Statistical Analysis of Functional MRI Data in the Wavelet Domain

Urs E. Ruttimann, *Member, IEEE*, Michael Unser,* *Senior Member, IEEE*, Robert R. Rawlings, Daniel Rio, Nick F. Ramsey, Venkata S. Mattay, Daniel W. Hommer, Joseph A. Frank, and Daniel R. Weinberger

***Abstract*— The use of the wavelet transform is explored for the detection of differences between brain functional magnetic resonance images (fMRI's) acquired under two different experimental conditions. The method benefits from the fact that a smooth and spatially localized signal can be represented by a small set of localized wavelet coefficients, while the power of white noise is uniformly spread throughout the wavelet space. Hence, a statistical procedure is developed that uses the imposed decomposition orthogonality to locate wavelet-space partitions with large signal-to-noise ratio (SNR), and subsequently restricts the testing for significant wavelet coefficients to these partitions. This results in a higher SNR and a smaller number of statistical tests, yielding a lower detection threshold compared to spatial-domain testing and, thus, a higher detection sensitivity without increasing type I errors. The multiresolution approach of the wavelet method is particularly suited to applications where the signal bandwidth and/or the characteristics of an imaging modality cannot be well specified. The proposed method was applied to compare two different fMRI acquisition modalities. Differences of the respective useful signal bandwidths could be clearly demonstrated; the estimated signal, due to the smoothness of the wavelet representation, yielded more compact regions of neuroactivity than standard spatial-domain testing.**

***Index Terms*— Functional magnetic resonance imaging, multiresolution analysis, statistical models, wavelet transform.**

I. Introduction

OVER the past few years, there have been numerous reports on the use of wavelets in medical imaging [1], [2]. One of the more traditional applications is data compression of image sequences and/or image volumes [3]–[5]. Wavelets have also been found to be useful for tomographic image reconstruction [6], [7], and in the design of new acquisition schemes for magnetic resonance imaging (MRI) [8]–[11]. Wavelet representations are also well suited for a variety of data processing tasks. For example, one of the earlier papers presented a simple threshold-based denoising method for MRI images [12], which was the precursor of Donoho's wavelet shrinkage method [13]. Perhaps most of the efforts in this area have been directed toward applying wavelets to digital mammography, for both image enhancement [14] and the detection of microcalcifications [15]–[17]. In Section II, a brief review of the basic concepts of the wavelet transform is presented, with emphasis given to the selection of basis functions, and to implementation issues regarding multiple dimensions and discrete realizations.

The present application of wavelets to the analysis of functional MRI (fMRI) data was inspired by our earlier work with functional image data obtained by positron emission tomography [18]. In functional imaging, the information of clinical interest is usually the difference between images of two different activation states of the brain controlled by some experimental paradigm. Because the images have a poor signal-to-noise ratio (SNR), arising from intrinsic biological heterogeneity and scanner-induced noise, averaging over several experimental trials (or subjects) is usually performed, yielding a mean difference image and its associated sample standard deviation (SD) image. The direct statistical analysis of these data in the spatial domain is problematic because of a poor SNR, the large number of pixels (i.e., test variables) that need to be investigated, and the often unknown and strong spatial correlation among the pixels. In order to mitigate some of these problems, statistical analysis in the wavelet domain, as described in Section III, is proposed. The motivation for this approach is derived from the fact that wavelets are efficient for the representation of a wide variety of signals. Consequently, if the signal to be detected is spatially relatively localized, it can be represented by a small number of strong local coefficients, while the power of white noise is uniformly spread throughout the wavelet space. It then follows that improved localized SNR conditions may arise in certain wavelet-space partitions, which can be exploited by the use of an orthogonal wavelet decomposition. Hence, our rationale was to locate such favorable space partitions and restrict the statistical testing for significant coefficients to these partitions. The benefits expected from such a strategy are improved SNR conditions, and a decrease of the detection threshold due to

Manuscript received March 24, 1997; revised November 18, 1997. The Associate Editor responsible for coordinating the review of this paper and recommending its publication was X. Hu. *Asterisk indicates corresponding author.*

U. E. Ruttimann deceased, was with the Laboratory of Clinical Studies, National Institute on Alcohol Abuse and Alcoholism, National Institutes of Health, Bethesda, MD 20892-1256 USA.

*M. Unser is with the Biomedical Imaging Group, Swiss Federal Institute of Technology, EPFL DMT/IOA, 1015 Lausanne, Switzerland (e-mail: michael.unser@epfl.ch).

R. R. Rawlings, D. Rio, and D. W. Hommer are with the Laboratory of Clinical Studies, National Institute on Alcohol Abuse and Alcoholism, National Institutes of Health, Bethesda, MD 20892-1256 USA.

N. F. Ramsey, V. S. Mattay, and D. R. Weinberger are with the Clinical Brain Disorders Branch, National Institute of Mental Health, National Institutes of Health, Washington, DC 20032 USA.

J. A. Frank is with the Laboratory of Diagnostic Radiology Research, Office of Intramural Research, National Institutes of Health, Bethesda, MD 20892 USA.

Publisher Item Identifier S 0278-0062(98)03010-9.

0278–0062/98$10.00 © 1998 IEEE

the reduced number of statistical tests, both contributing to a higher detection sensitivity without increasing type I errors.

Functional Magnetic Resonance Imaging (fMRI)

Functional MRI is a fast-developing technique for studying physiological processes in the brains of conscious human subjects. It measures blood oxygenation-level dependent (BOLD) signal changes caused by regional hemodynamic adjustments in response to changes in neuronal activity [19]. The current understanding is that an increase in local brain activity leads to hyperoxemia (decreased level of deoxyhemoglobin), which is due to an overcompensation of the local tissue perfusion in response to the increased energy demand in the activated neurons [20]. However, the exact mechanisms of these regulatory processes are not fully known. In particular, the interpretation for BOLD signal decreases is not yet established. Nonetheless, BOLD methods offer considerable advantages over other functional imaging modalities, in that they can be performed on widely available clinical scanners, do not require exogenous contrast agents or exposure to ionizing radiation, provide excellent spatial resolution, and can be registered with anatomical images acquired on the same machine.

Although the first reports of imaging in humans based on BOLD effects appeared in 1992 [21]–[23], important issues regarding sensitivity, reproducibility, and the nature of artifacts are still unsettled. Neuronal activity changes induced by various experimental stimuli typically result in signal intensity changes of 1%–5% in 1.5-T scanners, which are close to the scan-to-scan variability. While the SNR can often be improved by stimulus repetition with subsequent scan averaging, there is a practical limit to the number of scans that can be collected in a single human subject. Changes in physiological processes as a result of habituation, learning or fatigue, subject motion, and machine calibration drift, impose time constraints on the duration of an experiment. Consequently, scan acquisition time is an important factor. A major problem of BOLD methods is the presence of artifacts associated with head and/or vessel motion [24], as well as vascular inflow [25], [26] and drainage effects [27]. Most detrimental is that these artifacts are quite often correlated with the signal of interest, and thus resistant to simple image averaging. Hence, in attempts to overcome some of these technical difficulties, new image acquisition schemes are rapidly evolving where, unfortunately, formal performance comparisons are often missing (abstracts for the 1995 Annual Meeting of the Society for Magnetic Resonance yielded a list of 76 acronyms for different MRI acquisition techniques [28]).

Considering the rate at which new fMRI acquisition techniques emerge, the availability of objective, automated image comparison methods should be of great interest. The wavelet decomposition approach is eminently suited for such comparison tasks because it can be applied "blindly" to images acquired by different techniques, where knowledge of possible signal bandwidth limitations may not have been sufficiently well established [29]. In particular, no method-specific preprocessing is needed, and instead of requiring assumptions regarding signal bandwidth limitations, crude, yet unbiased estimates of the effective bandwidth are obtained as a byproduct of the statistical analysis in the wavelet domain.

The paper is organized as follows. Section II introduces a few basic concepts of wavelet transforms, presents extensions to multiple dimensions, and provides important detail for their digital implementation. In Section III we introduce the statistical model and develop inferential testing procedures for the detection of activation signals. Section IV details fMRI acquisition procedures and experimental neuroactivation paradigms. As an application of the developed methods, the performance of two particular fMRI acquisition techniques, PRESTO [30] and EPI [31] are compared in Section V, and conclusions are drawn in Section VI.

II. The Wavelet Transform

The usefulness of orthogonal transformations, such as the Fourier or orthogonal wavelet transforms, is that they project a signal onto a set of basis functions without altering the signal itself. If a key feature of a signal is well represented by a few basis functions in the set, while all other basis functions are orthogonal, then the presence of that feature can be easily detected by projecting the signal onto this orthogonal basis. However, in many applications, such as fMRI, the key signal features are not well known and the optimal basis functions cannot be specified in advance. For such applications, the multiresolution analysis made possible by the wavelet transform offers the benefit that signal features not known beforehand can be detected and extracted over many scales. A fusion of the extracted features at different scales then permits a succinct signal representation with emphasis on key signal properties.

