NeuroImage 128 (2016) 385–397

Contents lists available at ScienceDirect

NeuroImage

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

Representational dynamics of object recognition: Feedforward and feedback information flows

Erin Goddard [a,b,*], Thomas A. Carlson [a,b], Nadene Dermody [a], Alexandra Woolgar [a,b]

[a] Perception in Action Research Centre (PARC) and Department of Cognitive Science, Macquarie University, Sydney, NSW 2109, Australia
[b] ARC Centre of Excellence in Cognition and its Disorders (CCD), Macquarie University, Sydney, NSW 2109, Australia

ARTICLE INFO

Article history:
Received 18 June 2015
Accepted 5 January 2016
Available online 13 January 2016

Keywords:
Magnetoencephalography
Pattern classifier analysis
Feedback
Coarse-to-fine
Granger causality
Visual perception

ABSTRACT

Object perception involves a range of visual and cognitive processes, and is known to include both a feedfoward flow of information from early visual cortical areas to higher cortical areas, along with feedback from areas such as prefrontal cortex. Previous studies have found that low and high spatial frequency information regarding object identity may be processed over different timescales. Here we used the high temporal resolution of magnetoencephalography (MEG) combined with multivariate pattern analysis to measure information specifically related to object identity in peri-frontal and peri-occipital areas. Using stimuli closely matched in their low-level visual content, we found that activity in peri-occipital cortex could be used to decode object identity from ~80 ms post stimulus onset, and activity in peri-frontal cortex could also be used to decode object identity from a later time (~265 ms post stimulus onset). Low spatial frequency information related to object identity was present in the MEG signal at an earlier time than high spatial frequency information for peri-occipital cortex, but not for peri-frontal cortex. We additionally used Granger causality analysis to compare feedforward and feedback influences on representational content, and found evidence of both an early feedfoward flow and later feedback flow of information related to object identity. We discuss our findings in relation to existing theories of object processing and propose how the methods we use here could be used to address further questions of the neural substrates underlying object perception.

Visual object recognition and identification is an important task in everyday life, and the speed and accuracy with which we can identify objects are consistent with the visual system devoting considerable resources to this ecologically relevant process. Object perception involves a range of visual and cognitive processes, including a feedfoward flow of information along the 'ventral stream' of visual cortex (for example, see Tanaka, 1996; Grill-Spector et al., 2001), and also feedback from frontal and parietal areas such as prefrontal cortex. However, the way in which these areas interact to contribute to object perception remains unclear despite a growing experimental literature on the topic.

A range of theories of object perception (Bullier, 2001; Bar, 2003; Peyrin et al., 2010; Tapia and Breitmeyer, 2011; Horr et al., 2014; Hochstein and Ahissar, 2002) have hypothesized that very early top-down feedback, around 100–150 ms after stimulus onset (Bar et al., 2006), carries content regarding object identity from prefrontal cortex to the traditional 'bottom-up' dorsal and ventral visual pathways. These theories are based on results such as the speed with which humans can correctly respond to simple object categorisation tasks (for example, 120 ms after stimulus onset, Kirchner and Thorpe, 2006), combined with reports that early activity (130–150 ms after stimulus onset) from prefrontal sites can vary with object recognition (Thorpe et al., 1996; Bar et al., 2006).

The relative timing of feedforward and feedback flows of information in object perception has been suggested to depend on the spatial frequency content of the image, with the earliest information about object information coming from low spatial frequency (low-pass) image components (Bar et al., 2006; Chaumon et al., 2014; Fintzi and Mahon, 2014). This is broadly consistent with psychophysical results implying that low spatial frequency image components are processed prior to high spatial frequency components (Hughes et al., 1996; Parker et al., 1992, 1997; Schyns and Oliva, 1994; Neri, 2011). However, these effects are likely contingent on the relative usefulness of low and high spatial frequencies to the participant's task (De Gardelle and Kouider, 2010; Stein et al., 2014; Patai et al., 2013), which challenges the notion of clear segregation between rapidly propagated low-pass signals and slower 'High-pass' signals. Furthermore, neuroimaging studies providing evidence for differential processing of high and low spatial frequency stimuli in object processing (Bar et al., 2006; Chaumon et al., 2014; Fintzi and Mahon, 2014) have used stimuli varying in total spatial frequency content, and have not equated the stimulus types for low-level

* Corresponding author at: Department of Cognitive Science, Macquarie University, Sydney, NSW 2109, Australia.
E-mail address: erin.goddard@mq.edu.au (E. Goddard).

http://dx.doi.org/10.1016/j.neuroimage.2016.01.006
1053-8119/

properties such as overall luminance and contrast. This leaves open the possibility that their differential effects may be due to different image statistics across conditions.

Here we devised a new methodology to explore the timing of feedforward and feedback flows of object-related information based on magnetoencephalography (MEG) recordings. We applied a time resolved multivariate pattern classification analysis to magnetoencephalography (MEG) data (Carlson et al., 2011, 2013; Isik et al., 2014) and compared the object-related information in peri-occipital and peri-frontal areas at different time points. Using a novel extension of Granger causality analysis, we tested for evidence that the representational structure of object-related information in frontal regions predicted the representational structure of later responses in occipital areas.

Using this method, we compared the brain's processing of low and high spatial frequency object-related information. We used stimuli that had the same power at each spatial frequency in their Fourier amplitude spectra, and the same overall contrast, varying only the 'diagnostic' spatial frequencies (De Gardelle and Kouider, 2010). Stimuli in our 'Low-pass' condition had object signal in the low spatial frequencies while high spatial frequencies were phase-randomized, while for the 'High-pass' condition this was reversed.

