Statistical analysis of fMRI data using orthogonal filterbanks

Manuela Feilner, Thierry Blu and Michael Unser

Biomedical Imaging Group, DMT/IOA
Swiss Federal Institute of Technology Lausanne
CH-1015 Lausanne EPFL
Switzerland
e-mail: manuela.feilner@epfl.ch

ABSTRACT

Functional magnetic resonance imaging (fMRI) is a recent technique that allows the measurement of brain metabolism (local concentration of deoxyhemoglobin using BOLD contrast) while subjects are performing a specific task. A block paradigm produces alternating sequences of images (e.g., rest versus motor task). In order to detect and localize areas of cerebral activation, one analyzes the data using paired differences at the voxel level. As an alternative to the traditional approach which uses Gaussian spatial filtering to reduce measurement noise, we propose to analyze the data using an orthogonal filterbank. This procedure is intended to simplify and eventually improve the statistical analysis. The system is designed to concentrate the signal into a fewer number of components thereby improving the signal-to-noise ratio. Thanks to the orthogonality property, we can test the filtered components independently on a voxel-by-voxel basis; this testing procedure is optimal for i.i.d. measurement noise. The number of components to test is also reduced because of down-sampling. This offers a straightforward approach to increasing the sensitivity of the analysis (lower detection threshold) while applying the standard Bonferroni correction for multiple statistical tests. We present experimental results to illustrate the procedure. In addition, we discuss filter design issues. In particular, we introduce a family of orthogonal filters which are such that any integer reduction m can be implemented as a succession of elementary reductions m_1 to m_p where $m = m_1 \cdots m_p$ is a prime number factorization of m.

Keywords: functional imaging, fMRI, statistical analysis, medical imaging, wavelet application, filterbank

1. INTRODUCTION

fMRI (functional Magnetic Resonance Imaging) is a new technique that measures the activation of neuronal cells in the human brain indirectly. Neuronal activation involves changes in cerebral blood oxygenation that can be measured by a scanner, through a spin resonance technique.[1] This non-invasive detection of blood oxygenation level dependent contrast (BOLD) by fMRI offers new possibilities for neurophysiological and cognitive research, giving better insight into the functionality of the human brain.[2] The fMRI signals are intricate in their structure, which makes the evaluation of the data difficult. They contain a lot of noise and artifacts, originating not only from the acquisition, but also from physiological processes, e.g. blood pulsation.

Given such sequences of noisy fMRI data, the problem is to detect and localize the areas of activation in the brain. In particular, it is important to have a precise control of the probability of false detection so that we have a quantitative basis for reporting results in neuro-sciences. This means that one needs to know the significance level (or p-value) of the statistical tests that are used.

A method that performs the whole analysis process already exists. Its name is SPM (Statistical Parameter Mapping)[3]; it is available freely and is widely used by researchers analyzing functional images of the brain. SPM, however, has some limitations.[4] First, SPM is very intricate; researchers in the neuroscience, which are not necessarily well trained in statistics, tend to use it as a black box. There are many options that allow one to play with the data and change the outcome of the analysis. Second, SPM starts by applying a spatial Gaussian filter to the data which entails a loss of resolution. This filtering correlates the data which makes the statistical analysis more intricate; pixels can no longer be considered independently. Finally, the statistical inference in SPM is based on the theory of continuous Gaussian random fields[5,6]; it is complex mathematically and not entirely adequate because the data to which it is applied is discrete.

Here, we will propose to use orthogonal filtering and decimation as an alternative to the Gaussian filtering in SPM. Our primary motivation is to simplify the statistical analysis of fMRI data by not introducing correlation and make it more transparent.

The paper is organized as follows. In Section 2, we describe the fMRI-data and characterize the noise. Then, we propose a data model for the statistical analysis. With the result of this characterization, we also define a data-model. In Section 3, we present the orthogonal filterbank, wherewith we reduce the noise. In Section 4, we describe the statistical analysis, based on the above steps, to get the activation area. At the end, in Section 5, we present some results using real fMRI data.

2. FMRI-DATA

2.1. Data description

In this study, we are investigating simple fMRI block paradigms. A typical set of fMRI data consist of 8 repetitions of alternating blocks A and B. Symbol A stands for the brain volume with activation (e.g. left hand finger tapping) and B for the brain volume in the rest state.
In our case, a full volume of 30 slices of 128×128 pixels each was acquired every 6 seconds. An example of such a slice is shown in Figure 1.

