The Benefits of Rapid 3D fMRI

Martin A. Lindquist

Department of Statistics, Columbia University, New York, NY 10027

Received 23 September 2009; accepted 4 December 2009

ABSTRACT: Functional magnetic resonance imaging (fMRI) provides the ability to image blood dynamics through the entire brain with a high spatial resolution. However, the temporal resolution is much slower than the underlying neuronal activity one seeks to infer. Recent developments in rapid imaging allow 3D fMRI studies to be performed at a temporal resolution of 100 *ms*; a 10-fold increase compared to standard approaches. This increase in temporal resolution offers a number of potential benefits. First, it allows the focus of analysis to be shifted from changes in blood flow taking place 5–8 *s* after neuronal activity to more transient changes taking place immediately following activation. We argue that studying these changes provides valuable information about the relative timing of activation across different regions of the brain, which is crucial for inferring brain pathways. Second, rapid imaging allows for the efficient modeling of physiological artifacts without problems with aliasing; something that is difficult at standard resolutions. We illustrate how removal of these artifacts provides the increase in signal-to-noise ratio required for studying the subtle changes in oxygenation we are interested in. Finally, we show how high temporal resolution data provides the opportunity to focus the analysis on the rate of change in oxygenation rather than the level of oxygenation as is the current practice. The price of performing rapid imaging studies is a decrease in spatial resolution. However, we argue that the resolution is still comparable to the effective resolution used in most fMRI studies. We illustrate our approach using two fMRI data sets. Int J Imaging Syst Technol, 20, 14–22, 2010; Published online in Wiley InterScience (www.interscience.wiley.com). DOI 10.1002/ima.20217

Key words: rapid fMRI; Echo-Volumar imaging; single shot; initial negative dip; physiological noise; smoothing

I. INTRODUCTION

Functional magnetic resonance imaging (fMRI) provides the ability to image blood dynamics through the entire brain with a high spatial resolution. However, the temporal resolution is comparatively low which can lead to a number of possible problems. First, since the standard temporal resolution is slower than the underlying neural activity being studied, one faces the statistically intractable task of sorting out possibly unknown confounding factors influencing the ordering of the time of brain activity across different regions of the brain. This limitation impacts the ability to accurately estimate the brain pathways that are essential for understanding higher cognition. Second, currently used resolutions are not conducive to efficient modeling of physiological artifacts present in the fMRI signal (e.g., heart rate and respiration). As fMRI signal suffers from low signal-to-noise ratio (SNR) and physiological artifacts make up a large portion of the noise component, this is a serious impediment.

Functional MRI is most commonly performed using blood oxygen level dependent (BOLD) contrast (Ogawa et al., 1992) to study the vascular response in the brain to neuronal activity. The ability of BOLD imaging to infer the timing of neuronal activity is limited by the sluggish nature of the evoked hemodynamic response to a neural event, and analyses are generally focused on tracking changes in blood flow peaking 5–8 s after that neural activity has peaked. As the shape and timing of the response is known to vary across brain regions (Aguirre at al., 1998) it is difficult to make valid comparisons of the timing of activation between regions. However, a number of studies have shown evidence of the existence of an early transient component of the hemodynamic response, associated with a rapid increase in oxidative metabolic rate prior to the subsequent increase in blood flow (Cho et al., 1992; Ernst et al., 1994; Menon et al., 1995; Malonek et al., 1996; Yacoub et al., 1998). It has been hypothesized that this transient signal, referred to as the initial negative dip, may contain more precise information about both the location and timing of activity (Yacoub et al.,2001; Duong et al., 2000; Kim et al., 2000; Thompson et al., 2004). To date, this signal has proven difficult to detect in standard fMRI studies partly due to the coarse temporal resolution, and partly to low SNR.

In all fMRI studies, the observed signal is corrupted by random noise and nuisance components caused both by hardware reasons and the subjects' themselves. Heart-rate and respiration give rise to periodic fluctuations that are difficult to model due to violations of the Nyquist criteria, which states that it is necessary to have a sampling rate at least twice as high as the frequency of the periodic function one seeks to model. At standard temporal resolutions (e.g., TRs of 2 *s*) this is clearly violated and heart rate and respiration is often left unmodeled. Hence, aliasing will cause these fluctuations to be distributed throughout the time course giving rise to temporal autocorrelation.

Both issues outlined above can be effectively addressed through increases in the temporal resolution of fMRI studies. A 10-fold

Correspondence to: Martin Lindquist; e-mail: martin@stat.columbia.edu
Martin Lindquist's research is partially supported by NSF grant DMS-0806088.

increase would allow researchers to model and remove physiological artifacts, as well as study transient components of the hemodynamic response. However, such an increase necessitates a shift in the manner in which data is acquired. Typically, 3D imaging is performed by acquiring a stack of adjacent slices (e.g., 20–30) in quick succession. Since each slice is measured in sequence, and nuclei must be re-excited between slices, this places constraints on the time needed to acquire a whole brain volume. As an alternative, it is possible to design sampling schemes that directly measure points in 3D k-space (Mansfield et al., 1989; Mansfield et al., 1995; Lindquist et al., 2008a) in a single-shot; providing the opportunity for radical increases in temporal resolution.

In previous work, we have introduced an echo-volumar imaging (EVI) trajectory that allows 3D k-space to be sampled in a single-shot with a TR of 100 *ms* (Lindquist et al., 2008a; Lindquist et al., 2008b). We use the increased resolution to efficiently model physiological artifacts, thereby obtaining an increase in SNR. This, in turn, allows us to measure the initial decrease in oxygenated blood that takes place with activation, rather than the replenishment of oxygen to affected regions as is common practice. This allows us to move the statistical analysis significantly closer in timing to the underlying neuronal activity that we seek to understand. We believe this will enable researchers to accurately measure the cascade of activation in the brain, allowing for accurate estimation of the brain pathways essential for understanding higher cognition. This would make fMRI measurements more comparable to invasive methods of studying brain function used in animals (e.g., implanted electrodes). It also allows us to apply statistical techniques for modeling the hemodynamic response function across different brain regions as continuous functions. Using these techniques we can compute the rate of *change* in oxygenation level, which we believe provides valuable information not apparent when solely studying the oxygenation level. The price of increasing the temporal resolution is a sacrifice of spatial resolution. However, we argue that this sacrifice is relatively minor due to the manner in which fMRI data is typically analyzed. In addition, the use of parallel imaging promises to further bridge the gap in spatial resolution in the future.

This paper is organized as follows: In Section 2 we discuss our approach towards the acquisition and reconstruction of rapid 3D fMRI data. In Section 3 we compare the spatial properties of the resulting data set to those used in most standard fMRI studies. In Section 4 we discuss techniques for preprocessing and analyzing high temporal resolution data. In Section 5 we present results from two separate pilot studies performed using a visual-motor and an auditory-motor-visual paradigm, respectively. Finally, we conclude with a discussion of the pros and cons of rapid imaging.

