

The following content is provided under a Creative Commons license. Your support will help MIT OpenCourseWare continue to offer high quality educational resources for free. To make a donation or to view additional materials from hundreds of MIT courses, visit MIT OpenCourseWare at [ocw.mit.edu](http://ocw.mit.edu).

**PROFESSOR:** All right, so this brief reminder here is about the basic requirements in this course, which are that we have two written reports. And that's the main thing I want to remind you of, one in vision and one in audition. And I presume all of you have the syllabus that was a hand out. And in that syllabus-- and if you don't have it, we can provide you, of course, with a copy.

And on the second page of the syllabus, it says written report, vision part, and this report is one that is based on an article published many years ago, which is very influential article back in 1967 about the so-called accessory optic system. And what I would like you to do is to read that article but then proceed through the internet to see what have become the most recent discoveries about the so-called accessory optic system.

And that's what your report needs to be about. You can be reasonably summary about it so that you don't need to write 50 pages. But you know, you can write this whole thing up in four or five pages, taking into account as I've said what has been discovered since those days about the so-called accessory optic system.

The reason I thought this would be a good assignment because it will give you a good historical sense of how knowledge has expanded, in this case, since the 1960s, in uncovering yet another attribute that you have in the visual system that starts with a specialized set of cells in the retina and then proceeds through a special series of steps until it connects with the ocular motor system.

So anyway, that's the task. And that will consist of 10% of the report, of your total grade. And the same thing is 10% also for audition. Then the midterm exam is going to consist of 25% of your exams. And the final exam is a total 55, a which 15 is

vision and 40 is audition. So if you add them all up, you have an equal distribution in percentages for vision and for audition. So that then is the basic layout.

And so what I would like you to start thinking about is how you're going to put this report together. I was hoping that perhaps you could get it done by midterm. But I don't have a hard and fast rule about that. As long as you get it in before the final exam, that's fine. But it may make life easier if you can start working on that while we are talking about vision, rather than waiting until we are covering the auditory system in the course. Does anybody have any questions about this basic layout? OK.

All right, now today we are going to talk about form perception. Let me say one more thing, that next time as you have it on the syllabus, we are going to be covering illusions. I think that'll be fun. You'll see all kinds of interesting illusions and some inferences as to what we think about how those illusions come about by virtue of the fact that there are all sorts of interesting rules and facts about how the visual system works.

And then during the second half over the next lecture, this is this coming Wednesday, we are going to talk about visual prosthesis. All right, so that in essence what is we are going to cover next time. So now let's get back to today. We are going to talk about form perception.

This is a topic that, even at this stage, we still only have rather limited knowledge about how the brain is capable of carrying out the perception of form. And we are going to look at some inferences and some ideas about it. And I will also provide you with a brief historical background. Because this has been a topic of tremendous interest for people for centuries, trying to understand just how we are capable of seeing forms, shapes, patterns.

All right, so the first influential idea that emerged about this is called structuralism. And this idea actually took place in the late 19th century. And one of the prime proponents was a fellow called Titchener. And they had, in trying to come to grips with how we do this, had a rather basic idea, much like building a house with bricks.

The idea was that perception is an aggregate of simple elements. We just have a whole bunch of elements that you put together. And that generates a more complex percept, OK. The problem with this approach, well, the first problem was that they began to do experiments. And they asked subjects to try to divorce a personal impression of anything they looked at. But to sort of physically describe it.

So for example, if you looked at an apple, how do you have the idea that it's an apple? And so they said, well, it has say, four colors. And so the person would list four colors and said would have several shades. And put all those together, add them all up, and that equals apple.

Well, that was the idea then. And when they did this systematically, and the thing became almost ridiculous. Because they came up with more than 40,000 elementary sensations, thinking that somehow you puts these elementary sensations together, 40,000 of them, and that gives us a sense of an apple, or banana, or whatever.

Now as soon as this sort of became well known, a lot of opposition arose. And one of the most famous ones then was the opposition that was brought about by Gestalt psychologists. And the first consideration that I want to sort of bring up here is a rather simple picture. Here's a picture everybody looks at that. OK, what did you see there?

You saw a car, right? Now if you look at it, I mean this is actually a picture of a car, all right, but it's interrupted by all these vertical bars here. And yet we're able to infer that it's a car. And so the question comes up, how do you do this? I mean, you can't simply just add up a bunch of bricks and say it's a car. Because there's all this interruption.

And most things that we see in the world, we see in a discontinuous fashion. And as a result of this, the Gestalt psychologists, which happened sort of in the 1920s, began to think about this whole problem in a very different way. And they made a tremendous influence. And they came up with these new ideas of how we create the

perception of form.

The founder of the Gestalt psychology is a fellow called Max Wertheimer. So now, you can look this up in Wikipedia, by the way, if you look at them more closely, they came up with a few basic principles of organization. And they call this grouping. They argue that in the brain, somehow we group things up.

