CHANNELS FOR OBJECTS AND ENVIRONMENTS IN MACAQUE INFEROTEMPORAL CORTEX

by

Siavash Vaziri

A dissertation submitted to The Johns Hopkins University in conformity with the

requirements for the degree of Doctor of Philosophy.

Baltimore, Maryland

August, 2014

Abstract

Visual information processing in the primate brain is thought to occur along two major pathways, known as the ventral and dorsal pathways. The ventral ('what') pathway has been considered a system specialized for processing visual information about object structure/shape and identity. Evidence for this view comes from wide-ranging studies based on fMRI, behavioral effects of lesions and neural recording. In particular, recording studies in the macaque have revealed strong sensitivity to object shape information, ranging in complexity from 2D local contour orientation in early visual areas to 3D local surface curvature in the inferotemporal (IT) cortex, a region comprising the final stages of the ventral pathway.

While object vision has been the primary focus of research in the ventral pathway, fMRI studies relating to visual representation of places or environments (e.g. landscapes and interiors) have implicated dorsal pathway areas and the parahippocampal cortex. The goal of the research presented here was to test for the first time using neural recording whether neurons in anterior IT cortex of the macaque ventral pathway are in addition to objects also sensitive to environmental shape.

Neural responses were sampled using large-scale abstract visual stimuli that resembled landscapes and interiors. Specifically, visual stimuli spanned the entire scale

continuum from object-scale to environment-scale stimuli and an adaptive sampling approach was used to efficiently sample the virtually infinite span of shape space.

The study revealed a surprisingly strong sensitivity to environment-scale shapes in anterior IT. Additionally, there appeared to be an anatomical segregation in stimulus preference netween two IT sub-regions. Area TEd neurons were found to be predominantly selective for environment-scale stimuli whereas neurons in STSv were almost exclusively selective for object-scale stimuli. Extensive tests confirmed that neural selectivity for environment-scale stimuli is critically dependent on 3D shape.

Areas TEd and STSv have been recently described as forming separate processing channels in the ventral pathway based on anatomical connectivity evidence. Our study is the first to provide a functional distinction between these channels based on object versus environment processing.

Primary Reader: Dr. Charles E. Connor

Secondary Reader: Dr. James Knierim

Acknowledgments

I'm indebted to my parents for the sacrifices they made in their lives for my education. I'm grateful to my brother whose personality provided the perfect antidote to my often stress-induced crankiness. I'm especially fortunate to have met my wife during graduate school. Her companionship provided a critical booster for me to complete my research. I want to thank my advisor for providing the time and resources that allowed me to complete a challenging experiment in exactly the way I had first imagined it. Thank you to all the members of the MBI staff for maintaining an environment conducive to research. To all my lab-mates thank you for your friendship, I wish you all the best in your careers. YNWA.

Contents

List of Figures

Chapter 1

General introduction

1.1 Organization of primate ventral pathway

For over thirty years, the ventral visual processing pathway of the primate brain has been regarded as the neural substrate behind our capacity for rapid and robust object perception (Mishkin et al., 1983; Ungerleider and Haxby, 1994). This is a computationally difficult task given the variations of the retinal image possible for a given object and the virtually infinite size of object space (Thorpe et al., 1996). The original characterization of the monkey ventral pathway describes it as a serially staged hierarchy starting from the striate cortex (V1) and terminating in the inferotemporal cortex (TE). Intermediate processing stages are V2, V4 and TEO. The

gradually increasing complexity of neuronal response properties for visual stimuli, increasing receptive field sizes and lengthening of response latencies have been used as evidence for the presence of a hierarchical framework. In V1, neuronal responses can be characterized based on tuning for local orientation and spatial frequency (Hubel and Wiesel, 1962; De Valois and De Valois). Tuning for local orientation extends to illusory contours (Von der Heydt et al., 1984) in V2 neurons which also respond to more complex shape properties such as edge junctions (Hegdé and Van Essen, 2000). In V4, there appears to be strong sensitivity to contour curvature (Pasupathy and Connor, 1999), the first derivative of local contour orientation. As will be described in the next section, IT neurons respond to further complex properties of stimulus shape.

Based on a synthesis of neuroanatomical evidence, a recent analysis by Kravitz et al. (2013) has re-characterized the ventral pathway as a highly recurrent occipitotemporal network consisting of multiple processing channels. Specifically, macaque IT consists of three processing channels as shown in figure 1.1. The first channel runs along the ventral bank of the superior temporal sulcus (STSv). The second channel goes through area TEd in the dorsolateral half of the inferior temporal gyrus. The third channel goes through area TEv in the ventromedial half of the inferior temporal gyrus. These areas have distinct connections to output targets such as prefrontal cortex, basal ganglia and medial temporal structures. These distinct connectivity patterns point towards potentially major functional specializations within

IT cortex, a region that has mostly been studied for its role in object vision. The evidence presented here will highlight a major difference in function between 2 of these anatomically distinct channels (STSv and TEd) in terms of preferential visual processing for objects versus environments/places.

1.2 Macaque inferotemporal cortex: Object vision

The inferotemporal cortex is a high-level visual area considered to be critical for processing information about object form and identity. Bilateral removal of anterior regions of IT cortex (area TE) produces severe impairment in object discrimination tasks. In contrast, tasks involving evaluation of spatial relations between objects are only affected by bilateral removal of the posterior parietal cortex located in the dorsal visual processing stream(Mishkin et al., 1983; Mishkin, 1982; Pohl, 1973).

A lot has been learned about IT neural response properties since initial descriptions of larger receptive fields and more complex preferred visual stimuli (e.g faces) compared to edge-detectors of early visual cortex (Gross et al., 1972; Desimone et al., 1984). There is now strong evidence that IT neurons are sensitive to object form or shape (Tanaka, 1996; Kourtzi and Connor, 2011). Additionally, it has been reported that IT response selectivity for object shape is organized in a columnar fashion (Fujita et al., 1992).
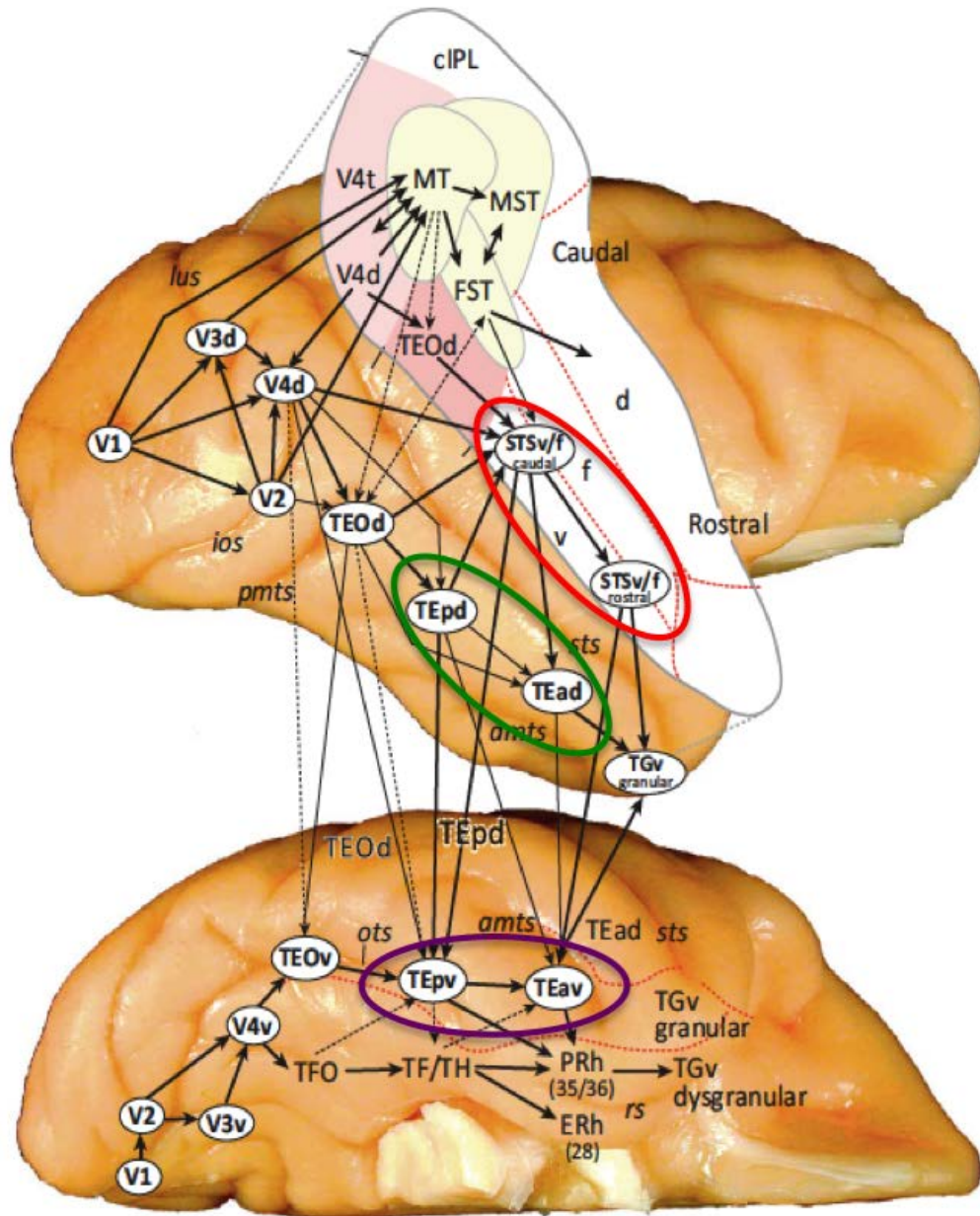
Figure 1.1: Processing channels in macaque inferotemporal cortex. Connectivity pattern of the occipitotemporal network in the macaque brain shown in lateral (top) and ventral (bottom) views. Strength of connections are indicated by line thickness. Based on these connections there appears to exist 3 parallel processing channels (highlighted by colored ellipses) going through the anterior inferotemporal cortex. These are STSv (red), TEd (green) and TEv (purple). Adapted from Kravitz et al. (2013).

Despite the challenge of adequate sampling in the enormous space of objects recent novel approaches have provided some insight into the nature of object representation in IT. A single-cell study using an evolutionary sampling algorithm has demonstrated tuning for explicit 3D object shape information in terms of curvature, relative position and orientation of local object features (Yamane et al., 2008). Quantitative analyses of shape tuning in IT has also demonstrated sensitivity to more abstract shape descriptors such as the medial axis structure of objects (Hung et al., 2012). These findings support structural theories of object shape (Marr and Nishihara, 1978; Biederman, 1987) which are based on a three dimensional representation of structural parts labeled, shape primitives, and their spatial relationships. Such coding schemes provide a more compact and explicit representation of object shape compared to the highly distributed and implicit representation in primary visual cortex.

Object representation in IT based on semantic or categorical information is an alternative theory to coding schemes based on structural information. Our ability to group objects using non-structural information such as animacy, behavior, utility and other conceptual associations suggests the presence of categorical representations (Kourtzi and Connor, 2011). It has been shown using multidimensional scaling analysis that the response pattern of a population of IT neurons differentiates between animate and inanimate object categories, as well as within category subdivisions involving faces, hands and bodies (Kiani et al., 2007). These categorical distinctions based on differential response patterns have also been observed in humans based on

imaging work (Kriegeskorte et al., 2008).

Object representation in IT also appears to be modular for certain object categories of evolutionary importance, such as faces. There appears to be a network of face patches dedicated for the processing of faces (Tsao et al., 2008). In similar fashion, there also exist regions that in IT that are particularly sensitive to color information (Komatsu et al., 1992,?; Lafer-Sousa and Conway, 2013). Therefore, IT cortex appears to show a high degree of functional specialization that is further corroborated by the results presented here.

1.3 Visual processing of places or environments

As the previous section highlighted, much of the research dedicated to the study of macaque IT has focused on object vision. However, objects in the real-world exist within places or environments, such as landscapes or interiors. Figure 1.2 shows two example real-world places in which the environment structure (green) has been distinguished from the objects (red). Objects can be described as entities that are relatively small in scale and are bounded within the visual field. In contrast, environments are composed of unbounded surfaces that are large-scale and extend beyond our visual field. This vast difference in scale and topology between object and environment structure forms the primary focus of the investigation presented here.

There have been many studies on place/environment visual processing based on more global techniques such as fMRI. These studies were initiated by the discovery of the human parahippocampal place area (PPA) (Epstein and Kanwisher, 1998), a cortical region in the parahippocampal gyrus that exhibits larger BOLD signals when subjects view images of environments versus objects. There is considerable debate over the nature of PPA selectivity for environments, ranging from arguments for sensitivity to the spatial layout of environments (Kravitz et al., 2011; Park et al., 2011), context-dependent representation of objects (Bar, 2004; Bar et al., 2008) and selectivity for high spatial frequencies present in images of environments (Rajimehr et al., 2011). In addition to PPA, other environment-selective regions have been identified in the retrosplenial complex and the transverse occipital sulcus, with parallel fMRI studies in macaques identifying potential homologs (Nasr et al., 2011). Lesions in the PPA lead to anterograde amnesia for scene layouts (Aguirre and D'Esposito, 1999). In this condition patients have difficulty orienting themselves in novel environments. This result combined with fMRI experiments that have pointed to a lack of a familiarity effect (higher responses for images of familiar versus unfamiliar scenes) suggests that activation in PPA does not reflect processes directly involved in recognition but might reflect the encoding of current spatial information about scenes (Epstein et al., 1999).

There has only been one neural recording study of environment processing in macaques (Kornblith et al., 2013). In this study fMRI was used to identify an envi-

ronment selective region (LPP: Lateral Place Patch) located in posterior regions of the ventral pathway (just anterior to V4). Single-unit recording in this area showed about 50% of the neurons had significantly higher responses to photographs of scenes compared to objects, textures and scrambled scene parts. Concurrent micro-stimulation of area LPP with fMRI allowed for an identification of a second potential scene selective region (MPP: Medial Place Patch) in a quite medial location, area TFO. Targeted neural recording in this area also confirmed selectivity for photographs of environments. It was also reported in this study that responses were modulated by texture, which, as will be demonstrated here, serves as an important depth cue.

Although, most areas implicated in environment processing are located outside of anterior regions of IT cortex (TE), a recent high resolution human fMRI study (Stansbury et al., 2013) has demonstrated scene-related activity in more anterior regions of the ventral pathway. The work presented here is the first neural recording study that comprehensively looks at sensitivity to both objects and environments in anterior inferotemporal cortex.
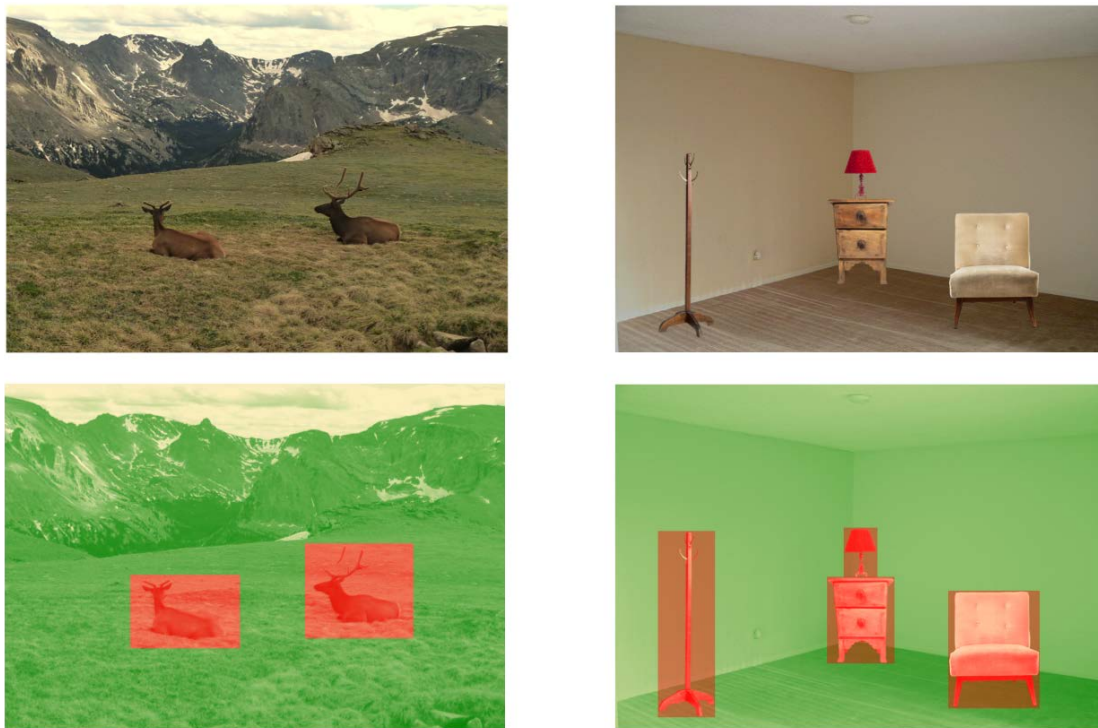
Figure 1.2: Objects and environments. Top panels show 2 example scenes. For each scene the corresponding panels at the bottom segregate the objects (red) from the background structure or *environment* (green). For the scene on the left the term environment refers to the landscape terrain and for the scene on the right it refers to the large-scale interior structures (walls, floor and ceiling).

1.4 Summary of chapters

In chapter 2, the methodology is described with particular emphasis on the design of abstract visual stimuli resembling objects and environments. The stereoscopic display setup that was used to show very large-scale visual stimuli will also be described. The particular implementation of an adaptive sampling algorithm used to probe neural responses will be explained with detailed information about the parameterization of scale and shape morphing techniques.

In chapter 3, the main neural recording results from the adaptive sampling algorithm in processing channels STSv and TEd will be presented. Specifically, analysis of shape scale sensitivity will demonstrate a biased preference for object-scale versus environment-scale visual stimuli between STSv and TEd. Analyses looking at temporal response properties and response distribution selectivity based on a sparseness measure are also presented. The consistency of the adaptive algorithm under different initial conditions is analyzed by comparing lineage-specific response patterns.

In chapter 4, the results from a series of auxiliary tests for low-level image properties will be presented. These include potential sensitivity to 2D stimuli, texture frequency, contrast patterns from different lighting conditions and peripheral visual stimulation, regardless of shape scale. In addition, results are presented from a test looking at the sensitivity of TEd neurons to shading, stereo and texture 3D depth cues.

In chapter 5, the implication of the results with respect to current understanding

of IT visual processing is discussed, with particular emphasis on its potential role in representation of environments. Future experiments and analyses that will help further clarify the nature of IT sensitivity to environmentsl visual information is described at the end.

Chapter 2

Methods

Recording and behavioral procedures are described in the following sections. The method for generation of abstract shapes that resemble objects and environments are explained in detail. This formed a critical part of the project. The display setup used in this project will be described. The display system was created to be as immersive as possible for conveying the difference in scale between object and environment structure. The particular implementation of an adaptive stimulus sampling algorithm and the stimulus transformations used in the algorithm are also covered in this chapter.

