INDIAN INSTITUTE OF TECHNOLOGY ROORKEE

NPTEL
NPTEL ONLINE CERTIFICATION COURES

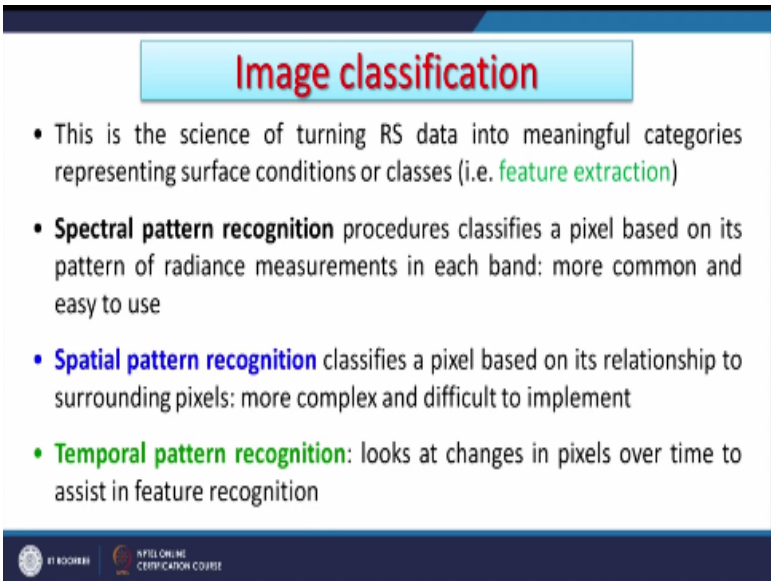
Digital Image Processing of Remote Sensing Data

Lecture – 15
Image Classification Techniques

Dr. Arun K. Saraf
Department of Earth Sciences
Indian Institute of Technology Roorkee

Hello everyone and welcome to this and new lecture topic which is image classification techniques of a digital image processing of remote sensing data course as you know that a image is satellite image are continuous data, and sometimes if we want to convert them into form some forms of maps so that the you know the reading quality of images becomes better and people those who do not do how to interpret images can also make sense out of that and reduce the number of classes in which are present and image to just few classes as per our requirements is called image classification technique so as we can see here that in this one image classification that this is the science of.

(Refer Slide Time: 01:15)



A turning remote sensing data into meaningful categories so we are from continuous we are going for discrete data sets, or meaningful categories which are representing surface conditions and classes and ultimate feature extraction, and in this also we say this spectral pattern recognition this procedures classifies a pixel base because it is a pixel base so this pattern of radiance measurements in each band more common and easy to use, so that if you are classifying this falls call composite you can do it.

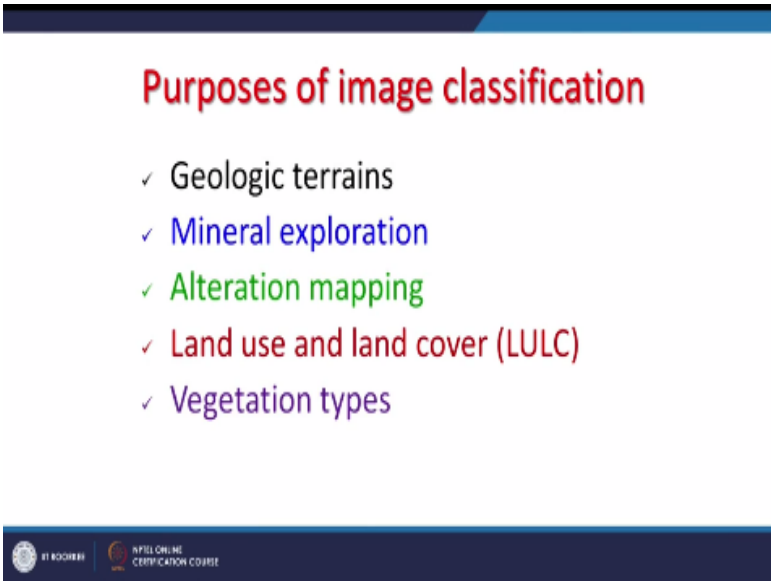
Three bands are there so three the classification will be performed on all three bands there are different classification techniques or divisions of classifications which we will discuss and like spectral pattern recognition there is also a special pattern recognition which classifies a pixel based on its relationship to surroundings pixels that neighborhood and the more complex and difficult to implement but anyway both are in a existing now days and third is that temporal pattern recognition as mentioned in my previous lecture.

That now- a - days availability of archival data is becoming the much more and like land said MSS data 1972 own word is available free of cost on that and therefore people are going for chain detection studies and in such studies lot of data times series data is analyzed and we want to detect the changes which has occurs in last 40 years, 45 years so if we are our study or application required that kind of thing then temporal pattern recognition that looks at the changes in pixel over time to assist in feature recognition so if you are having say and 10 image of after

years in each, we want to see how a particular feature area or ground has change a land from has change in all those years.

So that is temporal pattern recognitions many such powerful tools are becoming have a level some standard image processing installation digital image processing software nano tab but advance features are being added in these one, so purpose of as I have already mentioned the ultimate.

(Refer Slide Time: 03:46)



Aim is to use them the purpose of image classification in Geologic terrains so that we can identify a different wrong and different logs there ethnologies and also in mineral exploration alteration mapping very common one though in the this list I have kept at the last but this is the most common application of image classification is land use land cover classification also the changes in, land use in land cover because this is becoming a requirement in any large projects which in most large part of the land.

