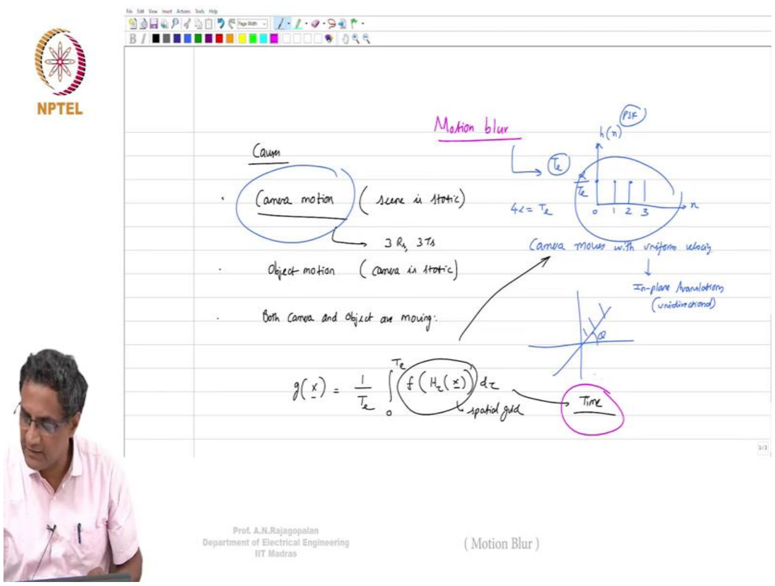
Image Signal Processing
Professor. A. N. Rajagopalan
Department of Electrical Engineering
Indian Institute of Technology, Madras
Lecture No. 35
Motion Blur

(Refer Slide Time: 0:16)



Applications where if you can understand this, this image formation this goes a long way in understanding many many things. So, you saying that you know during the time the camera exposure window is on is open and it is trying to gather intensities, so if if the camera was completely still rate and of course in this case here we are saying that the scene is static, so which that means that means that scene point will come and impinge on a particular pixel and area under intensity will be it will be simply that you would not see any blur, but let us say if this camera begins to move during the time of exposure then what will happen?

So, because of the fact that this camera moves so some other scene point will come and will come and hit the same pixel, more than one and you not be have two seen point, we do not know how many, so it is all a function of how how fast this guy moves and how it moves, whether it does a pure in plane translation or or or for example, you could just do something easy crazier than that, then whatever is going on in the scene.

So, it is like if you were sitting in that window, if you were the observer, then you would see so many scenes, you would see like I said it is one focused image, but then there are so many other sort of CC versions of that, that you are seeing because of the fact that this camera is moving.

And this averaging effect that you are seeing, see earlier days, I mean if you see for example until about 2000, people will will only talk about motion blur which was very simplistic, because only after your cell phones and all came, that this topic became a hot topic until then it was more the optical blur that let us say, you know everyone interested, but because of the fact that cell phones have become so light it is very hard to hold them stable, of course they also some inbuilt easy mechanism, but yet typically if you try to if you try to take something and especially if you have to take it very well it will go wrong.

Otherwise, I mean if you simply taking it right everything looks fine, but suppose if I have to say that oh, no, this has right this has to come well, then it is gone, that is what happens with me whenever I had to send some send some paper corrections to my student, he will say send me something sharp, I say yes, I will and then they will be a shake, other times it will be fine when let us say nobody wants it. So, it is like that. So, the idea is so earlier times people had a had a very very simple model also I will tell you that I will tell you a few things the differences between the optical blur and this.

So, motion blur of the motion blur was of this form, so if you were to look at a PSF, so so they will assume that the camera moves with a with a uniform velocity, this was a this was an assumption that they made, simply because they did not want to solve a harder problem than that so moves with uniform velocity, which simply means that you are kind of, you move like that or you move like that whatever if you are rotating rotate with the uniforms, in fact translation is what they use and most of the uniform velocity this was only in plane translation.

If you look at and and that too one direction, you know unidirectional then of course came, let us say some some methods that would say that it will still one-dimensional see earlier, people people assume that, that maybe if there is a motion it will be like this or maybe you would not move like that, you would not go around doing a lot of things, why would you.

So, they said let us assume it to be in plane translation, these are all in plane translations and typically unidirectional that means it is a simply a simply a one-dimensional kernel the PSF, it is simply is simply one-dimensional by which I mean that for example if I try to plot h f n for this, what kind of a PSF will I get? Then if you were to imagine, it could be that I am moving and since I am moving I am moving at a uniform velocity, I mean, it could be the during the time of exposure, all of this is during the time of exposure what all is going on.

