International Journal of Research
Available at https://edupediapublications.org/journals
e-ISSN: 2348-6848
p-ISSN: 2348-795X
Volume 04 Issue 17
December 2017

A Novel Decision Tree Algorithm for Land Cover Classification Using Hybrid Polarimetric Sar Data

K. V. Ramana Rao
Department of ECE, VIEW, Visakhapatnam, Andhra Pradesh, India – 530046
ramanadeepu@yahoo.co.in

Prof.P.Rajesh Kumar
Department of ECE,A.U.C.E, Visakhapatnam, Andhra Pradesh, India– 530003
rajeshauce@gmail.com

Abstract

The polarimetric information contained in polarimetric synthetic aperture radar (SAR) images represents great potential for characterization of natural and urban surfaces. However, it is still challenging to identify different land cover classes with polarimetric data. Hybrid polarimetric SAR data (RH, RV) from RISAT - 1 is found to be suitable for land cover classification of significant features that are well distinguished. The availability of high resolution hybrid polarimetric data from RISAT - 1 SAR systems supporting FRS -1 mode, made it possible to analyze the scattering mechanism for different land use and land cover features using the Raney decomposition (m-alpha, m-chi, and m-delta) techniques. Further to perform both supervised classification(parallelepiped, minimum distance, maximum likelihood and isodata classifiers) and machine learning (artificial neural net) classification also performed Decision tree classification.The proposed statistical Gumbel distribution model has been implemented and retrieves the threshold intensity values. In this proposedwork classification approach has been evaluated for RISAT-1 SAR hybrid polarimetric data of 21st October 2014 over an urban city, Visakhapatnam, in the state of Andhra Pradesh, India. Since the hybrid polarimetric radar data contains all the scattering information for any arbitrary polarization state, data of any combination of transmitting and receive polarizations can be synthesized, mathematically from hybrid polarimetric data. The RISAT-1 SAR hybrid polarimetric data were decomposed to retrieve the surface and volume scattering information. Both supervised classification and machine learning classification methods were appliedto land cover and few other land use classes based on ground truth measurements using maximum-likelihood (ML) distance measures that are derived from the complex distribution of SAR data at various polarization combinations. The results show that Decision tree classification accuracies for m-alpha, m-chi and m-delta methods were 99.743, 96.873 and 99.857 respectively. RISAT-1 hybrid polarimetric SAR data helps to classify land cover features efficiently.

Keywords and phrases: SAR, Hybrid Polarimetry, Image Processing, Classification, Decision Tree algorithm

1. Introduction

Nowadays land use and land cover classification of a particular part of many countries, especially city areas have got much importance for proper planning against the continuous alteration oftheearth surface. This is mainly due to that the land cover change is a result of many factors such as theglobal change that may occur with changingclimatic conditions, changes intheecosystem, geochemical cycles, biodiversity and various human activities. Different target decomposition techniques and as well as various classification algorithms have been proposed by Pottier, Lee et al., Cameron and Leung, Ferro-Famil et al. Ouarzeddine and Souissi, Fang et al., Park et al., Parks et al.Classification techniques are broadly divided into two types known as supervised classification and unsupervised classification techniques. Although a lotof research has been done in the field of SAR image classification, still there are certain limitations in each classificationtechnique due to the problem of discriminating of accurate features.

For example, besides being widely applicable, the major disadvantage ofthesupervised classification technique is that it is a single discriminative classifier which is applied to the individual pixel level or image objects (agroup of adjacent, similar pixels). If during training process any pixel is unidentified then supervised classifier cannot assign it to any class.

Also, supervised classifier is unable to recognize and represent unique categories which are not represented in training data. Similarly, unsupervised classification methods also suffer from certain limitations and disadvantages. In order to achieve more accurate results for land cover classification, it is advantageous to opt for more advanced classifier. Inrecentyearstheuseofdecision treeclassifierforlandcover classification of remotely sensed data has been increased considerably. Previous researchers show that decision tree algorithms consistently outperform supervised classification techniques. Decision tree classification is acomputationallyefficient algorithm. The other advantage of decision tree classifier includes its flexibility, simplicity, ability tohandle noisy and missing data, lack of dependence on probability distribution function of data (Lee, J. S. (2004). The decision tree rules for the classification were selected quantitatively on the basis of thestatistical Gumbel distribution empirical model and experimental investigation. The another objective of this proposed work is to improve our understanding about supervised classification to see, how they interact with training data, and how they affect cluster labeling for land cover classification if input parameters are SAR observables obtained by decomposition methods. In the proposed paper supervised classification methods, namely minimum distance, maximum likelihood, and parallelepiped, etc., are used for classification based on Raney decomposition. The parameters obtained by m-chi, m-alpha, and m-delta decomposition, are also taken as input parameters for the unsupervisedisodata classifier which is a state of art method used more often for classification.In the proposed work, it is observed thathybridpolarimetricSARsystemsaremorecapableof discriminating different land covers than Linear polarimetric SAR data. Thus in this paperRISAT-1hybridpolarimetric data has been usedfor the study.

2. Research Study Area

The researcharea is located in and aroundthegreater Visakhapatnam Municipal Corporation, Andhra Pradesh. The area covered in this investigation is about 621.52 sq.km of Visakhapatnam district,one of the North Coastal districts of Andhra Pradesh and it lies between 17^0 10' and 17^0 56' N latitude and 83^0 08' and 83^0 40' E longitude (Fig: 1). It is bounded on Northside partly by Orissa state and partly by Vizianagaram district, towards South by East Godavari district, towards West by Orissa state and towards EastbytheBay of Bengal.

