Terrestrial Cloud Classification

Koosh Doctor
DA-IICT, Gandhinagar
{201101221}@daiict.ac.in
Supervisor
Ranendu Ghosh

***Abstract*— Efficient and accurate interpretations of satellite imagery are very useful in weather forecasting and other meteorological operations. Terrestrial cloud classification is one such area which has been explored in this paper. Automatic cloud classification proves to be useful in such an area where large amount of data is dealt with and manual interpretation proves to be tiresome and more prone to errors. Hence, in order to have a significantly less time-consuming and error-free intervention, a computer-based cloud classification system will prove vital in larger projects where this system maybe a part of it. Visible and Infrared images from the Geostationary satellite Kalpana -1 VHRR are used as testing and training data.**
Coefficient of variance(COV) of each feature provides an intuitive estimate of the importance of the feature in cloud classification. The features having COV between 1% and 20% are likely to be the most significant in cloud classification.

***Index Terms*— Cloud Classification, Kalpana-1 VHRR, ReliefF algorithm,k-nearest neighbour (kNN) classifier,Coefficent of Variance**

I. INTRODUCTION

Real-time cloud classification systems have a high potential to prove to be an important asset in quick weather forecasting and meteorological analysis. The purpose of the paper is to extract cloud features from Visible and Infrared images which can be used by a combination of ReliefF algorithm and kNN classifier to classify a satellite image, containing gray levels of each pixel, into one of the 12 classes of clouds.
Previously, considerable research has been done in the area of cloud classification. Richard L. Bankert et al.(1992)[1] was able to classify the images into 10 classes of clouds with an accuracy of 79.8% by using the combination of Sequential Forward Selection Procedure and a Probabilistic Neural Network. Rashpal Kaur et al.(2007)[2] classified clouds into low, medium and high clouds by using threshold values of the Singular Value Decomposition matrix of 8x8 pixel images. The classification accuracy achieved in this case was 70-90%.A thorough study has been done by Bin Tian et al.[3], where they used a combination of PNN classifier, SVD matrix, Sequential Forward Selection Procedure and gray level co-occurrence matrix(GLCM) method for calculating the spectral features.

II. MOTIVATION

All the previously adopted approaches have resulted in cloud classification accuracy not more than 79.8% where the method classifies into 10 or more classes. There are cases where 90% classification accuracy have been achieved but with a smaller set of classes. This inspires a search for finding an empirical evaluation of different classification methods in order to improve the classification accuracy with a set having maximum number of classes.
The empirical evaluation worked upon by Zoran Bosnic et al.(2010)[4] have calculated the classification accuracy of different objects classified by the combination of each of the following feature selection procedure and classifiers :-
Feature Selection Procedures

1) ReliefF
2) Random Forest
3) Sequential Backward Selection
4) Sequential Forward Selection
5) Gini Index

Classifiers :-

1) Decision Tree
2) Random Forest
3) Support Vector Machines
4) Artificial Neural Network
5) Naive Bayes
6) *k* nearest neighbours

The results show that the combination of ReliefF Algorithm and *k*NN classifier provide the highest classification accuracy in most of the objects considered by the paper.

III. BACKGROUND

A. Cloud Classification Scheme

The 12 classes of terrestrial clouds considered in this paper are :-

1) Cirrus
2) Cirrocumulus
3) Cirrostratus
4) Altostratus
5) Nimbostratus
6) Stratocumulus
7) Stratus
8) Cumulus
9) Cumulonimbus
10) Clear
11) Altocumulus
12) Cumulus Congestus

B. Cloud Features

The features taken into consideration are textural, spectral and physical in nature which are taken from visible images(Channel 1) and infrared images(Channel 4). Below are the 204 feature components[1]:-

1) Textural Features
   a) Gray Level Difference Vector(GLDV)* and Sum and Difference Histogram(SADH)*(170)
      i) Mean
      ii) Standard Deviation
      iii) Angular Second Moment
      iv) Entropy
      v) Local Homogeneity
      vi) Contrast
      vii) Cluster Shade
      viii) Cluster Prominence
      ix) Correlation-SADH only

