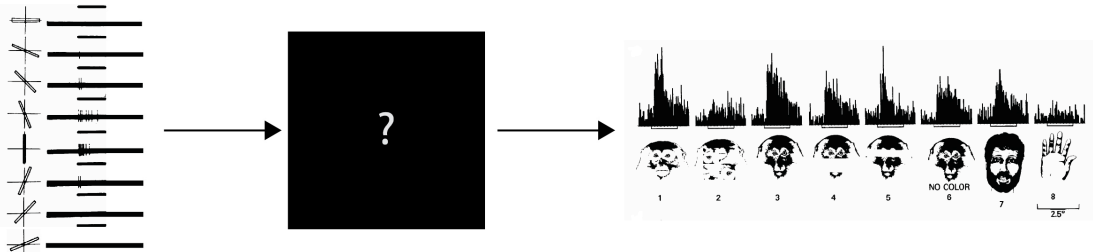


**BEWARE: These are preliminary notes. In the future, they will become part of a textbook on Visual Object Recognition.**

## Chapter VI. Adventures into *terra incognita*

In primary visual cortex there are neurons that respond selectively to lines of different orientation (**Chapter V**) (Hubel & Wiesel 1959, Hubel & Wiesel 1968). At the other end of the visual hierarchy, there are neurons that respond selectively to complex shapes such as faces, as we will discuss in **Chapter VII**. In between, there is a large expanse of cortex involved in the magic transformations that take simple stimulus features into our complex visual perceptions. How do we go from oriented lines to recognizing faces and cars and other fancy shapes (**Figure 6.1**)? Despite heroic efforts by a talented cadre of investigators to scrutinize the responses between primary visual cortex and the highest echelons of inferior temporal cortex, this part of cortex remains *terra incognita* in many ways. Visual information flows along the ventral visual stream from primary visual cortex into areas V2, V4, posterior and anterior parts of inferior temporal cortex. The cortical real estate between V2 and inferior temporal cortex composes a mysterious, seductive, controversial and fascinating ensemble of neurons whose functions remain unclear and are only beginning to be deciphered.

**Figure 6.1:** Through the cascade of computations along the ventral visual stream, the brain can convert preferences for simple stimulus properties such as orientation tuning into complex features such as faces.



### 6.1. Divide and conquer

**Table 6.1:** Response latencies in different areas in the macaque monkey (from Schmolesky et al 1998).

Area	Mean (ms)	S.D. (ms)
LGNd M layer	33	3.8
LGNd P layer	50	8.7
V1	66	10.7
V2	82	21.1
V4	104	23.4
V3	72	8.6
MT	72	10.3
MST	74	16.1
FEF	75	13

To solve the complex task of object recognition, the visual system seems to have adopted a “divide and conquer” strategy. Instead of trying to come up with a single function that will transform lines into complex shapes in one step, the computations underlying pattern recognition are implemented by a cascade of multiple approximately sequential computations. Each of these computations may be deceptively simple and yet the concatenation of such steps can lead to interesting and complex results. As a coarse

analogy, consider a factory making cars. There is a long sequence of specialized areas, departments and tasks. One group of workers may be involved in receiving and ordering different parts, others may be specialized in assembling the carburetor, others in painting the exterior. The car is the result of all of these sequential and parallel steps. To understand the entire mechanistic process by which a car is made, we need to dig deeper into each of those specialized sub-steps. To understand the mechanisms orchestrating visual object recognition, we need to inspect neuronal ensembles along the ventral visual stream.

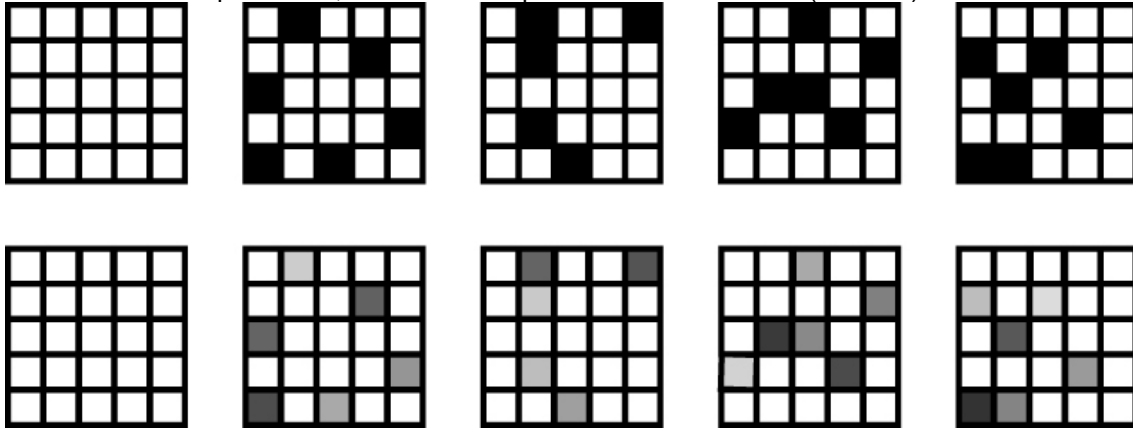
## 6.2. We cannot exhaustive study all possible visual stimuli

