CHAPTER

Sensory modeling: Understanding computation in sensory systems through image-computable models

2

Zvi N. Roth[a], Elisha P. Merriam[b], and Tomas Knapen[c,d,e]

[a]*School of Psychological Sciences, Tel Aviv University, Tel Aviv, Israel,* [b]*Laboratory of Brain and Cognition, NIMH, NIH, Bethesda, MD, United States,* [c]*Department of Movement and Behavioral Sciences, Vrije Universiteit Amsterdam, The Netherlands,* [d]*Spinoza Centre for Neuroimaging, Amsterdam, The Netherlands,* [e]*Netherlands Institute for Neuroscience, Amsterdam, The Netherlands*

Introduction to image-computable models

In this chapter, we focus on the application of image-computable models to fMRI. To clarify the scope of this chapter, we will first explain our definition of an image-computable model. A *model* is a formal, falsifiable instantiation of a theory often involving an explicit computational implementation. An *image-computable model* receives arbitrary visual stimuli as input, and outputs a transformation of those stimuli that can be compared to brain measurements. One of the most important characteristics of image-computable models is, as the name suggests, that they do not require any information in addition to the images themselves (Hermes et al., 2019). That is, they give rise to an end-to-end transformation from stimulus to measurement without reference to learning, annotations, or labels. Image-computable models can be productively combined with models of cognitive concepts such as attention (Klein et al., 2014; Kay et al., 2015; van Es et al., 2018), but here we eschew discussing those more elaborate models, and instead focus on the use of image computable models for studying the early visual cortex.

We also limit our treatment to one specific flavor of image-computable model; *mechanistic, neurally inspired processing models* that parsimoniously explain

Computational and Network Modeling of Neuroimaging Data. https://doi.org/10.1016/B978-0-443-13480-7.00007-7

fMRI data with as few parameters as possible. These models' operations resemble operations hypothesized to be carried out by the brain, with the individual model components performing biologically plausible computations. For instance, this definition excludes artificial neural network (ANN) models, for two reasons. First, due to their millions of parameters, ANNs have an extraordinarily high degree of complexity. So while such models produce impressive behavior, their complexity can often make it difficult to relate the computations they perform to the visual system. Second, although the overarching architecture and learning principles of ANN models can be regarded as neurally inspired, the specific parameter tunings that give rise to their operations are not intended to be biologically interpretable.

Another defining feature of image-computable models that we discuss in this chapter is that they are mechanistic; the goal is to understand local computations and the modeling is not explicitly geared toward statistical hypothesis-testing inference, as is discussed, for example, in Chapter 1. Moreover, image-computable models are used to explain the responses at the single-voxel level, as opposed to regions of interest. This means these models explicitly capitalize on the interpretation of voxel-level BOLD responses as sampling the joint activity of a population of some thousands of neurons. Fitting with this level of description, in David Marr's conceptualization of a hierarchy of models, these models live at the junction of the computational and algorithmic levels.

The reason we adhere to these specific scope-defining limitations here is that thusly defined image-computable models make for highly valuable scientific models. There are a number of reasons for this utility, centering on their transparency: not only can their sequence of operations be fully understood, their parameters are also interpretable both computationally and, often, biologically. This set of qualities makes image-computable models an important and highly successful tool to understanding how the living human brain's computations represent visual information from measured BOLD responses.

Examples of successful image-computable models

Two broad classes of image-computable model have been used to model responses in primary visual cortex: the population receptive field (pRF) model and the Gabor filter bank model.

pRF model

The pRF model starts with a simplifying assumption that the BOLD activity in a voxel represents the sum of the receptive fields of all of the neurons in the cortical tissue sampled by that voxel. Different neurons within a voxel's neural population

will have a range of RF centers (i.e., RF location scatter), orientations, and spatial frequencies (REFs, Hubel and Wiesel). Responses specific to orientation, eye-of-origin, spatial frequency, etc. are relatively diverse within the population of neurons sampled by a single voxel at standard resolution. But responses selective for spatial location are more consistent within this population: the dominant component of a voxel's population response is related to visual location. Thus, by disregarding the stimulus properties other than visual–spatial location, the pRF model is focused explicitly on mapping the retinotopic structure of the visual system and understanding the visual system's spatial computations. The main assumption here is that the population response may be well approximated by a circularly symmetric 2D Gaussian function defined in visual space. The main challenge is to find the position and size of the Gaussian that best approximates the population response (Fig. 1).

The pRF model is the primary successor to older techniques for mapping the retinotopic structure of the human visual cortex. These earlier methods used periodic stimulus designs to estimate the location in the visual field that produces an optimal response. Periodically rotating wedges and expanding/contracting rings would be used as stimuli to estimate voxels' polar angle and eccentricity preferences, respectively, based on the phase of periodic responses to visual stimulation (Sereno et al., 1995; Engel et al., 1997). The pRF modeling approach confers a number of advantages over these older techniques.

1. It is an explicit, image-computable model of visual responses in the spatial dimension: its estimated parameters are directly interpretable as relating to neural tuning inside our voxels.
2. Consequently, whereas these older techniques generally produced a point-estimate of spatial tuning, the pRF model also estimates the spatial extent of spatial tuning via its size parameter (the standard deviation of the circular Gaussian function).
3. The pRF modeling approach can simultaneously fit the shape of the hemodynamic response function as just an additional (set of) model parameter, which may do a better job of fitting population receptive fields in different tissue types (large veins, capillaries, different layers) (Kay et al., 2020).
4. Stimulus flexibility. Early retinotopic mapping techniques were limited to rings and wedges used to separately map out radial and angular components of the retinotopic map because of their dependence on a periodic stimulus design. Because the pRF model is image computable, it is more flexible and can run on arbitrary sequences of stimuli.

