Remote Sensing and GIS for Rural Development Professor Pennan Chinnasamy Centre for Technology Alternatives for Rural Areas (CTARA) Indian Institute of Technology, Bombay Week - 06 Lecture No - 02 Intro to digitization of images for raster data

(Refer Slide Time: 0:16)



Hello everyone, welcome to the NPTEL course on Remote Sensing and Gis for Rural Development, this is week 6, lecture 2. In week 6, we are looking at coordinate reference system, projections and other details of the map that are found on GIS data for download. However, some data may not be readily digitalized or readily input into the GIS system, for example a map that we discussed in the last class, it is the only piece of evidence that we have to show how the land was 60 years ago, 70 years ago, 100 years ago, because at that time there were no satellites and photographs were not that efficient.

So, there were extensive surveys done before the independence by the Britishers and after independence also by the survey departments. These have tremendous potential but also have a lot of errors because covering the entire country using surveys will have lot of assumptions, limitations and challenges not all regions are accessible, so that accessibility question is taken away by the satellite, drones and robotics. However, we also need to know what was the baseline scenario 70 years ago, 80 years ago to showcase the difference in land use land cover for example rural development, roads, transportation, electricity all those things.

So, there is very big information that we can collect from maps and that is what we will be looking at in today's lecture, how to tap into this tremendous information that is hidden in maps, when I say hidden it is not that you have to find very carefully but you will easily miss it, because there is so much information put in a map a paper map, so how are you going to extract it, how are you going to create new shape files, new data GIS data from your paper maps. So, let us look into that aspect in today's lecture.

(Refer Slide Time: 3:14)



So, the first aspect we will look at is converting images to raster, so we know that images already have a graded fashion, for example the image on the screen is a maps image it is a digital scan, so it was a paper and the paper image which is digitally processed or triangulated using survey data, it is scanned, it is a huge map, these are not small maps, it is a huge style and you have commercial scanners. So, where like for example posters you see that big of a scanner. So, when this scanner scans the details it the output is an image which is it grids it and every grid very high resolution puts a colour.

So, now this colour information is transported into your computer or hard disk as a image but there is no geospatial location attached, this is as same as taking my own image like my photograph and putting it here like my passport photograph for example, so both are same because my passport photograph does not have a geospatial location, it is just an image of myself and it could be taken in Mumbai, Chennai, US, Malaysia wherever it is the same me and that is always important, there is no need for a geospatial location. However, a map which depicts an area should have a geospatial location, because that is what it is representing, I am my image is representing myself, my photograph represents me whereas an image of a map represents a geospatial location.

So, I hope this is clear that there is a difference between scanning an image and georeferencing an image, scanning is part of geo-referencing, but all scanned images are not geo reference, same thing all photographs are not geo reference, if your camera, for example if your phone has a provision of putting the lat longs that we discussed in the previous class latitude longitude along with the image taken, then you have a geo reference image, it is heavy it because it has to attach a lot of files heavy as a not like a normal image it will get more information, so we will have to be careful in triangulating the data for this.

So, not all cameras are fitted with a geospatial location specific camera hardware but only some cameras have and it is expensive, these cameras are expensive than the normal cameras and you will find them easily in the market, but the point is we would like to showcase more open source systems as much as possible and also showcase how you could collect information and data through GIS in a cross-effective manner.

So, in today's lecture we will be looking at an image of a map that has been created by a government agency and we look at the specifics that one has to understand in detail before geo referencing or converting it into a raster. So, images can be made raster data with geospatial location all raster's are images, it has an image, it has pixels grids but not all images become raster data for the raster data you need to have geospatial location.

So, I am stressing on the geospatial location otherwise I will show you an example in the QGIS hands-on when we do this, I can open my own image in GIS it will open directly but it will not have a geo special location, so it will float, but where can it float, it has to be anchored to a particular location and that is what we will look at in this lecture.

So, moving on maps are still in the paper format as I said other tremendous amount of data in maps, if you want to do a long term analysis. So, I have been to a lot of these survey offices when I was a student and a researcher Junior researcher after my PhD and it is tremendous amount of knowledge it is not that easy to access it you have to write a letter as a student you can get letter from your head of the department or a professor saying that why you need to access and then you can walk into these survey offices give the letter and they will make it available.

