An analysis of a ring attractor model for cue integration

Xuelong Sun[1], Michael Mangan[2]*, and Shigang Yue[1]*

[1] Computational Intelligence Lab, School of Computer Science, University of Lincoln, Lincoln, UK 15612083@students.lincoln.ac.uk
http://www.ciluk.org/
[2] Sheffield Robotics, Dept. of Computer Science, The University of Sheffield, Sheffield, S1 4DP, UK
* joint last authors

Abstract. Animals and robots must constantly combine multiple streams of noisy information from their senses to guide their actions. Recently, it has been proposed that animals may combine cues optimally using a ring attractor neural network architecture inspired by the head direction system of rats augmented with a dynamic re-weighting mechanism. In this work we report that an older and simpler ring attractor network architecture, requiring no re-weighting property combines cues according to their certainty for moderate cue conflicts but converges on the most certain cue for larger conflicts. These results are consistent with observations in animal experiments that show sub-optimal cue integration and switching from cue integration to cue selection strategies. This work therefore demonstrates an alternative architecture for those seeking neural correlates of sensory integration in animals. In addition, performance is shown robust to noise and miniaturization and thus provides an efficient solution for artificial systems.

Keywords: ring attractor · cue integration · sensor fusion · optimal · Bayesian integration · head direction cells

1 Introduction

A fundamental principle underlying animal intelligence is the capacity to appropriately combine redundant sensory information (e.g. vision, olfactory and haptic) of the same percept (e.g. location of a sensory source) to achieve a more accurate and robust estimate [1, 2]. For example, both mammals and insects constantly track their pose using head-direction cells which combine information from external cues (e.g. from surrounding visual features) with self-motion cues (from path integration) to maintain a precise estimate of their current orientation [3, 4] (Fig.1). Yet, as all sensory information is subject to errors which can change drastically depending on the situation (e.g. relying on visual cues in a darkened room) animals must employ an adaptive cue combination strategy reflecting the known errors (variance) in the different sensory signals to achieve the optimal estimate of the desired environmental property.

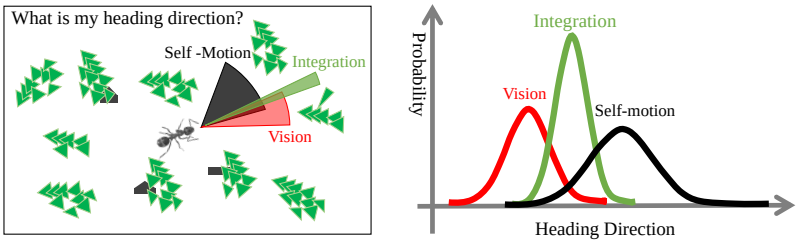

Fig. 1. The cue integration problem. Left: An example of an animal maintaining an estimate of it's current pose (green area) using different cues of varying certainty (Self-motion, black area, and vision, red area). Right: cues can be represented by conflicting Gaussian functions with the width describing the uncertainty of each and the optimal solution (green) given by weighting each cue according to their known variance as described by Bayes' rule.

Bayes' theorem (1) provides a mathematical framework describing the optimal way in which information from different sources should be combined, and it has been argued that animals have Bayesian brains [2, 5–7]. According to Bayes' Rule, the posterior probability $P\left(x_{true}|x_{cue}\right)$ (the probability of event x will happen when the cue about x is sensed) is proportional to the product of the prior probability $P\left(x_{true}\right)$ (the probability of event x happening based on prior knowledge) and the likelihood function $P\left(x_{cue}|x_{true}\right)$ (the probability of the cue when x truly happened, which represents the reliability of this cue). Assuming that the prior probability $P\left(x_{true}\right)$ is uniform and x_{cue} is corrupted by Gaussian noise with variance σ^2, then the posterior probability is proportional to $1/\sigma^2$. Therefore, when there are n cues all concerning x event and corrupted by Gaussian noise with variance $\sigma_i^2, i = 1, 2, ...n$, the optimal way to reduce the uncertainty of estimating x (i.e., the maximum the posteriori probability) is averaging the cues weighted by their reciprocal variances $1/\sigma^2$, as indicated by (2), which is identical with results calculated by the maximum-likelihood estimate (MLE) [6, 8]. The theorem asserts that cues with low variance (i.e more reliable) should be weighted more than those with high variance (i.e less reliable) as demonstrated in Fig.1.

