Modern Computer Vision

Prof. A.N. Rajagopalan

Department of Electrical Engineering

IIT Madras

Lecture-59

So, let me just start by one last thing right that remained, ok from the previous one was this in terms of this homography is this rotational homography, ok what we saw was actually a general homography but there is a special case. See as I said right in the earlier class I said that this planar homography and all this for planar scenes right. But suppose you had a 3D scene then one can ask a question can I sort of do something and still be able to stitch images together create a panorama, the answer is yes but under the constraint that you only purely rotate, ok you do not do anything else, ok. So that is like so your h min, so which basically means that means that you know see ok this equation that we have not kind of could have see you know shown this I know that and for a reason. So here right so this so another way to kind of look at it is, is that your min, so basically means that you have see translation should be 0, ok if you do not translate at all then you have something like kr k inverse, ok and that is one special homography where you can actually stitch images even if the scene is 3D. So basically the 3D queue itself comes only if you translate, if you do not translate there is no 3D information, ok if you simply keep rotating right there is no information about the 3D scene.

So in that sense right if you, if you want to stitch images like I said right if there is a parallax then there is no single homography that works but if you are only rotating then it is ok you can have a 3D scene you can still stitch images, ok this is a very special case but there are situations when let us say that people do use a rotational this one homography, ok so this is applicable even to see 3D scene something that we had excluded in the earlier conversation completely right we had said that things have to be planar, ok no I mean 3D business but here applicable even to see 3D scenes. So you can even be near ok does not matter it does not mean that right you had to be far off or anything. The other way to kind of look at it is if you are, if your D is actually you know how far away so if you are very far away then you can approximate even if you translate a little bit you can still say that most of it is only a rotation but that is another way to sort of say that you know if your translation is small relative to this depth of the scene that is again another that is a way to say that you are dealing with a rotational homography. Another way is to say that T is absolutely 0 I do not move at all I just rotate in which case I can be as close as possible I can still stitch ok there is no see 3D there is no parallax ok so this is just a special case ok.

Now let us move to the next thing right which we want to do that is about camera intrinsic and you know extrinsic, intrinsic so intrinsic means internal to the camera extrinsic means you know external to the camera ok. So now we will actually you know examine in more detail what that camera matrix is and so on right I mean till now we have been just assuming that right something will happen and then you will get an image and so on but now right we will actually look at the image formation model and see as to what things you need see for example right when somebody says that I have a camera and suppose I you know I could have image a 3D scene then kind of right can you tell for example if I have a point in 3D and right I mean suppose I have a camera here then can you tell where exactly right will this map on my image plane right that right to that level of detail we need now because till now we have been assuming that we take a camera we get images we said we can compute features all of that was never bothering about where the right how the image will get formed or where will it get formed where will the coordinates map right and all of that which is more like a forward thing this is like saying that given a scene where will the points come then once you understand that then we will ask the reverse question right given the images can you tell where did where did the points come from right so the first thing is to go the forward way because without that we cannot do the backward thing okay. So and this involves both I mean so for example where a point gets mapped in an image plane depends not only on the on the on where the where the coordinate is where the point is of course you know it will it is a function of where it is but really where it gets mapped right it depends upon depends upon the intrinsic of the camera that means what is the focal length of the camera and there are there are few other things also which are actually relevant for a camera you may think that probably focal length is only thing but there are few additional parameters that you need if that you may require I mean right if needed and then the other thing is in terms of reference right I mean you know for example for example right I mean you know if you kind of see change your reference and sort of shift your reference to a different coordinate system let us call it a world coordinate system instead of see typically the camera coordinates the coordinate system is centered you know around the camera center but if you sort of if you do not do that and if you say that you know I know only with respect to with respect to world coordinate system okay then you still need to be able to sort of do all your calculations and still be able to tell where this point will come if its coordinates are expressed in terms of a different coordinate system right. So both things we need so the other thing is external because that is like saying that I have a coordinate system it does not matter it has nothing to do with the internals of the camera the internals of the camera come in here when you want to say that right I mean a point P where will it come right on the you know image plane. So what this image plane and all that we will first see now so the way to kind of look at it is like this right so let us say let us say that we first kind of see draw the okay now let me first say that let me first say that I have the optical axis of the camera okay and the way we normally choose the X, Y, Z and all that is simply to make our own life easy.

