# Virtual Reality Engineering Prof. Steve Lavalle Department of Multidisciplinary Indian Institute of Technology, Madras

# Lecture – 11-1 Human Vision (motion perception)

I want to get into another (Refer Time: 00:16) another perception topic called motion perception. This is obviously very important for virtual reality, you can of course, look at static panoramas in virtual reality, but our perception of motion is really fundamental to the kinds of things that we that we do all over the place in virtual reality.

(Refer Slide Time: 00:35)



So, that is the next big topic motion perception. Remember as I said last time, it is kind of a template, you know there is x perception, where x may be all sorts of things depth scale, motion, color on and on and on. So, so we may also apply the same concept to hearing or touch. So, perception of different kinds of textures for examples, this is um, again as we go through these pathways from the sense organ, all the way up to the higher parts of the brain, there is some hierarchical processing going on and there is a perceptual process in there. And that is what we are trying to get an understanding of and trying to characterize. So, what are the purposes for motion perception, some of the main purposes?

Well of one of them is segmentation or segregation. In other words, we might like to know if we are looking at a scene in the in the jungle, let us say we might like to know that there is an animal that is moving there, all right. So, so the perception of that motion is very important, and you might imagine for evolutionary reasons it is very important, right, to know if there is something moving in the jungle, right or in the forest that is up that is about to I do not know potentially eat you let us say right. So, so, it is very important to identify that.

So, perceiving that motion among a static background is very important. And you might notice that, our eyes tend to fixate very quickly on something that is moving. In modern times, our ability to detect a moving car, or cycle or some other some other vehicle that may be a dangerous to us as we cross the road is also important right. So, as a similar kind of purpose, we are very good at extracting some object that is moving from a back-ground at stationary.

Um another purpose is to extract 3D structure, 3D structure of an object from the motion. So, for example, you know if I if I see this book moving around in some way, I get some idea about it is structure, this book does not have a very complex structure. But take something you know much more complicated like a chair with legs and such move it around and you get an idea of it is structure just from the way it is moving around on your retina, right.

There is some 2 dimensional image that is changing on your retina, but you can infer 3Dimensional structure you can learn a lot about that it is almost as good as if you walked up and started probing the object with your arms and use all of that information to try to build a 3D picture. So, we are very good at building 3D pictures, somehow in our minds of objects that we have not maybe ever seen exactly before, but we are using a lot of prior information again prior bias.

# (Refer Slide Time: 03:25)



And from the context, we have a lot of information about how the lights falling on it in a certain setting, that we are familiar with and we make judgments about it is 3Dimensional shape.

You can make illusions and fool the brain in some in some cases and get some false conclusions, but nevertheless we rely on this very carefully very frequently in a large number of settings. And finally, another key purpose is visual guidance of action, and this is exactly part of where I wrote on scale perception here one of the one of the important aspects reaching out and grabbing something. So, if we want to reach out and grab something, then we are using motion and we are using motion perception we are perceiving, the motion of our hand as we reach and grab something and that visual feedback helps us to accomplish the task. So, visual guidance of our actions so, that is useful for manipulation like grab a cup. general hand eye coordination, and may provide information about you know other kinds of self-motion.

Questions about that I am cover a couple more topics and then I am want get to one of the main aspects of this perception of motion which is going to be frames per second I really would like to talk about that that is; obviously, something that is on our brains a lot we talked about this in connection to the lab, the virtual reality lab and we would like to know how many frames a second are enough.

#### (Refer Slide Time: 05:10)



But I want to give a few more basic concepts first. So, one thing that I find fascinating is if we look at the neural circuitry; remember, we talked about the visual pathways from photoreceptors all the way up to the visual cortex. So, along that path and actually very early in the path, we have the following kind of circuitry so, suppose there is an object that is going to go moving across the retina or some object it goes moving across the retina. And imagine that there are neural detectors let us say. So, somewhere back here there is a detector A and A detector B. So, imagine this neuron fires whenever the object moves across this location, and as the object moves across the location here this one fires.

