Course on Image Signal Processing
Professor A.N. Rajagopalan
Department of Electrical Engineering
Indian Institute of Technology, Madras
Lecture 4
Basics of Images

(Refer Slide Time: 00:30)

So this is about the basics. So, when you see a picture, so she is actually Lena, her name is Lena and she is supposed to be a, you know she is the one whose images has been used, I do not know (())(0:40) times maybe by the image processing community right for free. I mean so in this world now probably it is unthinkable. So, her face apparently right has all the you know right features that you might need in order to evaluate algorithms.

No longer the case, but yes there was a time there was an extended period of time will people would simply say that if your algorithm works on Lena’s face then it is good, you might of course argue but then the fact is this image has been used over and over and over and in fact, like Lena was called I think in 2015, Vancouver I think, when ISIP 2015.

She was actually facilitated simply because I think, they just wanted to acknowledge the fact that her image is used so many times by the image processing community it was more like an acknowledgement.

So she came and then, so I believe. So now when you see an image like this, one of the things that you kind of notice is that when somebody talks about an image, then they say it is of certain size, they say its an m cross n image or it is a square image n cross n (())(1:47) 1024 by 1024 and so on. So, what those things really mean is what is called actually a pixel. A pixel is really you know a picture element when you say a picture you mean a really a picture element and when you see image is typically read these are the three fundamental kind of images that you will see.

One is the colour which is the first one this is a colour image, so this is typically made up of RGB components. The second is called really a greyscale image and then the third is called the binary image. Binary in the sense that it has only 1 and 0 levels this one. And this one is kind of grey again depending upon how many bits you use, we can have a greyscale image that has maybe 8 eight levels, 16 levels whatever 256 levels more number of levels and so on. And then you could also have really a colour image.

In our, in this course, the focus is going to be more on 2 and 3 the colour part we want to actually deal with except when we do some segmentation and so on. That maybe the only time when we will actually use the colour image, otherwise we would not actually, we would not really talk much about a colour image.

The greyscale and at the binary, are really the two images that will be greyscale mostly binary also, will be actually useful. In fact, a binary image is also very useful because sometimes you might just have some kind of cartoon and you are able to make out who it is. You do not even have to draw the entire tones and all of these, especially if somebody famous we know a cartoon diagram of that person you know that can be a so much information. So, in that sense binary is also a kind of say very-very relevant, so we will also see that.

(Refer Slide Time: 03:35)

So a pixel, in that sense, this will require a little plot so, so when we talk about a pixel, we really mean white and black, when we talk about a pixel, what we really mean is a picture element, a picture element and when you see a picture there are actually three things which you would immediately notice.

Ofcourse if it is a stat if it is a still frame, then there are two things that you notice, one is a brightness resolution one is really a brightness resolution, which actually means the number of grey levels which we are using which are there in that image.

So, the more the grey levels, the more smooths that particular, now smooth in the sense that the more natural it looks, number of grey levels in the image, this is what you mean by a brightness resolution. Normally it is about see 256, okay which means you use about 8 bits, but that is simply because of the fact that for a human being, beyond that you can ofcourse go higher.

You can go to 512 grey levels, you can even go higher and so on but then it does not seem to matter so much unless you have a specific application where you really this one quantize very fine, otherwise it is okay for a human being I think 256 grey levels you can appreciate what is there in a picture.

So, most of the images you will find are kind of you know, do I have a good 256 grey levels, but again if you try to see decrease the grey levels and so on you can start seeing some effects, like for example, in this case let me just show you an image with.

(Refer Slide Time: 05:23)

So, what would happen if you try to if you try to kind of play around with the grey levels, so here is a case where on the left, it is a same image, if you see the size and all it is exactly the same there is no change in the size but then on the left have we used about 256 grey levels.

Now, in the second one what you see is called the a false contouring, it looks like the face has these contours, especially if you see here and all there seems to be some kind of contouring that is happening which is actually not there, this is not at all supposedly there in the actual image.

So, this is called the false contouring so if we try to keep on decreasing the number of bits thinking that it is alright it may not be alright because somebody looks at the picture then he starts drawing wrong inferences. So, something like this is okay, 256 I mean you feel it is fine whatever is that to make of that face I can make out and similarly not just for faces this 256 is kind of a standard accepted now 256 grey levels.

Then the second one that you might want to look at when you see an image especially if it is a still image so one is this brightness resolution which ofcourse you know which is something that you immediately feel.