A. Multiresolution and Wavelet Decomposition of a One-Dimensional (1-D) Signal

For simplicity of the exposition, the 1-D case is considered first, which can be easily extended to multiple dimensions. An orthogonal wavelet transform is characterized by two continuously-defined functions: 1) the *scaling function* $\phi(x)$, and 2) its associated *wavelet* $\psi(x) = \sqrt{2}\sum_{k\in Z} g(k)\phi(2x-k)$, where $g(k)$ is a suitable weighting sequence. The scaling function ϕ is the solution of a two-scale equation

$$\phi(x) = \sqrt{2}\sum_{k\in Z} h(k)\phi(2x-k). \tag{1}$$

The sequence $h(k)$ is the so-called *refinement* filter. The wavelet basis functions are constructed by dyadic dilation (index j) and translation (index k) of the mother wavelet

$$\psi_{j,k} = 2^{-j/2}\psi(x/2^{-j} - k). \tag{2}$$

The sequences h and g—or, equivalently ϕ and ψ—can be selected such that $\{\psi_{j,k}\}_{(j,k)\in Z^2}$ constitutes an orthonormal basis of L_2, the space of finite energy functions [32], [33]. This orthogonality permits the wavelet coefficients $d_j(k)$ and approximation coefficients $c_j(k)$ of any function $f(x) \in L_2$ to be obtained by inner product with the corresponding basis functions

$$d_j(k) = \langle f, \psi_{j,k}\rangle, \quad c_j(k) = \langle f, \phi_{j,k}\rangle \tag{3}$$

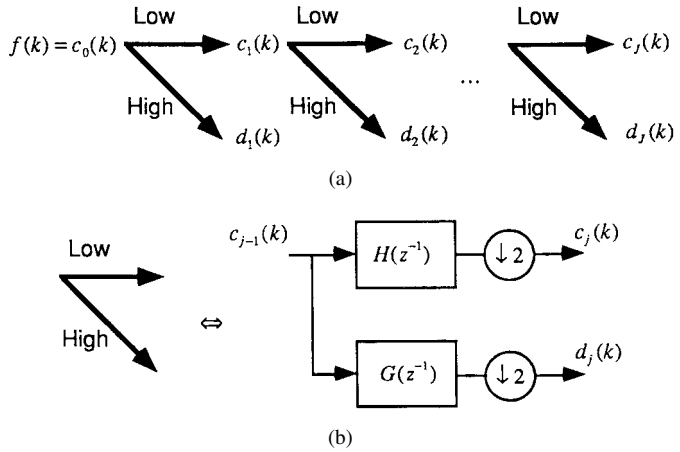

Fig. 1. Fast implementation of the wavelet transform. (a) Basic principle of the algorithm is the repetitive split of the sequences $c_{j-1}(k)$ into two halves using the low and high operators. (b) Implementation of the low (low-pass) and high (high-pass) operators using filtering and decimation by a factor of two.

where $\langle f, g\rangle = \int f(x)g(x)\,dx$ is the conventional L_2-inner product. In practice, the decomposition is only carried out over a finite number of scales J. The wavelet transform with a depth J is then given by

$$f(x) = \sum_{j=1}^{J}\sum_{k\in Z} d_j(k)\psi_{j,k} + \sum_{k\in Z} c_J(k)\phi_{J,k} \tag{4}$$

where $d_j(k)$ and $c_j(k)$ are defined in (3).

Although the synthesis and expansion formulas (4) and (3) are usually given for continuous signals [32], [33], equivalent expressions also exist for a purely discrete framework [34]. In the discrete formulation, which is the appropriate one here, these formulas can be rewritten in the following matrix form:

$$\mathbf{f} = \mathbf{W}^T\mathbf{d} \tag{5}$$

$$\mathbf{d} = \mathbf{W}\mathbf{f} \tag{6}$$

where $\mathbf{f} = (\cdots, f(k), \cdots)$ is the (infinite dimensional) signal (or image) vector, $\mathbf{W}$ the orthogonal wavelet transformation matrix, and $\mathbf{d} = (\cdots, d_1(k), \cdots, d_J(k), \cdots c_J(k), \cdots)$ the wavelet coefficient vector. The wavelet transform (6) is therefore an orthonormal transformation of the signal vector $\mathbf{f}$.

Rather than defining the transform matrix $\mathbf{W}$ explicitly, it is much easier to describe the underlying decomposition algorithm, which uses the two complementary filters h and g. In the orthogonal case, the low-pass filter h satisfies the so-called quadrature mirror filter (QMF) conditions

$$H(z)H(z^{-1}) + H(-z)H(-z^{-1}) = 2 \tag{7}$$

$$H(1) = \sqrt{2} \Leftrightarrow H(-1) = 0 \tag{8}$$

where $H(z)$ is the transfer function (z-transform) of h. The high-pass filter g is the modulated version of h given by

$$G(z) = z \cdot H(-z^{-1}). \tag{9}$$

The wavelet decomposition is implemented iteratively as in Fig. 1(a), by successive filtering and decimation using the QMF filterbank in Fig. 1(b). The iterative definition of $\mathbf{W}$ enables implementation of (5) and (6) for a signal vector of length N_0 by $O(N_0)$ operations, rather than $O(N_0^2)$. This makes the computation of the wavelet transform slightly more efficient than that of the standard FFT, which has complexity of $O(N_0 \log N_0)$.

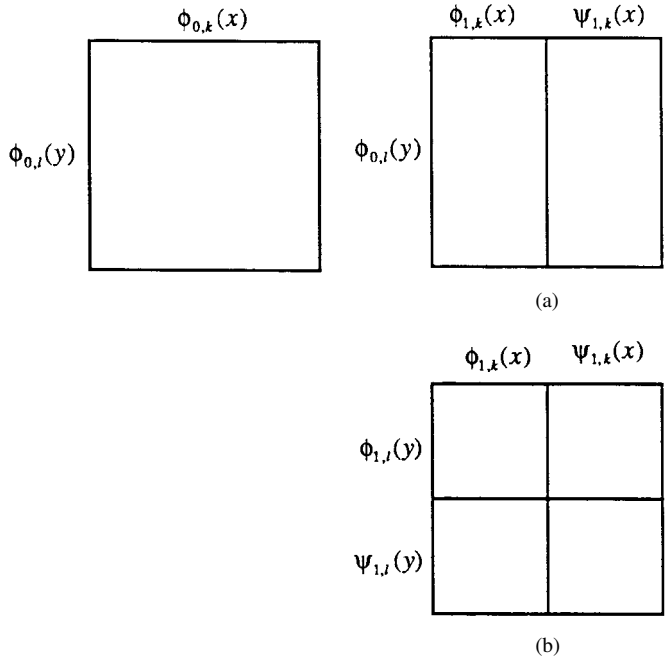

Fig. 2. (a) One iteration of the separable wavelet transform in 2-D. First, the basic 1-D algorithm is applied in the x-direction, which splits the columns of the data into two halves. Second, it is applied in the y-direction with (a) as input, splitting the rows into two halves. (b) The basis functions for each quadrant are obtained from the product of the corresponding basis functions in x and y. The procedure is then iterated on the upper left quadrant in (b).

B. Extension to Multiple Dimensions

The decomposition (3) is easily extended to two-dimensional (2-D) or three-dimensional (3-D) by using tensor product basis functions, which amounts to applying the 1-D decomposition algorithm [Fig. 1(b)] successively along the separate dimensions of the data. The effect of one iteration of this splitting process is illustrated in Fig. 2 for the 2-D case. In this way, one generates 2^q different types of basis functions in q dimensions. The corresponding qD separable scaling functions with $\mathbf{x} = (x_1, \cdots, x_q)$ are given by

$$\phi_{j,\mathbf{k}}(\mathbf{x}) = \prod_{i=1}^{q} \phi_{j,k_i}(x_i) \tag{10}$$

where we use the vector integer index $\mathbf{k} = (k_1, \cdots, k_q)$. The other $2^q - 1$ types of wavelet basis functions are obtained in a similar fashion by replacing one or several factors in (10) by a wavelet term of the form $\psi_{j,k_i}(x_i)$. Let $\mathbf{b} = \{b_1, \cdots, b_q\}$ a binary vector with $b_i = 1$ if ϕ_{j,k_i} is replaced by ψ_{j,k_i}, otherwise $b_i = 0$. By defining

$$\varphi_{j,k_i} = \begin{cases} \psi_{j,k_i}, & \text{if } b_i = 1 \\ \phi_{j,k_i}, & \text{otherwise} \end{cases} \tag{11}$$

the mixed tensor product wavelets then are

$$w^m_{j,\mathbf{k}}(\mathbf{x}) = \prod_{i=1}^{q} \varphi_{j,k_i}(x_i) \quad m = 1, \cdots, 2^q - 1 \tag{12}$$

with

$$m = \sum_{i=1}^{q} b_i 2^{i-1}. \tag{13}$$

Since ϕ is low pass and ψ high pass, the mixed tensor product wavelets will typically have a preferential spatial orientation along one (or several if $q > 2$) of the spatial directions. In this view, m assumes the role of a spatial direction indicator. For the 2-D case illustrated above, $w^m_{j,k}$ for $m = 1, 2, 3$ correspond to wavelets oriented along the vertical, horizontal, and diagonal directions. Note that by extending the definition of (12) to $m = 0$, it includes (10); i.e., $\phi_{j,k}(\mathbf{x}) = w^0_{j,k}(\mathbf{x})$.

The corresponding multidimensional coefficients

$$c_j(\mathbf{k}) = \langle f, \phi_{j,\mathbf{k}} \rangle \tag{14}$$

$$d^m_j(\mathbf{k}) = \langle f, w^m_{j,\mathbf{k}} \rangle \tag{15}$$

are in the digital implementation obtained iteratively by successive filtering and down-sampling by a factor of two as illustrated in Fig. 2 for the 2-D case.

C. Selection of Basis Functions

Spline wavelets have many desirable properties [35]. For our application, orthogonal spline (or Battle–Lemarié) wavelets [32], [36], [37] were selected for the following reasons: 1) orthogonality is required in the subsequent statistical analysis (Section III); 2) the resulting family of transforms has the advantage of using symmetrical basis functions; and 3) splines provide a simple way of reducing spectral overlap between resolution channels by increasing the degree of the spline n. Symmetric basis functions do not introduce phase distortions and thus maintain a more faithful signal localization in the wavelet domain; in particular, the location of strong wavelet coefficients due to a signal transition does not depend on the direction the signal is swept.

