JNeurosci
THE JOURNAL OF NEUROSCIENCE

Research Articles: Systems/Circuits

Behavioral evidence and neural correlates of perceptual grouping by motion in the barn owl

Yael Zahar[1], Tidhar Lev-Ari[1], Hermann Wagner[2] and Yoram Gutfreund[1]

[1]Department of Neuroscience, The Ruth and Bruce Rappaport Faculty of Medicine and Research Institute, The Technion, Haifa, Israel.
[2]Department of Zoology and Animal Physiology, Institute for Biology II, RWTH Aachen University, Aachen, Germany.

DOI: 10.1523/JNEUROSCI.0174-18.2018

Received: 21 January 2018

Revised: 7 June 2018

Accepted: 11 June 2018

Published: 2 July 2018

Author contributions: Y.Z., T.L.-A., H.W., and Y.G. designed research; Y.Z. and T.L.-A. performed research; Y.Z. and T.L.-A. analyzed data; Y.Z., T.L.-A., and Y.G. wrote the first draft of the paper; Y.Z., T.L.-A., H.W., and Y.G. edited the paper; Y.G. wrote the paper.

Conflict of Interest: The authors declare no competing financial interests.

This work was supported by a grant from the Israel Science Foundation to YG, a grant from the DFG to HW and YG and a research grant from the Rappaport Institute for Biomedical Research to YG. The authors declare no competing financial interests.

Corresponding author: Prof. Yoram Gutfreund, The Rappaport Faculty of Medicine, Technion, Bat-Galim, Haifa 31096, Israel, Email: yoramg@tx.technion.ac.il

Cite as: J. Neurosci ; 10.1523/JNEUROSCI.0174-18.2018

Title: Behavioral evidence and neural correlates of perceptual grouping by motion in the barn owl

Abbreviated title: Perceptual grouping in barn owls

Yael Zahar[1]*, Tidhar Lev-Ari[1]*, Hermann Wagner[2] and Yoram Gutfreund[1]

* Equal contribution

1 Department of Neuroscience, The Ruth and Bruce Rappaport Faculty of Medicine and Research Institute, The Technion, Haifa, Israel.
2 Department of Zoology and Animal Physiology, Institute for Biology II, RWTH Aachen University, Aachen, Germany.

Corresponding author:
Prof. Yoram Gutfreund
The Rappaport Faculty of Medicine, Technion,
Bat-Galim, Haifa 31096,
Israel
Email: yoramg@tx.technion.ac.il

Number of pages: 34

Number of figures: 7

Number of words:

Abstract 235

Introduction 563

Discussion 1552

Acknowledgment: This work was supported by a grant from the Israel Science Foundation to YG, a grant from the DFG to HW and YG and a research grant from the Rappaport Institute for Biomedical Research to YG. The authors declare no competing financial interests.

Abstract

Perceiving an object as salient from its surround often requires a preceding process of grouping the object and background elements as perceptual wholes. In humans, motion homogeneity provides a strong cue for grouping, yet, it is unknown to what extent this occurs in non-primate species. To explore this question, we studied the effects of visual motion homogeneity in barn owls of both genders, at the behavioral as well as the neural level. Our data show that the coherency of the background motion modulates the perceived saliency of the target object. An object moving in an odd direction relative to other objects attracted more attention when the other objects moved homogenously compared to when moved in a variety of directions. A possible neural correlate of this effect may arise in the population activity of the intermediate/deep layers of the optic tectum. In these layers the neural responses to a moving element in the receptive field (RF) were suppressed when additional elements moved in the surround. However, when the surrounding elements all moved in one direction (homogeneously moving) they induced less suppression of the response compared to non-homogenously moving elements. Moreover, neural responses were more sensitive to the homogeneity of the background motion than to motion-direction contrasts between the receptive field and the surround. The findings suggest similar principles of saliency-by-motion in an avian species as in humans, and show a locus in the optic tectum where the underlying neural circuitry may exist.

Significance statement

A critical task of the visual system is to arrange incoming visual information to a meaningful scene of objects and background. In humans, elements that move homogeneously are grouped perceptually to form a categorical whole object. We discovered a similar principle in the barn owl's visual system, whereby the homogeneity of the motion of elements in the scene allows perceptually distinguishing an object from its surround. The novel findings of these visual effects in an avian species, which lacks neocortical structure, suggest that our basic visual perception shares more universal principles across species than presently thought, and shed light on possible brain mechanisms for perceptual grouping.

Introduction

For humans, an object that is different from a uniform surrounding, such as a vertical bar embedded in an array of horizontal bars, tends to perceptually "pop out" (Treisman and Gelade, 1980; Wolfe and Horowitz, 2004). Perceiving an object as popping out relative to its neighbors often requires a preceding process of grouping in which the object and the background are grouped into perceptual wholes (Duncan and Humphreys, 1989; Kingstone and Bischof, 1999). In the early 20th century, Gestalt theorists attempted to explain factors that govern this organization, defining a set of laws of perceptual organization specifying how we construct simple individual elements into global wholes (Spillman and Ehrenstein, 2003).

Neurophysiological studies have sought for the neural correlates of Gestalt-like figure-ground-segregation mostly in primates (Lamme, 1995; Lee et al., 2002; Zhou et al., 2000; Zipser et al., 1996; Qiu and von der Heydt, 2005; Burrows and Moore, 2009), showing neurons that process contextual stimuli, which appear outside the classic receptive field (RF) of the cell but influence its response to a stimulus inside the RF. However, we hypothesize that the Gestalt principles derived from humans are a manifestation of visual mechanisms that evolved early in evolution as a means of breaking camouflage. Therefore, we expect to find similar principles in a wide range of animal species, beyond primates and mammals.

To explore this hypothesis, we studied an avian species, the barn owl, and focused on motion pop-out stimuli both at the perceptual and neuronal levels. It has been shown that this animal expresses pop-out perception for orientation and luminance stimuli (Orlowski et al. 2015; Orlowski et al. 2018). At the physiological level, it has been shown that tectal neurons in the barn owl respond stronger to an object moving in the RF when objects outside the RF move in an opposite direction compared to when the surrounding objects move in the same direction (Zahar et al., 2012). This is consistent with the proposed role of the optic tectum in saliency mapping (Mysore and Knudsen, 2011; Gutfreund, 2012). However, such a modulation from the surround does not necessarily imply perceptual grouping, but can arise from simple motion-contrast sensitivity between the RF and its surround (Hegde and Felleman, 2003). The pop-out effect in its classical interpretation involves global perception of a homogenous surround as a separate whole (Duncan and Humphreys, 1989; Hochstein and Ahissar, 2002). Thus, to address whether barn owls can utilize the homogeneity of motion for saliency mapping, it is necessary to use a paradigm that distinguishes between motion contrast and homogeneity.

To achieve this, we designed a paradigm in which responses are compared between when background elements move homogenously but in a direction that is less contrasting the direction of the target to when the background elements move non-homogeneously, in directions that are more contrasting the target's direction. We first demonstrate in behavioral experiments that an object moving relative to homogenously moving background elements is perceived by barn owls as being more salient compared to an object moving relative to a non-homogenous motion, consistent with perceptual grouping for figure-ground-segregation. In complementary neurophysiological experiments, we found that tectal neurons in the intermediate/deep layers similarly tend to respond preferentially to targets embedded in homogenous background motion compared to non-homogenous motion. Importantly, neural sensitivity to contrast and homogeneity matched the behavioral sensitivity of barn owls, suggesting a neural correlate of perceptual grouping by motion.

Materials and Methods

Animals

Seven adult barn owls (*Tyto alba*) were used in this study: five owls in electrophysiological experiments and two owls in behavioral experiments. The owls were hatched and raised in captivity, and kept in aviaries equipped with perching spots and nesting boxes. All experiments were performed in Haifa. All procedures were in accordance with the guidelines and approved by the Technion Institutional Animal Care and Use Committee. Surgical procedures were performed under isoflurane anesthesia, and in all recording sessions the animals were sedated with mixture of oxygen and nitrous oxide. During recording sessions no painful procedures were carried out.

Surgical procedures

Owls were prepared for repeated electrophysiological experiments in a single surgical procedure. First, the owl was anesthetized with isoflurane (2%) and nitrous oxide in oxygen (4:5). Lidocaine (lidocaine HCl 2% and epinephrine) was injected locally at the incision site. A craniotomy of 1 cm diameter was performed 0.6 cm lateral to the midline and 1.7 cm anterior from the anterior tip of attachment of the dorsal neck muscles to the skull. Then a recording chamber was cemented to the skull (Unifast dental cement mixed with cyanoacrylate adhesive) over the craniotomy. The chamber was filled with chloramphenicol ointment (5%) and sealed with a cap. After surgery, the animal was left to recover overnight in an individual cage and then released back to its home cage.

Electrophysiological recordings

Before each electrophysiological session, the owl was moved to an individual cage without food overnight. At the beginning of each electrophysiological session, the owl was anesthetized briefly with isoflurane (2%) and nitrous oxide in oxygen (4:5). Once anesthetized, the owl was wrapped in a soft leather jacket and positioned in a stereotaxic apparatus inside a double-walled, sound-attenuating booth (internal size 2.05 x 1.7 x 1.95 m). The head was bolted to the apparatus after aligning the visual axis using retinal landmarks ((Reches and Gutfreund, 2008). After the bird was fixed, the isoflurane was removed and the bird was maintained on a steady mixture of nitrous oxide and oxygen (4:5). Small weights were attached to the feathers on the owl’s eyelids to maintain eye opening throughout the recording session. The nictitating membrane was not restrained, allowing for spontaneous moistening of the cornea. The head chamber was opened, and a tungsten, parylene-coated or glass-coated micro electrode (0.5-1.5 MΩ; Alpha-Omega, Nazareth, Israel) was driven using a motorized manipulator. Since eye movements in barn owls are limited to a range smaller than ±2° (du Lac and Knudsen, 1990), we did not immobilize or control eye movements. The recorded electrical signal was amplified, digitized, and filtered

(313-5,000 Hz) using the AlphaLab SnR system (Alpha Omega, Nazareth, Israel). In each experiment, a threshold was set online to select the larger units in the recording sites and isolate action potentials from a small cluster of neurons (multi-unit recording). At the end of each recording session, the recording chamber was treated with chloramphenicol ointment (5%) and closed. The owl was then returned to its home flying cage.

Identification of the location of the recording site was based on stereotaxic coordinates and on the expected physiological properties: the OT was recognized by characteristic bursting activity and spatially restricted visual and auditory receptive fields. Position within the OT was determined based on the location of the visual receptive field (RF). Recordings were taken from all layers of the OT. The intermediate layers of the OT were located beneath the bursty layers and identified based on a transition from bursty activity to regular firing (Knudsen, 1982) (Netser et al., 2010). The electrode was advanced in small steps to search for sites with clear units and visual responses. Recording sites were separated by at least 300 μm. All recording sites were from the anterior part of the OT having visual receptive fields between left and right 20° and up and down 20° relative to the center of the visual field.

