Published in final edited form as:

Curr Opin Neurobiol. 2014 April ; 25: ix–xii. doi:10.1016/j.conb.2014.02.001.

The receptive field is dead. Long live the receptive field?

Adrienne Fairhall

Department of Physiology and Biophysics 1705 NE Pacific St. [?] HSB G424, Box 357290
University of Washington Seattle WA 98195-7290 USA Phone: 206.427.1557

Abstract

Advances in experimental techniques, including behavioral paradigms using rich stimuli under closed loop conditions and the interfacing of neural systems with external inputs and outputs, reveal complex dynamics in the neural code and require a revisiting of standard concepts of representation. High-throughput recording and imaging methods along with the ability to observe and control neuronal subpopulations allow increasingly detailed access to the neural circuitry that subserves these representations and the computations they support. How do we harness theory to build biologically grounded models of complex neural function?

Sensory neurophysiology is dominated by the concept of stimulus representation. Our senses are impinged upon by a variety of stimuli. The nervous system captures these stimuli and filters them to extract and encode a myriad of features. These features are thought to be assembled and integrated across sensory modalities to form representations of increasing complexity, specificity and invariance. These hierarchically organized representations then in principle become accessible to perception and enable the lifelong construction and updating of internal models of the world about which we reason and which provide a basis for invention and imagination.

The concept of a heterogeneous feature basis that becomes increasingly sophisticated as it is propagated hierarchically [1] has gained powerful traction because of the extraordinary finding that the responses of many individual sensory neurons are indeed intelligible: it is frequently possible to find stimulus parameters with respect to which the response of a neuron varies systematically and fairly repeatably. While most clearly elaborated in the visual system, this picture roughly recurs across almost all sensory domains [2,3]. Olfaction may be a counterexample [4], in the sense that representations of intermediate complexity do not appear to exist [5]. Much theoretical progress has been made in the past several decades in developing methods to mine input/output data to determine variations of cascade models, which identify linear filters that extract relevant stimulus components, and predict the firing rate as a nonlinear function of the filtered stimulus [6]. In some cases, such models

© 2014 Elsevier Ltd. All rights reserved.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

give excellent predictions of responses to restricted stimulus sets [7] from feature-based models.

Furthermore, theory has addressed not just what is encoded, but why the encoded features may assume the form they do. Two key principles have emerged: that these features may provide an efficient way to represent the specific statistical structure of the natural world [8], and that neural representations are sparse, in the sense that any natural input can be represented by the activation of relatively few neurons [9]. Further, it has been proposed that neural systems might use representations that facilitate computation [10,11] and that processes like adaptation can dynamically enhance the quality of representations [12,13,8]. The utility of such feature representation is validated by the rapid advance of “deep learning” networks in machine intelligence [14], which instantiate the principles of hierarchical feature selection learned from natural data, emergent high order features, and distributed and sparse representations. These advances have resulted in engineered networks that are now able to perform object and speech recognition tasks with unprecedented accuracy.

While this picture of sensory representation is compelling, there are many important caveats-- ones that will assume greater significance as more experiments move toward recording during natural behavior. The success of basic coding models in predicting responses is generally limited to certain stimulus regimes: models fitted using, for example, white noise do not generally accurately predict responses to natural inputs [15,16]. Even in the retina, the poster child for successful neural coding paradigms, the observation of complex feature selectivity such as sensitivity to figure/ground differences and multiple adaptation timescales has triggered the deployment of hybrid coding/dynamical models [16]. The hierarchical feature model, and its machine learning analog, is essentially feed-forward. In reality, feedback plays an enormous though not yet well-understood role in modulating responses by behavioral state, top-down effects and contextual cues [3], multimodality [17] and through interaction with signals of self-motion [2].

Given these complexities, is there an alternative way to think about neural representation? The diverse approaches to computational neuroscience that are expressed in this issue at times expose a tension between two paths to understanding brain function, one which might be seen as originating in computer science and the other in physics. In a computer science formulation, a circuit element implements an algorithmically defined function, a step in a logical chain. From a physical perspective, the state of such an element evolves according to dynamics specified by its interactions; neural circuits can be modeled as a set of differential equations driven by continuous and analogue inputs. This distinction could also be framed as one of function and mechanism. Of course, this somewhat artificial dichotomy between physics and computation becomes obvious when moving from the nervous system to the body: neural signals interface with biomechanics, which provide a fundamental contribution to the transformation from sensory inputs into behavior [18,19,2].

