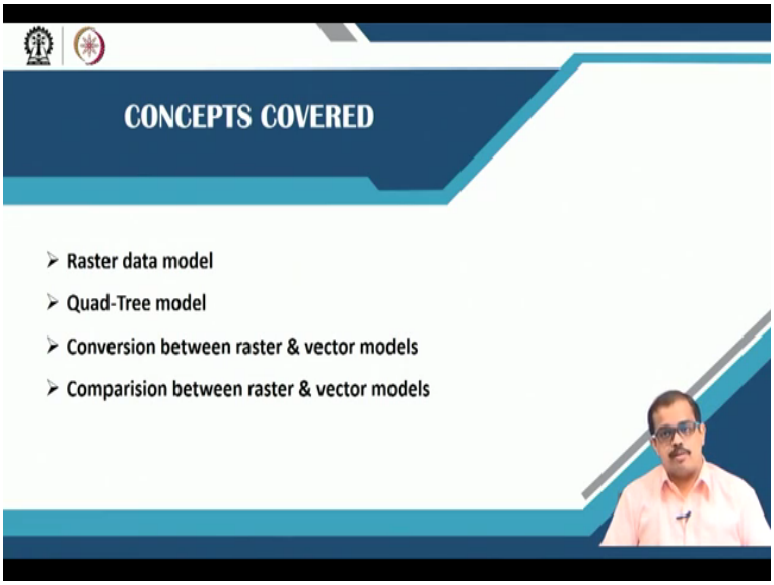
Geographic Information Systems
Prof. Bharath H Aithal
Ranbir and Chitra Gupta School of Infrastructure Design and Management
Indian Institute of Technology-Kharagpur
Module No. #03
Lecture No. #13
Representing the Real World (Continued)

Hello namaste, welcome back to the class, in today's lecture let us look at the second type of data model that is one raster data model.

(Refer Slide Time: 00:31)



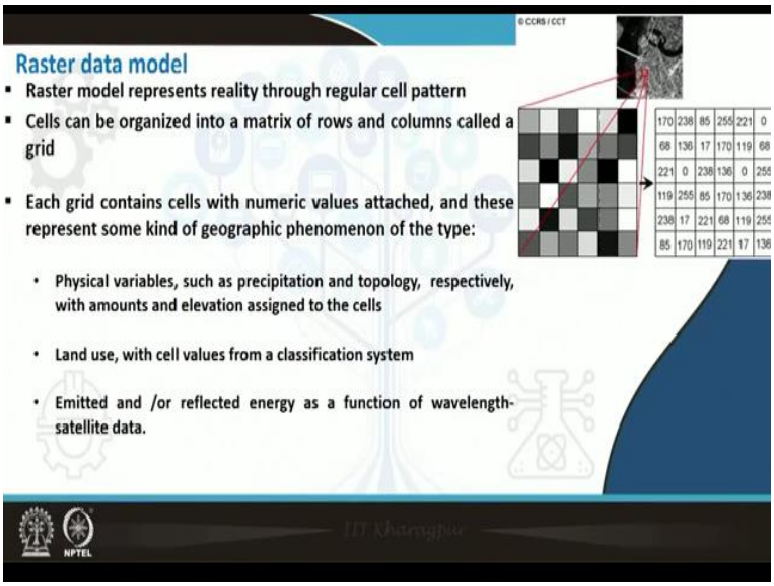
In my previous class we look at the vector data model, the 2 types of vector data model that are spaghetti data model and your topological data model. Spaghetti has its own disadvantages in terms of not having the topological information, whereas, where your topological data model has its topological information embedded into it. That is why the topological data model is quite better, but it is quite complex whereas spaghetti data model is quite simpler.

So, this is what we learned in the previous class, we also looked at what you mean by an attribute data and why do we need an attribute data and how do you store an attribute data, especially in terms of avoiding redundancy and how do I avoid redundancy it is a normalization, but in this class the continued class we look at what you mean by raster data model. The second type of data modern representation in GIS, we have 2types.

One is vector, the other one raster, so we look at raster data model today. How do we represent a raster data model, we will also look at a quatri model. It is a type of raster data model, will see what it talks about two tree or quad tree of quantity, so we look at this kind of model. How do you convert a raster and a vector data models there a way I mean, you can always convert raster model into vector model and vector into raster model me, we look at how it is.

And lastly will also end at this particular class with a comparison between a raster and a vector, it is extremely important because when you are generating a data the user should understand what you mean by raster and what you mean by a vector, what are the advantages of that model or the disadvantage of the model, without knowing this implementation of the data model it is extremely incorrect. So let us understand that in the same lecture.

