INDIAN INSTITUTE OF TECHNOLOGY ROORKEE

NPTEL

NPTEL ONLINE CERTIFICATION COURES

Digital Image Processing of Remote Sensing Data

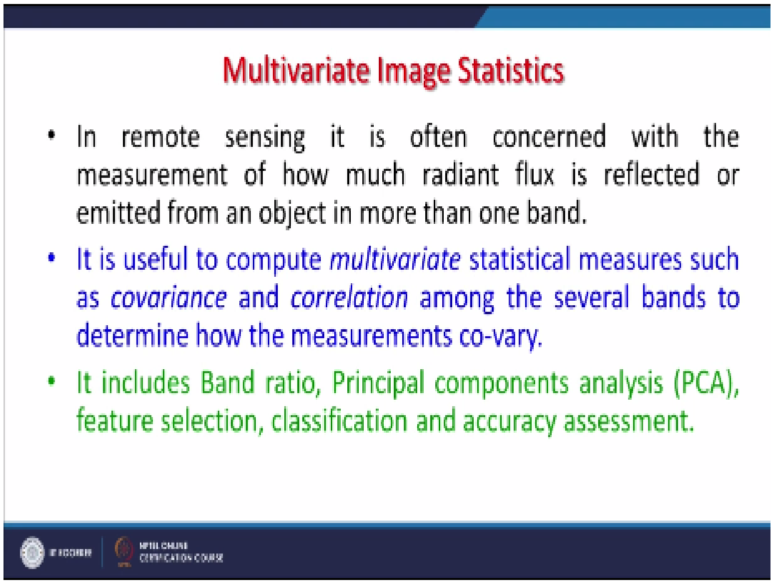
Lecture – 11
Multispectral Transform, Scatter Plot, Principal Component Analysis and Decorrelation Stretch

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

Hello everyone and welcome to 11 lectures of digital image processing of remote sensing data course and this we are going to discuss about how multi spectral analysis especially involving several bands that is multivariate analysis can be done digital images processing so in this we will be going through the scatter plot 2 dimensional plot band ratio analysis principal component analysis and deco relation stench.

So these are the topics for this first about is basically multivariate images statics because we know that the most of the remote sensing data is acquired in multi spectral form that means lot of bands data is coming and therefore we need to analysis the data so it is a often concern with the measurement of how much radiant flux is reflected or emitted.

(Refer Slide Time: 01:21)

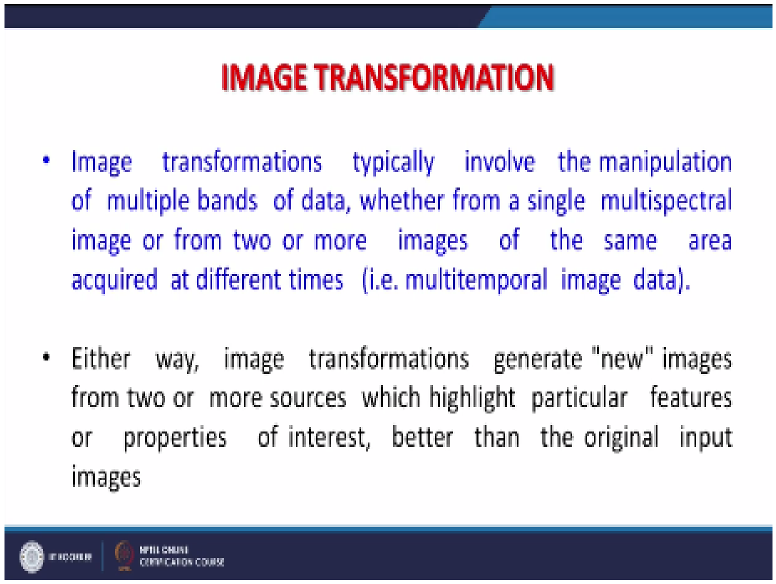


In case of a thermal infrared from an object in more than one band, because this is the band thing is not a single band analysis where histogram is more sufficient and stretching absence or also simple but when we over multi bands or multivariate analysis things becomes little complicated so it is useful to compute multivariate statistical measures such as covariance and correlation among several bands to determine how the measurements co vary.

This is required because if there are 7 or 8 bands in a one scanner or sensor then we many of these bands might be highly correlated to each other and so that whenever we are going for some color combination or composites we need to have those images which are different than each other and therefore the bands which are correlated to each other may be dropped in color composite. So this kind of analysis will allow us to understand and select the relevant bands.

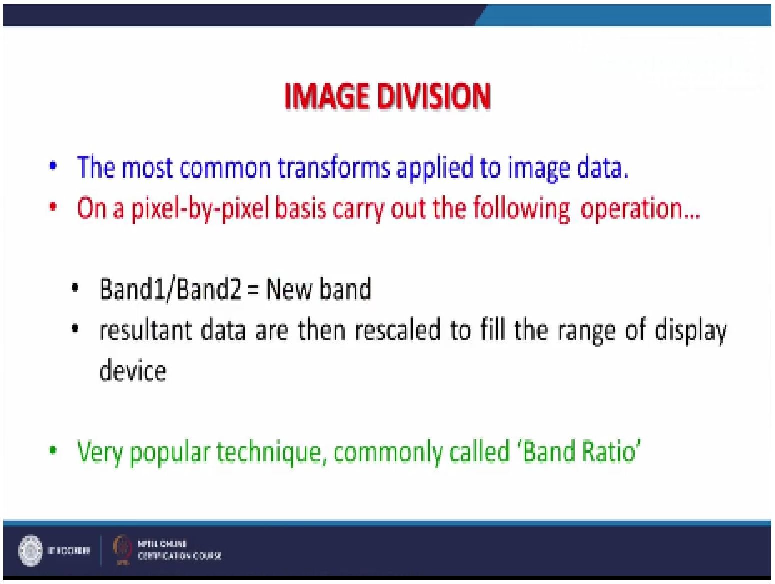
It includes a this multivariate analysis as I have mentioned earlier band ratio in techniques which we will discuss several limitations and advantages with band ratio principle component analysis very popular one PCA feature selection classification and accuracy assessment so mainly image transformation and that is multivariate that the image transformation typically involves the manipulation of multiple bands of data weather form a single multi spectral image

(Refer Slide Time: 03:12)



Or from 2 or more images of the same area acquired a different times that is multi temporal image data so that can also be analyses in this way and either way the image transformation generate new images from two or more sources which highlight particular features or properties of interest, better than the original input image in case of single band images most of the things which we have been doing though the values and may be in some case ne images may not be generated but in the multivariate analysis these things are there so first very simple multivariate analysis involving more than one band may be the division or band division.

