# Detecting and Infering Brain Activation from functional MRI by Hypothesis-Testing based on the Likelihood Ratio

Dimitrios Ekatodramis, Gábor Székely and Guido Gerig

Swiss Federal Institute of Technology Communication Technology Laboratory ETH-Zentrum, CH-8092 Zurich, Switzerland

{ekato,szekely,gerig}@vision.ee.ethz.ch

Abstract. For the measure of brain activation in functional MRI many methods compute a heuristically chosen metric. The statistic of the underlying metric which is implicitly derived from the original assumption about the noise in the data, provides only an indirect way to the statistical inference of brain activation. We propose an alternative procedure by presenting a binary hypothesistesting approach. This approach treats the problem of detecting brain activation at its origin by directly deriving a test statistic based on the probabilistic model of the noise in the data. Thereby, deterministic or parameterized models for the hemodynamic response can be considered. Results show that time series models can be detected even if they are characterized by unknown parameters, associated with the unclear nature of the mechanisms that mediate between neuronal stimulation and hemodynamic brain response. The likelihood ratio tests proposed in this paper are very efficient and robust in making a statistical inference about detected regions of brain activation. To validate the usefulness of our approach we present a simulation environment for functional MRI. This environment also serves as a testbed for comparative study and systematic tests.

## 1 Introduction

The human brain has a highly complex functional organization. Functional MRI techniques are specialized to code information about neurophysiological mechanisms of the brain in the image contrast. The analysis of functional MRI data aims at providing detailed maps depicting neuronal brain activation due to a specific sensory stimulation, so called *activation maps*.

Most analysis methods compute a metric upon which brain activation in functional MRI data is detected. This metric arises from a heuristically posed question, like 'how well does a model correlate with the time series data?', or 'what is the amount of signal power in the time series at the stimulation frequency?'. Answers to these questions indirectly lead to the conclusion about brain activation being present or not. In fact, a direct solution to the problem posed in functional MRI is the answer to the basic question 'is brain activation present or not?'. We present a binary hypothesis-testing approach to answer this question. For this, we make the assumption that each time series measurement represents either the hemodynamic response signal buried in noise or just pure noise. This binary hypothesis can be denoted as

 $H_1$  : Measurement  $\equiv$  Signal+Noise  $H_0$  : Measurement  $\equiv$  Noise

The signal may be a deterministic or parameterized model for the hemodynamic response, while the noise is characterized by a probabilistic model.

Validation of functional MRI results can be performed by comparison with findings using different modalities such as PET, MEG or EEG on the same person. Until now the exact nature of the transient behavior between onset of sensory stimulation, neuronal activation and hemodynamic response of the brain, remains unclear. This implies that validation can never become complete, since any underlying model or assumption in the analysis does not accurately represent the physical nature of functional MRI experiments. We present an environment where the hemodynamic response and the experimental parameters of a functional MRI experiment can be simulated in a realistic manner. Artificially produced data sets are necessary to evaluate the reliability of analysis methods that are applied on real data with similar experimental conditions.

## 2 Methods

## 2.1 Binary Hypothesis Testing

Having made an observation (time series) with functional MRI, one is faced with the necessity of making a binary decision between the two hypotheses  $H_1$  (signal present) and  $H_0$  (signal not present). Once a rule is selected to decide which of the two hypotheses is true for the present observation, the procedure is referred to as binary hypothesis testing. We present the Neyman-Pearson test as a binary hypothesis-testing approach.

