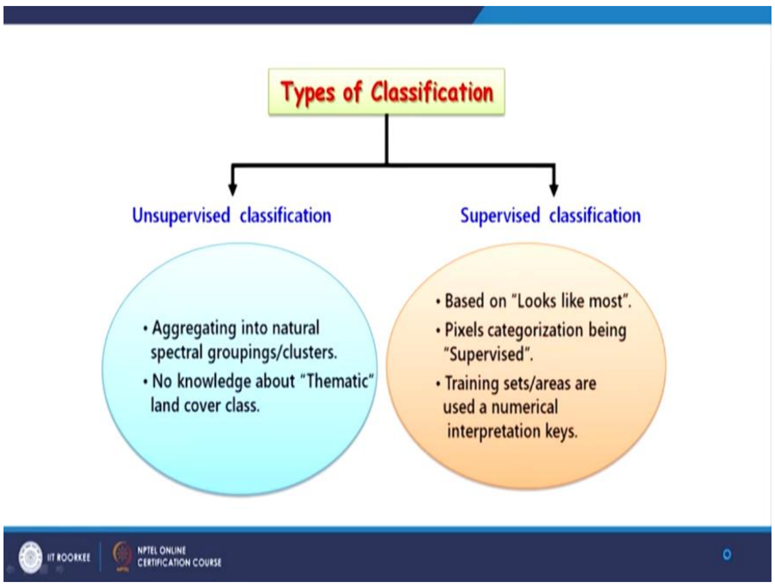
Remote Sensing Essentials
Prof. Arun. K. Saraf
Department of Earth Science
Indian Institute of Technology Roorkee

Module No # 07
Lecture No # 31
Supervised Image Classification Techniques and Limitations

Hello everyone and welcome to a new discussion and which we are going to have on supervised image classification techniques. And also we will be discussing about limitations of supervised classification or overall classification techniques as well. As we know that we have already discussed this part that there are 2 types of classification which we do in remote sensing or over the remote sensing images. The first one is unsupervised classification which we have already discussed and another one is the supervised classification.

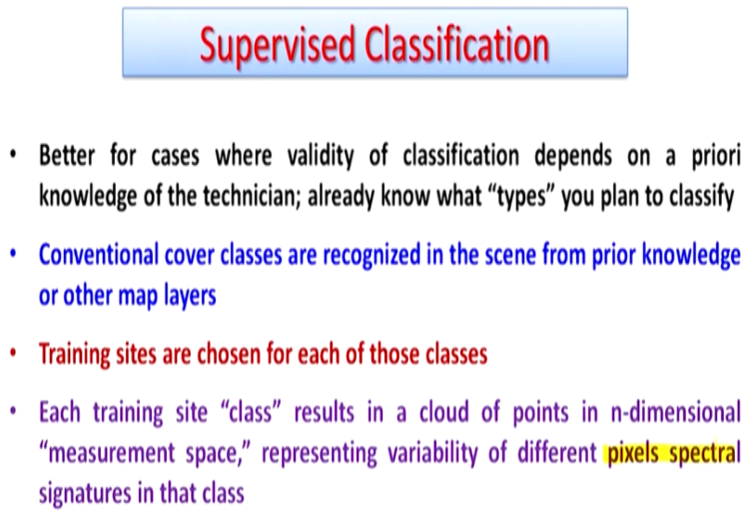
(Refer Slide Time 01:03)



As also we have discussed but very briefly I will mention the purpose of classification is to create a from a continuous change image data to create a map and that is better interpretable and easy to use by decision makers. So two techniques one is unsupervised that means fully based on computer no not much human intervention are there. And supervised classification that in each and every step human interventions are there and especially for training sets that is the most important thing here.

And based on the training sets and then corresponding statistics and then choosing appropriate classification technique we go for classification and finally also we go for accuracy assessment which I will be demonstrating to you in the next demonstration which we will have after this discussion on supervised classification.

(Refer Slide Time 02:13)



A supervised classification because there are human interventions are there and therefore a prior knowledge of that area is required and then only supervised classification can be performed. If Prior information about the area of which that satellite image belongs do if that information is not available then one should go for unsupervised classification rather than supervised classification

However if one is a very good on image interpretation techniques and having lot of experience of interpreting images then still one can go resort to supervise classification because while selecting training sets and which are the basis of supervise classification. If there is a confidence in selecting training sets in supervise classification then it possible that supervised classification can be achieved with good accuracy.

