

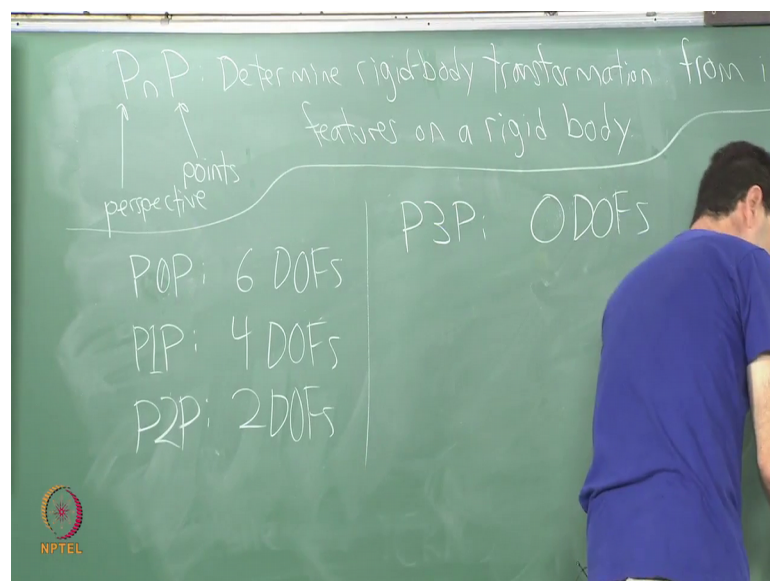
Virtual Reality Engineering
Prof. Steve Lavalle
Department of Multidisciplinary
Indian Institute of Technology, Madras

Lecture – 14
Tracking Systems (perspective n-point problem)

Welcome back, let us continue onward. So, in the last lecture we were talking about tracking systems and I mainly covered the case of orientation only tracking. This is very useful for a fully portable virtuality headset, and where you are completely relying on inertial measurements, inertial measurement units ah, which gives you gyroscope readings and (Refer Time: 00:38) readings and (Refer Time: 00:40) readings as well as. As some of you have asked, some you have asked, do not, you also need to take into account the position of the head.

So if you do the motion back and forth like this or in general if you are tracking other rigid body such as your hands. Then, you would like to have position as well as orientation. So, we talked about different cases. And remember we were mentioning that there is a case of line of sight visibility that was I think we finished on last time. So, visibility or line of sight you have some features that are in view of a camera, and then you would like to figure out where these are in the scene.

(Refer Slide Time: 01:17)



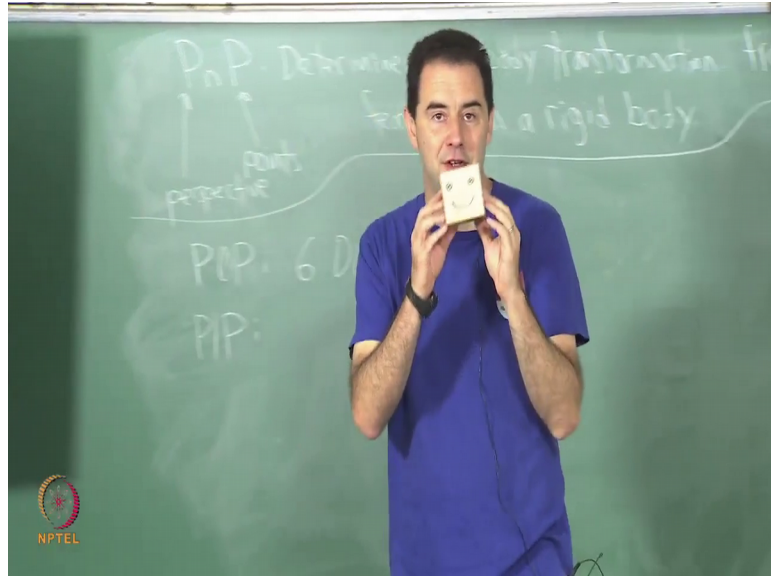
So, this leads to a very generic and well-studied problem called the PnP problem. The first P stands for perspective, which is just the model of the camera projection. Remember that, I said for each feature that you can see in the world, this corresponds to, as it strikes through the image plane, a ray that you can narrow down that feature to in the world frame, and then the second P corresponds to points. So, perspective end points problem determine. The rigid body transform.