      5 values computed for each measure:16x16 pixel region;maximum, minimum, mean, standard deviation of the 16 4x4 pixel regions within 16x16 pixel region
   b) Run Length*(10)
      i) Short Run Emphasis
      ii) Long Run Emphasis
      iii) Gray Level Distance
      iv) Run Length Distance
      v) Run Percentage
2) Spectral Features*(14)
   a) Maximum Pixel Value
   b) Minimum Pixel Value
   c) Range of Pixel Values
   d) Mode of Pixel Values
   e) Median of Pixel Values
   f) Mean of Pixel Values
   g) Standard Deviation of Pixel Values
3) Physical Features(10)
   a) IR Cloud Fraction
   b) Low Cloud Fraction
   c) Mid-Level Cloud Fraction
   d) Cirrus Cloud Fraction
   e) Multilayer Cloud Index
   f) Cloud Top Temperature
   g) Cloud Albedo
   h) Surface Temperature
   i) Visible Cloud Fraction
   j) Latitude

**AVHRR Channels 1 and 4*

C. ReliefF Algorithm

Since there are 204 different feature components, it is computationally expensive to calculate all the feature values in real-time system where for every small interval of time, a new image is captured and ready for the classification process.

ReliefF algorithm is an effective feature estimator. The algorithm can be viewed as a feature subset selection method which is applied before the actual classification.

1) Algorithm Working Principle: The principle of the ReliefF algorithm (Kira and Rendell, 1992[5]) is to estimate the quality of attributes on the basis of how well their values distinguish between instances that are near to each other. For example, given a randomly selected instance Ri, algorithm searches for its two types of nearest neighbors: one set from the same class, called nearest hit H, and the other from the different class, called nearest miss M. It updates the quality estimation W[A] for all attributes A depending on their values for Ri,M, and H. If instances Ri and H have different values of the attribute A then the attribute A separates two instances with the same class which is not desirable so we decrease the quality estimation W[A]. On the other hand if instances Ri and M have different values of the attribute A then the attribute A separates two instances with different class values which is desirable so we increase the quality estimation W[A]. The whole process is repeated for m times, where m is a user-defined parameter.

D. k *Nearest Neighbour Classifier(kNN)*

The ultimate objective of the paper is to classify an unknown image having a feature vector into one of the previously mentioned cloud classes. The kNN classifier is trained with the feature vectors of already classified images.kNN classifier operate on the premises that classification of unknown instances can be done by relating the unknown to the known according to some distance/similarity function. The intuition is that two instances far apart in the instance space defined by the appropriate distance function are less likely than two closely situated instances to belong to the same class.

1) Classifier Working Principle: The algorithm can be summarised as:
1.A positive integer k is specified, along with a new sample
2.We select the k entries in our database which are closest to the new sample
3. We find the most common classification of these entries
4. This is the classification we give to the new sample

IV. Data

Visible images and Thermal Infrared images taken from Indian Geostationary satellite Kalpana-1 VHRR are used for the working methodology. The images are 16x16 pixel matrices.

V. Working Methodology

A. Feature Extraction

For a particular time interval, the following features are extracted which results in the creation of a feature vector having 204 feature components[1].

1) GLDV Features: GLDV features were calculated from both Channel 1 and Channel 4.