And one of the things we group has to do with proximity. Another one has to do with similarity, another one with common motion, another one with closure, and another one with figure-ground perception. This last one I'm going to talk about at the end of today's presentation. But now I'm going to give you an example of these top two.

All right, so the most important conclusion that they had come to which was very much against structuralism is that the whole that you see is different from the sum of its parts. Somehow, there is an active process that creates our ability to see something that is not evidenced from its individual parts.

Now here is an example of what these principles are. Grouping by proximity, so here we have a bunch of dots. And if you'll put them closer together vertically, you see a bunch of vertical lines essentially. And if you put them closer together horizontally, you see a bunch of vertical lines, horizontal lines, vertical lines. Yeah? So we group things due to proximity.

Another reason we group things is shown here, is we group things according to shape, or similarity of shape, I should really say. And here what we have, you can readily see a group of nine disks here, and a group of nine triangles. And if they're not all the same as down here, then you have much more difficulty grouping it. So there's a strong tendency that we group things together that are similar.

So these are some general principles of how we organize our visual percepts. So now we can look at some more of these examples. In doing this, we are now going to try to be able to say a bit more about the brain itself. And in doing so, it's evident that the three major theories that try to deal with how the brain does it, according to the first of these, form perception is accomplished by neurons that respond

selectively to line segments of different orientations.

Now that theory, obviously, is the outgrowth of what we talked about when it was discovered that in V1, we have orientation selected neurons. And this orientation selectivity is something that you see in several progressively higher visual areas, as if everything in the world out there was broken down into oriented line segments, which are then somehow put together according to their orientation to enable us to see shapes. So that's one theory.

Another theory is that form perception is accomplished by spatial mapping of the visual scene onto the visual cortex. And I will elaborate on each of these. And the third theory is one that form perception is accomplished by virtue of Fourier analysis. So let us first then turn to form perception, supposedly based on breaking down the visual scene into oriented line segments.

Now, one of the big problems with this is that when you look at a famous art work that you see in every day in the *Wall Street Journal*, there is an artist who created this originally. And he created faces like this using only dots and varied the spatial frequency of the dots, as you can see here. This happens to be a person called Larry Poons. I'm sure none of you have heard of him.

But here was one of the pictures many years ago in the *Wall Street Journal*. And even today, as I say, you can see faces created this way. And you can readily recognize them, even though there are no oriented line segments here. And in fact, if you sort of squint up and you three quarter ways close your eyes, so you can't even see the dots anymore, you can still make out that face very, very clearly.

That's because the spatial frequency effect here is very important, and the degree of shading, meaning that this involves rather low spatial frequency analyses, rather than the analysis of particular orientations of line segments. Now this particular analysis has its counterpart in the observation that many of you have seen.

When you look at a person's face on television, and they want to prevent you from being able to see that face because it's confidential or something, what do they do?

Anybody remember? What you do is you put up a bunch of squares of different spatial frequencies. And each square is comprised of the mean illumination level of the actual face.

If you do that, this high frequency information is something that interferes with your ability to analyze face. So to give you an example of this, I'm sure all of you have seen this, but mainly I show you this. Here we have this example. How many of you can tell who these people are here? I guess nobody can, right?

Now then what we can do is we can increase the frequency of these. Here it is. And it's still very difficult to tell. And now I'm going to show you the actual photograph. And what I'm curious about is how many of these two people you actually know. I can tell you right away, they're extremely famous actors and actresses.

Who recognizes these people? How many of you recognize? Let me see your hand. Only a few people recognize it. That's how quickly time goes by. These were the most central, most exciting, best known people in the movies. And this is Humphrey Bogart, Ingrid Bergman.

OK, so these are these two very famous people whose faces were obstructed by this and by this using these high contrast edges that obscure your ability to smoothly analyze faces. So that then suggests that this idea that we extensively use oriented line segments to analyze faces at least is an insufficient explanation of how the brain processes shapes.

Now another idea is so-called topographic mapping. This idea I'm sure you can already reject on the basis of what I told you about how the visual system is laid out. But now I'm going to belabor that so that you can follow it closely. Here we have a runner, of course. And here we have a monkey brain to make it easy to understand.

And the idea here there is that this image, once it was discovered that the visual field is laid out topographically in the visual cortex, the idea was that what happens is somehow, and I may be an unfair to poke fun of it, but the idea is that the mind can look at the creation of this image on the cortical surface there. And thereby, it

can identify.

OK, it's almost like a photograph, looking at a photograph. In this case the mind looks at the photograph, so to speak, on the cortical surface. So they thought at the time is that this is what you have there, OK. That's the imprint of this image. And indeed, you say, oh my goodness, that's just like that. Therefore, I can recognize that person, and so on and so on.

Well, I mean that's a cute idea. But then when you take into account this has been discovered subsequently to these ideas, that the topographic layout in the visual cortex is actually not one to one because of the magnification factor. So let's look at that in some detail.

