Neurobiology

Dr. Nitin Gupta

Department of Biological Sciences and Bioengineering

IIT Kanpur

Week - 04

Lecture 5.3: Holistic visual perception

Hi everyone, welcome back to neurobiology. In this series of videos we have been talking about sensory perception. And in the last video we looked at the receptors, the first stage of processing. We saw that the receptors can be narrowly or broadly tuned and they can have large or small receptive fields. In this video we will look at various stages of sensory processing after the receptors and we will also discuss that the final perception is not simply a function of what was received at the receptors but it can be profoundly impacted by our prior expectations about the world. There is a lot of similarity between how our eyes function and how our camera functions.

So in both cases there is a lens in front through which the rays of light can be focused on a screen. So in the camera the screen can be a film or in modern cameras it could be a CCD or CMOS sensor. And in the case of eyes the screen is the retina. And these screens are lined up with photoreceptors.

So in the case of eyes these are the three kinds of cones and in the case of cameras these are three types of light detectors at each pixel. So each of these pixels has a receptive field. So any point in the visual space will map to a unique point on the screen both in the case of camera as well as in the case of retina. But after the screen the similarity begins to break down. So the camera just transfers this pixel by pixel information to the computer and that's what the final image is.

But in the case of eyes there is a lot of processing that happens. And what we finally see is not simply the pixel wise information that is projected on the retina. One important feature of our perception is that it is holistic in nature. So rather than viewing individual pixels in a visual image we tend to see objects that are familiar to us. This is also described in psychology literature as the idea of gestalt which basically refers to the idea that an organized whole can be more than the sum of its parts.

What we perceive finally can be very different from what is happening at the level of individual receptors. Here is an example of this. What do you see in this image? Most people describe this image as consisting of three black circles, one big triangle in between that has a black outline and one white triangle on top that is covering the first triangle and the three circles partially. But if we zoom into the pixels then we see that there are no big triangles here in reality. There are only three partial circles and three v's.

So the perception of two large triangles is actually a creation of our brain. And this happens because the brain interprets the images based on assumptions about what it expects in the world. These expectations have been set based on our prior experiences and also based on our evolutionary history over millions of generations that has gotten encoded in our neural wiring through our DNA. If you are not convinced with the previous example, let me try to give another example. What do you see in this image? Can you see anything meaningful being depicted here? Let me give you a hint.

This image contains some alphabets of English language. Let me give one more hint. There is actually just one alphabet that is repeated multiple times and present in various orientations. Are you able to figure it out now? Maybe not. Now what I am going to do is I will spill some ink over this image between the informative yellow parts.

I will not affect the yellow parts at all. And let's see if that makes any difference. Okay, so I have done that now. Now can you tell what is being shown in this image? Now you can easily tell that the image contains letter B present five times in different angles. And you can compare these two images to ensure that I have not cheated.

Ink has been added only in those parts that were gray. The yellow parts are still exactly the same in the left and the right image. And you can see that the region covered by the black ink is pretty much random. So why is it that we are able to decipher the letters easily after smearing the ink but not before that? This happens because our brain has become very good at separating foreground objects from background objects. Our brain does this computation a lot where one object covers another object.

And it assumes that some part of one object can be missing because they are covered and it can automatically fill for them. But when we see the same yellow portions without the black ink, this computation of filling in is not triggered. And so we are not able to recognize the characters. Let's look at one more example. So now we are seeing two images on left and right.

And both the images have a gray background. And there are some circular shapes in between. These shapes may appear as spheres that are coming out of the screen, or as cavities that are going into the screen. So which one appears to you as spheres and which one appears as cavities? Most people think that the left image has spheres that are coming out and the right image has the cavities that are going in. At least that's the first impression.

And the interesting part is that both these images are actually identical, except that the right one is rotated by 180 degrees. So let me show you that. I have rotated 180 degrees and now it is exactly identical as the left image. And now these shapes also appear the same as what we are seeing in the left image. They are probably looking like spheres coming out of the screen.

And if I rotate it again by 180 degrees, now it again becomes the original image and the spheres are turned into cavities, even though the image remains exactly the same except for the rotation. So what's going on here? The image is actually a 2D image and this perception of either spheres or cavities is a creation of the brain. And that is based on our assumption of where the light source is. So if you assume the light source on top, then the left image would be consistent with the spheres coming out because the top part would be exposed to light and the bottom part would show the shadow. And these images where the lower part is brighter, this would be consistent with cavities because now the light is falling on this part and the top part is in the shadow.

But if you assume light coming from bottom, then these would be consistent with cavities and these would be consistent with the spheres. So why is it that most people tend to think the left image as spheres rather than cavities? That's because most people automatically assume that the light is coming from the top and that has something to do with our past history because light sources, both natural such as sun or moon or artificial such as room lights or street lights are usually on top rather than being on the ground. And so our brain has gotten used to interpreting images accordingly. And this is something that is happening even without our conscious attention. We don't have to think and do all these calculations in our brain to figure out where the light is coming from and where the shadow should be and then interpret the image.

It has become a part of our sensory processing. Now here is a bit more entertaining example. I'm going to show you two images. Each image has two persons seated on chairs and you try to compare their sizes in the two images.

Here they are. So the person on the left has the same size in the two images but I'm sure you'll find that the person on the right looks much smaller here than here. But if we look in terms of pixels then they are exactly the same size. And let me convince you about that. So I'm going to make a rectangle around this person and now I'm going to shift this rectangle and you can see that they are exactly the same size in terms of pixels. So why is it that this person here looks much smaller than the person here? Again that has something to do with how our brain is automatically processing this image.

So it already analyzes the image and makes interpretation about the depth of each position and the perceived size of the person is calculated based on this assumed depth. So in this left image the brain assumes that the person is sitting at a farther distance and takes that into account to calculate the size. Whereas in the right image the brain interprets this image as the person sitting in front and so that correction for height is not made. The sensory processing in the visual system seems to be optimized for recognizing objects rather than paying attention to all the pixel-wise data. So it tries to extract those features that are helpful in identifying objects and the most important such feature are the edges.

So a line drawing like this that has very minimal information can still tell us very clearly that this image contains mountains and a house and a tree and an animal and a vehicle and maybe a sun and even birds. There is very little information in these scribbles here but our brain unmistakably interprets them as birds and this is possible because our brain has over time learned to interpret objects using their outlines rather than the finer details. The various features of the visual stimuli are extracted by different layers of the visual system. So here is a quick schematic of the visual system. The information from the retina goes through the optic nerve to a region within the thalamus called the lateral geniculate nucleus, LGN and from there it goes to the primary visual cortex V1.

Here is another representation of the same. So there are different types of cells in the retina, then the thalamus and then it goes to V1. The V1 is a cortex so it has multiple layers. There are some calculations happening in the V1. So various kinds of contrasts are calculated and then the information goes to higher cortical areas such as V2, V4, inferior temporal cortex and so on and on this side medial temporal cortex and other areas.

And at each of these areas different features of the visual stimuli are being calculated. And as we have seen that the prior experiences can play a role in these calculations so there is also a lot of feedback that is not depicted in this figure. So this feedback can also influence the processing at these layers. And as this information goes deeper and deeper more and more parts of the brain get involved. So a large part of our brain is involved in the visual perception. .