An Improved Algorithm for Hierarchical Classification of Remote Sensing Data

Vishal Srivastava[1*], Dheeraj Kumar[1], Arup Kumar Pal[2]

[1]*Department of Mining Engineering, Indian School of Mines, Dhanbad, India.*

[2]*Department of Computer Science And Engineering, Indian School of Mines,Dhanbad, India.*

Abstract - Feature extraction from the remote sensing images is one of the common operations in the field of Geosciences and Disaster Management. The remote sensing images contains huge amount of information, so there is a requirement for efficient and fast classification techniques. The major concerns issues in classification technique are to classify data into appropriate classes and to minimize the overall processing cost. In this work, the authors have presented a fast data classification technique that provides effective classify data for the remote sensing images. The classification has carried out in two phases, in the first phase the initial screening was done to select the significant points from the dataset based on the property of constructed histogram from the remote sensing images and in the subsequent phase, the classification was carried out on the reduced data set obtained from the first phase. This hierarchical approach reduces the overall computation cost. The scheme has been tested on some remote sensing images. The proposed scheme gives better performance than the results obtained from the ERDAS software, which is one of the popular software tools to analyze the remote sensing images in the field of Geosciences. After development of scheme some other parameter will be tested on data so that efficiency of proposed scheme will be measured on with respect to available tools and algorithms for classification. Other parameters are accuracy checking and confusion matrices, kappa statistics and computational time. Some other parameters like peak signal to noise ratio and mean square error also have been calculated to compare classified results with original data in each and every band. Accuracy calculated by NAÏVE BAYES classier, which is a probability based classification method.

Keywords - Feature Extraction, Remote Sensing, Hierarchical Classification.

I. INTRODUCTION

Classification of remotely sensed data is used to assign corresponding level with respect to group with homogeneous characteristics, with aim of discriminating multiple objects within the image.

The level is called class. Classification will be performed on the base of spectrally defined features such as density, texture etc. in the feature space. It can be said that classification divides the feature space into several classes based on decision rule. In many cases classification is undertaken using computer, with use of mathematical techniques. Classification will be made according to following procedure

1.1 process of classification

Depending upon objective and characteristics of image classes should be defined.

Selection of features

Sampling of training data

Estimation of the universal statistics

Decision rule for classification

Verification of results

Estimation of population statistics

1.2 Clustering

Clustering is group of data with similar characteristics. It is divided into hierarchical and non hierarchical clustering as follows

- Hierarchical clustering
  The similarity of cluster is measured distance measure. The minimum distance between the clusters will give merged cluster after repeated procedure from starting point of pixel-wise cluster to a final limited cluster. Distance will be calculated via following methods
  1. Nearest neighbor method
  2. Farthest neighbor method
  3. Centroid method
  4. Group average methods
  5. Ward method
- Non hierarchical clustering

At the initial stage arbitrary number of clusters is chosen. The member belonging to each cluster will be checked by selected parameters or distance and relocated in the cluster having higher separability. The ISODATA method and K-MEANS clustering are example of non hierarchical clustering.

The ISODATA procedure composed of following process

1. All members are relocated into closest cluster by computing distance between member and cluster.
2. Center of gravity of all cluster recalculated until convergence
3. If the number of clusters are within specified number and distance between them, and distance between the clusters met prescribed threshold, the clustering is considered complete.

