**First Year Project**

**Interactions Between Feature Representations are Well Predicted by a Hierarchical Model of Visual Cortex.**

**Or**

**Neural Substrates of Attention and Awareness.**

Dan Birman, PI: Justin Gardner

**Introduction**

In everyday life we feel a direct and undeniable connection between attending to something and our clear awareness of it. Despite this there exist a variety of laboratory situations in which attention and awareness appear to diverge. This apparent disconnect between experience and experimental findings has fueled a debate about whether *selective attention* and *awareness* are dissociable (Koch & Tsuchiya, 2007). In scene recognition experiments participants are able to identify the content of peripheral stimuli despite attending to a demanding fixation task (Li, VanRullen, Koch, & Perona, 2002). In contrast, many studies have found that participants are unable to respond about unattended stimuli, in particular for simple features such as shapes and colors (Mack, Arien & Rock, Irvin, 1998). One possible interpretation for the variability in results in these tasks is that stimulus features may be interacting in unexpected ways. The neural activity responsible for scene recognition may be modified only by certain types of attentional modulations. For example, attending to motion appears to result in diminished awareness of scene gist (Cohen, Alvarez, & Nakayama, 2011), while attending to other features may not cause in-attentional blindness. We propose that this feature-specific hypothesis may depend crucially on the layout of visual cortex. Our hypothesis is that modifying a neural representation through spatial or feature-based attention is only detrimental to awareness of other features when there are direct connections between the relevant neural representations. Contrast is well specified by neural activity in the earliest cortical visual area, V1. The human MT+ complex on the other hand is less sensitive to contrast, but considerably more sensitive to the presence or absence of motion in a stimulus. We propose to test our hypothesis by measuring behavioral and neural activity related to the representations of contrast and motion when participants are cued to a feature. Our hypothesis predicts that attention to contrast will be detrimental to the perception of motion, but not vice versa.

**Methods**

**Subjects**

Four human subjects (all male, ages 24-34) participated in the experiment. All subjects performed the behavioral experiment and one participant performed the functional MRI experiment. Subjects in the behavioral experiment performed one training session to become accustomed to the task, four to eight control runs (65 trials each), and between 12 and 24 task runs (100 trials each, 15% miscued). Subjects in the functional experiment performed a retinotopic mapping (1 hr, consisting of ten 4 minute scans) and four to five sessions of the main experiment (2 hrs each, consisting of ten 7 minute scans).

**Experimental Task**

Subjects performed a two-alternative forced choice discrimination task. On each trial participants were shown two patches of dots and asked to report which had a higher apparent contrast or motion coherence. Each dot patch was shown for 750 ms and followed by a 250 ms mask, generated by flashing random checkerboards at 55% contrast at 40hz. A random inter stimulus interval followed for between 200 and 500 ms after which the fixation cross turned white, indicating that the response period was starting. Participants had 1 second to respond. Each trial was followed by a random inter-trial-interval of 300 to 500 ms. Participants fixated a central cross (1 deg x 1 deg visual angle, 1 pixel wide, luminance ?!). During control runs participants were cued to attend to either motion or contrast and responded about the cued feature. On task runs participants responded about the miscued feature on 15% of trials. The trial timing was modified during scanning to improve estimation of the hemodynamic response, stimulus: 750 ms, mask: 250 ms, ISI: 200-1000 ms, resp: 1000 ms, ITI: 2000-10000 ms. During the behavioral experiment the screen updated at 100 hz and at 60 hz during scanning.

The dot patches appeared left and right of fixation, extending from 3.5 to 11 degrees horizontally and from -5 to 5 degrees vertically. The patches were displayed on a gray background (50% luminance) on a monitor with a linearized luminance scale. Each patch contained 1000 dots, half of which were darker than the background and half of which were equally brighter. The luminance difference between the dots and the background was defined as C / 2, where C is the contrast (0 to 1) on the current trial. A percentage of the dots, M, moved horizontally either right or left (randomly chosen on each trial) while the remaining 1 – M dots had random angles. M therefore reflects the motion coherence of the dot patch. All dots moved at a consistent speed of 3.25 degrees / s.

Contrast and motion discrimination performance was tested at a single pedestal intensity. Contrast was tested at 60% and motion coherence was tested at 10%. Both features were crossed such that neither feature was informative about the strength of the other feature on any given trial. For each feature a 1-up-3-down staircase (??) was used to set the increments in contrast or motion coherence that was added to the pedestal contrast on the target side. The independent staircases balanced task difficulty across the features so that subjects were always performing the task at a near-threshold level, eliminating any potential confound with task difficulty between conditions. During control runs pedestal values of 20/40/60/80% contrast and 0/10/20/40% coherence were used to allow estimation of the BOLD response across a larger range of feature intensities. For each pedestal an independent staircase was computed to maintain task difficulty near threshold. Pedestals were pseudo-randomly interleaved across trials.

