NeuroImage xxx (2016) xxx–xxx

Contents lists available at ScienceDirect

NeuroImage

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

The effect of spatial resolution on decoding accuracy in fMRI multivariate pattern analysis

Anna Gardumi *, Dimo Ivanov, Lars Hausfeld, Giancarlo Valente, Elia Formisano, Kâmil Uludağ

[a] Department of Cognitive Neuroscience, Faculty of Psychology and Neuroscience, Maastricht University, Maastricht, The Netherlands
[b] Maastricht Brain Imaging Center, Maastricht University, Maastricht, The Netherlands

ARTICLE INFO

Article history:
Received 11 September 2015
Accepted 10 February 2016
Available online xxxx

Keywords:
Multivariate pattern analysis
fMRI
7 T
Spatial resolution
Spatial smoothing
Auditory cortex

ABSTRACT

Multivariate pattern analysis (MVPA) in fMRI has been used to extract information from distributed cortical activation patterns, which may go undetected in conventional univariate analysis. However, little is known about the physical and physiological underpinnings of MVPA in fMRI as well as about the effect of spatial smoothing on its performance. Several studies have addressed these issues, but their investigation was limited to the visual cortex at 3 T with conflicting results. Here, we used ultra-high field (7 T) fMRI to investigate the effect of spatial resolution and smoothing on decoding of speech content (vowels) and speaker identity from auditory cortical responses. To that end, we acquired high-resolution (1.1 mm isotropic) fMRI data and additionally reconstructed them at 2.2 and 3.3 mm in-plane spatial resolutions from the original k-space data. Furthermore, the data at each resolution were spatially smoothed with different 3D Gaussian kernel sizes (i.e. no smoothing or 1.1, 2.2, 3.3, 4.4, or 8.8 mm kernels). For all spatial resolutions and smoothing kernels, we demonstrate the feasibility of decoding speech content (vowel) and speaker identity at 7 T using support vector machine (SVM) MVPA. In addition, we found that high spatial frequencies are informative for vowel decoding and that the relative contribution of high and low spatial frequencies is different across the two decoding tasks. Moderate smoothing (up to 2.2 mm) improved the accuracies for both decoding of vowels and speakers, possibly due to reduction of noise (e.g. residual motion artifacts or instrument noise) while still preserving information at high spatial frequency. In summary, our results show that – even with the same stimuli and within the same brain areas – the optimal spatial resolution for MVPA in fMRI depends on the specific decoding task of interest.

1. Introduction

Functional magnetic resonance imaging (fMRI) is currently the most popular non-invasive method to investigate human brain structure and function. It indirectly measures neural activity primarily via the blood oxygenation level-dependent (BOLD) effect. Standard univariate statistical analysis (i.e. general linear model (GLM) analysis) of the task-based fMRI data has been utilized to detect voxel-wise differences of BOLD activation levels and, thus, to infer which brain areas are involved in a certain task. In recent years, multivariate pattern analysis (MVPA) has been used in fMRI to extract information from spatially distributed activation patterns, which may go undetected in conventional univariate analysis. Reliable decoding of information from fMRI data acquired at 3 T has been demonstrated from activation patterns in different brain areas (Haxby et al., 2001; Cox and Savoy, 2003; Haynes and Rees, 2005; Kamitani and Tong, 2005; Kriegeskorte and Bandettini, 2007; Formisano et al., 2008). Different biophysical hypotheses have been proposed to explain the ability of MVPA on fMRI data to detect information inaccessible with GLM. It has been suggested that MVPA is sensitive to information encoded at the sub-millimeter scale of neuronal functional columns. Such information, even if sampled at the lower resolution of standard fMRI voxel sizes (e.g. $3 \times 3 \times 3$ mm^3), may be accessible by MVPA due to local variations and irregularities in the columnar organization, resulting in weak but consistent biases in fMRI responses of the different voxels (Boynton, 2005; Kamitani and Tong, 2005; Haynes and Rees, 2006; Kamitani and Tong, 2006; Kriegeskorte and Bandettini, 2007); this mechanism is, therefore, named *hyperacuity* or *voxel biased sampling*. Alternatively, the transposition from high spatial frequency components of columns preferences to lower spatial frequency of the fMRI signal may be attributed to the cortical vasculature. This hypothesis is based on the fact that, using the standard gradient echo (GE) MRI sequences, the fMRI signal stems mostly from veins draining blood from a given tissue volume (see Uludag et al., 2009). Thus, a specific vein could be more sensitive to one neuronal population than another introducing a spatial bias. Hence, this hypothesis is known as *biased draining regions* (Kamitani and Tong, 2005; Gardner et al., 2006; Kamitani and Tong, 2006; Kriegeskorte and Bandettini, 2007; Gardner, 2010; Kamitani and Sawahata, 2010; Kriegeskorte et al., 2010; Shmuel et al., 2010).

