Real Time Estimation of Population Receptive Fields Using Gradient Descent

Mario Senden^{1,2}

¹Department of Cognitive Neuroscience, Faculty of Psychology and Neuroscience, Maastricht University, P.O. Box 616, 6200 MD Maastricht, The Netherlands

²Maastricht Brain Imaging Centre, Faculty of Psychology and Neuroscience, Maastricht University, P.O. Box 616, 6200 MD Maastricht, The Netherlands

Correspondence to: mario.senden@maastrichtuniversity.nl

Abstract A real-time population receptive field mapping procedure based on gradient descent is proposed. Model-free receptive fields produced by the algorithm are evaluated in context of simulated data exhibiting different levels of temporally autocorrelated noise and spatial point spread. As with any model-free approach, the exact shape of receptive fields produced by the real-time algorithm depends on the stimulus. Nevertheless, estimated receptive fields show good correspondence with ground-truth receptive fields in terms of both position and size. Furthermore, fitting a parametric model to the previously obtained estimates approximates the exact shape of the true underlying receptive fields well.

Introduction

Real-time estimation of population receptive fields (pRFs; *Dumoulin and Wandell, 2008*) offers many practical advantages. It can, for instance, make population receptive fields immediately available for subsequent real-time decoding experiments reconstructing perceived (or imagined) visual stimuli (*Senden et al., 2017*). It might also allow for online optimization of the mapping procedure, for example by guiding stimulus placement to obtain ideal coverage of the visual field. Here, a real-time population receptive field (pRF) mapping procedure is proposed. The procedure employs a gradient descent on weights connecting the visual field to the cortex. The resulting weights (reflecting population receptive field estimates) can be directly used for subsequent real-time experiments. Alternatively, they can provide the basis for fitting a parametric pRF model. Model-free (raw) pRFs are evaluated in light of simulated data stemming from a model of V1 topography in the context of different levels of noise and point spreads of the blood-oxygen-level-dependent (BOLD) signal. Evaluations occur in terms of the ability of the procedure to localize pRFs, approximate their shape, and the ability of these pRFs to explain variance in model BOLD activity in response to a separate test stimulus.

Methods and Materials

Stimulus Design

The stimulus is an important contributor to the accuracy with which population receptive fields can be estimated (*Senden et al., 2014*). In the present study, a novel multi-scale bar stimulus is used. The stimulus is split into three concentric regions. The inner region spans a diameter of 7° and is traversed in twelve discrete steps by a bar ~ 0.58° in width. The intermediate region spans a diameter of 14° and is traversed in twelve steps by a bar ~ 1.16° in width. The outer region spans a diameter of 21° and is traversed in twelve steps by a bar ~ 1.75° in width. The upper row of *Figure 1* shows three representative stimulus configurations. As can be appreciated from the figure, bars in the three regions never overlap and can exhibit different orientations since they are presented in unrelated

Figure 1. A) Three representative configurations of the multiscale bar stimulus. Bars are oriented independently at the three scales and presented with independent sequences. **B)** Three representative configurations of nine honeycomb blobs. Checkerboards are shown for illustrative purposes only. All simulations and mapping procedures are based on binarized effective stimuli.

random sequences. Within each region, bars are presented in four different orientations with each orientation being presented six times. In total, there are 288 samples of stimulus presentation padded before and after by eight samples without stimulation leading to a total of 304 samples.

In addition to the multiscale moving bar stimulus used for pRF mapping, a second stimulus is used to evaluate the predictive power of estimated receptive fields. This second stimulus consists of 288 random combinations of nine blobs 1.33° in diameter arranged in a honeycomb pattern. The lower row of **Figure 1** shows three representative stimulus configurations. Both stimuli have a resolution of 150×150 pixels.

Model V1

Cortical Surface and Receptive Field Properties

A V1-like cortical sheet extending 40 mm along and approximately 36 mm orthogonal to the horizontal meridian in both hemispheres constitutes the basis of the model. 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 between cortical surface and visual field (*Balasubramanian et al., 2002; Schwartz, 1980*) with parameter values ($a = 0.7^\circ$, $\alpha = 0.9$; *Polimeni et al., 2005*). **Figure 2** shows the resulting eccentricity and polar angle maps across the cortical surface. The shape of model receptive fields is given by a 2-dimensional Gaussian

$$f(x_0, y_0, \sigma_x, \sigma_y, \theta) = \exp\left(-\frac{(x - x_0)^2 + (y - y_0)^2}{2\sigma^2}\right). \quad (1)$$

with (x_0, y_0) being the receptive field center and σ its size. Below an eccentricity of $E = 0.6^\circ$ all model vertices have a receptive field size of $\sigma = 0.125^\circ$ whereas they exhibit a linear relationship with eccentricity ($\sigma = 0.21E$) beyond this cutoff (cf. *Freeman and Simoncelli, 2011*).