II. LITERATURE REVIEW

Remote sensing classification either supervised or unsupervised, have numerous applications in disaster management and in land cover and land use applications. Recent advances are performed on color based approach because if color as a unit has been taken for feature extraction then lots of features can be classified on the user requirement (Jain, A., and Vailaya, A., 1996). Lots of software tools like: ERDAS, ENVI, GEOMATICA, MATLAB are present to classify the remote sensing images on the color basis. MATLAB is one of the premier software for such kind of analysis. K-means clustering is known for its simplicity to implement, so K-means clustering has been applied in segmentation as an important step to group the clusters. This algorithm suffers a problem of dynamically adjusted values of K, that is why result differs every time (Elona, T., and Koivistoinen, H., 2005). Since K-means clustering can be performed in any color space like: RGB, HSV, YcbCr, YUV, CIE L*a*b*, CIE L*u*v* color space. RGB, YCbCr, and YUV color spaces are commonly used in raw data. These color spaces do not provide more visually interpretable images (Chen, Tse-Wei et al. 2008). HSV color space have better results for image segmentation than RGB color space (Sural, S., Quian, G., and Pramanik, S., 2002). HSV color space has three elements that are Heu, saturation and value. Heu is color type in that particular area of image. Saturation is measurement of intensity of a particular color. Value represents the intensity in the image. It is a cylindrical or conical color space in which Heu represents angle of a pixel from red color, saturation is distance from central axis, and value is distance from bottom. Bottom considered as completely black and top is completely white space. Red color is at angle 0 degree, green is at angle 60 degree and blue is at angling 120 degree. All other colors presents in between of these three colors. Saturation increases from central axis to surface of cylinder or cone. Value increases from bottom to top. HSV color space is fundamentally different from widely known RGB color space because it differentiates intensity from the color chrome. Heu is also more important than saturation because it visually separate colors. So Heu and intensity has been taken as dominant feature (Kaufman, L., and Rousseeuw, P.J., 1990). Based on observation, a method that combine histogram bins for clustering is developed in this paper.

2.1 Preliminaries

In the present work, a clustering based algorithm has been used to classify data into clusters. K-means clustering algorithm has been used. Some other algorithm like C-means and fuzzy clustering algorithm may also be used but k-means algorithm is based on Euclidean distance so it is easier to implement.

- Remote Sensing image

  Remote sensing data is captured by satellite sensor. Satellite image gives information about earth behavior and geological changes. On the earth surface there may be number of features i.e. ground, vegetation, water bodies, seacoast, forest, and so on. Classification of images give required area of selected features from the image. This classified result can be used in land use, land cover, disaster management applications and various other geological applications.

- K-means Algorithm

  K-means clustering is conceptually very simple and easy to implement. This algorithm decomposes data into K clusters and provides centroid for each cluster center. Any data belongs to a cluster if the square Euclidean distance between corresponding cluster center is minimum with respect to other cluster centers. The algorithm runs iteratively and keep centroid track until termination criterion is achieved. The algorithm steps are given below.

- Algorithm (K-means clustering)

Step 1: Take N×D dimension training data where N is number of points in data set and D is dimension of each point.

Step 2: Pick K(K<<N) number of data points at random. These selected data points will be considered as initial cluster point or centroid.

Step 3: Map data point from training data set with respect to selected initial points from step 2. Mapping is done based on computation of Minimum Square Euclidean distance.

Step 4: Compute centroid from each cluster and find overall distortion.

Step 5: if difference in over all distortion between the last two successive iterations is smaller than some threshold value then stop the iterations otherwise go back to step 2 and update the initial center point with current center point or centroid from each cluster.

2.2 Research objective

1. Development of a computer based scheme for the efficient classification of remote sensing images.
2. Classify the data from other classification schemes and compare the classification results.
3. Finding out accuracy of both results.
4. Comparison of computational time of developed scheme and previously present scheme.
5. Results of developed scheme should be more interpretable and number of classes should be maximum.
6. For every classified result computation of kappa statistics and confusion matrix for getting inter rate agreement of classes in classified result
7. Calculation of peak signal to noise ratio and mean square error for quality assessment of classified image with respect to image in every band