Small spectral overlap results in good data decorrelation [18], which is important to achieve a high detection sensitivity. The excellent decorrelation property of orthogonal spline wavelets derives from the fact that splines with degree n provide $L = n + 1$ vanishing moments. A Taylor series expansion of the wavelet transform shows that the correlation of the coefficients between scales j_1 and j_2 decays proportional to $r^{(L+3/2)}$, where r is the ratio of the scales $r = 2^{(j_2 - j_1)} \leq 1$ [18], [29]. Hence, the correlation across scale can be controlled by choosing a sufficiently high polynomial order. In the limit, as L increases indefinitely, the spline wavelets tend to the modulated sinc-wavelet, and the correlation approaches zero [38]. However, as another ramification of the uncertainty principle, correlation suppression comes at the expense of a loss of spatial localization in terms of the decay rate of the filter coefficients, requiring larger spatial support for the implementation of higher order wavelet filters. Note, since the modulated sinc-wavelet is the limiting case, the spatial decay rate of the filter coefficients for low-order spline wavelets is always better than $1/x$ (i.e., powers of $1/x$). In practice, the extent of channel decorrelation must be balanced against a loss of spatial localization, and the issue of selecting the degree of the spline depends on the assumed smoothness of the signal to be detected and is essentially a matter of compromise. Fortunately, the assumption of signal smoothness is not very restrictive because smooth wavelet bases are (asymptotically) "near-optimal" for estimating signals that may contain some points of discontinuity, but are otherwise largely smooth [39], [40]. Higher order wavelets will yield higher detection sensitivity in areas of smooth signal changes, but also will give rise to "ringing" artifacts near discontinuities when the signal is reconstructed from a reduced number of coefficients, though much less than a suppression of Fourier coefficients would cause in a Fourier-base representation of the same signal [41]. Conversely, the representation of signal discontinuities with a reduced number of zero-order wavelets will be more faithful, at a loss of sensitivity in detecting smooth signal changes. Since the fMRI volumes in our application were spatially coregistered employing tricubic spline interpolation (Section IV-C), orthogonal cubic-spline wavelets were used for the statistical analyses.

D. Periodic Implementations

A simplified form for the transfer function of the low-pass refinement filter (Fig. 1) for an orthogonal spline of degree n is (cf. [42])

$$H^n(e^{j\omega}) = \sqrt{2}\cos^{n+1}(\omega/2)\sqrt{\frac{B^{2n+1}(e^{j\omega})}{B^{2n+1}(e^{j2\omega})}} \tag{16}$$

where $B^{2n+1}(e^{j\omega})$ is the Fourier transform of the discrete B-spline of degree $2n + 1$. Recursive formulas for computing $B^{2n+1}(z)$ can be found in [43]. For example, for piecewise linear splines $B^3(e^{j\omega}) \propto 2 + \cos(\omega)$ and for piecewise cubic splines $B^5(e^{j\omega}) \propto 1208 + 1191\cos(\omega) + 120\cos(2\omega) + \cos(3\omega)$. An expedient way to obtain these filter coefficients is to sample $H^n(e^{j\omega})$ at the frequencies $\omega_k = 2\pi k/M, k = 0, \cdots, M - 1$, and perform an M-point inverse FFT where M is chosen sufficiently large to avoid aliasing in the time domain. This yields symmetric low-pass filter coefficients $h(k)$, which are then truncated to $k \in -N, \cdots, +N$, with $N < M$, to yield finite impulse response filters of length $2N + 1$. The associated high-pass filter response is obtained from (9) by simply applying a shift of one to the low-pass filter coefficients and alternating the sign

$$g(k) = (-1)^k h(1 - k). \tag{17}$$

The function $f(x)$ is in practice a discrete sequence of finite length $N_0, \{f(k)\}_{k \in I_{N_0}}$, where $I_{N_0} = \{0, \cdots, N_0 - 1\}$. The inner products in (3) then take the form of circular convolutions; i.e.,

$$d(k) = \sum_{n=-N}^{N} g(n) f_p(k - n)$$

and

$$c(k) = \sum_{n=-N}^{N} h(n) f_p(k-n) \tag{18}$$

where $f_p(k) = f(k \bmod N_0)$ supplies the samples at the boundaries whenever $(k-n) \notin I_{N_0}$.

III. Statistical Analysis

Typically, statistical analysis of brain fMR images compares data from one subject acquired under two different experimental conditions, where the experiment has been replicated N times. In our particular application, each subject performed $N = 8$ on/off blocks of a finger tapping paradigm (Section IV-B). Eight block difference images were calculated by subtracting the means of the images acquired within a finger tapping "on" block, $g_i^{(1)}(\mathbf{n})$, from the means of the images acquired within the preceding "off" blocks, $g_i^{(0)}(\mathbf{n})$, yielding

$$g_i(\mathbf{n}) = g_i^{(1)}(\mathbf{n}) - g_i^{(0)}(\mathbf{n}), \qquad i = 1, \cdots, N \tag{19}$$

where $\mathbf{n} \in \mathbf{Z}^q$ is the equidistant sampling grid on $\mathbf{x}$. For fMR images, the number of dimensions q is typically two or three. The difference images are assumed to be characterized by the population model

$$g_i(\mathbf{n}) = f(\mathbf{n}) + e_i(\mathbf{n}) \tag{20}$$

where $f(\mathbf{n})$ is the unknown deterministic signal common to all replications, which we would like to recover, and e_i is a homogeneous random field of identically and independently distributed Gaussian noise, $iid \sim N(0, \sigma^2)$. Averaging over N uncorrelated replications improves the SNR and yields an estimate of the signal

$$\hat{f}(\mathbf{n}) = \frac{1}{N} \sum_{i=1}^{N} g_i(\mathbf{n}) \tag{21}$$

which is asymptotically $(N \to \infty)$ unbiased and consistent; i.e.,

$$E[\hat{f}(\mathbf{n})] = f(\mathbf{n}), \quad \mathrm{Var}[\hat{f}(\mathbf{n})] = \sigma^2/N. \tag{22}$$

The estimated sample variance at each pixel location $\mathbf{n}$

$$\hat{\Sigma}^2(\mathbf{n}) = \frac{1}{N-1} \sum_{i=1}^{N} [g_i(\mathbf{n}) - \hat{f}(\mathbf{n})]^2 \tag{23}$$

has $(N-1)$ degrees of freedom (DOF's). The use of (22) and (23) requires the assumption of uncorrelated replications of the random field, which was sufficiently well satisfied for the mean block images $g_i^{(1)}(\mathbf{n})$ and $g_i^{(0)}(\mathbf{n})$, because of the relatively long duration of the block cycle (Section IV-B). This assumption would definitely be violated if the volumetric images at each time point in the acquisition series were considered as replications, due to their high temporal correlation. If a time-series variance were to be used for statistical inference in the spatial domain, either specific time-series analysis methods that properly adjust for temporal correlation needed to be applied [44], or the effective DOF's for the residual variance in a general linear model incorporating the hemodynamic response function needed to be estimated [45].

Since e_i is homogeneous by assumption, an improved variance estimate can be obtained by pooling $\hat{\Sigma}^2(\mathbf{n})$ over all intracranial (**IC**) pixels, $n_{\mathrm{pix}} = \#\mathbf{IC}$, yielding an approximation of σ^2 with very large DOF [46]

$$\sigma^2 = \frac{1}{n_{\mathrm{pix}}} \sum_{\mathbf{n} \in \mathbf{IC}} \hat{\Sigma}^2(\mathbf{n}). \tag{24}$$

The recovery procedure of $f(\mathbf{n})$ is then cast within the framework of hypothesis testing, a well established practice in the life sciences. The null hypothesis H_0 postulates $f(\mathbf{n}) = 0$; i.e., there is no systematic difference between the images acquired under the two different experimental conditions. If the hypothesis is refuted by the data, then the inference is that the signal is nonzero at certain spatial locations. In that case, it is of particular interest to the neuroscientist to obtain both a good estimate of the spatial locations and shape of the signal at these locations.

An attractive method to apply this statistical framework is to assume that the images in question can be approximated by a continuous random field, where the pixel values are considered to be the realizations of a random field sampled on an equally spaced lattice $\mathbf{n}$ [46]. The relevant test statistics are then evaluated at each pixel and searched for local extrema that might indicate the presence of an activation signal. Formal statistical methods have been developed by Worsley *et al.* [46] to guard against false positive detections, who provided explicit expressions for the probability of excursion sets of Gaussian, t-, χ^2-, and F-fields. While elegant, potential drawbacks of the random field methods are that (a) a smoothness parameter, usually the full-width–half-maximum (FWHM) of the pointspread function of the imaging method, is required to be either known or imposed upon by filtering, and (b) the images be sufficiently finely sampled (FWHM/pixel size >2, see Section IV-A). Unfortunately, condition (b) does not hold for currently produced fMR images, and with regard to (a), the "proper" amount of further smoothing to be applied is often inextricably related to the research question itself, as is the case in our application. Since random field methods cannot be used in our context (at least not without prior image smoothing), the results of the wavelet-based analysis developed below will be compared to results obtained by a spatial-domain analysis, where the false positive detection rate (i.e., the significance level) is controlled by the Bonferroni correction. While this correction is somewhat conservative in the presence of spatial noise correlations (which are presumed to be relatively small due to the undersampling existing in present fMRI's), the method is certainly valid and, like the wavelet method, requires no presmoothing.

The multiresolution strategy offered by the wavelet decomposition contrasts from the random field methods, or other "traditional" signal analysis methods, in the way the question of "appropriate" image smoothing is approached. These methods consider the shape and/or size of the brain activation regions as known and typically apply low-pass filtering to the images, in an attempt to maximize the SNR.

The question then remains, what is the "best" smoothing filter to use? If the activations are highly focal, then only a little smoothing would be best. Conversely, if they are diffuse, more extensive smoothing would be appropriate. The problem is further compounded by the possibility that a particular brain stimulation task may elicit both types of activation patterns concurrently. Hence, a monoresolution strategy for these kind of applications is likely to be suboptimal. In an attempt to pick the "optimal" filter for each possible activation pattern, a proposal has been made to apply sequentially a set of spatially invariant and isotropic Gaussian low-pass filters with successively larger kernel widths, and then extend the search for activations over the 3-D location space as well as the 1-D filter scale space [47]. However, the decomposition into a set of low-pass-filtered images is both redundant and nonorthogonal. Consequently, the number of statistical tests required to locate activations is increased (search in four-dimensional (4-D) rather than 3-D space), and a higher detection threshold must be selected to protect the significance level, incurring a loss of detection sensitivity [48]. Although the idea of a scale-space search is also pursued in the wavelet-based estimation method, this search is statistically and computationally more efficient due to the lack of redundancy of the wavelet decomposition. In contrast to traditional, monoresolution signal detection techniques, wavelet detection is spatially adaptive and thus is able to deal in a direct and straightforward manner with signals that may have spatially heterogeneous smoothness properties, as well as a finite number of discontinuities. Hence, the question regarding the "best" smoothing need not be answered beforehand. In fact, answering that question is an integral part of the wavelet-based multiresolution strategy, and the answer may be complex in that different amounts of smoothing in different spatial neighborhoods may be required. It has been shown that if nothing is known regarding the smoothness (or lack thereof) of a signal, wavelets constitute a "near-universal" orthogonal basis [49]; i.e., they are, asymptotically, for a wide range of L_p-error norms and a variety of different smoothness classes, nearly optimal for recovering a signal buried in white noise.