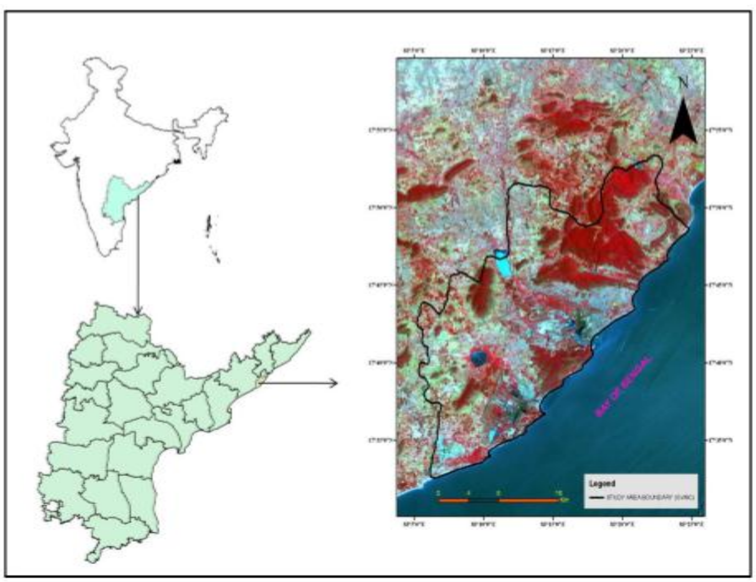


Figure 1. Location Map of the Study Area

3. Data Used

RISAT-1 data has been acquired in circular fine resolution strip map (cFRS) mode on October 21, 2014,over the city of Visakhapatnam an urban city of Andhra Pradesh state, India with the central latitude of 17.41N, central longitude of 83.23E and at an incidence angle of 38.780. The dataset has an azimuth resolution of 2.38m and ground range resolution of 2.87m. The scene was imaged during the ascending pass of the satellite with right looking (in international terminology, it is left looking) sensor orientation. Ground-truth parameters regarding soil moisture, urban, vegetation height and vegetation type, etc., were collected synchronously with the satellite passes. Ground truths are collected from various places covering the entire region under test. Around193 ground control points (GCP) were collected for training and 840 for testing the accuracy of classification map. Table-1 presents the training and control samples based on ground truth data. Based on ground truth information, six classes were identified: water (including sea water also), urban, vegetation (cropland, grassland, shrubs, trees,etc.), beach sand, road and bare soil surface.

Table 1. Ground Truth Survey Points

Class	Training samples	Test samples
Water	5	302
Urban	75	265
Vegetation	89	7
Beach Sand	4	137
Bare Soil	19	66
Road	1	45

4. Data Processing

RISAT-1 cFRS SLC data (level 1 product) was usedin the present study. Radiometric calibration was performed, and the data was multi-looked three times in range, and azimuth direction and a C2 matrix were generated from RH and RV data. A refined Lee filter was applied with a 5x5 window to suppress speckle noise. The filtered dataset was decomposed using m-χ decomposition and analysis was performed to discriminate various features through scattering mechanism. Until the generation of C2 matrix, own code is used,and then PolSARpro 5.0 is used for m-χ decomposition. Around 500 pixels were selected for each class for quantifying into even-bounce, odd-bounce and volume scattering components of PolSARpro 5.0 output. The values of the three componentswere normalized for each pixel using min, max value of the 500 pixels and the highest value of the three was considered to be the dominant scattering mechanism for that pixel. For the entire test area, the percentage of pixels showing dominant surface, double bounce and volume scatter was computed and plotted.

To classify the image,supervised classification was performed on the filtered image with a 5x5 window. The accuracy of the result of the classification was assessed by computing the confusion matrix from test areas and is shown in Table-2 for October 2014 data. Training and test areas for these studies were obtained from the ground survey and GPS data.

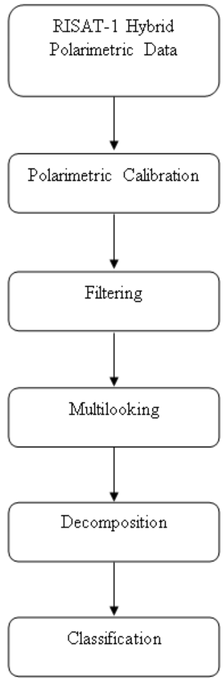


Figure 2. Methodology for land use land cover classification of RISAT-1 SAR data

5. Decision Tree Algorithm

Decision tree approach requires athorough knowledge of information-bearing features and their physical understanding. It has been already known that phase intensity is a function of the electromagnetic wave parameters such as wave frequency, its polarization, and its incidence angle, and it depends on the target characteristics such as surface geometry (size, shape, orientation distribution and spatial arrangement of objects), physical property (symmetry, non symmetry or irregularity of the target) and dielectric characteristics of the medium.The objective of the proposed work is to extract physical information from backscattering phase intensity behavior of various objects.The task of this work is to determine the class of each image pixel based on their features. The selection of proper features is important for classification. Decision trees are commonly used for variable selection to reduce data dimensionality in image analysis. Decision trees are used to predict membership of cases or pixels in the classes of a categorical dependent variable from their measurements on one or more predictor variables. In these tree structures, leaves represent classifications and branches represent conjunctions of features that lead to those classifications.

Decision tree algorithms have many advantages. They are white box model and simple to understand and interpret. If a given result is provided by the model, the explanation for the result is easily replicated by simple math, Decision trees are able to handle both numerical and categorical data, and requires little data preparation, they are robust and perform well with large data in a short time and Decision trees, performing univariate splits and examining the effects of predictors one at a time, have implications for the variety of types of predictors that can be analyzed.