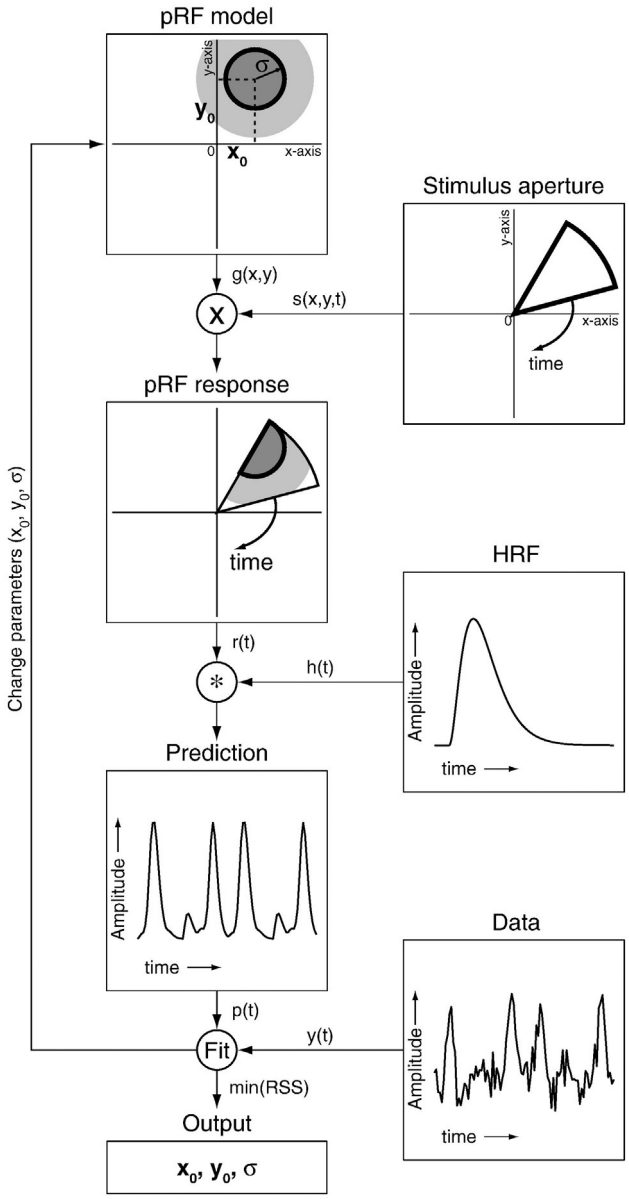

FIG. 1

pRF model and fitting procedure. The simple pRF model is defined by a Gaussian with three parameters: a center (x and y coordinates) and a standard deviation. The pRF fitting procedure finds values for these three parameters that best explains the response to the stimulus.

Stimulus designs

This stimulus flexibility freed researchers to use more irregular experimental designs, for example, stimuli with blank periods. Blank periods are effective when attempting to demonstrate spatially selective responses in brain regions with very large receptive fields (since stimuli without blanks can saturate responses in such areas) (Dumoulin and Wandell, 2008). The typical pRF mapping stimulus design is to scan a fixating subject while they view bar stimuli that traverse the visual field, in much the same way that Hubel and Wiesel used bars of light to map neuronal receptive fields in visual cortex (Hubel and Wiesel, 1962). This primary stimulus design of bars traveling across the screen combines sparseness of responses (for a given voxel, the bar only occupies its pRF for a relatively brief period, making response timing highly informative) with strength of responses (whenever the bar stimulus occupies a pRF, this is likely for a duration longer than 4 s—producing response of block-design strength). Generally, the textures inside these bars are optimized to evoke maximal response at a given location in the visual field. The texture in the bar can additionally be chosen to optimally activate brain regions tuned to specific visual categories, such as faces, bodies, scenes, etc. This has resulted in the use of bar apertures filled with full contrast counter-phase flickering checkerboard patterns as a generalized broadband stimulus (Dumoulin and Wandell, 2008; Amano et al., 2009; Fracasso et al., 2016), fast-alternating naturalistic visual stimuli, or even cartoon stimuli tailored to evoke stronger responses in high-level visual cortex (Silson et al., 2015; Benson et al., 2018; Kim et al., 2023).

Extensions to the linear pRF model

Early extensions to the basic, linear Gaussian pRF model captured two separate response signatures that the basic model cannot account for. These are surround suppression (depression of BOLD responses below empty-screen baseline whenever a stimulus impinges on a voxel's inhibitory penumbra), and response compression (sublinear increases of BOLD responses with increasing stimulus strength). Separate models were created to account for these different aspects of measured BOLD responses. Time courses displaying suppressive surrounds, predominantly appearing in V1, V2, and V3, are adequately modeled using a difference-of-Gaussians model. In this model, a larger, negative Gaussian's response is subtracted from the positive response of a smaller Gaussian with the same center location (Zuiderbaan et al., 2012). This subtractive response signature is still a linear one. However, BOLD responses also increasingly show nonlinear, compressive, response signatures as we move up the visual hierarchy. That is, as we increase the strength of a given pRF's stimulation, the response to this stimulation increases sublinearly, plateauing at higher stimulus strengths. The compressive spatial summation (CSS) model (Kay et al., 2013) produces this type of response signature by raising the response predicted by a linear model to a

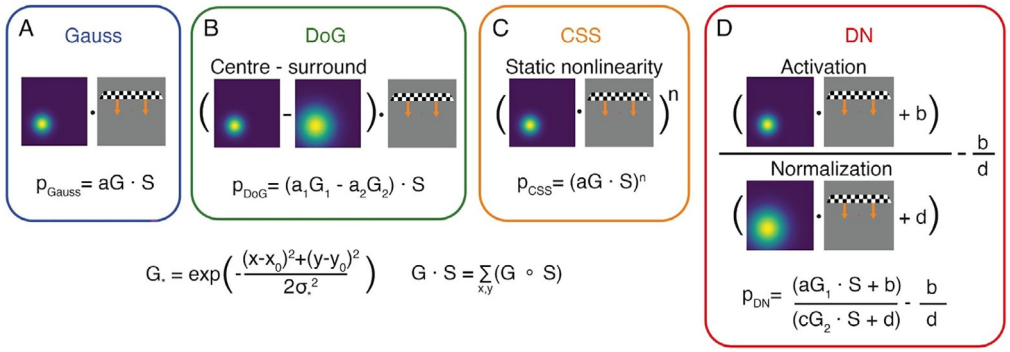

FIG. 2

Simple pRF model and elaborations. (A) The simplest pRF model is a Gaussian with three parameters (see Fig. 1). This model can be easily extended by adding additional parameters to describe phenomena such as (B) surround suppression (DoG) and (C) nonlinearities (CSS) in the neural response. (D) The DN model unifies both extensions by incorporating divisive normalization.

subunity power, generating a saturating, power-law input-response curve. A more recent model has attempted to unify these different models using a single parsimonious mechanism: divisive normalization (DN) (Heeger, 1992; Carandini and Heeger, 2011). The core concept of this canonical computational motif is that local, finely tuned responses are divided by the response of a more broadly tuned "normalization" pool of neurons. The computational motif of DN is able to explain a host of different findings across a range of different domains, from cross-orientation suppression to value-based decision-making. By implementing a divisive interaction between a larger "normalization" Gaussian pRF and a smaller "activation" Gaussian pRF, both centered on the same location, the DN pRF model is able to generate both suppressive and compressive response signatures (Aqil et al., 2021) (Fig. 2). Similar nonlinear mechanisms have been proposed to play a role in the temporal domain (Stigliani et al., 2017; Zhou et al., 2018). The normalization pool in this case has broader tuning in time: it contains a memory trace of past inputs. This allows this type of model to explain temporal subadditivity and repetition suppression. An additional testament to the parsimonious power of this approach is that the pRF modeling framework has proven effective as a scaffold for relating BOLD responses to presumed underlying electrophysiological responses, as measured using intracranial EEG (Harvey et al., 2013; Hermes et al., 2019; Groen et al., 2022). Another indication of the biological relevance of this type of image-computable modeling is that the distribution of DN pRF model parameters across the visual cortex is correlated with those of specific serotonergic (5HT) and inhibitory (GABA) receptor densities, as measured using positron emission tomography (Aqil et al., 2024).