Visual stimuli

The visual stimuli were computed in Matlab using Psych Toolbox extension (Brainard, 1997) and either displayed on a computer screen (17-in. LCD screen, at a refresh rate of 76 Hz) for the behavioral experiments or projected (refresh rate 72 Hz, XD400U; Mitsubishi, Japan) on a calibrated screen inside the sound attenuating chamber for the electrophysiological experiments (screen size 170 cm x 140 cm, 1.5 meter away from the owl). The projector was positioned outside the chamber, projecting the image through a double-paned glass window. Visual stimuli were dark dots presented on a gray background (luminance of background screen was ~20 cd/m^2 and luminance of dots was ~8 cd/m^2). In each recording site, we first estimated the visual RF by moving a visual stimulus on the screen (a dark dot about 1° in diameter) and listening to the neural discharge. The point that elicited the strongest neural discharge was chosen as the center of the RF. Typical width of RF in these layers was estimated in a previous study to be about 6^o-10^o (Zahar et al., 2012). After estimating the RF center, a test paradigm was applied in which a dark dot (about 1° in diameter) was positioned at the center of the RF (the target). The dot was embedded in an array of identical dots (the distractors) equally spaced at 10° intervals (see insets in Figs. 1 and 2). In several experiments the dots in the rectangle surrounding the target dot were omitted, thus increasing the distance between the target dot and its surrounding dots (see insets in Figs. 6 and 7).

In each trial, the initial frame of the dots array was displayed static for 1.5 s and then the target, sometimes together with the background elements, moved to the right for 600 ms on a straight horizontal line for a distance of 2.9^o. At the end of the movement, the last frame of motion was maintained static on the screen for 500 ms until the initiation of the next trial. In a previous study, no average difference was observed in the response properties of tectal neurons between leftwards and rightwards motions, and only weak modulations by direction were observed (Zahar et al., 2012). Therefore, to reduce the experimental trials, in this study we only studied responses to a target moving rightwards. The target dot (in the RF) was embedded in one of six possible contexts relative to its background dots: 1) the singleton condition in which the target moved rightwards while the distractors were static; 2) the uniform condition in which the target and distractors moved coherently rightwards; 3) the offset180° condition in which the distractors moved coherently in the opposite direction to target; 4) the offset90° condition in which the distractors moved coherently upwards (orthogonal to the target's rightwards movement); 5) the offset270° condition in which the distractors moved coherently downwards (orthogonal to the target's motion); and 6) the mixed condition in which the target moved rightwards while each of the distractors moved arbitrarily in one of the three directions – leftwards, upwards or downwards (see insets in Fig 2 for a graphical illustration of the six conditions). Offset135^{o} and offset -135^{o} conditions were also displayed (see insets in Fig. 6). In each test, conditions were interleaved randomly and repeated 15 times. In the mixed conditions, the dots moving leftwards, upwards or downwards were randomly reallocated in every trial.

Behavioral experiments

Two hand-raised barn owls (DO and DK) females of about 1 year of age were used to measure the behavioral responses. For the experiment, the owl was placed on a perch in a darkened room with a computer screen that was facing upwards in a pecking range below the owl. To track the owl's gaze, a lightweight wireless video camera (Owl-Cam, 30 frames per second, ~60º view angle) was mounted on the owl's head. The camera was self-assembled from a miniature micro-camera combined with a video broadcasting chip (900 MHz) and a rechargeable lithium-polymer battery (weight together with mounting unit was about 10.5 g). The Owl-Cam was attached to the head using a 3D printed attachment unit glued to the skull with dental cement. The unit was designed to maintain a fixed and reproducible relationship between the Owl-Cam and the head. Since barn owls lack substantial eye movements, a head-fixed camera can provide a reliable estimation of the owl's gaze position (Ohayon et al., 2008). Initially, the position of the gaze center (functional fovea) was calibrated for each owl by allowing the owl to fixate on multiple targets on the screen. The average position of targets on the video frame results in a single position corresponding to the point of gaze (Harmening et al., 2011; Hazan et al., 2015). Owls were pre-trained in a previous project (Lev-Ari and Gutfreund, 2017) to initiate a trial by fixating on a red dot on the center of the screen, waiting until it disappears and then search for Gabor patches on the screen. Thus, the owls were well-trained to initiate the trials and search the screen for rewarded targets, but they had never been trained for the specific task and stimulus at hand. In the current task, after fixation was achieved, the red fixation dot disappeared followed by one of the four stimulus conditions: singleton, offset90°, offset180° and mixed. All four conditions were interleaved randomly. In each test, the odd target was located randomly at one of the four positions corresponding to 4 cm above, 4 cm below, 4 cm to the right and 4 cm to the left of the screen center (Fig. 1A, upper panel; see also Videos 1 and 2). Owls typically searched the screen from a distance of about 25 cm. As in the electrophysiological experiments, dots were moved 1.3 cm on the screen (corresponding to a view angle of about 3^o) for a duration of 600 ms. However, unlike in the electrophysiological experiments, the movement was repeated continuously on the screen for up to 30 sec (once a dot reached the end of the path, it disappeared and reappeared immediately at the motion starting point; see Videos 1 and 2). Rewards (small chunks of chicken meat) were fed manually to the owl by the experimenter sitting behind a curtain. Food was given with forceps a few cm above the screen. The owls were rewarded about every second trial for initiating the trials and seeking the screen, but reward was not associated with a specific target. Owls performed about 20-40 trials a day. We tested the owls on consecutive days until they reached 35 repetitions of each condition (a total of 140 repetitions per owl for all four stimuli).

Data analysis and statistical testing

Owl-Cam data were analyzed using a custom Matlab GUI. Typically, owl search behavior consisted of stable fixation periods for 0.5-4 s durations terminated by rapid head saccades to a new fixation point (Videos 1 and 2, and (Hazan et al., 2015)). To estimate the fixation target, we defined a circular area with a radius of 50 pixels around the center of gaze as estimated from the calibration process described above. This corresponds to a viewing angle of about 8^{o}. Any dot within this area maintained for 10 consecutive frames was considered to be a target of gaze. The relatively large window was chosen to account for the relatively large area centralis of barn owls (Wathey and Pettigrew, 1989) and to include errors that may arise from differences in distance and gaze angle to screen plane. In each trial, the time from stimulus onset to the first gaze on the target (search time) and number of head saccades to reach the target were registered. In addition, for control, the number of head saccades and time to the first gaze on the dot opposite the target were registered (for example, if the target was above the center, the control dot was below the center, etc.). A trial was considered a success if during the 30 s window of stimulus presentation the owl fixated on target. Differences in success rates between conditions were tested using Mann Whitney test (non-paired). Differences between successes rates to target versus control were tested using Wilcoxon test (paired). In the search time (ST) analysis, we discarded trials with search times slower than three times the standard deviation of that test. This led to the exclusion of 5% and 6.5% of the trials for DO and DK, respectively. STs were tested using one way ANOVA with post-hoc Tuckey test.

Unit responses to a visual stimulus were calculated as the number of spikes in a given time window after stimulus onset minus the number of spikes during the same period of time immediately before stimulus onset (baseline activity). The duration of the time window for spike count was 600 ms, starting from the onset of motion. To observe the time course of the response, we generated post-stimulus time histograms (PSTHs) with 15 ms time bins. PSTHs were normalized to the maximum value achieved in each experiment and averaged across the population. For graphical display, curves were smoothed (5-point sliding average). The standard errors of the mean were depicted as the width of the PSTH curves. Differences between population responses were analyzed using one-way ANOVA and post-hoc Tuckey tests. To quantify the contextual modulation, we calculated the modulation index (MI) as follows: MI = $(R_{context1} - R_{context2})/(R_{context1} + R_{context2})$, where $R_{context1}$ is the response to the target motion in one surrounding context, and $R_{context2}$ is the response to the same target motion in a different context. Positive values of this index indicate a preference for context 1 over context 2. Distribution of MI indices was tested using a binomial sign test.

Results

Behavioral experiments

Behavioral measurements were conducted in owls spontaneously viewing displays of dot arrays on a computer screen. In all conditions, a single rightwards moving dot served as the target that could appear in one out of four locations (see Materials and Methods and Fig. 1A). Following trial initiation (fixation of a red dot), the barn owls typically scanned the computer screen and surrounding room with abrupt head saccades (Videos 1 and 2). If, during the 30 sec after the trial initiation, the target dot appeared within the gaze point window (see Materials and Methods), the trial was considered to be a successful trial in which the time (search time) and number of head saccades to gaze-on-target were registered. In the singleton condition, when only the target was moving while the rest of the distractors were stationary, both owls acquired the target in all trials (success rate of 1; Fig. 1B). Success rates dropped for the offset180^o condition (0.78 and 0.93 in DO and DK, respectively) and the offset90^o condition (0.76 and 0.73 in DO and DK, respectively), and further decreased for the mixed conditions to 0.31 and 0.4 in DO and DK, respectively (Fig. 1B). The success rate in the mixed condition was significantly smaller than in the other three conditions in both owls (Mann-Whitney test, $Z = -5.870$ $p < 0.001$, $Z = -5.249$ $p < 0.001$, singleton vs. mixed for DO and DK, respectively; $Z = -3.890$ $p < 0.001$, $Z = -4.664$ $p < 0.001$, offset180^o vs. mixed for DO and DK, respectively; $Z = -3.633$ $p < 0.001$, $Z = -2.696$ $p = 0.007$, offset90^o vs. mixed for DO and DK, respectively). In each of the four conditions, the success rate for fixating on a control dot (the dot opposite the target) was also measured (white columns in Fig. 1B). Success rates for reaching control dots were significantly lower in all conditions from reaching the target (Wilcoxon sign rank test; $Z = -5.657$ $p < 0.001$; $Z = -4.600$ $p < 0.001$ for DO and DK, respectively (singleton); $Z = -4.264$ $p < 0.001$; $Z = -5.568$ $p < 0.001$ (offset180^o) $Z = -5.477$ $p < 0.001$, $Z = -4.600$ $p < 0.001$ (offset90^o); $Z = -1.897$ $p = 0.029$, $Z = -3.051$ $p = 0.001$ (mixed) for DO and DK, respectively). These data suggest that the target in the mixed condition is perceived to be less salient to the owls compared to the homogenous conditions of both 180^o and 90^o offsets. However, even in the mixed conditions, the target attracts more gazing compared to the control targets (Fig. 1B, blue columns compared to corresponding white columns).