mean $\mu = \sum_m mP(m)$

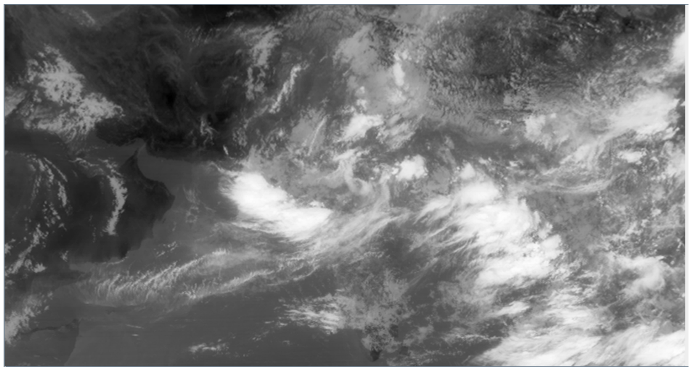

Fig. 1. TIR image provided by Space Application Centre,ISRO

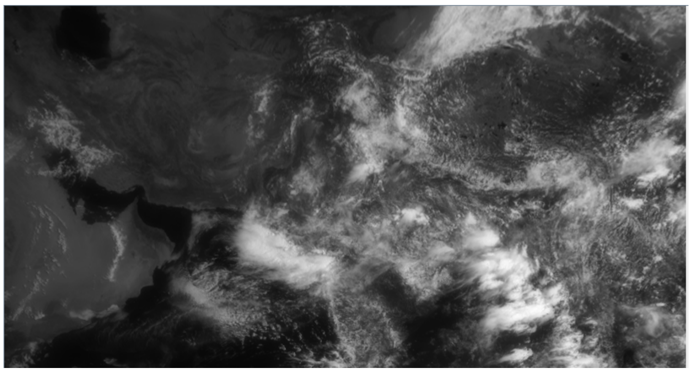

Fig. 2. VIS image provided by Space Application Centre,ISRO

standard deviation $\sigma = \sqrt[2]{\sum_m (m-\mu)^2 P(m)}$

angular second moment asm = $\sum_m [P(m)]^2$

entropy ent = $-\sum_m P(m)logP(m)$

local homogeneity lh = $\sum_m P(m)/(1+m^2)$

contrast con = $\sum_m m^2 P(m)$

cluster shade cs = $[\sum_m (m-\mu)^3 P(m)]/\sigma^3$

cluster prominence cp = $[\sum_m (m-\mu)^4 P(m)]/\sigma^4 - 3$

where m=I-J, the absolute difference of gray levels one pixel apart in a fixed direction.P(m) is the difference vector probability density function(estimated by gray level frequencies of occurrence/ total frequencies).

2) SADH Features: SADH features were calculated from both Channel 1 and Channel 4.

mean $\mu_S =$$\sum_K K P_S(K)$

standard deviation
sd = $\sqrt[2]{1/2[\sum_K (K-\mu_S)^2 P_S(K) + \sum_L L^2 P_D(L)]}$

angular second moment asm = $[\sum_K [P_S(K]^2 + \sum_L [P_D(L)]^2]$

contrast con = $\sum_L L^2 P_D(L)$

correlation
cor = $1/2[\sum_K (K-\mu_S)^2 P_S(K) - \sum_L L^2 P_D(L)]/sd^2$

entropy
ent = $-\sum_K P_S(K)log(P_S(K)) \sum_L P_D(L)log(P_D(L))$

local homogeneity lh =$\sum_L P_D(L)/(1+L^2)$

cluster shade cs = $[\sum_K (K-\mu_S)^3 P_S(K)]/sd^3$

cluster prominence cp = $[\sum_K (K-\mu_S)^4 P_S(K)]/sd^4 - 3$

where K=I+J and L=I-J. P$_S(K) and P_D(L)$ are the probability density function.

3) Run Length : Run Length statistics were calculated for both Channel 1 and Channel 4. The measurement of the features are based on the sets of adjacent pixels in a particular direction having the same gray level.

short run emphasis sre =$1/T_r \sum_i \sum_j P(i,j)/j^2$

long run emphasis lre = $1/T_r \sum_i \sum_j j^2 P(i,j)$

gray level distribution gld = $1/T_r \sum_i [\sum_j P(i,j)]^2$

run length distribution rld = $1/T_r \sum_j [\sum_i P(i,j)]^2$

run percentage rp = $1/T_p \sum_i \sum_j P(i,j)$

where :

i = 0 to N$_g - 1$

j =1 to N$_r$

N$_g$$-$number of gray level$s$

$N_r$$-$number of runs

$T_p$$-$number of image pixels

$T_r$$-\sum_i \sum_j P(i)(j)$

P(i,j) - number of occurrences of runs of length j having gray level i

4) Spectral Features: Spectral features include maximum, minimum, range, mode, median, mean and standard deviation of pixel values in channels 1 and 4.