But the pRF method also has its drawbacks. As the model is a purely spatial model without regard for additional stimulus features, one chief drawback is the need to convert the stimulus into an "energy" image. This image is created by estimating the local strength of responses to a given image, regardless of stimulus features. But the natural scenes we encounter in everyday life carry information about a multitude of cues: color, orientation, motion direction, etc. Capturing the feature-tuned responses to these types of stimuli requires going beyond spatial pRF models.

Gabor filter bank model

Although the dominant component of a voxel's response may be spatial selectivity, voxels in visual cortex, particularly in V1, have additional selectivity to cues such as spatial frequency and orientation. In the Gabor filter bank model, the image is sampled by oriented filters. The Gabor bank consists of layers of filters, each filter responding to a particular range of spatial frequencies. Within a layer, filters measure the energy at (or around) a particular orientation. In other words, the Gabor bank decomposes an image into spectral components, similar to a Fourier transform. Yet unlike a Fourier transform, this spectral decomposition is local, reflecting the amount of energy at each spatial frequency and orientation

at each position in the image: the Gabor bank performs wavelet decomposition on the image.

In addition to responding to a particular combination of spatial frequency and orientation, each Gabor filter is also centered on a particular location in the image. So in a sense, this model is an extension of the pRF model to include stimulus features beyond spatial position. The Gabor filters operate on any arbitrary image, without the need to first convert the image into an "energy" image. Each layer will essentially convert the original image into an energy image at that layer's spatial frequency and orientation. The main downside of such a model is that there is a large number of parameters to fit to each voxel. Each voxel is assumed to respond not only to a particular region of the image but also to a particular spatial frequency and orientation (Fig. 3). To speed up the fitting process without requiring a prohibitive amount of data, one can use regularization, such as lasso regression, ridge regression, or early stopping (Kay et al., 2008). Alternatively, one may choose to separate between fitting each voxel's spatial selectivity and spectral selectivity. For example, the spatial selectivity may be determined based on independent pRF estimates from a separate scanning session, while spatial frequency and orientation tuning may be fitted based on responses to natural scenes (Roth et al., 2022).

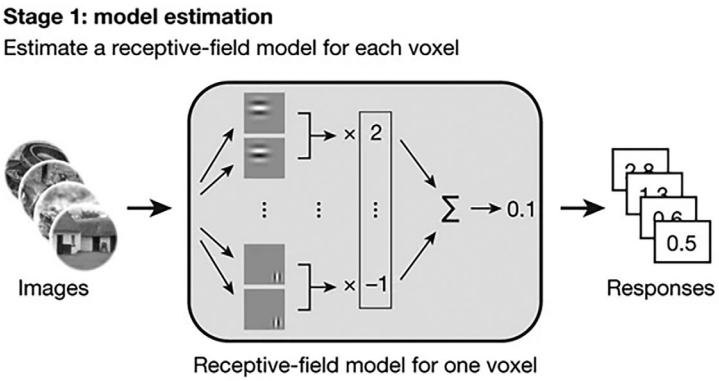

FIG. 3

The Gabor bank model. Each image is decomposed into Gabor filters at a range of orientations and spatial frequencies. Each voxel has weights for the Gabor filters, reflecting the voxel's orientation and frequency tuning. Predictions of a voxel's response amplitudes are generated by multiplying the filter weights with an image's decomposition into filter responses.

Extensions to the Gabor bank model

The Gabor bank includes filters to specific orientations and spatial frequencies, "low-level" features that V1 neurons are selective to. Extensions to the Gabor bank model include additional, more complex visual features. For example, a recent study (Henderson et al., 2023a,b) extended the Gabor bank to include mid-level visual

features, computed as correlation within and between different filters (following Portilla and Simoncelli, 2000). Such mid-level features may underlie selectivity in extrastriate cortical region V2 (Freeman et al., 2011, 2013; Ziemba et al., 2016). Another study replaced the entire Gabor bank and spatial selectivity components with a combination of both low- and mid-level visual properties (color, curvature, and texture components) and semantic object-identity labels (e.g., faces, places, bodies, words, and food) to model higher-order visual areas (Khosla et al., 2022). In contrast to image-computable models, semantic encoding models require semantic labels to be assigned to each image (see Box 1).

Box 1

Approaches related to, but distinct from, image-computable models

Here, we want to highlight additional families of models that share the approach and conceptual paradigms used in fitting image-computable models. We distinguish two subtypes: encoding models, and models of functional connectivity.

Encoding models. Encoding models describe how information is encoded by a population of neurons measured in single voxels using fMRI but that are not constrained by the demand that the model's predictions are image-computable. This loosened constraint widens the applicability of these models, while sacrificing end-to-end transparency. Powerful examples are encoding models based on the semantic information occurring in images or video material. One can then estimate semantic model parameters by fitting them to fMRI data (Huth et al., 2012). Research has shown that this type of modeling can reveal structure in the brain's high-level responses to semantic information and that this structure is shared across sensory modalities, i.e., it is similar between video-based and auditory narratives (Deniz et al., 2019). The transcription of semantic information may be done manually or using neural networks, but in both cases, the process that distills the semantic information from the visual primitives remains a black box, and hence, do not meet the requirements that we set forth in this chapter.

Functional connectivity. Similar modeling strategies can also be used to attempt to explain ongoing fMRI responses (the target) as a function of responses in a source region. This means that the input to these models is not a stimulus, but rather a pattern of BOLD responses. The so-called connective field models attempt to explain target voxels' responses as a function of a "connective field" defined on the surface of a source region, in direct analogy to the population receptive field model of visual-space processing (Haak et al., 2013). These models leverage the topographic structure of the source region to condition the model of functional connectivity. They have been used to discover the topographic structure across the brain in experimental paradigms that resists image-computable approaches, such as resting-state or naturalistic movie watching experiments (Gravel et al., 2014; Knapen, 2021).

Other studies have extended the Gabor bank approach to decompose images using the filters of deep neural networks (St-Yves and Naselaris, 2018). In a similar fashion to what we described earlier, each voxel can then be modeled as a weighted combination of the different filters. An additional pRF component may be added to sample all filters according to each individual voxel's spatial preference.

In principle, the entire visual system, i.e., both early visual cortex and high-level brain areas, can be modeled with image-computable models, which in turn can also

take into account task-effects (see below). Indeed, we have a solid understanding of the relevant filters with which V1 decomposes an image, and much progress has been made in determining the filters underlying V2 responses. But as we move up the visual hierarchy beyond V1 and V2, less is known about the computations being performed. Hence, we do not currently have an appropriate image-computable model for high-level visual cortex. Why is it so difficult to model high-level visual responses? Two reasons seem most likely. The first is that the relevant visual features become increasingly complex in higher order visual areas. In fact, artificial neural networks exhibit similar complexity in deep layers of the network, but we have yet to build image-computable models that successfully capture this complexity. While ANNs can capture this high-level visual complexity, they do it in an opaque manner, not by relying on explicit computations that are designed a priori by the researcher. Second, responses become increasingly influenced by task effects that are missed by models that focus on the stimulus alone.

