compute: 1) the inter-point distance (representing texture spacing), 2) the horizontal disparity between corresponding points in the left and right-eye images and 3) the horizontal displacement of points in a single eye's image resulting from the translation of the mid-line. The degree of change in these factors along the vertical axis of the image (slopes) are termed the monocular, binocular and motion parallax gradients, respectively, and they are the cues to surface slant we consider here. In general, the steeper the gradients the greater the perceived surface slant and, in the absence of distortion, a planar surface has linear gradients. In contrast, distortion introduces non-linearities (Fig. 1, top), i.e. surface curvature, which differentially biases local surface slant at different locations along the mid-line. To predict the direction and magnitude of local slant predicted by each cue, we computed the local gradients at three regions along the surface mid-line: near the top, near the bottom and in the middle (see Methods for details). In this scenario, all cues predict a pattern of local slant biases consistent with viewing a concave surface (Fig. 1, bottom): in the bottom portion of the image, surface slant is predicted to be overestimated (slanted further in the top-away direction). In the top portion of the image, surface slant is expected to be underestimated or even reversed (bottom-away). Lastly, in the middle portion, where distortion is mostly absent, perceived slant should be close to veridical.

4 GENERAL METHODS

4.1 Computational Model

Computational modeling was carried out in Matlab (2018b, MathWorks). We computed the perspective projection of a vertically oriented line positioned equidistant between a pair of off-axis stereo-cameras (FOV = 60°) with a separation of 0.062 m and focal plane at 1.5 m. The line was composed of 81 points spaced 0.025 m apart, and its center was set at a depth of 1.1 m from the mid-point between the cameras. The slant of the line was varied (0°-30° in 5° steps) along a rotational axis running through the center of the line and oriented perpendicular to the line-of-sight.

To determine the direction and magnitude of local slant predicted by each cue, we computed the local gradients at three regions along the surface mid-line within each camera's field of view: the "bottom region" was defined as the area spanning 30.55% to 36.11% to the top of the image, the "middle region" was defined as the area spanning 47.22% to 52.78% to the top of the image, and the "top region" was defined as the area spanning 63.89% to 69.44% to the top of the image. Within each of these portions of the image, the monocular, binocular and motion parallax gradients were linearly fit to determine best-fitting slopes. These, distorted, "local" gradient estimates were then converted to predicted slant by interpolating from the relationship between the non-distorted gradients and planar surface slant (determined numerically from simulations).

4.2 Apparatus

Psychophysical experiments were run on a Google Pixel 2 XL Android device (resolution: 2880×1440 pixels at 538 ppi) horizontally mounted in a Google Daydream View v2.0 (2017) VR HMD (FOV: 100°, lens-to-lens spacing: 64 mm). We selected this headset over more state-of-the-art headsets (e.g. HTC and Oculus) because the Google Daydream software permits customization of distortion correction parameters (see section 4.4 Software, image rendering and display). In these more sophisticated devices, however, distortion correction is automatically applied via the device drivers and cannot be readily overridden. Content was displayed stereoscopically (except for the monocular probe dots indicating the region of interest; see the section "Task"). The center of the display, dividing the left and right images, was aligned with the center of the HMD, which was equidistant from the lenses. Observers used a Google Daydream View 2.0 controller to register responses by pressing the clickable touch-pad. Observers' heads were stabilized in a chinrest during the experiment.

4.3 Task

In both experiments, observers completed a slant estimation task in which the local slant of a flat, textured surface was estimated at 3 different locations along the vertical meridian of the visual field: top, middle and bottom. In Experiment 1, the surface was either stationary or was moved side-to-side by the observer using the hand held controller. Images were either displayed without pre-distortion, such that the full extent of the lenses' pincushion distortion was visible, or with the level of barrel pre-distortion required to correct for the pincushion distortion as specified by the HMD manufacturer. In Experiment 2, observers moved the surface in the same way and three levels of distortion correction were assessed (none, intermediate and full correction). All experimental protocols were approved by the York University Research Ethics Board.