So where you know this this thing one has to really remember that prior knowledge is very much required because the experience is there of image interpretation then one can feel that different classes which are present in the image or scene can have a conventional like. For example conventional cover classes for a example means they might be forest different type of forest based on density.

We can say highly dense forest or open forest and then we might having agricultural land in that scene and we might be having built up land, we might be having bare grounds and you know wasteland kind of thing. So all these can be put in the conventional cover classes. If a image belongs to a hilly area then one has to be very careful. If image belongs to coastal area then again different kind of care is required.

However always necessary it is underlined here that the prior knowledge is required. If that is not there and then experience if that is not there then there might be some other source of information other map player can also be used for training the or creating the training sets. So training sites or training sets are chosen for each of those classes manually and this is done manually of coarse on a digital images which I will be just showing to you also.

And each training sites basically represents a class and it is not necessary that for each class only one training sets will be collected that can be many training sets for one class can be collected. For example if I am seeing and lot of areas are having forests then in different areas I will collect training sets and then say that if spectral characteristics of these training sets are same then classify as forest.

So each training sites or class results in a cloud of points which we will be seeing through a figure in a n dimension that is the measurement space through three dimensional scatter plots these clouds can also be shown. And then this represents the variability of different pixel spectral characteristics or signatures in that class. So whatever the variation within that class I take I took the example of forest.

So whatever the variability in the spectral signature of a forest class are there they are grouped as one.

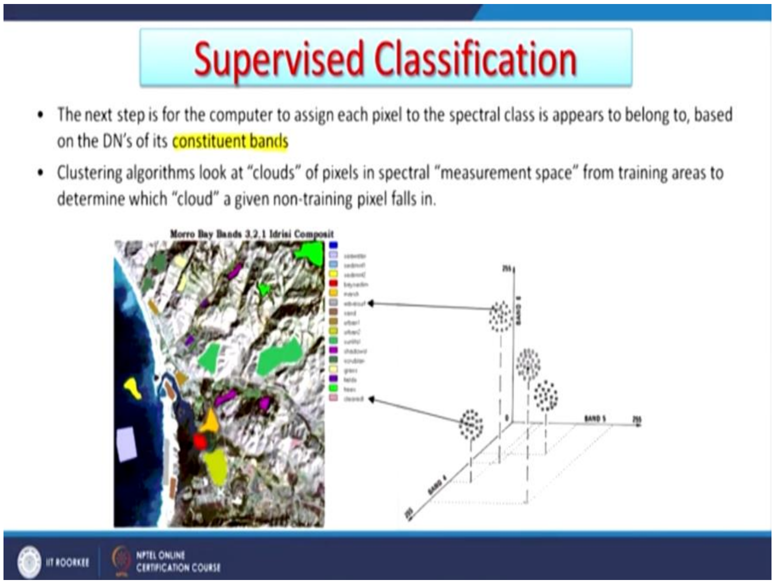
(Refer Slide Time 06:26)

Like for example here what we are seeing is a pre this is a side and one image is there of belonging to a coastal area and different training sets have been selected here. And if I focus on say on yellow areas then this is the you know one part in the water body has been selected and where it says their sediment to sediment that means might be some suspended or turbidity might be there.

But if you come very close to the coastal areas you are finding more sediments are there. So they are the side is selected here is different that is sediment one. And then also you know the sea water has been selected where we do not see any signatures of or evidence of sediments then it is the seawater it has been selected. So likewise in the whole area these things have been selected for a green areas two training sites have been selected which is a sanity cell and initially some names can be given. But later on these names can be changed it is for the training set.

And one more is there like here trees have been clearly identified and that so there are 2 training sets almost for each class at least there are 2 training sets for each classes are there. It is always a good practice and while doing this supervised classification that at least 2 or more training sets for each class must be selected within that image. So that the we get a better variability about the spectral characteristics against each class.