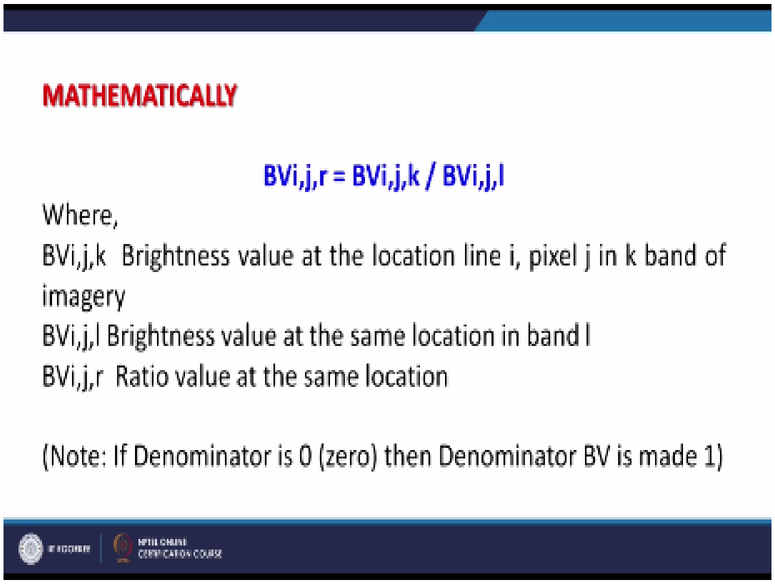
(Refer Slide Time: 04:00)



So the most common transforms applied to image data that is image division on pixel by pixel basis carried out the following operations for example band 1/ band 2 so a new band is generated and resultant data then rescale to fill the range of display because if you divide the number of a the range of pixel value will might reduce and therefore you want to rescale to the display range that is between 0 to 25 in most of the cases so this is very much required in most if these multivariate analysis this rescaling is the last stave which as to be done every time.

And this image division is applicable commonly also called a band ratio and there a very popular technique which we will see.

(Refer Slide Time: 04:52)

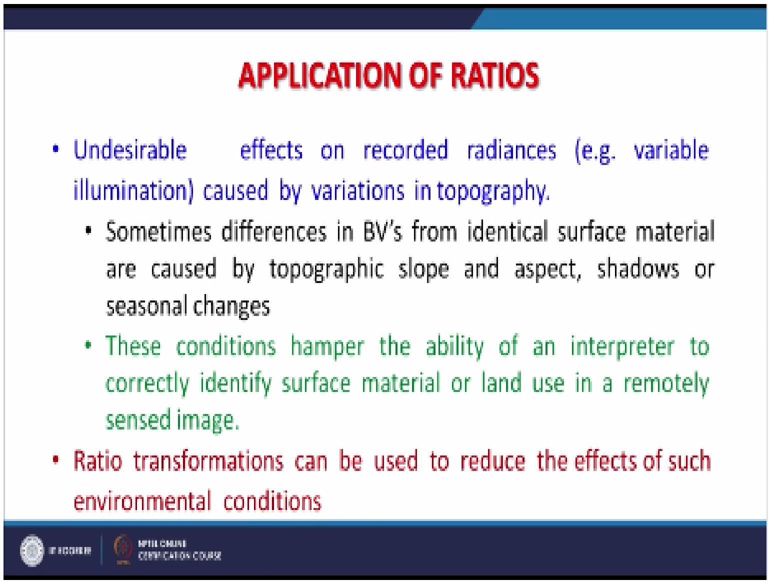


So mathematically basically if you look that a BVijr – BVijk/ BVijl and in this one BVijk that is the brightness value at the location line I pixel l j and k band of a measuring that means the band the band one in this case k = 1 we can consider so band 1 and each pixel basically this division will be made pixel by pixel so that is why location of pixels are given in all these two bands k and l and output band is the ratio image which is denoted here as r.

So this l is the brightness value of the same location in band l correspondent pixel will be divider this one as to remember and the ratio value at the same location that is the in case of the output image and in denominator in case of output image and in denominator is 0 if we end up the denominator BV is made 1 because there might be situation when this the band l that means may be band 2 when we create the ratio the value becomes finite so we reach scale 21.

And what are the application of a ratio measure 1 is to remove the schedule effects in this by the regent topography and to basically your ultimate aim in all these image processing techniques to improve the image interpretation so our image is becomes much better for interpretations this is what the undesirable effects on recorded radians that is the variable elimination caused by the variations the in the topography in day time if you are working in hilly region then one slope of the valley might be eliminated the other one may in the showdown and this kind of effects this differential solar eliminations can cause a difficulties by making image interoperations so if we imply when ratio techniques then these effects can be minimized that is the effects of topography due the differential solar eliminations.

(Refer Slide Time: 07:29)

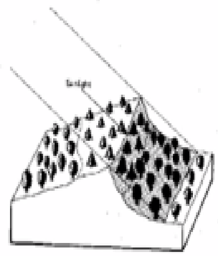


Sometimes differences in BV that is a band values from identical surface materials are caused topography slope and aspect, shadows or seasonal changes, so all these they are orientation of slope and reference to the position of the sun will pay major role here and these conditions hamper the availability of enterpriser as I have already mentioned to correctly identify surface material or land use in a remotes sensing.

And ultimately it will affect our interoperation in ratio transformations can be used to reduce the effects of such environmental conditions which are present in hilly region specially here is the example is given here.