So, first is to examine the paper maps which means they will not have time to go through all the paper maps let us say you want Chennai Mylapore map they will just say the number of the map but there are there could be multiple paper maps on Mylapore both temporarily and depending on what is needed, let us say a road map, there is a temple map because a lot of temples are there in Mylapore or a tank map.

So, all these are different, different maps, so you will have to go through the library of the maps and then physically look into all the papers and stuff and sometimes they preserve it well but still it is still a paper, so the old maps sometimes do get degraded. So, you have to be very careful, some people use gloves in foreign countries but because in those days there was no scanner, so they made the maps but they could not scan it now you can scan it but the resolution the detail is not the same. So, you should be appreciative of these maps it has a lot of data.

So, what people normally do is they take the information on a paper and then go back to the to the library it is like a library it is not books you have maps here and then you take the information and then give it back. So, they will they will give you an option to buy these maps from them in terms of copies and you cannot take some of the paper maps you cannot take out of the office, so you will have to request for an image or a scan copies, Xerox copy and they will make it available for a cost, because there is a lot of infrastructure they have to maintain for it, so there is always a cost.

So, need to be digitized for extracting data as I said maps are just paper or an image is just an image with data but no geospatial data, a map is just a paper. So, it has lat longs and information of the geospatial but it cannot just sit into just if an image cannot sit into the GIS platform. So, there is a difference in in using a GIS platform and a photo viewer in your desktop.

So, you have multiple software's per viewing image, so what is the difference between looking at an image in a photo viewer and GIS? In GIS you give location, whereas here there is no location needed a photo of a bus is a part of a bus, it need not have a location. So, same thing analogy you use here.

Tremendous amount of data can be extracted as I said when they do these surveys every single they love checklist, so every single data is input into the system and we will have to be careful in looking at what data we need and extracting it, if you extract this all this image data into a digital image, it will be too big, it will be too big to open it, even opening this image in

GIS is going to be hard, but what we are going to do is we are going to open it in GIS as a base layer and then from the base layer we are going to convert it into data that we require as a smaller amount.

So, once you down scale the image instead of taking all of it for example this image has airports, post offices, hospitals, police stations, temples, houses, roads, railways, etc, etc but we need only roads that are connecting rural to cities. So, for that part I do not need the other data that is present in the map.

So, the whole objective of this week is going to be how to extract information from these maps. So, there are tremendous amount of data that can be extracted as I said you will have to first make a checklist of what data you need because when they went to the field they have a large extension checklist of collecting data, this is very similar to census data census you may only want the amount the number of males and females in your district.

But if you look at the sensors download and look there will be multiple columns, they will download a lot of data for every aspect and you will use what aspects you need. So, please understand that this data can be used by data careful level.

Usual digitization is expensive as I said the scanners the commercial scanners that you saw is very expensive to convert digitally and mark digitally on the maps and then putting it into a geospatial data. So, GIS has come in as an inexpensive tool still you need this image, so the image is has to be scanned.

So, let us say this is a paper map and the paper map this is a newer map, so it was digitally done on a computer and then printed on a printer on a commercial printer, but the maps that I am talking about 1980s, 1970s, 60s, or even independence 40s, 50s, etc you will see that these maps were still paper maps and these paper maps have to be scanned.

So, the same printer that builds these maps can also scan, so you scan them and that scanning is very expensive because every single inch centimetre of the map has to be scanned. So, it goes very, very slow very high in definition it takes time not like your scanner like this.

So, even this scanner if you see an A4 paper sheet you will place on the tray, so on this tray you place an A4 sheet and then the scanning happens at every single centimetre throughout so that it is continuous data is continuous it does not stop in the top and then go back in the front and then take data the whole sheet is scan. So, that is what we need to accomplish here using the digitization of these maps.

So, let the scanner scan it so this is a kind of a printed map or you can also call as a scanned map from a digital device, however, it does not have geospatial location in the data on the map, there is I will show you where you can find the geospatial location on the map there is a lot of information we are going to extract it. So, this is a large scanners are needed as I explained and printers to make these maps.

