Modern Computer Vision

Prof. A.N. Rajagopalan

Department of Electrical Engineering

IIT Madras

Lecture-58

Now, what this means is the following okay, I think we spent too much time on that I was not planning to spend that much time okay. Now so now if I write down the first equation right, so λ XT =H11 XS + what was it H12 YS then H11 H12 okay + H13 right. Then λ YS on the left will be H21 XS + H22 YS + H23 and then λ right third will be this λ is H31 XS + H32 YS + H33 right this is what you have. Now if I want XT right, so XT I can get as λ XT by λ correct and of course all of this λ is assumed to be not equal to 0. So λ XT by λ, so that is like H11 XS + H12 YS + H13 by H1 sorry H31 XS + H32 YS + H33. Similarly I can say I need if I want my YT that will be λ YA this should be YT right not YS right λ YT.

So λ YT by λ that is H21 XS + H22 YS + H23 by again the same λ which is H31 XS + H32 YS + H33. So now if you simply do a X multiplication what do you get, then you get like H31 XT XS + H32 XT YS + H33 XT =H11 XS + H12 YS + H13 and then for YT you will get H31 XS YT + H32 YT YS + H33 YT =H21 XS + H22 YS + H23. What is then means is that see for example I can actually think about my unknowns as H11 if I stack them as a vector H12 H21 no H13 H11 H12 H13 H21 H22 H23 H31 H32 H33 right. This is my this is my characteristic vector then on the left hand side right you will get this is actually 9 X 1 right okay.

So you will get actually 2 columns right, so what this means is that so the standard way to kind of say represent this is like - so yeah so you see bring whatever is on the right hand side right here so it will be like - XS right. So that is what it is multiplying H11 then - YS that is multiplied I do not know whether I have space write this then H13 is - 1 and then H31 so it is like 0 0 0 right because you do not have anything for H21s I mean I am just I am just I am just rewriting this one right in the form of matrix vector and then oh wait a minute H31 right so that will be XT XS by XT into XS okay by the way H32 is XT YS and H33 is XT and similarly for the for the second coordinate what will I have so I will have H11 does not even come in right so I will have like 0 0 0 then I have H21 for which it is - XS then - YS then - 1 and then what is it H31 is XS YT okay I mean this became small okay but I hope you are able to follow and then you have got H32 is YT YS and finally YT this is okay right this is rewriting these two equations in a kind of a you know matrix vector form because this is the unknown this is the this is the homography matrix represented as a as a 9 X 1 vector. So now this matrix on the left is actually A and what you have what so what this means is that if you had if you had one point correspondence right between the two images that means right what this means is that see when I say XT YT XS YS that means that that for the source and the and so suppose I had image I1 then I had image I2 and suppose I wanted to know what is the motion between the two by motion I mean this homography H not the camera motion right how do I relate the two then what it means is given one point here and then if I know this one point correspondence I can get actually two equations correct given one point correspondence I can get you see two equations and and right because of the fact that because of the fact that I have only only I have only 8 unknowns in my H so what this means if I have 4 point correspondences this I write for one point corresponding give me one more point correspondence which you know which you know now to do you shift or something get shift or get one more get a third get a fourth okay then that it means that right I have already got you see 8 X 1 if you give me 4 point correspondences I have got 8 equations there which then means that I already have something like AH equal to 0 where this 0 will be what 8 X 1 okay this is 9 X 1 this is 8 X 9 okay. Now ideally right you should be able to solve this right I mean so and then right because of the fact that fact that right A is 8 X 9 right so it clearly means that there is there is one courtesy non-zero H which actually A should actually map to 0 right and that and that and that and that I mean what you call that is a non-zero H will be your will be this will be this is your homography and it can only be found out to a scale factor because A λ H is also equal to λ H that will also be equal to 0 right so the H you can only find up to a scale factor which is okay right that we know right from the start. So what this means is that if you had 4 point correspondences coming from whatever you have studied already if you could just run sift or something between the 2 images and if you have this point correspondence then it means that you can actually you know what you call de-warp right in a sense right so you have the other image which is warped version of this typically we use warping for more you know this one more what you call you know sort of a difficult case but let me just say that this is geometrically transformed and you can apply an inverse transformation right so that DC2 are aligned because why do you want them to be aligned I mean you might want to stitch them you might want to find a change whatever you want to do unless you align right you cannot do right any of that and therefore this alignment so as you can see right did we ever bother about what is the camera translation what is the original camera orientation we did not at all have to bother about that means there could be some people who are interested in all of that that will mean going back and solving an inverse problem but we are not here for that right we just want to know if I so that is what I said so by the end of this class you should be able to take your mobile capture images like this click click click right take 5, 6 images whatever you want and be able to stitch them all together to get a get a get a panorama and that will come provided you are able to run sift that is part of your next assignment and then you get the point correspondence and then you should be but then whether we do not stop here actually right this is not the way okay it is actually done so what is done is because of the fact that these point correspondences can be noisy see most of the time it is not exact because you are running a sift and then it could be it need not be exactly correct right so if you just take this this would be an exact solution provided those point correspondences are absolutely accurate then it is okay then you can do it this way but there is no guarantee that what you have is going to is going to give you exact point correspondences because that is coming from say another method some like sift or something that is going to return and then there is no guarantee that it is returning exact exact matches so if there are some outliers right then there is a problem because this does not handle outliers okay so to handle outliers there is something called what is called a ransack okay and what this means is random sample consensus and this has nothing to do with with homography and anything this is something you know which is which comes from the CS literature that is actually a nice idea what it says is anytime anytime that you have a set of points right and then and then you have a model and you and you want to and you want to find out you know you know a best fit to the model then what you should do is and so for example you know so for example ideally you should have had a straight line right which you have to fit but then but then it have these noisy points then how best to kind of see fit this line right and therefore if you knew what was the model of course you have to know what the model is then then say random this ransack what it does is it tries to tell you it tries to give you give you a proper sort of you know a systematic way to actually arrive at this arrive at this fitting of this model.