The Likelihood Ratio and the Neyman-Pearson Test. Interpreting the hypotheses  $H_1$  and  $H_0$  as a known source output (signal or '0') which has gone trough some probabilistic transition mechanism (addition of noise), the measurement can be described by a random variable X. Then the likelihood ratio  $\Lambda(x)$  provides the test

$$\Lambda(\boldsymbol{x}) = \frac{p_{\boldsymbol{X}|H_1}(\boldsymbol{x}|H_1)}{p_{\boldsymbol{X}|H_0}(\boldsymbol{x}|H_0)} \quad \begin{array}{c} H_1 \\ > \\ < \\ H_0 \end{array} \quad \gamma$$
 (1)

to decide which of the two hypotheses is true for a particular observation  $\boldsymbol{x}$  ([13]). The likelihood ratio is the ratio of the probability density functions of the measurement  $\boldsymbol{X}$ , assuming that  $H_1$  or  $H_0$  is true, respectively. For the likelihood ratio tests in this paper we determine the threshold  $\gamma$  using the Neyman-Pearson criterion. According to this criterion the false alarm probability  $P_f$  (deciding for  $H_1$  when  $H_0$  is true) is constrained to a specified value  $\alpha$ , while the detection probability  $P_d$  (deciding for  $H_1$  when  $H_1$  is true) is maximized for this constraint. Under the assumption of Gaussian white noise it is shown by Van Trees [13] that the threshold  $\gamma$  is then determined by solving the equation

$$P_f = P[\Lambda(\boldsymbol{x}) > \gamma | H_0] = \int_{\gamma}^{\infty} p_{\Lambda|H_0}(\lambda | H_0) d\lambda = \alpha$$
 (2)

where  $p_{A|H_0}(\lambda|H_0)$  denotes the probability density function of the likelihood ratio under  $H_0$ . Note that for a likelihood ratio test based on the Neyman-Pearson criterion (Neyman-Pearson test), the false alarm probability  $P_f$  and therefore the statistical inference is completely specified by the constraint  $\alpha$  for  $P_f$ . Therefore, whenever a Neyman-Pearson test can be designed, the user himself can determine the statistical inference to result from the binary decision problem.

**Detection of Known Time Series Signals.** A time series can be viewed as a series or collection of samples or observations taken from a continuous time signal. Having made an observation of a time series, we can make the following binary hypothesis:

$$H_1$$
:  $Y_i = s_i + n_i$ ,  $i = 1, \dots, N$   
 $H_0$ :  $Y_i = n_i$ ,  $i = 1, \dots, N$  (3)

Under  $H_1$  we assume that the  $Y_i$  of the observed time series consist of samples  $s_i$  of a known signal which is corrupted with additive Gaussian white noise. The noise samples  $n_i$  are assumed to be members of a Gaussian white noise process with zero mean and variance  $\sigma_n^2$ . They are statistically independent. The samples  $s_i$  can be thought of representing the model time series of the hemodynamic brain response.  $H_0$  makes the assumption that the  $Y_i$  of the observed times series consist of pure noise samples of the same type. N indicates the number of samples in the observed time series. The assumption of Gaussian white noise allows to formulate the likelihood ratio test

$$\sum_{i} y_{i} s_{i} \quad \stackrel{H_{1}}{\underset{H_{0}}{>}} \quad \gamma \tag{4}$$

By using equation (2) for the determination of the threshold  $\gamma$  according to the Neyman-Pearson criterion we obtain

$$\gamma = \sqrt{2}\sigma_n^2 \sum_i s_i^2 \cdot \operatorname{erf}^{-1}(1 - 2\alpha) \tag{5}$$

where  $\alpha$  is the constraint for the false alarm probability  $P_f$  and  $\operatorname{erf}^{-1}$  denotes the inverse of the error function. Using the threshold  $\gamma$  determined by (5) in the likelihood ratio test (4) results in a false alarm probability  $P_f = \alpha$ . Therefore, the statistical inference made from the test can be determined by the choice of  $\alpha$ .

The Generalized Likelihood Ratio Test for Unknown Parameters. If a hypothesis in a binary decision problem depends on an unknown parameter, the likelihood ratio test (1) is not directly applicable. This is due to the fact that the probability density function for that hypothesis, denoted as the likelihood function, also depends on the unknown parameter. Therefore, an estimate of the parameter is obtained under the assumption that the hypothesis which depends on this parameter is true. This estimate is then used in the formulation of a likelihood ratio test as if it was the correct value of the unknown parameter. A likelihood ratio test which is constructed with maximum likelihood estimates ([3, 13]) of some unknown parameters is called a generalized likelihood ratio test. Maximum likelihood estimates are obtained by minimizing the likelihood function using standard techniques of calculus.