The perceived saliency of the target is expected to be reflected also in the speed on which the target is fixated. Therefore, we analyzed search times (STs). Figure 1C shows the average ST for all four stimulus conditions. STs varied significantly in both owls, with the shortest average STs for the singleton conditions and the longest average STs for the mixed conditions (ANOVA, N=33, 26, 25, 11; $F_{3,91}$=6.552, p<0.001 for DO and N=32 ,31, 24, 14, $F_{3,96}$=10.427, p<0.001 for DK). The STs in the mixed conditions were significantly longer in owl DO than the other conditions and in owl DK significantly longer than the singleton and offset180^{o} conditions (post-hoc Tuckey test, p < 0.001, p = 0.037, p = 0.004 for owl DO and p < 0.001, p = 0.003, p=0.38 for owl DK).

Figure 1D shows the cumulative distributions of the number of fixations (head saccades) to reach the target. For both owls, the curves in the singleton conditions (black curves) were shifted leftwards and upwards compared to the rest of the curves. In more than 70% of the trials, the singleton target was reached in less than five saccades (in both owls). On the other hand, in the mixed conditions, five saccades to target were observed in less than 20% of the trials. The curves representing the offset180^{o} and offset90^{o} conditions were in between the singleton and mixed conditions, indicating fewer saccades performed before reaching the target in the homogenous conditions compared to the mixed conditions. Note that control curves (number of saccades to reach the control dot) in all cases were below the mixed condition curves (dashed lines in Fig. 1D). In summary, the results show that oddly moving dots were perceived to be more salient compared to dots moving coherently with other dots (control dots). However, the target dot attracted gaze faster, with less preceding saccades and more often when the background elements moved homogeneously compared to when they moved incoherently.

Electrophysiological experiments

In the first experiment, the responses of 99 multiunit recording sites from two owls were analyzed, 46 from the superficial bursty layers of the OT and 53 from the intermediate/deep layers (Ramon Y Cajal layers 10-14; (Knudsen, 1982; Netser et al., 2010). In each recording site, the position of the RF was estimated, and the target dot was then positioned at about the center of the RF. We tested six conditions randomly interleaved across trials: the four conditions tested in the behavioral experiments, as well as a uniform condition and an offset270^o condition (see insets in Fig. 2). In this study we analyzed population data and report modulations at the population level. We therefore, do not expect that restricting data to single-unit recordings would have qualitatively changed the results. Noteworthy, in previous studies where we isolated single units in the OT and compared with multiunits, the population results did not differ qualitatively (Reches and Gutfreund, 2008; Zahar et al., 2012; Wasmuht et al., 2017).

An example of the responses of a single recording site from the intermediate/deep layers is shown in Figure 2. In all conditions the neurons responded above base line to the motion of the dot within the RF. However, the response was highly modulated by the background context. Maximal responses were achieved in the singleton (50.3 spikes per stimulus) and offset180^o (39.3 spikes per stimulus) conditions (Figs. 2A and C, respectively). The response in the uniform condition was considerably smaller (13.8 spikes per stimulus, Fig. 2B). The responses in the upward (25.66 spikes per stimulus) and downward (28.06 spikes per stimulus) background motion (Fig. 2D and E respectively) were smaller compared to the offset180^o condition (*t*-test, n=15, p<0.001 for upwards and p=0.015 for downwards motion), however, they were larger than the uniform condition (*t*-test, n=15, p<0.001 for upwards and p<0.001 for downwards motion). Thus, this site responded mostly to motion contrast of 180^o, less to 90^o and the least to zero contrast. However, the response in the mixed condition

(17.93 spikes per stimulus) was smaller compared to the coherent upward and downward motions (compare Fig. 2F to Figs. 2D and 2E; *t*-test, n=15, p=0.017 for upwards and p=0.0105 for downwards motion). This single site example was typical of the population pattern shown below.

To compare the responses at the population level, we averaged the PSTHs from all recording sites in the intermediate/deep layers (n = 53). First, we compared the population responses in the singleton, uniform, offset180^o, offset90^o and offset270^o conditions (Fig. 3A). The singleton condition gave rise to the maximal average response followed by a lower, albeit still prominent, average response in the offset180^o condition. The uniform motion resulted in a dramatic reduction in average response strength (about 75% attenuation of peak response from the offset180^o context). This agrees with previous findings that tectal neurons robustly prefer opposing motion over uniform motion (Frost and Nakayama, 1983; (Zahar et al., 2012). The average PSTHs to the target motion embedded in a coherent upward or downward motion array were positioned in between the responses for opposite and uniform motion conditions (ANOVA, N=53, $F_{3,208}$=38, $p < 0.001$; post-hoc Tuckey test $p<0.001$). Next, we compared the population responses in the mixed condition with the offset90^o and offset270^o conditions (Fig. 3B). The average response to the mixed conditions was below the average responses to the offset90^o and offset270^o conditions and above the average response to the uniform condition (ANOVA, N=53, $F_{3,208}$=9.5, $p < 0.001$; post-hoc Tuckey test, p=0.02 and p=0.009 mixed compared to offset 90^o and offset270^o conditions, p=0.003 mixed compared to uniform condition). Note, in all stimulus conditions, the initial response was followed by a decrease of the average firing rates below the baseline level, suggesting an effect of inhibition. The crossing of the response curve to below the baseline was earliest for the uniform condition, followed by the mixed conditions, and latest for the offset90^{o} and offset270^{o} conditions (Fig. 3B).

For each of the four contrasting conditions between the RF motion and the surrounding motions (mixed, offset270^{o}, offset90^{o} and offset180^{o}), a modulation index (MI, see Materials and Methods) was calculated to quantify by how much the responses deviated from responses in the uniform condition. Figure 3C depicts the MIs for the offset270^{o} versus offset90^{o} condition. Both resulted with mostly positive MIs, indicating a preference to a target moving oddly in a direction orthogonal to the direction of the background elements compared to a target moving uniformly with its surrounding elements. Dots were distributed evenly around the center line (binomial sign test, p=1, n=53), indicating no average difference between modulation of the upwards versus downwards background motion. Therefore, in the following graphs we combined results for the offset90^{o} and offset270^{o} conditions to a single group of orthogonal offsets. Figure 3D shows the MIs for an orthogonal offset versus the offset180^{o} condition, showing a bias of distribution to larger MIs for the offset180^{o} condition (binomial sign test, p<0.001, n=106). The modulation indices of the offset180^{o} condition were significantly larger than the modulation indices obtained in the mixed conditions (Fig. 3E; binomial sign test, p<0.001, n=53). Similarly, the modulation indices of the orthogonal offsets were significantly larger than the modulation indices obtained in the mixed conditions (Fig. 3F; binomial sign test, p<0.001, n=106). Thus, neurons in the intermediate/deep layers of the OT tended to prefer homogenous over mixed background motion.

In this study, we distinguished between recordings from the superficial layers of the OT that receive direct retinal inputs and recordings from the intermediate/deep layers that receive visual inputs from the superficial layers and forebrain areas

(Luksch, 2003a). An example of a recording site from the superficial layers is shown in Figure 4. Except for the singleton condition, which shows a somewhat stronger albeit not significantly different response from the uniform condition (*t*-test, p=0.17, n=15), all other conditions produced responses that apparently did not differ from each other.

Figure 5 shows the population analysis of all 46 recording sites from the superficial layers. The average population PSTH in the singleton condition was significantly higher than the average PSTHs to the other conditions (Fig. 5A; ANOVA, $F_{4,225}$=32.4, p<0.001; post-hoc Tuckey tests, p<0.0001). The average response to the offset180^{o} was the second highest response and significantly larger compared to the uniform and orthogonal conditions (post-hoc Tuckey tests, P<0.001 for the uniform condition, p=0.002 for the offset90^{o} condition and p= 0.005 for the offset 270^{o} condition). The average responses to the rest of the conditions did not differ significantly from each other (Fig. 5B; ANOVA, $F_{3,180}$=0.51, p=0.6783). Thus, the main difference between the superficial and intermediate/deep layers was that on average the neurons in the superficial layers responded similarly to the uniform, orthogonal and mixed offsets, whereas in the intermediate/deep layers the neurons were significantly modulated by these offsets, responding stronger to the homogenous orthogonal conditions compared to the mixed condition (compare Fig. 3B with Fig. 5B). To directly compare recording sites in the superficial layers with results from recording sites in the intermediate/deep layers we calculated for each site the average difference between the responses to orthogonal conditions and the mixed condition. The difference was significantly larger in intermediate/deep sites compared to superficial sites (two-tailed t-test, df=97; p=0.0053).

In both the superficial and intermediate/deep layers, the MIs in the offset90^o condition did not differ systematically from the MIs in the offset270^o condition (Fig. 5C; binomial sign test, p=1, n=46). However, unlike in the intermediate/deep layers, the distribution of the MIs in the superficial layers was not significantly biased towards preferring the orthogonal background to the mixed background (Fig. 5F; binomial sign test, p=0.08, n=92).

The main conclusion from the electrophysiological results presented above is that the responses in the intermediate/deep layers of the OT to multiple elements cannot be simply explained by center-surround motion contrasts. The motion homogeneity of the surrounding elements plays a role in shaping the responses. Therefore, we performed a second experiment to address modulation by homogeneity, independent of contrast. For this, as before, the center dot moved to the right, however, the surrounding dots moved at offsets of either 135^o or -135^o (see insets in Fig. 6). By having two symmetrical offsets from the target we could control the homogeneity in the stimulus array (percent of dots moving in one direction) while maintaining the direction contrast between the center and the surrounding elements equal. In this paradigm we omitted the dots from the rectangle close to the target (see insets in Fig. 6B). Thus, the surrounding dots were not closer than 20^o from the target dot. We tested 49 multiunit recording sites in the intermediate/deep layers with this paradigm. Figure 6A shows raster plots of the responses from one recording site to five stimuli ranging in the percent of the surrounding dots moving in a direction of +135^o: 0% (0/21), 28% (6/21), 48% (10/21), 71% (15/21) and 100% (21/21). The average response was smallest for the mixed background (48%) and increased in both directions with increasing levels of homogeneity of the surround. At the population level, the homogeneity of the background elements significantly modulated the response (Fig. 6B; ANOVA, $F_{4,225}$=32.4, p<0.001). Both homogenous conditions (0% and 100%) elicited average responses that were significantly larger than the mixed condition (post-hoc Tuckey tests, n= 49, p<0.001). The intermediate conditions (28% and 71%) also elicited larger responses compared to the mixed condition (post-hoc Tuckey tests, n=49, p<0.05). Thus, the recorded population of neurons code the motion homogeneity of the elements in the surround.