So in this case our model is actually this is your homography okay that is what we want to find and we know that the minimum number of points that we need to fit a homography is 4 okay so we start from there so whatever model you fit you have to look at the minimum number of number of number of these point observations right that you need in order to be able to fit a model so in our case right we know that we know so for example imagine that imagine that right now I have let us say right a set of points for which the feature correspondences are available and I could have I could have say m number of them earlier we were taking m equal to okay well we have m by 2 feature correspondences because every correspondence gives you 2 equations or in other words if you had m point correspondence you will have you know 2 m equations in that matrix. Now we know that to kind of start with okay we said that if you had m equal to 4 we could solve the problem but now we are saying that we could have outliers right which can spoil the entire thing therefore what we will do is we will start with so I will just let us write down the steps quickly okay because should not take long what happened okay I was writing something so pick any 4 point correspondences okay correspondences and solve for H what would you then do next so you know that right I mean if I have 4 point correspondence I can solve for it what will you do next will you just accept it you do not want to accept it you want to check how good is this and to check how good is you should basically apply it on the other points right because you have so many of them right so you could have picked at random any 4 you can compute the homography you know so you solve that H we will see how to solve H equal to 0 right at the end you can do SVD or something and get the H vector which you can just you know reshape it in the form of a matrix which is H and then apply right then what do you do so use this H to calculate so you should focus on the fact that there should be a consensus right it is like a voting so use this H to calculate let me say Xi double ‘ =okay I think yeah okay so instead of using Xt Yt right so a notation here is like X X ‘ okay Y Y ‘ okay that is the reason why this ‘ is coming so it is like saying that in one in the left image or whatever in the in the in one of the images the coordinates you are representing it as XY and then the pair that that is the feature correspondence is X ‘ Y ‘ okay now what you are trying to do is so it is like this right so what you are trying to do is you are you have you have actually so you have taken 4 of these okay for which the point correspondences are known okay so you have taken these 4 you got the point correspondence you know the point correspondence you solved for H now what you are doing is you are applying it on these others here and you know that let us say right somebody already told you see Sif told you right that this ought to go here to some ‘ okay this is like let us say X5 ‘ okay let us say right so you have got X1 X2 X3 X4 let us say this is X5 this is X5 to X5 ‘ then somebody said this is actually X6 ‘ and so on right X6 going to X6 ‘ now what we are doing is by applying H on X5 or X6 or something we are we are trying to get to X double ‘ because we do not know whether we reach X ‘ right we may not reach X ‘ X5 ‘ right we may not reach it so which is the reason why we are kind of denoting it as Xi double ‘ because what where the homography takes it we do not ideally it should be taking it to Xi ‘ I mean if it was a very good homography it should take you directly and if the shift was correct right both things have to work right so use this Xi double ‘ is equal to H times Xi what is that X times Xi where these Xi's are the are those coordinates other than the ones that you have used for computing the homography these are not the same 4 they will go because that is how you computed H right not on them okay for the U =what is this and similarly Y and Yi double ‘ =H times Yi for the for the remaining Xi's Xi's and Yi's for the remaining Xi's Yi's remaining means other than the 4 okay that you had used for computing H. Then okay now you need to see normalize okay normalize Xi double ‘ and Yi double ‘ by what by dividing the first 2 elements so it is like this right so you have got Xi so you are calling Xi double ‘ Yi double ‘ and some λ right which could be sitting here so you have like λ λ λ and this is H applied on Xi Xi Yi 1 okay that is how you need to apply okay and once you apply it then this λ right could it may be 1 it may not be 1 so we will anyway scale these 2 guys by λ so that we get the actual image