ELSEVIER

NeuroImage

www.elsevier.com/locate/ynimg
NeuroImage 36 (2007) 361–369

# BOLD response analysis by iterated local multigrid priors

Selene da Rocha Amaral,* Said R. Rabbani, and Nestor Caticha

*Instituto de Física, Universidade de São Paulo, São Paulo, SP, Brazil*

Received 7 April 2006; revised 13 November 2006; accepted 22 November 2006
Available online 19 December 2006

**We present a non parametric Bayesian multiscale method to characterize the Hemodynamic Response HR as function of time. This is done by extending and adapting the Multigrid Priors (MGP) method proposed in (S.D.R. Amaral, S.R. Rabbani, N. Caticha, Multigrid prior for a Bayesian approach to fMRI, NeuroImage 23 (2004) 654–662; N. Caticha, S.D.R. Amaral, S.R. Rabbani, Multigrid Priors for fMRI time series analysis, AIP Conf. Proc. 735 (2004) 27–34). We choose an initial HR model and apply the MGP method to assign a posterior probability of activity for every pixel. This can be used to construct the map of activity. But it can also be used to construct the posterior averaged time series activity for different regions. This permits defining a new model which is only data-dependent. Now in turn it can be used as the model behind a new application of the MGP method to obtain another posterior probability of activity. The method converges in just a few iterations and is quite independent of the original HR model, as long as it contains some information of the activity/rest state of the patient. We apply this method of HR inference both to simulated and real data of blocks and event-related experiments. Receiver operating characteristic (ROC) curves are used to measure the number of errors with respect to a few hyperparameters. We also study the deterioration of the results for real data, under information loss. This is done by decreasing the signal to noise ratio and also by decreasing the number of images available for analysis and compare the robustness to other methods.**


*Keywords:* fMRI; Hemodynamic response; Bayesian data analysis; Prior information

## Introduction

The main effort in clinical and cognitive neuroscience using functional magnetic resonance imaging (fMRI) (Belliveau et al., 1992; Kwong et al., 1992; Ogawa et al., 1992) has focused most intensely on the localization of brain activity using the Blood Oxygenation Level-Dependent (BOLD) effect. Understanding brain function requires information not only of the localization of brain activity, but also of features that characterize this response (Friston et al., 1998; Buxton et al., 2004).

The translation of susceptibility into brain activity via $T_2^*$ weighted imaging is not an altogether clear process. It involves levels of blood oxygenation, regional characteristics of cerebral blood flow and volume, glucose and oxygen metabolism, running down to the level of neuronal activity. The MR signal carries, through this long chain of relations, information about neuronal activity. However, this coding occurs through an unknown and nonlinear, hopefully reasonably stable and monotonic process that depends on this activity. This information then bears, besides the spatial localization of the activity, on the regional dynamics governed by this function which is usually called the hemodynamic response (HR), which depends on time due to the dynamical nature of the neuronal activity. It has played a central role in understanding functional characterization of brain function (Buckner, 1998; Logothetis et al., 2001). Most analyzes model BOLD activity based upon some standard HR. However, several works showed that HRs widely vary from region to region, from task to task, and from subject to subject (Aguirre et al., 1998; Miezin et al., 2000; Handwerker et al., 2004).

These studies suggest that the activation results obtained using a generic HR may be different from those obtained using local hemodynamic responses. It can be case that HR significantly changes across regions of activation. In this case, the use of region-specific HRs may help to delineate more precisely the activated areas.

In this paper, we introduce a non parametric method for BOLD response analysis by iterated local multigrid priors (iMGP) a Bayesian method that deals with the multiscale properties of the problem. This is done by extending and adapting the Multigrid Priors (MGP) method proposed by Amaral et al. (2004) and Caticha et al. (2004). We choose an initial HR model which could be as simple as an ON/OFF square function and apply the MGP method to assign a posterior probability of activity to every pixel. This can be used to identify regions of activity. But it can also be used to construct the posterior averaged time series of activity for different regions. This permits defining a new model which is only data-dependent. Now in turn it can be used as the model behind a new application of the MGP method to obtain another posterior probability of activity, and so on in an iterative fashion. The method converges in just a few iterations and very importantly, is quite independent of the original HR model, as long as it contains some information of the activity/rest state of the patient.

* Corresponding author. Departamento de Física Geral, Rua do Matão 187, São Paulo, SP 05508-900, Brazil.
*E-mail address:* selene@if.usp.br (S. da Rocha Amaral).
**Available online on ScienceDirect (www.sciencedirect.com).**

1053-8119/$ - see front matter © 2006 Elsevier Inc. All rights reserved.
doi:10.1016/j.neuroimage.2006.11.058

We have applied our extended method to simulated as well as experimental fMRI data sets. We quantify the performance of the method, in a classical way, through receiver operating characteristic (ROC) curves which characterize the sensitivity and specificity relation. Different curves were obtained by choosing different parametrizations in terms of different hyperparameters, such as global prior probabilities, noise level, HF amplitude, the size of the region over which the data set is spatially averaged over the posterior of activity, or the iteration number. We also study the deterioration of the results for real data, under information loss. This is done by decreasing the signal to noise ratio and also by decreasing the number of images available for analysis and compare the robustness to Bayesian and Classical methods of the SPM2 software package.