5) Physical Features: This class of features include visible cloud fraction, albedo, surface temperature, cloud top temperature, infrared cloud fraction, low cloud fraction, low cloud fraction, mid-level cloud fraction, cirrus cloud fraction and multilayer cloud index.

B. ReliefF Algorithm

MATLAB(2012b) has a toolbox for ReliefF algorithm.
Syntax :-
rank=relieff(X,Y,K) where

1) X is the matrix having 204 columns(features) and N rows where each row represents 1 instance of a time interval when the image is taken.
2) Y is a matrix which contains 1 column and N rows each indicating the class of the corresponding to the rows in X.
3) K is the count of nearest neighbours in the same class taken into consideration.Since there are 12 types of clouds, the value of K must be the integral value of N/12. For example :- if there are 96 instances taken as input, the value of K should be 8 for optimum results.

The output from the algorithm would be the rank of each feature component according to its significance in determining the class of the image.

C. kNN Classifier

MATLAB(2012b) has a toolbox for *k*NN classifier. The adopted procedure is as follows :-

1) Divide the classified data into two types of data :-
   a) Training Data
   b) Testing Data
2) Include the first ranked feature in the feature vector and train the classifier with the training data on the basis of the selected feature vector. Syntax :-
   U=ClassificationKNN.fit(X,Y) where
   a) X is the matrix where each row represents an instance and each column represents the new feature vector
   b) Y is the class of the instance of the training data
3) Classify the testing data
   Syntax :-
   predict(U,Xnew) where
   a) Xnew is the new feature vector of the testing data
4) Calculate the classification accuracy
5) Add the next ranked feature into the feature vector and repeat the steps from step 2 where classification is done on the basis of a feature vector having one more feature.

Keep on adding a new feature to the feature vector until the maximum classification accuracy is achieved.
Hence, in the end we get a feature vector having N feature components which provide the highest classification accuracy with the least number of feature components.

VI. RESULTS

A. Feature Extraction

For 18 satellite imagery instances,Figures 3-9 , coefficient of variance is plotted for each feature.