To bridge function and mechanism, we suggest an embellishment of Marr's famous three-level schema of Computation, Algorithm, Physical Implementation to give physics a more

prominent role in further unpacking “implementation” into the true physical substrate and a comprehensible dynamical mechanism that can help to “explain” computation.

The three-part picture of Fig. 1 has been most extensively elaborated at the single neuron level. Experimentally well-founded conductance-based models describe the evolution of the voltage of single neurons as a function of inputs, depending on ion channel densities and morphology. These models can be reduced, analytically or numerically, to much simpler and highly predictive low-dimensional dynamical systems [20]. The resulting low-dimensional system is then amenable to analysis, leading to a coding model that expresses its computational properties [20,21,22]. Spike-triggering features approximately arise from the local linearization of the underlying nonlinear dynamical system. The threshold nature of excitability privileges certain stimulus components, reducing the dimensionality of the relevant feature space. Thus, quasi-linearity of the dynamical system establishes the system's filtering properties or feature selectivity, while the nonlinearity of spiking reduces the intrinsic dimensionality of the feature space [21].

The example of the single neuron, a fundamental unit of information encoding, highlights the duality between a dynamical system and a feature-selecting coding model. Any choice of coding model “queries” the dynamical system in order to generate an input/output relationship with respect to a specific variable or set of variables. Despite the sophisticated methods available to guide the selection of this variable set [6], the result is necessarily an impoverishment of the full behavior of the nonlinear system. An example is that of contrast gain in single neurons. When stimulated by inputs that vary over a certain range, the input/output function of many sensory systems depends on the stimulus range: the dynamic range of the response is matched to the input range [13]. Some single neurons show the same effect, demonstrating that the property can arise from intrinsic neuronal nonlinearities [22]. Identifying a low-dimensional model that matches experimental data allows analysis of the dynamics that lead to this coding property.

Extending such a multifaceted approach beyond single neurons is challenging; high-dimensional biophysical models will always be underspecified [23]. However, given that it is now possible to visually identify, record from and manipulate specific cell types, the motivation to incorporate this information in models is there. The appropriate mathematics to perform the necessary reduction of such high-dimensional systems is emerging [24,25]. Studies undertaken in this spirit are beginning to address important open problems, such as the role of diverse cell types [24,26,27], pharmaceuticals [28], neuromodulation [29,30,31] and the statistics of connectivity [32,24] in shaping circuit dynamics and computation. The goal of extracting computation from detailed modeling motivates the use of high-resolution imaging to determine not just a connectivity matrix but also the morphologies of dendritic arbors that influence small circuit computation [33,34,35].

On the output end, the motor field is currently carefully scrutinizing classical tuning curve concepts. Perhaps surprisingly, brain-computer interfaces using decoded motor cortex activity to drive robotic limbs can show considerable success with quite simple decoding algorithms [36,37], which could be viewed as the ultimate validation of a coding picture. As researchers emphasize [38], however, this may be an outcome of remarkable plasticity: as

brain-machine interface pioneer Eb Fetz showed decades ago, firing rates of individual neurons can be adjusted through feedback when connected to a decoder [39]. The success of simple linear decoders suggests that the network may adjust itself to the brain-machine interface if the decoding is easily learnable [40]. During normal (although overtrained) behavior, the complex, sometimes non-monotonic time dependence of multielectrode recordings of neuronal responses indicates that instantaneous tuning curves are certainly an impoverished description. Shenoy and colleagues [41] have argued for a return to a more fundamental dynamical systems approach. In this case, the high-dimensional description of Fig. 1 is not modeled but is represented by data; projecting this data into low-dimensional models may permit the inference of an effective dynamical model with explanatory power.