(Refer Slide Time: 08:14)

Ratios for Elimination of Topographic Effect



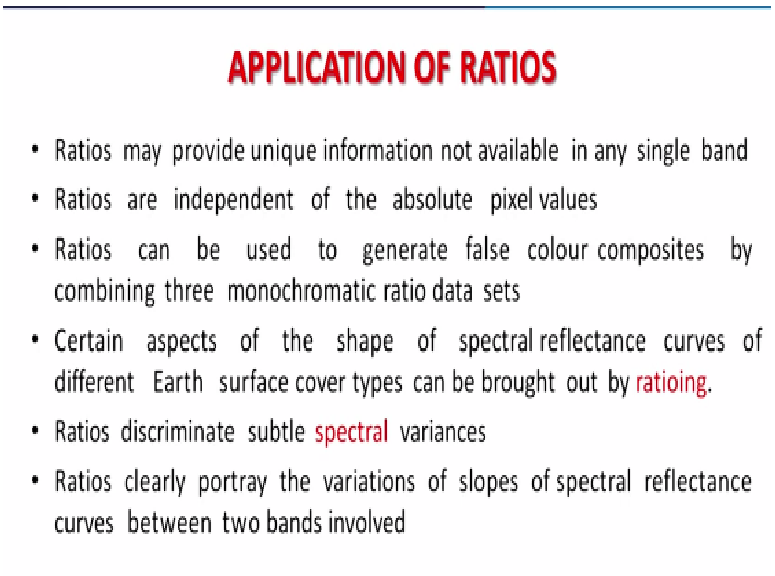
Landcover/ Illumination	Digital Number		Ratio
	Band A	Band B	
Decidous			
Sunlit	48	50	.96
Shadow	18	19	.95
Coniferous			
Sunlit	31	45	.69
Shadow	11	16	.69

That one side of the slope is a eliminated on this left side where as in the right side of the slope is in dark and therefore like in digital numbers of different features which are given here like for the slop which is band A might be having a pixel value of 48 in case of band B 50 but when they have create a ratio it becomes point 96 of course say the pixel value as to be in ratio image as to be integer value so it is rescale later on to look a two integer same like in case of shadow in band one band A or band one it is just 18 where as in band B it is 19.

So again ratio is a getting quite close as in case of sun light and shadow and therefore the affect of shadow is getting reduced another example this is the this was there form decides trees forest or vegetation the you are having examples of coni forest trees that here the in case of sunlit slope in case band 1 it is 13 pixel value in case of band in two or band B it is 45 so ratio comes 0.69 and in case of shadow again 11 and 16 ratio comes 0.69.

So when we will make the integers values then these values are coming very close here as in original bands these values where at very far and this is what the advantage we are removing the effects of shadow which is ahs been which has been caused due to the presence of rugged topography, so even if we are having same color type no problem and if we are having different covered type then a depended in the which features are there which lend cover are there that will reduce the effects of topography, radiance at shadow is only 50% of radiance of sunlight as in many cases we are seeing here generally and ratio nearly identical whereas the end result become identical and hence removal of topography effects. Now the applications of schedules may be that ratios.

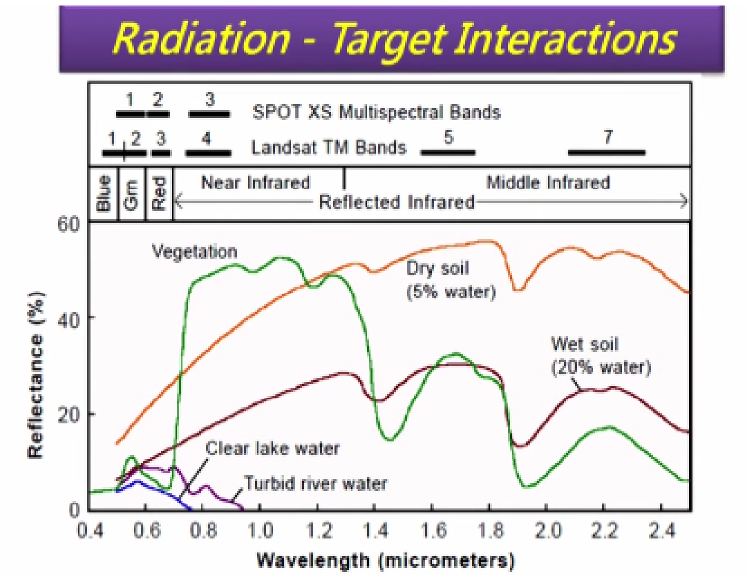
(Refer Slide Time: 10:44)



May provide unique information not available in any single band because two bands information are being used and ratio is simple mathematical technique and ratio discriminate certainly spectral variations and is contents aspect of shape and spectral reflectance curve of different earth surface covers can be brought out of by ratio, ratios are independent of absolute pixel value and ratios can be used to generate false color composites by combining three monochromatic ratio data sets.

That means if you are having many, many bands you can create ratios of two different bands having three output so three ratio output these three ratio outputs can be combined and can be created a color composites, so will say multi bands ratios so instead of involving this three bands in this ratio color composites you would be involving 6 bands and less effects of topography and therefore the image interpretation may become very reliable, so ratio clearly portray the variations of slopes of spectral reflectance curve between two bands involved.

(Refer Slide Time: 12:03)

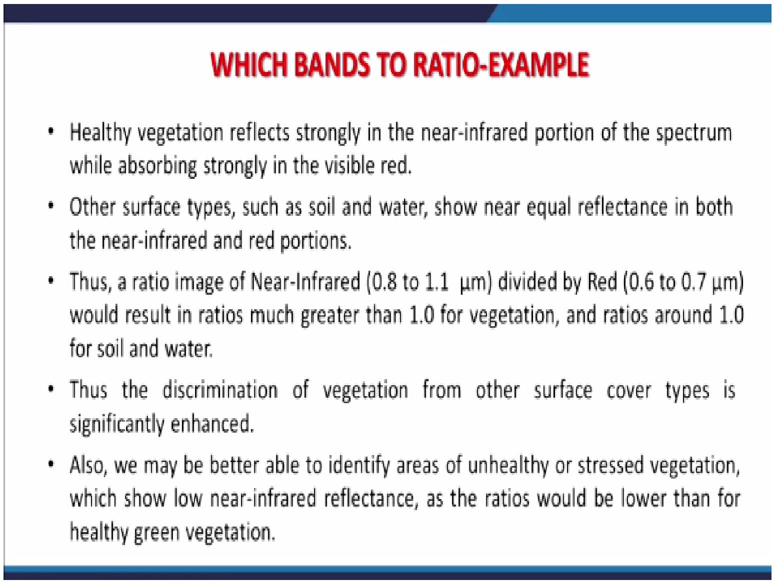


Here is the example here like in case of landsat TM you know that this they are covering first in the visible part then band food is in the near infrared part band 5 is the middle infrared far and then band 7 is you know in far basically far infrared, now band here band 6 is not mentioned because band 6 is thermal so there is not been considered in this particular slide, now different objects which are present on the earth.

Their spectral curves would be different in different bands this is what you are seeing, so if I create a ratio between say band4 and band 3 then I am not only removing the topographic affects but I am involving two bands and the same time because here the vegetation is making lesser change here. So in the spectral curve and therefore these bands this kind of band you are seeing will help us.