$$P\left(x_{true}|x_{cue}\right) = P\left(x_{cue}|x_{true}\right) P\left(x_{true}\right) / P\left(x_{cue}\right) \tag{1}$$

$$\hat{X} = \sum_{i}^{n} W_i X_i, W_i = \left(1/\sigma_i^2\right) / \left(\sum_{j}^{n} 1/\sigma_j^2\right) \tag{2}$$

Artificial systems must also solve the same problem although in robotics it is commonly known as sensor fusion. Sensor fusion for mobile robot navigation is a long standing issue and many statistical methods based on maximum *a posteriori*

and maximum-likelihood estimation have been applied to solve it [10]. Recent advances in deployment of robot systems such as cars with their suite of GPS, radar, cameras, and laser scanning sensors to estimate precise lane position, owe much to adoption of probabilistic integration of cues in line with the Bayesian formulation described above [11]. Yet, current SLAM methods [12], tend to be computationally expensive and unsuitable for application on small, cheap robot platforms. Learning from biology may bring significant benefits for solving these problems in artificial systems.

We therefore take a bio-inspired approach to firstly understand how animals resolve this task, which in turn may offer inspiration to engineers seeking efficient solutions. As a starting point, we use a classic neural network architecture known as a ring attractor network. Ring attractors can be constructed such that the output activity resembles a Gaussian profile that is maintained even in the absence of sensory input. When new sensory input is presented, the activity profile will shift towards and stabilize at the new location. If this sensory input is driven by orientation cues such as path integration or visual features then the Gaussian mean will naturally track the animal orientation. Such networks have been proposed to underpin the head-direction cells in animals [14, 15]. Further when more than one input signal is presented ring attractors can be constructed such that the output settles on the weighted average of the combined cues as required for optimal cue integration [16]. In a recent review, Jeffery et al [17] proposed that ring-attractor networks may provide a general architecture for optimal cue integration. Their biomimetic model (constrained by physiological data from rats) used a re-weighting mechanism to achieve optimal integration. Specifically, in the region where conflicting cues overlap, Hebbian learning rapidly strengthens local sysnapses causing peak activity to shift towards the position consistent with optimal integration.

In this study, we revisit the Touretsky [16] ring attractor network and assess its ability to combine conflicting cues of different strengths. Specifically we seek to assess how this network performs when given cues of different strengths and with different levels of conflict. Further, we wish to document if and when the network optimally integrates cues or if it adheres to a winner-takes-all (WTA) solution, or switches strategy depending on the situation.

Our results suggest that a Touretsky ring attractor network can integrate cues in a manner approaching optimal (i.e, consistent with MLE) for small conflicts. For larger conflicts the network switches to WTA mode, mirroring results of ethological experiments [18]. Performance is shown to be robust to noise and significant reduction in the network size, and thus provides a simple (no re-weighting mechanism required), compact ring attractor solution to cue integration that can provide inspiration for those seeking similar integration networks in animals or act as a bio-inspired method for optimal sensor fusion in robots.

2 Models and Methods

2.1 Touretsky Ring Attractor Model

Artificial Neurons The network is constructed using two populations of CTRNN (continuous time recurrent neural network) neurons which are simple nonlinear and continuous dynamical neurons suitable for simulating the subset of real numbers as required for our ring attractor model [19]. The average membrane potential c_i of a CTRNN i^{th} neuron is updated by the differential equation (3), where τ is the positive time constant and I_i is the total number of inputs into the neuron which equals the weighted sum of other neurons' outputs $O_j, j = 1, 2, ...n$ and the external inputs, as shown in (4), where W_{ji} is the weight matrix representing the connection strength from j^{th} to i^{th} neuron, g is the activation function and X_i is the external input. To acquire the nonlinear property of the network, the activation function of g should be a nonlinear function. Here we simply applied a semi-linear threshold function with a threshold defined by θ as indicated in (5).