It is likely that the relative success of coding and decoding strategies both at the sensory input and motor output side are reflections of the existence of relatively low-dimensional structure in the underlying dynamical system. In more central processing areas, the situation appears murkier. Neural recordings from prefrontal (monkey) and parietal (rodent) cortex during sensory decision making tasks reveal distributed, multimodal responses that depend on task [42,43,44]. How can one infer the computation that is taking place from these mixed responses? One route is again to search for low-dimensional structure in the data that separates these influences and hopefully elucidates their interaction [44]. Training recurrent neural networks to solve the task and using dimensionality reduction to analyse the resulting network activity is also an intriguing way forward [45,43].

Thus, there are two major routes toward reducing the intrinsic high dimensionality of neural circuits to low-dimensional models in the search for computation: quasi-analytical simplifications of high-dimensional dynamical systems [20,24,46], or data-driven approaches that discover low dimensional structure in high dimensional data [41,45]. The form of these low-dimensional systems should inform or potentially eventually replace coding models that allow for modulation by behavioral state and top-down effects, and incorporate the nonlinearities that account for adaptation and context dependence. Such an approach might permit a breaking out of conventional thinking about visual representation to encompass emerging alternative views of the true function of V1 [47].

While the feature or receptive field concept, broadly defined, still has much to contribute in helping to organize thinking about representation and its transformation in the brain, many important caveats apply both in the application of the concept and its existence in reality. Passive stimulation protocols favor a relatively simple feedforward perspective. Newer experimental techniques present the possibility, and indeed the necessity, of understanding the more complete dynamical system. Capturing the relationship between dynamically modulated representations and the underlying high-dimensional nonlinear dynamical system in the form of a low-dimensional model will be critical in deciphering complex computations. While some apparent correspondences are emerging between neuronal and artificial networks designed for object recognition, it remains to be seen whether these similarities will prevail under natural recording conditions during complex behavior. As machine learning approaches move beyond object recognition into the realm of natural behavior [48,23,10], the solutions that arise might inform our interpretations of central processing. Ultimately, the development of methods to map the dynamics of the physical

substrate onto the computational is the bottleneck in our ability to truly comprehend the biological mechanisms of intelligence.

Acknowledgments

We thank David Kleinfeld, Andre Longtin, Surya Ganguli, Dora Angelaki and the other attendees of the 2014 CIFAR pre-workshop, as well as Blaise Agüera y Arcas and Alison Duffy for stimulating discussions. We thank CIFAR for the opportunity to meet. This work was funded by CRCNS NIH grant R01DC013693-01. The views expressed in this article are not necessarily those of the NIH.