Assumptions of image-computable models

Image-computable models in general depend on several fundamental assumptions, and each individual image-computable model has its own specific set of assumptions. This is common to all explicit modeling endeavors: formalizing a theory into a concrete set of equations and/or algorithms forces a researcher to be highly explicit about the assumptions involved, and the domains in which the model is applicable. Model comparisons and simulations can help to investigate the limits of the assumptions used in a specific model. This aspect of image-computable modeling is a core strength of the approach, in that it is key in generating parsimonious explanations of computational principles. The assumptions shared across image-computable models of fMRI voxel-based responses are important too, because they are fundamental to how we may use fMRI to investigate neural computation. So, here we highlight several assumptions of this nature.

One fundamental assumption regards how fMRI voxels sample neural responses. Activity in fMRI voxels reflects the pooled activity of a large and diverse population of neurons. The fact that fMRI voxels are a pooled measure has several important implications for how fMRI responses can be used to infer properties of activity at the neural level. In order for voxels to exhibit any sort of tuning to stimulus properties, neurons must be organized in the brain in clusters, gradients, or some other form of coarse-scale pattern that will create a reasonably homogeneous pattern of activity over the spatial scale sampled by a voxel. For example, if neurons with some particular stimulus preference were scattered randomly throughout visual cortex and intermixed with other neurons with other preferences, each voxel would sample from a large number of neurons with different tuning properties, which could, in principle, cancel each other out. The resulting voxel response should have little or no tuning to the visual feature. It is interesting to consider how tuning at the voxel level depends on the chosen voxel size, and what this voxel size dependence implies about tuning

scatter in the sampled neural populations. In terms of position tuning, neurons in early visual cortex have clear position tuning (conceptualized as receptive fields) and are organized in a coarse-scale retinotopic map. In other words, nearby neurons respond to nearby locations in the visual field. As a result of a summing the activity of many nearby neurons, voxels in early visual cortex have clear position tuning as well. Similarly, nearby neurons in early visual cortex have similar spatial frequency tuning, and therefore, voxels in early visual cortex exhibit clear spatial frequency tuning (Aghajari et al., 2020; Broderick et al., 2022). But occasionally, different theories make different assumptions regarding how voxels sample neuronal activity, and these assumptions are reflected in details of the model used.

For example, the basic pRF model assumes the pRF is shaped as a symmetric Gaussian. This should be the case if neuron RFs sampled within a voxel are scattered homogeneously throughout the visual field. It is theoretically possible that neuronal RFs are scattered more in one direction (such as the radial direction, relative to fixation) than in the other perpendicular direction (the tangential direction), and this issue is currently debated (Silson et al., 2018; Lerma-Usabiaga et al., 2021).

Orientation preference in primary visual cortex is organized in cortical columns at a much finer scale than the retinotopic map (Hubel and Wiesel, 1968; Grinvald et al., 1986). Following the logic we laid out earlier, voxels may sample from neurons preferring the entire range of possible orientations, which should, in principle, result in voxels with no orientation tuning (Gardner and Merriam, 2021). Yet, fMRI decoding studies have repeatedly demonstrated the presence of orientation information in V1 voxels (Haynes and Rees, 2005; Kamitani and Tong, 2005). What is the source of decodable orientation information in these studies? One possibility, termed the "random bias" account, posits that each V1 voxel, despite sampling from columns of all orientations, happens to randomly sample from slightly more columns preferring a particular orientation. This would result in each voxel having a slightly larger response to a particular orientation (Boynton, 2005). This was an attractive conjecture because it suggested that fMRI could be used to study tuning properties of heterogenous populations of neurons within a voxel, allowing researchers to, in effect, peer past the spatial resolution limits of the fMRI measurements.

But, it turns out, this conjecture has been difficult to fully support or fully refute. Subsequent studies demonstrated that successful decoding of orientation relied on a coarse-scale map of orientation preference in V1 that corresponds to the retinotopic map, resulting in a "radial bias" across the cortical surface (Freeman et al., 2011). This finding highlights how difficult it can be to interpret the results of any decoding analysis: the mere ability to decode information from fMRI responses does not speak directly to the nature of the underlying neural computation.

Attempts to understand the neural computations that give rise to orientation decoding were further complicated by a theoretical paper that suggested that the orientation map was in fact a consequence of a subtle confound in the stimulus structure (Carlson, 2014), a finding confirmed by our lab in a subsequent empirical study (Roth et al., 2018). It turned out that stimulus edge effects have a powerful influence on the nature of the orientation map that can completely obfuscate measurement of

the underlying orientation tuning; orientation selectivity was found to be determined by the spatial position of the stimulus edges, and changing the edge can completely flip the orientation map (Roth et al., 2018). According to these findings, one might conclude that each voxel samples homogeneously from columns with all orientation preferences, resulting in voxels with no reliable tuning (Carlson, 2014). But if the edge effect simply obscures columnar orientation tuning, is it possible to measure orientation tuning with fMRI?

This question led our group to develop an approach for characterizing both edge effects and orientation selectivity using a pair of image computable models and statistical model comparison. Specifically, one model was built with the assumption that no orientation tuning was present in the fMRI measurements (beyond the edge effect described earlier), and a second model was built with the assumption of orientation selectivity, above and beyond what would have been induced by the stimulus edge. We fit both models to fMRI data and found that although they both perform well, the second model, which allowed for orientation selectivity, explained more variance in the data—variance that reflected true orientation tuning (Roth et al., 2022). This study demonstrates that once we are aware of the model's assumptions, we build that assumption into the image computable model to test whether those assumptions are correct.

A second major assumption is that image-computable models assume that neural responses depend entirely on the visual stimulus. This means that it should be possible, for example, to train a model on responses recorded in the morning and accurately predict responses recorded in the afternoon, tomorrow morning, or several months from now. The viewer, or participant, may be in an entirely different state of mind after a few hours, perhaps hungrier and grumpier, or less alert, but the way in which neurons respond to visual stimuli is assumed to be unchanged. According to this assumption, we can increase the signal-to-noise ratio of the measurements by averaging over responses measured on different days, since any difference between responses measured on different days is attributed to noise. The natural analogy is a computer program, which processes an image the exact same way, any time any day, regardless of any other processes the computer is executing (or has executed) at the time.

However, recent studies suggest that this is not the case that brains are not like computers in this way. Visual responses may be influenced by the internal brain states, which reflect a wide range of factors not directly linked to sensory input, and which may change dynamically over a range of time scales. In rodents, changes in arousal have been shown to modulate visual responses throughout the visual system (Niell and Stryker, 2010; Aydın et al., 2018; Savier et al., 2019; Schröder et al., 2020). Changes in arousal also affect primate (Sirotin and Das, 2009; Cardoso et al., 2012, 2019) and human visual cortex (Roth et al., 2020; Burlingham et al., 2022), and are likely to impact visual responses as well.