$$\tau \frac{c_i}{dt} = -c_i + I_i \tag{3}$$

$$I_i = \sum_{j=1}^{n} W_{ji} O_j + X_i = \sum_{j=1}^{n} W_{ji} g(c_j) + X_i \tag{4}$$

$$g(c) = max(0, \theta + c) \tag{5}$$

Network Geometry We implemented a variant of the classic ring attractor network [14] (Fig.2(a)) which replaces the inhibitory interneurons with a single global inhibitory (uniform inhibitory) neuron making the network easier to tune while giving the same performance [16]. Each excitatory neuron in the network has recurrent excitatory connections to all other neurons in the ring with weights decreasing with distance which is crucial for generating the bell-shape activation profile in stable state, as revealed in (6), where d_{ij} is the distance between the i^{th} and j^{th} neuron. Our network posses a single dynamic inhibitory neuron that sums inputs from the excitatory neurons and then proportionally inhibits the entire network. Note, for ease of understanding the recurrent connections from a single excitatory cell are shown in Fig.2 (a) but in reality each neuron has the same set of recurrent connections.

$$W_{ji}^{E \to E} = e^{\frac{-d_{ji}^2}{2\sigma^2}} \tag{6}$$

Fig.2 (b) shows the process by which the network combines input from multiple cues. We simulate cues of different strengths using Gaussian functions (see equation (7) where K is the scale factor, μ defines the peak position of the Gaussian curve (estimation of the certain property based on the cue) and σ^2 is the variance of the Gaussian function determining the reliability of the signal).

To have a corresponding connection with the integration neurons in the attractor, the cues are represented by the activation profile of N neurons with their preference p_i, and so the Gaussian curve is sampled by N points at intervals. This input is then passed to the integration population which is shown in unwrapped form in Fig.2 (b), and with the recurrent connections omitted for ease of reading. The integration population (and also population representing cue 1 and cue 2) has N neurons labeled with their preferences (for example, if these neurons represent the heading directions of the animal, the preferences will be the preferred directions evenly distributed around the entire 360° of possible directions). The inhibitory population has a single dynamic postulated inhibition neuron summing the activations from all integration neurons and which recurrently inhibits all integration neurons. Therefore, in accordance with equations (3) - (5), the average membrane potential of the output neurons (neurons in integration population) is computed by equation (8), where $X1$ and $X2$ represent the activation vectors of cue 1 and cue 2 respectively and u is the membrane potential of the uniform inhibitory neuron (calculated by equation (9)). Note that in order to maintain the nonlinear property and simultaneously guarantee the positive output of the model, we tuned the total input I to c and u using function g according to [16].

Note that in this paper, as an example, we use the ring attractor to represent the heading direction system so all the values have the unit-degree. But generally the unit could be other meanings when this model is applied to other specific contexts.

$$F(i) = \frac{K}{\sqrt{2\pi}\sigma} e^{-\frac{(p_i-\mu)^2}{2\sigma^2}} + \xi N(0,1), i = 1, 2, ...N \tag{7}$$

$$\tau \frac{dc_i}{dt} = -c_i + g\left(\sum_{j=1}^{n} W_{ji}^{E\to E} c_j + X1_i + X2_i + W^{I\to E} u\right) \tag{8}$$

$$\tau \frac{du}{dt} = -u + g\left(W^{I\to I} u + W^{E\to I} \sum_{k=1}^{n} c_k\right) \tag{9}$$

3 Results

Fig.3 (a) shows the response of our ring attractor network configured with 100 neurons when stimulated with two conflicting cues (65° apart) with different variances ($\sigma_{cue1} = 40°, \sigma_{cue2} = 35°$) shown by red and black lines. The response of the network (green line) approaches the MLE, i.e., the optimal integration (blue dashed line) rather than following the WTA solution. Fig.3 (b) shows that the network response is robust to noise with each cue corrupted by Gaussian white noise ((7) with $\xi = 0.01$). Finally, inspired by recent anatomical results showing that insects encode their heading direction using populations of only