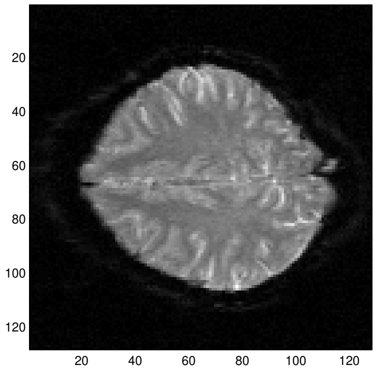


Figure 1. Example of a slice of raw data (upper part of the brain).

A time sequence of one voxel of an activated area is shown on the left hand side of Figure 2. It is obvious that a lot of noise is present in the data; the underlying signal which should follow the periodic activation pattern is hardly recognizable. The difference image of two slices (A-B) is shown on the right hand side of Figure 2. The amount of noise makes it almost impossible to detect any activation without resorting to statistical analysis.

The noise has two distinct origins: noise due to physiological activities (e.g. cardiac,[7] respiratory motions, task/stimulus related motions[8])[9] and noise of the measurement instrument (e.g. scanner).[10]
Also head motion cannot be completely avoided. Motion of the order of $\frac{1}{10}$ voxel may produce $1-2\%$ signal changes, which is not negligible. In comparison, fMRI BOLD effects are very small: $1-5\%$ signal changes.[11] This calls for the use of accurate image registration algorithms in a preprocessing stage.[12]

2.2. Characterization

To get a better understanding of the statistical nature of our data, we try to determine the main properties of the noise. This step is crucial if we are aiming at an accurate and rigorous statistical analysis of the results of fMRI experiments. The study of the noise proceeds with the same set of data as described above.
We make the general hypothesis that the noise is Gaussian distributed. This assumption is certainly reasonable when working with averaged finite energy data (due to central limit theorem), as this is the case here.

We question whether the variance is homogeneously distributed over space, or whether the variance differs significantly within the volume. The hypothesis for the following calculation is that the time sequence of each voxel is independent and normally distributed with variance σ; then, defining $s_k^2 = \sum_{t=1}^{n} \frac{(x_k(t)-\bar{x}_k)^2}{n-1}$, with $\bar{x}_k = \frac{1}{n}\sum_{t=1}^{n} x_k(t)$,

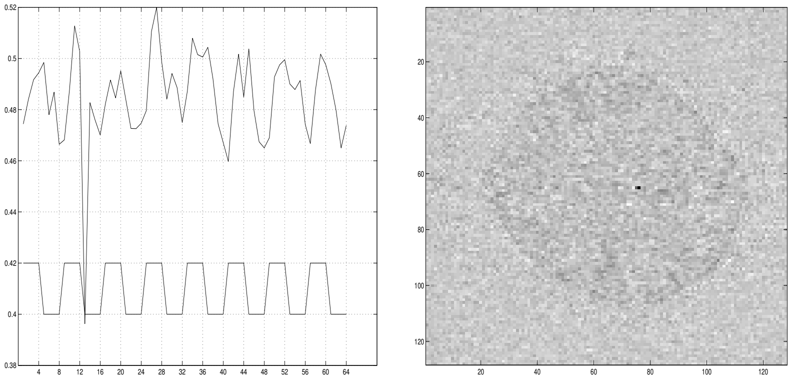


Figure 2. On the left hand side, we see a time sequence of an activated voxel. The upper plot is the sequence of the raw data (64 voxels). The lower plot gives the periodic rate of activated (up) state and rest (down) state. On the right hand side, we see the difference image of two slices (A-B). Without noise reduction, the activated area is not visible.

we know that $\frac{s_k^2}{\sigma^2}$ follows a χ^2 distribution with $n-1$ degrees of freedom, where k is a $3D$ integer index locating the voxel in the volume.
Therefore, we may write a confidence interval for this estimate as

$$P(a \leq (n-1)\frac{s_k^2}{\sigma^2} \leq b) = \alpha, \tag{1}$$

where P is the χ^2 cumulative probability.
"a" and "b" are to be computed in such a way as to obtain the smallest interval, depending on the confidence interval probability α. We observe that for a given α the σ^2 on a logarithmic scale has intervals of a constant size, which suggests a natural quantization step for the representation of the variance data s_k.

We thus get for each $\log(s_k^2)$ an uncertainty interval Δ, wherein the true $\log(\sigma^2)$ lies with the probability α (chosen to be close to 1). If two intervals are not adjacent, then we can assert that the corresponding estimates for σ^2 are significantly different.

Applying this to the slice we have already seen, we conclude that the variance is significantly not constant over space, contrary to what has often been stated in the literature.[13,14] The variances, however, appear to be strongly correlated over space—in other words, their variation is usually small within the neighborhood of a pixel.