* Corresponding author at: Department of Cognitive Neuroscience, Faculty of Psychology and Neuroscience, Maastricht University, P.O. Box 616, 6200MD Maastricht, The Netherlands.
E-mail address: anna.gardumi@maastrichtuniversity.nl (A. Gardumi).

http://dx.doi.org/10.1016/j.neuroimage.2016.02.033
1053-8119/

According to another hypothesis, MVPA may rely on large spatial scale non-columnar organization (Op de Beeck, 2010), such as radial preference maps (Freeman et al., 2011). Since MVPA represents a computational scheme to non-locally average the fMRI signal, in this framework, MVPA would be able to detect low spatial frequency information too weak to be detected with univariate analysis.

Note that these hypotheses are not mutually exclusive (see Shmuel et al., 2010; Swisher et al., 2010). Nevertheless, they do predict testable effects of spatial smoothing on decoding performance. Op de Beeck has shown that spatial smoothing does not deteriorate decoding performance of objects and orientations from activation patterns in lateral occipital cortex and V1, respectively (Op de Beeck, 2010). He interpreted these results as an argument against hyperacuity and in favor of large-scale organization. Further support for this hypothesis comes from the finding that it is possible to decode across experimental sessions performed in different days (Freeman et al., 2011). In contrast, several studies (Swisher et al., 2010; Alink et al., 2013; Misaki et al., 2013) demonstrated that spatial smoothing decreases decoding accuracies for orientation and ocular dominance from V1 data, suggesting relevant information content at the individual voxel level. The few investigations so far on the underlying mechanisms of MVPA on fMRI data and the effect of spatial smoothing have been limited to the early visual cortex. In addition, they have been restricted to a small set of stimuli and decoding tasks (e.g. decoding of orientation, ocular dominance, and direction of motion) and have yielded conflicting evidence.

The main goal of the current study is to investigate how information at different spatial resolutions contributes to MVPA decoding. We employed ultra-high field (7 T) fMRI to acquire high-resolution data (1.1 mm isotropic), which were then reconstructed at different effective spatial resolutions from original k-space data to evaluate the effects of spatial resolution on MVPA decoding performance. Based on an experimental paradigm and on stimuli that were used in a previous fMRI study at 3 T (Formisano et al., 2008), we presented speech stimuli (vowels) from different speakers and considered the single-trial decoding of vowels and speakers from auditory cortical response patterns. Compared to conventional 3 T fMRI, 7 T fMRI presents several advantages, such as higher signal-to-noise ratio (SNR) and contrast-to-noise ratio (CNR), and therefore the possibility of higher spatial resolution with lower partial volume effects and greater spatial specificity (Yacoub et al., 2005; Uludag et al., 2009; Polimeni et al., 2010). On the other hand, it presents challenges such as larger distortions, sensitivity to motion, and larger number of voxels to be handled by the decoding algorithm (Formisano and Kriegeskorte, 2012). Therefore, we also investigated the effects of temporal SNR, CNR, and head motion and of typical noise-reduction steps (spatial smoothing) on MVPA performances.

2. Material and methods

2.1. Subjects

Ten healthy volunteers (seven females, age range 25–32) with normal hearing took part in this experiment. Informed consent was obtained from all participants according to the approval by the Ethical Committee of the Faculty of Psychology and Neuroscience, University of Maastricht.

