Chapter 8: From inferior temporal cortex to cognition

In a basic attempt to simplify, categorize and organize our understanding of brain processes, we tend to use labels such as “vision”, “audition”, “memory”, “planning” or “decision making”. In some cases, these distinctions are warranted. In the case of the retina, we can safely assume that the responses are governed by the incoming visual stimuli. Since there is no feedback from the brain to the retina, we can assume with a certain degree of safety that cognitive demands such as the task at hand will not significantly impact the responses of retinal ganglion cells (except for eye movements of course). The distinctions become fuzzy as we ascend the visual hierarchy. Therefore, given the position of ITC at the top of the visual hierarchy, it should come as no surprise that the neurophysiological responses in ITC are strongly modulated by task demands and other non-visual constraints. In this chapter, we examine some of the effects that the spatial surround, the temporal surround, learning and task constraints can exert on the neurophysiological responses in ITC.

7.1 Timing in inferior temporal cortex, expanded

In Chapter 5, we argued that there is a progression in the response latencies of neurons along the ventral visual stream. The response latencies in ITC are, on average, longer than the ones in earlier visual areas, with latency values from ~80 to ~120 ms (e.g. (Hung et al., 2005; Richmond et al., 1990; Tovee et al., 1993)). These latencies are consistent with the overall behavioral estimates of the amount of time required for object recognition (see discussion in Chapter 6). These short latencies impose a strong constraint for the development of a biophysically plausible theory of visual object recognition. Neurons in ITC can also respond to very brief stimulus flashes, as short as 14 ms (Keysers et al., 2001).

The activity of a small population of ITC neurons in the initial ~100 ms after stimulus onset provides sufficient information to decode the identity (or category) of the stimulus (Hung et al., 2005; Tovee et al., 1993; Treves et al., 1999).

Some studies have argued that the latency of ITC neurons depends on stimulus characteristics. For example, Sugase et al showed macaque monkeys images that were divided into three different categories (monkey faces, human faces, and simple geometric shapes) and presented several different exemplars within each category. They argued that whereas the early responses (before 100 ms) in ITC were able to discriminate among the three categories, they did not
provide sufficient information to distinguish individual exemplars (Sugase et al., 1999). The finer information involved in identification was available more than 50 ms later.

### 7.2 The effect of spatial clutter and attentional modulation

Most of the examples that we have provided so far involve the presentation of isolated stimuli on a uniform background. In the real world, stimuli rarely show up on a uniform background and the visual system has to deal with the problem of clutter and identifying objects embedded amidst other objects and complex backgrounds. The difficulty of this problem is emphasized by games such as “Where is Waldo?” where you need to find an object surrounded by many similar stimuli. Many animal species astutely capitalize on the difficulty of recognizing objects in clutter by using camouflage.

Neurons in ITC (as well as neurons in earlier parts of ventral visual cortex) are significantly affected by the presence of other stimuli. The responses in area V4 (Connor et al., 1997; Ghose and Maunsell, 2008) and ITC (Chelazzi et al., 1998; De Baene et al., 2007; Li et al., 2009; Miller et al., 1993; Missal et al., 1999; Rolls and Tovee, 1995; Sato, 1989; Sheinberg and Logothetis, 2001; Zoccolan et al., 2005; Zoccolan et al., 2007) to the neuron’s preferred stimuli are suppressed by the addition of a second object within the receptive field. The degree of suppression varies substantially across recording areas and experimental conditions: some studies report strong suppression (e.g. (Rolls and Tovee, 1995)) or even that neurons may compute the average of the responses to the individual objects (De Baene et al., 2007; Zoccolan et al., 2005) while others report almost no suppression (e.g. (Gawne and Martin, 2002; Quian Quiroga et al., 2005)).

Given the (sometimes massive) response reduction observed in single neurons, how can we recognize objects in natural scenes at all? There are least two non-exclusive answers. First, small amounts of clutter can be overcome by the combination of multiple neurons (Agam et al., 2010; Li et al., 2009). Ultimately, even large neuronal populations will be impaired by heavy clutter. In those cases, the visual system uses attention to filter out parts of the visual input (Reynolds and Chelazzi, 2004). In the simplest instantiation, one can focus attention on a given part of the visual field, enhancing our discriminative power within the spotlight at the expense of discrimination outside. Several studies have demonstrated that spatial attention strongly enhances the responses of neurons (Fries et al., 2001; Luck et al., 1997; Moran and Desimone, 1985; Reynolds et al., 1999; Treue and Maunsell, 1999). The strength of spatial attention affects are particularly strong in ITC and seem to follow the reverse order of the visual hierarchy (Hochstein and Ahissar, 2002).