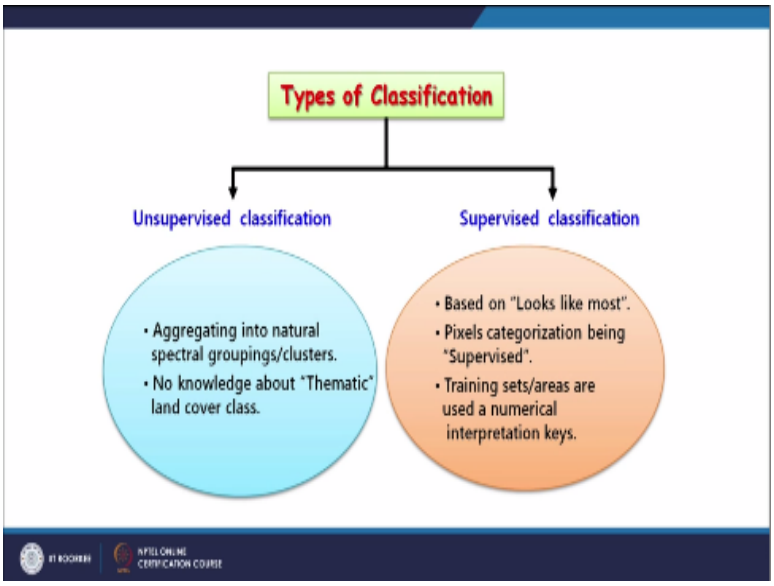
So that the requirements for environmental clearance is preparation of land use land cover maps so if we suppose in an areas a thermal power plant is proposed see in the year 2000 this is being considered, so now the in order to get environment clearance 10 years back data will be used a land use land map will be prepared because you are having through remote sensing data complete un base recordings of the land events in breathing.

So 10 years old data is classified the current data is classified and the after the coming of that thermal power plant again after 10 years again land use mapping is prepared and then we to see that how land has changed say in 10km radius or 50km radius around that power plant so because each project will bring changes on land, land cover and this is how it has to be access whether this change is really bad from environment point of view or not, so for that land use and land cover.

Is very, very common also for change detection studies otherwise to access certain things on the ground, the current things vegetation types it is extensively classification techniques are being used which are vegetation density based classification in forest cover mapping in others even in India for a survey of India is doing this, so the vegetation types and vegetation density base classifications are also being implied in case of vegetation cover in agriculture is where, this is none exist there are various applications.

Of each image classification as mentioned there are two basic types of classification techniques exists are known or have been implemented one is unsupervised classification.

(Refer Slide Time: 06:19)



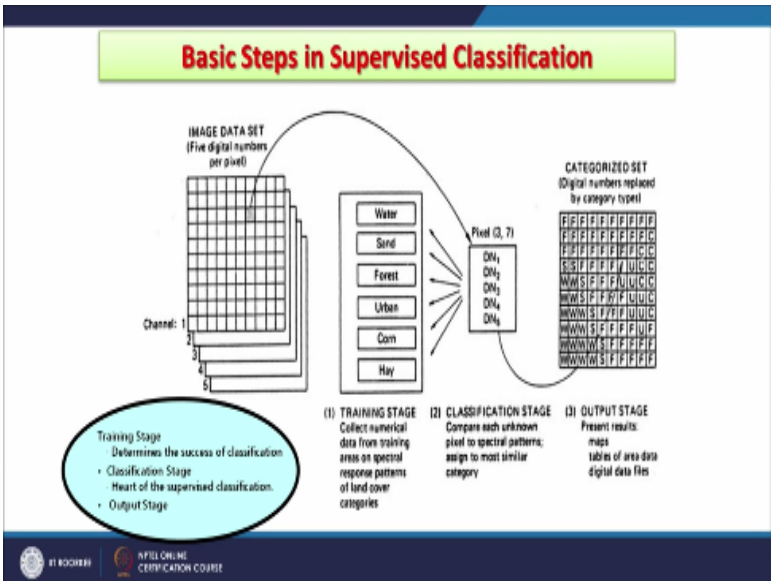
Where minimum human interventions are required by classification is being performed on a digital image processing software the other one is the supervise classification technique which where the training from a human interventions through human interventions are required be provide training and training set to the computer to the software to do the classification though it may be no recruit classification method, but it requires many inputs and to perform the classification whereas in case of unsupervised classification.

We have to only provide the number of classes of category we want to classify the particular image in two different classification categories, so this aggregation into natural spectral groupings or clusters so the clusters like if I am having the satellite image of a flatter in I want to classify in five categories but we are going to be five categories might be like forest area might be agricultural land might be waste land water body and built up land so likewise I am cluster the pixels those who are satisfying these in an three bands scenario.

Those who are satisfying this spectral characteristics they are assign one color or one class one category and likewise from continuous data which is originally as a continuous data you are making a discrete data that is classified image and no prior knowledge is required about the thematic classes, so if you are go down image interpretations just study that image as design decide how many classes can be there assign the classes and with the classification will be done, so it is quicker mean would be as accrued as supervise classification but large human interventions.

Are required many human interventions at different stages are required let see, so where then that looks like most and kind of concept pixel categorization being supervised the training sets I have to be provided and then you say that these number of classes for which I have fed the training set accordingly and based classified we will see all these details later so the basic in supervise classification is that this is the digital image.

