Spatiotemporal Searchlight Representational Similarity Analysis in EMEG Source Space

Li Su, Elisabeth Fonteneau, William Marslen-Wilson

Department of Experimental Psychology
University of Cambridge, Cambridge, UK
ls514@cam.ac.uk

Nikolaus Kriegeskorte

MRC Cognition and Brain Sciences Unit
Cambridge, UK

Abstract—**Time resolved imaging techniques, such as MEG and EEG, are unique in their ability to reveal the rich dynamic spatiotemporal patterning of neural activities. Here we propose a technique based on spatiotemporal searchlight Representational Similarity Analysis (RSA) of combined MEG and EEG (EMEG) data to directly analyze the multivariate pattern of information flow across the brain. This novel technique can recognize fine-grained dynamic neural computations both in space and in time. A prime example of such neural computations is our ability to understand spoken words in real time. A computational approach to these processes is suggested by the Cohort Model of spoken-word recognition. Here we show how spatiotemporal searchlight RSA applied to source estimations of EMEG data can provide insights into the neural correlates of the cohort model within bilateral frontotemporal brain regions.**

Keywords: MVPA, RSA, MEG, EEG, MNE, cohort model

I. Introduction

Multivariate Pattern Analysis (MVPA) has been successfully applied to fMRI (e.g. [1][2][3]) and to (quasi) time series data [4]. Unlike mass-univariate approaches such as SPM [5], MVPA is based on the pattern information that is naturally embedded in multi-channel recording of neural activations. Application of MVPA to time resolved imaging techniques (e.g. MEG and EEG) is, however, still very limited.

A particular type of MVPA is Representational Similarity Analysis (RSA) [6], which has been successfully applied to neuroimaging data from several different modalities. For instance, RSA has been used to relate the blood-oxygen-level dependent (BOLD) responses from human inferior temporal cortex (IT) to single-cell recordings (averaged spike rate over time) from monkey IT [6]. Here we extend RSA to apply to whole-brain source estimations of combined MEG and EEG (EMEG) data. This novel technique allows us to identify spatiotemporal neural signatures on the basis of the similarity structures they represent. To do this, we have generalized the spatial searchlight used in fMRI analysis [2], by adding a sliding time window to accommodate the additional temporal dimension. This spatiotemporal searchlight RSA is capable of generating systematic accounts of complex parallel distributed neuro-computational processes, operating on millisecond timescales. A prime example of such complex processes is characterized in the cohort model of spoken-word recognition [7][8][9]. This model describes how auditory speech input is mapped onto stored lexical representations in real-time. Specifically, when a speech segment enters the system, it activates a virtual cohort of word units, which code lexical representations that begin with this speech segment. These activated units compete to provide the best match to the accumulating input. As more inputs enter the system, the intended word should emerge as the best fitting candidate.

It is apparent that such complex dynamic processes involve large distributed neural networks working in parallel as the words are heard. A major challenge for cognitive neuroscience is to develop methods for characterizing these dynamic spatiotemporal patterns in EMEG data recorded while participants are listening to spoken language. In this paper, we will describe our methods for spatiotemporal searchlight RSA and show how they can be applied to capture cohort-based analysis processes as a spoken word is being heard.

II. Spatiotemporal Searchlight RSA

A. The Representational Dissimilarity Matrix (RDM)

The first step in the RSA process is the computation of similarity structures that express the dynamic patterns of neural activation over space and time. The primary data type that encodes this similarity structure is the representational dissimilarity matrix (RDM). Each entry in the RDM is a correlation distance (e.g. one minus the correlation value) between activation patterns elicited by a pair of experimental conditions (or stimuli) within a specific experimental condition. As shown in Figure 1, elements on the main diagonal of this matrix are zeros by definition. In the off-diagonal parts of the RDM, a large value indicates that the two conditions have elicited distinct spatiotemporal activation patterns, and vice versa for small values. RDMs computed using this method are symmetric about the main diagonal. Subsequent computations are therefore only performed on the lower-triangle portion of the matrix. RDMs can either be *brain-data* RDMs reflecting recorded brain activity, or *model* RDMs (see below) expressing specific theoretical hypotheses about the properties of this activity.

