Integration of Visual-Vestibular Self Motion:
Comparison of Landmark and Optic Flow Information

Markus von der Heyde & Heinrich H. Bülthoff

MAX-PLANCK-GESELLSCHAFT

Max Planck Institute for Biological Cybernetics - Cognitive and Computational Psychophysics Department - Tübingen, Germany
Markus.von.der.Heyde@Tuebingen.MPG.de - www.kyb.tuebingen.mpg.de

MPI FOR BIOLOGICAL CYBERNETICS

• Introduction

Turns in VR are frequently misperceived.

Perception of turn movements is crucial for self-localization and, consequently, for navigation. Yet in most virtual reality (VR) applications turns are misperceived, which frequently leads to disorientation. We compared the effects of optic flow information and reliable landmark information (see Fig. 1) on perceived turns, each in combination with vestibular information.

We investigated two sources of information: optic flow and landmarks.

Snapshots of the visual scenes

Fig. 1: The textured ground plane (a) provides optic flow information. The town scene (b) provides additional landmark information and absolute size cues. Subjects never saw a bird's eye view (c) of the scene.

• Methods

We presented visual and vestibular turns in VR.

We used a VR setup (Fig. 2) to provide both vestibular and visual turn information. The Motion-Lab enables us to investigate spatial updating and spatial cognition in virtual environments using multiple sensory modalities.

VR setup in the Motion-Lab

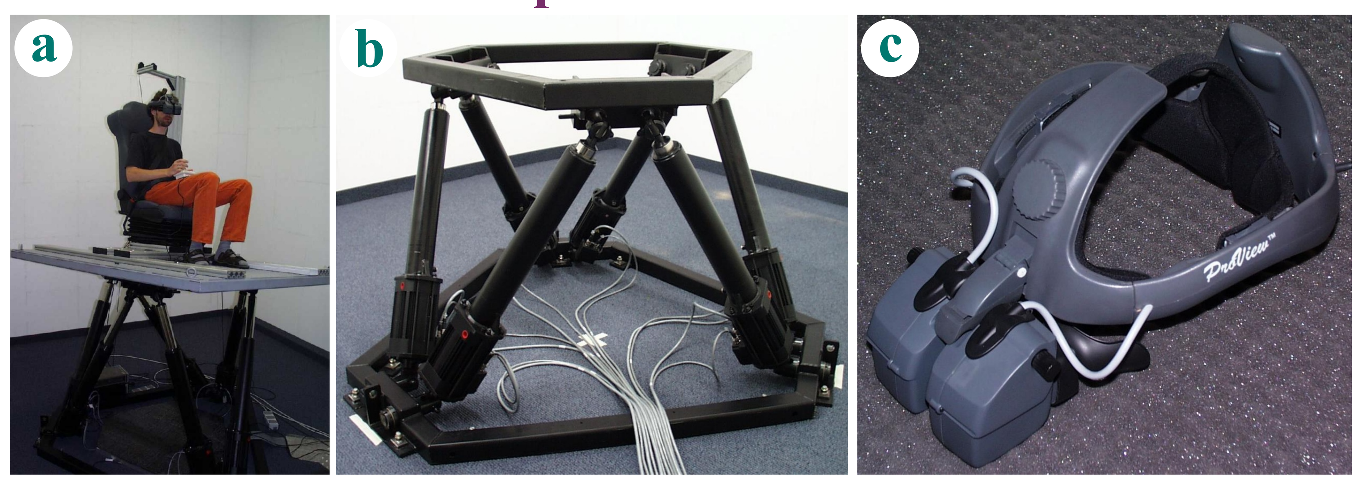


Fig. 2: The Motion-Lab setup (a) integrates a six degree of freedom motion platform (b: Stewart platform) and a high resolution (1024x768, 40x30 deg. FOV) head mounted display (c: Kaiser ProView 60 HMD).

Task: Memorize a trajectory with three turns and reproduce it with different gain factors.

The subjects' task was to learn and memorize a trajectory with three turns that included heading changes between 8.5 and 17 degrees. Subjects were visually translated with constant velocity of 1 m/s. The motion platform performed initial acceleration and final deceleration phases. The trials were between 40 and 60 seconds long. During a reproduction phase, the gain between the joystick control and the resulting visual and vestibular turns was independently varied by a factor of 1/sqrt(2), 1 or sqrt(2).

• Results I - varied angles, gain: visual = vestibular

Three angles were reproduced qualitatively correctly.