(Refer Slide Time: 08:57)

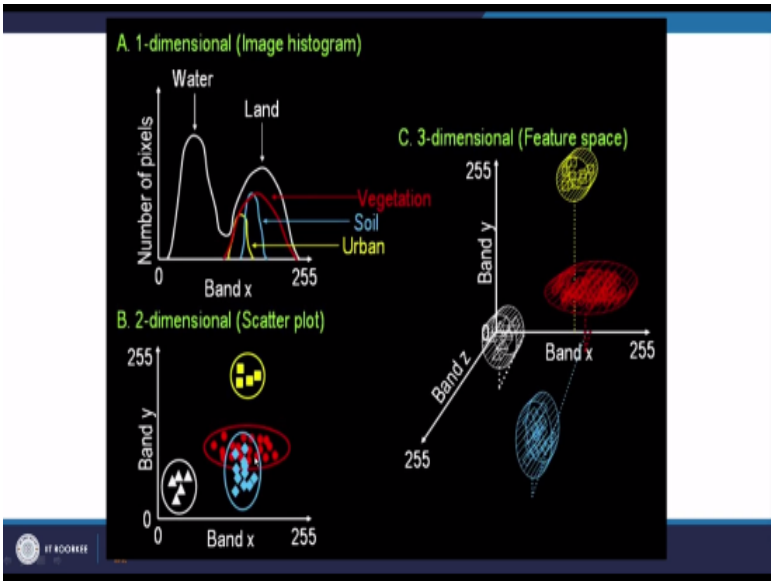


These are the training sets I have decided based on the my interpretation my understanding about the area having prior knowledge of the area that water bodies send for as urban cone fields agricultural fields and high fields and these then I there because we are having some digital numbers and group of numbers a range so that goes classification stage compare each unknown pixel and two spectral pattern assign to the most similar category and then finally the classification.

So here the legend is like f is stands for forest and what w stand for water body and so on so for, so that this continuous image as we now this trice through and supervise classification technique with human interventions so the training stage that is the training while providing the training set data and determines the success of classification how accurately you can said that training set in from the image there is so this classification maximum depends on viewed import, and that is the training sites and the classification stage, the hard of supervise classification which is very much

required, this is most essential thing and as accurate you can collect the training sites, as accurate your classifications results are going to be. Now here what we see that,

(Refer Slide Time: 10:31)

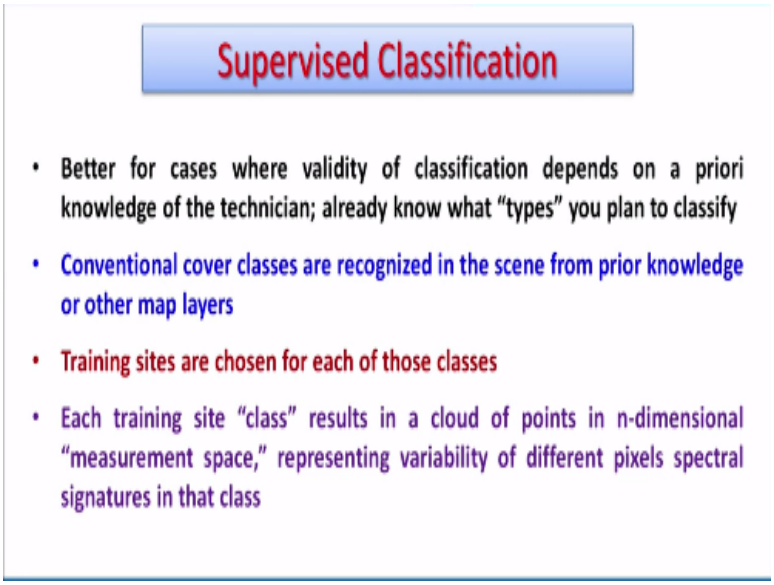


This is the spectrum curve of different natural features which are present on the surface on the natural objects, like water will appear and this is you know the frequency chart is there, so here you are having number of pixels and here you are having 0 to 255. So water is having generally less values, pure water will have less values and if we say infrared scenario vegetation is having, this situation.

So all features which are present will behave differently, in different parts of EM spectrum that we know, now when we plot a 2D scatter plot, what do you find? That all these land features or land objects are getting scattered at different location in a 2D plot, so this is systematic and if we go foe because, generally we go for classification having say a false color composite or a colored image, then you have to think dimensional scattered plot. There this can also be done very easily done but here because the schematic, colors coats or clustering or coating is very obvious.

But in natural, when you go for real operations, you may not find and so the distinct clustering, then the user has to decide it, that which pixels will go in which class. There is always a boundary cases and there the judgment an expert who knows something about the image interpretation will be required. So in supervise classification better for cases.

(Refer Slide Time: 12:23)



Where the validity of classification depends on a priori knowledge, so prior knowledge is there who ever doing classification, you can achieve a very good classification. Conventional classification classes are recognized in the scene from the prior knowledge or other map layers and if you are having, some other maps and looking those maps for corresponding areas, you can decide your training sheds as well

And training sites are chosen for each those free classes, where you find the desert vegetation area, this is water body, this is where soil and so on so far, and each training site class results in a cloud of point in say, take an example of three dimensional scattered plot and n dimensional measurement is spaced and representing variability of different pixels spectral signature in that class and like here this is the real example.

