Investigating Mice 3D Representation Capabilities from Exposure to Monocular Stimuli

Alberghi Riccardo, Calderoni Edoardo, Casiraghi Matteo, Colombo Alessio, Forzo Mattia MV

May, 2024

Abstract

In this study, we investigate the ability of mice to perceive the three-dimensional structure of their surrounding environment by analyzing two-dimensional images perceived with a single eye. Exposing them to monocular stimuli, we aim to elucidate the role of single-eye visual input in the depth perception of mice moving in a 3D world. Our results do not evidence clear patterns and the ones that emerged are subject to very strong limitations in their conclusiveness.

1. Introduction

Prior research has established that mice can perceive 3D structures when both eyes are used (binocular vision). However, it remains unclear whether mice can achieve a similar depth perception when relying solely on monocular cues. [1] [2] [4] [5]

The goal of this study is to investigate whether mice, which navigate a three-dimensional world, can infer the 3D structure of their environment from two-dimensional images viewed with one eye.

2. Data processing

To carry out our investigation, we use the allen brain observatory's ecephys project dataset, which dimension is almost 1TB. We investigated areas of the visual cortex and decided to only work on the sessions with the highest number of neurons in them and to study the neurons' responses to two types of stimuli: natural scenes and movie frames. As next step, we extract depth maps from the images. This has been achieved relying on a deep neural network based residual model called MiDaS. [3] An example is shown in Figure 1.

Our aim is understanding mice perception of depth, necessitating the exclusion of neuronal activity responses to other image features such as contrast or light intensity. Therefore, we crop the image to its center and select only those neurons whose receptive fields are focused within this defined area of interest. Finally, we also normalize the depth values using a classical Z-Score normalization and a MinMax scaling. The combination of this two methods has both interpretation and correctness purposes and in our case a double manipulation does not affect in any way the interpretability of the values. The whole depth-extraction process is pictorially shown in Figure 3.

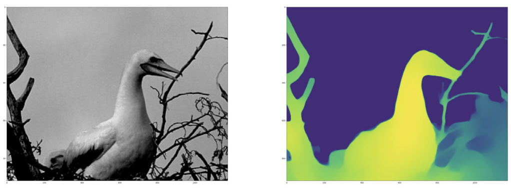

Figure 1: Example of depth estimation using MiDaS.

3. Analysis

Looking at the distribution of depths. We notice that intermediate values (ranging between 0.3 and 0.7) are far more represented than extreme values. The issue of unbalanced classes is taken into account in Section 5.

3.1. Visualizing correlations

After having checked presence of neurons with consistent firing rate via a raster plot, we look for relations between depth and firing rate. Unluckily, plotting correlation heatmaps, no general pattern seems to appear (Figure 4).

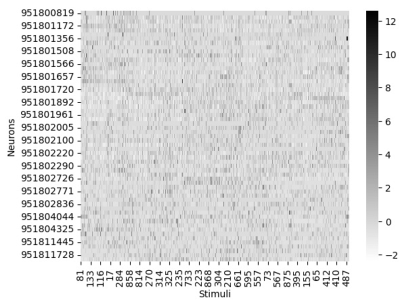

Figure 2: Neurons vs stimuli ordered by depth.

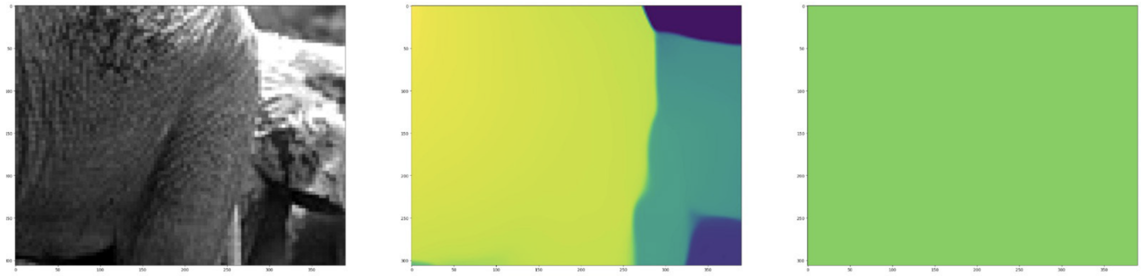

Figure 3: In order: Central crop of the image, depth of the crop, homogenization.

3.2. Selection of neurons with specific firing rate

Starting from the visualization of firing patterns for each neuron, we aim to select the ones whose firing rate looks promising for a regression task. In particular we look for neurons of mainly three different types:

1. *Peak Firing Rate Shifts*: neurons with shift in peak firing rate across different depth levels.
2. *Depth Tuned*: neurons that exhibit a clear preference for specific depth levels.
3. *Monotone Firing Behavior*: neurons whose firing rate responds monotonically to depth changes.

We extract neurons of the first two kinds fitting a Gaussian curve, while these of the third kind are detected by looking for a clear slope in the firing rate behavior.

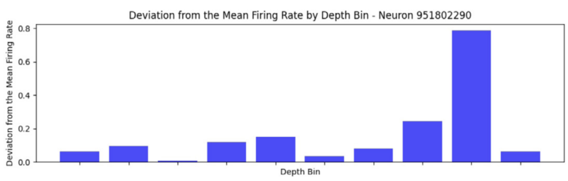

Figure 4: Example of selected neuron

3.3. Decoding Phase

We first implement a linear regression, and then a Support Vector Machine (SVM) with the aim to decode the depth value of a stimulus by looking at the neural responses. Our regression results are not at satisfying showing very low R^2 and not satisfying MSE.