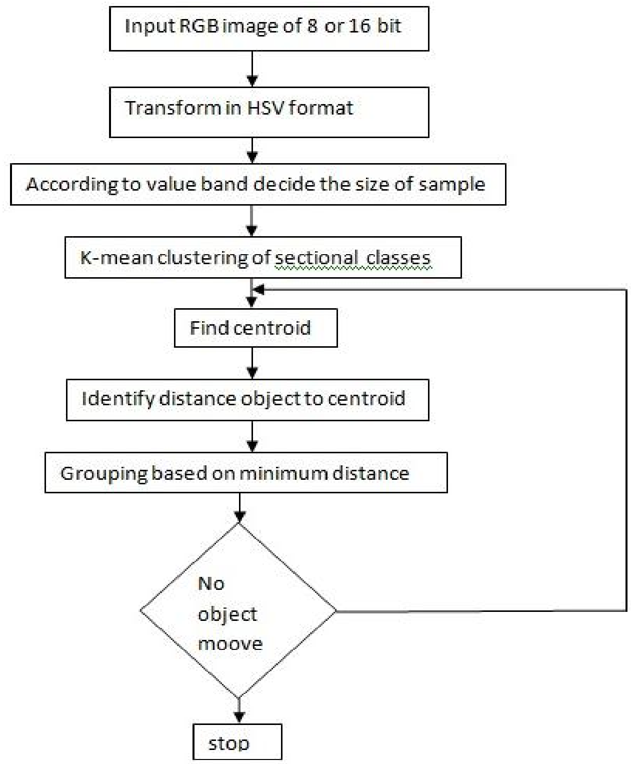


Figure 2.1 Flow chart of K-means clustering algorithm

III. METHODOLOGY

In figure 3.1, a 24 bit Lena image has been taken. In Heu band only some black and white portion shows which is giving very less information about original image.