**Stimulus Presentation**

Outside the scanner the visual stimuli were presented on a ViewPixx 22.5” LCD (VPixx Technologies) with a resolution of 1920 x 1200 pixels and a 100 Hz refresh rate at a distance of (?) cm from the subject’s eyes to obtain a field of view of ? x ?. Inside the scanner subjects used an adjustable mirror system to view an image that was rear-projected onto a fiberglass screen using an Eiki LC-WUL100L projector operating at 1920x1200, 5000 lumens, projected through a neutral density filter. The projector and LCD screen were calibrated to have linearized gamma scales using a PR650 Spectroradiometer (Photo Research Inc., Chatsworth, CA.). We dynamically adjusted the 10-bit gamma table to achieve the best luminance resolution possible (maintaining the linearized output) for displaying each dot patch. All stimuli were produced using MATLAB (The Mathworks Inc., Natick, MA, USA) and MGL (<http://gru.stanford.edu/doku.php/mgl/overview>).

**Eye Position Measurements**

An Eyelink 1000 eye tracking system (SR Research Ltd., Mississauga, ON, Canada) was used outside the scanner to confirm that subjects maintained fixation throughout the task. Eye tracking was not performed inside the scanner. The Eyelink system recorded corneal reflections of an external infrared light source and tracked the center of the pupil. A brief calibration was performed before each 5-minute run. Eye tracking setup was successful for all sessions. The calibration data was used to perform an affine transformation of the acquired eye tracking data to the position of the eye in degrees of visual angle.

Todo: eye position analysis

**Contrast and Motion Discrimination Functions**

Feature discrimination task performance was evaluated using feature-discrimination functions. A feature-discrimination function defined the relationship between the pedestal intensity (i) and the increment in intensity (deltai) required to obtain threshold-level performance. Feature-discrimination functions were computed separately for contrast and motion coherence. For each condition a maximum-likelihood procedure (Wichmann & Hill, 2001) was used to fit subject responses to a Weibull function (Weibull, 1951):

p (c ) = ….

Where p(delta i) is the probability of being correct given an intensity increment of delta I, lambda is the lapse rate, epsilon is the delta I for which the probability correct reaches 63% of the difference between chance and maximal performance, and m is the slope of the psychometric function. Subjects performed on average X psychometric functions with X trials each. A minimum of 50 trials were allowed per function, sufficient to estimate the discrimination threshold accurately (Kontsevich & Tyler, 1999). By running multiple staircases we were able to compute the variability of the threshold across runs

**MRI Acquisition and Preprocessing**

MRI data were acquired on a GE Discovery MR 750 on a Nova Medical 32ch head coil. Retinotopy experiments were collected on a Nova Medical 16ch visual array. For each subject we acquired a high-resolution 3D anatomical image (“canonical anatomy”) which was segmented via FREESURFER (http://surfer.nmr.mgh.harvard.edu) to generate white matter and gray matter segmentation (Dale, Fischl, & Sereno, 1999). We collected a single T1-weighted image (MPRAGE TR ??, TI ??, TE ??, FA ??, voxel size ??, matrix ??. Regions of interest were drawn on flattened representations of the cortical surface including the visual areas and the motion sensitive regions that defined hMT+. These regions of interest were constrained to voxels that intersected the gray matter. Analyses were conducted on original untransformed data while flattened representations were used for visualization.

Each functional experimental session consisted of a lower resolution T1-weighted image (“session anatomy”) (????) and multiple T2\*-weighted functional scans (multiband 8, TR 500 ms, TE 30 ms, flip angle ??, voxel size 2.5 x 2.5 x 2.5 mm, matrix 88 x 88). An automated procedure was used to find the best affine transform to align the session and canonical anatomy (Nestares & Heeger, 2000). The functional scans were aligned to the session anatomy directly using the coordinates measured by the scanner. Retinotopic mapping was performed using a T2\*-weighted functional scan (multiband 2, TR 1400 ms, TE 30 ms, flip angle ??, voxel size 2.5 x 2.5 x 2.5 mm, matrix size). Oblique slices were chosen to maximally cover the occipital visual areas, approximately perpendicular to the calcarine sulcus. For all subjects our functional sequences achieved full brain coverage.

fMRI images were analyzed through a pipeline using mrTools.

