Reading out the location being stored in spatial working memory with fMRI

G.J. Detre*; V.S. Natu; K. Schneider; K. DeSimone; S. Kastner; K.A. Norman

Department of Psychology and Center for the Study of Brain, Mind, and Behavior, Princeton University

642.11

Aims

To train a classifier to read out which of 12 locations is being maintained by subjects in their spatial working memory from an individual volume acquired over 2 seconds.

More precisely, to be able to read the (x,y) coordinate of the location being maintained.

To consider new ways of visualizing and quantifying how similar the neural representations of different conditions are to each other (their 'similarity structure')

To develop a new way of selecting voxels that takes their similarity structure into account

Task

1) Stare at fixation point in the center of the screen

2) Fixation point disappears, and target appears

3) Target disappears, and distractors appear - maintain location of target in spatial working memory

4) Distractors and target disappear – saccade to the target position and back

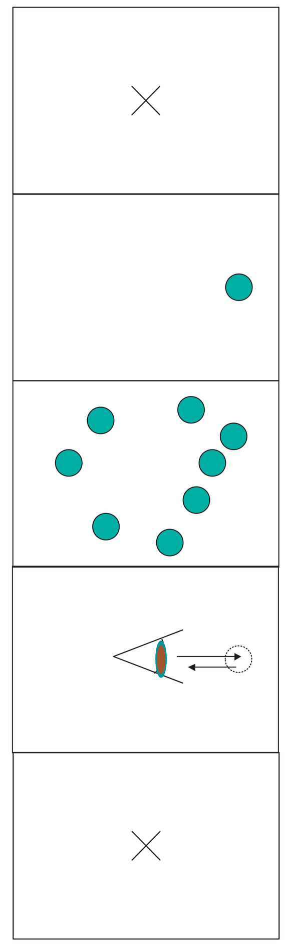
5) Repeat with stimuli moving around the clock (12 positions in total).

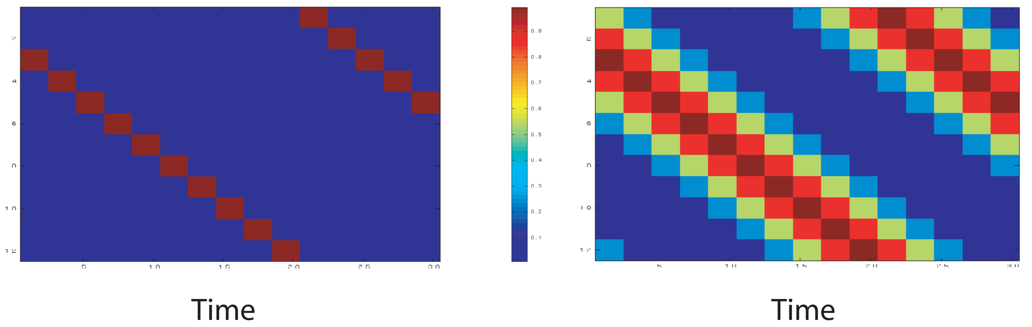
Conditions progress sequentially counter-clockwise around the clock. 5 seconds per position, 12 positions per cycle, 8 cycles per run, 6 runs.

Some spatial jitter added to target locations.

Siemens 3T Allegra functional scanning parameters: 128x128x20, 2x2x2 mm slice, 1 mm gap, axial slices, TR = 2000ms, TE= 40, Flip Angle = 90 degrees.

Data collected by Desimone et al. (2004).





Left: Schematic of the boxcar version of the regressors (rows = conditions, cols = timepoints). For analysis, we then convolved these with AFNI's gamma variate waveform haemodynamic response function.

Right: Schematic of the Gaussian bump regressors (before being convolved), reflecting assumption of Gaussian tuning curve coding. The peaks of the bumps match the boxcar regressors.

Summary of analysis procedure

AFNI pre-processing - motion correction, MGH undistortion, despiking, detrending (subtracting mean, linear and quadratic trends), 4mm smoothing

Import into matlab (using the AFNI Matlab library)

Create boxcar and Gaussian bump regressors, convolve them with HRF

N-minus-one no-peeking voxel selection (either Fourier analysis, multiple regression, or our new similarity structure method)

N-minus-one cross-validation classification, leaving out a single run each time (backpropagation neural network). See Polyn et al. (2004), Mitchell et al. (2004).