(Refer Slide Time: 02:22)



Now when we are looking at the raster data model raster model represents a reality to through a regular cell pattern, when I say regular cell pattern on your right-hand side or this if you see this is a satellite rate of a particular region. So, you have if you see zoom into this region and if you can maximize and you see this is a collection of nothing but a cells okay, these are called pixels. It is bounded boxes which has certain great value shades okay.

When I look at the black to white shades that there are. So, the each of the shades are representing certain value start being represented here, the maximum value is 255 and the

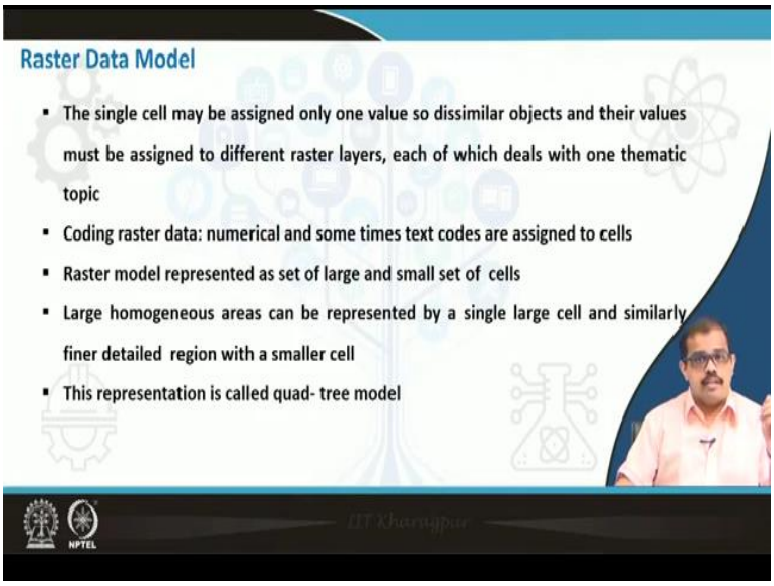
minimum value of 0 which means, say 0 is representing black and the maximum value is representing that is 255 representing y, so anything between 0 and 0 and 255 is represented between black and white okay.

So when you are representing, there is each of this value refers to the amount of reflection from the objects on the earth's surface, that is how we distinguish different objects on the earth's surface ok. So because it is a continuous data we refer to it as a field okay, so when we refer to this individual boxes, it is referred to a cells, cells or pixels okay, a cells can be organized into a matrix of rows and columns, and is also called as the grid, we had studied this in our previous module, it is a grid okay.

Each grid contains cell with the numeric values attached and these are present some kind of geographic phenomena, it can be a physical variables such as precipitation and topology with amounts in elevation assigned to the cells. So each if for example if it is topology it is it could have been , the values of the topological variations that are there on the earth's surface can be **a** is associated here.

If you are actually looking at any of the may be emitted or reflected energy as a function of wavelength of the satellite data, then this is the amount of reflected energy that is actually a captured by the satellite data okay. So it can be even land use with cell values from the classification system okay. So it can be any other thing that is being represented, but it is a continuous data okay.

(Refer Slide Time: 05:04)

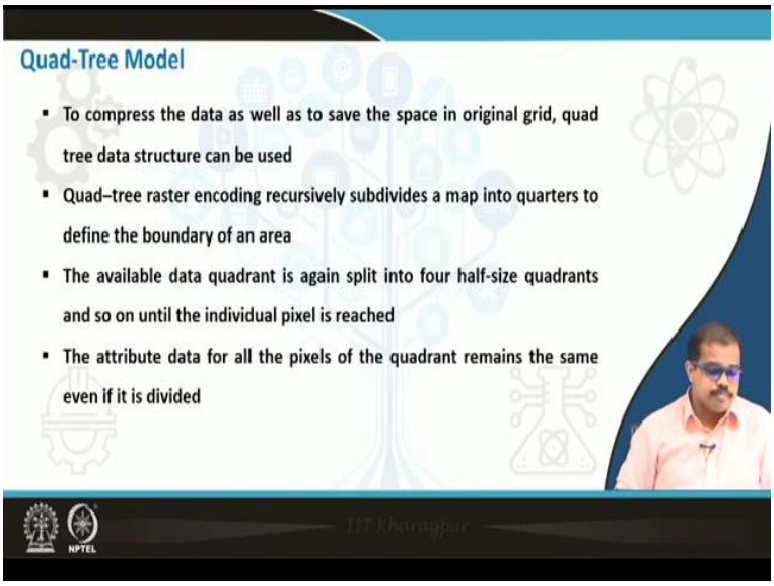


So when you look at a continuous data represented as discrete data, if you see this each of this value have a discrete value. it is always a continuous data that we obtain, but we convert it in a discrete cell values ok. Raster data model the single cell may be assigned only one value so dissimilar objects and their values must be assigned to different raster layers, each of which deals with one thematic topic always. Coding raster data is based on numerical and sometimes text codes are also sign fine.

