Extremely fast pRF mapping for real-time applications

Salil Bhat, Michael Lührs, Rainer Goebel, Mario Senden

ARTICLE INFO

Keywords: Population receptive field mapping Real-time fMRI Vision Stimulus encoding

ABSTRACT

Population receptive field (pRF) mapping is a popular tool in computational neuroimaging that allows for the investigation of receptive field properties, their topography and interrelations in health and disease. Furthermore, the possibility to invert population receptive fields provides a decoding model for constructing stimuli from observed cortical activation patterns. This has been suggested to pave the road towards pRF-based brain-computer interface (BCI) communication systems, which would be able to directly decode internally visualized letters from topographically organized brain activity. A major stumbling block for such an application is, however, that the pRF mapping procedure is computationally heavy and time consuming. To address this, we propose a novel and fast pRF mapping procedure that is suitable for real-time applications. The method is built upon hashed-Gaussian encoding of the stimulus, which tremendously reduces computational resources. After the stimulus is encoded, mapping can be performed using either ridge regression for fast offline analyses or gradient descent for real-time applications. We validate our model-agnostic approach in silico, as well as on empirical fMRI data obtained from 3T and 7T MRI scanners. Our approach is capable of estimating receptive fields and their parameters for millions of voxels in mere seconds. This method thus facilitates real-time applications of population receptive field mapping.

1. Introduction

The retinotopic organization of the human visual cortex has intrigued neuroscientists ever since the beginning of the nineteenth century when visual field maps were first discovered in soldiers suffering from occipital wounds (Holmes, 1918; Inouye, 1999). With the advent of functional magnetic resonance imaging (fMRI) in the early 1990s (Kwong et al., 1992; Ogawa et al., 1992), it became possible to map retinotopy non-invasively (Sereno et al., 1995; DeYoe et al., 1996; Engel et al., 1997). Sereno et al. (1995) pioneered a phase encoding procedure that allowed for the systematic investigation of polar angle and eccentricity distributions. More recently, Dumoulin and Wandell (2008) spearheaded the population receptive field (pRF) mapping approach which provided an expandable, parametric, model of receptive fields. This allowed researchers to study additional properties of receptive fields and their topography as well as relationships between receptive field properties.

The pRF approach has, for instance, enabled researchers to understand the relationship between eccentricity and the size of receptive fields along the visual hierarchy (Dumoulin and Wandell 2008; Amano et al., 2009; Harvey and Dumoulin 2011; Silva et al., 2018), to investigate neural plasticity and visual development from childhood to adulthood (Dekker et al., 2019; Gomez et al., 2018) and to study the dynamic changes of receptive fields in response to attention (Klein et al., 2014; Kay et al., 2015; Yo et al., 2017; Es et al., 2018). Furthermore, pRF modeling has aided researchers’ investigations of pathology such as Alzheimer’s disease (Brewer and Barton 2014), schizophrenia (Anderson et al., 2017), albinism (Ahmadi et al., 2019), (artificial) scotoma in patients with macular degradation (Baseler et al., 2011; Barton and Brewer 2015; Papanikolau et al., 2015) and in healthy populations (Binda et al., 2013; Haak et al., 2012; Hummer et al., 2018) and even blindness (Georgy et al., 2019). Additionally, the ability to estimate receptive field parameters is crucial for a number of applications. For instance, receptive fields can serve as a target for transcranial magnetic stimulation (Sack et al., 2009) or provide a spatial forward model for computational models (Peters et al., 2012). Furthermore, receptive fields can be inverted to provide a decoding model for reconstructing perceived, as well as imagined, visual stimuli (Thirion et al., 2006; Senden et al., 2019).

* Corresponding author at: Department of Cognitive Neuroscience, Faculty of Psychology and Neuroscience, Maastricht University, Maastricht, the Netherlands.

https://doi.org/10.1016/j.neuroimage.2021.118671.
Received 25 March 2021; Received in revised form 28 September 2021; Accepted 20 October 2021
Available online 26 October 2021.

1053-8119/© 2021 The Author(s). Published by Elsevier Inc. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/)
The latter has been suggested to pave the road towards pRF-based brain-computer interface (BCI) communication systems able to directly decode internally visualized letters from topographically organized brain activity (Senden et al., 2019). This is hindered, however, by the method’s immense consumption of computational time and resources. This issue largely remains unaddressed, although some recent work (Thielen et al., 2019) has proposed a fast deep-learning based mapping algorithm (DeepRF). The DeepRF method deploys a deep convolutional neural network (ResNet) which receives a time-series as input and predicts the corresponding pRF parameters. Once the network is trained, pRF parameters can be estimated simply using a rapid forward pass. This method is indeed faster than standard methods such as grid-search and achieves faithful estimation of pRF parameters with an average computational time of 0.01 to 0.03 seconds per voxel. However, the procedure requires the generated simulated data (for training) and the empirical data to have the same experimental design. Hence, for empirical data with a new experimental design, the network needs to be trained again and the training of the deep neural network can take up to several hours. Moreover, MRImi data typically contains a large number of voxels. The high spatial resolution at field strengths of 3 and 7 Tesla leads to hundreds of thousands or even millions of voxels. Therefore, despite achieving low computational time per voxel, the total computational time for all voxels is on the order of several minutes. This makes the approach unfeasible for real-time analysis. With the aim to enable estimation of receptive fields in real-time, we propose here a novel model-agnostic procedure which can be used offline (using ridge regression) as well as online (using gradient descent).

The method relies on regularized linear regression whose basis set is a hashed-Gaussian encoding of the stimulus-evoked response. Specifically, the stimulus space is exhaustively partitioned as a set of features where each feature uniquely encodes the stimulus by computing the overlap between the stimulus and a set of randomly positioned Gaus- sians. This type of encoding considerably reduces the memory require- ments and thereby accelerates calculations.

Using two previously acquired datasets from 3 Tesla and 7 Tesla MR systems, we show that the proposed approach works extremely fast. It is able to estimate receptive field shapes of millions of voxels within seconds. This allows the selection of visually responsive voxels through cross-validation and subsequent estimation of receptive field parameters within about one minute even if the data consists of more than 4 million voxels.

2. Methods

2.1. Fast mapping procedure (f-pRF)

2.1.1. Tile coding and hashing

To reduce computation time as well as to lower memory require- ments, we encode the stimulus using tile coding and hashing (Sutton and Barto, 2018). Tile coding is a linear function approximation used in reinforcement learning to deal with large and continuous state spaces. In tile coding, the state space is exhaustively partitioned into sub-regions called tiles. Usually, the presence of an entity within a tile (in this case, the presence of a stimulus in a region of the visual field) is encoded in a binary fashion. However, it is also possible to encode features using radial basis functions which have the additional benefit of varying smoothly. Memory requirements can be reduced further by hashing a group of individual, non-contiguous, tiles into a single tile. Fig. 1 de- picts tile-coding and hashing of sample stimuli. The presence of a stim- ulus is encoded as the extent of overlap between the stimulus and hashed tiles. For our purposes, we use a 2-D isotropic Gaussian as the radial basis function. Subsequently, we hash by combining five randomly se- lected Gaussians into a single tile leading to a total of 250 tiles. The 5 Gaussian tiles within a tiling may or may not overlap. We normalized each tile to ensure that the volume under its surface is equal to one.

2.1.2. Encoding stimuli

Using hashed-Gaussians as tiles, it is possible to encode retinotopic stimuli. First, an overlap between a binary indicator function and a tiling matrix $\Gamma$ (pixels-by-tiles) is computed. The binary indicator function $S$ (time-by-pixels) marks the position of the stimulus aperture at each moment in time. Subsequently, the computed overlap is convolved with a canonical two-gamma hemodynamic response function (HRF) function $(\psi)$ to obtain the encoded stimulus $\phi$:

$$\phi = ST \ast h$$

(1)

2.1.3. Ridge regression (f-pRF$_{ridge}$)

We use ridge regression for fast offline pRF mapping (i.e., after all functional volumes have been acquired). Specifically, the blood-oxygen-level dependent (BOLD) response is modeled by

$$B = \phi \theta + \epsilon$$

(2)

where $\theta$ are the estimated weights and $\epsilon$ denotes the residuals. Note that, prior to computing $\theta$, both $\phi$ and the BOLD data $B$ are z-normalized. In order to estimate $\theta$, the discrepancy between the measured and predicted BOLD response $(\phi \theta)$ needs to be minimized. Therefore, we define the error or the loss as

$$E = \frac{1}{2} \| B - \phi \theta \|^2 + \frac{1}{2} \lambda \| \theta \|^2$$

(3)

In order to avoid over-fitting, we use $L_2$ regularization and $\lambda$ denotes the regularization factor. The gradient of the error with respect to $\theta$ can be computed as

$$\frac{\partial E}{\partial \theta} = -\phi^T B + \phi^T \phi \theta + \lambda \theta$$

(4)

By setting $\frac{\partial E}{\partial \theta} \rightarrow 0$ and solving for optimal $\theta$, we get

$$\theta = (\phi^T \phi + \lambda I)^{-1} \phi^T B$$

(5)

Receptive fields can now be straightforwardly obtained by multiplying the tiling matrix with the estimated $\theta$: $W = \Gamma \theta$. These raw receptive fields are then subjected to post-processing as they tend to contain anomalous pixel intensities. These can be removed by first normalizing the raw receptive fields to the range $[0, 1]$ and then shrinking them by raising each entry to a power of some positive integer (shrinkage factor). This shrinks noisy pixel intensities close to 0 while leaving those close to 1 unaffected (Fig. 2), thus yielding cleaner receptive fields.