Similarly wherever the major changes are there in case of vegetation in case of dry soil or wet soil or water bodies turbid water bodies and so on, clear lake water bodies they all behave differently and a ratios can exploit that one and you can have a better output image for reliable interpretations image interpretations, for example which bands two ratio example this is what in the previous figure I was mentioning.

(Refer Slide Time: 13:52)



That the healthy vegetation reflects strongly in the near infrared because of high chlorophyll content and the spectrum of by absorbing strongly in the visible red. As you can see here that healthy vegetation will have a high reflectance in near infrared in case of landsat TM band 4 whereas in red band of landsat TM that is band 3 you do not have much reflection and this is what the healthy vegetation reflects strongly in near infrared portion of spectrum.

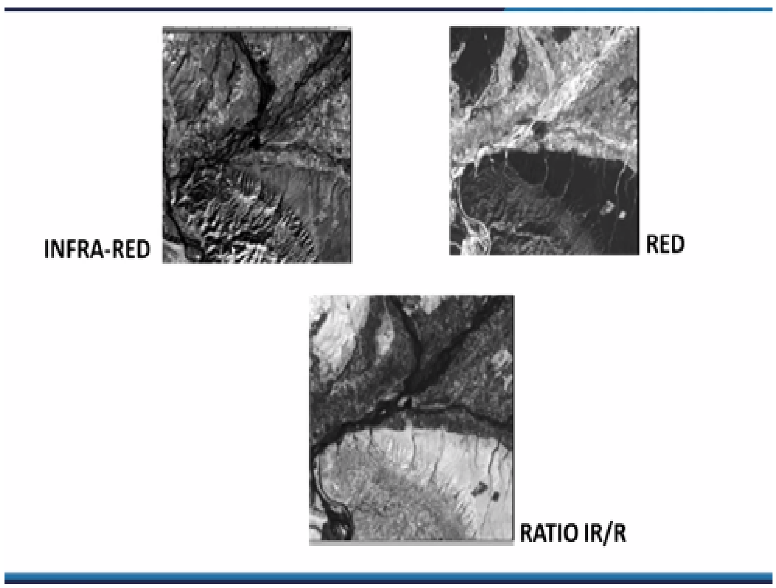
While absorbing strongly in the visible red and because of this sharp curve we can exploit so other surfaces types such as soil and water so near equal reflectance in both near infrared and red portions, it is not necessary that all objects or all kind of features which are present on the earth will behave differently, vegetation behaves differently water body, dry soil, wet soil with the all are having different spectral color.

So we need to understand the spectral curves and accordingly we should choose the band and then create the band ratio, thus the ratio image of near infrared divided by red would result in ratio much greater than one for vegetation and ratios around one for soil and water and this is that is why I was mentioning that where the curve has there is sharp change in the curve those two edges and bands will give you much better result.

This the dis-creation of vegetation from other surface curves type is significantly enhanced in such ratio outputs, and also we may give a better able to identify areas of unhealthy or stressed vegetation, vegetation suffering from lack of water or may be disease or something these by implying this ratio you can clearly identify the health of the vegetation we show nearly for a reflectance there is a shift then if a healthy vegetation less chlorophyll content.

It will the curve will shift towards the red and the ratio would be lower than the healthy green vegetation so implying such a band ratio technique even in case of vegetarian areas the health of vegetation can also be accessed DCT and this is then example of real example that this is infrared channel.

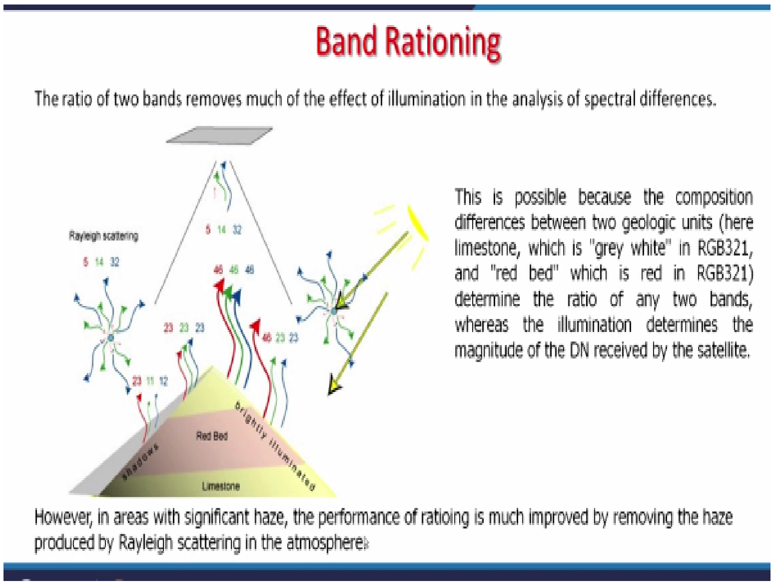
(Refer Slide Time: 16:23)



These are from IRS LISS 111 and this is the red channel and as you can see that this is the vegetative area which is coming dark in red channel whereas in inferred definitely it will have better illuminated part or a high brightness compared to red part but when we create a ratio then all the features which represent become much more clear second thing is, you will observe that the affect of topography especially in this part which is the Siva licks in this. Here they will that has been reduced and you see this image which is more or less quite representing reflector in.

But we know that it is not fair but for interpretation it might be better for certain application so that is big advantage that the discrimination power of the output image becomes much higher much reliable in compare to individual bands if I would have been using. So that is the advantage of band ratio, when the band ratio of two bands as have been mentioned removes much of the affect of illumination and analysis of a spectral.