Transform classifier output into (x,y) coordinates. Responses are considered correct if closer to the target location than any other.

Iterations \ Runs	1	2	3	4	5	6
1	test	train	train	train	train	train
2	train	test	train	train	train	train
3	train	train	test	train	train	train
4	train	train	train	test	train	train
5	train	train	train	train	test	train
6	train	train	train	train	train	test

Similarity structure voxel selection method

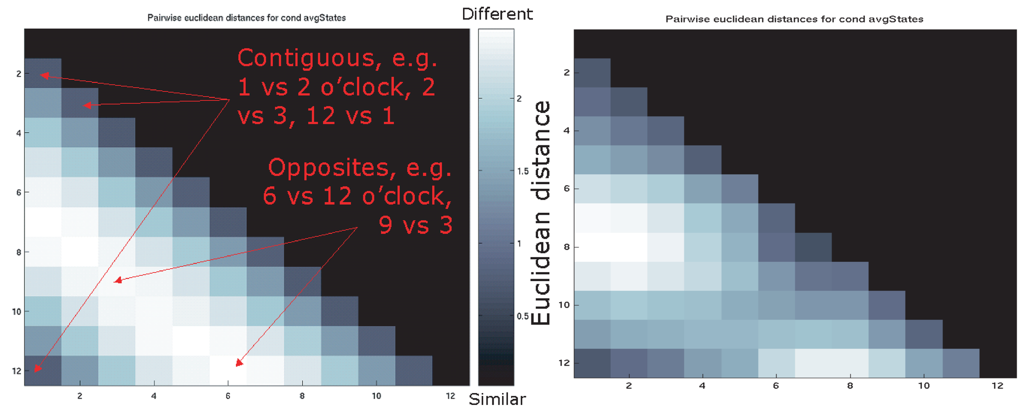
- Expect similar states of the world to elicit similar neural representations. To test this, we need a method that doesn't just treat the conditions as categorical (Haxby et al., 2001; Datta & DeYoe, 2004).

- Evidence for retinotopy of spatial working memory maps around the intraparietal sulcus, the frontal eye fields and the dorsolateral prefrontal cortex (Desimone et al, 2004; Sereno et al, 2001). We want to link this with an a priori model of how similar each condition is to every other condition, e.g. 12 o'clock is closer to 1 o'clock than to 6 o'clock. Predict that similar conditions (as measured by the physical distance of the targets) will elicit similar brain representations (in terms of the response patterns for each target location).

- Quantify the similarity structure of the conditions using the pairwise Euclidean distances between the (x,y) coordinates of each pair of target locations.

- Find the mean value for each condition, for each voxel. Calculate the pairwise distances between these twelve values, yielding a similarity structure triangle matrix for that voxel.

- Compare the similarity structure for each voxel to the similarity structure of the conditions using correlation.