Here is an actual reconstruction of the monkey area V1 here. And here is the visual field. And we are going to put these red arrows in the contralateral hemifield. Now remember what I told you before, that if you look at the visual scene, you can imagine your eyes being vertically cut in half. And you have a nasal and temporal hemiretina.

And the one which is contralateral hemifield crosses over and gets into this half of the brain. And this one crosses over to that half. So now if you do this and take the magnification factor into account, you look at these arrows, which are all identical in size, OK. There are one, two, three, four, five, six, seven of them. What you can see here is the actual impression on the cortical surface of the neurons that are being activated by this look like this, nothing like those arrows there.

And in fact, the central arrow is much bigger because of the huge magnification factor. Now this already creates a major problem in trying to believe in that particular theory, the topographic mapping theory. But now, if instead, you put those arrows halfway across each of the hemispheres like that, then whatever is on this side goes here. Whatever is on this side goes there. And this is the impression that is created.

And my goodness, you create that impression. And you say to yourself, my god. If

that's the case, how come I can see seven straight arrows of equal size when this is the impression that is being created in the brain? So therefore, obviously, they are not using a topographic map to analyze the visual scene at all.

Now this can be driven home even further. Let me make another point here. Here we have a monkey visual cortex again from the rear. Here is the visual field laid out. And if here we put in a bunch of dots in a circular fashion, this is the activation there. Now you say, oh, that's not bad. That looks like a circle.

But then if you put it along the midline, half and half, then this is the actual activation. But you still see a circle, even though the activation, in terms of the topography, is nothing like a circle. So that indeed created a much, much greater degree of skepticism among investigators to try to understand how we process shape.

Now let me make one other point here, which is a wonderful story, which I think will be good for you to remember. That's called the Giotto story, which says that when Pope Benedict, that was in the 12th century, or the 13th century, he set out to have the walls of the great cathedral of Saint Peter in Rome redecorated.

And so he sent out a bunch of messengers to various artists in Italy and asked them to provide some of their best work so he could evaluate it and could pick one guy to actually do the redecoration. Well, one of these messengers went to Ambrogio Bondone Giotto and asked him, he was a well known artist, and asked him to provide a painting of his, a drawing of his.

And Giotto said, oh, my goodness. I just don't have anything around. But I tell you what I'll do. He took out a red pen and drew a perfect circle, OK. And so the messenger took this to the pope. And the pope said, my god. I can't believe how incredible this is. And so Giotto got the job of redecorating the cathedral of Saint Peter.

To this day, there is this expression in Italy, in Tuscany in Italy, which says the round O of Giotto, OK. So this is the round O of Giotto, which somehow, in a sense,

denotes perfection, perfection in sight and in perfection in your ability of execution in terms of a drawing, for example.

All right, now to highlight this even further, let me point out this to you here. What we have here is a bunch of imperfect circles, one of which is perfect. So if you keep looking around, you should be able to spot which one it is. And you should tell me what letter denotes that circle. Which one is a perfect circle?

**AUDIENCE:** C.

**PROFESSOR:** Very good, all right. So we are incredibly good. There's a slight difference here, OK, or even slight difference here. And yet, we can see this very slight difference. And we can tell what a perfect circle is. I mean, that's incredible given how the impressions are made by those circles on the visual cortex. So that is an incredible puzzle of how we are capable of doing this. And I'm afraid, even to this day, we don't have a really good answer of how this happens.

OK, now the third theory that has become actually probably right now, one of the most successful ones, it claims that you analyze the visual scene by taking into account the spatial frequencies that are impinging on the retina, OK. This was created by Fergus Campbell and John Robson, very influential.

And they pointed out, first of all, that a very interesting finding, which I think I've mentioned once before, that if you vary the spatial frequency [? both, ?] as well as the contrast, we have this sensitivity function like this. I did show you a picture of it the last time. And then, furthermore, what they had shown is that you can create all kinds of complex precepts by varying the gratings. Make them simple gratings, compound gratings, and compound gratings with much lower contrast.

And this is simple different spatial frequencies. These are compound. And these, as I've said, are lower spatial frequencies. So then if you again squint, then you will see something much smoother. So they actually carried out a detailed mathematical analysis using Fourier analysis. And what they did was quite remarkable.

They would down break down a visual scene, like a photograph of New York with all

the skylights, OK, skyscrapers. And then they would convert that using Fourier analysis. And they could recreate the visual scene with a high degree of accuracy using that procedure. Now if indeed the visual system uses this, you have a number of basic logical requirements.

And those basic logical, let me come back to that, the basic logical requirements are that you need spatial frequency analysis. And I've shown you that already when we talked about V1, that neurons there are spatial frequency selective. Secondly, that there are contrast selective, which you know already. And of course, the orientation selective, and they can tell you about phase.

As long as you have these four attributes, you can perform a detailed Fourier analysis to reconstruct the visual scene. So now to stress this even further, they did a series of experiments, in which they asked a question, is it true that you have a particular spatial frequency analysis, that you can manipulate that?

