Modern Computer Vision

Prof. A.N. Rajagopalan

Department of Electrical Engineering

IIT Madras

Lecture-45

So again the idea is exactly the same, right. This was whatever we did. So how do we do this? So let I will call X star, Y star be the actual locations, okay I think I have to write bigger, be the actual locations corresponding to the maximum value, to the maximum value, okay. And then let us do a Taylor series expansion. Probably this course uses Taylor series the most I think, I do not know how many times we use Taylor series in other courses to this extent. I mean I do not know every other topic has Taylor series kicking in, right, about DC detected but this is more like an application, right, both the detected corner point X c, Y c.

So we have a corner detected X c, Y c and the actual corner we believe is at X star, Y star, right. So the corner response, okay let us actually write that down response function. Suppose right we would express it as a kind of a continuous function and this is some f of X star, Y star. So suppose we say that X star let X star is equal to X c + let us say ∂ X and Y star is equal to Y c + ∂ Y, okay.

That means in the neighborhood of X c, Y c. So this we can say is the value at X c, Y c and then account for how far away you are from that point. So ∂ X let us call this f X and this f is the corner response function by the way, okay, f X of X c, I mean so this evaluated at f c, Y c, you know X c, Y c + ∂ Y f Y X c, Y c + some second order terms and so on which we can ignore. Now what should we do next if you want to refine our estimates, our estimator X c, Y c, you want to refine it. So what do you think you will do? f is that m.

But now I expressed in a sort of a continuous form. So what shall we do? If X star, Y star was actually the maximum then something should happen there, right. So what is that? If you compute a gradient there that should be 0, okay. That is what we should obviously use. So we should say f X, so if I take the gradient value at X star, well because that is the maximum, you know, so its gradient should be 0.

So then this would be like f X of X c, Y c + ∂ X f X X of X c, Y c. And then similarly I also know that f Y should be 0 at X star, Y star. So I can write this as f of f Y X c, Y c + ∂ Y f X Y X c, Y c. There is something that I think probably missed out one minute. Yeah, yeah, here right I missed out now f X, okay, this one that I missed out now f X + ∂ X f X and then here also I mean this term I missed out, right.

So this should be, okay, so I think this should strictly be f Y X, right. I mean if you want to exactly follow what we have there f Y X and this will be f X Y, right. This will be f X Y X c, Y c +, which one? You know one minute, see f X, right, so that is this derivative + ∂ X f X X + ∂ Y f X Y X c, Y c. I mean it is actually δ by δ X, no, it is δ by δ X of this guy. So it will be f X Y is like δ by δ X of actually δ by, so it is like δ square f by δX δ Y.

No, no, what is the problem? There is a problem here? Down. Down, no, yeah, yeah, no, no, yeah, down I am not, no, the first one okay. Is that okay? Okay, now next one is with respect to Y, right. So we have f Y + f Y X + ∂ Y f Y Y, no. Sir, ∂ Y f Y Y + ∂ X f Y.

Ah, okay, oh, this one you are saying, sorry, sorry, yeah, is that what you are referring to? Okay. I thought, I thought the problem was with respect to the derivative, okay. This is okay? This is fine, right? Or is there a problem? No sir. Okay, ∂ Y f Y Y X c, Y c, okay, but this we know to be 0, right. This we know is 0, this we know is 0.

Therefore what would you, what would you do? So ∂, so we just have to write ∂ X ∂ Y in terms of, in a way to get a matrix vector form and get ∂ X ∂ Y. That is what we want, no? We want to know where is that ∂ X ∂ Y. Okay, so I think that I am just going to write down this matrix, okay, which you can verify is okay. F X X of X c, Y c, F X Y of X c, Y c and F Y X of X c, Y c, but most of the cases, right, they will just assume F X Y to be equal to F Y X, F Y Y, X c, Y c and then into ∂ X ∂ Y that is the unknown to be equal to –, so this sign and all that we need not worry because ∂ X could itself have been a negative quantity about ∂ X and ∂ Y, but anyway if you write it in this form it comes out to be F X X c, Y c, F Y X c, Y c. So all these quantities are known.

Now but this, this how do I calculate? I need, I need a second derivative right along X. Yeah, so we can use a kernel right, so what kernel would you go for? What would be the entries of that kernel? Applause will give you sum, whereas I just need F X X right, so how will you do that? We did that no? 1 – 2, – 2 1. I mean that is how we got right? Then along, see for X it was like 1 – 2 1, for Y it was 1 – 2 1, we added the board, we added board to get 1 1 1 1 – 4, now we are just using them individually right? So what you will have is, it is actually a 1D kernel right, so 1 – 2 1, all 0s right elsewhere. So this will give me F X X and for F Y Y right, I should use 1 – 2 1 and then 0s elsewhere. And for this guy it will be 1 1 – 1 – 1 and then for this one right, which is, check why is this so? It should be easy for you to find out, this will be, this is what will give you F X F X Y, this kernel.

