Supplementary Materials

Inverso et al.:

Amp-lifying Evoked Potentials**:** Decomposing V1 and V2 sources using retinotopy constrained EEG source reconstruction without fMRI

1. Supplementary Methods

1.1 Integrated modeling method

1.1.1 GLM in the current study

James's (2003) multifocal analysis used a model that made minimal assumptions about the form of kernels in the model. This avoided the possible bias involved in restricting the form of system kernels, but at the cost of high dimensionality of the model, with potentially large variance for each of the huge number of parameters estimated. With visually evoked potentials (VEPs) the number of parameters in the model becomes very large and problems with high variance are incurred. In particular, a combinatorial explosion in numbers of parameters occurs when exploring interactions between inputs.

This method was then elaborated to fit the waveforms as linear combinations of a small set of basis functions, with the basis functions also being estimated for the data set being fitted. This can hence be termed an empirical basis function approach. It can also be termed a bilinear model; that is, with the fitted values being linear in relation to a set of coefficients when the basis functions are held fixed, and also being linear in relation to the parameters defining the basis functions when the coefficients are held fixed.

A further way of considering this approach is that the system model corresponds to the decomposition that is obtained by doing a singular value decomposition of a set of system kernels considered as functions on time and space, but the fitting is done directly to the recorded compound response signals over time, rather than to the derived entity being the set of estimated system kernels (Goh, 2008).

1.1.2 GLM Formulation

The General Linear Model (GLM) is formulated as:

$$y = X\beta + e \qquad 1$$

y is the observed response (dependent variable), X are the explanatory variables (independent variables), β are the unknown regression coefficients, and e contains the errors or noise.

At a high-level we are attempting to determine the brain signal related to the stimulus. Ideally the true signal is the response recorded, however, a major source of noise is other cortical activity that is not related to the stimulus. For example, a fatigued participant may have alpha bursts that perturb the VEP to a stimulus. Evoked potential analysis requires that these other cortical activities are not synchronized relative to the stimulus, so that averaging over repeats, or in our case regressing on the stimulus signal, leads to the stimulus related activity dominating. Muscle artifact, 50 or 60Hz electrical mains frequency, and capacitance changes from participant movement, are additional noise components that can affect the recording signal.

The responses recorded can be modeled as:

$$y = \hat{y} + e \qquad 2$$

Here y is the signal recorded, $\hat{y}$ is the expected value of response, e is the noise in the system and is assumed to have expected value zero (average value in limit over many identical replicates), and have stationary second-order statistics, i.e. potentially with covariance between two time points, but dependent only on the interval between the two times.

Because we record over a length of time (e.g. four minutes), and multiple channels (e.g. 64), y is a $n \times m$ matrix, representing time points by channels, having elements $y(i,j)$ where i indexes time points($i = 1 \dots n_{maxTime}$), and j indexes channels ($j = 1 \dots m_{maxChannels}$).

In multifocal VEPs we seek to find the response associated to the stimuli that appeared on screen. In this case, the pattern pulse presentations within regions of a dartboard. If we index the 84 stimulus regions by $r = 1 \dots nreg$, then the stimulus sequences are represented by $x(i,r)$, where $i = 1 \dots n_{maxTime}$ and $r = 1 \dots nreg$. In the pattern pulse stimuli used, x is zero when the stimuli is absent and one at the start time of a stimulus pulse presentation (a pulse may last for multiple frames, however, the fitted waveform is the response to the whole pulse, not the response 'per frame').

1.1.3 Kernel formulation as a linear model

The following demonstrates how the usual system kernel formulation is expressed as a linear model within the regression framework; although, in this study it is elaborated to the basis function approach, defined subsequently, using a model in terms of kernels convolved with the stimulus signals to give the fitted response yields:

$$\hat{y}(i,j) = \sum_{r=1}^{nregion} \sum_{k=lagmin}^{lagmax} x(i-k,r) \times g(k,r,j) \qquad 3$$

k indexes over a time-window of integration from $lagmin$ to $lagmax$, e.g. 0–300ms for a VEP response to pattern pulse presentation.

