Modern Computer Vision

Prof. A.N. Rajagopalan

Department of Electrical Engineering

IIT Madras

Lecture-65

So, let us choose 2 points on the ray, okay on the ray x is z. So, this is actually 4 X 1, okay when I say when I say x of z I mean I mean that z k inverse x ~ or what a kl inverse x ~ 1, okay that is that is x of z. So, the one point right, so 1 at z equal to 0 that is the camera center. See the idea is that you want to take you want to take see 2 points which are probably right easy to handle rather than right taking up anything which is arbitrarily there on that line. So, 1 at z equal to 0 and the other at z equal to infinity that is a point at infinity, on that ray right on that on the receive back with this one projected ray. So, for so for z equal to 0 right your your x of z will look like 0 0 0 1, right because because your your z is bang at the your point is bang at the camera center.

And for z equal to infinity right your x of z will look like k inverse x ~ again at this x ~ is this and this entry has to be 0, okay. Why is this so? Because of the because of the fact that a point at a point at right infinity. So, see this is also the reason why the why you kind of resort to a homogeneous coordinate. See I mean so in order to write I mean express a point at point at infinity right what you will typically do is forget about this this particular x y and all.

So, you so if you have right if you have if you have if you want to express a point at infinity the last element should be 0. Now, now you know in this case right in this case this is the k inverse x ~ stays because of the fact that the the relation that this whole thing should map to x ~ is right should still be there right. We cannot have something arbitrary sitting there because we know that any point projected from the on that ray from that ray right if I have to project it to the image plane it should map to map to x y 1. And and then to indicate the fact that right z is at infinity in the sense that the point is at infinity the last element right is actually 0, okay. This is the way you actually represent a point at infinity the x ~ that is sitting here that is still to convey to you that this is not an arbitrary ray.

This is an arbitrary ray on which any point projected from this ray on to the left camera will always map to map to this one x y, okay. That is the reason why it takes this one, okay. Now okay now suppose I apply so so now I want to know where these two points map on to my on to my second image right on to the on to the right. So that will be like P right and then I need to put this x z right. So x z if I put as 0 0 0 1 right.

So this is where this is where the first point right which is right which is at which is at origin right is going to go and that will that will be like what P r is k l i 0 and then what is this 0 0 0 1, no no not i 0 it is r t no. Yeah, yeah, no no I wrote this wrong this should be r t no there is there is and k r yeah this is not k left this is k r, k r r t 0 0 1 and therefore, right this will be like k r into t right. What is what what will be this point actually? A P pole. A P pole yeah this is where this is where the center is actually intersecting you know the other image plane this is where this is where this A P pole is. And the other point right will come like P r acting on or though this should be k l inverse by the way I do not know right where I I should consistently maintain that k l inverse otherwise there will be confusion k l inverse x ~ 0 and that will go like k r r t k l inverse x ~ 0, 0 is a scalar by the way okay.

So this will be k r r k l inverse x ~. What is the inverse of an upper triangular matrix? Is it upper triangular or is it lower triangular? Lower triangular. Lower triangular no it is upper triangular anyway does not matter just wanted to. I was curious okay so k r r k l inverse hey wait a minute what did I do x ~ yeah right this is how it is right okay. Now we have right one point is this right this this is the first point in the second image plane this is the second point in the again in the say right camera and the line passing through them is what we want right is what we are interested in right that is the X product of these two chaps right.

So so the so that will be your right see epipolar line for that point okay again right it is not for for for anything yeah it is for x ~. So the right epipolar line for x ~ for x ~ is then given by and given by let us say what did we call represent L , right that is what we said okay L , is equal to what did I have k l k r i okay. So what is that k r i k r t oh sorry yeah it is a translation. k r t X product what is the other one k r k r r k l inverse x ~ right x ~. Now k r is actually actually a matrix right what is a 3 X 3 right and there is and there is a result in X product right which is a standard result.

