1. Introduction

Human visual perception of motion has been a central research topic in cognitive psychology and neuroscience over a century. Most previous research examining human motion processing has been undertaken with the head and the body stable in space. While maintaining head and body stability can be achieved in controlled laboratory settings, under natural viewing conditions, the head and the body are frequently in motion. A consequence of this “self-motion” is that retinal image motion is confounded, and must be considered as a product of both object motion and self-motion. For successful environmental interactions, the visual system must resolve the problem of distinguishing between these two contributions to retinal image motion.

To date, the manner in which object motion influences the perception of self-motion has been systematically examined (e.g. [1–3]), and neurophysiological models have been proposed to account for the computation of self-motion extraction in the presence of object motion (e.g. [4,5]). However, not many studies have examined the converse situation of how self-motion affects the judgment of object motion (see [6]). Although it has recently been reported that global processing of optic flow plays a fundamental role in the recovery of object motion during self-motion (e.g. [7]), it remains in question how accurately the visual system removes the flow generated during body or head movements from the retinal motion to recover object motion.

Several studies that examined self-motion in the real world or in a virtual environment found that concurrent physical body movements impair visual perception of related three-dimensional (3D) motion [8–12]. These findings support Barlow’s [13] proposal that highly correlated events (such as walking and an expanding optic flow pattern) mutually specify each other. The perceptual system uses this correlation to inhibit connections between simultaneously active neural units to enhance neural responses for novel stimuli in the world (see also [14]). Durgin et al. [10] further showed that the operation underlying the inhibitory connections is reflected by a bias in response rather than a change in sensitivity.

1.1. Head-referenced object motion perception during head movements

For object motion perception during head movements, Probst and Wist [15] presented the motion of a spot of light to observers via a head mounted display (HMD) and found increased reaction...
times for detecting such head-referenced object motion during head rotation. Recently, Li et al. [16] presented observers with a horizontally oscillating checkerboard image motion via an HMD and measured their judgment of the head-referenced image motion magnitude during concurrent horizontal (yaw) or vertical (pitch) head movements. They found that both yaw and pitch head movements reduced observers’ perceptual sensitivity to head-referenced image motion compared with when the head was still. Furthermore, compared with when the head was still, both yaw and pitch head movements also caused an overall underestimation of image motion by 10%. Taken together, these findings indicate that concurrent head movements suppress head-referenced object motion perception, similar to the reported suppression effect of physical translation on related 3D motion perception. However, different from what was reported by Durgin et al. [10], in the study by Li et al., the suppression of head-referenced object motion perception during concurrent head movements entailed both a sensitivity reduction as well as an underestimation response bias.

1.2. World-referenced object motion and the field of view

In most natural situations, head turns trigger the vestibulo-ocular reflex (VOR) which initiates compensatory pursuit eye movements within 10 ms to maintain the fixation on the object of interest [17]. The extra-retinal signals of pursuit eye movements (such as the efference copy) are used by the brain to tell which part of the retinal motion is due to eye movements and which is due to object motion in the world (e.g. [18–20]). However, for the studies that presented object motion stimuli via an HMD to observers [15,16], the stimulus motion was head-referenced requiring no compensatory eye movements to keep the gaze on the stimuli during head turns. This technique has been neurophysiologically shown effective in suppressing VOR-triggered compensatory eye movements in squirrel monkeys [21]. Thus, the suppression of object motion perception during concurrent head movements reported by the HMD studies supports the proposal of a sensory object motion inhibition process using primarily vestibular and proprioceptive signals generated during head turns (e.g. [22]). It still remains in question whether concurrent head movements can have a similar suppression effect on world-referenced object motion when VOR-triggered compensatory eye movements are made.

Furthermore, previous studies that presented object motion stimuli via an HMD prevented observers from seeing optic flow generated by their own head movements. As the visual system normally removes the global flow experienced during self-motion from the total retinal motion to recover object motion (e.g. [7]), it is possible that the reported suppression effect of head movements on image motion perception from the HMD studies is due to a confusion of retinal motion components during head turns. Allowing observers to view the flow generated by their own head movements thus might help accurately perceive image motion during concurrent head movements through flow parsing. Given that a larger field of view (FOV) provides more global flow information especially about self-rotation (e.g. [23,24]), it might lead to more accurate perception of self-generated flow and thus more accurate object motion recovery during concurrent head turns.