Each trial began with a green "fixation" dot ($\sim 2.5°$ wide) displayed in the center of the visual field until the observer clicked a hand controller button. Immediately following the offset of the fixation dot, a red "probe" dot ($\sim 2.5°$ wide) was displayed monocularly to the left eye for 1s in one of 3 possible locations along the vertical meridian: top ($\sim 18°$ above center), middle (at the center of the visual field) and bottom ($\sim 18°$ below center). Observers were told that they would be asked to judge the slant of the surface in the region indicated by the probe dot. The probe dot was presented monocularly so that its depth could not be determined and used as a reference to infer surface slant. Once the probe dot disappeared, a Voronoi-textured slanted surface, spanning the entire visual field, was displayed until the observer pressed a hand controller button. In the static condition, observers listened to a metronome ticking at 60 beats per minute (1Hz) while viewing the surface; they were instructed to press the hand controller button once they heard 4 clicks, ensuring an exposure time of approximately 4 seconds. In the motion condition, observers were instructed to sway the hand controller side-to-side (keeping their arm stationary while rotating their wrist) in synchrony with the metronome, which, in turn translated the surface in a periodic motion between end-points at ± 25 cm (the motion was constrained within this 50 cm range). They were instructed to move the surface to each end-point in synchrony with 4 clicks of the metronome (i.e. 4 seconds) to achieve ~ 2 cycles of lateral surface translation, after which they pressed the hand controller button and the surface disappeared. During the surface movement, the surface center was yolked to the controller. Following the offset of the surface, an adjustment disc appeared which could be rotated about the x-axis (horizontal axis) via rotation of the hand controller about the same axis; observers were asked to match the slant of the disc to the perceived local slant in the region of the surface initially specified by the red probe-dot. Observers registered their estimate and initiated the next trial by pressing the hand controller button. See Fig. 2 for an illustration of the events of a given trial.

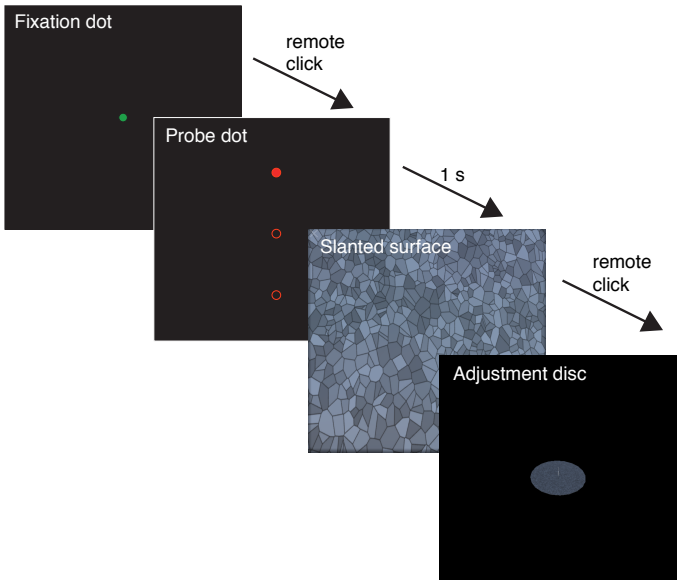

Figure 2: An illustration of the events in a trial: 1) the green fixation dot appears and, following a button press, disappears. 2) A red probe dot displayed monocularly to the left eye, in one of three possible positions appears for 1 s (top probe position is visible, middle and bottom positions are shown as open points for illustration purposes only). 3) the slanted surface appears (static or translated side-to-side by the observer) and after 4 beats of a 1 Hz metronome the observer presses a button to advance. 4) The adjustment disc appears, and the observer matches its slant to that of the surface at the probe dot location.

The motion and distortion conditions (see Methods specific to Experiments 1 and 2 for more details) were run in separate blocks. In a given block, each combination of the 7 surface slants (0-30° in 5° steps) and 3 probe locations was tested 5 times for a total of 105 trials. Each block's duration was approximately 15-20 minutes.

4.4 Software, image rendering and display

The experiment/virtual scene was developed as an Android application in Unity (2018.3.14) and C# with the Google VR SDK package (v1.200.1). Unity's default stereo-camera model was used to render images stereoscopically (60° FOV), with the inter-camera distance set to 62mm (which is close to the population average inter-pupillary distance). The cameras were equidistantly placed on either side of the origin of the world-coordinate space, and all other objects were placed relative to the origin. The fixation point (green disc) was placed at (x =0, y =0, z=1 m) so it appeared in the center of the visual field; the probes (red discs) were placed at a depth of 0.6 m and at the same horizontal position as the fixation point, but offset vertically (top: 0.2 m offset, middle: 0 m offset, and bottom: -0.2 m offset). The surface was a square plane (120×120 cm $= 14,400$ cm^2) with its center aligned with the origin point and located 1 m from the eyes in depth; slants were rendered by rotation of the surface along the x-axis running through this point. The plane was textured with a Voronoi pattern. The adjustment disc had a diameter of 0.07 m and appeared at the centre of the display area. A 0.06 m $\times$ 0.002 m line extended from the center of the disc, along the normal to the surface, provided an additional cue to the orientation of the disc.