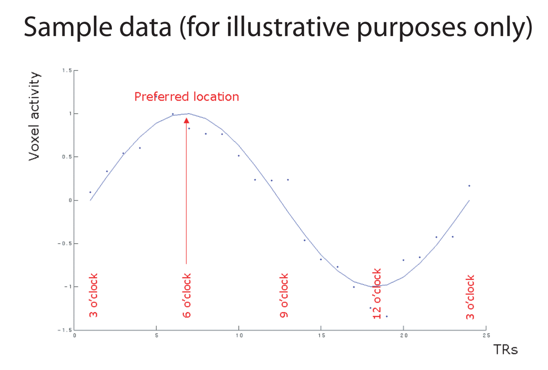


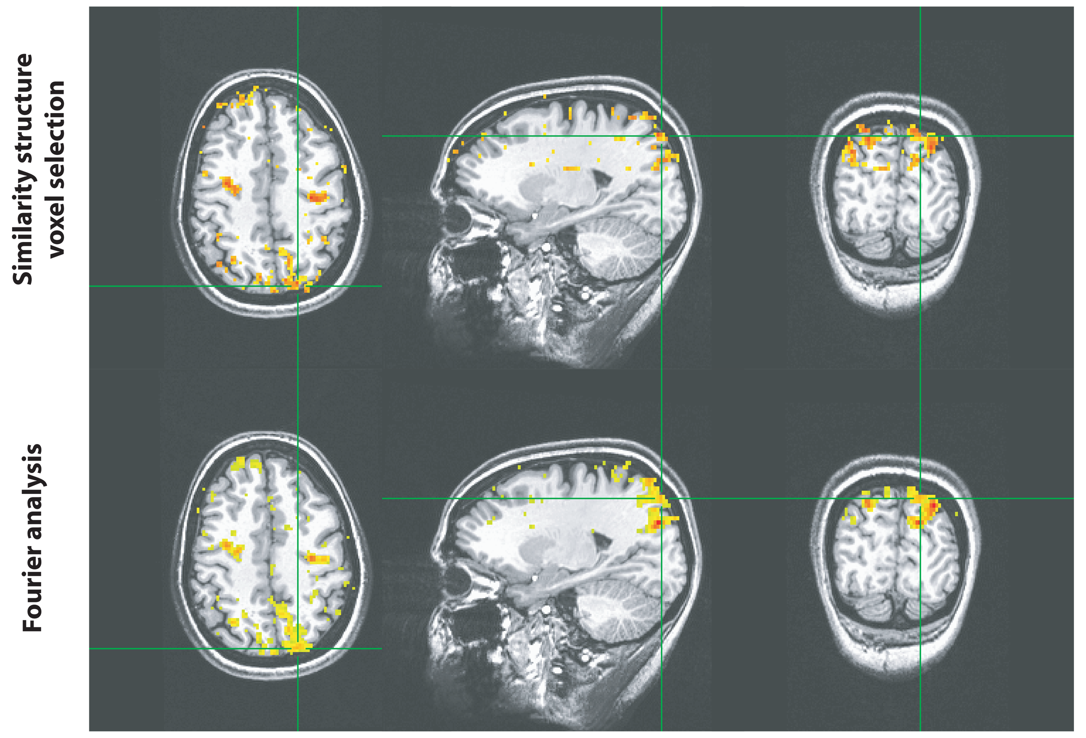
Left: Similarity structure matrix for conditions (i.e. our model of what the distances should look like)

Right: Similarity structure for the best 200 voxels

Benchmarking the similarity structure voxel selection method against a Fourier analysis

Voxels with Gaussian-tuning curves should peak in activity during presentation of their preferred location, with diminished activity before and after for neighbouring locations, resulting in a roughly sinusoidal activity timecourse. Sereno et al. (2001) described how to use a Fourier analysis on each voxel to measure its coherence at a frequency of 1/cycle length, and showed how this yields retinotopic brain areas, colored by the phase of the best-fitting sine wave and thresholded by the correlation.



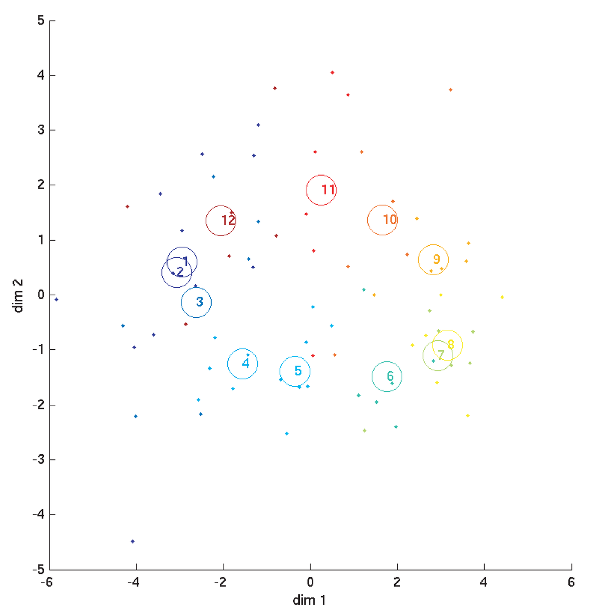


Above: comparison with Fourier analysis. Comparisons with multiple regression (using different data) also look promising.