So, which one is going to fire first well first a fires then B fires what is interesting is there is extra neural circuitry that looks like the following, it goes through here I was a box that I am going to describe in a minute, and then these feet together to some third a neural structure called, let us say C, and inside of this box is a delay say some delta t perhaps it is 50 milliseconds. So, what happens here is this C fires if both A and B fire, but only if a fired a little bit ago. So, a fires a little while ago and B is firing right now, the superposition of those in time right a little while ago and then B right now will fire C. So, that is a very, very simple motion detector. So, we have those all over the place for varying speeds and various distances of neighboring distances. And so, one interesting outcome of that is that you are able to fool this sense by something called the wagon wheel effect.

### (Refer Slide Time: 07:01)



Which if you have ever made some kind of pattern like the following, let us take a we put some stripes around the rim of a wheel. And then you spin the wheel it may start to look like it is rotating backwards, even though it is going forward. So, this is interfering with the basic operation of this neural circuitry it is easy to fool it into thinking that it is going in the opposite direction. Because it is just based on very simple timings and the flashings of neighboring, let us say the kind of in some sense the flashing or signal appearing in various spots along your retina and those photoreceptors responding. So, so you can very easily fool it into thinking reverse directions. Unless, there is a more coherent picture of moving from place to place to place to place along a very long sequence.

# (Refer Slide Time: 08:06)



So, this is very interesting I think also, one thing to think about is object motion versus observer motion. This is one of the key ingredients to avoiding simulator sickness if you are not careful. So, if I just were to look at the images, if I could somehow make a recording of the images that hit your array of photoreceptors, right? And we imagine we are just going to play a video back, now of what is hitting your photoreceptors.

I might see some kind of motion and occurring there, but how do I know if that motion is due to you moving, or some object in the scene moving, could be some combination of the 2, but let us suppose it is one or the other, how would I distinguish between the 2 if I were just watching the video I do not think there is any way to fully distinguish that right unless there is one maybe you would start to guess from the context what is probably happening, and use your powerful brains to make an inference. But if you just look at the raw data there may not be enough there to really make the judgement let me give a simple example.

Let us suppose they have an eye looking upward, all right? So, I will draw the cornea sticking out here, and I have an object that moves from right to left. So, it is moving along like this; so, if I draw out the retina of the eye, I guess when the objects over here on the right, then there is an image of it, over here on this part of the retina. And then when it moves across it ends up over on this part of the retina. I am not drawing too careful of a diagram here, but I think you will understand the principle getting really curvy

here. So, as the object moves from right to left, imagine the eye is held fixed, and then the image goes from left to right makes sense all right. So, that is one possibility this is the object moves, and the observer and eye are fixed, all right. So, the observers head and the eye let us say are fixed, and then there is another case which in which I will hold the object in one place, and then I have an eye here and I move the eye over here.

(Refer Slide Time: 10:42)



So, when I start off on this side, can I have the retina coming around ? Draw it again over here, should be the same I just moving across from left to right. And I start off over here I get the image on this side. And want to come over to this side I get the image on the other side, way off there I see all right. So, in this case, the eye moves and the objects fixed. So, these 2 cases, I just want to say if you imagine how the image looks along the retina, they should more or less correspond to the same thing.

Now one thing I have I have done here I should admit this is a bit fictitious of a scenario, because it is actually very hard to do this in practice. So, if I if I want to you know try to look at an object let us say that is moving across my field of view, it is very hard to prevent your eyes from tracking it and doing smooth pursuit it is very hard for you to disable your vestibule ocular reflex if you are moving around while you are fixated on something so in fact, if I go back and forth like this I am looking at the camera right, now it is very hard for me to keep from continuing to look at the camera, it is like I have to put my fingers on the sides of my eyes like this and then I can keep it from rotating that

is about the only thing I can do not try that at home I guess maybe that is hazardous to your health.

But so, it is very hard to stop these reflexes. So, given if that is the case given that your eyes are going to be rotating anyway. That is going to make these 2 cases even look very similar to the case of no movement at, all right and because your eye is rotating to keep a stable image on the retina. So, I just want to point out all these cases like even if your eyes are not rotating these 2 are somewhat indistinguishable; however, when your eyes do rotate, in one case if you are doing this will invoke the vestibule ocular reflex, which you are moving your head back and forth, you can also do with pure rotation which is also translating your eyes through space, and then this corresponds to smooth pursuit if this object is moving slowly enough. So, those are different modes of operation, which is part of the key of how the brain is going to distinguish between these.

