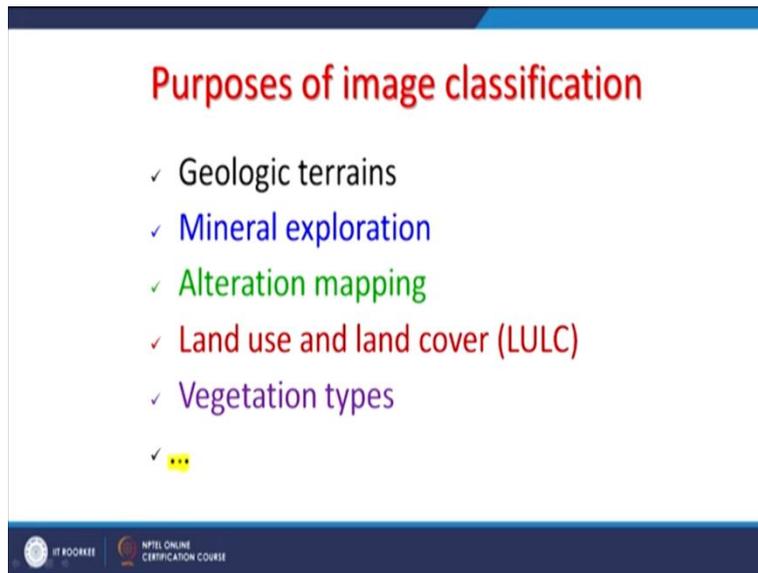


Remote Sensing Essentials
Prof. Arun. K. Saraf
Department of Earth Science
Indian Institute of Technology Roorkee

Module No # 06
Lecture No # 30
Unsupervised Image Classification and Density Slicing Techniques

Hello everyone and welcome to a new discussion and today we will be discussing unsupervised image classification and density slicing techniques. You know that classification image classification and its main purpose of that is to you know to reduce the number of variation in among pixel values to just few.

(Refer Slide Time 01:10)



So that we can use these images once they are classified directly to various purposes like geologic terrain, mineral exploration, alteration mapping, land use and land cover which is or vegetation mapping or many other things and we can use this so this is non exhaustive list.

(Refer Slide Time 01:21)

Image classification

- This is the science of turning RS data into **meaningful categories representing surface conditions or classes (feature extraction)**
- **Spectral pattern recognition** procedures classifies a pixel based on its pattern of radiance measurements in each band: more common and easy to use

Classification image classification is generally 2 types one is a supervise classification and another one is the unsupervised classification. In unsupervised classification what we do we submit the image to the basically to the software which is having some inbuilt algorithms. And based on that it classify images and there are different techniques for unsupervised classification techniques. So one can choose that I want to classify my image using this algorithm and once I choose then image is classified.

So human interventions or human intelligence is used in a very limited way. However in case of supervise classification through out the classification process the human intervention plays very important role. And some times you get a better output using going for supervise classification in terms of accuracy then unsupervised classification. So unsupervised classification is basically exploiting the spectral variability which is in the pixels in a color composite and using a particular algorithm you classify.

So image classification is a basically science of you know converting your remote sensing data or images into a meaningful categories representing surface condition or classes. An image is basically a continuous data and if I talk about a single band scenario and 8 bit scenario then the values within an image can vary between 0 to 255. This is the best possible scenario that values. However as you have seen through various example that generally histograms of raw image our input image does not occupy the full dynamic range.

So in this image classification what we do we categorize instead of having a full dynamical range of 2 roughly you can say in a single band image you are having you may have 255 classes. So you are reducing those number to just maybe 5, 6, 7 categories of maybe land uses maybe land cover maybe lithology or many other things which we use in different applications. So that is why we call as a meaningful categories representing surface conditions or in sort we also call classes.

And I some way we can also call as a feature extraction because after all these are features maybe a forest feature, maybe a built up line feature, maybe a waterboarding feature, maybe a river, an mountain and so on. So ultimate aim is of image classification either supervised or unsupervised is to extract features and or sometimes we say object extractions also. So this is a basically spectral pattern recognition which produces classify a pixel of an image based on its pattern of radiance measurements after all an image is having this radiant measurement maybe in terms of reflection or emission.