References

1. Poggio, T.; Mutch, J.; Leibo, J.; Rosasco, L.; Tacchetti, A. The computational magic of the ventral stream: sketch of a theory (and why some deep architectures work).. MIT Technical Report, MIT-CSAIL-TR-2012-35. 2012. <http://cbcl.mit.edu/publications/ps/MIT-CSAIL-TR-2012-035.pdf>
2. Maravall M, Diamond M. Current Opinion in Neurobiology. 2014; 25
3. Fritz J, Shamma S. Current Opinion in Neurobiology. 2014; 2
4. Segundo L, Snitz K, Sobel N. The perceptual logic of smell. Current Opinion in Neurobiology. 2014; 25:107–115. [PubMed: 24440370]
5. Miura K, Mainen ZF, Uchida N. Odor representations in olfactory cortex: distributed rate coding and decorrelated population activity. Neuron. 2012; 74(6):1087–98. [PubMed: 22726838]
6. Sharpee TO. Computational identification of receptive fields. Annu Rev Neurosci. 2013; 36:103–20. [PubMed: 23841838]
7. Pillow JW, Shlens J, Paninski L, Sher A, Litke AM, Chichilnisky EJ, Simoncelli EP. Spatio-temporal correlations and visual signalling in a complete neuronal population. Nature. 2008; 454(7207):995–9. [PubMed: 18650810]
8. Sharpee TO, Calhoun AJ, Chalasani SH. Information theory of adaptation in neurons, behavior, and mood. Current Opinion in Neurobiology. 2014; 25:47–53. [PubMed: 24709600]
9. Olshausen BA, Field DJ. Sparse coding of sensory inputs. Current Opinion in Neurobiology. 2004; 14:481–487. [PubMed: 15321069]
10. Cox D. Current Opinion in Neurobiology. 2014; 25
11. DiCarlo JJ, Zoccolan D, Rust NC. How does the brain solve visual object recognition? Neuron. Feb 9; 2012 73(3):415–34. [PubMed: 22325196]
12. Adibi M, McDonald JS, Clifford CW, Arabzadeh E. Adaptation improves neural coding efficiency despite increasing correlations in variability. J Neurosci. 2013; 33(5):2108–20. [PubMed: 23365247]
13. Wark B, Lundstrom BN, Fairhall A. Sensory adaptation. Curr. Opin. Neurobiol. 2007; 17(4):423–9. [PubMed: 17714934]
14. Bengio Y. Learning deep architectures for AI. Foundations and Trends in Machine Learning. 2009; 2(1)
15. Rieke F, Rudd ME. The challenges natural images pose for visual adaptation. Neuron. 2009; 64(5):605–16. [PubMed: 20005818]
16. Kastner DK, Baccus SA. Insights from the retina into the diverse and general computations of adaptation, detection, and prediction. Current Opinion in Neurobiology. 2014; 25:63–69. [PubMed: 24709602]
17. Seilheimer RL, Rosenberg A, Angelaki DE. Models and processes of multisensory cue combination. Current Opinion in Neurobiology. 2014; 25:38–46. [PubMed: 24709599]
18. Roth E, Sponberg S, Cowan NJ. A comparative approach to closed-loop computation. Current Opinion in Neurobiology. 2014; 25:54–62. [PubMed: 24709601]
19. Cohen N, Sanders T. Nematode locomotion: dissecting the neuronal-environmental loop. Current Opinion in Neurobiology. 2014; 25:99–106. [PubMed: 24709607]
20. Brunel. Current Opinion in Neurobiology. Apr.2014 25