Next, we asked how many elements are required for an opposing effect to take place. For this we performed an experiment in which the number of dots in the surrounding area varied between 0 to 21. In each trial, the number of dots and their positions on the screen were chosen randomly and either moved uniformly with the target dot (in the RF) or opposite the target. Possible positions for the dots were chosen from the dots array as in the experiment above (10^{o} spacing). Again, we omitted the dots from the rectangle close to the target (see insets in Fig. 7B). Data were collected from 54 multiunit recording sites in the intermediate/deep layers. A single dot moving inside the RF induced a vigorous response in the site shown in Figure 7A (lower raster and gray bar). Adding a second dot (somewhere in the surrounding array) resulted in a suppression of the average response. The suppression seemed independent of whether the motion was uniform to or opposite the target's motion. Similarly, 2, 4 or 6 dots in the surrounding area suppressed the response relative to the singleton response, independent of direction of motion (uniform or offset180^{o}). Stronger responses to opposing over uniform motion began to emerge when eight dots were displayed in the surrounding area and continued with additional dots (compare blue bars with red bars in Fig.7A).

Across the recorded population (n=54, Fig. 7B), the suppression of the response to a target in the RF by the additional dots in the surround is clear for both uniform and opposing conditions. However, in the uniform condition (red curve), suppression increased gradually reaching about 75% suppression at 21 surrounding dots. In the offset180^o condition (blue curve) the downwards inclination stopped at around 4 surrounding dots, and suppression level was kept at about 50% throughout (Student's t test comparing regression slopes, t_{860}=4.57, p<0.001). Thus, for the neurons to respond stronger to a motion contrast between the target and its surrounding area several coherently moving elements are required.

Discussion

Tectal neurons are known for their sensitivity to local motion. A small moving object gives rise to strong tectal responses if moving relative to a static background (Verhaal and Luksch, 2015) or if moving in a direction opposite to its background (Frost and Nakayama, 1983). By contrast, if an object moves in the same direction as the background the neural responses can be highly suppressed and sometimes completely eliminated (Frost et al., 1981; Frost and Nakayama, 1983; Dellen et al., 2004; Mysore et al., 2010; Zahar et al., 2012). This robust property of tectal neurons, which has been observed in the OT of fish, birds and mammals (Davidson and Bender, 1991; Zahar et al., 2012; Ben-Tov et al., 2015), is thought to allow rapid detection of localized motion and is consistent with the proposed role of the OT in the selection of the most salient stimulus (Dutta and Gutfreund, 2014; Mysore and Knudsen, 2011). Tectal sensitivity to opposing motion over uniform motion has also been associated with the ability to ignore self-induced motion cues (Frost et al., 1990) and with figure-ground-segregation (Frost et al., 1988; Davidson and Bender, 1991). However, sensitivity to opposing motion between the RF and its surrounding area is not sufficient for motion-based figure-ground-segregation. For this, it is essential to respond preferentially to targets moving oddly relative to a homogeneous motion in the background (Hegde and Felleman, 2003). Because this requirement necessarily contains motion contrasts between the receptive field and the surrounding area, it is not trivial to experimentally differentiate sensitivity to local motion contrasts from figure-ground sensitivity per se.

In this study, we addressed to what extent tectal neurons are modulated by homogeneity of the background. Previous studies have addressed a similar question in visual cortical areas by making use of conjunction stimuli (Hegde and Felleman, 2003; Burrows and Moore, 2009). Using this approach, it was shown that V1 neurons in monkeys are mostly sensitive to RF - surround contrasts rather than pop-out per se (Hegde and Felleman, 2003). Sensitivity to the homogeneity of the surround rather than to local discontinuities between the receptive field and the surround seems to require a higher level of visual processing (Burrows and Moore, 2009). Here, we used a somewhat different approach, testing one sensory feature, the direction of motion. We compared responses to stimuli where background motion is contrasting and homogenous (similar to pop-out stimuli) with responses to stimuli where background motion is contrasting but not homogenous (a mixed combination of three possible directions, two orthogonal and one opposite the direction of the target). The advantage in this design is that the target to background difference is defined by one feature, and can therefore be quantified easily as the average difference to target across all elements. Thus, the motion direction contrast between the target object and the surrounding objects followed the following order: offset180^o > mixed > offset90^o = offset270^o > uniform. Interestingly, the average population neural response in the intermediate/deep layers followed a different order: offset180^o > offset90^o = offset270^o > mixed > uniform. The responses to the orthogonal conditions outrun the mixed condition even though in the later, the target direction differed more from the directions of the background elements. Thus, sensitivity to center-surround motion contrast does not provide a good description of the responses. The homogeneity or regularity of the surrounding area enhances the responses to the target, consistent with motion-based grouping for figure-ground-segregation. This was also shown in an experiment where the homogeneity of the surround was varied while maintaining a constant background to center contrast (Fig. 6). The responses observed here are reminiscent of pop-out perception in humans (Duncan and Humphreys, 1989). In most visual search tasks, the strongest pop-out effect (shortest detection latencies and shallower search slopes) are observed when background elements are similar and the target is dissimilar. Pop-out strength scales down in a continuous manner as the similarity between the target and distractors increases, and scales down further as the similarity between the background elements decreases.

Our behavioral experiments show that barn owls, as well, perceive a target contrasting a homogenous background as salient compared to a target contrasting a mixed background (Fig. 1). Note that in our behavioral experiments, the owls were not trained to search for the odd target. Reward was given to encourage search behavior, but the target selection was spontaneous. This suggests that motion contrasting a homogenous background is an innate and powerful salient feature for barn owls.

Comparing the neural responses with the behavioral results we find that the population neural responses in the OT qualitatively matched the behavioral responses. The mixed conditions, which gave weaker neural responses compared to orthogonal and opposing motions, also gave slower responses and lower success rates at the behavioral level. However, at the neuronal level, the gap between the population responses of the offset180^{o} and the orthogonal conditions was larger than the gap between the responses of the orthogonal conditions and the mixed condition (Fig. 3A and B), whereas at the behavioral level, particularly in owl DO, the average difference between the responses to the offset180^{o} and orthogonal conditions was smaller relative to the difference to the mixed conditions (Fig. 1B-C). Thus, it seems that at the behavioral level the effect of the motion homogeneity over the mixed condition is stronger compared to at the neural level. This may indicate processing that takes place downstream from the OT to further separate between the mixed and the homogenous conditions, and/or, that the read-out for perception is from a sub-population of the recorded neurons. Interestingly, the spontaneous search of the odd target in the mixed condition, even though resulted in poor performance relative to the homogenous conditions, was still significantly above chance level. This, again, agrees with the neural responses in the intermediate/deep layers, which on average were greater in the mixed conditions than in the uniform condition (Fig. 3B).

The Gestalt principles for perceptually organizing the visual scene have been established by human psychologists. However, birds can demonstrate remarkably similar principles. For example, barn owls have been shown to perceive subjective contours (Nieder and Wagner, 1999), and pigeons are capable of grouping by shape and color (Cook et al., 1996). Our finding adds to these previous findings, proposing the hypothesis that human Gestalt principles are manifestations of general neural mechanisms evolved to cope with common ecological needs of visually foraging animals. This raises the intriguing possibility that birds and mammals share similar neural mechanisms for perceptual grouping.

The intermediate/deep layers of the OT provide the major output pathways of the OT (Luksch, 2003b). Neural responses in these layers have been shown to be highly context-dependent, modulated by other modalities (Mysore et al., 2010; Zahar et al., 2009), by stimulation history (Reches and Gutfreund, 2008; Netser et al., 2011) and by stimuli well outside of the RF (Mysore et al., 2010; Zahar et al., 2012). The findings of neural correlates of perceptual grouping in these layers agree with the emerging notion that the intermediate/deep layers of the OT form a priority map to represent the most relevant stimulus for the control of orienting behavior (Mysore and Knudsen, 2011; Gutfreund, 2012). This evolutionary role of the OT seems to be conserved in vertebrates all the way to primates (Kardamakis et al., 2015; Boehnke and Munoz, 2008). Neurons in the monkey's superior colliculus are also highly modulated by the surroundings and the history of stimulation (Davidson and Bender, 1991; Boehnke et al., 2011), and can discriminate between the selected target and distractors in visual feature and conjunction search tasks (McPeek and Keller, 2002; Shen et al., 2011).

It is possible that the processing for obtaining selective modulation by a homogenous contrasting background takes place in the retina (Olveczky et al., 2003). Our results do not support this. The superficial layers of the OT, which receive direct retinal input and relay visual information to the intermediate/deep layers, did not show selectivity to homogenous versus mixed backgrounds. However, if the neural responses are shaped by the motion direction contrasts between the RF and its surround, independent of the homogeneity, the prediction is for the responses to the mixed conditions to be significantly larger than the responses to the orthogonal conditions. This prediction is fulfilled neither in the superficial layers (Fig. 5B) nor in the intermediate/deep layers (Fig. 3B). Therefore, a basic effect of homogeneity can also be traced in the superficial layers. The effect increases in the intermediate/deep layers where the responses were found to code the level of homogeneity independent of the motion direction contrast (Fig. 6).

Previous studies have revealed an extensive lateral inhibitory network in the avian OT. This network contains a feedback loop through the isthmi complex, which enables the more powerful stimulus to suppress responses to the less powerful stimulus, and thus give rise to competitive interactions (Wang, 2003; Mysore and Knudsen, 2013; Marin et al., 2007). However, the lateral inhibition mediated by the isthmi complex seems to be nonspecific to direction or orientation of the stimulus (Maczko et al., 2006; Saha et al., 2011) and therefore cannot explain selective modulation. Consistent with nonspecific lateral inhibition, we observed nonspecific suppression by the surrounding elements when only one or two elements in the surround were shown. The sensitivity to motion contrasts that we observed in the OT seems to require a group of homogenously moving elements (Fig. 7). The neural circuitry to achieve this important property is yet to be discovered.

References

Ben-Tov M, Donchin O, Ben-Shahar O, Segev R (2015) Pop-out in visual search of moving targets in the archer fish. Nat Commun 6:6476.

Boehnke SE, Munoz DP (2008) On the importance of the transient visual response in the superior colliculus. Curr Opin Neurobiol 18:544-551. Epub 2008 Dec 2006.

Boehnke SE, Berg DJ, Marino RA, Baldi PF, Itti L, Munoz DP (2011) Visual adaptation and novelty responses in the superior colliculus. Eur J Neurosci 34:766-779.

Brainard DH (1997) The psychophysics toolbox. Spatial Vision:433-436.

Burrows BE, Moore T (2009) Influence and limitations of popout in the selection of salient visual stimuli by area V4 neurons. J Neurosci 29:15169-15177.

Cook RG, Cavoto KK, Cavoto BR (1996) Mechanisms of multidimensional grouping, fusion, and search in avian texture discrimination. Animal Learning & Behavior 24:150-167.

Davidson RM, Bender DB (1991) Selectivity for relative motion in the monkey superior colliculus. J Neurophysiol 65:1115-1133.

