Exp Brain Res
DOI 10.1007/s00221-014-4136-1

RESEARCH ARTICLE

Optimal visual–vestibular integration under conditions of conflicting intersensory motion profiles

John S. Butler · Jennifer L. Campos · Heinrich H. Bülthoff

Received: 22 May 2014 / Accepted: 20 October 2014

Abstract Passive movement through an environment is typically perceived by integrating information from different sensory signals, including visual and vestibular information. A wealth of previous research in the field of multisensory integration has shown that if different sensory signals are spatially or temporally discrepant, they may not combine in a statistically optimal fashion; however, this has not been well explored for visual–vestibular integration. Self-motion perception involves the integration of various movement parameters including displacement, velocity, acceleration and higher derivatives such as jerk. It is often assumed that the vestibular system is optimized for the processing of acceleration and higher derivatives, while the visual system is specialized to process position and velocity. In order to determine the interactions between different spatiotemporal properties for self-motion perception, in Experiment 1, we first asked whether the velocity profile of a visual trajectory affects discrimination performance in a heading task. Participants performed a two-interval forced choice heading task while stationary. They were asked to make heading discriminations while the visual stimulus moved at a constant velocity (C-Vis) or with a raised cosine velocity (R-Vis) motion profile. Experiment 2 was designed to assess how the visual and vestibular velocity profiles combined during the same heading task. In this case, participants were seated on a Stewart motion platform and motion information was presented via visual information alone, vestibular information alone or both cues combined. The combined condition consisted of congruent blocks (R-Vis/R-Vest) in which both visual and vestibular cues consisted of a raised cosine velocity profile and incongruent blocks (C-Vis/R-Vest) in which the visual motion profile consisted of a constant velocity motion, while the vestibular motion consisted of a raised cosine velocity profile. Results from both Experiments 1 and 2 demonstrated that visual heading estimates are indeed affected by the velocity profile of the movement trajectory, with lower thresholds observed for the R-Vis compared to the C-Vis. In Exp. 2 when visual–vestibular inputs were both present, they were combined in a statistically optimal fashion irrespective of the type of visual velocity profile, thus demonstrating robust integration of visual and vestibular cues. The study suggests that while the time course of the velocity did affect visual heading judgments, a moderate conflict between visual and vestibular motion profiles does not cause a breakdown in optimal integration for heading.

Keywords Multisensory integration · Self-motion · Maximum likelihood estimation · Optimal integration · Acceleration · Visual · Vestibular

J. S. Butler · J. L. Campos · H. H. Bülthoff (✉)
Max Planck Institute for Biological Cybernetics,
Spemanstrasse 38, 72076 Tübingen, Germany
e-mail: heinrich.buelthoff@tuebingen.mpg.de

J. S. Butler (✉)
Trinity Centre for Bioengineering, Trinity College Dublin,
Dublin 2, Ireland
e-mail: jslbutler@gmail.com

J. L. Campos
iDAPT, Toronto Rehabilitation Institute, UHN, Toronto, Canada

J. L. Campos
Department of Psychology, University of Toronto,
Toronto, Canada

H. H. Bülthoff
Department of Brain and Cognitive Engineering,
Korea University, Seoul, South Korea

Journal : Large 221 | Dispatch : 28-10-2014 | Pages : 11
Article No : 4136 | ☐ LE | ☐ TYPESET
MS Code : EBR-14-0313 | ☑ CP | ☑ DISK

Introduction

As a car travels down 5th Avenue in New York, it accelerates and decelerates every 100 m in response to the traffic signals, while avoiding vehicle and pedestrian traffic and staying within the lane. In order to achieve this successfully, several movement parameters must be effectively perceived, including heading direction and velocity. These movement parameters are perceived using multiple sensory systems, including the visual and vestibular systems. Previous research has shown that both heading (Warren and Hannon 1988; Royden et al. 1992, 1994; Butler et al. 2010; Fetsch et al. 2010a) and relative ego-velocity (Gibson 1950; Frenz and Lappe 2005) can be perceived using visual information alone in the absence of physical cues to motion. In contrast, while there has been a great deal of research assessing the capacity to visually judge the accelerations of external objects (e.g., Brouwer et al. 2002; Schlack and Albright 2007; Schlack et al. 2007, 2008), far fewer studies have considered the capacity of humans to judge accelerations/decelerations in the context of self-motion. Festl et al. (2012) have shown that humans can extract information from optic flow in order to discriminate motion profiles specifying different rates of acceleration. However, this study also concluded that absolute judgments of ego-accelerations are not possible without visual scaling of the scene provided through depth cues (and/or without the scaling provided through non-visual cues).

In terms of vestibular contributions to self-motion perception, a collection of recent research by our group and others has shown that vestibular information alone can also be used to perceive heading in humans and non-human primates (Telford et al. 1995; Butler et al. 2006; de Winkel et al. 2010; Fetsch et al. 2010a; MacNeilage et al. 2010; Crane 2012; Nolan et al. 2012; Cuturi and MacNeilage 2013). During translational motion, the main function of the vestibular system is as an acceleration detector. As the vestibular system responds to changes in acceleration, it cannot directly estimate velocity and therefore becomes uninformative during constant velocity motion (Benson et al. 1986).

Recent inquires have now begun to focus on how information from the visual and vestibular systems are integrated to perceive different aspects of self-motion. Because each of these sensory systems have different reliabilities for perceiving different movement parameters (i.e., vision more sensitive to changes in position and velocity and vestibular more sensitive to detecting accelerations) and for different movement conditions (e.g., vision more sensitive for slow and vestibular more sensitive for fast ego-motion (Berthoz et al. 1975; Zacharias and Young 1981)), thus the optimal combination of the cues should take into account the reliability of the single cue percepts. Indeed, within the context of heading perception, behavioral and neurophysiological studies have shown that when visual and vestibular cues are presented synchronously, they elicit a more reliable behavioral and neuronal response than either cue alone (Butler et al. 2010, 2011a; Fetsch et al. 2010a, 2012; Gu et al. 2011). These results have been embedded into the theoretical framework of maximum likelihood estimation (MLE), such that predictions about the reliability and the weights of the unisensory cues can be used to make predictions in combined cue conditions (see Ernst and Banks 2002; Ernst and Bülthoff 2004). Neurophysiological recordings in non-human primates have revealed heading direction sensitive neurons in areas such as MT, medial superior temporal (MST) and ventral intraparietal (VIP). Models of this neurophysiological data have indicated that vestibular responses are driven mainly by velocity (VIP and MSTd) and acceleration (VIP) components, whereas visual responses were driven mainly by velocity (VIP) (Fetsch et al. 2010b; Chen et al. 2011).

Far fewer studies have attempted to quantify the relative contributions of visual and vestibular cues to the perception of linear acceleration (Berger and Bülthoff 2009; Berger et al. 2010). Berger et al. (2010) asked observers to judge how "believable" a forward accelerating movement was under different visual–vestibular conditions. Their results indicated that a simulated visual acceleration that coincided with a physical backward pitch was rated as most believable. However, the range of physical pitch movements that were rated as believable when combined with the same visual acceleration profile was quite broad. Further, the reasonably high tolerance for physical motions that differed from the visual motions did not seem to be affected by whether the observer was consciously aware of the discrepancy, indicating that this integration is likely quite robust.

