

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

Chapter 13: Towards the neuronal correlates of visual consciousness

We have accompanied and witnessed the adventures of information processing along the ventral visual stream, starting with photons impinging on the retina all the way to the remarkable responses of neurons in inferior temporal cortex. Throughout this cascade of processes we found neurons with increasing degree of similarity to our recognition capabilities. Furthermore, we have been argued that there have been major strides towards developing machines that can recognize objects using algorithms that are inspired by biological circuits. Along the way, we have perhaps forgotten about a major aspect of our visual experience, namely, the subjective feeling of seeing and experiencing the visual world. Even if we can write biologically inspired code that will recognize objects at human level performance, does that imply that computers will “feel” the way we do when we recognize those objects?

The question of subjective awareness in the context of visual perception is part of the grander theme of consciousness. The age-old question of how a physical system can give rise to consciousness has been debated by philosophers, clinicians and scientists for millennia. Over the last decade, there has been increased interest in using modern Neuroscience techniques to further our understanding of the circuits and mechanisms by which neurons may represent and distinguish conscious content. Two particularly good books that discuss the history and progress in the field include (Crick, 1994; Koch, 2005).

13.1 A non-exhaustive list of possible answers

Multiple answers have been proposed over the years in an attempt to explain how a physical system can give rise to consciousness. I will not have time to do justice or discuss them in detail here. Instead, I would like to group them and list some of the main ones.

(1) “Religious” answers. These are non-scientific explanations that often invoke the need for a soul, a homunculus, or some form of communication between physical systems and other entities. Several variants of these explanations abound including the writings of Plato, the bible, Aristotle, Thomas Aquinas, Karl Popper, Sigmund Freud and even top-notch scientists such as John Eccles.

(2) The “mysterian” approach. Proponents of this notion argue that science simply cannot understand consciousness.

(3) It’s just an illusion. Some philosophers (e.g. Dennett) argue that there is no such real phenomenon as consciousness. The feeling of consciousness is just an illusion.

(4) Epiphenomenon. A related version of this is the notion that consciousness is an epiphenomenon. As soon as multiple neurons and complex networks are connected, the feeling of consciousness arises but it does not serve any purpose

(in the same way that a computer may heat up during function but this heat does not really serve any computational purpose).

(5) New physics. Others (e.g. Roger Penrose and more recently Christof Koch) argue that we need new (as yet undiscovered) laws to explain consciousness.

In stark contrast with the above explanations, several neuroscientists have become interested in the arguably simpler notion that consciousness arises from the specific function of neuronal circuits. Which circuits, when and how remains to be determined through scientific investigation without invoking new laws of physics or souls. This approach involves several working assumptions:

(1) We are conscious (i.e. consciousness is a real phenomenon and it is not an illusion or an epiphenomenon)

(2) Some other animals are also conscious (which enables us to probe for consciousness in animal models)

(3) We start with simple questions that we can try to study rigorously. We start with vision. Hopefully, we will be able to extrapolate some of what we learn from vision to other sensations (e.g. pain, smell, self-awareness)

(4) We need an explicit representation. Only parts of the brain will correlate with the contents of consciousness. We search the *neuronal correlates of consciousness* (NCC)

The strategic decision to start by investigating a rather reduced domain, that of investigating the neuronal correlates underlying visual awareness, clearly leaves many fascinating topics out. These include dreams, lucid dreaming, out of body experiences, hallucinations, meditation, sleep walking, hypnosis, self awareness, the so-called notion of qualia and feelings. This does not imply that these are not interesting and relevant topics; it merely reflects a strategic decision of how to approach a difficult scientific question.

13.2 The search for the NCC, the neuronal correlates of consciousness

The NCC (neuronal correlates of consciousness) is defined as a *minimal* set of neuronal events and mechanisms that are jointly *sufficient* for a *specific conscious percept*.

It is critical to define some of the highlighted terms. The NCC is defined as a *minimal* set. A solution such as “the whole healthy human brain can experience consciousness” is not very informative. The neural mechanism should be “sufficient” (not just necessary) to represent a conscious percept. This clause leaves out so-called “enabling” factors such as the heart or the cholinergic systems arising in the brainstem. We are seeking for the correlates for “specific conscious percepts” such as seeing a face (as opposed to generic aspects such as being conscious/unconscious).

It is quite clear that not all brain activity is directly linked to conscious perception at any given point. To clarify, this does not mean that those brain processes are not important or interesting. For example, significant resources

and neurons are devoted to controlling breathing, posture, walking, etc. With some exceptions, most of the time we are not aware of such processes. A particularly striking documentation of relatively sophisticated brain processing that does not reach awareness is given by a patient studied by Goodale and Miller (Goodale and Milner, 1992). This patient had severe damage along the ventral visual stream and the dorsal stream was relatively unimpaired. The patient could not recognize shapes but could still act on those shapes with relatively sophisticated precision. For example, the patient could not report the orientation of a slit but could place an envelope on the slit rather accurately. The search for the NCC concerns investigating which neuronal processes and mechanisms correlate with conscious content and which ones do not.