Okay and then once you have this, then you have, then you have this matrix now right and then you know this anyway, this is what you computed, I mean no, F X of course you need a derivative. So this is also then, for this you will need a, you will need a first order sort of a differencing along X and this will be a first order differencing of the corner function along Y. And then ∂ X ∂ Y is simply invert this, multiplied with this vector and that will give you the ∂ X ∂ Y that you need to, that you need in order to know exactly, exactly where the, where the corner is right, to what extent or how close you are to X C Y C, is that okay? At most be 1 in this case yes, yeah that is what I have, I mean my guess is that it will automatically give a ∂ X ∂ Y such that it cannot be greater than 1, that is what I am thinking, but that occurred to me also. I was just thinking we are not constraining ∂, that is what you are saying right, we are not constraining it, but my, but my feeling is in the immediate neighbor, I mean because of the fact that ∂ X ∂ Y are small, that is why we are ignoring all the higher order terms and all right, I mean in that expansion, so I am saying, I am thinking that probably this automatically comes that way. This invariance right, the covariance and the invariance that we talked about the other day, what were the 2 things we are worried about, one was photometry, another was geometry right.

Now let us first, let us first kind of see, kind of you know worry about, worry about geometry. Now is a corner, so let us first look at the simplest right, translation. So is a corner covariant with respect to a translational operation, it is right, so corner as a feature, yes, is actually covariant. What about, what about the, what about the strength of the corner, that is actually invariant, strength is invariant. What about rotation, again corner feature will, will is actually covariant.

What about the, what about the strength? Strength is not invariant. Strength is not invariant, how many, how many people agree with that? Let me, let me give you guys some, some time to think about that. So what will happen I mean, so, so you, so you, so you have a, so the corner response right, if you actually rotate, why would, why would your, why would your lambda 1 and lambda 2, why would, I mean just, just the Eigen vector orientation would change right, just the orientation would change, why would the strength change? The feature is covariant, but why is the strength, strength, why is the strength not invariant? 2D, 2D in plane. Lambda 1 and lambda 2 will change. No, Eigen vector directions will change.

Strength is invariant. Check this out know, if you are, if you are still δbtful, take a, take a corner rotate. What about scaling? Scaling the entire image. Scaling the entire image, like zoom in or zoom out right, something like zooming in or zooming out, what will happen? A corner is covariant.

Is not. Why? I am saying it is not even covariant, even the corner feature is not actually covariant with scaling. Why, why, why would you think that this, why, why is that with scaling we have a problem? Because a, because a corner right, if you zoom in, how will a corner look like? See. It may look like an edge something like that. Exactly, see a corner like this, when looked at a certain scale, it looks very sharp.

Suppose zoom in right, it will start to look like this okay, and, and then what will happen is certain some of these things would get flagged as edges right, and, and this, and this, and this, and this point itself right that you are trying to say declare as a corner as you keep zooming in right, this is the in fact that they will look more and more like edges.

Because see something like this right is going to, is going to look more like an edge and not like a corner right. So, so, so depending upon the scale of the zoom, I mean it is not like the, the, the strength will remain invariant and then the corner will remain, corner feature will remain covariant no. Okay, so, so far as the scaling and scaling which basically means that you know anything above that you know you can do a fine whatever you want to do right, which is like you know involving you know scaling along different directions whatever you want to do. Scaling and above, so the simplest things that you can hold against is a translation and simply you know a rotation. The moment you go beyond that right you do some scaling and all right then, then this, then of course the even strength is also not invariant, the corner is itself.

So I think let me just remove this and say is not covariant okay rather than writing like it is not covariant. Photometric what can we say? Photometric so photometric let us say simple things what can it hold against? Suppose I increase intensity globally can it sort of hold against it? Why? No, no that is right but then okay all of them but that is, that is like a verbal because, because you are taking a gradient and gradient will knock off that DC you know okay that has to be reason right. I mean of course intuitively it looks like I mean everything goes up then it should not matter but then that is true there is so, so if I have I xy + a right invariant. That is I x and I y will be the same. What happened? I x and I y will be the same.

I x and no, no, no, no, no I mean the, I mean the image intensity not the gradient. Yes sir but that I x and I y value should be same. I x and I y value. They will be same. Yeah, yeah I x and no what do you mean by I x should be equal to I y? No, no, no, no, no, I said that I x and I y values will remain same.