Detection of a Cosine Waveform with Unknown Phase. In functional MRI observations of brain physiology are repeated many times in a single experiment. Since in these cycles the subject is switching from one condition to another, a periodic response in the time series of activated brain areas is expected. For this reason, in many functional MRI evaluation procedures a sinusoidal form of the hemodynamic response is proposed ([7, 8]). However, the phase of such sinusoidals which is associated with the delay of the hemodynamic response function is arbitrarily chosen or empirically assumed ([1, 10]). Therefore, we present a detection procedure for a cosine waveform time series with unknown phase. For this, we make the binary hypothesis

$$H_1$$
:  $Y_i = \cos(\omega t_i + \phi) + n_i$ ,  $i = 1, \dots, N$   
 $H_0$ :  $Y_i = n_i$ ,  $i = 1, \dots, N$  (6)

Under  $H_1$  the  $Y_i$  of the observed time series are assumed to be samples of a cosine function with known frequency  $\omega$  at time instants  $t_i$ . However the phase  $\phi$  is not known. Suppose that  $\omega = 2k\pi/p$ , where k is an integer number  $\neq 0$  and p is an integer number > 0. Further, assume that  $N \geq p$  and that N is a multiple of p. If we denote that  $t_i = i$ , then we have the conditions

$$\sum_{i=1}^{N} \cos\left(\frac{2k\pi}{p}i + \phi\right) = 0 \quad \text{and} \quad \sum_{i=1}^{N} \cos^2\left(\frac{2k\pi}{p}i + \phi\right) = \frac{N}{2}$$
 (7)

for any possible value of  $\phi$ . This means that the cosine function in (6) is sampled over an integer number of periods, and with the same number of samples per period. The observation under  $H_1$  is assumed to be corrupted with Gaussian white noise of zero mean and variance  $\sigma_n^2$ . Under  $H_0$  only pure noise of the same type is assumed.

After determining a maximum likelihood estimate for the unknown phase  $\phi$  considering the conditions in (7) we can formulate the generalized likelihood ratio test

$$\left(\sum_{i} y_{i} \cos \omega t_{i}\right)^{2} + \left(\sum_{i} y_{i} \sin \omega t_{i}\right)^{2} \quad \underset{H_{0}}{\overset{H_{1}}{\geq}} \quad \gamma \tag{8}$$

The determination of the threshold  $\gamma$  is again based on the Neyman-Pearson criterion. Using equation (2) we obtain

$$\gamma = \frac{N}{2}\sigma_n^2 K_2(\alpha) \tag{9}$$

where  $\alpha$  is the constraint for false alarm probability  $P_f$  and  $K_2(\cdot)$  denotes the cumulative  $\chi^2$ -distribution with 2 degrees of freedom. Using the threshold  $\gamma$  determined by (9) in the likelihood ratio test (8) results in a false alarm probability  $P_f = \alpha$ . Therefore, the statistical inference made from the test can be determined by the choice of  $\alpha$ .

## 2.2 Verification of the Gaussian Noise Model in functional MRI Data

The development of the likelihood ratio tests (8) and (4) was based on a Gaussian assumption about noise. Application of these tests on real data is legitimate only after verification of the Gaussian white noise model in real data. We have performed this verification empirically and with the help of the  $\chi^2$ -distribution and the statistical  $\chi^2$ -test. We have made use of a dummy data set acquired from a volunteer. The data set consists of 512 functional images with an image resolution of  $128 \times 128$  pixels, acquired with EPI on a Bruker TOMIKON S200. Only pixels within brain tissue were incorporated in the verification process. The volunteer was not exposed to any specific sensory stimulation.