**Retinotopy**

Visual fields were determined based on a retinotopy performed in a separate scanning session. High-contrast radial checkerboard patterns were presented either as an expanding or contracting ring or a 90\* rotating wedge. Each scan consisted of 10.5 cycles (24 s per cycle) of the ring expanding/contracting or the wedge completing a full rotation with a sampling rate of 17 volumes per cycle (178 volumes per scan). In addition four presentations of a sweeping bar stimulus were made. Each session therefore consisted of two scans of the ring stimulus (one expanding, one contracting), four scans of the wedge stimulus (two each clockwise and counter-clockwise), and four scans of the bar stimulus. A generative model of voxel responses (the Population Receptive Field model, ??) was fit to each voxel, identifying the Gaussian response field parameters that best fit the recorded response data. Visual fields were then defined according to published criteria.

**Feature-response Functions**

To compute the feature-response functions, a deconvolution analysis (for details see: (Gardner et al., 2005)) was used to determine the mean hemodynamic response to each dot patch in the contralateral visual cortex. The average time-course in each visual area for each grating location was computed and the response following stimulus presentations for 20 s was calculated, assuming linear summation for responses that temporally overlapped. These responses were calculated separately for each combination of feature (contrast, motion coherence) and cueing condition (cued, miscued) at every intensity increment, rounded to the nearest 10%. This resulted in 36 total conditions (contrast: 8 intensities x 2 cueing + coherence: 10 intensities x 2 cueing). A gamma function was fit to this deconvolved response and the amplitude of this function determined the magnitude of response. These response magnitudes were then plotted as a function of stimulus intensity to yield the contrast-response function for each visual area and cue condition.

**Results**

**Discussion**

Our hypothesis is that the hierarchical organization in visual cortex is sufficient to predict the asymmetrical interactions that can occur between features in visual tasks. For example, it is well known that under demanding attention conditions our ability to report the ‘gist’ of a scene remains largely unimpaired (Li et al., 2002). But this effect breaks down under specific experimental conditions, for example when the attention demanding task requires motion tracking (Cohen et al., 2011, p. 201). We propose that reconciling these results requires an understanding of how neural activity in visual cortex is interconnected. In our design we asked participants to report responses about a single feature, either stimulus contrast or motion coherence, while simultaneously manipulating the intensities of both features. This allowed us to estimate the impact of being cued and asked to respond about different features (task: miscued), compared to responding about the cued feature. We found behavioral results that supported our hypothesis, showing that miscuing coherence is more detrimental to performance than miscuing contrast. We recorded BOLD responses in visual cortex to investigate whether specific voxel populations might be predictive of this asymmetrical cueing relationship. We found that ??!? I don’t know!?!?

**References**

Cohen, M. a, Alvarez, G. a, & Nakayama, K. (2011). Natural-scene perception requires attention. *Psychological Science*, *22*(9), 1165–72. http://doi.org/10.1177/0956797611419168

Dale, A. M., Fischl, B., & Sereno, M. I. (1999). Cortical surface-based analysis: I. Segmentation and surface reconstruction. *Neuroimage*, *9*(2), 179–194.

Gardner, J. L., Sun, P., Waggoner, R. A., Ueno, K., Tanaka, K., & Cheng, K. (2005). Contrast adaptation and representation in human early visual cortex. *Neuron*, *47*(4), 607–20. http://doi.org/10.1016/j.neuron.2005.07.016

Koch, C., & Tsuchiya, N. (2007). Attention and consciousness: two distinct brain processes. *Trends in Cognitive Sciences*, *11*(1). http://doi.org/10.1016/j.tics.2006.10.012

Kontsevich, L. L., & Tyler, C. W. (1999). Bayesian adaptive estimation of psychometric slope and threshold. *Vision Research*, *39*(16), 2729–2737.

Li, F. F., VanRullen, R., Koch, C., & Perona, P. (2002). Rapid natural scene categorization in the near absence of attention. *Proceedings of the National Academy of Sciences of the United States of America*, *99*(14), 9596–601. http://doi.org/10.1073/pnas.092277599

Mack, Arien, & Rock, Irvin. (1998). *Inattentional Blindness.* The MIT Press.

Nestares, O., & Heeger, D. J. (2000). Robust multiresolution alignment of MRI brain volumes. *Magnetic Resonance in Medicine*, *43*(5), 705–715.

Weibull, W. (1951). A Statistical Distribution Function of Wide Applicability. *Journal of Applied Mechanics*.

Wichmann, F. A., & Hill, N. J. (2001). The psychometric function I: Fitting, sampling, and goodness of fit. *Perception & Psychophysics*.