So raster data model represented as a set of large and small set of cells. It can be large set of cells or a small set of cells, that is how you represent normally a raster data. Large homogeneous area can be represented by a single large cell and similarly the finer detail region will with the smaller cells. So that is how we group the larger region versus a smaller region. This kind of representation is called as quad tree, will see how it is done.

But as of now when you look at the raster data model can be numerical. it can be some type the type of text codes it must be assigned to different raster layers. For example, if you have number of layers each of these layers are representing the some dissimilar values in the entire the real world model. If you have a real world model each of these layers are representing I did show you in the previous class.

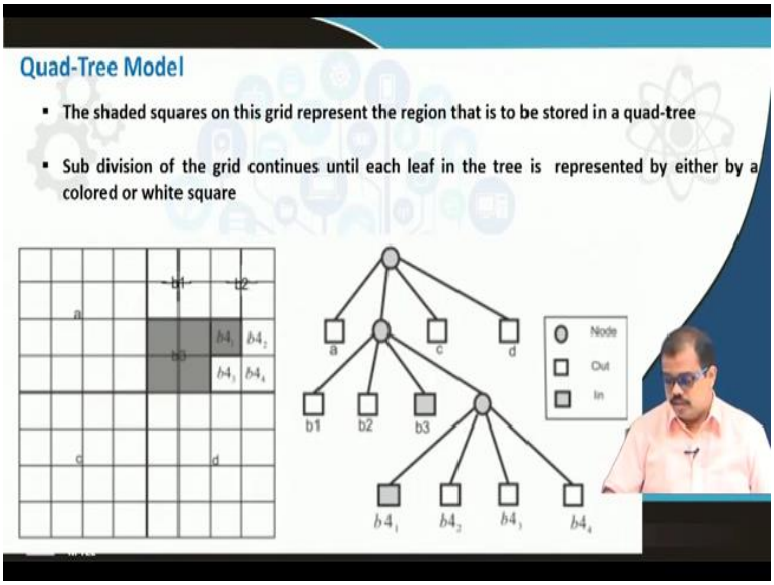
(Refer Slide Time: 06:43)



So when you look at the quad-tree quad-tree is extremely good in terms of compressing the data as well as saving space in the original grid, quad tree data structures can be used efficiently in terms of saving the data, quad tree raster encoding recursively is a subdivide some map into quarters to define the boundary, or an area, it means every map is divided into a quarters.

I will explain this in my next slide how does it divide and how will understand the division, the the available data quadrant is again split into 4 half size quadrant and so on until the individual pixels is reached.

(Refer Slide Time: 07:28)



For example, when we look at the next this one, if this is the entire image that we have

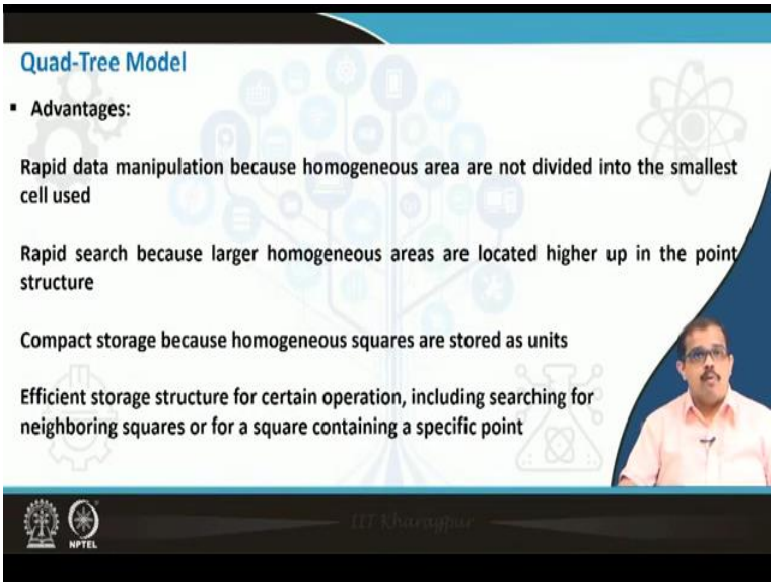
considered, if this is I am representing this image in terms of quad-tree raster data model okay, first what it does, it draws it into 4 equal quadrants, that is a b c and d and, now you have 4 different quadrants. Now, once we have 4 different quadrants what it does, for example, let us take this particular quadrant of b.