So you can just take this for granted m a X m b where m is a matrix and a and b are vectors and does anybody know this property okay I do not know how many of you remember all this I do not remember but this is what it is a X b okay. So this k r being actually common to both so what we can do is we can write this as k r minus minus transpose okay this is not that vector okay this is transpose yeah exactly m inverse transpose so it is minus t so it is like k r inverse transpose okay k r inverse transpose and then a X b right so that will be t X r k l inverse x ~. See x ~ is actually 3 X 1 and therefore, right you can see that right I mean everything matches okay dimension wise but this itself right I mean you know we saw right that okay right that there is a that there is a way to kind of say write this so yeah exactly so we can we will just go and use that and we will write this as t X and which is which is that skew symmetric matrix that you can derive out of the 3 components of t right into this vector right which is r k l inverse x ~ okay. This is this is actually 3 X 1 this is 3 X 3 this is 3 X 3 and therefore, l , is equal to 3 X 1 right that is that l 1 l 2 l 3 whatever for that it be polar sign. Now see if this is all k okay now if you this product right this T x into r that you see here T x into r this is called this essential matrix whereas, this product right which is like k l or whatever k r minus T T r k l inverse right this is called the fundamental matrix it is called the fundamental matrix or in other words right you can write this as you can write this as l , is equal to the fundamental matrix x multiplying x ~.

So, you see right so what it is doing is x ~ is a point in the see in the your left image left camera right you act if you knew the fundamental matter we do not know F as yet it is basically 3 X 3 matrix and this is 3 X 3. Now if you had the fundamental matrix with you then if you simply applied it on a point on the left image then it will tell you to which a people or line it maps to in the say right image it does not give you the point in the right image that is why I said homography relates a point to a point here this will take a point map it to a line right which is the which is right epic polar line corresponding to that point which also means that if you are looking for a correspondence search you should be searching along that line and that line can be at any arbitrary angle in the second dimension right. And one more thing that which you also understand is that if the see right I mean in this case we said that x ~ , is actually a corresponding point of this one x ~ right in your say in your right image what does that mean that actually means that means that if you took dot product of x ~ prime right with this line ok that should be 0 because that point should lie on this epipolar line if I knew the correspondence right. If I knew the correspondence that x ~ maps to x ~ prime then x ~ prime should be somebody that lies on this on this epipolar line right because it is it is like right one of those points is my correspondence I and I actually know the correspondence. If I knew the in fact, in fact, right this would be true for any point on that line any point x ~ x ~ , that lies on that line will have to satisfy this in particular if if that is a corresponding point point of say right x ~ then yeah right I mean so what you can write is you can write x ~ , if x ~ , is actually a corresponding point of x ~ right then F x ~ is equal to 0 right because this is a transpose x ~ transpose F x ~.

So, this is you write dot product of you know right of mean of of I mean simply a dot product of line with that point right. So, it is just making sure that if that point lies on that line then it is a dot product with that line should be 0 ok and and and this equation that is why I said that we should not think that right it is actually a relation between 2 points or something it is more a relation between between right between between an epipolar line and a point ok. And and this also right leads to certain certain right what you call you know interesting properties for the fundamental matrix and and in the same way right I mean if you wanted if you wanted if you wanted to come from the yeah by the way right I mean if you wanted to come from the right to the left then what would happen I mean if I had a point there let us say x ~ prime and I wanted to find out right what is that is epipolar line I have the right fundamental matrix. So, so right one way to look at it is take the transpose of this right. So, that will be like x ~ transpose what is this f transpose x ~ , right and you can think of x transpose x ~ , as as the right as the epipolar line in the in the left image plane right and and this dot product with x ~ right which is a corresponding point should be 0.

So, in that sense right. So, this is f is a f is a fundamental matrix going from if you look at the ordering f is like p l to p r right f is like f takes a point from the left image plane from the left camera maps it to an epipolar line on the right camera. So, in that sense I mean if you want to just remember the ordering right f is like p l to p r whereas, f transpose is like p r to p l. So, f transpose will be like taking a point at the in the in a right image plane and then get a right mapping it to an epipolar line in this left image plane. So, both so right this is also fundamental matrix that is also fundamental matrix except f transpose operates between p r and p l f operates between p l and p r.

There are there are there are a couple of properties of this right which yeah which I think which I which we can go through. So, first claim is that the the epipole right is in the is in the see is in the null space of f. See f is already a fundamental matrix and and and right because it has been it has been arrived at using homogeneous coordinates f itself can only be only be found up to a scale factor. That means even though there are 9 elements sitting inside it actually only 8 are unknowns because you can only find it up to a scale factor. But now there is one more thing right which is now entering which is that the epipole in this is in this is in the this one null I mean it is a basically a null vector of vector of f right.