So, let me give you how the brain distinguishes between these cases. Let me just list out the information that is being used it is not just looking at a raw video signal; let us say from the array of photoreceptors. It is got more sources of information coming together.

(\*

(Refer Slide Time: 13:50)

So, the brain uses more information to distinguish. We talked last time during (Refer Time: 14:24) motions there is saccadic masking or suppression, which when this happens, it suppresses the motion detectors. So, while we are performing saccade, it is that the motion detectors that, I just erased one of them right that when I did the a B and C

neural circuit those outputs are suppressed. So, already it is not going to be feeding motion detection information to your brain, if your brain knows that you are performing saccade so, very interesting. So, there is already some inhibiting that occurs there another thing that the brain uses is eye movement commands.

So, that is very interesting the somewhere there is a command issued to let us say commands issued to contract or expand the muscles that move your eyes and rotate them right. So, if that is the case it is a matter of sending those signals to the right place where perception is being performed or the motion perception is being performed and using that information. This is sometimes called by neuroscientist efference copies by the motion commands or motor commands very powerful information, right?

The same thing happens in engineering if I give a command to a robot that says move forward. And then I observe what the robots doing you can make a lot better inference if you have access to the command oh I know that the robot was commanded to go forward, but if you program the robot yourself you have that information, if you did not program the robot you do not know what is inside of it then you do not have that information, right. So, so this is extra information that is inside the brain, throughout our neural system let us say and I can be used here.

And then finally, one more piece of information is large scale motion. So, if the entire scene is moving what is usually the case, is it usually the case that everything around us is moving simultaneously is usually the case that in that event it usually means that we are moving right. So, if I if I almost sudden notice a huge change, I go like this, right it is very unlikely all of you are rotating, and in some kind of role all right. So, that is prior information if there is a enormous large-scale motion across the retina, probably you are the one that is moving.

And so, that is a falling back you might remember, I talked about the big swing illusion, where in an amusement park over 100 years ago, there were a bunch of participants sat down in a swing that was actually stationary, and in the entire room started rocking. That is such an unusual scenario, that would correspond to you know it corresponds to in reality everything is moving around you, but your brain interprets that is I must be moving and everything else is stationary. And that is what made people very sick and quickly believing that the chair that these chair that are in a spinning all the way around that the swing goes all the way around 360 degrees.

Because it was tricking this information basically, giving you an artificial scenario where the large-scale motion assumption is in fact incorrect yeah.

Student: (Refer Time: 17:46) what cases (Refer Time: 17:51).

Um let us see, if we are I am saccadic masking occurring. So, let us see if if we are holding stationary, and we are looking at an object. Well, let us make nothing to find the scenario first. So, is it is it let suppose there is a stationary object. And we are looking around the object, right, like reading a book for example, in that case there is a lot of change across the retina, but the motion detection is suppressed, right? There is another scenario where let us say the object is moving slowly, in that case there is no saccade it is smooth. Pursuit there is a scenario where there is the objects moving very fast, and some additional (Refer Time: 18:30) are being inserted in order to keep up. But I think at that point there is already been enough motion detected I believe from the smooth pursuit times in between the (Refer Time: 18:38) I am no expert on neuroscience time trying to speculate as best as possible for those different scenarios.

Student: (Refer Time: 18:43) the both cases identical except the last (Refer Time: 48).

Um no, I think the I think the saccade motions we do, most often the seccade motions that we are doing another, normal saccade mode is to point the fovea or oriental fovea. So, that it is so, that you are getting higher acuity images of a single fixed target. So, those particular motions are interpreted because of the saccadic masking as that must be motion of the eye just in that mode all right. So, in the case of smooth pursuit it is usually not the case that you are doing extra saccade to try to take in all the information of the object by pointing the fovea, while it is also moving usually in the case of smooth pursuit is trying to keep a stable image on the retina. So, it is not a saccade case hm.

Student: (Refer Time: 19:35) that can also be a information (Refer Time: 19:41).

That is a very good question, yeah, I think that is I think that is reasonable, why not add the vestibular sense here? I think that is reasonable very good yeah absolutely if the head is moving there is going to be additional vestibular information. If the head is fixed then it uses eye movement commands. So, I like that very much yeah. So, vestibular input if head moves. Or I guess your entire body could be moving without your head moving if you are on some kind of moving carts, let us say in a car for example, then the head itself it is still technically moving with respect to the world frame, but may not move with respect to your collarbone for example. All right? Very good any other questions or comments ? Yeah.