II. THE ACQUISITION AND RECONSTRUCTION OF RAPID 3D fMRI DATA

In this section we discuss the acquisition and reconstruction of high temporal resolution fMRI data. Much of the methodology has been developed in past work. The EVI trajectory we use is introduced in two recently published papers (Lindquist et al., 2008a, Lindquist et al., 2008b), while the reconstruction algorithm and the use of prolate speheroidal wave functions is described in a separate series of papers (Yang et al., 2002; Lindquist, 2003; Lindquist et al., 2006; Lindquist and Wager, 2007).

A. Acquiring Single-shot 3D EVI Data. The process of designing new k-space sampling trajectories is an interesting mathematical problem. The goal is to find a trajectory $\mathbf{k}(t) = (k_x(t), k_y(t), k_z(t))$ that moves through k-space and satisfies a series of constraints. The trajectory is defined as a continuous curve along which measurements are made at uniform time intervals determined by the sampling bandwidth of the scanner. The trajectory starts at (0,0,0) and its subsequent movement is limited by constraints placed on both its speed (gradient) and acceleration (slewrate). In addition, there is a finite amount of time the signal can be measured before the nuclei need to be re-exited and the trajectory returned to the origin. Finally, the trajectory needs to be space-filling. This implies that each point in the lattice contained within some region (e.g., cubic or spherical) around the center of k-space needs to be visited long enough to make a measurement. The size of this region determines the spatial resolution of the subsequent image reconstruction. For a more complete formulation of the problem we refer readers to Lindquist et al. (2008b).

Our goal is to determine a sequence of k-space points, measured in a single shot, that optimize the spatial resolution of the image subject to machine constraints. Due to convenience in image reconstruction and statistical analysis we currently limit ourselves to EVI trajectories; a class of trajectories previously considered in work by Mansfield and colleagues (Mansfield et al., 1989; Mansfield et al., 1995; Harvey and Mansfield, 1996). In our notation image-space is denoted by Ω and k-space by Ω^*. Mathematically, the measurement of the MR signal at the j^{th} time point of a readout period can be written

$$S(\mathbf{k}(t_j)) \approx \int_{\Omega} f(\mathbf{v}; t_j) \exp(i\mathbf{v}'\mathbf{k}) d\mathbf{v}, \tag{2.1}$$

where $t_j = j\Delta_t$ is the time of the measurement, $f(\mathbf{v}; t_j)$ the spin density over $\mathbf{v} \in \Omega$ at time t_j, and $\mathbf{k}(t_j)$ is the k-space point at which the signal is measured. Let $\mathbf{g}(t_j)$ and $\mathbf{s}(t_j)$ be the gradient and slew rate of the trajectory, i.e.

$$\mathbf{g}(t_j) = \frac{\mathbf{k}(t_j) - \mathbf{k}(t_{j-1})}{\gamma \Delta_t}, \quad \mathbf{s}(t_j) = \frac{\mathbf{g}(t_j) - \mathbf{g}(t_{j-1})}{\Delta_t},$$

where γ is the gyromagnetic ratio. Any trajectory must adhere to machine dependent constraints

$$\|\mathbf{g}(t_j)\| \leq G_0 G/\text{cm}, \quad \|\mathbf{s}(t_j)\| \leq S_0 G/\text{cm/ms} \tag{2.2}$$

which places constraints on the speed and acceleration of the trajectory.

In our application we are only interested in measuring the low-frequency portion of k-space. Hence, we limit sampling to a cubic (or hyper-rectangular) lattice with span $\Delta_k = 2\pi$ / FOV centered around (0,0,0), which we denote A. Given the total time T of a readout period and the shape and span of the k-space lattice A to be sampled, the problem is to maximize the size of A covered by a smooth curve $\{\mathbf{k}(t), 0 \leq t \leq T)\}$ satisfying $\|\mathbf{k}'(t)\| \leq \gamma G_0$ and $\|\mathbf{k}''(t)\| \leq \gamma S_0$ and the boundary conditions $\mathbf{k}(0) = \boldsymbol{k}'(0) = (0, 0, 0)$. We considered the class of all such curves composed of line segments parallel to the z-axis in k-space and half circles that change the *x-y* position among the different line segments; see Figure 1 for an example. We proved (Lindquist et al., 2008a) that the optimal solution is to set the diameter of the half circles to Δ_k, and to maximize the acceleration (deceleration) toward (from) the middle of each line segment in the z-direction, provided the condition $\|\mathbf{k}'(t)\| \leq \gamma G_0$ for the gradient holds. Since we wish to sample the center of

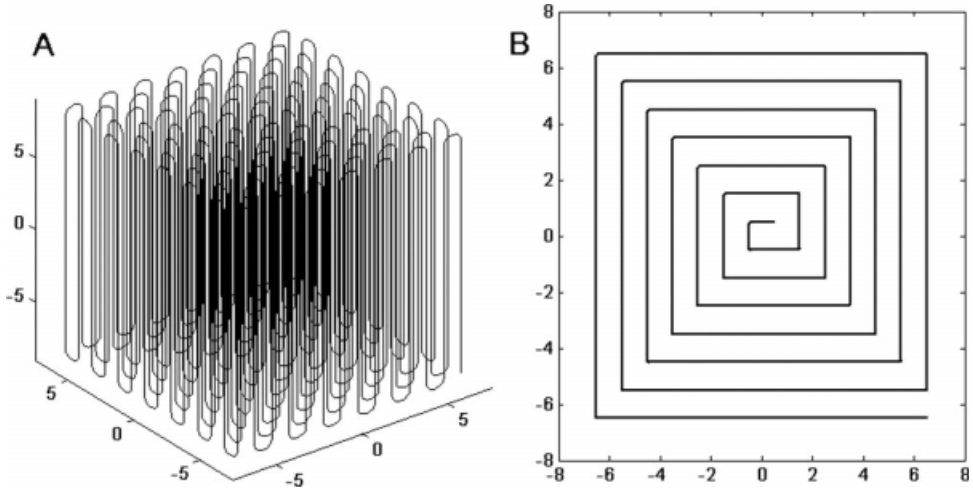

Figure 1. (A) An implementation of the echo-volumar imaging trajectory. (B) The echo-volumar imaging trajectory shown in (A) projected onto the xy-plane.

k-space at the beginning of the readout period while the SNR is high, the projection of our EVI trajectory to the *x*-*y* plane is a square spiral beginning from the origin. This EVI trajectory visits 11796 points in one readout period of 47 *ms* to cover a $14 \times 14 \times 17\ \Delta_k^3$ hyper-rectangle in the center of k-space, with FOV $= 2\pi\Delta_k = 24$ cm. Data is measured in a $14 \times 14 \times 46$ array representing measurements in 196 line segments of length $17\ \Delta_k$. Since the trajectory accelerates and decelerates in the line, the position of the 46 points in the *z*-direction are symmetric about $z = 0$, but not equally spaced. This trajectory is shown in Figure 1, while the individual k_x, k_y and k_z components are shown in Fig. 2.