B. Moving Spatiotemporal Searchlight

Signals in EMEG source space have a complex spatial and temporal distribution. In addition, the signal-to-noise ratio (SNR) is low in a single EMEG time series. To increase SNR, the conventional Event Related Potential (ERP) approach reduces noise that is not time-locked to the stimuli by averaging across many trials and recordings from multiple channels are often averaged. In contrast, RSA and other MVPAs can potentially increase the level of detectable signal by integrating information from multiple channels using more sophisticated multivariate statistics, which extract additional

information from the patterns than simply the mean amplitude of all channels. This advantage of MVPA over its univariate counter-parts has been demonstrated in the field of fMRI. We argue that this is also true for EMEG, and have extended the concept of spatial searchlight developed for fMRI to include an extra temporal dimension. This is done by combining a spatial searchlight with a sliding time window under the assumption that neuro-cognitive representations are realized in the continuous spatiotemporal patterns of EMEG data.
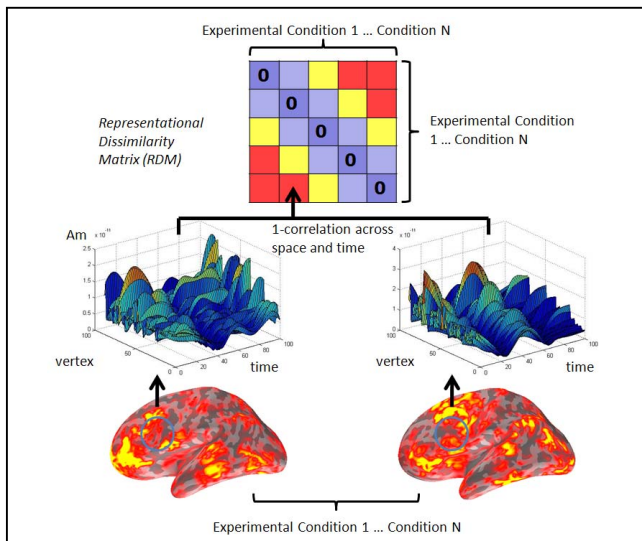

Figure 1.   Computing brain-data RDMs from spatiotemporal patterns

Specifically, at each spatial location (or *vertex*) on the source estimated cortical surface (see Section III), the searchlight covers a hexagonal cortical patch (radius = 20 mm), which includes 128 vertices. The searchlight also extends in time for 20 ms – a time period chosen on the basis of estimates of the timing with which neural circuits are updated [10]. The time step of the sliding window is set to 5 ms. The RDM computed for each spatiotemporal searchlight location is assigned to the center vertex of the hexagon. We allow the searchlight to overlap in space and in time, on a whole-brain basis across the length of the recording epoch for each word, resulting in a brain-data RDM at each vertex and each time point. Sampling with overlapping spatiotemporal searchlights enables us to detect distributed and transient representations that might otherwise straddle the boundary between adjacent cortical patches or successive temporal bins and fail to be analysed as a single pattern.

C. Comparison of Brain-data RDMs to Theoretical Models

At the second step, theoretical models are also represented by RDMs, which define hypotheses about the data. Such hypotheses can either be described informally, for example, images of objects can be categorized as either living or nonliving, or they can be formally derived from a computational model. A computational model can be a set of closed-form equations that compute certain properties of the stimuli, such as the power in a particular frequency band within a sound sequence. However, more sophisticated cognitive models can also be used, e.g. the distributed representation in a

connectionist network can form the basis of model RDMs. Here, based on the cohort model, we computed a function for each word in the experiment (see Section III) describing the dynamic change of cohort size over time. We then derived a model RDM from these functions by computing pair-wise cross correlation distances between them (see Figure 2). Finally, the model RDM is compared to the brain-data RDMs by computing a Spearman correlation as the searchlight moves in space and in time. The output of the searchlight mapping indicates when and where in the brain the model RDM fits to the observed patterns of neural activity. The Spearman correlation value (ρ) is assigned to the center vertex at each searchlight testing point. This results in a spatiotemporal ρ-map for each participant.
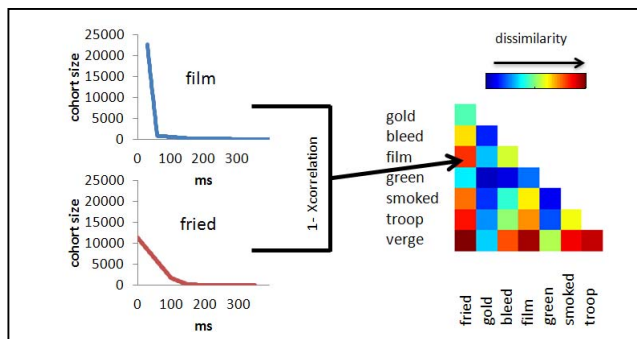

Figure 2.   Computing a theoretical model RDM based on the cohort model (*film* and *fried* are examples of stimuli used in the experiment)