The level of distortion correction (barrel distortion pre-warping) was manipulated using Google VR's viewer profile generator (https://wwgc.firebaseapp.com/) to set the distortion coefficients of the function that maps distorted to undistorted image points:

$$r = r_d + k_1 r_d^3 + k_2 r_d^5 \quad (1)$$

where r_d is the radial distance from the optical axis to a point in the distorted image and r is the corresponding point's radial distance from the optical axis in the original image. k_1 and k_2 are co-efficients that determine the direction and magnitude of distortion. To achieve the full pincushion distortion produced by the lenses, we set the coefficients to the lowest possible value allowed by the profile generator, $k_1 = k_2 = 0.0001$. We used Google VR's provided default viewer profile for the Daydream View v2.0 (2017) to achieve the appropriate (full) level of correction; the reported coefficients for this setting are $k_1 = 0.4331$ and $k_2 = -0.085$. This amounts to a barrel pre-distortion of $\sim 29\%$ (the diagonal radius of the pre-distorted image relative to the original image). For an intermediate level of distortion correction, we used coefficient settings for the Daydream View v1.0 (2016) which has lower magnification lenses than the second generation HMD; the reported coefficients for this setting are $k_1 = 0.01$ and $k_2 = 0.67$. This amounts to a barrel pre-distortion of $\sim 23\%$ (the diagonal length of the pre-distorted image relative to the original image). See Fig. 3 for screenshots showing each of the pre-distortion conditions.

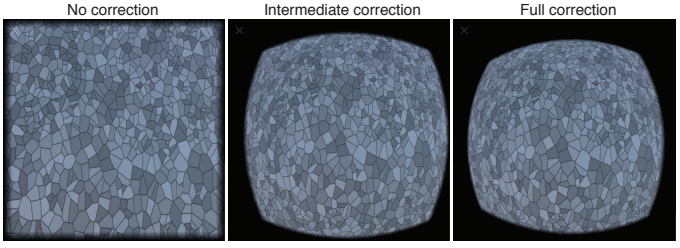

Figure 3: Different levels of distortion correction that we applied to images. No correction and Full correction in Experiment 1; No correction, Intermediate correction and Full correction in Experiment 2. Only the left eye image of the slanted surface is shown here.

5 Experiment 1

The aim of Experiment 1 was to assess the effects of surface motion and lens distortion on local slant estimation.

5.1 Methods

5.1.1 Conditions

Eight observers (3 males, 5 females; age range: 20-41, median age: 24) participated in 4 blocks, each comprising one of the combinations of 2 levels of surface motion (static and moving) by 2 levels of distortion (no correction and full correction). For each observer, the motion factor (static or moving surface) was consistent for the first 2 blocks and then changed to the other condition for the last 2 blocks, while the distortion factor (no correction or full correction) alternated between conditions from block-to-block. Therefore, there were 4 possible orderings of blocks; each was assigned randomly to pairs of observers (total of 8 observers). Each block consisted of 105 trials (see General Methods 4.3 Task), for a total of 420 trials.

5.1.2 Analysis

Linear regressions were performed for each observer's mean slant estimate as a function of the true slant angle. This was performed separately for each combination of distortion condition, motion condition and probe location. We then submitted the best-fit y-intercepts to a three-way ANOVA with probe position (top, middle and bottom), distortion condition (corrected and uncorrected) and motion (static vs. moving) as factors. To assess motion and distortion-induced biases in slant estimation, we tested for differences in intercept values (paired t-tests) between probe locations for each block (i.e. combination of motion and distortion conditions). All significance levels were subject to Holm-Bonferroni correction for multiple comparisons.

5.2 Experiment 1 Results

Irrespective of the motion condition (moving vs. static), estimates for distortion correction conditions were well fit by linear regression as indicated by a mean R^2 value of 0.910 ($SD = 0.07$) for the full correction/moving condition and a mean R^2 value of 0.907 ($SD = 0.11$) for the full correction/static condition. In contrast, in the absence of distortion correction the estimates were less directly related to true slant with lower R^2 values: for the no correction/moving condition, the mean R^2 value was 0.69 ($SD = 0.30$) and for the no correction/static condition the mean R^2 value was 0.66 ($SD = 0.34$). These differences in fits between corrected and non-corrected conditions could reflect the shallower slopes of observers' performance curves due to biases in the no correction condition driving estimates to ceiling and floor levels. Alternatively, the poorer fits could be due to increased variability in estimates in the presence of distortion. Fig. 4 shows the mean performance for all conditions.