So generally we perform image classification on the images which are representing reflected values. And that is the more common things. However we can also have a spatial pattern recognition and which classifies a pixel based on its relationship to the surrounding pixel. So in this spatial pattern recognition and the neighborhood pixels which are in neighborhood are also considered however this kind of pattern recognition because image classification is also a kind of pattern recognition.

So this kind of spatial pattern recognition rather than spectral patter recognition is relatively very complex and difficult to implement and therefore we and do not see this part of understanding has been implemented in normal software's. And if somebody would like to do it then one has to really search for the best possible options or otherwise by himself he has to create a program. So spectral pattern recognition or you know which is based on a pixel on his pattern or radiant measurements is easy and has been implemented by a many software's.

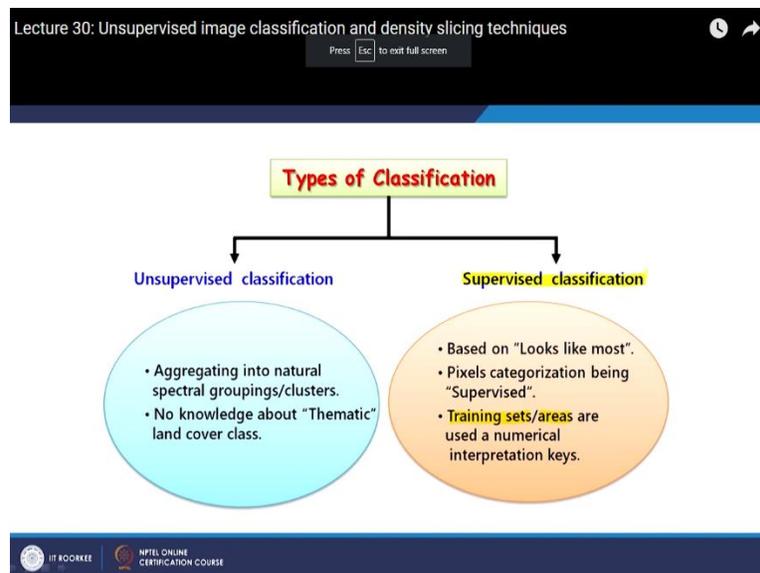
And since as you know we have been discussing that then after the inception of a Landsat 1 in 1972 lot of a data or archives are available of all those about 47, 48 years of data is there and the data is all available free of cost. So people also would like have started using this archive or

these archives of different sensors or this remote sensing data in temporal pattern recognition how things are changing?

Lot of now currently the major use of remote sensing data and especially this archive data is in trend analysis. The change detection studies that means temporal pattern recognition. So that kind of recognition that kind of work is going on especially because lot of changes are taking place have already taken place due to climate change or global warming. And therefore this old data 1972 onwards of remote sensing data has become a really very good assets and that too as you know that remote sensing images records things unbiasedly because after all there is an instrumental or a sensor which scans and record the things.

There is no human interventions or influences in the images when these are being recording. So these are the true recordings unbiased recording of the events, unbiased recording of the features or you know classes or surface condition of that time. If these have been recorded in 1972 this was the situation in 1970. So if I want to compare between 1972 to 2019 I may use these 2 images and can see the temporal pattern changes or change detection I can do. So that is another big application of archived remote sensing data is in temporal change pattern. This part I have just already mentioned that is unsupervised classification and supervised classification two techniques are there.