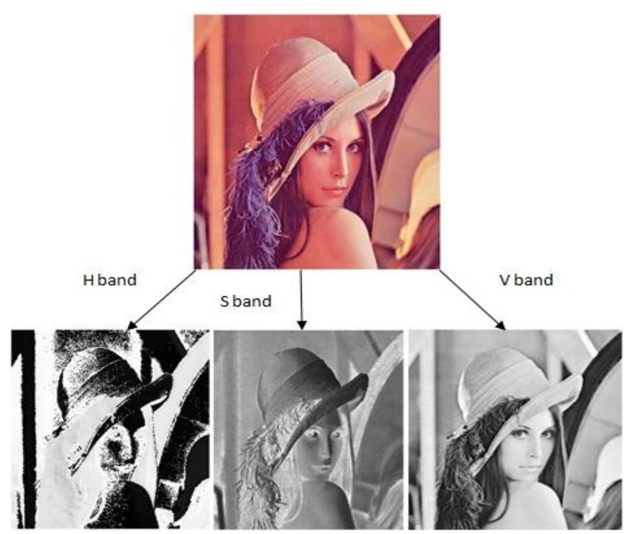


Figure 3.1 24 bit (512×512) Image in HSV color space

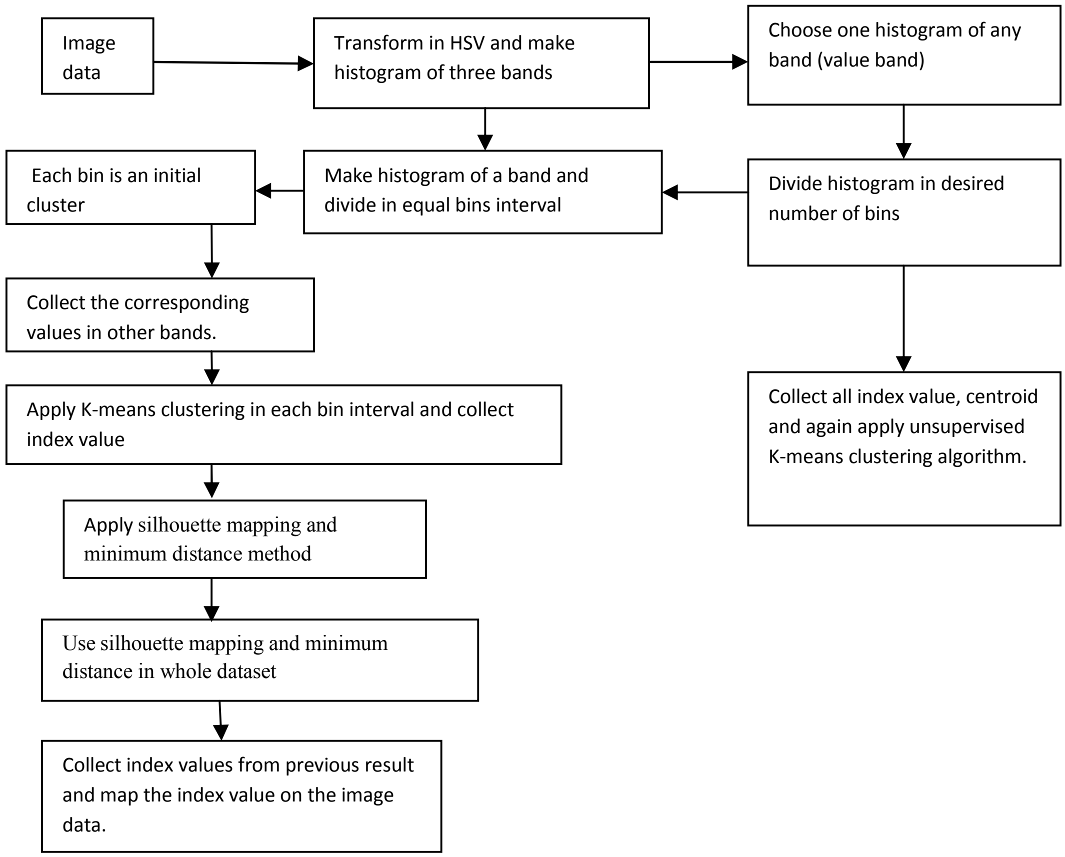


Figure 3.2 Methodology of proposed scheme

Saturation band is giving some more information. But the value band contains most of the information from original image. In present work value band has been divided into number of small clusters. Other band values have been extracted according to value band for clustering. In Figure 3.2 block diagram shows flow of algorithm. For testing of scheme quickbird satellite has been used.

Figure 3.2 shows flow of algorithm by which optimum number of classes can be determined. Above scheme is giving variable results, so some manual testing has been performed based on local minima and silhouette then refinement of the scheme has been done by cluster centroid distance formula which is basically Euclidean distance between two points in multi-dimensional space.

IV. EXPERIMENTAL RESULTS

A 24 bit RGB image of sunder ban of either has been used. 24 bit image contains Intensity values vary from 0-255. Transform RGB in HSV color format. Separate the H, S and V band from image.

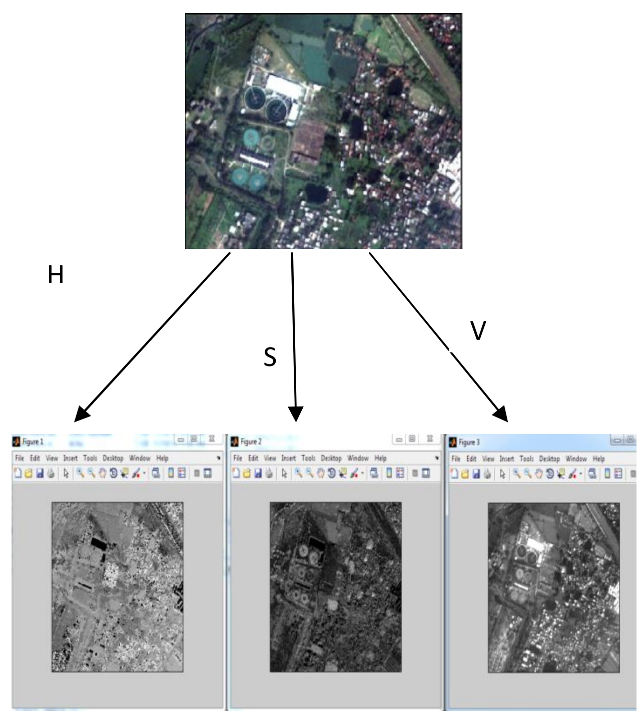


Figure 4.1 24 bit (267×275) sunder ban urban HSV image

Value band has divided in 4 intervals of 64 bins from the histogram.

Remotely sensed sunder ban image contains road, water bodies, buildings, vegetation, ground etc. Sometimes saturation band also gives good information since it is color nature of the image. Intensity is variation of grey values between 0to1. Hue band values are angle of any color from central axis so this band is not so useful to detect information.