(Refer Slide Time: 06:30)

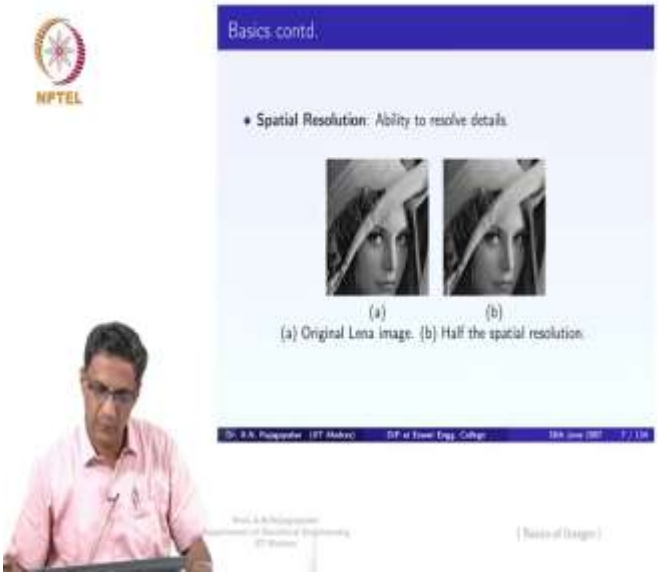
The second one is what is called a spatial resolution and these two are completely independent of each other. Spatial resolution has to kind of say do with how many sensor elements do you really pack in an image, so it is like saying how many, so for example if you have the number of sensor elements per unit area which is higher, then it means you are actually sampling a scene, you know in a sort of a higher with a higher spatial density.

Similar to the way you sample sequences in time, similarly if you want a good spatial resolution, we will have to sample a scene closely enough and if you have the sensors that are spaced too far away then it means you are not sampling enough.

So how many sensor elements are packed, how many sensor elements in a let us say in a unit area or in a sense you are talking about what is it density that you have. So the higher is your density then you would be able to see, you know so the idea is that finer details will start to appear, otherwise if you have a resolution which is spatial, so now it is independent of the other one.

So the brightness thing is independent of this, you could have the spatial resolution of whatever and then you could have a brightness resolution which can be 4 bits, it can be 8 bits whatever so these two aspects are totally independent of each other. And the spatial resolution simply gives you a sense for how much fine details are you able to see in that image.

(Refer Slide Time: 08:06)



So here, for example let me just go to that picture again and if you see the next one, so here is a case where you see a spatial resolution effect on the left what you have is the original Lena image and on the right what you have is actually a spatial resolution which is that half the size which actually means that what we have done is we have just expanded it to the same size as the original so that you can appreciate as to what is going on what you are losing.

When you see the two, what you see is, I mean they both have exactly the brightness resolution is identical in both, but the image on the left seems to be conveying more information in terms of fine details and so on. Whereas, the image on the right when you see it has this jaggedness you can see that the hat appears a little more jagged then what it is and so on.

So you can see that this is not probably something that you would like. Again, if certain problems if you feel that this resolution is enough you might actually go with it. Sometimes what also happens is that you might actually try to solve a problem where somebody gives you a low resolution image like this and asks you, can you kind of go to from here to there?

By low resolution what I mean is if this guy is let us say 512 cross 512 then this image is actually 256 by, well it is a 256 it is like a resolution down by 2, but it is been scaled up so that the two you can actually compare.

Now, you can actually ask that can I go from 256 cross 256 to 512 cross 512 that is also something that is of interest. So when you have a camera that is kind of cheap and you want to implement some signal processing algorithms that will work on the image and take you back to a high resolution grid.

Okay this is typically not just an interpolation problem there is more than that. So you can try interpolation and all but interpolation works basically with whatever it already has, you cannot bring anything new. Interpolation works with whatever is already around, whereas unless ofcourse you have already see oversampled or something which we are not assuming.

So what is normally done is, should I, can I do it with just one image, can I go like single image from a single image can I go to a high resolution which looks like very ill-pose problem or you might ask can I have a bunch of images taken all at low resolution but then they should be taken in a certain way so where they convey together, they actually convey information that can take me from that low resolution grid to high resolution grid, it can saves you money, it can saves you on lot of things.

Now, the third kid of and again the spatial resolution is independent of like I said the brightness.

(Refer Slide Time: 10:33)

There is one more thing which is actually what is called the temporal resolution, this is not a factor if you have a still image but it is a factor when you have a video. So, when you have a video somebody says that the video that is playing at 30 frames per second, somebody says I have a video that is playing say at a 1000 frames per second and so on.