So, so so what would happen is, since you are moving at a uniform velocity, so they will so it boils down to saying that saying that you have kind of you know, these these values at let us say zeroth pixel which is like which you just started to move you did not move at all, then you move by 1 pixel then removed by 2 pixel you move by 3 pixel, you can also have fractional sort of a pixel but then people were not worried about all that, because if there is 1 pixel motion, then the see clearly you can see blurring effects anything less than that they would just ignore.

And and all of these will have a weight, so if you had a total exposure time as Te, if your exposure time of the of that camera is Te then each one of them will have a uniform exposure which is like alpha by Te or in other words say if you have like 0 1 2 3, so it will be like 4 alpha is equal to Te, so that so that these weights sum to 1. Again similar to any other PSF, just as just as you have an optical PSF, it simply means that simply means that the the weight that you are associating to every motion every translation is all uniform means the same.

Now, this weight has a has a very very interesting interpretation, I will when you when you look at the larger picture, it well actually hit you, now, it will simply look like they all have a because it is moving uniformly, therefore all of them have and have the same sort of, share of weight, therefore each one is like alpha by Te and therefore 4 alpha should be equal to Te and therefore so so so it is like each one each one gathers equal amounts of it densities.

Now, this was a very very simple model and then people thought that it is enough and then when people went on to kind of talk about the you know on a plane it could also have let us say let us say one-dimensional kernel like that, we should be at an angle at some sort of the theta angle. But still it largely remained 1D, so the only explanation that came was instead of assuming that that you know, the basically it is along just Te, the X axis, so they would say it could also be along X and Y axis, in which case in which case you will have translation along both directions.

And therefore, you would have some angle theta and typical is overall assumed to be known and people were solving simple problems. But then one equation was known which which I am going to write, which we will interpret now, the actual larger picture, so g of x is equal to 1 by Te integral 0 to Te. So, Te is a is exposure time, so the time for which is the windows open and then f, why I am why I am saying this is something that you can easily relate to because you have done homographies that I want to get a bring that portion back again in the year, something that you did for really a pinhole camera and right now we are also looking at really a pinhole camera.

We are saying that everything is focused, we are not saying there is any optical blur, but now the homography notion will again bring it back in here. So, it is like f of H tau, H tau of x d tau. This is the spatial grid or in other words, what does this actually effectively saying is that see for example, if the camera moved during the time that the window was open so so you know it could be that it was here, then it went somewhere, then it went somewhere, then it went somewhere and then a long time it is so every time every time you can think of it as not doing some kind of a homography on f.

So, here there is a time notion, tau is your time, so as time goes on from 0 to Te, so this camera is kind of see it kind of moving around and in a continuous case you would sort of think that you are accumulating all the warped versions, by warp I mean I mean the whatever if this camera moved by some amount there must have been some H tau with respect to that and that H tau is being applied on the spatial grid of f to give you a wrap deaf. It could be a simple translation to be rotation, so the real equation is this a simplification of this is that, where you simply assume that H tau is pure translation in plane unidirectional other.

But in general, this is how it will look, but at this equation by itself never gave too much insight into what you can play around with and one of the issues with this was that then you know it kind of say tied down to time, so it is interesting when somebody somebody writes this equation to surround for such a long time and then everyone would just look at it and say fine we have that simpler model go and use it.

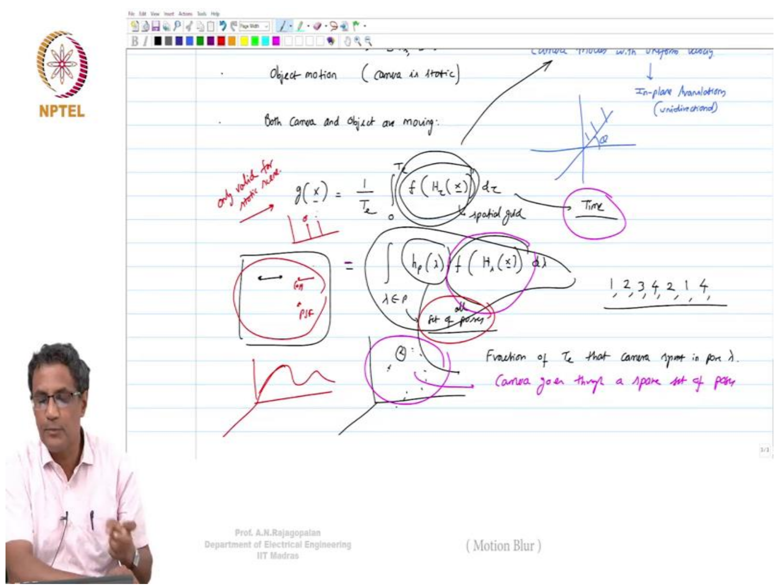
But now if you take a cell phone, especially and now if you have a motion, you cannot say that it is going to be pure in plane translation and all that, it could be a crazy kind of motion. Now, what is happening is so the blurred image that you are seeing g is but several warped versions of what you would have ideally is like to see as a focused image warped versions of warped by which I mean a homography being applied.