2.3. Deterministic model

In order to make a statistical inference, we first need to specify a model.
Here, we assume to have independent (over time and space), but (non-stationary) Gaussian noise. The corresponding data model is:

$$x_k(t) = \mu_k(t) + e_k(t) \cdot \sigma_k,$$

where μ_k is the average hemodynamic function, e_k is an i.i.d. normalized noise ($< e_k >= 0, < e_k^2 >= 1$) and the variance σ_k varies over the pixels, but changes only slightly.
The hypothesis is that $\mu_k = \mu_1$ during A and $\mu_k = \mu_2$ during B. Thus, the detection of activation uses a statistical test that determines on a pixel by pixel basis, whether or not μ_1 and μ_2 are equal.
Our next goal will be to apply a spatial transformation to the data such that the model is still approximatily valid in the transformed domain.

3. ORTHONORMAL FILTERBANK

Since fMRI is so noisy, we need to reduce the noise in the spatial domain. As an alternative to the Gaussian filtering of SPM, our proposal for spatial noise reduction is to use a downsampled filterbank. The idea is that smoothing suppresses high frequencies so that the data may as well be represented with less samples. In addition, we try to preserve the independence of the noise as much as possible in order to simplify subsequent statistical analyses.

Our solution is to apply an orthonormal filter, which has the property of transforming white noise into white noise. We will also show that the independence assumption in the model in Section 2.3 remains approximately valid, under the assumption that σ_i^2 changes only slightly over space.

3.1. Orthonormal filter with decimation by a factor m

The downsampled filter is shown in Figure 3; $h_m(k)$ is a symmetric, orthonormal impulse response and m the downsampling rate.

This sequence of operations can be expressed as a descrete inner product

$$y(l) = < x(k), h_m(ml - k) >_{\ell_2} . \tag{2}$$

The definition of orthogonality is:

$$< h_m(k), h_m(k - ml) >= \delta_l \quad \Leftrightarrow \quad \sum_{l=0}^{m-1} |H_m(e^{j\omega - j\frac{2\pi l}{m}})|^2 = m,$$

where $H_m(z) = \sum h_m(k) z^{-k}$ is the z-transform of h_m.
If the object to filter is of dimension p, we build the corresponding p-D filter by tensor product.

For our application, the requirements on the filters are:

i) symmetry: needed in order to avoid phase distortion when localizing activated voxels.

ii) approximation order L (reproduction of polynomials of degree $\leq n = L - 1$): this property is beneficial for a better approximation of signals with most energy in the low frequency part.

iii) all possible integer decimation factors.

iv) consistent family with respect to chaining (for the simplicity of the design $\rightarrow$ modular library of transforms):

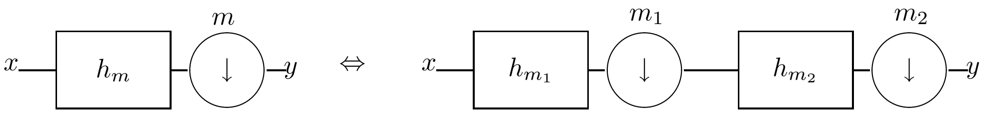


Figure 3. Signal flow diagram, where $m = m_1 m_2$.

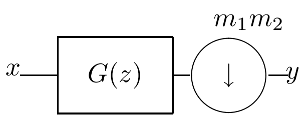
We will use the orthonormal spline filters of degree n, which are given by

$$H_m(e^{j\omega}) = \left(\frac{\sin(m\frac{\omega}{2})}{m \sin(\frac{\omega}{2})} \right)^{n+1} \sqrt{\frac{A(e^{j\omega})}{A(e^{jm\omega})}} \tag{3}$$

$$A(e^{j\omega}) = \sum_{k \in \mathbb{Z}} \left| \frac{\sin(\frac{\omega}{2})}{(\frac{\omega}{2} + k\pi)} \right|^{2(n+1)} \tag{4}$$

Proposition 1. *The spline family satisfies the requirements i)—iv).*

Proof. ii) This property is valid, where $L = n + 1$ in (3), because splines satisfy Strang-Fix conditions.[15]
iv) We start from the right hand signal flow diagram in Figure 3. In the first step, we exchange the decimation factor m_1 with the filter h_{m_2} and apply a *noble identity*[16] to get the system in Figure 4. The equivalent filter $G(z)$ is:



$$G(z) = H_{m_1}(z) H_{m_2}(z^{m_1})$$

Figure 4. Signal flow diagram after the exchange.

