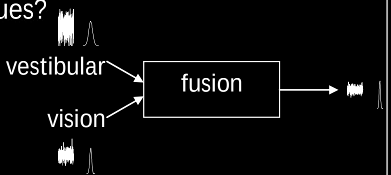
Perceptual roll tilt thresholds demonstrate visual-vestibular fusion

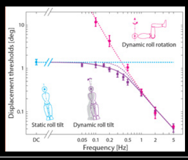
Faisal Karmali, Koeun Lim, Adil Adatia, Romain Claret, Keyvan Nicoucar, Daniel M. Merfeld
Jenks Vestibular Physiology Lab, Massachusetts Eye & Ear Infirmary and Harvard Medical School, Boston, USA.

Introduction

- How does the brain perform sensory fusion of vestibular and visual cues?
- Is visual-vestibular fusion consistent with optimal linear Maximum Likelihood Estimation (MLE)?
- Which predominate as a function of frequency - visual or vestibular cues?

- Estimation requires fusion of noisy and sometimes ambiguous signals.
- Sensory recognition thresholds, which are proportional to the precision of a sensory cue, provide a method of examining how sensory signals are fused.

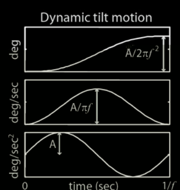




- Cross-modality fusion of tactile and vision cues follows Maximum Likelihood Estimation (MLE), in which cues are weighted proportionally to their precision (Ernst and Banks, 2002).
- We (Lim et al., 2009) showed that roll tilt direction recognition thresholds depend on the fusion of SCC and otolith cues.

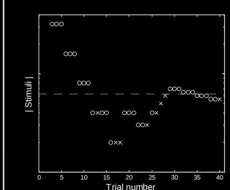
Methods

- Perceptual thresholds were measured during roll tilt motion at a range of frequencies ($T(f)$) in three conditions. Although cues such as tactile and somatosensory are also available, we previously have shown they provide a minor contribution.
- Thresholds were measured using an adaptive staircase (3-down 1-up) in which subjects tilted either left or right and indicated perceived tilt direction.



Experimental Conditions	
Vestibular only	Subject rotated in the dark
Visual only	Subject still while observing a rotating scene
Vestibular+visual	Subject rotated with the lights on

- Dynamic motions consisted of a single-cycle of sinusoidal angular acceleration yielding a sigmoid-like displacement.
- A range of frequencies (.05, .1, .2, .5, 1, 2, 5 Hz) allowed us to use differing dynamic properties to separate the contribution of each cue (canal, otolith, vision).
- Subjects were screened for normal vestibular function (4 female, 3 male, mean age 39, range 29-61).
- Predictions of vestibular+visual thresholds based on individual thresholds were made using MLE, with standard deviations calculated using a bootstrap.



$$\frac{1}{(T(f)_{visual+vestibular})^2} = \frac{1}{(T(f)_{visual})^2} + \frac{1}{(T(f)_{vestibular})^2}$$

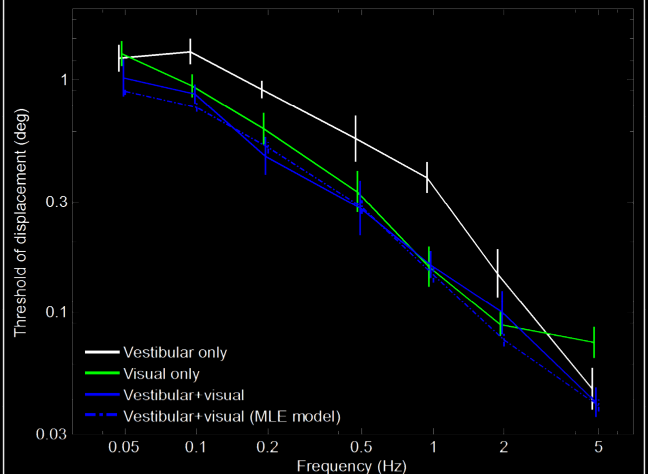
Results

Vestibular cues predominate over visual cues at 5 Hz. Visual cues predominate over vestibular cues at 0.1-2 Hz. Vestibular and visual cues contribute equally at 0.05 Hz.