(Refer Slide Time 08:55)



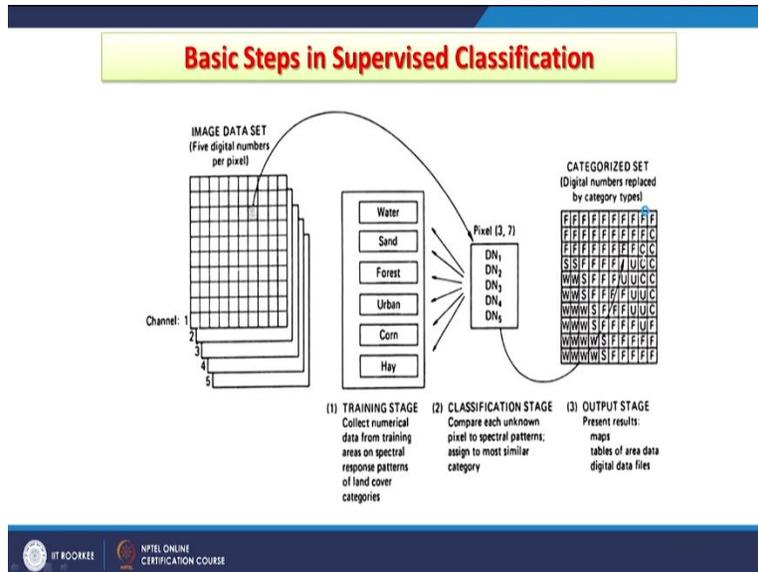
Here what we do in unsupervised classification which we will be discussing further here in in this particular discussion that aggregating into natural spectral grouping or cluster or categories because after all image is as I mentioned is a continuous. Continuous bearing pixel values are there in a 2 dimensional matrix and what in classification we are trying to do? We are putting these pixels into different categories or clusters. And in this one because this is unsupervised completely being done based on algorithm by a software on computer.

So there is a no knowledge of about thematic land cover class at this stage only when we go and do unsupervised classification over and say false color composite. What we say that I want to classify in 7 classes and I want to use this algorithm which we will be discussing. So once you have chosen these 2 things the computer or a software will classify. But there without having any knowledge or intelligence human interventions are only up to these 2 stages. The number of classes you declare and which algorithm which method you want to use that is.

Where as in case of supervised classification which is based on looks like most that probably this these group of pixels are falling in this category. And therefore through human interventions through selecting the training sets and training areas either you and you put certain inputs from your own based on your interpretation and intelligence and then rest is done by the computer. And when we go for supervised classification though it is time consuming and it requires a prior knowledge of that area of which the image is being classified.

Once that information is there prior knowledge is there then we can create a very good output through supervised classification rather than on supervise classification. So that what we say accuracy part of course would be much better with supervised classification because of human intelligence will be used. So here in supervised classification pixels categorization beams supervised that is why it is called supervise classification through human interventions. Whereas in case of unsupervised classification everything is being done through computers.

(Refer Slide Time 11:40)

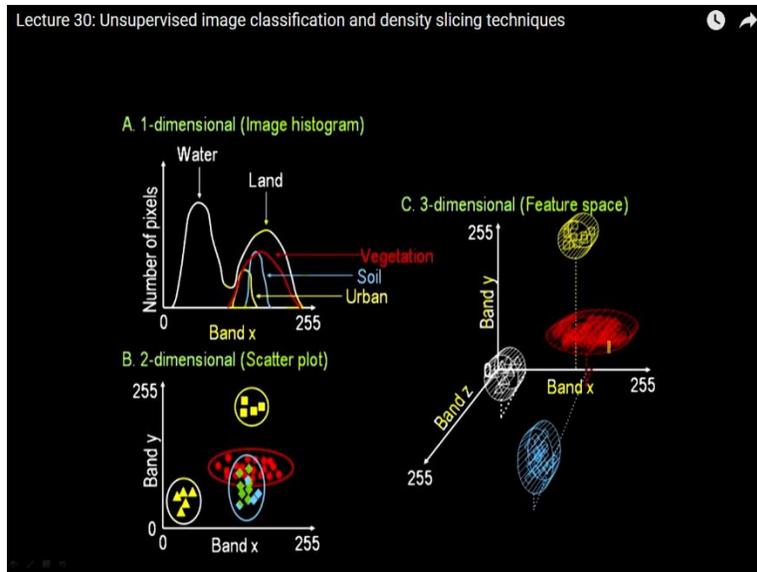


Now when we see the basic steps in supervised classification which I have just mentioned the selecting training sets. So here one example is given that this is my input image and I have selected a training set and saying that this number of this pixel belongs to the water body this is of course the schematic this belong to the sand, this forest, this is urban, this is corn, this is hay. So once the this is the training I have given now training to the computer to recognize the different features in the image and once that is done then compare each unknown pixel to spatial pattern which I have provided the input.