(Refer Slide Time: 17:41)

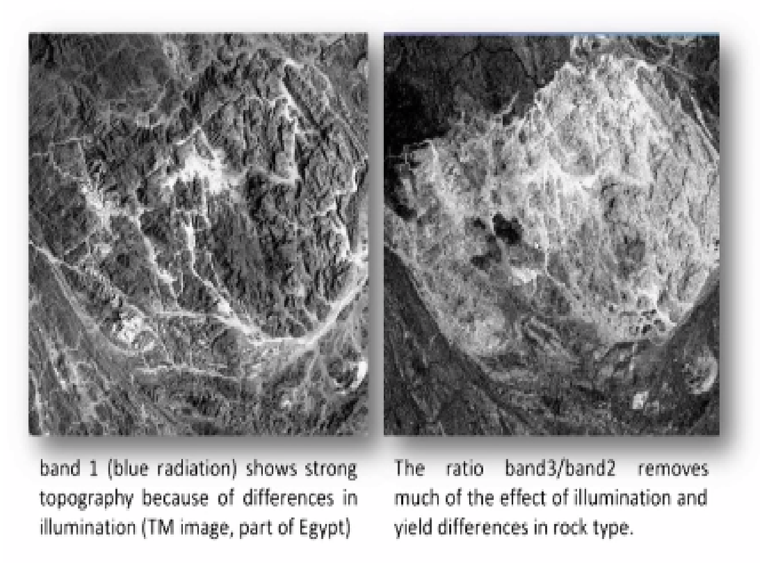


Differences and this is possible because the composition between two geologic units for example lime stone which is grey white and in red bed which is the red determine the ratio of nay two bands whereas the elimination determination magnitude of the DN received either satellites and here is the example this is the sun and energy is coming it is getting reflected so you are having this is the illuminated part.

So and this is the red band this is the lime stone and they are giving different reflections as you can see 46, 23, 23 and likewise and however urea which are having LISS or a atmospheric distortions or a having this because of relay scattering and this the numbers from these rocks or reflection from these rocks is reduced significantly roughly half so in the significant however in the areas we with significant days the performance of ratio is much improved by removing the age produced by Rayleigh scattering in the atmosphere.

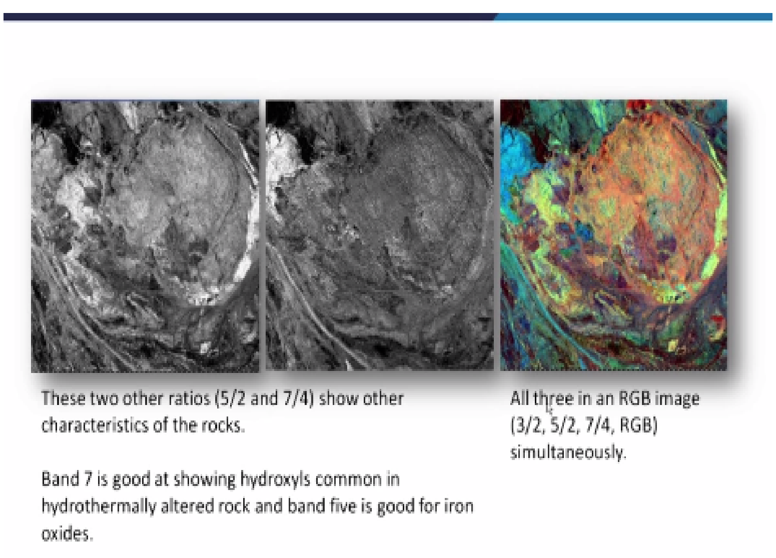
And therefore band ratio showing when we find that there is a hiss then we can imply again band rationing to improve our images, so the product image that ratio image we have the less affects of scattering. Because if that is causing and the hiss so release scattering affects can also be minimized and this is the example here and that this is the band one blue radiation source strong topography.

(Refer Slide Time: 19:24)



Effects whether ratio band 3/band 2 removes much of the affect of illumination and yield differences of rock and here you cannot identify different rocks very clearly and the effect of topography is very much dominating but here the effect of topography as reduced whereas the discrimination between two rocks or many rocks have increased, so that is the advantage of band rationing.

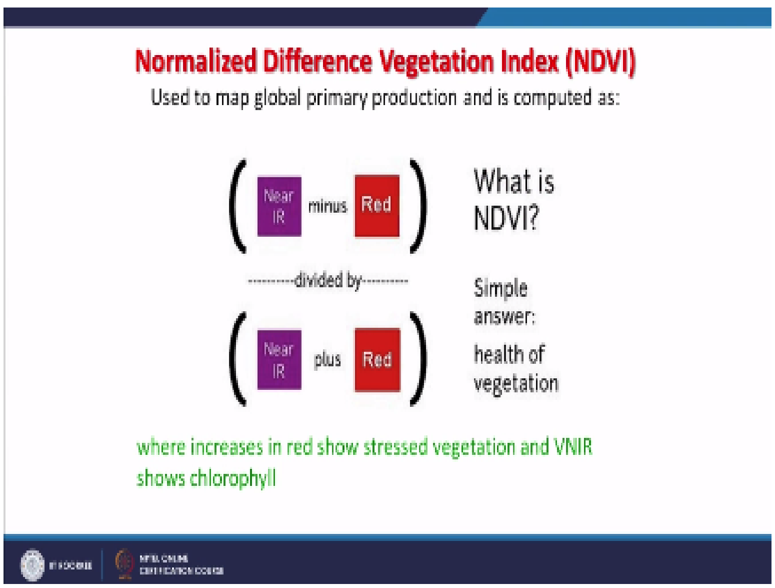
(Refer Slide Time: 19:51)



Another example about the color composites so these are the two band ratios 5/2 and 7/4 involving land sat TM data and here that all these three R and RGB that is the color composite of band radios so 3/2 which goes in red channel, 5/2 goes in green channel, 7/4 goes in blue channel and when we put we get a output image now you compare this output image either with 5/2 single or 7/4 which are evolver but you do not get that kind of discrimination power of different features which are present then in say color composite of band ratios.

So it is a very good tool in image processing to exploit the maximum variations which are present in different bands and can, can create a new product here band 2, band 3, band 5 and band 4 and 7, so 2,3,4,5 and 7. 5 bands characteristics where relations have been incorporated in this color composite in normally we can only incorporate about 3 in a standard falls composite. So that is why it is giving much better results.