$$G(e^{j\omega}) = \left(\frac{\sin(m_1 \frac{\omega}{2})}{m_1 \sin(\frac{\omega}{2})} \right)^{n+1} \sqrt{\frac{A(e^{j\omega})}{A(e^{jm_1\omega})}} \left(\frac{\sin(m_1 m_2 \frac{\omega}{2})}{m_2 \sin(m_1 \frac{\omega}{2})} \right)^{n+1} \sqrt{\frac{A(e^{jm_1\omega})}{A(e^{jm_1 m_2\omega})}} \tag{5}$$

$$\Leftrightarrow \quad G(e^{j\omega}) = \left(\frac{\sin(m_1 \frac{\omega}{2})}{m_1 \sin(\frac{\omega}{2})} \frac{\sin(m_1 m_2 \frac{\omega}{2})}{m_2 \sin(m_1 \frac{\omega}{2})} \right)^{n+1} \sqrt{\frac{A(e^{j\omega})}{A(e^{jm_1\omega})} \frac{A(e^{jm_1\omega})}{A(e^{jm_1 m_2\omega})}}. \tag{6}$$

After cancellation of common factors, we get:

$$G(e^{j\omega}) = \left(\frac{\sin(m_1 m_2 \frac{\omega}{2})}{m_1 m_2 \sin(\frac{\omega}{2})} \right)^{n+1} \sqrt{\frac{A(e^{j\omega})}{A(e^{jm_1 m_2\omega})}} = H_{m_1 \cdot m_2}(e^{j\omega}). \tag{7}$$

$\square$

Note that this property is also true in the other direction: splines are the only "realizable" (i.e., $H_m(z) = \frac{P_m(z)}{Q_m(z)} \frac{B(z^m)}{B(z)}$, where P_m, Q_m are of finite degree and B is arbitrary) filters that satisfy requirement iv!
The proof of this result is very technical and won't be included here.
The proof that an orthogonal filter preserves the white noise is given in the Appendix. We also show that the model given in Section 2.3 remains approximately valid: If the variance varies only slightly over the size of the filter, then the noise is still independent.

4. HYPOTHESIS TESTING

Before starting with the statistical analysis, we omit the first volume of each block, because it contains a transition signal that would make the steady state analysis more difficult (note: it is possible to keep this first transient measurement by introducing a "hemodynamic" function in the model[17–19]).

We now assume that we have reduced spatial noise by filtering. To find the activation, we calculate the mean image of all activation state volumes A and of all rest state volumes B. Then the difference of these two mean values is made. We then obtain a volume of difference images. The difference will be denoted as $\bar{y}_{D_k} = \bar{y}_{A_k} + \bar{y}_{B_k}$. All subsequent analyses are done at the pixel level.

To decide which pixel is activated, we use hypothesis testing. We denote the null hypothesis—the pixel under consideration is not activated—by H_0. Under this hypothesis, $\bar{y}_{D_k} = 0$ and the distribution for the mean value of each voxel relative to its standard deviation is known (t-distribution). If a voxel exceeds some expected bound for $\bar{y}_{D_k}$, the null hypothesis is rejected and the alternative hypothesis H_1 is accepted; we conclude that the pixel is activated. The decision is made by setting a threshold.

Given some probability value α, we compute the threshold t such that the probability that $|t_k|$ exceeds t is less than or equal to α, assuming that H_0 is true, i.e, assuming that $\bar{y}_{D_k} \sim t$ distribution

$$P(t_k > t) = \alpha.$$

In other words, α is the probability of observing higher values t than by chance, given that the null hypothesis is true. This probability α is called the *level of significance* of the test. For all $t_k > t$, $k =$ voxels, we thus reject the null hypothesis and accept the alternative hypothesis H_1.

1 Null hypothesis H_0: $\bar{y}_{D_k} = 0$ (no activation)

2 Alternative hypothesis H_1: $\bar{y}_{D_k} \neq 0$ translates to $\bar{y}_{D_k} \geq T$, where T is a significance threshold as explained in Section 4.1.

The p-value is the probability of observing a test statistic at least as extreme as the one we observed, given the null hypothesis is true. The smaller the p-value, the stronger the evidence against H_0.
The test on the distribution has to be adapted to the set of data, which transforms the pixel values into p-values. Since the variance σ^2 changes with k, we apply a t-test (Student-t-distribution) with $n-1$ degrees of freedom.[20] The corresponding test statistic is

$$t_k = \frac{\bar{y}_{D_k}}{\frac{s_{D_k}}{\sqrt{n}}} \tag{8}$$

with the pooled variance estimated as

$$s_D^2 = s_A^2 + s_B^2 \tag{9}$$

$$s_A^2 = \sum_{i=1}^{n} \frac{(y_A(i) - \bar{y}_A)^2}{n-1}; \qquad s_B^2 = \sum_{i=1}^{n} \frac{(y_B(i) - \bar{y}_B)^2}{n-1}, \tag{10}$$

where s_A^2 and s_B^2 are the variances for the states A and B, respectively.