2.1.4. Similarity metric

In order to compare the receptive fields obtained from ridge regres- sion with corresponding ground-truth/grid-search receptive fields, we use the Jaccard Index (JI; or Jaccard Similarity). The Jaccard Index is bound to $[0, 1]$ (give that the quantities that are being compared are in the same range), with 1 indicating that receptive fields are identical. Since the Jaccard Index is a conservative metric that is highly sensi- tive to residual anomalous pixel intensities, we aid interpretation of our
results by providing a **Null-model** generated from a re-sampling procedure. Specifically, for each estimated receptive field, we pair it with a random ground-truth/grid-search receptive field and compute the JI. The average over these pairs is the JI of one randomization. We repeat this procedure 1000 times to obtain a Null-distribution of randomized JIs. We refer to the mean of the Null-distribution as the **baseline**.

### 2.2. Online gradient descent (f-pRF$_{gd}$)

For online pRF mapping we use gradient descent to iteratively update $\theta$ with each acquired volume. In this case, we define the loss function as

$$E = \frac{1}{2} (B - \phi \theta)^T (B - \phi \theta)$$

The gradient of the loss function with respect to the parameter $\theta$ is

$$\frac{\partial E}{\partial \theta} = -\phi^T B + \phi^T \phi \theta$$

At each time point, $\theta$ is updated by a factor (learning rate $\eta$) of the gradient. Note that, unlike ridge regression, a regularization term is not needed in this case, as gradient descent is effectively regularized by the learning rate (see Appendix A). Considering the $n^{th}$ time point, the update can be computed as

$$\theta_n = \theta_{n-1} + \eta (\phi^T B - \phi^T \phi \theta_{n-1})$$

Similar to the offline method, prior to tile coding and hashing, the stimulus needs to be convolved with the HRF. Furthermore, both the BOLD signal $B$ and encoded stimulus $\phi$ need to be z-normalized. However, in an online setting this needs to be performed in real-time. Real-time z-normalization requires real-time estimation of the mean and variance of a signal which can be done using Welford’s online algorithm (Welford, 1962). Once the current mean $\bar{x}(t)$ and variance $\sigma^2(t)$ have been estimated, the current z-score can be estimated as

$$Z_n = \frac{x_n - \bar{x}}{\sigma_n}$$

### 2.2.1. Voxel selection

Since not all measured voxels are visual, and hence may not carry significant information, a voxel selection procedure is desirable. We evaluate voxels in terms of the cross-validated Pearson correlation coefficient (fitness) between their predicted and measured BOLD responses. To account for temporal autocorrelation in the BOLD response, we use a blocked cross-validation procedure (Roberts et al., 2017). Specifically, the data is split into $p$ windows along the time axis. Ridge regression is performed on window 1 and the estimated $\theta$ values are used to predict the BOLD response for the remaining $p - 1$ windows. This is followed by ridge regression on windows 1 and 2 and predicting the BOLD response in the remaining $p - 2$ windows. This procedure continues until ridge regression is performed on windows 1 to $p - 1$ and the BOLD response is predicted for the $p^{th}$ (last) window. The overall fitness for each voxel is then given by the mean of fitness values computed for each split. The data used in this paper has 304 time points. We split the data into 4 windows of equal length and retain voxels whose fitness falls within the top 1%.

### 2.3. Fast pRF parameter estimation

Post-processed receptive fields obtained from our ridge regression and gradient descent methods can be readily used to estimate parameters of an isotropic Gaussian pRF model (i.e. the x-location, y-location and size) using a fast procedure. Since peak pixel intensity of a Gaussian receptive field is at its center, we estimate the x- and y-coordinate of pre-processed model-free receptive fields by finding the location of their peak pixel intensity. To estimate the size of receptive fields, our procedure utilizes the relationship between the standard deviation, eccentricity and the mean pixel intensity in an isotropic Gaussian embedded in a finite image. Specifically, given a Gaussian at a fixed location, mean pixel intensity increases as a function of its standard deviation. Furthermore, in a finite image and assuming a fixed size, mean pixel intensity decreases as the Gaussian is progressively moved toward the edge of an image. Therefore, for a given image size, we generate isotropic Gaussians with 25 different standard deviations, located at 25 eccentricities along an axis of 45°, and compute their mean pixel intensities. This can be utilized to perform a linear regression with mean pixel intensity and eccentricities predicting the receptive field size. We then use the resulting regression weights together with previously estimated locations and mean pixel intensity of our receptive fields to obtain an estimate of their size.

### 2.4. Grid search (GS-pRF)

In order to compare our approach with a standard grid-search based approach, we use the grid-search described in (Senden et al., 2019). In order to search for pRF location, this approach splits the visual field into 100-by-100 circular grid points whose density decays exponentially with eccentricity. Close to the fovea the grid is thus denser than in the periphery, taking cortical magnification into account. Since there is a well established linear relationship between eccentricity and the size of pRF (Freeman and Simoncilli, 2011), the grid-search explores a range of slopes (from 0.1 to 1 in steps of 0.1) of the size-eccentricity relationship rather than pRF sizes directly. This effectively allows for exploration of a greater range of receptive field sizes. The code for this approach (as well as our new method) is part of the Computational Neuroimaging Toolbox (available at https://github.com/ccmnmaastricht/CNI_toolbox).

### 2.5. Data

#### 2.5.1. Simulated data

We simulate fMRI data for a V1-like cortical sheet extending 55 mm along and approximately 40 mm orthogonal to the horizontal meridian in both hemispheres. Since such a sheet is akin to a flattened cortical mesh, model units are referred to as vertices rather than voxels. Each vertex in the model is a 0.5 mm isotropic patch whose receptive field center is directly related to its position on the surface in accordance with a complex-logarithmic topographic mapping (Schwartz, 1980; Balasubramanian et al., 2002) with parameter values ($\alpha = 0.7, a = 0.9, Polimeni et al., 2005$). The shape of model receptive fields is given by a 2-dimensional Gaussian

$$f(x, y) = \exp - \frac{(x - \mu_x)^2 + (y - \mu_y)^2}{2\sigma^2}$$

Fig. 2. The effect of shrinking a raw receptive field. a, Raw receptive field displaying undesirably large pixel intensities. b, The receptive field after shrinkage with a factor of 9. c, The corresponding ground truth receptive field.
with \((\mu_x, \mu_y)\) being the receptive field center and \(\sigma\) its size. Below an eccentricity of \(\sigma = 2.38\) all model vertices have a receptive field size of \(\sigma = 0.5\) whereas they exhibit a linear relationship with eccentricity \((\sigma = 0.21e)\) beyond this cutoff (c.f. Freeman and Simoncelli, 2011).

A simulated fMRI signal (sampled at a rate of 0.5 Hz) for each vertex is obtained by first performing element-wise multiplication between the receptive field of a vertex and the effective stimulus presented per time point, summing the result and subsequently convolving the obtained signal with the canonical two gamma hemodynamic response function. Two sources of distortion are added to the signal. First, a spatial smoothing kernel is applied to simulate the point-spread function of BOLD activity on the surface of the striate cortex (Shmuel et al., 2007). Second, autocorrelated noise generated by an Ornstein-Uhlenbeck process with variance \(\sigma^2_{noise} = 0.5\) is added. The smoothing kernel is independently applied to the clean signal and the noise before the two are combined. We simulate both 3T- and 7T-like signals by adjusting the full-width at half-maximum of the spatial smoothing kernel (3.5 mm and 2 mm for 3T and 7T, respectively; c.f. Shmuel et al., 2007) and the time constant of the Ornstein-Uhlenbeck process (2.25 s and 1 s for 3T and 7T, respectively).