B. Reconstructing 3D EVI Data. Once 3D k-space has been sampled, the data needs to be reconstructed for statistical analysis. A standard approach towards reconstructing nonuniformly sampled k-space data is to interpolate the data onto a Cartesian grid (Jackson et al., 1991) and thereafter apply the fast Fourier transform (FFT). As an alternative, we instead use the prolate-spheroidal wave function (PSWF) to recover images from k-space data, while simultaneously reducing truncation artifacts and spatially smoothing the data (Shepp and Zhang, 2000; Yang et al., 2002; Lindquist, 2003; Lindquist et al., 2006; Lindquist and Wager, 2007). The efficiency of the PSWF for handling under-sampled k-space data has been confirmed using both simulations and experimental data in previous work (Yang et al., 2002; Lindquist et al., 2006).

Let B be a convex region in Ω, whose shape and size will be described below, and A the $14 \times 14 \times 46$ set in k-space on which the data was measured. Now, let $w(\mathbf{k})$ be the function on A which maximizes

$$\lambda_B = \max \int_B \left| \sum_{\mathbf{k}\in A} w(\mathbf{k}) \exp(-i\mathbf{v}'\mathbf{k}) \right|^2 d\mathbf{v} \qquad (2.3)$$

subject to the constraint $\int_\Omega | \sum_{\mathbf{k}\in A} w(\mathbf{k}) \exp(-\mathrm{i}\mathbf{v}'\mathbf{k})|^2\ d\mathbf{v} = 1$. The PSWF is defined as

$$\phi(\mathbf{v}) = \sum_{\mathbf{k}\in A} w(\mathbf{k}) \exp(-i\mathbf{v}'\mathbf{k})/C_w, \qquad (2.4)$$

where $C_w = \int_\Omega \sum_{\mathbf{k}\in A} w(\mathbf{k}) \exp(-i\mathbf{v}'\mathbf{k})$. It follows that the PSWF is normalized to $\int_\Omega \phi(\mathbf{v}) d\mathbf{v} = 1$ and concentrated on B with $\lambda_B = \int_B |\phi(\mathbf{v})|^2 d\mathbf{v} \,/\, \int_\Omega |\phi(\mathbf{v})|^2\ d\mathbf{v} \approx 1$. Hence, the PSWF is the function that maximizes the signal concentration over the region B, for all functions having a Fourier transform with support on A.

Using the PSWF we can estimate the spin density as

$$\hat{f}(\mathbf{v}) = \sum_{\mathbf{k}\in A} w(\mathbf{k}) S(\mathbf{k}) \exp(-i\mathbf{v}'\mathbf{k})/C_w, \quad \mathbf{v} \in \Omega. \qquad (2.5)$$

It follows from (2.1) and (2.4) that the above estimator is approximately the convolution of the spin density and the PSWF:

$$\hat{f}(\mathbf{v}) \approx \sum_{\mathbf{k}\in A} w(\mathbf{k}) f^*(\mathbf{k}) \exp(-i\mathbf{v}'\mathbf{k})/C_w = \int_\Omega f(\mathbf{u})\phi(\mathbf{v}-\mathbf{u}) d\mathbf{u}. \qquad (2.6)$$

Since the PSWF concentrates in B with $\lambda_B \approx 1$, the size of B can be viewed as the spatial resolution of the reconstructed image (2.5). Typically, we choose B to be a circular region whose radius is chosen to be as small as possible while still retaining the property $\lambda_B \approx 1$. Compared with a slower full scan of k-space, the loss of spatial resolution in (2.6) is similar to smoothing the image post data acquisition, which is typically done anyway in fMRI data analysis. The main difference is the optimality of the kernel $\phi(\cdot)$ in (2.6). Ultimately, the spatial resolution using our approach is roughly equivalent to that obtained after applying a Gaussian filter with a FWHM of 12 *mm* to an image with dimensions $64 \times 64 \times 64$ and a *FOV* of 20 *cm*. Hence, the PSWF approach provides a means for performing both regridding and spatial smoothing as part of the reconstruction process.

III. SPATIAL PROPERIES OF 3D EVI DATA

The main drawback to our approach is that we sacrifice spatial resolution to increase temporal resolution. However, as alluded to above, a similar sacrifice is made whenever spatial smoothing is employed in the analysis of fMRI data. The process of smoothing is equivalent to applying a low-pass filter to the sampled k-space data prior to reconstruction. This implies that much of the acquired data

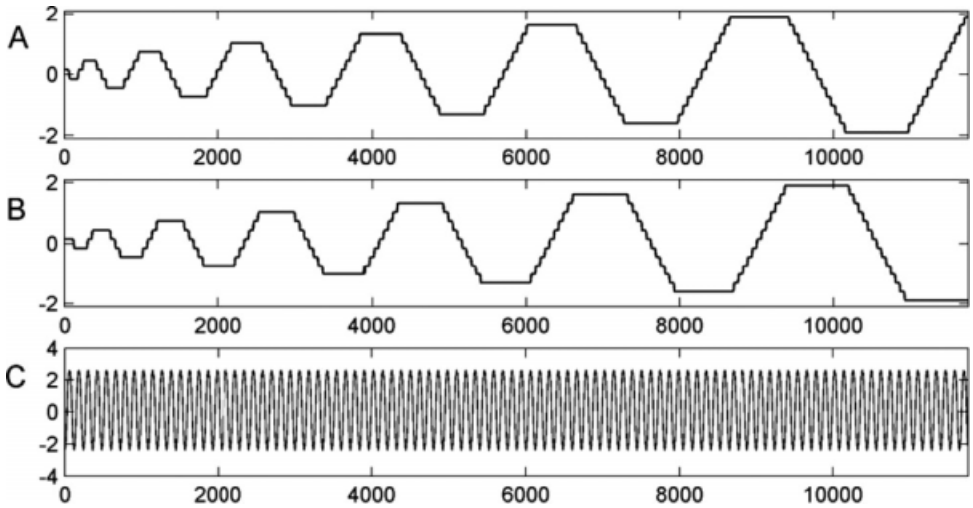

Figure 2. (A-C) Plots of k_x, k_y, k_z for the echo-volumar imaging trajectory shown in Figure 1.

is discarded as a byproduct of smoothing and spatial resolution is sacrificed. However, this sacrifice is made *after* the experimental k-space data has been collected and therefore does not allow for the possibility of gaining temporal speed-up as compensation. Additionally, acquiring an image with high spatial resolution and thereafter smoothing the image does not lead to the same results as directly acquiring a low resolution image. The signal-to-noise ratio during acquisition increases as the square of the voxel volume, so acquiring small voxels means that signal is lost that can never be recovered. Hence, it is optimal in terms of sensitivity to acquire images at the desired resolution and not employ smoothing. Our acquisition schemes have been designed to acquire images at the final functional resolution desired, thus allowing for much more rapid image acquisition as time is not spent acquiring information that will be discarded in the subsequent analysis.