4.1. Determining significance levels

We have different possibilities to set the threshold. Some of them are discussed below.

4.1.1. Bonferroni-correction

If we look at all pixels jointly, we have to correct for multiple testing. Since we don't want to have globally more than α% of wrong activation decisions, we have to divide α by the number of pixels in the image. This is known as Bonferroni-correction for multiple testing:

$$\alpha' = \frac{\alpha}{N},$$

where N = number of pixels (or wavelet coefficient) in the image.
We have applied it to some of our data, see Figure 5 and 6.

4.1.2. Probability of the maximum

Another way of dealing with multiple variables is to look at the maximum of a Gaussian vector. We choose the threshold by computing the probability that only the maximum of N random variables is greater or equal to this threshold (11).

$$\begin{aligned} P(\max(|x_n|) \geq T) &\leq P(|x_1| \geq T) + P(|x_2| \geq T) + ...P(|x_n| \geq T) \\ &\leq Np \end{aligned} \tag{11}$$

where

$$p = P(|x_0| \geq T).$$

It follows that $p \geq \frac{\alpha}{N}$; the lower bound is the same result as with Bonferroni. When the data are independent, the lower bound is achieved exactly meaning that the approach of the probability of the maximum and the Bonferroni-correction are essentially equivalent. For correlated data, one could in principle find a $p > \frac{\alpha}{N}$, which yields a lower detection threshold.

4.1.3. Binomial law

Another possibility for multiple correction is to determine the number of false detections (n_0) that is expected from our significance level α. The probability to observe n_0 false detections in a random field without activation follows a binomial distribution:

$$P(n = n_0) = \binom{N}{n_0} p^{n_0}(1-p)^{N-n_0}, \tag{12}$$

with $p = P(t_k > t)$ and N = number of pixels in the image. Since we know how many wrong detections we should expect in the image, we are able to compare them with the effectively observed ones. If there are significantly more than expected, then we have activation within the image. Yet, we have not localized it, because this is a global test only.

In the case of Bonferroni-correction, $n_0 = 0$ and the mean value $\langle n \rangle$ of appearance of a false detection is equal to $Np = \alpha'$ (usually $\alpha' = 0.05$). Consequently Bonferroni's approach is very conservative.

Applying a global test, we may also want to know the location of the detection. Since it is most probable that the activation appears in clusters while the false positives appear scattered, we could check for clusters of a given size after the global testing.

5. RESULTS

To illustrate the technique, we present some examples of real data analysis (cf. Figures 5 to 6). We see that without any filtering, we hardly get any detection if we apply the Bonferroni-correction, see Figure 5. After orthogonal filtering, pixels tend to vary more smoothly. If we now apply the Bonferroni-correction, we detect activity in the motor cortex. We also calculated the corrected threshold of SPM[6] and applied it to the Gaussian-smoothed image with a sigma corresponding to the smoothing factor of the orthogonal filter (Figure 5). In our method, after decimation, the activated area is conserved as well, see Figure 6.

6. CONCLUSION

With this proposal, we hope to have provided a viable alternative to SPM, the de facto standard in the field. In particular, we have simplified the statistical analysis of the data by preserving its independence in the transformed domain. Additional advantages of our decimated orthogonal filterbank approach are speed and straightforward control of resolution. Here, we use decimation as a alternative to SPM's not-so-straightforward notion of RESEL.

The present method is also applicable, with minor modifications, for the detection of activation in the wavelet transform domain, following the initial proposal of Ruttimann et al. We presently feel that a non-stationary model as described here is more appropriate for fMRI than the white noise model considered in.[14] In other words, it seems more appropriate to use t-tests rather than z-tests as in.[14]

APPENDIX A. DECORRELATION PROPERTY

Proof. We first prove that an orthogonal filter preserves white noise:
input signal: $x_n = \mu_n + \sigma_n e_n$, where $< e_n >= 0$ and $< e_n e_k >= \delta_{n-k}$
output signal: $y_n = \sum_k h_{nM-k} x_k$

$$y_n = \underbrace{\sum_k h_{nM-k}\mu_k}_{\hat{y}_n} + \sum_k h_{nM-k}\sigma_k e_k$$

$z_n = y_n - \hat{y}_n, < z_n >= 0$
$< z_n z_{n'} >=< \sum_k h_{nM-k}\sigma_k e_k \sum_k h_{n'M-k'}\sigma_{k'} e_{k'} >$
$= \sum_k \sum_{k'} h_{nM-k} h_{n'M-k'} \sigma_k \sigma_{k'} \underbrace{< e_k e_{k'} >}_{\delta_{k-k'}}$
$= \sum_k h_{nM-k} h_{n'M-k} \sigma_k^2$

