NeuroImage 109 (2015) 217–231

Contents lists available at ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/ynimg

Large-scale Probabilistic Functional Modes from resting state fMRI

Samuel J. Harrison [a,b,c,*], Mark W. Woolrich [a,b], Emma C. Robinson [a], Matthew F. Glasser [d], Christian F. Beckmann [a,e], Mark Jenkinson [a], Stephen M. Smith [a]

[a] Oxford Centre for Functional Magnetic Resonance Imaging of the Brain (FMRIB), Oxford, UK
[b] Oxford Centre for Human Brain Activity (OHBA), Oxford, UK
[c] Life Sciences Interface Doctoral Training Centre (LSI-DTC), Oxford, UK
[d] Department of Anatomy and Neurobiology, Washington University, Medical School, St. Louis, MO, USA
[e] Donders Institute for Brain, Cognition and Behaviour, Radboud University Nijmegen, The Netherlands

ARTICLE INFO

Article history:
Accepted 1 January 2015
Available online 15 January 2015

Keywords:
Resting state fMRI
Functional parcellation
Bayesian modelling
Subject variability
ICA

ABSTRACT

It is well established that it is possible to observe spontaneous, highly structured, fluctuations in human brain activity from functional magnetic resonance imaging (fMRI) when the subject is 'at rest'. However, characterising this activity in an interpretable manner is still a very open problem.

In this paper, we introduce a method for identifying modes of coherent activity from resting state fMRI (rfMRI) data. Our model characterises a mode as the outer product of a spatial map and a time course, constrained by the nature of both the between-subject variation and the effect of the haemodynamic response function. This is presented as a probabilistic generative model within a variational framework that allows Bayesian inference, even on voxelwise rfMRI data. Furthermore, using this approach it becomes possible to infer distinct extended modes that are correlated with each other in space and time, a property which we believe is neuroscientifically desirable.

We assess the performance of our model on both simulated data and high quality rfMRI data from the Human Connectome Project, and contrast its properties with those of both spatial and temporal independent component analysis (ICA). We show that our method is able to stably infer sets of modes with complex spatio-temporal interactions and spatial differences between subjects.

Introduction

Using resting state fMRI it is possible to generate enormously rich data sets that capture some of the complexity of the brain's intrinsic dynamics and connectivity. However, generating representations that meaningfully simplify the data, while still capturing these dynamics, is an immensely challenging problem.

Initial analyses of rfMRI data focused on finding regions of highly correlated activity (Biswal et al., 1995), with spatial independent component analysis (sICA) coming to prominence as a robust method for extracting regions consistent with knowledge from task analyses (Kiviniemi et al., 2003; Smith et al., 2009).

Recently, there has been much interest in techniques which analyse functional connectivity across the brain, including the potentially time-varying or non-stationary nature of these connections (E.A. Allen et al., 2014; Baker et al., 2014; Cribben et al., 2012; Seghier and Friston, 2013). However, for all but the simplest analysis techniques it is necessary to work in a lower dimensional space than the hundreds of thousands of voxels in a typical rfMRI data set. This is typically achieved either by extracting parcels from an anatomical atlas, or using high-dimensional sICA (Kiviniemi et al., 2009; Smith et al., 2013a). However, it is well known that "[i]nconsistent or imprecise node definitions can have a major impact on subsequent analyses" (Fornito et al., 2013), which again throws the question of how best to generate meaningful representations of resting state activity into sharp relief.

Therefore, an aim has become to find an interpretable and robust way of representing rfMRI data, at the same time as capturing as much of the complex temporal dynamics as possible.

Definitions

For this paper, we will use the following definitions. We will take a *network* to be a set of interacting elements—synonymous with the mathematical formalism of a graph as a set of nodes and edges. Functional connections, that is to say the edges between nodes, may vary in their presence and strength over time.

* Corresponding author at: Oxford Centre for Functional Magnetic Resonance Imaging of the Brain (FMRIB), Oxford, UK.
E-mail address: samuel.harrison@ndcn.ox.ac.uk (S.J. Harrison).

http://dx.doi.org/10.1016/j.neuroimage.2015.01.013

We define a *parcel* to be a set of voxels acting with a single representative time course. These are often derived from a 'hard' parcellation of grey matter into multiple non-overlapping regions (Rubinov and Sporns, 2010; Yeo et al., 2011; Craddock et al., 2012). However, given the trend for using components from a high-dimensional sICA for connectivity analyses (E.A. Allen et al., 2014; Kiviniemi et al., 2009; Smith et al., 2013a), we relax this definition slightly. In the spatial domain, a parcel is taken to represent a set of positive weights, potentially varying in magnitude, with limited overlap between different parcels. The definition we have given therefore allows, for example, blurred boundaries or parcels that contain bilaterally paired regions.

We define a *mode* as any spatial distribution over the brain that shares a common time course. This is similar to a parcel, but the definition is wider as this imposes no restrictions on the spatial properties. For example, multiple modes can be highly overlapping, and individual modes can include anti-correlated regions (meaning that some regions within the mode have a negative spatial weight and others have a positive one). A mode—as an extended spatial distribution having common temporal dynamics—can be defined either in terms of a spatial voxelwise map, or as a weighted set of spatial parcels.

In general, it is possible to take the time courses from either parcels or modes and use these as the nodes to examine in a subsequent network analysis, but we will focus on modes here.