D. Group Statistics

At the group level, one can choose between fixed-effects and random-effects tests. Here we use a random-effects test, in the form of a one-sample t-test performed across subjects over their ρ-maps, resulting in a group t-map that summarizes the effect across the group. Both the ρ-map and the group t-map are descriptive statistical maps, from which we can draw inferences. However, in the source estimations of EMEG data, there can be tens of thousands of vertices (here we down-sampled to 10,242 vertices per hemisphere) multiplied by hundreds of time points. The descriptive statistical maps will therefore contain very large numbers of individual comparisons. This creates a massive multiple-comparisons problem, potentially resulting in high proportions of false positives. In the next section, we propose a nonparametric correction method based on cluster-level randomization testing, from which we can derive a p-value, while controlling the false-positives rate across the whole brain.

E. Correction for Multiple Comparisons

Most standard statistical tests, such as t- or F-tests, require assumptions to be met about the underlying distribution. However, these assumptions do not always hold in neuroimaging, in particular when certain nonstandard statistics are used. Such nonstandard statistics include classification accuracy derived from MVPA or different distance measures used in RSA and other machine-learning approaches. A versatile approach that applies to such statistics is randomisation testing [11][12][13]. Here, the multiple-comparisons problem in RSA for EMEG is addressed by randomisation testing of cluster-level statistics. This controls
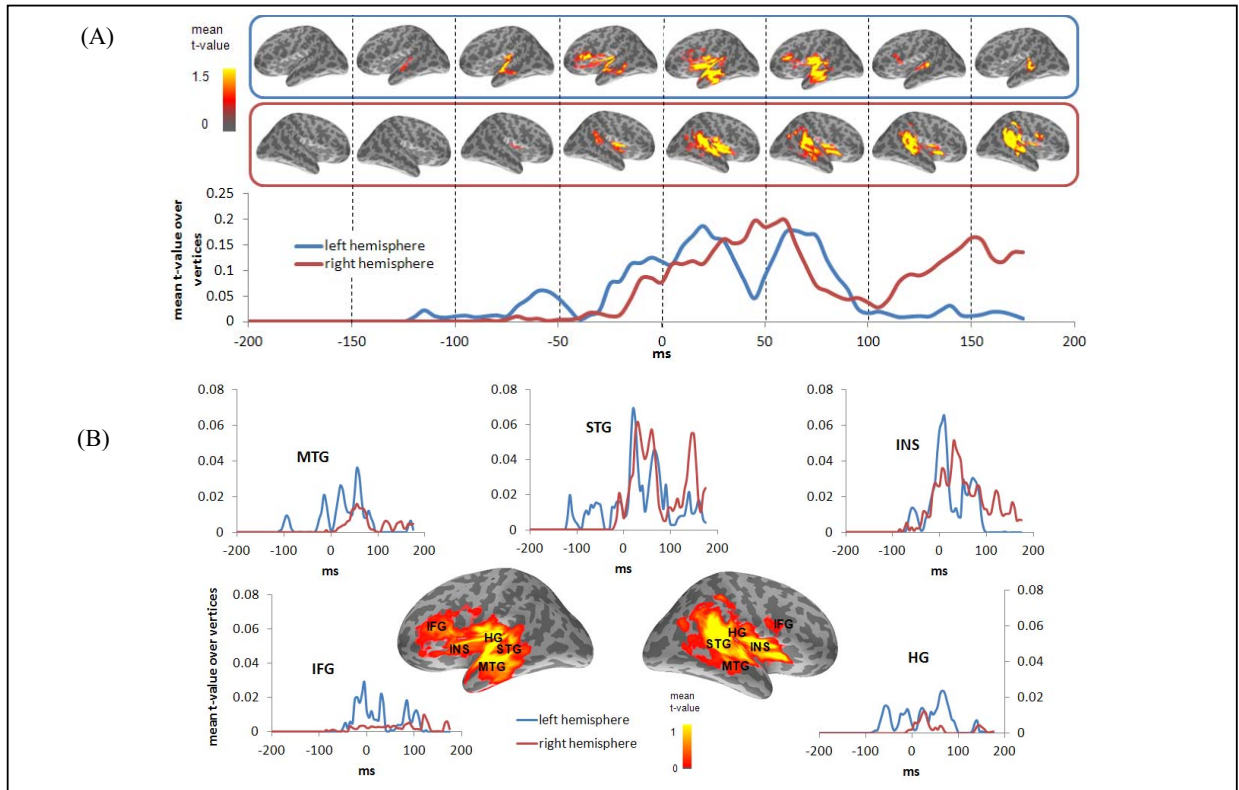
Figure 3. (A) Significant clusters showing brain areas that carry information about cohort competition over space and time. (B) Effect time courses for individual brain areas. MTG- middle temporal gyrus, STG – superior temporal gyrus, INS – insula, HG – Heschl's Gyrus, IFG – inferior frontal gyrus.

for false positives arising from multiple comparisons without making any assumptions about the distribution of the data. The procedure is as follows.