One of the challenges to investigate the function and preferences of neurons in cortex is that we have a limited amount of recording time. Given current techniques, it is simply impossible to examine the large number of possible combinations of different stimuli that might drive a neuron. Consider a simple scenario where we present black or white pixels within a 5x5 matrix (**Figure 6.2**). There are  $2^{25}$  such stimuli. If we present each stimulus for 100 ms and we do not allow for any intervening time in between stimuli, it would take more than 5 weeks to present all possible combinations. There are many more possibilities if we allow each pixel to have a grayscale tone between 0 and 255. We can typically hold extracellular recordings with single (non-chronic) electrodes for a couple of hours at best.

## 6.3. Response latencies increase along the ventral stream

It takes time for signals to propagate through the brain. Part of this time has to do with the speed of propagation of signals along axons and dendrites. The main component of these delays, however, is the multiple computations and integration steps in each neuron. Response latencies to a stimulus flash within the receptive field of a neuron increase from ~45 ms (LGN) to ~100 ms in inferior

**Figure 6.2:** With current techniques, we cannot exhaustively sample all possible stimuli. Here we consider a 5x5 grid of possible binary images (top) or possible grayscale images (bottom). Even for such simple stimuli, the number of possibilities is immense (see text).



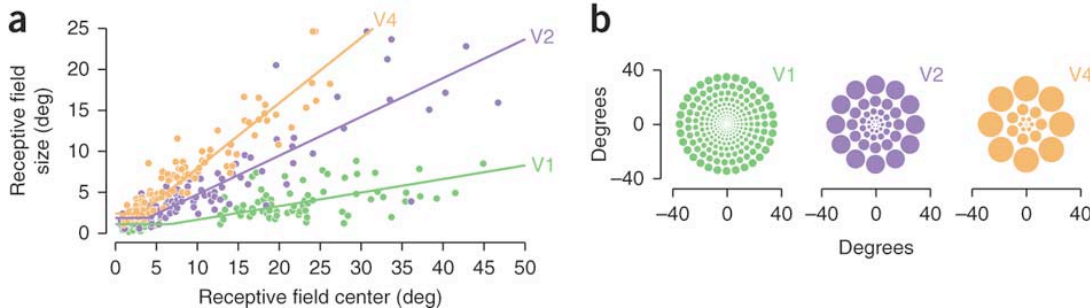
temporal cortex (Hung et al, 2005, Schmolesky et al 1998) (**Table 6.1**). It should be emphasized that these are only coarse values and there is a lot of neuron-to-neuron variability within each area. For example, an analysis in anesthetized monkeys by Schmolesky and colleagues show latencies ranging from 30 ms all the way to 70 ms in primary visual cortex. Because of this variability the distribution of response latencies overlap and the fastest neurons in a given area (say V2) may fire before the slowest neurons in an earlier area (say V1). All in all, each additional processing stage along the ventral stream adds an average of ~15 ms of computation time. The response latencies constrain the number of computations required to perform computations along the visual hierarchy.

#### 6.4. Receptive field sizes increase along the visual hierarchy

As we ascend through the visual hierarchy, receptive fields become larger (**Figure 6.3**). Receptive fields range from below one degree in the initial steps (LGN, V1) all the way to several degrees or even in some cases tens of degrees in the highest echelons of cortex (Kobatake & Tanaka 1994, Rolls 1991). The range of receptive field sizes within an area also increases with the mean receptive field size. The distributions are relatively narrow in primary visual cortex but investigators have described a wide range of receptive field values in V4 or inferior temporal cortex. A significant factor determining this range is the eccentricity. There is a significant increase in receptive field sizes with eccentricity within an area. Receptive fields near the fovea are smaller than in the periphery. The scaling factor between receptive field size and eccentricity is more pronounced in V4 than in V2 and in V2 compared to V1.