The design matrix, X, then has columns for each combination (k,r), and the kernel values g are catenated in the coefficient matrix β. The design matrix X and coefficients β are augmented with blocks to fit nuisance parameters; in this study a constant and linear trend, and sine and cosine terms for the electrical mains frequency and its harmonics. When a recording session is broken into discrete segments (e.g. four 1 minute recordings with breaks in between), the nuisance components have separate coefficients for each continuous segment.

1.1.3.1 Basis Function and Bilinear Model

The kernel model has few assumptions, however, it is overparameterized, as it allows completely free values at each lag, region, and channel combination, incurring relatively large variance for each parameter. The basis function method presented in Goh (2008) reduced the dimensionality of the model using assumptions that are reasonable given the known neuroelectrophysiology of VEPs. A family of evoked potential waveforms, including the VEP, can be approximated reasonably well by linear combinations of two component waveforms (Bair et al., 2003; Dandekar et al., 2007; Maier et al., 1987; Zhang and Hood, 2004). This is possible because most of the waveform's response power is generated by two cortical sources, which combine in different proportions across the electrode array, as evidenced by Principal Component Analysis (PCA) and singular value decomposition (SVD) of raw response waveforms (James, 2003; Lesevre and Joseph, 1979; Maier et al., 1987). With the simplification that the response is generated by two cortical sources the kernel can be constrained to be a

bilinear combination of basis functions. A basis function model is fitted directly to the MFVEP overlapping recorded signal. Goh (2008) terms this a bilinear model because it is a small number of basis functions over time that form a bilinear combination that are weighted by separate coefficients for each region-channel combination:

$$g(k,r,j) = \sum_{b=1}^{nbf} B(k,b) \times C(b,r,j) \qquad 4$$

Where B is an array of basis functions, for lag $k = lagmin..lagmax$, indexed by $b = 1..nbf$. Three basis functions are used, V1, V2 and a third to avoid bias, which provides an extra dimension relative to the typical dimensionality PCA produces. The bilinear form for the kernels thus has the same form as a singular value decomposition (SVD) of the set of kernel waveforms, however, the basis functions and coefficients are fitted directly to the signal data, rather than to a set of estimated kernels. As initial values, the three basis functions are taken as a gamma function and its first and second time-derivatives. The fitting procedure then optimizes the basis functions along with the waveform coefficients for that particular data recording. Because the basis functions are estimated from the data, the method is an empirical basis function technique as compared to SPM, which uses a fixed hemodynamic response function.

In GLM form$(y = X\beta + e)$, the fitted values are modeled as:

$$\hat{y}(i,j) = \sum_{r=1}^{nreg} \sum_{b=1}^{nbf} \left[\sum_{k=lagmin}^{lagmax} x(i-k,r) \times B(k,b) \right] \times C(b,r,j) \qquad 5$$

For estimation of the coefficients $C(b,r,j)$, the design matrix X has as columns the stimulus sequences convolved with the basis functions, for each combination of r and b. By holding the basis functions fixed, the coefficients are estimated by Weighted Least Squares (WLS) regression, with a great reduction in dimensionality compared to the unconstrained kernel model.

1.1.3.2 Noise

The noise in our EEG VEP recordings is largely from uncontrolled physiological activity, instead of electronic noise mainly because the preamplification and driven- right-leg techniques the BIOSEMI EEG amplifier employs greatly reduces and often eliminates external noise. Unfortunately, the physiological noise has significant low-frequency power

that largely overlaps the frequency band of interest. As in (Goh, 2008), a noise specification is employed that assumes the noise signal e is a low-pass filtered white noise signal. The noise signal is then modeled as a recursive digital filter's output, that is an Infinite Impulse Response (IIR) filter, with a Gaussian white noise input sequence (an independent identically distributed (iid) sequence of Gaussian variates). This sequence is termed the innovations or prediction error sequence.