Equal Variance over Space and Time. A histogram of all pixel values within brain tissue in the time series of the dummy data set is shown in figure 1a. We see that this histogram is well approximated by a Gaussian distribution function with variance equal to the variance estimate over all pixels under investigation. The variance estimate has been obtained using the standard estimation calculus for the variance over all pixels, after subtracting the mean from each time series. This estimate, denoted as the overall variance, is assumed to be the true noise variance. Figure 1b shows a histogram of the normalized variance estimate from the time series within brain tissue (solid line). The computed histogram is well approximated by the theoretically expected  $\chi^2$ -distribution (dotted line). The plot in figure 1b therefore supports our assumption that the measured overall variance is well the true noise variance, and that it is the same at any pixel location. To corroborate our assumption about equal variance over space we have performed the statistical  $\chi^2$ -test for any two pairs of time series, to test if they are drawn from different distributions ([12], p. 489). 95\% of the significance levels obtained from the  $\chi^2$ -test were > 0.1. Therefore we can say that there are very insignificant differences in the distribution of practically any two time series in the dummy data set. This in turn is an indication that our assumption about the variance being the same in any time series, i.e. equal over space, holds.

To verify that the variance in the time series of the dummy data set is equal over time, we have observed the distribution of the normalized variance estimate in time series within a sliding time window. In the data set with 512 time samples, we have chosen a window width of 128 time instants. The computed mean and standard deviation of the normalized variance estimate for any time section are shown in figure 1c and 1d (solid line), together with the mean and standard deviation of the theoretically expected  $\chi^2$ -distribution (dotted line). We observe that the mean and the variance remain well the same over time. On the basis of the results in figure 1c and 1d it seems justifiable to conclude that the assumption of the equal variance over time is satisfied.

Gaussian Distribution of the Time Series Samples. The assumption we make, is that the time series data are drawn from a Gaussian distribution with zero mean and variance equal to the measured overall variance. Note that the mean of the assumed Gaussian distribution is zero, because the mean of each time series has been subtracted. To verify our assumption we have performed the statistical  $\chi^2$ -test for the distribution of the samples of each time series and the assumed Gaussian distribution ([12], p. 488). 95% of the significance levels obtained from the  $\chi^2$ -test were > 0.1. Therefore, we can say that there are very insignificant differences in the true distribution of the data and the Gaussian distribution. This in turn is an indication that our assumption about the time series data being Gaussian distributed holds.

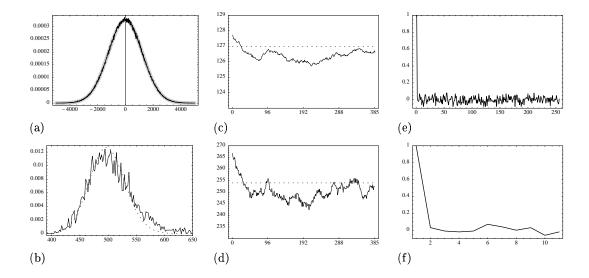


Fig. 1. Verification of the Gaussian white noise model in functional MRI data. (a) Histogram with 1024 bins of the time series data within brain tissue (solid line), well approximated by a Gaussian distribution (thick grey line) with variance equal to the variance estimate from the data. (b) Histogram with 128 bins of the normalized variance estimate from the time series within brain tissue (solid line), well approximated by the theoretically expected  $\chi^2$ -distribution with 511 degrees of freedom (dotted line). (c) Mean and (d) standard deviation of the normalized variance estimate from windowed time series data (solid line). The dotted line indicates the theoretically expected values. The abscissa indicates the starting point of the time window in the dummy data set of 512 time samples. (e) Example of the autocorrelation coefficients calculated from a time series of the dummy data set. (f) A clipping from the plot in (e).

