



yuvakshētra[®]

Institute of Management Studies (YIMS)
Ezhakkad, Mundur, Palakkad - 678631, Kerala.

ACCREDITED BY NAAC WITH B+ GRADE (1st CYCLE)

Affiliated to the University of Calicut & Managed by the Diocese of Palghat

DEPARTMENT OF COMPUTER SCIENCE

ICACRI

**INTERNATIONAL CONFERENCE FOR ADVANCED
COMPUTATIONAL RESEARCH AND INNOVATIONS**



2024

VOLUME I , ISSUE I

CONFERENCE PROCEEDINGS

English Language
Title of the Book : International Conference for Advanced Computational Research and Innovations (ICACRI - 2024)
Editor : JIBIN JOY
Published by : Yuvakshetra Institute of Management Studies
Address : Ezhakkad, Mundur, Palakkad, 678600
Rights Reserved
First Edition : FEBRUARY 2024
Cover Design : JIBIN JOY
Printed at : Jim Offset, Palakkad
Publishers : Yuvakshetra Publications, Ezhakkad, P.O, Mundur, Palakkad
E mail : yimspublication@yuvakshetra.org,
: yuvakshetra@gmail.com
Website : www.yuvakshretra.org
Tel : 9400012368, 8714345789
Distributors : Yuvakshetra Publications, Ezhakkad, P.O, Mundur, Palakkad
E mail : yimspublication@yuvakshetra.org,
: yuvakshetra@gmail.com

No part of this publication may be reproduced or transmitted in any form or by any means without prior written permission of the author.

ISBN : 978-81-968246-5-5.

UNSUPERVISED LEARNING FOR HYPERSPECTRAL IMAGE CLASSIFICATION

Dr.A.Gokilavani

Assistant Professor & Head

Department of Computer Science

KG College of Arts & Science,

Coimbatore – 641035

Email: a.gokilavani@kgcas.com

Contact no.: 9789498209

ABSTRACT

Out of all the data sources accessible to geographic information systems (GIS), remote sensing is one of the most crucial ones. Remote sensing is the process of gathering data about the surface of the Earth without physically being there. It detects radiations that are emitted and reflected and are normally captured by sensors that are installed on an aircraft or a satellite. Modeling and monitoring activities on the Earth's surface as well as identifying elements in the land cover by analysing spectral characteristics collected by sensors are the two main goals of remote sensing. When a specific sensor device collects and processes information from the electromagnetic spectrum, it is known as Hyperspectral Imaging (HSI). The data it generates is a goldmine of information. Use this data to solve a variety of problems in a variety of applications. A digital image's pixels are divided up into groups using hyperspectral imaging classification. These techniques were used to classify hyperspectral images using unsupervised hyperspectral image classification algorithms. K-Means and ISODATA algorithms are employed. ENVI is used to apply two algorithms to a hyperspectral image of Washington DC, USA. The accuracy of the procedure was assessed using Principle Component Analysis (PCA) and K-Means or ISODATA algorithm in this paper. The ISO-DATA algorithm outperforms the K-Means algorithm in terms of precision. Since The K-Means algorithm has a classification accuracy of 78.3398 percent, whereas the ISODATA algorithm has a classification accuracy of 81.7696 percent. When the number of classification iterations increased, so did the processing time.

Keywords: Unsupervised classification; K-Means algorithm; ISODATA algorithm; ENVI.

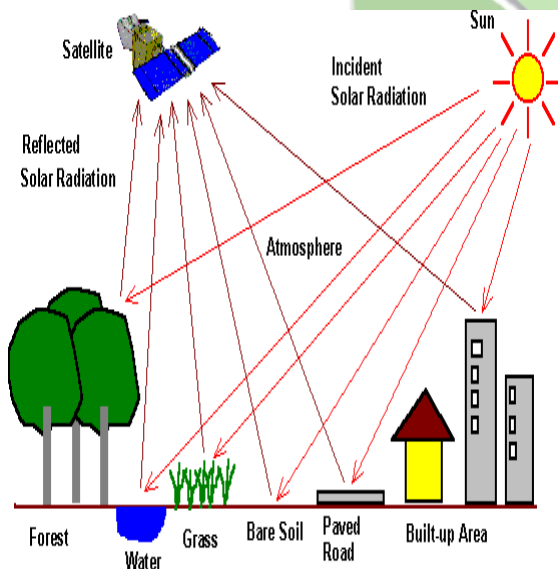
I. INTRODUCTION

Hundreds of continuous spectral bands are often present in hyperspectral remote sensing - based pictures, allowing for the exact discrimination of the many spectrally comparable land cover classifications. The "curse of dimensionality" (also known as the "Hughes effect") is caused by the fact that

such high-dimensional data also contains strongly correlated and useless band information. The classification accuracy of the hyperspectral pictures is significantly decreased by the presence of unimportant and highly correlated spectral bands. Therefore, dimensionality reduction is a more difficult method to improve the hyperspectral image's accuracy of classification.