Dellen BK, Clark JW, Wessel R (2004) Motion-contrast computation without directionally selective motion sensors. Phys Rev E Stat Nonlin Soft Matter Phys 70:031907.

du Lac S, Knudsen EI (1990) Neural maps of head movement vector and speed in the optic tectum of the barn owl. J Neurophysiol 63:131-146.

Duncan J, Humphreys GW (1989) Visual search and stimulus similarity. Psychol Rev 96:433-458.

Dutta A, Gutfreund Y (2014) Saliency mapping in the optic tectum and its relationship to habituation. Front Integr Neurosci 8:1.

Frost BJ, Nakayama K (1983) Single visual neurons code opposing motion independent of direction. Science 220:744-745.

Frost BJ, Scilley PL, Wong SC (1981) Moving background patterns reveal double-opponency of directionally specific pigeon tectal neurons. Exp Brain Res 43:173-185.

Frost BJ, Cavanagh P, Morgan B (1988) Deep tectal cells in pigeons respond to kinematograms. J Comp Physiol A 162:639-647.

Frost BJ, Wylie DR, Wang YC (1990) The processing of object and self-motion in the tectofugal and accessory optic pathways of birds. Vision Res 30:1677-1688.

Gutfreund Y (2012) Stimulus-specific adaptation, habituation and change detection in the gaze control system. Biol Cybern 106:657-668.

Harmening WM, Orlowski J, Ben-Shahar O, Wagner H (2011) Overt attention toward oriented objects in free-viewing barn owls. Proc Natl Acad Sci U S A 108:8461-8466.

Hazan Y, Kra Y, Yarin I, Wagner H, Gutfreund Y (2015) Visual-auditory integration for visual search: a behavioral study in barn owls. Front Integr Neurosci 9:11.

Hegde J, Felleman DJ (2003) How selective are V1 cells for pop-out stimuli? J Neurosci 23:9968-9980.

Hochstein S, Ahissar M (2002) View from the top: hierarchies and reverse hierarchies in the visual system. Neuron 36:791-804.

Kardamakis AA, Saitoh K, Grillner S (2015) Tectal microcircuit generating visual selection commands on gaze-controlling neurons. Proc Natl Acad Sci U S A 112:E1956-1965.

Kingstone A, Bischof WF (1999) Perceptual Grouping and Motion Coherence in Visual Search. Psychological Science 10:151-156.

Knudsen EI (1982) Auditory and visual maps of space in the optic tectum of the owl. J Neurosci 2:1177-1194.

Lamme VA (1995) The neurophysiology of figure-ground segregation in primary visual cortex. J Neurosci 15:1605-1615.

Lee TS, Yang CF, Romero RD, Mumford D (2002) Neural activity in early visual cortex reflects behavioral experience and higher-order perceptual saliency. Nat Neurosci 5:589-597.

Lev-Ari T, Gutfreund Y (2017) Interactions between top-down and bottom-up attention in barn owls (Tyto alba). Anim Cogn.

Luksch H (2003a) Cytoarchitecture of the avian optic tectum: neuronal substrate for cellular computation. Rev Neurosci 14:85-106.

Luksch H (2003b) Cytoarchitecture of the avian optic tectum: neuronal substrate for cellular computation. Rev Neurosci 14:85-106.

Maczko KA, Knudsen PF, Knudsen EI (2006) Auditory and visual space maps in the cholinergic nucleus isthmi pars parvocellularis of the barn owl. J Neurosci 26:12799-12806.

Marin G, Salas C, Sentis E, Rojas X, Letelier JC, Mpodozis J (2007) A cholinergic gating mechanism controlled by competitive interactions in the optic tectum of the pigeon. J Neurosci 27:8112-8121.

McPeek RM, Keller EL (2002) Saccade target selection in the superior colliculus during a visual search task. J Neurophysiol 88:2019-2034.

Mysore SP, Knudsen EI (2011) The role of a midbrain network in competitive stimulus selection. Curr Opin Neurobiol 21:653-660.

Mysore SP, Knudsen EI (2013) A shared inhibitory circuit for both exogenous and endogenous control of stimulus selection. Nat Neurosci 16:473-478.

Mysore SP, Asadollahi A, Knudsen EI (2010) Global inhibition and stimulus competition in the owl optic tectum. J Neurosci 30:1727-1738.

Netser S, Ohayon S, Gutfreund Y (2010) Multiple manifestations of microstimulation in the optic tectum: eye movements, pupil dilations, and sensory priming. J Neurophysiol 104:108-118.

Netser S, Zahar Y, Gutfreund Y (2011) Stimulus-specific adaptation: can it be a neural correlate of behavioral habituation? J Neurosci 31:17811-17820.

Nieder A, Wagner H (1999) Perception and neuronal coding of subjective contours in the owl. Nat Neurosci 2:660-663.

Ohayon S, Harmening W, Wagner H, Rivlin E (2008) Through a barn owl's eyes: interactions between scene content and visual attention. Biol Cybern 98:115-132. Epub 2007 Dec 2008.

Olveczky BP, Baccus SA, Meister M (2003) Segregation of object and background motion in the retina. Nature 423:401-408. Epub 2003 May 2011.

Qiu FT, von der Heydt R (2005) Figure and ground in the visual cortex: v2 combines stereoscopic cues with gestalt rules. Neuron 47:155-166.

Reches A, Gutfreund Y (2008) Stimulus-specific adaptations in the gaze control system of the barn owl. J Neurosci 28:1523-1533.

Saha D, Morton D, Ariel M, Wessel R (2011) Response properties of visual neurons in the turtle nucleus isthmi. J Comp Physiol A Neuroethol Sens Neural Behav Physiol 197:153-165.

Shen K, Valero J, Day GS, Pare M (2011) Investigating the role of the superior colliculus in active vision with the visual search paradigm. Eur J Neurosci 33:2003-2016.

Spillman L, Ehrenstein WH (2003) Gestalt factors in visual neuroscience. In: The Visual Neurosciences (Chalupa LM, Werner JS, eds), pp 1573-1589. Cambridge, MA: MIT Press.

Treisman A, Gelade G (1980) A feature-integration theory of attention. Cognitive psychology 12:97-136.

Verhaal J, Luksch H (2015) Processing of motion stimuli by cells in the optic tectum of chickens. Neuroreport 26:578-582.

Wang SR (2003) The nucleus isthmi and dual modulation of the receptive field of tectal neurons in non-mammals. Brain Res Brain Res Rev 41:13-25.

Wasmuht DF, Pena JL, Gutfreund Y (2017) Stimulus-specific adaptation to visual but not auditory motion direction in the barn owl's optic tectum. Eur J Neurosci 45:610-621.

Wathey JC, Pettigrew JD (1989) Quantitative analysis of the retinal ganglion cell layer and optic nerve of the barn owl Tyto alba. Brain Behav Evol 33:279-292.

Wolfe JM, Horowitz TS (2004) What attributes guide the deployment of visual attention and how do they do it? Nat Rev Neurosci 5:495-501.

Zahar Y, Reches A, Gutfreund Y (2009) Multisensory enhancement in the optic tectum of the barn owl: spike count and spike timing. J Neurophysiol 101:2380-2394. Epub 2009 Mar 2384.

Zahar Y, Wagner H, Gutfreund Y (2012) Responses of tectal neurons to contrasting stimuli: an electrophysiological study in the barn owl. PLoS One 7:e39559.

Zhou H, Friedman HS, von der Heydt R (2000) Coding of border ownership in monkey visual cortex. J Neurosci 20:6594-6611.

Zipser K, Lamme VA, Schiller PH (1996) Contextual modulation in primary visual cortex. J Neurosci 16:7376-7389.

Figure Legends

Figure 1. Behavioral paradigm and results. **A**, The owls initiate trials by fixating a circle at the center of the screen. In a trial, an array of dots appears on the screen with

the target dot positioned randomly in one of four positions around the center. Possible target positions are designated by red dashed circles in the upper panel. The lower panels illustrate the four stimulus conditions. From left to right: singleton, offset180°, offset90° and mixed. **B**, The success rates (proportion of trials fixating the target) are shown for owls DO and DK separately. Colors designate the four tested conditions. The white numbers show successes per trials. The internal white columns show the corresponding proportion of trials fixating the control dot (opposite the target). Horizontal lines indicate a significant difference between groups. **C,** The average times to fixation on target (search time) are shown separately for the two owls DO and DK. Colors indicate the different conditions. Horizontal lines indicate significant differences between groups (Post-hoc Tuckey test, $p < 0.05$). Error bars indicate SEMs. **D,** Normalized cumulative distributions of number of fixations to reach the target (solid lines) and to reach the corresponding control (dashed lines). A rightward shift indicates a tendency for more fixations before reaching the target and downward shift less successful trials.

Figure 2. Example of the responses of a single recording site in the intermediate/deep layers. **A,** Responses to 15 repetitions of the stimulus are shown in the raster plot. The corresponding PSTH is shown below. The red vertical line designates the onset of stimulation. The inset shows a scheme of the singleton stimulus condition. The red circle encloses the target dot which is moving rightwards (direction indicated by the small arrow). **B,** Responses to the uniform stimulus. The inset shows a scheme of the uniform stimulus condition where all dots are moving coherently rightwards. Format as in A. **C,** Responses to offset180° stimulus. The inset shows a scheme of the offset180° stimulus condition where the target is moving rightwards and all other dots are moving leftwards. Format as in A. **D,** Responses to offset90° stimulus. The inset

shows a scheme of the offset90^{o} stimulus condition where the target is moving rightwards and all other dots are moving upwards. Format as in A. **E,** Responses to the offset270^{o} stimulus. The inset shows a scheme of the offset270^{o} stimulus condition where the target is moving rightwards and all other dots are moving downwards. Format as in A. **F,** Responses to the mixed condition. The inset shows a scheme of the mixed stimulus condition where the target is moving rightwards and all other dots are moving either leftwards, upwards or downwards. Format as in A.

Figure 3. Summary of population results from recordings made in the intermediate/deep layers of the OT. **A,** Population average PSTH smooth curves comparing responses of singleton, uniform, offset180^{o}, offset 90^{o} and offset270^{o} conditions. The width of the curves designate the SEM. **B,** Same as in A but comparing population average curves between uniform, offset90^{o}, offset270^{o} and mixed conditions. **C,** the scattergram shows the MIs in offset270^{o} versus the MIs in offset90^{o}. The red diagonal line designate the midline. **D,** The scattergram shows the MIs in the orthogonal conditions versus the MIs in the offset180^{o} condition. Format as in C. **E,** The scattergram shows the MIs in the mixed condition versus the MIs in the offset180^{o} condition. Format as in C. **F,** The scattergram shows the MIs in the mixed condition versus the MIs in the orthogonal conditions. Format as in C.