Subjects (six in each condition) were able to reproduce the learned training angles (Fig. 3). In both conditions subjects showed a slight overshoot.

Reproduced angles

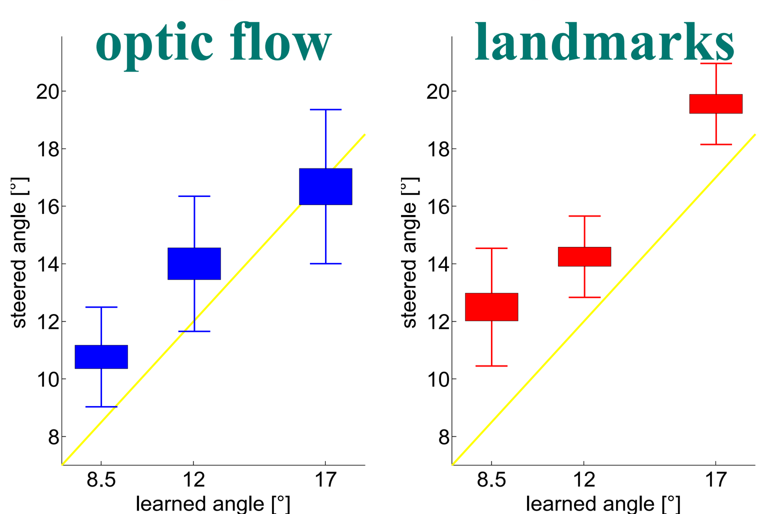


Fig. 3: Turn angles were in general overestimated. Only the conditions with identical gain factor are plotted here.

• Results II - varied gain factors, average turn angle

Optic flow: Dominant effect for bigger gain factor.

When landmark information was provided, subjects followed a purely visual strategy, thus ignoring conflicting vestibular information. With reduced visual information (optic flow only), the modality with the bigger gain factor had a dominant effect on the reproduced turns (Fig. 4).

Scene: Landmarks lead to visually dominated responses.

Visual-vestibular interaction

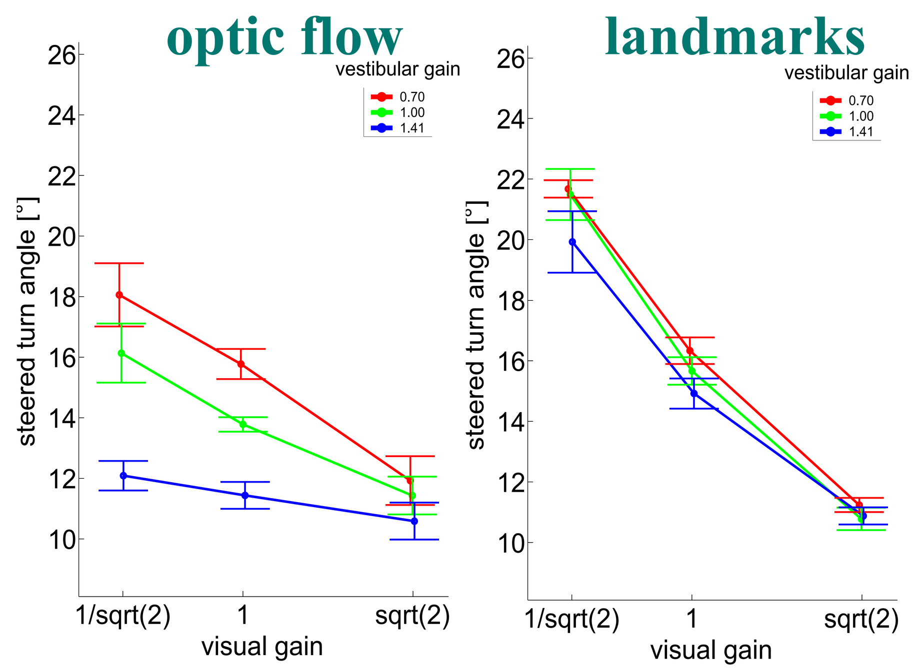


Fig. 4: Motor response (turned angle using a joystick control) for optic flow and landmark condition plotted against the varied visual gain factor. The vestibular gain factor is plotted in different colors. The error bars correspond to the standard error of the mean.

• Model predictions

We identified two main effects: visual and vestibular influence.