There remains a gap in knowledge within this field of research regarding how different spatiotemporal characteristics of a self-motion profile, including information about heading and velocity interact. Given that self-motion perception is not well represented by a discrete event, but takes place over space and time, it is likely that there is a continuous integration of information throughout a given trajectory. This integration may therefore be dependent on different dynamic spatiotemporal parameters and may be contingent on individual sensory cue reliabilities. For instance, it is not yet clear whether the precision of heading perception could be affected by the velocity profile of the motion trajectory. It is also not clear whether the process by which visual and vestibular cues are integrated is affected by whether the visual and vestibular motion profiles are different.

Therefore, in order to first evaluate whether the velocity profile of a motion trajectory affects heading

Journal : Large 221 | Dispatch : 28-10-2014 | Pages : 11
Article No : 4136 | LE | TYPESET
MS Code : EBR-14-0313 | CP | DISK

perception, Experiment 1 evaluated heading discrimination performance on a purely visual heading task for different velocity profile conditions. Specifically, heading comparisons were made for either constant velocity motion profiles or raised cosine velocity profiles. If visual heading discrimination thresholds are purely based upon displacement, there should be no performance differences between the two visual motion conditions. As velocity is the first derivative of displacement, it is impossible AQ3 to change one without changing the other. Therefore, in order to create discrepancies in velocities between the two motion profiles (visual and vestibular), while keeping displacement equal (important to control for in the context of heading perception), average velocity was equated across profiles."

In contrast, if higher derivatives (acceleration) do play a role in visual heading perception, this could lead to several different predictions. First, given that a raised cosine velocity profile is the more natural and commonly experienced visual motion profile and is typically coincident with actual physical self-motion, it is possible that the discrimination thresholds will in fact be lower in the raised cosine velocity trials compared to the constant velocity trials. Another reason why the raised cosine velocity trials might result in a more reliable response is that participants may not integrate information across the whole profile but rather over a shorter window of time than the one second trials which could result in a higher velocity and hence a more reliable heading percept (Crowell and Banks 1993).

However, when considering the visual alone conditions, the vestibular cue is indicating no changes in velocity which could introduce an intersensory conflict. This conflict may be greater for the raised cosine visual motion profile given as there is a change in velocity which is in conflict with the vestibular cue which might result in a less reliable discrimination of heading.

Experiment 2 helped to further test the role of the visual motion profile on self-motion by assessing multisensory heading perception when visual and vestibular cues were combined. Specifically, this experiment evaluated whether visual and vestibular cues would be optimally combined, even when the motion profiles of each differ. Again, the same two visual motion profiles were included during multisensory conditions, leading to cue combinations that were either congruent (both raised cosine velocity profiles) or incongruent (constant velocity visual motion and raised cosine velocity vestibular motion). This also allowed us to evaluate whether visual and vestibular inputs continued to be optimally integrated during self-motion (as has already be shown several times previously using this paradigm) when the visual velocity profile and the physical/vestibular velocity profile are discrepant.

Methods

Participants

Fifteen participants (five males) completed Experiment 1, and six new participants (two males) completed Experiment 2. The average age was 24 years (19–31). All participants had normal or corrected-to-normal vision, including normal stereo vision. Participants were compensated with pay of 8 Euros per hour and were naïve to the purposes of the experiment. Participants gave their informed consent before taking part in the experiment, which was performed in accordance with the ethical standards specified in the 1964 Declaration of Helsinki.

Apparatus

Experiment 1

In Experiment 1, participants were seated 1.0 m in front of a large back-projected screen. The screen had a field of view of 102° × 82° with a resolution of 1,280 × 1,024 pixels and a refresh rate of 60 frames per second.

Experiment 2

The experimental setup and stimuli were identical to those described previously (Butler et al. 2010, 2011a). This experiment was conducted in the Motion Lab at the Max Planck Institute for Biological Cybernetics, which consists of a Maxcue 600, six degree of freedom, Stewart platform manufactured by Motion-Base PLC, UK (Fig. 1a).

To mask the noise of the platform, participants wore noise-cancellation headphones with two-way communication capability while white noise was played. To mask the vibrations of the platform motors, somatosensory vibrations were produced by subwoofers installed under the seat and foot plate. A foam head rest was used to keep head movements to a minimum. The visuals were displayed on a projection screen with a field of view of 86° × 65° and a resolution of 1400 × 1050 pixels and a refresh rate of 60 frames per second. Participants viewed the projection screen through an aperture, which reduced the field of view to 50° × 50°, thereby increasing immersion and avoiding conflicting information provided by the stability of the frame around the screen and the visual motion information being projected on the screen. In both experiments, participants freely viewed the stimulus.

Participants responded using a four-button response box. The stereoscopic image was generated using red–cyan anaglyphs. All experiments were coded using a graphical real-time interactive programming language (Virtools™, France).

Journal : Large 221 | Dispatch : 28-10-2014 | Pages : 11
Article No : 4136 | ☐ LE | ☐ TYPESET
MS Code : EBR-14-0313 | ☑ CP | ☑ DISK

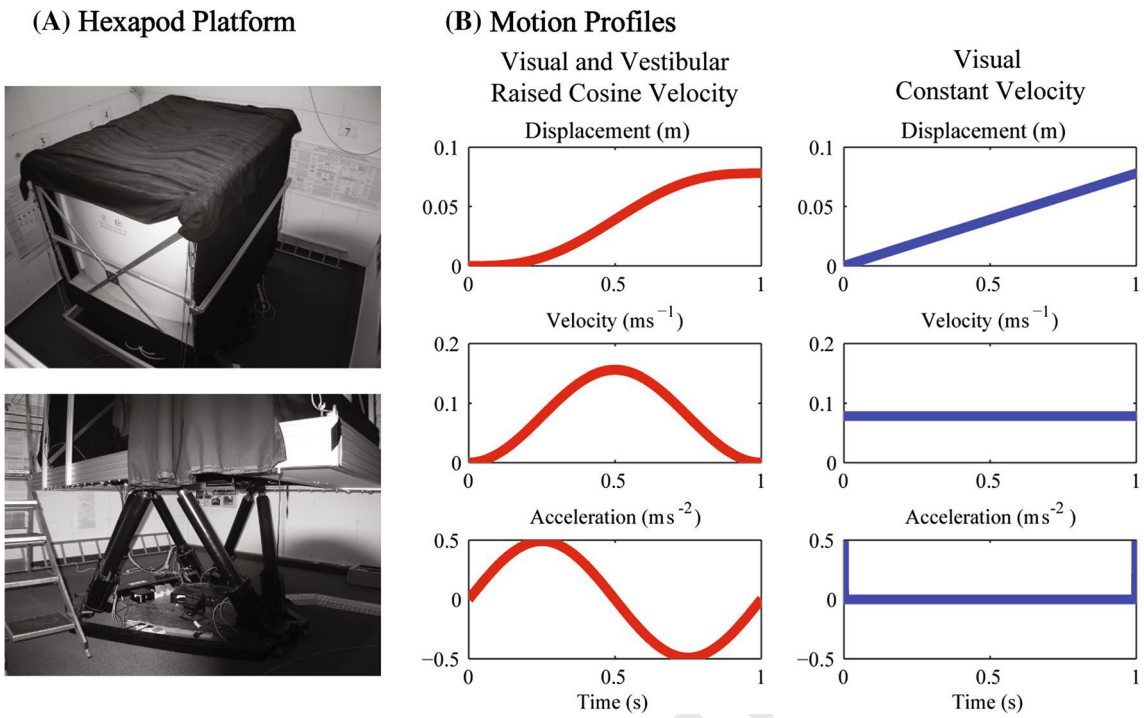

Fig. 1 **a** *Top view* of the hexapod platform. *Bottom view* of the hexapod platform, **b** displacement, velocity and acceleration of the motion profiles. *Left*, raised cosine velocity profile for the vestibular and visual conditions (*red*). *Right*, constant velocity profile which is only for the visual condition (*blue*)

Stimuli

The visual stimulus in both Experiments 1 and 2 consisted of a limited lifetime Gaussian starfield. Each star was a circle that had an average diameter of 0.15° which were Gaussian smoothed in contrast to minimize any sharp contrast changes and hence avoid aliasing and a lifetime in the range of 0.5–1.0 s. The maximum number of stars on the screen at one time was 200, and the minimum was 150. Each dot subtended between 0.1° and 0.2°, which depended on their virtual distance ranging from 2 to 2.5 m. The starfield was presented in stereo on a gray background to facilitate the fusing of the red and cyan images.