Now, this particular quadrant of b is then divided again equally into again 4 different quadrant that is b1 b2 b3 and b4. Now, once it has not reached an individual cell, but it is a collection of cells if you take b1, b1 is a collection of 4 cells 1 2 3 4 okay. Now if you do go into the b 4 quadrant. So now b4 as b4 1 b4 2 b4 3 b4 5 b4 5.

So, this once it has done b1 b2 b3 b4 then it takes individual quadrants and again divides into 4 different quadrants. So, it becomes in then drills down to this particular thing, which is a pixel. So, from an image to a pixel is how this quad tree model works, you have a node okay node is this entire image okay. So this node is then divided into 4 quadrants which is a b c and d. okay.

Now, if I take the b quadrant is again divided into b1, b2, b3, and b 4 which is here and each of these b1, b2, b3 b4 is then divided into b1 to b1 2 b1 b1 3 b 1 4 similarly b4 1 b4 2 b4 3 b4 4 okay. So now b4 1 is representing 1 pixel b4 2 is representing 1 pixel b4 3 is representing 1 pixel b4 4 is representing a pixel and these are individual pixels. So, this is how the quad tree model actually stores data okay.

(Refer Slide Time: 09:46)

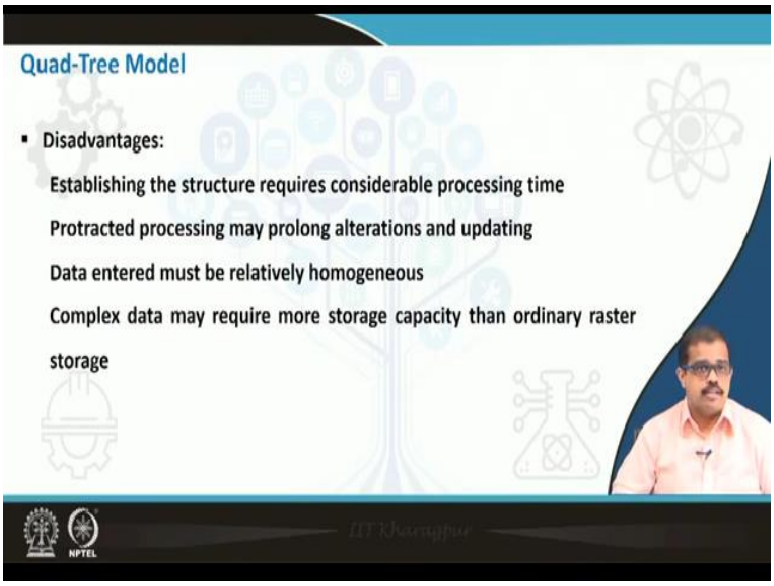


What are the advantages of if you look at the quad tree model, there are huge amount of advantages are there. For example, rapid data manipulation because of homogeneous areas are not divided into smaller cells used. So you can look at data manipulation in a very fast way, rapid search because large homogeneous areas are located higher up at the points. So, if you look at the first homogeneous areas are somewhere here okay.

But these are heterogeneous spaces that you may find, so that is how you that is one of the advantages of a quad tree model, then you have a compact storage because homogeneous squares are stored as units. So, when there is a unit it is the storage is extremely efficient that is by the quad tree model is quite advantages efficient storage structures for certain operations including searching for neighbouring squares or for a square containing a specific point.

If you are looking at a specific cell then, for example let us say someone wants to look at a specific cell which actually representing a specific quantity or quality on the earth's surface. So in order to point out that specific cell it is much easier in a quad tree done looking at searching in the entire database.

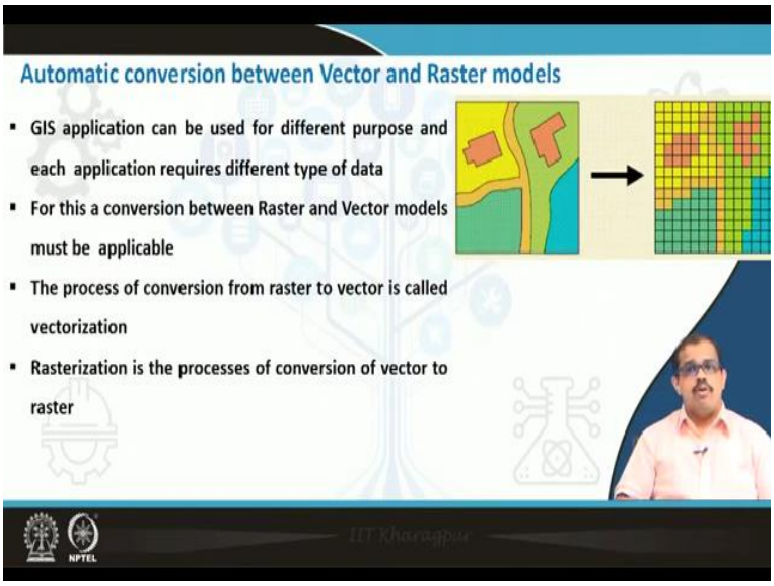
(Refer Slide Time: 11:07)



