Quantifying spatial uncertainty of visual area boundaries in neuroimaging data

Dean Kirson
Section of Neurobiology, Center for Perceptual Systems, Imaging Research Center, and Institute for Neuroscience, The University of Texas at Austin, Austin, TX, USA

Alexander C. Huk
Section of Neurobiology, Department of Psychology, Center for Perceptual Systems, Imaging Research Center, and Institute for Neuroscience, The University of Texas at Austin, Austin, TX, USA

Lawrence K. Cormack
Department of Psychology, Center for Perceptual Systems, Imaging Research Center, and Institute for Neuroscience, The University of Texas at Austin, Austin, TX, USA

Functional magnetic resonance imaging (fMRI) of the human brain has provided much information about visual cortex. These insights hinge on researchers’ ability to identify cortical areas based on stimulus selectivity and retinotopic mapping. However, border identification around regions of interest or between retinotopic maps is often performed without characterizing the degree of certainty associated with the location of these key features; ideally, assertions about the location of boundaries would have an associated spatial confidence interval. We describe an approach that allows researchers to transform estimates of error in the intensive dimension (i.e., activation of voxels) to the spatial dimension (i.e., the location of features evident in patterns across voxels). We implement the approach by bootstrapping, with applications to: (1) the location of human MT+ and (2) the location of the V1/V2 boundary. The transformation of intensive to spatial error furnishes graphical, intuitive characterizations of spatial uncertainty akin to error bars on the borders of visual areas, instead of the conventional practice of computing and thresholding p-values for voxels. This approach provides a general, unbiased arena for evaluating: (1) competing conceptions of visual area organization; (2) analysis technique efficacy; and (3) data quality.

Keywords: fMRI, retinotopy, statistics, bootstrapping


Introduction

Functional magnetic resonance imaging (fMRI) has provided a new means for investigating the functional properties of the human brain. Perhaps the greatest success of this relatively new technology has been the identification of visual cortical areas by means of stimulus selectivity and retinotopic mapping. Stimulus selective areas such as the human MT-complex (MT+) are routinely identified as regions of interest for further study (Tootell et al., 1995). Likewise, representations of the visual field corresponding to visual areas such as V1, V2, and V3 are often identified using standard retinotopic mapping techniques (DeYoe, Bandettini, Neitz, Miller, & Winans, 1994; Engel, Glover, & Wandell, 1997; Engel et al., 1994; Sereno et al., 1995). However, ongoing research and debate surrounds the presence and form of additional functionally specific and retinotopically organized visual areas in dorsal, ventral, and lateral regions of extrastriate cortex (Brewer, Liu, Wade, & Wandell, 2005; Hadjikhani, Liu, Dale, Cavanagh, & Tootell, 1998; Hagler, Riecke, & Sereno, 2007; Hagler & Sereno, 2006; Hansen, Kay, & Gallant, 2007; Larsson & Heeger, 2006; Schluppeck, Glimcher, & Heeger, 2005; Sereno & Huang, 2006; Sereno, Pitzalis, & Martinez, 2001; Smith, Greenlee, Singh, Kraemer, & Henning, 1998; Swisher, Halko, Merabet, McMains, & Somers, 2007; Tootell & Hadjikhani, 2001; Wade, Brewer, Rieger, & Wandell, 2002).

Despite the acceptance and prevalence of using fMRI to identify and study the function of human visual areas, methods of analyzing fMRI data remain a topic of active development. The vast majority of the resulting techniques are ultimately concerned with testing the statistical significance of voxel activation, i.e., expressing the uncertainty along the intensive dimension of measurement. While this makes perfect sense in some contexts, in the case of identifying visual areas the primary concern is the location of some feature of the data, rather than its intensity per se (Nickerson, Martin, Lancaster, Gao, &
Fox, 2001). It is often important to be able to confidently assert not just whether or not a particular voxel’s response passes some statistical threshold, but whether a spatial pattern of responses occurs in a particular location or exhibits some systematic structure. An appreciation of spatial uncertainty would strengthen such claims about the location of functional specialization and/or visual field organization. It thus seems desirable if not imperative to adopt a method of quantifying and displaying the spatial uncertainty associated with spatial features in fMRI data.

The importance of estimating spatial uncertainty becomes clearer in the context of specific examples. Conventional analysis of visual mapping data focuses on the pattern of intensity of each voxel over time in response to a visual stimulus. In the case of an experiment designed to localize a functionally specific visual area, this amounts to testing whether each voxel in the expected anatomical location exhibits a correlation with the alternation between two stimulus conditions. Voxels whose response falls above some statistical threshold are identified as composing the visual area. Likewise, in the case of phase-encoded retinotopic mapping, analysis focuses on estimating the timing (phase) of each voxel’s response in relation to the periodic visual stimulus that sweeps across the visual field. Voxels that fall at a “phase reversal” are interpreted as lying at the intersection of two abutting maps and hence are indicated as the border between retinotopic areas.