Two motion profiles were used in this study, a raised cosine velocity profile (acceleration/decelerations at start/end of motion) and a constant velocity profile (step function with a single impulse in acceleration at the start and of the motion). If visual heading discrimination thresholds are purely based upon displacement, there should be no performance differences between the two visual motion conditions. The raised cosine velocity profile was

$$s_{\text{Raised}}(t) = 0.49\frac{(2\pi t - \sin(2\pi t))}{4\pi^2}, \quad 0 \le t \le 1s. \qquad (1)$$

This profile had a maximum displacement, velocity and acceleration of 0.078 m, 0.156 ms^{-1} and 0.49 ms^{-2}, respectively (Fig. 1b Left), which is above the detection threshold for blindfolded accelerations (Benson et al. 1986).

The constant velocity profile, which was only employed in the visual condition, was

$$s_{\text{constant}}(t) = 0.078t, \quad 0 \le t \le 1s. \qquad (2)$$

This profile has a maximum displacement and velocity of 0.078 m and 0.078 ms^{-1}, respectively. The acceleration profile is 0.0 ms^{-2} except for an impulse at the very start and end of the profile (Fig. 1b Right).

General procedure

Participants performed a two-interval forced choice task (2-IFC) in which they were asked to judge "in which of the two intervals did you move more to the right" (see Fig. 2). Each trial consisted of two heading motions, the standard and the comparison (counterbalanced across trials). The standard angle was always fixed at 0° (straight ahead), while the eight comparison angles were −20° −10° −5° −2°, 2° 5°, 10° and 20°. All trials were initiated with a beep to indicate that the participants could start the trial with a button press. After pressing the start button, there was a 0.75-s pause between the stimulus appearing on the screen and the onset of the motion. Between intervals, there

Journal : Large 221 | Dispatch : 28-10-2014 | Pages : 11
Article No : 4136 | LE | TYPESET
MS Code : EBR-14-0313 | CP | DISK

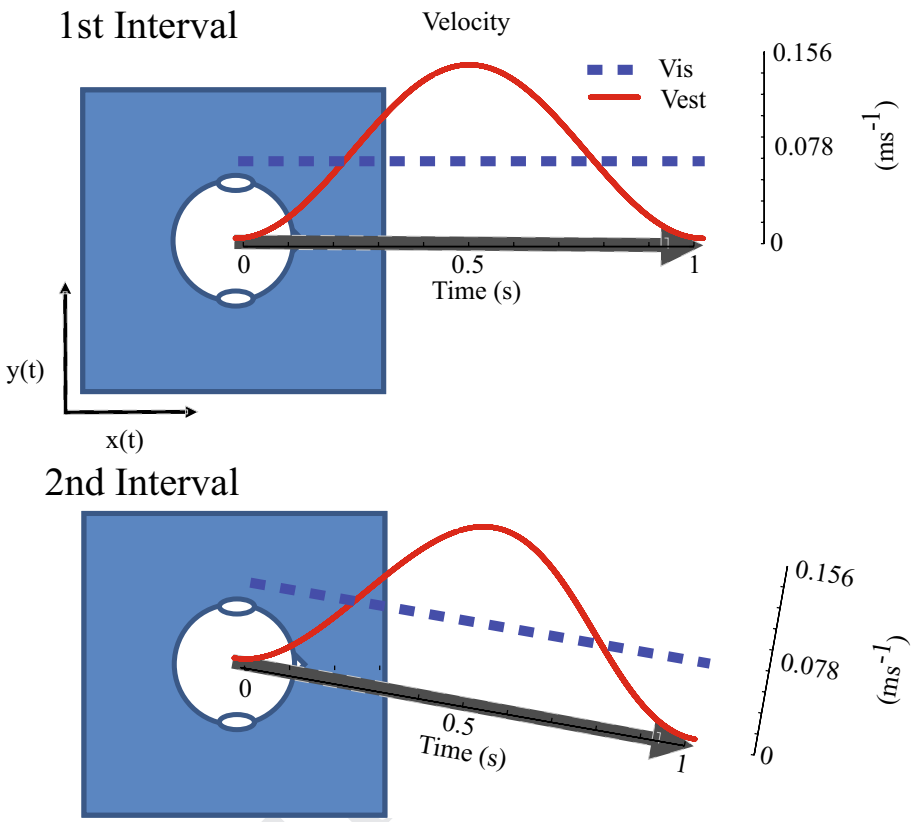

Fig. 2 Top-down schematic of a 2-IFC combined incongruent trial, in which the motion profile for the visual cue is a constant velocity (*blue*) and the vestibular cue is a raised cosine (*red*)

was a 1-s pause. Participants responded after the second interval. In the vestibular and visual–vestibular conditions in Experiment 2, after responding, the participants were passively moved back to the start position with no visual stimulus on the screen at a subthreshold vestibular velocity of 0.025 ms^{-1}, for about 5 s prior to commencing the next trial.

Experiment 1

Participants completed a visual alone heading task. The heading task was blocked with respect to the different motion profiles, a constant velocity profile (C-Vis) and a raised cosine velocity profile (R-Vis), which were counterbalanced across participants. For each condition, there were 30 repetitions of each of the eight comparison angles (240 trials total). These trials were divided into three blocks with each block lasting approximately 7 min. In total with breaks, the experiment lasted 1 h.

Experiment 2

Six naïve participants completed six blocks of visual alone trials (3 C-Vis, 3 R-Vis), three blocks of vestibular alone trials and six blocks of visual–vestibular trials (3 C-Vis/Vest, 3 R-Vis/Vest). Therefore, in the combined cue trials (visual–vestibular), half were congruent (R-Vis/R-Vest), and half were incongruent (C-Vis/R-Vest) (see Fig. 2). Each block consisted of eight comparison angles repeated ten times (240 trials per condition). The experiment was run over five one and a half hour experimental sessions.

Data analysis

Individual participant's data for each condition were fit with a Gaussian psychometric function using the psignifit toolbox (Wichmann and Hill 2001a, b). From the calculated fit, the just noticeable differences (JNDs) were determined, which is proportional to the standard deviation of the distribution. The JND value is inversely related to reliability, and thus, the higher the JND, the higher the discrimination threshold and the lower the reliability. For all analyses, the significance level was set at 0.05.