Visualizing the similarity structure at a high level

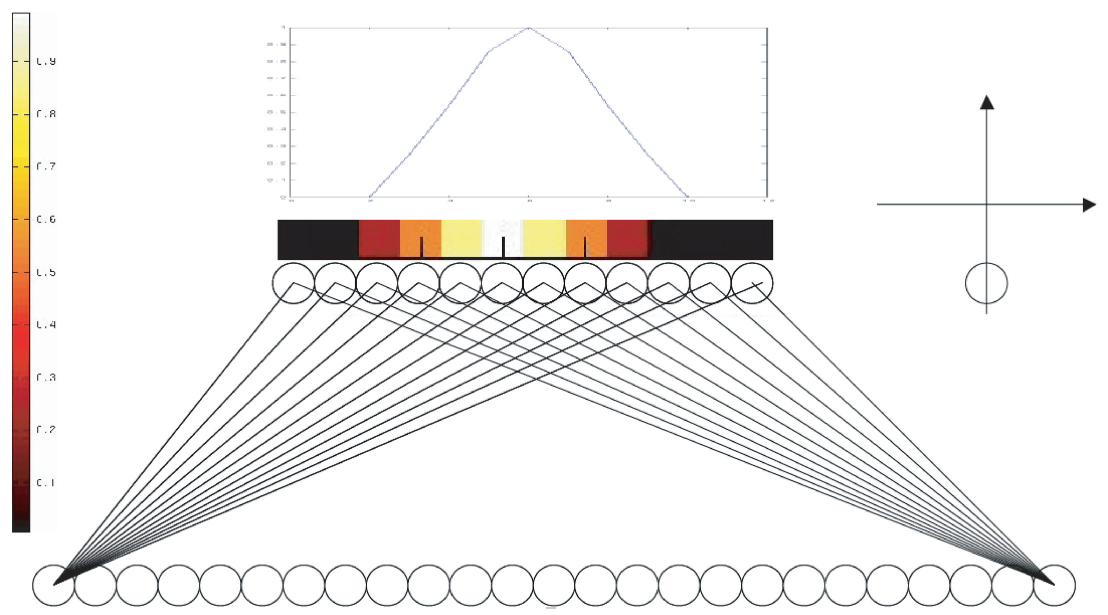
Multi-dimensional scaling tries to find a low-dimensional space which preserves distance relationships. An MDS run on the pairwise distances between conditions would yield a perfectly circular ring, with the conditions proceeding sequentially, in order to maximize the proximity of similar conditions.

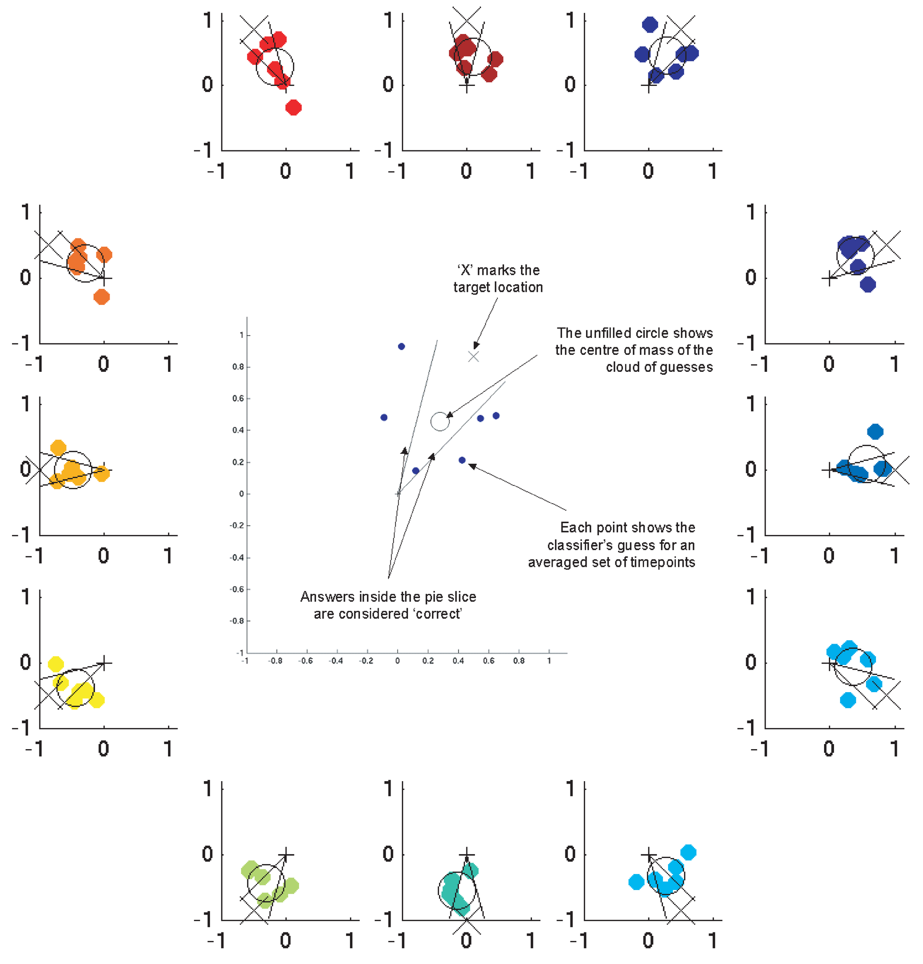
Right: the results of an MDS run on the similarity structure of the 500 best voxels. Dots = individual timepoints, circles = centre of mass for a given condition.