To assess distortion-induced biases, we carried out a three-way ANOVA on the best-fit y-intercept with probe position (top, middle and bottom), distortion condition (corrected and uncorrected), and motion (static vs. moving) as factors. The ANOVA on best-fit y-intercepts revealed a significant effect of probe location ($F_{2,14} = 11.66, p = 0.001, \eta_p^2 = 0.6248$) and a significant distortion by probe location interaction ($F_{2,14} = 11.80, p = 0.0001, \eta_p^2 = 0.6287$).

Given the significant distortion by probe interaction, and the lack of a motion effect, we carried out a set of paired t-tests comparing the y-intercepts between probe locations for each block (combinations of motion and distortion conditions). In the no correction condition, all pairwise differences were significant after Holm-Bonferroni corrections: Bottom vs. Middle, $p = 0.014$ and 0.014 for moving and static surfaces, respectively; Bottom vs. Top, $p = 0.033$ moving and 0.014 static; and Middle vs. Top, $p = 0.043$ moving and 0.014 static. In the full correction condition none of the pairwise differences were significant after Holm-Bonferroni corrections: Bottom vs. Middle, $p = 0.71$ and 0.95 for moving and static surfaces, respectively; Bottom vs. Top, $p = 0.92$ moving and 0.95 static; and Middle vs. Top, $p = 0.92$ moving and 0.18 static. In summary, when distortion was

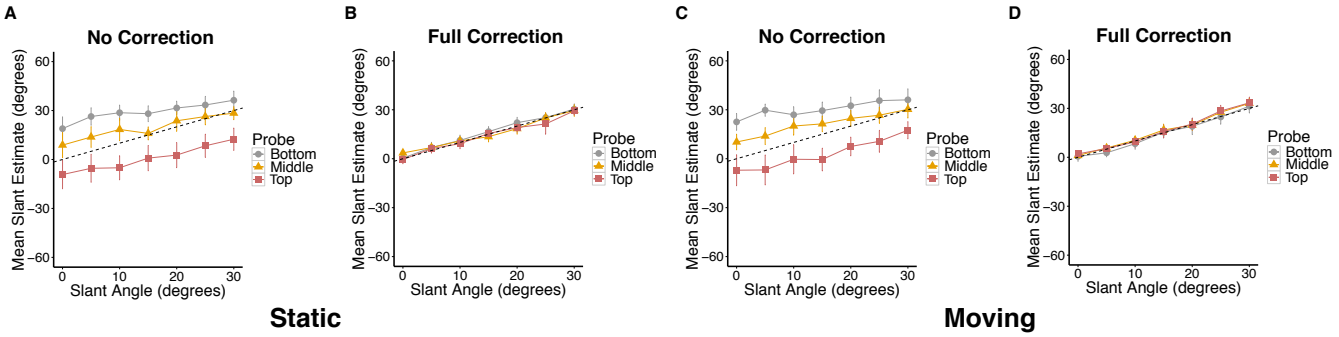

Figure 4: Mean performance in local slant estimation task for Experiment 1 (error bars = $\pm SEM$) with A) no distortion correction and static surfaces, B) full distortion correction and static surfaces, C) no distortion correction and moving surfaces, D) full distortion correction and moving surfaces. In each case the dashed line represents ideal performance.

appropriately corrected, there were no significant pairwise differences in y-intercepts between the probe locations. Conversely, when distortion was fully present, there were significant differences in y-intercepts between every pairwise comparison of probe locations (see Fig. 5).

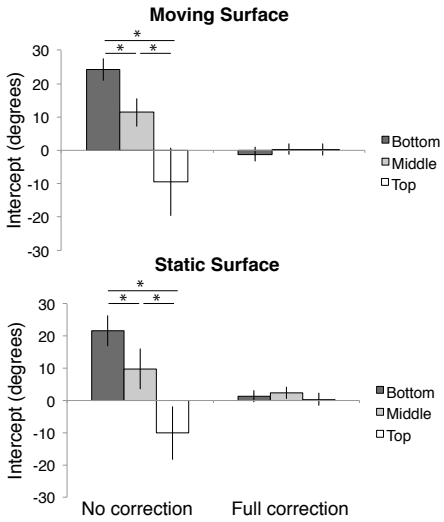

Figure 5: The mean best-fit y-intercepts (error bars = $\pm SEM$) plotted for each motion and distortion condition in Experiment 1. Asterisks signify significant pairwise differences (p <0.05).

