Motion parallax links visual motion areas and scene regions

Andreas Schindler *, Andreas Bartels *

Vision and Cognition Lab, Centre for Integrative Neuroscience, University of Tübingen, Tübingen 72076, Germany

ARTICLE INFO

ABSTRACT

When we move, the retinal velocities of objects in our surrounding differ according to their relative distances and give rise to a powerful three-dimensional visual cue referred to as motion parallax. Motion parallax allows us to infer our surrounding's 3D structure as well as self-motion based on 2D retinal information. However, the neural substrates mediating the link between visual motion and scene processing are largely unexplored. We used fMRI in human observers to study motion parallax by means of an ecologically relevant yet highly controlled stimulus that mimicked the observer's lateral motion past a depth-layered scene. We found parallax selective responses in parietal regions IPS3 and IPS4, and in a region lateral to scene selective occipital place area (OPA). The traditionally defined scene responsive regions OPA, the para-hippocampal place area (PPA) and the retrosplenial cortex (RSC) did not respond to parallax. During parallax processing, the occipital parallax selective region entertained highly specific functional connectivity with IPS3 and with scene selective PPA. These results establish a network linking dorsal motion and ventral scene processing regions specifically during parallax processing, which may underlie the brain's ability to derive 3D scene information from motion parallax.

© 2015 Elsevier Inc. All rights reserved.

Introduction

During egomotion the apparent relative motion of objects in our surrounding provides a strong three-dimensional visual cue referred to as motion parallax. Motion parallax provides cues about two important, yet mostly independent pieces of behaviorally relevant information: ego-motion on one hand and scene-layout on the other. Most studies examined the former, and used random dot fields to contrast ego-motion compatible coherent flow, such as expanding, contracting or rotating dot patterns, to moving dot fields inconsistent with ego-motion. These studies found several dorsal visual areas responsive to coherent optic flow; i.e. V6 (Cardin and Smith, 2010; Pitzalis et al., 2010; Fischer et al., 2012b) human MST, CSv (Smith et al., 2006; Wall and Smith, 2008) and several visual areas in the parietal cortex (Bremmer et al., 2001; Konen and Kastner, 2008). A similar network of motion-responsive regions was found in studies examining structure from motion using dot patterns with relative velocities, such as induced by rotating objects (Orban et al., 1999; Murray et al., 2003; Klaver et al., 2008; Beer et al., 2009) and these studies often also reported co-activation in ventral visual areas involved in object processing (Orban et al., 1999; Beer et al., 2009).

However, almost no prior study examined motion parallax in context of its role in providing cues for the spatial depth layout of our surrounding. When walking through a scene, motion parallax and velocity gradients provide direct cues about the distances of surrounding objects and thus about the spatial layout of the surrounding scene. In accord with this, behavioral evidence points to an important role for parallax in navigation (Frenz et al., 2003; Medendorp et al., 2003). Even though it has recently been shown that scene responsive visual regions such as the parahippocampal place area (PPA) (Aguirre et al., 1998; Epstein and Kanwisher, 1998) and the occipital place area (OPA also known in anatomical terms as transverse occipital sulcus (TOS)) (Grill-Spector, 2003; Dilks et al., 2013) but not the retrosplenial cortex (RSC) (Maguire, 2001) respond to visual motion (Korkmaz Hacialihafiz and Bartels, 2015), it is unclear whether these regions also respond to parallax cues.

Here we used fMRI to examine neural processing of motion parallax as induced by lateral observer motion across a visual scene (Fig. 1a). Our stimulus comprised rectangles of different sizes whose relative velocities were chosen such that the observer had the impression of horizontally moving left- and rightwards in front of a depth-layered tunnel scene. Several low-level control conditions matched the relative motion between the rectangles as well as their absolute retinal motion. The controls destroyed the tunnel percept as well as any self-motion cue but preserved all low-level properties. Additionally, we presented a condition involving a moving planar tunnel stimulus without relative motion. This stimulus preserved high-level motion cues compatible to planar motion but lacked parallax cues.