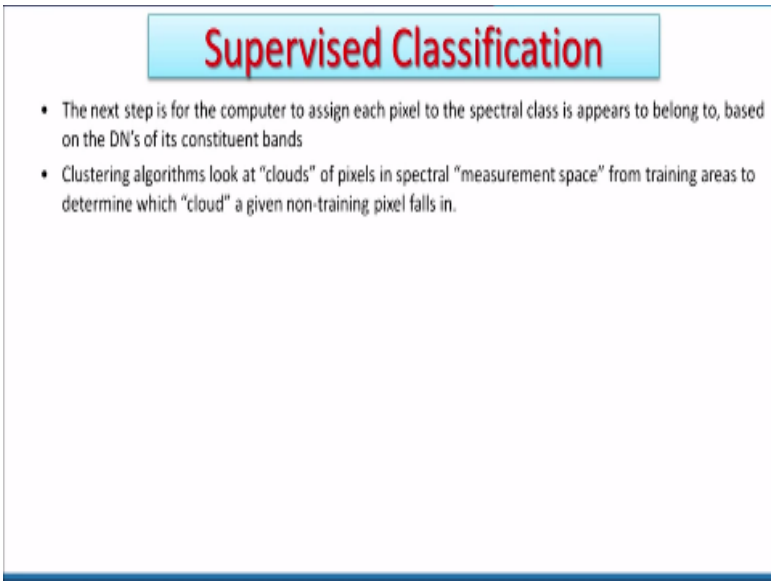
(Refer Slide Time: 13:26)

That these are the based on the prior knowledge and the spectral cores statics and knowledge of interpretation techniques, one has classified these are these are the training sites, the final classification images are not here but these are the training sites I have given and for different clustering or say if I take this one then I am having, this is some materials, so I am having this green color, this is not necessary but only one training site for each classes to selected.

You can have your 10 training sites for one class, so declaring that this belongs, this green area belongs to the same class and likewise, I have given so many classes over here, layer on you will realize that there is not much difference, so you can group easy classes that means number of classes you can further reduce few classes that is also possible. Here in the same way the water body, two types of classes has been given one mainly for having sediment or interrogatory which is having very high close to the coast.

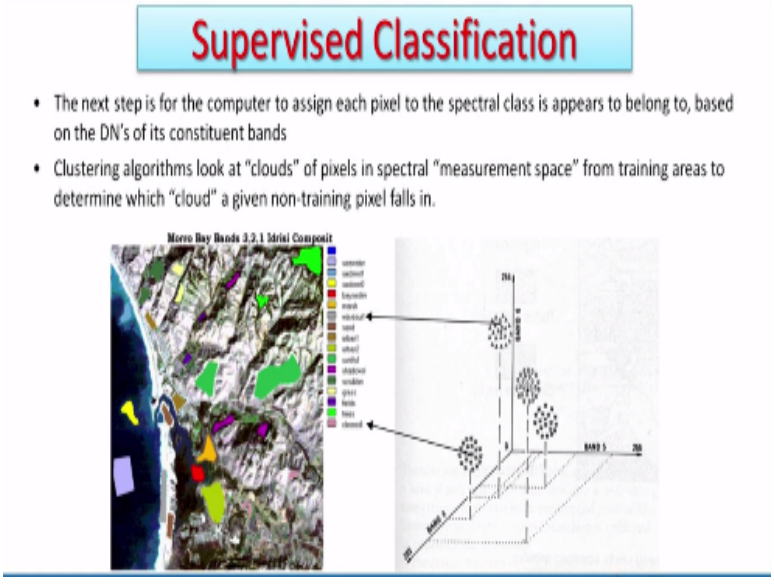
And in deep water you are having more clear water and accordingly different class has been chosen. Later on if you realize, that I just want to classify it has a water body, I do not want to discrete water body, between the water body. Whether it is having termite or not, then you can rebook and reduce number of classes, so that is always there in the time of post processing. So the next system here once you have collected the training sites, next system is to computer.

(Refer Slide Time: 15:12)



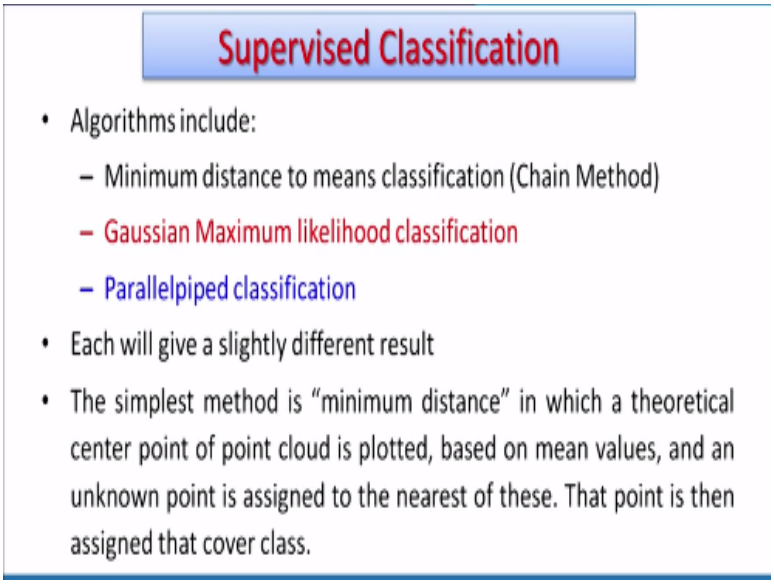
To assign each pixel to the spectral class is appears to belonging to the area and then the clustering algorithm look at clouds of pixels in spectral measurement space that is n dimensional scattered plot, from training areas to determine from which clouds is giving training pixel falls and like here this through schematic relation, as we have shown here.

(Refer Slide Time: 15:37)



Cluster of pixels of three bands are having training site of this area, like wise you know these different clusters are shown here and so once theses clusters are been identified then this is done for the entire image and then you see in the final results.