The art and science of gathering data from a distance is known as remote sensing. For the purpose of interpreting and managing the Earth's resources and surrounding environment it is seen as the measurement and analysis of electromagnetic radiation that is transmitted through, reflected from, or absorbed and dissipated by the atmosphere, the hydrosphere, and materials on or near the land surface. It is possible to take images of the Earth's surface using optical remote sensing, which uses visible, near infrared and short-wave infrared sensors to observe the solar radiation reflected from targets on the background. At different wavelengths, different materials reflect and absorb differently. Spectral reflectance fingerprints in remote sensing photos can so distinguish the targets [1][2][3].

Airborne Imaging Spectro-radiometer for Applications (AISA) and other hyperspectral sensors made it possible to create a continuous reflectance spectrum for each pixel in the picture. It is possible to use these



methods to make distinctions between various types of earth surface features. While humans can only perceive

visible light in three bands (red, green, and blue), hyperspectral imaging divides the spectrum into more than a dozen different bands (infrared, RGB, and UV). This analogy is useful in understanding the concept of hyperspectral imaging better (see Fig. 2) [5][6].

If you look at the definition of hyperspectral, you'll see that it refers to the fact that there are a lot of wavelength bands. Spectral specificity means hyperspectral imaging gives extensive spectrum information to help distinguish and identify materials that are unique in spectral make-up. When compared to other remote sensing data types, hyperspectral imaging offers the potential for more precise and comprehensive information extraction [7]. As a result of its improved capabilities, it is better able to detect potentially hazardous materials and to provide the additional data required to identify and classify these substances. Material's spectral properties are shown by the HSI pixels, which form spectral vectors [10].

Several drawbacks to hyperspectral imagery include a lack of precision in

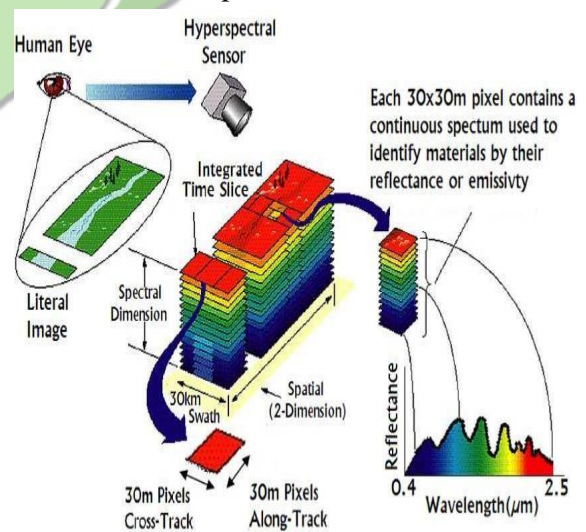


Fig 2. Hyperspectral imaging

attributes such as directions and distances, due to the earth's sphere. For example, shadows may obscure a particular area to be studied, or the brilliance of the light may be exaggerated in a particular region [11]. An important constraint of hyperspectral imaging is the pixel size, which might be large enough to include many properties but be difficult to categorise or small enough to have no features that can be classified [12]. [13, 14].

II. CLASSIFICATION

The primary goal of satellite imagery categorization is to accurately assess terrain features and extract relevant information [13]. It's important to note that unsupervised and supervised algorithms are the two most used categorization methods. Fig. 3 shows an example of unsupervised classification. K-Means method and ISODATA are employed in an unsupervised classification chain. Unsupervised Classification is depicted in the following figure.

A. K-Means Classification

Using the K-means method is a simple way to get the average of a collection of K-sets. The K-Means algorithm's goal is to minimise cluster variation [15][16][17]. The pseudo-code for the K-means method can be found here [18].

B. Algorithm for Iterative Self-Organizing Data Analysis (ISODATA)

As one of the most commonly employed algorithms for unsupervised classification, the ISODATA algorithm (see Fig. 5). [19][20][21]: The ISODATA clustering steps are as follows: This is a screenshot of the ISODATA Classifier.