2.1   Behavioral task

Single-unit recording was performed on two male rhesus monkeys (Macaca mulatta) while they performed a passive fixation task. They were trained to maintain

fixation within 1^o(radius) of a 0.1^o diameter spot for trials lasting 4 seconds each in order to obtain a juice reward. Eye position was tracked using a 1000 Hz infrared eye-camera (EyeLink). The monkeys were seated at a distance of 55 cm from the visual display screen. The head was immobilized with a titanium post mounted in an acrylic cap attached to the skull with orthopedic screws. All animal procedures were approved by the Johns Hopkins Animal Care and Use Committee and conformed to US National Institutes of Heath and US Department of Agriculture guidelines.

2.2 Electrophysiological recording procedures

The electrical activity of well-isolated single neurons was recorded with epoxy-coated tungsten electrodes (Microprobe or FHC) that were guided into the brain using a custom-built micro-drive assembly placed on a recording chamber over the craniotomy. The custom electrode drive had two degrees of rotational freedom allowing access to the entire extent of the inferotemporal cortex from a single circular craniotomy of 1 cm diameter. The electrodes were inserted through a transdural guide tube. Action potentials of individual neurons were amplified and electrically isolated using a Tucker-Davis Technologies recording system. Individual neurons were isolated online using a custom-built software implementing a level trigger and multiple-window discriminator.

Recording locations were targeted stereotaxically using structural magnetic resonance images obtained prior to the start of experiments. For each desired recording target identified in the MRI image the required angles and depth of electrode penetration were determined based on geometrical information about the recording chamber. The determination of recording location was further refined based on sequences of sulci and response characteristics observed while lowering the electrode. Inferotemporal neurons were sampled in the posterior-anterior range of +8 to +22 mm (with respect to the inter-aural line) in the ventral bank of the superior temporal sulcus (STSv) and area TEd.

2.3   Trial paradigm

During fixation visual stimuli were flashed on the screen for a presentation period of 750 ms and with an inter-stimulus interval of 250 ms. 4 stimuli were displayed per trial. The fixation target appeared at screen center, at one of 4 right/left image disparities corresponding to fixation at 55 cm (screen depth), 100 cm, 10 m, or 100 m. The stimulus was always tangent to the fixation point, at 0 disparity, so the monkey was always fixating a point on the stimulus surface. The fixation depth for stimuli within a trial was kept the same to avoid rapid changes in fixation point disparity.

2.4   Visual stimulus display setup

A visual display system was set up with the aim of conveying as much as possible the vast difference in scale between environment and object structure. Therefore, in order to maximal realism the required display system needed to be capable of showing very large-scale stimuli. To achieve this a stereoscopic projection system onto a large screen was used as shown in figure 2.1. The screen subtended 77 degrees of visual angle horizontally and 61 degrees vertically at a viewing distance of 55 cm from the monkey. Stereo depth cue was conveyed using differential polarization of the left and right eye stereo images (rendered in openGL) back-projected from two precisely aligned high-resolution (1400 X 1050) color projectors. Alignment of projectors was performed by superimposing two projected rectangular grids (one from each projector) using electronic lens shift features. Linearly polarized filters with mutually perpendicular axes of polarization were positioned in front of each projector lens. Matching filters were positioned in front of the monkey's eyes. The screen (Stewart - FilmScreen 150) was a rear-projection screen that maintained light polarization. Prior to each experimental session, binocular fusion and stereoscopic depth perception were verified with a random dot stereogram saccade target detection task. In addition to stereo, visual stimuli were rendered with texture and shading depth cues which unlike stereo are important for conveying depth structure at far distances.
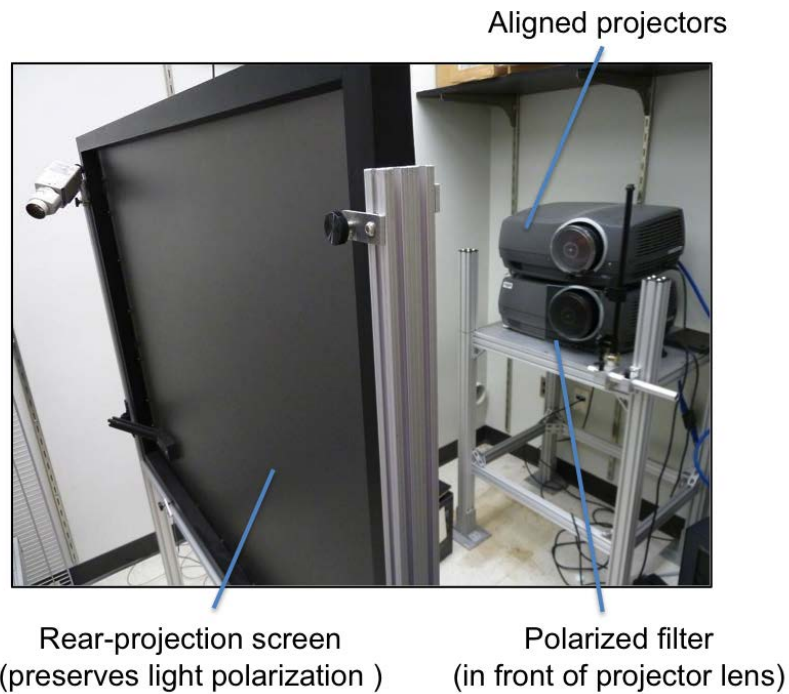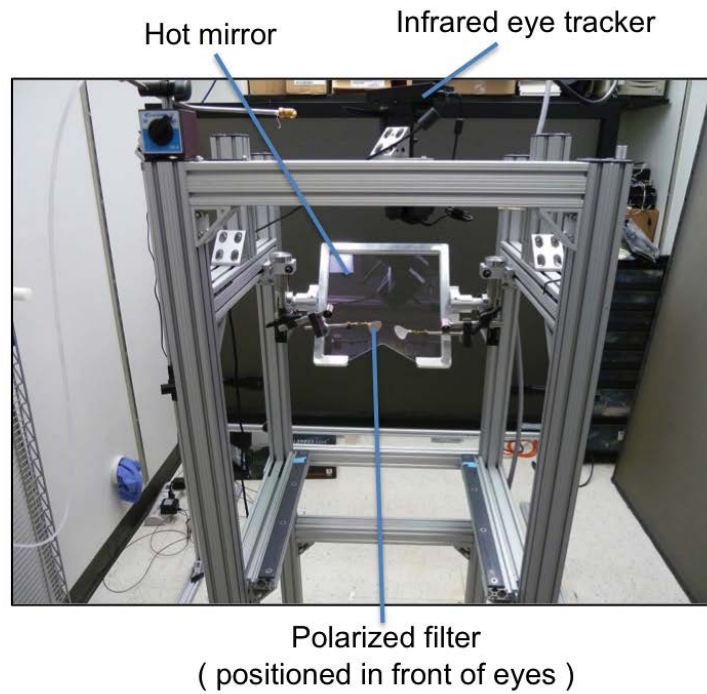
Figure 2.1: 3D display setup. Frontal (top panel) and rear view (bottom panel) of the visual display system. Different components of the setup have been labeled. The aluminum frame in the top panel is the docking location for the monkey chair. Im-

ages were displayed in stereo using a passive polarization technique. High luminance output projectors were used to offset brightness reduction due to filter and screen media. Optic fiber cables positioned at the corners of the screen image were used to detect precise onset and offset of visual stimuli. As indicated an infrared eye camera was used to track eye position.

2.5   Visual stimulus shape generation

Software was created capable of generating unique abstract shapes resembling objects and environments. These abstract shapes served as visual stimuli to probe neural responses. As shown previously for the domain of objects (Yamane et al., 2008; Hung et al., 2012), abstract stimuli make it possible to primarily focus on shape/form information because they lack the semantic confounds associated with natural image stimuli. Here we tested abstract stimuli that ranged in scale from object-like shapes subtending a few degrees of visual angle to environment-like shapes that extended beyond the 77 degree boundary of the display screen, resembling landscapes and interiors (figure 2.2). Shapes of intermediate scale that resembled very large objects were also explored. In fact, as demonstrated in figure 2.3, the stimulus-generation software was capable of sampling in a parameterized way the entire scale continuum from object to environment-scale stimuli. The use of custom designed abstract shapes allowed for the parameterization of scale which is the target of a quantitative analysis presented in chapter 3.

Each stimulus was derived from a spherical B-spline surface represented as a polar grid of control points. There are several advantages to using spline-based visual stimuli. OpenGL, which was used to render stimuli, has a very fast and stable algorithm to evaluate B-spline vertices. Fast generation of stimuli was important since as will be explained in section 2.6 an adaptive sampling algorithm was used that requires online generation of stimuli based on neural feedback. Most importantly spline based

stimuli make it relatively easy to define a large variety of shapes by simply applying various transformations to the control points.

The control points of the starting sphere were arranged as a series of latitudinal rings between the two poles of the sphere. Initially, the spherical spline surface was set to a particular scale based on the angular subtense of its diameter which could range from 4 degrees to 142 degrees, encompassing the object to environment scale continuum. Note that since the boundary of the screen was limited to 77 degrees, stimuli derived from spheres of very large diameters (beyond 77 degrees) only had their visible surface displayed on the screen. The global orientation of the sphere was set by randomly applying rotations about the cardinal axes. Next the sphere was randomly modified in shape to produce each unique stimulus. These shape variations were produced by applying specific changes to the position of the control points. Large-scale shape variation was produced by randomly changing the positions, orientations, sizes and aspect ratios of 2-5 latitudinal control point rings forming the spline sphere. Smaller scale variations were based on defining points and paths along the surface and applying random height perturbations creating structures resembling hills and valleys in the case of large-scale stimuli. Surface smoothness was also parameterized. This paramterization covered the range of curvilinear structures found in natural landscapes to rectilinear structures of man-made environments (e.g interiors). Note in the case of environment-like stimuli that resembled interiors the implicit position of the template sphere was such that the viewer was located inside the sphere.
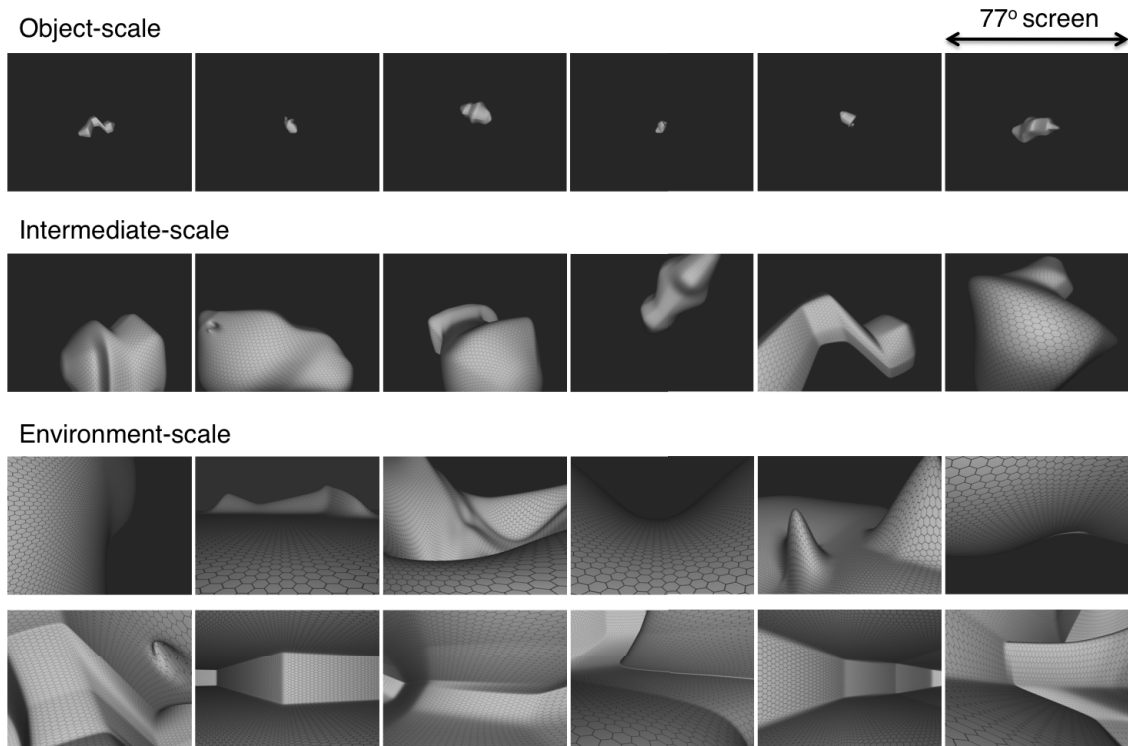
Figure 2.2: Visual stimulus paradigm covering objects and environments. Example abstract visual stimuli generated using a custom-made spline-based shape generation software. Stimuli have been arranged into three categories of shape scale: Object-scale, Intermediate-scale and Environment-scale. Large-scale shapes resembling landscapes and interiors exceed the 77 degree boundary of the display setup. In addition to stereo depth cues, texture and shading add to 3D realism.
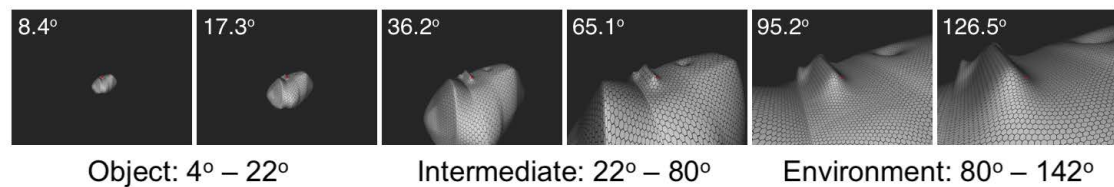
Figure 2.3: Shape scale continuum. Example stimulus shown at six different scales sorted in increasing magnitude from from left to right. Angular values indicate the visual angle subtended by the sphere diameter used to create each stimulus. The fixation point (red dot) remains at the same point on the stimulus surface. As scale increases, the stimulus extends beyond the screen borders, and the visible portion becomes landscape-like. The range of scale tested was between 4 and 142 degrees.

20

2.6   Adaptive stimulus sampling algorithm

The main challenge to neural recording studies of complex visual form is that the visual stimulus domain is virtually infinite. Even if the scope of visual stimuli is constrained to shapes of limited complexity adequate sampling would require stimulus set sizes far beyond the practical limits (greater than 1000 stimuli) of recording experiments. Therefore, sampling methods based on a random selection of stimuli could easily fail to detect the maximal response range of the neuron. This would be especially problematic in this investigation which involves the comparison of neural responses to object and environment-like stimuli. Preference may be incorrectly attributed to one of the two stimulus categories (object versus environment) if the random stimuli of the truly preferred category fail to sample the high response range of the neuron.

As an improvement over random sampling, adaptive search methods use neural responses to guide sampling towards the high response region of the shape domain for the particular neuron. Previously, this method has been successfully applied to object shape coding (Yamane et al., 2008; Hung et al., 2012). Given a limited number samples, adaptive methods provide a more focused sampling in the relevant response region of the neuron compared to random methods. Therefore, adaptive sampling was used here to allow for a more fair comparison between optimized objects and optimized environment stimuli.

2.6.1 Specific implementation of adaptive sampling

Initially responses are obtained to a set of randomly generated shapes. These are called the first generation stimuli. Importantly, stimuli in the object, intermediate and environment-scale range are equally sampled. Each stimulus is repeated 5 times and the order of stimulus presentation is randomized. For the next generation, 20% of the stimuli are random shapes and 80% are partially morphed versions of stimuli (ancestors) from generation one. In section 2.7 the morphs will be described in more detail. The selection of ancestor stimuli from generation 1 to produce morphed descendants was probabilistic. High response stimuli had a greater probability of producing morphed descendants. Specifically, ancestors were rank ordered based on their response and divided into 5 percentile ranges. Thirty percent of ancestors were from the top range (90-100% of maximum response), twenty percent each from the three middle ranges (70-90, 50-70, 30-50%), and ten percent from the bottom range (0-30%). It is important that descendants from low to medium response stimuli are collected to prevent an early convergence of the algorithm to a local minimum. Sampling from the lower response stimuli also helps with constraining of shape-tuning models.

This procedure is repeated for multiple generations for as long as the neuronal isolation can be maintained, typically 5-8 generations for a total of 400-640 stimuli. Each generation consists of both random shapes and morphed descendants generated from potential ancestors across all previous generations. Critically, for each cell two

lineages of our adaptive algorithms are performed where each lineage begins with a different set of random shapes and independently evolves across multiple generations. Running these 2 lineages allows verification of convergence under different initial random shapes. Each generation in this adaptation procedure comprises 40 stimuli in each of the 2 lineages, for a total of 80 stimuli.

The particular values used for the proportion of random versus morphed stimuli and the values specifying probability distributions for ancestor selection were chosen based on previous experiments. The challenge is to find values that provide a balanced trade-off between fuller exploration of the stimulus search space and refinement of existing solutions.

As mentioned earlier, equal number of random stimuli were sampled from the object-, intermediate- and environment-scale range. Object-scale stimuli were based on spline spheres with diameters subtending 4 - 22 degrees. Intermediate-scale stimuli subtended 22 - 80 degrees and the environment-scale range subtended 80 - 142 degrees. In order to select stimulus scale *within* each scale-range, uniform sampling was performed. An alternative method for random selection of stimulus scale would have been to uniformly sample from the *entire* 4 - 142 degree range without first subdividing intro specific intervals. In the latter approach it was observed that only a small proportion of stimuli visually appeared object-like relative to large-scale stimuli. Therefore, in order to avoid this under-sampling of object-like stimuli, the scale-range was divided into specific intervals (unequal: Objects [4 - 22 degrees] and environ-

ments [80 - 142 degrees]) and then equal numbers of stimuli were generated from each of them using uniform sampling. The choice of 80 degrees as the transition from intermediate- to environment-scale stimuli was based on the fact that at that scale stimuli began to extend well beyond the boundaries of the display screen, resembling unbounded landscapes. As will be shown in the results (3.6), responses of neurons in the TEd channel of IT greatly increased near this scale transition. The choice of 22 degrees for the transition from object- to intermediate-scale stimuli was based on the upper limit of stimulus sizes used in previous studies looking at object vision.

2.7   Morphing transformations

Descendant stimuli in the adaptive algorithm were created by morphing the shape, position, orientation, and/or scale of the ancestor. Each descendant was produced based on the application of 1-3 morphing operations (here referred to as morphs) to the ancestor. The number and magnitude of the morphs were probabilistic functions of response rate, to produce more alteration of low response ancestors and less alteration of high response ancestors.

Figure 2.4 demonstrates some of the morphs applied to an example environment-scale ancestor shape. The following enumerates the morphs and their probability of application:

- Scaling: 35%

- Rotation: 10% (equal probability for x-, y-, and z-axis rotation)

- Latitudinal ring modification:

  - Displacement: 5%

  - Change in elliptical aspect ratio: 5%

  - Change in curvature: 5%

- Surface feature modification:

  - Changes to height function of surface distortion: 5%

  - Changes to path shape of a surface distortion: 5%

  - Changes to position of a surface distortion: 2.5%

  - Removal of a surface distortion: 2.5%

- Translation: 10%

- Other morphs:

  - depth change: 5% (Closer fixation depths were prohibited when they would cause the stimulus surface to intersect the surface of the projection plane)

  - Longitudinal scaling of the spline grid: 5%

  - Global longitudinal curvature: 5%

The probabilities of specific morphing transformations were heavily weighted toward scaling in order to ensure greater sampling across the object to environment scale continuum.

