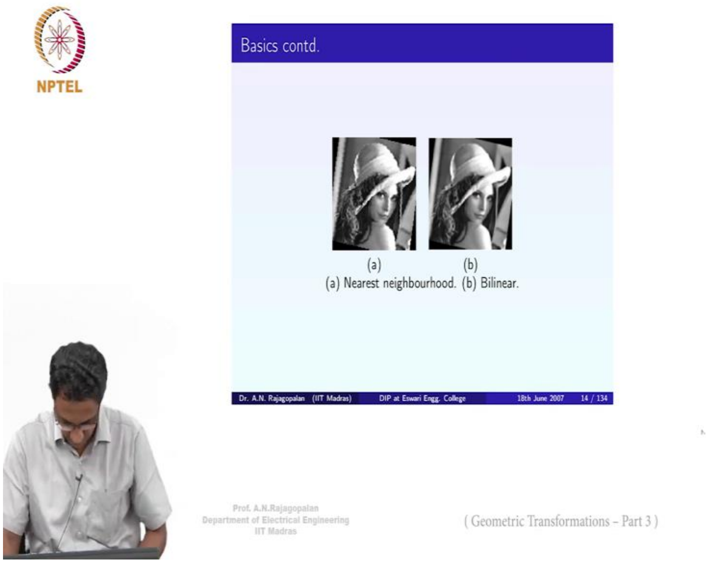
Image Signal Processing
Professor A. N. Rajagopalan
Department of Electrical Engineering
Indian Institute of Technology, Madras
Lecture 9
Geometric Transformations – Part 3

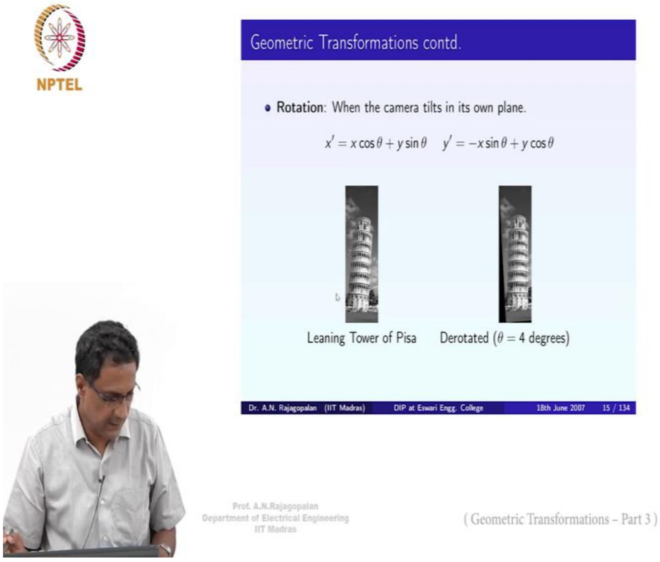
(Refer Slide Time: 0:23)



If you are really wondering how this translation and all will look like. So, I thought I will go down and show you some of those things. As if you see if you do a nearest neighbour kind of interpolation is all a target to source. Those are all target to source the first is Lena but we have done a nearest neighbour kind of interpolation.

So, hopefully you can see that jaggedness might on the hat and other places that is, if you do a bi-linear interpolation, it looks a lot more lot more get nice to the eyes. Gives you a smoother feeling and gives you a sort of a better feel.

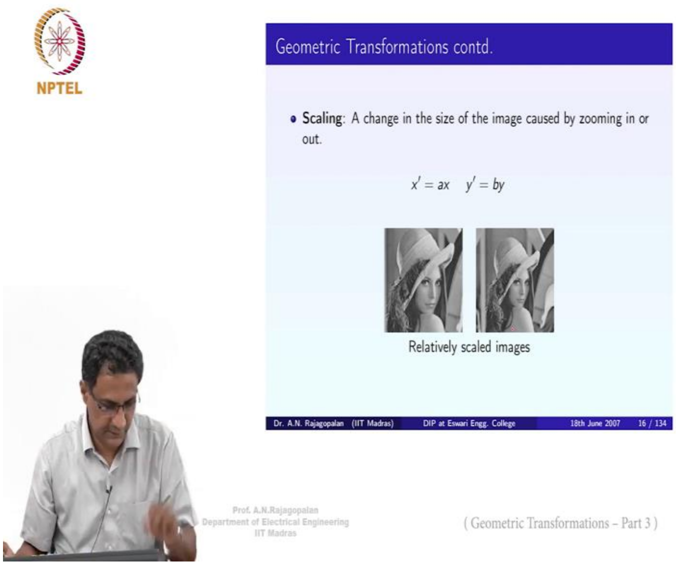
(Refer Slide Time: 0:59)