Despite the seeming rigor of the visual neuroscience region-of-interest approach, the identification of these regions is often rather qualitative. Functionally specific responses are identified on the basis of approximate expected anatomical location, combined with the application of an arbitrarily set statistical threshold. Retinotopic maps are often ambiguous, even to expert eyes, and much debate has surrounded the interpretation of qualitative patterns. In both of these contexts, one should be tempted to ask “Where are the error bars?” The spatial uncertainty associated with the location of a visual area’s borders is not directly provided by whether a particular voxel’s pattern of intensity places it above or below a statistical threshold, or exhibits a particular phase of response relative to the stimulus. To characterize and quantify the error of interest—the spatial uncertainty—one must transform error in the dimension of voxel intensity to error in the spatial dimension.

In this paper, we further explain the importance of transforming intensive error to spatial error and propose a method for doing so based upon the increasingly popular statistical technique of bootstrapping (Efron & Tibshirani, 1993). Although other approaches could be used to perform this transformation, the bootstrap does so without relying on or enforcing various statistical assumptions and can be utilized to analyze existing data in a straightforward and intuitive manner across a variety of specific applications. The bootstrap can be used in conjunction with any approach to the processing and analysis of imaging data and thus also provides an unbiased arena to compare the outputs of various analysis schemes.

In the following sections, we begin by briefly describing the process of transforming intensive to spatial error in the context of simulated one-dimensional data. We then describe this spatial uncertainty framework in greater detail, in the context of a real example: the identification of human area MT+ based on its motion responsivity. This example further explains the spatial uncertainty approach and the implementation with real data. This exercise yields a quantitative characterization of the spatial uncertainty associated with the localization of MT+. We then apply the approach to the identification of the V1/V2 border from standard phase-encoded retinotopy experiments, as well as the retinotopic representation of visual field eccentricity. This application demonstrates how the estimate of spatial uncertainty from intensive error can be generalized to any particular type of experiment or method of data analysis, introducing a way for quantifying spatial uncertainty associated with various features of retinotopic maps.

Here we introduce the spatial uncertainty framework in the context of a simplified one-dimensional example. Consider a 1-D row of voxels, which can be thought of as a row taken from a 2-D map that resulted from the analysis of fMRI data. Along this 1-D row, we have simulated a bell-shaped blob of activity perturbed by noise. Figure 1A shows the intensity of the fMRI response as a function of space; thin colored segments show ten different simulated runs and the thick black curve shows the mean of those runs. Throughout the figure panels, the dashed curve shows the expectation of the Gaussian profile used to generate this simulated activation.

Figure 1B shows the effects of applying a threshold along the intensive dimension. Voxels whose mean response surpassed the intensity threshold are still shown, and voxels that fell below threshold are set to zero. This is akin to the standard “statistical parametric map” representation typically used to describe the results of an fMRI experiment. In this representation, suprathreshold voxels are pseudo-colored to represent locations of significant activity, and subthreshold voxels are rendered transparent, revealing the corresponding anatomy below.

The critical point here is simply that the conventional representation shown in Figure 1B does not indicate the amount of uncertainty in space associated with the location of the activated region. In fact, the intensity-based threshold only serves to exaggerate the crispness of the location of significant activity, as it transforms a bell-shaped blob of
activity (dashed curve) to a sharp-edged mesa of high intensity (solid curve). Unfortunately, the error bars (shown in Figure 1B for a few sample voxels) are of little help in estimating spatial characteristics of the activation, because they exist only within the intensive dimension (i.e., they are oriented vertically).

In many applications, it would be preferable to estimate the uncertainty in space, not in intensity, of the activated region. To do so requires a transformation from the intensive dimension ($y$-axis in Figure 1) to the spatial dimension ($x$-axis in Figure 1). A simple way to implement this transformation is the application of a resampling technique such as the bootstrap. We describe the details of one particular bootstrapping implementation in later sections where we consider real fMRI data. For the purposes of this simulated example, it is sufficient to: (a) resample from the original ten runs with replacement and calculate the average response; (b) apply the same intensity threshold as applied in Figure 1B; and (c) repeat the resample and threshold steps ($a$ and $b$) multiple times, recording the resulting locations of suprathreshold voxels identified on each iteration.

Figure 1C shows the results of several such resampled iterations. Note that the location of the suprathreshold region of “significant activity” exhibits some variability in space, i.e., along the $x$-axis. One can thus characterize this variability in the spatial dimension, which amounts to quantifying spatial uncertainty. Figure 1D shows the result: the mean activated region, along with horizontally oriented error bars that indicate the spatial uncertainty (95% confidence intervals) associated with the location of suprathreshold activity.

This exercise illustrates the basic notion of transforming intensive error into spatial uncertainty. This transformation provides information that is not explicit in simple post-threshold statistical maps and which in fact is often obscured or distorted by thresholding in the intensive dimension. It therefore becomes important to further consider the relation between intensive error and spatial error. Indeed, if spatial error is generally a simple function...
of intensive error, there would be little need in practice to bother with actually doing the transformation. On the other hand, if the relationship between intensive and spatial error is not always straightforward, this estimation process becomes informative (and arguably imperative) in situations where spatial error is the primary concern.