Figure 2. Eccentricity (top) and polar angle (bottom) maps representing the topographic organization of the V1 model. Following radiological convention, the left hemispheric surface map is shown on the right and vice versa. Centrally presented circles provide color codes for eccentricities and polar angles.

Simulated BOLD Signal

A simulated BOLD signal (sampled at a rate of 0.5 Hz) for each vertex is obtained by first performing elementwise multiplication between the receptive field of a vertex and the effective stimulus presented per time point, summing the result, and subsequently convolving the thusly obtained timecourse with a hemodynamic response function. The hemodynamic response function of each vertex is given by a two-gamma function

$$h(t) = \left(\frac{t^{\alpha_1-1} \beta_1^{\alpha_1} e^{-\beta_1 t}}{\Gamma(\alpha_1)} - c \frac{t^{\alpha_2-1} \beta_2^{\alpha_2} e^{-\beta_2 t}}{\Gamma(\alpha_2)} \right) \quad (2)$$

with parameters α and β controlling the shape and scale, respectively, and c controlling the ratio of response to undershoot. For each vertex, the five parameters are drawn from normal distributions around their canonical values. Specifically $\alpha_1 \sim \mathcal{N}(6, 0.5)$, $\alpha_2 \sim \mathcal{N}(16, 0.5)$, $\beta_1 \sim \mathcal{N}(1.1, 0.1)$, $\beta_2 \sim \mathcal{N}(1.1, 0.1)$, and $c \sim \mathcal{N}(0.167, 0.1)$.

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 striate cortex (Shmuel et al., 2007). Second, noise following a second-order autoregressive AR(2) model is added ($w_{t-1} = 0.8$, $w_{t-2} = -0.2$). The smoothing kernel is independently applied to the clean signal and the noise before the two are combined. The proposed mapping procedure is evaluated for a range of point spreads (2.25 mm to 4 mm in 8 steps) and SNRs (0.25 to 2 in 8 steps). Furthermore, performance results are evaluated in more detail (including fitting a parametric model) for a point spread of 2.5 mm full-width at half-maximum (FWHM; Shmuel et al., 2007) and a signal-to-noise ratio (SNR = $\sigma_{\text{signal}}^2 / \sigma_{\text{noise}}^2$) of 1.25.

pRF Mapping Procedure

Gradient Descent

The receptive field of each individual vertex i is considered to be a weight row vector $\mathbf{w}_i$ of length p connecting all pixels in the visual field to the vertex. The weight vectors of all vertices can then be combined into the v -by- p weight matrix $\mathbf{W}$. At any given moment in time, the predicted BOLD activity ($\hat{\mathbf{y}}_t$) given this matrix and a column vector representing the effective stimulus $\mathbf{s}_t$ is then given by $\mathbf{y}_t = \mathbf{W}\mathbf{s}_t$ and the discrepancy between the predicted and observed (z-normalized) BOLD activity ($\mathbf{y}_t$) can be measured by

$$D_t = \frac{1}{2}(\mathbf{y}_t - \hat{\mathbf{y}}_t)^T(\mathbf{y}_t - \hat{\mathbf{y}}_t). \quad (3)$$

82 The gradient of the discrepancy with respect to the weight matrix is given by

$$\frac{\Delta D_t}{\Delta \mathbf{W}} = -(\mathbf{y}_t - \hat{\mathbf{y}}_t) \mathbf{s}_t^T \quad (4)$$

83 leading to the update of the weight matrix against this gradient according to

$$\mathbf{W} \leftarrow \mathbf{W} + \eta(\mathbf{y}_t - \hat{\mathbf{y}}_t) \mathbf{s}_t^T \quad (5)$$

84 where η is a learning rate (1×10^{-5}). The effective stimulus matrix $\mathbf{S}$ (of dimension p -by- t) is the result of convolving
85 the timecourse of each pixel with the canonical two-gamma hemodynamic response function and z-normalizing
86 the result. Weights are thus updated online with each time point constituting a new training example. Before
87 training, initial weights are drawn from a normal distribution $\mathbf{W} \sim \mathcal{N}(0, 1 \times 10^{-5})$.