Then then the next thing is rotation. So, this is a tower that I talked about the Pisa tower answer is sitting there so, you just ignore that. While says four degree but maybe even even do better than that. So, here is that here is that rotation example which you are supposed to do and if you notice this zeros that are actually coming in here this is simply because we are doing it as an implementation in the lab.

But in a real situation you would not get this 0s and all when you actually tell some new information will come in. But since in this case that is not that is not anything outside of this outside of this mesh. So, we simply include zeros. Does not mean that in a real case you know you get zeros and all. This is simply an implementation aspect.

(Refer Slide Time: 1:49)

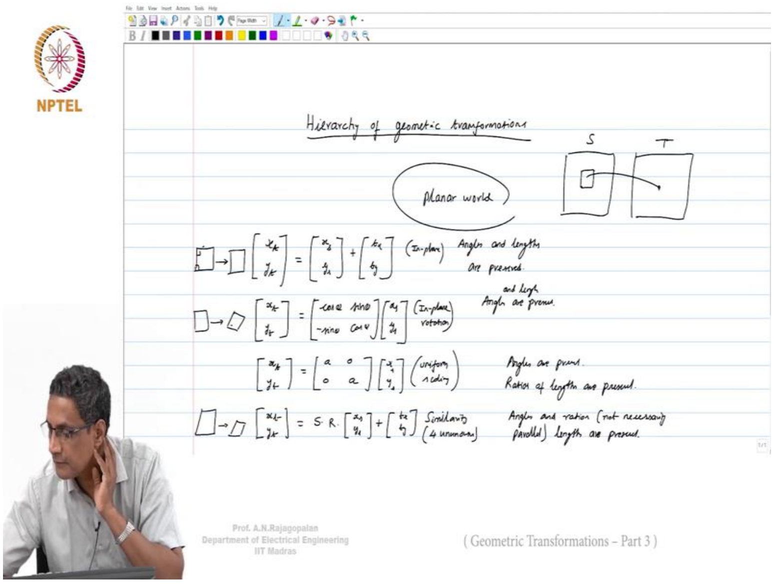


Then scaling again that is you see here this is like a zoom out. If you look out the look at the second one related to the first it is like a zooming out if you did an actual lab implementation you will get zeros outside of outside of this Lena's face. Here you know we have taken (())(2:08) implemented in a way that that you know he knew what was outside this and he copied it I was telling you yesterday.

You can also do that if you had access to information outside of this because in this case we knew what lay what lay outside of the phase we could actually copy but in the reality when you implement it do not assume all that just so, if you do this if you zoom out you will get zeros here because you do not know what lies outside the outside this grid. Then let us see what is next.

Prior to going to other things I do not think there is anything. So, and in all of this rate please kind of remember that that that you know when you when you do this target to source does not mean that they are not target intensities or not.

(Refer Slide Time: 2:58)



The target to source it simply means that we want we still want to go from the grid wherein we have no information. But the thing is what we are trying to do is, instead of going from source to target, you are picking a coordinate here which is a location that you know, it is an integer location.

You are applying the inverse transform and then then you are landing somewhere here in the source, fertilizer which you are using let us say, the image at 4 neighbours if you are doing a bilinear interpolation and whatever is that intensity that you get by this weighted averaging, that intensity will get assigned here to this location to the target.

Starting from target to source, it does not mean that we know target intensities or something, no. both cases, we do not know the target grid. Target is a image if we want to maybe if you would like you arrive at. It is it is it is supposed to be generated, generated given the source. So, (())(3:45) information that we have is in the is regarding the source. From the source we want to go to a target, whether you go from go from source coordinates to target coordinates, or whether you come from target to source, is what is what matters.