From the previous step, we obtain a map of summary statistics, such as a ρ- or t-map from the group. A primary cluster-forming threshold is applied to these maps. The primary threshold controls the relative sensitivity of the analysis to low, but extended versus high, but focal effects. In this paper, the primary threshold is chosen so as to select the top 5% of all ρ- or t-values across the brain and across all time points. After applying the primary threshold, clusters are identified by spatiotemporal contiguity. We allow the cluster to travel in space between different time points. From each cluster, a cluster-level statistic is calculated as the exceedance mass [14] across space and time. We call these clusters *observed* clusters.

A key step for nonparametric testing is the simulation of the null distribution. In order to control the family wise error rate for our spatiotemporal maps, we need to simulate the distribution of map-maximum cluster statistics under the null hypothesis that there is no difference between any conditions in the experiments. Thresholding the null distribution of map-maximum cluster statistics to select the top 5%, will ensure that we will have a risk of merely p = 0.05 of detecting any cluster as significant if the null hypothesis is true. We simulate the null hypothesis by randomisation of the condition labels, based on their exchangeability under the null hypothesis [13]. In the fixed-effects test, condition labels in the brain-based RDMs are randomly permuted 1,000 times, and the group ρ-map is recomputed each time. In the random-effects test, the sign of ρ-

maps from each subject are randomly flipped 1,000 times and a group t-map is recomputed for each random flipping. This is based on the assumption that the ρ values are symmetrically distributed about zero under the null hypothesis.

For each map generated by simulation, we applied the same primary threshold used to form observed clusters to identify *simulated* spatiotemporal clusters. We compute the cluster-level statistics for all simulated clusters, and then build the null distribution, by selecting the cluster from each run of the simulation with the largest cluster statistics. A corrected p-value is given to an observed cluster by looking at where its cluster statistics locate in the distribution of simulated cluster statistics. This procedure controls the false-positives rate for the whole brain and whole time period.

III. EXPERIMENTAL DETAILS

17 healthy, right-handed native English speakers have participated in the experiment. Participants listened passively to English words (e.g. *fried*, *film*, *dream*) and occasionally performed a 1-back memory task. Combined MEG and EEG data was collected at the MRC Cognition and Brain Sciences Unit using a 306-channel Vectorview MEG and 70-channel EEG system. After removing artifacts, trials were aligned to the onset of the last phoneme of the word, which was set as time zero. The time window of interest was defined as the period from 200 ms prior to the alignment point to 200 ms after it. Baseline correction was based on a 100 ms window preceding the time window of interest. MRI images were obtained on a 3T Siemens Trio with 1 mm isotropic voxels. From the MRI

data, a representation of the cerebral cortex was constructed using FreeSurfer. The forward model was calculated with a three-layer Boundary Element Model (BEM) using the outer surface of the scalp as well as the outer and inner surfaces of the skull identified in the anatomical MRI. Anatomically constrained source estimations were created with Minimum-Norm Estimation (MNE) [15] combining MRI, MEG, and EEG data.

IV. RESULTS

Figure 3 (A) shows the extent to which the brain-data RDMs from the brain fit the model RDM. There are significant clusters in both hemispheres after correcting for multiple comparisons (random effect, $p < 0.001$). The results of spatiotemporal searchlight RSA are effectively information-rich movies. To give a summary overview of the effects, the top panels in Figure 3 (A) show the spatial distribution of significant clusters averaging across each 50 ms period. The bottom panel in Figure 3 (A) shows the averaged t-values over the significant clusters by hemisphere, which provides a global view of the temporal evolution of the effects. Even at this coarse level, we can see that the left hemisphere responds to cohort competition as early as 125 ms prior to the alignment point, and that the effects on the right occurred around 100 ms later than on the left.

Figure 3 (B) unpacks the global view of the results into several key brain areas (anatomically defined using FreeSurfer) that are critical to lexical processing. In general, effects were first shown in bilateral temporal regions (MTG, STG), and then in left inferior frontal regions (IFG). The early effects in MTG and STG indicate a rapid access to lexical information and arguably reflect the activation of neural systems engaged by the presence of cohort competition. The later effects in left IFG, beginning at around -40 ms, may reflect neuro-cognitive decision processes associated with the selection of the intended words from among multiple competing lexical candidates.