First case: uniform variance: $\sigma_k = \sigma$: $< z_n z_{n'} >= \sigma^2 \underbrace{\sum_k h_{nM-k} h_{n'M-k}}_{\delta_{n-n'},\ \text{if } h_n \text{ is orthogonal}}$

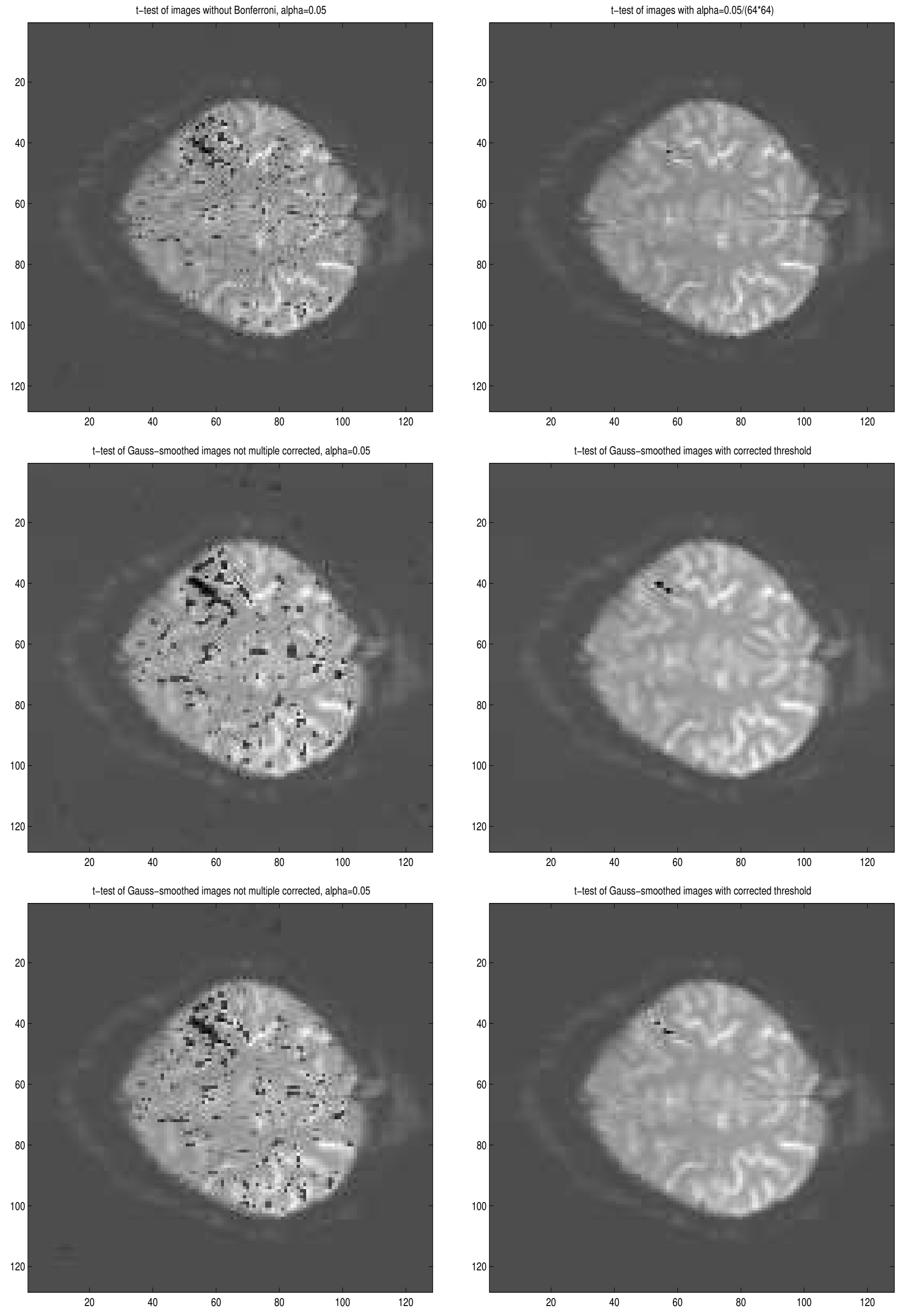


Figure 5. In the first two images, the test is done without filtering the data. On the left hand side, the significance level is $\alpha = 0.05$. This corresponds to a threshold $t = 1.7139$ (without Bonferroni-correction). The detected pixels have different grayscales, because they are represented here as p-values. In the map on right hand side the Bonferroni-correction is applied. In the second two maps, the images are filtered with the orthogonal spline filter, scale 2, but not downsampled, just to compare with the nonfiltered image. In this case, the detections are more clustered, and the "true" activation is more visible. In the bottom images, we filtered with a Gaussian kernel like in SPM.