In this study, Gumbel distribution statistical model was used to implement the LULC classification. The Gumbel distribution statistical model is to help for discriminating the threshold values from the satellite imagery by using band math in ENVI. The decision tree algorithm was performed by using the statistical model threshold values for classification. A decision tree can be created based on training samples using Gumbel distribution statistical model. After the decision tree is constructed, it can be used to identify the class of other unknown cases.

e-ISSN: 2348-6848
p-ISSN: 2348-795X
Volume 04 Issue 17
December 2017

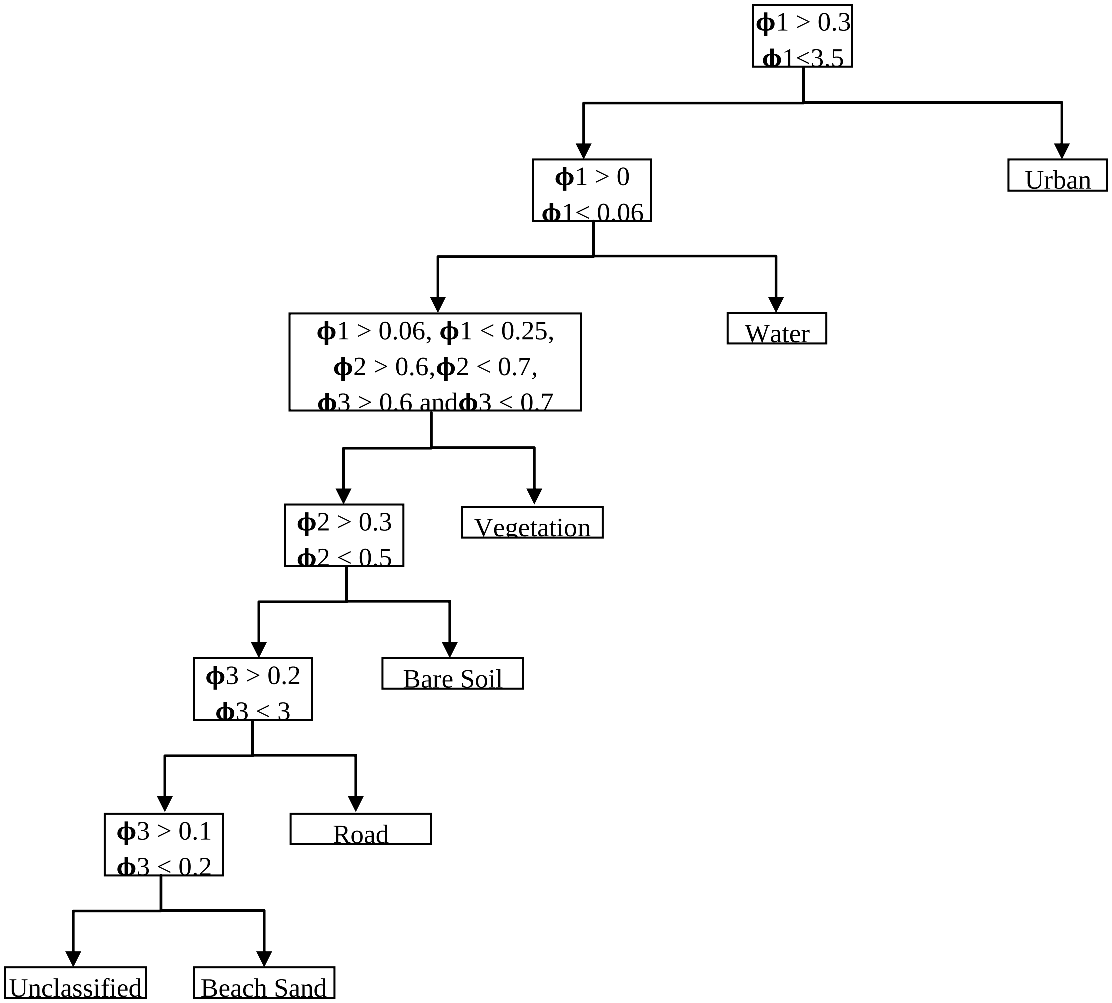


Figure 3. Algorithm for Decision Tree Classification, ϕ1, ϕ2 &ϕ3 are Even, Diffuse & Odd bounce respectively

6. Results and Discussions

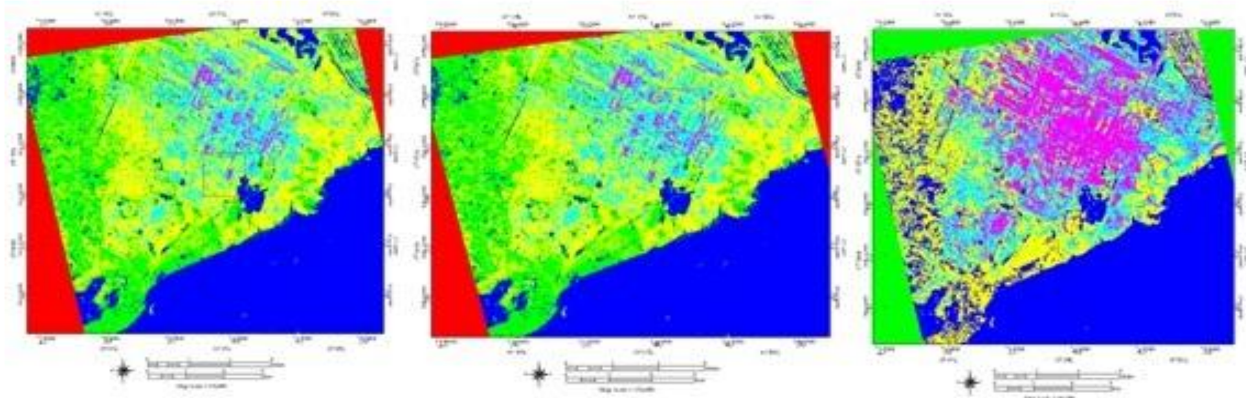
In this section,theclassification results that are obtained from various classifiers like Decision Tree classification, Artificial Neural Net (ANN) classification, supervised classifications such as minimum distance(MD), maximum likelihood(ML), parallelepiped based onRaney decomposition and also unsupervised isodata classification based on Raneydecomposition. Using the confusion matrixwhich is also called error matrix, overall accuracy, kappa coefficient of all the classification results are calculated with the help of ground truth information. All the results are obtained by polSARpro and ENVI processing tools.