2.7.1 Probabilistic selection of morphing operation magnitude

As mentioned in the previous section, the chosen magnitude of a given morph to be applied to a selected ancestor was a probabilistic function of the response rate evoked by that ancestor stimulus. The goal was to bias probabilities to generate smaller sized morphs for already high response stimuli and larger morphs for low response stimuli. This was implemented by defining a separate morph-size probability distribution for each of the 5 response percentile groups (described in section 2.6.1). Therefore, depending on the percentile group of the ancestor, one of 5 probability distributions was randomly sampled using an inverse transform method to generate a value for the size of a morph to be applied to the ancestor.

The probability densities for groups 3, 1 and 5 are shown in figure 2.5. Note that these probability densities were defined on a bounded range of morph size values normalized to a range of -1 to 1. The sign indicates directionality (e.g whether the change is a decrease or increase) and the densities are symmetrical about zero. As shown in figure 2.5(a), for ancestors that fell in group 3 (medium response range), the morph-

Ancestor stimulus

Scaling | Rotation | Latitudinal ring displacement | Surface feature modification | Translation

Morphed descendants
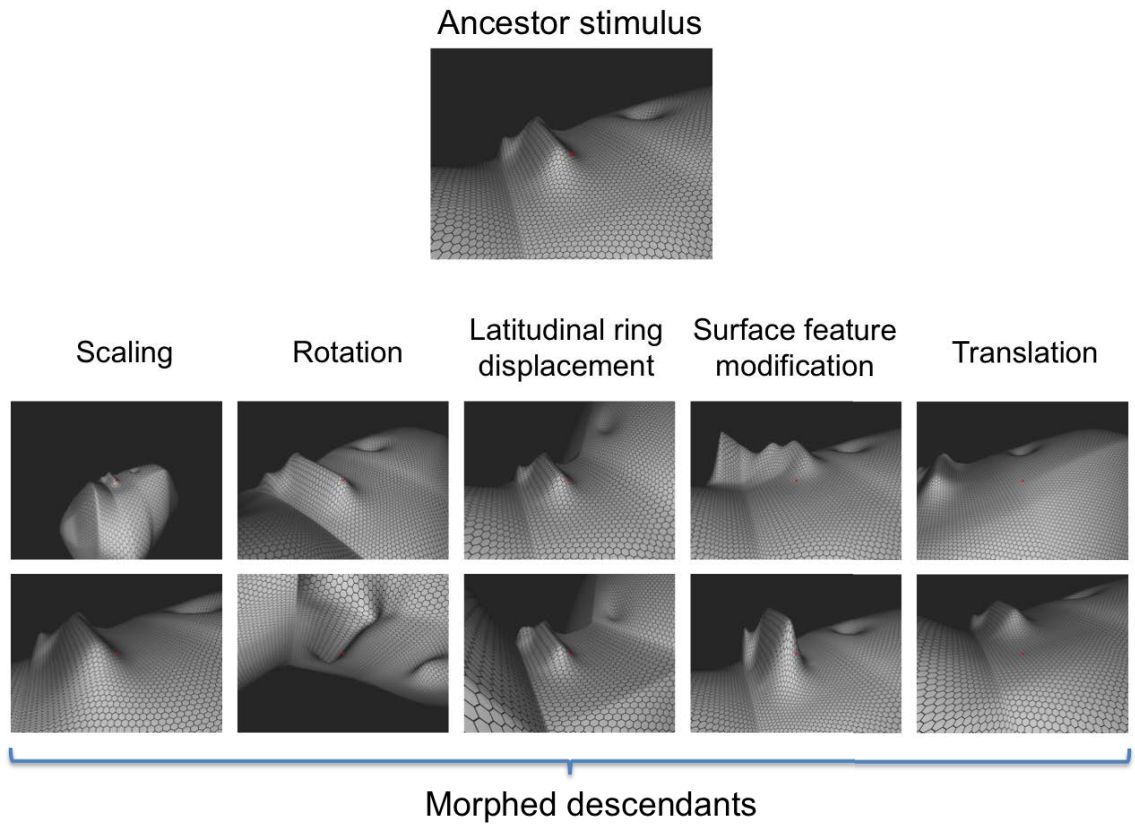
Figure 2.4: Shape morphing examples. An example ancestor stimulus is shown at the top. The bottom stimuli are example descendants generated under different morphing transformations of the ancestor. For each transformation type 2 example descendants arranged vertically are displayed.

size probability distribution was uniform and therefore all size ranges were equally probable. For group 1 (lowest response group) stimuli the morph-size probability distribution was defined such that the largest change had a density value that was 5 times larger than the smallest change. This probability density is shown in figure 2.5(b). Opposite to group1, group 5 (high response stimuli) ancestors were morphed based on a probability distribution that favored smaller changes (figure 2.5(c)). Distributions for group 2 and 4 ancestor stimuli were also biased (to a lesser extent than group 1 and 5) to favor smaller and larger morph changes respectively.
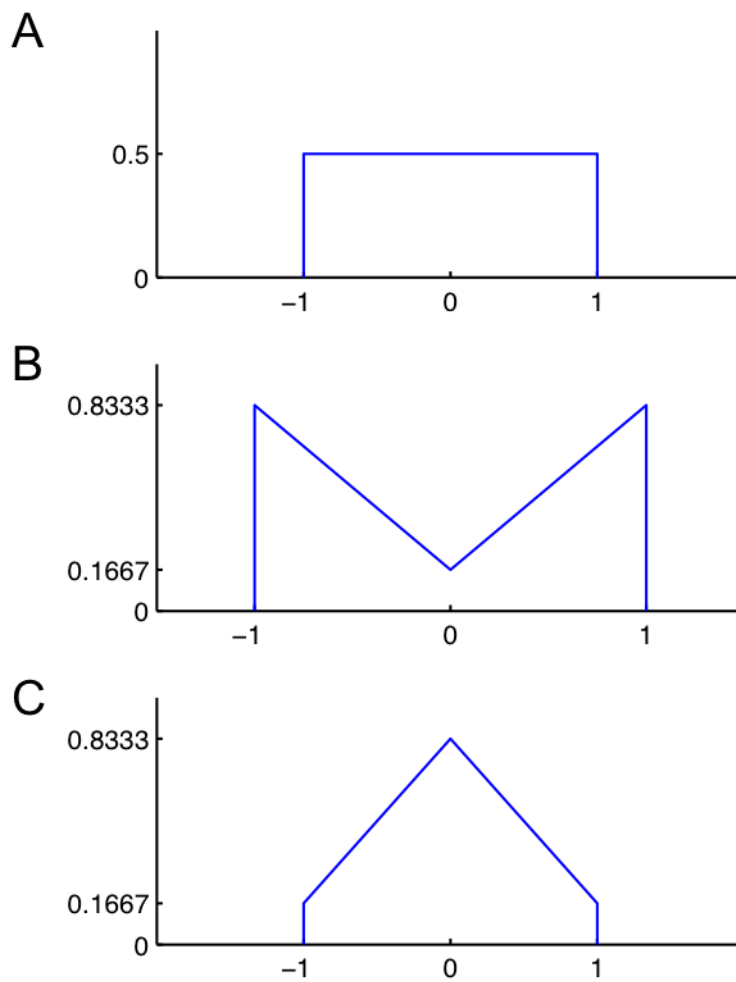
Figure 2.5: Probability distributions for random selection of morph size. X-axis indicates morph size, normalized to -1 to 1 range. (A) Density used for response group 3 ancestors. (B) Density for response group 1 ancestors. (C) Density for response group 5 ancestors.

Chapter 3

Shape scale tuning in IT cortex

The adaptive sampling algorithm was applied to 141 single-units (76 from TEd and 65 from STSv) across 2 monkeys (M1 = 108 units, M2 = 46 units). For each neuron two lineages of the algorithm progressed simultaneously to test for convergence. Example neuron data from TEd and STSv will be presented along with analyses that are subsequently applied to the population data. The main focus of the analysis in this chapter is on stimulus scale sensitivity. In particular results will be presented that indicate biases in scale tuning between the STSv and TEd processing channels. Results were combined across the 2 monkeys since the observed biases in scale-tuning preference were consistent between them.
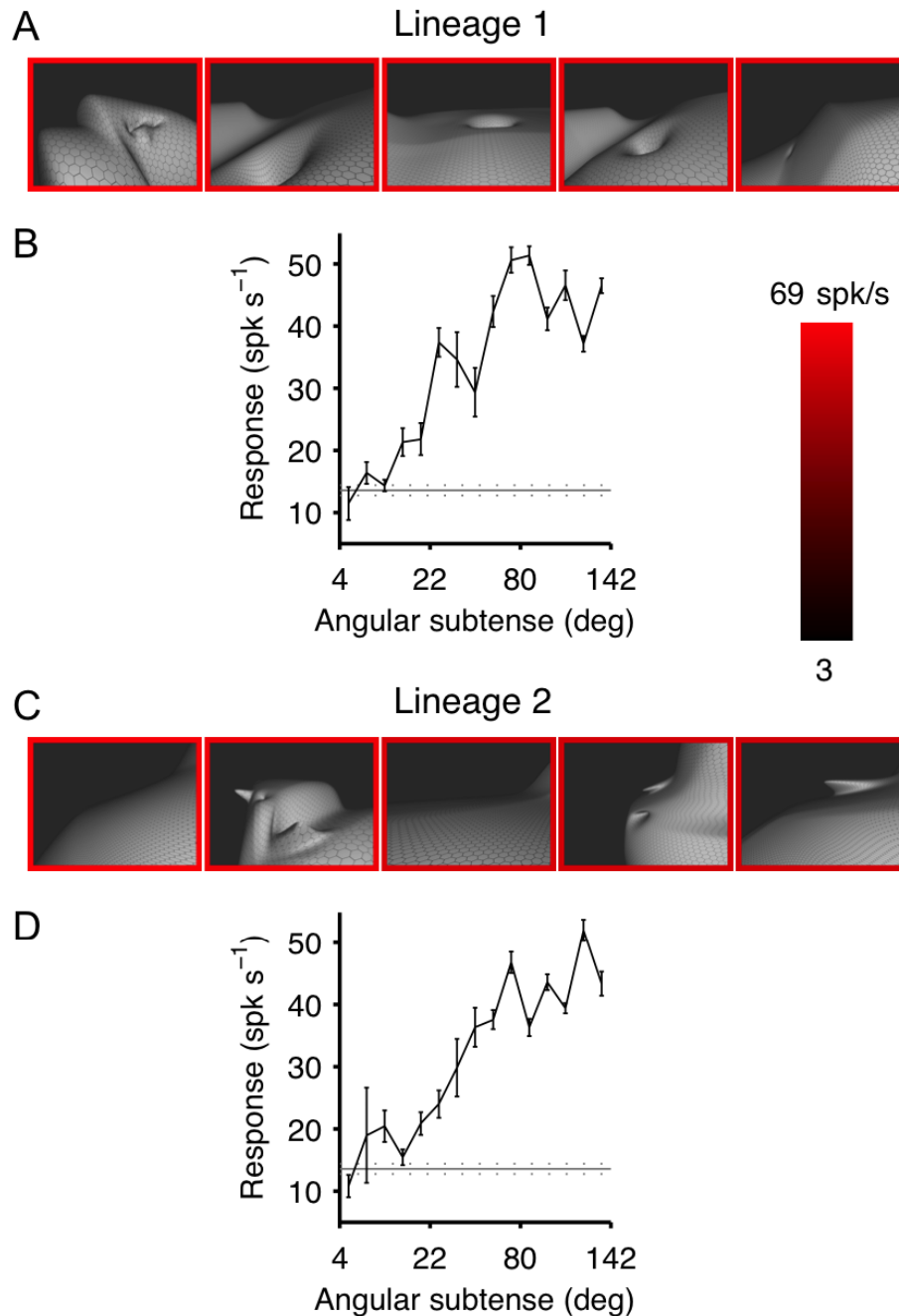
Figure 3.1: Example TEd neuron. (A and C) Top five stimuli that evoked the highest responses per lineage (vertically arranged) in the adaptive sampling algorithm. The color of the frame around each stimulus indicates average response (750 ms presentation, 5 repetitions) of the cell according to the color bar. (B and D) Average response (+/- SEM) of the example TEd neuron as a function of stimulus scale in each lineage. Baseline activity (gray line, +/- SEM) was averaged across randomly scheduled null stimulus presentation periods during the fixation task.

3.1   TEd example neuron

Figures 3.1(a,c) show the top five stimuli that evoked the highest responses in each lineage for an example neuron in TEd. For this neuron 10 generations of the adaptive sampling algorithm was performed providing responses to 800 visual stimuli (400 per lineage). The figure shows that for both lineages, high response stimuli consist of environment-scale shapes. Figures 3.1(b,d) show scale-tuning plots for each lineage. The plots were created by binning the scale range into 15 bins and calculating the average response to stimuli falling into each bin. In both cases responses to environment-scale (80 deg - 142 deg) shapes were near 50 Hz and fell to near baseline activity as scale decreased to object-scale levels (4 deg - 22 deg). The scale-tuning plots were strongly correlated ($r = 0.83$, $p < 0.0002$) indicating a convergence of the adaptive algorithm to shapes of similar scale under different initial conditions. In order to statistically test for scale preference, a Wilcoxon rank-sum test was performed between the top 10 environment-scale and top 10 object-scale stimuli. The results confirmed the strong bias for environment-scale stimuli over object-scale stimuli ($p < 0.0001$).

3.2   STSv example neuron

The results for an example neuron in STSv are shown in figure 3.2. Six generations of the adaptive algorithm involving 480 stimuli were performed. Here, in contrast to
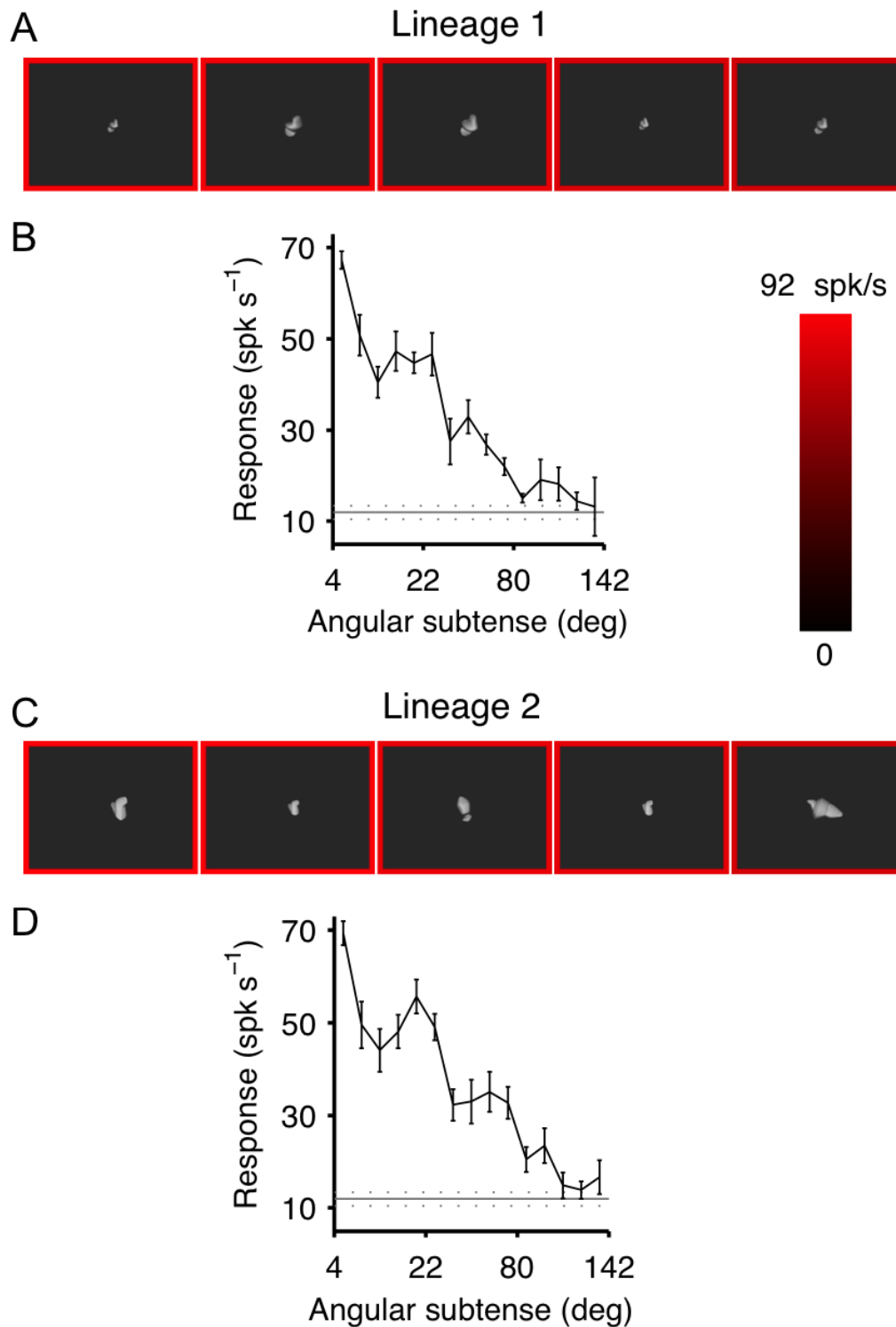
Figure 3.2: Example STSv neuron. (A and C) The five highest-response stimuli from each lineage (columns) are shown. Details as in 3.1(a,c). (B and D) Lineage specific scale response function for the example STSv neuron. Details as in 3.1(b,d)

the TEd example, high response stimuli are object-scale shapes (<22 degrees). This is consistent with previous studies of object coding in STSv. Scale-tuning plots for each lineage show that responses are high in the object-scale range and decrease to baseline levels for environment-scale stimuli. The scale-tuning pattern between the lineages was highly correlated (r = 0.96, p < 0.0001) confirming a consistent evolution of the algorithm under different initial conditions. The bias towards object-scale over environment-scale stimuli was highly significant (p < 0.0001, rank-sum test).