88 Real Time Signal Processing

89 In order for the described gradient descent approach to be used in a real-time setting, the effective stimulus
90 needs to be convolved with a hemodynamic response in real-time. Furthermore, both the BOLD signal and the
91 effective stimulus need to be z-normalized in real-time. Real-time convolution is achieved via a running convolution
92 algorithm based on fast Fourier transformation (FFT) using an overlap-add scheme (*Wefers, 2015*). Specifically, as
93 a preparatory step, the hemodynamic response function of length N is transformed using an FFT. Subsequently
94 at each point in time t , the incoming signal of length 1 is zero-padded to length N and likewise transformed. The
95 transformed signal and hemodynamic response are then multiplied elementwise and the result transformed back
96 to the time domain by an inverse FFT. The final convolved signal is the result of adding up these overlapping partial
97 convolution results (x_{partial}):

$$x_{\{t, \dots, t+N-1\}} = x_{\{t, \dots, t+N-1\}} + x_{\text{partial}}. \quad (6)$$

98 For real-time z-normalization, both mean and variance of the signal are estimated using online algorithms:

$$\bar{x}_t = \bar{x}_{t-1} + \frac{x_t - \bar{x}_{t-1}}{t} \quad (7)$$

$$s_t^2 = \frac{t-2}{t-1} s_{t-1}^2 + \frac{(x_t - \bar{x}_{t-1})^2}{t}, t > 1 \quad (8)$$

99 to subsequently obtain the z-score

$$z_t = \frac{x_t - \bar{x}_{t-1}}{s_{t-1}}. \quad (9)$$

100 Parametric Model Fitting

101 A parametric receptive field model can be fit to the estimated weight vector $\mathbf{w}_i$ of each vertex i . Here, in accor-
102 dance with the ground-truth receptive fields, the parametric model is an isotropic Gaussian with parameter set
103 $\psi = (x_0, y_0, \sigma)$. Using the Levenberg-Marquardt algorithm (*Levenberg, 1944; Marquardt, 1963*) implemented in
104 MATLAB's (2015a, The MathWorks, Natick, MA) optimization toolbox, the parameter set ψ for each vertex is found
105 that solves the problem

$$\min_{\psi} ||f(\psi, X, Y) - \mathbf{w}_i||_2^2 \quad (10)$$

106 where f is the Gaussian function (cf. *Equation 1*) and X and Y are x- and y-coordinates of each pixel in the visual
107 field. The initial conditions ψ_0 for each vertex are the x- and y-coordinate corresponding to the peak value in its
108 weight vector and $\sigma_0 = 0.1$. To improve the fitting procedure in light of potentially noisy raw pRF estimates, each
109 weight vector $\mathbf{w}_i$ is squared and subsequently normalized with respect to its maximum value before applying the
110 Levenberg-Marquardt algorithm.

Figure 3. Pearson's correlations between predicted and model pRFs and BOLD activity as a function of point spread and signal to noise ratio. **A)** Correlations between estimated and model eccentricity. Except for SNRs below 0.75, estimated and model eccentricity maps are virtually identical regardless of spatial point spread. **B)** Circular correlations between estimated and model polar angle. Except for SNRs below 0.75, estimated and model polar angle maps agree very well ~ 0.9 . Point spread has a negligible effect. **C)** Correlations between estimated and ground-truth receptive field shapes (weights). The estimated and model pRF shapes show good agreement. Agreement improves with higher SNR but is hardly affected by point spread. **D)** Correlations between estimated and model BOLD activity obtained for the test stimulus (see **Figure 1**). As for the shapes, the estimated and model BOLD activity show good agreement. Agreement improves with higher SNR. This is expected since the model BOLD activity in response to the test stimulus become less noisy with increasing SNR. Point spread has again a negligible effect.