m-alpha m-chi m-delta

Figure 4. Raney Decomposed Images with Ground truth points of RISAT-1 21st October 2014

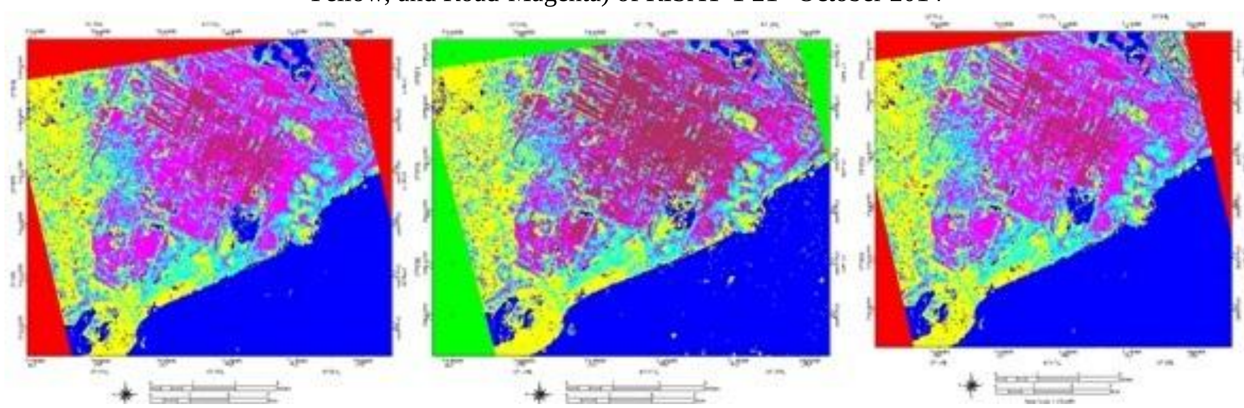
e-ISSN: 2348-6848
p-ISSN: 2348-795X
Volume 04 Issue 17
December 2017

Unsupervised Classification



m-alpha m-chi m-delta

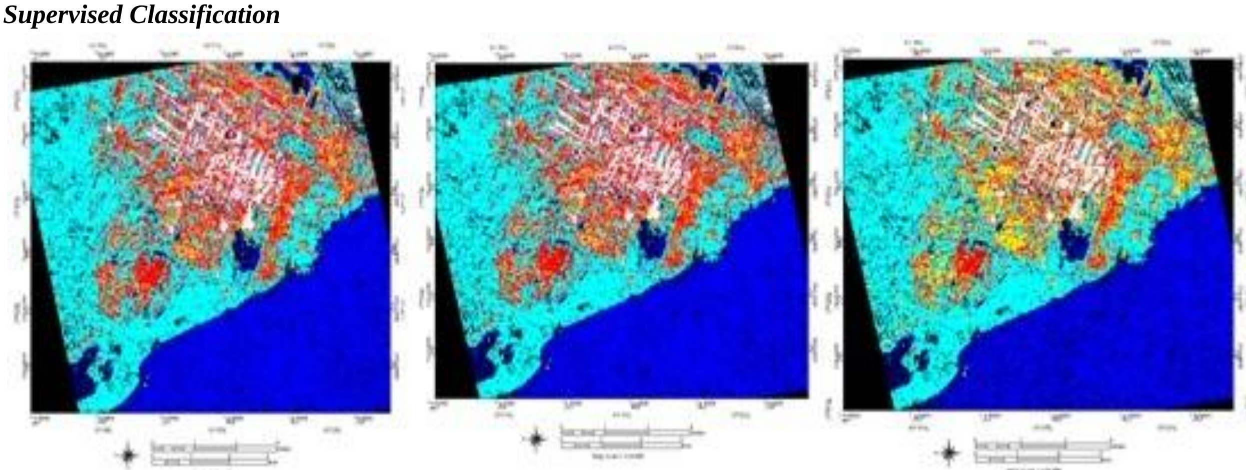
Figure 5. Isodata classified Images (Urban-Cyan, Vegetation-Green, Water-Blue, Beach sand–Light green, bare soil-Yellow, and Road-Magenta) of RISAT-1 21st October 2014



m-alpha m-chi m-delta

Figure 6. K-Mean classified Images (Urban-Brown, Vegetation-Magenta, Water-Blue, Beach sand–Yellow, Bare soil-Cyan and Road-Light Brown) of RISAT-1 21st October 2014