2.2. Stimuli and task

We used the same auditory stimuli as in the study of (Formisano et al., 2008) consisting of three vowels (/a/, /i/, /u/) spoken by three different speakers (sp1: female, sp2: male, sp3: male). For each of these 9 conditions, three different tokens were included in order to introduce acoustic variability. All stimuli were equated in length to 230 ms and in sound intensity by matching their root mean square amplitude. For more details about the stimulus properties, please see (Formisano et al., 2008). Prior to the functional experiment, participants were familiarized with the stimuli and were able to recognize the corresponding vowels and speakers. During the fMRI experiment, subjects were instructed to attentively listen to the stimuli while fixating a white cross in the center of the screen. The stimuli were presented in the silent gap between two subsequent image acquisitions (see below).

In order to ensure the engagement of the participants in both listening and fixating tasks, the participants performed a one back-task on the speaker dimension irrespectively of the spoken vowel: 10% of the total number of trials were catch-trials (signaled to the participants by the fixation cross turning red for 100 ms), in which the subjects were asked to report whether the speaker of the last heard sound was the same as the previous one. Subjects performed the task by pressing a button with either the index ("Yes"-answer) or the middle ("No"-answer) finger of the right hand. Catch-trials were excluded from all subsequent analyses.

The sounds were played according to a slow-event related design with a variable interstimulus interval (ISI) of 6–8 TRs (TR = 2500 ms, average ISI 17.5 s). At the beginning of the fMRI session, the volume of the stimuli was adjusted to a comfortable intensity level. The stimuli were presented in the 500 ms silent gap via MR-compatible earphones (Sensimetrics S14, Malden, MA, USA). After the experiment, all subjects reported a clear hearing of the stimuli. Every run consisted of 5 trials for each of the 9 stimulus conditions and 5 catch-trials, resulting in a total of 50 trials and a run duration of approximately 15 min. The order of stimulus presentation was randomized within and across runs. Four functional runs were acquired, leading to a total of 200 trials in the whole experiment.

2.3. Data acquisition

Functional and anatomical images were acquired with a 7 T Siemens Magnetom scanner using a 32-channel Nova Medical head coil. Four high-resolution (1.1 mm isotropic voxel size) functional runs were acquired using a gradient-echo (GE) EPI sequence (Moeller et al., 2010) with the following parameters: TR 2500 ms, TE 22 ms, Partial Fourier 5/8, GRAPPA 2, delay in TR 500 ms, multi-band acceleration factor 2 with blipped-CAIPIRINHA (1/FOV shift 4; Setsompop et al., 2012). The sequence was optimized to maximize tSNR in the auditory cortex. In two separate pilot runs of 140 volumes (~6 min, resting state), we acquired the sequence with these parameters and additionally a variant with GRAPPA 3, Partial Fourier 6/8 and TE 24.4 ms. The latter showed less distortions and signal dropout only in the anterior and posterior parts of the brain albeit with a lower tSNR in the auditory cortex (23.27 versus 34.02, respectively).

In addition to the magnitude images, phase images were collected in order to allow image reconstruction with lower voxel resolution (see below for details). 48 slices were acquired centered approximately on the superior temporal gyrus, covering the auditory cortex. One high-resolution (0.7 mm isotropic voxel size) anatomical image covering the whole brain was collected using MP2RAGE sequence (Marques et al., 2010).

2.4. Data analysis: Preprocessing and univariate analysis

Functional and anatomical data were preprocessed and analyzed in BrainVoyager QX 2.8.2 (Brain Innovation). The four functional runs were 3D motion corrected and coregistered to the first volume of the first run through rigid-body transformation (3 translational and 3 rotational parameters). Neither nonlinear transformation nor distortion correction algorithm were applied to avoid interpolation confounds in our comparison across resolutions. We visually inspected every coregistered run and no large motion was observed. Linear and low-frequency non-linear drifts up to 7 cycles per time course were removed via temporal high-pass filtering. This cut-off frequency, corresponding to a cut-off period of ~128 s, was adequate to the stimulus design and analyses here employed (as estimated though spectral analysis of the class stimulus design). For each subject, the anatomical image was segmented at the gray-white matter boundary via an automatic procedure.

Please cite this article as: Gardumi, A., et al., The effect of spatial resolution on decoding accuracy in fMRI multivariate pattern analysis, NeuroImage (2016), http://dx.doi.org/10.1016/j.neuroimage.2016.02.033