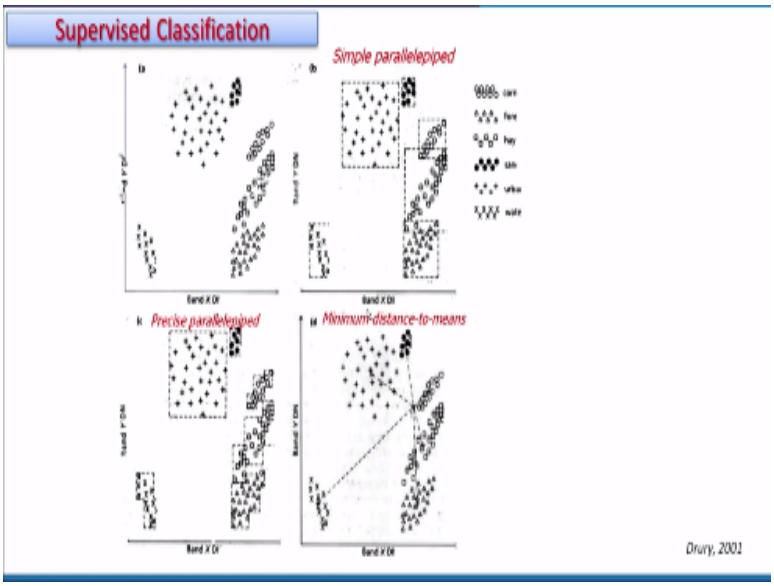
(Refer Slide Time: 16:01)



Then the supervised classification algorithm includes and there are various options are available during the supervise classification, one is called minimum distance to means classification or also called chain method, another one is the Gaussian maximum like hood classification, parallelepiped classification and each these options under this supervise classification will give a slightly different result. Which we will see why these results are different because the grouping of the clusters.

In which class you will put the clusters that decision will be made differently, if we apply these different techniques, that is the minimum distance to me the question class plan and parallelepiped. The simpler method is considered, is the minimum distance in which the theoretical centre point of point cloud is plotted, basically it is the pixels of three bands based on mean value and the unknown point assigned to the nearest of these and that point is assigned the cover class.

(Refer Slide Time: 17:11)



So here few examples are given for these scatter plots, two dimensional scatter plots examples here, this is the original scatter plot having 2 bands in board, now the clustering here is done like this. This is how these are getting plotted based on their respective characteristics in two different bands. If I go for minimum distance to mean then a point is plotted here, and the such is made that which cluster is having minimum distance and within in that all these pixels will be classified as one category, this one is schematic it is the waste land.

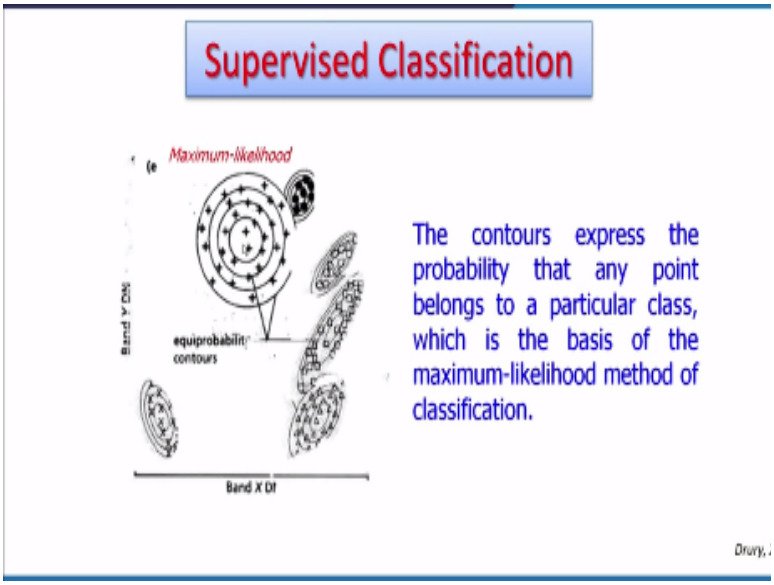
For this cluster this is the sand body and this may be the water body and likewise in the minimum distance classification is made. When we go for parallelepiped classification here a box, it is also sometimes called the box classified. A box that is square plotted along the cluster

in the long poring, so all those pixels are falling they will fall on one category. So rather than having a minimum distance to mean here we are just blocking or creating boxes around these.

But what will happen like here there are scenarios, where the some pixels are falling at two boxes, maybe in three boxes together and therefore we have a drawn calcification because of overlapping characteristics. So you do not have, there might be an option which is the précised parallelepiped, instead of just creating always box, you can create and surround the cluster but they not only the box and can have separate classes for each cluster then will you have.

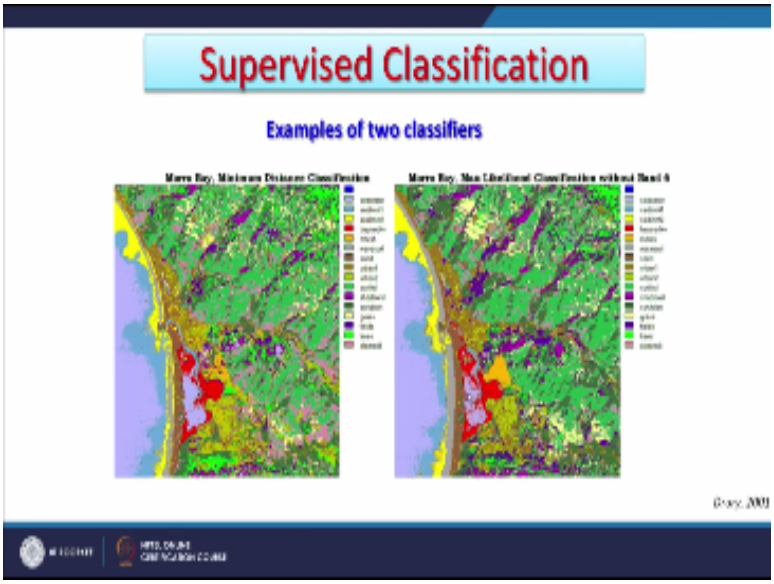
So if you go for this, kind of classification on this image and go for this then definitely you will have different results, so different methods of spectral classification can be represented diagrammatically here by reference to various plots.