The figure shows the roll tilt displacement threshold for recognition of direction at different frequencies in three conditions. In the vestibular only condition (black), gravitational tilt measured by the otoliths dominated at low frequencies while the SCC contribution increased at higher frequencies. As frequency increased, threshold was reached with smaller tilt angles, dropping from 1.24˚ at 0.05 Hz to 0.05˚ at 5 Hz. In the visual only condition (green), thresholds decreased from 1.31˚ at 0.05 Hz to 0.07˚ at 2 Hz. The inflection at 2 Hz is consistent with previous studies (Nakayama 1981). These results show that visual cues are more precise than vestibular cues at intermediate frequencies, less precise at 5 Hz, and similar at 0.05 Hz. In the vestibular+visual condition (solid blue), in which visual and vestibular cues were available to the brain, thresholds were 36% smaller than vestibular only thresholds; differences were significant at 0.1 Hz-1 Hz ($p<0.05$; one-sided *t*-test). Vestibular+visual thresholds were also 13% smaller than visual only thresholds, although this difference was only significant at 5 Hz ($p<0.05$; one-sided *t*-test).

Fusion of visual and vestibular cues is consistent with optimal linear weighting (MLE).

The predicted vestibular+visual threshold (dashed blue) based on combining the two individual thresholds was calculated using maximum likelihood estimation (MLE). The predicted and empirical thresholds are not significantly different at any frequency (ranging from $p=0.19$ to $p=0.97$; *t*-test) or across all frequencies ($p=0.23$; ANOVA). These results suggest that fusion of visual and vestibular cues is consistent with MLE weighting.

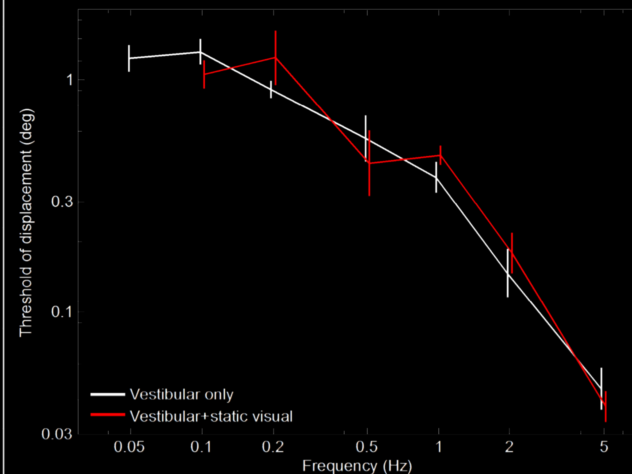


Secondary experiment: Static visual cues only

In a secondary experiment, we measured thresholds during combined visual-vestibular stimulation, but only provided visual cues before and after motion; during motion the lights were off.

Vestibular cues predominate over static visual cues.

Figure 2 shows the roll tilt position thresholds for recognition of motion for which the lights were off during motion and on before and after motion, providing visual orientation cues, but not visual motion cues. The thresholds (red) are not different from vestibular only (black) thresholds ($p>0.2$ for all frequencies; *t*-test). We conclude that during self motion vestibular cues are more precise than visual cues presented only before and after motion.



Conclusions

1. The predominance of vestibular and visual cues varies with frequency
2. Roll tilt motion thresholds are smaller when vestibular and visual cues are combined, consistent with MLE
3. Vestibular cues predominate over static visual cues

Ernst MO, Banks MS (2002) Humans integrate visual and haptic information in a statistically optimal fashion. Nature 415: 429-433.
Lim, K., Nicoucar, K, and Merfeld, D. M. (2009) Perceptual direction-detection thresholds for whole body roll tilts about an earth-horizontal axis. Society for Neuroscience Abstracts.
Nakayama K, Tyler CW (1981) Psychophysical isolation of movement sensitivity by removal of familiar position cues. Vision Res 21: 427-433.

Authors gratefully acknowledge the support of NIH/NIDCD R01-DC04158