3.3 Differential selectivity for shape scale between STSv and TEd

Analysis of results at the population level confirmed differential responses to object-scale and environment-scale stimuli between STSv and TEd. Figure 3.3 depicts scale-tuning for each neuron in the population. The vertical axis indicates shape scale, increasing in magnitude from bottom to top as demonstrated by the example stimuli on the side. Scale-tuning for each neuron is displayed as a vertical colored strip where positions along this strip, that is along the scale axis, are assigned a color and brightness. The color is a redundant cue for scale, from red to green, indicating progression from object-scale to environment-scale stimuli. The brightness of the color indicates the average normalized response strength of the neuron to stimuli at that particular scale, where brighter colors indicate higher response strengths. There-
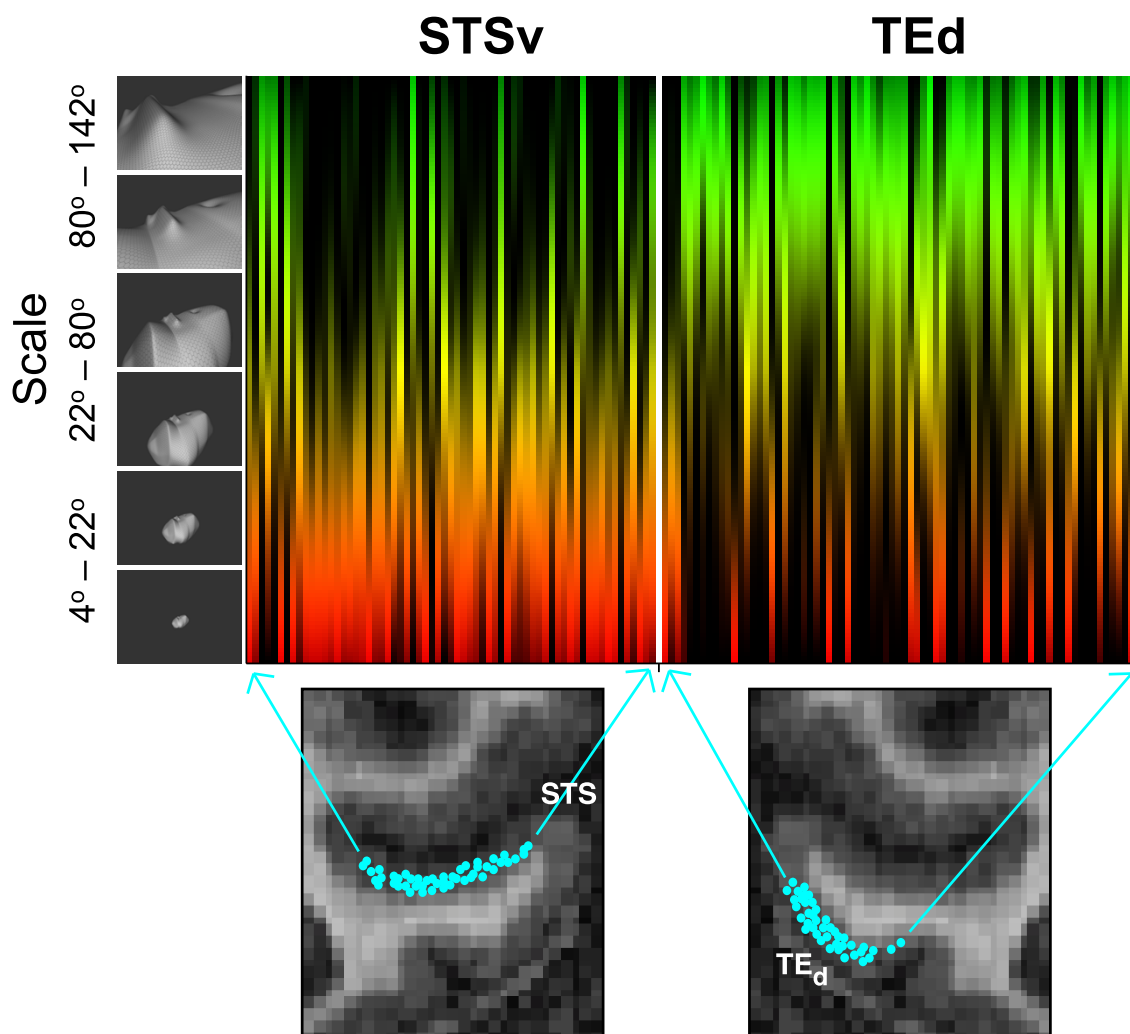
Figure 3.3: Population scale tuning results. Response profile of each neuron (n=141) in the population as a function of shape scale (vertical axis). Each column shows the normalized scale-tuning profile of a single neuron as a color strip. Positions along the strip have been assigned a hue and intensity. The hue component of the color (red to green gradient) is a redundant cue of the y-axis indicating shape scale. Red indicates small objects and green indicates large environments. The intensity component of the color reflects the average response strength of the neuron to all stimuli (in upper half of neuron's response range) at that particular shape scale. Data from the two lineages have been combined in this plot (refer to figure 3.5 for lineage comparison). The position of a given column along the horizontal axis indicates the recording location of the neuron around the circumference of IT as indicated by the two coronal MRI images. Note recording positions have been projected onto a single coronal slice. Refer to figure 3.7 for detailed postero-anterior recording positions.

fore bright green indicates high average responses to environment-scale stimuli and bright red indicates high average responses to object-scale stimuli. The ordering of these colored strips along the horizontal axis is based on the location of the recorded neuron as shown by the coronal MRI images. Strips to the left of the white vertical gap identify neurons recorded in STSv, arranged in a medial to lateral direction from left to right. Strips to the right of the white line belong to neurons recorded in TEd, arranged in a lateral to medial direction from left to right. This figure shows a clear segregation in shape-scale preference between STSv and TEd consistent with the example cell results. High response object-scale stimuli predominate in STSv whereas high responses to environment-scale stimuli are more prevalent in TEd.

3.3.1   Shape scale bias: statistical significance

In order to confirm the statistical significance of this trend, figure 3.4(b) shows the results of rank-sum tests for all the neurons in the population. As explained for the example neurons, rank-sum tests were performed between the top 10 environment-scale and top 10 object-scale stimuli for every neuron. The value of this test statistic is being plotted against the y-axis. The circular data-points identify each neuron in the population and the arrangement along the x-axis exactly matches the anatomical ordering in the scale-tuning plot (3.4(a)) replicated here from the previous figure 3.3 for clarity. The horizontal dotted lines in the plot 3.4(b) indicate the lower and upper critical values of this test at a significance value of 0.05 (two-tailed). Green
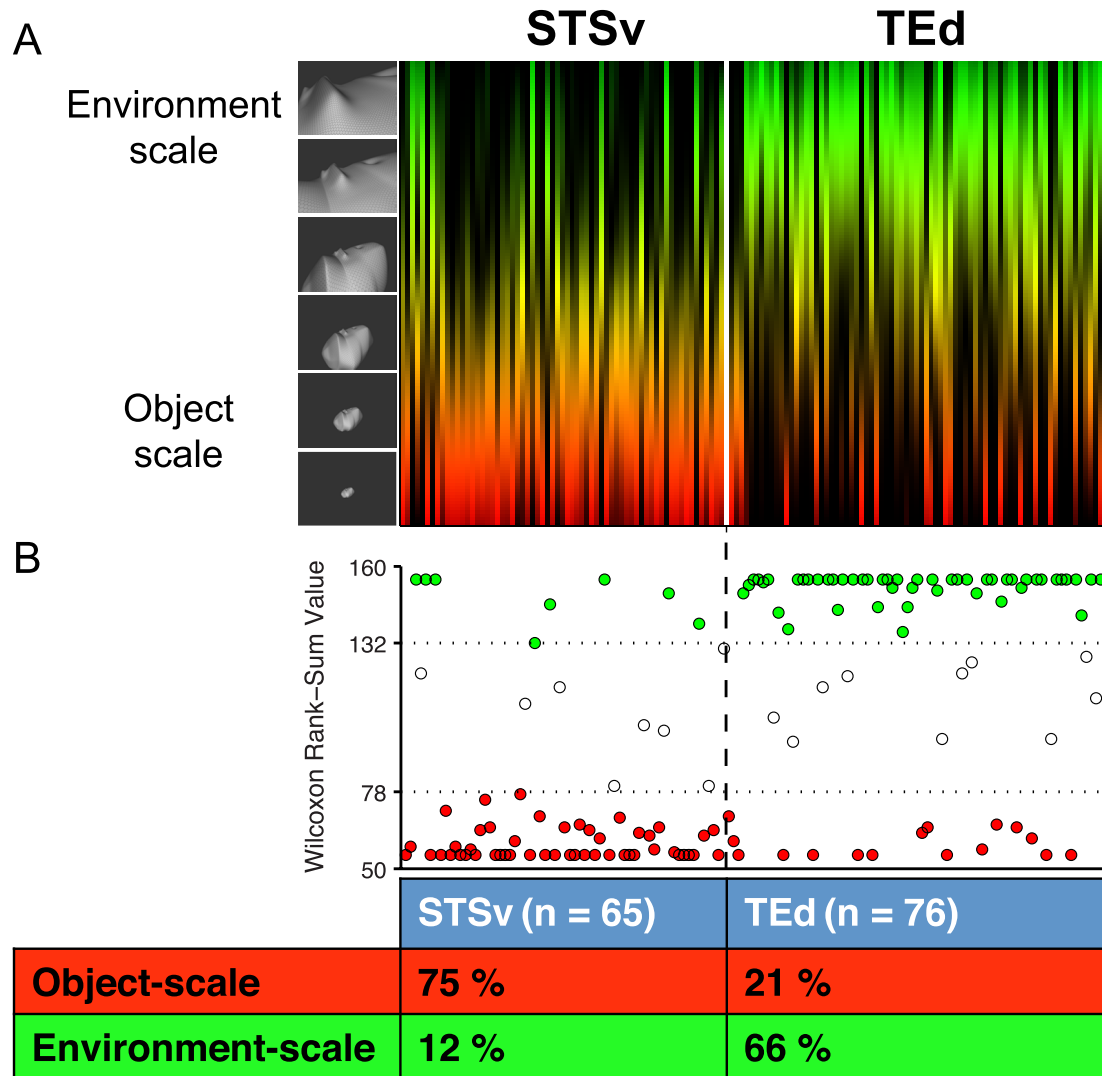
Figure 3.4: Statistical significance of shape-scale selectivity. (A) Population scale tuning plot replicated from figure 3.3 for comparison purposes. (B) Plot displays the results of a Wilcoxon rank-sum test for each cell to determine its selectivity for object- versus environment-scale shapes. The rank-sum test was performed between the top 10 environment-scale (angular subtense 80 - 142 degrees) stimuli that evoked the highest responses and the top 10 object-scale (angular subtense 4 - 22 degrees) stimuli evoking the highest responses. Each circle in the plot indicates the rank-sum test value (y-axis) for one cell. The ordering of the cells along the horizontal axis matches their order in Panel A and the vertical dashed line marks the division between STSv cells on the left and TEd cells on the right. The two horizontal dotted lines indicate the critical-values for the Wilcoxon rank-sum test using 2 samples of size 10 at a two-tailed significance value of 0.05. Rank-sum values above the the

upper critical value (132) indicate significant preference for environment-scale shapes. Points falling in this range have been filled green to match the color scheme in panel A. Rank-sum values below the lower critical-value (78) indicate significant preference for object-scale shapes and points in this region have been filled red. Data points falling between the two threshold show no statistically significant preference for either category (points are unfilled).

data points identify neurons with significant preference for environment-scale stimuli (above the upper critical value), red data points identify neurons with significant preference for object-scale stimuli (below the lower critical value) and unfilled data points are neurons that did not a show statistically significant preference for either scale (intermediate rank-sum values). The majority of STSv neurons (49/65; 75%) were significantly more responsive to object stimuli. Only 8 STSv neurons were significantly more responsive to environmental stimuli. In contrast, most TEd neurons (50/76; 66%) were significantly more responsive to environmental stimuli and only 16 TEd neurons were significantly more responsive to object stimuli. A chi-squared test of independence between recording channel (STSv or TEd) and scale-tuning preference (object, environment or non-significant) was performed. The result (chi-squared (2, N = 141) = 46.8, p < 0.0001) rejects the null hypothesis that recording location and stimulus scale preference are independent.

3.3.2   Shape scale bias: algorithm convergence

It is critical to test whether the adaptive sampling algorithm will under different initial conditions produce consistent scale-tuning patterns. As explained earlier, 2 lineages of the adaptive algorithm that were initialized with completely different random shapes were used and evolved independently. Correlation analysis was performed between the lineage specific scale-tuning functions of every neuron. The distribution of these Pearson correlation values is shown in figure 3.5 for the STSv and TEd

populations separately. As shown the correlation values were strongly positive and significant (filled bars) for both populations. This shows that the observed biases in shape-scale preference are reproducible under different initial conditions in the sampling algorithm.

3.3.3 Shape scale bias: non-linear tuning

The scale-tuning functions of the example neurons shown in figures 3.1 and 3.2 appears to be mostly linear. This, in fact, was atypical as most neurons had fairly steep sigmoidal responses as a function of scale. Figure 3.6 shows the average population response function for STSv and TEd neurons. The response function for TEd is almost step-like with an inflection point at about 80 degrees. This is roughly the scale at which stimuli began to exceed the boundaries of the display system (77 degrees), forming unbounded environment-like shapes. On the other hand, the average response function of STSv neurons reached its minimum at this scale. These results show that scale-tuning functions were on average strongly non-linear and therefore at the population level can provide categorical (object versus environment) information about the observed stimulus.
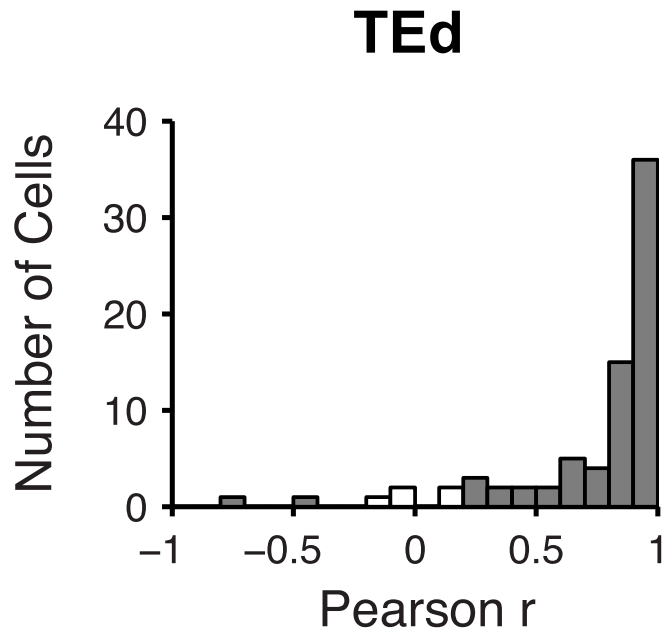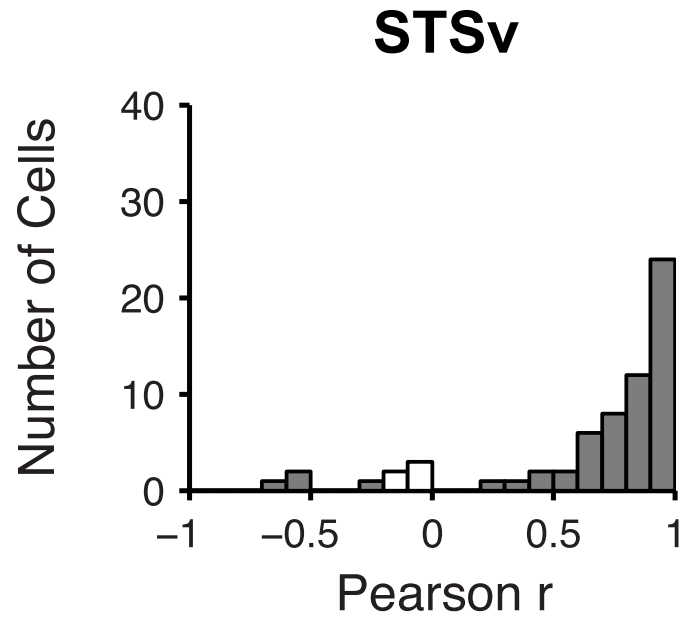
Figure 3.5: Convergence of adaptive algorithm. Bar plots show the distribution of Pearson correlations between the lineage specific scale-tuning profiles of each cell in the STSv population (top plot) and TEd population (bottom plot). Bars that are filled gray contain cells with statistically significant correlations (p < 0.01).
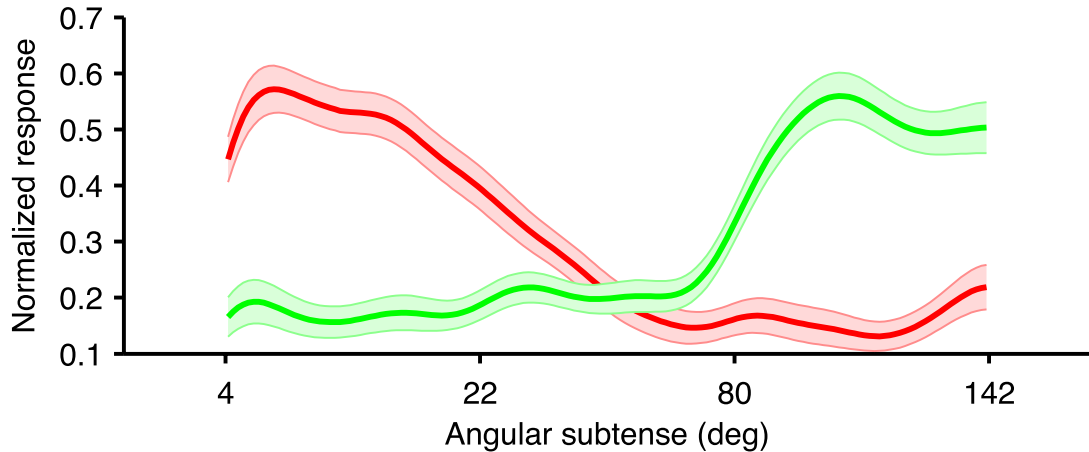
Figure 3.6: Average normalized population response. Average normalized response levels in STSv (red) and TEd (green) as a function of shape scale. The shaded regions represente the 95% confidence interval of average response values.