But GIS provides a simple and open source solutions for all of this, so instead of asking someone to do it who you have to pay, we are going to make mappers through this lecture. So, I will be teaching you in this week how to digitize this map and then extract data which honestly I have seen a lot of consultants doing for a very, very high price for rural NGOS. For example a map like this to extract and give data I have seen NGOS charge 1 lakh NGOS pay, I am sorry for consultant and that is literally half a day's work or even two hours work not more than that.

So, why they give it is because the capacity is not built the whole aim of this NPTEL course is also to provide capacity to all students to use these software's and also use it for a particular course, here the rural development is used maybe I will have a GIS and remote sensing for urban development, but for today this lecture it is rural development.

The principles are same but the applications are different and much more sub objectives and principles will be discussed if the course goes as an urban development course, I think a lot of people are there for urban development, but for rural it is very, very less you do not see many courses for rural. So, I am very happy that a lot of students have registered for this course, it is a very important course and I have travelled in the previous lectures why rural development is important.

So, as I said GIS provides simple open source solutions and open source because we are using open source data plus open source software, if you use a proprietary software then the price goes up and some people are very towards the proprietary software because it is easier to use, as again if you have funds you can use it otherwise open source free QGIS is still good enough, it is as good as any other software. Let us take a look at an example closely as I said this is the image map that we are going to use for the next 6 week lecture and throughout you will see how we use this map in detail.

(Refer Slide Time: 17:57)





But first let us have a look at a closer look at this map. So, you could see that there are multiple, multiple information in this map are we able to digest everything, it is hard difficult but let us see what we can extract. So, you could see here there are published by the director Survey of India and then where it is Department of Science and Technology who funds it and then you also have the scale, datum, projections, we look at some of them and the legend so this is what is called the legend or what does each colour or symbol represent.

So, if you see here you have this as boundaries and then here we have railway, roads, train, tracks etc on this part and then you have airports you have helipads and then you have post office. So, all these are marked on this map, how much of this is rural development that depends on you because you are the one who is going to extract the data, now we will look into how we are going to extract data.

The most importantly there are sheet numbers, so this is the sheet number for this image and you could see that in Bengaluru Karnataka, so the major area is Karnataka and then it is in the Karnataka tumkur region and tumkur Bangalore Urban. So, there is Bangalore Rural and there is Bangalore Urban I am just going to show here Bangalore Urban because it is on the border of rural and urban.

So, you will see a couple of boundary effects here between rural and urban and you will see the tile number. So, this is the tile number D43 R12 and then R11 is there, R 15 is there, so there is a schema they use to mark the numbers, let us not get into that but every sheet every tile has a number. So, D43 may be representing the entire belt and within that there are sub pixels sub grids which are labelled as R12, R1, R11, etc. It is also important to understand that these are gridded, so the entire India is gridded into sheets and then within the sheets there are some grids and here D4312 is there and within that you have the lat long. So, you have the latitude which is here a 77 it starts and then goes on from 77 degrees 30 minutes to 32 seconds 32 minutes and 30 seconds.

So, 77, so here is also 77 but it does not come as some sense they want less, so you will see some writings are very half written why, because they want to not overcrowd the map with writing, so that is why you will see a short form of writing here. For example this is also 77 degrees 32 minutes and 30 seconds this is 77 degrees 13 minutes 0 seconds but you do not see the 77 here, all throughout is 77, same here 13 degrees 15 minutes, this is 13 degrees 12 minutes 30 seconds, this is 10 minutes. So, how does it differ? It differs by 2 minutes and 30 seconds there is a grid.

So, now you go 10 and then you go to 730, so 10 minus 230 is 730, so all this is there. So, it is minutes and seconds similar to your clock, so 30 seconds is half a minute which is 60 seconds is 1 minute and that is how the grid is developed. So, that is one part.

And then let us look at source of the map the source of the map is given on the top and also on the bottom you saw that the source of the map is a Survey of India which is a government agency established in 1787. So, as I said kind of British era also 1787 and 67 and then you come here at the bottom also, you can see that published under the director of Survey of India and then which state which location it was done. So, all this is there and the director Karnataka geospatial data centre Survey of India block Bangalore. So, all of this does relate back to your Survey of India products Survey of India.