It is another thing whether homography is is applicable for that kind of a scene or not, but let us assume that it is applicable in which case you are taking different different you are applying different different homography’s I mean they are not applying, it is just that what is going on and all of that is getting (())(09:16) so you can so you can think of f as this and then maybe maybe at tau equal to tau and it had another homograph, so it made f to be a rotated version, then maybe at you know tau be another and then it could go back and forth, you see here, you do not assume that it has be in one direction and all.

Along time, whatever happens you know you could go back and forth, but then along time whatever you are doing all the all the warped kind of all the warps just say average them and that should that should be see theoretically yield you a G, which makes sense, this makes sense, but then the hard part is this time notion.

And and we all understand that why should I need time and I am going to talking about averages, if have like 1 2 2 some 3 4 5 5 and if I had to find the average I mean to say look at how many time each one of them occur in those numbers and simply add them up in a kind of a different way. Why should I so?

(Refer Slide Time: 10:16)



So it is like this, so here is where in 2000 or and I think 2000 or 2004 around that time I think even later than that there is a paper that is kind of came out that see equivalent wrote this down I will write down what they wrote it wrote it down as.

So, they said this all the myth which was obvious I mean if there is anything great about that, but then that changed the way you to (())(10:32) interpret the whole thing, so they said this is equal to integral lambda exactly equal to, there is no approximation nothing equal to hp of lambda f of H lambda H of lambda of x and then d lambda, where p is the is the set of poses, set of all poses. Now, what this what is equally means, so what will you interpret this hp of lambda as now?

See, it is like saying that see what it is saying is if a camera went from here went there went there and suppose suppose it was brought back for some reason came back and landed again in the same homography that is was in originally in, then it then it went somewhere else, so it is like saying that if you look at the sequence suppose I write it down as 1, 2 then then maybe write maybe 3 and then you have 4, suppose each is pose, I mean pose 1, pose 2, pose 3, pose 4 but then after some time I come back to pose 2 maybe I come back to pose 1, then I jumped to pose 4, it depends on what is good is going on there.

So, if you were to look at a look at a look at a time trajectory, sometimes way it may be important for you, but as far as blur is concerned why should I worry that 1 occurred first and then 2 and then 3 and then 4, I would rather say that 2 occur twice, 3 occurred only once in that is in the see total exposure time 4 occurred twice since 1. So, this is hp of lambda is really the fraction of the exposure time fraction of Te that camera spent in a pose lambda camera spent in pose lambda and this makes equal sense.

So, you are just you are just saying that saying that you know turn off, for me and say time is not at all relevant, averages, it does not matter what sequence sequence I have those numbers whichever order I put them I will get the same average. But now what you are saying is, you would rather rewrite this equation you will throw away this notion of time and simply write it in this form.

This sort of this this this enables you do you know think about whole problem in a sort of a new way. For example, if I had asked you a point spread function, see if you imagine that imagine that imagine that here you know your camera is undergoing see for example, if it was a pure translation, which I which I drew earlier you know I kind of drew a kernel.

Now, if I asked you, how would you find out if I told you where the camera was in every instant of time, which means that you would have to know what was pose of the camera and what you would do is if I asked you in the image, can you tell me what will be the point spread function, at some place, if you knew the camera motion, you can actually find it, how? Because you would simply take the I mean so for example here I mean I can explain that to you from this, this will be easier to explain if you go back to the time notion it will be hard for you.

See in this notion, what this means is that, you see you agree that your PSF could be either time way spatially varying or it could be just a constant, if it is pure translation in plane like I said the first case that will be invariant all over, because you are simply translating the whole image you are translating the whole image you are translating the whole image there for all pixels are of the same amount of blur, but then if you if you have to think a little deeply, so how do I then extract a point spread function at a pixel? How do I evaluate it?

So, what do you do? You come to a pixel know the location, apply the pose on that particular location, see where it goes and then you know the weight for that post, so the weight is something like this, so it is like saying that when you are trying to gather photos if you just stay in one pose for a certain time, but then if you stay longer, if you happen to stay longer in certain pose then it means that you are acquiring acquiring more photos.

It could also be that you know you are fleeting I mean you do not spend that much time, but then you know, you kind of see come back again and again that either case you are accumulating more photos, either you just stay in that pose and then you just gather as much as you can then you move.