Methods

Participants

Twelve participants (nine female, ten naïve to the purposes of the study) took part in an initial psychophysical experiment used to calibrate the visual images for their low and high spatial frequency content. Nine participants (five female, eight naïve to the purposes of the study) completed the second psychophysical experiment and the MEG experiment. All had normal or corrected to normal vision, and naïve participants were paid for their time. All participant recruitment and experiments were conducted with the approval of the Macquarie University Human Research Ethics Committee.

Visual stimuli

Visual stimuli were generated and presented using Matlab (version R2013a) and routines from Psychtoolbox (Brainard, 1997; Pelli, 1997). In all experiments we used the same set of 24 images, which were selected from a set of 92 supplied by Nikolaus Kriegeskorte (described in Kriegeskorte et al., 2008). All images were segmented real world objects on a gray background. In both psychophysical and imaging experiments, participants judged whether each presented object was smaller or larger than a shoebox. We chose the 24 images such that for both the 'smaller than a shoebox' and 'larger than a shoebox' groups there were six animate and six inanimate objects.

Each image was 175 × 175 pixels, and subtended 15 degrees visual angle (dva) in each experiment. We converted each original color image to grayscale by setting the RGB coordinate of each pixel to the average of the R, G, and B coordinates of that pixel in the original image. In order to equate all images for their power at each orientation and spatial frequency, we set the amplitude matrix of each image to the average amplitude matrix of all images. To find the average amplitude matrix across the 24 images, we performed a two-dimensional discrete Fourier transform of each image, which yielded an amplitude and phase matrix for each image, and then for each point in amplitude matrix we used the average amplitude across the 24 images. We used the same amplitude matrix for every stimulus image, varying only the phase matrix that was used in the two-dimensional inverse discrete Fourier transform to generate a given stimulus.

The phase matrix of each stimulus was derived from one of the 24 images, with varying amounts of phase randomization introduced to the phase matrix. The four stimulus conditions, along with the pattern of phase randomization in each case, are illustrated in Fig. 1. Phase randomization was introduced to one or more of three spatial frequency bands: low (< 0.90 cycles/dva), medium (≥ 0.90 cycles/dva and ≤ 1.03 cycles/dva) and high (> 1.03 cycles/dva). In the 'Low-pass' condition, the phase of spatial frequencies in the high and medium bands was randomized, and in the 'High-pass' condition the low and medium bands were randomized. In the 'Strong signal' and 'Weak signal' stimulus conditions, the phase of all spatial frequencies in the medium band was randomized, along with varying proportions of the frequencies in the low and high bands, as detailed below.

Since every image contained at least some phase randomization, we were able to repeat the randomization process and generate different versions of the same image with the same object signal, such that a new image was used for every trial. Also, since all images had the same amplitude matrix, the objects could not be distinguished when phase randomization was complete across all spatial frequencies. This ensured that objects could not be identified based on the orientation/spatial frequency profile of the randomized images.

Psychophysical experiments

We conducted two psychophysical experiments in order to measure the detectability of the objects in the different conditions. Stimuli were generated and displayed on a Dell OptiPlex 9010 desktop computer driving an AMD Radeon HD 7570 graphics card to draw stimuli to a 60 × 33 cm Samsung SyncMaster SA950 Full HD 3D LED monitor, refreshed at 120 Hz. Experiments took place in a darkened room and the monitor was viewed from a distance of .64 m.

In the first experiment, we included only the 'Low-pass' and 'High-pass' conditions, and measured the detectability of each object in these two conditions as a function of the amount of phase randomization. At the start of each session the participant chose the keys on a keyboard they would use for their responses ('smaller' and 'larger' than a shoebox) and after these responses the experiment commenced. Each trial began with a central fixation marker (a small gray cross) that was displayed on a black background for 250 ms, after which the stimulus image was displayed on a black background for 500 ms before being replaced by the fixation marker. Participants were given an unlimited amount of time to respond. Following the participant's response, they would receive feedback on their decision (displayed as 'correct' in green, or 'incorrect' in red) for 500 ms, and then the next trial would commence.

Each of the 12 participants completed 8 sessions of 15–20 minutes each, consisting either of only 'High-pass' or only 'Low-pass' stimuli, and including either the first 12 or the second 12 images in the set. The order of sessions was counterbalanced across participants. Each session included 12 randomly interleaved adaptive psychophysical staircases (one for each of the 12 images) (Kontsevich and Tyler, 1999) consisting of 30 trials each. The adaptive staircase set the degree of phase randomization on each trial in order to reliably estimate the detection threshold of each image (the point at which the participant was 75% correct in identifying whether the object was smaller or larger than a shoebox). At the completion of the 8 sessions, we had two estimates of detection threshold of each image in both the 'Low-pass' and 'High-pass' conditions. Results and stimuli from the first psychophysics experiment are included in the Supplementary Material.

The average detection thresholds across participants were used to generate low and high-pass versions of each object that were of approximately equal detectability. We found the maximum possible signal for which the signal in the low and high-pass images were equal multiples of the average detection threshold, and used these maximum matched signal values to define the 'Low-pass' condition and the 'High-pass' condition of the MEG experiment. Images in the 'Strong signal' condition were defined by setting the signal in both the low and high spatial frequency bands to these maximum matched values.

Finally, in the second psychophysical experiment, we calibrated the signal level in the 'Weak signal' condition individually for each of the nine participants who went on to complete the MEG experiment.