coordinates want image coordinates no by dividing the divide by the because the third row whatever it yields we do not know exactly if it is 1 then of course you do not want to do anything by dividing this third element of you know by dividing the first 2 elements of HXi by its third element right. Then what you have to do we still do not know how good is this we have to still check the check right whether we are done exactly so we have to compute some error now right so find so what will you find root of Xi ‘ - Xi double ‘ square + Yi ‘ - Yi double ‘ square should not be Xi Yi right because you are going to the other image where the feature correspondence is Xi ‘ Yi ‘ right I mean here no you are going here right so you want to compare it with that with that original whatever shift gave you would have said that this goes to Xi ‘ Yi ‘ the homography is saying that it goes to Xi double ‘ Yi double ‘ ideally the 2 should be the same they may not be the same okay though find this right and if this number okay let us call this right this is equal to epsilon then if epsilon is less than a threshold okay now this is all not in terms of any intensity these are all in the spatial coordinate now okay we are not talking about intensity we are talking about how close we are and then therefore if you set a threshold let us say if you say that anything less than 1 pixel distance I am okay with it that is something that right you should take a call on so once you decide a threshold right then you say then you add it to a consensus set which means that it satisfies that means this Xi Yi satisfies my so this homography is able to push this Xi Yi in the right manner within a tolerance. So then update your update your update your see consensus so then if this happens then increment this is the consensus set increment the consensus set this consensus set is for whom? H yeah exactly consensus set for that H that you that you derived from those 4 point correspondences right and now and if the total number of consensus points if the total number of consensus points is greater than a threshold again right so these are some hyper parameters that that you should have but these are pretty robust okay than a threshold okay and so this like I mean out of those points whether you want to say that 60% of them whether they whether they are whether they are obeying the consensus this kind of tolerance then we can stop okay else continue for a fixed number of iterations until this else continue for a fixed number I mean sometimes that if you are not able to but actually you can show that you know you can do pretty well but you know in case let us say you are still going on and on and on you may just want to stop after a fixed number of iterations but take the one whichever gave you the maximum consensus not like the last one does not mean that wherever you end you take that right that need not be a good one so you should take the one that whatever worked out the best continue for a fixed number of iterations is that clear and choose the H with maximum yeah one minute let me just finish this and choose the H with the I mean with the maximum consensus and one more thing is that right you do not simply you do not simply stop so what is what is okay what you also additionally do is you do not simply take that H directly right so you know that you know that out of all of these points there are 4 points for which the H seems to be the best then what you do is when you solve for this AH equal to 0 you do not directly use that H okay I mean you have already done this AH equal to 0 got some H know what you do is you now put all the points which are actually within that set within the consensus set and then solve you resolve makes sense know because rate I mean you could I mean if you wanted some kind of you know some sort of a stability against noise and so on what will you do I mean you will try to use as many observations as you want but now these are not outliers these are all inliers right because these have all satisfied that tolerance criterion so it is like saying that I have got 4 points but then I find that out of my 100 points there are 40 more which are actually good so why should I just use 4 I will use all the 44 I will use the original 4 + the other 40 put them all into this A matrix and solve for H again okay that H is going to be you know that is what that is the H that I will eventually use not the one that I got originally just with those 4 points but I now know which H to I mean so I know I know where to go to kind of get that is a consensus set it is saying which point should I choose to get at by model right is that okay okay.