Despite the fact that Bayesian methods strive to attain objectivity, a wealth of different Bayesian methods have appeared in the literature due to the fact that different practitioners choose to codify prior information in different manners. In particular, the realization that spatial dependencies should influence the prior has been used in several studies (e.g. Gossl et al., 2001; Woolrich et al., 2004; Penny et al., 2003; Selene da Rocha et al., 2004; Penny et al., 2005). There is a variety of multiscale methods that can be borrowed from different areas of analysis and physics. In particular, wavelet decomposition (Turkheimer et al., 2000) has brought in an extension of spatio-temporal Fourier analysis. Multigrid techniques, originally designed to deal with multiscale phenomena described by partial differential equations (Fedorenko, 1964), have been found useful to accelerate Monte Carlo simulations (Goodman and Sokal, 1989). While related, by the multiscale theme, these are different in practical details, as are different from another relative, the Renormalization Group (Swendsen, 1979). Our method, inspired by multiscale analysis, as the Renormalization Group, technically resembles more a multigrid type, since it avoids dealing with the calculation of costly marginals. While the differences between these multiscale theoretical tools are clear, the similarities and applicability scopes have not yet been clearly determined. Maximum Entropy methods have been shown to be very useful in dealing with problems of Fourier inversion in the presence of noisy data and reduced K-space information. It is tempting to speculate that in this case Bayesian and the related method of Maximum Entropy, with a multigrid flavor may be brought to complement wavelet methods.

The explosion of Bayesian applications in the last decades is partly due to the fact that efficient Monte Carlo methods have permitted integrating over large dimensional spaces. Therefore many times Bayesian methods are feared for the heavy computational load they might imply. As we will show bellow, our method does not need Monte Carlo integration over nuisance parameters or equivalent mean field approximations and so it is very fast, taking no more than 2 min per slice on a small personal computer.

## Methods

In this section we explain how to infer the hemodynamic response as a function of time. Since this is based on an iterated extension of the Multigrid Grid Priors method, we now briefly review it.

### *fMRI analysis by multigrid priors*

The main idea behind the MGP algorithm (see Selene da Rocha et al., 2004; Caticha et al., 2004 for details) stems from the fact that a Bayesian hypothesis test gives better results the richer the prior information used. A Bayesian approach must address two main issues: the prior probability and the likelihood which is based on the model of the physical effect and the knowledge about the noise process that affects the data. Forgetting to model the prior, or just choosing a uniform prior leads to the maximum likelihood method. In Amaral et al. (2004), we showed, in a realistic numerical simulations, that the maximum likelihood-based hypothesis test had 300% larger errors than a simple Bayesian single pixel test that used correct global activity to guide the choice of the prior. Of course, it can be claimed that this global activity is unknown. To answer this, we can think of two possible strategies. The first is to give up: prior knowledge of this type is not available and therefore such improvement can't be reached. The second, much more optimistic, way of facing this, is to realize that new methods are needed that address the prior. In fact it strongly suggests seeking any information that had not been used before. We devised a method in which the prior information is to be found in the spatial correlations of the data.

We are looking for activity which, although fairly localized, appears in regions that are typically much larger than a single voxel. This piece of information about the spatial correlation of activity can be translated into useful prior distribution through a multiscale approach. The prior is constructed by looking sequentially at different spatial scales.

We first deal with the geometry. Define the multigrid: a hierarchical sequence, labeled by $q=0, 1,\ldots, Q$, of square lattices of size $2^{Q-q}\times 2^{Q-q}$. Label the sites of the lattices by $p^q=(p_x^q, p_x^q)$ with $p_x^q, p_x^q=1,\ldots, 2^{Q-q}$. This can be easily extended to three (voxels instead of pixels) or more dimensions. Each lattice is composed of sites which can be thought of as coarse grained pixels or $q$-pixels. Each $q$-pixel comprises a region $B_q$ made up of $2^q\times 2^q$ original pixels (0-pixels). For $q=0$, the lattice is just the original lattice of the individual pixels. For $q=Q$ the lattice is reduced to a single $Q$-pixel.

Now we describe the data. The data are made up of the set of time series $D^0=\{Y_{p^0}\}_{t=0,\ldots,T}$ one time series, of length $T+1$ for each pixel at site $p^0$. The data are used to define coarse grained $q$-scale data sets: $D^q=Y_{pq}(t)$, defined for each of the scales labeled by $q$, by spatial averaging in the spirit of a Renormalization Group coarse grain transformation:

$$Y_{p^q}(t)=\frac{1}{2^2}\sum_{p^{q-1}\epsilon B_{p^q}} Y_{p^{q-1}}(t)=\frac{1}{2^{2q}}\sum_{p^0\epsilon B_{p^q}} Y_{p^0}(t) \tag{1}$$

To implement the Bayesian approach, we need a model for the data and the noise, that is, a function of time, $M_{p^q}(t)$ for each pixel $p$ when it is active. Since its determination is the main topic of this paper, we will describe this problem later. Consider for now that $M_{p^q}(t)$ is given, to be specific consider that it is an ON/OFF square function, we can then by a method analogous to that of Eq. (1) define a coarse grained $q$-model:

$$M_{p^q}(t)=\frac{1}{2^2}\sum_{p^{q-1}\epsilon B_{p^q}} M_{p^{q-1}}(t)=\frac{1}{2^{2q}}\sum_{p^0\epsilon B_{p^q}} M_{p^0}(t) \tag{2}$$

At any scale $q$ we define the set of hypotheses, one for each $q$-pixel: $H_1^q$: "There is activity in the $q$-pixel $p^q$"

From models $M_{pq}(t)$, data $D^q=\{Y_{pq}(t)\}$ and a choice of a noise process we construct a likelihood: $P(D^q|H_1^q)$, a probability that codifies the information that the data would have been observed if the $q$-pixel were active.