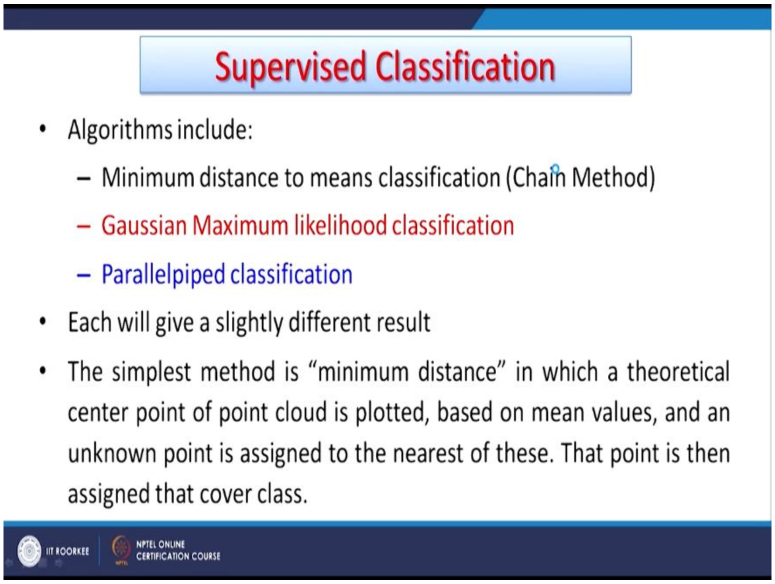
(Refer Slide Time 08:18)



And one say and these there have been and that means the computer has been trained on these. The next step is to assign each class to a spectral class and that appear or belongs to based on their pixel values and within that there you know because we will be performing on the you know color images or false color composite so on constituent bands. And then there are clustering algorithms which are also used to which looks these clouds of pixels in a n dimensional space.

And then the measurement is also done depending on what kind of algorithm one uses. And the pixels which are not falling in any of these classes are classified as you know unclassified or giving no training sets there. For example here a 3 dimensional three bends have been used in this schematic and a what it is showing band 4 band 5 and band 6. And these are the clusters which have been used here for example they might be clusters for one area and other clusters may be the waves of area and another cluster is here and so on so forth.

(Refer Slide Time 09:55)

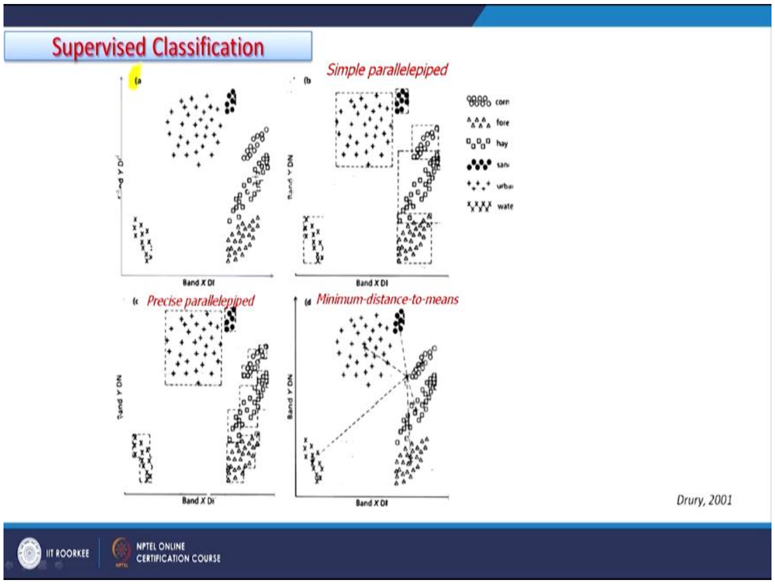


Now how which cluster will be assigned how this will be done and that we will see that the algorithms based on the different algorithms which are used in the measurement space to fix that this cluster belongs to a particular class. So algorithm which are used are the very that is minimum distance to main classification and so that we find that you know the clustering how the clustering is there if it is very open cluster then this minimum distance to mean is going to be large.

But if there are tight clusters then this can be one of the options of among the algorithms. Another algorithm can be a Gaussian maximum likelihood classification that the probability is based on the probability. And then parallelly piped classification is also there and we will be seeing soon a figure through which we will try to understand all these 3 well known classification techniques which are used in supervise classification.

And each if we if we put if we use the same training set of the same image but when we classify using these three different algorithms it is highly likely that we might be getting 3 different results but and might be very close to. For example the simplest method which is the first one that is the minimal distance is to mean is a in which the theoretical center point of the cluster or point cloud is plotted. And based on the mean values of that cluster and unknow point is assigned to the nearest of these and that point is then assigned the cover class.

(Refer Slide Time 11:52)



So likewise and now we will be seeing this and that in a 2 dimensional space rather than in 3 and just for understanding and this is of coarse estimating. But all three method which I have just mentioned have been used. So this is here the first one top left A is showing the different clusters of different classes and classes are shown here like circles are corn and this for hay and other urban areas and wasteland.

Now when we go for the first technique which is simple one the minimum distance to mean this is how many distance is calculated a point is the overall mean is faster here calculated and then it is found that which cluster a having minimum distance from the mean and then the mean of each these clusters are then their clusters are assigned different classes and then you go for classification.