Results

Mapping Performance

The proposed gradient descent approach to pRF mapping is evaluated for a V1 model with circular receptive fields. The performance of the algorithm is evaluated in terms of how well it localizes receptive fields as measured by the correlation between model and estimated eccentricity as well as polar angle maps, in terms of how well it reproduces receptive fields as measured by the correlation between model and estimated weights, and finally in terms of how well the BOLD activity predicted by the estimated weights for a novel stimulus protocol correlates with model BOLD activity observed for the same stimulus. **Figure 3** shows the performance of the algorithm as a function of SNR and point spread. For all considered measures SNR affects performance whereas point spread does not, at least within the considered range.

Next, a more in depth evaluation is presented for $\text{SNR} = 1.25$ and a point spread equal to 2.5 mm. **Figure 4** shows eccentricity and polar angle maps obtained from the gradient descent approach. The agreement with the underlying maps shown in **Figure 2** is visually striking. Indeed correlations between the model and estimated maps are $r = 0.98$ and $r_{\text{circ}} = 0.91$ for eccentricity and polar angle, respectively. **Figure 5** shows three representative ground-truth (panel **A**) and estimated (panel **B**) receptive fields at different eccentricities (and hence of different sizes). Correlations between model and estimated pRFs are $r = 0.15$, $r = 0.48$, and $r = 0.68$, at the three eccentricities respectively. These results suggest that correspondence improves with receptive field size. **Figure 6** shows the correspondence between ground-truth and estimated receptive fields as a function of model pRF size. Panels **(A,C)** show the correlation between model and estimated receptive fields for raw estimates stemming directly from the gradient descent **(A)** and after fitting an isotropic Gaussian **(C)**. Panels **(B,D)** show the correlation between BOLD activity in response to the test stimulus based on ground-truth and estimated pRFs, again for raw estimates **(B)** and after fitting an isotropic Gaussian **(D)**. As can be appreciated from the figure, correlations for raw estimates

Figure 4. Eccentricity (top) and polar angle (bottom) maps resulting from the proposed gradient descent receptive field mapping procedure. Receptive field locations of each vertex in Cartesian coordinates are given by the location of the peak value in its weight vector and are transformed to polar coordinates.

Figure 5. Representative receptive fields at three eccentricities (0°, 4°, and 8°). **A)** Ground-truth isotropic Gaussian receptive fields of the V1 model. **B)** Raw receptive field estimates resulting from the gradient descent mapping procedure. **C)** Receptive fields resulting from fitting an isotropic Gaussian model to the estimates.

Figure 6. Performance as a function of ground-truth receptive field size. **A)** Correlation between model and raw estimated receptive fields for each vertex. **B)** Correlations between model BOLD activity in response to the test stimulus and predicted BOLD activity in response to the stimulus based on raw estimated receptive fields. **C)** Correlation between model and Gaussian-fit receptive field for each vertex. **D)** Correlations between model BOLD activity in response to the test stimulus and predicted BOLD activity in response to the stimulus based on Gaussian-fit receptive fields.

increase with receptive field size. This is likely due to non-zero residual weights covering a larger proportion of the visual field for small as compared to large receptive fields. Consequently, the relationship between correlation and receptive field size is reduced after fitting an isotropic Gaussian to the raw estimates. At the same time, correlations are generally larger both for receptive fields and predicted BOLD activity. Indeed, correlations between ground-truth and estimated pRFs for the three previously described receptive fields increase to $r = 0.59$, $r = 0.98$, and $r = 0.97$, respectively, after fitting a parametric model. Finally, ground-truth and estimated receptive field sizes correlate well with $r = 0.95$.

Computational Performance

For the proposed mapping procedure to be practically useful in a real-time setting it is essential that all necessary computations can be carried out at a rate that matches fMRI signal acquisition which typically ranges between 0.33 Hz (TR = 3 s) 0.67 Hz (TR = 1.5 s). The speed with which the proposed algorithm performs real-time signal analysis and updates the 22 500 pixels by 11 520 vertices weight matrix used here is estimated on two systems. The first is an HP® Z440 workstation with an Intel® Xeon® E5-1650 v3 (3.5 GHz, 15 MB cache, 6 cores) processor and 32 GB DDR4 RAM running Linux Ubuntu 14.04 as operating system. The second is a desktop computer with an Intel® Core i5-6600K v6 (3.5 GHz, 6 MB cache, 4 cores) processor and 16 GB DDR4 RAM running Microsoft® Windows 10 Professional as operating system. The algorithm has been implemented in MATLAB (2015a, The MathWorks, Natick, MA). On the first system, median and maximum code execution times are 0.51 s and 0.61 s, respectively. On the second system, median and maximum code execution times are 0.83 s and 1.08 s, respectively. Furthermore, the algorithm requires a total of 4.1 GiB of RAM, 2 GiB of which are used to store the weight matrix. These measures indicate that it is indeed feasible to use the proposed algorithm for real-time mapping of population receptive fields. Subsequent fitting of the parametric model to the weights obtained from the real-time mapping procedure takes ~ 3.7 min and ~ 2.9 min on the first and second system, respectively.