To implement the wavelet-based detection procedure, the average difference image (21) is subjected to a multiresolution decomposition according to (14) and (15). At each level j, the wavelet coefficients $\hat{d}_j^m(\mathbf{k})$ are computed by (15), with f replaced by $\hat{f}$, and where $w_{j,k}^m$ are the qD tensor product wavelets (12) along the directions $m = 1, \cdots, 2^q - 1$. Under the null hypothesis $g_i(\mathbf{n}) = e_i(\mathbf{n})$, and $w_{j,k}^m$ perform orthonormal linear transformations on the means of N *iid* Gaussian variates with variance σ^2, resulting from (22) in the distribution of $\hat{d}_j^m(\mathbf{k})$ as *iid* $N(0, \sigma^2/N)$. Hence, standardizing the wavelet coefficients by the standard error

$$\sigma_N = \sigma/\sqrt{N} \tag{25}$$

with σ obtained from (24), yields for each of the m directional channels at the resolution level j

$$\hat{d}_j^m(\mathbf{k})/\sigma_N \sim iid\, N(0,1) \tag{26}$$

and

$$\left(\hat{d}_j^m(\mathbf{k})/\sigma_N\right)^2 \sim iid\, \chi_1^2. \tag{27}$$

Properties (26) and (27), in conjunction with the orthogonality of the decomposition permit a two-stage approach to the estimation of f, which reduces the overall number of statistical tests that need to be performed. The first stage addresses the question as to whether there is significant signal power in any of the $(2^q - 1)J$ direction-oriented resolution channels. The approximation coefficients $\hat{c}_J$ represent the extreme low-pass and dc information in the images and, as is common practice [13], are routinely left unprocessed for inclusion in the subsequent signal estimation by the inverse wavelet transform. In the second stage, only channels with significant signal power are further examined to determine the spatial location of the signal. Hence, based on (27), the sum of the squared, standardized coefficients in each channel is under the hypothesis H_0 a chi-square variate with DOF equal to the number of summation terms. This provides the rationale for the first-stage test procedure, which jointly tests in each resolution channel the significance of the coefficients. Channels where H_0 is accepted are discarded as carrying only noise, yielding the reduced coefficient set, shown in (28) at the bottom of the page, where

$$\theta_j = \sigma_N^2 \chi^2_{n_j^m;\alpha} \tag{29}$$

is the threshold at resolution j obtained from the $(1 - \alpha)$ probability cutoff of the chi-square distribution with DOF equal to the number of intracranial wavelet coefficients at level j, $n_j^m = \#\mathbf{IC}_j$. If each of the $(2^q - 1)J$ tests is performed at a significance level of $\alpha = p/(2^q - 1)J$ (Bonferroni adjustment), the overall significance per volumetric image is maintained at the specified p-level. To facilitate the follow-up testing in the second stage, the index pairs (j', m') of channels with significant power are entered into a look-up table $\mathbf{T}_\alpha$.

The second stage procedure follows from (26) as a two-sided z-test of only the coefficients in the channels determined to carry significant power

$$\tilde{\hat{d}}_j^m(\mathbf{k}) = \begin{cases} \tilde{d}_j^m(\mathbf{k}), & \text{if } \left|\tilde{d}_j^m(\mathbf{k})\right| > \tau \\ 0, & \text{otherwise,} \end{cases} \quad (j,m) \in \mathbf{T}_\alpha \tag{30}$$

where

$$\tau = \sigma_N z_{\alpha'} \tag{31}$$

is the threshold for the standardized normal variate with the significance level α' adjusted for the total number of follow-up tests performed in the wavelet search space, i.e., the total

$$\tilde{d}_j^m(\mathbf{k}) = \begin{cases} \hat{d}_j^m(\mathbf{k}) & \text{for } j = j' \quad \text{and} \quad m = m', \quad \text{if } \sum_{\mathbf{k}\in\mathbf{IC}_j} \left[\hat{d}_{j'}^{m'}(\mathbf{k})\right]^2 > \theta_j \\ 0, & \text{otherwise,} \quad j = 1, \cdots, J, \ m = 1, \cdots, 2^q - 1 \end{cases} \tag{28}$$

number of intracranial coefficients contained in the channels considered for follow-up testing

$$\alpha' = p \Big/ \sum_{j,m \in \mathbf{T}_\alpha} n_j^m. \tag{32}$$

Equation (32) constitutes the Bonferroni adjustment for multiple testing, requiring an elevation of τ relative to the single-test threshold. The associated penalty in terms of loss of statistical power (increased rate of signal misses) is ameliorated to the extent that $\sum_{(j,m \in \mathbf{T}_\alpha)} n_j^m < n_{\mathbf{pix}}$.

The final step in this signal recovery procedure consist in applying the inverse discrete wavelet transform (5) to the surviving coefficients, i.e.,

$$\mathbf{f}^* = \mathbf{W}^T \tilde{\tilde{\mathbf{d}}}. \tag{33}$$

Hence, the estimate f^* is obtained as the sum of wavelets with coefficients that exceeded a statistically determined noise threshold. This process can be viewed as a kind of adaptive noise filtering, where the filter passband is determined by the SNR levels in the various resolution channels. Frequency bands of the dyadic resolution channels are combined into the overall signal bandpass definition, depending on whether the estimated power level in a channel exceeds the noise threshold. This results in a crude estimate of the effective bandwidth of the activation signal f^*. It is obtained from the first testing stage as 2^{-j^*} times the sampling rate, with f^* the smallest j (highest resolution level) where at least one of the $2^q - 1$ directional channels carries significant signal power.

It is to be emphasized that the principle of wavelet thresholding as described above does not depend on the use of thresholds based on parametric statistical distributions. Nonparametric thresholds may alternatively, and perhaps more validly, be derived from the corresponding empirical null distributions generated by randomization procedures applied to the original images [50]. Such procedures could be generalized to establish spatially variant thresholds for the detection of activation signals in the presence of spatially inhomogeneous noise (i.e., different sample SD's at different pixel locations). While randomization methods provide "exact" p-values, nothing is known regarding the statistical power of the corresponding tests, and they impose a heavy computational burden to accumulate the null distributions. For these reasons, randomization procedures are mainly relegated, at least currently, to validate existing parametric methods, particularly if there is doubt regarding the tenability of some of the underlying assumptions.

IV. Image Acquisition and Experimental Procedures

A. Imaging Parameters

All studies were performed on a clinical 1.5-T scanner (SIGNA, General Electric Medical Systems, Milwaukee, WI). The PRESTO method used a GE quadratic head coil and standard shielded gradients with a maximum amplitude of 10 mT/m and maximum slew rate of 17 T/m/s. As previously described [30], the pulse sequence had the following parameters: $\mathrm{TR} = 24$ ms, with TE ranging from 29.6 to 40.4 ms for five consecutive gradient echoes, and a flip angle of 11°. This produced susceptibility (T_2^*-weighted) images, captured in a data matrix of $64 \times 50 \times 24$ pixels in transverse orientation. The field of view (FOV) was $240 \times 187.5 \times 90$ mm, which, after reconstruction resulted in isotropic pixels with dimension 3.75 mm, and an isotropic resolution of 4.5-mm FWHM. Because of the limitations of the gradient field homogeneity, only the middle 14 contiguous slices of the imaged volume were used for analysis. This provided a sampled volume with a total axial width of 5.25 cm that included the primary sensorimotor cortex as the area of primary interest, and extended from the top of the brain to the middle segment of the lateral gyrus.

Gradient-echo EPI scanning was performed with a combined radio-frequency (RF) and gradient insert coil capable of generating 20 mT/m with a rise time of 100 ms (Medical Advances, Milwaukee, WI). Fifty-six sagittal (3.75-mm-thick) interleaved sections were acquired (TR/TE = 5 500/60 ms, with a flip angle of 90°), with a FOV of 240×240 mm [31]. This provided full-head coverage in a data matrix of $64 \times 64 \times 56$ with 3.75-mm isotropic pixel size and an isotropic FWHM after reconstruction of 4.5 mm.

B. Subjects and Tasks

Data are presented from seven healthy (mean age 30 yr), strongly right-handed volunteers, who gave informed consent to submit to MRI scanning by both the PRESTO and the EPI technique. The research protocol was approved by the National Institute of Mental Health Human Studies Review Board.

During scan acquisition, the subjects had their eyes closed and performed a finger tapping task with their right hand. The tapping was self-paced (about 2 Hz) and consisted of touching the thumb sequentially with each one of the digits, repeating this sequence for 33 s. Eight alternating blocks of 33 s rest were followed by 33 s of finger tapping. Six fMRI scans were acquired during each block to yield a total of 96 images. The first and last image in each block of six scans were discarded after conclusion of the experiment to avoid transients related to the delay in the vascular responses to neuroactivity changes, which is in the order of a 4–6 s [22], [23]. The remaining four images in each block were considered steady state representations, and averaged within each block to yield mean "on" block images, $g_i^{(1)}(\mathbf{n})$, and mean "off" block images, $g_i^{(0)}(\mathbf{n})$, $i = 1, \cdots, 8$, used in (19).

C. Image Preprocessing

In an attempt to reduce artifacts due to small involuntary head motion, all images of each subject were registered off-line to the first image of the data set such that the squared gray-level difference summed over all brain pixels was minimized. This was achieved by applying a modified version of the Marquardt–Levenberg algorithm to a cubic-spline resolution pyramid representation of the respective volumes [51]. Since only intrasubject registration was required, the transformation parameters to achieve registration were restricted to 3-D rotations and translations.