2.5.2. Three tesla empirical data

This dataset, previously described in (Senden et al., 2014), comprises a retinotopy run obtained from three participants (all male, age range = 27–35 years, mean age = 32 years). During this run a bar aperture (1.5° wide) revealing a flickering checkerboard pattern (10 Hz) was presented in four orientations. For each orientation, the bar covered the entire screen in 12 discrete steps (each step lasting 2 s). Within each orientation, the sequence of steps (and hence of the locations) was randomized and each orientation was presented six times. Furthermore, within each presentation four bar stimuli were replaced with mean luminance images for four consecutive steps. These data were acquired on a Siemens 3T Tim Trio scanner equipped with a 32-channel head coil (Siemens, Erlangen, Germany) using a gradient-echo-planar imaging sequence (31 transversal slices; TR = 2000 ms; TE = 30 ms; FA = 77°; FoV = 216 × 216 mm²; 2 mm isotropic resolution; no slice gap; GRAPPA = 2). Preprocessing consisted of slice scan time correction, (rigid body) motion correction, linear trend removal, and temporal high-pass filtering (up to 2 cycles per run).

For visualization purposes, we also include anatomical data for subject 1. Anatomical data were collected with a T1-weighted MPRAGE imaging sequence (192 sagittal slices; TR = 2250 ms; TE = 2.17 ms; FA = 9°; FoV = 256 × 256 mm²; 1 mm isotropic resolution; GRAPPA = 2). In the anatomical images, the gray/white matter boundary was detected and segmented using the advanced automatic segmentation tools of BrainVoyager 20 which are optimized for high-field MRI data. A region-growing approach analyzed local intensity histograms, corrected topological errors of the segmented gray/white matter border, and finally reconstructed meshes of the cortical surfaces (Kriegeskorte and Goebel 2001; Goebel et al., 2006). The data is available at 10.34894/ITWFZP. The dataset is available at doi:10.5061/dryad.mb8h6.

2.5.3. Seven tesla empirical data

This dataset, previously described in (Senden et al., 2019), comprises retinotopy as well as passive viewing of letter stimuli obtained from six participants (2 female, age range = 21–49 years, mean age = 30.7 years). During the retinotopy run a bar aperture (1.33° wide) revealing a flickering checkerboard pattern (10 Hz) was presented in four orientations. For each orientation, the bar covered the entire screen in 12 discrete steps (each step lasting 3 s). Within each orientation, the sequence of steps (and hence of the locations) was randomized and each orientation was presented six times. During the passive viewing run four letters ('H', 'T', 'S' and 'C') were presented in a 8° by 8° bounding frame for a duration of 6 s and their shape was filled with a flickering checkerboard pattern (10 Hz). These data were acquired on a Siemens Magnetom 7T scanner (Siemens; Erlangen, Germany) equipped with a 32-channel head-coil (Nova Medical Inc.; Wilmington, MA, USA) using high-resolution gradient echo-planar imaging sequence (82 transversal slices; TR = 3000 ms; TE = 26 ms; generalized auto-calibrating partially parallel acquisitions (GRAPPA) factor = 3; multi-band factor = 2; FA = 55°; FoV = 186 × 186 mm²; 0.8 mm isotropic resolution). In addition, this dataset includes five functional volumes acquired with opposed phase encoding directions to correct for EPI distortions that occur at higher field strengths (Anderson et al., 2003). Preprocessing further consisted of (rigid body) motion correction, linear trend removal, and temporal high-pass filtering (up to 3 cycles per run).

For visualization purposes, we also include anatomical data for subject 3. Anatomical data was acquired with a T1-weighted magnetization prepared rapid acquisition gradient echo (Marques et al., 2010) sequence (240 sagittal slices, matrix = 320 × 320 m, voxel size = 0.7 mm isotropic, first inversion time TI1 = 900 ms, second inversion time TI2 = 2750 ms, echo time (TE) = 2.46 ms repetition time (TR) = 5000 ms, first nominal flip angle = 5°, and second nominal flip angle = 3° Anatomical images were interpolated to a nominal resolution of 0.8 mm isotropic to match the resolution of functional images. In the anatomical images, the gray/white matter boundary was detected and segmented using the advanced automatic segmentation tools of BrainVoyager 20 which are optimized for high-field MRI data. A region-growing approach analyzed local intensity histograms, corrected topological errors of the segmented gray/white matter border, and finally reconstructed meshes of the cortical surfaces (Kriegeskorte and Goebel 2001; Goebel et al., 2006). The data is available at 10.34894/ITWFZP.

2.6. Real-time processing

To mimic a real-time scenario, we limited the preprocessing of functional data to trilinear 3D rigid body motion correction which was applied in a simulated real-time setup using Turbo-BrainVoyager (TBV; v4.0b1, Brain Innovation B.V., Maastricht, The Netherlands). The data was accessed directly from TBV using a network interface providing fast transfer speed suitable for real-time applications. The receiver was implemented in MATLAB (version 2019a, The Mathworks .inc, Natick, MA, USA) using JAVA based TCP/IP interfaces.

3. Results

All experiments were performed using MATLAB (version 2019a, The Mathworks .inc, Natick, MA, USA) running on an HP Z440 workstation with an Intel Xeon Processor (E5–1650 v4, 32GB RAM) and an Ubuntu 20.04 operating system. The set of hyperparameters (learning rate \(\eta = 0.1\), shrinkage factor = 6 and \(FWHM = 0.15\)) remain the same for all experiments, except for reconstruction of perceived letter shapes where a shrinkage factor of 9 was used. This set of hyperparameters was chosen using Bayesian Optimization performed on a simulated dataset with Jaccard Distance as loss function. All the figures generated using MATLAB (including the parts of Figs. 1 and 2) were generated using export.fig (Altman, 2020). Our methods are available as part of the Computational Neuroimaging Toolbox (https://github.com/ccnmaastricht/CNN_toolbox). Example scripts for using our approach are available at https://github.com/ccnmaastricht/real-time_pRF.

3.1. Fast mapping procedure

3.1.1. Simulated data

The fast, ridge-based, mapping procedure (f-pRFridge) was first tested on simulated data to investigate whether it faithfully recovers known population receptive field shapes and their parameters. Overall, the mean Jaccard Similarity (JS) between the estimated and ground-truth receptive field shapes was 0.3452 (95% CI [0.3409, 0.3495]) and 0.3920 (95% CI [0.3877, 0.3963]), for simulated 3T and 7T data, respectively. For comparison, corresponding Null-model JS values were 0.0418 and
This page discusses the comparison of ridge-estimated and ground-truth receptive fields. It mentions that Bayesian optimization maximized fidelity of the receptive field shape rather than pRF parameters. In the case of this dataset, a larger shrinkage factor would result in better pRF size estimates at a cost of somewhat worse estimates of receptive field shapes. This indicates that there is slight trade-off between these two goals but appropriate choices of the shrinkage parameter allow for faithful estimates of either.

Next, we evaluated the f-pRF$_{ridge}$ approach in terms of its computational performance. To that end we measured both memory consumption and the computation time required for the mapping procedure itself as well as for subsequent parameter estimation. Computation times were estimated using MATLAB’s stopwatch utility. The execution time measured using this utility can be affected by many unknown variables pertaining to memory, processor, memory caching, MATLAB’s just-in-time compiler, etc. This may influence the execution time measurement each time a subroutine is executed. Therefore, we report computation times as a mean over 100 runs. Memory requirements were estimated using GNU/Linux’s `pmap` command. The memory requirements reported here are calculated as $\text{mem}_{\text{max}} - \text{mem}_0$, where $\text{mem}_{\text{max}}$ is the maximum amount of memory consumed during the procedure and $\text{mem}_0$ is the memory occupied by MATLAB before starting the procedure.

### Table 1

Correlations between estimated and ground-truth pRF parameters.

|          | X-coordinate  | Y-coordinate  | $\sigma$ ($\sigma$) |
|----------|---------------|---------------|---------------------|
| 3T       | 0.9913 (95% CI [0.9909, 0.9916]) | 0.9871 (95% CI [0.9865, 0.9876]) | 0.9674 (95% CI [0.9660, 0.9686]) |
| 7T       | 0.9958 (95% CI [0.9957, 9960]) | 0.9949 (95% CI [0.9946, 0.9951]) | 0.9681 (95% CI [0.9668, 0.9693]) |

0.0410, respectively. There is thus good correspondence between estimated and ground-truth receptive field shapes which is also apparent from the sample receptive fields shown in Fig. 3. Next, we examined the correspondence between receptive field parameters obtained with the two methods. While receptive fields mapped using the f-pRF$_{ridge}$ are not exactly Gaussian, estimated parameters nevertheless show an excellent correspondence with ground truth parameters for both simulated 3T and 7T data (see Figs. 4 and C.20 respectively as well as Table 1).