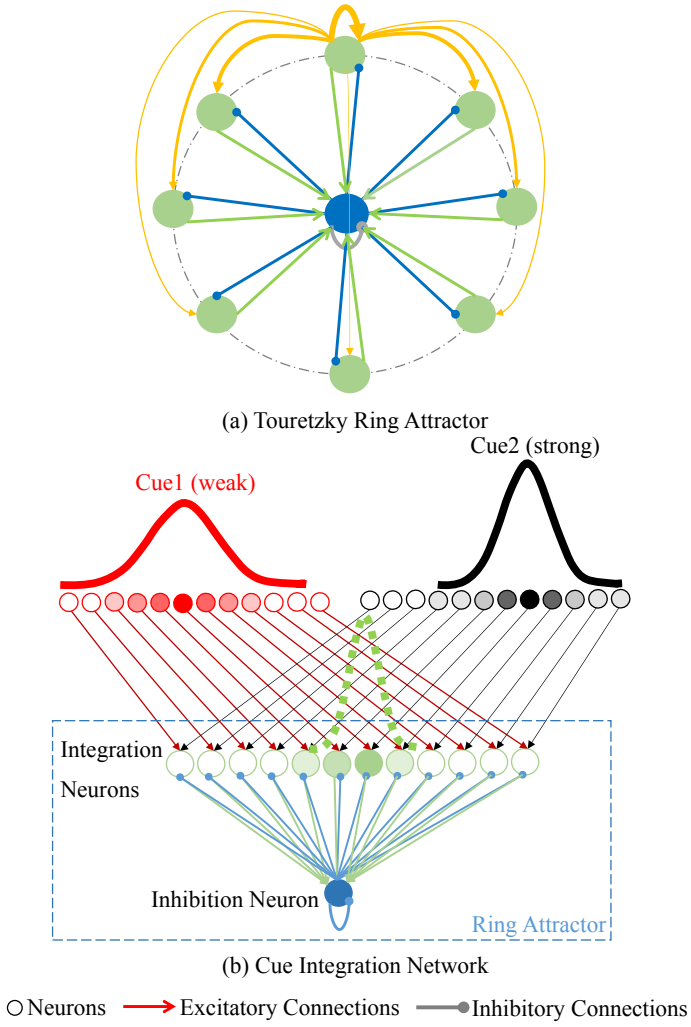

Fig. 2. The implemented Touretzky ring attractor network. (a) Excitatory neurons are shown by green circles, and the global inhibitory neuron depicted by the blue circle. The recurrent excitatory interneurons are shown by orange arrows with connection strength decreasing with distance between neurons. Excitatory and inhibitory connections between the global inhibitory neuron are also shown in blue and green respectively. (b) The full integration network shown in unwrapped form (minus recurrent connections for ease of reading) with example inputs and optimal output overlaid.

8 directional neurons in each hemisphere of the central complex [20, 21] we reduced the number of neurons in our integration network from 100 to 8. Fig.3 (c) demonstrates that the cue integration properties of the network remained stable despite the obvious loss of resolution in the Gaussian functions.