Figure 4. Example of responses from a single recording site in the superficial layers of the OT. **A,** responses to 15 repetitions of the singleton stimulus are shown in the raster plot. The corresponding PSTH is shown below. The red vertical bar designates the onset of stimulation. The inset shows a scheme of the stimulus condition. The red circle indicates the target dot in the RF which was moving rightwards (indicated by the small arrow). **B,** Responses to the uniform stimulus. The inset shows a scheme of the uniform stimulus condition where all dots are moving coherently rightwards.

Format as in A. **C,** Responses to offset180^o stimulus where the target is moving rightwards and all other dots are moving leftwards. Format as in A. **D,** Responses to offset90^o stimulus where the target is moving rightwards and all other dots are moving upwards. Format as in A. Responses to the offset270^o stimulus where the target is moving rightwards and all other dots are moving downwards. Format as in A. F, Responses to the mixed condition where the target is moving rightwards and all other dots are moving either leftwards upwards or downwards. Format as in A.

Figure 5. Summary of population results from recordings made in the superficial layers of the OT. **A,** Population average PSTH curves comparing responses of singleton, uniform, offset180^o, offset90^o and ofset270^o conditions. The width of the curves designate the SEM. The dashed line designate baseline level **B,** Same as in A but comparing population average PSTH curves between uniform, offset90o, offset270^o and mixed conditions. **C,** the scattergram shows the MIs in offset270^o versus the MIs in offset90^o. The red diagonal line designate the midline. **D,** The scattergram shows the MIs in the orthogonal conditions versus the MIs in the offset180^o condition. Format as in C. **E,** The scattergram shows the MIs in the mixed condition versus the MIs in the offset180^o condition. Format as in C. **F,** The scattergram shows the MIs in the mixed condition versus the MIs in the orthogonal conditions. Format as in C.

Figure 6. Effects of changing the homogeneity of the elements in the surround. **A,** Raster plots showing the responses of a single recording site. The percent of the dots moving at an offset of 135^o (Northwest) was varied in 5 steps between 0-100%. The red line designates the onset of motion which lasted 600 ms. The columns on the right designate the average responses at each of the five steps. Error bars indicate SEM. **B,** A histogram showing the average population responses (n=49, in the

intermediate\deep layers) as a function of the percent of dots moving north-west. The error bars indicate SEM, The red column at the right indicates the population average response to the uniform condition. The lines at the top indicate results of statistical analysis (Tuckey post-hoc test). The average response of the uniform condition was significantly smaller from all other conditions (Tuckey post-hoc test, $p<0.01$). The insets show an illustration of the visual display at 0%, 48% and 100% Northwest motion.

Figure 7. Changing the number of distractors. **A,** Raster plots showing responses of a single site. The number of distractors vary from singleton (lower raster) to 21 distractors (upper rasters). Numbers on the left designate the number of distractors. The left column shows results from when the distractors moved in the same direction as the target (uniform). Right column shows results from when the distractors moved in the opposite direction of the target (offset180). Red vertical line designate stimulus onset. The bar plots on the right show the average responses ± SEM to the uniform conditions (red bars) and the offset180 conditions (blue bars). **B,** The population (intermediate/deep layers) average normalized responses as a function of the number of distractors in the surround, shown separately for uniform motion (red curve) and offset180 motion (blue). Error bars designate SEMs. Single asterisk designates a significant level of 0.05 and double asterisk a significant level of 0.001(one-tailed paired t-test). The insets show examples of stimulus displays for singleton, 4, 12, 14 and 21 distractors. The dashed circle designates a hypothetical RF with the target dot inside.

Figure 1

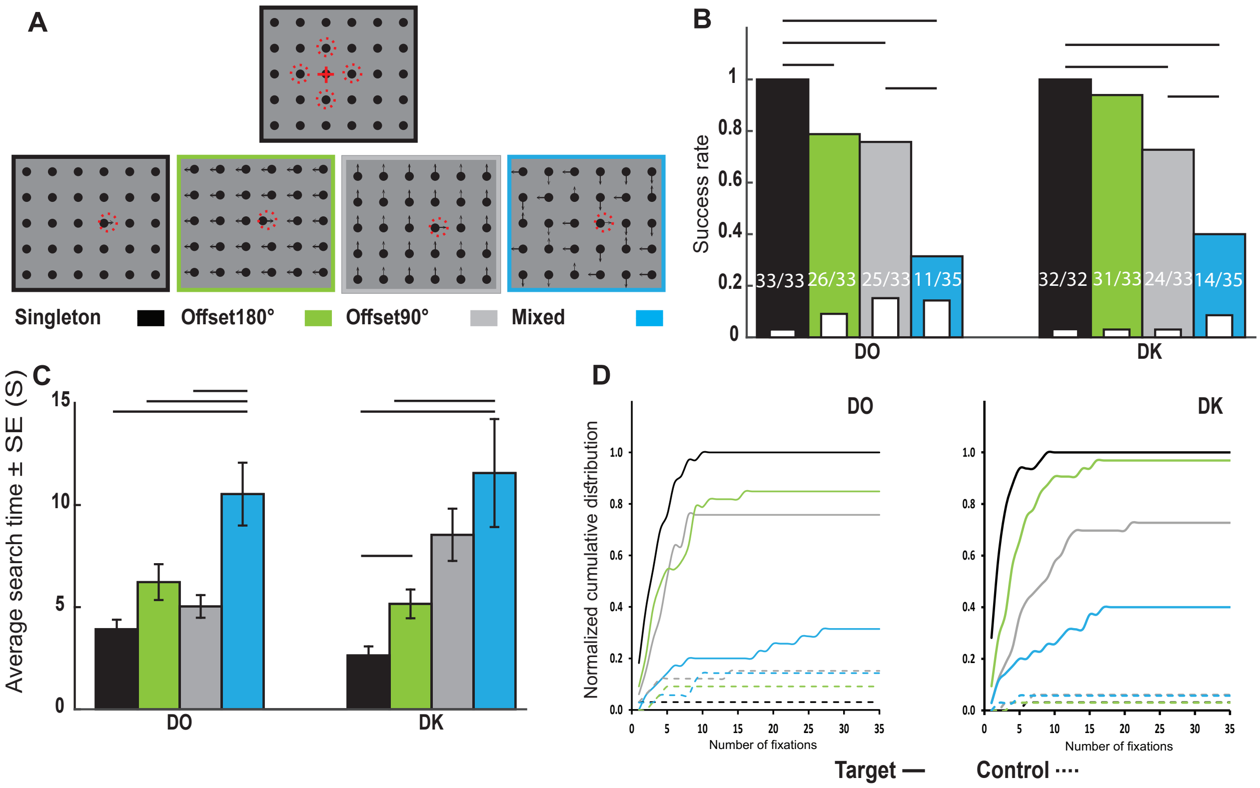

A
Singleton
Offset180°
Offset90°
Mixed
B
Success rate
1
0.8
0.6
0.4
0.2
0
33/33
26/33
25/33
11/35
32/32
31/33
24/33
14/35
DO
DK
C
Average search time ± SE (S)
15
10
5
0
DO
DK
D
Normalized cumulative distribution
DO
DK
1.0
0.8
0.6
0.4
0.2
0.0
0
5
10
15
20
25
30
35
Number of fixations
Target —
Control ····

Figure 2

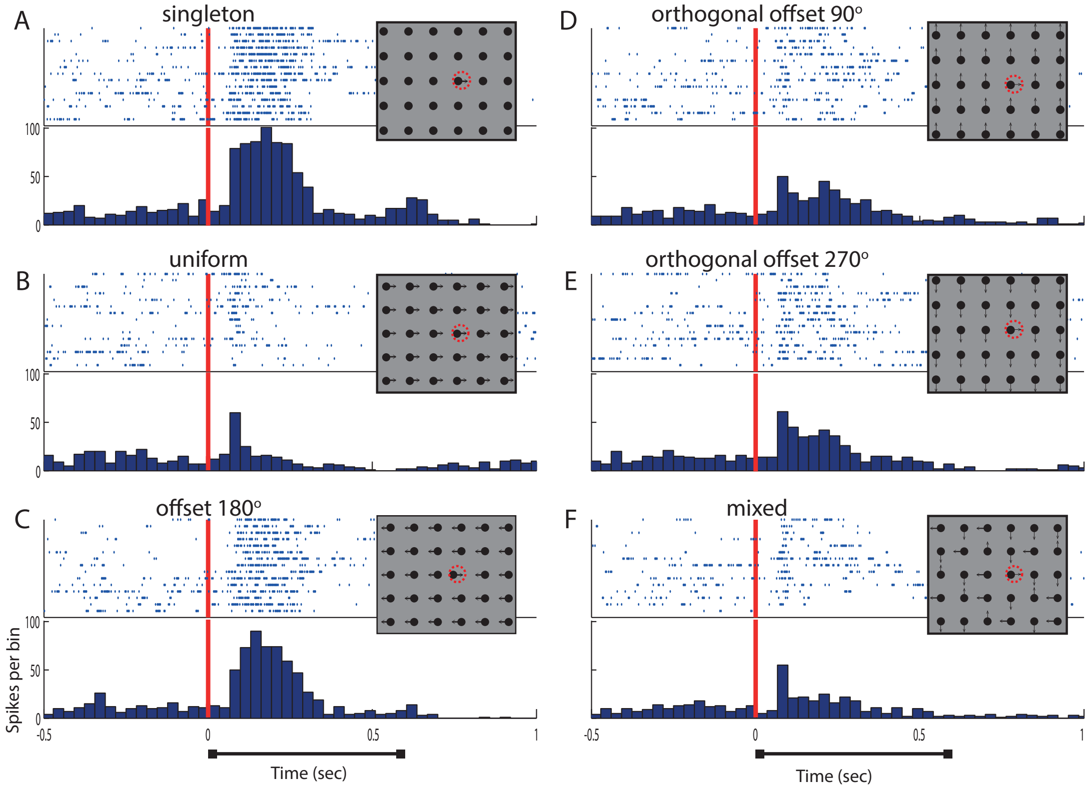

A
singleton
B
uniform
C
offset 180°
D
orthogonal offset 90°
E
orthogonal offset 270°
F
mixed
Spikes per bin
100
50
0
-0.5
0
0.5
1
Time (sec)
Time (sec)