In general, when performing multi-subject analysis a large amount of smoothing is almost always performed. This is done to overcome limitations in the spatial normalization by blurring any residual anatomical differences. In these situations, Gaussian kernels with widths of up to 12 *mm* FWHM are typically applied to the data. When applying a kernel of this size to standard fMRI data (e.g., 64 × 64, 20 *cm FOV*) only roughly a third of k-space data is actually used, with no increases in temporal resolution to show for it. For this reason alone, we believe that rapid 3D imaging should become the norm when performing group studies, as it provides a 10-fold increase in temporal resolution without any real lose in spatial resolution.

IV. ANALYZING RAPID 3D FMRI DATA

In this section we discuss the statistical analysis of high temporal resolution fMRI data. We begin by describing how efficient modeling of physiological artifacts can be used as a means of increasing the SNR, and continue by discussing several statistical procedures

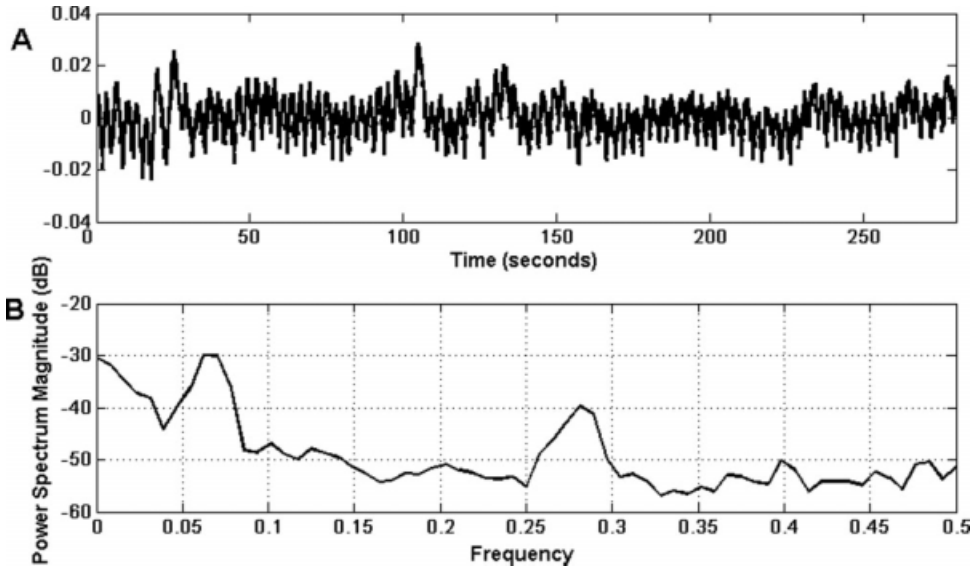

Figure 3. (A) A sample time course acquired using our EVI trajectory. (B) The periodogram of the time course shows clear spikes in frequencies associated with respiration and heart-rate. Removal of these frequencies will remove the periodic fluctuations apparent in the time course, thereby increasing SNR.

for studying transient components of the hemodynamic response such as the initial negative dip.

A. MODELING PHYSIOLOGICAL ARTIFACTS

The observed fMRI signal is corrupted by random noise and nuisance components caused both by the hardware and the subjects' themselves. In particular, physiological noise due to patient motion, respiration and heartbeat cause fluctuations in the signal across both space and time. The heart beats between 60–90 times per minute and respiration occurs roughly 15–20 times per minute, giving rise to periodic fluctuations in the signal. However, these fluctuations are difficult to model at standard temporal resolutions (e.g., 2 *s*) as the Nyquist criteria states it is necessary to have a sampling rate at least twice as high as the frequency of the periodic function one seeks to model (Lindquist, 2008). Therefore, aliasing will cause these signal components to be distributed throughout the time course giving rise to temporal autocorrelation and increased variability.

In high temporal resolution studies it is possible to completely circumvent problems related to aliasing. With measurements every 100 *ms* it is straight-forward to model and remove noise due to heart rate and respiration. One approach is to estimate both components and include them in a subsequent model (Lindquist et al., 2008b; Lindquist, 2008). Another option is to remove them through application of a properly designed band-pass filter. Figure 3 illustrates the latter approach. The top row shows a time course acquired with a TR of 100 *ms* using our rapid fMRI technique. There is a clear periodic component present in the signal. The bottom row shows the periodogram for the same time course. There are two obvious spikes in the spectrum corresponding to heart rate and respiration. By removing these frequencies one can effectively eliminate these components from the signal. As long as one is careful not to design the task to lie in the same frequency range, this will provide an increase in SNR.

B. Studying Transient Components of the BOLD Response. Functional magnetic resonance imaging (fMRI) is typically based on studying the vascular response in the brain to neuronal activity. Neural activity leads to an increase in both the cerebral metabolic rate for oxygen ($CMRO_2$) and the supply of oxygen via the cerebral blood flow (CBF). The positive BOLD response is believed to be the result of a transient uncoupling between $CMRO_2$ and the supply increase, causing a reduction in paramagnetic deoxyhemoglobin in the capillaries and venules. Most fMRI methods use the positive BOLD response to study the underlying neural activity. However, this approach is limited by the sluggish nature of the underlying evoked hemodynamic response to a neural event, which peaks 5–8 s after that neural activity has peaked. Several studies have shown that CMRO2 increases more rapidly than CBF in the time immediately following neural activity, giving rise to a decrease in the BOLD signal in the first 1–2 s following activation, called the initial negative dip. The amplitude of the dip is smaller than that of the positive BOLD signal, and there is also evidence that it is more localized to areas of neural activity (Yacoub et al.,2001; Duong et al., 2000; Kim et al., 2000; Thompson et al., 2004). Due in part to these reasons, the negative response is difficult to detect, but if observable it would offer the potential to more accurately track rapid neural events as it occurs in a time scale closer to the neural activity. In order to accurately detect the dip increases are required both in temporal resolution and SNR. However, the two are intrinsically linked as rapid imaging allows for efficient modeling of physiological artifacts, which in turn increases