And as I said, source to target is not really a good idea. So, go from target to source, so that you can actually visit every location in the target and come into the source, land wherever and then use whatever interpolation that you are familiar with. In this case, I just showed you how you can do a bilinear interpolation and you assign that that intensity here and then move on to the next location, do the same thing again assign and intensity and that is how you will fill up this, that is how all those images have been filled up.

Now, as I said these are actually a simple transformation, the one that we have seen including in plain rotation, in plain translation, they uniforms well, uniform scaling. These are all very very simple examples. Now, what I will do is, I will write down what is called the hierarchy of this geometric transformations. There is there is an hierarchy, there is a hierarchy of geometric transformations. And we will see from the simplest to the most complex one.

And the idea is the idea is to know actually that eventually it is like saying that from moving camera, what do I see? if I see the world. All these geometric transformations and all, what they are trying to tell you is, if you had a camera right with you if you able to move, how would the how would the world look like? How will the world map on to your image? Now, what you can show is, in all these transformations and all, you can do for simple cases.

It does not mean that for every situation you can actually do this. You can if you if you want to do that, then it may not be as simple as this, not that it is impossible because people do align whatever, any scene any image of any scene to another image of the same scene and so on. but the goal behind doing this kind of geometric transformation is to actually understand as to when you can do this a simple one global linear transformation that can take you from one image to the other.

That is a very very very elegant and special case because all that you have to do is, get some matrix which you have to estimate. And then ones that matrix is known, then you can simply so that law holds for the entire image. And that you can apply on that image to get the kind of say get the other one. Now, this need not always happen. But whenever it happen, it is actually very nice case.

That is why whenever we talk about geometric transformations, there is something called a planar world. The planar world occupies a special place. Even though we may not think of the planar world to be that interesting, for us a 3D world is much more interesting a planar is also in a sense but they effectively it is the is still a plane. So, a planar world is what will be central to our argument and alone the way, I will also tell when you can actually relax this relax this as assumption.

It is like saying that so the whole goal is like this, I take a camera and I move, I want to know when I can relate the images across two views through a single linear transformation. That is the goal. Now, the whether you can do whether no can the seen completely arbitrarily in the sense suppose I take this classroom and say it is a complete 3D world and I and I take a camera and translate.

But if I take an image from here, I take an image, is it true that I can I can relate the two, the answer is no but if I were to stay right there and if I simply rotate the image without translating the camera, can I actually, so it is like saying that I take one image like this and then I rotate about the centre of the camera without translating the camera. If I take another image, can I relate, the answer is yes.

So, we want to know as to what is that, what is it that? So, the whole idea is, taking a camera and moving and then trying to trying to relate images that you capture along the way. When can you relate them and when can you relate them in a straight forward manner. By straight forward manner, I mean one transformation, one linear transformation that will take all the pixels here to the to the other grid.

Each pixel can move in its own way, that is alright, we are not saying that they should all move by same amount. For example, if you do a simple rotation, rotation as you understand about the origin that pixel never moves actually, whereas as you go outward, radially outward, the guys at the boundary actually move a lot. The guys that are closer to the origin do not move at all. In fact, the guy at the origin does not move at all.

All that is okay, we do not mind each one moving by its account, but that should be one transformation that should take them all. We do not want for example one transformation for this, for this part of the scene another transformation for that part of the scene. That we do not want. Want just one transformation that will take us from here to there, does not mean that pixels will all move by the same amount, no.

If it is a translation, maybe yes but if it is a rotation it does not happens, scaling it is not true. Different pixels move you see differently depending upon where they are in the image. So, the goal is to sort of, arrive, actually we want to get off in our head we want to be able to think about what camera motion and what scene will lead to will lead to these kind of say, transformations.

For example, you can ask when I said yesterday that you can simply get a translation like what is that what is that some tx ty. You might wonder now when does that happen? I can I can simply take a take a camera, translate here, take two images and claim that, no these two image will be related by a simple shift of tx and ty for all the pixels, things like that. Or for example, can you have a plane that is inclined and this camera is looking at it, then what happens? Does it have to be front parallel, can that be inclined?