So this is the optical axis of the camera okay what the optical axis is the optical axis is something that is that is that is actually perpendicular to the image plane. So it kind of goes through the goes through the center of the image plane so the image plane is where the is where you get your image okay and your and your see camera center right could actually be could actually be somewhere here that means a pinhole right could be somewhere here okay right and you are actually watching some scene right which is there let us say right I will say I have a point P okay then what will happen is as you know right of course and I have not drawn this very very correctly but then right what you can imagine is imagine this is a ray right that comes from P and then kind of hits the image plane here right. So this is called the image plane by which we mean that right this is the plane on which the image is formed and this is your F this is the focal length of the camera so that is why you get this telephoto and all those right different different lenses that depends upon how far away is the pinhole from a Z is the image plane from the lens center from the optical center right this is called pinhole this is also called the camera center this is also called the optical center and so on okay that is the one through which the ray is actually passing okay. Now the point is so okay so right this is your F and that is called the this is the focal length. Now what we do is right instead of actually you know looking at an inverted image we do not see inverted images in our camera right so what we will normally do is so we will we will actually okay now the X and X and Y coordinates are chosen such that such that right I mean you know we have let us say right I mean so such that your XY plane is exactly parallel to the image plane so I mean I hope I hope I can draw that reasonably well so it is so so right it will be it will be like this so it is like it is like you have chosen your XY plane like this your Z axis will align with the optical axis okay that is how we choose it just for just for convenience okay so this is your X okay let us say this is XC okay I have a notation which will follow and let us say this is YC and this we will call as ZC and the ZC okay aligns with the optical this aligns with the optical axis or we align it so we kind of treat our Z to be to be kind of perpendicular perpendicular to the right image plane and the image plane itself is exactly parallel to the XC, YC plane that we have right so you have this in your head now right you have an image plane and then you are to say XYZ is like this X and Y on this plane parallel to the image plane and Z is sticking out okay.

Now when you have when you have a situation like this right so what we okay now as I said that we do not so instead of looking at an inverted image we just kind of think of the image plane as being on the front but I said in a real system you do not put the image plane on the front it will block all the rays right so this is only for convenience so we so we will actually look at the image plane as being here okay that is F in the front exactly F away okay instead of being on the back we will kind of put it in the front because then it makes our analysis easier right everything looks the way we see them okay. Now the point is you have you have a coordinated X, Y here right that is the image coordinate right and you want to be able to able to kind of relate it to this guy which let us say let us call this some XC, YC, ZC okay what this means is that the coordinate the coordinate of this point P with respect to this camera center right so your coordinate system is now is now centered around around the optical center okay that is where the coordinate system is now and with respect to that right this guy is like XC, YC, ZC which means that XC, YC so you can imagine right that is on a plane XC, YC and then ZC is this right from here to here is ZC how to draw that so let us say from here to here is actually ZC that is that capital ZC. So that is also very typically we infer that as depth right I mean how far away right that is like that is our notion of depth actually typically and of course you know eventually okay we want to we want to look at depth and all that right that is the goal but then before we do that you know this is this is important okay so we have something like ZC and therefore right I mean you can simply do a I mean right what is that you can do a simple similarity of triangles right and you can see that right okay now I mean I will just do it for say one dimension I mean I can only plot on 1D right so you can imagine it is all happening on a 2D. So what you can say is so this X by F right is equal to XC by YC ZC right so this angle right that you are seeing okay that is formed by these 2 triangles right one coming like that if I draw like that you cannot you can see right so one is this another is this right so if you look at it so this angle is tan of that is of course XC by F which is also the same as capital XC by ZC or in other words the image coordinate X is F into XC by ZC right and similarly right you can also write that Y and this X and Y are image coordinates Y is F into YC by ZC okay and a lot of the things that I said last time right in terms of how the scene should be and all of the all of that relates to how the ZC behaves okay but you know we will not go into that what happens is ZC is a constant for a scene and all that that will actually dictate the way the image gets formed but this is the most general thing. So that we each point will have its own XC YC ZC coordinates and this is these 2 equations right this is called the perspective projection equation right and this equation is the one that makes you know parallel lines appear to kind of meet in an image plane and so on right when you see an image by railway track right if you capture it looks like they are converging right all that effect is coming right because of this.

This is the real model a perspective kind of see projection you can there are different camera models okay but the most general is this. Then okay now again right like we did last time we want to be able to relate the image coordinate to the scene coordinate right through a matrix kind of a form therefore let us again right go back to our homogeneous kind of a notation.