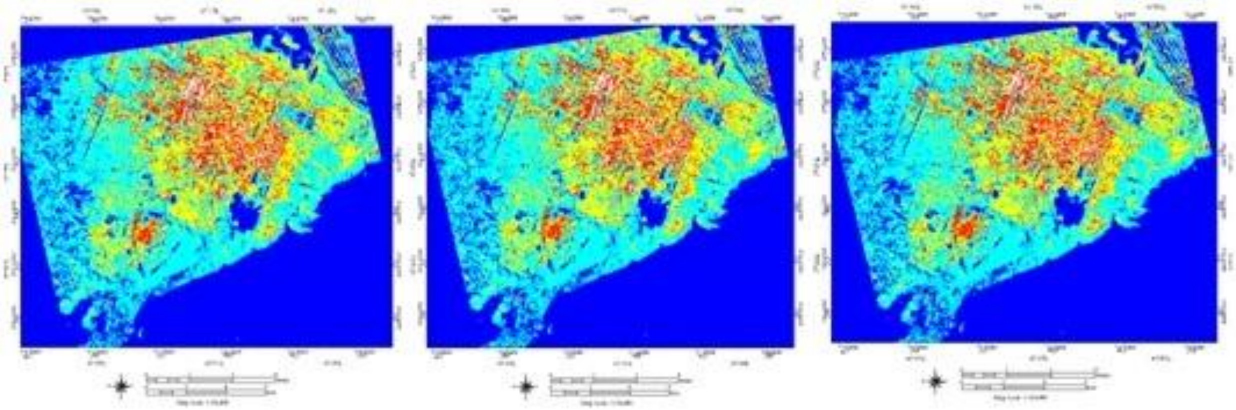
Supervised Classification



m-alpha m-chi m-delta

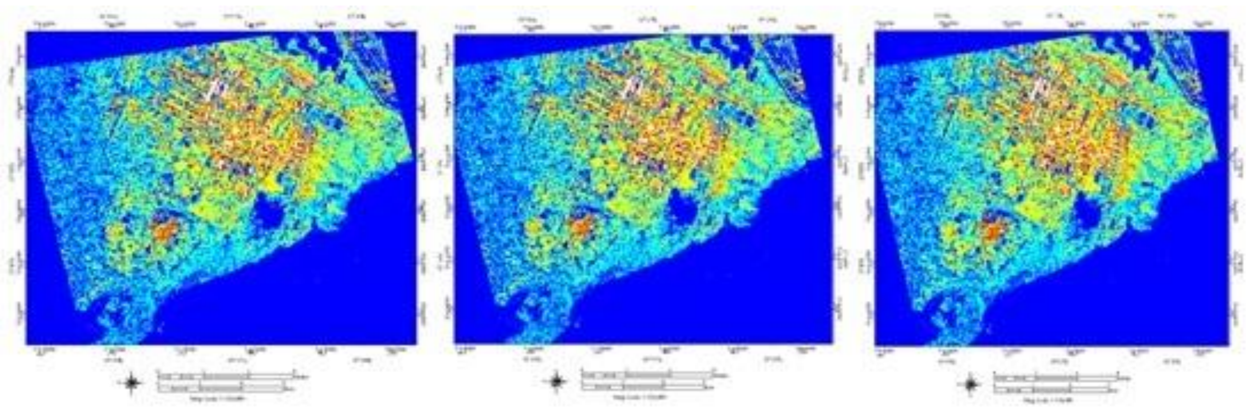
Figure 7. Parallelepiped classified Images (Urban-White, Vegetation-Red, Water-Blue, Beach sand–Cyan, Bare soil-Yellow, and Road-Magenta) of RISAT-1 21st October 2014

e-ISSN: 2348-6848
p-ISSN: 2348-795X
Volume 04 Issue 17
December 2017



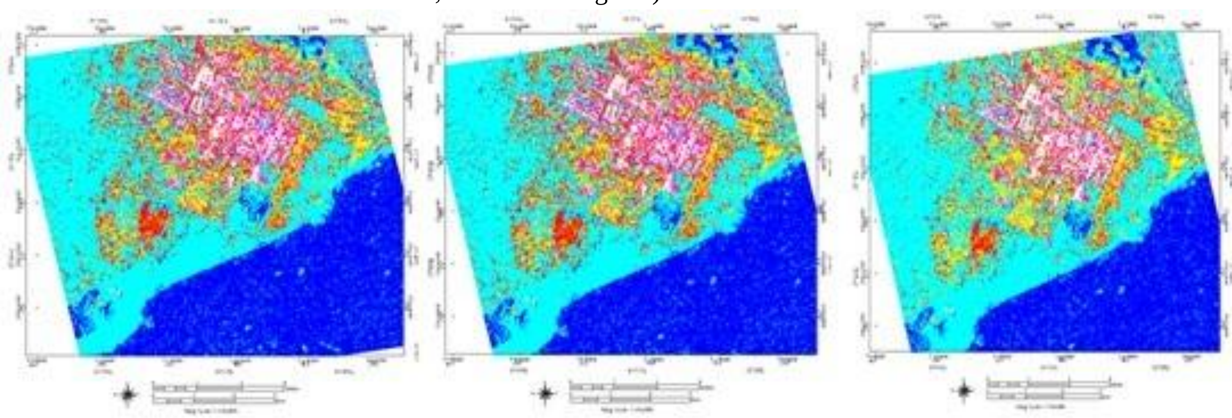
m-alpha m-chi m-delta

Figure8. Minimum Distance classified Images (Urban-White, Vegetation-Red, Water-Blue, Beach sand–Cyan, Bare soil-Yellow, and Road-Magenta) of RISAT-1 21st October 2014



m-alpha m-chi m-delta

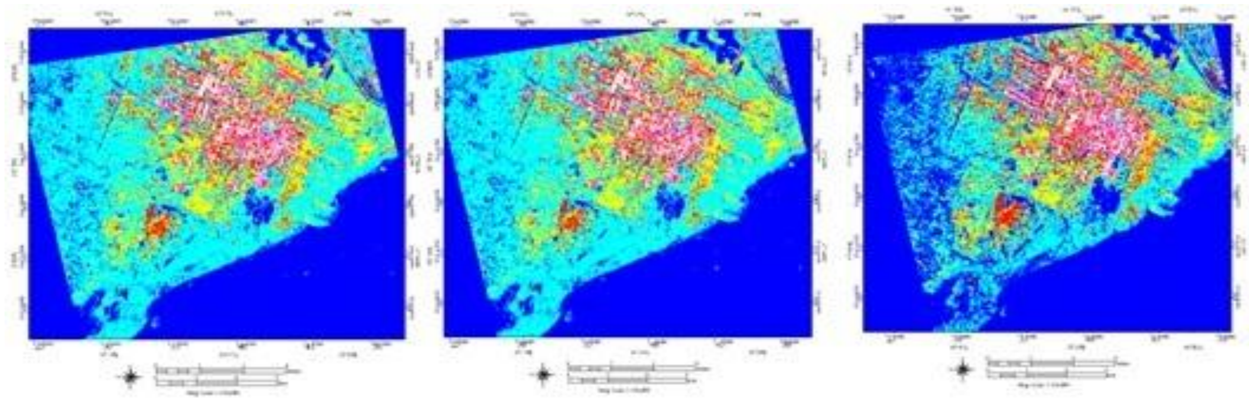
Figure 9. Mahalanobis Distance classified Images (Urban-White, Vegetation-Red, Water-Blue, Beach sand–Cyan, Bare soil-Yellow, and Road-Magenta)of RISAT-1 21st October 2014