(Refer Slide Time: 21:18)

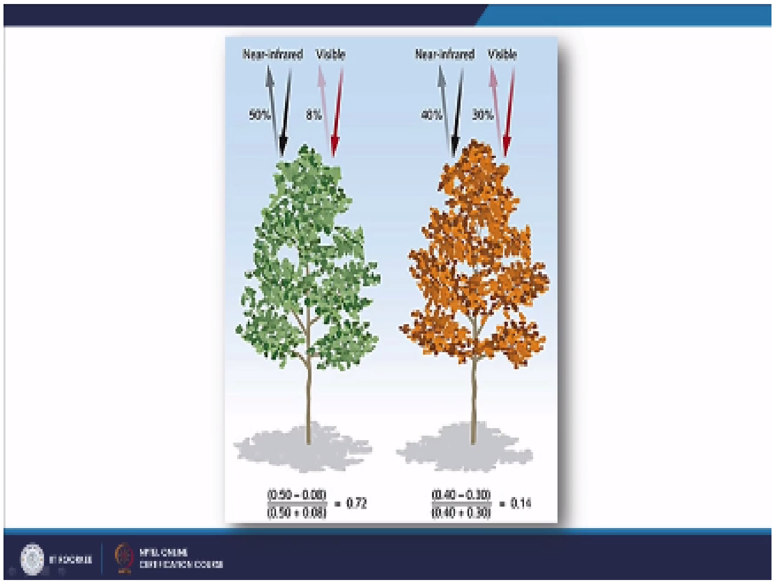


Now little advance band swing techniques and that the end of just by dividing but we making more different combinations or bringing little more arithmetic in it and that is very popular one is called normalized difference vegetation index or NDVI and it is for using course relatively course resolution data like NYVHR or modish data, at global scale NDVI maps are being use so we can assess that what is happening to the vegetation cover globally.

And in this one the near infrared band is subtracted by red band and this whole is divided by near infrared plus red band and that it will what NDVI will do it will bring highlight the health of vegetation, so when increases in red show stresses vegetation and VNIR that is very infrared band shows chlorophyll, and that is why these two bands are in more in this normalized difference vegetation index.

Here is the example like this is the healthy vegetation on right side you are seeing the vegetation which is not having much chlorophyll

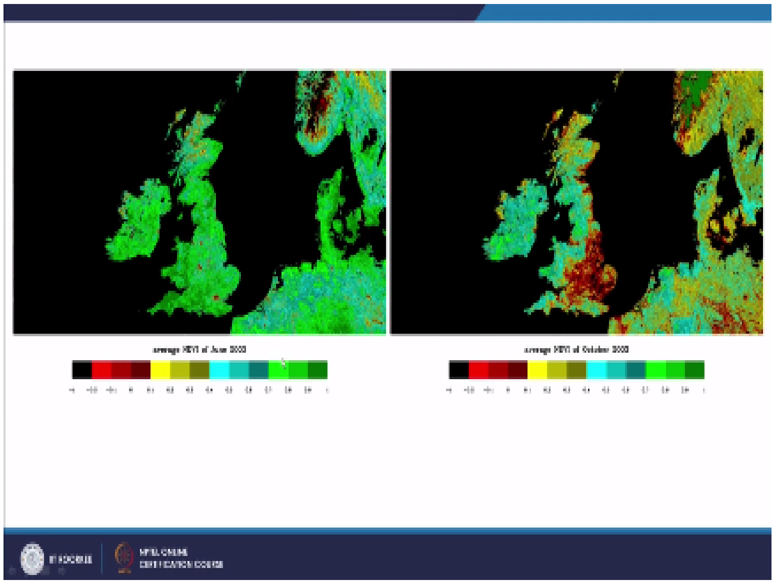
(Refer Slide Time: 22:42)



So here in infrared the reflection is coming about 50% and invisible just 8% but when we go for a dryly leaves or vegetation which is suffering from stress many lack of water or some disease then the near infrared deflection has reduced here by roughly 20% and whereas visible reflection has increased very significantly from 8 to 30. And if we see the ratios then see this is healthy vegetation ratios 0.72.

Whereas in case of not healthy vegetation under distress is just 14, so that is the advantage that you assess in a vegetated area, the vegetation which is suffering from stress and they will the destination between healthy and vegetation suffering digress will be enhanced as we can see here. And this, these are the examples of at counting Intel scale or country scale UK and Northern Island and some part of Europe as here.

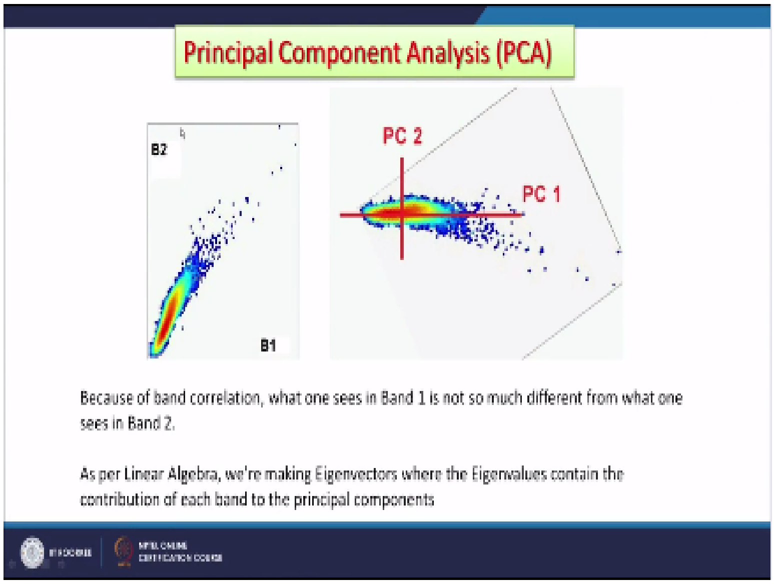
(Refer Slide Time: 23:51)



So this is the NDVI of June 2003 this is the NDVI of October 2003, you can see that how this NDVI has increase or decrease so in October because now on this is the one set of winter and you are having very less NDVI value, but in case of June which is you are having full of green view and NDVI is stretching even up to 1. So that is the you can assess that how vegetation cover at if we because generally we imply course resolution dates where country scale or continuality scale this vegetation conditions can be assessed.

Now further to this there is further analysis which is called PCA or principle component analysis, where we use the scatter plot two dimensional histogram we have already discussed when we discuss histogram and scatter plots so we will just directly go for this.