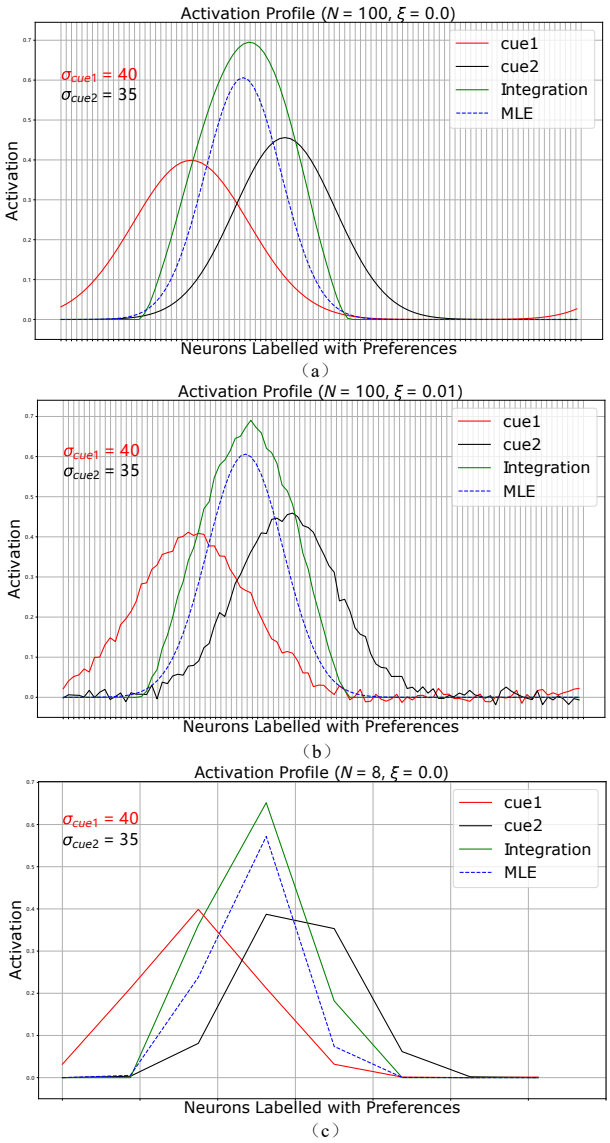

Fig. 3. Integration of conflicting cues by a ring attractor network. Activation profiles of cues are shown by the red and black curves, the output profile of RA (the ring attractor) by the green line, and the MLE by the blue dashed line. (a) shows the results for a noise-free network with 100 neurons, (b) shows the results of the same network with added white noise, and (c) show the results when the number of neurons is reduced to 8.

To assess the performance of the network across likely scenarios we performed two more experiments using the noise free network with 100 neurons. Firstly, we

assessed the performance of the network when presented with cues that were increasingly disparate. Cue 1 ($\mu_{cue1} = 0°, \sigma_{cue1} = 40°$) was presented at the same position throughout the tests, while Cue 2 was presented at increasingly distant positions (from 0° to 180° in 5° steps). We performed this analysis under three conditions: (a) cues with identical variance ($\sigma_{cue1} = \sigma_{cue2} = 40°$); (b) cues with slight differences in variance ($\sigma_{cue1} = 40°, \sigma_{cue2} = 35°$); and (c) cues with significantly different variance ($\sigma_{cue1} = 40°, \sigma_{cue2} = 20°$). Fig.4 (a-c) shows the peak response of the network (green line) overlaid on the MLE (blue line) and WTA (red line) solutions. With cues of equal variance (Fig.4 (a)) the network response approaches (though never very precisely matches) the MLE solution but changes to WTA-like responses when cue-conflict exceeded approximately 100°. With small differences in variance (Fig.4(b)), the network again weights cue in an approximately optimal manner but shifts to a WTA response at higher values ($> 110°$). In contrast when more significant differences in variance were presented (Fig.4 (c)), the network changes from the MLE to WTA response at much smaller conflicts ($> 60°$).

Fig.4 (d) shows data from a previous cue-combination experiment in rats [18] (black line) overlaid with the bioligically constrained ring attractor network with re-weighting mechanism [22] (orange line) as the cue conflict is increased as in our experiment. We note that our model response (Fig.4 (a)) adheres closely to the animal data.

Secondly we assessed the performance of the network when the certainty of one cue was altered while the other was held constant. Specifically, cue 1 and cue 2 were presented 90° apart. While cue 2 variance was held at 40°, the variance of cue 1 was increased from 5° to 200°. Fig.5 shows the peak position of the activation profile of the network changes from a WTA state for uncertainty of cue 1 below 15° and above 160° but performs a weighted average when uncertainty of cue 1 in the range 20° to 155°. Thus, although not acting in a truly optimal manner the switch from WTA to weighted-average and back again follows the general profile of the MLE prediction.

4 Discussion

In this work, we re-visited the classic ring-attractor network described by Touretsky [16] to assess if it could be configured for optimal cue integration, and if so whether this might give inspiration for those seeking such networks in animals or provide a biologically-inspired solution for robotics.

We report that that our implementation of the classic Touretsky ring attractor network perform optimal-like cue integration when presented with conflicting cues rather than tending to a winner-takes-all solution as often cited. The network output is also shown to be robust to noise on the sensory input and to reduction to 8 neurons encoding direction (as in insects). Our sweep tests showed that both the variance and distance between conflicting cues strongly affect the network properties. With equal or small differences in variance of cues the network performs a weighted average for small cue conflicts, but switches to