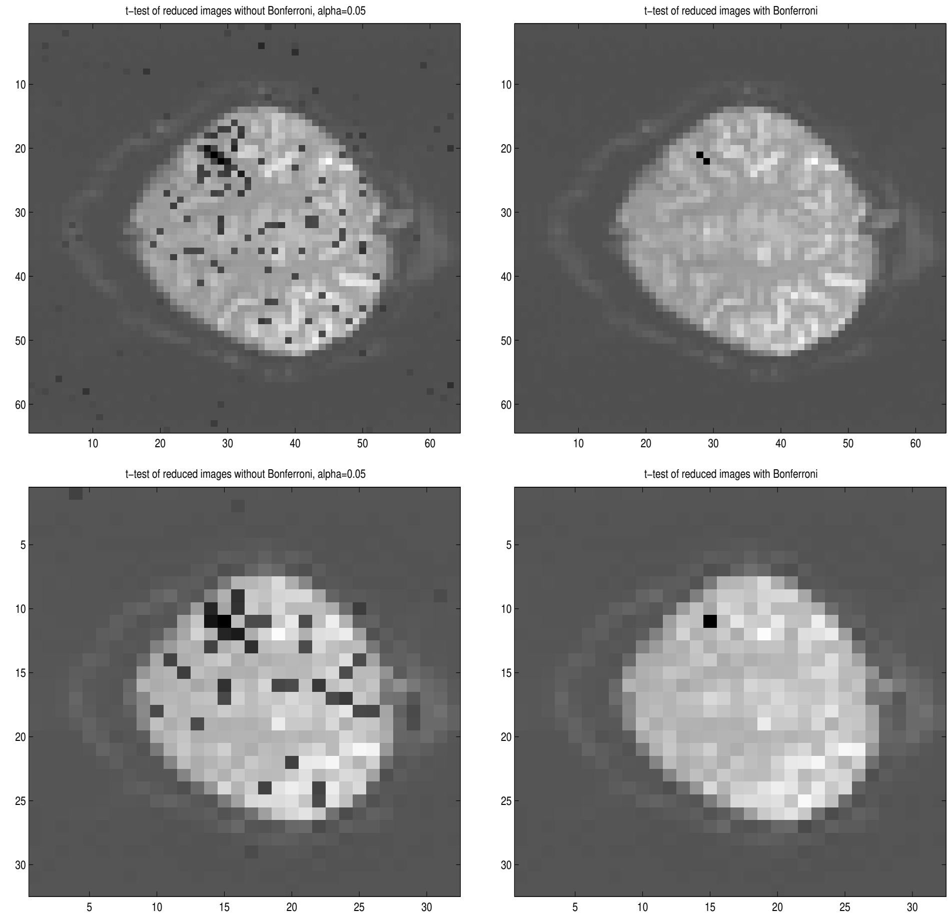


Figure 6. On the left hand side, the upper map is obtained by filtering and downsampling with the orthogonal spline filter with scale 2. The significance level is $\alpha = 0.05$ (uncorrected). On the right hand side, the upper image is obtained in the same way, but the threshold was corrected by the Bonferroni-correction. In the lower images, we filtered and downsampled with scale 4

Second case: close to uniform variance: $\sigma_k = \sigma(1 + \varepsilon_k)$, $\varepsilon_k \ll 1$

$$< z_n z_{n'} >= \sum_k h_{nM-k} h_{n'M-k} \sigma^2 (1+\varepsilon_k)^2 = \sigma^2 \delta_{n-n'} + 2\sigma^2 \underbrace{\sum_k h_{nM-k} h_{n'M-k} \varepsilon_k}_{r} + \sigma^2 \sum_k h_{nM-k} h_{n'M-k} \varepsilon_k^2$$

Using Cauchy-Schwartz, we get $r \leq \sup |\varepsilon_k| \sum_k |h_{nM-k} h_{n'M-k}| \leq \underbrace{\sqrt{\sum_k h_{nM-k}^2 \sum_k h_{n'M-k}^2}}_{=1} \sup |\varepsilon_k| \leq \sup |\varepsilon_k|$.