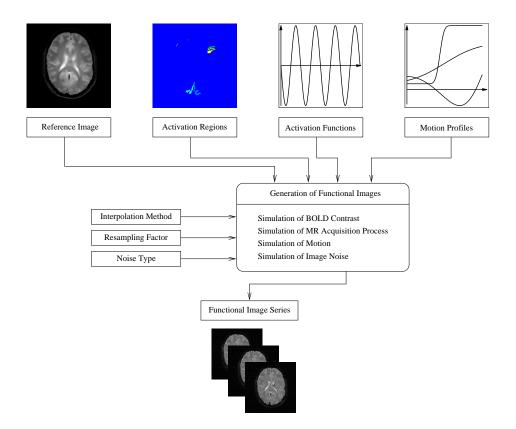
Independent Time Series Samples. The assumption of the noise being white requires the verification that the samples in the time series are independent. For this verification we have made use of the Box-Pierce test for white noise ([6,8]). The Box-Pierce test statistic calculates the sum of squared autocorrelation coefficients at lags k = 1..K. An example of autocorrelation coefficients calculated from a time series of the dummy data set are plotted in figure 1e,f.

Under the null hypothesis that the time series in question is serially independent, or white noise, the Box-Pierce statistic is assumed to be  $\chi^2$ -distributed with K degrees of freedom. Therefore, we can again use a  $\chi^2$ -test for significant dependencies in the time series samples. 90% of the significance levels obtained from the  $\chi^2$ -test were > 0.1. This means that there are very insignificant dependencies in the time series samples. From this test result we can deduce that our assumption of the noise being independent holds.

### 2.3 Simulation of functional MRI Data

To fulfill the requirements for simulation a software environment for the generation of artificial functional images, as shown in figure 2, was designed and implemented using AVS as a development tool ([14, 9]. In the implementation the user determines different physiological and experimental parameters of a functional MRI measurement. This mainly consists of the outlining of brain activation regions on a reference image and of the assignment of a hemodynamic response to any of these regions. Further parameters, such as image resolution, number of images and the noise level in the images, can be determined. The software also supports the simulation of motion artifacts. For this, interpolation procedures for the spatial transformation of images are provided.

In functional MRI the amplitude of the hemodynamic response may vary from experiment to experiment ([11, 2, 4]). Taking into account a constant noise level in the images, analysis methods for functional MRI data may therefore take advantage of high response amplitudes or may suffer from weak responses. For this reason, we have considered important to investigate the influence



**Fig. 2.** Functional imaging simulation environment. The user supplies a high resolution structural reference image, activation regions and functions, and motion profiles. Further, he selects an interpolation method, a resampling factor and a noise type. The software outputs a low resolution functional image series. The information of the activation functions, contained in the time series at pixels within the activation regions, is corrupted by the displacement of the images to each other according to the supplied motion profiles and by the specified noise type.

of the signal-to-noise ratio on the performance of methods applied to functional MRI data by the use of the simulation environment. For this investigation, the signal-to-noise ratio

$$SNR \equiv \frac{\hat{A}}{\sigma_n} \tag{10}$$

in artificial functional data sets was defined as the ratio of the amplitude  $\hat{A}$  of the activation function and the standard deviation  $\sigma_n$  of the noise in the functional images.

#### 3 Results

# 3.1 Comparison of Analysis Methods using Simulated Data

The performance of five different methods for functional MRI analysis with respect to the signal-to-noise ratio (10) was investigated using simulated test data. These methods were:

- The averaged difference. This simple method averages the images taken under rest and stimulation conditions and calculates a difference image between these two averages.
- The correlation method [1]. This method calculates the linear correlation coefficient between each time series and a reference pattern.
- The Fourier method [5]. In the Fourier transform of the time series, the amplitude at the stimulation frequency is used as an index to responsiveness to the stimulation.
- Principle component analysis (PCA). This method identifies activated brain regions in the eigenimages obtained by singular value decomposition of the data.

- The likelihood ratio test (8) proposed herein.