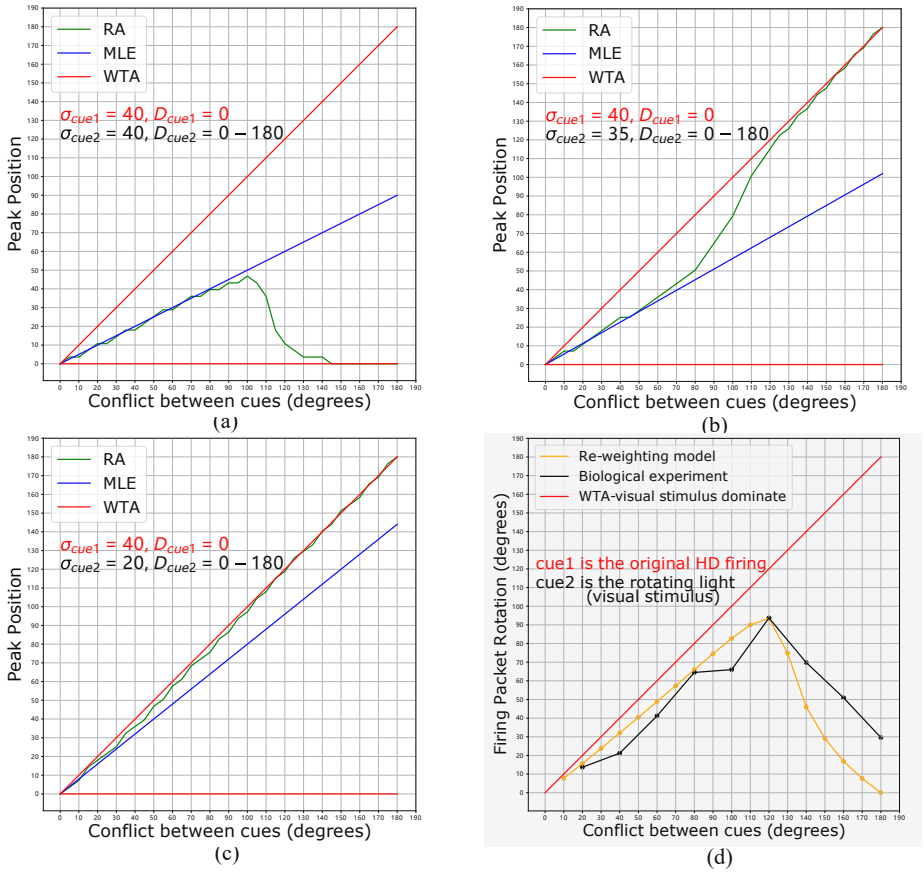

Fig. 4. Network performance with increasing cue conflict. Cue 1 was presented in the same location while cue 2 was presented at increasing distances. For (a), (b) and (c), the response of the RA (ring attractor) is shown by the green line, the WTA prediction by the red line, and the MLE by the blue line. (a) Cues of equal variance (b) Cue 1 with slightly higher variance than Cue 1, (c) Cue 1 with significantly higher variance than cue 1. (d) data from similar cue combination study in rats (black line) [18] and alternative re-weighting model (orange line) [22] (data provided with thanks by Dr. Hector Page and Prof. Kate Jeffery).

a WTA response for larger conflicts. For larger differences the network switches to WTA responses at much small conflicts. This changing of response is akin to meta-Bayesian decision making where it is highly sub-optimal to integrate two hugely conflicting cues e.g. one should not go West, when one cue states North and the other South. Instead one should choose the best single option, but how does the agent know when to apply each strategy? We show that the [16] ring attractor network inherently possesses this capacity.