The low pass filter mapping white-noise innovations to recorded noise is modeled as an all-pole filter, so that the inverse filter mapping noise to innovations is an all-zero filter, i.e. a Finite Impulse Response (FIR) filter. The innovation sequences $z(i,j)$ are then given by:

$$z(i,j) = \sum_{k=0}^{nd} d(k) \times e(i-k,j) \qquad 6$$

$d(0)$ is fixed at 1 and the coefficients $d(1) \ldots d(nd)$ are estimated.

1.1.3.3 Estimation Procedure

The estimation algorithm minimizes residual sum of squares by iterating between two stages (Goh, 2008). The first stage fixes the basis functions and pre-whitening coefficients, and uses WLS to estimate the coefficients $C(b,r,j)$. Regression design size is manageable because each channel is estimated in sequence, this is possible because the maximum likelihood estimates for a multivariate regression with common regressors and weighting is equivalent to a series of univariate regressions.

The second stage fixes the coefficients C and updates the basis functions and pre-whitening coefficients. The update performs a Gauss-Newton step projecting onto the tangent plane giving a local linear approximation to a curved model surface. Channels are pooled with weighting inverse to the estimated channel's noise variance.

1.1.4 Source Current Estimation

Source current estimation attempts to map the voltages recorded at the scalp with EEG to the currents at the neural sources.

1.1.4.1 V2 Ventral Dorsal Split

It is important to note that while the visual field is mapped continuously across V1, V2 is split ventrally and dorsally. The split causes an inversion of waveform polarity as

different banks of the calcarine sulcus are activated; and varies between individuals and across eccentricity. This inversion line is on average 15° below the horizon, and ranges from 30° below the horizon to just above the horizon (Clark et al., 1995). It is therefore possible that stimuli as sampled here in the visual field span across the horizontal meridian and are represented partially on both V2 dorsal and V2 ventral.

Few researchers have noted the partial V2 dorsal ventral split (Aine et al., 1996; Clark et al., 1995). Di Russo et al. (2001; 2005) did account for it by positioning stimulus at polar angles of 25° above and 45° below the horizontal meridian to stimulate approximately the opposite banks of the calcarine sulcus without spanning the meridian.

Goh (2008) accounted for it in dipole source localization in the same way done here, by mapping the horizontal meridian into both V2 dorsal and ventral. Figure 1E depicts the mapping of dartboard regions, Figure 1B, onto a right hemisphere flat map. Notice regions 6, 18, 30, 42, 54, 66 & 78 and 7, 19, 31, 43, 55, 67, & 79 are repeated in both V2 Dorsal and V2 Ventral. These regions correspond to 15° above and 15° below the center of the dartboard, which should encompass the horizon for most participants. Three cases are possible, 1. the horizon is entirely on V2 Dorsal, in this case the V2 ventral map will be insignificant, 2. the horizon is entirely on V2 ventral, in which case the V2 dorsal map will be insignificant, and 3. the horizon zigzags regions and is partially on both. Given anatomical findings case 3 is expected.

1.1.5 Integrating Patches: P and N

1.1.5.1 Equivalent Normal Vector

Figure 1E shows the dipole source patch map. Each patch is represented by an equivalent dipole. The position for dipole summarizing source s is denoted by the variable $P(s)$, representing X, Y, Z coordinates in mm. The intention is to fit common activation waveforms across multiple patches to account for area and folding each patch is represented by an equivalent normal vector. The vector is determined by the mean orthogonal direction to the cortical surface integrated over a given cortical patch. The equivalent normal vector's length is in mm^2 and represents the effective dipole strength accounting for any cancellations from cortical folding. The normal vector length can be thought of as the area of a flat plane patch that has the equivalent dipole effect. The

equivalent normal vector for source s is denoted by $M(s)$, and is a vector with dimensions for 3D space, XYZ, and units of mm^2.

To fit the measured signal of a given EEG channel, the equivalent normal vector is dual-paired with a gain covector for the source location and channel, in μV/nAm, estimated by the forward modeling algorithm. In our case, the gain is calculated on a 3mm volume grid that covers the areas of interest, and is evaluated with tricubic interpolation for each of the fitted dipole positions, $P(s)$.