And so they did an experiment, which called a frequency-specific adaptation experiment. They would present to a subject this display and have the subject look at this for a couple of minutes without having to fixate. And then they would look at each of those. And they found that this, which is the same spatial frequency, they had difficulty seeing because of this adaptation.

And so then they did a series of careful studies and carried out this analysis and showed that you could get any kind of spatial frequency to lose your sensitivity for it if you had been pre-exposed to it. So that's what that looked like. And by doing this systematically, they came up with the idea that what you have in the visual cortex is a series of channels that are spatial frequency-selective.

And I showed you we talked about V1, that indeed, there are neurons there that are selected to particular spatial frequencies. And they proposed that you have a series of channels like this, OK, that peak at different spatial frequencies. And by activating them selectively, you can reconstruct virtually anything out in the visual scene using Fourier analysis.

Now OK, I think that then is the essence of that theory. And I can tell you that some people avidly believe in it. And there are some people who are highly skeptical. I'm not sure where I stand at this stage about that.

But now, people began to study our shape, the ability to see shapes by recording various visual areas. And I want to tell you next about some studies that had been done primarily in Japan looking at inferotemporal cortex, which is [INAUDIBLE] already mentioned that to be involved in the analysis of shapes and in particular, in the analysis of faces as well.

So what these investigators did is they would record to individual neurons in inferotemporal cortex. And when they did that, they would present various stimuli to see how those cells responded. And they found, they claimed to find, that there was some incredible specificity in inferotemporal cortex for shapes.

And so what they did, here's an example. Here, we have the neurons' responses, histograms. And on top, above it, we have the particular shape that was presented repeatedly. So this particular cell responded vigorously to this, but very poorly to that, and so on down the line.

But you can see that it responded to several different shapes, quite well to this one, reasonably well to that one, so a whole range of them. So it wasn't like this particular neuron responded only to one particular shape. But that idea that you have neurons which are specific for certain shapes did take on a lot of attraction on part of investigators. And some of them indeed thought that you have these neurons which are selective to individual elements in the visual scene.

And those subsequently the critics refer to as your having grandmother cells. Somewhere in the brain there's a cell that represents your grandmother, all right. Now that idea was disabuse subsequently. But it took very stronghold in many investigators' minds.

And here's another example of yet another inferotemporal neurons, in this case again for various shapes, showing that this shape was, elicited a lot of responses,

as well as did this. The rest of them didn't respond, didn't elicit as much of a response. So these kinds of experiments then, they tried to systematize. And they came up with an idea which, at least to my mind, borders on the absurd, which is shown here.

Here is a section of inferotemporal cortex. And they argue that they are columns there, and that these columns represent different percepts. So here we have a bunch of percepts that process the monkey. This is a monkey, of course, that process monkey faces. And by inference, therefore, there must be some areas in inferotemporal cortex that process human faces in humans. And then others process different aspects of the visual scene.

Now the big problem, technical problem, with this approach is that when you record from individual neurons, all right, typically in these experiments, you can study a single neuron for a fairly limited time, not like for months on end, just maybe a few hours or something.

And so because of that, you can only present a limited number of visual stimuli. Now there are millions of visual stimuli out there. And so to really establish how specific this particular neuron is, they say well, these are the only shapes it's showing us. What if you use the cross, or who knows, many, many other things, something that is three dimensional? How would these cells respond?

And so these cells respond to different degrees, to hundreds and hundreds and hundreds of different stimuli. And the real fact then is that anytime a stimulus appears out there, you're activating tens of thousands of neurons, maybe even more, in the visual system. And each of those neurons, especially in inferotemporal cortex, responds to different degrees to the different stimuli.

It is the compendium of these many, many neurons firing into different degrees to different stimuli that gives you sort of an overall computational ability to say what that stimulus is. Now to be able to analyze that, it is that complicated, that really takes a lot of effort.

Now some people are now trying to do this. And the way you try to do this is that you're recording from hundreds and hundreds of neurons with multiple electrodes, present various scenes. And you see how these neurons respond as an aggregate, so that you can determine whether or not there is indeed the potential for some sort of computation that takes place in gaining the impression of a particular individual face.

Now one of the very important facts here is, of course, that as you look around the visual scene, let's imagine you're looking at a face. You're looking at a face head on. You're looking at a face in a profile. You're looking at a face tilted. You're looking at a face close. You're looking at a face far. And it's still all the same face.

And yet the impressions that face makes on the visual system, in the retina, in the geniculate, in the visual cortex, varies a great deal. And yet we can come up with constancy, which is sort of a higher level process. And I'll come to that in short order.

All right, so now, therefore, we need to start to talk about what we call intermediate level vision. What is intermediate level vision? So far, we've talked almost exclusively about basic visual capacities, color, brightness, pattern, texture, motion, depth, the very, just the very basic types.