(Refer Slide Time: 24:57)

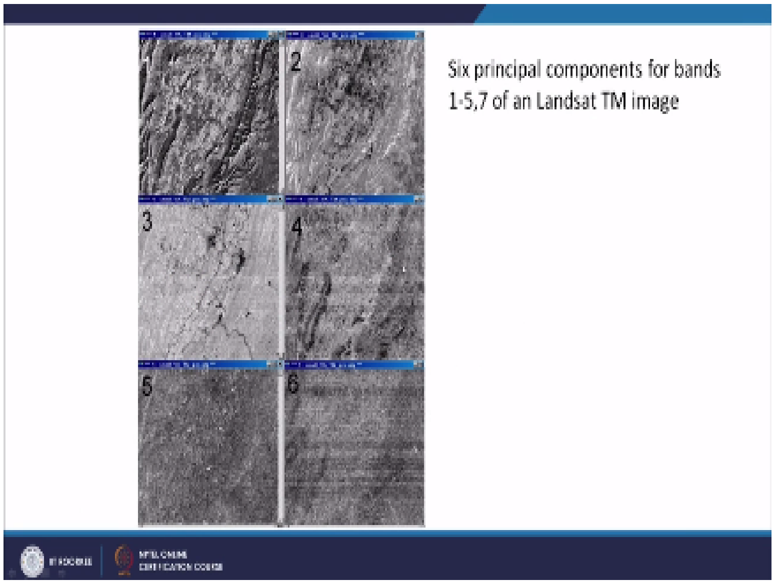


And I will explain in this graphics that we plot band 1 and band 2 like this, now we are co-related like this, this is showing that they are very much co-related so then this origin is shifted in the center and this rotation is there, so this becomes so the origin instead of here is shifted here and the principle component 1 this is aligned with the maximum variations which are present between these two bands.

And the next one is the PC2 which is perpendicular to PC1 and I think that if you are having multi band data or band spectral data so you can graphically I can represent mainly three components PC1, PC2, PC3 but in with using digital image processing software's we can go for many income not only three but may be 4,5,6 depending what you are trying to achieve. So here only I am showing 2 so because of band correlation which you are seeing here that one sees in band 1 is not so much different from what we see on the band 2.

And as per linear algebra we are making Eigen vectors where the Eigen values contain the contribution of each band to the principle component.

(Refer Slide Time: 26:31)

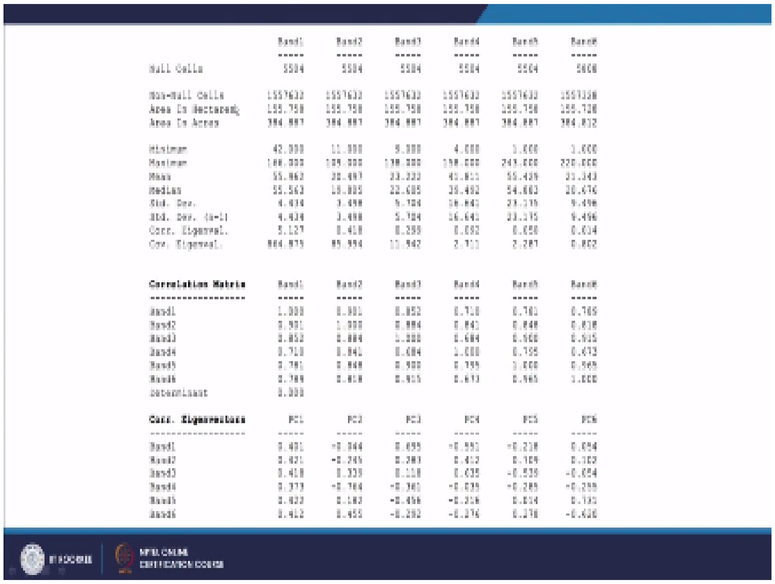


And this is what we see here that there are the six principle components of band 1, 1 to 5 and 7 of landsat TM 6 is missing because 6 is thermogenoal so that is involved in the analysis. And this is what you are seeing that principle component 1 will always have the maximum variations and the principle component which is last in this case is 6 which will have maximum noise and as you can see here.

So all noise which is having not much relation with any bands will go in principle component 6 in this particular example and maximum variations which are present in the image in these images of 6 channels, 6 bands they are all has been accumulated basically if I can use this word with principle component 1, so principle component, the component 1 image will have maximum variations the last one will have maximum noise nothing else.

And now one more step is also possible because now I know that this is their the other principle components are highly de-correlated they are not correlated so I can also make principle composers of principle components so I can use principle component 1,2,3 and assign say red, green, blue in this thing and I can create a color composite of principle component and such output image is not only involving 5 or 6 bands in this case in this example 6 bands.
But when we create a color composite they are creating much more better images for of course for reliable interpretations this is how the generally statics which you will get.

(Refer Slide Time: 28:31)



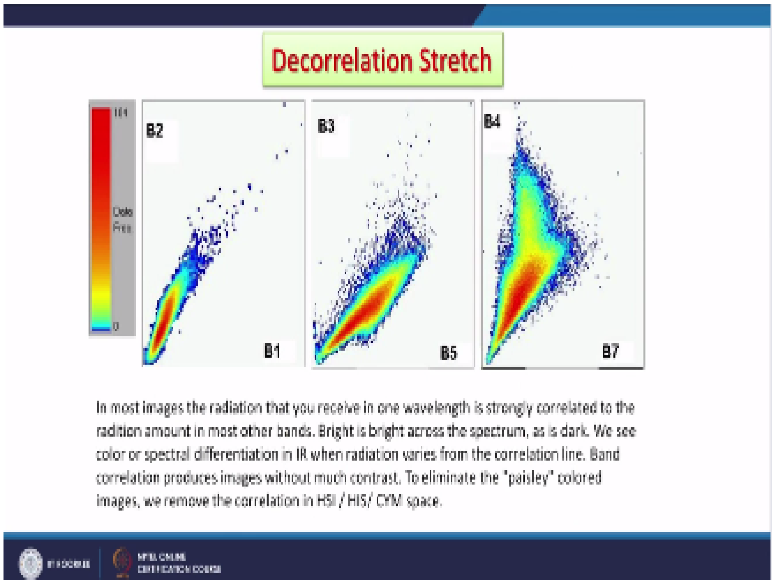
	Band1	Band2	Band3	Band4	Band5	Band6
Null Cells	5504	5504	5504	5504	5504	5608
Non-Null Cells	1557632	1557632	1557632	1557632	1557632	1557328
Area In Hectares	155.758	155.758	155.758	155.758	155.758	155.728
Area In Acres	384.887	384.887	384.887	384.887	384.887	384.812
Minimum	42.000	11.000	5.000	4.000	1.000	1.000
Maximum	188.000	109.000	138.000	158.000	243.000	220.000
Mean	55.962	20.497	23.222	41.811	55.439	21.343
Median	55.563	19.885	22.605	38.492	54.883	20.476
Std. Dev.	4.434	3.498	5.704	16.643	23.175	9.496
Std. Dev. (n-1)	4.434	3.498	5.704	16.643	23.175	9.496
Corr. Eigenval.	5.127	0.418	0.259	0.092	0.050	0.034
Cov. Eigenval.	884.875	85.994	11.942	2.711	2.287	0.802