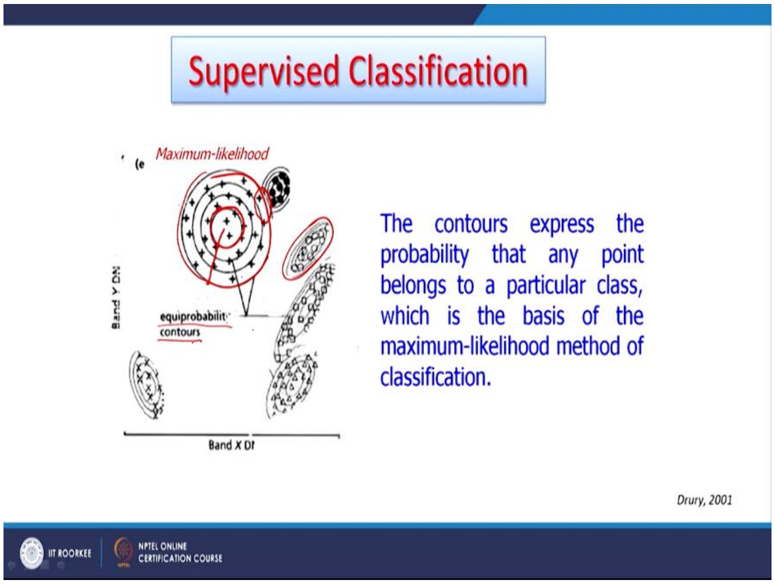
So if we look that the again that the minimum distance which theoretical center point or point cloud is plotted. So this is what the theory theoretical center point of the clouds is plotted here. And based on the mean values of these clusters an unknow point is assigned to the nearest of these and the point then assigned the cover class. So this is how the mean distance minimum distance to mean classification or algorithm work.

When we go for the second one like simplified or parallelepiped that you know rather than maximum likelihood the third one we will take that one. So in a simply parallelepiped what we are finding and that different kind of and instead of mean and they are boxes or demarcation is done on a two dimensional space. And then whichever the pixels which are find are getting within these demarcations are assigned to a one particular class and this algorithm is called simple parallelepiped.

There is imprecise parallelepiped which you can see that the large clusters are fine but when we focus here we find that like for this class example hay here a large area has been taken and therefore there are overlaps also. So in order to avoid those overlaps and one like these are the overlaps which you are seeing here and these are the overlaps and they will bring errors in our classification.

So in a precise one these overlaps are avoided and rather than just creating rectangles or you know squared to cover each class a very precise boxes are created like this here as you can see here and likewise. So this will give a more precise classification and that is why it is called precise classification same example here is also in this part which is stands for waste water or maybe sorry maybe the water class. So likewise this can be so different methods of a spectral classes and can be represented here and in the bivariate scatter plot.

(Refer Slide Time 15:35)



Now the second option which was a there the second algorithm which is the maximum likelihood it is based on of coarse probability. So the contours say are drawn and as you go away of the center of the cluster the probability reduces of getting those pixels in that particular class from the center. So the contours basically express the probability that any point which belongs to a particular class if I take this as example then then in class which is the basis of the maximum likelihood odd method of classification.

So the pixels which are having these values are having the highest possibility of probability of being classified more accurately and as I go away from the center the probability reduces. So there might be some like on the border area and there might be some problem because these contours of other class might be also overlapping there. It is not necessary that these contours have to be circular these contours can have electric electrical shape also. And which we say equiprobability contours also so likewise classification can be done.

So there are you can say there are overall 4 classification technique so for which are algorithms which are used in supervised classification one is minimum distance to mean then you know the standard this parallelepiped classification then precise parallelepiped classification and the last one is the maximum likelihood which is based on the based on the probability. Now as earlier mentioned that is you can see that the different algorithms are choosing different classes based on certain methods.