But how do you how do you arrive at this? So, sorry yeah yeah yeah, but you need to show that right I mean you need to you need to show that right I mean f acting on on e right should be actually 0 right. So, now, in order to show that right so, so what what this means is that let us say right for so, this is the so, for any point let us just kind of look at this so, for any point X ~ other than other than this epipole ok. That means in this in the in the left image plane right if you take any point other than this epipole E ok. We know that the epipolar line the epipolar line is L , is equal to f X ~ and this is fine. And we also know that what do we know see for example, for any any point right X ~ other than the epipole so, so right.

So, I am here right and and let us say right I mean I have my I have my camera here right and then and then and then you have the other camera here. So, I have my I have my right epipole here and then similarly right you have something like an epipole here ok. So, you have you have this and when you actually cut this right I am not I am not showing it precisely, but then right wherever wherever this cuts this image this diagram is not good the other one was was better ok. Now, you have this E right now what it is saying is we know that the right epipolar line L , is equal to is equal to write f X ~ contains contains what? Right epipolar. The right epipolar it should be E , right contains the epipolar contains epipolar E , right.

Why why is this this true because because for example, right I mean right I mean you know right any point you take here it will it will actually give you even say epipolar line here and because of the fact that all rays from right wherever they come. So, for example right I mean you have the camera center here there is a ray that is coming here because of which you have a point here there could be another there could be right I mean you know another ray right which comes from there that also that also has to come through the through the through the camera center. Therefore for this point right you may have an epipolar line for the other point you may have an epipolar line, but all these all these epipolar lines right I mean because of the fact that because of the fact that they all have to kind of say converge right at they all have to actually converge at this E , right because they all have to contain E , right. Because all of them meet at this camera center right and therefore, one of the points that you can always project from C to C , right is this guy is this E , and therefore, E , will sit in every one of them right. So, whichever x and we are kind of deliberately avoiding taking E because sorry right we want to show it in general.

So, for example, right I mean x ~ other than the epipolar we know that the epipolar line contains this epipolar epipolar this one right epipolar E , or in other words E , transpose right f x ~ right is equal to 0 for all x ~. If something like that happens right what does that mean that basically means that E , f transpose right should be 0 or in other words you can write this as right this is f transpose right E , right is equal to is equal to 0. So, which basically means that means that right E , is in the say null of f transpose similarly right if you just operate the other way right you can show that f E is equal to 0 or E is in the say null of f. So, what this what this is now further saying is that that your f right originally you had 3 X 3 that means there were 9 unknowns first of all homogeneous representation therefore, only say only say 8 unknowns and and right because of because of the fact that because of the fact that right the see epipolar should be in this say null of f therefore, it actually becomes rank 2 now and because of which your unknowns actually reduced to 7 instead of instead of 8 for f. Now one one interesting thing right which you can actually do is the do is the following right I mean see I mean what this also means is that see some people talk about this view synthesis right.

So, it is like saying that I mean once you understand this geometry what it actually means is that right suppose suppose I have a camera here and and you know and I have a scene right and yeah. So, let us say right I have some scene point here and then here is my camera. Now suppose I asked you right and ok right assume assume that I have a second camera of course, we still have not shown how to compute depth right we have to do that, but, but then right assume that I have the stereo pair and I can compute you know a disparity we will show that right we have not shown that yet, but yeah right we just have to search along the day before line to kind of see get that this is disparity. Now once you once you get actually a depth map right which which which which basically means that means that right from this from this camera center right what is the what is the what is the z value of of right to each of these points ok. That is what you want right I mean eventually that is what you mean by actually depth.