Under what situation, what will happen is what we want to see. So, so but I thought initially we will so the reason was we will start with a very simple case where we would not really worry be what was the camera motion, what was the scene and all? We simply say, this is the this is the transformation that works. But now we want to we want to step back now and we want to analyse this to what motion and what scene will lead to what transformation?

And again, the transformations are not limited to the 3 that we saw yesterday, that is why this is an entire hierarchy. You got this? So, we will kind of go through so that I just want to set the tone for why we are doing this. So, this hierarchy, the simplest is what you say yesterday. Because in the first one, well is what I will call as so here what you have is this xt yt is equal to, this is xt yt is equal to xs ys plus tx ty.

This was the one that we saw yesterday, this is one of the simplest and what kind of what kind of things will kind of keep intact in the in terms of the image features, angles and lengths, angles and lengths are kept intact are the preserved or the kind of say, remain intact. So, it is your angels do not change. So, what this means is that if you had let us say, square like that and if you if you apply the transformation, it will remain a square except that it just shifted.

None of these angles change, these remains as right angles, then they remain as right angle, the lengths do not drain, nothing happens. The next one is actually rotation which is xt and all of this is all the red image plane. This is all in this all motion in the image plane, tx ty is all on the image plane. When we down the line, we will see when the when the camera motion enters into the picture.

xt yt is equal to cos theta sin theta if you are looking at a clockwise rotation minus sin theta cos theta and then xs ys. This is this is what we called as in plane translation, last time itself I wrote down, in plane this is also in plane rotation in plane rotation and what does it keep angles are actually preserved are preserved and of course there are angles as well as lengths are preserved. So, that is common to both in fact.

So, you have this, then you might actually end up seeing something like that. All these are right angle, so as this and so on. Those angles and lengths are kind of say, preserved. Then let us look at uniform scaling which is xt yt and I and I going to deliberately tell uniform. So, here I am going to use a 0 0 a and then I have xt yt. This is uniform scale xs ys. xs ys this is uniform scaling.

And here angles are angles are preserved and ratios of lengths are preserved, ratios of lengths are actually preserved. So, the idea is what do these kinds of transformations do to your image? So, that is what we are interested. Then the then up the hierarchy the next one that comes up is xt yt is equal to let us say scale which is actually uniform scaling. If I call this as s, this uniform scaling, then rotation, this is in plane rotation R and then here we have xs ys plus tx ty.

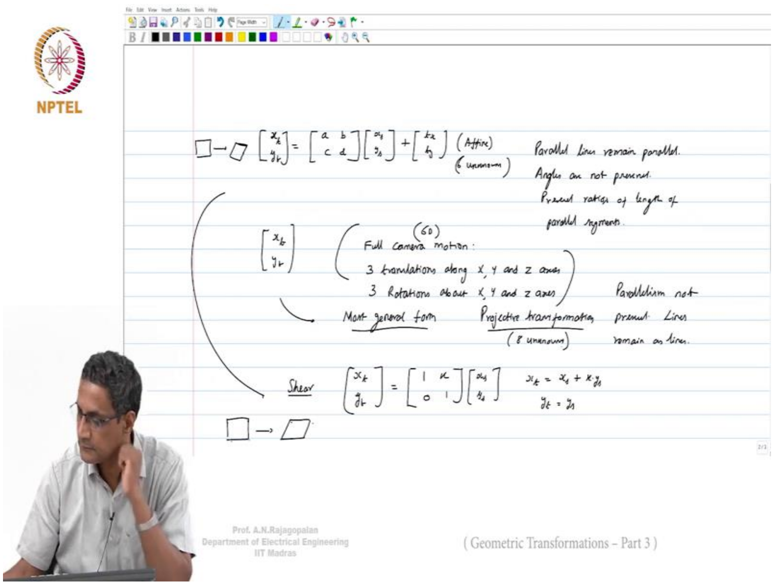
So, if you if you look at the number of unknowns as you going, so here you have two, here you have only one unknown, here you have one unknown, but in the case of similarity transformation, there are one unknown coming up here, 1 plus 1 2 plus 2 4. So, we got like 4 unknowns and this called a similarity transform. And a special case of this are your uniform scaling and all.