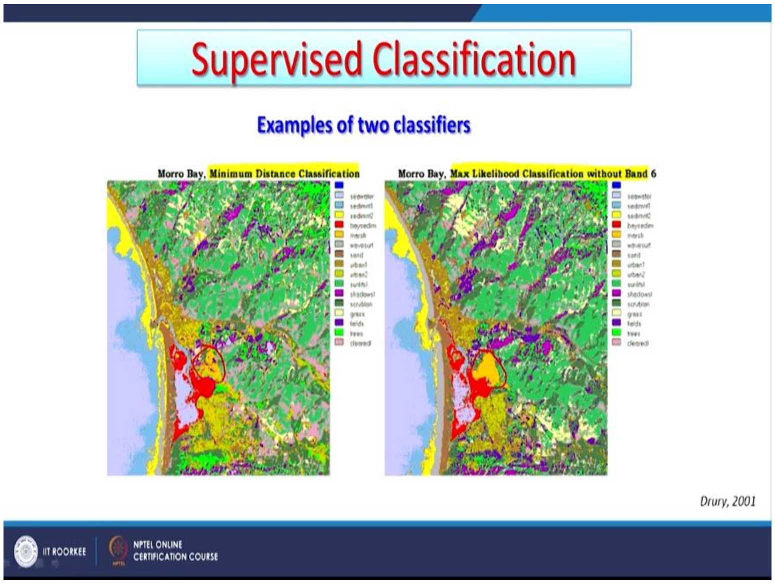
Some is using mean some is using you know the boundaries to define different classes in some this one like maximum likelihood is based on the probability. And therefore it is inevitable that if I put all different algorithms on the same image with same training sets I am bound to have different results. Now there may be a question and that which is the best and that can only be decided based on once if we are having prior knowledge or after the classification through all these techniques when we go ground truthing.

And we find that a particular algorithm has given the best result then for that particular area one can say that this classification technique or this classification algorithm has been formed good. But on the same image if the season is changed a different date images they are different year date image there you may get different results. So in one particular image minimum say minimum maximum likelihood maybe a more may give more better results.

And in some other date or season of the same area or maybe of the same sensor if I use again the same way I may find that minimum distance to mean MDM is giving better results. So this is a some base highly subjective and the classification the best thing is to find the accuracy part later.

Prior information or prior knowledge of that area will always help and will give the better results in the classification.

(Refer Slide Time 19:44)



Now I am going to show you the examples of 2 classification or using 2 classifies one is the first option that is the minimum distance to mean which is there which you can see here. And another one is the probability based method on this a classification that is maximum likelihood classification. And what we find that thou the number of classes is have been kept the same but if you start focusing on area wise or different areas you may find that some are getting better class better classified through one classification some classes and some are getting better classified using another class.

For example here I find that here in it is getting a better maximum likelihood that the maximum likelihood is giving matter more smooth results. And though the training sets are the same which have been used for these 2 classification algorithms. So there is in the same image same training sets 2 different algorithm 2 different results. That is the now which one is correct again as I mentioned that it has to be checked accuracy assessment has to be done then only one can say that a particular classification for that particular area have been found better or more suitable.

So this says as I have been mentioning that this classification is very subjective. Secondly so far I have not touched about the spatial resolution. The coarser the spatial resolution and better the results one may have because the difference in spectral characteristics and within training sets may not be the much as compared to high spatial resolution satellite images. So more you know the border in in terms of its spatial resolution and same in terms of a spectral resolution as well.

That if you are having a broader bends like in earlier Landsat MSS or even Landsat TM you may get you know very smooth results rather than speckles in your image in the classified map. But if you go for higher resolution satellite images and then the spectral characteristic for different training sets will vary very widely and therefore the output may not be as accurate that one would like to have it.

So that means what we can say at this stage that these classification algorithm also depends on a spatial resolution as well as on the spectral resolution or characteristic. So coarser these things are the better results one can obtain however but at that scale might be good. But we go for higher spatial resolution high spectral resolution our expectation through classification is much more and therefore we have to be very careful while choosing the training sets on high spatial resolution satellite image.

More care is required with the high spatial resolution as compared to relatively coarse resolution satellite images. So this is what a lot of the class accuracy of classification will depend on lot many things especially I am talking about supervised classification and because it depends on the spatial resolution it depends on which band you have chosen and then it depends on your the training set how accurately, how reliably you have collected these training sets.

And then finally assessing the accuracy part of your classification accuracy can be assessed through the statistical technique and accuracy should be assess if one would like to develop a highly reliable say land use map or forest cover map are out of through this classification technique or a lithological map and then ground truthing is very much required. And if a ground in the ground if it has been checked like an image it has been classified as a dense budgie or a dense forest but when we go in the field, we find that it is open forest not dense forest.

Then again and the training sets have to be changed and classification has to redone again so that it matches with the ground truthing. So there are lot of subjectivity are there non the less the main purpose here is to reduce the number you know the I say continuous image and through these classification techniques we are reducing into certain categories and discretizing the whole continuous data set into just few classes. So this brings to end of this brief discussion on supervise classification thank you very much.