Results

Experiment 1

Figure 3a shows data from a representative participant for the visual alone raised cosine velocity condition (red) and visual alone constant velocity condition (blue) and cumulative Gaussian fits. The average visual alone heading thresholds for the R-Vis and C-Vis were

Journal : Large 221 | Dispatch : 28-10-2014 | Pages : 11
Article No : 4136 | LE | TYPESET
MS Code : EBR-14-0313 | CP | DISK

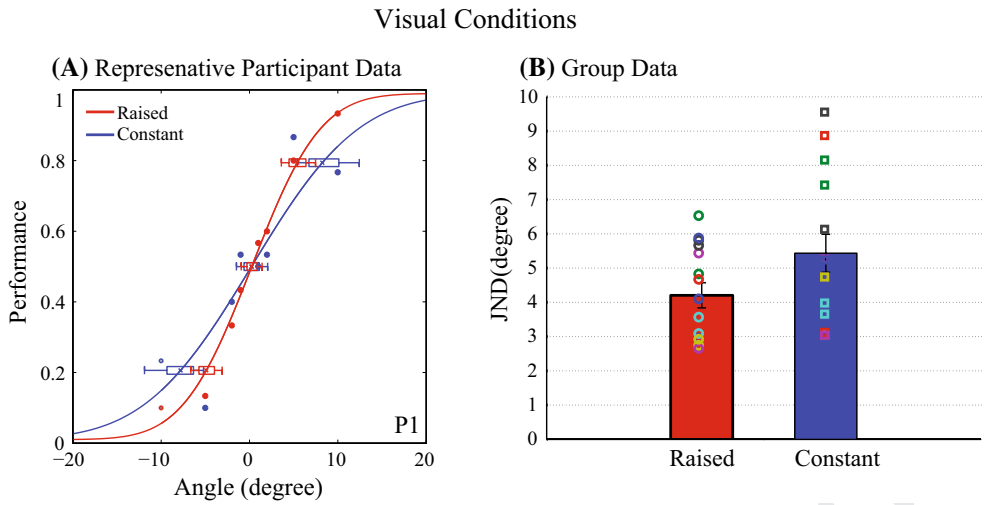

Fig. 3 Experiment 1 results. **a** Data for raised cosine visual alone (*red*), constant visual alone (*blue*) and for a representative participant. The data show the proportion of perceived "more rightward" responses as a function of heading angle. *Solid lines* represent the cumulative Gaussian functions that were fitted to the data. *Box plots* *whiskers* indicate the bootstrapped confidence intervals at −2, −1, 1 and 2 standard deviations. **b** Visual alone group average heading JNDs for the different motion profiles. *Error bars* represent standard error of the mean. *Open circles/squares* represent individual participants

$\mathrm{JND}_{\text{R-Vis}} = 4.20° \pm 0.37°$ and $\mathrm{JND}_{\text{C-Vis}} = 5.43° \pm 0.55°$, respectively (Fig. 3b). The results of a paired two-tailed *t* test revealed that the JNDs for the R-Vis condition were significantly lower than the C-Vis condition ($t_{15} = -3.298$; $p < 0.005$). These results indicate that the type of motion profile affects the discrimination of heading. Furthermore, the more natural yet complex visual motion profile (R-Vis) yielded the most reliable responses.

Experiment 2

Based on MLE models and supported by past studies, a combination of congruent visual and vestibular inputs should result in a reduction of variance compared to unisensory conditions (Fetsch et al. 2009, 2010a; Butler et al. 2010, 2011a). What is unknown is whether optimal integration will occur if there is conflicting information provided by the visual and vestibular inputs in the form of different velocity profiles and how these results compare to visual only conditions. If the combined visual and vestibular cues are integrated irrespective of whether the motion profiles are congruent or incongruent, the unimodal JNDs, $\mathrm{JND}_{\text{Vis}}$, $\mathrm{JND}_{\text{Vest}}$, can be used to predict the optimal $\mathrm{JND}_{\text{Pred}}$ of the visual–vestibular condition (Ernst and Bülthoff 2004).

$$\mathrm{JND}^2_{\text{Pred Vis-Vest}} = \frac{1}{1/\mathrm{JND}^2_{\text{Vis}} + 1/\mathrm{JND}^2_{\text{Vest}}} = \frac{\mathrm{JND}^2_{\text{Vis}}\mathrm{JND}^2_{\text{Vest}}}{\mathrm{JND}^2_{\text{Vis}} + \mathrm{JND}^2_{\text{Vest}}} \quad (3)$$

If the cues are not optimally combined, as has been the case in some previous studies under certain conditions (de Winkel et al. 2010; Butler et al. 2011a), then the JND in the combined conditions would be no less than the JND of the most reliable unimodal cue.

$$\mathrm{JND}_{\text{Vis-Vest}} \geq \min(\mathrm{JND}_{\text{Vis}}, \mathrm{JND}_{\text{Vest}}) \quad (4)$$

The 95 % confidence intervals were calculated for each participant's JND by a bootstrap procedure with 1999 repetitions (for details see Wichmann and Hill 2001b). The predicted multisensory 95 % confidence intervals Δ_{Pred} were calculated by the propagation of error formula (Taylor 1997)

$$\Delta_{\text{Pred}} = \left|\frac{\partial \mathrm{JND}_{\text{Vis-Vest}}}{\partial \mathrm{JND}_{\text{Vis}}}\right| \Delta_{\text{Vis}} + \left|\frac{\partial \mathrm{JND}_{\text{Vis-Vest}}}{\partial \mathrm{JND}_{\text{Vest}}}\right| \Delta_{\text{Vest}}. \quad (5)$$

from the first derivatives of the predicted visual–vestibular $\mathrm{JND}_{\text{Pred}}$ and the visual Δ_{Vis} and vestibular Δ_{Vest} confidence intervals.

Unimodal conditions

The average visual alone heading JNDs for the R-Vis and C-Vis were $\mathrm{JND}_{\text{R-Vis}} = 5.5° \pm 0.54°$, $\mathrm{JND}_{\text{C-Vis}} = 7.3° \pm 0.83°$, thereby replicating the results of Experiment 1 with a different group of participants. This also demonstrates that even though there were differences in the visual displays in Experiments 1 and 2, including different FOVs and an aperture, the result that the velocity profile had an impact on the reliability of the visual alone heading response was unchanged. For vestibular alone heading, the $\mathrm{JND}_{\text{Vest}}$ was $5.6° \pm 0.76°$ (Fig. 4a). An ANOVA was performed on the heading discrimination performance (JND) for the R-Vis

Journal : Large 221 | Dispatch : 28-10-2014 | Pages : 11
Article No : 4136 | LE | TYPESET
MS Code : EBR-14-0313 | CP | DISK