Current methods

Many techniques have been proposed to identify modes or parcels. Perhaps the simplest is to extract time courses from labelled regions in a pre-defined anatomical atlas, though the validity of this has been called into question as the correspondence between anatomical landmarks and functional regions is unclear (Fornito et al., 2013). The obvious alternative is to use a pre-defined atlas containing regions based on previous functional studies, an approach which is likely to have a higher validity.

However, the arguable weakness of atlas-based approaches is their reliance on the registration process to enforce consistency across subjects. There is an enormous amount of interesting structure present in rfMRI data, and it seems reasonable to assume that this could be harnessed to inform the specification of functional regions. In fact, one of the key assertions we make in this paper is that it is possible to attempt to use the characteristics of the rfMRI data to correct for subject mis-alignments.

There have therefore been a large number of strategies proposed that attempt to infer functional regions from the data—so called 'data-driven' approaches. Temporally consistent co-activation is the implicit assumption that defines both parcels and modes, but by itself this does not lead to a unique decomposition. Therefore, it is necessary to add additional constraints to make the inference problem identifiable.

The most widely used data-driven approach is to look for modes that are independent using ICA. Due to the large numbers of voxels and relatively few time points of early studies, spatial ICA gave the most robust decompositions and therefore became the dominant approach. However, almost as soon as it was introduced, concerns were raised. Given that "[distinct] large scale neuronal dynamics can share a substantial anatomical infrastructure" (Friston, 1998; Smith et al., 2012), it is unclear how well sICA will decompose extended modes that spatially overlap. These concerns were allayed to some extent by Beckmann et al., who showed that in the presence of noise ICA components can still contain strong residual dependencies, and highly correlated maps can be recovered by a simple thresholding step (Beckmann et al., 2005). What is perhaps less clear is what, if any, biases are introduced when the data has a high SNR or when large groups are analysed, both cases where the inferred maps are expected to contain very little noise.

An alternative approach is to look for temporally independent modes, which has recently become possible as studies of large cohorts acquired at low TR have generated enough time points for temporal ICA (tICA) to operate robustly (Smith et al., 2012), albeit still most likely requiring the concatenation of several fMRI data sets to achieve reasonable reproducibility. This allows spatially overlapping modes to be identified, at the expense of placing restrictions on the global temporal dynamics—as well as this being a concern in and of itself, this restriction will also limit any subsequent network analyses of the mode time courses. As Smith et al. discuss, temporally independent functional modes (TFMs) are forced to have orthogonal time courses, meaning that further analysis of the temporal interactions between different modes is not straightforward (Smith et al., 2012).

As well as the choice of spatial or temporal independence, various extensions have been proposed to extract meaningful subject-specific information from group ICA decompositions (Damoiseaux et al., 2006; Filippini et al., 2009; Varoquaux et al., 2010; Erhardt et al., 2011).

While each ICA strategy has its own advantages, the fundamental issue with all ICA-based approaches is that "it is not clear that, from a neuroscientific point of view, independence is the right concept to isolate brain networks, as no functional system is fully segregated" (Varoquaux et al., 2010). What is perhaps surprising is how demonstrably well ICA approaches work, given that their central assumptions are often violated (Hyvärinen, 2013); for example, forms of ICA have been developed that explicitly incorporate information derived from the residual statistical dependencies between components (Hyvärinen and Hoyer, 2000; Hyvärinen et al., 2001). Therefore, while ICA approaches have been particularly useful for characterising fMRI data, one would hope that a less restrictive set of assumptions could engender decompositions with even higher validities.

Other data-driven approaches suggested have had varying degrees of success. Many are based on machine learning techniques, where the key assumptions underpinning the algorithms are only loosely related to the expected properties of rfMRI data. These include clustering approaches (Yeo et al., 2011; Craddock et al., 2012), regularised variants of principal component analysis (PCA) (G.I. Allen et al., 2014), non-negative matrix factorisation (Lee et al., 2011), image gradient detection in correlation maps (Cohen et al., 2008) and hidden Markov models (Eavani et al., 2013) to name but a few.

Finally, there are a few approaches which try to explicitly model rfMRI data. The multi-subject dictionary learning (MSDL) approach of Varoquaux et al. (2011) forms a model that explicitly looks for modes/parcels, and there are some conceptual similarities with our approach. Their algorithm contains a hierarchical model for spatial subject variability, a constraint favouring simultaneously smooth and sparse spatial distributions as well as the ability to capture the temporal correlations between modes.

Due to the similarities between our approaches, we will give a brief description of the most recent version of their model, of which more details can be found in the work of Abraham et al. (2013). Their spatial model at the group level is detailed, simultaneously enforcing non-negativity, sparsity and spatial contiguity. The subject maps are modelled by including a set of additive, Gaussian-distributed deviations from the group maps. Their time series model specifies that there should be a consistent between-mode correlation structure but does not restrict the form of the time series; therefore, it does not model any haemodynamic processes. Finally, these constraints are combined with a noise model, and the resulting cost-function governing their decomposition is solved with a computationally efficient stochastic gradient descent approach. The most recent results they report are on an rfMRI data set consisting of 48 subjects.

In this paper, we develop an analysis technique that explicitly models some of the key properties of resting state modes within a Bayesian framework. The Bayesian approach allows very flexible models to be constructed in a principled manner; crucially, we solve the system using a variational approach, thereby making the algorithm efficient enough to work on full fMRI data sets.