m-alpha m-chi m-delta

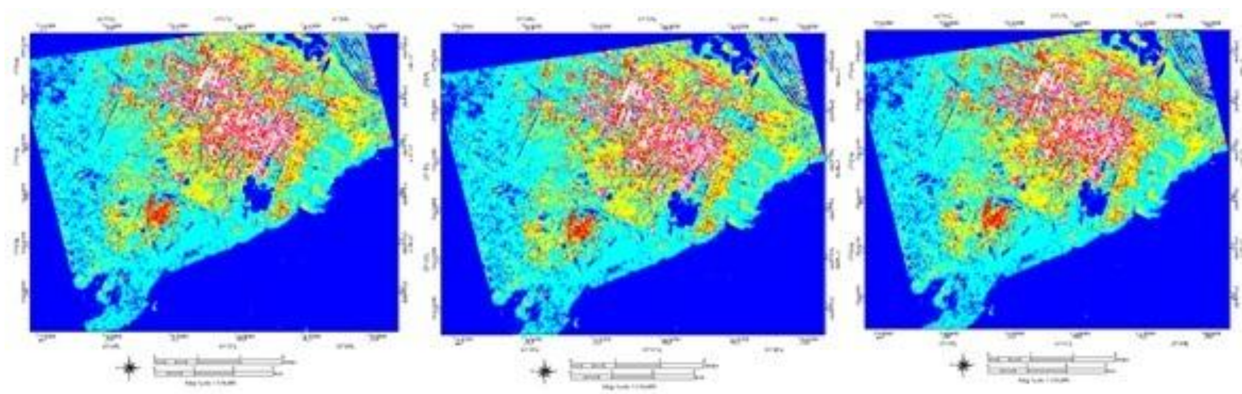
Figure10.Maximum likelihood classified Images (Urban-White, Vegetation-Red, Water-Blue, Beach sand–Cyan, Bare soil-Yellow, and Road-Magenta)of RISAT-1 21st October 2014

e-ISSN: 2348-6848
p-ISSN: 2348-795X
Volume 04 Issue 17
December 2017



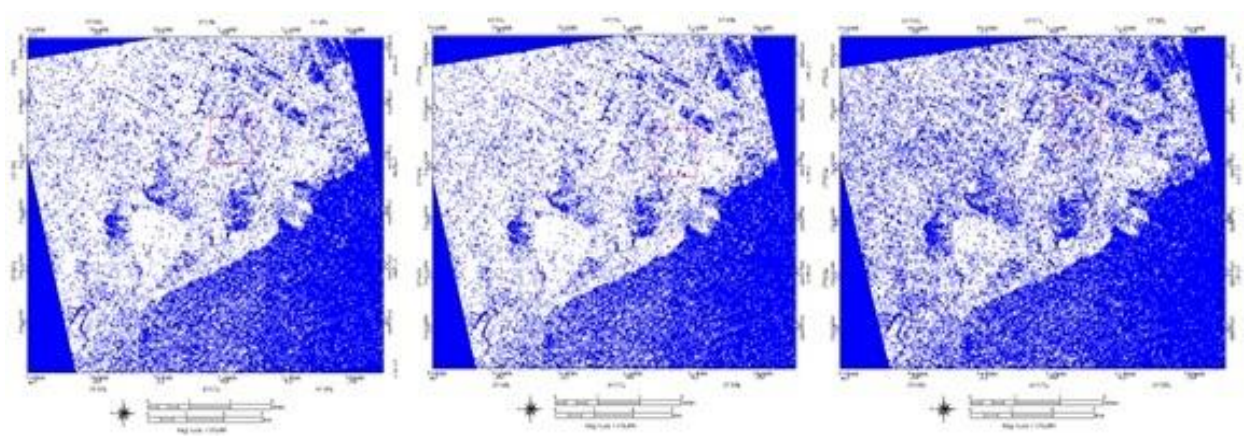
m-alpha m-chi m-delta

Figure 11.Artificial Neural Net classified Images (Urban-White, Vegetation-Red, Water-Blue,Beach sand–Cyan, Bare soil-Yellow, and Road-Magenta)of RISAT-1 21st October 2014