III. ASSESSMENT OF THE RESULTS AND CONCLUSION

ENVI was used to do an unsupervised classification on a hyperspectral image. The image of Washington, DC that was used in the application of hyperspectral imaging is the hyperspectral dataset that was applied to the image. Principle Component Analysis (PCA) and K-Means or ISODATA methods are applied to the hyperspectral image. A categorised image is the outcome of applying the K-Means algorithm and ISODATA algorithm. When the number of iterations to get the categorised image grew, the process duration rose. Since each pixel in the image has been classified into a category other than "Unclassified Class," statistical information derived from the image data and verified using the K-Means and ISODATA algorithms is reliable. Because ISODATA algorithm has an overall classification accuracy of 81.7696 percent whereas K-Means algorithm has an overall classification accuracy of 78.3398 percent, ISODATA algorithm is superior in terms of accuracy.

An introduction to ENVI (Environment of Visualizing Images)

Images are processed using ENVI. It was made to process data collected by remote sensing. It is capable of seeing and analysing large amounts of image data in great detail. In terms of scientific data formats, it can handle a wide range [22][23].

B. Cases and Research

- Comparison of the hyperspectral image's outcomes after applying various RGB bands. Classification accuracy can be improved by increasing or decreasing the number of classification iterations.

- A comparison of the first time using K-Means and the second time using ISODATA. A comparison is made between the first time the PCA algorithm was used and the second time the K-Means algorithm was used, with results from both methods being used.

PCA and K-means were applied to a hyperspectral image (Washington, DC) using distinct RGB bands (see Table I) for each band. Fig. 3 displays the PCA results obtained using the image and the test values from Table I. The total number of courses

Experiment Description	Overall Accuracy	Class Percentage						
		Unclassified	Roof	Grass	Land	Trail	Road	Water
Comparison between the result of a classified image with only one iteration (Ground Truth) and a classified image with three iterations.	62.7079%	0.0	100	100	100	100	83.29	21.93
Comparison between the result of a classified image with three iterations (Ground Truth) and a classified image with ten iterations.	87.6110%	0.0	100	100	100	100	90.34	8.87
Comparison between the result of a classified image with only one iteration (Ground Truth) and a classified image with ten iterations.	57.8171%	0.0	100	100	100	100	89.59	27.06
Comparison between the result of a classified image with three iterations (Ground Truth) and a classified image with fifteen iterations.	87.6110%	0.0	100	100	100	100	90.34	8.87

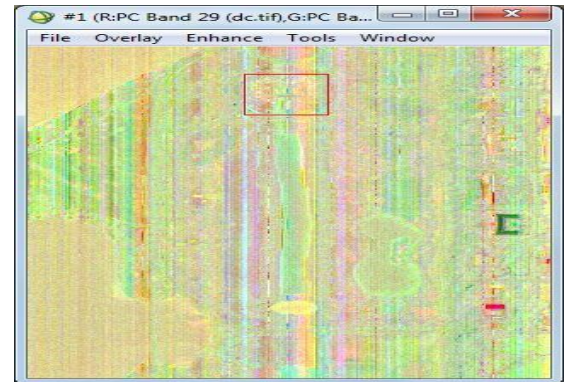


FIG 3. K-MEANS CLUSTERING



Figure 4 depicts the results of applying K-Means method on the output pictures of PCA, where the number of iterations is 3 and the number of clusters is 6. A hyperspectral image's classification accuracy rises as the number of iterations is raised; conversely, it falls as the number of iterations is decreased (Case Study 2). Based on the results of the tests described above, we can say that this is correct. Table II summarises the experiment's findings.

A hyperspectral image of Washington, DC was subjected to the K-means and Iterative Self-Organizing Data Analysis Technique Algorithm algorithms in Case Study 3. (ISODATA). There are two ways to do this: first using PCA, and then using either K-

Means or ISODATA as a second stage in the PCA process. PC Band 172 for R, PC Band 86 for G, and PC Band 24 for B have been selected. K-Means parameters include the number of classes (six) and the number of iterations (three). K-Means method results are shown in Fig. 8-a. The ISODATA settings specify a class size of 4–6 and a maximum iteration count of three. Using ISODATA yields the results shown in Fig. 8-b.

A 4) Case Study 4: It's all about comparing the results of implementing PCA and K-Means, as demonstrated in Fig. 9-a and Fig. 9-b, in terms of classification accuracy. The hyperspectral image is used in this investigation (Washington DC).