But now when we talk about intermediate visual capacities, we talk about constancy. One example of constancy, of course, is that the face that you're looking at, whether it's profile, or head on, or near, or far, it's still the same face. So we get constancy out of it. Then there's an important necessity to be able to select various aspects out of the visual scene, to be able to select if you recognize things, to induce transposition and variance, and to be able to make comparisons, and also lastly, to be able to say where things are in space.

So those are these intermediate capacities. And we'll talk a little bit about them next. All right, so here's an example of constancy. What we have here is a bunch of words for doubt in different sizes, and orientations, and some handwritten, and so on. And yet all of them say the same thing, doubt. We extract that, OK.

Now to make this even more difficult, I'm going to tell you that there's one of these words here which is not doubt and see who can find the word that is not doubt. Everybody find it yet? OK, right there, not doubt but doubts. And so we are able to extract the difference in this kind of visual scene, even though it is incredibly subtle.

And yet, we are also able to say that this doubt, and this doubt, and this doubt are all the same, even though they're different size or different in print. We can extract from that the common element. So this is an incredible capacity on part of the visual system. And of course, the big question comes up, how is this achieved? And at this stage, we only have very preliminary answers to that.

OK, so now I want to show you a couple other cute things here. This is another famous artist, Hirschfeld, He is no longer with us, unfortunately. He published hundreds of these little cartoons in *The New Yorker*, all right. Now, you see his name here, Hirschfeld. What do you see at the end of his name? What's that-- what is this? What, that? Huh? A number three.

Why does it say Hirschfeld three? His name isn't Hirschfeld three. And then if you've seen several of these, how many of you have seen these before? Oh, you missed out on something very good. At least look it up sometimes again on the internet. Just type in Hirschfeld. And some of these will come up.

So he says three. And some of his pictures may have one. Some of his pictures may have five or six at the end. And you say, well, what on earth is that? Anybody know? Yes.

**AUDIENCE:** Doesn't he have other names hidden in the images? So maybe that's the [INAUDIBLE].

**PROFESSOR:** Ah ha, you're thinking well. OK, that's getting there. OK, so he's telling us something that there are three of in this picture. That's what he's telling us. OK. So what is the three of that he has in this picture? Ah ha, anybody want to come up with? It

**AUDIENCE:** The line orientation?

**PROFESSOR:** Huh?

**AUDIENCE:** Like the orientation of the line?

**PROFESSOR:** Orientation? No. Mhm. OK, I'll give you a clue. Hirschfeld had a daughter--

**AUDIENCE:** Nina.

**PROFESSOR:** Who he was very fond-- hey, fantastic, Nina. OK. So had a daughter, Nina. And he decided in his display to put her name up there. And now he said three, that he put her name up there three times. Now who can find all three of them?

**AUDIENCE:** [INAUDIBLE] the arm [INAUDIBLE].

**PROFESSOR:** OK, here's one. Here is another one. OK, and here is the third one, in the arms. And so if you, that's one of the fun things you can do. Back when he was doing this actively, every time he had a cartoon like this, I would spend some time trying to see where were the Ninas.

[LAUGHTER]

So that I think is quite remarkable and further highlights the complexities and the remarkability of our visual system. OK, so now, I'm showing you yet another picture. In this case again, you have to be knowledgeable about something. Does anybody know who this is? Oh, all right.

Well, let me see if I make it smaller. Can you see that now? Who is that? OK, that's Voltaire. All right. Now Voltaire, 18th century, is a guy who had outdone all of us. And he had written 2,000 books. Can you believe that? The guy had written 2,000 books.

But that's not the prime point here. Most of us who know a little bit about history immediately recognize him as, oh yeah, that's Voltaire. But then if you go back here and you look at this more closely, and so what I'm going to do here is just to tell you what a remarkable artist. The guy who did this is Salvador Dali. This, by the way, is his wife, OK.

And so what I'm going to do here, I'm going to blow up the center portion of this figure. There it is. Can you see this? This is two faces, two nuns. And this is their outfit. So there's nothing there that is really Voltaire. But he's such a remarkable artist that he could create Voltaire by playing around with this, OK.

So let me go back to the display again. OK, so this is Salvador Dali. And certainly, it's an artist who is remarkable. He has many paintings in which this kind of a double confusing thing, playing on your ability to see. So because of that, for people like me who's interested in how we see things, certainly I'm very intrigued by his artwork. And you may enjoy looking that up again on the internet.

OK, so now I'm going to tell you yet another interesting aspect of art involved in this. And this one has to do with a book, a very clever book that was written by David Hockney. Has anybody ever heard of him? I'm not surprised since he's not that well known. But the book he wrote is called *Secret Knowledge*.

Now what was that all about? Well, he analyzed how artists created paintings way, way, way back when and in subsequent years up to the present. And now I'm going to show you some of his pictures and give you a sense of what this is all about.

Here is a picture, OK, by Masolino da Panicale in 1425. This was a typical kind of picture back then. Artists had a very poor sense of how to create depth perception in a painting. That was before they came up with vanishing point. And so this kind of looks flat and fairly expressionless. This was in 1425.