We therefore examined the relationship between intensive error and spatial error by extending the simple 1-D example described above. Figure 2 shows the results of this analysis, in which we varied the degree of the intensive error and measured the resulting spatial error. Intensive error was manipulated by injecting a variable amount of Gaussian-distributed noise to the voxel intensities. Spatial error was calculated by performing the same transformation described in Figure 1 and recording the resulting 68% confidence interval (i.e., the standard deviation) of the location of the edge of the activated region. We repeated this analysis for three different levels of intensive threshold (low, middle, and high).

The results demonstrate that, even in a simple situation, spatial error need not be a simple or intuitive function of intensive error. As shown in Figure 2, knowing the degree of error in the dimension of voxel intensity (e.g., the $p$-values associated with particular voxels) does not allow one to directly infer the degree of error in the spatial domain (e.g., the precision of the localized activation) simply because the functions are nonlinear. Moreover, note that the form of this nonlinear dependence varies greatly with different thresholds, further emphasizing that transforming intensive error to spatial error provides novel information not easily recovered from visualization of thresholded statistical maps. Given that there is no universal standard for threshold levels across all types of experiment and analysis, this dependence further motivates the direct calculation of spatial error. The potential complexity of the relationship revealed by this simple exercise underscores the importance of expressing spatial errors when addressing spatial questions.

In real neuroimaging data, there are likely to be many other factors that affect the relationship between intensive and spatial errors. In light of the complexity of this relationship even in a simple simulation, it appears critical to estimate spatial error, as well as to use an estimation technique that does not make strong parametric assumptions. In the following sections, we illustrate the utility of the intensive-to-spatial transformation on actual fMRI data.

Sample fMRI data and analysis

Here we briefly describe the sample data used in the following two applications.

fMRI

Data collection

All data were collected on a 3T GE scanner at the Imaging Research Center, The University of Texas at Austin, except for the retinotopy data, which were collected on a 3T GE scanner at Stanford University. Human subject procedures were approved by the respective Institutional Review Boards. Specific data sets were selected simply on the basis of graphical clarity, convenient numbers of repetitions, and general representativeness. We confirmed that the analyses described in this report could be applied to similar data sets, yielding similar results.

The MT+ localizer data (Application 1) consisted of 10 identical repetitions of a simple blocked alternation experiment, each lasting 240 sec. The subject viewed 10 24-sec cycles of a pattern of dots that switched between moving (12 sec) and stationary (12 sec). An fMRI volume was collected each 3 sec along a series of planes (slices) through the subject’s brain.

The retinotopy data (Application 2) consisted of 4 repetitions of a rotating wedge stimulus or an expanding ring stimulus (described in more detail in Application 2, below), each lasting 114 sec. The subject viewed 6 36-sec cycles of the stimulus. An fMRI volume was collected each 1.8 sec.

For the initial MT+ localizer example, the results of fMRI data analysis are shown overlaid on anatomical images acquired on the same slices. For the retinotopy example, the fMRI data were coregistered to a
high-resolution anatomical volume. The gray matter of this anatomical volume was segmented into gray and white matter, and a computationally flattened view (“flat map”) is shown, with corresponding results of fMRI data analysis overlaid.

Data analysis

fMRI data were analyzed using standard methods described in detail elsewhere (Brewer et al., 2005; Dougherty et al., 2003; Engel et al., 1994, 1997; Larsson & Heeger, 2006; Wade et al., 2002). We emphasize that the bootstrapping implementation (described in detail in a later section) does not dictate any specific analysis method. The particular analysis method used in this paper was chosen simply for simplicity and illustrative purposes, in addition to the fact that many of the techniques are relatively standard.

In brief, the fMRI response at each voxel was fit with a sinusoid with the same frequency as the stimulus. This yielded three numbers: the amplitude and phase of the best-fitting sinusoid; and coherence, the correlation between the data and this best-fitting sinusoid. Coherence is a standard measure of signal-to-noise ratio. For the purposes of this exercise, it can be thought of as loosely equivalent to any of the standard signal-to-noise-based statistics considered in most voxel-based analyses (e.g., t-statistics, F-values, z-scores, etc.).

Use of the bootstrap to implement the transformation from intensive error to spatial error

The proposed spatial uncertainty framework exploits the fact that multiple repetitions have been acquired, resampling them to infer the sampling distribution of the ROI across space. To appreciate the specifics, we begin by describing each run, r, as a four-dimensional x–y–z–t block of data (3 spatial dimensions of the images, fluctuating over time). Multiple (n) runs are acquired, and so the ith repetition is referred to as ri. Critically, the bootstrapping procedure will resample from this collection of rs, meaning that each complete x–y–z–t block of data (ri) is treated as the core unit. This preserves all the complex spatiotemporal correlation structure that is present in fMRI data, whatever its exact form may be. Thus, the bootstrap makes no assumptions about the particular form of the relation between intensive error and spatial error (e.g., those shown in Figure 2).