A binary brain mask image, encompassing all intracranial pixels, was created by applying an intensity threshold to the average image computed from all registered volumes for a subject. T_2^*-weighted images display usually very little tissue contrast [30], [52], rendering their intensity histograms strongly bimodal (background or tissues). Hence, the threshold was selected as the intensity level corresponding to the "valley-point" separating the two modes in the histogram. The resulting masks were trimmed subsequently to ensure homogeneity of variance as stipulated in the model (20). For that purpose, the mean "on" block images $g_i^{(1)}(\mathbf{n})$ were subtracted from adjacent mean "off" block images $g_i^{(0)}(\mathbf{n})$, and the resulting mean and sample variance computed at each pixel within the mask. As in previous analyses [30], [52], a chi-square test with $N = 8$ DOF's (nominal Bonferroni corrected significance level $p = 0.01$ per 2-D slice) was applied to eliminate pixels not satisfying the homogeneity assumption, resulting in a trimmed intracranial mask (IC). Only data from pixel locations within this mask were tested subsequently for the presence of activation signals. Pixels eliminated this way constituted <5% of the brain volume and were located mainly along the brain boundary (subarachnoidal cerebrospinal fluid, sagittal sinuses). High variance at these locations was most likely due to motion and/or gradient artifacts, as well as classification errors in the initial generation of the brain mask due to partial volume effects. Mean and sample variance of the difference images were then recomputed based on the trimmed intracranial mask **IC**, because this mask defines the region of homogenous variance, and pixel locations outside this region should not contribute to the estimates (5) and (6) of the random field parameter σ^2.

V. Results

Most currently applied fMRI acquisition methods are based on echo planar imaging (EPI), which involves the repetitive scanning of a series of 2-D slices throughout the brain [21]–[23]. Direct 3-D scanning methods of contiguous volumes would have the potential of yielding less ambiguous registrations to anatomical MRI's, because the time skew inherent in sequential slice acquisitions is eliminated. Also, 3-D scans are less susceptible to inflow effects [53] than sequential 2-D methods. However, conventional 3-D gradient-echo methods require prohibitively long acquisition times because of the restriction that TR>TE, while TE must be long for sufficient susceptibility (T_2^*) weighting. To overcome this restriction, echo shifting methods have been developed in which the echoes are shifted beyond the next RF pulse [53], [54], permitting TR>TE and resulting in much shorter scanning times. The specific 3-D method to be compared with EPI combines echo shifting with the acquisition of multiple gradient echoes per RF excitation [30], [55] and carries the acronym PRESTO (principles of echo shifting with a train of observations).

The quality of the PRESTO and EPI images were compared by a previously used index [30], defined as $\sigma/\bar{g}$, where σ is the pooled SD of the difference images (24), and $\bar{g}$ is the average MR signal intensity, pooled over all brain pixels n_{pix}

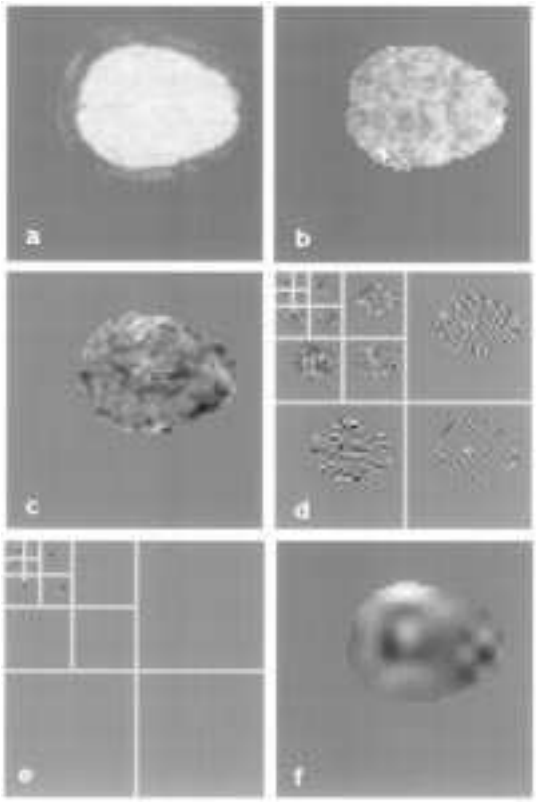

Fig. 3. PRESTO modality. (a) Average of all fMRI's, (b) sample SD of on/off difference images, (c) sample average of on/off difference images, (d) wavelet coefficients for four-level transform, (e) significant ($\alpha = 0.05$) wavelet coefficients, and (f) inverse wavelet transform of significant coefficients.

and all images

$$\bar{g} = \frac{1}{2Nn_{\text{pix}}} \sum_{i=1}^{N} \sum_{\mathbf{n}\in\mathbf{IC}} g_i^{(1)}(\mathbf{n}) + g_i^{(0)}(\mathbf{n}). \tag{34}$$

Note, the smaller this index, the better the image quality. For the seven subjects the index value ranged from 0.74% to 1.19% for the PRESTO method, and between 0.81% and 1.28% for the EPI method. This index was statistically not different between the two modalities, and its range was similar to that in a previous study (0.8%–1.2%) [30].

Examples of the wavelet-based signal detection procedure applied to a PRESTO and an EPI brain slice are shown in Figs. 3 and 4. The PRESTO method yields a transversely oriented volume from which a set of transaxial slices, 3.75-mm-thick, were selected that included the area of the primary sensorimotor cortex (23 mm from the top of the brain). These are presented in Fig. 3 with the front of the brain oriented toward the right. The EPI method provided sagittal slices, 3.75-mm-thick, located 38 mm to the left (ipsilateral to the active hand) of the midsagittal plane, with the front of the brain oriented toward the right (Fig. 4). In Fig 4(a) the average fMRI's before subtraction are shown, and illustrate the very low tissue-contrast (T_2^*) that is typically obtained in functional imaging. In Fig 4(b) and (c) the respective sample SD images (square root of $\hat{\Sigma}^2(\mathbf{n})$ in (23), for $\mathbf{n}$ within the brain mask) and the mean difference images (21) are shown. The display

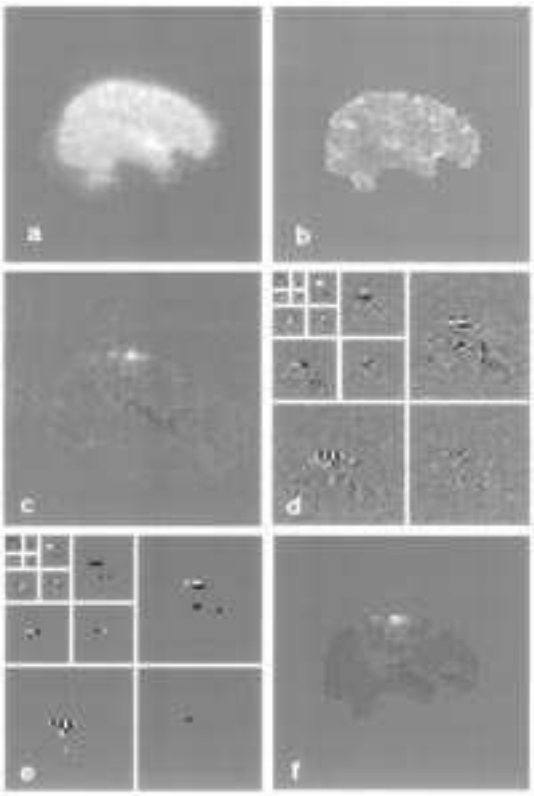

Fig. 4. EPI modality. (a) Average of all fMRI's, (b) sample SD of on/off difference images, (c) sample average of on/off difference images, (d) wavelet coefficients for four-level transform, (e) significant ($\alpha = 0.05$) wavelet coefficients, and (f) inverse wavelet transform of significant coefficients.

has been scaled to the highest positive (white) gray level. Shown in Fig. 4(d) are the 2-D wavelet coefficients before thresholding and in Fig. 4(e) after thresholding (30), for an overall significance level per volume of $\alpha = \alpha' = 0.05$. The resolution channels are organized in the usual pyramid representation, as illustrated in Fig. 2 for the first resolution level. The results for the PRESTO images are typical in that they had in general only a few significant coefficients, with none of them at the finest resolution level. Resyntheses of the difference images by inverse wavelet transform of only significant coefficients are shown in Fig. 4(f). For both modalities, a large amount of noise suppression was attained because most of the coefficients deleted were from high-resolution channels. However, since generally a larger number of coefficients from the high-resolution channels were suppressed in PRESTO images, the resulting bandwidth reduction was more severe for this modality than for EPI.

Table I summarizes the results of the statistical procedures applied to the particular images shown in Figs. 3 and 4. For PRESTO, none of the variance ratios (i.e., total channel power/noise power) at the two highest resolution levels exceeded the critical threshold (28). In particular, only five out of the total of 12 channels [see Fig. 3(e)] carried significant power and required follow-up testing by z-tests (30), (31). This resulted in five significant coefficients, or 0.55% of the total number of 904 pixels in the brain mask. In contrast, all channels for the EPI slice [see Fig. 4(e)] exceeded the noise threshold, resulting in 55 significant coefficients after follow-up testing by z-tests, or 8.36% of the total number of 658 pixels in the brain mask.

TABLE I
STATISTICAL TEST RESULTS FOR IMAGES IN FIGS. 3 AND 4

		PRESTO			EPI		
Channel[a]		VR[b]	VR_{thr}[c]	n_c[d]	VR	VR_{thr}	n_c
	V	0.45		0	5.53		9
1	H	0.75	1.36	0	5.88	1.42	14
	D	0.24		0	1.52		1
	V	1.55		0	24.69		7
2	H	2.39	1.80	0	13.53	1.94	5
	D	1.36		0	5.31		2
	V	6.02		1	28.10		4
3	H	9.74	2.75	2	15.83	3.09	4
	D	4.55		1	16.22		4
	V	6.75		0	13.98		1
4	H	32.72	8.12	1	15.93	8.95	2
	D	7.39		0	8.99		2

[a] V: vertical; H: horizontal; D: diagonal resolution channels.
[b] Variance ratio: $\mathrm{Var}(\hat{d}_j^m(\mathbf{k})/\sigma_{\mathrm{N}})$.
[c] Threshold value of VR for overall significance per volume of $\alpha = 0.05$, $VR_{\mathrm{thr}} = X^2_{n_j^m;\alpha}/n_j^m$.
[d] Number of coefficients remaining significant after follow-up test (30).