That is why it need, but there are certain disadvantages always, as always establishing the structure requires considerable processing. So, as I explained the entire structure and its extreme processing capability, then protracted processing may be prolonged alterations and updating data entered must be relatively homogeneous. This is one very important disadvantage data entered must be relatively homogeneous.

Then you have complex data that may require more storage capacity than the ordinary raster data. So when you have the complexity that increases in the raster data then you may need certain amount of storage higher storage capacity but always quad tree model is quite good in representing your entire or minimizing the storage requirement for your entire data to be stored raster data to be stored.

(Refer Slide Time: 12:05)

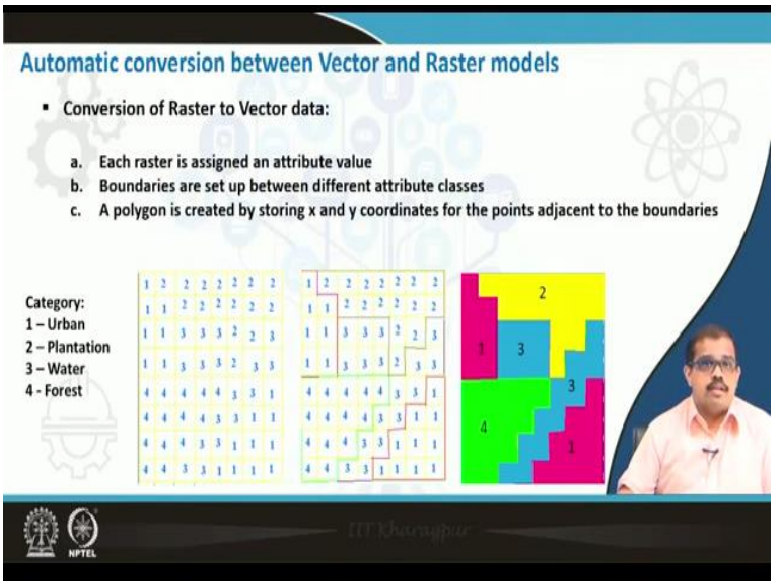


Now, once we have stored both, once we have understood both a raster model and a vector model. So, the next thing is, can we convert the vector to the raster or raster to vector, here when we are looking at it this particular image on your left hand side is representing a vector data model. If you see this is a road okay, this is a building ok building is a polygon, there is a set of points that is represented here okay.

And this is another polygon and this is another polygon here okay. So this particular house or this particular house is also in between this particular green polygon. So that is about the vector data. Now, if we have to convert this vector data so you need to know a process called vectorization okay. So, if you are converting a vector data to raster data it is called vectorization. If you are converting a raster data into a vector data it is called rasterization okay.

So when you are looking at this kind of conversion you need certain things to be kept in mind that is what we look at. So, when for example, you have this to be converted pixel. So, you should know what is actually the resolution of that particular conversion, you are actually looking at.

(Refer Slide Time: 13:33)



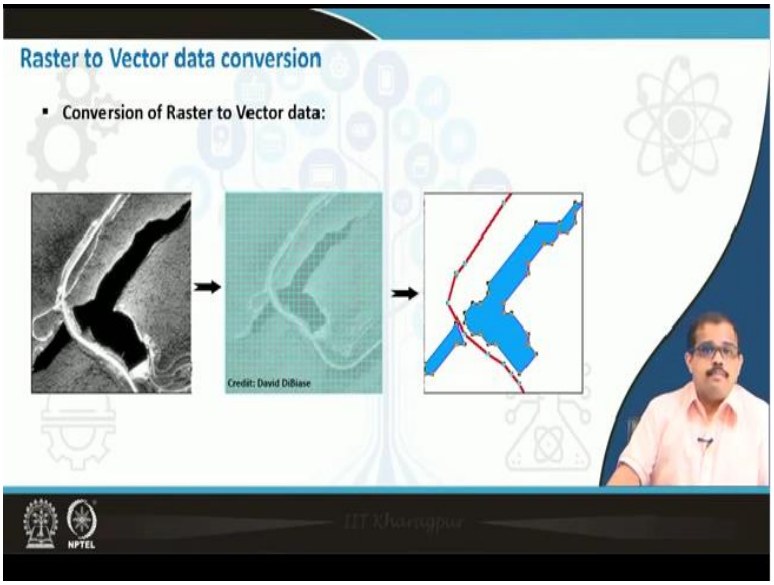
For example I will give you a complete details of how you convert raster to vector and raster to raster. Let us say this is a raster data that you have considered 1 to 2 all of these are representing different language classes in this particular image okay. Let us say we have 4 land-use classes, one is urban, one is plantation, other one is water, the other one is forest. So please keep this in mine because many of them in converting from a raster to vector, vector raster you go wrong when you are looking at how do you convert a vector to raster okay.