Please note that despite the high correlation, the receptive field size tends to be slightly overestimated by our method. Fig. 5 shows that the size estimations can be improved by using a different shrinkage factor (in this case, a higher one). The shrinkage factor we obtained from Bayesian optimization maximized fidelity of the receptive field shape rather than pRF parameters. In the case of this dataset, a larger shrinkage would result in better pRF size estimates at the cost of somewhat worse estimates of receptive field shapes. This indicates that there is slight trade-off between these two goals but appropriate choices of the shrinkage parameter allow for faithful estimates of either.

Fig. 3. Comparison of ridge-estimated and ground-truth receptive fields. a) Small (top) and large (bottom) estimated and ground-truth receptive fields for simulated 3T data ($TR = 2000\text{ms}$). b) Small (top) and large (bottom) estimated and ground-truth receptive fields for simulated 7T data ($TR = 3000\text{ms}$).

Fig. 4. Estimated vs ground-truth pRF parameters for a) f-pRF$_{ridge}$ and b) GS-pRF. A line with a slope of 1 is included as a reference. Voxels whose receptive fields lie outside the field of view were ignored for estimating pRF parameters. Results are from simulated 3T data. Results for simulated 7T data are comparable (see supplementary Fig. C.20).

Fig. 5. Estimated vs ground-truth pRF sizes (standard deviation) using f-pRF$_{ridge}$ with a shrinkage factor of 9.

Next, we evaluated the f-pRF$_{ridge}$ approach in terms of its computational performance. To that end we measured both memory consumption and the computation time required for the mapping procedure itself as well as for subsequent parameter estimation. Computation times were estimated using MATLAB’s stopwatch utility. The execution time measured using this utility can be affected by many unknown variables pertaining to memory, processor, memory caching, MATLAB’s just-in-time compiler, etc. This may influence the execution time measurement each time a subroutine is executed. Therefore, we report computation times as a mean over 100 runs. Memory requirements were estimated using GNU/Linux’s `pmap` command. The memory requirements reported here are calculated as $\text{mem}_{\text{max}} - \text{mem}_0$, where $\text{mem}_{\text{max}}$ is the maximum amount of memory consumed during the procedure and $\text{mem}_0$ is the memory occupied by MATLAB before starting the procedure.
(which includes loading of data into memory and other background processes occupying memory). Memory consumption during the procedure was logged every 0.1 second using GNU/Linux’s watch command. Note that since here we are only interested in computational performance, we test the mapping procedure on randomly generated data of the size 304-by-voxels. Memory consumption was averaged over 100 repetitions of the procedure. As can be appreciated from Fig. 6 the f-pRF procedure is extremely fast (less than 1 s for more than 3 million voxels) when compared to grid-search (GS-pRF), which may take up to several hours. The computation time only starts to increase as the needed memory exceeds the available memory. As a consequence, virtual memory gets consumed which slows down the mapping procedure. Memory consumption scales linearly with the number of voxels and allows for estimation of ~1.75 and ~3.5 million voxels on systems with 8 GB and 16 GB of RAM, respectively.

### 3.1.2. Empirical data

Following up on simulation results, we tested the f-pRF procedure on previously acquired empirical data. Similar to the simulated data, we assess our method in terms of its ability to estimate pRF shapes and their parameters as well as computational performance. Since ground truth receptive field shapes and parameters are not known for empirical data, we assess our method on its ability to produce estimates that are consistent with a GS-pRF mapping procedure (as described in Section 2.4). Sample receptive fields estimated in the 3T and the 7T empirical data are shown in Figs. 7 and 8, respectively. Retinotopic surface maps for a representative subject in the 7T and 3T datasets are shown in Fig. 9. These results qualitatively indicate a good agreement between receptive field location parameters estimated by our method and the GS-pRF approach. Quantitatively, we observe that the Jaccard similarity between receptive fields estimated using the f-pRF and GS-pRF methods consistently exceed those expected based on the Null model (Table 2). The JS is particularly high (with respect to the baseline) for subjects 3, 5 and 6 for the 7T empirical dataset. This is also apparent from scatter plots showing the correspondence between pRF parameters obtained from our method and GS-pRF in representative subjects (see Figure 10a and Figure 10b for the 3T and 7T dataset, respectively). In order to compare our approach with the GS-pRF in terms of predictability, we report mean, cross-validated, correlation coefficients (over top 1% voxels obtained using crossvalidation), between the predicted and the acquired BOLD signals, in Table 3. We also report the mean correlation coefficients over top 1% voxels obtained from GS-pRF (not crossvalidated) in Table C.9. It is clear from both the Tables 3 and C.9 that our method yields better correlation coefficients across subjects and datasets. It is to be noted that there exists a disagreement between the top 1% voxels selected using f-pRF and GS-pRF. We report the percentage overlap between the top 1% voxels obtained from f-pRF and the ones obtained from GS-pRF in Table C.10. The percentage overlap is in general high for the majority of the subjects. The most overlap we observed was 80.49% (subject 3, 7T data) and the least we observed was of 44.12% (subject 1, 7T data).

### Table 2

Mean Jaccard Similarities between receptive fields estimated with the fast procedure and those obtained from GS-pRF for (a) 3T and (b) 7T empirical data.

|    | Jaccard Similarity | baseline |
|----|-------------------|----------|
| a) |                   |          |
| S01| 0.2781 (95% CI [0.2721, 0.2840]) | 0.0781   |
| S02| 0.2607 (95% CI [0.2554, 0.2661]) | 0.0957   |
| S03| 0.3050 (95% CI [0.2993, 0.3107]) | 0.0905   |
| b) |                   |          |
| S01| 0.1939 (95% CI [0.1903, 0.1976]) | 0.0605   |
| S02| 0.3088 (95% CI [0.3044, 0.3132]) | 0.0673   |
| S03| 0.4996 (95% CI [0.4961, 0.5035]) | 0.0701   |
| S04| 0.3333 (95% CI [0.3295, 0.3371]) | 0.0707   |
| S05| 0.4527 (95% CI [0.4489, 0.4565]) | 0.0786   |
| S06| 0.4564 (95% CI [0.4532, 0.4595]) | 0.0808   |
Fig. 8. Comparison of f_pRFridge and ground-truth receptive field parameters for 7T data. a) Small (top) and large (bottom) estimated and GS-pRF receptive fields for subject 1. b-f) Same as panel a for subjects 2–6, respectively. The axes are centered at pRF location estimated using GS-pRF. The axes are centered at pRF location estimated using GS-pRF.

Fig. 9. Exemplary eccentricity and polar angle maps in both hemispheres of a) S03 from 7T data and b) S01 from 3T data. The upper row shows maps obtained using our fast parameter estimation procedure whereas the bottom row shows maps obtained using the GS-pRF procedure. In accordance with the correlation results between maps (see Table 5b), the two polar angle and eccentricity maps are visually highly similar. Surface maps for S05 from 7T data are shown in C.23.
estimation for validation, mapping estimation.

We obtain the second voxels.

We evaluate the memory consumption. We estimate both based on 100 runs for each dataset. Since each subject has a different number of voxels, for each run a subject was chosen randomly. The computation time is computed separately for f-pRFridge, cross-validation and parameter estimation. The mean computation times or execution times for both datasets are reported in Table 4a and b.

The computation times suggest that our algorithm is extremely fast in mapping receptive fields. The actual mapping procedure happens within a second for the 3T dataset and in a few seconds for the 7T dataset. Cross-validation, which selects the best voxels, finishes in a couple of seconds for the 3T dataset and takes less than a minute for the 7T dataset. The estimation of pRF parameters (for all voxels) also takes only a few seconds for both datasets. This means that receptive fields and their pRF parameters are readily available for further analysis.

3.2. Online gradient descent (f-pRFgd)

To demonstrate the capability of f-pRFgd to work in a real-time setting, we mimicked a real-time scenario using TurboBrainVoyager (as described in Section 2.6). We show in Appendix A that f-pRFridge and f-pRFgd yield similar receptive fields through hyperparameter sharing. Therefore, we do not provide an evaluation of the ability of the method to reliably estimate receptive field shapes and parameters. Instead, we evaluate its performance in terms of whether estimated receptive fields are suitable for projecting cortical activity back into the visual field. For that purpose we utilize data acquired as subjects passively viewed letter shapes previously described in (Senden et al., 2019). The reconstructions obtained from our approach (see Fig. 11) are recognizable and comparable to those obtained from receptive fields resulting from GS-pRF.