Label	*Precision*	*Recall*	f1 score
Background	0.82	0.82	0.82
Center	0.75	0.81	0.78
Foreground	0.69	0.59	0.63

Table 1: SVM classification report of fig. A.5.

On the other hand, classification results seem more promising, especially when considering frames from the movie. Using a binary classifier we obtain good results on natural scenes and excellent ones on movie frames. Then, we increased the number of classes to three. The obtained results are very bad on natural scenes and still promising on on the movie frames.

4. Limitation to our study

The main limitations to our study are:

1. Receptive Field Identification
2. Mean-Depth Approximation
3. Dataset Limitations
4. Unbalanced Classes
5. Model Specificity
6. Distinction between Fore- and Background
7. Unintended measurements

A detailed explanation of them is to be found in the Appendix.

5. Conclusions

In conclusion, this study aimed to explore the capability of mice to perceive three-dimensional structures using monocular cues. Despite the sophisticated methodologies employed, including depth mapping with a pre-trained neural network, correlational analysis of neuron firing rates and depth, and various predictive modeling techniques such as linear regression and SVM classification, no definitive patterns were identified that suggest a clear perception of depth when monocular stimuli are used. These findings suggest a potential limitation in the sensory capabilities of mice or possibly in the methods available to measure such perceptions accurately. Further research may need to explore alternative approaches to uncover subtler aspects of depth perception under monocular vision conditions.

References

[1]H. C. Boone, J. M. Samonds, E. C. Crouse, C. Barr, N. J. Priebe, and A. W. McGee, ≪Natural binocular depth discrimination behavior in mice explained by visual cortical activity≫, Current Biology **31**, 2191–2198.e3, ISSN: 0960-9822 (2021) https://doi.org/10.1016/j.cub.2021.02.031.

[2]A. L. Chioma, T. Bonhoeffer, and M. Hübener, ≪Area-specific mapping of binocular disparity across mouse visual cortex≫, bioRxiv, 10.1101/591412 (2019) 10.1101/591412.

[3]dwofk, *Multiple depth estimation accuracy with single network*, https://github.com/isl-org/MiDaS.

[4]J. Poort and A. F. Meyer, ≪Vision: depth perception in climbing mice≫, Current Biology **31**, R486–R488, ISSN: 0960-9822 (2021) https://doi.org/10.1016/j.cub.2021.03.066.

[5]J. M. Samonds, V. Choi, and N. J. Priebe, ≪Mice discriminate stereoscopic surfaces without fixating in depth≫, Journal of Neuroscience **39**, 8024–8037, ISSN: 0270-6474 (2019) 10.1523/JNEUROSCI.0895-19.2019.

Appendix A. Additional Information

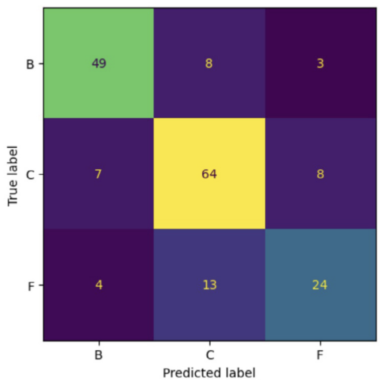

Figure A.5: Confusion matrix of SVM classification with 3 classes on movie frames.

Appendix A.1. Explanation of study limitations

1. *Receptive Field Identification*
   The study focused on neurons with receptive fields closely aligned with specific image crops. Future studies could benefit from a more refined approach to identifying these receptive fields, perhaps by weighting the neuronal response by its intensity across various image crops.
2. *Mean-Depth Approximation*
   Assigning a mean depth to each crop is not the most refined solution, even if in our case was effective enough for many frames. Future methodologies could either reduce the crop size to ensure depth uniformity or assign depth based on the most represented value within the crop.
3. *Dataset Limitations*
   The limited number of neurons with the desired receptive fields used in the study could have affected the robustness of the findings.
4. *Unbalanced Classes*
   The representation of extreme depths in the dataset was limited, which might have skewed the results.
5. *Model Specificity*
   The reliance on linear models may have constrained the study's ability to capture more complex relationships between neuronal firing rates and perceived depth. Employing more sophisticated models, such as non-linear regression or machine learning techniques, could potentially uncover patterns not detectable by linear methods.
6. *Distinction between Fore- and Background*
   The arbitrary threshold set between foreground, center, and background in the study is a notable limitation. Although cross-validating this threshold might introduce bias, considering alternative methods to define this distinction could lead to more accurate classification and understanding of depth perception.
7. *Unintended measurements*
   Probably the main limitation of this analysis lies in the possibility to conclude that the good results obtained using 3-classes SVM directly come from depth perception. Our findings may also be due to other features present in the image (contrast, light intensity, blurring, shape) and not solely to depth.

Appendix A.2. Code snippets

As follows an example of code to extract neurons with specific firing behavior.

```
def get_neuron_shift(params, t = 0.1):
    neurons = []
    for i, (mean, _, _) in enumerate(params):
      if i > 0:
        prev_mean = params[i - 1, 0]
        if abs(mean - prev_mean) > t:
          neurons.append(i)
    return neurons
```

Listing 1: Example of code to get specific neurons.