The same transform we have been doing all along here from identified observed features on a rigid body. So, in other words, let us start with rigid body. So, in the meanwhile go back to this cube head that I had before. So, I may have particular features on here and you know, perhaps it corresponds each one of my feature may correspond to a corner of this, while I can imagine putting a bright LED on it, and then based on where these LEDs appear in the image I have the problem of this figure or this, this rigid body had moved somewhere, I want to figure out what its position and orientation is, what is given to me is exactly each one of the LEDs.

I know the coordinates of the LEDs in the body frame and I also have labels for them. This is LED number 1, may be on this corner. This is LED number 2 on this corner and so, forth I have labels on all of them and when I see them in the image, I can recover the labels; that is what identified means. So, how can this be done in practice when I could have different colored LEDs right that would be one way to distinguish them. Another way to do it, is to have them flashing in some way and make some kind of code overtime. So, if I see them in multiple frames, they could flash between a light mode and a dark mode or they could switch between two different frequencies, and you can give some kind of coded signal to identify them over several frames. So, if things like that could be done.

Let us think about different versions of this problem and I like to, I like to talk about degree of freedom. So, how many degree of freedom do we have for the rigid body before we see any features 6, full 6 right? Alright, so I guess I could say that if we just have $6 - n$, where n is the number of points. So, if I have zero points seen, then I guess I remain with 6 degrees of freedom for the object. If I observe a single point in the image I just want a reason about the degrees of freedom now.

(Refer Slide Time: 04:52)



So, for example I have this object and I can state that one feature is fixed. So, suppose its this corner, I have observed this corner, I have identified it in the image. So, the question is, what can this object do now, while keeping this corner fixed in an image. How many degrees of freedom have I lost?

Student: 3.

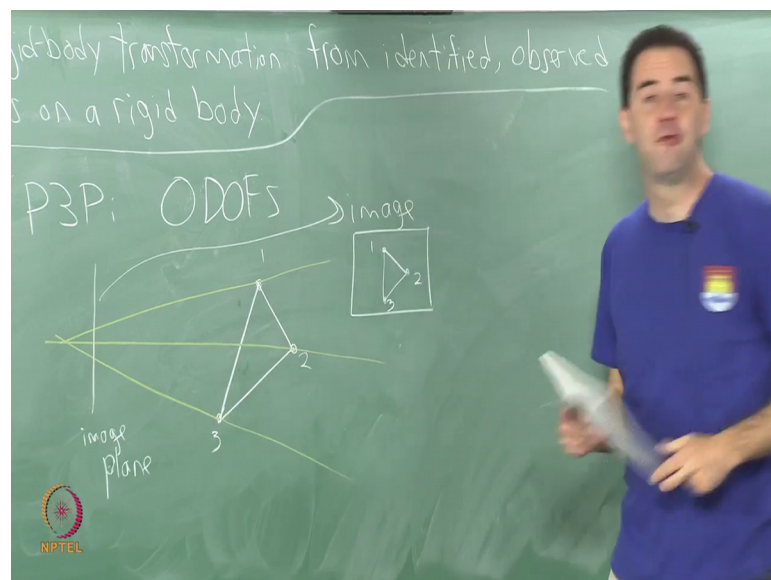
Have I lost 3, I see. So, can I do any rotation of this? Now not any rotation (Refer Time: 05:18), this would be blocked right,, but let us just say in terms of degrees of freedom analysis, we just do small prohibition. So, I should be able to do any yaw pitch and roll. So, that is 3 degrees of freedom, still intact correct, but what if I make this thing closer or further from the camera, moving exactly along the perspective projection line that goes through the point. Can I do that too? Yes, I think I can do that. So, that leaves 4 degrees of freedom for this.

So, what is changed is that this one point. Did I get it right? This one point can no longer go this way and can no longer go this way, because it has been fixed by the pixel in the image. So, in some sense you can imagine the I J coordinates or the two constraints right. The I J coordinates of this point in the image or the two constraints each one of those drops has degree freedom. So; that means, that there are 4 degrees of freedom left after one point. So, what about perspective two point problem? How many degree of freedom are going to be left? Who thinks that there is a pattern here maybe?