Schematic visual and vestibular effects

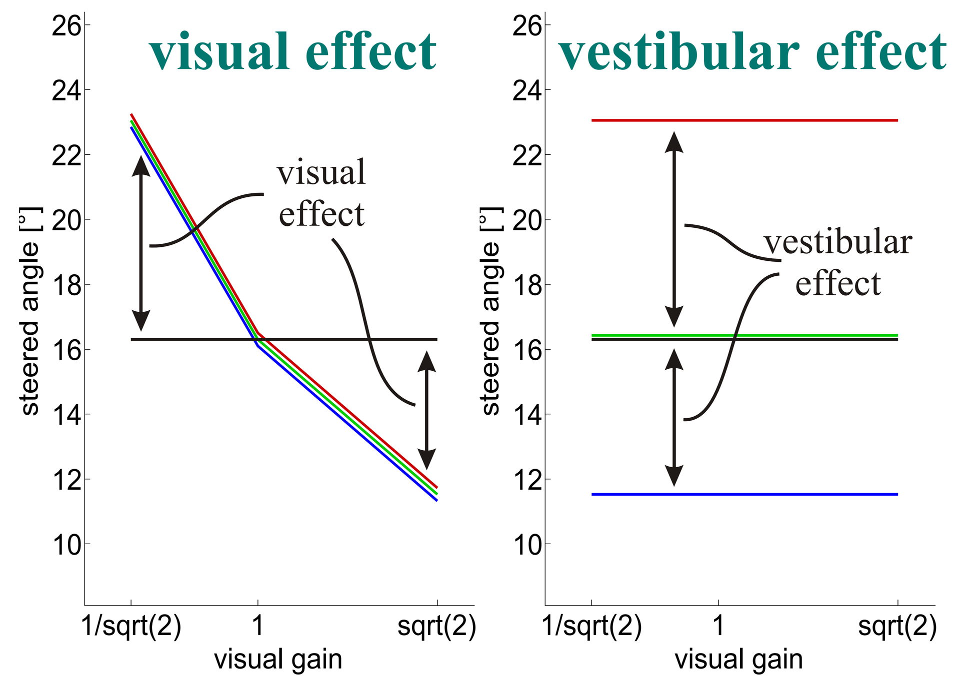


Fig. 5: Here we schematically show both main effects (visual and vestibular) each plotted by means of motor response. For a small gain factor one has to turn further and for a large gain factor one has to stop earlier.

Three simple models are proposed to fit the data pattern.

We propose three simple models for the combination of visual and vestibular information: Additive, multiplicative, and the "max-rule" model.