m-alpha m-chi m-delta

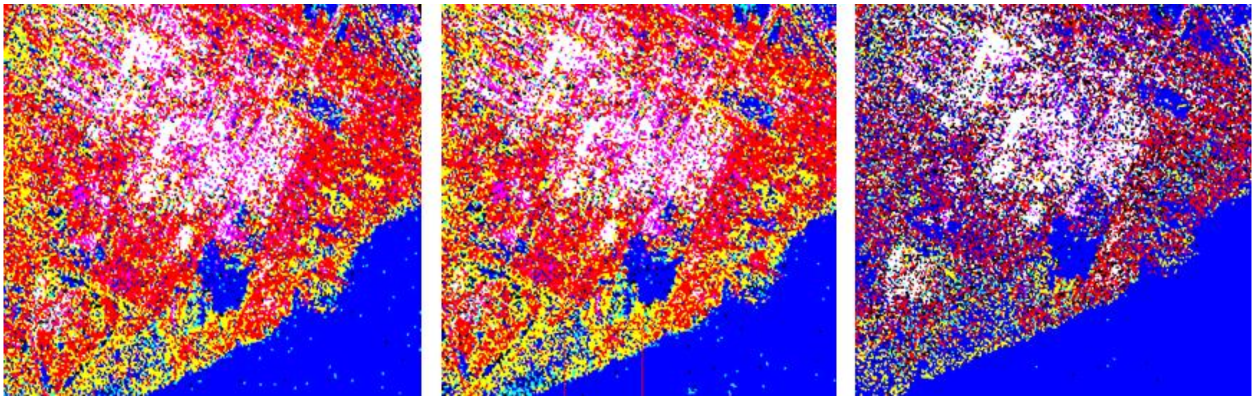
Figure12. Support Vector Machine classified Images (Urban-White, Vegetation-Red, Water-Blue, Beach sand–Cyan, Bare soil-Yellow, and Road-Magenta)of RISAT-1 21st October 2014.



m-alpha m-chi m-delta

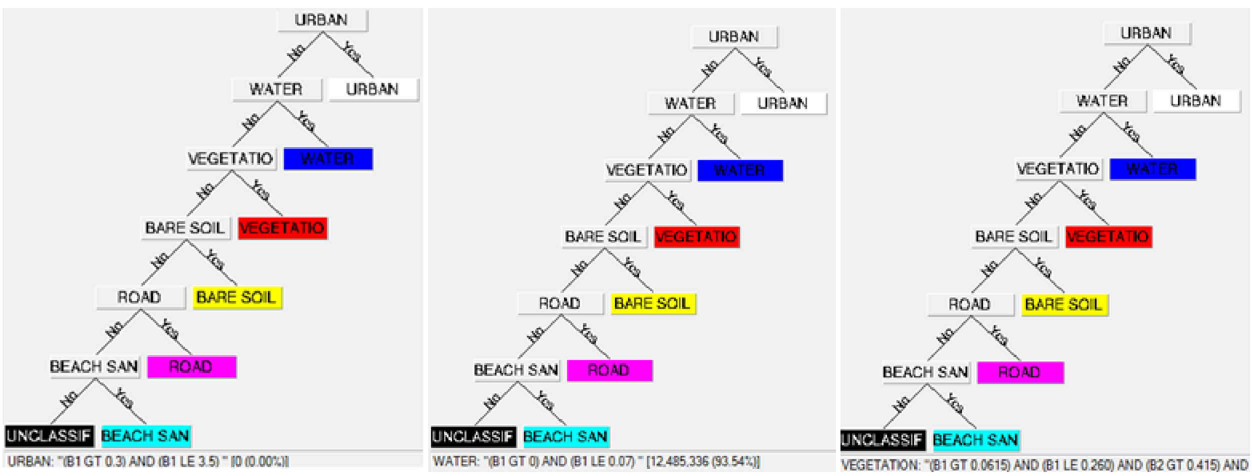
Figure 13. Binary Coding classified Images (Urban-White, Vegetation-Red, Water-Blue, Beach sand–Cyan, Bare soil-Yellow, and Road-Magenta) of RISAT-1 21st October 2014

International Journal of Research
Available at https://edupediapublications.org/journals
e-ISSN: 2348-6848
p-ISSN: 2348-795X
Volume 04 Issue 17
December 2017



m-alpha m-chi m-delta

Figure 14. Decision Tree classified Images (Urban-White, Vegetation-Red, Water-Blue, Beach sand–Cyan, Bare soil-Yellow, and Road-Magenta) of RISAT-1 21st October 2014.



m-alpha m-chi m-delta

Figure 15. Decision Tree feature branches

The classification results showing overall classification accuracyand kappa coefficient for different methods are summarizedin table2. The results show that decision tree classifier performsbetter than all the parametric supervised classification techniques(maximum likelihood(ML),theminimum distance(MD) and parallelepiped) and unsupervised isodata classification. This is because this classifier is implementedwiththoroughexpert knowledge-based data obtained by Gumbel distribution statistical model andexperimental validation which does not require anyassumptionsabout the frequency distribution properties of the data. In the case of non-parametric classifiers, the maximum likelihood (ML) classifier performs better than the minimum distance (MD)and parallelepiped classifiers. This is because the maximum likelihood (ML) classifier is based on the Bayesian probability theory developed on second order statistics.On the contrary Minimum distance (MD) and parallelepiped classification methodsare basedonfirst order statistics. This maximum likelihood classifier assumes that input data (training data) is normally distributed and independent. These classification tests based on Raney decomposition and it gives not such good results for all three classification algorithms.

This is because the removal of the speckles done with boxcar filter which resultsinblurring of sharp edges and over-filtering effects. Isodata and k-mean classification also could not give good results.