21. Hong SSH, Aguera y Arcas B, Fairhall AL. Single neuron computation: from dynamical system to feature detector. *Neural Comput.* 2007; 19(12):3133–72. [PubMed: 17970648]
22. Mease RA, Famulare M, Gjorgjieva J, Moody WJ, Fairhall AL. Emergence of adaptive computation by single neurons in the developing cortex. *J Neurosci.* 2013; 33(30):12154–70. [PubMed: 23884925]
23. Chris Eliasmith, Oliver Trujillo. The use and abuse of large-scale brain models. *Current Opinion in Neurobiology.* Apr.2014 25:1–6. [PubMed: 24709593]
24. Wolf F, Engelken R, Puelma-Touzel M, Florez Weidinger JD, Neef A. Dynamical models of cortical circuits. *Current Opinion in Neurobiology.* 2014; 25
25. Schaffer ES, Ostojic S, Abbott LF. A complex-valued firing-rate model that approximates the dynamics of spiking networks. *PLoS Comput Biol.* 2013; 9(10):e1003301. [PubMed: 24204236]
26. Gjorgjieva, M. Famulare; Mease, R.; Moody, WJ.; Fairhall, AL. Implications of single-neuron gain control for information transmission in networks.. *Computational and Systems Neuroscience (COSYNE)*. 2011. Available from Nature Precedings, <http://precedings.nature.com/documents/5860/version/1>
27. Harris K, Mrsic-Flogel TD. Cortical connectivity and sensory coding. *Nature.* 2013; 503:51–58. [PubMed: 24201278]
28. Ching S, Brown EN. Modeling the dynamical effects of anesthesia on brain circuits. *Current Opinion in Neurobiology.* 2014; 25:116–122. [PubMed: 24457211]
29. Bargmann CI, Marder E. From the connectome to brain function. *Nat Methods.* 2013; 10(6):483–90. [PubMed: 23866325]
30. Nakahara H. Multiplexing signals in reinforcement learning with internal models and dopamine. *Current Opinion in Neurobiology.* 2014; 25:123–129. [PubMed: 24463329]
31. Fee M. *Current Opinion in Neurobiology.* 2014; 25
32. Hu Y, Trousdale J, Josic K, Shea-Brown E. Motif statistics and spike correlations in neuronal networks. *J. Stat. Mech.* 2013 doi:10.1088/1742- 5468/2013/03/P03012.
33. Takemura SY, Bharioke A, Lu Z, Nern A, Vitaladevuni S, Rivlin PK, Katz WT, Olbris DJ, Plaza SM, Winston P, Zhao T, Horne JA, Fetter RD, Takemura S, Blazek K, Chang LA, Ogundeyi O, Saunders MA, Shapiro V, Sigmund C, Rubin GM, Scheffer LK, Meinertzhagen IA, Chklovskii DB. A visual motion detection circuit suggested by *Drosophila* connectomics. *Nature.* 2013; 500(7461):175–81. [PubMed: 23925240]
34. Helmstaedter M, Briggman KL, Turaga SC, Jain V, Seung HS, Denk W. Connectomic reconstruction of the inner plexiform layer in the mouse retina. *Nature.* 2013; 500(7461):168–74. [PubMed: 23925239]
35. Elyada YM, Haag J, Borst A. Different receptive fields in axons and dendrites underlie robust coding in motion-sensitive neurons. *Nat Neurosci.* 2009; 12(3):327–32. [PubMed: 19198603]
36. Wander, Jeremiah D.; Rao, Rajesh PN. Brain-computer interfaces: a powerful tool for scientific inquiry. *Current Opinion in Neurobiology.* Apr.2014 25:70–75. [PubMed: 24709603]
37. Velliste M, Perel S, Spalding MC, Whitford AS, Schwartz AB. Cortical control of a prosthetic arm for self-feeding. *Nature.* 2008; 453(7198):1098–101. [PubMed: 18509337]
38. Carmena JM. Advances in neuroprosthetic learning and control. *PLoS Biology.* 2013; 11(5):e1001561. doi:10.1371/journal.pbio.1001561. [PubMed: 23700383]
39. Fetz EE. Operant conditioning of cortical unit activity. *Science.* 1969; 163:955–957. 571. [PubMed: 4974291]
40. Legenstein R, Chase SM, Schwartz AB, Maass W. *J Neurosci.* Jun 23; 2010 30(25):8400–10. [PubMed: 20573887]
41. Shenoy, K. M. Sahani; Churchland, M. Cortical control of arm movements: a dynamical systems perspective. *Annu. Rev. Neurosci.* 2013; 36:337–59. [PubMed: 23725001]
42. Roy JE, Riesenhuber M, Poggio T, Miller EK. Prefrontal cortex activity during flexible categorization. *J. Neurosci.* 2010; 30(25):8519–28. [PubMed: 20573899]
43. Mante V, Sussillo D, Shenoy KV, Newsome WT. Context-dependent computation by recurrent dynamics in prefrontal cortex. *Nature.* 2013; 503(7474):78–84. [PubMed: 24201281]

44. Kaufman MT, Raposo D, Churchland AK. Dynamics of decision and action in rat posterior parietal cortex. *Society for Neuroscience*. 2013:668.04.
45. Sussillo D. *Current Opinion in Neurobiology*. 2014; 25
46. Hedrick KR, Cox SJ. Structure-preserving model reduction of passive and quasi-active neurons. *J Comput Neurosci*. 2013; 34(1):1–26. [PubMed: 22714391]
47. Keller GB, Bonhoeffer T, Hibener M. Sensorimotor mismatch signals in primary visual cortex of the behaving mouse. *Neuron*. 2012; 74(5):809–15. [PubMed: 22681686]
48. Mnih V, Kavukcuoglu K, Silver D, Graves A, Antonoglou I, Wierstra D, Riedmiller M. Playing Atari with Deep Reinforcement Learning. arXiv:1312.5602v1 [cs.LG].

Figure 1.

Taking computation here to subsume Marr's computation and algorithm, that is, a description of the system's function, a complete understanding of the neural mechanisms of computation should work at several levels. Data and detailed modeling provide a high-dimensional description of the system. To understand how this concrete implementation carries out a computation, it is useful to develop a low-dimensional description in which the fundamental mechanism of the computation is exposed. The transformation from high-dimensional implementation to low-dimensional model captures the parameter invariance or robustness of the implementation.