13.3 In search of an explicit representation

Upon seeing an object, the retinae are activated. In fact, stimulating each of the photoreceptors in the same pattern and magnitude evoked by a given object should elicit the object's percept. Does this imply that the retinal photoreceptors constitute the desired NCC? Not quite. Those neurons in the retina activate neurons in the LGN, which in turn activate primary visual cortex, which in turn transmit the information to higher areas within ventral visual cortex. Several lines of evidence suggest that the activity in early visual areas from the retina to primary visual cortex cannot be the locus of the NCC. I will not discuss here the discussion and argumentation against a correlation between early visual activity and cortex; for a lucid discussion, see (Crick and Koch, 1995). A critical aspect of the NCC is that the representation of visual information must be "explicit".

One way to define an explicit representation is that it should be possible to decode the information via a simple linear classifier. A computer may hold a representation of the information for a given object in a digital photograph. However, as we have discussed in Chapters 11 and 12, decoding such information requires a cascade of multiple computations. Information about objects is not explicitly represented in a digital photograph. Similarly, the retina does not hold an explicit representation of objects.

Several visual illusions acutely point out the need for explicit representations. Consider the Kanizsa triangle illustrated in Chapter 1. We perceive strong edges defining the triangle even in parts of the image where there is no visual information (i.e. there is no real edge). Such a perception of an edge implies that there should be neurons that represent that subjective edge. Neurons in the retina do not respond to such illusory contours.

13.4 An experimental approach to study visual consciousness

The discussion of visual illusions suggests a promising path to investigate the neuronal correlates of visual consciousness by investigating which neuronal processes coincide with subjective perception. A particular type of visual illusion that has been quite fruitful in this regard involves the use of bistable percepts. A famous example of a bistable percept is the Necker cube. The same visual input can be seen in two different configurations. In the case of the Necker cube, it is possible to voluntarily switch between the two possible interpretations of the same input.

Such volitional control is not possible in the case of binocular rivalry. Under normal circumstances, the information that the right and left eyes convey is highly correlated¹. What would happen if you show two completely different stimuli to the right and left eyes? Under these conditions, we perceive either one or the other stimulus in a seemingly random fashion, a phenomenon called binocular rivalry (Blake and Logothetis, 2002). Extensive psychophysical investigation has provided a wealth of information about the conditions that lead to perceptual dominance of visual stimuli, what can or cannot be done with the information that is being suppressed and the dynamics underlying perceptual alterations (Blake, 1989; Blake and Logothetis, 2002). What is particularly interesting about this phenomenon is that, to a reasonably good first approximation, the visual input is constant and yet subjective perception alternates between two possible interpretations of the visual world. Investigators then ask: what are the neuronal changes that correlate with these subjective transitions? Several studies have shown that only a small fraction of neurons in early visual areas follow the subjective changes whereas most neurons in higher visual areas are strongly modulated by the immediate contents of visual awareness (Leopold and Logothetis, 1999).

This type of experiments may pave the road to an initial understanding of certain circuits and neuronal activity changes that correlate with subjective perception. What would constitute evidence of understanding the NCC? We argue that four conditions should be met. (1) We should be able to model and predict neuronal responses given a perceptual state. (2) Conversely, we should be able to predict perceptual states from neuronal responses. (3) We should be able to elicit a percept by activating the corresponding neuronal patterns (e.g. via electrical stimulation). (4) We should be able to inactivate or repress a perceptual state by modifying the neuronal activity patterns. There still seems to be a long way to understand the neuronal correlates of visual consciousness by meeting these four conditions. Yet, nowadays, these questions and themes have become a major area of research and we may be surprised to observe major progress in the field in the years to come.

13.5 References

¹ It is not identical, though. The small differences between the input from the right eye and left eye provide strong cues to obtain 3D information. 3D movies specifically

- Blake R (1989) A neural theory of binocular rivalry. *Psychological Review* 96:145-167.
- Blake R, Logothetis N (2002) Visual competition. *Nature Reviews Neuroscience* 3:13-21.
- Crick F (1994) *The astonishing hypothesis*. New York: Simon & Schuster.
- Crick F, Koch C (1995) Are we aware of neural activity in primary visual cortex ? *Nature* 375:121-123.
- Goodale M, Milner A (1992) Separate visual pathways for perception and action. *Trends in Neurosciences* 15:20-25.
- Koch C (2005) *The quest for consciousness*, 1 Edition. Los Angeles: Roberts & Company Publishers.
- Leopold DA, Logothetis NK (1999) Multistable phenomena: changing views in perception. *Trends in Cognitive Sciences* 3:254-264.