And then Hockney noticed that just five years later, another artist came up with a picture, just five years later, that looks almost like a photograph. And he said, what on earth has happened? And that's why the book is called the *Secret Knowledge*. So what happened? Anybody know?

Ah, all right. Well, let me tell you what happened. What happened is that at that time, in the 1400s, they came up with the lens, OK. And so they created a device called the camera obscura. Does anybody know what the camera obscura is? It's essentially very similar to a camera, OK.

So here we have it. Here's a camera obscura. And what they did, these artists, they would create a building. So it will be dark inside. And then they would put the lens here. And they would put some object out here that they wanted to make a painting of. And that would be reflected onto a piece of canvas here.

Of course, it would be upside down, right? And then they would paint it. And so once it was painted, they could turn it around and finish it and sell it. Now, the reason they did this, the prime reason they did this, is because it was much much, much quicker to create a portrait, for example, by using this procedure than to actually look at a person and paint them on the canvas while looking at it. OK.

So that was done. And of course, this was sort of a no, no thing. And because of that, it was kept secret. All these artists who did this, and some very famous artists did so, were very careful never to disclose to the public that they did this kind of thing. Because it was conceived to be kind of a cheating thing, OK.

So what happened then was that all these paintings were created. And here's an example by van Eyck in 1436. Again, this looks much like a photograph. But some other thing is a bit distorted. So Hockney undertook a careful, detailed analysis of how we could tell whether a painting was a real painting or one that used the camera obscura method. And that's also a real painting in a sense, but a painting using the camera obscura method.

And he came up with a series of criteria which are listed in the book. But I'm only going to deal with one of them. OK, so here is an example. This is in 1597. This is by Caravaggio, OK. And what is notable about this person?

Well, to Hockney, what was notable is that this person is holding the wine glass in his left hand. Yeah? And he said, huh. That's curious. And then he looked at a whole bunch of other paintings. And he had one in which there were three people on the painting. And all three of them were left handed. Yeah. And he said, my god, that is really curious.

And he said, well, let me analyze what happens when you do this kind of stuff with the camera obscura method. So here's the example of this. This is the original image, he claims, meaning the person is right handed, is not left handed. Then you put this person through the lens and put him up here upside down like that, OK.

He's painted on the canvas like that. And then what you do is you rotate this 180 degrees. And when you do that, lo and behold, the person becomes left hand. And to make this clear, I added the F here. So this is a normal F. This is what is projected onto the canvas. And this is when you rotate it 180 degrees.

So you reverse the left according to Hockney, reverse the image. And so it was a dead giveaway that most of the people who appeared left handed in his paintings, in these paintings, I should say, not just Caravaggio, but several other people, used the camera obscura method for creating the painting.

All right, so that is the process. And then I want to show you one more example of this. OK, this is a very famous painting also, all right. This is the so-called marriage of Giovanni Arnolfini. And this by van Eyck, again 1493, way, way, way, way, way, way back then.

Now this is a famous painting. Now there's one interesting thing first of all I want to point out. You see this here? That's proof that in those days they had come up with the lens, yeah. Now what is wrong with this picture? This guy who is about to marry this woman is holding her with his left hand. I mean, that's unacceptable. You're supposed to hold it with the right hand, yeah.

Now the reason he's holding it with the left hand is because the artist, van Eyck, used the camera, according to Hockney, used the camera obscura method to take this, to paint this picture, OK. And then he rotated 180 degrees and became, he became left handed as a result.

Now if instead of having done that, you would go back, you don't have to go back. Go back to here, and let's go back to this. If you take this guy here and you create the picture. But then instead of rotating it, if you could flip it, then you would, he will

remain right handed.

But of course, we can't do that because it's on a canvas. It's not on a transparency. So that then is the interesting story of artwork that was created using the camera obscura system that further highlights the amazing interestingly complicated manner in which we can analyze the visual scene for shapes.

All right, now another factor that is in a similar vein has to do with the recognition of faces. Lots of experiments have been done, including several in our department here, that has recognized that, unfortunate use of terms, recognition, recognize, became aware of the fact that facial recognition depends very heavily on seeing faces right side up.

When faces are upside down, you have great difficulty telling who is who, OK. So let's do that. Here are a bunch of faces. And I bet you can tell this one, right? Who is it? And who is this? Very good. And those two you don't know. The problem is that I am going to flip this over now. And you still don't know who those two are.

OK, so this one is Norbert Wiener. Now all of you know about Norbert Wiener. He is one of the great geniuses of our time. He came up with the digital code. He used to be a professor at MIT. And this here is Chuck Vest. Anybody recognize him? Chuck Vest was a president of MIT for what? For eight years, I believe, maybe more, 12 years? I forget. And he has sunk into obscurity, even though he was incredibly visible for many, many years.