And then find out the similar characteristics of pixels in the entire image and assign different categories like here it has been done. So F is of course forest here and then maybe C for corn cover or corn field and W for water and so on so forth. So output stage that is present here which is showing number of classes which are here number of classes we are chosen what total 6 and your output will be also 6 but human interventions are there.

So at training stage we determines the basically the be success of classification better the trained training we provide to the computer to recognize a similar pixels based on the their respective characteristics higher the accuracy we will achieve. And that is greater training is basically hard of supervise classification.

(Refer Slide Time 13:28)



Now in a when I have been discussing about histogram at that time I are also the spectral curves I said that these are the fundamental of remote sensing digital image processing. And those things are revisiting again for different discussion. Now here this is image histogram and while looking of course this is schematic but by looking these we can identify different you know different features which are present within an image in a particular band.

Generally water will have a less reflection so as we are seeing and land open land without means you know just a land part we may have a higher pixel values and different vegetation may have different soil may have other things. But the same histogram we can also plot in a 2 dimensional or that in that way we call as a scatter plot. And when we do this thing because in single image histogram wherever the overlap is there like in this part then we it is difficult to disconnect or pick the front of features which are present in image.

But when we use the 2 bands so now from single band scenario to multispectral when we you go for multi spectral our discrimination of different objects becomes much easier. And so when these features are plotted are one are in yellow color, and blue are in soil, red is the vegetation and these strangles which you are seeing at the land part. Then now it is easier except for overlapping between vegetation and soil. Except that part these it is easier to discriminate between this and the urban through a 2 dimensional histogram.

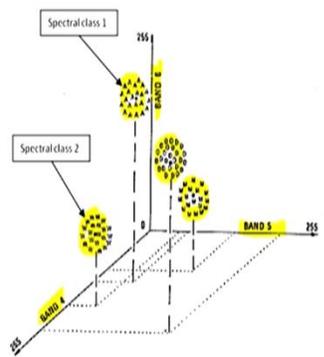
But if we go for a three dimensional histogram that is called the feature space. What we see here that the discrimination now 3 bands are in a band x band y and band z. So when these three are used then the discrimination between different objects becomes much more easier and the overlap which we have been seeing in the vegetation and soil in a two dimensional histogram or a scatter plot is no more there.

And we can discriminate different objects variously. So single band scenario very difficult to identify different object which are present in the image different features. But when we go for 2 dimensional that is scatter plot some objects can be discriminated can be isolated very easily. Whereas still some overlap might be there. But when we go for three dimensional then things become because three bands instead of 1 band, 2 band. Now 3 bands are being used and then discrimination and feature space becomes much easier.

(Refer Slide Time 16:49)

Unsupervised Classification

- Assumes **no prior knowledge**
- Computer groups all pixels according to their spectral relationships and looks for natural clusterings
- Assumes that data in different cover class will not belong to same grouping
- Once created, the analyst assesses their utility and can adjust clustering parameters