When let us say something is some R is identity and you can only scaling tx ty 0, then you get uniform scaling. So, you can derive the previous three special cases of similarity transform. So, this is similarity transformation. This is called and it have 4 unknowns. So, these unknowns are the ones for which when you need to solve, you need those feature point correspondence and all that.

And here what it say is angles and ratios are not necessarily parallel segments, not necessarily parallel lengths or not necessarily parallel lengths are preserved. So, if I want to draw some kind of a diagram, what it might mean is, if I want to indicate something like this, then there could be a rotation and so on but I don't know whether I am drawing it exactly but something like this is okay.

So, your so even through like these two these two segments are not parallel, it does not matter. As long as there is uniform scaling, that is what that is what special about the similarity transform. Involves a uniform scaling, involves rotation, in plane and then a translation is again in plane. So, four unknowns or what it can do is, it can actually keep these things intact. You can go above that what is called an affine.

(Refer Slide Time: 15:45)



The next one up the hierarchy is what is called what is called an affine transformation and that will look somewhat like this. So, you have xt yt is equal to a b c d. Some a b c d, then xs ys plus tx ty. This is what is called affine and most of time this seems to be the one let us say people think usually happens, but of course something else above this that can also happen but affine is very very common.

What it can do? Parallel lines remain parallel. Which means that, let me finish these parallel lines remain parallel, angels are not preserved so, so what it can do is you can take something like this and do something like that. so, parallel lines remain parallel, but then the angles can change. If this was right angle, this need not be right angle anymore. Angles are not preserved are not preserved.

It preserves actually what it does is it preserves what is called a cross ratio. I mean preserves ratios of lengths of the parallel segments, it preserves, but they are enough. Before that let me write this, preserves ratios of lengths. See earlier case, what I had written was it will actually preserve ratios of lengths of not necessarily parallel. Here, this will preserve ratios of lengths of the parallel segments.

The earlier one if you see, there will be a small difference. There I had said, weight are not necessarily parallel. Here, if you have an affine transformation, then the angles can angles can change but then what it will do is, parallel lines remain parallel. If you go higher than this, there is something else that can happen, but hen that involves what is called what is called a full camera motion.

These are all, see whatever we have seen till now is all kind of restricted camera motion. When you have a full camera motion, what is that mean? That means if call it a 6D motion, a 6D motion by which we mean you can have 3 translations involves 3 translations along x y and z axes and 3 rotations about x y and z axes which is which is so this is the this is the most general form, most general form and you know, in order to even write this kind of a mapping between the source and the kind of target we have to actually, we have to actually go to go to a different sort of a coordinate system.

I will talk about it now. So, this is the most general form, this is called really a projective transformation, it is called a projective trans... It has a name, the earlier one was affine and this is called a projective transformation. And this is the most general form in which in which you can relate images of the scene where the camera can go, camera can undergo arbitrary motion, arbitrary 60.

Now, what does what does it? It kind of say retain so lines remain lines thankfully. Parallelism not preserved. That is in affine you had that parallel lines remain parallel. Here, parallelism not actually preserved. Lines remain as lines. It is supposed to suppose say, supposed to be preserved, what is called a cross ratio. (())(19:56) all that all that is effectively means is that the midpoint of the segments and all they do not change. The midpoint remain where it is.

We do not have to we do not have to get to much involved into those aspects. But just remember that the you might ask, can it ever happen that that a line could change or something? You use a camera, have you ever seen that something like an edge becomes a curve or something? Because I because I said lines remain lines. I mean, why the hell would line turn into a curve or something? Have you ever seen that happen? Or can you think of situations when a line can become actually bend or something?

So, rolling shutter is one such example where it can happen, where it happens because of the camera motion. See there are other cases when if you have noticed spherical aberrations and all. Those are lens aberrations. We are talking about when does camera motion lead to something like that, align or something becoming bend, so you can imagine that if were to take a camera and suppose let us say suppose you move and suppose this is not this is not your usual camera, it is like a rolling shatter camera, in fact all, you should actually you should actually check this out.

