Modern Computer Vision

Prof. A.N. Rajagopalan

Department of Electrical Engineering

IIT Madras

Lecture-41

Let me go to an associated topic what is called line detection. Now, let us talk about line detection. So until now we know that we have got these right edge pixels if you use a canny or something and of course it is not really deep network I am again kind of switching back to a traditional thing. So, the thing is right I mean okay so for example you could have an image right where let us say right you have these right these are your edge pixels and you know what you would think that all of these probably fallen aligned right and therefore you know the output should perhaps be more or less aligned right which I should be able to pass through I mean which in a sense right which I kind of believe that you know that these points kind of lie on a line and therefore right I mean so for example if I see an image right if I see a wall here right if I see a wall there then instead of seeing them as edge pixels I would want to kind of draw a line through them draw a line through them draw a line through them automatically right and because a line right you know would give you a lot more information than simply having those say edge you know this one the pixels. The other thing is that edge pixels do not be restricted to lines for example you know you could have these edge pixels sitting on some kind of an ellipse or something or they could be sitting on some circle or something to us immediately right it may not be you know the algorithm it is still an edge pixel sitting in isolation but then when they are brought together right they could actually give out a much more meaningful information. For example if I gave you a human face you know somebody in an image against a background and I asked you to pick out the human face what will be the simplest thing that you would do? What would be the simplest thing you would do? What will be a simplest idea? Ellipse exactly right so you would say that you know get me an ellipse whichever object looks like because the face is somewhat ellipse right elliptical so instead of looking at facial features and worrying about all that I will say right pick all those objects that can be fit into and fit into and say right elliptical elliptical equation right I mean something like that you can do so that is a kind of a you know a quick you know entry into to seeing where the faces are rather than saying you know construct something for a face and all that so I am just saying so where they can be applied how they can be applied right is all right I mean you know so I think it is all up to us.

So the idea is that that right you might want to so in this case right I mean you know I am just I am just you know taking line as an example because that is the simplest which you

can ask so there is something called a Huff transform okay you know it goes after a person whose name is I think some CV Huff or something okay so it is called the Huff transform and the Huff transforms goal is this right and you can have a Huff transform for line detection you also have Huff transform for circle detection ellipse detection we do not want to go into all of that I just want to give one example so that you understand the basic you know philosophy right so this is kind of a voting method okay this Huff transform is something like a kind of a voting method and Huff transform for line detection okay so we are looking at line detection. So it starts with this simple idea that for example right I mean if I have let me say right I mean I have a line here okay now one way to of course one of the things that this Huff transform needs is there should be an equation right a parametric equation for let us say whatever object we are actually looking for okay so it is like saying that so suppose I have the slope intercept form what is that I mean y is equal to mx plus c right this is a kind of a slope intercept form right now what I can actually you know tell us if I actually know the so for example I mean so for example if I know the if I know say m and c right then if I kind of walk along x right if I walk along x then I can find out all those y's right which actually sit on this line correct so as I can say which are all the points that actually satisfy this equation of the line so all that I need to do know is the m and c right and then for that I will as I keep varying my x I can find out what are all the y's that actually satisfy this equation. Now the right interesting thing right about this is that I mean right just to kind of right now what you say give you an intuition is that you know when you see in an image right what you do you do not see the line directly right the canny or something does not understand line so what you see is points so you see a point there you see a point there right something else in the most ideal case right they are all of course you may not see all of them right you might have actually thresholded something something would have vanished maybe after this you get a point there but there is still some point there. Now the point is right for you know through each point right I mean if I just take this point right then I can think of infinite lines right now which can actually pass through that point right so which means that there are so many m and c's right I mean so if you can think of so many lines which actually pass through that but there is something which is actually common to these points in the sense that each one of them right so for example each one of them certainly has this has this one and one is m and m and m and c which is this line right right which I have drawn the point is how do you how do you actually how do you actually tap that information right so how do you tap the information that these fellows are all actually falling on a line.