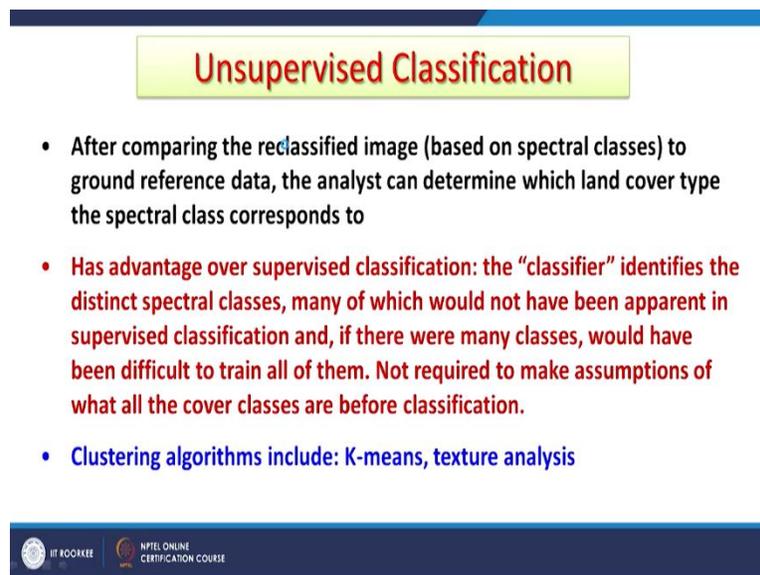


Now in unsupervised classification basically this is what is done as you can see the 3 dimensional histogram is shown here or you know feature space is shown band 4, band 5, and band 6 of Landsat TM has been plotted different features that is a class 1, class 2, class 3, class 4 all are there. And easy to discriminate and once we declared to the system that I want to classify this image a colored image with 3 band image into 7 classes and I want to use this method of classification then classification becomes much easier.

Though it is unsupervised may not be as accurate as supervised but for supervised classification the prior knowledge is very much required of that area of that in which the image belongs to so this one has to remember. So here no in case of unsupervised classification no prior knowledge is required and computer basically categorizes or groups all pixels according to their spectral relationships and looks for natural clustering as you have are seeing in this feature space or 3 dimensional histogram.

And assumes through this clustering that different land cover classes will not be belong to the same grouping. And once create the analyst for the user assesses their utility and can adjust clustering parameters.

(Refer Slide Time 18:31)



The slide features a title 'Unsupervised Classification' in red text on a yellow background. Below the title are three bullet points: the first is black text describing the process of comparing reclassified images to ground reference data; the second is red text highlighting the advantage of unsupervised classification over supervised methods; the third is blue text listing clustering algorithms like K-means and texture analysis. The slide footer includes logos for IIT Roorkee and NPTEL Online Certification Course.

Unsupervised Classification

- After comparing the reclassified image (based on spectral classes) to ground reference data, the analyst can determine which land cover type the spectral class corresponds to
- **Has advantage over supervised classification: the “classifier” identifies the distinct spectral classes, many of which would not have been apparent in supervised classification and, if there were many classes, would have been difficult to train all of them. Not required to make assumptions of what all the cover classes are before classification.**
- **Clustering algorithms include: K-means, texture analysis**

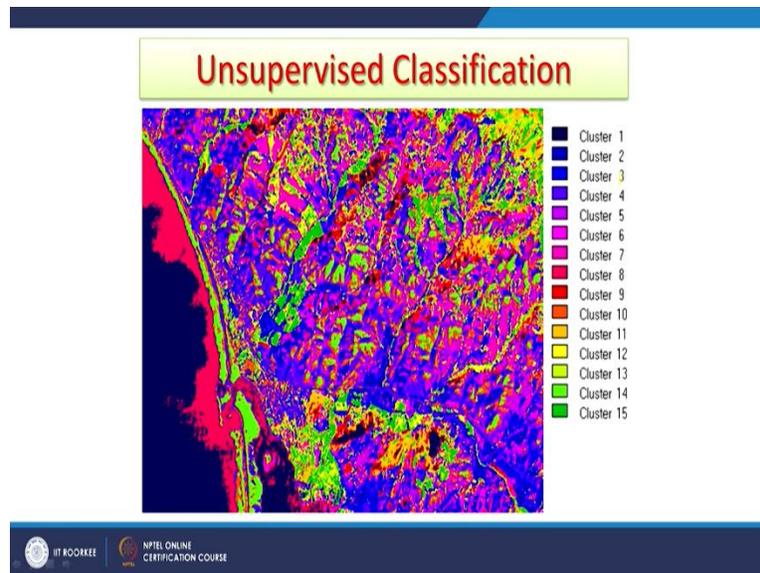
And we also showing an example also through a image classified subjected to unsupervised classification. So after comparing the reclassified image based on the spectral classes because this is what it is being done here in unsupervised to ground reference data if you are having the user can determine which land cover type the spectral classes correspond or through our own experience of image interpretation if I have classified an image in 4, 5 classes and by seeing a false color composite I can identity that there is a water body, there is a forest, there is a built up land, there is a agriculture land and there is a bare ground.