1.3. The current study

In the current study, we examined the effect of the FOV on perceiving world-referenced image motion during concurrent head movements. The goal was to extend the earlier findings from the study by Li et al. [16] to further our understanding of the underlying mechanism responsible for the inhibitory effect of concurrent head movements on object motion perception.

Specifically, instead of presenting the visual stimuli to observers via an HMD, we presented horizontally oscillating checkerboard image motion (Fig. 1) on a large rear-projected screen that was physically fixed in the world. Observers sat in front of the screen and viewed the image motion while making yaw or pitch head movements, or holding their head still. They were then asked to rate the magnitude of the image motion. As the checkerboard image motion was world-referenced, observers’ head movements would trigger VOR-induced compensatory eye movements to keep their fixation on the checkerboard image during head turns. This allowed us to examine the contribution of compensatory eye movements to image motion perception during concurrent head movements. We tested a large (73°H × 73°V) and a small (25°H × 25°V) FOV for the checkerboard image motion to examine whether providing global flow would help the visual system parse out retinal motion due to head rotation for accurate image motion perception. If so, participants should estimate the magnitude of the checkerboard motion more accurately during concurrent head movements with the large than the small FOV display.

2. Methods

2.1. Participants

Twelve participants (three males, nine females) aged between 20 and 39 viewed the large FOV display and 10 participants (five males, five females) aged between 21 and 30 viewed the small FOV display. All participants were students at the University of Hong Kong (HKU) and provided informed consent in accordance with guidelines from HKU Human Research Ethics Committee. All had normal or corrected-to-normal vision and were naïve to the specific purpose of the study.

2.2. Visual stimuli

A red checkerboard image pattern was generated using a Dell Precision Workstation 670n with an NVIDIA Quadro FX 1800 graphics card, and was then rear-projected on a large screen (110°H × 94°V) with an Epson EMP-9300 LCD projector (native resolution: 1400 × 1050 pixels) at the frame rate of 60 Hz. The image pattern was presented at the center of the large screen which was calibrated to correspond to the cyclopean eye of the participant at the beginning of the experiment. As in the study by Li et al. [16], the bright red squares in the checkerboard pattern were set to 8-bit RGB triples of (100, 0, 0), and the dark red squares were (50, 0, 0). The Michelson luminance contrast between the two square colors was 58%. The center of the checkerboard pattern had a central cross composed of five dark red squares. Each square in the checkerboard pattern subtended a visual angle of 6.8° on a side. Two FOV sizes of the checkerboard image were tested in the experiment: the large FOV subtended a visual angle of 73°H × 73°V (Fig. 1a) and the small FOV subtended a visual angle of 25°H × 25°V (Fig. 1b). For both FOV sizes, the checkerboard image was Gaussian masked (σ = 18.3° and 6.3° for the large and small FOVs, respectively) to obscure its sharp edges on the screen and the related relative motion cues.

In each trial, the checkerboard image oscillated sinusoidally from side to side at 0.6 Hz at one of seven amplitudes: 0°, 0.94°, 1.88°, 2.82°, 3.76°, 4.70°, and 5.64° (one-half of peak-to-peak displacement). The displacement amplitudes were the same as in the study by Li et al. [16], which resembled image motion errors on an HMD observed by users during head turns when the system had seven equally spaced head-tracking latencies ranging from 0 to 120 ms [25,26]. The use of such sinusoidal oscillatory motion thus

Please cite this article in press as: K.L. Yeung, L. Li, Effect of the field of view on perceiving world-referenced image motion during concurrent head movements, Displays (2012), http://dx.doi.org/10.1016/j.displa.2012.08.003
has a direct relevance to the head-tracked virtual environment system design.