6 Experiment 2

The purpose of Experiment 2 was to test the effects of different levels of distortion on the local slant estimation of moving surfaces.

6.1 Methods

6.1.1 Conditions

Six observers (2 males, 4 females; age range: 24-35, median age: 29) two of whom completed Experiment 1, participated in this study. The 3 distortion levels (full correction, intermediate correction and no correction) were blocked and each observer received a different order. Each block consisted of 105 trials (see General Methods 4.3 Task) for a total of 315 trials. In all conditions the surfaces were in motion, as described in Experiment 1.

6.1.2 Analysis

Linear regressions were performed for each observer's mean estimate of slant as a function of the true slant angle. This was done for each distortion condition and probe location. We then submitted the y-intercepts to a two-way ANOVA with probe position (top, middle and bottom) and distortion condition (fully corrected, partially corrected and uncorrected) as factors. To test for significant biases in local slant estimation, we carried out a set of paired t-tests comparing the y-intercepts between probe locations for each of the 3 distortion conditions. Again, all significance levels were subject to Holm-Bonferroni correction.

6.2 Experiment 2 Results

Local slant estimation with full distortion correction gave the best linear fits, with a mean R^2 value of 0.93 ($SD = 0.06$). Intermediate distortion correction also resulted in good linear fits, with a mean R^2 value of 0.90 ($SD = 0.08$). Linear fits were the weakest in the "no correction" condition, with a mean R^2 value of 0.70 ($SD = 0.24$). These differences in goodness of fit between the different distortion conditions likely reflect shallower slopes and increasing response variability, specifically when the probe is not centred. See Fig. 6 for plots of average performance in each distortion condition.

To test for distortion-induced biasing of local slant estimation, we submitted the best-fit intercepts to a two-way ANOVA with probe position (top, middle and bottom) and distortion condition (full correction, intermediate correction and no correction) as factors. The ANOVA revealed a significant effect of probe location ($F_{2,10} = 46.4, p = 8.7 \times 10^{-6}, \eta_p^2 = 0.9028$) and a significant distortion by probe location interaction ($F_{4,20} = 16.7, p = 3.6 \times 10^{-6}, \eta_p^2 = 0.7702$).

Given the highly significant distortion by probe interaction, we carried out a set of paired t-tests (corrected for multiple comparisons) comparing the y-intercepts between probe locations for each of the 3 distortion conditions. In the full correction condition, none of the pairwise comparisons were significantly different, with Holm-Bonferroni corrected p-values ranging from 0.079 (Middle vs. Top) to 0.71 (Middle vs. Bottom). In the intermediate correction condition, all of the pairwise differences were significant, with Holm-Bonferroni corrected p-values ranging from 0.0082 (Bottom vs. Top) to 0.015 (Bottom vs. Middle & Middle vs. Top). Lastly, in the no correction condition, all of the pairwise differences were significant, with Holm-Bonferroni corrected p-values ranging from 0.0008 (Bottom vs. Middle) to 0.0127 (Middle vs. Top). In summary, no significant differences in best-fit y-intercepts between probe locations were found in the fully corrected condition. Conversely, both intermediate and no correction conditions resulted in significant differences in y-intercepts for all pairwise comparisons between