(Refer Slide Time: 19:29)



Maximum like hood is the other scenario over here, which is based on the probability that any point belongs to a particular class which is the basis of the maximum like hood method of classification. So this is another very popular supervise classification technique, as you go up away from the center of the cluster, the probability or the equation will reduce, the probability of falling in that cluster reduces but at one stage you have to say, now this is these beyond these pixels beyond in giving this cluster plot they do not belong to this class. So the limit has to be achieved, so this is the relied is called maximum likelihood method. Beyond that you do not have, so two classifiers examples are given the real one is based on the minimum distance.

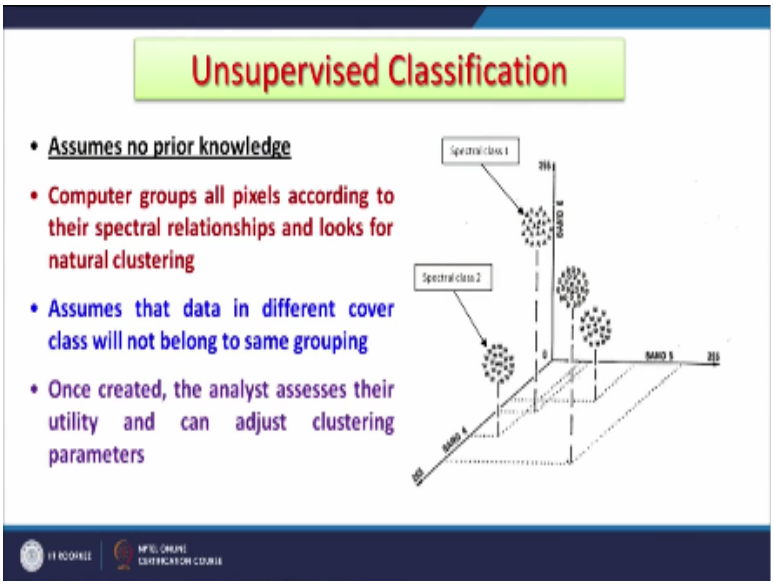
(Refer Slide Time: 20:23)



To mean classification technique and again maximum likelihood classifier. When you look very carefully you would find that both techniques are classifying the same image with different results. Now it is up to you to except and how you would except you can assess the errors if you are having the ground information you can verify that and then come back again, improve the classification and clustering and choosing the area and then you can reclassify the thing.

So there is another advantage but each classification options within the supervised classification technique will produce slightly different results. Now in case of one supervised classification the input is only in number of classes in user is required a human depth to only their extend human intervenes are required. So the same cluster plot is shown here we say this belongs two spectral class one is spectral class two three four likewise.

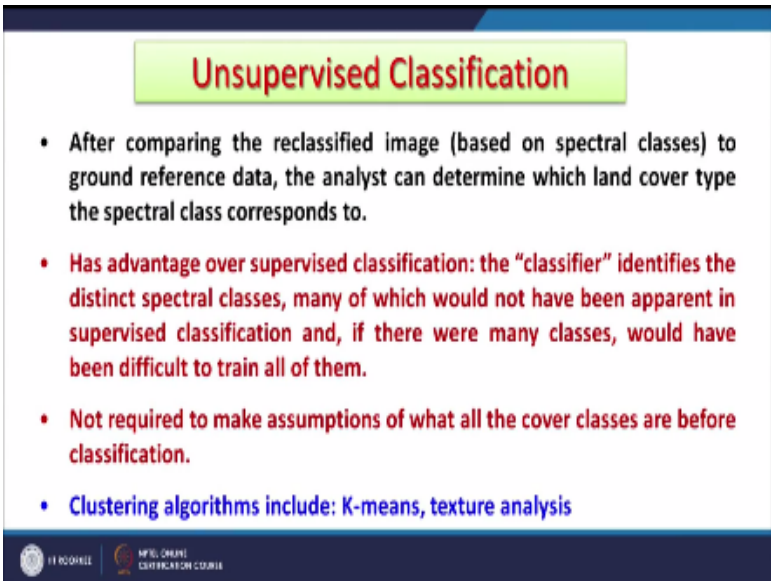
(Refer Slide Time: 21:30)



So assume it is no prier information knowledge is required computer groups all pixels according to their spectral relationships and looks for natural clustering. Assumes the data in different cover will not belong to same group that means the two ground different features will not have same type of spectral characteristic, this is the assumption in un supervised classification. Once cleared it the analyst assesses their utility and adjust clustering parameters that means number of classes you can change.

Suppose I have initially I have gone for ten classes I find that the output is unrealistic it is not really showing the very accurate classification, so I may group few classes and then go for say six I make at a very realistic classification results. So after comparing the reclassified image based on spectral classes to the ground reference data.