If you have that right you can even ask for example, right I mean you know if I had virtually right if I actually place a camera somewhere here I do not even have a camera there, but I can virtually place a camera here and I can ask what will this actually seen look like right that is called get a say novel view synthesis right. So, it is like it is like a view right that you have not seen before the camera has not seen before, but you want to get a say do that I mean how would you do that. See you go like this right. So, so you have right what do you have when I write I mean X ~ right you have like P left right I 0 right X X ~ which is actually a 3D coordinate and right and this guy right which you which you want some X ~ ,. So, you want to know right where this point will go see what does it what does it effectively mean this is not a homography anymore right this involves a depth now this involves a full blown 3D now you cannot use a homography to get there, but you can still ask the same question there we did know we said that we wanted to synthesize a new image can I rotate can I do all that right that we did.

A similar question you can ask here can I can I can I go from one view to another view, but then the scene is 3D now. So, those are now the point is right. So, the point is I want to know where will this X ~ map that is like say X ~ , and I want to know what is the intensity to assign to it. The intensity assignment looks kind of straight forward because right because of my I know the I know the intensity at this X ~ and if I know right where it maps then I can actually map that intensity there, but if the point is right how do I get there. So, if you if you look at it right now if I know the R and T because right that is something that I have to give I have to say that I keep the camera at some R and T right that is not unknown it is like view synthesis.

So, I say that I pick my R and T I say that I keep my camera there. So, which basically means that my this is k l this is not p l this is k l and then in this case right what do you have you have like k R and then R T and then you say right X ~ ok. Now, given that given that right I mean there are actually right two ways of doing it I mean one the I mean the let me just give you the intuitive feel for this. So, what this actually means is that I mean X ~ because you now know right because of the fact that 3D is already computed right. So, you know your you know your say X Y Z corresponding to every X ~ here this is if you have the depth map I am saying this is not like without the depth map that is why I said I have a stereo pair of computed my say depth map and having computed the depth map it allows me to now synthesize new views now.

So, what this means is that means is that right I mean no no if I if I kind of if I kind of right have this have this is the have this have this is the X ~ right here and if I if I kind of know these values which I know in this case right because I know right what is my X Y Z and I know my R and T I know my I know my say right camera intrinsic I know where exactly where exactly it maps and then the and then the and then the intensity that I can get get because of the fact that I know the intensity at actually X ~. You do not have to go through actually 3D I mean right you do not have to go through X ~ you can actually map them such that you can go from go from go from you know go aX aX these coordinates through the image coordinates also you can do, but I just wanted to mention this that right once you have a 3D right it gives you so much power that is the reason why let us say people do people do all kinds of rendering and all right they show that right I mean the object will look like this from here look like that from here all that is possible because of this right all that you need is the pose of the camera right. So, once you have a depth map of something right then all that you in order to generate a new view corresponding to or generate an image corresponding to new view all that you need is a depth map on the pose with respect to some sort of you know this one a reference position. If you have a reference camera with respect to which you have a depth map then you can always seek a new image corresponding to a new view all that you need is a depth map and this pose ok. If you can if you can do that then you can actually synthesize new views and all which probably you would have seen somewhere read somewhere, but this is the this is the underlying map that actually governs it and I leave it to you to show that you can actually relate them through through X ~ and X ~ prime also it does not necessarily have to come through X ~ even though right I mean you can kind of see that right from here you can I mean all that you have to do is you know write up X ~ as X Y Z 1 and then then it will become like R X and then you can invert it right you can do all that and then just substitute for X Y Z and then you will get everything related in terms of in terms of that X ~.

So there at what you what you tell is X ~ prime it will automatically give you it will automatically give you X ~ prime expressed in terms of X ~, but through the pose matrix and through the through the camera intrinsics that you will give you will kind of directly get X ~ prime without even going through you are indirectly going through X Y Z because you have that right that is the reason why you are able to do all of this, but you do not have to do it through this you can directly relate just like you had a homography thing where a two image coordinates even for this called this called this called view synthesis right. So, even view synthesis and this is actually called warping it when if you kind of remember I had used this warping and I had said as loosely that you know actually when you do homography it should not ideally be called warping right warping actually means 3D and pose that is when people when when somebody refers to warping is meaning that there is a 3D world and you are trying to try to warp one image to another and that warping cannot happen with a simple homography. That warping can only happen like this it can only happen with the camera poses and then the and then actual actual they see right the I mean 3D structure of this you have the geometry then you can do ok. We will stop here and then you know I just want to take some special cases which I thought we will do which we will do next class.