Figure 3

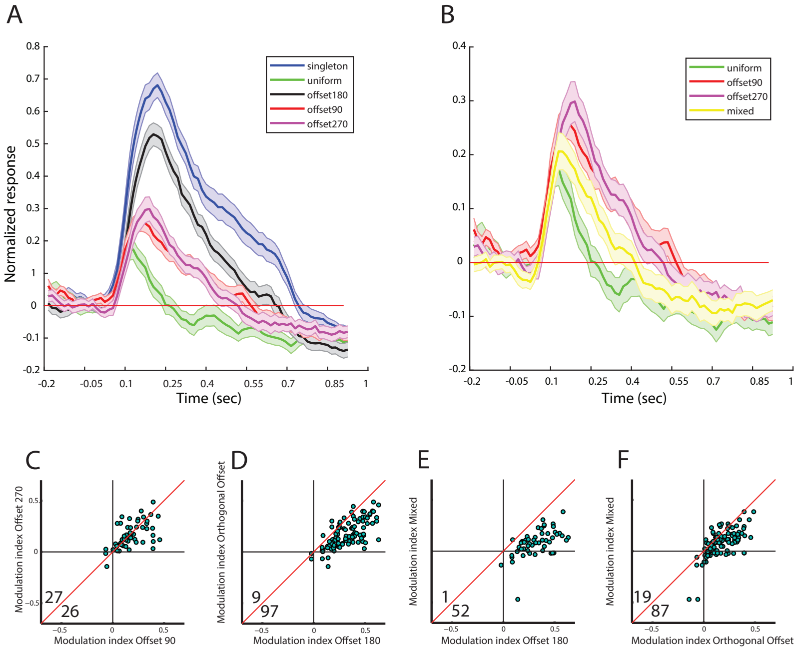

A
B
singleton
uniform
offset180
offset90
offset270
mixed
Normalized response
Time (sec)
-0.2
-0.05
0.1
0.25
0.4
0.55
0.7
0.85
1
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
-0.1
C
D
E
F
Modulation index Offset 270
Modulation index Offset 90
Modulation index Orthogonal Offset
Modulation index Offset 180
Modulation index Mixed
Modulation index Offset 180
Modulation index Mixed
Modulation index Orthogonal Offset
-0.5
0
0.5
27
26
9
97
1
52
19
87

Figure 4

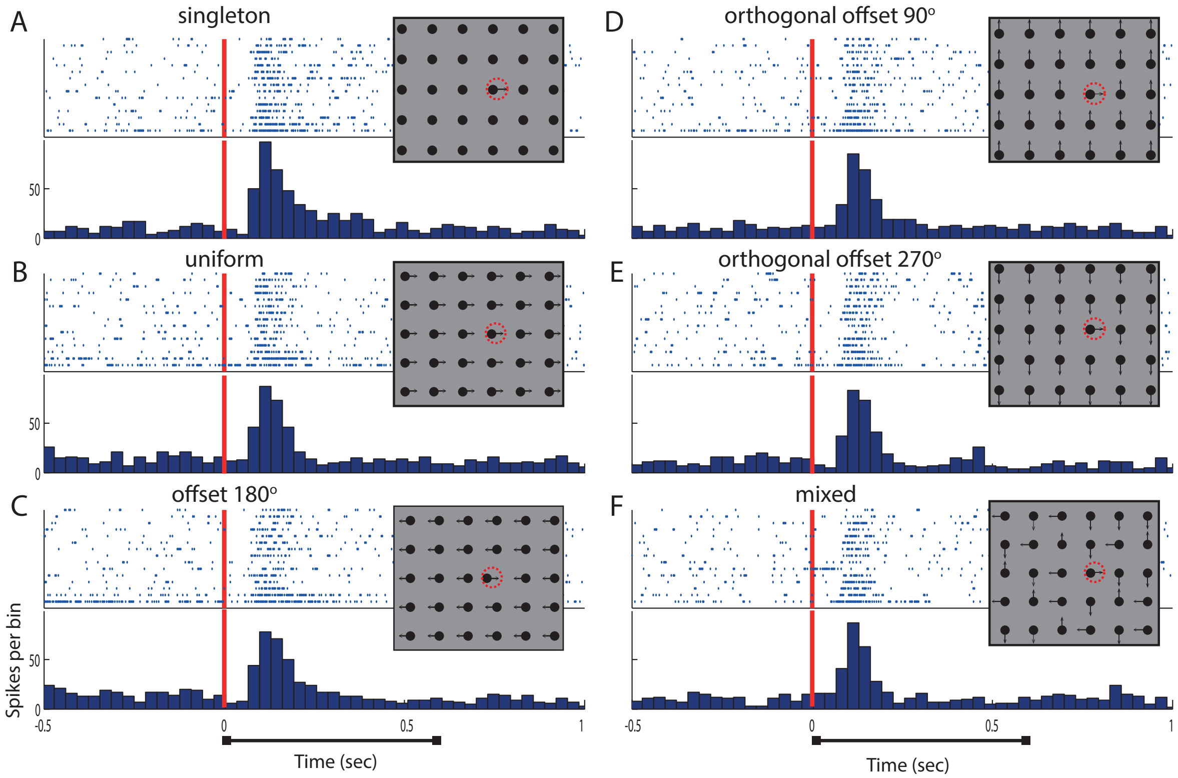

A
singleton
B
uniform
C
offset 180°
D
orthogonal offset 90°
E
orthogonal offset 270°
F
mixed
50
0
Spikes per bin
-0.5
0
0.5
1
Time (sec)

Figure 5

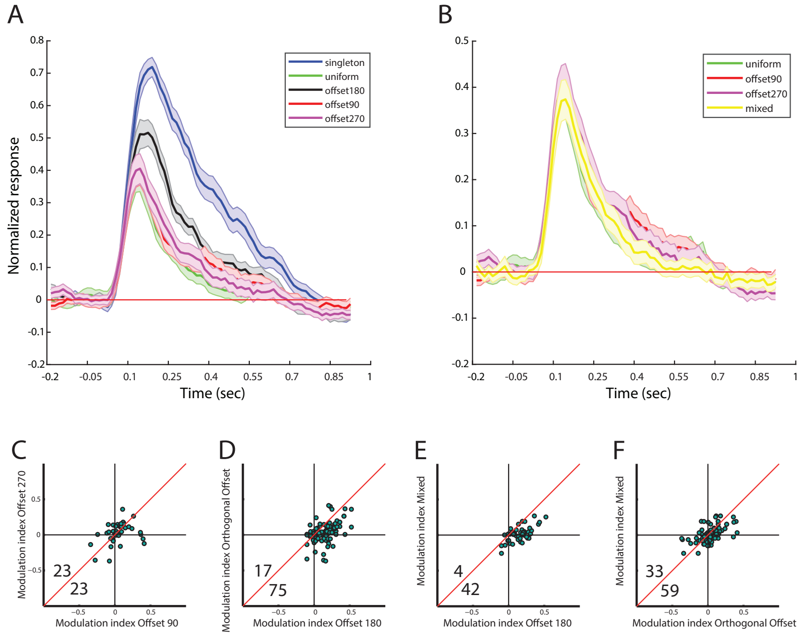

A
Normalized response
Time (sec)
singleton
uniform
offset180
offset90
offset270
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
-0.1
-0.2
-0.2
-0.05
0.1
0.25
0.4
0.55
0.7
0.85
1
B
Time (sec)
uniform
offset90
offset270
mixed
0.5
0.4
0.3
0.2
0.1
0
-0.1
-0.2
C
Modulation index Offset 270
Modulation index Offset 90
23
23
D
Modulation index Orthogonal Offset
Modulation index Offset 180
17
75
E
Modulation index Mixed
Modulation index Offset 180
4
42
F
Modulation index Mixed
Modulation index Orthogonal Offset
33
59
0.5
0
-0.5