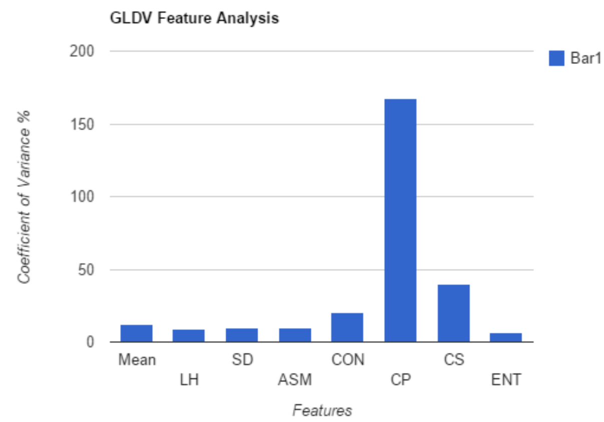

Fig. 3. GLDV features Analysis of VIS images

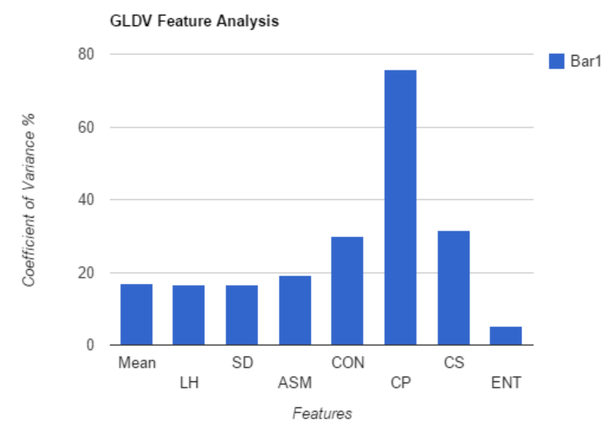

Fig. 4. GLDV features analysis of TIR images

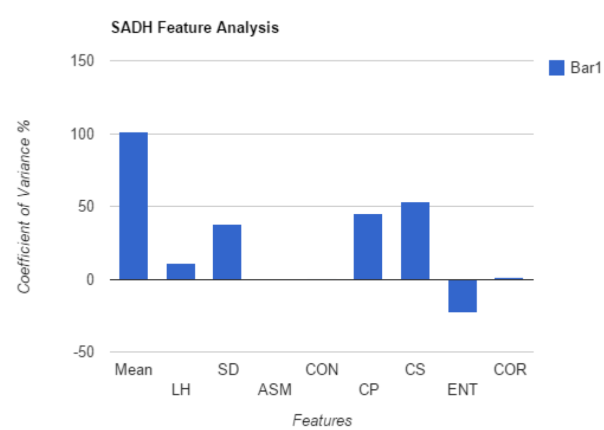

Fig. 5. SADH features analysis of VIS images

Coefficient of variance = standard deviation/mean

Coefficient of variance gives an intuitive estimation of the significance of a feature. A feature must not vary too

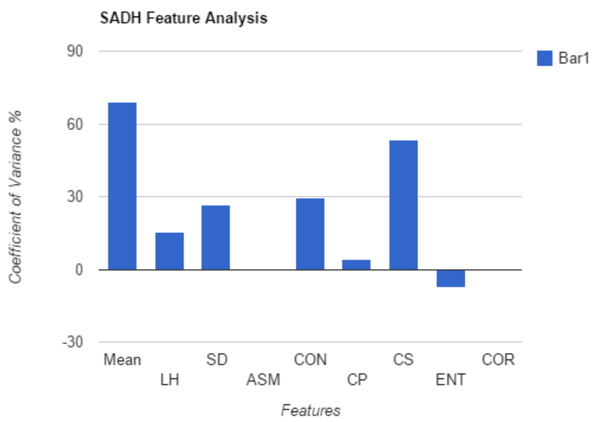

Fig. 6. SADH features analysis of TIR images

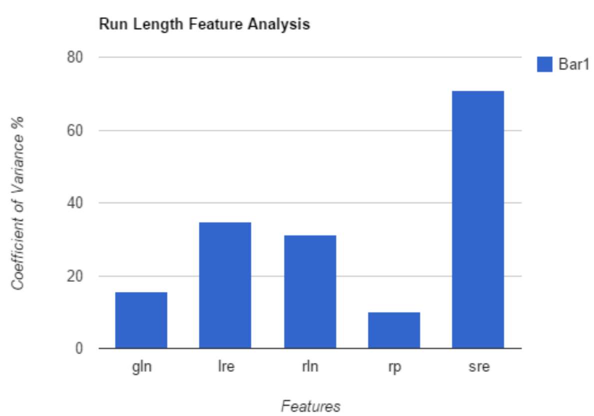

Fig. 7. Run Length Features analysis of VIS images

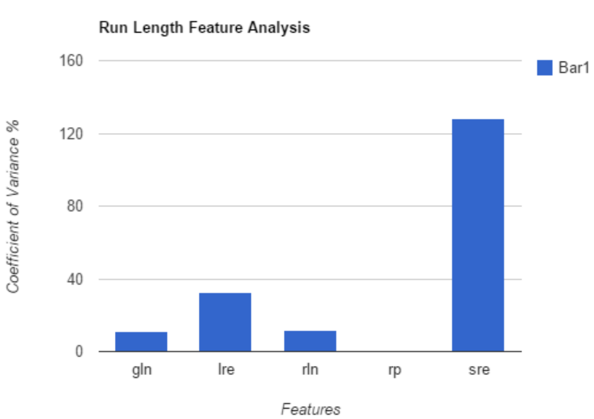

Fig. 8. Run Length Features analysis of TIR images

much since if it varies at a larger scale among images which belong to the same class, the significance of the feature in classification decreases. Additionally, the feature should not remain static since if the values of that feature is similar