In simulated functional images the amplitude  $\hat{A}$  10 of a cosine waveform with phase  $\phi = \pi/2$ used as an activation function was varied from a value of 40 to 1000 in 25 equidistant steps, and from a value of 1000 to 5000 in another 20 equidistant steps. Using a noise standard deviation  $\sigma_n$  10 of 1000, totally 45 image data sets were produced. Each data set contained 64 images of  $128 \times 128$  pixels. The five different methods were used to produce activation maps from the generated image data sets with SNR increasing from 0.04 to 5. Statistical inference was automatically provided by the simulation environment by counting pixels inside and outside the simulated activation regions. For each method the false alarm probability  $P_f$  was constrained to a specified value while the detection probability  $P_d$  was evaluated for each SNR level. The plot in figure 3a shows the values of  $P_d$  for constrained  $P_f = 0.05$ . According to figure 3, the performance depending on the SNR shows more or less the same progression for different analysis methods. For SNR < 0.2 all the methods failed almost completely in detecting areas of activation, whereas for SNR > 3 all methods almost perfectly succeeded in the detection task, when the false alarm probability was constrained to 0.05. The correlation method is observed to be superior to the others. However, the improvement with respect to the other methods is not remarkable. The detection performance of the likelihood ratio test (8), the only method assuming the phase of the cosine waveform as unknown, is comparable to the other methods and justifies its application on real functional MRI data.

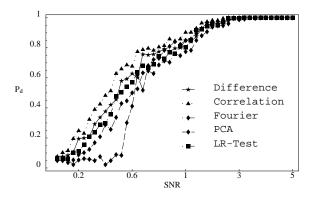


Fig. 3. Performance of different methods on simulated functional MRI data depending on the SNR. Detection probability  $P_d$  for constrained false alarm probability  $P_f = 0.05$ .

#### 3.2 Detection and Inference of Brain Activation by Likelihood Ratio Tests

A functional image data set consisting of 64 images ( $128 \times 128$  pixels) was acquired from a volunteer with a Bruker TOMIKON S200 scanner (University Children's Hospital, Zurich) using a gradient echo imaging sequence. Goggles with a flicker display identical to those used in electrophysiological examinations were used to visually stimulate the volunteer. During the stimulation state the volunteer was exposed to flickering red light at 8 Hertz. This flickering light was switched off in the rest state. Both states had a duration of 30 seconds. The experiment started with one cycle of rest and activation state without data acquisition so that the volunteer became accustomed to the stimulation. Then 8 images under the rest state were acquired followed by 8 images acquired under stimulation conditions. This stimulus alternation was totally repeated for four cycles. We have chosen a cosine waveform with phase  $\phi = \pi/2$  as a reference pattern for the hemodynamic response.

Figures 4a-c show the results obtained by the correlation method and the two likelihood ratio tests at a significance level of 0.05. In the activation map in figure 4a, obtained by the

correlation method, the result is not satisfactory because of many spurious pixels and regions all over the map. The map in figure 4b, obtained by the likelihood ratio test (4) gives a good visual impression about the extents of the activated brain areas. The spurious pixels in the background are reduced to a much greater degree than the result obtained by the correlation method. At the same time, there is no loss of pixels or regions within the head. In the activation map from figure 4c, obtained by the likelihood ratio test (8), almost no spurious pixels appear in the background. Although the phase of the cosine waveform was assumed to be unknown, the test still has succeeded to detect pixels forming activation regions similar to those in the map from figure 4b.

Figures 4d-f show the results obtained by the correlation method and the two likelihood ratio tests at a significance level of 0.005. The activation map obtained with the correlation method still shows some residual spurious pixels in the background. The maps obtained with the likelihood ratio tests show practically no pixels in the background.

Figures 4g-j illustrate the location of the detected regions from the maps 4d-f in an overlay with an original functional image. This visualization can help to identify falsely detected pixels lying in the background. It also depicts the extent of the regions consisting of detected pixels. In our example the activated regions lie within the cortical brain region which is known to be responsible for visual neuronal processing. No clear interpretation of the isolated pixels in brain areas far from the expected can be made. Usually, such pixels are considered as false positives on an empirical basis and can be removed from the activation map.