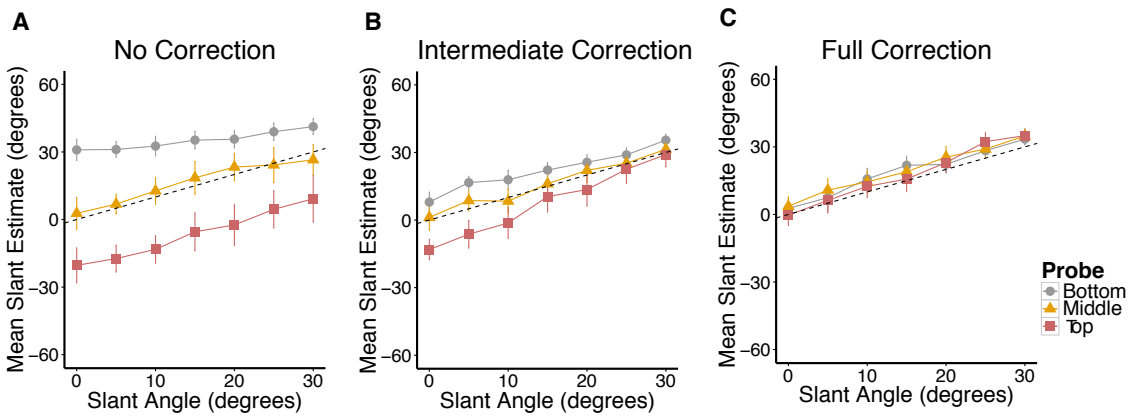

Figure 6: Mean local slant estimates for each distortion condition in Experiment 2 (error bars = $\pm SEM$): A) No correction, B) intermediate correction and C) full correction. Dotted line represents ideal performance.

probe locations (see Fig. 7). The y-intercepts estimate the amount of apparent slant in an objectively frontal (0º) surface and it is clear that significant distortion is perceived in the uncorrected and partially corrected cases.

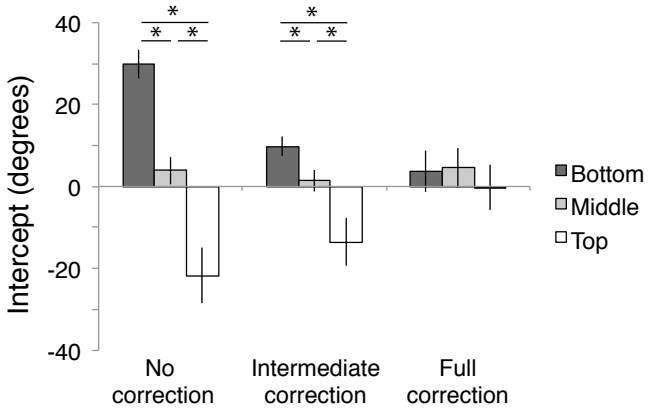

Figure 7: Mean best-fit y-intercepts plotted for each distortion condition in Experiment 2 (error bars = $\pm SEM$). Asterisks represent significant pairwise differences.

7 Discussion

We have shown that pincushion distortion induces biases in local surface slant estimates, irrespective of whether the surfaces are static or moving. The pattern of biases is consistent with a distortion-induced perception of concave curvature in planar surfaces. That is, slant of surfaces objectively slanted top away is overestimated in the bottom portion of the display, while in the top of the display, perceived slant is reversed or underestimated. When distortion is accurately compensated for, slant estimation is veridical for both moving and static surfaces. Furthermore, we have shown that an intermediate level of lens distortion resulting from insufficient pre-distortion correction produces the same, but diminished, pattern of biases.

The dominance of vision over the kinesthetic senses might partially explain why we did not find an effect of motion parallax on perceived surface shape. That is, observers controlled the movement of the surface via a hand-held controller. Therefore, the interpretation of motion parallax was linked to the observer's own sense of kinesthetic motion in virtual space, which in turn may have been dominated by vision [13, 14].

We anticipated that in Experiment 1 participants would see textures along the surface warping as they moved through the visual field. This would make it clear that the distortion was linked to the display (and not to the rendered surface). A potential consequence of this would be to enable perceptual dissociation of the surface slant and the distortion, resulting in a more veridical percept i.e. less curvature. However, this appears not to be the case in that perceived surface curvature induced by distortion was just as salient in moving surfaces as in static surfaces at all distortion levels. Thus it appears that the residual distortion cannot be segregated and discounted when judging the attitude of surfaces in virtual environments. This has potential implications for situations in which users must interact with surfaces in virtual environments particularly within working space, for example in training environments or gaming. If distortion correction is insufficient, as in the intermediate distortion condition in Experiment 2, residual distortion could lead to marked errors in performance. Therefore, inverse distortion models and their coefficients must be determined accurately, preferably using objective calibration techniques, for specific combinations of lenses and viewing parameters in a given HMD.

We used the Google Daydream (2017) headset to capitalize on the ease with which the level of distortion correction can be adjusted for this device. In more advanced devices such as the HTC Vive, or Oculus Rift, distortion correction is automatically applied via device drivers and is difficult to override. Furthermore, for the simple scene used in these experiments advanced features of such state-of-the-art headsets would not offer a substantial advantage. Nevertheless, our results are applicable to most VR headsets with high-magnification optics that induce pincushion type distortion, especially if distortion correction is not appropriately applied. A previous study showed that as little as 5% of residual distortion has perceptual consequences in surface slant discrimination [25].