In order to assess the convergence of our algorithm, we keep track of how gradients (Fig. 12a) and Jaccard Similarity (Fig. 12b) change over time. The analysis was performed using simulated data. It can be seen from Fig. 12a that the relative change in weights converge very early during stimulus presentation and the Jaccard Similarity converges gradually over the time. There are some occasional peaks in the gradient change which are the result of an occasional overshooting. Our method facilitates tracking of gradients in real-time. Researchers can possibly use this to track changes in gradients or (hashed)weights in real-time and decide when to stop presenting the stimulus. Furthermore, the gradient-descent method can be used in conjunction with simulations to optimize experimental design.

Table 3
Mean correlation coefficients between predicted and acquired BOLD signal obtained from (a) 3T and (b) 7T empirical data using top 1% cross-validated voxels.

| a) | f-pRFridge | GS-pRF |
|---|---|---|
| S01 | 0.4424 (95% CI [0.4396,0.4452]) | 0.3800 (95% CI [0.3753,0.3846]) |
| S02 | 0.4413 (95% CI [0.4386,0.4441]) | 0.3750 (95% CI [0.3706,0.3794]) |
| S03 | 0.4494 (95% CI [0.4460,0.4528]) | 0.3850 (95% CI [0.3797,0.3903]) |

| b) | f-pRFridge | GS-pRF |
|---|---|---|
| S01 | 0.4237 (95% CI [0.4227,0.4246]) | 0.2369 (95% CI [0.2348,0.2390]) |
| S02 | 0.3855 (95% CI [0.3847,0.3864]) | 0.1772 (95% CI [0.1745,0.1799]) |
| S03 | 0.5501 (95% CI [0.5484,0.5519]) | 0.5312 (95% CI [0.5280,0.5345]) |
| S04 | 0.4115 (95% CI [0.4103,0.4127]) | 0.2962 (95% CI [0.2940,0.2985]) |
| S05 | 0.4797 (95% CI [0.4782,0.4812]) | 0.4277 (95% CI [0.4246,0.4307]) |
| S06 | 0.4928 (95% CI [0.4913,0.4943]) | 0.4206 (95% CI [0.4178,0.4234]) |

Table 4
Mean computation times in seconds for (a) 3T and (b) 7T empirical data. The average data size per subject reflects the amount of data processed by the algorithm at a time.

| a) | Mean | Average data size per subject |
|---|---|---|
| f-pRFridge | 0.2055 (95% CI [0.2010,0.2100]) | 199,931 Voxels x 304 Volumes |
| Cross-validation | 1.3495 (95% CI [1.3162,1.3828]) | 199,931 Voxels x 304 Volumes |
| Parameter estimation | 0.4494 (95% CI [0.4460,0.4528]) | 199,931 Voxels x 304 Volumes |

| b) | Mean | Average data size per subject |
|---|---|---|
| f-pRFridge | 7.0519 (95% CI [6.1177,7.9861]) | 4,200,164 Voxels x 304 Volumes |
| Cross-validation | 51.4157 (95% CI [46.7570,56.0744]) | 4,200,164 Voxels x 304 Volumes |
| Parameter estimation | 7.4357 (95% CI [7.2950,7.5764]) | 4,200,164 Voxels x 304 Volumes |
Table 5
Correlation coefficients between the pRF parameters obtained from the fast procedure and those obtained from GS-pRF for (a) 3T and (b) 7T datasets, respectively.

|   | X-coordinate | Y-coordinate | Standard Deviation |
|---|--------------|--------------|-------------------|
| a) |              |              |                   |
| S01| 0.7557       | 0.7381       | 0.2287            |
|    | (95% CI [0.7436, 0.7674]) | (95% CI [0.7252, 0.7505]) | (95% CI [0.2023, 0.2548]) |
| S02| 0.6906       | 0.7526       | 0.2809            |
|    | (95% CI [0.6758, 0.7048]) | (95% CI [0.7404, 0.7644]) | (95% CI [0.2552, 0.3062]) |
| S03| 0.7551       | 0.7270       | 0.1758            |
|    | (95% CI [0.7429, 0.7667]) | (95% CI [0.7137, 0.7398]) | (95% CI [0.1488, 0.2026]) |
| b) |              |              |                   |
| S01| 0.6411       | 0.6042       | 0.0386            |
|    | (95% CI [0.6294, 0.6525]) | (95% CI [0.5916, 0.6165]) | (95% CI [0.0190, 0.0581]) |
| S02| 0.7038       | 0.6238       | 0.0298            |
|    | (95% CI [0.6937, 0.7135]) | (95% CI [0.6116, 0.6356]) | (95% CI [0.0104, 0.0102]) |
| S03| 0.9723       | 0.9645       | 0.5603            |
|    | (95% CI [0.9712, 0.9734]) | (95% CI [0.9632, 0.9659]) | (95% CI [0.5359, 0.5834]) |
| S04| 0.8332       | 0.7761       | 0.0970            |
|    | (95% CI [0.8271, 0.8391]) | (95% CI [0.7682, 0.7838]) | (95% CI [0.0776, 0.1164]) |
| S05| 0.9066       | 0.8976       | 0.3332            |
|    | (95% CI [0.9030, 0.9100]) | (95% CI [0.8937, 0.9013]) | (95% CI [0.3157, 0.3505]) |
| S06| 0.9216       | 0.9249       | 0.3229            |
|    | (95% CI [0.9186, 0.9245]) | (95% CI [0.9220, 0.9277]) | (95% CI [0.2941, 0.3295]) |

Table 6
Mean computation time (in seconds) per volume per subject for the real-time mapping technique performed on (a) 3T and (b) 7T empirical data. For each subject, data for 304 vol was recorded. Since the first iteration (corresponding to the first volume) is usually abnormally high, it is not included for computing mean and standard deviation.

|   | Mean | First Iteration | Number of Voxels |
|---|------|-----------------|------------------|
| a) |      |                 |                  |
| S01| 0.1439 | 0.2648         | 229,125          |
| S02| 0.1172 | 0.2007         | 177,684          |
| S03| 0.1249 | 0.2181         | 192,984          |
| b) |      |                 |                  |
| S01| 2.3974 | 5.1320         | 4553,058         |
| S02| 2.0619 | 4.2419         | 3957,690         |
| S03| 2.4027 | 4.9889         | 4564,020         |
| S04| 2.0082 | 3.9555         | 3830,696         |
| S05| 2.1258 | 4.5136         | 4023,250         |
| S06| 2.2441 | 4.7994         | 4272,268         |

Fig. 11. Reconstructions of perceived letter shapes.

Fig. 12. Change in (a) the gradient (relative to weights) and (b) Jaccard Similarity over time. The Jaccard Similarity was computed between estimated and ground-truth receptive fields. The plot was generated using simulated 3T data.

Fig. 13. Memory and computation time requirements for the real-time pRF as a function of the number of voxels. Data points corresponding to 0 s reflect < 1 Kb of memory consumed and < 0.01 seconds required for execution, respectively.

The mean computation times per volume acquisition per subject for the 3T and 7T datasets are reported in Table 6a and b, respectively. MATLAB uses a just-in-time compiler, which has to be executed the first time and has to first load the subroutine into memory and compile. This often causes the first iteration to be slower. Therefore, we exclude the execution time of the first acquired volume while computing the mean and standard deviation and report it separately. The average computation time per acquired volume is less than the repetition time (2000ms for 3T and 3000ms for 7T), which means that the receptive fields are updated before the next volume is acquired. This is especially useful in a real-time setting where the mapping occurs while the data is being acquired. Fig. 13 depicts how memory requirements scale with computation time. The computation time only starts to increase when the
required memory exceeds the available memory. Generally, up to 1 million voxels can be comfortably estimated within less than 1500 ms and requiring less than 2 GB of RAM.

3.3. Poorly estimated receptive field size

At larger eccentricities our approach shows poor correspondence with the GS-pRF algorithm in terms of receptive field size. This is surprising given the good correspondence between estimated and ground-truth receptive field sizes for simulated data. One potential reason for the discrepancy between our (model-free) and the GS-pRF approach is that the latter assumes receptive fields to have a circular shape. If receptive fields are not circular, a GS-pRF method may estimate receptive field sizes inaccurately. Elongation of receptive fields have been observed in several studies (Greene et al., 2014; Silson et al., 2018; Lee et al., 2013; Merkel et al., 2018) rendering this a viable explanation for the discrepancy. An alternative explanation, assuming receptive fields are generally circular, is that the model-based GS-pRF procedure can accurately capture sizes of receptive fields located beyond the visual field of view (the region of the visual field covered by the stimulus) whereas our model-free procedure cannot. Indeed, our model-free procedure would produce a smaller, elongated, receptive field located within the field of view if a large receptive field is located outside the field of view. Below we explore both possibilities.