So one way to one way to kind of right do it is you actually you know create create you know a discretized space simply like you know where you can have a grid where you have you see m and c values and then and then start to see and then right I mean you know start to start to kind of right I mean you know fill that grid in a particular way I won't I won't kind of go into that because something something something more something else is

actually done because this has one problem right this slope interested form has one problem what is that problem so if you if you if you were to think about if you are if you have to think about some kind of a voting scheme right where you want to see I mean how many of these points vote for one particular m and c okay I mean if you if you have to do that then that will sort of give you an indication that there are enough number of points voting for one m and c and therefore probably there is a line that is passing through that but then there is there is a problem with that what is the problem the problem is mainly in terms of a line right for example if you have let us say y equal to a constant right. So for example if you if you if you had when what do you say so sorry I mean right I mean if you had let us say x equal to what is this so so right so so so if you have any equation like this right so that is like what x equal to 5 right that would be that would be right you know equation of this line right now an equation so a line like this will have actually infinite slope right so this m m m can m m right you know typically can be right unbounded I mean if I if I start right making this line right vertical and vertical and vertical there will because you will have lines like that also no you cannot say that no no vertical lines do not even exist in my image you cannot say that in fact many a time when you look at a wall there are so many vertical lines and therefore you have to be able to right deal with them therefore this slope intercept form is not the one that let us say people use one of the main reasons is because the slope can become unbounded and this guy cannot handle it what is the what is the what is the alternate form for a line that you are familiar with I think I am sure I mean that you must all I mean any eleven twelve somewhere you would have done it no what is called a normal form right what is the so you have something called a normal form of a line what is that what is that normal form what is the normal form of a line anybody exactly exactly right so so for this line right so from the origin right I mean you find out find out a point on the line you know which is which is the closest to the origin which is which is of course which is which is actually the line which is orthogonal to that line which is a perpendicular oops wait a minute which is actually a perpendicular and and this you call as some length rho and this you call as an angle theta right so with respect to the x axis then what you can do is you can actually write down the normal form which I am not going to show and all it is all very straightforward right you can you know this is all just a tenth eleventh or whatever rho is equal to I think what is it I mean x cos theta plus y sin theta okay this is called a normal form of a line of a line okay this is what this is what the Huff transform uses not the slope interceptor why because right I mean you know in this okay the I mean instead of m and c rho and this angle theta actually play a role okay but in the right nice thing is see for example I mean see for example if I wanted x equal to 5 which here right in order to x plus x equal to 5 I need an infinite slope whereas if I wanted x equal to 5 what would I do here what will be my my this one theta and what will be my rho rho will be 5 and then theta will be 0 right so it means that we just finite so wherever I go right whichever line I want I can always say represent even if I want x equal to 0 right I mean so if I want to want to represent the y axis which is what x equal to 0

right okay and if I want the x axis that will be like y equal to 0 whatever it is any line on this plane it can be can be expressed with just finite values of rho and theta they never go unbounded okay so that is why a normal form makes makes actually a lot more sense and the and the same thing applies see for example if I if I fix a rho and actually a theta right that means that means that means right this rho okay which I have fixed and then I have this this courtesy theta and then if I vary my x then then all those points right that lie on this line are the ones are the ones whose y will satisfy this equation correct. Now the the main thing right which this allows you to do it is now similarly you can go for some kind of a voting but now the advantage is that you do not have to you do not worry about unbounded values and so on of course you know there are some you know there is a slight variation in the way let us say right people do it so usually the standard form is to let theta go from go from let theta go from 0 to pi okay that means that means you go all the way from 0 to pi and then let let rho can be both positive as well as as well as see negative I mean you could have had theta theta going from 0 to 2 pi in which case you know rho rho right up and it could have been simply a positive number but the most standard thing to do is actually to let you know theta go from 0 to pi and then and then so for example let so your rho if it is on this side then and then then we will take actually negative sign for that okay. So this rho can rho is like this I mean this is just this one you know a construction aspect but but right what is what is what is more more what you call more interesting is this see suppose for example so suppose for example right I have this grid okay let us say that let us say that right on this along this axis I have my rho values and along this axis I have I have you know the theta values let us say I simply put the rho as 1 and then 0 what will be the range of rho by the way yeah if I have an m n cross n image what can be the maximum rho that I can go to m by m by m by root 2 and then m by 2 you know m by root 2 right so you can go like so so here I have just put it as 1 and minus 1 typically it will be like you know m by root 2 to actually minus m by root 2 okay that will be the range of rho if I have an m cross m image right because because you have this image and then and then if you want to find out which is the farthest diagonal which is the longest diagonal right okay which you can have then you have like you know m by 2 m by 2 and root whatever right and that will give you m by root 2 therefore this rho in general can go from m by root 2 to minus m by root 2 and then theta will go from 0 to pi okay. Now now if I take a point okay if I take a point see I think so the so the way to way to kind of write look at this is it is actually interesting it is like you know actually sort of you know dual a duality here the I mean right duality is in terms of what you what you fix see for example right I mean what to so for example right in this case suppose I suppose I say that okay let me just let me just go back to that equation rho is equal to X cos theta plus see Y sin theta right.