Correlation Matrix	Band1	Band2	Band3	Band4	Band5	Band6
Band1	1.000	0.901	0.852	0.710	0.781	0.789
Band2	0.901	1.000	0.884	0.841	0.848	0.838
Band3	0.852	0.884	1.000	0.684	0.900	0.915
Band4	0.710	0.841	0.684	1.000	0.795	0.673
Band5	0.781	0.848	0.900	0.795	1.000	0.965
Band6	0.789	0.838	0.915	0.673	0.965	1.000
Determinant	0.000					

Corr. Eigenvectors	PC1	PC2	PC3	PC4	PC5	PC6
Band1	0.401	-0.344	0.695	-0.551	-0.218	0.054
Band2	0.421	-0.245	0.283	0.412	0.709	0.102
Band3	0.418	0.339	0.118	0.635	-0.539	-0.054
Band4	0.373	-0.784	-0.361	-0.035	-0.285	-0.259
Band5	0.423	0.182	-0.456	-0.316	0.034	0.731
Band6	0.412	0.455	-0.292	-0.276	0.278	-0.626

That band 1 here, band2, band3 and band basically here in this shown as band 6 band 7 basically and this is correlation matrix and this is correlation Eigen vectors are there. So band1 definitely it is 100% correlated so 1 is there and similarly you can see that how they are correlated band1 to band2 is 0.9, so it is a very high correlation and therefore principle component analysis is would be much useful in such cases.

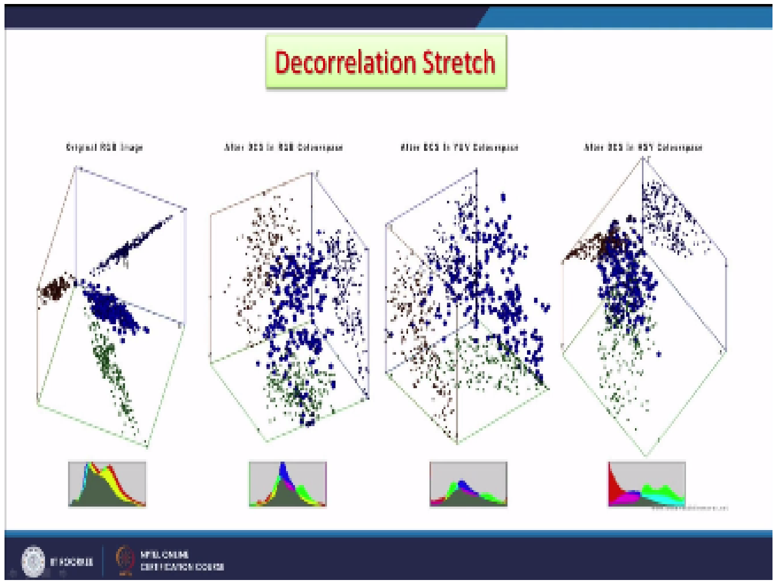
(Refer Slide Time: 29:08)



Now the next step in principle component analysis is de-correlation stretch which is again based on the principle component analysis like example is shown here that band1 and band2 scattered plot is showing highly correlated data and what we can do we can shift the origin as in case of principle component analysis and then a stretch alone these axis to occupy them full range which is available for display.

And why we are de-correlating each component within each component and at the end if we see this is scatter plot we are occupying a very large area which was not possible with normal thing, so this band 1 band2 is correlated highly correlated is there whereas band4, band7 are less correlated, so in this is the situation.

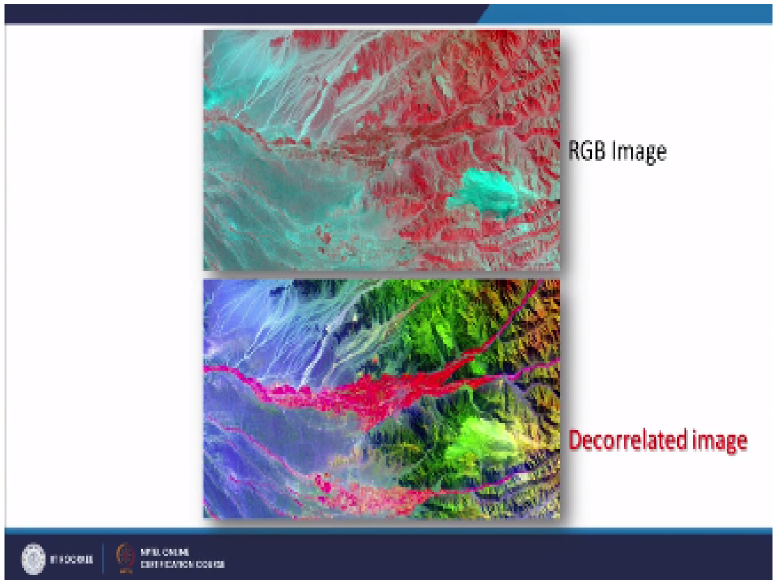
(Refer Slide Time: 30:12)



And if we plot and see this you know cube here this is the original RGB this is how the pixel values are distributed. When we go for de-correlations stretch DCS in RGB color space they are occupying very large area or space within that color cube which we have been discussing also and this is another color combination where if I go then I am, we are occupying much more area which is available within that color cube and HSY that the hue saturation and yellow color space then we are occupying little less.

So this de-correlation stretch and in RGB combinations many times will give you much better results with full of colored images and example here is that this is simple RGB image but when the principle component.

(Refer Slide Time: 31:20)



And then we did the de-correlation of different components the principle component PCs and then made a color composite this is what the result is and see that the quality of this combination. Here I can discriminate different rocks, different ground features comparatively very easily then on simple RGB image this is the advantage of going for multivariate analysis in digital image processing, and this brings to the end of this lecture thank you very much.

For Further Details Contact
Coordinator, Educational Technology Cell
Indian Institute of Technology Roorkee
Roorkee – 247 667
E Mail: etcell,iitrke@gmail, etcell@iitr, ernet,in
Website: www.iitr.ac.in/ETC, www.nptel.ac.in

Web Operations
Dr.Nibedita Bisoyi
Neetesh Kumar
Jitender Kumar
Vivek Kumar
Production Team
Sarath. K
Pankaj Saini
Arun. S

Camera
Mohan Raj
Online Editing
Jithin. K
Graphics
Binoy. V. P

NPTEL Coordinator
Prof. B.K. Gandhi

An Educational Technology Cell
IIT Roorkee Production