Although the detected regions look quite similar, in detail we observe slight differences. A visual inspection of the detected time series can empirically give an impression about the fidelity of the results. Therefore we have plotted some time series which were detected by the three methods within brain tissue in figures 4k-m. These time series were all selected from within the detected activation regions in the occipital lobe. The plot in figure 4k shows a time series which was detected by all three methods. The plots in figure 4l and 4m show time series detected only by one of the two likelihood ratio tests (4) and (8). In figure 4k we see that the time series signal nicely follows the switching from the rest state to the activation state (indicated by the grey bars). This effect is visually not so impressive in figure 4l, but still observable. Figure 4m shows a time series which seems to be related to the stimulus in an opposite way. In fact a difference in phase of 180° with respect to the time series in figure 4k,l may be suspected. It is up to a neurophysiologist to relate time series as shown in figure 4m to a hemodynamic response evoked by neuronal activation due to the visual stimulus.

# 4 Conclusion

We have applied the likelihood ratio tests (4) and (8) also on data sets with a large number of images comprising visual stimulation. Thereby, we have obtained similar results as those presented in figure 4. Differences in location between detected pixels were only encountered in small detail. The example demonstrated is therefore representative for the results we obtained from all analyzed data sets. This may lead to the conclusion that activation signals evoked by a neuronal stimulus show no phase shift between onset of stimulation and hemodynamic response. From the time series example in figure 4m however, we see that the likelihood ratio test (8) has detected a signal which obviously represents a phase shift of 180° between stimulation onset and response. Although this effect was detected only at a few pixel positions, it shows that detection procedures like the test (8) may contribute to the elaboration of unclear behaviour of hemodynamic effects.

We conclude that the likelihood ratio tests perform well on real data. Providing coequal results as other methods their power lies in allowing for the assumption of unknown parameters. The derivation of the likelihood ratio statistic (and the subsequent test procedures) is based on the probabilistic model of the noise in the data and therefore provides a direct approach to the detection problem. Other methods are based on the probabilistic consideration about a

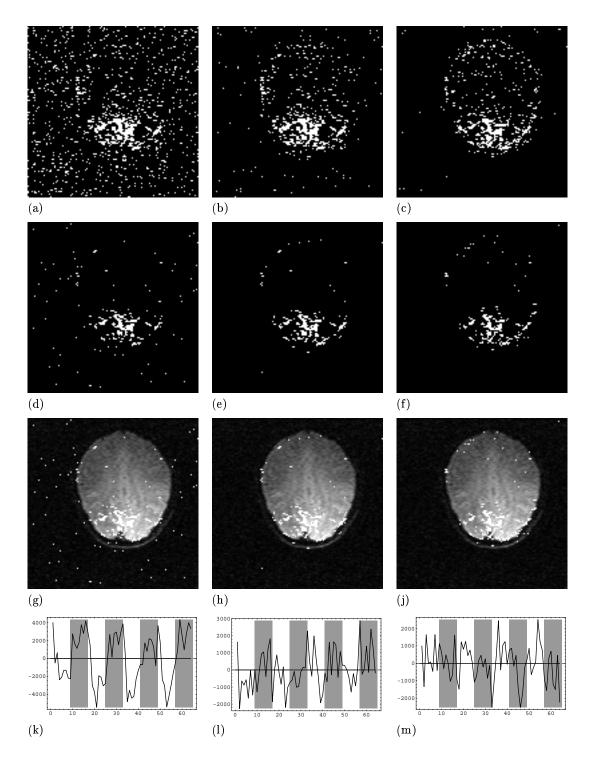


Fig. 4. Comparison between the activation map obtained by the correlation method and likelihood ratio tests (4) and (8), assuming a cosine waveform as a reference pattern. (a)-(c) Activation maps at a significance level of 0.05. (d)-(f) Activation maps at a significance level of 0.005. (g)-(j) Overlay of Activation maps (d)-(f) on a functional image from the investigated data set. (k) Time series detected by all three methods, (l) only by the test (4), (m) only by the test (8).

chosen metric (averaged difference, correlation, etc.) and therefore address the detection problem indirectly. Consequently, in a signal detection task based on hypothesis testing with the likelihood ratio, the statistical inference is made about the assumed signal model, and not about a metric computed from the model.