Now if I take a point right so let us kind of go back to that image that I had right I had so suppose I take some X of course you know one of the things is okay that is leave it so if I take a point X, Y okay I take a point X, Y because that is what I have on my image right I

have so many X, Y's there so I take one X, Y right from that so I have my image right wherein I have done you know a canny edge detector or something I have lots of these guys right these pixels some of them of course you know seem to lie along a line so I am taking one such pixel which has an which has some X, Y right which is this X, Y right which I am so let us think about what this line is this line right. Now when I kind of see take this X, Y and if I change my angle theta so that means my X and Y are frozen here right I have taken an X and Y if I change my angle theta what kind of a curve was what kind of a what kind of a thing will I see here what will I plot what will I end up plotting I have to plot rho know I am varying theta I have X and Y fixed that means what rho will change what law will it follow it will follow a sinusoidal law right because it is it is going to cos theta plus Y sin theta right so it is so what will happen is right so I mean I do not have the exact figure but something like this is what you will get for that one point and so the one point traces a sinusoidal curve of cosine whatever it I mean a curve on the so this is called the half space this is called the half space or the or the see theta rho theta space which is of course a discretized space by the way there is some see there are there is lot of studies on how you should how best to quantize and so on let us not the idea is not to enter into all that okay but the idea is that you have you know a half space right and in which in which this guy traces a curve now suppose I took I took one more point right so it is like another point on this edge I do not know whether it is there on that edge but I have just taken one more point which is which is another XY now I plot this what will happen that will also be a sinusoidal that will also be a sinusoidal curve right because I am going to fix another X, Y and I am going to vary my theta from 0 to pi I will plot my rho what will happen yeah exactly right so what will happen is so the other guy right might actually come like this and then go off like that but then they will actually intersect here because that because of the fact that they lie they both lie on lie on one line with a rho and actually a theta right so among all the rho's and theta's that they can span there will be one rho and theta which with they both will actually vote for that is why this voting comes so for example it is also so if I take another point right it might be it might be like that right it might go like that but that will also cross that cross that point and go because of the fact that you know they are all lying on the same line so if you now do this what will happen is what will happen is right I mean if you if you actually accumulate right so if you accumulate these votes that you are getting now if something is lying on a line then all of them will try to will try to vote for let's say one bin right if you think about this as a discretized space okay so actually ideally what you have is actually you know this one a discretized kind of space okay I mean you can actually do a discretization whatever I mean there are various strategies for that that's not the that's not the intent of this of this lecture okay so then then the idea is that okay this guy right will get flagged as a very high number and you can imagine no you know right in an image if I tell you tell you that I have a line there can imagine how many how many of these points lie on that right if there is a line actually so then kind of right what will happen is then that so when you when you get a

right eventually look at sort of a local maximum okay then right you may have you may have one here and then the right of another you might have so it's not true that you will have just one line right in an image no but these guys will actually stand out you may also have what else can happen are these only two things you will see I mean like let's say I have two lines okay in that line okay there I will say two lines huh now what will my half space look like will this will these so let's say for the other line there is some other rho theta right that gets flagged okay that's also very significantly high number this is also significantly high number and when I know that okay that there are actually two lines now okay this is it anything else I mean this will it be as clean as this this this space this half space what else can happen exactly right so so basically what will also happen is I mean you have got these two lines but then there are also points going I mean you can also think of lines going through let's say two points right two points I mean you can always draw a line no so you will also have some little little votes here and there but those have to be ignored right because those are not really when you say that you have a line it doesn't mean write two points I mean that's the way we might but that's not the way this interprets it it says there should be a dominant line in the image with lots of h pixels supporting it okay that's why that's why this is what you have this is voting this in mechanic otherwise you will end up you know drawing so many lines lines right through the image which won't even make sense so so when we say this we we actually you know we understand that that you know it's not like right these are the only two things you know that are going to get flagged but these are the if there are two lines that those two will get predominant predominantly high number of votes right and which will stand out and you will know that okay these are these are the lines okay so right let me just write this down just for this benefit of you know because these things sometimes right I mean you will have to search you know all over the place huh yeah only after we have actually yeah we have flagged the so that's like the output of a canny edge detector we have now and we want to we want to tell what are all the lines through this image now I think I think I started by saying that yeah so the so the huff transform right so the way to kind of think about the huff transform is like this so so first is create you know a 2D 2D quantized r theta space as a rho theta space then what then what you do is next is for each for each edge point for each edge point so it's already there okay it has to be an edge point let's say xi comma yj all the all the all the histogram bins okay I mean I mean the even histogram you know right what is histogram the histogram bins in rho theta in rho theta space that correspond to the that correspond to the ht of transform that is the sinusoidal curve for a point for each point all the wait a minute for each for each edge point oh increment okay I think you know that's what I was thinking for each edge point xi comma increment okay there should be a word increment here and that's missing increment all the all the histogram bins in rho theta space that actually correspond to the sinusoidal curve for a point at that at that location so what this means is and that is a location so it's like you know each of so when you say when you say you know histogram means what you're trying to say is how many votes right each