So, there was a price for this 77 rupees, so some data is free open, so 77 rupees is not big as I said it is low cost the maps may be some of them are expensive also, so you will look at some of the examples.



So, then the first important thing is the source of a map that we have seen, then we jump into the scale, as I said scale is how much is 1 unit representing the real life length, let us say 1 centimetres how much is 1 centimetre on the map equivalent to on the ground. So, if you look at this closely here the scale is given as 1 is to 50,000. So, every unit that you measure on the map is equal to 50,000 times on the ground.

So, if you look at this example let me zoom in you are zooming into this lake part, so Keres kind of a lake and if you say I put my centimetre scale is 1 centimetre, 1 centimetre in length this part let me put my pointer, so that you could see what I am showing, so this part. So, this lake I am looking at and this distance, so distance from the boundary to this boundary let us assume it is 1 centimetre. So, what does that mean? The length is 1 centimetre and that equals to 50,000 centimetres on the ground.

So, now you know that 100 centimetres is a meter and then a 1000 meters is a kilometre, so if you do the conversions you can convert it to kilometres or meters as per your need. So, that is how you scale up and then look at the distance how big this lake is, it is a smaller lake compared to the other bigger lakes that we have down here, Yelahanka lake is pretty big.

So, moving on then you do the survey number as I explained earlier every sheet is marked by a survey number. So, there is an overall sheet number which is giving you the Karnataka region, for example and within the Karnataka region you can choose a smaller region and then take an image.

So, one of my the TA of this course and my PhD student Pranadh is working on this area, so we had to buy these maps and that is why I am showing how to extract data from these maps. Again these were if you look at this region especially in this region 50 years ago, 60 years ago, it was the rural region now it has become Urban.

So, there is a lot of shift from rural to Urban we do not call that as a development we call it conversion, development is rural development when we talk about rural development is how does development happen within the area for rural entities, let us say farmers and ecosystem services, animals plants biodiversity, etc. Converting rural to urban is not called development here, it is conversion.

Whereas, if you increase the land under productivity that is rural development giving more access to people for water, energy, health in rural regions that is civil development not converting it totally into an urban thing which means the livelihood should still be rural, here in an urban setting the livelihood is very different industries, offices, etc.

So, north arrow and direction, so the map if you zoom out of this style you will see that it will be given a north arrow and here it is not this upwards which means I will draw it, so north this like this. So, this is not normally we see the N is symbol and once you know the N is simple the others can be taken up opposite to the north is South and then north you have to the right is your East and then you have your West.

So, there is a understanding why West Indies was call West Indies, so because from Europe they wanted to find India and when they sailed this was a very, very historic when sailors were trying to find the route of India, so that they wanted to do trade by gold, spices, silk, etc. So, they found accidentally, so instead of Europe they had to come to East to find India that is why it is called East trading.

So, from Europe you know Europe is here and India is here. So, from Europe they had to go around Africa and then come to East, East is where India was there not West but accidentally the sailors went to West side rather than East. So, they travelled this side and then they landed in a land which is very similar to India, so the climate the people, the food everything was very similar but they thought something that oh I do not see the gold I do not see the silk I do not see the Palaces that people said and then they labelled it West India or that is why West Indies is called.

So, this is a small historic story people share we need to look at the actual but saying that the sailors found these regions from Europe and instead of going east and west you can see how a map this north arrow is very, very important. So, if the Sailor did not have this correct instead of east it was pointing west, so the sailor went west side and then they found a different land. So, here this can be avoided by just using a proper arrow marks.

And then the year when was this year done this is very important as I said we need temporal analysis you can find the year on the map by zooming in as I said there is a lot of data in this map you will have to find where this map is having data and what data it has. So, somewhere here you have 2005 in the copyright you have 2011, 2001. So, that is what the year we will be using as the year of the map.

So, there is also references on photo reference and then yeah we will use the year when we open the full map we will find the year of this map also I have truncated the map, so we

would not find it here but yeah here first edition is 2011, so you have the 2011 as the map here.