We chose squares of distinct sizes as building blocks of our stimuli rather than random dots for several reasons. First, the response profiles of higher-level category selective areas are still under debate (Grill-Spector and Weiner, 2014; Andrews et al., 2015). It seems however that scene selective regions like PPA are primarily concerned with representation of 3D local space, be that scenes or scene diagnostic

* Corresponding authors at: Vision and Cognition Lab, Centre for Integrative Neuroscience, University of Tübingen, Otfried-Müller-Str. 25, 72076 Tübingen, Germany.
E-mail addresses: andreas.schindler@tuebingen.mpg.de (A. Schindler), andreas.bartels@tuebingen.mpg.de (A. Bartels).
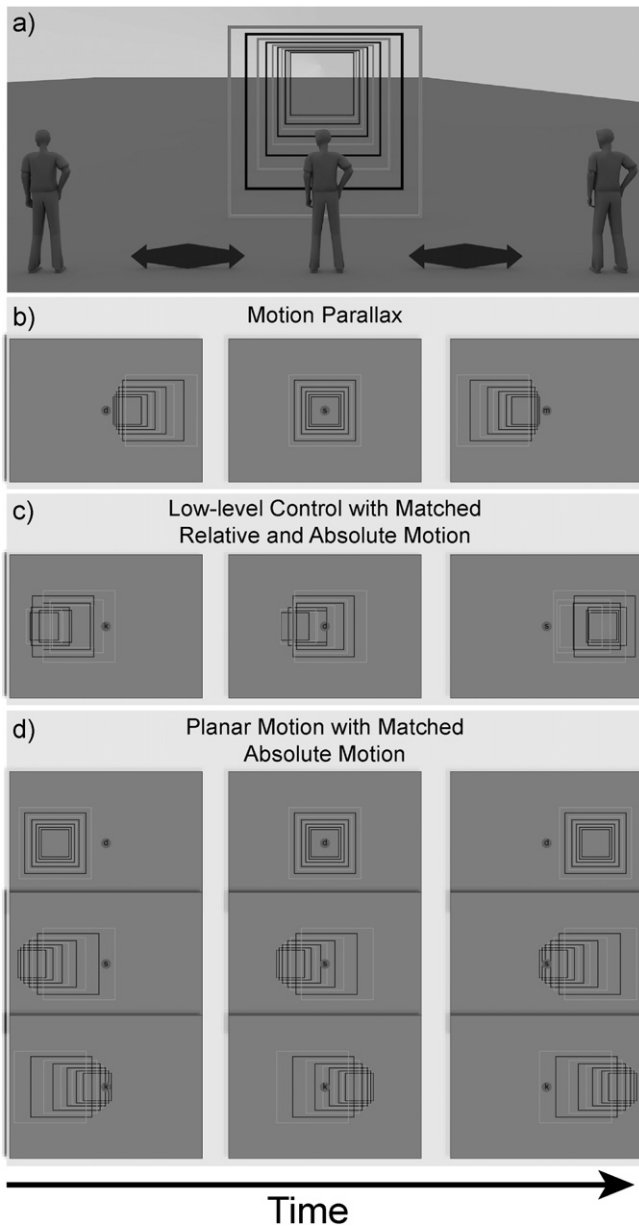
Fig. 1. Stimuli used in the present study. (a) Illustration of the scene layout including the moving observer based on which the motion parallax stimulus was designed. (b–d): Illustration of the three main conditions. (b) Motion parallax condition: the observer had the impression of moving in front of a tunnel in left- and rightwards directions. (c) Low-level control condition: the amount of relative and absolute retinal motion was matched to that of the motion parallax stimulus, but the trajectories of the rectangles were phase randomized such that the tunnel percept vanished. (d) Planar motion control condition: one of three different snapshots of the motion parallax stimulus moved left- and rightwards. Absolute motion was matched to the motion parallax stimulus motion but relative motion between rectangles was absent. Not illustrated are two additional conditions that used phase randomized motion as in (c), yet that matched only relative or absolute motion. Finally, there were two static conditions: the static scene condition presented the tunnel snapshots shown in (d), the static control condition presented snapshots of (c).