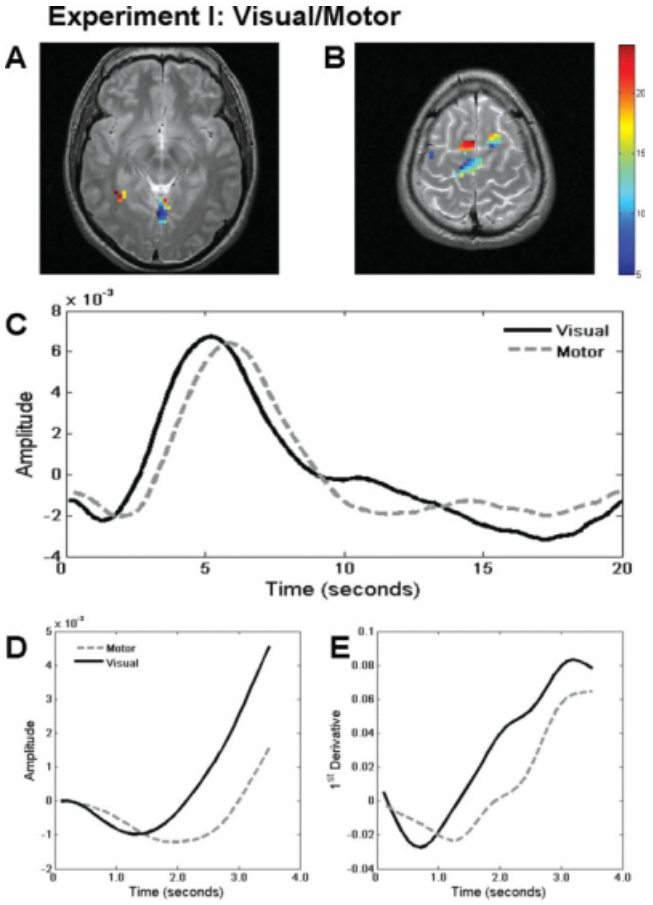

Figure 4. (A-B) Maps of the time-to-dip in voxels with significant initial negative BOLD response over two slices in the visual-motor experiment. A slice containing the visual cortex is shown to the left and one showing the motor cortex to the right. (C) The estimated hemodynamic response over the visual (bold) and motor (dashed) cortices. (D) The first 3 s of the response shows a dip appearing in the visual cortex prior to the motor cortex. (E) The rate of change in the visual cortex exceeds that in the motor cortex during the first second following stimulation. This implies that activation is taking place in the visual prior to the motor cortex as expected. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

SNR. In addition, without removal of these physiological artifacts, our experience is that it is difficult to detect the initial dip even at high temporal resolutions.

We perform statistical analysis voxel-wise using a multi-step procedure. We begin by removing physiological artifacts due to heart-rate and respiration. Next, we use the general linear model (GLM) approach (Worsley and Friston, 1995) to detect regions in the brain with significant positive BOLD response. These regions would typically be categorized as having task-induced neuronal activation in a standard fMRI analysis. We will ultimately be more concerned with detecting regions with significant negative BOLD signal. However, we make the assumption that regions having a significant dip will also have a subsequent positive rise in BOLD signal if they are involved in the task. Therefore, this step works as a screening process to remove voxels showing no signs of task-induced activation, and allows for a more data-intensive inspection of voxels of interest. Next, a statistical test based on the bootstrap

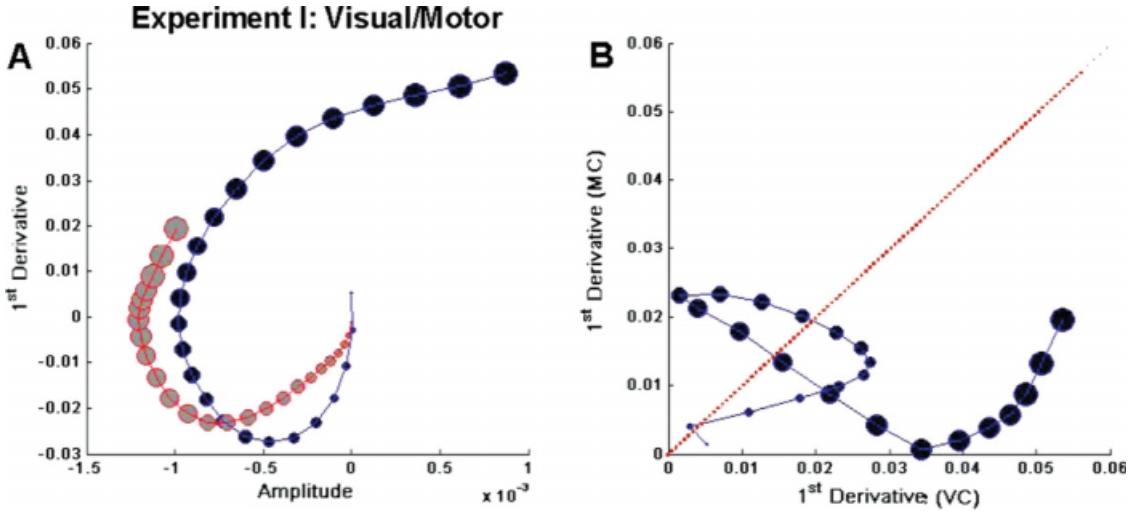

Figure 5. (A) The amplitude of the response plotted against the rate of change is shown in the visual (black) and motor (gray) cortices for the first 3 s after stimulation. The size of the data points increase as a function of time, indicating their temporal ordering. (B) The absolute rate of change in the visual cortex is plotted against the motor cortex. Plots below the dotted line show greater absolute change in the visual cortex, while points above show greater change in the motor cortex. These plots together indicate a greater absolute rate of change in the visual cortex compared to the motor cortex in the first second after visual stimulus, as would be expected by the task design. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

method (Efron and Tibshirani, 1998) is performed to determine whether the dips are statistically significant. Using re-sampling methods a bootstrap distribution is calculated for both the peak amplitude of the negative BOLD response and the time-to-dip. This is done by taking a sample of size m with replacement from the m repetitions of the stimuli to create a new time course of the same length as the original. The time course is averaged over the m repetitions, and the amplitude and time-to-peak is recorded. This procedure is repeated a large number (e.g., 1000) times to create a bootstrap distribution for the amplitude, as well as the time-to-dip. Using the bootstrap distribution for amplitude, tests can be performed to determine whether the amplitude is significantly different from zero. For voxels with significant dips, a bootstrap distribution for the pair-wise difference in time-to-dip is used to determine whether the time-to-dip was significantly different between voxels across the brain. This allows us to determine the order in which the dip occurs across different regions associated with the task. Correction for multiple comparisons is performed using an FDR controlling procedure (Benjamini and Hochberg, 1995).

As a final step, we model the hemodynamic response in active brain regions. We use a basis set approach that allows us to model the response as a continuous function; an idea commonly used in the emerging field of functional data analysis (FDA) (Ramsay and Silverman, 2005). The benefit of this approach is that it simultaneously provides a model for the derivatives of the response. The first derivative of the hemodynamic response represents the *change* in oxygenation induced by a stimulus as a function of time. We believe this is ultimately a more interesting metric then studying the level of oxygenation as we want to understand changes in oxygenation across the brain in the time immediately following activation. We further explore different ways of presenting the results. One approach is to plot the derivative against the original function. We believe this can be informative as the differentiation may expose effects that are difficult to spot by simply looking at the original function.

V. EXPERIMENTAL DESIGN

A pilot study consisting of two separate experimental paradigms was performed to confirm the feasibility and efficiency of our approach. The first paradigm consisted of fifteen cycles of 20 s periods. At the beginning of each period a 100 ms light flash was presented. The subject was instructed to press a button with their right thumb immediately after sensing the flash, thereby leading to activation of the motor cortex. During the 20 s time period, images were acquired rapidly every 100 ms using our cubic EVI trajectory. The sequence was repeated fifteen times, each replication producing a dynamic data set of 3D images at 200 temporal points. The second activation paradigm also consisted of fifteen cycles of 20 s periods. At the beginning of each period a tone was sounded through headphones which the subject was wearing. The subject was instructed to press a button with their right thumb immediately after hearing the tone. Upon pressing the button a 100 ms light flash was presented, leading to activation of the visual cortex. Hence, in this paradigm the relative timing of the onset of activation in the visual and motor cortices is flipped compared to the first paradigm.