TABLE II
PERCENTAGE OF WAVELET CHANNELS WITH POWER SIGNIFICANTLY ABOVE NOISE

resolution level	PRESTO mean %	PRESTO range %	EPI mean %	EPI range %
1	0	0-0	71.4	52.4-90.5
2	25.7	9.5-38.1	72.4	47.6-90.5
3	54.3	38.1-76.2	55.7	19.0-76.2
4	38.1	4.8-57.1	49.5	19.0-85.7

Table II summarizes the accumulated results from all seven subjects regarding the test for significant channel power (28) and (29). At the highest resolution level, none of the channels in the PRESTO images were significant, while 71.4% of these channels carried significant power in EPI ($p < 0.0001$). Also at the second resolution level were the proportion of useful channels higher in EPI than in PRESTO ($p < 0.0001$), and only at the third level were the percentages for the two modalities comparable ($p = 0.769$). At the lowest resolution level, there was a slightly larger proportion of channels with significant power in EPI than in PRESTO ($p = 0.018$). These results show that, viewed from a multiresolution perspective, EPI scans contained signal features at a finer resolution than PRESTO scans. Consequently, the bandwidth limitations im-

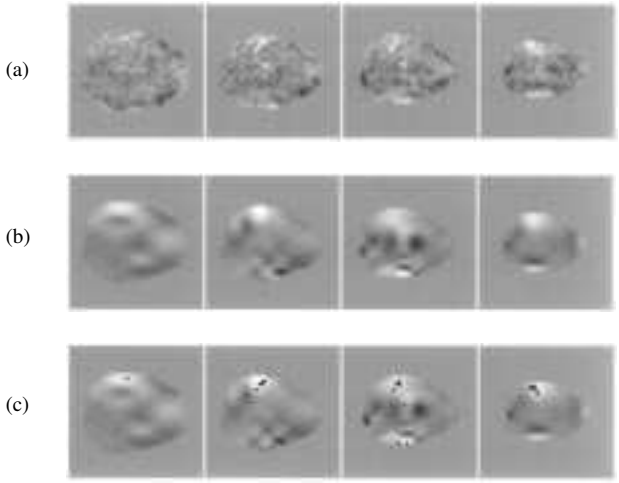

Fig. 5. Four contiguous PRESTO slices. Row (a) sample average of on/off difference images, row (b) signal response estimated by wavelet detection, and row (c) regions of significant signal response ($\alpha = 0.05$) detected by spatial-domain analysis (in black), superposed over wavelet detection results.

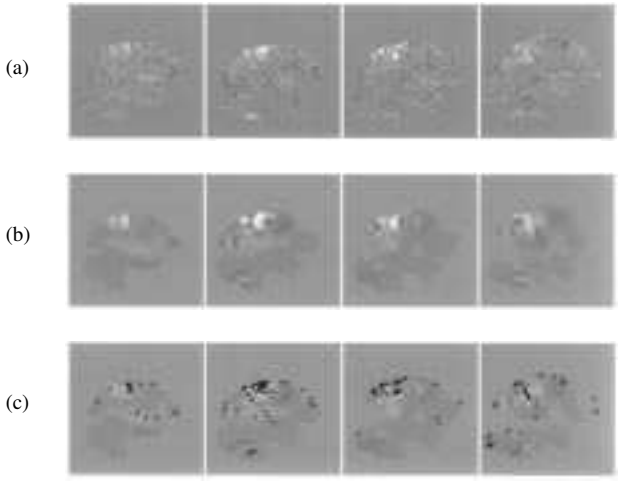

Fig. 6. Four contiguous EPI slices. Row (a) sample average of on/off difference images, row (b) signal response estimated by wavelet detection, and row (c) regions of significant signal response ($\alpha = 0.05$) detected by spatial-domain analysis (in black), superposed over wavelet detection results.

posed implicitly by resynthesizing activation images from only significant wavelet coefficients are less severe for EPI than for PRESTO. Since the finger tapping tasks were identical for both imaging modalities, one can reasonably assume the spatial smoothness of the respective activation areas to be very similar. Hence, the smoother appearance of the estimated activation signals for PRESTO relative to EPI [Figs. 3(f) and 4(f)] is for the largest part due to the lower SNR conditions in the difference images formed from the PRESTO scans.

Figs. 5 and 6 compare the signal estimates derived by wavelet analysis in a sequence of contiguous slices (columns) to the results of direct significance testing in the spatial domain. For the spatial domain analysis, the model as defined in (20) with independent, white noise at the pixel level was assumed. Rows (A) present the signal estimates (21) obtained by averaging, and rows (B) show the signal estimates resulting from the wavelet analysis (33). The significance of the spatial-domain signal estimates was established by z-tests on the ratio of the average difference image (21) to the pooled standard error (25). The Bonferroni adjustment was applied to maintain the overall significance level per volume at $\alpha = 0.05$; i.e., the p-level for the pixelwise tests was set to $p = 0.05/n_{\mathrm{pix}}$, $n_{\mathrm{pix}} = \#\mathbf{IC}$. Rows (C) show in black, superposed over the wavelet-based signal estimates, the pixels found to be significant by these z-tests. The signal estimates shown in rows (A) were at these pixel locations significantly different from zero. The wavelet-reconstructed signal estimates reflect the smoothness of the selected wavelets, resulting in regions of estimated neuroactivity that are generally larger than the discontinuous regions obtained by significance testing and thresholding in the spatial domain. This is most evident for PRESTO images. At certain locations, only one pixel exceeded the threshold for significance in the spatial-domain testing, whereas the corresponding signals reconstructed from significant wavelets indicated more diffuse areas of increased activity.

In comparison to PRESTO, the SNR in images generated by EPI (Fig. 6) was sufficiently high in the regions of the primary sensorimotor cortex, resulting in the inclusion of wavelet coefficients at the highest resolution level (see also Table II). This yielded signal estimates with the smallest amount of smoothing imposed by the wavelet representation. Therefore, the signal extent estimated from the wavelet representation overlapped largely with the signal extent derived by pixelwise statistical testing of the original (not additionally smoothed) image. However, this was only true in regions of the sensorimotor cortex where brain activation was expected to occur. Outside these regions, the spatial detection method found additional significant responses, which were mainly of single-pixel extent and most often did not coincide with wavelet resynthesized responses. The local SNR in these outlying areas was not sufficient to yield wavelet coefficients of significant strength at the highest resolution level. Whether these outlying responses are "true" or technical artifacts is beyond the scope of this work. If we assume them to be artifacts, then some local smoothing of the EPI data, as automatically implemented by the retention of only significant wavelet channels, is indeed indicated.

In order to further differentiate the signal estimates obtained by wavelet reconstruction and by spatial significance testing, a 3-D region of interest (ROI) was defined that included the left-hemispheric primary sensorimotor cortex in both modalities. This ROI contained 1 125 pixels (59.3 cm^3), of which four consecutive 2-D sections for the same slices shown in Figs. 5 and 6 are displayed in rows (a) of Fig. 7. In rows (b), the wavelet-derived signals were thresholded at a 0.5% of the average image intensity $\bar{g}$ (27). With this threshold setting slightly below the best observed scan-to-scan stability (0.74%), most of the activation signal (maximum about 5% in BOLD methods [20]) was captured. Rows (c) show pixels (white) where significant activity was detected by spatial thresholding [corresponding to black pixels of Figs. 5 and 6 (c)]. A qualitative comparison shows that in PRESTO images the thresholded signal estimates derived from significant wavelet coefficients produced larger clusters than spatial thresholding of the mean difference images, while for EPI the two results were similar.

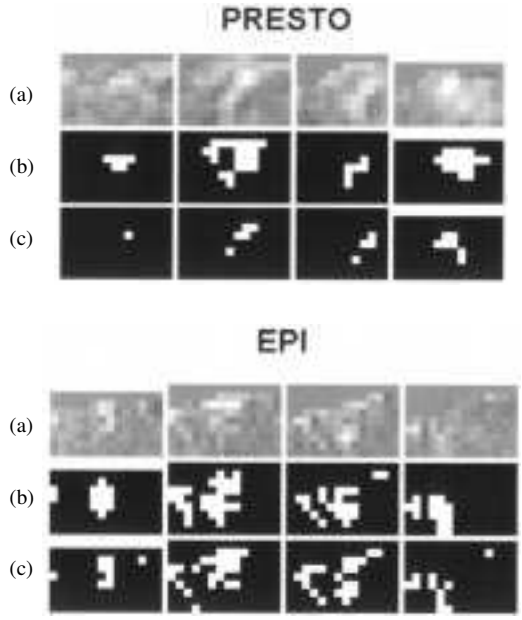

Fig. 7. Four contiguous slices from ROI of left primary sensorimotor complex for (top) PRESTO and (bottom) EPI modality (corresponding to Figs. 5 and 6). Rows (a) sample averages of on/off difference images, rows (b) wavelet-detected responses, thresholded at 0.5% of the average fMRI intensities, and rows (c) regions of significant response detected by spatial-domain analysis.

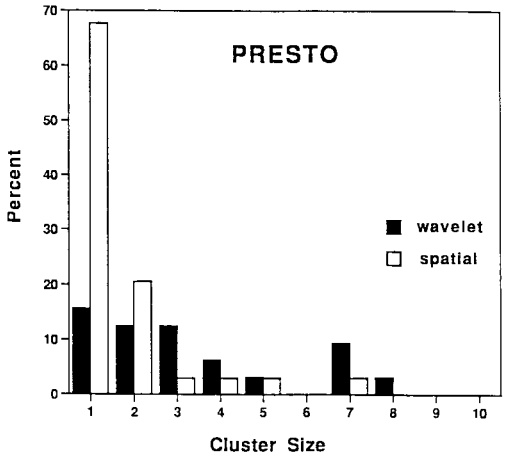

Fig. 8. Cluster-size distributions within ROI for PRESTO. Filled bars, regions detected by wavelet analysis and subsequent thresholding at 0.5% of the average fMRI intensities. Open bars, significant regions detected by spatial-domain analysis.