3.3.1. Anisotropic model

We investigate the ability of our approach to capture elongated receptive fields by generating simulated data (similar to 2.5.1) using anisotropic Gaussians as ground-truth receptive fields:

\[
    f(x, y) = \exp\left(-\frac{(x - x_0)^2}{2\sigma_x^2} + \frac{(y - y_0)^2}{2\sigma_y^2}\right)
\]

We vary \(\sigma_y\) as a ratio of \(\sigma_x\) such that the ratio between \(\sigma_x\) and \(\sigma_y\) increases with eccentricity. We first obtain \(\sigma_y\) as described in Section 2.5.1. We then compute \(\sigma_y = \rho \sigma_x\); where \(\rho\) is \(\sigma_y\) rescaled in the range [0.5, 3]. We generate simulated 3T and 7T data with this anisotropic model with the remaining simulation parameters remaining the same as described in Section 2.5.1. We define standard deviation \(\sigma\) of such anisotropic receptive fields as the geometric mean of \(\sigma_x\) and \(\sigma_y\), that is, \(\sigma = \sqrt{\sigma_x \sigma_y}\).

Using the geometric mean ensures that the area of an ellipse with semi-minor axis \(\sigma_y\) and semi-major axis \(\sigma_x\) is the same as a circle with radius \(\sigma\).

To examine whether our approach reliably captures the shape of the receptive fields, we visually inspect them. Figs. 14 and 15 show that our approach is able to generally capture anisotropic receptive field shapes and sizes rather well. The corresponding correlation coefficients are reported in Table 7. However, as receptive fields become more elongated, our method tends to slightly underestimate their size. Interestingly, the GS-pRF method assuming isotropic receptive fields tends to somewhat overestimate receptive field sizes at large eccentricities. In conjunction, these effects can account for the discrepancy between the f-pRFridge and GS-pRF mapping procedure. In order to analyze our approach quantitatively, we compute the JS between estimated receptive fields, ground truth receptive fields and the receptive fields obtained from GS-pRF (see Table 8). Note that the GS-pRF method yields pure Gaussians containing no anomalous activations whereas our method yields anomalous activations surrounding the receptive field. Even slight anomalies get penalized in the JS thus accounting for overall better fit observed for the GS-pRF method.

3.3.2. Receptive fields beyond the field of view

Next, we examine to what extent our approach fails to effectively map the receptive fields that (partially) lie beyond the field of view. We generate simulated data with 1000 isotropic receptive fields. All receptive fields are centered at an eccentricity of \(5\sqrt{2}\). This would mean, in the case of visual field of maximum radius of 5, the receptive fields are all centered at the top right corner (edge of the stimulus space). Fig. 16 depicts how the estimated pRF size parameter is affected by an increase in the measured visual field. For comparison, we estimate the receptive field size parameter from receptive field shapes obtained with f-pRF ridge using three different methods (a) our fast parameter estimation procedure (b) Levenberg-Marquadt (Levenberg 1944) with pRF location parameters estimated using our fast procedure and (c) an exhaustive search for pRF parameters for an initial guess followed by maximum likelihood estimation (Mineault, 2021). Lastly, we use our grid search method without the slope (eccentricity vs size) assumption; i.e., exploring a fixed set of pRF sizes (we term this vanilla grid-search). It is clear from Fig. 16 that it is a limitation of our approach (f-pRFridge) that receptive fields whose center is at the border or even exceeds the visual
Table 7
Correlation coefficients between fast and GS-pRF estimated pRF parameters for simulated (a) 3T and (b) 7T data based on anisotropic ground-truth receptive fields.

| a) | X-coordinate | Y-coordinate | Standard Deviation |
|---|---|---|---|
| f-pRF$_{ridge}$ vs. ground truth | 0.9900 (95% CI [0.9895, 0.9903]) | 0.9886 (95% CI [0.9881, 0.9890]) | 0.9588 (95% CI [0.9571, 0.9604]) |
| GS-pRF vs. ground truth | 0.9950 (95% CI [0.9948, 0.9952]) | 0.9962 (95% CI [0.9960, 0.9963]) | 0.9163 (95% CI [0.9130, 0.9195]) |
| f-pRF$_{ridge}$ vs. GS-pRF | 0.9909 (95% CI [0.9905, 0.9913]) | 0.9891 (95% CI [0.9887, 0.9896]) | 0.9122 (95% CI [0.9088, 0.9156]) |

| b) | X-coordinate | Y-coordinate | Standard Deviation |
|---|---|---|---|
| f-pRF$_{ridge}$ vs. ground truth | 0.9915 (95% CI [0.9912, 0.9919]) | 0.9940 (95% CI [0.9938, 0.9942]) | 0.9615 (95% CI [0.9599, 0.9630]) |
| GS-pRF vs. ground truth | 0.9971 (95% CI [0.9970, 0.9972]) | 0.9972 (95% CI [0.9971, 0.9973]) | 0.9215 (95% CI [0.9184, 0.9245]) |
| f-pRF$_{ridge}$ vs. GS-pRF | 0.9923 (95% CI [0.9920, 0.9926]) | 0.9927 (95% CI [0.9924, 0.9930]) | 0.9115 (95% CI [0.9081, 0.9149]) |

Fig. 16. Change in the estimation of the pRF size parameter with respect to change in field of view. All the receptive fields are centered at an eccentricity of $5\sqrt{2}$. The maximum radius increases (left to right) in the range $[5, 8, 12, 15, 19]$. The receptive fields sizes (1 standard deviation) are in the range $[0.5, 20]$. The plots are generated using a) f-pRF$_{ridge}$ (b) Levenberg-Marquadt with pRF location parameters estimated using f-pRF$_{ridge}$ and (c) an exhaustive search for pRF parameters for initial guess followed by maximum likelihood estimation (Mineault, 2021). (d) vanilla grid search. Note that the discrete steps in the grid search result from the resolution of the grid. As more parameter values are explored, this curve would become increasingly smooth but estimation would take more time.

field of view are poorly mapped. It can also be seen from the figure that estimations get better as the field of view becomes larger. Hence, in order to faithfully estimate receptive fields that are centered at high eccentricities, we recommend designing the experiment with a larger field of view.

Subsequently, we also investigated if there exists a trade-off between the size of the field of view and the reliability of estimation of the receptive fields. For this purpose, we generate simulated data with 1000 isotropic receptive fields that lie at the center of field of view (eccentricity of 0). We choose eccentricity of 0 to eliminate the possibility of having border effects as described previously. The methods for estimation of pRF size parameter are the same as mentioned previously. It can be seen from Fig. 17, that initially (left most figures) the size estimations are poor for large receptive fields since they are extremely large for a relatively small field of view. As the field of view increases, the estimations get better. However, after a certain point, relatively small receptive fields get estimated poorly. This is because, in the case of relatively large field of view, the bar aperture needs to be wider to cover the field of view with the same number of steps. As a consequence, the stimulus becomes too coarse to reliably estimate smaller receptive fields. In order to account for smaller receptive fields while also measuring a large field of view, a narrow bar would need to traverse the field of view with a larger number of steps. However, this would require longer runs. Our method thus demands a trade-off between field-of-view, stimulus-size and run-length. In cases where fast estimation of pRF parameters is not essential, we thus recommend to use a grid-search method since it is less affected these effects.
4. Discussion

We propose a fast approach for receptive field mapping and pRF parameter estimation that is suitable for real-time applications. A linear voxel-to-pixel map typically is a huge data matrix which requires a large amount of memory in order to be stored and to be operated on; rendering operations slow. To reduce data by more than 90%, we encode the stimulus using tile coding and hashing. This lowers memory requirements and hence strongly reduces computational time. We evaluated our approach on simulated as well as real empirical data in terms of computational times, correlations between predicted and acquired BOLD signals, fidelity of estimated receptive field shapes and parameters and the suitability of estimated pRF shapes for projecting cortical activity back into the visual field.

We find that our approach is extremely fast at mapping receptive field shapes and estimating their parameters with computational times on the order of seconds and ~1 minute, respectively. Specifically, because our approach can successfully estimate receptive field shapes for large amounts of voxels in mere seconds, it is straightforward to identify visually responsive voxels by conducting a quick cross-validation procedure. This allows limiting parameter estimation to these voxels and thus to keep computational time low for this process as well. This also eliminates the need of using a pre-defined mask; though such a mask may be used in addition. Furthermore, cross-validation is performed in batches and we provide the option of adjusting the batch size which can further speed up parameter estimation. The only other approach, to date, that facilitates fast mapping of pRFs is based on training deep neural networks to predict pRF parameters (DeepRF; Thielen et al., 2019). Our method is fundamentally different from DeepRF and does not assume initial shape of the pRFs. In terms of computation time, our approach remains significantly faster. Average computational time required per voxel is less than a second for DeepRF. It has been reported in Thielen et al. (2019), that the computation time required per voxel is on the order 10^-2 seconds (excluding the time required for training). In the case of $f_{pRF_{ridge}}$, the computational time required per voxel for parameter estimation is on the order 10^{-3} seconds.