If I had shown this to you say eight or 10 years ago, you would immediately recognize him. Because he, at that time, was a president at MIT. So that then is an interesting fact that even though we are capable of using intermediate level vision in a very sophisticated way, it does not seem to work that well for upside down faces. You guys did pretty good with those. But it takes a while.

If I had flashed that on in a tachistoscope, you wouldn't have had a vague idea of who those people were. But once it's on for a while and you can analyze it, you can

eventually tell even an upside down face. OK, now in the same vein as we are talking about, our ability to process shapes based on contours and whatnot, one of the interesting set of experiments people had done is to look at what is called subjective contours.

And so let me give you a couple of examples of that. Here's an example. Almost instantly, here you can see a disk. And here you can see a square rotated 45 degrees, a diamond if you will. But if you analyze this carefully, you can see that about 80% of this border here is not a border. There's no border here. There's no border here, here, here, here, here. And yet we can see a square.

So there's some strange ability in part of the visual system to complete inferred borders. All right, I'll come back to that in just a minute. Now another example of this is shown here. Can you make out what it says here? If you look a little bit, you should be able to see. What is it? Visual system, very good. It was difficult to see. But as soon as I turn this into color, you have no trouble at all.

And that highlights another important reason why color vision is so useful and has evolved. Because it enables us to see borders, where under certain conditions and lighting conditions, borders would not be visible on the black and white. OK, so now that's another interesting example, has to do with further subjective contours.

And here's an example of one with a high contrast, where you can readily see a cube. Does everybody see a cube there? All right, but if you look here, it's next to impossible to see that. Because here the stimuli are isoluminant. So here you eliminated contrast, which is a very important aspect of being able to analyze the visual scene in various ways, including three dimensions.

All right, so now here's a very interesting discovery that was made by recording an area, V2. The recordings were then made to see whether those neurons perceive subjective contours. So here's a receptive field. And you take this bar and move it back and forth across it. And you can see it gives a vigorous response.

Now you do the same thing. But you don't, here you have only a subjective contour.

And then you move this back and forth across. Somehow the information is added up from other areas, so that this cell responds, not that well, but responds reasonably well to the subjective contour. So it's said that in V2, you can carry out some of these higher level processes that enables you to complete figures even when they are incomplete like that.

OK, here's another example of that. This is even more dramatic. In this case, we take this bar back and forth across. That's a vigorous response. And then you do the same thing. You create a bar here simply by these continuous horizontal bars. And still, the cell responds quite well, as if it were seeing an edge here.

So we have this kind of completion in area V2. So this is sort of an initial hint then that already in area V2, you begin to process higher level events, among which is the fact that you can complete incomplete contours. All right, so now, the next thing we're going to turn to, we're going to ask the question, when we deal with these so-called intermediate level visual capacities, what happens when you take out such areas as V4 and MT to your ability to see these intermediate visual capacities?

So let me describe to you some of these, how one would do experiments like this. First of all, monkeys are trained. This is done with monkeys, of course, because we can't just take a human and remove V4 MT. And so what you do here is you first present a fixation spot.

Once the monkey fixates, you present a shape, in this case a square. And then you present a whole bunch of other ones, only one of which is the same as the original. And the monkey has to make a [INAUDIBLE] there to be rewarded. So he has to be able to detect identity, which is an intermediate visual capacity. So that's what this is like.

And so then what you can do here is to vary the amount of information you can provide, again by reducing the amount of contour information that you can provide. You can do this in several ways. First of all, before we do that, I'll come to that in a second, let's just see, how does the monkey do when you do those very similar shapes I just have shown you after you take out area V4?

What you find here is that the monkey initially, after you take out the area V4, can't even do the regular task, which is this one, right? So this is identical to this one in this case. And when you do that, he does very poorly to begin with and then gradually, over many, many days, improves a great deal. Then if you put a new figure in, then it takes a while for it to learn that, all right.

Now then, what you can do, now I come to this business of having these same figures. But you can vary them now, so that you have to do a transposition to do an intermediate visual task. In this case, what you do is that you vary the size. In this case, it's identical just like it was.

In this case, you see that this is smaller than this. But you say, oh, I have to go find the circle. I'm not looking for identity. I'm looking for something that's the same looking. And here we have even bigger case. This is a triangle. And so the monkey makes a [INAUDIBLE] to that one here to this one. So that has to do a test position in size.

Another thing you can do that we talked about, you can vary the amount of contour information by doing this kind of occlusion. And the degree of occlusion you can vary by varying the spatial frequency of the display. And lastly, you can also decrease or increase the amount of contour information that you provide.

Now if you do with this, you get a huge effect after V4 lesion. This was the normal condition with the varied object size. This is the occlusion. And here is the varied contour information. You can see that there is quite a dramatic loss, not a total, but quite a notable loss, in the monkey's ability to perform this task.

And this is also reflected in the huge increase in the latencies with which the monkey can perform the task. So area V4 seems to play an important role in these intermediate visual capacities. And I come to some more examples of that in just a minute.