So, if I hold two points fixed, maybe I take these opposing corners here and say that they are fixed. I do not know if I can hold on to this very well. So, I guess in that case, it looks like I could spin it like this right. Is that only one degree of freedom remaining then, Yes, it is somewhat difficult to see,, but I should be able to also do some kind of transformation, where I move this back and forth. So, that the angle between the two rays here right that stays fixed. So, I am kind of, as if these two features are moving along rails right. And I can go further back while reorienting this; that is another degree of freedom. So, that seem ok. Try it at home with your own setup.

I think I could do it here with a piece of chalk if I just. So, it is like this, it should be, I could have it like this right and I can go like this correct right. So, that is the actual degree of freedom. Try it again here if that is a, I can go like this. So, we get this (Refer Time: 07:33) degree of freedom um. So; that means, that there are two dots left here. So, we get $P \ 3 \ P$. And as you might imagine we are down to zero degrees of freedom remaining. So, there no more degree of freedom remaining. In that case I have a picture that looks something like this.

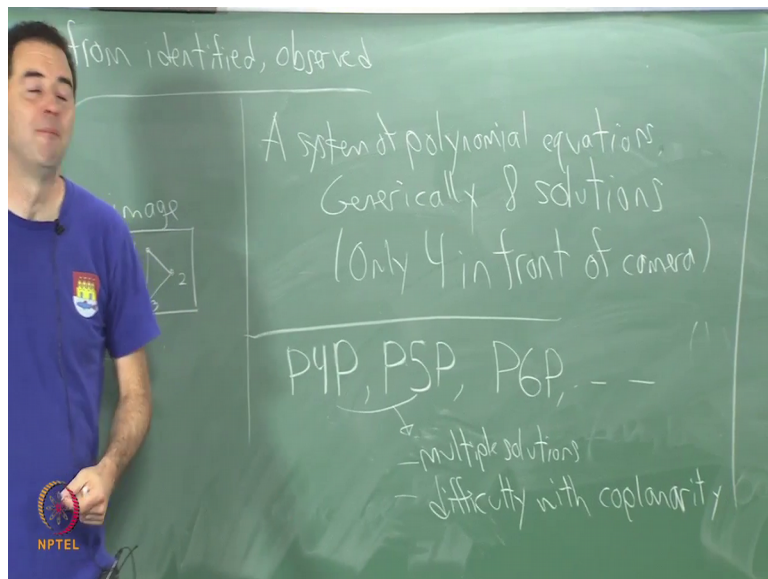
(Refer Slide Time: 08:09)



I have imagine a rigid triangle that has my, that has the features on the corners 1 2 3 and then (Refer Time: 08:18) all being observed to some image. So, I try to make these lines on me here. So, there is an image plane where these are being observed. So, I get these three points observing the image. So, somewhere here in the image.

All I am really imagining is that, is detection of 1 2 and 3. So, they are labeled, I know which point is which and then I can pin this down, except it is not the complete picture. So, I know to figure out exactly where this triangle is located, involves solving some equations. So, you have a system of polynomial equations to solve. People have solved this one um, you would find solutions all over the internet and books, but the interesting thing that comes up when you try to solve this, to figure out the solution, is that you find out that there are generally 8 solutions.

(Refer Slide Time: 09:38)



So, you get a system of polynomial equations and there are generically 8 solutions. It turns out that 4 of the solutions would be on the other side of the lines here. So, um. So, I have 4 behind the focal point and 4 in front of the focal point. So, in reality you get only 4 in front; that is if you set up lines in your system of equations here. So, only 4 in front of the camera. So; that means, that if I have a fixed rigid triangle and I make a kind of a house for that, a cage for that, these yellow lines (Refer Time: 10:37) means that there are 4 different ways that I could stick the triangle on there. So, that the vertices of the triangle touch these lines. So, you can also try that for the home project. If you like make a pyramid out of straight sticks, and see if you can fit a triangle in therefore different ways. So, the order that works out for that in order to remove redundancy.