Participants sat in a chair at 0.56 m away from the screen and viewed the stimuli monocularly with their dominant eye in a light excluding booth. The screen edges were covered in matte black cloth to minimize their visibility and any relative motion of the checkerboard image with respect to the screen edges. Participants’ head position was tracked by an InterSense InertiaCube3 sensor mounted on a helmet that they were wearing (Fig. 2). Participants were instructed to fixate the center of the cross in the checkerboard image pattern and to estimate the magnitude of the checkerboard image motion while rotating their head from side to side (yaw) or up and down (pitch), or holding their head still. As the image motion was world-referenced, fixating the center of the cross while making head movements elicited VOR-triggered compensatory eye movements to maintain gaze on the cross. Both yaw and pitch head movements were paced by the beep of a computer-generated metronome at 0.5 Hz. The head movement frequency was chosen to be slightly different from the image oscillation frequency (0.6 Hz) to prevent participants from synchronizing their head movements with the image motion. Head movements started from the left (yaw) or the bottom (pitch). An alarm based on the head angle readings of the InterSense sensor reminded participants to limit their yaw head movements to 15° relative to the straight ahead, and their pitch head movements to 10° above and 20° below the eye level, respectively. The use of an asymmetric range for pitch head movements was due to the greater ease of making downward than upward head motion. Fig. 3 shows sample raw and pitch head movement data from a participant with the large FOV display, indicating that participants in general could follow the instructions and make accurate head movements as required.

2.3. Procedure

For the trials in which participants made head movements, the checkerboard image motion was presented 1 s after they started yaw or pitch head movements and lasted for 5 s until the end of the 6-s head movement interval. For the trials in which the head was still, the image motion was presented for 5 s. The screen was otherwise black. The phase of the oscillating checkerboard image motion was randomized in each trial. At the end of each trial, participants were asked to verbally report their perceived magnitude of the checkerboard image motion as a percentage relative to standards shown at the beginning of each experimental condition. The experimenter, who was generally blind to the specific experimental condition, recorded the response and proceeded to the next trial.

The experiment was composed of four head movement condition blocks, with each block containing 35 randomized trials (five trials × seven motion amplitudes). Both groups of participants tested with the large and the small FOV display completed the head still condition first to get the baseline judgment scale. They then completed the yaw and pitch head movement conditions in a counterbalanced order. Participants completed a final head still condition to examine whether there was any shift in their judgment criteria. To anchor their rating responses, participants were shown 0°, 2.82°, and 5.64° checkerboard motion amplitudes while holding their head still for at least three times at the beginning of each head movement condition. They were told that these motion
amplitudes corresponded to 0%, 50%, and 100% of the maximum checkerboard image motion, and they could view them as many times as needed prior to testing. Each head movement condition block lasted about 10 min, and participants completed the experiment within 1 h.

2.4. Data analysis

As in the study by Li et al. [16], we performed a linear regression analysis of estimated motion magnitude ($\Psi_e$) against actual motion magnitude ($\Psi_b$) to measure the effect of concurrent head movements on image motion perception. The linear regression equation is given as:

$$\Psi_e = K\Psi_b + A,$$

where $K$ is the slope indicating the proportional change of estimated motion magnitude for given input motion stimuli, and $A$ is the intercept representing an overall estimation bias. In other words, a decrease in $K$ indicates a reduction in perceptual sensitivity to the change in the presented image motion, and a decrease in $A$ indicates an overall underestimation of the input motion stimuli, which is independent of the change in $K$.

Specifically, we first converted input image motion amplitudes to percentages with 100% corresponding to the largest motion amplitude (5.64%). After computing the median of the five motion estimates that participants gave in the five trials tested for each input motion amplitude, we regressed the medians of the motion estimates against the six non-zero input image motion amplitudes for each participant for each head movement condition. We analyzed the zero input motion amplitude separately for an accurate assessment of the intercept in the linear regression analysis. Fig. 4 shows sample motion estimation data from the yaw head movement condition with the large FOV display along with the regression line. A regression with unity slope and zero intercept corresponds to perfect estimation. To compare how perceptual sensitivity ($K$) and estimation bias ($A$) changed across the different head movement conditions and the FOV sizes, we conducted separate mix-design ANOVAs on the regression slopes and intercepts.