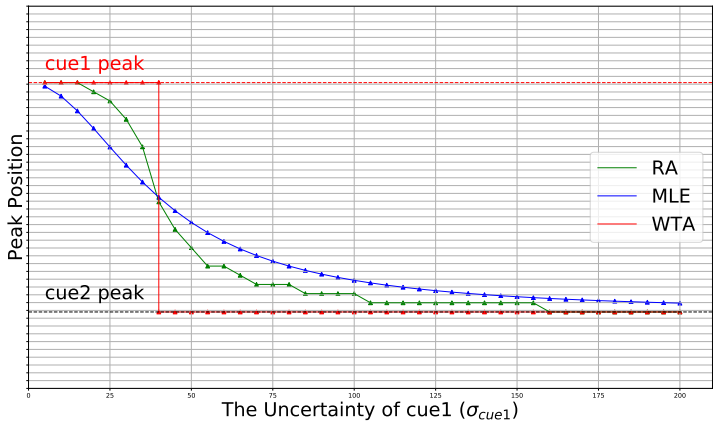

Fig. 5. Network performance with changes in cue variance. Cue 1 and cue 2 were presented at the same location 90° apart. The variance of cue 2 was kept at 40° while cue 1 changed from 5° to 200° in intervals of 5°. The position the peaks of activation profile of cues 1 and 2 are shown by the dashed red and black lines respectively; the WTA response by the solid red line; the MLE by the blue line; and the RA (ring attractor) output by the green line.

Over two decades ago the ring attractor network was proposed as a possible solution underpinning the head direction cells in mammals that integrate directional cues from different sensory modalities to maintain an accurate readout of their current orientation [14]. Recent models of the head direction cells of mammals have moved away from the original ring attractor architecture because the physiology does not mirror the excitatory interconnections required by the original model [22, 23], and the belief that ring attractors will tend to a WTA outcome over the weighted-average observed in behavioral experiments [17]. Through augmentation of these models with a re-weighting mechanism (Hebbian learning) [22] it has been proposed that this network architecture may be a ubiquitous neural circuit underlying optimal cue integration across many functions [17]. Here we provide new evidence that the original ring attractor network can also perform weighted cue integration (closely matching the performance of the re-weighting network [22] Fig.4 (d)), or cue selection in a manner closely approximating data from rats [18] (Fig.4 (d)).

Direction cells have recently been revealed in insects (*Drosophila*) with so-called E-PG neurons forming a bump of activity that moves in response to both rotation of vision cues and self-motion and combines in both cue selection or cue integration like a averaged weighted [4]. These E-PG neurons have also been shown to have ring attractor dynamics [24]. Biomimetic models constrained by the anatomy of the animal have successfully recreated the activation phenomena of behavioral experiments [25, 26] but have not, as yet, been extended to the broader cue integration problem discussed here and in [17]. We note that [25] showed that fixed connection weights are sufficient to track the self-motion and

visual cues well with the dynamic re-weighting with slower learning rates giving improved performance describing a trade-off between computational complexity and required robustness.

By analyzing the Touretzky ring attractor network, we show that it should still be considered a biologically plausible mechanism to achieve cue integration of the animals. Although not well suited to describe the head-direction system of mammals due to physiological constraints, it is an open question whether other areas of animal brains that perform cue integration may use this ring attractor architecture. For instance, the lateral accessory lobe (LAL) of insects brain, which is a converging point of sensory information and has inputs from sensory lobes, mushroom body and the central complex [27, 28] provides a candidate to search for such network architectures. To date, we know little about how different cues (like vision memory from mushroom body and path integration from central complex) might be integrated in this area and wherein ring attractors may also play crucial roles. As a bio-inspired neural network, the ring attractor is a compact but efficient model to solve the similar problems in sensor fusion and its anti-noise and stable performance with only 8 neurons endow it the advantage of implementation on robots with limited computation resources.

Acknowledgments

This work was supported by EU FP7 projects HAZCEPT (318907), HORIZON 2020 project STEP2DYNA (691154). We also thank Prof. Kate Jeffery and Dr. Hector Page for provision of data shown in Fig.4 (d).

References

1. Shettleworth, S. J. (2010). Cognition, evolution, and behavior. Oxford University Press.
2. Ernst, M. O., & Knoblich, G. (2006). A Bayesian view on multimodal cue integration. Human body perception from the inside out, 131, 105-131.
3. Blair, H. T., & Sharp, P. E. (1996). Visual and vestibular influences on head-direction cells in the anterior thalamus of the rat. Behavioral neuroscience, 110(4), 643.
4. Seelig, J. D. & Jayaraman, V. (2015). Neural dynamics for landmark orientation and angular path integration. Nature, 521,186191. DOI: 10.1038/nature14446
5. Cheng, K., Shettleworth, S. J., Huttenlocher, J., & Rieser, J. J. (2007). Bayesian integration of spatial information. Psychological bulletin, 133(4), 625.
6. Kording, K. P., & Wolpert, D. M. (2004). Bayesian integration in sensorimotor learning. Nature, 427(6971), 244.
7. Kording, K. P. (2014). Bayesian statistics: relevant for the brain?. Current opinion in neurobiology, 25, 130-133.
8. Ernst, M. O., & Banks, M. S. (2002). Humans integrate visual and haptic information in a statistically optimal fashion. Nature, 415(6870), 429.
9. Cox, I. J. (1991). Blanche-an experiment in guidance and navigation of an autonomous robot vehicle. IEEE Transactions on robotics and automation, 7(2), 193-204.