For a quantitative comparison of the wavelet based analysis to the spatial detection method, cluster-size distribution analyses were applied to the thresholded signal representations derived by either method. The results are summarized in Figs. 8 and 9 for cluster sizes of up to ten pixels. For PRESTO (Fig. 8), the proportions of clusters with one or two pixels obtained by the spatial detection method greatly exceeded those obtained by the wavelet detection method. Equality of the two distributions was clearly rejected ($p < 0.001$) by the Kolmogorov–Smirnov test [56], with median cluster sizes for the spatial and wavelet methods of one and six, and maximum detected cluster sizes of seven and 28, respectively. A corresponding cluster-size distribution analysis for EPI (Fig. 9) indicated a similar tendency, although less pronounced, for the spatial detection method to yield disproportionately smaller cluster sizes. The median (maximum) cluster sizes for the spatial method was 2 (32), as compared to three (34) for the wavelet method, and the Kolmogorov–Smirnov test formally rejected distribution equality ($p < 0.02$). Hence, the results of these cluster-size analyses are in concordance with the rankings of image smoothness for the various approaches. The spatial method applied no smoothing in addition to the one imposed by the volume coregistration procedure (Section IV-C), while the smoothing implemented by the wavelet method was dependent on the inherent local SNR conditions, and was typically minor for EPI and quite extensive for PRESTO images.

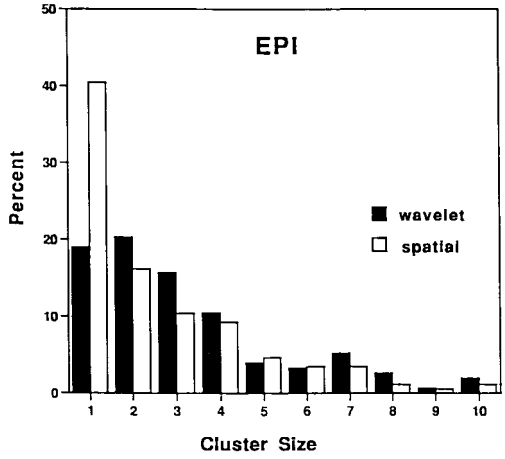

Fig. 9. Cluster-size distributions within ROI for EPI. Filled bars, regions detected by wavelet analysis and subsequent thresholding at 0.5% of the average fMRI intensities. Open bars, significant regions detected by spatial-domain analysis.

TABLE III
DETECTION THRESHOLDS[a] FOR WAVELET AND SPACE DOMAIN ANALYSIS WITHIN AN ROI

	PRESTO		EPI	
	mean	SD	mean	SD
% of channels requiring follow-up	5.32	3.04	38.93	5.33
% of coefficients tested	1.83	1.59	60.27	7.39
standardized wavelet threshold	3.22	0.09	3.87	0.03
standardized spatial threshold	4.59	0	4.59	0

[a] Standardized with respect to the standard error (25), overall significance per ROI of $\alpha = \alpha' = 0.05$.

Table III compares for the selected 3-D ROI the statistical efficiency of the wavelet-based analysis to that of the Bonferroni-adjusted spatial thresholding. The first row shows for the two modality images the space reduction attained by implementing the first step of the wavelet-based procedure that tests for significant channel power (28), (29). There were 120 wavelet channels (10 slices $\times$ 4 resolution levels $\times$ 3

orientations) to be tested in this ROI, of which, on average, about 6 (5.32%) for PRESTO and about 47 (38.39%) for EPI were found to carry significant power and to require follow-up testing. The second row expresses the resultant search space reduction in terms of the number of coefficients that needed to be tested in the second stage (30). Out of the total of 1 125 coefficients (equal to the number of pixels in the ROI) only about 20 (1.83%) required follow-up testing for PRESTO and about 678 (60.27%) for EPI. The large difference between the two modalities arises from the fact that for PRESTO nearly all channels with significant power were at low resolution levels (Table II), containing progressively smaller number of wavelet coefficients, while for EPI significant power was distributed into high-resolution channels that carry many coefficients. This difference in the number of follow-up tests effected the respective Bonferroni adjustments (24) and resulted in standardized thresholds of 3.22 and 3.87 for PRESTO and EPI, respectively. Hence, compared to the standardized spatial threshold of 4.59 (1 125 tests), the search space reductions achieved by the wavelet method yielded for both modalities a substantial threshold reduction. The lowest standardized threshold for $p = 0.05$ is attained if only a single z-test were required, yielding a value of 1.96. The standardized thresholds for PRESTO and EPI were about at 48% and 73%, respectively, of the difference between the standardized spatial threshold and this lower limit.

VI. Conclusion

In applications where no realistic assumptions regarding the shape, smoothness and location of the signals to be detected can be made, wavelet-based multiresolution methods are eminently suitable. Such applications arise when new functional imaging modalities with unknown characteristics emerge, and/or the signals of interest are largely unspecified due to either ill-controlled experimental paradigms or unknown underlying physiology. All of the above apply to brain fMRI's of human subjects. Wavelet-based methods permit the investigator to avoid the familiar problem of selecting the "best" monoresolution smoothing filter by using instead the more appropriate multiresolution point of view. By making use of the orthogonality and regularity conditions imposed on the wavelet bases, a fully automated procedure can be developed that applies standard parametric tests of significance directly to the wavelet coefficients. This then answers the posed question as to the "best" filtering by producing almost optimal, spatially adaptive filters. In contrast to the monoresolution approach, the resulting adaptive filters do not need to be specified beforehand, but rather are automatically derived from the prevalent SNR conditions at the various resolution levels of the image itself. Consequently, wavelet-based methods are particularly indicated if nothing specific can be assumed regarding the shape and spatial extent of the activation signals to be detected. In our application, a good estimate of the fastest spatial change of the signal of interest that could be expected was available, because image smoothness was imposed by the cubic-spline interpolation process required for the coregistration of the serial fMRI volumes. Hence, this knowledge was taken into account in the choice of cubic-spline wavelets for an efficient signal representation. Since both PRESTO and EPI scans were represented in the same wavelet basis, the relative comparison of the two acquisition techniques was not biased by this choice.

The two-stage signal detection approach as implemented here, answered one or two heuristic questions asked in sequence: 1) Are there any activation signals with a certain spatial extent present in the image? If affirmative, then: 2) Where are they located? This sequence achieves economy in statistical testing by limiting the search for spatial localizations to orthogonal signal-space partitions (i.e., resolution scales) with favorable SNR's. Hence, wavelet-based procedures have a potential advantage over spatial detection methods in that the SNR's prevalent in a limited number of statistically selected resolution channels typically exceed a monoresolution SNR, and that this selective resolution limitation results in a lower number of statistical tests and thus, a lower detection threshold for a given overall significance level. These advantages were shown to be practically realizable in direct comparisons to monoresolution spatial-domain detection for both PRESTO and EPI scans. Although both scans were acquired with equivalent nominal image resolutions, the wavelet based analysis clearly showed that, compared to EPI, current PRESTO techniques yielded generally a lower SNR in the difference images between baseline and task performance, and a lower resolution bandwidth by one to two octaves of the estimated activation signal. Clearly, this comparison of the two acquisition methods was only with respect to performance indexes characterizing signal detection, and both methods are still evolving and being improved.

References

[1] M. Unser and A. Aldroubi, "A review of wavelets in biomedical applications," *Proc. IEEE*, vol. 84, pp. 626–638, Apr. 1996.
[2] A. Aldroubi and M. Unser, *Wavelets in Medicine and Biology*. Boca Raton, FL: CRC, 1996.
[3] B. J. Lucier, M. Kallergi, W. Qian, R. A. DeVore, R. A. Clark, E. B. Saff, and L. P. Clarke, "Wavelet compression and segmentation of digital mammograms," *J. Digital Imag.*, vol. 7, no. 1, pp. 27–38, 1994.
[4] A. Nosratinia, N. Mohsenian, M. T. Orchard, and B. D. Liu, "Interframe coding of magnetic-resonance images," *IEEE Trans. Med. Imag.*, vol. 15, pp. 639–647, Oct. 1996.
[5] J. Wang and H. K. Huang, "Medical image compression by using three-dimensional wavelet transformation," *IEEE Trans. Med. Imag.*, vol. 15, pp. 547–554, Aug. 1996.
[6] T. Olson and J. DeStefano, "Wavelet localization of the Radon transform," *IEEE Trans. Signal Processing*, vol. 42, pp. 2055–2067, Aug. 1994.
[7] F. Peyrin, M. Zaim, and R. Goutte, "Construction of wavelet decompositions for tomographic images," *J. Math. Imag., Vision*, vol. 3, no. 1, pp. 105–121, 1993.
[8] J. B. Weaver, X. Yansun, D. M. Healy, and J. R. Driscoll, "Wavelet-encoded MR imaging," *Magn. Reson. Med.*, vol. 24, no. 2, pp. 275–287, 1992.
[9] D. M. Healy and J. B. Weaver, "Two applications of wavelet transforms in magnetic resonance," *IEEE Trans. Inform. Theory*, vol. 38, pp. 840–860, Mar. 1992.
[10] L. P. Panych, "Theoretical comparison of Fourier and wavelet encoding in magnetic resonance imaging," *IEEE Trans. Med. Imag.*, vol. 15, pp. 141–153, Apr. 1996.
[11] N. Gelman and M. L. Wood, "Wavelet encoding for 3-D gradient-echo MR-imaging," *Magn. Reson. Med.*, vol. 36, no. 4, pp. 613–619, 1996.
[12] J. B. Weaver, X. Yansun, D. M. Healy, and L. D. Cromwell, "Filtering noise from images with wavelet transforms," *Magn. Reson. Med.*, vol. 21, no. 2, pp. 288–295, 1991.