Student: (Refer Time: 10:37).

Yes.

Student: (Refer Time: 20:41).

Smooth pursuit versus vestibule ocular; well, the vestibule ocular is using your vestibular signal.

Student: Yeah.

Which I guess that is where it got used for one of the cases right. So, it is it is using exactly your self-motion in order to compensate. The smooth pursuit case is not your emotion, it is the motion of the object you are trying to see right. And so, it is based purely on expectation of the motion of the object, and trying to keep that stable across your retina, right? Both are trying to do some kind of let us say stable image on your retina, right? These are fundamental differences and whether or not the vestibular signals getting used and evoked because of the motion, or you are just trying to track a moving target, and that is where the difference is are I am just trying to say that kind of geometrically in terms of just raw images these look identical. But there is so much more information going being obtained by the brain that it can distinguish these 2 cases.

Student: That is seems like without there are extra information (Refer Time: 21:40).

That is what I want you to think, yes, that is right. Like, wow, it does not have you know if just from images. So, if we were to think like engineers and we just turn on a video signal, and look at it well how does the brain know right. So, I am saying yes, the brain knows there is a lot more information here, all right. Very good these are very good comments very good ways to think about things there is one more piece of one more piece of information that is very naturally extracted from the images that change across your retina and this is called optical flow this is this ends up being very important and it is one of the most important aspects in relation to motion sickness or simulator sickness, I should say not necessarily motion sickness.

Because the VR participant is not necessarily the one moving, all right? Usually we say motion sickness for something you might get from moving in a vehicle, let us say moving in a car while being a passenger especially that is motion sickness, simulator sickness would refer to what you get from a virtual reality experience for example, where you may perceive your own motion, but you are actually not moving. So, those 2 are kind of in versus motion sickness versus simulator sickness.

(Refer Slide Time: 22:51)



So, optical flow ; so, what happens is in this case? The brain can keep track of images on the retina or features let us say on the retina. So, there may be some particular feature, I see a red dot above the camera there. So, I can I can move around in some way, and that dot is going to appear somewhere on the retina and I can keep track of that, right? And if there are several dots, then if I get closer to them the dots will start spreading apart, if I go further away they will start contracting if I go to the left then will move to the right. So, this is called optical flow information, where features are moving in the image. So, our brain perceives them to be the exact same object or element in the world there are a bunch of them and we perceive a kind of general flow because of it.

So, it is tracking movement of features on the retina, and we end up with a kind of vector field, kind of velocity field on the image plane or roughly the image sphere if you think about the way the retinas shaped it is fine.

(Refer Slide Time: 24:27)

So, if we think about various self-motions, as I just illustrated, if I am going forward, then the flows tend to be outward from the center, correct? Also, the flows tend to be faster at the periphery, then in then straight in front of me. So, not only are there directions in the optical flow, but there is also a magnitude all right. So, these are these are full full-fledged velocities in the image plane or image sphere if you would like to view it that way. If we go backward, then these reverse, and if we go to the left, then we see flow to the right. So, when I move myself to the left, I see you flowing to the right. Also, if I turn myself counter clockwise, I also see flowing to the right. So, I see also counter clockwise.

# (Refer Slide Time: 25:47)



Of course, I will get flow in the opposite direction by going right or a clockwise rotation, our ability to perceive optical flow turns out to be strongest as I look forward right about here. I think, somewhere around 20 to 30 degrees off; this is the place where you have a very high density of rods if you remember when we looked at photoreceptor density. And generally, it is very strong in the periphery, right? So, that means, that if there is some kind of motion in the periphery we respond very quickly to that turn our heads could be a predator right maybe some evolutionary reasons for that. So, just generally speaking it is stronger than periphery I believe it peaks around 30 degrees or so, and then it starts trailing off again because of the low density of photoreceptors that we have as we gradually go to the side gradually look to the side.