### 7.3 Learning
The responses of ITC neurons are not fixed. Rather, learning has a significant effect on the neuronal preferences in ITC. It is tempting to conjecture that, at least in part, learning to recognize novel shapes and objects depends in the long term on adjustments to the firing properties of ITC neurons.

A series of elegant studies that speak to this property of ITC neurons was carried out by Miyashita (Higuchi and Miyashita, 1996; Miyashita, 1988; Miyashita and Chang, 1988). He presented sequences of fractal patterns to a monkey while recording the activity of ITC neurons. Day after day, the temporal sequence was the same and he observed that some of the neurons started to develop tuning for objects depending on the presentation order. The tuning of ITC neurons can also be sharpened by experience (Freedman et al., 2005). Furthermore, it is also possible to observe rapid learning effects whereby the neuronal preferences can be modified in an unsupervised fashion during a single recording session (Li and Dicarlo, 2008).

**7.4 Does it matter to you? Then it matters to ITC as well**

In addition to attentional modulation (mentioned above), neuronal activity in ITC is influenced by any aspect of cognition that you may think of. ITC neurons are interested in what you are doing with the visual information, what you are paying attention to, whether you want to retain the information, whether you have perceived the stimulus.

For example, imagine a neuron that responds vigorously to object A and not to object B and consider the following simple task. In some trials, the monkey has to search for and saccade to target A whereas in other trials the monkey has to search for and saccade to target B. The neuron will strongly enhance its activity when A is the target compared to those trials when B is the target even when the visual stimulus is identical in both conditions (Chelazzi et al., 1998). Furthermore, following up on the same example, in each trial, the monkey is instructed that the target is A or B, and there is a delay before the image with the two objects shows us. During the delay, even though there is no visual stimulus, the neuron fires more strongly when A is the target. In other words, the neuronal response correlates with the short-term memory required to solve the task. It is tempting to think of these neuronal responses as a correlate of visual imagery (Miyashita, 1993).

An extreme example of response modulation in ITC is provided by the phenomenon of binocular rivalry. If you present a stimulus A to the right eye and a stimulus B to your left eye, most of the time you do not perceive a mixture of A and B. Instead, your perception alternates in a seemingly random fashion between brief periods where you see A and brief periods where you see B (Alais and Blake, 2005; Blake and Logothetis, 2002). Given that the stimulus is constant (except for very small eye movements), what brain processes correlate with your alternating percepts? In a remarkable study, Sheinberg and Logothetis
showed that almost all neurons in ITC strongly respond when the monkey is perceiving the preferred stimulus and they are essentially shut down while the monkey perceives the non-preferred stimulus (Sheinberg and Logothetis, 1997). In other words, a neuron selective to object A will fire if and only if the monkey perceives A, even though A is still present on the screen when the monkey perceives B.

7.5 What is in the brain of a neuron?

One day, we may be able to inject a tracer in a neuron in ITC, follow all its inputs and thus construct a detailed quantitative model of what types of stimuli drive its responses. Until that day, let us speculate about a possible simple scenario. The simplest possible description is that the responses of neurons in ITC represent the (non-linear, weighted) sum of its inputs in the same way that we can think of orientation selectivity in primary visual cortex as arising from a suitable combination of LGN-type receptive fields. Neurons in ITC may respond to complex shape combinations (e.g. (Connor et al., 2007)). Several anthropomorphic approaches use two-dimensional renderings of objects such as cars or shoes. While these objects can certainly drive neurons in ITC, they do not necessarily imply that ITC neurons care about cars, shoes or any such shape. The specific features preferred by ITC neurons may be differentially represented in those complex objects leading investigators to describe a “car” preferring neuron. Care should be taken in the interpretation of such statements.

References

Higuchi, S., and Miyashita, Y. (1996). Formation of mnemonic neuronal responses to visual paired associates in inferotemporal cortex is impaired by perirhinal and entorhinal lesions. PNAS 93, 739-743.