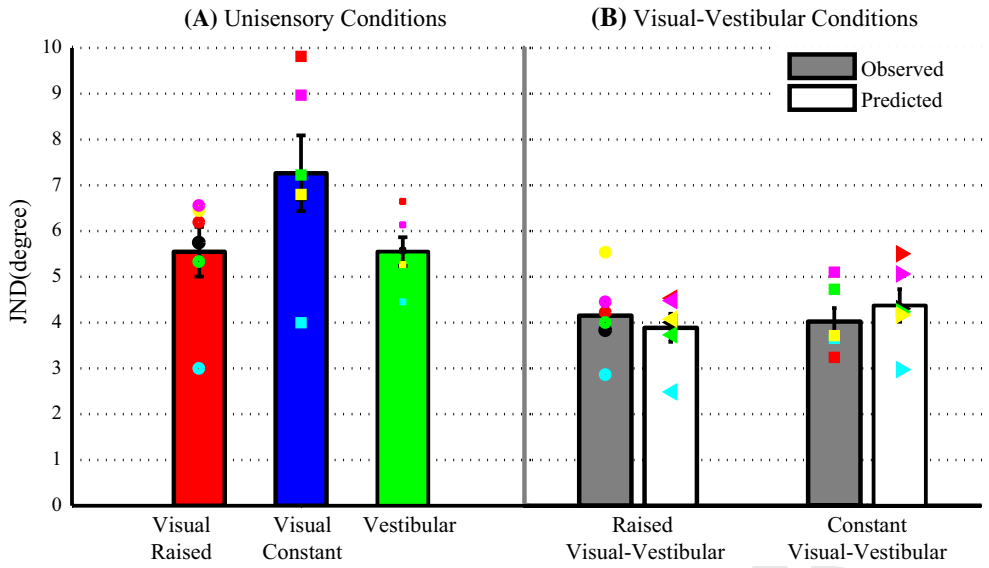

Fig. 4 Experiment 2 results. **a** Unimodal average heading JNDs. **b** Bimodal observed and predicted average heading JNDs for the congruent and incongruent motion profile conditions. *Error bars* represent standard error of the mean. *Overlaid symbols* represent individual participants for each condition

alone, C-Vis alone and Vest alone for all participants. The analysis revealed a significant main effect of condition ($F(2, 10) = 9.086$, MSE $= 0.646$, $p < 0.01$). Three two-tailed paired-sample t tests were conducted to compare the unimodal conditions, R-Vis vs. C-Vis, R-Vis vs. Vest and C-Vis vs. Vest. As in Experiment 1, the R-Vis and C-Vis alone were significantly different ($t_5 = -3.552$, $p < 0.016$). The C-Vis condition was significantly different from the Vest alone condition (C-Vis: $t_5 = -3.22$, $p < 0.025$), but the R-Vis was not significantly different from the Vest alone condition ($p = 0.993$).

Bimodal conditions

The observed and predicted average JNDs for the bimodal congruent motion profiles were $\text{JND}_{\text{Obs\_R-Vis/R-Vest}} = 4.15° \pm 0.35°$, and $\text{JND}_{\text{Pred\_R-Vis/R-Vest}} = 3.88° \pm 0.31°$, respectively. The observed and predicted average JNDs for the bimodal incongruent motion profile were $\text{JND}_{\text{Obs\_C-Vis/R-Vest}} = 4.03° \pm 0.29°$, $\text{JND}_{\text{Pred\_C-Vis/R-Vest}} = 4.37° \pm 0.35°$ (Fig. 4b), respectively.

To compare the multisensory conditions with the visual conditions and the role of motion profile, the data were submitted to a 2 sensory condition (Vis alone vs. Vis–Vest) × 2 motion profile (C-Vis vs. R-Vis) repeated measures ANOVA. The analysis showed a significant effect of sensory condition ($F(1,5) = 18.115$, MSE $= 1.779$, $p < 0.01$), but no significant effect of motion profile ($F(1,5) = 3.768$, MSE $= 0.710$, $p = 0.069$) and a significant interaction effect ($F(1,5) = 5.105$, MSE $= 0.567$, $p = 0.03$). The significant interaction is driven by the statistical difference between the profiles in the visual conditions (see above) but not for the combined condition (Post hoc t test: C-Vis–Vest vs. R-Vis–Vest: $t_5 = 0.296$, $p = 0.776$).

Follow-up paired-sample t tests to investigate the effect of sensory condition revealed a lower JND in the Vis–Vest condition than the Visual Alone for both the raised cosine and constant velocity profile (R-Vis vs. R-Vis–Vest: $t_5 = 3.917$, $p < 0.015$ and C-Vis vs. C-Vis–Vest: $t_5 = 4.463$, $p < 0.01$). To compare the multisensory conditions with the vestibular alone condition, the vestibular JND was submitted to a paired t tests with the raised Vis–Vest JND and the constant Vis–Vest JND which revealed a lower JND in the Vis–Vest conditions (Vest vs. R-Vis–Vest: $t_5 = 3.808$, $p < 0.015$ and Vest vs. C-Vis–Vest: $t_5 = 3.34$, $p < 0.025$). This shows that there was increased reliability of the heading response in the visual–vestibular condition for both the congruent and incongruent conditions.

In order to compare the observed results with model predictions (based on MLE optimal integration), a 2 (observed vs. predicted) × 2 motion profile (C-Vis vs. R-Vis) repeated measures ANOVA was performed. The analysis revealed that observed and predicted JNDs were not significantly different ($F(1,5) = 0.494$, MSE $= 0.022$, $p = 0.89$), nor was there a main effect of motion profile ($F(1,5) = 0.0556$, MSE $= 0.346$, $p = 0.49$) or an interaction effect ($F(1,5) = 2.085$, MSE $= 0.275$, $p = 0.21$). The results show that for both the congruent condition (R-Vis/R-Vest) and the incongruent condition (C-Vis/R-Vest), an

Journal : Large 221 | Dispatch : 28-10-2014 | Pages : 11
Article No : 4136 | LE | TYPESET
MS Code : EBR-14-0313 | CP | DISK

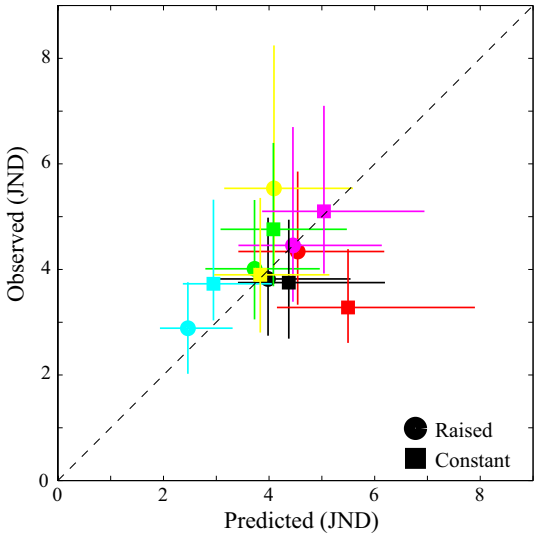

Fig. 5 *Scatterplot* of predicted visual–vestibular JNDs versus observed visual–vestibular JNDs. *Different symbols* represent different visual motion profiles (*circles* = R-Vis/R-Vest and *squares* = C-Vis/R-Vest). *Different colors* represent individual participants. *The dashed line* indicates the ideal data. *Error bars* represent 95 % bootstrapped confidence intervals

optimal reduction in variance was observed that was in line with MLE model predictions (Fig. 5).

Discussion

The results of both Experiment 1 and 2 demonstrate that visual velocity motion profiles can affect visual heading discrimination. The most typical and natural form of visual motion that occurs when changing from a stationary to a moving position includes an initial acceleration component followed by a deceleration component prior to the final resting state. Indeed, the current results show that it is this type of profile that resulted in the least variable visual heading estimates. In contrast, when the visual motion started suddenly (constant velocity), which is an event that does not typically occur during real world interactions, higher thresholds were observed. This suggests that despite the fact that the instructed task at hand (heading perception) was mainly contingent on detecting changes in direction, other spatial–temporal features of the motion profile (i.e., velocity/acceleration in this case) affected the sensitivity of responding. It is not entirely clear based on these initial findings what the nature of this relationship is. It is possible, for instance, that having a natural acceleration component to the visual motion helps to disambiguate self-motion from object motion (Festl et al. 2012). It is also possible that having an acceleration component to the visual motion generates a more compelling sense of "vection" or illusory self-motion, which could improve measures of self-motion perception. Indeed, there are many examples demonstrating that ratings of vection are higher when acceleration components are added to a purely visual motion stimulus (e.g., viewpoint jitter) compared to when constant velocity visual motion is presented (Palmisano et al. 2000, 2003, 2007, 2008, 2011). Notably, vection onsets typically occur after longer periods of constant velocity than was afforded by the stimulus duration used here. This is presumably because illusory visual motion becomes most compelling once the conflicting expectation of the vestibular onset cue to motion (i.e., the acceleration component) subsides (Palmisano et al. 2008). However, the current results demonstrate that even within the 1-s stimulus presented here, significant differences in the sensitivity to visual acceleration profiles are clearly observed.