It is not so hard. Take your cell phone and then you kind of look at look at the edge of a wall and then you go like this fast, go like this and come back. So, what will happen is, if your motion is fast enough, then what happens is because all sensors are not exposed at the same time, so what will happen is, some so the so the top few rows would have would have would have observed this motion which means which means that that the line will kind of say, tend to tend to bend this way and then and then because you are coming back and forth, you are coming back so then the rest of the portion will try to bend inwards.

And we see that image that will that will have some kind of a curve which actually means that it is introducing something which is actually in the scene. That that could complicate matters. On top of that you may have blurring and all that. Blurring never changes the changes a line into a curve. Blur will only smear a line. If you have seen, if you blur a sharp line, what will that will happen is, it will smear around it.

It would not become bend. So, rolling shutter is very very special case. But then of course most cameras have it, so it is kind of it is also important to important. Now one more think that I forgot is, there is something called shear. This is a special case of affine which is called the shear. Let me write that down here. Now, what is called a shear? It is actually a special case of affine. See a b c d is affine. Now, special case is something like this, what we call is x shear.

So, this will look like 1 k 0 1 xs ys. So, what this means is that, your xt will be like xs plus k times ys and yt will remain equal to ys. This is called a shear. So, you see this is a special case of that a b c d. k can be 1 k can be greater than 0, no it can also be positive negative whatever. If k is 0 of course, then case is not interesting and this this will kind of so this is called an, this is in fact called an x shear.

So, what it will do is, it will take your guy like this and it will do this. So, your y I mean, so if you see here, x changes. If you think of this is 0 comma 0 whatever, 0 comma 1, 1 comma 0, 1 comma 1, you try to plot it, then what will happen is, here y coordinate is, y has not changed. This is at let us say this height is 1 so, this also at 1. So, all these coordinates, the y coordinate will be 1. The y coordinate for this entire segment will be 0. But the x coordinate will change depending upon what is a y, whatever is the y coordinate.

For 0 comma 0 it will stay there, then for let us say, 0 comma 1, it will it will move, x will move like 0 plus whatever. If k is 1, it will be like 1 times 1. So, it will become 1. So, this point will go here, this point will go there, this point is here and this point is here. These two do not move because y is 0 for those two points. So, this is called a shear. I do not know this is something that that you actually experience.

Think of think of something like you move you move in you move inside a car, you are going travelling by car and assume that assume that amuse that on the road, if you had a loy camera, if you just look out, you are not the driver hopefully, you are not looking out of the window and driving. Assume that assume that you have somebody who is driving the car, car is going this way and you look out of the window and suppose you saw a zebra crossing on the road.

And suppose this car is translating this way, then the zebra crossing will get sheared in your eyes you will you can see that shearing effect or you just take a camera and if you capture it, next time when you travel do that, watch out for that zebra crossing and you take the picture not the driver. So, that is what, rather you can also have a y shear which is like 1 0 k 1. So, this is called shear. This is a special case.

And all of these, I want to come back later to be able to we should be able to answer us so when these things happen? Now, I am just writing down some matrices and saying that this is a shear, this is that, but the ideally you will get to know a nicer insight a better insight if you were to now visualize what camera motion and what scene will cause this, this matric to occur. Any doubts at this point? This seems to be problematic.

Now, what I am going to do is, I am going to I am going to I am going to look at a projective transformation because that is the most general and then everything else will be a spin up of that. Ones you understand how you are how we can arrive at really a projective transformation, it has actually 8 unknowns by the way. Affine will have 6, a b c tx ty.

I will write that. Let me write it down. I think I forgot. Where is that affine thing? Here, so it has 6 unknowns. So, all these unknowns are things that we might have to solve for I because nobody will tell us what is that, what is the a b c d. They may just give you two images and they will say align them, then we should find out what is this a b c d and tx and ty. This this projective transformation has 8 unknowns. We will see why and how the line. And that is the most general form.