A healthy male volunteer participated in the study after giving informed consent in accordance with a protocol approved by the Stanford Institutional Review Board. In both paradigms the first cycle was thrown out and the resulting data consisted of 14 cycles with a total of $14 \times 200 = 2800$ time units. The data was acquired with an effective TE 30 ms, flip angle 20 degrees, field of view 200 $\times$ 200 mm^2, slice thickness 185 mm and bandwidth 250 kHz. The experiment was performed on a 3.0 T whole body GE scanner. T2-weighted FSE scans were obtained for anatomic reference (TR/TE/ETL = 3000 ms/68 ms/12, 5 mm interleaved contiguous slices, $FOV = 24$ cm, 256 $\times$ 128 matrix).

After data collection, the raw k-space data was reconstructed into images of size 64 $\times$ 64 $\times$ 64 for each of the 2800 time units, and each voxel-wise time course was analyzed for activation. First, a band-pass filter was used to remove heart-rate and respiration from each time course. Next, the GLM approach was used to

determine regions with significant positive BOLD response. The design matrix consisted of three columns corresponding to a quadratic trend model for the signal drift and two columns corresponding to the canonical HRF and its temporal derivative (Hensen et al., 2002). For each voxel deemed significant in the GLM analysis, bootstrap tests were performed to detect significant dips. Finally, the hemodynamic response was modeled over the visual and motor cortices. Because of the periodic nature of the task design the HRF was modeled using a Fourier basis set, whose period was equivalent to one trial of the stimulus (20 *s*). In order to obtain a smooth representation of the HRF we limit ourselves to using the first 40 basis functions. Once the model was fit, we calculated the first derivative and used it to study changes in oxygenation over time.

VI. RESULTS

The results of the first paradigm are shown in Figure 4. Panels A-B show significant voxels for two slices, one centered in the visual and the other in the motor cortex. The time-to-dip was estimated for each active voxel and results are superimposed onto the maps. Figure 4C shows the estimated HRF for the visual and motor cortices, respectively. We clearly see that the HRF estimated from the visual cortex proceeds the one from the motor cortex throughout the course of the 20 s run, as makes sense since the visual cortex is logically the first region of the brain that begins to work on the image. Figure 4D shows a close-up of the first 3 s following activation. A negative dip appears first in the visual cortex, followed after a few hundred milliseconds delay by a negative response in the motor cortex. Studying the derivatives (Figure 4E) shows the rate of change in the time immediately following activation is substantially higher in the visual compared to the motor cortex. Hence, it appears that the dip contains valuable information about the relative timing of activity in the visual and motor cortices. It should however be noted that similar information is contained in the rise as well.

Figure 5A shows scatter plots of the amplitude of the response against its first derivative, over both the visual and motor cortices, for the first 25 time points. Figure 5B plots the absolute values of the derivative from the visual and motor cortices against each other. In both plots the radius of the points in the scatter plot increase as a function of time to illustrate the temporal aspect of the data. It is interesting to note that the absolute change in the visual cortex dominates that in the motor cortex throughout the first 2.5 s following activation. It is only in the interval between 1–1.5 *s* that the motor cortex has a larger absolute level of change, but this is caused by the fact that the response in the visual cortex changes sign a half second earlier than the motor cortex in order to begin its rise.

Figures 6–7 show results for the second paradigm. Here there appears to be confounding in the timing of the peaks in the visual and motor cortices when studying the rise. This can be seen most clearly in Figure 6C where the signal over the motor cortex appears to peak *after* the signal in the visual cortex, even though the motor cortex is logically activated first. However, studying the dip alleviates this confounding, as can be seen in Figure 5D-E. A bootstrap test finds that the difference in the time-to-dip between the visual and motor cortices is not statistically different from zero. However, this is hardly surprising as the visual stimulus appears almost immediately (30 *ms* delay) after the button press.

Figure 7 shows derivative plots for the second paradigm. It is interesting to note that they appear quite different to those presented in Figure 5 for the visual/motor paradigm. From Figure 7B it appears that the rate of change is greatest in the motor cortex for

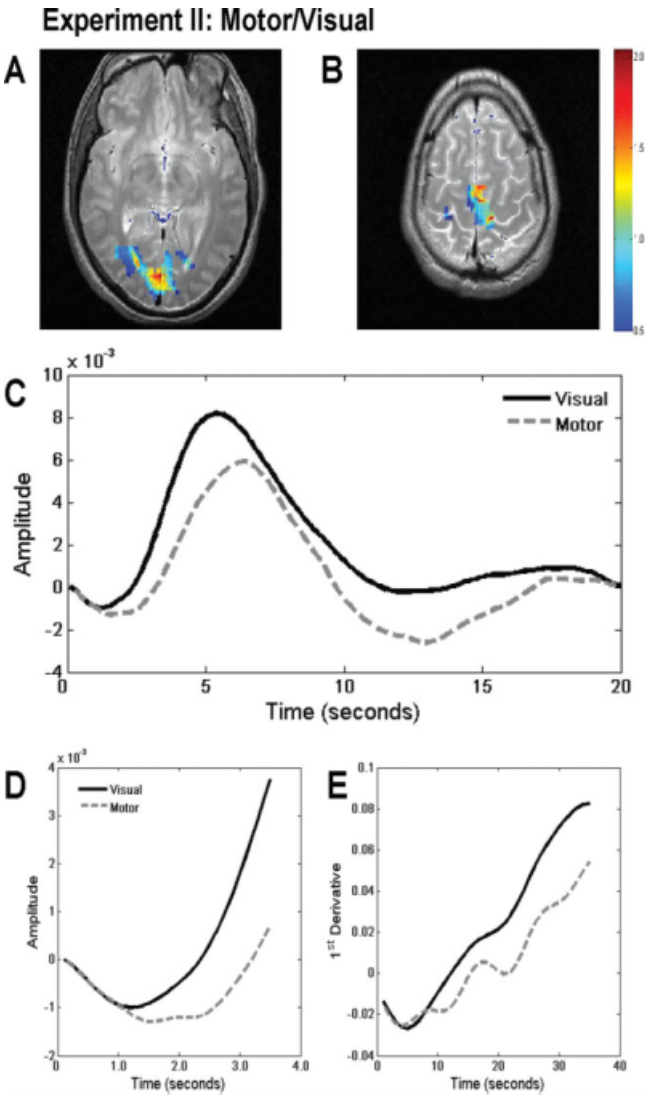

Figure 6. (A-B) Maps of the time-to-dip in voxels with significant initial negative BOLD response over two slices in the auditory-motor-visual experiment. A slice containing the visual cortex is shown to the left and one showing the motor cortex to the right. (C) The estimated hemodynamic response over the visual (bold) and motor (dashed) cortices show the peak rise occurs in the visual cortex before the motor cortex, even though the motor cortex is activated first. (D) The first 3 s following stimulation for the two time courses. (E) The rate of change for the two time courses appearing in (D). Studying the dip appears to alleviate the confounding. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

the first five time points, followed by 5 time points in which the change is greater in the visual cortex. At this point the response in the visual cortex changes direction leading to 5 additional time points where the motor cortex shows a greater change than the visual cortex. Thereafter, the rate of change in the visual cortex dominates that of the motor cortex. Hence, it appears that the first 0.5 *s* indicate a more immediate change in oxygenation level in the motor cortex compared to the visual cortex. The results imply that more accurate information regarding the temporal information may be obtained by studying the negative dip rather than the positive BOLD response as is the common practice. However additional studies are required to show reproducibility.