Even when the internal brain state is not experimentally manipulated, visual responses may change nonetheless. A phenomenon termed "representational drift" has been demonstrated in rodents (Deitch et al., 2021; Marks and Goard, 2021),

and more recently also in humans (Roth and Merriam, 2023). In representational drift, neural responses to the same stimulus exhibit systematic changes over time that are not linked to any change in the stimulus. Such "drift" implies that an image-computable model trained on data collected on day 1 will be suboptimal at predicting responses on day 2, and the difference in responses between data collected at different time points will grow monotonically with time.

Changes in neural tuning have been demonstrated with changes in attention (Womelsdorf et al., 2006), but neural tuning can also change with task (Kay et al., 2023). In a simple case, a change in task will modulate some parameters of the neural tuning, such as a shift in the response gain, or a change in the preferred stimulus. In theory, it is possible that in some cases, neural tuning changes completely with the task, such that completely different models are needed to explain the data when subjects are engaged in different tasks. This is likely the case, for example, when trying to model responses in frontal cortex (Mante et al., 2013). This type of tuning conflicts with the assumption of image-computability, as it indicates that neural responses are best explained by a combination of both image-based computation and task.

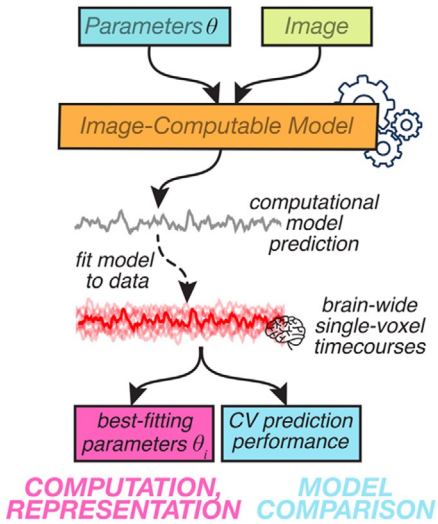

FIG. 4

The analysis paradigm used for image-computable modeling. A model converts an image and some set of model parameters into an explicit model prediction of BOLD responses. This means we can estimate optimal parameter values for all voxels in parallel, and interpret these best-fitting parameters as reflecting computational and representational properties of neural responses. Moreover, by performing cross-validation, models can be compared in terms of their ability to predict left-out data, regardless of the specific parametrizations of the models in our comparison.

Building, testing, interpreting image-computable models

An image-computable model needs to go through multiple steps to produce response predictions (Fig. 4). The initial stage of many models is to posit a filter or set of filters, and taking the dot product of the filter(s) and the stimulus. This produces an idealized linear response. The simplest models use this linear response as a signal prediction; however, more complex models apply additional steps such as rectification, exponentiation, division, etc. to perform a desired response transformation. The model-design choices made here are usually informed by mathematical considerations and relevant results from animal electrophysiology. The predicted neural signal that the model produces is then usually convolved with a hemodynamic response function to produce a prediction at the level of the measured signal.

Fitting image-computable models

There are several complementary strategies for fitting image-computable models to our neuroimaging data. We here separate these into two broad categories; direct optimization of interpretable parameters, and (penalized) regression that delivers parameter values as beta-weights. In both strategies, we are interested in two outcomes: (1) the values of best-fitting model parameters, and (2) the fit quality as assessed by cross-validated fraction of variance explained.

In direct parameter optimization, one specifies parameter values that dictate the behavior of the model and then uses the presented stimulus as input to the model to produce a predicted response—a simulation. By varying the values of the model parameters, one can produce different predicted responses. For instance, in the pRF model, these parameters are horizontal and vertical position in the visual field, and the pRF's sampling extent (the standard deviation of the pRF's Gaussian window). We can then employ the arsenal of curve-fitting routines to find the optimal parameter values, that is, the parameter values that best explain the measured signal. Thus, researchers are required to make many choices concerning whether to use grid-fitting or iterative fitting, what the appropriate cost function is, what search algorithm is appropriate, and possible combinations thereof. For instance, in many cases, parameters are fit using a coarse-to-fine approach starting with a grid search, simulating many different model predictions and using them to explain the measured signals in turn. For every voxel, the best-fitting combination of parameter values from the grid stage is then used as a starting point for an iterative search procedure. This combination of approaches ensures both that a full parameter domain is sampled and global minima within this domain are likely found while also allowing detailed search for the final parameter values. The choice of numerical optimization procedure (or sequential combination thereof) is dependent on things such as a researcher's speed-accuracy trade-off, a fitting problem's tendency to get stuck in local minima, our knowledge about the parameter space, implementational details, etc.

When fitting models using regression, one constructs a design matrix composed of separate model predictions. For instance, in the Gabor wavelet model, one prediction would be generated for each filter in the filter bank. This design matrix is then used to estimate beta weights for each of the regressors using gradient descent or general linear model fitting (see Chapter 1). Depending on the structure of the model, design matrix, and data quality and amount, one can use ordinary least squares regression (in the case of large amounts of high-quality data and/or a high-rank design matrix) or move to regularization to avoid overfitting (for example, when there are many regressors relative to the number of data points). In the case of regularized fitting, cross-validation is required to set the appropriate value for the penalization hyperparameter(s), and a separate validation set is generally used to compute explained variance of the optimized model.

The two fitting strategies outlined here both require cross-validation to compare different models, especially when these differ in their number of parameters. Their differences are mainly in the way by which the researcher imposes constraints on the model's ability to explain data. For instance, the Gaussian pRF model limits the set of possible spatial patterns with which a stimulus is compared to a very specific set: coherent, unitary, and smooth patches following a Gaussian profile defined in visual space. This very reasonable restriction of the possible outcome space can be seen as a form of regularization. In comparison, it is also possible to fit a pRF model using penalized ridge regression, using all screen pixels' time courses, convolved with an HRF, as regressors (Lee et al., 2013). In this case, the ridge penalization implements a smoothness constraint on the shape of the resulting spatial filter. This example serves to illustrate that both strategies can be used to constrain the space of possible outcomes; by a priori definition in the case of parameter optimization and based on statistical considerations in the case of penalized regression.

Additional considerations for choosing a Gabor bank model

When building a Gabor bank model, one can tweak the specific model by adjusting the architecture and the parameter values. For example, the model will have a specific number of orientations and of spatial frequency levels, and the filters will have specific orientation and spatial frequency bandwidths. These values are dependent on each other, so once a choice of filters bandwidth has been made, the number of filters will be determined by the image size.

One elegant version of Gabor bank is the Steerable Pyramid model (Simoncelli et al., 1992; Simoncelli and Freeman, 1995). This model consists of shiftable filters, similar to Gabors, that are convolved with the image. The model is designed such that the (squared) filters span Fourier space uniformly. This eliminates subtle artifacts that can be produced by a Gabor bank (Clifford and Mannion, 2015; Roth et al., 2018).

Pitfalls in image-computable modeling

As with all modeling approaches, image-computable modeling is subject to several potential pitfalls. Here, we discuss a nonexhaustive list of common sorts of errors.