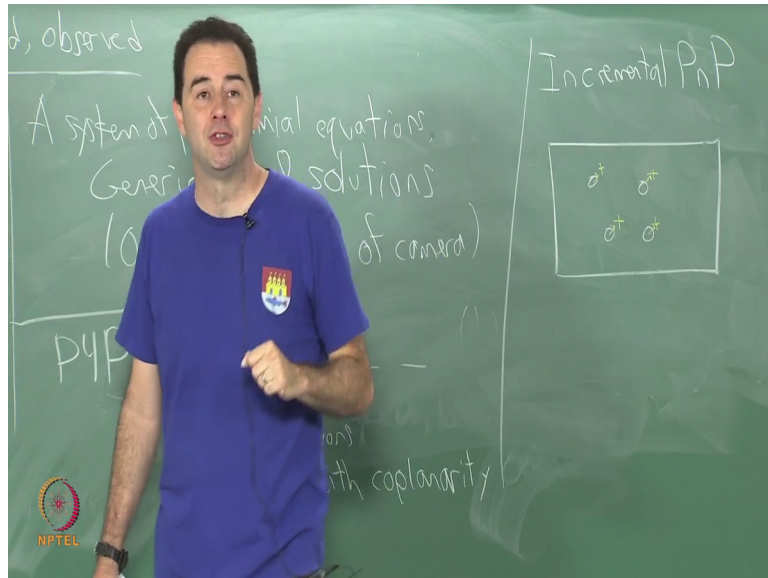
You just go further um. So, if we get up to P 4 P P 5 P. There is still some redundancy. So, there is still some multiple solutions and problems with co planarity, and there is a

special be true in actual systems, where observing this points or an image due to (Refer Time: 11:52) and noise you may end up with several plausible solutions out of very close in terms of image coordinates,, but they might be quit far away in terms of the orientation of the triangle. So, has also been almost solutions there, absolon solutions in this range, and so eventually you get up to P 6 P and higher. The more points that you observe the better and the further away from being coplanar the better.

So, this provides greater distinguishing power. So, on the (Refer Time: 12:26) said two heads set for an example here. There are I belief about 40 LEDs and they are distributed all over the place not necessarily co planar um. So, um. So, this gives a lot of discrimination power in terms of resolving the position and orientation of the head set, when it is in front of the camera um, you do not see the LEDs, because they are hidden behind infrared transparent plastic. So, we cannot see through the plastic, but if you could see the infra-red spectrum then you would see the LEDs you can also look at it with an IR camera. If you want directly and show the LED lighting. Alright this makes sense.

So, we get enough information, we get enough equations to solve um, I do not want to go into what happens to um, you know how to solve each one of these equations. You start entering into a subject called computational real algebraic geometry um, it is sort of like the polynomial generalization of linear algebra. So, you know linear (Refer Time: 13:27) using your algebra, the whole method for polynomial system solving. They become very useful in various fields in engineering (Refer Time: 13:36) motion planning. For example, they show up um, they also show up here in these case of these kinds of these solutions um. One thing to pay attention to is you call incremental PnP which is suppose I am looking at the image and in one coordinate frame.

(Refer Slide Time: 13:47)



I have the features and I have an estimate of the position and orientation, and then in the next coordinate frame I notice that these features move by some small amount and I can still, I still have my identification going. So, these features move by some small amount. All I have to do is, slightly update my estimate of the position and orientation correct, but because in one frame time. Let us suppose your camera is running at 60 frames a second. So; that means, your heads moving, it cannot move too far in 16.67 milliseconds. So, there is this tiny change in the image here and all I need to figure out, what is the change in position and orientation. I have the gyroscope and (Refer Time: 14:46) which I can use as well to make a good estimate, as far as how much change there has been in the transformation, and so to do the final bit of correction, I could just perform a very simple local optimization.

(Refer Time: 15:01) this object I could imagine, imagine using the perspective transformation that I am just (Refer Time: 15:08) this object in the virtual world that I am trying to track which is the head. I am trying to track the head and I want to just want to do perturbation to that the new position orientation I am estimating lines up perfectly with the observation. So, I can just proceeding incremental fraction like this, like an incremental optimization while just do a little bit of perturbation in each step, and it ends up being fairly straight, straight forward, because it is not as hard as completely from scratch. I show you the head set where am I have to figure it out by solving equations from scratch. Once I know the solution it does not change by very much from one frame to the next. So, that is what that part is questions about that. So, that gives me a new

measurement. Now it is another kind of drift correction information which is the updated position. So, a position estimate that is based on line of sight visibility. So, now, I can use that again in a filtering method. Let me just give a little bit of description of how that works and then we can move on to another tracking technology called light house and then I will be done with tracking methods.