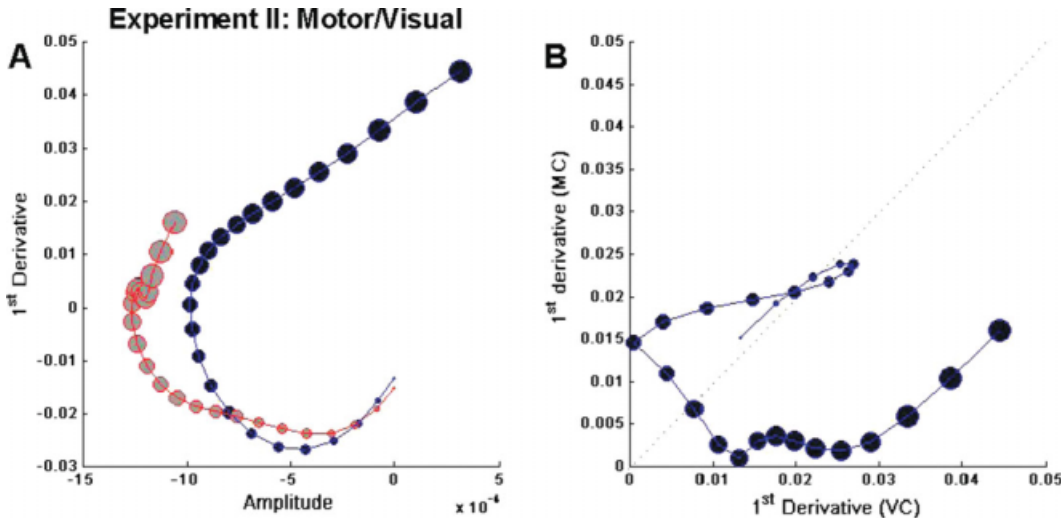

Figure 7. (A) The amplitude of the response plotted against the rate of change is shown in the visual (black) and motor (gray) cortices for the first 3 s after stimulation. The size of the data points increase as a function of time, indicating their temporal ordering. (B) The absolute rate of change in the visual cortex is plotted against the motor cortex. Plots below the dotted line show greater absolute change in the visual cortex, while points above show greater change in the motor cortex. The plots indicate a greater rate of change in the motor cortex compared to the visual cortex in the first 0.5 *s* after the auditory stimulus, as would be expected by the task design. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

Finally, we compared four different metrics for measuring the timing of activation: the onset of the dip, the time-to-peak dip, the onset of the rise and the time-to-peak rise. Each metric was manually read off from the signal over the visual cortex for each individual cycle and compared to corresponding reaction time data. The reaction time data gives a measure of the shift in the onset of the visual stimulus in each cycle and should be reflected in the various measures of latency. Figure 8 shows a scatter plot of the reaction time (in *ms*) against each of the four latency measures. The correlation was highest between time-to-dip and reaction time ($r = 0.473$). In fact, this was the only correlation that was significantly different from 0 ($p < 0.05$). This indicates that the dip appears to be providing important information about the onset of the underlying visual stimulus, especially compared to the rise.

VII. DISCUSSION AND CONCLUSION

In this paper we have attempted to highlight the benefits of performing rapid fMRI studies. The main benefit is that it allows researchers to study early transient components of the hemodynamic response related to local oxygen metabolism following activation, i.e., the initial negative dip. We do not believe this can be done efficiently in a standard fMRI experiments with a temporal resolution on the order of seconds. We show in two pilot studies that the dip contains valuable information about the timing of activation across the brain and that this information can be confounded when studying the rise. Hence, we believe our technique can provide invaluable information for inferring brain pathways.

Another benefit is that it allows for the efficient removal of physiological noise due to cardiac and respiratory effects. Since respiration gives rise to a periodic function, with a period length of $\sim$3 s, the Nyquist criteria does not allow us to model the signal in a standard low resolution fMRI experiment. Rapid fMRI circumvents this issue and allows for efficient reconstruction of the underlying signal. This is beneficial as it allows us to significantly clean up the fMRI signal prior to analysis and obtain more accurate estimates of the hemodynamic response function. In our experience, without performing this pre-processing step it is difficult to study the dip because of its relatively low SNR.

In addition, rapid imaging alleviates issues related to the fact that spatially separate regions of the brain are sampled at different times, thus negating the need for slice-time correction. It may also allow for more accurate correction of subject motion, as movement occurring during the acquisition of each individual volume will be reduced. However, these issues are not explored further at this time.

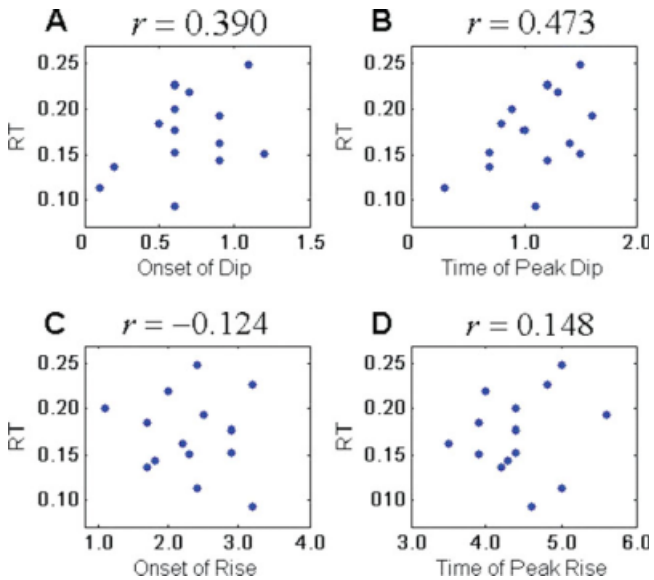

Figure 8. Scatter plots of the reaction time (RT) across the 14 runs of the second experimental paradigm against (A) the onset of the dip, (B) the time-to-peak dip, (C) the onset of the rise, and (D) the time-to-peak rise. Bootstrap tests show that only the time-to-peak dip is significantly correlated with RT ($r = 0.473$, $p < 0.05$.) [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

The main drawback to our approach is that we sacrifice spatial resolution in order to increase temporal resolution. However, advances in multi-coil techniques (Sodickson et al., 1997; Pruessmann et al., 1999) promise to bridge this gap in the future. In these techniques, multiple k-space measurements can simultaneously be made and prior knowledge about RF field distributions or the image sensitivity of the coils can be utilized to construct images from under-sampled k-space data. We have recently implemented two new trajectories that allow us to obtain images with a temporal resolution of 100 *ms* and a spatial resolution on the order of $25 \times 25 \times 17$ and $46 \times 46 \times 17$, respectively. With the latter spatial resolution we are quickly approaching the resolution used in standard fMRI experiments, but with a 10-fold increase in temporal resolution.