# (Refer Slide Time: 26:50)

(\*

So, based on this optical flow, we have a big problem in virtual reality, a big problem, you have experienced in the lab already I am sure; which is the illusion of self-motion from optical flow. This is called vection. I have said this before ah, if you are sitting stationary in a train and the train next to you starts moving you are looking out the window maybe you are not even paying a lot of attention. In fact, it is your peripheral vision that notices a that there is some motion to the side, you get a feeling like you are going backwards all of a sudden right.

So, it is the illusion of self-motion, when you grab on to a controller or you hit the right keys on the keyboard and you move your avatar around in virtual reality, you perceive an optical flow that gives you the illusion of self-motion, every time you do that, that is a mismatch between your vestibular signals which can measure the true acceleration, and your eyes which are trying to also measure the acceleration. And your eyes are being fooled, but your vestibular organ is not that mismatch causes a significant amount of nausea in people.

And in fact, it is very easy to make horribly nauseating experiences I believe one of the worst things you can do is to make it appear that you are virtually rotating. And then put some kind of sinusoidal oscillation into it at about a couple of hertz. And people will get sick very fast like I do you want to have a fun project to do in the lab try to try to make the most sickening experience you can it should be some kind of rotational thing experi-

ment with different rates, and you know compete with your friends further for the prize of the most sickening experience.

So, that is the problem is that the vision system, and vestibular system are in conflict. Again, I talked about decision theory your brain is trying to fuse information from multiple sources normally in the real world there is a consistent picture, now there is a big inconsistency. So, being nauseated is not necessarily a surprising response to that, right? Because maybe another thing consistent with that is you have been poisoned right. So, so it is good to eject right who knows.

So, so it is hard to know exactly what all the mechanisms are, and why this is happening, but it may be that you get sickness because the body believes that something bad has been ingested right and. And so, why else would there be such a strange conflict, it would not be able to take into account, modern engineering as one of the plausible hypotheses. Unless you experience enough of this and then eventually you may be able to train yourself to overcome this. Mismatch does not may be a lot of work to overcome some people may learn it easily some may not could be bit could depend heavily on age gender many other kinds of factors.

Um I should give one simple example of this we did some experiments, and I think it is worth considering just to try to alleviate these kinds of problems, this is more and someone has more to do with user interface design or avatar control or movement things like that, but I want to give you a very simple example of this problem, and how it tends to behave so.

# (Refer Slide Time: 30:19)



So, a very simple example is suppose you want to ramp up, your avatar or character motion. In other words, suppose you believe you are in a virtual world, right, into the beautiful world you would like to move forward you grab onto a controller, and you want to just start walking forward I want to ask you, which of the following 2 are best to do. So, you would like to get up to some fixed walking speed. And so, this is your speed here, this is time, and so one scenario is I mean you would like to get up to this speed let us say. So, one scenario is you just gradually ramp up. And then get to your speed that you would like to be at.

(Refer Slide Time: 31:07)



And then the other choice is you just decide that you are going to rather make a continuous change you just discontinuously jump to the speed you would like to be at and go forward.

So, this one corresponds not at all to the real world, because normally you cannot accelerate at an infinite rate right. So, so, but this is what you can do. So, you just pulse and all of a sudden bam you are you are walking at full speed ; which one of these is a more comfortable experience for virtual reality, let us call this scenario A scenario B. Who thinks scenario a would be more comfortable, right? So, this corresponds very closely to the real world, you gradually ramp up your speed just more comfortable. scenario B is this more comfortable ok. I have seen I think 0 hands being raised, but some little head nods here and there ah. So, here is what happens.

Student: (Refer Time: 32:10).

What is that B? I am hearing more B's now, right? Right, I should mean you know professors are always trying to trick you right. So, it cannot be the you know that this looks like the obvious, kind of comfortable sort of answer just make it maybe like it real bit would be in the real world I start off I just gradually get up to a speed and then I am comfortably going along. Well, what happens here in the first case is there is an interval of time between these lines here. During this interval of time, there is a mismatch, right? There is a vestibular mismatch.

So, your brain has the opportunity to latch on to a mismatch over some period of time. And it builds a lot of confidence right you imagine the brain is going oh yes, I am certain there is a mismatch it is been going on for a second or 2, right whereas, in this case there is a mismatch only at some tiny fraction of a second right. So, just a if you want to look at the limiting mathematical case, then maybe it is only a single 0 length interval impulse, right bam there is some mismatch.