And like wise I will identify and group and put their names also. So that is the advantage this has over the supervise classification that because the classifier identifies the distinct spectral classes.

So the bias which human can have is not will not prevail in case of unsupervised classification. So many of which would not have apparent in supervised classification and if they were many classes would have been difficult to train all of them.

Of course better results when we get when we are having lot of groupings and not many classes present in an image. But if there is a lot of a heterogeneity present in an image then whether you go for supervised or unsupervised classification the accuracy will reduce significantly. So but if it has to be performed then it has to be performed. And here also clustering algorithm available maybe the K means and texture analysis in case of unsupervised classification.

(Refer Slide Time 20:33)



So an image has been subjected to unsupervised classification before classification it was declared that cluster you know create 15 classes. Now I can regroup these classes and can create a better image rather than having 15 classes. So that can be that way I can create a much better map. Now I should not call as a image now a map which may be a land use map using the satellite image.

So these number of clusters or classes can be reduced instead of 15 can be reduced to 7, 8 by regrouping implying or you know applying some human interventions like the class which is showing the water body an adjacent class which is showing the shallow water. So if I do not want in my map shallow water and deep water together separately then what I will do I will merge these 2 classes into one as a water body. And why which I will reduce number of classes

and my map which is an output through unsupervised classification becomes much more useable to many applications.

(Refer Slide Time 20:55)

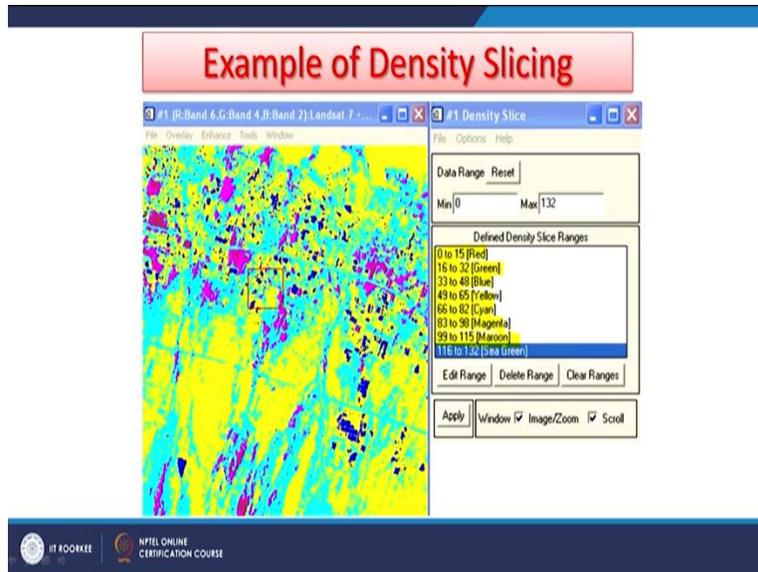
Slicing / Density Slicing

- Gray levels (DN) distributed along the x-axis of an image histogram are divided into a series of analyst-specified intervals (slices)
- All DNs falling within a given interval are displayed at a single DN in the output image
- The process of converting the continuous grey tone of an image into a series of density intervals, or slices, each corresponding to a specific digital range

Now we come to another discussion which is also a unsupervised way of classifying image through density slicing or in some literature you may find they say image slicing. So density slicing image slicing and though the here is it is generally done on a single band. So when you are having pixels which are distributed along x axis on a main histogram. They are divided into series of user specified intervals or slices. So you are having a histogram you can slice down histogram into different slices like in a loaf of bread you can create the slices.