3. Results

All 48 linear regressions (four head movement conditions × 12 participants) for participants tested with the large FOV display and all 40 linear regressions (four head movement conditions × 10 participants) for participants tested with the small FOV display were statistically significant ($r \geq 0.47, p < 0.05$). During yaw and pitch head movements, five of the 24 medians of motion estimates for zero input motion amplitude with the large FOV display had non-zero magnitude ranging from 1% to 20%, and four of the 20 medians for zero input motion amplitude with the small FOV had non-zero motion magnitude ranging from 5% to 30%. Overall, this indicates that regardless of the FOV size, participants could correctly perceive image stability during head movements as reported by previous studies (e.g. [27]).

3.1. Slope

For both participant groups tested with the large and the small FOV display, a paired $t$-test showed that the mean slope for the “before” head still condition (mean ± SE: 0.89 ± 0.06 and 0.96 ± 0.06 for the large and the small FOV, respectively) was not significantly different from that for the “after” head still condition (0.96 ± 0.06 and 0.92 ± 0.08, $t(11) = -1.46$, $p = 0.17$ and $t(9) = 0.54$, $p = 0.60$, respectively). We thus averaged the slope data from the “before” and “after” head still conditions to generate one slope measure of the combined head still condition. The mean slope averaged across participants is plotted against head movement amplitude in Fig. 5a. For both FOV sizes, the mean slope of the combined head still condition was not significantly different from unity (0.92 ± 0.06 and 0.94 ± 0.06, $t(11) = -1.29$, $p = 0.22$ and $t(9) = -0.95$, $p = 0.37$ for the large and the small FOV, respectively), indicating that their perceptual sensitivity to image motion was veridical when their head was stationary.

Please cite this article in press as: K.L. Yeung, L. Li, Effect of the field of view on perceiving world-referenced image motion during concurrent head movements, Displays (2012), http://dx.doi.org/10.1016/j.displa.2012.08.003
To examine the effect of head movement and the FOV on the regression slopes, we conducted a 3 (head movement condition) × 2 (FOV size) mixed-design ANOVA and found that only the main effect of head movement condition was significant ($F(2, 40) = 8.28$, $p < 0.001$). The main effect of FOV size and the interaction effect of head movement condition and FOV size were not significant ($F(1, 20) = 0.24$, $p = 0.63$ and $F(2, 40) = 1.01$, $p = 0.37$, respectively). Tukey HSD tests showed that across the two FOV sizes, the mean slope of the combined head still condition (0.93) was significantly larger than those of the yaw and pitch head movement conditions (0.72 and 0.78, $p < 0.01$ and $p = 0.05$, respectively), and the latter two were not significantly different from each other ($p = 0.54$). This indicates that both yaw and pitch head movements reduced participants’ perceptual sensitivity to image motion, and this reduction in sensitivity was independent of the display FOV size.

### 3.2. Intercept

For both the large and small FOV participant groups, a paired t-test showed that the mean intercept for the “before” head still condition (mean ± SE: $-2.6\% ± 3.3\%$ and $-0.7\% ± 4.2\%$ for the large and the small FOV, respectively) was not significantly different from that for the “after” head still condition ($-2.1\% ± 2.3\%$ and $-5\% ± 1.9\%$). $t(11) = -0.25$, $p = 0.81$ and $t(9) = 0.95$, $p = 0.37$, respectively. We thus averaged the intercept data from the “before” and “after” head still conditions to generate one intercept measure of the combined head still condition. The mean intercept averaged across participants is plotted against head movement condition for the two FOV sizes in Fig. 5b. For both FOV sizes, the mean intercept of the combined head still condition was not significantly different from zero ($-2.3\% ± 2.6\%$ and $-2.8\% ± 2.4\%$, $t(11) = -0.91$, $p = 0.38$ and $t(9) = -1.18$, $p = 0.27$ for the large and the small FOV, respectively), indicating that participants did not show a significant motion estimation bias when their head was still.

A 3 (head movement condition) × 2 (FOV size) mixed-design ANOVA on the regression intercepts showed that only the main effect of FOV size was significant ($F(1, 20) = 5.71$, $p < 0.05$). The main effect of head movement condition and the interaction effect of head movement condition and FOV size were not significant ($F(2, 40) = 0.66$, $p = 0.53$ and $F(2, 40) = 2.16$, $p = 0.13$, respectively). Tukey HSD tests showed that across the three head movement conditions, the mean intercept for the small FOV group ($-7.1\%$) was significantly different from that for the large FOV group ($-0.64\%$, $p < 0.05$), and a separate t-test showed that the latter was not significantly different from zero ($t(35) = -0.43$, $p = 0.67$). This indicates that the reduction of the FOV from 73°H × 73°V to 25°H × 25°V caused participants to have an overall underestimation of image motion by about 7% across the head movement and head still conditions.

