We would like to thank the reviewers for their insightful and constructive suggestions. Our reply to reviewers is below (with original reviewer comments in blue).

Reviewer #1: In this paper, the authors use computational methods to investigate the performance of neurons that are thought to mediate a range of complex behaviours in Drosophila, including pattern recognition. They use receptive field data that has recently been obtained experimentally to investigate the information that would be provided by these neurons, particularly when viewing different patterns. The authors match their findings to those of behavioural experiments and find that the output of their simulations agree well with the flies' behaviour. They then use neural network modelling to investigate how well the output of these neurons can be used to identify visual features.

This is a clear, well-written paper with a style and organisation that is easy to follow. The findings are interesting and of value to anyone interested in understanding the relationship between sensory input and behavioural output in biological systems.

I have one major comment

Page 9: First line under the title 'What information is preserved in this simple neural code'. I do not agree that the stimuli were complex scenes. I am missing a discussion on how the output of the neurons would perform in natural scenes and an elaboration of the relationship between the output of these neurons in a natural context. All of the experimental data has been acquired in artificial situations with highly simplified stimuli and it leaves me wondering about how these neurons would respond to more natural objects with the image statistics of a natural background. I understand that the lack of experimental data using such stimuli limits their ability to provide a clear answer to this question but I still think that it should be addressed in some way in the discussion. For example, could you address the question of whether it is likely that using naturalistic objects on a naturalistic background would change the output of the neurons? Could motion parallax provide additional shape information that is not present in the 2D stimuli presented?

We have made a number of amendments in relation to the reviewer's major comment. Firstly, we have changed the mentioned sentence that uses the phrase "complex scenes" so as not to imply that we understand how complex scenes would be encoded by these cells. We have introduced a new figure which shows an illustration of how the R4d neurons respond to natural images, indicating that there would be useful information given about the positions of the trees. Although, as the reviewer notes, the scope of this paper was principally to examine the RF outputs in relation to the artificial stimuli used in behavioural experiments, we agree that the broader question of the role these cells could play in a more natural context is an interesting one. Therefore, in addition to the figure we have added an extra section to the discussion, which discusses this issue.

Minor comments on the text:

Page 3, 4th line from the bottom: add 'the' before 'direction' or reorganise the end of this sentence for clarity

Page 7, 3rd line from the top: add 'B' to (Figure 2) for clarity

Page 7, 3rd line from the bottom: add 't' to 'though'

Page 14, 6th line after the heading: there is a word missing here 'Each point on the is initially...'

Page 22, Figure 1 legend: it would be helpful to the reader if it were made explicit that the blue crosses are in 1A rather than 1B

Figure 1B: the contrast between the blue and green is not very high, consider using different colours. The legends are too large and cover the data (I understand that this is symmetrical but it would help the reader to get a clearer picture of the responses to each edge).

All minor corrections were carried out.

Reviewer #2: Neuronal coding in the visual system of Drosophila melanogaster: How do small neural populations support visually guided behaviours?

General Comments:

The main contribution of this paper is the demonstration that a small population of neurones in the central complex are well suited for encoding information about size, position and orientation of objects. The details related to the extraction of kernels, calculation of R4d and R2, and the replication of behavioural experiments are well presented. The simulation tests are fairly detailed and easy to follow, and the results appear accurate. However, the motivation behind the author’s model is not well justified.

1. Previous models, based on the same data (Seeling and Jayaraman 2015) have already been proposed (e.g. A Cope et.al., reference #30). Although the authors refer to alternative models, the article lacks a systematic comparison between their model and previously proposed models. More specifically, it is not clear, why the authors have not used a ring attractor.

There appears to be some confusion here. We are modelling the data in Seelig and Jayaraman's 2013 paper, which looked at the visual receptive fields of ring neurons cf. their 2015 paper which looked at a ring attractor circuit in the ellipsoid body (both in Nature). We have, however, added a section in the discussion which discusses our modelling in the context of central complex organisation. This brings together the papers and talks about the inputs to ring neurons and their role in the ring attractor etc.