The slices may be of equal size or slices can be of different sizes depending on your requirements. So all these pixels values falling within a given interval are displayed in a single color or single value into output image and why which what you are doing through density slicing also that the variability which was present in image is reduced to only few slices or few classes by which you can convert an image easily in 2 different slices and different classes. So this process converting of continuous and grey tone. When grey tone when mentioned it means the single band scenario and image into a series of density intervals or slices and each this slice will represent a specific digital range.

(Refer Slide Time 23:40)



For example here we are seeing an input image is here and it has been classified that pixels which are having value between 0 to 15 are assigned red color. A category name will also be assigned in a your region that values which are 0 to 15 might be water bodies. Now values pixel values which 116 to 132 are green maybe your vegetation and so on so forth. So here the slicing is done like this. Now here what is basically is done in this example is sort of slicing of a you know almost of same thickness.

Sometimes it is not necessary to slice down an image into same thickness like a loaf of bread as I said. All slices need not to be of same thickness you can as per your requirements you can change and can create a direct output from a grey level map.

(Refer Slide Time 24:49)

Pixel-based vs. Object-oriented classification

- In the past, most digital image classification was based on processing the entire scene pixel by pixel. This is commonly referred to as **per-pixel (pixel-based) classification**.
- **Object-oriented classification** techniques allow the analyst to decompose the scene into many relatively homogenous image objects (referred to as patches or segments) using a multi-resolution image segmentation process.
 - The various statistical characteristics of these homogeneous image objects in the scene are then subjected to traditional statistical or fuzzy logic classification.
 - Object-oriented classification based on image segmentation is often used for the analysis of high-spatial-resolution imagery (e.g., 1×1 m Space Imaging IKONOS and 0.61×0.61 m Digital Globe QuickBird).



Now there are some other classification techniques which I will discuss very briefly and that is a pixel based which we have been we have been discussing so far versus object oriented classification because these classifiers are also become popular and these have been object oriented classifiers are becoming popular and these have also been implemented in various software's. So in most of these image classifiers a classification which were based on processing the entire scene pixel by pixel. In a pixel base and this is common we say per pixel or pixel based classification.

However it is now possible to do it with object oriented classification which allows user to basically decompose segment the image in seen into many relatively homogeneous image objects. The pixels which are having almost same values are considered as one object referred as it patches or segments using a mutli resolution image segmentation process. So here instead of multispectral you can also have a multi resolution image segmentation.

So various statistical characteristics of these homogeneous image objects in the scene are then subjected to traditional statistic statistical or fuzzy logic classification. The advantage because why we are moving from pixel based classification to object oriented classification because for the same area you may have may be having images of different resolution. And If I want to you know use these multi resolution images of the same area for a better classification then this is the approach that go for object-oriented classifier.

Though relatively understanding wise in software point of the coding point of view it is little difficult. However this objective oriented classification which is based on image segmentation or decomposition is often used for the analysis of high spectral resolution imagery. So unsupervised classification and generally is good for moderate or you know coarse resolution images. But since a spatial resolution is improving day by day of all sensors on different satellites. So new classification techniques have to be evolved and this object oriented classification is one of them.

For example here a 1 by 1 meter space imaging that is a IKONOS here is the name of the satellite on which panchromatic data of one meter resolution became possible or was available. Then we had a quick word satellite by the digital globe and this is 61 centimeter 0.61 meter resolution. So as we move towards higher and higher resolutions a spatial resolutions the conventional classification techniques cannot be applied.

And therefore one has to move towards object oriented classification. So this brings to end of this discussion especially about three things which we have discussed here. And one is about your unsupervised classification and what are the limitation also and what are the advantages. And the second is the density slicing a very common technique to reduce the number of ranges in the pixels to into four groups of categories.

And third one for high spatial resolution satellite images object oriented classification. This part we have done in a very brief manner. But this brings just for comparison that how things are moving and what are the new developments are taking place. So this brings to the end of this discussion thank you very much.