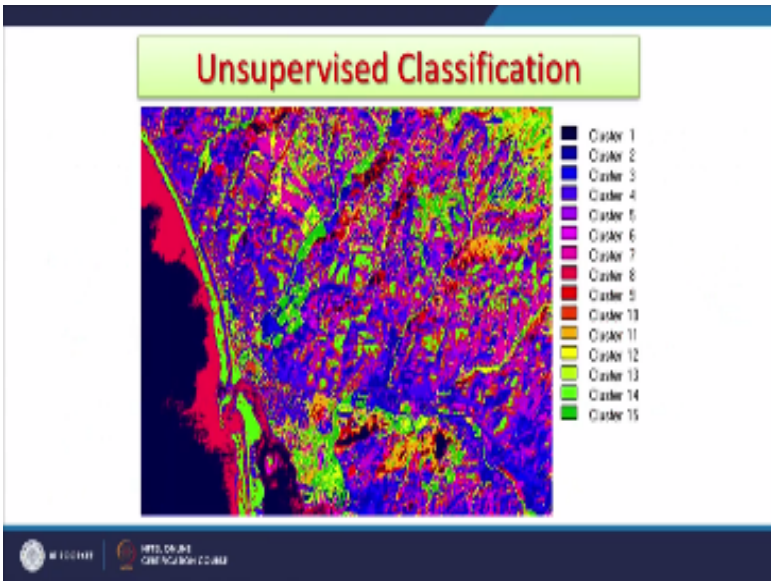
(Refer Slide Time: 22:33)



The analyst can determine which land cover type they spectral classes corresponds to, and then you can and the reason as you can assign whether it is forest agriculture land water body based land and so on. So had at one stage over supervised classification because the classifier identify he distinct spectral classes automatically many which would not have been apparent in supervised classification and if they were many classes would have been difficult to train all of them.

So because here everything is being done by computer for this clustering and a grouping, so you it may sometimes it may providing good results may not but in supervised classification you are having better controls. And it is not required to make assumptions of what all cover classes are before classification and clustering algorithm include, K means structure analysis and so on.

(Refer Slide Time: 23:31)



So there are also different options are available with unsupervised classification, so one example the same image that the number of classes initially are 15 clusters and they are assign different colors. Now based on the ground knowledge like for example I know this is water body but having turbidity, so if in say this is the color 7 then I assign a water turbid water, and water which is dark blue like this one then I may assign this one or this may be the cluster two I assign as a clear water.

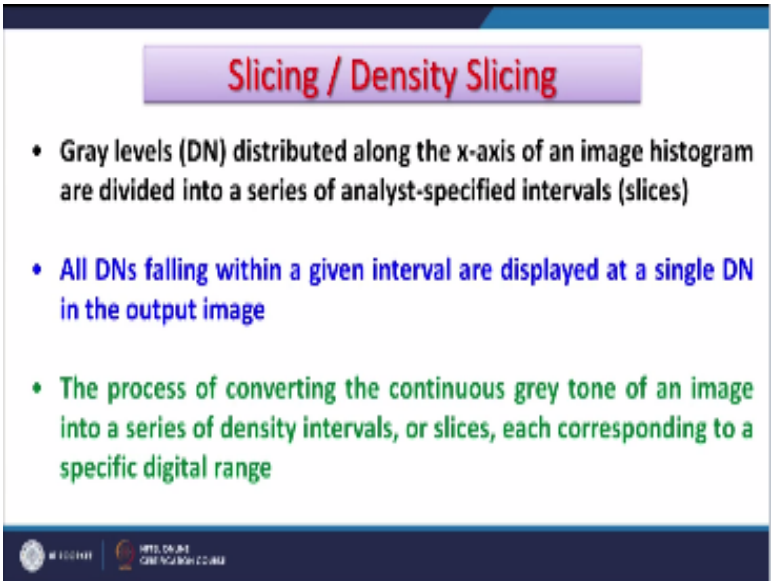
And likewise I can may be the send body is here so this instead of writing cluster 14 I will write a send body or beach. So this type of classification is much quicker of doing classification both options supervise and unsupervised are available and with both options different further options are available for classification but every classification will have some errors because after all from a continuous data nature is also continuous in that sense.

From continuous data you are discrediting you are reducing number of space, suppose if I am having 8 bit image that means I may have 256 classes in an image. Now you are reducing those classes to 8 and that means there while going do this kind of greater reduction I may introduce some errors, and those errors if they check with the ground data can be minimized and we can have a very realistic classification maps based on the satellite image.

But as I mention this is extensively use specially for land use land cover classification forest in agriculture in geology and so on. Now in this first part I have discuss mostly in the multi spectral domain that when you are have multiple channels you do the classification is specially

on falls cluster composite and so on. But if you are having the single band then these options are may use to you then the best thing is the simpler one is called the density slicing the density slicing is suited for single band classification.

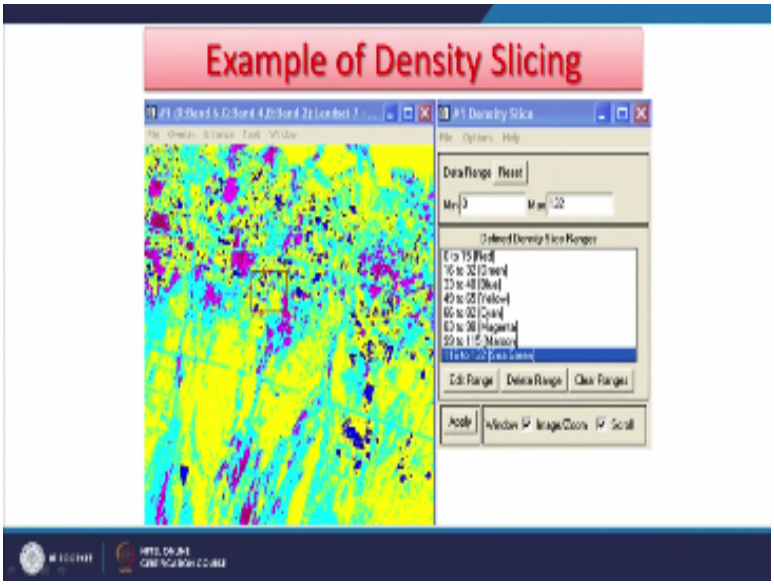
(Refer Slide Time: 26:03)



Here the pixel values distributed along axes of an image histogram are divided into series of analyst as per user define intervals or slices like you are having a low for bread you can slice so the thickness of each slice can be decided by the user itself so number of slices you want to create out of a beard low how thick it is going to be or whether the thickness has to be uniform or it can you can have very thickness that all can be chosen by the user itself.