Integration of visual and vestibular information

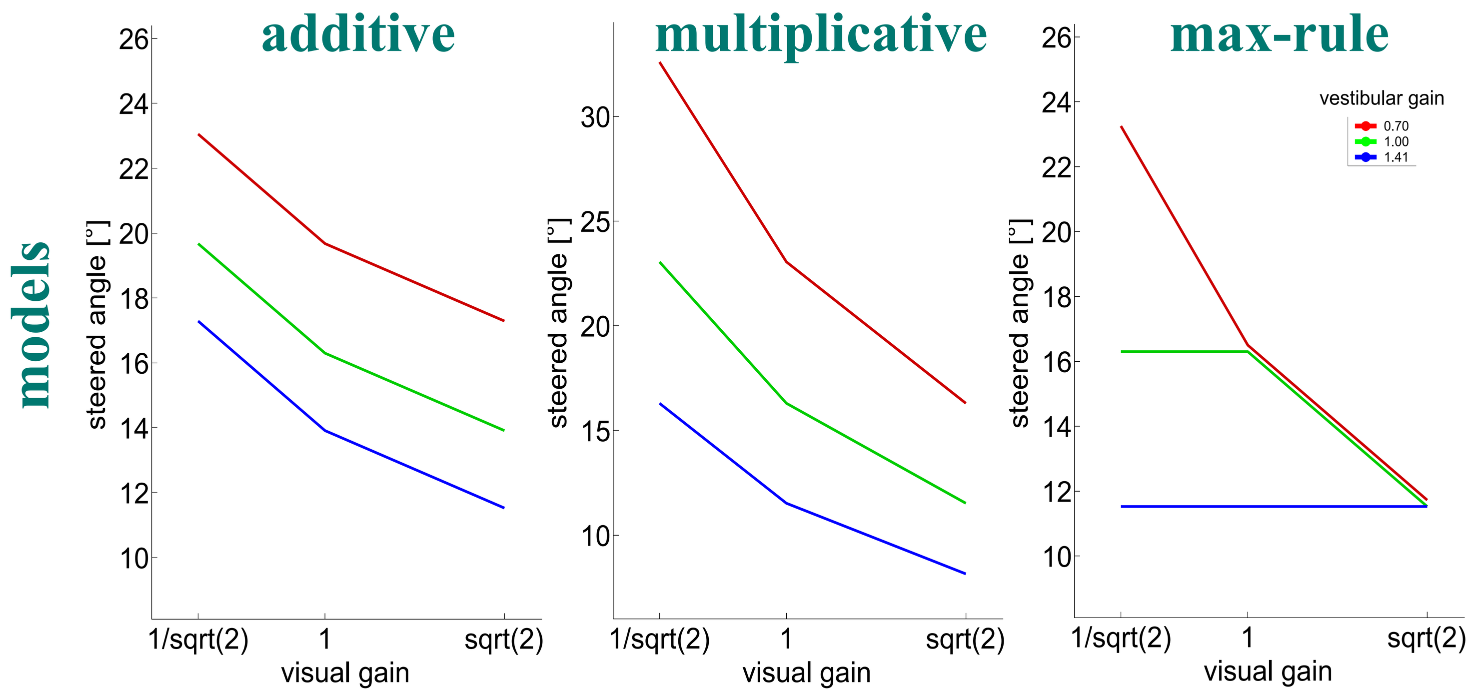


Fig. 6: For the additive model the difference between the graphs is consistent across visual gain. The multiplicative model on the other hand leads to different slopes and offsets. The max-rule model predicts for the bigger gain factor an exclusive effect, e.g., turns are executed until one modality reaches the memorized turn angle. Therefore, the responses are constant for maximal gain factors.

• Data fit for simple models

Only the max-rule explains the quite similar responses for maximal gain factors.

The additive model fits the scene data perfectly.