To implement a bootstrapping procedure, the blocks of data r are resampled with replacement to create a bootstrapped data set of length n. Data analysis is then performed (in this example, simple calculation of coherence followed by thresholding), and the results (e.g., locations of suprathreshold voxels composing the ROI) are recorded. This process is then repeated for a large number of times (200 or more), with each iteration operating upon another n samples from r selected with replacement. The resulting distribution of the results acquired across all bootstrapped iterations can then be
analyzed and visualized in any number of ways. Figure 3 shows a graphic schematic of the spatial uncertainty framework and bootstrapping implementation.

Estimates of MT+ boundary location uncertainty

We applied this bootstrapping procedure to the localization of visual area MT+. As described above, we started with the original data that consisted of several runs of a standard MT+ localizer. Figure 4 (top left) shows the averaged data, analyzed using standard methods and then thresholded.

Figure 4 (right column) shows the bootstrapping process used to transform the intensive error into spatial error. On each of the 200 bootstrapping iterations, we resampled the same number of runs from the original data, with replacement. We then averaged each bootstrapped data set and performed a standard correlational analysis: identifying voxels whose intensity time course exhibited coherence with the moving vs. stationary stimulus. Voxels that were above a particular coherence threshold were identified as distinct regions using a generic algorithm. On each bootstrapping iteration, we recorded the locations of the voxels falling within the suprathreshold region. The right column shows the results from 3 of the 200 bootstrap iterations: thresholded average data (upper) and resulting identified regions (lower). Note that, although the results of each iteration are similar, there is some degree of variability across them. This variability results from the process of resampling with replacement and is the critical element of the bootstrapping process.

By summing over all bootstrapped replicates, one can calculate the spatial sampling distribution of the boundary. Figure 4 shows two representations of the resulting variability, which serve as complimentary estimates of the spatial uncertainty associated with the location of MT+ (middle and lower panels, left column). The top panel shows the original post-threshold data, arrived at via conventional analyses. The middle panel shows the same data, superimposed with green dots indicating the midpoint of the estimated boundary location distribution (i.e., the 50th percentile of the distribution of boundaries). Shown in blue are 1-D error bars that represent the bootstrapped 95% confidence intervals about the horizontal or vertical location of the edge (i.e., the 2.5th and 97.5th percentiles; these error bars are analogous to those shown in Figure 1D). These 4 error bars were chosen in illustrative locations: note that they are of very different widths (e.g., the lowest error bar is smaller than the plotting symbol). This demonstrates that the spatial pattern of activity is not well approximated by a 2-D Gaussian and implies that nonparametric estimation of spatial uncertainty is important.

The bottom left panel again shows the original post-threshold data, now with a trio of contours that depict the complete 2-D profiles of spatial uncertainty. The green middle contour represents the median location of the entire 2-D boundary of MT+, and the inner and outer iso-certainty contours correspond to the 2-D 95% confidence interval. This 2-D confidence interval represents the region within which 95% of the estimated MT+ boundaries fell. In this paper, we adopt the convention of referring to bootstrapped 68% and 95% confidence intervals as ±1 and ±2 standard errors about the mean (SEM).
If one wanted to adopt a conservative definition of MT+, one could select only the voxels that fell within the inner boundary (corresponding to an alpha level of 0.025). An intermediate proposal would be to pick the voxels that were more likely than not to fall within MT+ (e.g., within the green boundary). A more liberal definition of the area would include all voxels that fell within the outer boundary.

Figure 4. Quantification of spatial uncertainty about the location of human area MT+. Left column illustrates ways of representing spatial uncertainty. Top panel shows post-threshold responses from average of 10 MT+ localizer runs. Note that this traditional analysis focuses on the intensity of each voxel and does not represent spatial error. Middle panel shows same post-threshold mean data, with four one-dimensional representations of spatial uncertainty about the boundary of the suprathreshold activity region. Green points indicate the 50th percentile of edge location, and blue error bars show the 95% confidence interval about this median edge position in either the horizontal or vertical dimension. The locations of these horizontal and vertical error bars were chosen to illustrate the inhomogeneity in spatial uncertainty surrounding different parts of the region (e.g., the lower error bar is smaller than the plotting symbol). Bottom panel shows the full 2-D pattern of spatial uncertainty. Green contour indicates the 50th percentile of the edge location, and the blue contours represent the 2.5th and 97.5th percentiles, i.e., the 2-D 95% confidence interval. Right column illustrates the bootstrapping implementation to generate these spatial confidence intervals. Runs of an MT+ localizer were resampled with replacement to create new replicated data sets. These data sets were averaged, analyzed for coherence, and thresholded to create mean data samples. Regions of interest and subsequent borders of these ROIs were identified in each resampled average and summed over the total number of bootstrapped replicates to assess spatial error.
These error bars and boundaries serve as intuitive, graphical, and quantitative representations of the spatial uncertainty associated with the location of a functionally defined visual area. We emphasize that these estimates of spatial uncertainty are distinct from estimates of error in the intensive dimension and can be calculated using resampling techniques that do not require one to impose or assume a particular form of error in either the intensity or spatial dimensions. Finally, these estimates of spatial uncertainty provide information not present in the activation maps alone that can support quantitative claims and principled arguments about the location of a functional brain area.