The design of a likelihood ratio test based on the Neyman-Pearson criterion does not require any knowledge about a priori probabilities or costs and is therefore well suited for the detection of brain activation in functional MRI. Further, the statistical inference is completely determined by the design (i.e., the constraint for the false alarm probability). This allows a user who is not familiar with thresholds to decide himself about the inference made from test results by supplying a probability value.

Taking a look at figure 4, the likelihood ratio tests are superior and more robust in making a statistical inference about detected regions compared with the correlation method [1]. The thresholds in our likelihood ratio tests all depend on the variance of the noise in the data. We have verified that the Gaussian white noise holds well in functional MRI data. We can therefore assume the value of the true noise variance from an estimate over a huge amount of pixels ( $\approx 4 \cdot 10^6$ ) within brain tissue, and use this estimate for the determination of the threshold in the likelihood ratio tests. Therefore, statistical inference is made by accounting for the noise variance for time series within brain tissue. The correlation method relates the threshold for the calculated correlation coefficients to a Gaussian distribution with zero mean and variance 1/N, N denoting the number of samples in the time series, which is an approximation to the true distribution of correlation coefficients. Consequently, the threshold for the correlation method does not account for the true noise variance. By comparing figure 4a and 4b it is seen that this can lead to an incorrect inference of the results.

# References

- 1. P. A. Bandettini, A. Jesmanowicz, E. C. Wong, and J. S. Hyde. Processing strategies for time-course data sets in functional MRI of the human brain. *Magnetic Resonance in Medicine*, 30:161–173, 1993.
- 2. P. A. Bandettini, E. C. Wong, R. S. Hinks, et al. Time course EPI of human brain function during task activation. *Magnetic Resonance in Medicine*, 25:390-397, 1992.
- 3. M. Barkat. Signal Detection & Estimation. Artech House, Boston, 1991.
- 4. A. M. Blamire, S. Ogawa, K. Ugurbil, et al. Dynamic mapping of the human visual cortex by high-speed magnetic resonance imaging. In *Proc. Natl. Acad. Sci. USA*, volume 89, pages 11069–11073, 1992.
- 5. K. Boulanouar, I. Berry, J. F. Demonet, et al. A monofrequential approach to the detection of activated pixels in fMRI. In NeuroImage, Second International Conference on Functional Mapping of the Human Brain, June 17-21, 1996, Boston, MA, page 51, 1996.
- G. E. P. Box, G. M. Jenkins, and G. C. Reinsel. Time Series Analysis. Prentice Hall, New Jersey, 3rd edition, 1994.
- 7. G. M. Boynton, S. A. Engel, G. H. Glover, and D. J. Heeger. Linear systems analysis of functional magnetic resonance imaging in human V1. *Journal of Neuroscience*, 16(13):4207-4221, 1996.
- 8. E. Bullmore, M. Brammer, S. C. R. Williams, et al. Statistical methods of estimation and inference for functional MR image analysis. *Magnetic Resonance in Medicine*, 35:261-277, 1996.
- 9. D. Ekatodramis, G. Székely, E. Martin, and G. Gerig. A simulation environment for validation and comparison of fMRI evaluation procedures. In NeuroImage, Second International Conference on Functional Mapping of the Human Brain, June 17-21, 1996, Boston, MA, page 57, 1996.
- K. J. Friston, C. D. Frith, R. Turner, and R. S. Frackowiak. Characterizing evoked hemodynamics with fMRI. NeuroImage, 2:157-165, 1995.
- 11. K. K. Kwong, J. W. Belliveau, et al. Dynamic magnetic resonance imaging of human brain activity during primary sensory stimulation. In *Proc. Natl. Acad. Sci. USA*, volume 89, pages 5675–5679, 1992.
- 12. W. H. Press, B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling. *Numerical Recipes in C.* Cambridge University Press, Cambridge, 1988.
- 13. H. L. Van Trees. Detection, Estimation, and Modulation Theory. John Wiley, New York, 1968.
- 14. AVS Developer's Guide. Advanced Visual Systems Inc.