Recent research suggests that precise distortion correction should take into account the direction of gaze relative to the optical axis of the HMD [12, 19]. We instructed observers to maintain fixation at the location of an initial fixation dot (see fig. 2), but it is possible that eye movements were made on occasion. However, from post-test observer reports, and our observations during pilot testing, any change in the distortion field when gaze direction was shifted was imperceptible (irrespective of distortion condition). This observation is supported by our results, which show that with full-distortion correction, there is no difference in perceived slant across the tested locations along the midline (slant was veridically perceived). If

observer gaze shifts caused systematic changes in the distortion field in the current experiments, then we should have seen systematic biases, even with distortion correction.

Here, we have focused on the impact of pincushion distortion on the perception of surface shape, as this is the most common distortion induced by the magnifying lenses in HMDs. However, in a previous study we reported that barrel distortion also has negative impacts on the perception of slant: global surface slant is underestimated in the presence of barrel distortion, although to a lesser degree compared to pincushion distortion, and overall slant perception is less precise in the presence of either type of distortion [25]. Planar surfaces with barrel distortion appear convex, rather than concave. Therefore, with barrel distortion, the biases in local slant perception would be in the direction opposite of what we measured here: overestimation of slant near the top of the display, underestimation of slant near the bottom of the display and near veridical slant perception near the middle. Barrel distortion is likely to be encountered if the lens distortion in HMDs is over-corrected. Furthermore, barrel distortion can also be induced by fish-eye lenses typically used in the wide-angle cameras of video see-through Augmented Reality (AR) devices [3, 18]. The images displayed in these type of AR headsets must be corrected by the application of pincushion distortion. Again, our results show that insufficient or improper correction will likely lead to misperception of slant in the physical scene, making interaction with the real world (i.e. walking, navigating and placing objects on surfaces) inaccurate and potentially dangerous.

8 Conclusion

We have demonstrated that local surface orientation is biased in the presence of lens distortion in typical HMDs. This perceptual bias is driven by the curvature of visual space due to non-uniform lens magnification. The resultant surface curvature and local slant biases remain significant when imagery is in motion relative to the viewer. The degree of slant bias varies directly with the amount of residual distortion and can be mitigated by careful application of inverse distortion.

Acknowledgment

This work was funded by an NSERC Collaborative Research and Development (CRD) grant, in partnership with Qualcomm Canada Inc, and the Canada First Research Excellence Fund (CFREF) for the Vision: Science to Application (VISTA) program.

References

[1] R. S. Allison, L. R. Harris, M. Jenkin, U. Jasiobedzka, and J. E. Zacher. Tolerance of temporal delay in virtual environments. In *Proceedings IEEE Virtual Reality 2001*, pages 247–254, Mar. 2001.

[2] R. S. Allison and I. P. Howard. Temporal dependencies in resolving monocular and binocular cue conflict in slant perception. *Vision Research*, 40(14):1869–1885, 2000.

[3] M. Bajura and U. Neumann. Dynamic registration correction in video-based augmented reality systems. *IEEE Computer Graphics and Applications*, 15(5):52–60, 1995.

[4] D. C. Brown. Decentering distortion of lenses. *Photogrammetric Engineering and Remote Sensing*, 32:444–462, 1966.

[5] P. Drap and J. Lefèvre. An Exact Formula for Calculating Inverse Radial Lens Distortions. *Sensors*, 16(6):807, 2016.

[6] A. Fitzgibbon. Simultaneous linear estimation of multiple view geometry and lens distortion. In *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. CVPR 2001*, volume 1, pages I–I, Dec. 2001. ISSN: 1063-6919.

[7] L. R. Harris, M. R. Jenkin, D. Zikovitz, F. Redlick, P. Jaekl, U. T. Jasiobedzka, H. L. Jenkin, and R. Allison. Simulating self-motion I: cues for the perception of motion. *Virtual Reality*, 6(2):75–85, 2002.

[8] E. Hecht. *Optics*. Addison-Wesley, 3rd edition, 1998.

[9] J. M. Hillis, S. J. Watt, M. S. Landy, and M. S. Banks. Slant from texture and disparity cues: Optimal cue combination. *Journal of Vision*, 4(12), 2004.