3.4 Anatomical trends in scale tuning within STSv and TEd

Results were analyzed for trends within the STSv and TEd regions. Figure 3.7 shows the stereotaxic recording location of neurons in STSv and TEd in the horizontal plane. Each point corresponds to one neuron and is color-coded according to its shape-scale preference based on the previously described rank-sum test. Correlation analyses were performed between recording position (either posterior-anterior or medio-lateral) and the rank-sum test statistic value in both regions. Note large rank-sum test values indicate stronger preference for environment-scale stimuli and smaller values indicate stronger preference for object-scale stimuli. There were no

42

significant posterior-anterior (TEd: r = -0.12, p = 0.31; STSv: r = -0.04, p = 0.75) or medio-lateral trends (TEd: r = -0.075, p = 0.52; STSv: r = -0.10, p =0.43) in scale-tuning in either STS or TEd. This figure also highlights the fact that recording mostly spared the lip of the superior temporal sulcus where the anterior-lateral (AL) face patch has been discovered (Freiwald and Tsao, 2010). In figure 3.7 the lip of the sulcus is approximately delineated by the area between the solid and dashed black lines in each plot. The AL face patch is located at around +12mm and, as shown, at most two neurons were sampled from this region. Both of these neurons were selective for object-scale stimuli, which is consistent with the description of this region as a face patch. Therefore, recording regions in STSv and TEd did not overlap the AL face patch for most neurons.
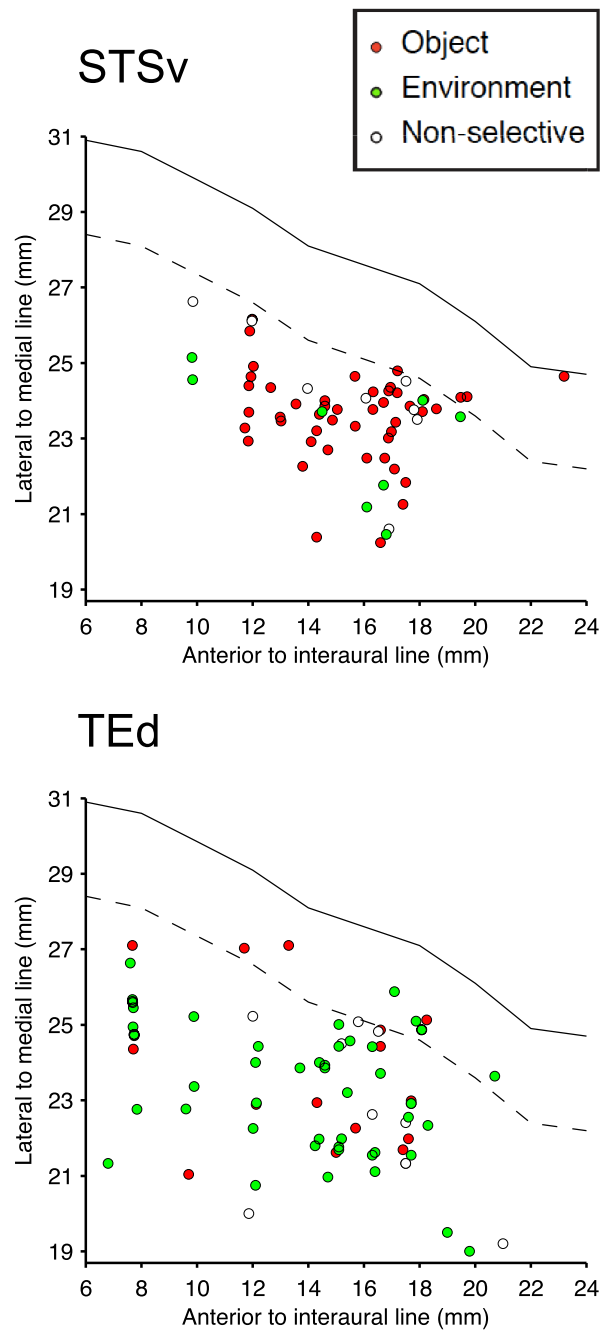
Figure 3.7: Stereotaxic recording positions. Plots show the stereotaxic recording positions of cells from STSv (top) and TEd (bottom). X-axis indicates the postero-anterior direction relative to the inter-aural line. Y-axis indicates the medio-lateral position relative to the midline. In each panel, the area between the dashed and solid lines corresponds to the lip of the STS. The color of the circles signifies the cell's preference for environment- versus object-scale shapes according to the rank-sum test (3.4).

3.5 Response selectivity based on sparseness analysis

Most neurons in both STSv and TEd had highly selective visual responses. Strong responses were evoked for only a small percentage of the total number of random stimuli tested indicating fairly high sparseness. Figures 3.8(a,b) show the response distribution across random stimuli for the example neurons of TEd and and STSv, respectively. As shown, for both neurons most random stimuli had responses in the low range and there were only a few stimuli that evoked particularly strong responses. In order to quantify sparseness, the following standard measure of response distribution density (RD) was used (Rolls and Tovee (1995); Vinje and Gallant (2000)):

$$RD = \frac{\left(\frac{1}{n}\sum_{i=1}^{n} R_i\right)^2}{\frac{1}{n}\sum_{i=1}^{n} R_i^2}$$

$$sparseness = \frac{1 - RD}{1 - \frac{1}{n}}$$

R_i is the base-line subtracted response of the neuron under consideration to its i'th stimulus (resulting negative responses are clipped to zero). The value of RD can achieve a minimum of $\frac{1}{n}$ and a maximum of 1 where n is the total number of visual stimuli used to probe the neuron. Sparseness is defined as $1 - RD$ and is normalized by $1 - \frac{1}{n}$ to force a maximum value of 1 (when RD is at its minimum). A sparseness value of 1 indicates a 'grandmother' cell, which has a non-zero response to only one of

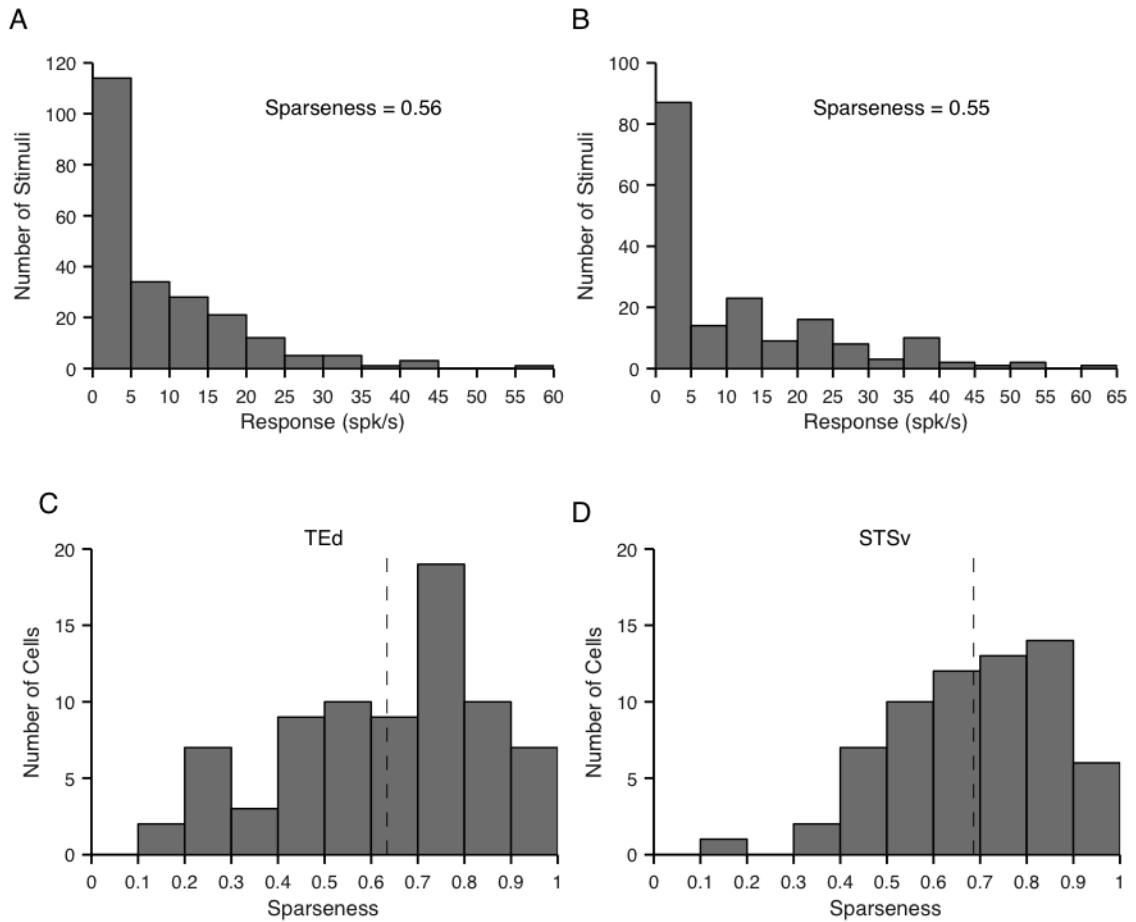Figure 3.8: Sparseness analysis to measure selectivity of responses in TEd and STSv. (A) Response distribution for the figure 3.1 example neuron from TEd. (B) Response distribution for the figure 3.2 example neuron from STSv. (C) Distribution of sparseness values for 76 TEd neurons. Average sparseness was 0.63 (vertical dashed line). (D) Distribution of sparseness value for 65 STSv neurons. Average sparseness was 0.69.

the stimuli tested. A sparseness value of 0 indicates a perfectly uniform response to all stimuli. The computed sparseness values for the response distribution of the TEd and STSv examples are 0.56 and 0.55, respectively. The distribution of sparseness values for neurons in the 2 regions is shown in figures 3.8(c,d). Average sparseness for TEd (0.63) and STSv (0.69) were not significantly different (t-test, p = 0.12).

3.6   Temporal response profiles

In the preceding analyses the response rates used for stimuli were calculated averages over a 750 ms presentation period. For a group of TEd neurons, figure 3.9 shows PSTHs computed for the top stimulus (5 repetitions) of each of the neurons. As shown the preferred stimuli are environment-scale and evoke high phasic responses. The time-course of responses was used to compute the time at which selectivity for objects and environments emerged.

Figure 3.10(a) shows the population averaged temporal response profiles for the STSv and TEd populations. Note, for the STSv population, only cells that were selective for object-scale stimuli (based on rank-sum test) were used in the computation of the response profiles. Analogously, for TEd, only cells that were selective for environment-scale stimuli were used. For each population (STSv or TEd), a separate response profile was computed for object-scale (red trace) and environment-scale (green trace) stimuli. For example, to compute the 'object' temporal response pro-

file for STSv, for each neuron in the STSv population, a response density was first computed for every object-scale stimulus (based on 5 repetitions and a Gaussian smoothing of 5 ms standard deviation). Next, these computed response densities were averaged to produce a temporal response profile of one neuron for its object-scale stimuli. These response profiles were then averaged across neurons in the STSv population to produce the red trace shown in figure 3.10(a) for STSv.

Figure 3.10(b) shows the distribution of scale-tuning onsets for the STSv and TEd populations. Scale-tuning onset of a given neuron refers to the time at which its stimulus-scale preference for either objects or environment stimuli became differentiated. This was computed by finding the first time bin (1 ms) at which the response profiles for the object and environment stimuli became significant different from one another (t-test , $p < 0.05$) and remained significantly different for all subsequent time bins in the stimulus presentation period. The distribution of these scale-tuning onsets for both STSv and TEd is fairly broad. However, the average scale-tuning onset in STSv of 113.6 +/- 6.4 ms is significantly smaller than the average scale-tuning onset in TEd of 132.5 +/- 6.2 ms ($t(97) = -2.13$, $p < 0.05$). The slower emergence of environment-scale selectivity could reflect the large visual region over which environmental stimuli must be integrated or the lower acuity of peripheral visual processing needed in environment shape perception. There were no significant trends in scale-tuning onset in either the anterior-posterior direction (STSv object-selective neurons: $r = -0.085$, $p = 0.57$; TEd environment-selective neurons: $r = -0.020$, $p = 0.89$) or

48

the medio-lateral direction (STSv object-selective neurons: r = 0.11, p = 0.46; TEd environment-selective neurons: r = -0.078, p = 0.59).

3.7   Chapter conclusion

The analysis presented in this chapter shows that there exist differential responses to shapes of different scales between STSv and TEd. Consistent with previous results STSv neurons were found to be selective for stimuli in the object-scale range (4 - 22 deg). The novel finding presented here is that effective stimuli for a large proportion of TEd neurons were large-scale shapes (80 - 142). Neurons were found to be highly selective as demonstrated by sparseness of their response distributions. Scale-tuning functions were found to be non-linear and in the case of TEd the average scale tuning function was sigmoidal with high steepness at the transition point between large bounded objects (visible borders) and large unbounded environment-like stimuli. This study provides the first quantitative analysis of 3D shape scale sensitivity encompassing the large domain of objects and environments as well as intermediate-scale shapes.
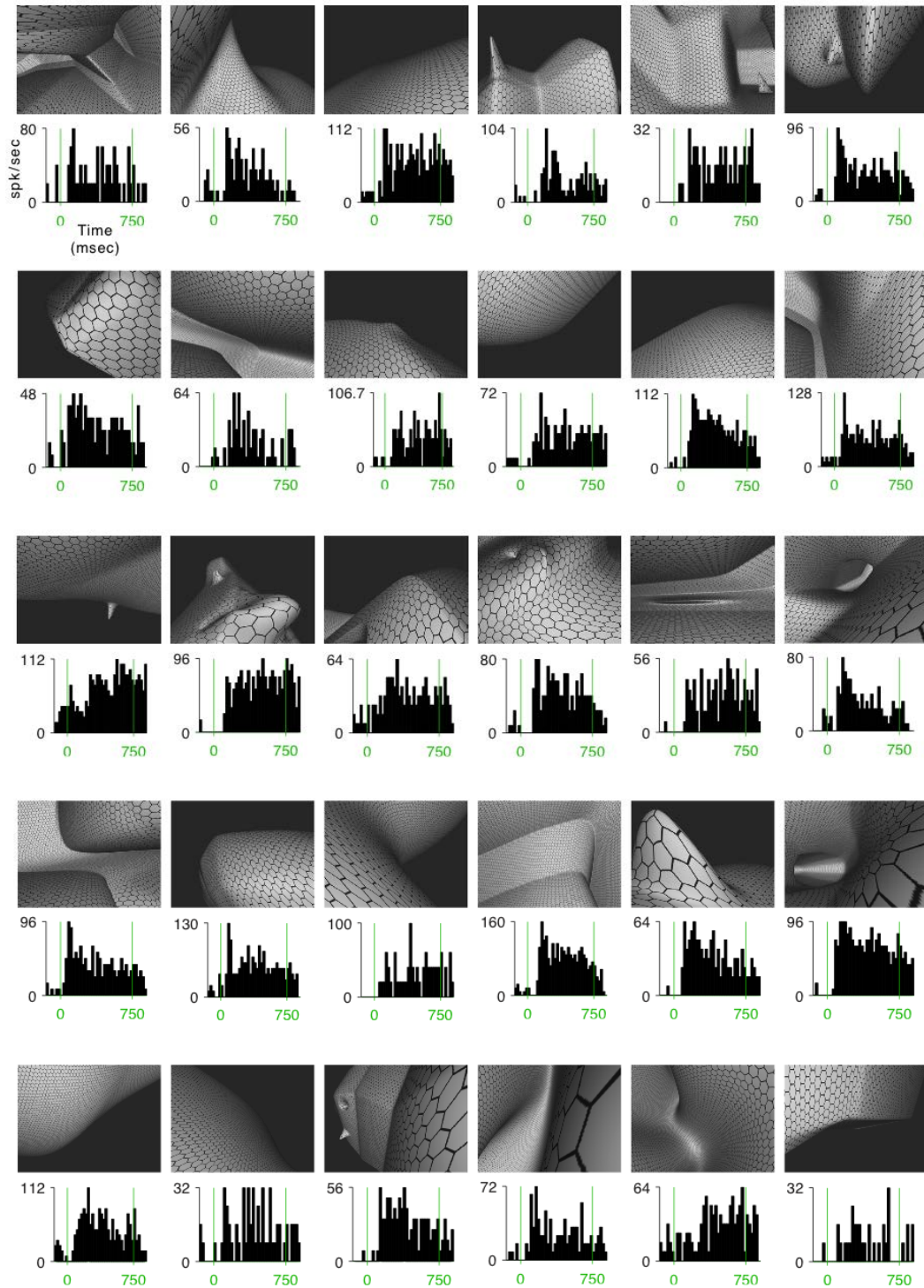
Figure 3.9: For each neuron the preferred stimulus along with the corresponding PSTH (below stimulus) is shown. Green vertical bars in the PSTH plots indicate stimulus offset and onset times.

50
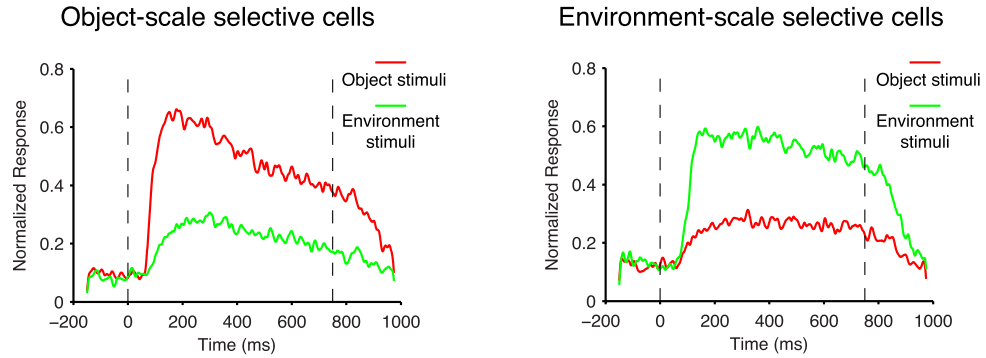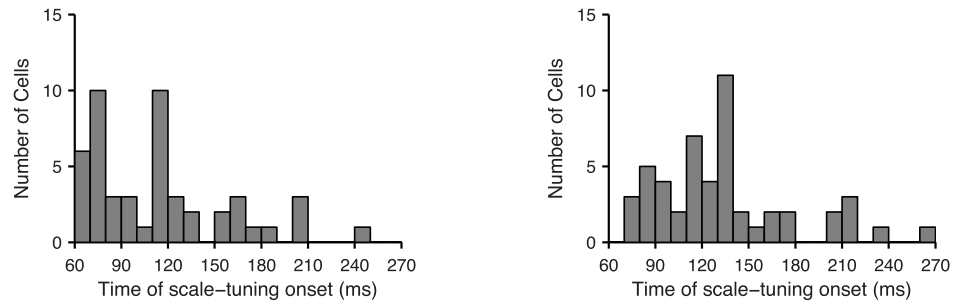
Figure 3.10: Temporal evolution of object selectivity in STSv and environment selectivity in TEd. (A) Average normalized temporal profiles of responses is STSv and TEd neurons, selective for objects and environments, respectively. Red trace in both plots indicates average response profile to object-scale stimuli, Green trace indicates average response profile environment-scale stimuli. Vertical lines indicate stimulus onset and offset. (B) Scale-tuning latency distributions for the STSv (left) and TEd populations (right). Scale-tuning latency for a given is the time at which resposnses profiles to objects and environment became significantly different and remained differentiaed for the rest of the stimulus presentation period.

Chapter 4

TEd selectivity for shape-in-depth

In this chapter, results from a series of auxiliary tests will be presented. These were controls designed to test whether various low-level stimulus factors such as spatial frequency of texture pattern, peripheral receptive field stimulation or lighting direction can account for the observed (see chapter 3) selective responses to environment-scale shapes in TEd. The tests were performed at the end of the main adaptive sampling experiment whenever recording time permitted. The stimuli used in these control tests were the optimal environment-scale shapes found from the adaptive sampling algorithm.