Overall Accuracy = (307844/392960) 78.3398%								
Kappa Coefficient = 0.7373								
Ground Truth (Percent)								
Class	Unclassified	Roof	Grass	Tree	Trail	Road	Water	Total
Unclassified	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Roof	0.00	100.00	1.59	0.00	14.92	0.00	0.00	13.15
Grass	0.00	0.00	93.34	17.02	0.15	0.00	0.00	24.96
Tree	0.00	0.00	5.07	46.06	32.74	0.00	0.00	14.39
Trail	0.00	0.00	0.00	33.14	51.08	6.26	0.00	14.94
Road	0.00	0.00	0.00	3.77	1.11	84.79	0.00	17.73
Water	0.00	0.00	0.00	0.00	0.00	8.95	100.00	14.82
Total	0.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00

The statistics of the C class

1) Using the K-means technique to compute class statistics: For the Washington DC picture results, the Means and Standard Deviation for all classes are shown in Fig. 10-a and Fig. 10-b, respectively, which illustrate a correlation between band number and

value. This graph illustrates the minimum, maximum, and average values for each band of the Tree class. The standard deviation for the Tree class is shown in Figure 10-d. Table III provides a description of the class distribution, and Table IV presents the ground truth image-based confusion matrices. Figure 11 shows a total class mistake.

2) Applying the ISODATA technique to calculate class statistics: Analyze Washington DC hyperspectral image data using the ISODATA technique, which is shown in Fig. 12-a. Means for all classes are shown, as well as their standard deviations (Fig. 12-b). Fig. 12-c displays the minimum, maximum, and mean values for each band in the Tree class, as well as the standard deviation. The standard deviation for the Tree class is shown in Fig. 12-d. Table VI displays the degree of misunderstanding among the various classes, as seen in Table V. Matrixes based on the original image's data. A complete error of classification is shown in Figure 13.

TABLE II. Class Distribution Summary (K-Means Algorithm)

Unclassified	0 points (0.000%)
Roof	51,689 points (13.154%)
Grass	98,094 points (24.963%)
Tree	56,529 points (14.385%)
Trail	58,722 points (14.944%)
Road	69,678 points (17.732%)
Water	58,248 points (14.823%)

TABLE III. Accuracy Class Percentage of Experiment Description

IV. CONCLUSION

Hyperspectral images contain a wide range of spectral information that can be used to detect and separate materials with a distinct spectral fingerprint. In order to classify a hyperspectral image, it is necessary to

identify items with similar features. After doing Principle Component Analysis (PCA) using ENVI, unsupervised classification algorithms (K-Means algorithm and ISODATA algorithm) are utilised.

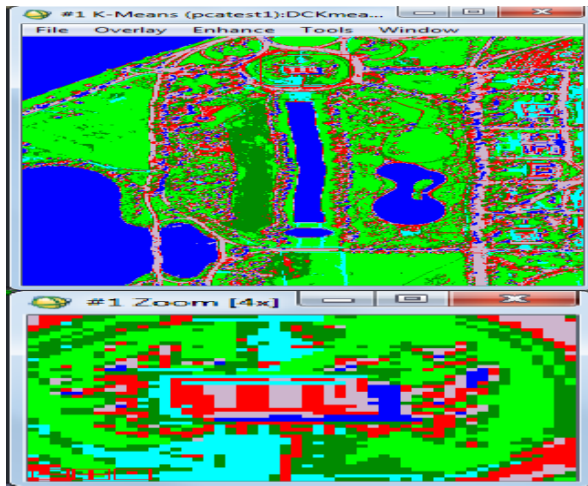


Fig 4. Result for K-means

Overall Accuracy = (321322/392960) 81.7696%								
Kappa Coefficient = 0.7726								
Ground Truth (Percent)								
Class	Unclassified	Trail	Grass	Tree	Road	Water	Roof	Total
Unclassified	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Trail	0.00	36.54	0.10	0.00	0.00	0.00	2.32	1.70
Grass	0.00	0.00	99.81	21.07	0.00	0.00	10.95	30.40
Tree	0.00	0.00	0.09	60.90	14.40	0.00	8.88	16.02
Road	0.00	0.00	0.00	18.03	78.73	4.96	0.00	19.40
Water	0.00	0.00	0.00	0.00	1.53	95.04	0.00	17.85
Roof	0.00	63.46	0.00	0.00	5.34	0.00	77.85	14.62
Total	0.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00

TABLE IV. Confusion Matrix Using Isodata Algorithm

PCA is used in data analysis before classification to minimise the dimensionality of hyperspectral images. A typical site from the research area in the United States capital of Washington DC is used to test these methods. The K-Means classification strategy had an overall accuracy of 78.3398 percent,

whereas the ISODATA classification approach had an accuracy of 81.7696 percent. Both the K-Means and ISODATA algorithms produce reliable findings, however ISODATA performs better on images of the test region.