10. Kam, M., Zhu, X., & Kalata, P. (1997). Sensor fusion for mobile robot navigation. Proceedings of the IEEE, 85(1), 108-119.
11. Thrun, S., Burgard, W., & Fox, D. (2005). Probabilistic robotics. MIT press.
12. Fuentes-Pacheco, J., Ruiz-Ascencio, J., & Rendón-Mancha, J.M. (2015). Visual simultaneous localization and mapping: a survey. Artificial Intelligence Review, 43(1), 55–81
13. Lambrinos, D., Mller, R., Labhart, T., Pfeifer, R., & Wehner, R. (2000). A mobile robot employing insect strategies for navigation. Robotics and Autonomous systems, 30(1-2), 39-64.
14. Zhang, K. (1996). Representation of spatial orientation by the intrinsic dynamics of the head-direction cell ensemble: a theory. Journal of Neuroscience, 16(6), 2112-2126.
15. Skaggs, W. E., Knierim, J. J., Kudrimoti, H. S., & McNaughton, B. L. (1995). A model of the neural basis of the rat's sense of direction. In Advances in neural information processing systems (pp. 173-180).
16. Touretzky, D. S. (2005). Attractor network models of head direction cells. Head direction cells and the neural mechanisms of spatial orientation, 411-432.
17. Jeffery, K. J., Page, H. J., Stringer, S. M. (2016). Optimal cue combination and landmark-stability learning in the head direction system. The Journal of physiology, 594(22), 6527-6534.
18. Knight, R., Piette, C. E., Page, H., Walters, D., Marozzi, E., Nardini, M., ... & Jeffery, K. J. (2014). Weighted cue integration in the rodent head direction system. Philosophical Transactions of the Royal Society of London B: Biological Sciences, 369(1635), 20120512.
19. Beer, R. D. (1995). On the dynamics of small continuous-time recurrent neural networks. Adaptive Behavior, 3(4), 469-509.
20. Pfeiffer, K., Homberg, U. (2014). Organization and functional roles of the central complex in the insect brain. Annual review of entomology, 59, 165-184.
21. Heinze, S. (2017). Unraveling the neural basis of insect navigation. Current opinion in insect science.
22. Page, H. J., Walters, D. M., Knight, R., Piette, C. E., Jeffery, K. J., & Stringer, S. M. (2014). A theoretical account of cue averaging in the rodent head direction system. Philosophical Transactions of the Royal Society B: Biological Sciences, 369(1635), 20130283.
23. Page, H. J., Walters, D., & Stringer, S. M. (2015). Architectural constraints are a major factor reducing path integration accuracy in the rat head direction cell system. Frontiers in computational neuroscience, 9, 10.
24. Kim, S. S., Rouault, H., Druckmann, S., & Jayaraman, V. (2017). Ring attractor dynamics in the Drosophila central brain. Science, 356(6340), 849-853.
25. Cope, A. J., Sabo, C., Vasilaki, E., Barron, A. B., Marshall, J. A. (2017). A computational model of the integration of landmarks and motion in the insect central complex. PloS one, 12(2),e0172325.
26. Kakaria, K. S.de Bivort, B. L. (2017). Ring attractor dynamics emerge from a spiking model of the entire protocerebral bridge. Frontiers in behavioral neuroscience, 11.
27. Homberg, U. (1994). Flight-correlated activity changes in neurons of the lateral accessory lobes in the brain of the locust Schistocerca gregaria. Journal of comparative physiology A, 175(5), 597-610.
28. Barron, A. B., & Klein, C. (2016). What insects can tell us about the origins of consciousness. Proceedings of the National Academy of Sciences, 113(18), 4900-4908.