Yeah exactly no what is that? What about, what about scaling? Yeah it will be invariant but strength will be. And strength will all but then it will all go up by a because the gradient right multiply the gradient will also get scaled by a. So really right nothing much will happen. So in that sense right it can actually hold on against these two. Other things right we will handle it.

We will scale that also by a no? Done. Okay now what about the feature part I see right this is only about finding the corners right. So we are only telling where the corners are but then we have to be able to match them. So matching is not easy okay with respect to Harris. It is not there is no systematic way right by which you can do matching but the thing right typically that is done is you take a patch around the corner and you match.

So it is like saying that it is like saying that one image now you are now you are crossing across two images now right. Now assume that assume that you did a Harris corner kind of detector and you found a bunch of corner points. You did a Harris corner detector on the second image which is supposed to be of the same scene okay but then taken from a different view point let us say and then you get again a bunch of corners. Now you want to be able to so for each one of them you know how to do all of this. You have come you have arrived at all the corner point now you have to match right.

I mean that is how we said right if you want to do the stereo if you want to do object matching recognition panorama right everything requires a match. So now so the point is that we need to be able to able to tell right that this point is that this corner is that. So one way to do it is one way that is recommended is you take a patch here and then you take an image patch here and try to match them but then the but then there is something right that can go wrong because it is a rotation and all then there is no point matching just like that right. So what is typically done is because you know that the because you know that the eigenvectors are also rotated by a certain angle you know that angle right. So what you do is you kind of de-rotate this patch which is I mean we are I mean let us assume that it can be done right I have not talked about how you how you do geometry transformations and images but it is not such a big deal.

So you can actually de-rotate the de-rotate the one of the patches right with respect to that theta angle and then you do a correlation and again this correlation and all is done in a fairly sophisticated way what is called a normalized grayscale correlation and all I mean the idea was not to kind of go into all that but just to and then some people in fact reduce use the corner use the ratio of the lambdas also to sort of to be more sure I mean not just the cross correlation because you anyway anyway have the lambdas right and and as we said if it is if it is a pure rotation or something right then then you know even the ratio of the eigenvalues or the eigenvalues themselves that you could also use them to be to be a little more sure but the key thing is actually patch matching in fact it is stereo right this is what is done right. So this correlation is what is done right so you have a point here right one image and then you have a second view but then in stereo right it is much more much more what you call you know it is a it is a more it is a it is a I mean you know exactly where to look for and along that row you can actually search right you can do a cross correlation to find out but here right you are actually flagging automatically corners which are like which are like stanδt points right which are which are already screaming for attention right in a sense and therefore right and typically you know people reduce a search and all you know people do it in a smart way they do not go around looking at all the corners right just assume that probably this much is all that would have happened geometrically and therefore around that region right you look for something and around that region there can be only a two or three corners not like too many of them so you try to pick which one of them is most likely correlate intensities right and then get a see the site and this normalized grayscale correlation you should typically do because you know one could have a slightly higher intensity it could have been taken on a different day or a different time so there could be illumination issues therefore one does what is called a normalized grayscale correlation okay that is how you do and you know that is that is how you can do a coronary detection okay. Now corners are not the only points right which are actually interesting because anytime that you look at an image right I mean what we call what we sort of read in a loosely define as blobs okay so you know in an image that suppose I showed you a flower you know so suppose I showed you a flower field which is a standard example which I think I will also show you in the next class I showed you a field right containing flowers right and then you can see that you know as you go back right the flowers kind of shrink in size but then the ones in the front are kind of big in size and each one of them looks interesting right.

So you need so when you flag interest points so one is a corner is more like an interest point whereas blobs are more like more like see region points well yeah I mean instead of interest points they are called as region points I mean it may look a sound a little contradictory they are saying region on the one hand and you are saying point at the other hand but that is the way it is okay so they are called like region points what it means is within that region something interesting is going on okay. Now how to how to see capture that and how to get of know declare things like that is it because that is what in an image it is not like an image has corners all over the place right so you cannot just it is it depend on corners alone and also the fact that Harris corner and all has some inherent weaknesses so the idea was to actually idea is to come up with something so there was a lot of work during that time in fact between the 90s and the 90s and 2000 a lot of work was going on in terms of what could be you know what could be you know what you call strong features good features so right so people came up with so this one guy came up with what is called SIFT right scale invariant feature transform which is what I will talk about next class and that is actually robust to a lot of things I mean it is robust to translation, rotation, scaling, it is even illumination changes to a certain extent it is assigned the way the way they build it up right is also very systematic and that that I think that is what I think we will do in the next class.