[10] A. D. Hwang and E. Peli. Stereoscopic three-dimensional optic flow distortions caused by mismatches between image acquisition and display parameters. *Journal of Imaging Science and Technology*, 63(6):60412–1, 2019.

[11] O. Janeh, N. Katzakis, J. Tong, and F. Steinicke. Infinity Walk in VR: Effects of Cognitive Load on Velocity during Continuous Long-Distance Walking. In *ACM Symposium on Applied Perception 2019*, SAP '19, pages 1–9, Barcelona, Spain, Sept. 2019.

[12] J. A. Jones, L. C. Dukes, D. M. Krum, M. T. Bolas, and L. F. Hodges. Correction of geometric distortions and the impact of eye position in virtual reality displays. In *2015 International Conference on Collaboration Technologies and Systems (CTS)*, pages 77–83. IEEE, 2015.

[13] L. Kohli. Redirected touching: Warping space to remap passive haptics. In *2010 IEEE Symposium on 3D User Interfaces (3DUI)*, pages 129–130. IEEE, 2010.

[14] L. Kohli, M. C. Whitton, and F. P. Brooks. Redirected touching: The effect of warping space on task performance. In *2012 IEEE Symposium on 3D User Interfaces (3DUI)*, pages 105–112. IEEE, 2012.

[15] R. Konrad, A. Angelopoulos, and G. Wetzstein. Gaze-contingent ocular parallax rendering for virtual reality. *ACM Transactions on Graphics (TOG)*, 39(2):1–12, 2020.

[16] G. A. Koulieris, K. Akşit, M. Stengel, R. K. Mantiuk, K. Mania, and C. Richardt. Near-eye display and tracking technologies for virtual and augmented reality. In *Computer Graphics Forum*, volume 38, pages 493–519. Wiley Online Library, 2019.

[17] S. A. Kuhl, W. B. Thompson, and S. H. Creem-Regehr. HMD Calibration and Its Effects on Distance Judgments. *ACM Trans. Appl. Percept.*, 6(3):19:1–19:20, 2009.

[18] M. Lee, H. Kim, and J. Paik. Correction of Barrel Distortion in Fisheye Lens Images Using Image-Based Estimation of Distortion Parameters. *IEEE Access*, 7:45723–45733, 2019.

[19] J. Martschinke, J. Martschinke, M. Stamminger, and F. Bauer. Gaze-dependent distortion correction for thick lenses in hmds. In *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pages 1848–1851. IEEE, 2019.

[20] D. Pohl, G. S. Johnson, and T. Bolkart. Improved pre-warping for wide angle, head mounted displays. In *Proceedings of the 19th ACM Symposium on Virtual Reality Software and Technology - VRST '13*, page 259, 2013.

[21] W. Robinett and J. P. Rolland. A Computational Model for the Stereoscopic Optics of a Head-Mounted Display. *Presence: Teleoperators and Virtual Environments*, 1(1):45–62, 1992.

[22] B. J. Rogers and T. S. Collett. The appearance of surfaces specified by motion parallax and binocular disparity. *The Quarterly Journal of Experimental Psychology Section A*, 41(4):697–717, 1989.

[23] J. P. Rolland and T. Hopkins. A method of computational correction for optical distortion in head-mounted displays. Technical Report TR 1993/45, University of North Carolina at Chapel Hill. Department of Computer Science, 1993.

[24] A. Thaler, S. Pujades, J. K. Stefanucci, S. H. Creem-Regehr, J. Tesch, M. J. Black, and B. J. Mohler. The Influence of Visual Perspective on Body Size Estimation in Immersive Virtual Reality. In *ACM Symposium on Applied Perception 2019*, SAP '19, pages 1–12, Barcelona, Spain, Sept. 2019.

[25] J. Tong, R. S. Allison, and L. M. Wilcox. The Impact of Radial Distortions in VR Headsets on Perceived Surface Slant. *Journal of Imaging Science and Technology*, 63(6):60409.1–11, 2019.

[26] X. M. Wang, A. Thaler, S. Eftekharifar, A. O. Bebko, and N. F. Troje. Perceptual distortions between windows and screens: Stereopsis predicts motion parallax. In *2020 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*, pages 685–686. IEEE, 2020.

[27] L. Wilcox, R. Allison, S. Elfassy, and C. Grelik. Personal space in virtual reality. *ACM Transactions on Applied Perception (TAP)*, 3(4):412–418, 2006.