It should also be noted that the depth scaling information provided by the starfield stimuli that was used in the current study was only provided through stereo cues, and therefore, estimates of absolute velocity/acceleration were not possible (relative comparisons only). It is conceivable that having additional scaling information from richer depth cues may also impact the extent to which visual velocity profiles affect heading discrimination.

The second main finding of this study was that when the different visual velocity motion profiles were presented simultaneously with vestibular inputs, optimal cue integration was observed irrespective of whether the cues were congruent (R-Vis and R-Vest) or incongruent (C-Vis and R-Vest). This suggests that the integration of visual and vestibular cues is quite robust and tolerant of spatiotemporal conflicts. In the context of multisensory integration for other sensory combinations (e.g., visual-tactile, visual auditory), a wealth of evidence has shown that optimal integration is affected by cue coincidence in space and/or time (Hartcher-O'Brien et al. 2014). That is, when information from two sensory systems occur at locations that are too far apart or occur at moments that are too far separated in time, these inputs are not perceived as being associated with one event and thus optimal integration fails. Specifically, this can lead to the reliability of the responses under multisensory conditions being equal to or worse than either of the unisensory conditions (Wallace et al. 2004; Gepshtein et al. 2005; Roach et al. 2006; Kording et al. 2007; Bentvelzen et al. 2009; Wozny et al. 2010; Wozny and Shams 2011). It is worth noting that in our study while the temporal/velocity profiles were incongruent, other characteristics of the stimuli were consistent (e.g., the onset of the stimuli and the heading information which were always congruent across motion profiles). This could explain why participants combined the visual and vestibular inputs in this case, which is unlike other examples showing a breakdown of optimal integration in the face of intersensory conflicts.

Journal : Large 221 | Dispatch : 28-10-2014 | Pages : 11
Article No : 4136 | LE | TYPESET
MS Code : EBR-14-0313 | CP | DISK

However, there are other recent examples that also demonstrate robust cue integration under conflicting sensory conditions. Raposo et al. (2012) and Sheppard et al. (2013) showed that both humans and rats combine audio and visual information over time in an optimal fashion even when the stimuli were not presented synchronously, thus illustrating that the time course of stimuli do not have to align in time for optimal integration. This result agrees with our findings such that even when the visual and vestibular motion profiles are incongruent, participants combine the information in an optimal fashion. In a recent paper, Drugowitsch et al. (2014) have proposed a model of visual–vestibular integration using a novel diffusion model in which visual velocity information and vestibular acceleration information is accumulated over time to make a decision. This model does not require that the stimuli are synchronous but that the information for the decision is combined and accumulated over time.

The optimal integration of incongruent visual and vestibular motion profiles could be explained in relation to the brain acting as a moving average (low pass) filter of the sensory signals (Werkhoven et al. 1992). Thus, while the motion profiles are inconsistent at the start and the end of the movement, the period around the peak velocity could be perceived as similar and hence integrated. Furthermore, this could account for the more reliable visual alone heading responses in the raised cosine velocity profile condition. While over the one second movement, the displacement and the average velocity are identical for both motion profiles if the heading discrimination judgment is made over a smaller window of time centered around the maximum velocity than the raised cosine velocity profile would have a larger displacement and average velocity than the constant velocity profile and as higher velocities can result in a more accurate discrimination of heading (Crowell and Banks 1996).

Currently, very little is known about the tolerance of visual–vestibular integration to spatial or temporal conflicts. There is evidence that visual–vestibular integration remains optimal under spatial discrepancies (Butler et al. 2010; Fetsch et al. 2010b) and that the believability of forward accelerations is maintained under a range of physical pitch angles (MacNeilage et al. 2007; Berger et al. 2010). Ash and Palmisano (2012) have also demonstrated that vection is enhanced when visual and vestibular inputs are combined during oscillating head movements, even if the physical oscillations are not spatially congruent with the visual oscillations. In general, however, the range of motion parameters under which visual–vestibular conflicts lead to non-optimal integration is unknown. One possible limitation of this study is that we cannot speak to the relative weights of the visual and vestibular cues. Ideally, future research will sample from a wider and more comprehensive range of spatial and temporal offsets, but it should be noted that in our previous work, there was a bias toward the vestibular cue which was not accounted for the by reliability of the cue which will have to be taken into consideration (Butler et al. 2010).

There is, however, reason to believe that the optimal integration of visual and vestibular inputs may be particularly resistant to cue conflicts, because unlike other cue combinations, these two cues maintain a very tight, causal relationship (MacNeilage et al. 2007; Frissen et al. 2011; Campos et al. 2012; Prsa et al. 2012). (Prsa et al. 2012) recently argued that mandatory fusion between visual and vestibular cues is observed during ego-rotations. Specifically, they argue that unisensory estimates of rotations from visual and vestibular inputs are not retained once integration has occurred. While mandatory fusion has been demonstrated for intramodal cue integration (e.g., visual cue integration—see Hillis et al. 2002), this is the first reported demonstration in the context of cross-modal integration. Future experiments would test the limit of the combination of the cues by including larger incongruencies between the visual and vestibular motion profiles. A caveat to the mandatory fusion argument is that multisensory integration can be stimulus and task specific; for example, auditory and tactile cues exhibit multisensory integration for frequency discrimination (Yau et al. 2009; Butler et al. 2012) but not for duration or intensity discrimination (Yau et al. 2010; Butler et al. 2011b). This could also be the case for visual–vestibular integration that for a different task mandatory integration might breakdown.

There are several applied implications for these findings in the context of both fixed-based and motion-based simulation (Teufel et al. 2007; Bles and Groen 2009; Barnett-Cowan et al. 2012). For instance, it is not uncommon for purely visually based simulations (i.e., fixed-base) to refrain from using visual accelerations in an attempt to avoid the experiences of motion sickness that can be caused by sensory conflicts (Wallis et al. 2002). However, the use of such motion profiles could have more global effects on perception and performance across a variety of parameters (i.e., velocity and heading perception). In the context of motion-based simulators, it is typically the case that the range of movements being simulated extends beyond the motion capabilities of the platform, and therefore, motion cueing algorithms must be used to create the illusion of extended motion (Grant and Reid 1997). One of the key necessities of these algorithms is to ensure an acceptable congruency between visual and physical motion cues. The results of the current study suggest that the inclusion of physical motion leads to more precise estimates of self-motion and that this precision persists even when conflicts are present between visual and physical inputs. Understanding the range of tolerable conflicts will help to better

Journal : Large 221 | Dispatch : 28-10-2014 | Pages : 11
Article No : 4136 | LE | TYPESET
MS Code : EBR-14-0313 | CP | DISK

define motion cueing algorithms in future and maximize the capabilities of such simulations.