Pitfall #1: A realistic appreciation of fit quality. If the model is fit using an insufficient amount of data, or if the data signal-to-noise ratio is not sufficiently high, the model predictions will be inaccurate, and hence explain only a small proportion of the total variance in the signal. To determine the reliability of the model fits, it is common to use cross-validation: we use one portion of the data to fit the model, and test how well the fitted model predicts or corresponds to the held-out portion. This is a good procedure with which to guard against overfitting and to compare models against one another, but it is important to compare their performance to the noise ceiling, which specifies the amount of explainable variance in the signal (for instance, as gauged by test–retest reliability of responses) (David and Gallant, 2005). When we observe clear divergence, this might mean there are additional aspects of the data that need modeling and indicate that one may expand either data acquisition and/or the complexity of the models used. As this type of data-model mismatch will be evident in the structure of the model fit residuals, we strongly advocate detailed inspection of model residuals as an integral stage in model development. This process will, over the course of model development, enable one to tune model complexity and generally lead to models that more parsimoniously explain data.

Pitfall #2: Implicit biases in stimuli. Because image-computable models operate on the stimulus, such approaches are susceptible to small but systematic biases in the stimuli themselves. This has been appreciated in electrophysiological approaches to receptive field mapping for some time and was part of the motivation for developing white noise RF mapping through methods such as m-sequences (Reid et al., 1997).

Here, we list a number of biases identified in image-computable models that may have varying degrees of impact on results. First, the standard traveling bar pRF design, despite its evident benefits, imposes certain appreciable biases on model outcomes. First, the standard experimental design shows bars in a circular aperture, which causes edge effects. This aperture interacts with the very predictable spatiotemporal stimulation inherent in these designs. These effects are minimized by using the so-called multifocal mapping paradigms that present multiple stimuli at different locations and spatial extents in a trial-based format, analogous to white noise mapping stimuli used in electrophysiology (Binda et al., 2013), although these designs suffer in terms of effective signal to noise and have biases of their own—for instance, in estimates of pRF size (Infanti and Schwarzkopf, 2020). As mentioned earlier, the content of the traveling bar stimuli biases in which visual regions responses will be maximal. For instance, filling bar apertures with scenes will evoke stronger responses in scene-selective cortex, allowing researchers to research spatial selectivity in those regions (Silson et al., 2015). Similar reasoning holds for different visual categories, with

generalized naturalistic bar content used to evoke stronger visual–spatial responses throughout high-level visual cortex (Benson et al., 2018; Kim et al., 2023). Yet, for any stimulus category, the image statistics can potentially bias the model fits. For example, spectral statistics of natural images may bias models fit to responses to natural scenes, particularly when regularization is used to improve the model fits.

Pitfall #3: Overestimating specificity of localization. Many early efforts at building image-computable models of the visual system specify a restricted area of visual/image space for which a voxel's response is evoked. This has led to a common understanding of responses as highly localized. We want to point to two specific, recently developed reasons that call this understanding into question. First, some researchers have assumed that a voxel responds only in the area of the Gaussian standard deviation that was used to model the response. However, a Gaussian response profile implies that the voxel responds beyond the standard deviation border, albeit at a lower response amplitude. For example, if a voxel with a pRF size of 2.5 degrees is centered 3 degrees away from a stimulus, then the voxel will respond to the stimulus at approximately 50% of its maximal response amplitude. This property of pRFs has implications for stimulus vignetting: all voxels that respond to the edge of a grating will have orientation information, even if the distance from stimulus to pRF center is greater than the pRF standard deviation. Second, the results from the divisive normalization pRF model show that the normalization penumbra of a V1 voxel's response extends six- to eightfold further than its "classical," linear Gaussian pRF's extent (Aqil et al., 2021). These two examples show that responses, even at the level of V1, are much more spatially integrated than local models assume, with important implications for experimental paradigms and analyses that depend on locality. For example, selecting voxels whose responses to a specific spatial stimulus are significantly positive excludes many voxels that indeed carry information about that stimulus, for example, in their negative response deflections.

Pitfall #4: Code and method complexity. Image-computable models place a high burden on the researcher in terms of their technical process and ability to develop and maintain complicated processing pipelines. The codebase associated with an image-computable modeling project is likely to span thousands of lines of code, behooving the researcher to invest in the adherence to principled coding standards. But inherently, all modeling code of realistic complexity is likely to include bugs. Without performing additional checks, such errors in the model's implementation can lead to wrong conclusions. These additional checks should consist of detailed simulations of model predictions, the interactions between model parameters and noise levels, and analyses centering on parameter recovery (Wilson and Collins, 2019). Luckily, the field has recently started moving toward reproducible pipelines for model comparisons (Lerma-Usabiaga et al., 2020) meant to combat these types of implementation-based errors. In general, the move toward open datasets (Benson et al., 2018; Chang et al., 2019; Allen et al., 2021) and open software/methodologies promises to lead to improvements in common practice in the field.

Open challenges and future directions

In this chapter, we have described how image-computable models embody explicit, testable hypotheses regarding the neural computations carried out by the visual system. Image-computable models place a large burden on the researcher to formalize and implement the models, in contrast to an alternative approach involving artificial neural networks. In this manner, image-computable models exemplify a specific and highly valuable approach to fMRI studies of both the visual system specifically and sensory systems more generally. Task effects are increasingly seen as an interesting component of stimulus-evoked activity and highlight the inherently limited scope of image-computable models. Future generations of models should combine image-computable models with explicit models of task effects to explain brain responses in a more general domain. This will allow us to make increasingly more general models of processing: trickling the value of image-computable models up into the brain.

Take-home points

- Image-computable models are transparent: the user defines the underlying computations.
- Image-computable models require no labeling.
- Image-computable models facilitate model comparison.
- Components of image-computable models may correspond directly to neural properties.
- Image-computable models are (currently) limited to stimulus processing, excluding task effects.

References

Aghajari, S., Vinke, L.N., Ling, S., 2020. Population spatial frequency tuning in human early visual cortex. J. Neurophysiol. 123, 773–785.

Allen, E.J., St-Yves, G., Wu, Y., Breedlove, J.L., Prince, J.S., Dowdle, L.T., Nau, M., Caron, B., Pestilli, F., Charest, I., Hutchinson, J.B., Naselaris, T., Kay, K., 2022. A massive 7T fMRI dataset to bridge cognitive neuroscience and artificial intelligence. Nat. Neurosci. 25, 116–126.

Amano, K., Wandell, B.A., Dumoulin, S.O., 2009. Visual field maps, population receptive field sizes, and visual field coverage in the human MT+ complex. J. Neurophysiol. 102, 2704–2718.

Aqil, M., Knapen, T., Dumoulin, S.O., 2021. Divisive normalization unifies disparate response signatures throughout the human visual hierarchy. Proc. Natl. Acad. Sci. U. S. A. 118 (46), e2108713118.

Aqil, M., Knapen, T., Dumoulin, S.O., 2024. Computational model links normalization to chemoarchitecture in the human visual system. Sci. Adv. 10, eadj6102.