In terms of fidelity, we observe excellent correspondence between estimated pRFs and ground-truth pRFs both in terms of shapes and parameters for simulated data. For empirical data, we observe excellent
correspondence between pRF locations estimated from our procedure and from GS-pRF (as described 2.4). However, for pRF size, results of the two methods correspond less well. In particular, for larger eccentricities correspondence is poor. We observed that the receptive fields at high eccentricities; i.e., located at or beyond the border of the stimulated region of the visual field, tend to be anisotropic (elongated) and mostly underestimated. It has been reported in Greene et al. (2014) that pRFs located in the proximity of this border tend to be underestimated. Studies such as Silson et al. (2018), Merkel et al. (2018, 2020) further report elongation of receptive fields and Silson et al. (2018) report that the aspect ratio of receptive fields increase with eccentricity. In order to investigate how reliable our mapping procedure is when the underlying receptive fields are elongated and when the receptive fields are located near the stimulus border we conducted additional, unplanned, analyses.

First, we simulated data based on elliptical receptive fields. We observed excellent correspondence between pRF parameters from our approach and ground-truth parameters. Our algorithm estimated the size of such elliptical receptive fields better than the GS-pRF method. This means that our method is more flexible and freer in capturing the shape of the receptive fields than model-based methods. As such, our method is in principle able to capture the true shape of a receptive field. However, an analysis of how receptive field size estimates are effected by their eccentricity and the visual field of view revealed that estimates are only accurate within a certain region of the visual field of view. Our analysis revealed that a major limitation of our approach is the inability to reliably estimate receptive fields that are located at or beyond the stimulus border (lie outside the visual field of view). In order to accommodate such receptive fields, we recommend setting up the experiment with a larger field of view (Fig. 16). Our analysis also suggests that there exists a trade-off between the size of the field of view and reliable estimation of receptive fields. Having a large field of view may hamper reliable estimation of relatively small receptive fields (Fig. 17), if the stimulus needs to be enlarged to cover the expanded field of view with the same number of steps. In order to reliably estimate small receptive fields with a large field of view, we recommend increasing the number of steps in which the stimulus sweeps the space (thereby increasing the length of the experiment). Alternatively, if fast estimation of pRF parameters is not an issue, one could consider utilizing a vanilla grid-search method as it is less susceptible to these effects (Figs. 16d and 17d).

The flexibility of our method comes thus at the cost of an inability to deal with large, circular, receptive fields that lie beyond the field of view (i.e. outside the region of stimulation). This is in line with the observation that linear encoding methods (such as ridge regression) fail to reliably estimate large receptive fields (Lage-Castellanos et al., 2020); or rather those receptive fields that partially lie beyond the field of view. In the pRF validation framework established by Lerma-Usabiaga et al. (2020), the authors report that pRF size estimates significantly depend on the HRF model, and varies significantly from implementation to implementation. Similar observations have been made by Lage-Castellanos et al. (2020). However, both studies report negligible differences in location estimates. This is in line with our findings.

Nevertheless, for the purpose of projecting cortical activations back into the visual field, the true shape of receptive fields at the edges of the visual field do not matter. Indeed, as can be seen from Fig. 11, our model-free approach faithfully reconstructs the letter shapes from their associated BOLD activity. Recognizable reconstructions of these shapes were possible even though data underwent real-time preprocessing which is generally considered being of lower quality than offline preprocessing. This highlights that our method is suitable for real-time applications such as content-based BCI letter-speller systems.

In that context it is also important to highlight that the results reported here were obtained using a single set of hyperparameters (learning rate, FWHM and shrinkage factor) except for reconstruction of mental imagery where we used a higher shrinkage factor. While hyperparameters were obtained using Bayesian Optimization on the simulated dataset (refer to Appendix B), the set of hyperparameters used here produced robust results for empirical data obtained from different participants, field strengths and pre-processing procedures. We provide an example script for Bayesian Optimization as well as for generating synthetic data to allow potential users of our method to adjust hyperparameters to their specific experimental setups and needs (https://github.com/ccmasta/richt/real_time_pRF).

We also encourage users to report their stimulus, setup and the hyperparameters that they found optimal for their dataset/experiment (https://forms.gle/3XkD4D76ixeG7J5u6). This information will be compiled to offer optimized user-recommendations on the wiki of the Computational Neuroimaging Toolbox.

In conclusion, we present an extremely fast and flexible pRF mapping approach which can be either used in parallel with data acquisition (f-pRF) or after the data has been fully acquired (f-pRF\textsubscript{adj}). This opens the door for real-time applications that rely on pRF estimates such as BCI speller systems. We also propose a fast method to estimate pRF parameters. A limitation of this method, and model-free approaches in general, is that receptive fields partially lying beyond the stimulus space are dealt with poorly. This can be remedied by combining fast estimation of receptive fields for voxel selection with subsequent grid-search to obtain their pRF parameters.

Acknowledgments

This project has received funding from the European Union’s Horizon 2020 Research and Innovation Programme under Grant Agreement Numbers 945539 (HBp SGA3) and 779860 (ERC-2017-PoC).

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.neuroimage.2021.118671.

References

Ahmadi, K., Herbik, A., Wagner, M., Kanowski, M., Thieme, H., Hoffmann, M.B., 2019. Population receptive field and connectivity properties of the early visual cortex in human almondinim. Neuroimage 202, 116105.

Altman, Y. 2020. “export\_fig.” GitHub. https://github.com/altman/export\_fig/releases/tag/v3.14.

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 (5), 2704–2718.

Anderson, E.J., Tibber, M.S., Sam Schwarzkopf, D., Shergill, S.S., Fernandez-Egea, E., Rees, G., Dakin, S.C., 2017. Visual population receptive fields in people with schizophrenia have reduced inhibitory surrounds. J. Neurosci. 37 (6), 1546–1556.

Anderson, J.R.R., Skare, S., Ashburner, J., 2003. How to correct susceptibility distortions in Spin-Echo-Echo-Planar Images: application to diffusion tensor imaging. Neuroimage 20 (2), 870–888. doi:10.1016/S1053-8119(03)00326-7.

Balasubramaniam, M., Polimeni, J., Schwartz, E.L., 2002. The V1–V2–V3 complex: qua-siconfocal dipole maps in primate striate and extra-striate cortex. Neural Netw. 15 (10), 1157–1163.

Barton, B., Brewer, A.A., 2015. fMRI of the rod scotoma elucidates cortical rod pathways and implications for lesion measurements. Proc. Natl. Acad. Sci. 112 (16), 5201–5206.

Baseler, H.A., Gouws, A., Haak, K.V., Racecy, C., Crossland, M.D., Tufail, A., Rubin, G.S., Cornelissen, F.W., Morland, A.B., 2011. Large-scale remapping of visual cortex is absent in adult humans with macular degeneration. Nat. Neurosci. 14 (5), 649–655.

Binda, P., Thomas, J.M., Boynton, G.M., Fine, I., 2013. Minimizing biases in estimating the reorganization of human visual areas with BOLD retniotopic mapping. J. Vis. 13 (7), 13.

Brewer, A.A., Barton, B., 2014. Visual cortex in aging and Alzheimer’s Disease: changes in visual field maps and population receptive fields. Front. Psychol. 5, 74.

Dekker, T.M., Schwarzkopf, D.S., Haas, B., Nardini, M., Sereno, M.I., 2019. Population receptive field tuning properties of visual cortex during childhood. Dev. Cogn. Neurosci. 37, 100014.

DeYoe, E.A., Carman, G.J., Bandettini, P., Glickman, S., Wieser, J.O.N., Cox, R., Miller, D., Neitz, J., 1996. Mapping striate and extrastriate visual areas in human cerebral cortex. Proc. Natl. Acad. Sci. 93 (6), 2382–2386.

Dumoulin, S.O., Wandell, B.A., 2008. Population receptive field estimates in human visual cortex. Neuroimage 39 (2), 647–660.

Engel, S.A., Gary, H.G., Brian, A.W., 1997. Retinotopic organization in human visual cortex and the spatial precision of functional MRI. Cereb. Cortex 7 (2), 181–192 New York, NY: 1991.