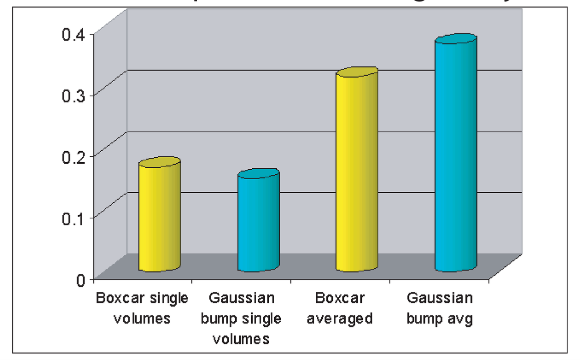
Classification

Schematic of simple 'Gaussian bump' backpropagation network (no hidden layer). Note: actual analyses were conducted with a 20-unit hidden layer.





Classification performance (single subject)



Extensions

A secondary aim of developing the similarity structure voxel selection method was to find a way of selecting voxels in a multivariate way, to match the multivariate classification analysis. The similarity structure method has so far only been applied to voxels one by one, but could equally be applied to groups of voxels at a time. This becomes a much harder search problem, because of the combinatorics of trying permutations of sets of voxels, but we plan to start by moving a spherical spotlight around the brain to choose clusters of voxels at a time as a first approach.

Multi-Voxel Pattern Analysis (MVPA) toolbox

All of the analyses described were implemented using the Princeton Multi-Voxel Pattern Analysis (MVPA) toolbox in Matlab (Polyn et al., 2005), which is available online at **http://www.csbmb.princeton.edu**. This facilitates import/export of data, simple pre-processing, and a variety of voxel selection and classification algorithms within an n-minus-one no-peeking framework.

Conclusions

We can classify which of the 12 clock-locations is the target better than chance for single volumes, and much better than chance when averaging multiple volumes from the same condition together. Classifier architectures designed to take the similarity between conditions into account (Gaussian bump) perform better than simple 1-of-n classifiers that treat the conditions as categorical, for averaged data.

The new voxel selection method that takes into account the similarity structure of the conditions performs comparably to more standard Fourier analysis and could be extended in the future to work multivariately on sets of voxels at a time.

References

Datta R, DeYoe EA (2004). I know where you are secretly attending! Reading the focus of attention from the brain with fMRI. *Soc Neurosci Abstr.*

Desimone K, Schneider K, Pinsk M, Norman K & Kastner S (2004). Response properties of a putative LIP area in human parietal cortex. *Soc Neurosci Abstr.*

Haxby JV, Gobbini MI, Furey ML, Ishai A, Schouten J L, & Pietrini P (2001). Distributed and overlapping representations of faces and objects in ventral temporal cortex. *Science*, 293, 2425–2429.

Mitchell TM, Hutchinson R, Niculescu RS, Pereira F, Wang X, Just M, Newman S (2004). Learning to decode cognitive states from brain images. *Machine learning* 57:145-175.

Polyn SM, Cohen JD & Norman KA (2004). Detecting distributed patterns in an fMRI study of free recall. *Soc Neurosci Abstr.*

Polyn SM, Detre GJ, Takerkart S, Natu VS, Benharrosh MS, Singer BD, Cohen JD, Haxby JV, Norman KA (2005). A Matlab-based toolbox to facilitate multi-voxel pattern classification of fMRI data. *Meeting of the Organization for Human Brain Mapping.*

Sereno MI, Pitzalis S, Martinez A (2001). Mapping of contralateral space in retinotopic coordinates by a parietal cortical area in humans. *Science*, 294, 1350-1354.

For a downloadable version, please visit: http://compmem.princeton.edu/publications.html