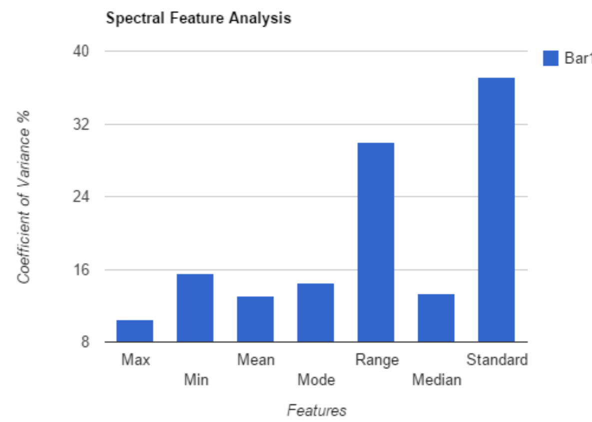

Fig. 9. Spectral Features analysis of VIS images

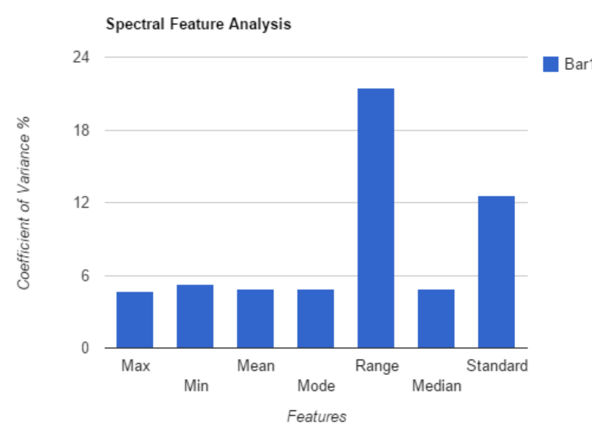

Fig. 10. Spectral Features analysis of VIS images

for images belonging to different classes, its significance decreases in classification.

B. *ReliefF algorithm and* kNN *classifier*

Since this intervention requires expertly classified images, the results of this part of the simulator are not available.

VII. Scope in Future

1) With the presence of manually classified images to train the classifier, this approach of cloud classification can be validated and has a high potential to have a higher accuracy than contemporary methods
2) By analyzing the current cloud patterns, a probabilistic approach can be employed in order to predict cloud patterns for the future time intervals.

VIII. Conclusion

A new approach of real-time cloud classification has been partially implemented. Spectral and Textural features of large number of images are extracted. Since classified images are an integral part of the classifier and are not available due to some unfortunate reasons, the system could not be completed.

IX. Acknowledgment

I would like to thank my mentor,Dr. Ranendu Ghosh for his unconditional support and guidance in this field of meteorological classification.
I would also like to express special thanks to Dr. V.Sathiyamoorthy,Scientist,Space Applications Centre (ISRO) for attempting to provide appropriate training data for the BTP.

References

[1] Richard L Bankert. Cloud classification of avhrr imagery in maritime regions using a probabilistic neural network. *Journal of Applied Meteorology*, 33(8):909–918, 1994.

[2] Rashpal Kaur and A Ganju. Cloud classification in noaa avhrr imageries using spectral and textural features. *Journal of the Indian Society of Remote Sensing*, 36(2):167–174, 2008.

[3] Bin Tian, Mukhtiar A Shaikh, Mahmood R Azimi-Sadjadi, Thomas H Vonder Haar, and Donald L Reinke. A study of cloud classification with neural networks using spectral and textural features. *Neural Networks, IEEE Transactions on*, 10(1):138–151, 1999.

[4] Luka Čehovin and Zoran Bosnić. Empirical evaluation of feature selection methods in classification. *Intelligent Data Analysis*, 14(3):265–281, 2010.

[5] Kenji Kira and Larry A Rendell. A practical approach to feature selection. In *Proceedings of the ninth international workshop on Machine learning*, pages 249–256, 1992.