Further applications and insights

It is important to note that the particulars of the above example are not critical for the application of this bootstrapping process. So long as the fMRI data are resampled properly (i.e., as \(x\)-\(y\)-\(z\)-\(t\) blocks from the original data), any statistical analysis can be applied to the data, and any particular way of deciding and recording activated voxels or clusters can be implemented. In fact, the results of the bootstrapping process can be used to gauge the relative efficacy of particular analysis schemes or steps.

For example, one can consider the effects of different statistical thresholds on the resulting spatial uncertainty of localization. Given that statistical thresholds are general conventions for accepting a particular amount of possible error and noting that many assumptions underlying classical statistics are violated in the specific case of fMRI data, choices of thresholds are often viewed as particularly arbitrary in the analysis of neuroimaging data. In the case of mapping visual areas, one might decide that a less arbitrary or assumption-laden approach would be to pick the statistical threshold that minimized spatial uncertainty surrounding key features of the data.

One might expect (or hope) that, over some range, raising the statistical threshold might reduce the spatial uncertainty of an area boundary by collapsing the region of uncertainty around the “true” boundary of the area. However, an analysis of the effects of thresholding reveals that this is not the case, at least for the data being considered. Figure 5 shows the 95% confidence contours...
surrounding the activated patch of motion-responsive cortex considered in Figure 4. Each row shows the same data analyzed identically, except for the application of an increasingly high threshold on the coherence statistic (top to bottom).

The outer and inner blue contours indicate the 2.5% and 97.5% confidence levels associated with the edge of the activated region, and the green contour shows the median edge location. The outer contour loosely tracks the locations of the suprathreshold voxels, gradually shrinking as the coherence threshold is raised (top to bottom). Ideally, the location of the green contour would remain roughly fixed, at least for some range of thresholds, and the size of the inner blue contour would increase, approaching the median as threshold increases. In fact, the inner region of high spatial confidence is relatively large at a low coherence threshold and gradually shrinks as the coherence threshold is increased.

There are many implications of this, but we wish to emphasize two here. The first is that the spatial certainty of a region varies with the applied threshold, as does the estimated size of the region. Knowing the uncertainty associated with the estimated edge location thus provides an additional key piece of information to consider when selecting a threshold for data display and further analysis. In this instance, for example, thresholds of 0.3 or lower might be suspect because large portions of the resulting edge distribution fall outside the gray matter.

The second is that there might be situations in which, over some threshold range, the median location of an area boundary does in fact remain relatively fixed while the error distribution shrinks around it. In such a case, a very principled decision could be made about what threshold gives the best estimate of the true size and location of the putative area.

In addition to threshold level, other data processing steps can similarly be varied (or included and excluded) to assess their effects on the certainty of spatial localization. In general, this approach allows for the characterization of spatial uncertainty in data and how various analysis steps affect it, without requiring assumptions or assertions (e.g., spatial smoothing). Various statistical models and/or preprocessing operations can be assessed in a similar way. The bootstrapping implementation provides an unbiased arena for these comparisons, as it does not make parametric assumptions that could bias the results in favor of one particular approach.

Although we have used 200 bootstrap iterations to estimate the sampling distributions of key spatial features, we note that it is becoming feasible to perform larger numbers of iterations (e.g., $10^3 - 10^4$). We confirmed that the results shown in this paper do not change appreciably when greater than 200 iterations (i.e., the differences fall within measurement error). However, future applications that involve larger numbers of free parameters and/or larger numbers of runs that can be resampled may indeed benefit from more iterations.

### Advantages of the bootstrapping implementation

Resampling approaches have previously been applied to fMRI data for purposes including characterization of voxel-by-voxel false positive rates and the assessment of particular analysis schemes (Auffermann, Ngan, & Hu, 2002; Baumgartner et al., 2000; Biswal, Taylor, & Ulmer, 2001; Bullmore et al., 2001, 2003; Friman & Westin, 2005; Jahanian, Hosseinzadeh, Soltanian-Zadeh, & Ardekani, 2004; McIntosh, Chau, & Protzner, 2004; McKeown & Hanlon, 2004; Nandy & Cordes, 2007; Sendur, Suckling, Whitcher, & Bullmore, 2007; Strother et al., 2002); bootstrapping has also been applied to the analysis of diffusion tensor imaging (Jones, 2003; Pajevic & Basser, 2003) and EEG data (Benar, Gunn, Grova, Champagne, & Gotman, 2005). There has also been work investigating the relation between spatial smoothness of fMRI data and thresholds placed on the size of an activated cluster, including applications to cortical maps rendered on computationally flattened gray matter (Hagler, Saygin, & Sereno, 2006).