Discussion

A real-time population receptive field estimation procedure based on gradient descent has been developed and evaluated in light of circularly symmetric ground-truth receptive fields in a model of V1 topography. The algorithm showed good performance in estimating receptive field location and size, robustness in the face of spatio-temporal distortions, and good computational performance. Fitting a parametric model to the raw estimates obtained from the algorithm further improves receptive field estimates and allows for accurate prediction of BOLD activity in response to new stimuli. While this fitting cannot be achieved in real-time, it is sufficiently fast to be computed during a short break for the subject (or potentially during a resting state or anatomical measurement) and to subsequently be used for any real-time applications.

It is important to note that the presented approach is merely a proof of concept and as such rather simple. However, it is vastly extendible. For instance, while in the present study the gradient descent approach is used to adjust weights linking the visual field directly to the cortex, it is possible to interpose complex filtering steps (cf. *St-Yves and Naselaris, 2017*). That is, a fixed set of spatial filters could be continuously applied to the effective stimulus with adjustable weights linking this filter set to the cortex. In this manner, more complex response profiles of voxels could be estimated in real time.

Acknowledgments

The author would like to thank Koen Frolichs for kindly providing the honeycomb test stimulus. The author would further like to thank Carmine Gnolo for his helpful comments.

References

- Balasubramanian M**, Polimeni J, Schwartz EL. The V1–V2–V3 complex: quasiconformal dipole maps in primate striate and extra-striate cortex. *Neural Networks*. 2002 apr; 15(10):1157–1163.
- Dumoulin SO**, Wandell BAA. Population receptive field estimates in human visual cortex. *Neuroimage*. 2008; 39(2):647–660.
- Freeman J**, Simoncelli EP. Metamers of the ventral stream. *Nature Neuroscience*. 2011 may; 14(9):1195–1201.
- Levenberg K**. A Method for the Solution of Certain Non – Linear Problems in Least Squares. *Quarterly of Applied Mathematics*. 1944; 2:164–168.
- Marquardt DW**. An Algorithm for Least-Squares Estimation of Nonlinear Parameters. *Journal of the Society for Industrial and Applied Mathematics*. 1963 jun; 11(2):431–441.
- Polimeni JR**, Hinds OP, Balasubramanian M, van der Kouwe A, Wald LL, Dale AM, Fischl B, Schwartz EL, Others. The human V1–V2–V3 visuotopic map complex measured via fMRI at 3 and 7 Tesla. In: *Society for Neuroscience Abstracts*; 2005. .
- Schwartz EL**. Computational anatomy and functional architecture of striate cortex: a spatial mapping approach to perceptual coding. *Vision research*. 1980 jan; 20(8):645–669.
- Senden M**, Reithler J, Gijzen S, Goebel R. Evaluating Population Receptive Field Estimation Frameworks in Terms of Robustness and Reproducibility. *PLoS ONE*. 2014 dec; 9(12):e114054.
- Senden S**, Emmerling T, Frost M, Goebel R. Faithful reconstruction of imagined letters from 7T fMRI measures in early visual cortex. In: *OHBM 2017* Vancouver, Canada; 2017. https://files.aievolution.com/hbm1701/abstracts/37566/1536_Senden.pdf.
- Shmuel A**, Yacoub E, Chaimow D, Logothetis NK, Ugurbil K. Spatio-temporal point-spread function of fMRI signal in human gray matter at 7 Tesla. *NeuroImage*. 2007 apr; 35(2):539–552.
- St-Yves G**, Naselaris T. The feature-weighted receptive field: an interpretable encoding model for complex feature spaces. *bioRxiv*. 2017 apr; p. 126318. doi: 10.1101/126318.
- Wefers F**. Partitioned convolution algorithms for real-time auralization, vol. 20; 2015. <http://publications.rwth-aachen.de/record/466561>.