So, please listen to it carefully, each raster is assigned an attribute value I said attribute information is extremely important okay. And that attribute information can be urban, plantation, water and forest, as simple as that okay. Now, boundaries have set up between different attribute classes. So, if I consider 1 2 3 4 as attribute classes. so, you draw boundary something like this okay.

Number of polylines, this is the first polyline from the starting node, second part polyline, third polyline, 4 polyline, 5th polyline and 4th and the final polyline, all of this polyline will be only set up as a boundary where you have only similar information. If you let say if all of these are 1. so, this polyline camp could have been done something like this, but there are a set of 1s here, there is set of 1s here. So, both of these will belong to the same vector data or the vector entity okay which is nothing but urban.

So you draw boundaries across 1 boundaries across 2 boundaries across 3, wherever it is found, if it is found in a continuous not a pixels or a continuous draw it as a single boundary for it is different you have 3 2s here 2 2s here. So, each of this will become a polygon okay, so based on that you drop a boundary. Once you have drawn this boundary this each of this can be converted as a polygon line or a point. So polygon is created by storing x and y coordinate for a point, a point adjacent to the boundaries okay.

(Refer Slide Time: 16:06)



So that is how you do it, let me give you some example of how do you convert up a raster data to a vector data. So, let us say that you have this particular image of a raster image or a satellite image. So here let me go back and come to this. So, if you see there is the raster data then you have a vector data under this put here, then you start having the same for this polyline I am trying to draw across okay.

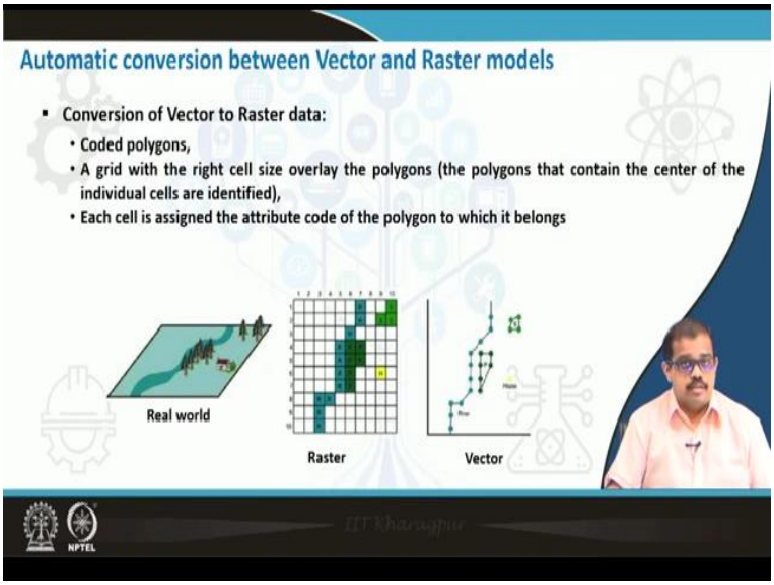
Once you have drawn this particular polyline, then I start looking at the number of points another polygon, see this is a similar point I am drawing across a similar point okay. So it is number of polylines from this polygon, you have another polygon here detached okay. Now it becomes a it is nothing but your entire vector to raster. Let me show it to you again. So you have this, this is the first feature that I am trying to digitize, it is a road across here this is a road okay.

You have 2 polygons, 1 polygon here, 1 polygon here. So, you draw this polygon first and you

draw the second polygon, we can also draw this as another road connecting road something like this okay. So that is up to you, but I have just shown 2 parts of it having a line under polygon okay. So I hope everyone have understood how do you convert a raster data into a vector date. So just let me show it is again. So, this is a number of grids that ah there.

So, you start placing it like this and you go on digitizing it. I spoke about how you digitize in my previous class, then you are digitizing it this lake, this water body here okay, this is the first water body and you have the second water body that has been digitized. So, you have 2 different polygons that have been digitized and can be used as a vector date fine.