We have proposed to apply the bootstrap in a distinct manner, using it to provide a simple means for explicitly estimating a spatial confidence interval, thus transforming the analytic focus from voxel intensity to spatial distributions of active voxels. This implementation does not typically require the collection of additional data nor does it involve additional forms of data processing to facilitate the analysis. In the following section, we show how this procedure can be applied to characterize the spatial uncertainty associated with other types of spatial features in neuroimaging data.

### Application 2: Characterization of spatial uncertainty along the border between retinotopic visual areas V1 and V2

We now apply the framework to the identification of human primary and secondary visual cortices from phase-encoded retinotopy experiments. In these experiments, a spatially restricted pattern of changing contrast is shown to the subject in one of two forms. In this example, we first consider the mapping of visual field angle, which is critical for defining the boundaries between abutting visual field maps. Polar angle (with reference to the vertical meridian, with the center of fovea as the origin) is mapped using a flickering checkerboard wedge that rotates periodically in a clockwise or counterclockwise direction through the visual field. As the wedge rotates, it creates a wave of activation throughout retinotopically organized visual areas, successively and systematically.
stimulating portions of each map. In this way, the entire visual field is represented by a time-dependent pattern of activity across space.

The time series from each voxel is analyzed to estimate the part of the visual field to which it is maximally responsive. Each voxel’s time series is fit with a sinusoid, and the coherence, phase, and amplitude of the sinusoid are extracted. The phase is an indicator of the timing of the voxel’s response, which thus corresponds to a particular position of the stimulus in the visual field. A pseudocolor representation of phase is commonly shown and used to delineate the borders between retinotopic maps (Figure 6, top).

However, because the border is represented by a color corresponding to a phase reversal, the exact location of the border within this spatial band of reversal phase is ambiguous. Most voxels in this region will correspond to clearly being in visual area V1 or visual area V2. Voxels at the border, however, cannot be confidently assigned to either V1 or V2. Thus, classifying voxels that do not reliably fall within either V1 or V2 provides a principled definition of the V1/V2 border that does not ignore uncertainty. Here we apply the bootstrapping framework to this simple example, but we emphasize that it could be applied to the assessment of borders between areas in which the retinotopic maps are less clear, or are subject to theoretical debate, or both, and thus are more deserving of quantitative assessments of spatial certainty.

To demonstrate this application, we performed bootstrapping to quantify the spatial certainty associated with phase-encoded retinotopic mapping. Figure 6 shows the results of a conventional phase-encoded retinotopy experiment and the resulting estimates of spatial uncertainty generated by the bootstrapping framework. The top panel of the figure shows a standard pseudocolored map of visual field representation: different colors indicate different temporal phases; different temporal phases indicate responsiveness to different parts of the visual field. In the

![Retinotopic phase angle map](image1)

![Standard error contours](image2)

Figure 6. Spatial uncertainty associated with the V1/V2 boundary as determined by retinotopic mapping of visual angle. Left column (top) shows the retinotopic map of phase angle, generated using a rotating wedge stimulus, on a flat map of the anatomy. Black lines were hand-drawn at visually determined borders between areas V1 and V2, and areas V1 and V2 are labeled, as per common practice. Left column (bottom) shows spatial confidence intervals about the location of the V1/V2 dorsal and ventral boundaries on the retinotopic map of phase angle. White contour lines represent one standard error of the mean around the phase reversal known to be the V1/V2 boundary. Right column shows the bootstrapping implementation used to define standard error contours. Regions of interest were found by resampling with replacement from the original rotating wedge runs and restricting phase to values around the phase reversals denoting the V1/V2 border. Phase reversal regions were summed over all bootstrapped replicates to assess spatial error.
center of the map, the smooth color/phase/position transitions from cyan to blue to magenta to red indicate a systematic hemifield representation, indicative of V1.

As per common practice, hand-drawn borders between V1 and the dorsal and ventral components of V2 are shown with black lines (Figure 6, top panel). These borders are identified at the points of phase reversal, the direction of color change makes an abrupt flip, indicating the existence of an abutting, mirror-symmetric retinotopic map. Semi-automated approaches to the identification of retinotopic maps have also been described and typically rely on fitting some form of a template to the data. Regardless of the particular method used to identify maps, none of these approaches are typically extended to provide a straightforward estimate of the spatial certainty associated with the locations of boundaries between areas. In other words, there is not usually an error bar shown about the proposed locations of boundaries.

We quantified the spatial certainty associated with the V1/V2 boundary using a bootstrapping procedure as follows. The procedure was very similar to that used to characterize MT+ in Application 1, except that we were primarily interested in identifying the location of a boundary defined by a phase reversal (as opposed to simply a suprathreshold response level).

On each bootstrapping iteration, we resampled runs of the retinotopic mapping experiments with replacement from the original data. We averaged these resampled data and created a bootstrapped phase map similar to that shown at the top of Figure 6. We then identified the voxels with phases consistent with the point of reversal between maps, corresponding to the dorsal and ventral borders between these areas. We recorded the location of these phase reversal regions. Each bootstrapped replicate yielded slightly variable estimates of the V1/V2 phase-reversal border (Figure 6, right column), and this variability again allows us to quantify the spatial uncertainty associated with these measurements.