This the brain may rule as an outlier. It may say there is always outliers in the data a kind of hiccup it is gone. So, this seems to be perceived as an outlier; like, what was that some big flash some huge mismatch probably junked discarded this one the brain is very certain and the brain is right you know, there is some very consistent mismatch happening here. So, so this ends up being very uncomfortable A, B ends up being the one that is more comfortable. This is some of the read it is one of the reasons why you will see this as well in some of the apps where if you are rotating your character, you may notice it is some of them when you press the button it rotates a fixed number of degrees each time. Rather than smoothly rotating smoothly rotating is very nauseating, but jumping discreetly is not as bad.

But if you jump discreetly 90 degrees each time, then you are not even sure which way you are facing anymore. So, you have difficulty of being disoriented, you are not really sure where you are facing anymore it would be great to just point and click where you would like to go and just be there. If we could do that in the real world it wouldn't be; so, bad either right. So, so that is fine, but you also start to get confused about your where you are in space. So, that is a trade-off you have to deal with.

So, we want to minimize this kind of discomfort, due to this vestibular mismatch with your visual cues. But at the same time, we also want to make sure that you can get that your spatial reasoning and awareness remains intact that is a challenge we face. If you can make progress on that for your class project I would be delighted, you know, there is a great great great challenge, right? Questions about that ?

Student: If there is an upper limit to the constsnt velocity at which (Refer Time: 34:52) does not appear?

Um you may be may be a lower limit, like, if then this match is sort of small enough there is another limit which could be that maybe I make this slope so low that it is so gradual that it may be comfortable right. So, so that limit could exist.

Student: I am asking about the limit to the speed itself.

Yeah.

Student: So, upper limit to the speed.

(Refer Time: 35:21)

Student: Yeah will it cause.

I think it will tend to be worse, but there are interesting questions about going up to speeds that are familiar to us, right. It is like we may become if we believe we are in a car, we may be accustomed to going up to some speed and then remaining at that, and if we see stimulus it is like a bunch of roads around that may be comfortable. If we are in a building and we go up to some speed that is quite outrageous faster than we could run in the building that may be uncomfortable. But maybe it does not correspond to anything that we have ever seen before. So, maybe its in fact, not uncomfortable I do not know you know there is a lot of things going on here there are a lot of prior cues our brain is basing it is conclusions this kind of unhappiness that comes up because of the mismatch on a lot of prior information.

So, that is why you know I wonder about when I showed you the birdly demonstration of the flying you go flying over virtual san Francisco, for example, that does not correspond to an experience that we have ever had before in the real world, there is also a motion platform it is helping your vestibular cues, but there is a lot of vection going on here I think nevertheless.

But that is not the same as this scenario like here where you are just doing ordinary everyday walking what your brain is quite familiar with and there is a very obvious mismatch. So, it may be helping it to be to be worse by doing things at exactly the same kinds of speeds that were normally comfortable with. So, it is providing everything in perfect alignment except this one mismatch. And the brain we have anytime if a lot of things are mixed up maybe it would not you know maybe it will not be as sickening, I really do not know you have to do the experiments on a case by case basis.

Student: (Refer Time: 36:57) in virtual reality in the most of the points are discrete and you cannot have smooth speeding. So, it will be like you are jumping from one point of the other at various speeds. So, that might one point at suddenly it is at an another point so.

Yea, very good, I think that is a very good point in fact, it is going to lead into my next topic which is apparent motion. So, if you have something that is really if you have a sequence of images, that are being generated by a digital system, and you know under what conditions, do you perceive that as a continuous motion under what conditions does that feel like the real world you perceive that like the real world. And so, so there is a nice sort of gradual entry way based on your question in into the next topic. You know, it is fundamentally I do not really make too much of a distinction between continuous versus discrete if the discretization intervals are so small that it very closely approximates con-

tinuous way below any sensible levels that could be detected by the body then the 2 are essentially the same. The universe itself is a discrete or continuous well; I think it is more discrete actually than continues, when you look at it carefully when you move beyond Newtonian mechanics, and into quantum mechanics and things. Though so, the discrete verson continuous itself isn't extremely important distinction this context, but when it is discrete what exactly are the rates, and are those perceptible, and that ends up being very interesting.