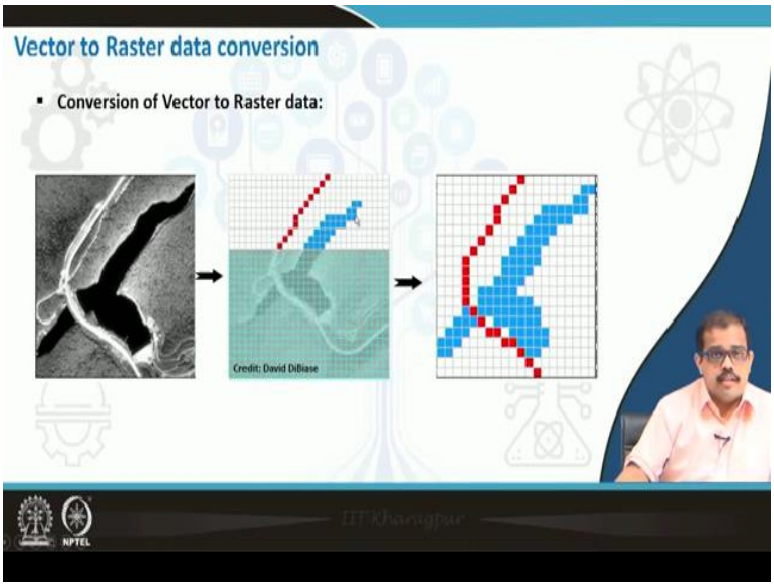
(Refer Slide Time: 18:10)



So now if you have to convert vector to a raster data, raster is a coded polygon's, sorry vector is a coded polygon, a grid with the right cell -sized overlays the polygons the polygons are contain the centre of the individual cells are identified. For example, if this is the real world okay and this is a vector information. Now you overlays the part of the cells that on a polygonc. The polygon here the size of the vector information is equal to the size of a raster cell.

By each cell is assigned to an attribute code of a polygon to which it belongs okay, I have already shown how it is done but just to signify how that the attribute code is belongs. For example, first we have converted it as a boundaries, once you have converted it as a boundary then you can convert is as a polygon fine.

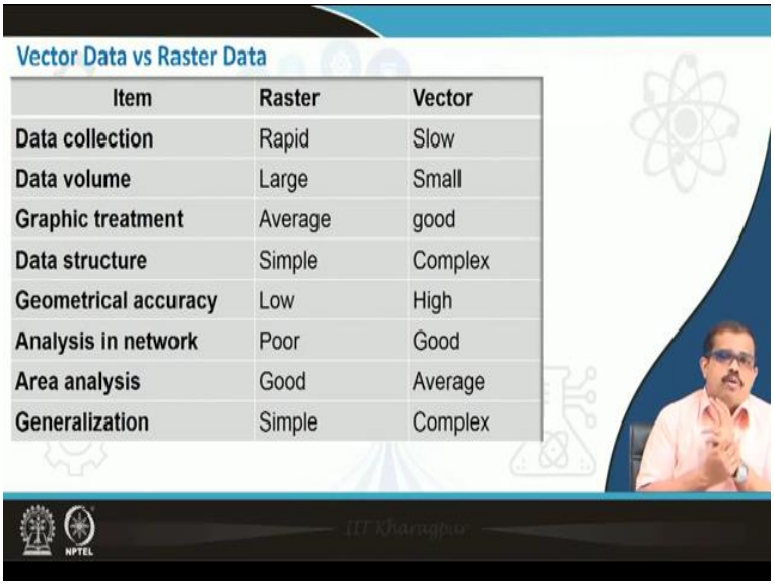
(Refer Slide Time: 19:09)



So to give you how do you convert a vector data to a raster data. For example, if a something like this okay, so you this is a vector data, if you remember, this was a vector data now, this was a line, now you overlay that particular as a grid, see these are different cells as a great and both of them will have the same connect system or same protection and the same way of looking at it.

So now this particular line is then converted into a feature that this in form of a pixel, you can see here every line is nothing but pixel that has been put here okay that is how you convert a vector data to a raster data okay.

(Refer Slide Time: 20:03)



So now when we look at vector data and a raster data, there are certain issues that is very important that you understand, by example when you look at our data collection always a raster data is extremely rapid in terms of data collection okay, you have various source of data the you you have information is that are easily accessible. So it is easier a fairly easier in order to collect a raster data, but when you look at the vector data is extremely slow.

Because, you should understand the quantities that are there on the ground and collect each and every quantity in the terms of point, line and polygons and the each in a different layer. So that becomes extremely difficult in order to have a vector data, but when you look at the data volume raster data has huge volume that are the storage space that is required in your entire database, whereas vector data is quite small. it is quite easily storable.