This gives the predicted scalar gain, $\mathrm{G}_s(s,\mathrm{j})$, in units of μV/(nAm/mm^2), predicting response in μV in channel j, for a given current dipole density (in nAm/mm^2) of activation over the patch comprising source s

$$G_s(s,j) = \langle G\big(j, P(s)\big), M(s)\rangle \qquad 7$$

The angle brackets represent sum of products of vector and gain covector over the three dimensions of space.

1.1.5.2 Integrated Modeling: Calculating Normal Vectors and Dipole Positions

Once the equivalent normal vectors and dipole positions are determined estimating the waveform coefficients is a relatively easier linear regression problem. Determining the normal vectors and positions are the more difficult tasks. The starting point in the current study follows the method of Goh (2008) where fMRI is used to constrain the dipole positions, and a 2D flat map is used to determine a grid on the cortical surface.

Surface reconstruction software and manipulation programs, such as BrainSuite, FreeSurfer, and CARET, can produce 2D flat maps of the 3D surface. BrainSuite produces these maps in a coordinate system they term U and V, which we follow here. The estimates of the cortical patches for each source are parameterized by the patch corners in UV coordinates separately for each hemisphere. The dipoles are modeled with 14 rows over polar angle in V1 and V2, and the inner 5 rings of the 7 ring stimulus (Figure 1E). This gives 360 parameters, 90 corners $(14+1)\times(5+1)$, a U and V per corner, and 2 hemispheres: $90_{corners} \times\ 2_{UV} \times 2_{hemispheres} = 360$.

The equivalent normal vector is calculated by numerical integration over the corresponding cortical surface patch in 3-dimensional XYZ space. The benefits are that it accounts for cancellation from cortical folding and possibly differing patch areas across

the stimulus' rings. The approach of integrating over patches contrasts with Ales et al. (2007; Ales et al., 2010) who used a single normal vector that was allowed to shift position within bounds defined by an fMRI retinotopy to optimize the fit to the observed response, and Hagler et al. (2009), who used a single normal vector that was fixed in its fMRI defined position.

The dipole sources are represented on an HSV colormap Figure 1E. The dartboard sectors are coded by hue, ring by value, and V1, V2 dorsal, and V2 ventral by saturation.

1.1.5.3 Calculating normal vectors given patch corners

One difficulty with calculating the dipoles on a patch is the 2D flat maps in UV space are triangulations vastly complicating the computation of patch integrals. To simplify calculations triangulations are converted to a grid system, which is easier to integrate over. The triangulation in UV space is converted to a fine grid also in UV space. Now mapping from UV to XYZ space (3D surface) can be rapidly evaluated at arbitrary points with bilinear interpolation, and the corresponding partial derivatives are also immediately available.

We are interested in the area of the UV map that corresponds to V1 and V2, the flat map itself can represent the entire cortical surface or a partial area of the cortical surface, e.g. occipital cortex, depending on how the experimenter constructs it. To assist in defining the V1 and V2 areas, a second coordinate system termed AB is created to represent V1, V2 dorsal, and V2 ventral.

The benefit of the elaborate mapping scheme of AB → UV → XYZ is it can be rapidly evaluated on a square grid of AB points, and each patch's equivalent dipole can be quickly estimated from the summation over the square of AB grid points within each patch.

The AB coordinate space is a continuous coordinate system corresponding to the row and ring respectively of the dartboard stimulus (14 rows and 5 rings). The A coordinate is identical to the polar angle of the visual hemifield represented within V1 area, and thus ranges from -90 to +90 degrees. V2 cortical patches are simply extrapolated from this range, thus dorsal V2 corresponds to -90 to -180 degrees, however, it is extended to -210 to allow for the split patch representation over the horizontal meridian. V2 ventral is thus +90 to +210.

Following the stimulus scheme representation the B coordinate corresponds to the ring's eccentricities. The ring centers are represented by integral values 1, 2, 3, 4, and 5, making B range from 0.5 to 5.5 for the first five rings of the dartboard. The B coordinate is therefore a monotonic mapping of visual field eccentricity, and the scale corresponds to the rings' spacing. The implication is the B coordinate is approximately linear with distance on the cortical sheet—this linearity is as accurate as the ring's spacing is faithful to the variation in cortical magnification with eccentricity.