Accuracy Assessment

Table 2.Overall Accuracy (O.A) and kappa coefficient estimates for all the features obtained by different decomposition and classification techniques

Classification	M-Alpha		M-Chi		M-Delta	
Unsupervised Classification	Overall Accuracy	Kappa Coefficient	Overall Accuracy	Kappa Coefficient	Overall Accuracy	Kappa Coefficient
Isodata	31.523	0.171	36.078	0.212	36.078	0.212
K-Mean	56.562	0.473	36.439	0.243	56.562	0.472
Supervised Classification	M-Alpha		M-Chi		M-Delta	
Parallelepiped	77.689	0.730	77.689	0.730	74.447	0.691
Minimum Distance	84.380	0.812	84.380	0.812	85.589	0.826
Mahalanobis	72.826	0.674	72.826	0.674	75.656	0.708
Maximum Likelihood	97.247	0.967	97.247	0.967	97.221	0.967
Artificial Neural Net	96.732	0.961	96.732	0.961	97.427	0.970
Support Vector Machine	94.262	0.931	94.262	0.931	94.853	0.938
Binary Coding	23.623	0.087	23.623	0.087	21.616	0.063
Decision Tree	99.743	0.997	96.873	0.962	99.857	0.998

7. Conclusion

Them-delta decomposition for decision tree classification shows promising results for different features such as odd, even and volume scattering. Thedifferent classes of urban land, water, and vegetation are decomposed as even bounce, odd bounce and volume scatterers.The decision treeclassifier detects all land cover features more accurately from Gumbel distribution model image pixel than other artificial neural net, support vector machine, parallelepiped, the minimum distance(MD), maximum likelihood(ML)andisodata classifier. In this proposed paper, decision tree algorithm is appliedto C-band data, and it works well for this band. In order to apply to different band data, we have to check parameter values again for trainingexpert knowledge basedbecause it is a well-knownfact that radar response to various land covers is polarization and frequency dependent.

8. References

[1] Cameron, W. L., and L. K. Leung. "Feature motivated polarization scattering matrix decomposition." *Proceedings of IEEE I international lRadar Conference*, May 1990.

[2] Cloude, S.R and E. Pottier. "A review of target decomposition theorems in radar polarimetry." *Geoscience and Remote Sensing* 34, no. 2 (1996): 498-518.

[3] Cloude, S.R, and E.Pottier. "An entropy based classification scheme for land applications."

[4] Cloude, S.R.and E.Pottier. "An entropy based classification scheme for land applications of polarimetric SAR." *IEEE* (Geoscience and Remote Sensing) 3, no. 1 (1997): 68-78.

[5] de'ath, G. and K. E.Fabricius. "Classification and regression trees: A powerful yet simple technique for ecological data analysis." *Ecology* 81, no. 11 (2000): 3178-3192.

[6] Enderle, D. I. M. and R. C. Weih Jr.,. "Integrating supervised and unsupervised classification methods to develop a more accurate land cover classification." *Journal of the Arkanas Academy of Science* 59 (2005): 65-73.

[7] Fang, C., H. Wen and W. Yirong. "An improved Cloude-Pottier decomposition using H/R/span and complex Wishart classifier for polarimetric SAR classification." *International Conference on Radar*, 2006.

[8] Ferr0-Famil, L. and E. Pottier. "Dual frequency polarimetric SAR dara classification and analysis." *Progress in Electromagnetics Research* 31 (2001): 247-272.

[9] Ferro-Famil, L., E. Pottier, and J. S. Lee. "Unsupervised classification of multi-frequency and fully polarimetric SAR images based on the H/A/Alpha-Wishart classifier." *IEEE Transactions on Geoscience and Remote Sensing* 39, no. 11: 2332.

[10] Friedl, M. A. and C. E. Brodley. "Decision tree classification of land cover from remotely sensed data." *Remote Sensing of Environment* 61 (1997): 399-409.

[11] Hui, Y., et el.,. "Extracting wetland using decision tree classification." *Proceeding of the 8th WSEAS International Conference on Applied Computer and Applied Computational Science*, 2004: 240-245.

[12] Krogager, E.,. "A new decomposition of the radar target scattering matrix." *Electronics Letter* 26, no. 18 (1990): 1525-1526.

[13] Lee, J. S., M. R. Grunes, and E. Pottier. "Quantitavie comparison of classification capability: Fully polarimetric versus dual and single-polarization SAR." *IEEE Transactions on Geoscience and Remote Sensing* 39, no. 11 (2001).

[14] Lee, J. S., M. R. Grunes, E. Pottier, et el. "Unsupervised terrain classification preserving polarimetric scattering characteristics." *IEEE Transactions on Geoscience and Remote Sensing* 42, no. 4 (2004): 722-731.

[15] LEE, J. S., M. R. Grunes, T. L. Ainsworth, et al.,. "Unsupervised classification of polarimetric SAR images by applying target decomposition and complex Wishart distribution." *IEEE Transactions on Geoscience and Remote Sensing* 37, no. 5 (1999): 2249-2258.

[16] Lee, J.S., M. R. Grunes, and R.Kwok. "Classifications of multi-look polarimetric SAR imagery based on the complex Wishart distribution." *International Journal*

of Remote Sensing 15, no. 11 (1994): 2299-2311.

[17] McNaim, H., J. shang, X. Jiao, and C. Champagne. "The contribution of ALOS PALSAR multipolarization and polarimetric data to crop classification." *IEEE Transactions on Geoscience and Re-Progress In Electromagnetics Research B* 47, no. 12 (2009): 3981-3992.

[18] Ouarzeddine, M. and B. Souissi. "Unsupervised classification using wishart classifier." *USTHB, F.E.I.*, (EI Alia Bab Ezzouar), 2007.

[19] Pal, M. and P. M. Mather. "An assessment of the effectiveness of decision tree methods for land cover classification." *Remote Sensing of Environment* 86 (2003): 554-565.

[20] Park, S. E. and W. M. Moon. "Unsupervised classification of scattering mechanisms in polarimetric SAR data using fuzzy logic in entopy and alpha plane." *IEEE Transactions of Geoscience and Remote Sensing* 45, no. 8 (2007): 2652-2664.

[21] Pottier, E.,. "Radar target decomposition theorems and unsupervised classification of full polarimetric SAR data." *IEEE transactions of Geoscience and Remore Sensing Symposium*, 1994: 1139-1141.

[22] Praks, J., E. C. Koeniguer, et el.,. "Alternatives to target entropy and alpha angle in SAR polarimetry." *IEEE Transactions on Geoscience and Remote Sensing* 47, no. 7 (2009): 2262-2274.