Now the only thing only thing right which remains is how do you solve this now okay now that we know that okay now let me also write that step maybe write just that you know so that do I have okay so then you know resolve I will just write this down resolve for H using all the say inliers and this and this RANSAC is something that you can use anywhere okay it is not just a nomenclature any problem that you have in fact in fact there are some latest additions you know updates to RANSAC that can even do with 5% of inliers you can still fit a model you get a 95% outliers okay with just 5% of inliers you can still fit a fit a good model okay. Now the point is right so now what you have is like AH equal to 0 this is what you solve and now A is some m X 9 this is like 9 X 1 this is now a vector right m X 1 and this m is m is much larger than 8 okay this is like now at least squares kind of thing right so now write you have a lot of observations and now there is also no guarantee that you can satisfy this now right because of noise so what do you do you will find out you will try to minimize norm of AH square with respect to H right I mean you will just find that H which will look at minimize the norm because you know you can no longer expect that that right you will be able to find an H that will actually satisfy AH equal to 0 if it happens it is good but then there is no guarantee there could still be noise right in your final shift correspondence there could still be some noise this is just in layers right it means that within a tolerance level they were all good that means there is still some scope for noise in there therefore what would you do so which means that it is like AH transpose AH right which is like H transpose A transpose AH right and therefore what you want to find is find that H which will actually minimize this is a length of AH right that is what you want to do and therefore what are the standard ways to do it is use what is called SVD which you all know singular value decomposition which is like you know taking any rectangular matrix for that matter you can decompose it as U sigma V transpose so let us say in this case you have got M X 9 therefore U is actually an orthogonal matrix M X M this will be M X 9 this will be 9 X 9 and U U transpose is identity V V transpose is identity but this will have a different size of course this is like M X M like 9 X 9 okay. And then once you plug that in here right then you will get H transpose A transpose AH if you actually plug in the singular value decomposition expression for A then you will get H transpose then A transpose is V sigma transpose U transpose then AH right so A is U sigma V transpose so U transpose U is identity right so these are all you know these are actually inverses of each other and therefore you get like H transpose sigma sorry V sigma transpose sigma V transpose and sigma transpose sigma is what sigma is M X 9 so this is 9 X 9 right and what it contains are what are called this the singular values of A okay this is something that if you do not know just going to say read it up okay if you are not familiar just read it up it contains what are called the singular values sigma 1 square sigma 2 square sigma 9 square what are the singular values actually what are they? They are the Eigen values of actually A transpose A okay they are just the basically the root of the Eigen values okay root of the Eigen values of A transpose A but square I mean I am indicating these are sigma 1 square sigma 2 square up to sigma 9 square that is what this guy will be and therefore and you can order them right you can order the columns of V such that the sigma singular values are in the order of the you know top one being sigma 1 square and the smallest one being sigma 9 square it can so happen that sigma 9 square can also be right can also be 0 okay it depends upon how your point correspondences are but that will be the smallest value okay then what do you do? So for example if you now take H to be let us say V 9 okay that is the last column of V right so if you take H to be you need an H right so if you take H to be V 9 right you can show that that is what it like minimizes now because then you have like H transpose which will then become V 9 transpose and then V and then you have got like a diagonal matrix here and then V transpose and then it is going to multiply V 9 and we know that V transpose V is identity therefore this will give you a vector which will have all 0's except for the last entry which will be 1 because V transpose V is identity you know right so V 9 is by the last column of V so V transpose is just that row right so row column that will give you this is okay right this is straight forward so I am not kind of trying to explain this and all in great detail this just follows and therefore if you multiply a diagonal with this what will you get? You will get another vector which will be 0 0 that is sigma 9 square at the end if you if you multiply this with this vector and yeah exactly and now so now V into this is what so 0 into the first column of V + 0 into the second column of V + sigma 9 square into the last column of V last column of V is V 9 therefore you will get V 9 transpose sigma 9 square V 9 but then this is just a scalar right and V 9 transpose V 9 is 1 because V transpose V is 1 therefore you get sigma 9 square and we have ordered this whole thing such that sigma 9 square is actually the is actually this the you know smallest singular value right and therefore if you pick if you do singular value decomposition of A whatever be the size and then if you take the take the take the vector the last column of V that will that is your H and then you can reshape it now right because I really want it to be a homography matrix therefore you can reshape it so the H that you got you write it back in that form H 1 1 H 1 2 H is like H 1 1 it is a vector right so write it back in the form of a matrix you want to ask something. Well see for example right because because H can be found out only up to a scalar in this case it is like H transpose H is 1 the even norm is 1 that is the constraint V transpose V you know that is the constraint because it can be only found out I mean you can use any constraint but but going by HVD whatever will actually mean that you are putting a constraint that you know norm of norm of H is 1 and then the idea is that say once you have this H right so what this means is if you wanted to stitch a panorama now what would you do right so for example okay we are almost done now so I have let us say write let us say 5 images okay which you have which you have captured with your camera take the one which is at the center okay so that the motion is you know is not too much on either side if you take the first one that this motion is going to be too large right with the 5 with the fifth one so you take the middle one and then and then you should actually create a canvas which is on which on which right you want to get a say right create this effect and and then the and then this central image right is going to be here and now what you do is now with respect to all of this right you have your H 2 1 whatever right you have your H 2 1 you have your H 3 1 you have your whatever right I mean you know if you want to call this H 4 1 then you have your H 5 1 you have all these matrices all these homographies which you can find now you know how to find the homographies given the point correspondences you can you have all of this now all that it means is go from this canvas right when you know apply apply each one and then see where it goes right I mean sometimes what can happen is there could be an overlap that means a seen a point here could map here inside as well as it could map here inside that means it is appearing in in both the views right in which case you can take simply a median of those values it could could come in multiple views no because some overlap can also happen it is not like this overlap there is an overlap then you will see that applying the inverse homography will actually take you to the same seen point in let us say right multiple images you can either take the average of those values or you can take the median of those values and then and then plant them here you continue doing this then you can actually fill up the whole canvas and that is how you get a get this white picture that you have right that you see in your phone ok and and you saw that nowhere did we need to know to answer you right we never needed to know what was the camera motion we do not need to we do not want to know ok.