Modern Computer Vision

Prof. A.N. Rajagopalan

Department of Electrical Engineering

IIT Madras

Lecture-48

Now, read what is actually actually actually done is now is when this notion of scale space comes in. I think you know somewhere in an earlier class I did just loosely mention it, but now comes the idea of what is called a scale space. So, the space has to do with actually you know the this one image grid and the scale has to actually do with σ. So, when you say scale space I mean you can relate space to your image grid and then scale to the actual σ rate which you are going to use. So, now, I hope now I am going to well let me just say that let me tell you what is that σ, σ is R by root 2, but this you have to show. Now, you can actually now you can actually imagine that if I had an image right where let us say you had those flowers right at various different scales and you are trying to examine it.

Now, clearly if I just use one *σ* and try to analyze it what will happen only those blobs that fit to that σ will fire right nothing else will fire. So, the way to do it is to actually to keep the image of course fixed right that is all I have and then and then I should examine it at various scales of my log right. So, that is what that is why you get what is called a scale space right. So, scale space is like is like having I mean it is like having right different scales.

So, which which basically means that means the rate when you have a σ 1 for this you have a σ 2 for this you have a σ 3 for this you keep on varying varying your say σ. So, you can go from a high to low I mean you can go from a low to high right. So, let us say right from a low σ to a high σ and the and the idea is that I mean actually actually right there is a there is a lot of you know what you call you know psychophysical theory and all that has gone into this where let us say right people have proved that a Gaussian is apparently is the best right among all functions which you can think of. So, if you are examining something something at let us say various different scales what are the things that you want is what? Ok yeah right I mean. So, I will I will actually I will you know because it is not very obvious at what it what it should be.

So, it is like this right I mean if you if you are actually applying a function at different different scales that are examining something you do not want new artifacts entering right you do not want a function that introduces artifacts right a Gaussian does not. So, so the idea behind using scale is that you know is that you know a broader. So, what it means is the finer structure should get suppressed right because you are going to a coarser scale right higher σ means that you are examining something at a sort of you know a coarser scale right because you are kind of looking at large objects. So, you have a big σ you are actually having a big probe you are trying to examine a blob right. So, when you go up it is like having a coarser and coarser scale.

So, coarser and coarser scale the finer guy should just drop off right because I mean those should those should not any more we could say relevant. And another thing is this operator that you are using whichever function you are using to scale up and down that should not introduce new things which were not there at an earlier scale right. I mean because if something suddenly surfaces ok right then there is a problem separately among the various things that you can think about but there is a lot of theory right I mean you know even I am not an expert in that. But a lot of work has been done and then a Gaussian is apparently the best. This is a very famous paper Lindbergh or somebody if you read I think that is a 60 page paper where he analyzes the whole thing you know and ok the I mean nice thing is that a Gaussian is the best and of course a log right when it comes straight from straight from that a Gaussian.

So, what is now done is you examine at various scales right. So, you go from a low σ then a higher σ and then you hope that you know at some place something when it actually matches the scale it will it will you know it will what you call no it will glow right in a sense. Now, now what what is typically done is in order to get the max and of course you are looking at the magnitude response ok because it could be a minimum it could be a maximum both are ok. See when you say a blob right it could also be something like you know 1 going to 0 or it could also be a 0 and then a 1 right both are blobs right except that in one case it will be just the opposite the extrema in one case it is a maximum in another case actually is actually you know a minimum both are relevant for us. So, it is not just this blob right it is also something like this is also important for us 0 to 1 that is also a blob for us.

So, both are equally relevant in one case you will have a maxima in another case you will have actually a minima therefore what is analyzed is a strength of the scale normalized normalized response ok to and all these operations are simply a convolution operation right I mean I am not going to be talking about how best you can implement this and so on right there are very good ways to implement this very quick ways to do this and so on. But the idea is that what you do is you know if you are at a point right if you are at a scale right suppose suppose I am here and I want to check right should I kind of consider this as a as an extremum or not then what you do is you know you kind of look around a 3 cross 3 neighborhood ok at the scale above one scale above one scale below and also I also read at your own scale. So, then how many neighbors have you got 20 what is that 26 right 26 neighbors you have got. So, this guy has got say 26 neighbor 8 at its own scale then 9 above it and 9 below ok 9 at a higher σ 9 at a at a at a at a higher and immediately lower σ and if this turns out to be the highest right among all of them then you say that right this is a local extremum. And what you saw in those circles right is actually that that σ see the important thing is I mean when I write I mean if I simply said that I mean here is an here is a point you would be able to know at what scale right it was found no you do not know at what scale it matched that is why you saw those circles there right those circles are not really I mean that is that is only a only indicative of the scale at which it was found to be interesting.

It was a region that was found to be interesting at a certain scale therefore, it if you kind of work through this scale space then you will see that something fires at a let us say very high σ then around that point right I mean see like I said right this is a space right and this is the scale. So, on the space you have a grid right where you know where you know the x y location where it happened, but it is not enough if you simply point the location you also have to tell at what scale was that examined right and that is why you have those circles showing up because they are telling at what σ right did that fire right and that is why that is how that is how you actually right I mean you are able to you are right you are able to kind of you know do the this one. So, if you go back to this figure right. So, that is what that is what has happened here. Here the equation itself is given I think.

So, that makes it simple for you. So, here is a butterfly example right. So, you can see that. So, you can see this blobs right. So, at some is it like high σ or low σ no see this picture right before picture then later.

So, you are actually picking up the big blobs right that means your σ should be high right these are the course guys that are so the brightest ones are the ones that have a high response here ok. Now, this is ok 26 neighbors right this is how you do and find maxima squared Laplacian response in the scale space. So, see here right. So, you see so many interesting points and again I mean you may not want to use all of them right you may want to finally, right choose only a certain number of them you can actually you can actually do a do a thresholding further thresholding, but yeah with the right this how it looks. So, it is not very easy to interpret right I mean it is like saying that something of interest right has flared up has fired up and what is of interest is whatever is this operator thinks is interesting.

See what you know the I mean right nice thing about let us say about you know a traditional way is there is still a lot of insight right which you can gather whereas in deep learning right I mean as far as the knowledge gained is concerned it is shallow right because the network is deep, but then what we understand is very shallow whereas traditional is the other way around right it is not at all deep, but then at least the kind of see thought process that in fact there was a guy called Marr David Marr.