Aydın, Ç., Couto, J., Giugliano, M., Farrow, K., Bonin, V., 2018. Locomotion modulates specific functional cell types in the mouse visual thalamus. Nat. Commun. 9, 4882.
Benson, N.C., Jamison, K.W., Arcaro, M.J., Vu, A.T., Glasser, M.F., Coalson, T.S., Van Essen, D.C., Yacoub, E., Ugurbil, K., Winawer, J., Kay, K., 2018. The human connectome project 7 Tesla retinotopy dataset: description and population receptive field analysis. J. Vis. 18, 23.
Binda, P., Thomas, J.M., Boynton, G.M., Fine, I., 2013. Minimizing biases in estimating the reorganization of human visual areas with BOLD retinotopic mapping. J. Vis. 13, 13.
Boynton, G.M., 2005. Imaging orientation selectivity: decoding conscious perception in V1. Nat. Neurosci. 8, 541–542.
Broderick, W.F., Simoncelli, E.P., Winawer, J., 2022. Mapping spatial frequency preferences across human primary visual cortex. J. Vis. 22, 3.
Burlingham, C.S., Ryoo, M., Roth, Z.N., Mirbagheri, S., Heeger, D.J., Merriam, E.P., 2022. Task-related hemodynamic responses in human early visual cortex are modulated by task difficulty and behavioral performance. Elife 11, e73018.
Carandini, M., Heeger, D.J., 2011. Normalization as a canonical neural computation. Nat. Rev. Neurosci. 13, 51–62.
Cardoso, M.M., Sirotin, Y.B., Lima, B., Glushenkova, E., Das, A., 2012. The neuroimaging signal is a linear sum of neurally distinct stimulus- and task-related components. Nat. Neurosci. 15, 1298–1306.
Cardoso, M.M.B., Lima, B., Sirotin, Y.B., Das, A., 2019. Task-related hemodynamic responses are modulated by reward and task engagement. PLoS Biol. 17, e3000080.
Carlson, T.A., 2014. Orientation decoding in human visual cortex: new insights from an unbiased perspective. J. Neurosci. 34, 8373–8383.
Chang, N., Pyles, J.A., Marcus, A., Gupta, A., Tarr, M.J., Aminoff, E.M., 2019. BOLD5000, a public fMRI dataset while viewing 5000 visual images. Sci. Data 6, 49.
Clifford, C.W., Mannion, D.J., 2015. Orientation decoding: sense in spirals? Neuroimage 110, 219–222.
David, S.V., Gallant, J.L., 2005. Predicting neuronal responses during natural vision. Network 16, 239–260.
Deitch, D., Rubin, A., Ziv, Y., 2021. Representational drift in the mouse visual cortex. Curr. Biol. 31, 4327–4339. e4326.
Deniz, F., Nunez-Elizalde, A.O., Huth, A.G., Gallant, J.L., 2019. The representation of semantic information across human cerebral cortex during listening versus reading is invariant to stimulus modality. J. Neurosci. 39, 7722–7736.
Dumoulin, S.O., Wandell, B.A., 2008. Population receptive field estimates in human visual cortex. Neuroimage 39, 647–660.
Engel, S.A., Glover, G.H., Wandell, B.A., 1997. Retinotopic organization in human visual cortex and the spatial precision of functional MRI. Cereb. Cortex (New York, NY: 1991) 7, 181–192.
Fracasso, A., Petridou, N., Dumoulin, S.O., 2016. Systematic variation of population receptive field properties across cortical depth in human visual cortex. Neuroimage 139, 427–438.
Freeman, J., Brouwer, G.J., Heeger, D.J., Merriam, E.P., 2011. Orientation decoding depends on maps, not columns. J. Neurosci. 31, 4792–4804.
Freeman, J., Heeger, D.J., Merriam, E.P., 2013. Coarse-scale biases for spirals and orientation in human visual cortex. J. Neurosci. 33, 19695–19703.
Gardner, J.L., Merriam, E.P., 2021. Population models, not analyses, of human neuroscience measurements. Annu. Rev. Vis. Sci. 7, 225–255.

Gravel, N., Harvey, B., Nordhjem, B., Haak, K.V., Dumoulin, S.O., Renken, R., Curčić-Blake, B., Cornelissen, F.W., 2014. Cortical connective field estimates from resting state fMRI activity. Front. Neurosci. 8, 339.

Grinvald, A., Lieke, E., Frostig, R.D., Gilbert, C.D., Wiesel, T.N., 1986. Functional architecture of cortex revealed by optical imaging of intrinsic signals. Nature 324, 361–364.

Groen, I.I.A., Piantoni, G., Montenegro, S., Flinker, A., Devore, S., Devinsky, O., Doyle, W., Dugan, P., Friedman, D., Ramsey, N.F., Petridou, N., Winawer, J., 2022. Temporal dynamics of neural responses in human visual cortex. J. Neurosci. 42, 7562–7580.

Haak, K.V., Winawer, J., Harvey, B.M., Renken, R., Dumoulin, S.O., Wandell, B.A., Cornelissen, F.W., 2013. Connective field modeling. Neuroimage 66, 376–384.

Harvey, B.M., Vansteensel, M.J., Ferrier, C.H., Petridou, N., Zuiderbaan, W., Aarnoutse, E.J., Bleichner, M.G., Dijkerman, H.C., van Zandvoort, M.J., Leijten, F.S., Ramsey, N.F., Dumoulin, S.O., 2013. Frequency specific spatial interactions in human electrocorticography: V1 alpha oscillations reflect surround suppression. Neuroimage 65, 424–432.

Haynes, J.D., Rees, G., 2005. Predicting the orientation of invisible stimuli from activity in human primary visual cortex. Nat. Neurosci. 8, 686–691.

Heeger, D.J., 1992. Normalization of cell responses in cat striate cortex. Vis. Neurosci. 9, 181–197.

Henderson, M.M., Tarr, M.J., Wehbe, L., 2023a. Low-level tuning biases in higher visual cortex reflect the semantic informativeness of visual features. J. Vis. 23, 8.

Henderson, M.M., Tarr, M.J., Wehbe, L., 2023b. A texture statistics encoding model reveals hierarchical feature selectivity across human visual cortex. J. Neurosci. 43, 4144–4161.

Hermes, D., Petridou, N., Kay, K.N., Winawer, J., 2019. An image-computable model for the stimulus selectivity of gamma oscillations. Elife 8, e47035.

Hubel, D.H., Wiesel, T.N., 1962. Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. J. Physiol. 160, 106–154.

Hubel, D.H., Wiesel, T.N., 1968. Receptive fields and functional architecture of monkey striate cortex. J. Physiol. 195, 215–243.

Huth, A.G., Nishimoto, S., Vu, A.T., Gallant, J.L., 2012. A continuous semantic space describes the representation of thousands of object and action categories across the human brain. Neuron 76, 1210–1224.

Infanti, E., Schwarzkopf, D.S., 2020. Mapping sequences can bias population receptive field estimates. Neuroimage 211, 116636.

Kamitani, Y., Tong, F., 2005. Decoding the visual and subjective contents of the human brain. Nat. Neurosci. 8, 679–685.