#### 6.5. What do extrastriate neurons prefer?

**Figure 6.3:** Receptive field increases within eccentricity for a given area and receptive field increases along the ventral visual stream at a fixed eccentricity. **a.** Experimental measurements based on neurophysiological recordings in macaque monkeys. **B.** Schematic rendering of receptive field sizes in areas V1, V2 and V4. Reproduced from Freeman and Simoncelli, 2013.



There have been a few systematic parametric studies of the neuronal preferences in areas V2 and V4. These studies have clearly opened the doors to investigate the complex transformations along the ventral visual stream. Despite multiple interesting studies comparing responses in V1, V2 and V4, there isn't yet a clear unified theory of what neurons "prefer" in these higher visual areas. As discussed in Chapter 3, the notion that V1 neurons show a preference for orientation tuning is well established, even if this only accounts for part of the variance in V1 responses to natural stimuli (Carandini et al 2005). There is significantly less agreement as to the types of shape features that are encoded in V2 and V4. There have been several studies probing responses with stimuli that are more complex than oriented bars and less complex than everyday objects. These stimuli include sinusoidal gratings, hyperbolic gratings, polar gratings, angles formed by intersecting lines, curvatures with different properties, among others (Hegde & Van Essen 2003, Hegde & Van Essen 2007, Kobatake & Tanaka 1994, Pasupathy & Connor 2001). Simple stimuli such as Cartesian gratings can certainly drive responses in V2 and V4. As a general rule, neurons in V2 and V4 can be driven more strongly by more complex shapes. As discussed above in the context of latency, there is a wide distribution of stimulus preferences in V2 and V4.

## 6.6. Illusory contours

Neurons in area V2 respond to illusory contours (to a much larger extent than neurons in area V1) (von der Heydt et al 1984). The responses to illusory contours are remarkable because there is no contrast change within the neuron's receptive field. Hence, these responses must indicate a form of context modulation that is consistent with human perception of borders.

## 6.7. A colorful V4

Neurons in area V4 are particularly sensitive to stimulus color (Zeki 1983). Neurons in area V4 demonstrate sensitivity to color properties that are more complex than those observed in earlier areas such as LGN parvocellular cells or V1 blobs. Neurons in V4 have been implicated in the phenomenon of color constancy whereby an object's color is relatively insensitive to large changes in the illumination.

### **6.8. Attentional modulation**

In addition to the bottom-up visual input, the activity of neurons along extrastriate visual cortex can be strongly modulated by context, task goals and other variables. A prime example of this type of modulation involves spatial attention. Neurons typically show an enhancement in the responses when their receptive field is within the locus of attention. The magnitude of this attentional effect follows the reverse hierarchical order, being significantly stronger in area V4 compared to area V1.

## References

- Carandini M, Demb JB, Mante V, Tolhurst DJ, Dan Y, et al. 2005. Do we know what the early visual system does? *Journal of Neuroscience* 25: 10577-97
- Hegde J, Van Essen DC. 2003. Strategies of shape representation in macaque visual area V2. *Vis Neurosci* 20: 313-28
- Hegde J, Van Essen DC. 2007. A comparative study of shape representation in macaque visual areas v2 and v4. *Cerebral cortex* 17: 1100-16
- Hubel D, Wiesel T. 1959. Receptive fields of single neurons in the cat's striate cortex. *Journal of Physiology (London)* 148: 574-91
- Hubel DH, Wiesel TN. 1968. Receptive fields and functional architecture of monkey striate cortex. *The Journal of physiology* 195: 215-43
- Hung CP, Kreiman G, Poggio T, DiCarlo JJ. 2005. Fast Read-out of Object Identity from Macaque Inferior Temporal Cortex. *Science* 310: 863-66
- Kobatake E, Tanaka K. 1994. Neuronal selectivities to complex object features in the ventral visual pathway of the macaque cerebral cortex. *J Neurophysiol* 71: 856-67
- Pasupathy A, Connor CE. 2001. Shape representation in area V4: position-specific tuning for boundary conformation. *J Neurophysiol* 86: 2505-19
- Rolls E. 1991. Neural organization of higher visual functions. *Current opinion in neurobiology* 1: 274-78
- Schmolesky M, Wang Y, Hanes D, Thompson K, Leutgeb S, et al. 1998. Signal timing across the macaque visual system. *Journal of Neurophysiology* 79: 3272-78
- von der Heydt R, Peterhans E, Baumgartner G. 1984. Illusory contours and cortical neuron responses. *Science* 224: 1260-62
- Zeki S. 1983. Color coding in the cerebral cortex - The reaction of cells in monkey visual cortex to wavelengths and colors. *Neuroscience* 9: 741-65