4.1   2D versus 3D

As explained in chapter 2, stimuli in the main test were rendered in 3D using stereo, texture and shading cues. Potentially, neurons in TEd may not be sensitive to shape-in-depth information and would be equally selective for large 2D shapes or large 2D texture patterns. This result would argue against the interpretation that TEd neurons are selective for environments. The large-scale stimuli shown in figure 2.2 resemble landscapes and interiors because they contain shape-in-depth information associated with these natural environments. In the case of STSv, results from a previous study (Yamane et al., 2008) have already shown that neurons selective for object-like shapes are very sensitive to 3D shape information. The study showed that selective responses for optimal object-scale stimuli were largely abolished when a 2D silhouette version of the original stimulus is displayed. A similar test was performed in this study, to see whether TEd neurons that show selectivity for environment-scale stimuli actually care about shape-in-depth information. Responses were obtained to 2D equivalent versions of the preferred environment-scale stimuli that lacked 3D depth cues.

Figure 4.1(a) shows the preferred environment-scale stimulus (labeled 3D) of an example TEd neuron, and next to it are five associated 2D conditions for which responses were obtained. These are a simple 2D silhouette of the original stimulus without texture, 2D versions where the texture pattern was flattened onto the projection plane at 2 different spatial frequencies, and 2D versions where the orientation

of the texture lines were randomized, again at 2 different spatial frequencies. Note that stereo and shading were removed for all these five 2D conditions. In this test, in addition to 2D versions of the preferred environment-scale stimulus, responses to 2D versions of a non-preferred environment-scale stimulus were also obtained. This non-preferred stimulus was chosen from the low-response range of the neuron's response distribution.

The goal was to quantify the degree to which stimulus selectivity for the preferred shape is affected by removal of 3D cues. To quantify stimulus selectivity, the following measure of response modulation strength (ModS) was computed for 3D and 2D conditions:

$$ModS = \frac{R_{prefStim} - R_{nonPrefStim}}{R_{max}}$$

In the 3D case modulation strength was computed as the response difference between the preferred and non-preferred environment-scale stimuli shown in 3D divided by the maximum responses of the neuron. In the 2D case, since there are 5 different 2D versions (silhouette, 2 fronto-parallel texture, 2 randomized texture), a separate modulation strength was computed for each of the 5 conditions. In each case the modulation strength was the response difference between the preferred and non-preferred environment-scale stimuli shown in the particular 2D condition under consideration divided by the maximum response of the neuron. The largest value out of the 5 measures was selected to represent 2D modulation strength in order to report the largest

amount of selectivity that can be explained under 2D conditions.

The plot in figure 4.1(b) compares 2D (y-axis) and 3D (x-axis) modulation strengths for 27 TEd neurons that according to the rank-sum (3.4) test were selective for environment-scale stimuli. The data-point in cyan identifies the example neuron that was tested with stimuli shown in figure 4.1(a). As shown, 3D modulation strengths are mostly clustered around the maximum value of 1 indicating high selectivity. The average 3D modulation strength was 0.86. Critically, the corresponding 2D modulation strength (average = -0.041) falls to near zero and in some cases becomes slightly negative indicating a reversal of the stimulus preference (paired t-test, $p < 0.0001$).

This results shows that when 3D cues conveying shape-in-depth information are removed, differential responses to environment scale stimuli are largely abolished. Therefore, the presence of large 2D shapes or textures do not account for the observed selectivity for environment-scale shapes. In addition, the results argue against a possible explanation of the results based on stimulation of peripheral parts of the visual field since the 2D shapes were just as large. Even though a large receptive field is necessary for sensitivity to large-scale shapes, these results indicate that peripheral stimulation in this large receptive field using 2D shapes is not sufficient to explain high responses to environmental stimuli. This point was further tested in section 4.2.
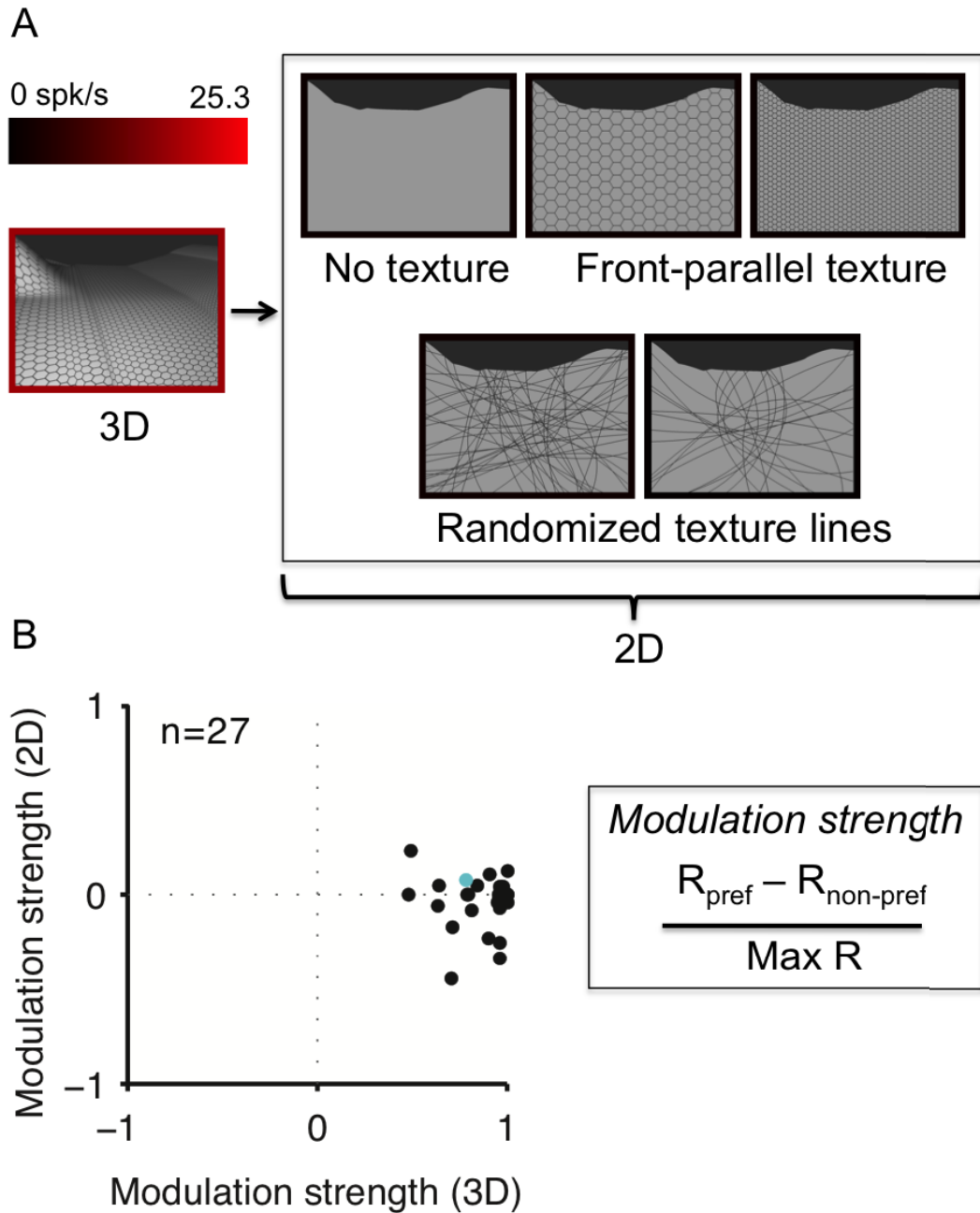
Figure 4.1: Sensitivity to 3D shape-in-depth vs. 2D shape. (A) For an example TEd neuron the preferred environment-scale stimulus is shown (labeled 3D) along with equivalent 2D versions consisting of stimuli with either fronto-parallel hexagonal texture, random line texture, or no texture (silhouettes). Also, unlike the original 3D stimulus, the 2D versions did not have shading or disparity cues either. Borders indicate response rate (see scale bar). (B) For 27 TEd cells with significant selectivity

for environmental stimuli, one high-response and one low-response environmental stimulus were selected from the adaptation experiment. Modulation strength is the response difference divided by the maximum. 3D modulation strength (x-axis) is based on the original stimuli with all depth cues. 2D modulation strength (y-axis) was based on one of the 5 different 2D conditions, whichever produced the highest modulation value. The cyan data point identifies the neuron with stimuli shown in panel A.

Figure 4.2 shows the modulation results for a group of neurons in STSv. In this case modulations were computed as a response difference between a preferred and non-preferred object-scale stimulus divided by the maximum response of the cell. As in figure 4.1(b) for TEd, removal of 3D cues largely abolishes differential responses indicated by the near zero values for 2D modulation strength (average 2D modulation = 0.095). The average 3D modulation strength was 0.87 (paired t-test, $p < 0.0001$). These results confirm previous findings (Yamane et al., 2008; Janssen et al., 2000) that STSv is sensitive to 3D shape-in-depth information and also shows comparable sensitivity to shape-in-depth observed for TEd neurons.

4.2 Peripheral receptive field test

Since environment-like stimuli were large-scale, covering more than 70 degrees of visual angle, they provided visual stimulation of the peripheral field which was absent for object-scale stimuli presented centrally. It may be the case that neurons in TEd have more peripheral receptive fields compared to STSv and therefore object-scale stimuli presented in the periphery might be just as effective in eliciting high responses compared to environment-scale stimuli. This possibility was tested in a control where optimized object-scale stimuli from the adaptive sampling algorithm were displayed at various peripheral locations. Specifically, responses were obtained at 9 positions on a 3 by 3 square grid centered at the central fixation point with 30
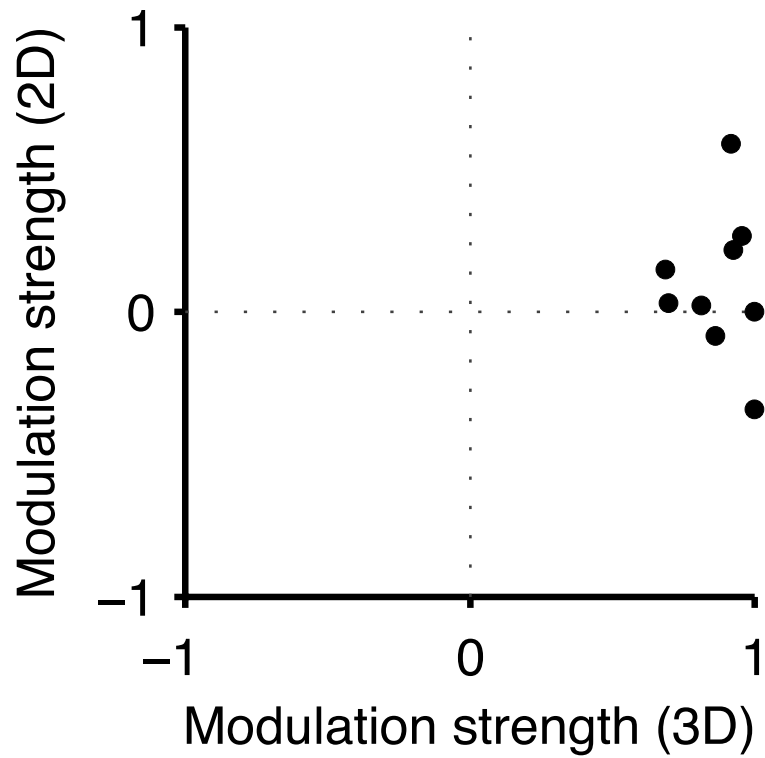
Figure 4.2: Sensitivity to 3D shape-in-depth vs. 2D shape for STSv neurons. As in figure 4.1 one high-response and one low-response object stimulus were selected from the adaptation experiment. Modulation strength is the response difference divided by the maximum. 3D modulation strength (x-axis) is based on the original stimuli with all depth cues. 2D modulation strength (y-axis) was based on stimuli with no disparity cues, no shading, and either fronto-parallel hexagonal texture, random line texture, or no texture (silhouettes), whichever produced the highest modulation value.

degree spacing in the horizontal and vertical direction. Figure 4.3 shows the results for 8 neurons in TEd by comparing responses to the preferred environment-scale stimulus and the highest response elicited by one of the peripheral object conditions. As shown, responses to the environment-scale stimulus are significantly larger than responses to the peripheral object stimulus for all neurons. Therefore, peripheral stimulation alone using optimized object stimuli does not fully account for the high responses to environment-scale shapes.

4.3 Comparison of 3D depth cues: texture, shading and stereo

Section 4.1 established that 3D cues are critical for TEd responses. The relative importance of the specific cues consisting of shading, stereo and texture was the focus of this auxiliary test. A full factorial test where each cue was a factor was performed on the preferred environment-scale stimulus of 22 TEd neurons. The level of each factor was binary indicating the presence or absence of that particular depth cue in the rendered stimulus. Figure 4.4 shows multiple scatter plots where each plot compares the responses to the environment-scale stimulus with all depth cues present (plotted against x-axis) versus one of the conditions of the factorial test (plotted against y-axis) consisting of no depth cues, individual depth cues and combinations of depth cues. To analyze these results a full model 3-way repeated measures ANOVA

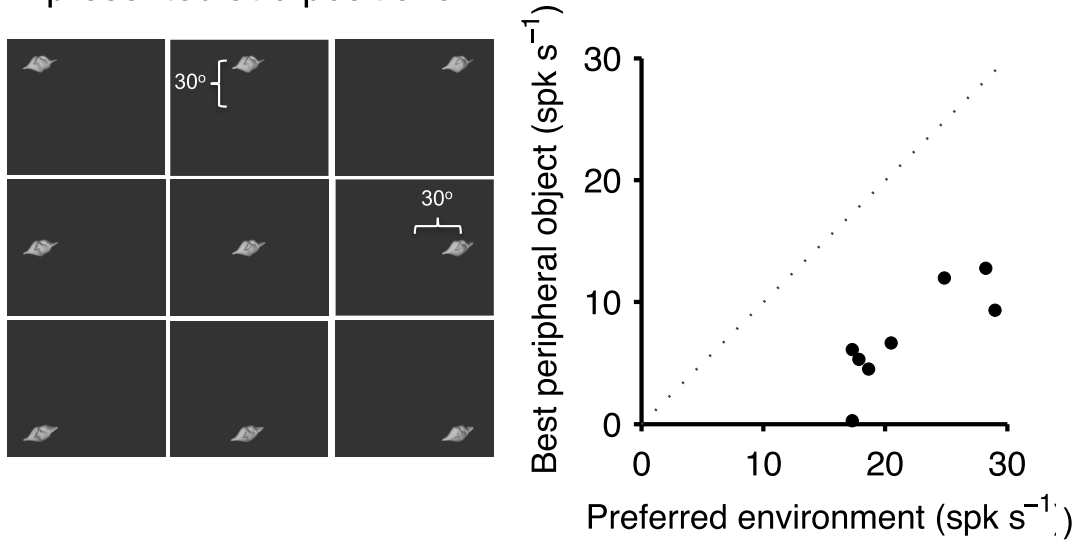**Best object-scale stimulus presented at 9 positions**

Figure 4.3: Peripheral object responses of TEd neurons. For 8 cells in TEd (selective for environmental stimuli) response to an optimal environment stimulus (horizontal axis) is compared to the maximum response to an optimal object stimulus tested at 9 positions in a square grid (centered at fixation with 30 x and y spacing) (vertical axis).

was performed. The results show a strong significant main effect for texture ($F(1,21)$ = 16.8; $p < 0.002$), a significant main effect for shading ($F(1,21)$ = 10.4; $p < 0.005$) and no significant main effect for stereo ($F(1,21)$ = 1,8; $p = 0.194$). Interaction terms consisting of combinations of depth cues did not show a significant effect. The finding that stereoscopic depth cues do not contribute to responses for environment-scale stimuli is consistent with that fact binocular disparity differences become negligible at far depths present in environments.

4.4   Texture spatial frequency test

A potential low-level stimulus property that might account for selectivity of TEd neurons for environment-scale stimuli might be tuning for a specific spatial frequency range contained in the texture pattern. A previous study (Rajimehr et al., 2011) argued that selectivity for natural images of scenes can be explained away based on sensitivity for high-spatial frequencies. In order to evaluate sensitivity to spatial frequency content, responses of 17 TEd neurons were obtained to their preferred environment-scale stimulus with texture patterns that were either halved or doubled in density. Figure 4.5 compares responses to the original environment-scale stimulus (plotted against y-axis) with versions containing texture patterns with halved or doubled densities (plotted against x-axis). The data points shown in cyan identify the neurons with the stimuli shown near the plot axes. As the scatter plots show
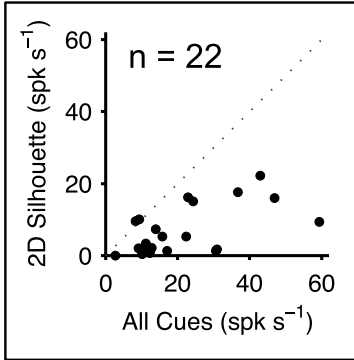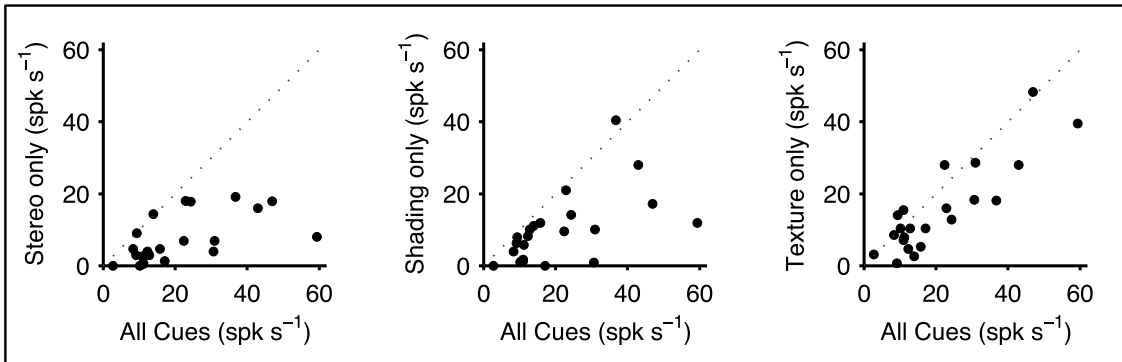
Figure 4.4: Dependency of TEd environment responses on 3D shape cues. For 22 TEd neurons, response to an optimal environment stimulus rendered with stereoscopic disparity, shading and texture (horizontal axis) compared with various (7) sub-combinations of those cues (vertical axes).

the majority of points are located nearer the diagonal indicating that changes in the texture spatial frequency do not affect responses. A one-way repeated measures ANOVA confirmed that there is no statistically significant effect of texture density ($F(2,32) = 0.16$; $p = 0.85$).
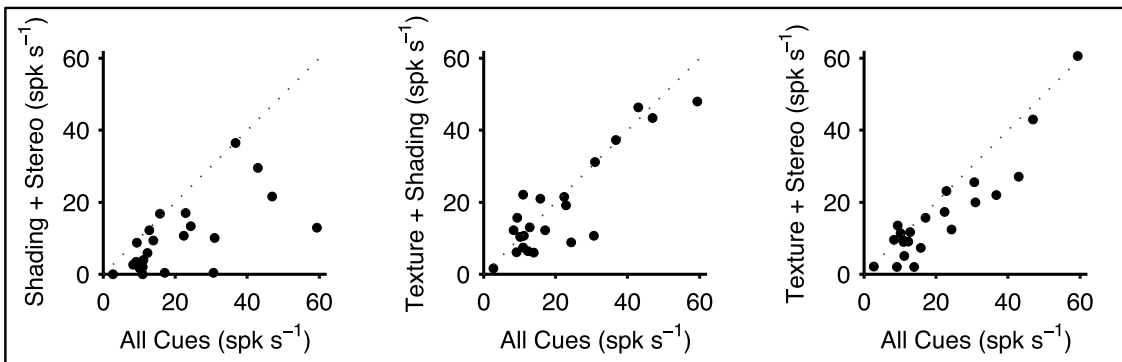
4.5   Effect of lighting direction