Especially if you have the sports kinds of events, so you need a frame rate is very high because dynamics is very high and there is a ball that is flying around and you need to have a camera that is kind of capturing all that is going on.

So you need a very-very high frame rate, so in that sense that is called a temporal resolution. So it is like temporally how finely red eye is really sampling the scene, correct? So on one hand spatial means on that spatial grid how frequent how closely are you sampling, temporal is like you may have a spatial resolution that you have already fixed, temporal is like you know how finely.

So for example if you are doing a 30 frames per second that means every 1 by 30 second you have a frame which you are capturing. Now in between is something happens right then obviously you do not see that because you are kind of a grid such that you are sample at let us say t naught and then you will sample at say t naught plus will say 1 by 30 then you will sample at t naught plus 2 by 30 and so on that is what you are going to sample, which means that any activity that happens in between or you do not see it.

Now in some cases, in some cases it may not matter because when you, if it is a especially if it is not a very high action video it is okay you can play at that rate and typically at 30 frames per second is normal, it is what most of your camera is typically have 30 frames per second and also all this comes with kind of trade-off, you might ask what stops me by in terms of doing higher frame rate and so on which I will talk in a minute.

But the whole idea is that where specially you know when you want to go for a higher frame rate this means that what you wish to capture high dynamics in the scene. So for example, sometimes you might have somebody fires the bullet, so if you want to capture the real essence of how it is traveling and so on, you might need a very-very high frame rate camera.

There are also algorithms that talk about given a low frame rate video can actually be used at to build the high frame rate. Again, I means all these are you know on the one hand we could have a hardware which actually supports it, the sense that now you simply invest more money, you pay lets us say more money and then you say from 30 frames per second camera I do not want that I need something which can do at 1000 frames per second.

On the other hand, you might say that you know I just have this 30 frames per second kind of a video but then if I wanted to do a higher frame rate, can I do some kind of a signal processing which will kind of interpolate the frames ideally.

Ideally what you want something more than something new which happen in between but then it looks impossible to seen what might have happened weather now in many cases, it is like saying that if you saw a bullet go from here to there, then you would assume that over the bunch of frames you can get a see visualize that there must be a you know a trajectory you might interpolate.

And then the interpolation might also give you a feel as though you are able to see the actual bullet but in between if something had happened if that particular bullet has gone somewhere and come back, then ofcourse you would not be able to see but then physics is so good that you know these things do not happen that is why you can still survive, right you can do some interpolation and you know make somebody think that alright you have done actually a good job.

But whenever such things happen where you do not have you have not sensed anything right in between then you are only hoping that based upon whatever you have whatever things you are building up is able to fill in those gaps.

So again, so but all these factors one important thing is all these factors that you know all that three above are you know independent of each other like you could have spatial resolution, arbitrary whatever of high size, temporal could be something else all these are independent of each other, all the brightness, the spatial and temporal independent of each other.

Now, if you ask I mean if I had a spatial resolution, if I had to make it higher, let us say I have 512 cross 512 and then suppose you say I want to go up to say 5000 cross 5000, what stops me, is it the hardware that stops me really is it like I cannot make a sensor which is that small because what it will mean is we want to sense something finer and finer that means your sensor area should actually shrink, right because of the sensor area is itself so big that means whatever hits it, I mean you might only see the average of all those intensities.

I mean you know think of a spatial resolution like this, now if I, okay now let me draw here, if I had a box, if I had a sensor which was of this size, then whatever are these light rays that are going to fall on this I would just get an average of all of that ray, I will see some value.

Now instead if I had sensors that are let us say you know that are smaller in size in which case I can actually divide this into 4 let us say and I have a sensor which is only of this box or this box, now I can actually independently sense each intensity, there will be bunch of rays that will fall on this I will have a value for that, I will have a value for this, I will have a value for this, I will have a value for this.

That means I am able to sense the scene, in a more kind of finer way so instead of just saying what is the average of these 4 I might actually tell what each of these 4 is. Furthermore, I can go even finer I can say that if I had a sensor, okay if I had something that could do something like that, then maybe now if I could do that then maybe I can go even finer and so on. But then, in the process what happens is your sensor size is actually shrinking, your sensor size has to shrink.

Now the more you shrink it, even though you might have the hardware to shrink it, you might be able to fabricate something like that all that is okay, but there is one thing which kind of fundamentally limits how much lower you can go and that fundamental limit is actually post what is called short noise.