Then the projection and the coordinate system this we have discussed in the previous lecture and that has been given here. So, the projection is UTM it is a type of a projection there are multiple types, so for this region the Indian subcontinent and Karnataka central region UTM is pretty good. So, the projection has a light bulb experiment I showed is good for this part.

And then you have the datum where is the centre for the coordinate system and that is given as WGS 84, WGS 84 again is a type of coordinate reference system CRS with a particular datum in the centre. So, all this information is given.

And then you do have your lat long as I discussed earlier the latitude is the line that runs from top to down vertical and it is given as 13 degrees. So, here it is 13 degrees it starts and ends within the 13 degrees, it does not go beyond and the same longitude is at 77 degrees, so 77 degrees is where this style is located and then there is subdivisions.

Legends in important places where here I have shown it when I zoomed in, so these are the legends and important places where we need to collect data and stuff. The cost is given as I said for a particular reason we had to buy this map and the price is 77 rupees, again 77 rupees is still affordable by many for rural data extraction and the area covered here is pretty big. So, we are not looking at a small area we are looking at a large area for analysis.



(Refer Slide Time: 31:30)



Moving on I also would like to open this data portal where you can take the data, there is multiple, multiple things that you can download in terms of geospatial data, let me quickly open it.

So, if you click this link you will get into this webpage you may open it for you and we will see that you come to this webpage when you open this online survey map of India, it is an online portal by the government of India, department of science the technology minister of science and technology and it has a lot of products, some are very, very costly 6900 rupees for a particular geo database high resolution etc, whereas, you can also get open source maps.

So, you can say free maps and then it will format, it is just searching for all the free maps that are available. So, open street maps is there you can click to download the product it will open a page and then ask you to give the mobile number password etc, for which you should register. So, always have this registered you can sign in and then do it.

But let us go back to the Survey of India and you could see all the products here you have all products topographical maps which are the DM and then a digital elevation models you have the digital elevation models and administrative boundaries which is the administrator boundary of states, districts, sub districts, villages talukas, blocks, etc.

DM is digital current model we will be using this in this course, and geo-reference colour raster and then digital geographical map, village boundary database, village boundaries the names the attributes a lot of people are asking questions in the forums I need this data this is where you are going why I am promoting this is because government of India and the boundaries are very, very accurate because it is promoted by the government of India, open series map it is a free Pdf. So, this is where you can collect the open series map I am just going to click it we need to take out the other digital products and other maps are also there.

(Refer Slide Time: 34:01)







So, here when you go to open series map you will have to give the number of let me first go home and then look at the products. So, those who search for it on a different Google search or other search database you will come to the home portal and from here you go to products I have given the product link digital product link in the link itself.

So, this is our digital products open yeah, so open series maps and free other maps you can click the free other maps and then you see that open series map OSM is there which will be also using in this course. But if you look at the boundary maps you can see that there are a lot of other boundary maps let me go through their maps. So, you come down this digital vector database 68000, so you read it with 68000 and the size is also there 1 is to 1 million sheet.

And then you have the administrative databases these are free and I highly recommend you to use it. So, what is this is the entry entire country taluka boundary, taluka boundary with the names look at the boundaries, this is the boundary that you will have to use you cannot use different boundaries that other foreign Nationals, foreign agencies are showing it has to be accurate by law it has to be accurate. So, please use this map for any development of India reports and stuff.

And then you have the entire country district level entire country taluk up to taluk level with headquarters some differences there between these two maps and then the district level map with headquarters. So, taluk, districts, state and then the whole boundary this is a digital terrain elevation model, the elevation boundaries we will be using an open source but still people who want to use it can use this map.

And then geo reference colour raster 1000 rupees, digital geographical map you have different, different maps that are produced by agencies through the through the database and then you have village boundaries as I said earlier and then open series open series map. So, this is where you will click to buy it, when you want to buy it you will have to click download and give the sheet number or enter the enter first you have login and go and then you will enter the store.

So, this is where you could get your data and information from these maps, so I will also go through in the next lecture of geo-referencing the map and extracting the data, until then I will conclude this lecture and see you in the next lecture. Thank you.