### 4. Discussion

The slope and intercept data from the current experiment show that both concurrent yaw and pitch head movements reduced participants’ perceptual sensitivity to world-referenced image motion by about 20% compared to the veridical levels when the head was still, but neither yaw nor pitch head movements caused any motion estimation bias. These results are comparable with those at a similar image oscillation and head movement frequency (0.5 Hz) of the study by Li et al. [16] in which participants judged head-referenced image motion via an HMD that suppressed compensatory eye movements during head movements. The similar pattern of data of the two studies shows that concurrent head movements in general suppress observers’ perceptual sensitivity to image motion regardless of VOR-triggered compensatory eye movements.

While reducing the display FOV from 73°H × 73°V to 25°H × 25°V (i.e., about 1/10 of the surface area) did not affect perceptual sensitivity to image motion during concurrent head movements, the intercept data show that it caused an overall underestimation of image motion by 7% across the head movement and head still conditions. This is consistent with the findings that reducing the central FOV to smaller than 60° caused observers to underestimate the translation speed of an optic flow field [28].

In summary, the results from the current study show that a large FOV does not improve perceptual sensitivity to world-referenced image motion during concurrent head movements, although overall motion estimation is close to veridical with a large FOV display. The suppression of perceptual sensitivity to world-referenced image motion during concurrent head movements with a large FOV display is a stable effect which we also observed in a separate experiment in which we asked observers to use a mouse to adjust the distance between a pair of vertical bars to indicate the range of the checkerboard image motion they perceived during head movements. Combining the findings from the current study and those from Li et al. [16], we draw the conclusion that the suppression effect that concurrent yaw and pitch head movements have on the perception of both head- and world-referenced horizontally oscillating image motion is not due to the confusion of retinal motion components. Instead, it is due to a general inhibition of object motion signals performed by a compensatory mechanism that strives to stabilize the perception of the world during head turns [29]. Wallach [29] argued that the inhibition of object motion signals can be a mechanical stabilizing process triggered by VOR signals, and the findings from the current experiment support the proposal of the neural attenuation of object motion signals triggered primarily by vestibular and proprioceptive cues generated during head motion (e.g., [22]).
As the horizontally oscillating checkerboard image creates retinal motion more similar to what would be expected in yaw than pitch head movements, the similar suppression effect that yaw and pitch head movements have on perceptual sensitivity to world-referenced image motion observed in the current study does not appear to agree with Barlow’s [13] inhibition theory that emphasizes sensory correlation. Nevertheless, Li et al. [16] tested a range of image oscillation (0.25–2 Hz) and head movement frequencies (0.25–1 Hz) and found that across the range of the tested frequencies, while both yaw and pitch head movements reduced perceptual sensitivity to head-referenced image motion compared with when the head was still, the reduction in sensitivity with pitch head movements (about 25%) was about half of that with yaw head movements. This supports Barlow’s inhibition theory by showing that the suppression of perceptual sensitivity to horizontal motion is more for the highly related yaw than the non-related pitch head movements. The current study tested only one image oscillation (0.6 Hz) and head movement frequency (0.5 Hz), thus future research is needed to test a range of frequencies to find out whether the findings from Li et al. [16] can be generalized to world-referenced image motion.

The potential applications of the findings of the current study include the design of countermeasures for virtual environment systems to mitigate user perception of spurious image motion arising from head tracking latency. The findings of the current study and from Li et al. [16] together suggest that spurious motion would be less noticeable during head motion than when head is still regardless of whether the image motion is presented to users via an HMD or on a large screen. Furthermore, as the FOV does not appear to affect perceptual sensitivity to image motion during concurrent head movements, increasing the display FOV may not help solving the problem.

Acknowledgements

This study was supported by a grant from the Research Grants Council of Hong Kong (HKU 7478/08H) to L. Li. The authors thank Diederick Niehorster for his helpful comments on a previous draft of the article.

References