REFERENCES

- [1] Sachin T. Jadhav and Mandar M. Kolap, "A Review on Methodologies of Remote sensing, GPS And GIS", IOSR Journal of Electronics & Communication Engineering (IOSR-JECE), ISSN : 2278-2834, ISBN : 2278-8735, PP. 1-6.
- [2] Foudan Salem, Tarek El-Ghazawi, and Menas Kafatos, "Remote Sensing And Image Analysis For Oil Spill Mitigation In The Red Sea" January 2001.
- [3] Foudan Salem and Menas Kafatos, "Hyperspectral Image Analysis For Oil Spill Mitigation", 22nd Asian Conference on Remote Sensing, Singapore, November 2001.
- [4] "Hyperspectral imaging", Fleurus (Municipality, Province of Hainaut, Belgium): CRW Flags Inc., 2010.
- [5] Principles Of Remote Sensing, ITC Educational Textbook series, ©ITC 2001. [Online]. Available: <http://www.gdmc.nl/oosterom/PoRSHyperlinked.pdf>
- [6] "HyperSpectral Image: How it happens". <http://www2.brgm.fr/mineo/RapportFinal/Image16.gif>
- [7] P. Shipper, "Introduction to Hyperspectral Image Analysis," An International Electronic Journal, 2003. [Online]. Available: <http://spacejournal.ohio.edu/pdf/shippert.pdf>. [Accessed 10 october 2013].
- [8] G. Shaw and H. Burke, "Spectral

- imaging for remote sensing," Lincoln Laboratory J, p. 3–28, 2003.
- [9] J. B. Campbell, "Introduction to Remote Sensing" (3rd Edition), London: Taylor and Francis, 2002.
- [10] S. Peg, "Introduction to Hyperspectral Image Analysis", Earth Science Applications Specialist Research Systems, Inc.
- [11] SpecTIR, "Hyperspectral Image," SpecTIR, 2012. [Online]. Available: <http://www.spectir.com/technology/hyperspectral-imaging/>. [Accessed 10 Nov. 2013].
- [12] S. Dr Ropert, "Introduction To Remote Sensing," New Mexico State University.
- [13] [Online]. Available: <http://disi.unitn.it/seminars/281>. [Accessed 22 September 2013].
- [14] Tan, "The K-Means algorithm," December 2006. [Online]. Available: <http://www.cs.uvm.edu/~xwu/kdd/Slides/Kmeans-ICDM06.pdf>. [Accessed 2013].
- [15] A. Villa, J. Chanussot, J.A. Benediktsson, C. Jutten c, R.Dambreville, "Unsupervised methods for the classification of hyperspectral images with low spatial resolution", Journal of Pattern Recognition, Volume 46 Issue 6, pp. 1556-1568, June 2013.
- [16] S.Praveena S.P.Singh I.V.Murali Krishna, "An Approach for the Segmentation of Satellite Images using K-means", KFCM, Moving KFCM and Naive Bayes Classifier. International Journal of Computer Applications, Volume 65, No.20, 2013.
- [17] M.Priyadharshini,R.Karthi,S.N.Sangeethaa, Premalatha, k.S.Tamilselvan, "Implementation of Fuzzy Logic for the High- Resolution Remote Sensing Images with Improved Accuracy", IOSR Journal of Electrical and Electronics Engineering (IOSR-JEEE) e-ISSN: 2278-1676,p-ISSN: 2320-3331, Volume 5, Issue 3, 2013, PP 13-17.
- [18] "K-Means Clustering," Home.deib.polimi.it, [Online]. Available: http://home.deib.polimi.it/matteucc/Clustering/tutorial_html/kmeans.htm.
- [19] "ERDAS Field Guide", 2005 Leica. Geosystems Geospatial Imaging, LLC, 2005.
- [20] Mather, Paul, "Computer Processing of Remotely-Sensed Images", 2004 John Wiley & Sons, Ltd
- [21] Memarsadeghi, N., Netanyahu, N.S., LeMoigne, J., "A Fast Implementation of the ISODATA Clustering Algorithm", International Journal of Computational Geometry and Applications, IJCGA, Vol. 17, No. 1, PP. 71-103, 2007.
- [22] "Multispectral Processing using ENVI's Hyperspectral Tools," 10 December 1997. [Online]. Available: <http://www.ltid.inpe.br/tutorial/tut14.htm>. [Accessed 6 November 2013].
- [23] ENVI, "ENVI User's Guide", Research Systems, Inc., 2004.