Conclusion

The study demonstrates that, while visual velocity profiles can affect the sensitivity of visual heading discrimination, a conflict between visual and vestibular velocity profiles does not cause a breakdown in optimal integration for heading. This suggests that optimal visual–vestibular integration is robust in the face of some spatiotemporal conflicts and future work will help to define the extent to which these types of conflicts are tolerated before integration breaks down.

Acknowledgments This research was supported by the Max Planck Society and by the Brain Korea 21 PLUS Program through the National Research Foundation of Korea funded by the Ministry of Education. We are grateful to the participants for their time. We are also grateful to Paul "Pogen" MacNeilage, Julian Hoffmann, Edel Flynn, Karl Beykirch, Stuart Smith and Jack Loomis for their help at different stages of this project.

References

Ash A, Palmisano S (2012) Vection during conflicting multisensory information about the axis, magnitude, and direction of self-motion. Perception 41:253–267

Barnett-Cowan M, Meilinger T, Vidal M, Teufel H, Bülthoff HH (2012) MPI CyberMotion Simulator: implementation of a novel motion simulator to investigate multisensory path integration in three dimensions. J Vis Exp. doi:10.3791/3436

Benson AJ, Spencer MB, Stott JR (1986) Thresholds for the detection of the direction of whole-body, linear movement in the horizontal plane. Aviat Space Environ Med 57:1088–1096

Bentvelzen A, Leung J, Alais D (2009) Discriminating audiovisual speed: optimal integration of speed defaults to probability summation when component reliabilities diverge. Perception 38:966–987

Berger DR, Bülthoff HH (2009) The role of attention on the integration of visual and inertial cues. Exp Brain Res 198:287–300. doi:10.1007/s00221-009-1767-8

Berger DR, Schulte-Pelkum J, Bülthoff HH (2010) Simulating believable forward accelerations on a Stewart motion platform. ACM transactions on applied perception 7(Artn 5). doi:10.1145/1658349.1658354

Berthoz A, Pavard B, Young LR (1975) Perception of linear horizontal self-motion induced by peripheral vision (linearvection) basic characteristics and visual–vestibular interactions. Exp Brain Res 23:471–489

Bles W, Groen E (2009) The DESDEMONA motion facility: applications for space research. Microgravity Sci Technol 21:281–286. doi:10.1007/s12217-009-9120-1

Brouwer AM, Brenner E, Smeets JB (2002) Perception of acceleration with short presentation times: can acceleration be used in interception? Percept Psychophys 64:1160–1168

Butler JS, Smith ST, Beykirch K, Bülthoff HH (2006) Visual vestibular interactions for self motion estimation. In: Driving simulation conference, Paris

Butler JS, Smith ST, Campos JL, Bülthoff HH (2010) Bayesian integration of visual and vestibular signals for heading. J Vis 10:23. doi:10.1167/10.11.23

Butler JS, Campos JL, Bülthoff HH, Smith ST (2011a) The role of stereo vision in visual–vestibular integration. Seeing Perceiving 24:453–470. doi:10.1163/187847511X588070

Butler JS, Molholm S, Fiebelkorn IC, Mercier MR, Schwartz TH, Foxe JJ (2011b) Common or redundant neural circuits for duration processing across audition and touch. J Neurosci 31:3400–3406. doi:10.1523/JNEUROSCI.3296-10.2011

Butler JS, Foxe JJ, Fiebelkorn IC, Mercier MR, Molholm S (2012) Multisensory representation of frequency across audition and touch: high density electrical mapping reveals early sensory-perceptual coupling. J Neurosci 32:15338–15344. doi:10.1523/JNEUROSCI.1796-12.2012

Campos JL, Butler JS, Bülthoff HH (2012) Multisensory integration in the estimation of walked distances. Exp Brain Res. doi:10.1007/s00221-012-3048-1

Chen A, DeAngelis GC, Angelaki DE (2011) Representation of vestibular and visual cues to self-motion in ventral intraparietal cortex. J Neurosci 31:12036–12052. doi:10.1523/JNEUROSCI.0395-11.2011

Crane BT (2012) Direction specific biases in human visual and vestibular heading perception. PLoS One 7:e51383. doi:10.1371/journal.pone.0051383

Crowell JA, Banks MS (1993) Perceiving heading with different retinal regions and types of optic flow. Percept Psychophys 53:325–337

Crowell JA, Banks MS (1996) Ideal observer for heading judgments. Vis Res 36:471–490

Cuturi LF, MacNeilage PR (2013) Systematic biases in human heading estimation. PLoS One 8:e56862. doi:10.1371/journal.pone.0056862

de Winkel KN, Weesie J, Werkhoven PJ, Groen EL (2010) Integration of visual and inertial cues in perceived heading of self-motion. J Vis 10:1. doi:10.1167/10.12.1

Drugowitsch J, DeAngelis GC, Klier EM, Angelaki DE, Pouget A (2014) Optimal multisensory decision-making in a reaction-time task. Elife:e03005 doi: 10.7554/eLife.03005

Ernst MO, Banks MS (2002) Humans integrate visual and haptic information in a statistically optimal fashion. Nature 415:429–433. doi:10.1038/415429a

Ernst MO, Bülthoff HH (2004) Merging the senses into a robust percept. Trends Cogn Sci 8:162–169. doi:10.1016/j.tics.2004.02.002

Festl F, Recktenwald F, Yuan C, Mallot HA (2012) Detection of linear ego-acceleration from optic flow. J Vis 12 doi: 10.1167/12.7.10

Fetsch CR, Turner AH, DeAngelis GC, Angelaki DE (2009) Dynamic reweighting of visual and vestibular cues during self-motion perception. J Neurosci 29:15601–15612. doi:10.1523/JNEUROSCI.2574-09.2009

Fetsch CR, Deangelis GC, Angelaki DE (2010a) Visual–vestibular cue integration for heading perception: applications of optimal cue integration theory. Eur J Neurosci 31:1721–1729. doi:10.1111/j.1460-9568.2010.07207.x

Fetsch CR, Rajguru SM, Karunaratne A, Gu Y, Angelaki DE, DeAngelis GC (2010b) Spatiotemporal properties of vestibular responses in area MSTd. J Neurophysiol 104:1506–1522. doi:10.1152/jn.91247.2008

Fetsch CR, Pouget A, DeAngelis GC, Angelaki DE (2012) Neural correlates of reliability-based cue weighting during multisensory integration. Nat Neurosci 15:146–154. doi:10.1038/nn.2983

Frenz H, Lappe M (2005) Absolute travel distance from optic flow. Vis Res 45:1679–1692. doi:10.1016/j.visres.2004.12.019

Frissen I, Campos JL, Souman JL, Ernst MO (2011) Integration of vestibular and proprioceptive signals for spatial updating. Exp Brain Res 212:163–176. doi:10.1007/s00221-011-2717-9

Journal : Large 221 | Dispatch : 28-10-2014 | Pages : 11
Article No : 4136 | ☐ LE | ☐ TYPESET
MS Code : EBR-14-0313 | ☑ CP | ☑ DISK

Gepshtein S, Burge J, Ernst MO, Banks MS (2005) The combination of vision and touch depends on spatial proximity. J Vis 5:1013–1023. doi:10.1167/5.11.7

Gibson JJ (1950) The perception of the visual world. Houghton Mifflin, Boston

Grant PR, Reid LD (1997) Motion washout filter tuning : rules and requirements. American Institute of Aeronautics and Astronautics, Reston