In the main test stimuli were rendered using a frontal lighting direction. Neurons in TEd might be sensitive to the specific contrast pattern in environment-scale stimuli produced by this particular lighting direction. A previous study (Yamane et al., 2008) demonstrated that IT neurons selective for object-scale stimuli are invariant to changes in lighting direction. The effect of lighting direction on responses to environment-scale stimuli was investigated in this control test for 22 TEd neurons. Four additional lighting directions consisting of top, bottom, left and right were tested on the preferred environment-scale stimulus of each neuron. The scatter plots in figure 4.6 show the results as comparisons between responses to the original frontal lighting direction (plotted against y-axis) and the other 4 lighting directions (plotted against x-axis). The responses appear invariant to changes in lighting direction, as confirmed statistically by a one-way repeated measures ANOVA that did not show a main effect for lighting ($F(4,84) = 0.64$, $p = 0.64$). As before the data points in cyan identify the neuron with the stimuli shown next to the axes.
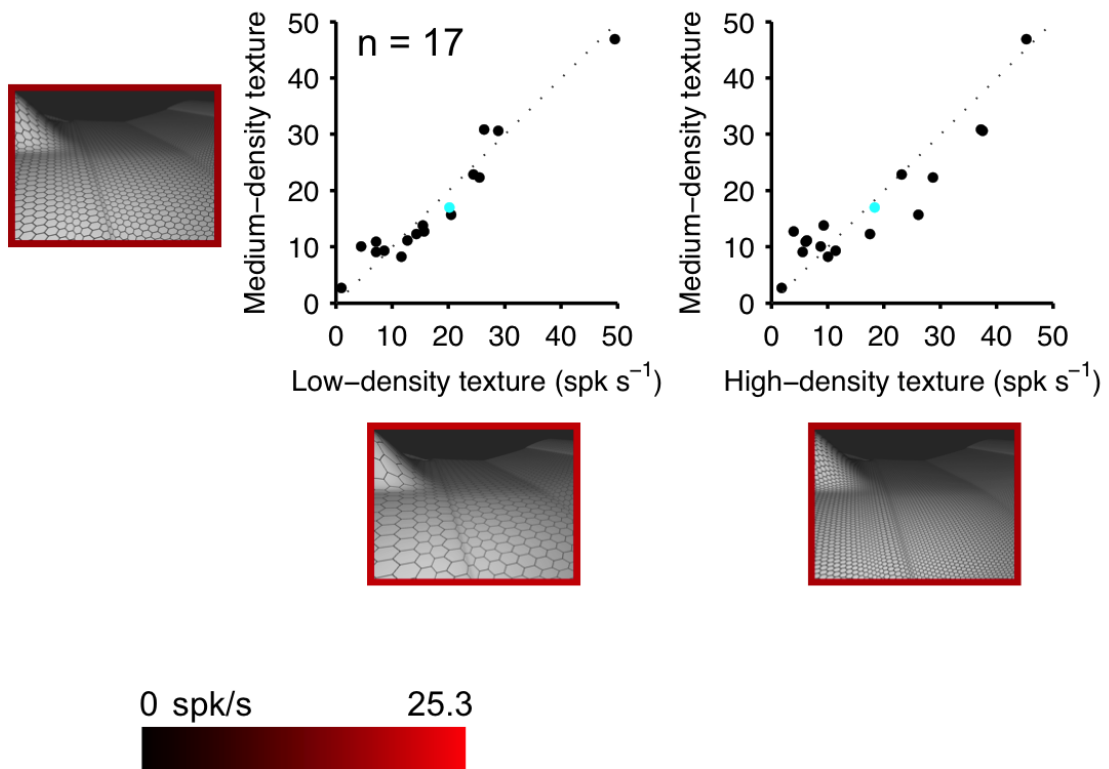
Figure 4.5: Dependency of TEd environment responses on texture spatial frequency. For 17 TEd cells, response (vertical axis) to the preferred environment-scale shape shown in the original texture density (medium) is compared to responses (horizontal axis) to the same shape shown in a decreased (halved) texture density (left plot) and an increased (doubled) texture density (right plot). The stimulus images next to the axes belong to the neuron shown in cyan.
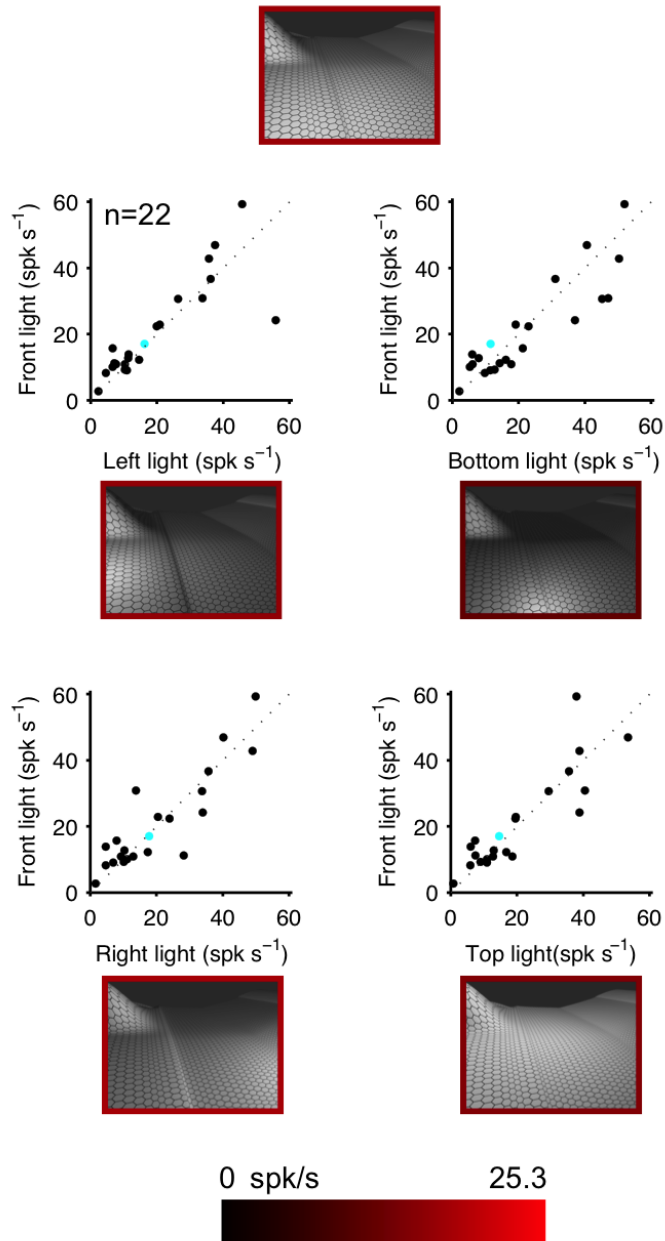
Figure 4.6: Consistency of response across lighting directions. Four scatter plots compare the responses of TEd cells to their preferred shape under the default frontal lighting condition (y-axis) to responses under lighting direction (x-axis) from the right, left, bottom and top. The stimulus images show the preferred shape of the example cell in cyan under the different lighting conditions.

4.6 Scale tuning test for single shapes

As described in section 2.7 scaling morphs were applied at a high probability in the adaptive algorithm and therefore stimuli of various shapes were extensively tested across the scale range as the algorithm evolved. However, it is conceivable that there may have been high response shapes that by chance only evolved within a limited scale range. For example, a high response object-scale shape that is selected to produce morphed descendants may by chance never receive a scaling morph large enough to produce descendants in the environment-scale range. These hypothetical descendants may actually evoke high responses comparable to the object-scale ancestor given that IT responses are known to show some size-invariance (Ito et al., 1995). Therefore, in this example the results may be biased in showing object-scale selectivity due to a lack of broader sampling in the scale range for the preferred shape.

In order to identify whether such biases occur 2 separate runs of the adaptive algorithm are performed and, as shown in figure 3.5, for the majority of cases (131/141) there were strong significant correlations between the lineage-specific scale tuning functions. The consistency in scale-tuning patterns between 2 different runs of the algorithm makes it less likely that the observed shape scale preference are the result of a high response stimulus that by chance only evolved within a limited scale range.

As an additional test to address this concern, for some neurons in both STSv and TEd responses to their most effective stimulus from the main test were obtained across a large portion of the scale range. Figure 4.7 shows the responses of an example

TEd neuron to a single preferred environment-scale shape at 6 different scales. The original stimulus from the main adaptive algorithm is the large-scale stimulus at the far-right. Responses are much stronger at large scales in the environment-scale range. For this neuron the rank-sum test described in section 3.3.1 had shown selectivity for environment scale stimuli consistent with the responses observed here.

Figure 4.8 shows results of this test for other neurons in TEd. In each plot the black line indicates responses to the optimal environment stimulus shape chosen from the main test shown at different scales. For each neuron, as scale decreases to the object-scale range responses to this optimal shape also decreases to near baseline level showing consistent scale-tuning compared to the results of the main test. For some of these neurons, responses to their most effective object stimulus shape (sub-optimal compared to the best environment) was also tested at different scales (shown in blue). In this case responses remained low even at larger-scales indicating that these neurons are selective for shape as well as scale.

Figure 4.9 shows the results of similar tests for some neurons in STSv. For these neurons, the responses (shown in black) to their optimal object-scale stimulus shape decreases as scale increased to environment-scale range. Some neurons were also tested with their most effective environment stimulus shapes (sub-optimal compared to the best object). In this case (shown in blue) responses remain low even at scales in the object-scale range indicating that there is selectivity for shape as well as scale.
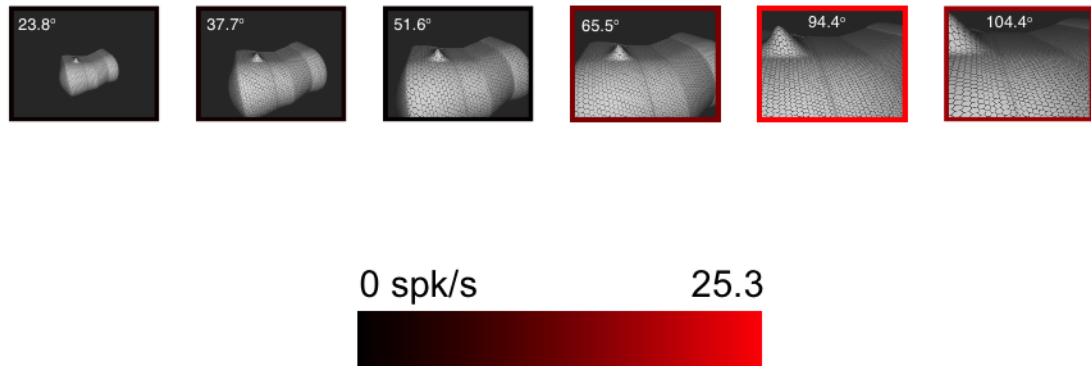
Figure 4.7: Scale-tuning test for an example TEd cell. The high-response environmental stimulus for an example TEd neuron was presented at 6 scales, ranging from object to environmental. The original stimulus from the adaptive experiment was the largest scale, at far right.
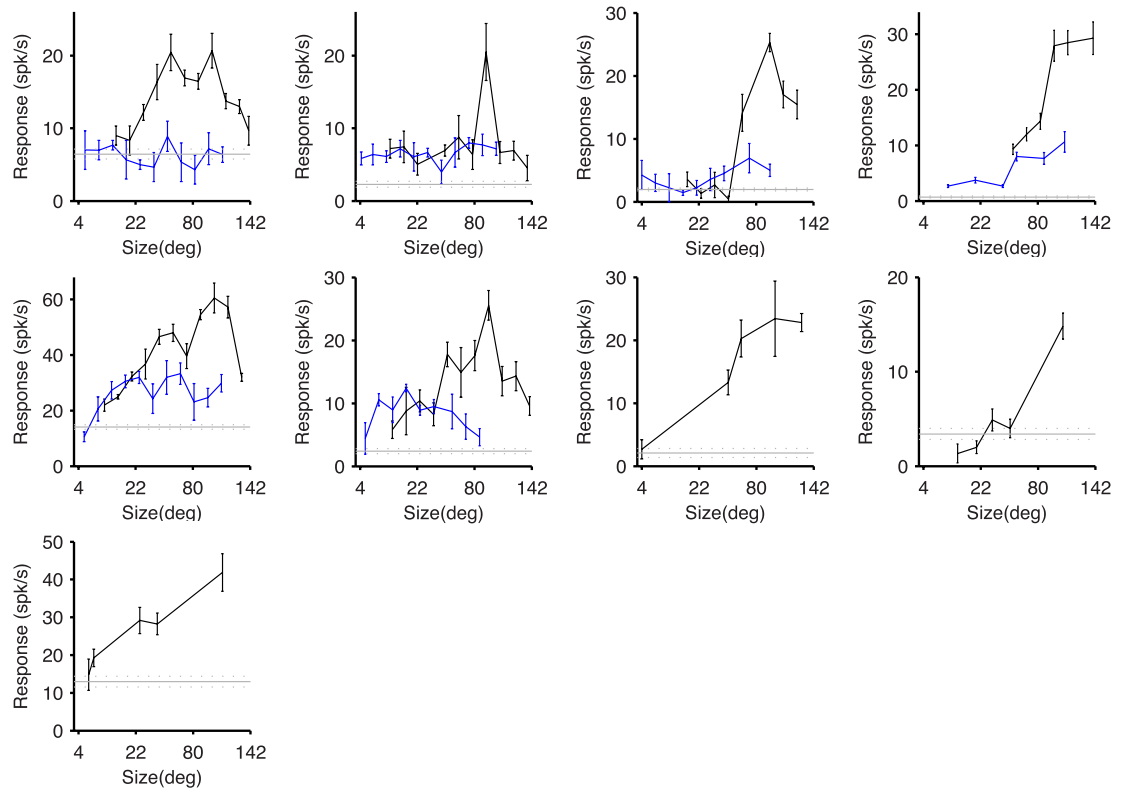
Figure 4.8: Scale-tuning test: TEd. Responses (+/- s.e.m.) of TEd neurons selective for environmental shapes (Wilcoxon rank sum test, two-tailed, p < 0.05) as a function of stimulus scale. The optimal environment stimulus shape (black) and in some cases the optimal object stimulus shape (blue) were presented at scales ranging from objects to environments.
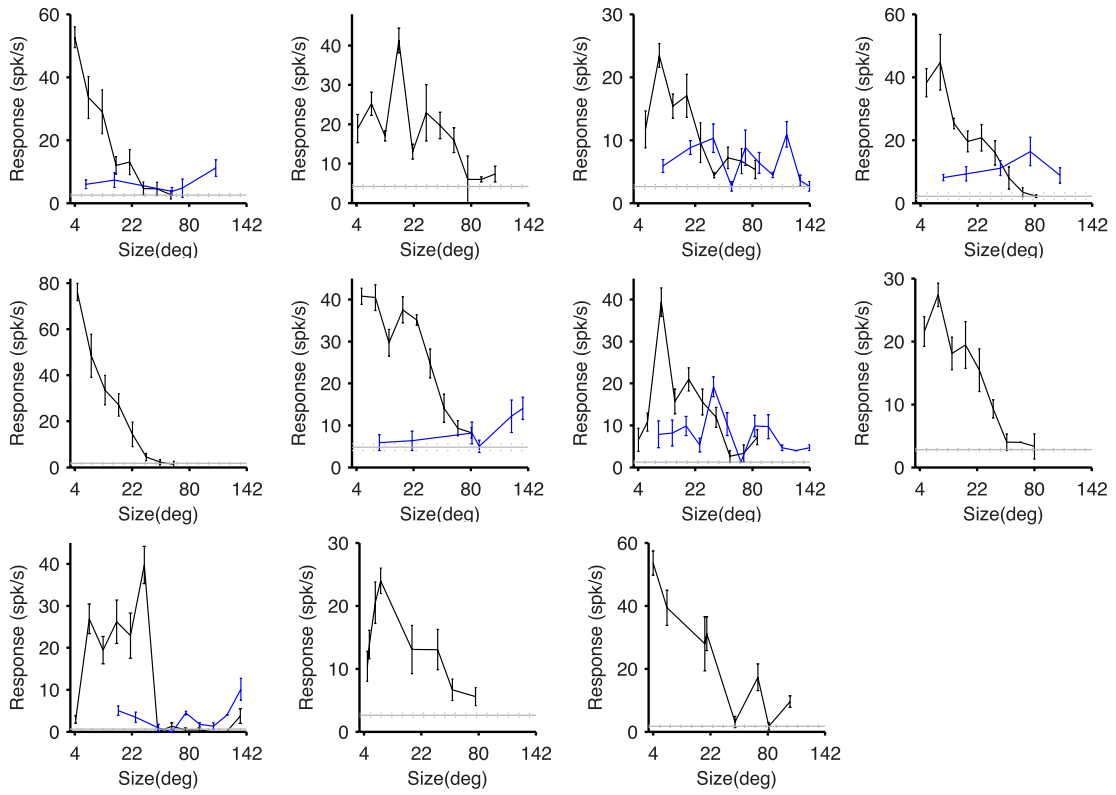
Figure 4.9: Scale-tuning test: STSv. Responses (+/- s.e.m.) of STSv neurons selective for object shapes (Wilcoxon rank sum test, 2-tailed, p < 0.05) as a function of stimulus scale. Details as in 4.8.

4.7 Chapter conclusion

The analysis in this chapter confirms that the strong selectivity for environment-scale shapes reported in TEd (see chapter 3) cannot be explained based on associated low-level stimulus properties. Firstly, TEd responses are critically dependent on shape-in-depth information conveyed by 3D depth cues. An interesting finding reported in this chapter is that non-stereoscopic cues, especially texture, serve as the important cues for shape-in-depth information contained in environment-scale shapes. Auxiliary tests presented here show that responses are invariant to changes in lighting direction and spatial frequency. Finally, additional tests for scale sensitivity applied to a single shapes showed consistent biases in tuning between STSv and TEd compared to the results of the main test presented in chapter 3.