Figure 6

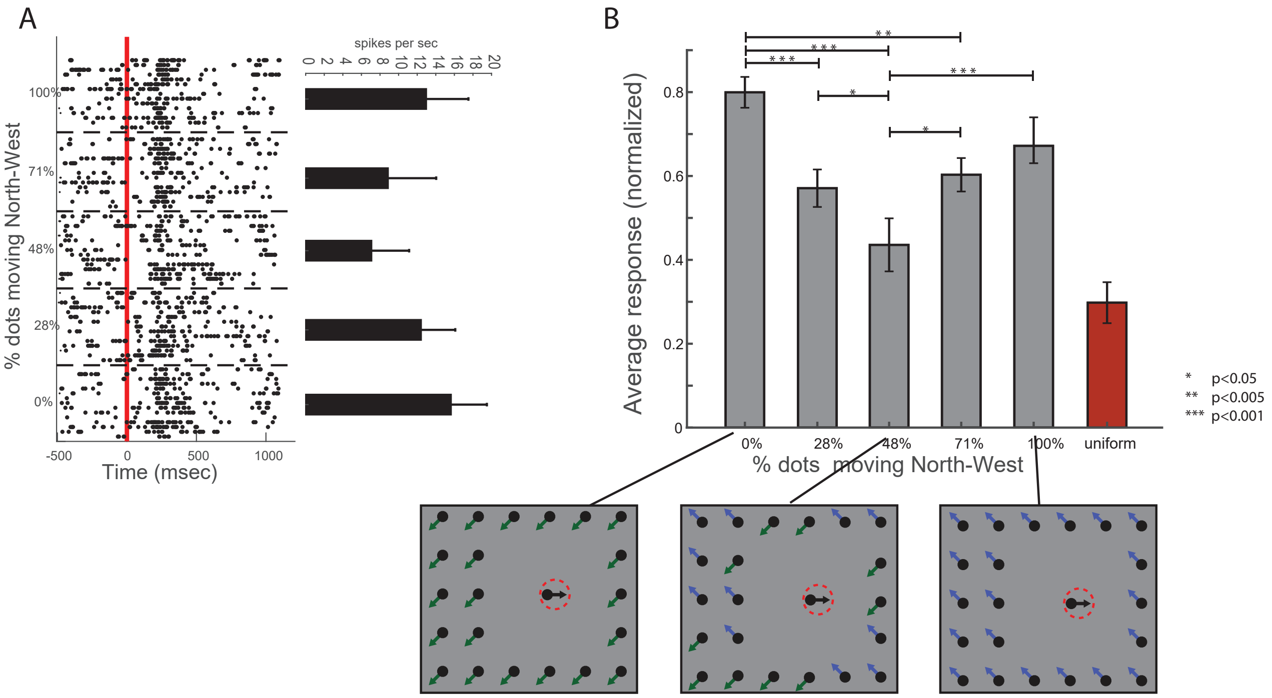
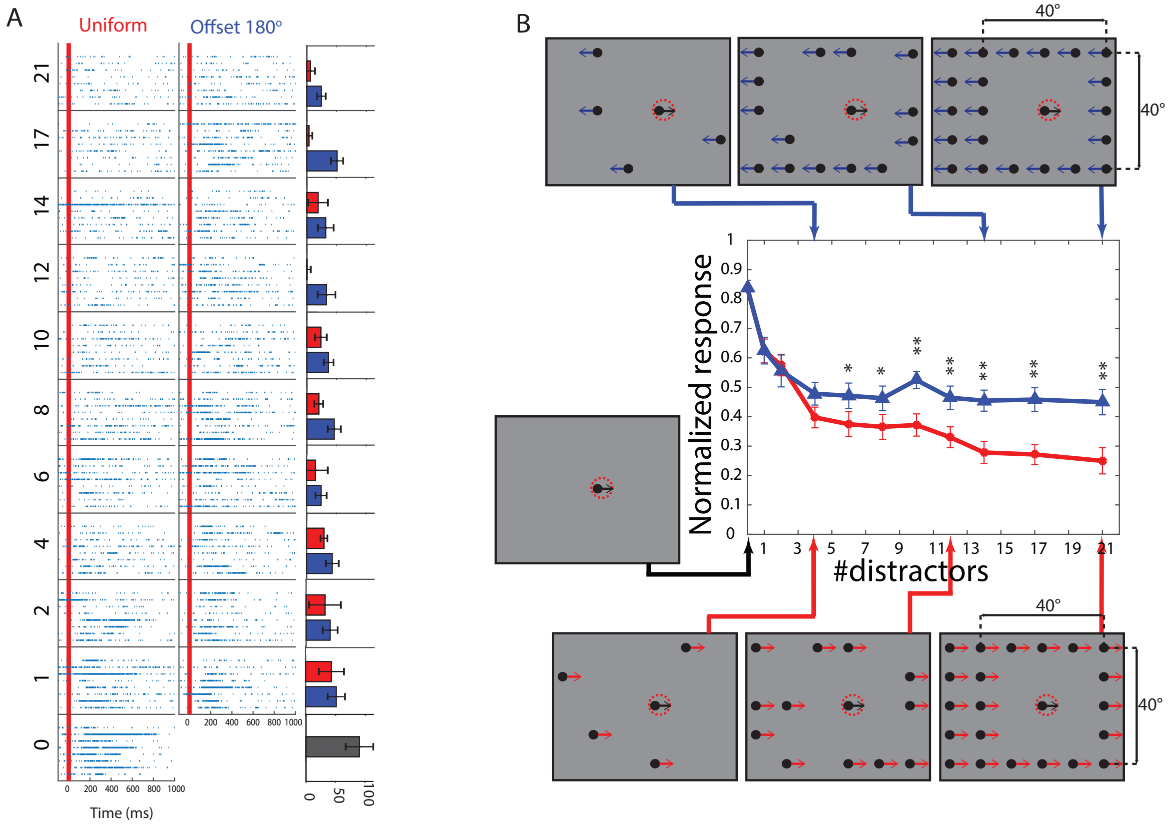
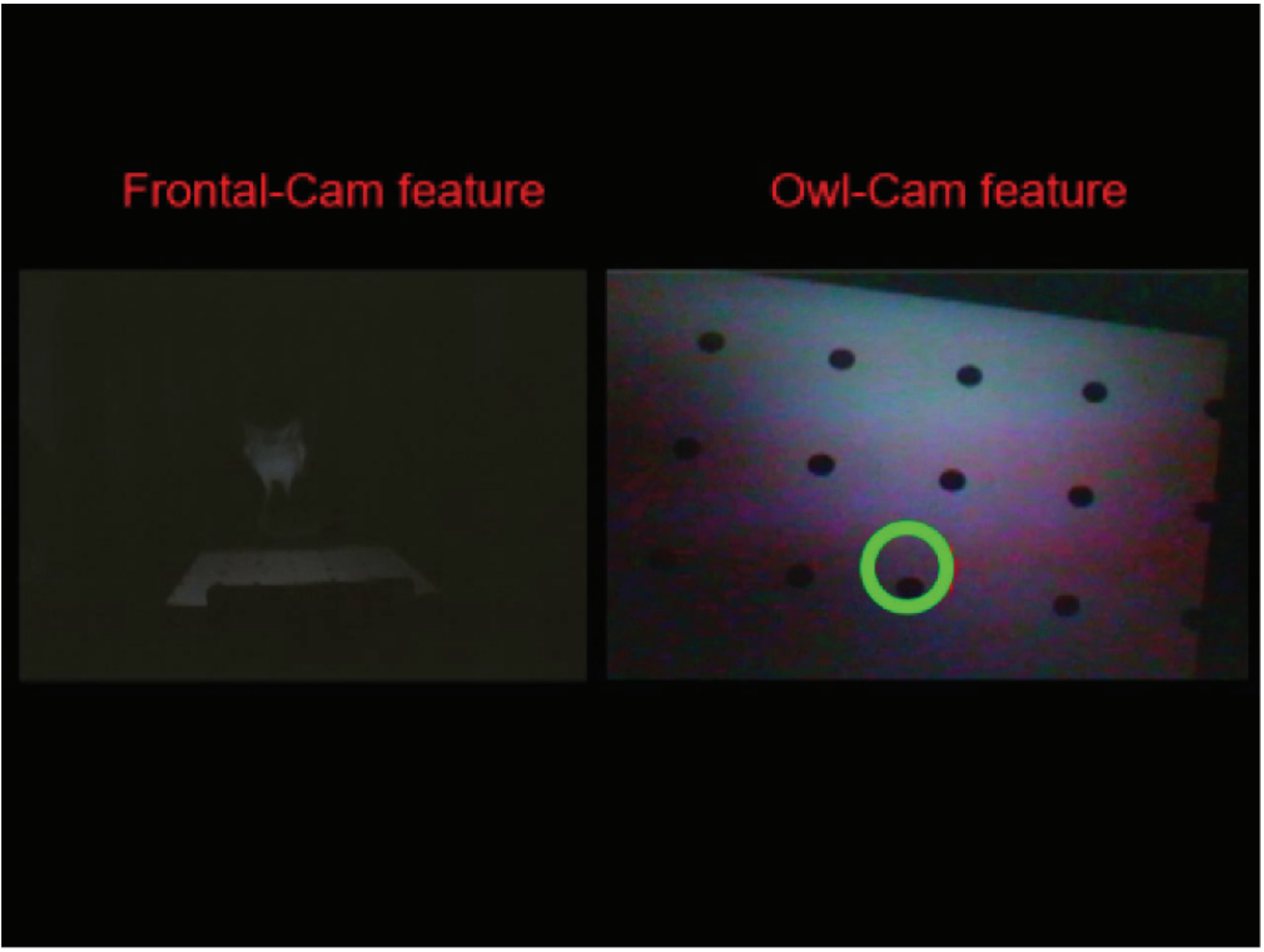
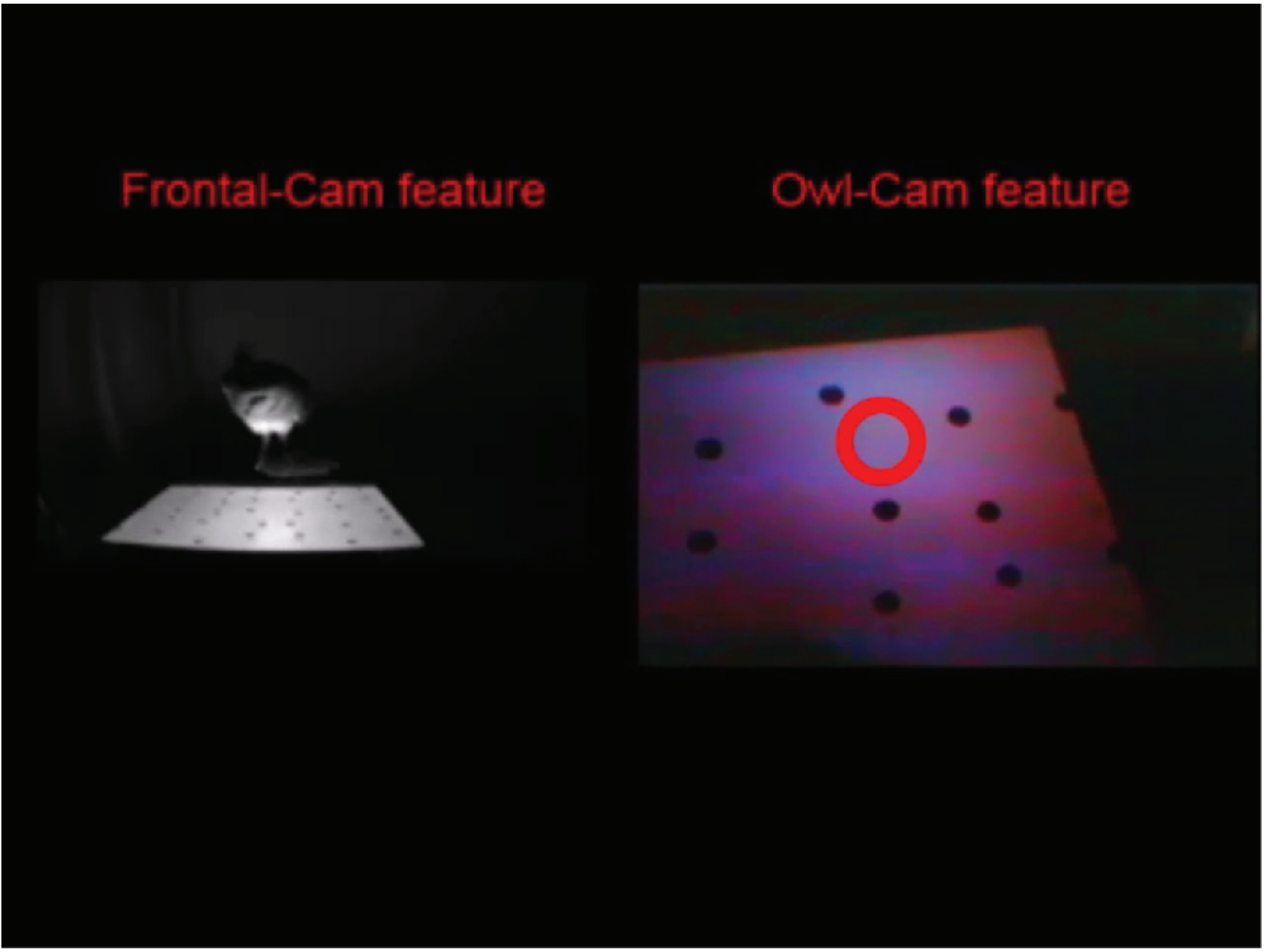

A
spikes per sec
0
2
4
6
8
10
12
14
16
18
20
100%
71%
48%
28%
0%
% dots moving North-West
-500
0
500
1000
Time (msec)
B
Average response (normalized)
0
0.2
0.4
0.6
0.8
**
***
***
*
***
*
0%
28%
48%
71%
100%
uniform
% dots moving North-West
* p<0.05
** p<0.005
*** p<0.001

Figure 7


A
Uniform
Offset 180°
21
17
14
12
10
8
6
4
2
1
0
0 200 400 600 800 1000
Time (ms)
0
50
100
B
40°
40°
Normalized response
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
1 3 5 7 9 11 13 15 17 19 21
#distractors
*
*
**
**
**
**
**
40°
40°


Frontal-Cam feature
Owl-Cam feature


Frontal-Cam feature
Owl-Cam feature