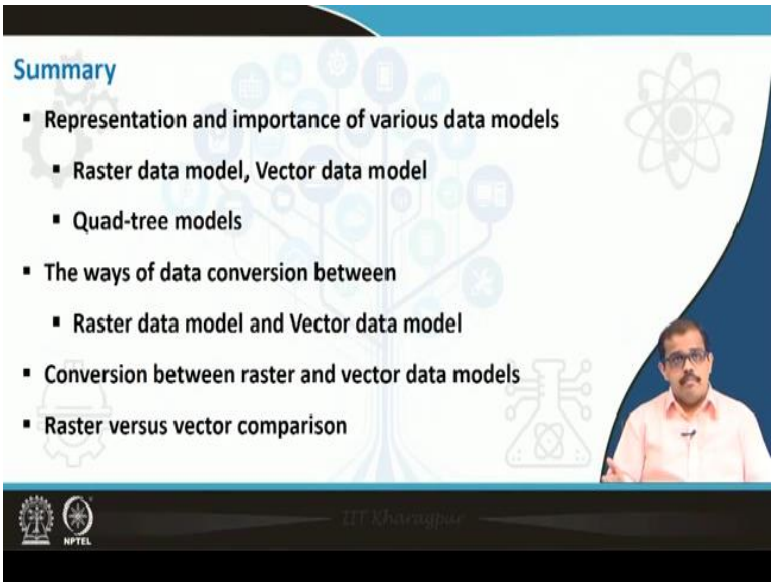
So that is why you have, a large amount of our vector data, in most of the data is stored as a vector data raster data. And when you look at graphic treatment of example, I was giving you an example of land-use okay in my previous slide, if you look at that particular slide you can understand that it is easier or it is much easier to understand the graphic treatment in a vector data than in a raster data because raster data is not very smooth it doesnot give you distinction between 2 objects or 2 phenomena.

whereas a vector data gives you a clear distinction. and, in terms of data structure raster is quite simple. whereas vector is a complex, we have spoken about quad tree, and we have also look at the topological data model. Geometrical accuracy always raster data has to load geometrical accuracy whereas vector data is quite highly accurate data collected from the ground, if it is a primary source analysis in the network.

It is quite poor because raster data itself as a leader great or data volumes were loading, should you should have a high computational capability and extremely good network, otherwise analysis in the network is quite of pork whereas vector data analysis in the network is quite good because it is small data volume and much easier to analyses and whereas area analysis is much easier in terms of a raster data, area analysis can be done.

Whereas vector data. its average , it can even save us are now below punched many times, but when you are looking at generalization raster data is simple, if you are generalizing the raster data then it is quite simple, but as a vector data, it is extremely complex in terms of generalization.

(Refer Slide Time: 23:03)



And coming back to the summary we looked at representation importance of various data model is very importantly, we look that raster data models and the vector data model, but in a h in this particular session we look at raster data model, we said it is a collection of number of pixels or areas and these area in the form of a matrix with m cost and number of rows and columns and this can be represented in a quad tree data model very efficiently used.

So how does a quarterly work, we first it divides the entire image into our entire raster into 4 quadrants then gets into that quadrant divides again into 4 different quadrants then gets in to that particular quadrant and digs deep until it reaches a single pixel, so that is how a quad tree data model works, then we have looked at how do you convert a raster data to a vector data and vector data to a raster data.

So it is extremely easy to convert to any kind of data that you need it depends on how you have to represent your data okay. So if you are working completely on the raster database or having entire dataset in raster data it is much useful to have a raster data for every values if you are

working on a vector database then it is much useful to have a vector database in the entire model.

So when we are doing conversion between a raster and vector data model you analyzed how it converts of a raster to vector data. So raster to vector data is in the form of a boundary builds a boundary, then it converts based on that particular whatever the attribute values that I give an example of urban , vegetation water and other categories. So, it converted into 4 category vector data then you have a vector to raster conversion, wherein you can see that each of this vector sorry raster to vector, vector to raster each of this area since then converted as pixels.

Now, then you have raster to vector conversion wherein each of this raster you will have to digitize and put it and give it an attribute value or give it a value and then it is converted from a raster data to a vector data. So, this is what we learnt in today's class, we also learnt about how raster data is is better in certain ways and how the vector data is better in certain ways.

So it is up to the user to choose what kind of data model that he or she wants to use in their analysis. So this is about different data models, how to use a data model, what are the different ways of handling a data model. So in the next class we look more into these aspect and let us meet in the next class thank you very much.