Chapter 5

Discussion

5.1 Selectivity for environmental visual stimuli in anterior IT

The presence of such strong sensitivity for 3D environment-like visual stimuli in anterior IT (TE) is very surprising. For a long time this region has been exclusively regarded as a high-level visual area for object perception (Desimone et al., 1984; Mishkin et al., 1983; Tanaka, 1996). Consistent with these previous studies 21% of the 76 neurons recorded from TEd were selective for object stimuli. However, a much larger proportion (66%) of neurons were found to be selective for environments. Previous studies in the macaque have reported environment selectivity in more posterior regions of the ventral pathway, located near TEO (Rajimehr et al., 2011) and anterior

V4 (Kornblith et al., 2013). The results presented here indicate that processing for environments extends anteriorly all the way to TE. These areas may potentially form a processing network analogous to the multiple face processing modules that also have posterior and anterior components (Tsao et al., 2003).

In previous macaque fMRI studies that have tried to find scene selective regions, anterior IT (e.g TEd) has not been reported to show significant selective activity based on standard scene localizers.(Kornblith et al., 2013; Rajimehr et al., 2011). These studies have identified discrete selectivity in posterior IT and ventral V4. There is however some evidence from one fMRI study in macaques (Nasr et al., 2011) showing a broader swath of activity that extends from TEO into anterior IT. An imaging study in humans (Stansbury et al., 2013) based on high-resolution fMRI has also shown the presence of some environment-related activation in more anterior regions of the ventral pathway. It may be the case that low signal-to-noise issues due to interference from the ear canal contributed to the detection failure in some of these previous studies. The study by (Kornblith et al., 2013) provides a precedent in which neural recording in the medial place patch revealed selectivity for scenes even in the absence of activations based on standard fMRI place localizers. Another possibility for the lack of reported activity in previous studies might be due to a difference in the visual stimulus type. Unlike the present study which uses abstract stimuli defined purely by their structural information, these previous fMRI studies use photographic images of natural scenes that contain strong semantic-level information.

Since semantic/categorical representations in the brain can be independent of structural measures including scale (Konkle and Oliva, 2012), scene localizers based on natural images might have a greater tendency to identify areas with stronger sensitivity to semantic-level representation of environments.

5.2   TEd selectivity for 3D environmental shape relies on non-stereo cues

An important finding based on the results of chapter 4 was that TEd neurons are especially sensitive to shape-in-depth information provided by texture cues. A previous recording study (Janssen et al., 2000) in TEd based on a limited set of object-scale shapes defined stereoscopically reported that neurons in this region, unlike STSv, are not sensitive to stereo depth cues. The authors concluded that TEd is selective for 2D shapes. The results presented here encompass these findings by showing that stereo depth cues have little effect on TEd responses; however the neurons display strong selectivity for 3D environment-scale stimuli defined by shading and texture depth cues. This result is highly consistent with the fact that binocular disparity changes become negligible at far depths found in environments, providing little information about shape-in-depth. Auxiliary tests confirmed that differential responses of TEd neurons are largely abolished when 3D cues are removed, similar to cells in STSv. Therefore, TEd neurons are in fact very strongly sensitive to 3D shape that

is defined by non-stereoscopic cues, especially texture. These results are consistent with texture-based modulation of responses to natural scene images observed in a posterior environment-selective regions (Kornblith et al., 2013).

A potential low-level explanation of the texture cue effect based on high spatial frequency (Rajimehr et al., 2011) is dispelled in control tests that show little effect on responses due to changes in the texture density. Likewise, the contribution of shading cues cannot be explained by sensitivity to particular contrast patterns since changes in lighting direction had no significant effect on environment selective responses.

5.3 Anatomically distinct processing for objects and environments

Importantly, the results here indicate there is a segregation of object and environment processing along STSv and TEd, respectively. Based on neuroanatomical evidence STSv and TEd have been characterized as parts of separate processing channels in the inferotemporal cortex (Kravitz et al., 2013). STSv provides most of the IT input to ventrolateral prefrontal areas involved in object working memory (Saleem et al., 2008). STSv, unlike TEd, provides inputs to the orbitofrontal cortex which is implicated in processing object value (Webster et al., 1994). These distinct output targets are consistent with the observed stronger selectivity for object stimuli in STSv compared to TEd.

The segregated processing of object and environment information is another example of modularity in the ventral pathway. Previous studies relating to face (Moeller et al., 2008) and color (Lafer-Sousa and Conway, 2013) processing have revealed discrete processing centers for each attribute along the inferior temporal lobe. The neuron sampling presented in this study overlaps in the posterior-anterior range with the AL face patch. However, the majority of recordings were outside of this face patch, which is located at the lip of the superior temporal sulcus. Therefore, the results expand on the existing modularity previously described in IT.

The anatomically distinct selectivity for object- and environment-scale visual stimuli in anterior IT might originate from retinotopic biases propagated from earlier visual areas (Kravitz et al., 2013). In both humans (Levy et al., 2001) and monkeys (Lafer-Sousa and Conway, 2013), there is evidence for eccentricity biases that are in close association with face (overlap with central field bias) and place (overlap with peripheral field bias) selective regions. Therefore, the relatively ventral location of environment-selective TEd compared to object-selective STSv might reflect the ventral origin of peripheral upper field inputs from early visual cortex. In natural conditions, where objects in the near vicinity are foveated at ground level, the upper field periphery is likely going to contain the most information about background environment structure. In a control test, TEd neurons were tested for explicit retinotopy by showing optimized objects at various peripheral locations. The failure of these tests to evoke strong responses comparable to those observed for environment stim-

uli argues against explicit retinotopic representations. As suggested by (Lafer-Sousa and Conway, 2013), retinotopically biased inputs from earlier visual areas are not just extended to anterior regions but have adopted functional specializations associated with object/face processing requiring high-acuity central representations and environment/place processing involving peripheral representations.

5.4   Connection to human imaging studies

The segregation of object and environment information found here might be related to a recent human fMRI study reporting anatomically separate activation for objects of different size (Konkle and Oliva, 2012). It was reported that smaller objects maximally activate regions in the superior/lateral ventral regions and larger objects maximally activate regions in the inferior/medial regions of the ventral pathway. This organization roughly parallels the findings here, where selectivity for large-scale shapes resembling environments is common in the inferior channel (TEd) and selectivity for smaller-scale shapes resembling objects is common in the superior channel (STSv). However, in the human fMRI study the discovered size-based organization depended on the real-world size knowledge of the stimuli and not on their retinal size. It would be interesting to test TEd responses for environment-like stimuli that actually subtend a limited angular size. However, since the experiment here is performed on monkeys it is difficult to assess whether 'real-world' knowledge of stimuli

78

is consistent with humans.

5.5   Potential functional roles of TEd

Given that IT is known to be important in form discrimination, TEd may represent spatial information needed for recognition of familiar environments independent of object content. It may also be the case that TEd has a more advanced role in which it performs integration of environmental spatial information coming from more posterior modules (Kornblith et al., 2013) with object information from the STSv channel to create a holistic representation of a scene. The connections from STSv may provide the object information and as reported earlier the study here did find object-selective cells in TEd. The idea of an advanced role for TEd in environment processing would also be consistent with a human imaging study that showed LOC (homologous to macaque IT) appears to combine information about scene and scene-related object information (MacEvoy and Epstein, 2011).

5.6   Future directions

The analysis presented here has focused on stimulus scale sensitivity differences between the STSv and TEd channels of anterior IT. Given previous studies that specifically looked at shape tuning in the object domain (Yamane et al., 2008; Carlson et al., 2011; Hung et al., 2012), similar shape focused analyses may reveal the

presence of neural representation of structural information for environments. Potentially, model-free spike-triggered analyses can be used to test whether there exist population biases in terms of structural measures such as curvature or orientation that are related to natural image statistics. For example, a prediction for an area that represents landscape environments would be a bias in the preference of the neural population for surface orientations that point upward. In contrast, areas that represent interior environments are not expected to show a bias in surface orientation preference, as interiors contain a more complete sampling of the orientation domain. It has been shown previously that object sensitive areas show a bias in preference for surface orientations in the visible hemisphere as opposed to the self-occluded hemisphere (Yamane et al., 2008). In terms of curvature, objects and environments contain surface curvatures that overall are more in the convex range. In contrast, interior environments are more globally concave. These biased curvature distributions might be reflected in the curvature sensitivity of STSv and TEd populations.

Model-based analyses performed for every neuron incorporating explicit structural information allow for more precise characterization of shape tuning (Hung et al., 2012; Yamane et al., 2008). Models that characterize stimulus shapes based on a configuration of local surface patches described by their curvature, orientation and position values have previously been successful in modeling STSv neural responses to object-like stimuli (Yamane et al., 2008). Similar model-based characterizations can be applied to environment-scale stimuli used in this study to demonstrate ex-

plicit coding for surface shape. A potentially interesting model comparison would be between models describing position information based on a stimulus-define reference frame versus models based on an egocentric representation. Representations of environment structure for the purpose of navigation might require a more egocentric-coding of environment-structure. The successful use of models to characterize neural responses requires dense sampling in the response region of the neuron to constrain multi-parameter mathematical models. Given that this experiment encompasses a much larger space covering both objects and environments, low-sample counts may place a limitation on the use of modeling analyses.

Finally, in order to directly relate the reported results in TEd to natural images of environment, future experiments need to be performed involving a large set of natural images or possibly an innovative combination of natural images and 'equivalent' abstract shapes that just capture the structural information. These experiments will allow a comparison of the observed results in this study with previous work based on photographic natural images of environments (Kornblith et al., 2013).

Bibliography

Aguirre, G. K. and D'Esposito, M. Topographical disorientation: a synthesis and taxonomy. *Brain*, 122(9):1613–1628, 1999.

Bar, M. Visual objects in context. *Nature Reviews Neuroscience*, 5(8):617–629, 2004.

Bar, M., Aminoff, E., and Schacter, D. L. Scenes unseen: the parahippocampal cortex intrinsically subserves contextual associations, not scenes or places per se. *The Journal of neuroscience*, 28(34):8539–8544, 2008.

Biederman, I. Recognition-by-components: a theory of human image understanding. *Psychological review*, 94(2):115, 1987.

Carlson, E. T., Rasquinha, R. J., Zhang, K., and Connor, C. E. A sparse object coding scheme in area v4. *Current Biology*, 21(4):288–293, 2011.

De Valois, R. and De Valois, K. Spatial vision. 1988. *Oxford: Oxford University Press*, 10:145–175.

Desimone, R., Albright, T. D., Gross, C. G., and Bruce, C. Stimulus-selective properties of inferior temporal neurons in the macaque. *The Journal of Neuroscience*, 4(8):2051–2062, 1984.

Epstein, R. and Kanwisher, N. A cortical representation of the local visual environment. *Nature*, 392(6676):598–601, 1998.

Epstein, R., Harris, A., Stanley, D., and Kanwisher, N. The parahippocampal place area: recognition, navigation, or encoding? *Neuron*, 23(1):115–125, 1999.

Freiwald, W. A. and Tsao, D. Y. Functional compartmentalization and viewpoint generalization within the macaque face-processing system. *Science*, 330(6005):845–851, 2010.

Fujita, I., Tanaka, K., Ito, M., and Cheng, K. Columns for visual features of objects in monkey inferotemporal cortex. *Nature*, 360(6402):343–346, 1992.

Gross, C. G., Rocha-Miranda, C., and Bender, D. Visual properties of neurons in inferotemporal cortex of the macaque. *Journal of neurophysiology*, 1972.

Hegdé, J. and Van Essen, D. C. Selectivity for complex shapes in primate visual area v2. *J Neurosci*, 20(5):61–66, 2000.

Hubel, D. H. and Wiesel, T. N. Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *The Journal of physiology*, 160(1):106, 1962.

Hung, C.-C., Carlson, E. T., and Connor, C. E. Medial axis shape coding in macaque inferotemporal cortex. *Neuron*, 74(6):1099–1113, 2012.

Ito, M., Tamura, H., Fujita, I., and Tanaka, K. Size and position invariance of neuronal responses in monkey inferoterotemporal cortex. 1995.

Janssen, P., Vogels, R., and Orban, G. A. Selectivity for 3d shape that reveals distinct areas within macaque inferior temporal cortex. *Science*, 288(5473):2054–2056, 2000.

Kiani, R., Esteky, H., Mirpour, K., and Tanaka, K. Object category structure in response patterns of neuronal population in monkey inferior temporal cortex. *Journal of neurophysiology*, 97(6):4296–4309, 2007.

Komatsu, H., Ideura, Y., Kaji, S., and Yamane, S. Color selectivity of neurons in the inferior temporal cortex of the awake macaque monkey. *Journal of Neuroscience*, 12(2):408–24, 1992.

Konkle, T. and Oliva, A. A real-world size organization of object responses in occipitotemporal cortex. *Neuron*, 74(6):1114–1124, 2012.

Kornblith, S., Cheng, X., Ohayon, S., and Tsao, D. Y. A network for scene processing in the macaque temporal lobe. *Neuron*, 79(4):766–781, 2013.

Kourtzi, Z. and Connor, C. E. Neural representations for object perception: structure, category, and adaptive coding. *Annual review of neuroscience*, 34:45–67, 2011.

Kravitz, D. J., Peng, C. S., and Baker, C. I. Real-world scene representations in high-level visual cortex: it's the spaces more than the places. *The Journal of Neuroscience*, 31(20):7322–7333, 2011.

Kravitz, D. J., Saleem, K. S., Baker, C. I., Ungerleider, L. G., and Mishkin, M. The

ventral visual pathway: an expanded neural framework for the processing of object quality. *Trends in cognitive sciences*, 17(1):26–49, 2013.

Kriegeskorte, N., Mur, M., Ruff, D. A., Kiani, R., Bodurka, J., Esteky, H., Tanaka, K., and Bandettini, P. A. Matching categorical object representations in inferior temporal cortex of man and monkey. *Neuron*, 60(6):1126–1141, 2008.

Lafer-Sousa, R. and Conway, B. R. Parallel, multi-stage processing of colors, faces and shapes in macaque inferior temporal cortex. *Nature neuroscience*, 2013.

Levy, I., Hasson, U., Avidan, G., Hendler, T., and Malach, R. Center–periphery organization of human object areas. *Nature neuroscience*, 4(5):533–539, 2001.

MacEvoy, S. P. and Epstein, R. A. Constructing scenes from objects in human occipitotemporal cortex. *Nature neuroscience*, 14(10):1323–1329, 2011.

Marr, D. and Nishihara, H. K. Representation and recognition of the spatial organization of three-dimensional shapes. *Proceedings of the Royal Society of London. Series B. Biological Sciences*, 200(1140):269–294, 1978.

Mishkin, M. A memory system in the monkey. *Philosophical Transactions of the Royal Society of London. B, Biological Sciences*, 298(1089):85–95, 1982.

Mishkin, M., Ungerleider, L. G., and Macko, K. A. Object vision and spatial vision: two cortical pathways. *Trends in neurosciences*, 6:414–417, 1983.

Moeller, S., Freiwald, W. A., and Tsao, D. Y. Patches with links: a unified system for processing faces in the macaque temporal lobe. *Science*, 320(5881):1355–1359, 2008.

Nasr, S., Liu, N., Devaney, K. J., Yue, X., Rajimehr, R., Ungerleider, L. G., and Tootell, R. B. Scene-selective cortical regions in human and nonhuman primates. *The Journal of Neuroscience*, 31(39):13771–13785, 2011.

Park, S., Brady, T. F., Greene, M. R., and Oliva, A. Disentangling scene content from spatial boundary: complementary roles for the parahippocampal place area and lateral occipital complex in representing real-world scenes. *The Journal of Neuroscience*, 31(4):1333–1340, 2011.

Pasupathy, A. and Connor, C. E. Responses to contour features in macaque area v4. *Journal of Neurophysiology*, 82(5):2490–2502, 1999.

Pohl, W. Dissociation of spatial discrimination deficits following frontal and parietal lesions in monkeys. *Journal of comparative and physiological psychology*, 82(2):227, 1973.

Rajimehr, R., Devaney, K. J., Bilenko, N. Y., Young, J. C., and Tootell, R. B. The parahippocampal place area responds preferentially to high spatial frequencies in humans and monkeys. *PLoS biology*, 9(4):e1000608, 2011.

Rolls, E. T. and Tovee, M. J. Sparseness of the neuronal representation of stimuli

in the primate temporal visual cortex. *Journal of Neurophysiology*, 73(2):713–726, 1995.

Saleem, K. S., Kondo, H., and Price, J. L. Complementary circuits connecting the orbital and medial prefrontal networks with the temporal, insular, and opercular cortex in the macaque monkey. *Journal of Comparative Neurology*, 506(4):659–693, 2008.

Stansbury, D. E., Naselaris, T., and Gallant, J. L. Natural scene statistics account for the representation of scene categories in human visual cortex. *Neuron*, 79(5): 1025–1034, 2013.

Tanaka, K. Inferotemporal cortex and object vision. *Annual review of neuroscience*, 19(1):109–139, 1996.

Thorpe, S., Fize, D., Marlot, C., et al. Speed of processing in the human visual system. *nature*, 381(6582):520–522, 1996.

Tsao, D. Y., Freiwald, W. A., Knutsen, T. A., Mandeville, J. B., and Tootell, R. B. Faces and objects in macaque cerebral cortex. *Nature neuroscience*, 6(9):989–995, 2003.

Tsao, D. Y., Moeller, S., and Freiwald, W. A. Comparing face patch systems in macaques and humans. *Proceedings of the National Academy of Sciences*, 105(49): 19514–19519, 2008.

Ungerleider, L. G. and Haxby, J. V. whatand wherein the human brain. *Current opinion in neurobiology*, 4(2):157–165, 1994.

Vinje, W. E. and Gallant, J. L. Sparse coding and decorrelation in primary visual cortex during natural vision. *Science*, 287(5456):1273–1276, 2000.

Von der Heydt, R., Peterhans, E., and Baumgartner, G. Illusory contours and cortical neuron responses. *Science*, 224(4654):1260–1262, 1984.

Webster, M. J., Bachevalier, J., and Ungerleider, L. G. Connections of inferior temporal areas teo and te with parietal and frontal cortex in macaque monkeys. *Cerebral cortex*, 4(5):470–483, 1994.

Yamane, Y., Carlson, E. T., Bowman, K. C., Wang, Z., and Connor, C. E. A neural code for three-dimensional object shape in macaque inferotemporal cortex. *Nature neuroscience*, 11(11):1352–1360, 2008.

Curriculum Vitae

Born in Tehran, Iran, on February 12, 1981, Siavash Vaziri received his B. S. and M. S. E. degree in Biomedical Engineering from Johns Hopkins University in 2003 and 2005, respectively. Next, he enrolled in the Biomedical Engineering Ph.D. program at Johns Hopkins University in 2005. As a M. S. E. student at Johns Hopkins, his research showed that the brain performs optimal integration of visual information in a reaching task. The results of the work were published in the Journal of Neuroscience. As a Ph.D. student at Johns Hopkins, Siavash has researched the neural visual representation of large-scale structures that resemble environments (e. g. landscapes and interiors). The results of his research show separate processing channels for object and environment shape in the macaque inferotemporal cortex. He has presented his findings in the Society for Neuroscience conference in San Diego. Siavash was inducted as a member of the Tau Beta Pi Engineering Honor Society and Alpha Eta Mu Beta Biomedical Engineering Honor Society. Starting September of 2014, he will begin postdoctoral research at Johns Hopkins Unversity.