Figure Ground Modulation in the rodent visual system

Francisco Luongo[1], Lu Liu[1], and Doris Tsao[1,2]

Caltech, Pasadena, CA
1. Division of Biology and Biological Engineering, Caltech **2.** The Howard Hughes Medical Institute

INTRODUCTION

- Visual circuits should be able to generate concise descriptions of visual scenes to appropriately guide behavior (e.g., visual objects and their properties) but how the visual system might generate such a description remains an open question.

- Two such useful representations are the segmentation of visual objects into surfaces (e.g. Figure vs. Background) and the assignment of edges to specific objects (e.g. Border ownership).

- We propose to study the mechanisms by which the brain constructs these representations in the rodent visual system.

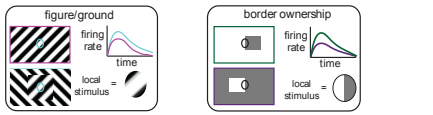
- We assayed primary and higher order visual areas for figure ground and border ownership modulation using a second order figure motion stimulus.

- We find strongest border ownership modulation in V1 and strongest figure ground modulation in higher order area LM.

- Neural networks can be used to capture some of the nonlinear responses of neurons.

- Mice can perform a texture-invariant figure identification task

BACKGROUND



*Qiu and von der Heydt, 2007 // Zhou, Freedman, and von der Heydt 2000 // Jeurrisen, Self, Roelfsema 2013 // Lamme 1995

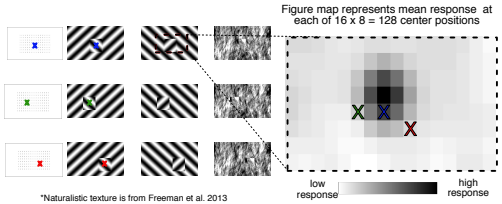
MAPPING VISUAL AREAS

Retinotopic visual areas identified via widefield imaging with either a Thy1::GCamP6 or CamkII::ttA x TetO::GCamp6 transgenic animal.

Reference vasculature images were used to guide subsequent optical (2-photon) or electrical (64 channel silicon probe) recordings

STIMULI

Figure presented at 16x8 = 128 locations with continuous motion of both background and the figure. Figure motion is anti-phase to background providing a second order motion cue.



*Naturalistic texture is from Freeman et al. 2013

SCHEMATIC RESPONSES

Model responses from a hypothetical simple cell, figure/ground cell, and border ownership cell. Notice the simple cell responds to the local stimulus statistics while the figure ground and border ownership cells have a preferred position of figure independent of texture

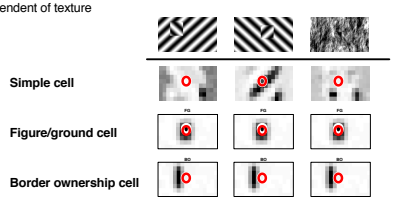
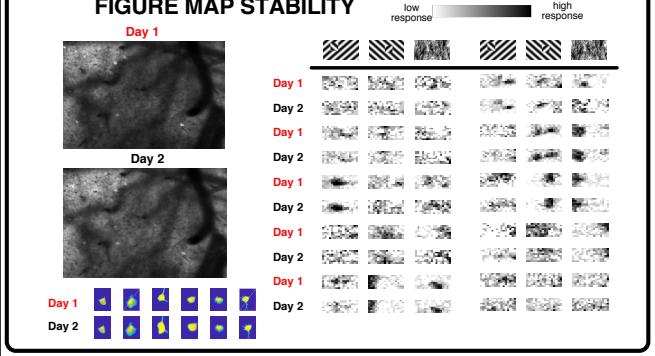
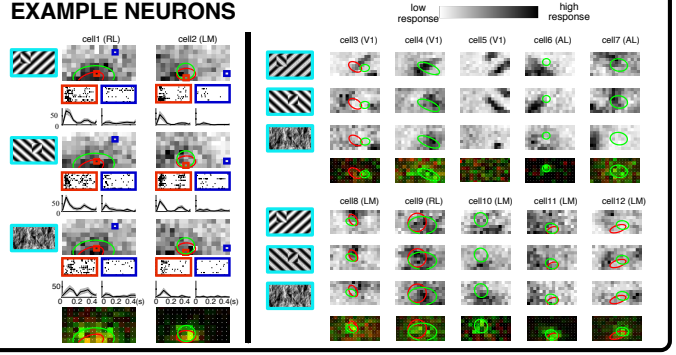


FIGURE MAP STABILITY

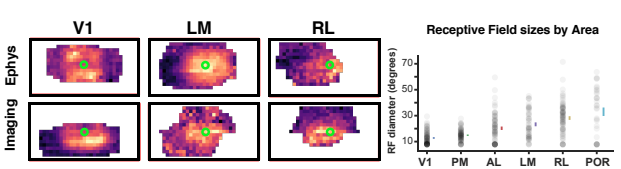


EXAMPLE NEURONS



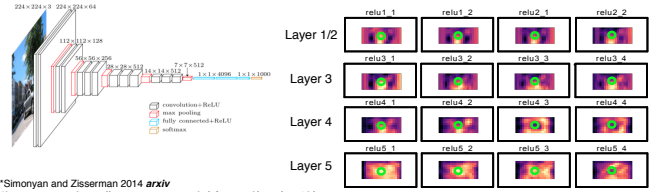
RECEPTIVE FIELD CENTERED MAPS

We combined maps across neurons by binning based on the receptive field location and computing a population level map which shows a clear figure/ground signal in LM that is maximal at the receptive field center (green circle) whereas in V1, neurons seem to prefer edges in their receptive fields