Kay, K.N., Naselaris, T., Prenger, R.J., Gallant, J.L., 2008. Identifying natural images from human brain activity. Nature 452, 352–355.

Kay, K.N., Winawer, J., Mezer, A., Wandell, B.A., 2013. Compressive spatial summation in human visual cortex. J. Neurophysiol. 110, 481–494.

Kay, K.N., Weiner, K.S., Grill-Spector, K., 2015. Attention reduces spatial uncertainty in human ventral temporal cortex. Curr. Biol. 25, 595–600.

Kay, K., Jamison, K.W., Zhang, R.Y., Uğurbil, K., 2020. A temporal decomposition method for identifying venous effects in task-based fMRI. Nat. Methods 17, 1033–1039.

Kay, K., Bonnen, K., Denison, R.N., Arcaro, M.J., Barack, D.L., 2023. Tasks and their role in visual neuroscience. Neuron 111, 1697–1713.

Khosla, M., Ratan Murty, N.A., Kanwisher, N., 2022. A highly selective response to food in human visual cortex revealed by hypothesis-free voxel decomposition. Curr. Biol. 32, 4159–4171. e4159.

Kim, I., Kupers, E.R., Lerma-Usabiaga, G., Grill-Spector, K., 2023. Characterizing spatiotemporal population receptive fields in human visual cortex with fMRI. J. Neurosci. 44 (2), e0803232023.
Klein, B.P., Harvey, B.M., Dumoulin, S.O., 2014. Attraction of position preference by spatial attention throughout human visual cortex. Neuron 84, 227–237.
Knapen, T., 2021. Topographic connectivity reveals task-dependent retinotopic processing throughout the human brain. Proc. Natl. Acad. Sci. U. S. A. 118 (2), e2017032118.
Lee, S., Papanikolaou, A., Logothetis, N.K., Smirnakis, S.M., Keliris, G.A., 2013. A new method for estimating population receptive field topography in visual cortex. Neuroimage 81, 144–157.
Lerma-Usabiaga, G., Benson, N., Winawer, J., Wandell, B.A., 2020. A validation framework for neuroimaging software: the case of population receptive fields. PLoS Comput. Biol. 16, e1007924.
Lerma-Usabiaga, G., Winawer, J., Wandell, B.A., 2021. Population receptive field shapes in early visual cortex are nearly circular. J. Neurosci. 41, 2420–2427.
Mante, V., Sussillo, D., Shenoy, K.V., Newsome, W.T., 2013. Context-dependent computation by recurrent dynamics in prefrontal cortex. Nature 503, 78–84.
Marks, T.D., Goard, M.J., 2021. Stimulus-dependent representational drift in primary visual cortex. Nat. Commun. 12, 5169.
Niell, C.M., Stryker, M.P., 2010. Modulation of visual responses by behavioral state in mouse visual cortex. Neuron 65, 472–479.
Portilla, J., Simoncelli, E.P., 2000. A parametric texture model based on joint statistics of complex wavelet coefficients. Int. J. Comput. Vis. 40, 49–70.
Reid, R.C., Victor, J.D., Shapley, R.M., 1997. The use of m-sequences in the analysis of visual neurons: linear receptive field properties. Vis. Neurosci. 14, 1015–1027.
Roth, Z.N., Merriam, E.P., 2023. Representations in human primary visual cortex drift over time. Nat. Commun. 14, 4422.
Roth, Z.N., Heeger, D.J., Merriam, E.P., 2018. Stimulus vignetting and orientation selectivity in human visual cortex. Elife 7, e37241.
Roth, Z.N., Ryoo, M., Merriam, E.P., 2020. Task-related activity in human visual cortex. PLoS Biol. 18, e3000921.
Roth, Z.N., Kay, K., Merriam, E.P., 2022. Natural scene sampling reveals reliable coarse-scale orientation tuning in human V1. Nat. Commun. 13, 6469.
Savier, E.L., Chen, H., Cang, J., 2019. Effects of locomotion on visual responses in the mouse superior colliculus. J. Neurosci. 39, 9360–9368.
Schröder, S., Steinmetz, N.A., Krumin, M., Pachitariu, M., Rizzi, M., Lagnado, L., Harris, K.D., Carandini, M., 2020. Arousal modulates retinal output. Neuron 107, 487–495. e489.
Sereno, M.I., Dale, A.M., Reppas, J.B., Kwong, K.K., Belliveau, J.W., Brady, T.J., Rosen, B. R., Tootell, R.B., 1995. Borders of multiple visual areas in humans revealed by functional magnetic resonance imaging. Science 268, 889–893.
Silson, E.H., Chan, A.W., Reynolds, R.C., Kravitz, D.J., Baker, C.I., 2015. A retinotopic basis for the division of high-level scene processing between lateral and ventral human occipitotemporal cortex. J. Neurosci. 35, 11921–11935.
Silson, E.H., Reynolds, R.C., Kravitz, D.J., Baker, C.I., 2018. Differential sampling of visual space in ventral and dorsal early visual cortex. J. Neurosci. 38, 2294–2303.

Simoncelli, E.P., Freeman, W.T., 1995. The steerable pyramid: a flexible architecture for multi-scale derivative computation. In: Proceedings International Conference on Image Processing. IEEE, pp. 444–447.

Simoncelli, E.P., Freeman, W.T., Adelson, E.H., Heeger, D.J., 1992. Shiftable multiscale transforms. IEEE Trans. Inf. Theory 38, 587–607.

Sirotin, Y.B., Das, A., 2009. Anticipatory haemodynamic signals in sensory cortex not predicted by local neuronal activity. Nature 457, 475–479.

Stigliani, A., Jeska, B., Grill-Spector, K., 2017. Encoding model of temporal processing in human visual cortex. Proc. Natl. Acad. Sci. U. S. A. 114, E11047–e11056.

St-Yves, G., Naselaris, T., 2018. The feature-weighted receptive field: an interpretable encoding model for complex feature spaces. Neuroimage 180, 188–202.

van Es, D.M., Theeuwes, J., Knapen, T., 2018. Spatial sampling in human visual cortex is modulated by both spatial and feature-based attention. Elife 7, e36928.

Wilson, R.C., Collins, A.G., 2019. Ten simple rules for the computational modeling of behavioral data. Elife 8, e49547.

Womelsdorf, T., Anton-Erxleben, K., Pieper, F., Treue, S., 2006. Dynamic shifts of visual receptive fields in cortical area MT by spatial attention. Nat. Neurosci. 9, 1156–1160.

Zhou, J., Benson, N.C., Kay, K.N., Winawer, J., 2018. Compressive temporal summation in human visual cortex. J. Neurosci. 38, 691–709.

Ziemba, C.M., Freeman, J., Movshon, J.A., Simoncelli, E.P., 2016. Selectivity and tolerance for visual texture in macaque V2. Proc. Natl. Acad. Sci. U. S. A. 113, E3140–E3149.

Zuiderbaan, W., Harvey, B.M., Dumoulin, S.O., 2012. Modeling center-surround configurations in population receptive fields using fMRI. J. Vis. 12, 10.