So all DN values falling within a given interval within one slice or displayed at a single DN in the output image so from a continuous image like a loaf of bread it is continuous you are slicing it. Maybe at equal interval, or may be at varying intervals and then the processing of converting the continuous gray level tone to an image is series of density interval or discrete classes slices each corresponding to a specific digital image or area.

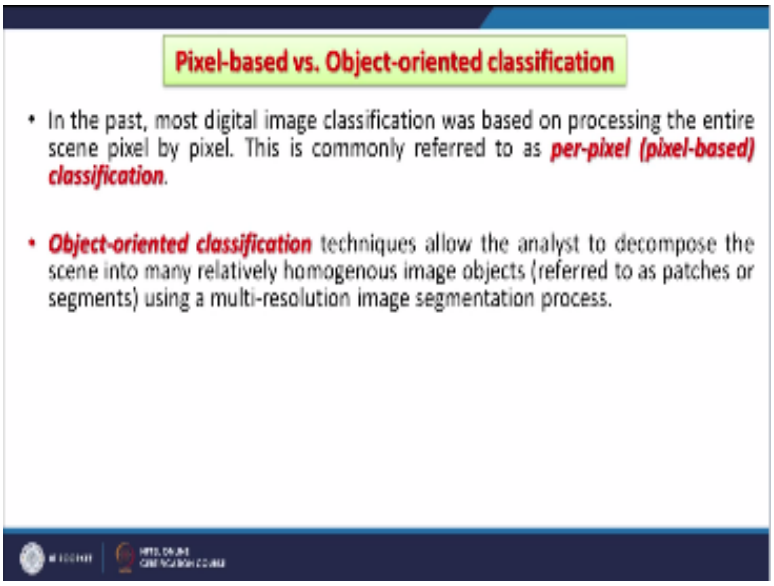
(Refer Slide Time: 27:10)



Here is the example that you are having an image I have classified in too few classes to pixel values say falling between 0 15 are assign red color, red color might be a type of land class or feature which are present on the ground pixel 16 to 32 are green and so on so far. And one say you have done the classification like this really equal interval classification or you can imply a certain classification techniques instead of having it equal interval you can have a equal area classification you can a quintal classification you can have natural break classification on single band image and likewise you can create classes here.

Mainly it is suited for single band for multispectral then you choose supervise or unsupervised classification.

(Refer Slide Time: 28:00)



There are because if you look the publication which are coming in digital account digital image processing in remote censuring generals and other generals. You would find that lot of development is taking place the classification techniques so apart from basic supervise and unsupervised and then different options are associated with that people are developing new and newer technique based on new mathematical approaches.

How these clusters will be decided that they belong to this particular class and not so there are various techniques are coming one is another one is based object oriented classification so in pass most digital classification was based on the processing of entire scene pixel by pixel and this is commonly referred as the per pixel, pixel based classification, so all these so far whatever the classification techniques which we have discussed based on pixel, pixel by pixel or spectral classes.

Whereas object oriented classification techniques allow the analyst to de compose the scene into many relatively homogeneous image objects, homogeneous areas and that will have one class definitely. As referred as a patches or segments as be went when we are discussing the data remote censoring or digital data compression techniques we will also looking the homogeneous area.

So her also the homogenous areas are searched, and then they are assigned a particular class using a multi resolution image segmentation process. So this one so there are various as mention that various options are available with the object oriented classifications and a different images,

because the problem here is that we are going from a initially we had a relatively as per today reference we relate we have relatively closer resolution images like land set MSS or even TM. But now we are having very high resolution images so the conventional classification techniques may not suit now high resolution images.

So we have to involve in new classification techniques we have discussed one more object oriented classification techniques are there classification based on neuron network or fuzzy logic all are be in either being tested or implemented in to various image processing software. But as mention earlier classification no matter which one you apply well produce a different output and more complication becomes when for the same area if you apply say classification which is based on neuron network.

So if I am having an image pre man shown image and post mason image but my classification technique remain same still I will have to different results, not only the because the classes the land features have changes the registration cove has change water what a size is change not only that but the classification results the error part will also be different. So it becomes virtually a same based classification technique that the, you know the classification which you perform on one scene when it is changed may be of different season your classification will change.

So the but all kinds of options are now becoming available for different types of classification so this brings to the end of the discussion thank you very much.

For Further Details Contact
Coordinator, Educational Technology Cell
Indian Institute of Technology Roorkee
Roorkee – 247667
E Mail: etcell.iitrke@gmail.com. etcell@iitr.ernet.in
Website: www.iitr.ac.in/centers/ETC, www.nptel.ac.in

Web Operators
Dr. Nibedita Bisoyi
Neetesh Kumar
Jitender Kumar
Vivek Kumar

Production Team
Sarath. K
Pankaj Saini
Arun. S

Camera
Mohan Raj

Online Editing
Jithin. K

Video Editing
Jithin. K

Graphics
Binoy. V. P

NPTEL Coordinator
Prof. B. K. Gandhi

An Educational Technology cell
IIT Roorkee Production