Because AB directly represent V1 and V2 the mapping of AB to UV will be adjusted to optimize the fit of the observed evoked waveforms. This approach allows the adjusting of the mapping of visual space to V1 and V2 quadrants. It assumes the 2D and 3D surfaces accurately map the cortical sheet (UV and XYZ), and the forward model's gain matrix is accurate. The AB to UV mapping performed by Goh (2008) were set to initial estimates from MultiFocal fMRI (MFfMRI) retinotopy, and then adjusted as needed because initial retinotopy mapping can be partially incomplete and require manual adjustment before use as a constraint (Goh, 2008). The current research developed a method that removes the fMRI requirement and uses an interactive mapping approach.

The mapping from AB to UV is parameterized by the corners of the dipole source patches, i.e. each square patch in the grid has four corners, and each square patch represents a dipole source. The continuous AB–UV mapping is approximated on a fine grid of AB locations with bicubic spline interpolation.

1.1.5.4 Area Cost

The non-linear regression procedure used to optimize the AB to UV mapping incorporates a cost on patch area. The area cost stabilizes the model fitting with regard to human anatomical visual cortical patch areas. A prior distribution of cortical patch areas was used from Schira et al.'s (2007) formula for areal magnification integrated over the dartboard stimulus' regions' eccentricity ranges. The area cost is the deviation of the patch areas in the mapping from this prior.

2. Supplemental Figures

Supp. Fig. 1. Schematic workflow from data acquisition to dipole model generation and waveform decomposition.

Supp. Fig. 2. Initial retinotopic-layout (RL) with the custom MATLAB software to produce a Retinotopy Constrained Source Estimation and dipole model. The user roughly positions the RL centering V1 in the calcarine sulcus and the unconstrained minimization optimization routine finishes positioning the map. Upper left: RL Map on 2D flat map with cortical annotations from FreeSurfer shaded with curvature to show depth. Annotations of note: light blue: calcarine sulcus, red cuneus, tan: lingual, maroon: occipital pole (annotations imported from Freesurfer aparc.a2009s.annot atlas). Upper right: current fitted V1 V2 decomposition (V1 red V2 green). Lower left: retinotopy on 3D surfaces. Lower right: costs of the current RL Map (clicking an entry “undoes” to that RL map). Bottom: gray bar reports the software’s current mode to the user. Supp. Fig. 3 shows the final after optimization.

Supp. Fig. 3. Final retinotopic layout and dipole model after fminunc optimization.

Supp. Fig. 4. Histogram of dipole 3D vector difference (‘angle error’) across all patches between EEG-RL and fMRI-RL broken out by subject, V1, V2, V2d, V2v. Rows from V2 +-15° are excluded as they are deweighted in the fitting and thus less robustly fit. Subject s001 overall has a smaller vector difference than s002, this also evident in Figure 2 as s001 has a tighter waveform correlation than s002.

Supp. Fig. 5. Cortical Curvature versus Angle Error between EEG-RL dipoles and fMRI-RL dipoles for combined s001 and s002 both hemispheres. Curvature is the amount of the curve’s curvature towards the surface normal measured as a ‘rate of change’ mm^{-1}. A small inverse correlation is evident, with a decrease in error to an increase in cortical curvature. Flat areas across a sulcus can cancel more readily than dipoles on a curve,

however, more data and further investigation is needed. **A** V1 and V2 R=-0.23 n=200, **B**, V1 R=-0.22 n=120, **C** V2 R=-0.26 n=80, **D** V2v R=0.04 n=40 **E** V2d R=-0.17 n=40.

Supp. Table 1. s001 Left Hemisphere's EEG-RL vs fMRI-RL vector differences by rows and rings.

Supp. Table 2. s001 Right Hemisphere's EEG-RL vs fMRI-RL vector differences by rows and rings.

Supp. Table 3. s002 Left Hemisphere's EEG-RL vs fMRI-RL vector differences by rows and rings.

Supp. Table 4. s002 Right Hemisphere's EEG-RL vs fMRI-RL vector differences by rows and rings.