Now, yet another important factor in analyzing the visual scene occurs when it is our task to find something out there that is less noticeable. Remember, we talked a little

bit about using camouflage, in which case, you have to find something lesser to survive.

Now in this case, what we can do here, do a similar experiment in a monkey. On the left side, the target, the one the monkey's supposed to select, has a much higher contrast than the distractors. And you can vary the degree of difference, but always the target is brighter than the others, so that it stands out.

But then, you must be equally able, whether you're an animal or a human, to be able to pull out something that's lesser in the visual field. And you have to be able to do this if you are going to survive in your environment. Now here then, the task is to go to this lesser stimulus, because it's the odd stimulus.

So what you're extracting is the odd stimulus. Here it's easy to extract, and here it's difficult to extract. And so now the question is what happens in the visual cortex? What area plays a role in this?

And so it was discovered that area V4 is very important for this. And let me explain this to you then. You do an experiment in which you remove area V4 and see what happens. And here what we have is when the star gets brighter, you vary the luminance difference.

And you can see there's a mild deficit with the V4 lesion, highly significant but still fairly mild. On the other hand, when you make it dimmer, the monkey is practically staying at the probability. He cannot do the task at all. So somehow, V4 plays a very important role in being able to ferret out some subtle things in the environment, lesser things.

All right, so this then brings us to yet another way of analyzing this. We can, in this case, go to the larger target, in this case to the smaller target. And you can recognize yourself that this is certainly easier than that. But a normal monkey can do both of these quite well, shown here.

This is the normal monkey's performance, the target larger, here the target smaller. He does extremely well. But with a V4 lesion when the target is smaller, it's totally

devastated, so as if area V4 were involved in the analysis of subtle things, things which are lesser than, rather than being reflex like and going to the brightest, biggest thing in the world. OK, so that then indeed highlights the fact that we have some of these areas, including V4 and of course, therefore MT, that involved in these much more subtle types of visual analyses that we need to perform.

Now capitalizing on these kinds of subtle things that we are capable of doing, artists, in addition to the ones that I've shown you before, also created all kinds of percepts, I should say paintings, sketches, precepts, that cause confusion by playing around with these factors.

One of these here is a very well known audience called Escher. This was so near the end of the 19th century. And here what we have is this is sort of a figure-ground confusion. Here what we have is a bunch of birds that fly to the right, and also a bunch of birds that fly to the left.

And so it's confusing. It's alternating. You don't know which one is which. And it has to do with a very, very clever creation of figure-ground confusions. And actually, next time we talk about illusions, I will bring in some more of these kinds of curious effects.

And then here's another one. And it's very hard for you to tell are the stairs going up? Are they going down? What's going on? It's the same kind of play with the paintings to create confusion in your perceptions. And then here is yet another one, where you don't know is water running up, or is the water running down here? That's again an interesting confusion that Escher has created.

All right, so that then is the essence of what I wanted to cover today to highlight the fact that our ability to extract shape information in the world is absolutely incredible. And it has triggered not only experiments to try to understand that by scientists, but it has created a tremendous amount of artwork that played with these because it was so enjoyable. And certainly, the several paintings that I have shown you must have given you a pretty good sense of how artists capitalize on these limitations in ability to extract intermediate and high level vision capacities from the visual scene.

All right, so to summarize them, first of all, I mentioned three major theories that have to do with form processing. The one, the orientation of line segments. And I pointed out to you that that theory is not particularly powerful. And then the one, the topographic theory, which turned out to be bordering on the ridiculous. And finally, Fourier analysis, which seems to have a lot of power, even though it also has created a lot of skeptics.

All right, then I pointed out to that are the V2, V4, and inferotemporal cortex, play important roles in intermediate vision, these intermediate visual capacities we talked about in detail. Then I pointed out to you also that in V2, it was discovered that there are some neurons that respond to subjective contours, indicating that already in area V2, we can perform these incredibly higher level abilities to extract information when it is unclear in the visual scene.

Then I pointed out to you that recognition of objects transformed in the various ways is compromised by V4 and inferotemporal lesions. And V4 lesions also produce major deficits in learning and selecting lesser stimuli, which is a very important attribute for us to be able not to respond in reflex-like fashion to what is most obvious out there, but to be able to extract the subtle things from the visual scene.

Some inferotemporal neurons are selective for objects including faces. But most respond to a variety of objects whose recognition is based on the differential activity of a great many neurons. OK, so that then brings me to the last point, which is how we process and deal with ambiguities in perception, unfortunately still remains a mystery.

And so there's a lot of space here for new investigators to come up with exciting, interesting new findings about how the brain performs these kinds of subtle analyses of the visual scene. OK, I think I'll leave this until next time. I'll talk about this next time.

OK, does anybody have any questions? All right, I hope that you can take a little bit of time out, look at some of these artists on the internet, and look at Hirschfeld and

the Ninas that I had shown you.