V. CONCLUSIONS

In this paper, we have introduced a novel approach to the analysis of spatiotemporal patterns in EMEG source space based on searchlight RSA. This method can capture fine-grained dynamic neural computations in the brain, and at the same time can do so on a large scale – encompassing the whole brain. We described nonparametric procedures that address the spatiotemporal multiple-comparisons problem. This represents a significant advance over current applications [16], which largely use techniques first pioneered some 40 years ago in the studies using MVPA for EEG analysis [17]. In addition, whole brain source analysis, which is arguably the neuroscientifically preferable approach, is not used routinely (though see [18]).

The ability to directly analyze pattern information in both space and time from neural activity in the brain enables us to generate and validate computational models and cognitive theories in a natural and informative way. The RDM encapsulates the similarity structure of the patterns of variation in the EMEG signals, and provides an abstract means of characterising neural activity. As shown in our study of aspects of the cohort model (a computational account of spoken word recognition), RSA is able to closely relate dynamic patterns at the neuronal level, measured electrophysiologically, with patterns derived from higher-level cognitive theories. In particular, we have identified a bilateral neural network and its temporal evolution in frontotemporal regions that supports the processing of cohort-based spoken word recognition.

REFERENCES

[1] Haxby J.V. et al. (2001) Distributed and overlapping representations of faces and objects in ventral temporal cortex, Science, 293:2425-2430.

[2] Kriegeskorte N. et al. (2006) Information-based functional brain mapping, PNAS, 103:3863-3868.

[3] Haynes J.D. and Rees G. (2006) Decoding mental states from brain activity in humans, Nature Reviews Neuroscience, 7:523-34.

[4] Mourão-Miranda J. et al. (2007) Dynamic discrimination analysis: A spatial-temporal SVM, NeuroImage, 36:88-99.

[5] Friston K.J. et al. (1995) Statistical parametric maps in functional imaging: A general linear approach, Human Brain Mapping, 2:189-210.

[6] Kriegeskorte N. et al. (2008) Matching categorical object representations in inferior temporal cortex of man and monkey, Neuron, 60:1-16.

[7] Marslen-Wilson W.D. (1984) Function and processing in spoken word recognition: a tutorial review. In H. Bouma & D.G. Bouwhuis (Eds.), Attention and Performance X: Control of Language Processing.

[8] Gaskell M.G. and Marslen-Wilson W.D. (1997) Integrating form and meaning: a distributed model of speech perception. Language and Cognitive Processes, 12:613-656.

[9] Marslen-Wilson W.D. and Tyler L.K. (2007) Morphology, language and the brain: the decompositional substrate for language comprehension, Philosophical Transactions of the Royal Society B: Biological Sciences, 362:823-836.

[10] Panzeri S. et al. (2001) Speed of feedforward and recurrent processing in multilayer networks of integrate-and-fire neurons. Network: Computation in Neural Systems, 12:423-440.

[11] Bullmore E. et al. (1996) Statistical methods of estimation and inference for functional MR image analysis. Magnetic Resonance in Medicine, 35:261-277.

[12] Brammer M.J. et al. (1997) Generic brain activation mapping in fMRI: a nonparametric approach. Magnetic Resonance Imaging, 15:763-770.

[13] Nichols T.E. and Holmes A.P. (2001) Nonparametric Permutation Tests for Functional Neuroimaging: A Primer with Examples. Human Brain Mapping, 15:1-25.

[14] Bullmore E.T. et al. (1999) Global, voxel, and cluster tests, by theory and permutation, for a difference between two groups of structural MR images of the brain. IEEE Transactions on Medical Imaging, 18:32–42.

[15] Hämäläinen M.S. and Ilmoniemi R.J. (1994) Interpreting magnetic fields of the brain: minimum norm estimates, Medical & Biological Engineering & Computing, 32:35-42.

[16] Chan A.M. et al. (2011) Decoding word and category-specific spatiotemporal representations from MEG and EEG. NeuroImage, 54(4):3028-39.

[17] Izawa S. and Uchiyama T. (1973) Application of Multivariate Analysis to Quantitative Classification of EEG Patterns of the Mentally Retarded. RIEEC Research Bulletin 2.

[18] Besserve M. et al. (2011) Improving quantification of functional networks with EEG inverse problem: Evidence from a decoding point of view, NeuroImage, 55(4):1536-1547.