So that $< z_n z_{n'} >= \sigma^2 (\delta_{n-n'} + \rho_{n,n'}), \text{where} |\rho_{n,n'}| \leq 2 \sup |\varepsilon_k| + \sup \varepsilon_k^2 \ll 1$.
Thus, in both cases, z_n remains uncorrelated with $z_{n'}$ for $n \neq n'$. $\square$

Everywhere we have homogenous zones, this result is valid. On the border of such zones, like the border of the brain against the background, it is not valid anymore over a band that has the width of the "efficient" length of the filter G.

ACKNOWLEDGMENTS

We thank Arto Nirkko for providing the fMRI data, and Jan Kybic for sitting in the scanner.

REFERENCES

1. K. Shung, M. Smith, and B. Tsu, *Principles of Medical Imaging*, Academic Press, 1992.
2. S. Ogawa, R. Menon, D. Tank, S. Kim, H. Merkle, J. Ellerman, and K. Ugurbil, "Functional brain mapping by blood oxygenlevel dependent contrast," *Biophysical J.* **64**, pp. 803–812, 1993.
3. R. Frackowiak, K. Friston, C. Frith, R. Dolan, and J. Mazziotta, *Human Brain Function*, Academic Press, 1997.
4. S. Gold, B. Christian, S. Arndt, G. Zeien, T. Cizadlo, D. Johnson, M. Flaum, and N. Andreasen, "Functional MRI statistical software packages: A comparative analysis," *Human Brain Mapping* **6**, pp. 73–84, 1998.
5. K. Worsley, "Local maxima and the expected Euler characteristic of excursion sets of ξ^2, t and f fields.," *Journal of Applied Probability* **26**, pp. 13–42, 1994.
6. K. Worsley, A. Evens, S. Marrett, and P. Neelin, "A three dimensional statistical analysis for cbf activation studies in human brain.," *J. Cerebral Blood Flow and Metabolism* **12**, pp. 900–918, 1992.
7. M. Dagli, J. Ingeholm, and J. Haxby, "Localization of cardiac-induced signal change in fMRI," *NeuroImage* **9**, pp. 407–415, 1999.
8. E. Bullmore, M. Brammer, S. Rabe-Hesketh, V. Curtis, R. Morris, S. Williams, T. Sharma, and P. McGuire, "Methods for diagnosis and treatment of stimulus-correlated motion in generic brain activation studies using fMRI," *Human Brain Mapping* **7**, pp. 38–48, 1999.
9. S. Ogawa, R. Menon, S. Kim, and K. Ugurbil, "On the characteristic of functional magnetic resonance imaging of the brain," *Annu. Rev. Biophys.* **27**, pp. 447–474, 1998.
10. T. Budinger, F. Wehrli, *et al.*, *Mathematics and Physics of Emerging Biomedical Imaging*, National Academic Press, 1996.
11. R. Cox, "A two day workshop on functional MRI," in *Proc. Colloque de Physique C2*, June 1996.
12. P. Thevenaz, U. Ruttimann, and M.Unser, "A pyramid approach to subpixel registration based on intensity," *IEEE Trans. Image Processing* **7**(1), pp. 27–41, 1998.
13. D. Ekatodramis, *A Hypothesis-Testing Approach to Functional Inference of Brain Activity from Magnetic Resonance Image Series.* PhD thesis, Swiss Federal Institute of Technology Zurich, Zürich, 1998.
14. U. Ruttimann, M. Unser, R. Rawlings, D. Rio, N. Ramsey, V. Mattay, D. Hommer, J. Frank, and D. Weinberger, "Statistical analysis of functional MRI data in the wavelet domain," *IEEE Trans. Medical Imaging* **17**(2), pp. 142–154, 1998.
15. G. Strang and G. Fix, "A Fourier analysis of the finite element variational method," in *Constructive Aspect of Functional Analysis*, Cremonese, ed., pp. 796–830, Rome, 1971.
16. M. Vetterli and J. Kovačević, *Wavelets and Subband Coding*, Prentice-Hall, 1995.
17. P. Bandettini, A. Jesmanowicz, E. Wong, and J. Hyde, "Processing strategies for time-course data sets in functional MRI of the human brain," *Magnetic Resonance in Medicine* **30**, pp. 161–173, 1993.
18. S. Rabe-Hesketh, E. Bullmore, and M. Brammer, "The analysis of functional magnetic resonance images," *Statistical Methods in Medical Research* **6**, pp. 215–237, 1997.
19. J. Rajapakse, F. Kruggel, J. Maisog, and v. Y. D. Cramon, "Modeling hemodynamic response for analysis of functional MRI time-series," *Human Brain Mapping* **6**, pp. 283–300, 1998.
20. C. Chatfield, *Statistics for Technology*, Chapman & Hall, 1983.