2. A tremendous amount of information is already known about the visual system of Drosophila. However, the authors have modelled R4d and R2 neurones with simple linear filters, by using kernels. It will be beneficial to at least discuss why the use of linear filters is appropriate.

Seelig and Jayaraman used the RFs in this way as linear convolutions (to verify the RF forms). We have added some discussion of this to Materials and methods. We have additionally justified the use of linear filters in the main text (in the Results section).

3. The authors compare the discriminability of patterns using a distance measure between Rfs. This approach is relatively similar to methods suggested by the author themselves in other articles for place recognition or navigation (e.g. Dewar, A.D. et. al, 2014. What is the relationship between visual environment and the form of ant learning-walks? An in silico investigation of insect navigation. Adaptive Behavior). It is frustrating that authors do not link their present model to other existing models of place recognition and pattern recognition.

Also, the authors often refer to which patterns flies can or can not discriminate. It should be noted that the discrimination capabilities of animal heavily depend on the training method (see for example A Dyer et. al. 2004 Bumblebees (Bombus terrestris) sacrifice foraging speed to solve difficult colour discrimination tasks. J.Comp.Physiol.A 190, doi:10.1007/s00359-004-0547-y) .

Finally, in the last paragraph of ‘Pattern discrimination in flies and ring neuron population codes’, the authors suggest that the addition of Rfs improve the patterns discrimination. However, the data are not shown without apparent reasons.

We have now discussed our choice of the root mean square difference function in the methods, where it is also noted that there are other possibilities. We have not cited Dewar et al. as we did not look at place homing in this work and so we felt the connection was a little tenuous. We did previously carry out some different modelling work looking at homing that made use of the ring neuron RFs (Dewar et al., Navigation-specific coding in the visual system of Drosophila, Biosystems, 2015) and this is mentioned in the text. In the new discussion section, we now discuss previous models of pattern recognition (e.g. Horridge).

We have now clarified (in the Results section) that when we talk of the ‘discriminability’ we are using the term in a narrow sense, but that for real flies it will depend on task etc.

Data showing a boost in discriminability for certain pattern pairs with the addition of extra RFs, which was not previously shown, now appears as two new panels in Fig. 2.

Specific Comments:

Page 4 (Results, first paragraph): The author should give an intuitive explanation for the use of average filters. The sentence ‘Though this averaging … we examine’ is not justified properly.

Justification for averaging added to text. See also final point.

Page 6 (Results: Pattern discrimination in flies and ring neurone population codes, paragraph 2): The root mean square is used as discriminability measure without motivation and therefore sounds arbitrary. It should be noted that other metrics exist: absolute difference, kullback-leibler, bayesian information criterion, etc.

It is true that the choice of difference metric (r.m.s. difference) was somewhat arbitrary, though we didn't select this measure because it performs better. We have added a comment about this to the methods section.

Page 13 (Discussion, Do flies recognise patterns?, paragraph 3): The author used we in ‘… we have shown that the information ...’ although referring to published work from different authors, reference #2 previously

Corrected.

Page 14 (Materials and methods, Turning visual receptive field data into visual filters). The authors should ask Seeling and Jayaraman for their original data instead of extracting values from figures for better precision and clarity.

Responses to individual comments are listed below, using the reviewer's original numbering system.  
We have not obtained the original data for the RFs. We were already averaging across flies so small changes in the accuracy of individual RF recordings would be unlikely to make much difference. There's sufficient averaging in the process that getting the originals would not add much. Similarly, other processing steps rely on averaging and reduce the significance of using original data. Finally, in the course of modelling research for our other paper (Dewar et al., Navigation-specific coding in the visual system of Drosophila, Biosystems, 2015), variants on the averaging process were used and have shown negligible effect on results.