Finally, we accumulated these bootstrapped phase reversal regions over all bootstrapped iterations, and then visualized the results in a manner similar to that used in Application 1. Figure 6 (left bottom panel) shows the 2-D confidence interval contour associated with the location of the phase reversal between V1 and V2. Voxels falling within the white contour were identified as part of the V1/V2 boundary frequently; \( \geq 16\% \) of the time). This corresponds to the outer edge of a 1 standard error region around the border (i.e., the 16th and 84th percentiles form a 68% confidence interval, which is equivalent to one standard error). In other words, voxels falling within the contour are within one error bar of the boundary; voxels falling outside cannot be confidently assigned to the boundary and can be assigned to V1 or V2.

Several aspects of the forms of these confidence boundaries around the V1/V2 border are worth considering. First, the confidence boundaries are not infinitely narrow, confirming that there is of course substantial spatial uncertainty associated with the precise location of the V1/V2 border.

Second, the confidence boundaries have somewhat idiosyncratic shapes, demonstrating that the shape of the boundary itself may be irregular and that the corresponding degree of spatial confidence is not smooth. Some parts of the boundary are identified with higher certainty than others. This is especially evident in the boundary surrounding the magenta-colored phase reversal.

Third, and perhaps most importantly, the degree of certainty as quantified by the boundaries does not perfectly map on to the qualitative sharpness of the changing pattern of pseudocolor shown on the map. It is anecdotally known that the choice of colors can have profound effects on the apparent sharpness of the phase reversals. Thus, different visualization formats can have strong effects on how compelling a particular map might appear. In contrast, the quantitative confidence intervals are immune to changes in display format and are thus an improvement on subjective estimates of how quickly and cleanly the colors appear to shift.

This analysis demonstrates that, regardless of the degree of uncertainty one is willing to accept in distinguishing two retinotopic maps, the boundary between these visual areas should not be thought of as a crisp edge. Rather, analysis of spatial uncertainty reveals a phase reversal (border) region, a set of voxels that cannot be confidently assigned to either abutting map. If one wanted to be particularly conservative in distinguishing the two areas (perhaps for the purposes of testing for additional functional differences between these two areas), selecting V1 and V2 as the respective regions that end at the outer contour surrounding the V1/V2 boundary would be appropriate. Other sizes of confidence intervals (e.g., 5% or 95%) could be used to perform more conservative or liberal distinctions between maps and their border.

We emphasize that we are not proposing the specific method of defining a phase-reversal band as a preferred method for identifying the V1/V2 (or any other) border, or for quantifying the associated spatial uncertainty. This method was simply intuitive to use and easy to implement. Rather, we wish to emphasize that the spatial uncertainty framework (and the bootstrapping implementation) can be used to transform voxel (intensive) uncertainty into spatial (extensive) uncertainty for any border or region finding method. In fact, we would welcome a demonstration of a more precise method for finding the V1/V2 border using the spatial uncertainty framework. Our proposed framework could in fact be used to compare the spatial precision of various techniques. One last point that can be made is that even the relatively course border identification technique we employed revealed pleasingly narrow estimates of spatial uncertainty.

This approach can also be used to assess other spatial features of retinotopic maps. For example, the representation of eccentricity is also critical for defining visual area maps. First, finding the foveal representation provides a
starting point from which to expect radiating lines of iso-angle representation. Second, when maps are ambiguous or contentious, it is often helpful to confirm that the representations of angle and eccentricity are orthogonal.

We thus considered the spatial uncertainty associated with the cortical representation of visual field eccentricity. Eccentricity was mapped using a phase-encoding procedure analogous to that for the polar angle representation, except the stimulus was a checkerboard ring that expanded through the visual field, periodically cycling from the fovea to far eccentricities. Figure 7 shows the results of estimating spatial uncertainty. To do this, we used a bootstrapping procedure similar to that used for the analysis of the polar angle representation, except that we identified voxels whose phases were consistent not with phase reversal borders but with particular eccentricity ranges, such as the fovea and a more peripheral (i.e., farther eccentric) representation (shown by the purple and yellow ranges in Figure 7, respectively).

The results of this analysis confirm that the fMRI data yield precise estimates of eccentricity. Confidence boundaries are drawn around the foveal representation (purple region, marked with an asterisk) and an eccentric representation (yellow region). Note that the confidence boundaries confirm cortical magnification (i.e., the foveal representation is expanded relative to the more eccentric representation) and also reveal that the shapes of these representations may be irregular. One notable observation is how narrow the confidence boundary around the more eccentric representation is, as compared to the representation of angle shown in the previous figure. This suggests that the representation of eccentricity may be more precise than the representation of angle, at least in the data set considered here.

This quantitative comparison serves as an example of how this framework can allow one to compare the quality of various data sets. One could use these analyses to guide future work investigating whether various stimulation protocols yield lower spatial error, and/or whether competing proposed maps are more or less precise than one another. One could also use these results to decide the relative number of repeats for angle and eccentricity mapping within a session, if equal amounts of spatial uncertainty are desired.