Finally, it is worth noting that acquiring an image with high spatial resolution and thereafter smoothing the image does not lead to the same results as directly acquiring a low resolution image. The signal-to-noise ratio during acquisition increases as the square of the voxel volume, so acquiring small voxels means that signal is lost that can never be recovered. Hence, it is optimal in terms of sensitivity to acquire images at the desired resolution and not employ smoothing. Our acquisition schemes can be designed to acquire images at the final functional resolution desired allowing for much more rapid image acquisition as time is not spent acquiring information that will be discarded in the subsequent analysis. Hence, as smoothing is often used to improve inter-subject registration and overcome limitations in the spatial normalization it would be natural to use rapid imaging techniques whenever performing group studies. We believe this should become the norm in the future.

ACKNOWLEDGMENTS

The author thanks Cun-Hui Zhang, Larry Shepp and Gary Glover for their help on this work.

REFERENCES

S. Ogawa, D. Tank, R. Menon, J. Ellerman, S. Kim, H. Merkle, and K. Ugurbil, Intrinsic signal changes accompanying sensory simulation: functional brain mapping and magnetic resonance imaging, Proceedings of the National Academy of Sciences 89 (1992), 5951–5955.

G. K. Aguirre, E. Zarahn, and M. D'Esposito, The variability of human, BOLD hemodynamic responses, NeuroImage 8 (1998), 360–369.

Z. Cho, Y. Ro, and T. Lim, Nmr venography using the susceptibility effect produced by deoxyhemoglobin, Magnetic Resonance in Medicine 28 (1992), 25–38.

T. Ernst and J. Hennig, Observation of a fast response in functional mr, Magnetic Resonance in Medicine 32 (1994), 146–149.

R. Menon, S. Ogawa, X. Hu, J. Strupp, P. Andersen, and K. Ugurbil, Bold based functional mri at 4 tesla includes a capillary bed contribution: echo-planar imaging mirrors previous optical imaging using intrinsic signals, Magnetic Resonance in Medicine 33 (1995), 453–459.

D. Malonek and A. Grinvald, The imaging spectroscopy reveals the interaction between electrical activity and cortical microcirculation: implication for optical, pet and mr functional brain imaging, Science 272 (1996), 551–554.

E. Yacoub, T. Le and X. Hu, Detecting the early response at 1.5 tesla, NeuroImage 7 (1998), S266.

E. Yacoub, A. Shmuel, J. Pfeuffer, P. Van De Moortele, G. Adriany, K. Ugurbil, and X. Hu, Investigation of the initial dip in fmri at 7 tesla, NMR in Biomedicine 14 (2001), 408–412.

T. Duong, D. Kim, K. Ugurbil, and S. Kim, Spatio-temporal dynamics of the bold fmri signals: Toward mapping columnar structures using the early negative response, Magnetic Resonance in Medicine 44 (2000), 231–242.

D. Kim, T. Duong, and S. Kim, High-resolution mapping of iso-orientation columns by fmri, Nature Neuroscience 3 (2000), 164–169.

J. Thompson, M. Peterson, and R. Freeman, High-resolution neurometabolic coupling revealed by focal activation of visual neurons, Nature Neuroscience 7 (2004), 919–920.

P. Mansfield, A. Howseman, and R. Ordidge, Volumar imaging using nmr spin echos: echo-volumar imaging (evi) at 0.1 t, Journal of Physics E 22 (1989), 324–330.

P. Mansfield, R. Coxon, and J. Hykin, Echo-volumar imaging (evi) at 3.0 t: First normal volunteer and functional imaging results, Journal of Computer Assisted Tomography 19 (1995), 847–852.

M.A. Lindquist, C. Zhang, G. Glover, and L. Shepp, Rapid three-dimensional functional magnetic resonance imaging of the negative bold response, Journal of Magnetic Resonance 191 (2008a), 100–111.

M.A. Lindquist, C.-H. Zhang, G. Glover, and L. Shepp, Acquisition and statistical analysis of rapid 3d fmri data, Statistica Sinica 18 (2008b), 1395–1419.

Q. Yang, M.A. Lindquist, L. Shepp, C.-H. Zhang, J. Wang, and M. Smith, The two dimensional prolate spheroidal wave function for mri, Journal of Magnetic Resonance 58 (2002), 43–51.

M.A. Lindquist, Optimal data acquisition in fmri using prolate spheroidal wave functions, International Journal of Imaging Systems and Technology 13 (2003), 803–812.

M.A. Lindquist, C.-H. Zhang, G. Glover, L. Shepp, and Q. Yang, A generalization of the two dimensional prolate spheroidal wave function method for non-rectilinear mri data acquisition methods, IEEE Transactions in Image Processing 15 (2006), 2792–2804.

M.A. Lindquist and T.D. Wager, Spatial smoothing in fmri using prolate spheroidal wave functions, Human Brain Mapping 29 (2007), 1276–1287.

P. Harvey and P. Mansfield, Echo-volumar imaging (evi) at 0.5 t: First whole-body volunteer studies, Magnetic Resonance in Medicine 35 (1996), 80–88.

J. Jackson, C. Meyer, D. Nishimura, and A. Macovski, Selection of a convolution function for fourier inversion using gridding [computerised tomography application], Medical Imaging, IEEE Transactions on 10 (1991), 473–478.

L. Shepp and C.-H. Zhang, Fast functional magnetic resonance imaging via prolate wavelets, Applied and Computational Harmonic Analysis 9 (2000), 99–119.

M.A. Lindquist, The statistical analysis of fmri data, Statistical Science 23 (2008), 439–464.

K. J. Worsley and K. J. Friston, Analysis of fMRI time-series revisited-again, NeuroImage 2 (1995), 173–181.

B. Efron and R. Tibshirani, An Introduction to the Bootstrap, Chapman & Hall/CRC, 1998.

Y. Benjamini and Y. Hochberg, Controlling the false discovery rate: A practical and powerful approach to multiple testing, Journal of the Royal Statistical Society, Series B. 57 (1995), 289–300.

J. Ramsay and B. Silverman, Functional Data Analysis, 2nd Edition, Springer, 2005.

R. Henson, C. Price, M. Rugg, R. Turner and K. Friston, Detecting latency differences in event-related bold responses: application to words versus nonwords and initial versus repeated face presentations, NeuroImage 15 (2002), 83–97.

D. Sodickson and W. Manning, Simultaneous acquisition of spatial harmonics (smash): fast imaging with radiofrequency coil arrays, Magnetic Resonance in Medicine 38 (1997), 591–603.

K. Pruessmann, M. Weiger, M. Scheidegger, and P. Boesiger, Sense: sensitivity encoding for fast mri, Magnetic Resonance in Medicine 42 (1999), 952–956.