Gu Y, Liu S, Fetsch CR et al (2011) Perceptual learning reduces interneuronal correlations in macaque visual cortex. Neuron 71:750–761. doi:10.1016/j.neuron.2011.06.015

Hartcher-O'Brien J, Di Luca M, Ernst MO (2014) The duration of uncertain times: audiovisual information about intervals is integrated in a statistically optimal fashion. PLoS One 9:e89339. doi:10.1371/journal.pone.0089339

Hillis JM, Ernst MO, Banks MS, Landy MS (2002) Combining sensory information: mandatory fusion within, but not between, senses. Science 298:1627–1630. doi:10.1126/science.1075396

Kording KP, Beierholm U, Ma WJ, Quartz S, Tenenbaum JB, Shams L (2007) Causal inference in multisensory perception. PLoS One 2:e943. doi:10.1371/journal.pone.0000943

MacNeilage PR, Banks MS, Berger DR, Bülthoff HH (2007) A Bayesian model of the disambiguation of gravitoinertial force by visual cues. Exp Brain Res 179:263–290. doi:10.1007/s00221-006-0792-0

MacNeilage PR, Banks MS, DeAngelis GC, Angelaki DE (2010) Vestibular heading discrimination and sensitivity to linear acceleration in head and world coordinates. J Neurosci 30:9084–9094. doi:10.1523/JNEUROSCI.1304-10.2010

Nolan H, Butler JS, Whelan R, Foxe JJ, Bülthoff HH, Reilly RB (2012) Neural correlates of oddball detection in self-motion heading: a high-density event-related potential study of vestibular integration. Exp Brain Res. doi:10.1007/s00221-012-3059-y

Palmisano S, Gillam BJ, Blackburn SG (2000) Global-perspective jitter improves vection in central vision. Perception 29:57–67

Palmisano S, Burke D, Allison RS (2003) Coherent perspective jitter induces visual illusions of self-motion. Perception 32:97–110

Palmisano S, Bonato F, Bubka A, Folder J (2007) Vertical display oscillation effects on forward vection and simulator sickness. Aviat Space Environ Med 78:951–956

Palmisano S, Allison RS, Pekin F (2008) Accelerating self-motion displays produce more compelling vection in depth. Perception 37:22–33

Palmisano S, Allison RS, Kim J, Bonato F (2011) Simulated viewpoint jitter shakes sensory conflict accounts of vection. Seeing Perceiving 24:173–200. doi:10.1163/187847511X570817

Prsa M, Gale S, Blanke O (2012) Self-motion leads to mandatory cue fusion across sensory modalities. J Neurophysiol 108:2282–2291. doi:10.1152/jn.00439.2012

Raposo D, Sheppard JP, Schrater PR, Churchland AK (2012) Multisensory decision-making in rats and humans. J Neurosci 32:3726–3735. doi:10.1523/JNEUROSCI.4998-11.2012

Roach NW, Heron J, McGraw PV (2006) Resolving multisensory conflict: a strategy for balancing the costs and benefits of audio-visual integration. Proc Biol Sci 273:2159–2168. doi:10.1098/rspb.2006.3578

Royden CS, Banks MS, Crowell JA (1992) The perception of heading during eye movements. Nature 360:583–585. doi:10.1038/360583a0

Royden CS, Crowell JA, Banks MS (1994) Estimating heading during eye movements. Vis Res 34:3197–3214

Schlack A, Albright TD (2007) Remembering visual motion: neural correlates of associative plasticity and motion recall in cortical area MT. Neuron 53:881–890. doi:10.1016/j.neuron.2007.02.028

Schlack A, Krekelberg B, Albright TD (2007) Recent history of stimulus speeds affects the speed tuning of neurons in area MT. J Neurosci 27:11009–11018. doi:10.1523/JNEUROSCI.3165-07.2007

Schlack A, Krekelberg B, Albright TD (2008) Speed perception during acceleration and deceleration. J Vis 8(9):1–11. doi:10.1167/8.8.9

Sheppard JP, Raposo D, Churchland AK (2013) Dynamic weighting of multisensory stimuli shapes decision-making in rats and humans. J Vis 13 doi: 10.1167/13.6.4

Taylor JR (1997) An introduction to error analysis : the study of uncertainties in physical measurements. University Science Books, Sausalito

Telford L, Howard IP, Ohmi M (1995) Heading judgments during active and passive self-motion. Exp Brain Res 104:502–510

Teufel H, Nusseck H-GG, Beykirch K, Butler JS, Kerger M, Bülthoff HH (2007) MPI motion simulator: development and analysis of a novel motion simulator. In: AIAA modeling and simulation technologies conference and exhibit, South Carolina, pp 1–11

Wallace MT, Roberson GE, Hairston WD, Stein BE, Vaughan JW, Schirillo JA (2004) Unifying multisensory signals across time and space. Exp Brain Res 158:252–258. doi:10.1007/s00221-004-1899-9

Wallis G, Chatziastros A, Bülthoff H (2002) An unexpected role for visual feedback in vehicle steering control. Curr Biol 12:295–299

Warren WH, Hannon DJ (1988) Direction of self-motion is perceived from optical flow. Nature 336:162–163

Werkhoven P, Snippe HP, Toet A (1992) Visual processing of optic acceleration. Vis Res 32:2313–2329

Wichmann FA, Hill NJ (2001a) The psychometric function: I. Fitting, sampling, and goodness of fit. Percept Psychophys 63:1293–1313

Wichmann FA, Hill NJ (2001b) The psychometric function: II. Bootstrap-based confidence intervals and sampling. Percept Psychophys 63:1314–1329

Wozny DR, Shams L (2011) Recalibration of auditory space following milliseconds of cross-modal discrepancy. J Neurosci 31:4607–4612. doi:10.1523/JNEUROSCI.6079-10.2011

Wozny DR, Beierholm UR, Shams L (2010) Probability matching as a computational strategy used in perception. PLoS Comput Biol 6 doi: 10.1371/journal.pcbi.1000871

Yau JM, Olenczak JB, Dammann JF, Bensmaia SJ (2009) Temporal frequency channels are linked across audition and touch. Curr Biol 19:561–566. doi:10.1016/j.cub.2009.02.013

Yau JM, Weber AI, Bensmaia SJ (2010) Separate mechanisms for audio-tactile pitch and loudness interactions. Front Psychol 1:160. doi:10.3389/fpsyg.2010.00160

Zacharias GL, Young LR (1981) Influence of combined visual and vestibular cues on human perception and control of horizontal rotation. Exp Brain Res 41:159–171

Journal : Large 221 | Dispatch : 28-10-2014 | Pages : 11
Article No : 4136 | LE | TYPESET
MS Code : EBR-14-0313 | CP | DISK

Journal:	221
Article:	4136

Springer
the language of science

Author Query Form

Please ensure you fill out your response to the queries raised below and return this form along with your corrections

Dear Author

During the process of typesetting your article, the following queries have arisen. Please check your typeset proof carefully against the queries listed below and mark the necessary changes either directly on the proof/online grid or in the 'Author's response' area provided below

Author Proof

Query	Details Required	Author's Response
AQ1	Please check and confirm inserted city name "Toronto" in the affiliation is correct .	
AQ2	Please confirm the section headings are correctly identified.	
AQ3	Please check for the missing quotes in the sentence 'Therefore, in order…equated across profiles'.	