objects (Henderson et al., 2008; Mullally and Maguire, 2011; Harel et al., 2013). In accord with this, our parallax stimulus comprises a high-level stimulus defining a three-dimensional local space. However, in contrast to natural scene or object images our stimulus was well controllable in terms of low-level visual motion features as well as for static high-level Gestalt cues (i.e. the 'tunnel' percept). We also favored our stimulus over random dot kinetograms, as with the latter we would not have been able to control high-level effects such as static scene cues. Finally, recent evidence shows that cardinal orientations and right angles, such

as present in our stimuli, provide optimal stimulation for scene responsive regions (Nasr and Tootell, 2012; Nasr et al., 2014).

We found that only parietal regions IPS3/4 and an occipital region, abutting scene responsive area OPA laterally, responded specifically to motion parallax. Connectivity analyses revealed two highly selective functional links during motion parallax — one between IPS3 and the occipital parallax responsive region, and one between the latter and scene selective PPA.

Materials and methods

Observers

Fifteen observers took part in this study (nine males, six females, age 22–34 years). All gave written informed consent prior to participation. The study was approved by the ethics committee of the University Clinic Tübingen.

Stimuli and paradigm

The main experiment was a block design in which we presented 49 stimulation blocks per run, made up of 7 repetitions of 7 conditions, presented in a counter-balanced and pseudo-randomized fashion. Each stimulus block lasted 18 s. Five runs were collected for each observer.

All stimuli were composed of the same set of 10 square-outlines that differed in their sizes (edge lengths: 8.2°, 7.0°, 6.1°, 5.4°, 4.8°, 4.4°, 4.0°, 3.7°, 3.4°, 3.2°). The square outlines alternated in gray and white (see Fig. 1). A main motion parallax stimulus evoked the percept of a 3D tunnel scene, and six control stimuli were designed to control for different high- and low-level stimulus aspects of the main parallax stimulus while not giving rise to a parallax percept. This led to a total of 7 conditions described below in detail.

Stimuli were gamma corrected and projected (resolution: 1024 × 768) on a screen positioned behind the observers' head, viewed at 92 cm distance and spanning 22 × 16 visual degrees. Observers viewed the stimuli binocularly via a mirror mounted on the head coil. The paradigm was programmed with Psychtoolbox 3 (http://psychtoolbox.org/), and run on MATLAB 2010 (MathWorks) on a Windows PC.

Motion parallax condition

The relative velocities of the rectangles were chosen such that the observer had the impression to move laterally in front of a tunnel in left- and rightwards directions, while fixating a point behind the tunnel (Fig. 1b). The screen positions, sizes and velocities of the rectangles were calculated using a three-dimensional virtual reality model with the following parameters in virtual space: each square had an edge-length of 5 m, and was 5.4 m distant to the next square. The observer was 35 m in front of the nearest square, and fixated in infinity. The observer moved laterally within a range of 10.7 m, with a maximal velocity of 8 m/s (28 km/h) and a sinusoidal velocity profile. When projected onto the screen, this led all squares to move in-phase with the others, yet covering different distances and moving at different velocities. The front-most square moved within a range of 17.4 visual degrees, with a peak-velocity of 13.6 °/s, the rear-most square moved within a range of 6.8 deg, with a peak-velocity of 5.3 °/s.

We derived six control conditions from this main parallax stimulus.

Control conditions for relative and absolute motion

To control for visual low-level features, we considered two main factors: absolute velocity and relative velocity. Absolute velocity was defined as the mean velocity of all rectangles between subsequent frames across the entire stimulus block and weighted by the rectangles' edge length (yielding a measure for average pixel-velocity on the screen). Relative velocity of a given rectangle was defined as the mean pairwise velocity difference of that rectangle compared to each of the others, averaged across the stimulus block. This was averaged across