[13] D. L. Donoho and I. M. Johnstone, "Ideal spatial adaptation by wavelet shrinkage," *Biometrika*, vol. 81, no. 3, pp. 425–455, 1994.
[14] A. F. Laine, S. Schuler, F. Jian, and W. Huda, "Mammographic feature enhancement by multiscale analysis," *IEEE Trans. Med. Imag.*, vol. 13, pp. 725–740, Aug. 1994.
[15] R. Strickland, "Wavelet transforms for detecting microcalcifications," *IEEE Trans. Med. Imag.*, vol. 15, pp. 218–229, Apr. 1996.
[16] H. Yoshida, K. Doi, R. M. Nishikawa, M. L. Giger, and R. A. Schmidt, "An improved computer-assisted diagnostic scheme using wavelet transform for detecting clustered microcalssifications in digital mammograms," *Acad. Radiol.*, vol. 3, no. 8, pp. 621–627, 1996.
[17] B. Y. Zheng, W. Qian, and L. P. Clarke, "Digital mammography—Mixed feature neural-network with spectral entropy decision for detecting microcalcifications," *IEEE Trans. Med. Imag.*, vol. 15, pp. 589–597, Oct. 1996.
[18] U. E. Ruttimann, M. Unser, D. Rio, and R. R. Rawlings, "Use of the wavelet transform to investigate differences in brain PET images between patients," in *Proc. SPIE, vol 2035, Mathematical Methods in Medical Imaging II*, San Diego, CA, 1993, pp. 192–203.
[19] S. Ogawa, T. M. Lee, A. S. Nayak, and P. Glynn, "Oxygenation-sensitive contrast in magnetic resonance image of rodent brain at high magnetic fields," *Magn. Reson. Med.*, vol. 14, pp. 68–78, 1990.
[20] C. T. W. Moonen, "Imaging of human brain activation with functional MRI," *Biological Psych.*, vol. 37, pp. 141–143, 1995.
[21] S. Ogawa, G. W. Tank, R. S. Menon, J. Ellermann, S. Kim, H. Merkle, and K. Ugurbil, "Intrinsic signal changes accompanying sensory stimulation: Functional brain mapping with magnetic resonance imaging," in *Proc. Nat. Acad. Sci. USA*, 1992, vol. 89, pp. 5951–5955.
[22] P. A. Bandettini, E. C. Wong, M. S. Hinks, R. S. Tifofsky, and J. S. Hyde, "Time course EPI of human brain function during task activation," *Magn. Reson. Med.*, vol. 25, pp. 390–397, 1992.
[23] K. K. Kwong, J. W. Belliveau, D. A. Chesler, I. E. Goldberg, R. M. Weisskoff, P. B. Poncelet, D. N. Kennedy, B. E. Hoppel, M. S. Cohen, R. Turner, H. M. Brady, and B. R. Rosen, "Dynamic magnetic resonance imaging of human brain activity during primary sensory stimulation," in *Proc. Nat. Acad. Sci. USA*, 1992, vol. 89, pp. 5675–5679.
[24] J. V. Hajnal, R. Myers, A. Oatridge, J. E. Schwieso, I. R. Young, and G. M. Byder, "Artifacts due to stimulus correlated motion in functional imaging of the brain," *Magn. Reson. Med.*, vol. 31, pp. 283–291, 1994.
[25] J. H. Duyn, C. T. W. Moonen, R. W. de Boer, G. H. van Yperen, and P. R. Luyton, "Inflow versus deoxyhemoglobin effects in BOLD functional MRI using gradient echoes at 1.5 T," *NMR Biomed.*, vol. 7, pp. 83–88, 1994.
[26] S. Kim, K. Hendrich, X. Hu, H. Merkle, and K. Ugurbil, "Potential pitfalls of functional MRI using conventional gradient-recalled echo techniques," *NMR Biomed.*, vol. 7, pp. 69–74, 1994.
[27] S. Lai, A. L. Hopkins, E. M. Haacke, D. A. Wasserman, P. Buckley, L. Friedman, H. Meltzer, P. Hedera, and R. Friedland, "Identification of vascular structures as a major source of signal contrast in high resolution 2-D and 3-D functional activation imaging of the motor cortex at 1.5 T: Preliminary results," *Magn. Reson. Imag.*, vol. 30, pp. 387–392, 1993.
[28] R. L. Savoy, in *Rec., A Two Day Workshop on Functional Magnetic Imaging*, Boston, MA, June 16–17, 1996, p. 12.
[29] U. E. Ruttimann, M. Unser, P. Thévenaz, C. Lee, D. Rio, and D. W. Hommer, "Statistical analysis of image differences by wavelet decomposition," in *Wavelets in Medicine and Biology*, A. Aldroubi and M. Unser, Eds. Boca Raton, FL: CRC, 1996, pp. 115–144.
[30] P. van Gelderen, N. F. Ramsey, G. G. Liu, J. H. Duyn, J. A. Frank, D. R. Weinberger, and C. T. W. Moonen, "Three-dimensional functional magnetic resonance imaging of human brain on a clinical 1.5 T scanner," in *Proc. Nat. Acad. Sci. USA*, 1995, vol. 92, pp. 6906–6910.
[31] V. S. Mattay, J. A. Frank, A. K. S. Santha, J. J. Pekar, J. H. Duyn, A. C. McLaughlin, and D. R. Weinberger, "Whole-brain functional mapping with isotropic MR imaging," *Radiology*, vol. 201, pp. 399–404, 1996.
[32] S. G. Mallat, "A theory of multiresolution signal decomposition: The wavelet representation," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 11, pp. 674–693, July. 1989.
[33] I. Daubechies, *Ten Lectures on Wavelets*. Philadelphia, PA: Society for Industrial and Applied Mathematics, 1992.
[34] O. Rioul, "A discrete-time multiresolution theory," *IEEE Trans. Signal Processing*, vol. 41, pp. 2591–2606, Aug. 1993.
[35] M. Unser, "Ten good reasons for using spline wavelets," presented at *Proc. SPIE, vol 3169, Wavelet Application in Signal and Image Processing V*, San Diego, 1997, pp. 422–431.
[36] G. Battle, "A block spin construction of ondelettes. Part I: Lemarié functions," *Commun. Math. Phys.*, vol. 110, pp. 601–615, 1987.
[37] P.-G. Lemarié, "Ondelettes á localisation exponentielles," *J. Math. Pures et Appl.*, vol. 67, no. 3, pp. 227–236, 1988.
[38] A. Aldroubi and M. Unser, "Families of multiresolution and wavelet spaces with optimal properties," *Numer. Functional Anal., Optimizat.*, vol. 14, nos. 5–6, pp. 417–446, 1993.
[39] R. A. DeVore, B. Jawerth, and B. J. Lucier, "Image compression through wavelet transform coding," *IEEE Trans. Inform. Theory*, vol. 38, pp. 719–746, 1992.
[40] D. L. Donoho, "Unconditional bases are optimal bases for data compression and for statistical estimation," *Appl. Comput. Harmonic Anal.*, vol. 1, pp. 100–115, 1993.
[41] ———, "De-noising by soft-thresholding," *IEEE Trans. Inform. Theory*, vol. 41, pp. 613–627, 1995.
[42] M. Unser, A. Aldroubi, and M. Eden, "A family of polynomial spline wavelet transforms," *Signal Processing*, vol. 30, no. 2, pp. 141–162, 1993.
[43] ———, "B-spline signal processing. Part I: Theory," *IEEE Trans. Signal Processing*, vol. 41, pp. 821–833, Feb. 1993.
[44] E. Bullmore, M. Brammer, S. C. R. Williams, S. Rabe-Hesketh, N. Janot, A. David, J. Mellers, R. Howard, and P. Sham, "Statistical methods of estimation and inference for functional MR image analysis," *Magn. Reson. Med.*, vol. 35, pp. 261–277, 1996.
[45] K. J. Worsley and K. J. Friston, "Analysis of fMRI time-series revisited—Again," *Neuroimage*, vol. 2, pp. 173–181, 1995.
[46] K. J. Worsley, S. Marrett, P. Neelin, A. C. Vandal, K. J. Friston, and A. C. Evans, "A unified statistical approach for determining significant signals in images of cerebral activation," *Human Brain Mapping*, vol. 4, no. 1, pp. 58–73, 1996.
[47] J. B. Poline and B. M. Mazoyer, "Analysis of individual brain activation maps using hierarchical description and multiscale detection," *IEEE Trans. Med. Imag.*, vol. 13, pp. 702–710, Aug. 1994.
[48] K. J. Worsley, S. Marrett, P. Neelin, and A. C. Evans, "Searching scale space for activation in PET images," *Human Brain Mapping*, vol. 4, no. 1, pp. 74–90, 1996.
[49] D. L. Donoho, I. M. Johnstone, G. Kerkyacharian, and Picard, "Wavelet shrinkage: Asymptopia?," *J. Roy. Statistical Soc., Series B*, vol. 57, pp. 301–369, 1995.
[50] A. P. Holmes, R. C. Blair, J. D. G. Watson, and I. Ford, "Non-parametric analysis of statistic images from functional mapping experiments," *J. Cereb. Blood Flow Metab.*, vol. 16, pp. 7–22, 1996.
[51] P. Thévenaz, U. E. Ruttimann, and M. Unser, "A pyramid approach to sub-pixel registration based on intensity," *IEEE Trans. Image Processing*, vol. 7, pp. 27–41, Jan. 1998.
[52] N. K. Ramsey, B. S. Kirkby, P. van Gelderen, K. F. Berman, J. H. Duyn, J. A. Frank, V. S. Mattay, J. D. van Horn, G. Eposito, C. T. W. Moonen, and D. R. Weinberger, "Functional brain mapping of human sensorimotor cortex with 3-D BOLD fMRI correlates highly with $H_2{}^{15}O$ PET rCBF," *J. Cereb. Blood Flow, Metab.*, vol. 16, no. 5, pp. 755–764, 1996.
[53] J. H. Duyn, V. S. Mattay, R. H. Sexton, G. S. Sobering, F. A. Barrios, G. Liu, J. A. Frank, D. R. Weinberger, and C. T. W. Moonen, "3-dimensional functional imaging of human brain using echo-shifted FLASH MRI," *Magn. Reson. Med.*, vol. 32, pp. 150–155, 1994.
[54] C. T. W. Moonen, G. Liu, P. van Gelderen, and G. S. Sobering, "A fast gradient-recalled MRI technique with increased sensitivity to dynamic susceptibility effects," *Magn. Reson. Med.*, vol. 26, pp. 184–189, 1992.
[55] G. Liu, G. S. Sobering, J. H. Duyn, and C. T. W. Moonen, "A functional MRI technique combining principles of echo-shifting with a train of observations (PRESTO)," *Magn. Reson. Med.*, vol. 30, pp. 764–768, 1993.
[56] M. Hollander and D. A. Wolfe, *Nonparametric Statistical Methods*. New York: Wiley, 1973.