Conclusions

In summary, we have described a framework for transforming intensive error to spatial error associated with spatial features in fMRI data. We also proposed a method for implementing this transformation using bootstrapping, thus providing an implementation that does not require making parametric assumptions, nor does it rely on filtering data to conform to convenient assumptions. This framework and implementation allow for any type of data analysis to be performed and afford an unbiased arena to evaluate and compare the results of different data analysis approaches.

In this paper, we have chosen two simple analysis techniques to illustrate the spatial uncertainty approach, but any form of data analysis can be performed on the resampled data, and the results of the analysis provide a quantification of the spatial certainty of their output. This is a key strength of the bootstrapping implementation, given that there is rarely a single standard way to analyze the results of a particular fMRI experiment and given that elements of particular analyses like the choice of statistical threshold are often contentious or nearly arbitrary.

Why quantify spatial uncertainty?

The spatial uncertainty framework allows one to characterize not just where an active region may be, but
the spatial pattern of uncertainty associated with this localized region. Standard practice in reporting locations of regions involves thresholding a statistical map to yield punctate “hot spots” of activation that are “significant” (in the null-hypothesis-testing sense) along the dimension of activation but says nothing explicit about the certainty of actual location on the cortical surface. This amounts to reporting the mean (of the spatial activation) without reporting the standard error or confidence interval (in space). Such a practice would generally be frowned upon in conventional data reporting and is particularly troublesome in fMRI analysis, given that the choice of statistical threshold is often arbitrary, or is based on various parametric assumptions or post hoc corrections, or both. Thus, nonparametric estimates of spatial uncertainty provide a critical piece of information when reporting the spatial distribution of activity. We propose that such a method of reporting uncertainty in space should be customary.

The second advantage applies to the identification of retinotopic maps in cases when the existence and form of such spatial representations are contentious and ambiguous. For example, much ongoing work and some significant debate have surrounded the existence and form of maps around human V4, dorsal extrastriate areas like V3A/V3B/V7, the intraparietal sulcus, and the lateral occipital (LO)/MT+ complex (Brewer et al., 2005; Hadjikhani et al., 1998; Hansen et al., 2007; Huk, Dougherty, & Heeger, 2002; Larsson & Heeger, 2006; Schluppeck et al., 2005; Sereno et al., 2001; Swisher et al., 2007; Tootell & Hadjikhani, 2001; Wade et al., 2002). In some of these cases, maps have been proposed by one group but reinterpreted or called into question by other groups. Although certain qualitative criteria can be applied to evaluate some proposals (e.g., the constraint that representations of polar angle and eccentricity be approximately orthogonal), a quantitative method for assessing uncertainty seems important. This is especially the case given that retinotopic maps in later visual areas are often noisy, different data analysis and rendering schemes may differentially accentuate or obscure various elements of this noise, and examples may be handpicked post hoc to bolster a particular interpretation.

The spatial uncertainty approach, in conjunction with the bootstrapping implementation, provides a rigorous means for evaluating the confidence associated with various proposed maps, as well as a means for assessing the quality of data from different groups and from different analysis schemes (e.g., Dumoulin & Wandell, 2008; Duncan & Boynton, 2003; Larsson & Heeger, 2006; Schira, Wade, & Tyler, 2007; Sereno et al., 1995). Within a study, one could use this framework to test whether proposed boundaries between noisy maps are statistically reliable and/or whether representations of angle and eccentricity are orthogonal. Likewise, quantitative comparisons across studies could test whether competing accounts of retinotopic organization are supported by comparable amounts of spatial certainty, potentially revealing differences in data quality or the precision of analyses.

In summary, we propose that the characterization of spatial uncertainty associated with the localization of visual areas should be considered as a complement to the standard focus on error in the intensities of individual voxels. It provides a quantitative tool that can strengthen and disambiguate the typically qualitative inferences made about spatial patterns of activity. This approach can be implemented in a way that does not require additional data, can be applied to any specific data analysis scheme, and can also be used both to perform power analyses and to compare the efficacy of various analysis schemes. We have shown that this approach can provide intuitive, graphical descriptions of spatial uncertainty, as well as quantitative confidence intervals in space. We propose that reporting of such nonparametric spatial uncertainty should be customary and suggest that such estimates may facilitate the resolution of ongoing debates associated with functional localization and retinotopic mapping.

Acknowledgments

This work was supported by NSF CAREER Award BCS-0748413 and a Mind Science Foundation research grant to A.C.H., and UT-Austin Imaging Research Center pilot grants to A.C.H., and L.K.C. We thank David Ress for discussions and assistance with programming.

Commercial relationships: none.
Corresponding author: Lawrence Cormack.
Email: cormack@mail.utexas.edu.
Address: Center for Perceptual Systems, 1 University Station, A8000, University of Texas, Austin, TX 78712, USA.

References


fMRI methods: Calculation of the uncertainty in the location of an activation site in a PET image. *Neuroimage*, 14, 194–201. [PubMed]