Es, D.M., Theeuwes, J., Knapen, T., 2018. Spatial sampling in human visual cortex is modulated by both spatial and feature-based attention. Elite 7, e36928.
Free man, J., Simoncelli, E.P., 2011. Metamers of the ventral stream. Nat. Neurosci. 14 (9), 1195.

Georgy, L., Jans, B., Tamietto, M., Pito, A., 2019. Functional reorganization of population receptive fields in a hemispherectomy patient with blindsight. Neuropsychologia 128, 198–203.

Goebel, R., Espposito, F., Formisano, E., 2006. Analysis of Functional Image Analysis Content Test (FIAC) data with brainvoyager QX: from single-subject to cortically aligned group general linear model analysis and self-organizing group independent component analysis. Hum. Brain Mapp. 27 (5), 392–401.

Gomez, J., Natu, V., Jeska, B., Barnett, M., Grill-Spector, K., 2018. Development differentially sculpts receptive fields across aging and high-level human visual cortex. Nat. Commun. 9 (1), 1–12.

Greene, C.A., Dumoulin, S.O., Harvey, B.M., Res, D., 2014. Measurement of population receptive fields in human early visual cortex using back-projection tomography. J. Vis. 14 (1), 17.

Haak, K.V., Cornelissen, F.W., Morland, A.B., 2012. Population receptive field dynamics in human visual cortex. PLoS ONE 7 (5), e37686.

Harvey, B.M., Dumoulin, S.O., 2011. The relationship between cortical magnification factor and population receptive field size in human visual cortex: constancies in cortical architecture. J. Neurosci. 31 (38), 13604–13612.

Holmes, G., 1918. Disturbances of vision by cerebral lesions. Br. J. Ophthalmol. 2 (7), 353.

Hummer, A., Ritter, M., Woleitz, M., Ledolter, A.A., Tik, M., Dumoulin, S.O., Holder, G.E., Schmidt-Erfurth, U., Windschberger, C., 2018. Artificial scotoma estimation based on population receptive field mapping. Neuroimage 169, 342–351.

Inouye, T., 1909. Die Sehstörungen bei Schuβverletzungen der körperl cortical Sehsphäre: nach Beobachtungen an Verwundeten der letzten japanischen Kriege. Engelmann, pp. 1–114.

Kay, K.N., Weiner, K.S., Grill-Spector, K., 2015. Attention reduces spatial uncertainty in human ventral temporal cortex. Curr. Biol. 25 (5), 595–600.

Klein, B.P., Harvey, B.M., Dumoulin, S.O., 2014. Attraction of position preference by spatial attention throughout human visual cortex. Neuron 84 (1), 227–237.

Kriegeskorte, N., Goebel, R., 2001. An efficient algorithm for topologically correct segmentation of the cortical sheet in anatomical MR volumes. Neuroimage 14 (2), 329–346.

Kwong, K.K., Belliveau, J.W., Chesler, D.A., Goldberg, I.E., Weisskoff, R.M., Poscette, B.P., Kennedy, D.N., Hoppel, B.E., Cohen, M.S., Turner, R., 1992. Dynamic magnetic resonance imaging of human brain activity during primary sensory stimulation. Proc. Natl. Acad. Sci. 89 (12), 5675–5679.

Lage-Castellanos, A., Valente, G., SENden, M., De Martino, F., 2020. Investigating the reliability of population receptive field size estimates using fMRI. Front. Neurosci. 14, 825.

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.

Lerm-Usabiga, 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 (6), e1007924.

Levenberg, K., 1944. A method for the solution of certain non-linear problems in least squares. Q. Appl. Math. 2 (2), 164–168.

Marques, J.P., Kobert, T., Krueger, G., Zwaag, W., Moortele, P.F.V., Gruetter, R., 2010. Mip2rage, a self-bias-field corrected sequence for improved segmentation and T1 mapping at high field. Neuroimage 49 (2), 1271–1281.

Merckel, C., Hopf, J.-M., Schoenfeld, M.A., 2018. Spatial elongation of population receptive field profiles revealed by model-free fMRI back-projection. Hum. Brain Mapp. 39 (6), 2472–2481.

Merckel, C., Hopf, J.M., Schoenfeld, M.A., 2020. Modulating the global orientation bias of the visual system changes population receptive field elongations. Hum. Brain Mapp. 41 (7), 1765–1774.

Mineault, Patrick, 2021. Auto Gaussian & Gabor fits (https://www.mathworks.com/matlabcentral/fileexchange/31485-auto-gaussian-gabor-fits). MATLAB Central File Exchange.

Ogawa, S., Tank, D.W., Menon, R., Ellermann, J.M., Kim, S.G., Merkle, H., Ugurbil, K., 1992. Intrinsis signal changes accompanying sensory stimulation: functional brain mapping with magnetic resonance imaging. Proc. Natl. Acad. Sci. 89 (13), 5951–5955.

Papanikolaou, A., Keliris, G.A., Lee, S., Logothetis, N.K., Smirnakis, S.M., 2015. Nonlinear population receptive field changes in human area V5/M+ of healthy subjects with simulated visual field scotomas. Neuroimage 120, 176–190.

Peters, J.C., Reithler, J., Goebel, R., 2012. Modeling invariant object processing based on tight integration of simulated and empirical data in a common brain space. Front. Comput. Neurosci. 6, 12.

Polimeni, J.R., Hinds, O.P., Balasubramanian, M., van der Kouwe, A., Wald, L., Dale, A.M., Fischl, B., Schwartz, E.L., et al., 2005. Two-dimensional mathematical structure of the human visuotopic map complex in V1, V2, and V3 measured via fMRI at 3 and 7 Tesla. Journal of Vision 5, 888 8.

Roberts, D.R., Bahn, V., Ciuti, S., Boyce, M.S., Elith, J., Guillauro-Arroita, G., Hausenstein, S., et al., 2017. Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure. Ecology 40 (6), 913–929. doi:10.1111/ecog.02881.

Sack, A.T., Cohon Kadish, R., Schuermann, T., Moore, M., Walsh, V., Goebel, R., 2009. Optimizing functional accuracy of TMS in cognitive studies: a comparison of methods. J. Cogn. Neurosci. 21 (2), 207–221.

Schwarz, EL, 1980. Computational anatomical and functional architecture of striate cortex: A spatial mapping approach to perceptual coding. Vision Res. 20 (6), 645–669.

Senden, M., Emmerling, T.C., Hoof, R.V., Frost, M.A., Goebel, R., 2019. Reconstructing imagined letters from early visual cortex reveals tight topographic correspondence between visual mental imagery and perception. Brain Structure and Function 224, 1167–1183 3.

Senden, M., Reithler, J., Gielen, S., Goebel, R., 2014. Evaluating population receptive field estimation frameworks in terms of robustness and reproducibility. PLoS ONE 9 (12), e114054. doi:10.1371/journal.pone.0114054.

Sergio, 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 (5212), 889–893.

Shmuel, A., Yacoub, E., Chainmow, D., Logothetis, N.K., Ugurbil, K., 2007. Spatial-temporal point-spread function of fMRI signal in human gray matter at 7 tesla. Neuroimage 35 (2), 539–552.

Silson, E.H., Reynolds, R.C., Kravitz, D.J., Baker, C.L., 2018. Differential sampling of visual space in ventral and dorsal early visual cortex. J. Neurosci. 38 (9), 2294–2303.

Silva, M.F., Brancamp, J.W., Ferreira, S., Castelo-Branco, M., Dumoulin, S.O., Harvey, B.M., 2018. Radial asymmetries in population receptive field size and cortical magnification factor in early visual cortex. Neuroimage 167, 41–52.

Sutton, R.S., Barto, A.G., 2018. On-policy Prediction with Approximation. In: Reinforce- ment learning: An introduction. The MIT Press, pp. 217–222.

Thielen, J., Guciu, U., Guchturk, Y., Ambrogioni, L., Bosch, S.E., van Gerven, M., 2019. DeepRF: ultrafast population receptive field mapping with deep learning. bioRxiv, 752290.

Thirion, B., Duchesnay, E., Hubbard, E., Dubois, J., Poline, J.-B., Lebihan, D., Dehaene, S., 2006. Inverse retinotopy: inferring the visual content of images from brain activation patterns. Neuroimage 33 (4), 1104–1116.

Vo, V.A., Sprague, T.C., Serences, J.T., 2017. Spatial tuning shifts increase the discriminability and fideity of population codes in visual cortex. J. Neurosci. 37 (12), 3386–3401.

Welford, B.P., 1962. Note on a method for calculating corrected sums of squares and products. Technometrics 4 (3), 419–420.