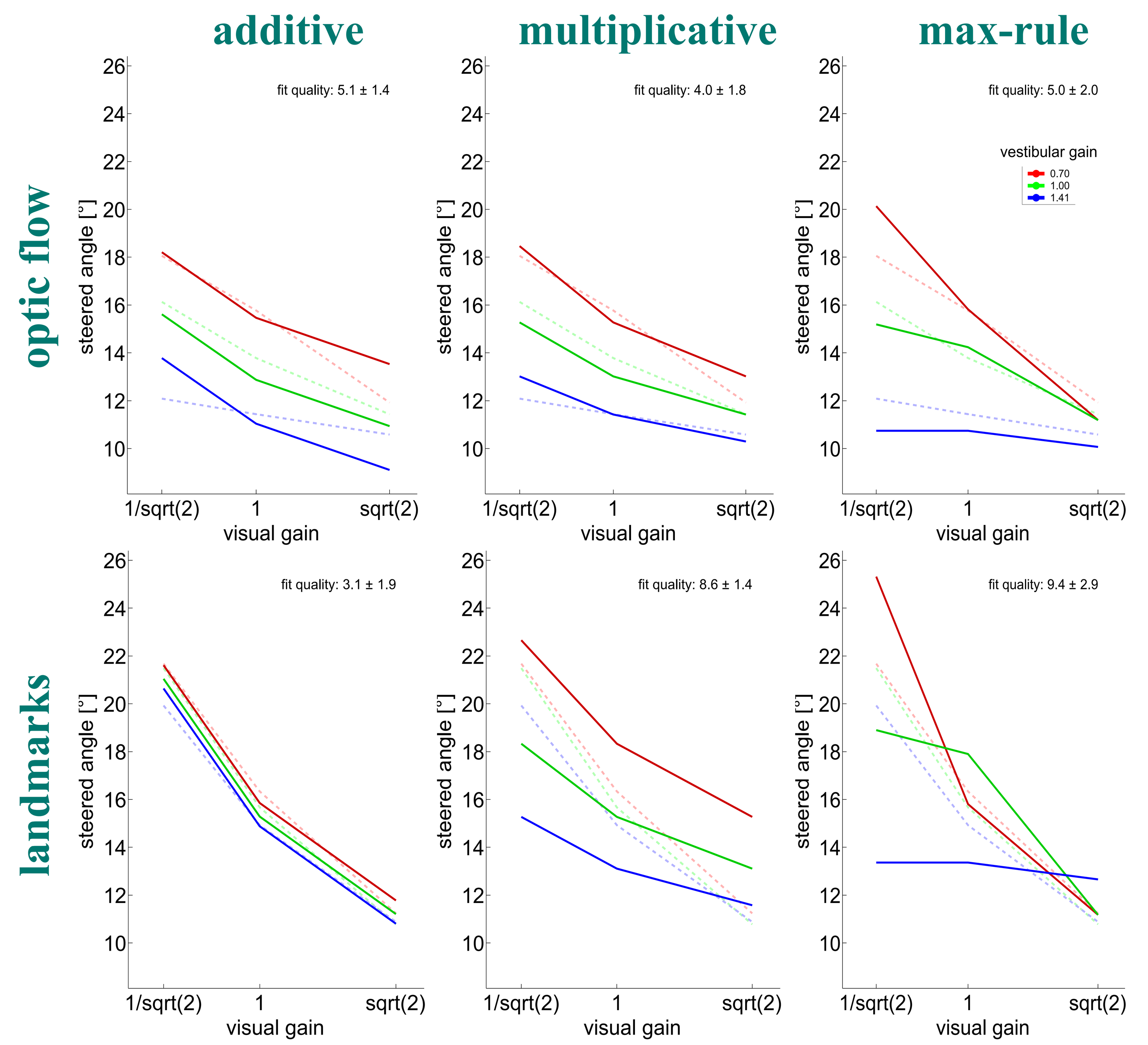


Fig. 7: Data fit for the optic flow (top row) and landmark (bottom row) condition. The average of the fits for the individual subjects is superimposed onto the actual data. The fit quality is the average quadratic difference between model and data.

None of the simple models explain both conditions.

- Qualitatively, the max-rule explains the optic flow data best.
- Only the additive model fits the landmark condition well.
- None of the models can explain both conditions.

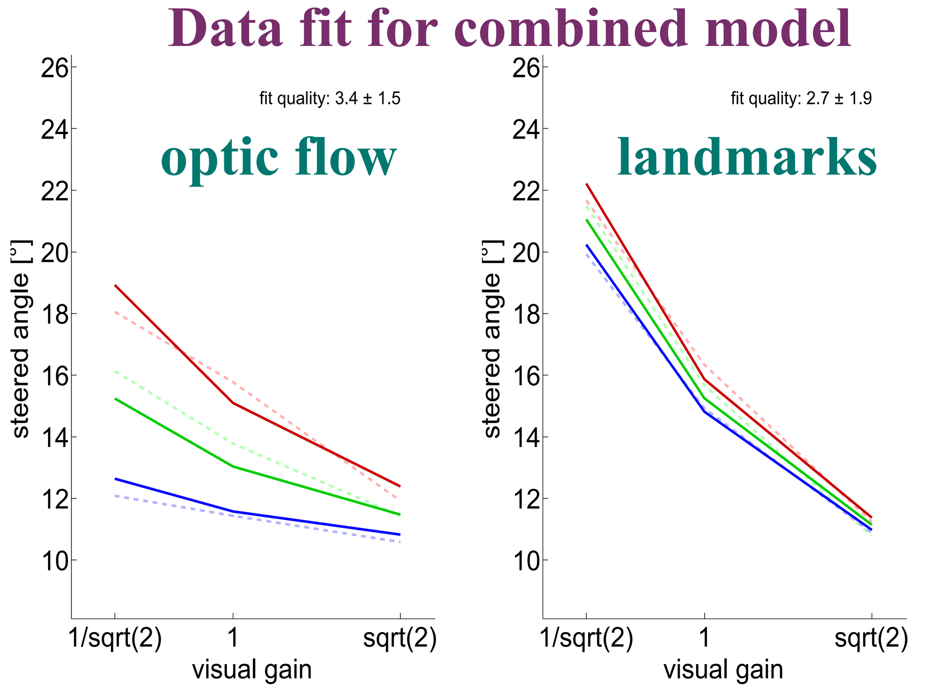
Consequently, the integration of visual and vestibular cues is more complex and may use two different strategies depending on the information available.

• Conclusions

A combined additive and multiplicative model fits the data, but the coefficients may not be meaningful.

Fig. 8: A combined model with an additive and a multiplicative component is used for a final data fit. The fit quality is far better in comparison to the other models, but the coefficients can no longer be interpreted in a meaningful way due to instability problems; small changes in the data result in large differences in the weighting given to the additive or multiplicative component.

Data fit for combined model



The integration of visual and vestibular cues may use two strategies depending on the information available.

- For the integration of optic flow and vestibular cues a combined model with **additive *and* multiplicative** components can fit our data.
- Landmark information is very robust and changes the cue integration to a **visual only strategy**.
- Does the max-rule still apply? Yes, if visual landmarks result in a very high gain or weighting.