So, the idea is that, you do not have to worry about whether the camera moving uniformly at what is going on at all, you just have to worry about where is it at and then you know in which particular pose, so you can see if you can think of the 6D sort of a trajectory, (())(14:58) if I were to think of it, a some kind of a 6D figure which I am trying to draw here, so you can think of the camera pose or something right that that it travels in that pose, that kind of a pose space.

So, initially it might have been here, here, but then in time we do not even care whether it was here first or here first, we just want to know in this pose how much time did it spend, whether it spent at one go that much time or whether it spent initially then came back came back again it does not matter, I just want to know how much it gathered by sitting there and that is you're your hp of lambda, hp of lambda is that weight that you want to you want to assign for that pose are you able to when this is something that we are able to able to understand I mean it just requires us to stretch our imagination little bit more.

So, it is like saying that, so since you are going to be gathering more photos if you spend more time in a pose, so when you finally compute the averages you should way that particular guy more, so here for example, so here if you see if for some homography if an hp of lambda is high that means you actually spent a lot of time in that pose and we do not care how you spent it, we are not interested that you first came and then let kind of say revisited you are not worried.

But then how to find this is another issue, just saying that this is all equivalently this, how one finds hp of lambda in order to be able to explain a blurred image is something else, but right now I just wanted you to understand the image formation process because even with this, so the one thing that you also notice is that, the camera goes through goes through a sparse set of these poses.

But there are some cases where where where you actually rotate it like at a very high speed for example submarines and all, when they have a camera there to come up so they kind of send this camera up within within milliseconds, it needs to swing around at a very high speed just to know what is on the surface and then it kind of then comes back and then somebody has to make sense of whether that blurring and all is very heavy because because the speed at which it actually rotates is so high that if you look at those images you cannot make out anything.

But then you have to make sense out of that either a human has to make sense of that or you have to do one of these things, in order to be able to de blur, not able to remove the blurring effect and show us to what that scene might have looked like. But I am saying normally our normal camera when you use like a cell phone and all, it only travels a few of these poses you do not spend so you know you do not walk all all over this place this go somewhere go somewhere go somewhere and you are done.

And what is equation saying is if you can find out what these things are at each at each particular pose, how much how much weight should be assigned, which is the fraction of the exposure time, it is simply warp the image by that homography rated by the term by that homography and simply, keep adding them, if it is a discrete case, then you will have a discrete set of poses that is what you eventually do, you cannot do a continuous case, so you will actually do now a discretization of the poses.

And then and then you will have a table that will say for this pose this much is the weight for that pose that much is the weight and then you simply wrap multiply by that the weight and see add those whole thing up. This is not valid if the object moves and all, like I said, no, see only for this you can have a simple model, this would not work I mean if there is an independently moving object and now, you are not accounting for that.

Static seen moving camera that is what I said, static anyway, let me write that down, so nothing maybe that is a valid dot, so only valid, so this equation is valid for static seen, so all all this motion supposed, the only other case where it is valid is if the whole scene is occupied by one object and that object moves in this way, then it is, you can imagine, suppose the whole scene is occupied because one object and that guy is kind of say you can have your camera still and that object could be moving, then also it would apply.

But normally we do not assume that one full object will cover the whole image. But suppose I asked you, at this point, what is a PSF? Assume that it is all spatially varying and how do you find it? You do not even need the image, you just need the location of the coordinate and you you basically apply I mean and just as just as you could have say did your homography, what did you do?

You apply to all homography and say each of those locations right except that instead of one homography you got a bunch of them now, so you would apply let us say whatever 1h lambda and you see where this point goes and there so for example, for example, you can you can you can get some kernel on a kind of you know a 2D space, you can get something, you know, which looks which looks like which we can have any shape for that matter.

And unlike optical blur you cannot a priori tell anything about this, an optical we said it is nice isotropic and all that, it typically because you know there is a nice aperture symmetric and therefore the blur and all has some symmetry nothing nothing of that then you know will this guy obey, because you could have moved it this way you could have moved it in another way, nobody knows a priori, how this kernels might look like.

But if somebody gives tells you that these are the camera poses that it went through then you can actually you can you can without even looking at the image I can tell exactly what should be the kernel here, what should be the kernel? Because all that I will do is apply every pose to this point see where it goes and then wherever it goes, so so in that kernel for that point I have a kernel and and at that particular location I will actually put so because this is like 0 comma 0 for me.

In the kernel and therefore if it goes here under some homography then then you see whatever is that value hp of lambda that that will that will access it here and then you apply the apply the next homography maybe takes it somewhere else then whatever it is that you are you whatever is the weight hp of lambda corresponds to the and since hp of lambda sums up to 1, all these weights will also sum up to 1, whenever will you exceed the sum beyond 1, this will be automatically.