each let's say one of the one of those bins that you have in that half space right how much how much of votes right is it able to get so when you take an xi comma yj then you find that sinusoidal law then you go and you know increment each one of those bins wherever it falls. Then, repeat step was repeat the above step let's say call this step one repeat step one for all the for all these edge points for all the edge points and to find the find the line segments what you will do is search the rho versus theta histogram for a local maxima or for for a local maxima or peaks because it doesn't have to be just one right so maxima peaks then generate lines corresponding to to the to the rho theta coordinates coordinates of the of the local maxima that means for all those rho theta bins for which you have maximum value or they which are which turned out to be peaks generate lines corresponding to them so what does it mean to generate I mean simply join all those points to the rho theta coordinates of the of the even local local maxima okay yeah that's it and then I think there is some some pictures that I wanted to show you so where is that so you go and see this right anyway okay these are again from the past one. Threshold will yeah so what you normally do is you know so for example you look at I mean you do this non maxima kind of a suppression so on that grid right when you wherever you are sitting if you have local maxima then you will also look around okay it may happen that right depending upon how these points are right it may happen that you may also have something else which is also which is also not so not so low then you just because you don't expect two lines to be that close right what you will do is you will suppress one again that's like you could actually be wrong by doing that but then you just look at the numbers between those two because again that this is actually you know a different discretized space right so quantized space so some approximations have to happen it's not like it's an entirely clean thing but I mean see the idea is that it's a very neat way to actually do it right I mean everything will have nothing is like you know 100% you know this one but yeah more or less right it will actually work well I will show you some examples right so for example so right if you had an image like this right then and thresholding assume that we have performed some edge detection right and so for example you know so here is where you have those edge pixels you still don't know which of them are lines but of course to your visual eye right you can immediately see that it may know these are those predominant lines and so on but then after you pass it through a half transform right then you see that you know that that actually picks up the ones that are the most significant right and again when right depending upon your application you may want to go back and re-tweak and then allow a few more and so on but then the idea is that idea is that right in that half space there is still so much information out there for example it might look like something like this so it's not like entirely clean I mean it's not like this is a one peak or something so for example right here this spanner so if you see right so if you see after the after the after the after kind of see detecting the lines right now it's passing a line straight through right those two those two are your lines which are your output and even if you have noise and all it can still do it I mean that's what is actually nice about it I mean which

is why anyway I will tell you some other day so here it says these are top 20 most voted lines okay so you can say can I have these top 40 and then maybe there if you see that too many too many unnecessary lines are coming then maybe you will just stop it that's another that's a hyper parameter then you have to stop right you have to kind of you have to do a through a thresholding somewhere no say that only these are significant advantages conceptually simple easy to implement handing handles missing and occluded data what does that mean a discontinuity yeah so something is obstructing something for example right I have a line but then but then suppose I place an object you know on top of it now I can't see these at all on these pixels I don't see something else has come and sat on it but then but then a half transform allows me to draw that line through the occlusion because because it knows that these are all supposed to be on a line therefore something should have been underneath which has been occluded or we just get a partially visible whatever and so it can actually handle that gracefully and then can be adapted to many types of forms so now you can think right if I had circles and maybe right you can have a half space for circles you can have space for ellipse which is all there by the way then the disadvantage is that they are the more complex object becomes there is then then you know this kind of a parametric form right can can be more and more involved this was a line so maybe my my idea in you know talking about is line was to simply convey what is what is underlying what is you know a principle of the half transform but where all it can be applied you can always right escalate it and then go to other objects and looks already one single type of object in the sense that you know there is only lines which you are looking for at a time and then the length and the the the where the line exists I mean not not that I mean you know so I mean you cannot tell what is the what is the kind of say length of the line right because it will just it will draw a line all through that's it and then a collinear segments right it cannot actually separate right because it will think that the whole thing is a line see for example if you had if you had a line like this you had a line like this in between if you did not have anything it might still believe that right that there was a line through it now you will have to you will have to take a call on whether I should treat this in between as empty that there was nothing there or whether there is something whether I want to consider something was there just that my edge pixel when I did a canny right those things got lost and those are all weak pixels they probably went away but then they ideally should exist there and therefore right I will actually allow this line to run through but this cannot tell on its own it just draws a line through and through okay so we will stop here for today and then meet whatever next week.