NEURAL RESPONSES IN VGG19

We analyzed responses in VGG-19 to our figure stimuli and looked for spatial structure of responses (averaged over presentations of orthogonal gratings). Like our experimental data, some neurons respond to figure edges independent of local stimulus characteristics (e.g. layer 4 similiar to V1) and some neurons display robust figure ground modulation (e.g. layer 5 similiar to LM)



*Simonyan and Zisserman 2014 ***arxiv***
image source: https://www.cs.toronto.edu/~frossard/post/vgg16/

FGM/BO INDEX

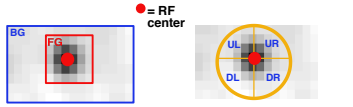
We compute a figure/ground (FG) and border ownership (BO) index as defined below.

$$BOM = max(UD_{assymetry}, LR_{assymetry})$$

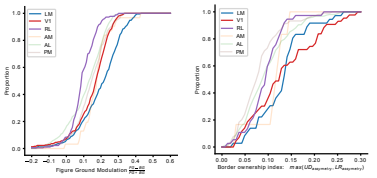
$$FGM = \frac{FG - BG}{FG + BG}$$

$$UD_{assymetry} = \frac{(UL + UR) - (DL + DR)}{UL + UR + DL + DR}$$

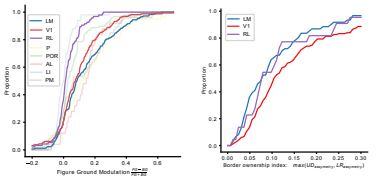
$$LR_{assymetry} = \frac{(UL + DL) - (UR + DR)}{UL + UR + DL + DR}$$



2-photon calcium imaging

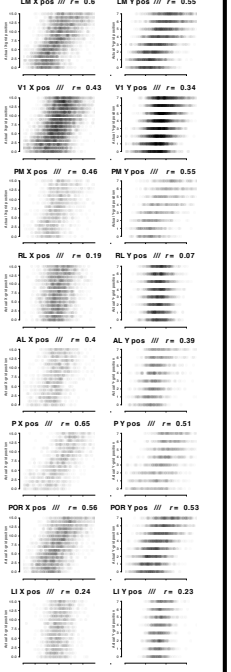


electrophysiology



LINEAR MODEL

We train linear regression to predict X and Y position across all 3 stimulus conditions using populations recorded in the same experiment.

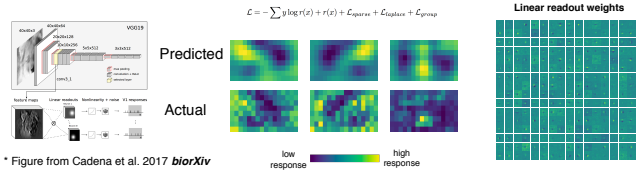


MODELING RESPONSES USING VGG19 TRANSFER LEARNING

Cadena et al. 2017, have proposed to model neuronal responses as a linear combination of intermediate representations from convolutional neural networks combined with non-linearity in poisson noise. This can be thought of as a GLM where the inputs are intermediate layer representations from a convolutional neural network thus enabling these models to capture non-linear feature combinations.

We fit responses using 10,000 natural images and pick regularization parameters based on cross validation within natural images then predict responses to the figure stimuli.

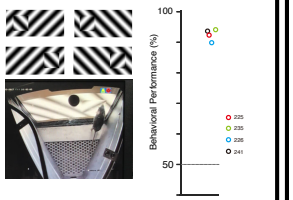
$$\mathcal{L} = -\sum y \log r(x) + r(x) + \mathcal{L}_{sparse} + \mathcal{L}_{laplace} + \mathcal{L}_{group}$$



* Figure from Cadena et al. 2017 ***bioRxiv***

BEHAVIOR

We trained animals using a touch screen paradigm to always choose the side of the screen with a texture defined figure. Mice were forced to choose between two different texture combinations (orthogonal gratings and two locations of the figure)



CONCLUSIONS

- Strongest figure ground modulation is present in LM.

- Strongest border ownership modulation is present in V1.

- Neurons in higher order visual areas display strong spatially structured responses to 2nd order figure motion.

- Neuronal populations in V1 prefer edge-related second order (shearing) motion.

- Neuronal populations in LM prefer figure-related second order motion.

- A neural network model trained for image recognition (VGG19), displays a similar trend in preference for edges and figures through the network hierarchy.

- Mice can perform texture invariant figure/ground task opening the door to understanding the full circuit by which distributed segmentation signals enable representation of a discrete object which can then trigger a motor response.

FUNDING