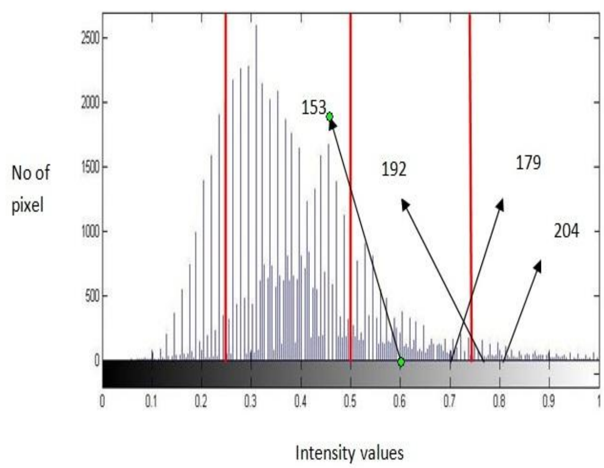


Figure 4.2Variation of intensity values in the image data

Histogram is divided in 64 bin intervals. For 255 values there is 4 sections. K-mean clustering has been performed in intervals of 0-64, 65-128,129-192,193-255. After getting optimal classes in all the intervals, k-mean clustering algorithm has been performed in all the clusters. By this way computational time has reduced and classes are optimal in numbers. In histogram after 192 spectral values, the distribution of pixels is very flat. In k-mean clustering it is creating an empty cluster .So the interval from 140 to 255 has been be adjusted. The value band has been divided in four clusters of 64 bins each. Data is of eight bit so, four clusters of 64 bins has been created.

First bin interval (0-64):

In first bin interval number of clusters started from a big and random value which is 30. For selection of classes the silhouettes has been formed. Silhouettes which are better separated and having peak value at similar levels have been selected as optimum classes. In first bin interval two silhouettes sets are very close.

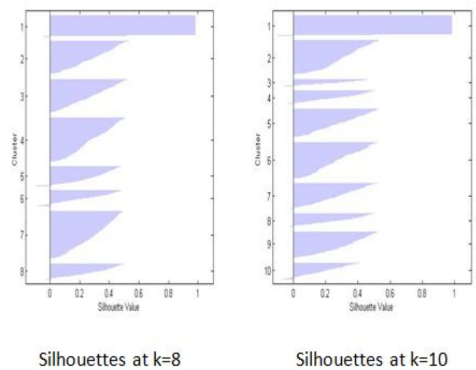


Figure 4.3 Silhouettes of first cluster

8 clusters silhouettes are better separated than 10 clusters silhouettes in first bin interval. If sum of distance is equal for any two cases, then the case having more number of classes has been considered.

iter	phase	num	sum
1	1	12302	2203.41
2	1	2018	1864.65
3	1	1851	1458.41
4	1	1180	1420.59
5	1	907	1394.85
6	1	566	1382.79
7	1	607	1369.42
8	1	478	1360.41
9	1	460	1349.41
10	1	364	1342
11	1	306	1338.95
12	1	137	1338.18
13	1	80	1337.93
14	1	130	1337.19
15	1	112	1336.96
16	1	34	1336.94
17	2	0	1336.94

17 iterations, total sum of distances = 1336.94

Figure 4.4.1 Sum of cluster distance at k=8 is 1336.94

10	1	514	1226.27
11	1	208	1225.17
12	1	235	1223.36
13	1	202	1221.54
14	1	128	1220.74
15	1	120	1219.17
16	1	285	1216.62
17	1	127	1216.08
18	1	133	1215.61
19	1	144	1215.14
20	1	90	1214.77
21	1	87	1214.61
22	1	104	1214.24
23	1	51	1214.14
24	1	84	1213.92
25	1	43	1213.84
26	1	156	1212.21
27	1	128	1211.08
28	1	96	1210.31
29	1	60	1209.89
30	1	79	1209.51
31	1	63	1209.33
32	1	43	1209.27
33	1	40	1209.23
34	1	16	1209.2
35	1	24	1209.18
36	1	1	1209.18
37	2	0	1209.18

37 iterations, total sum of distances = 1209.18

Figure 4.4.2 Sum of cluster distance at k=10 i 1209.18

- *Other less favoured cases in first bin interval:*

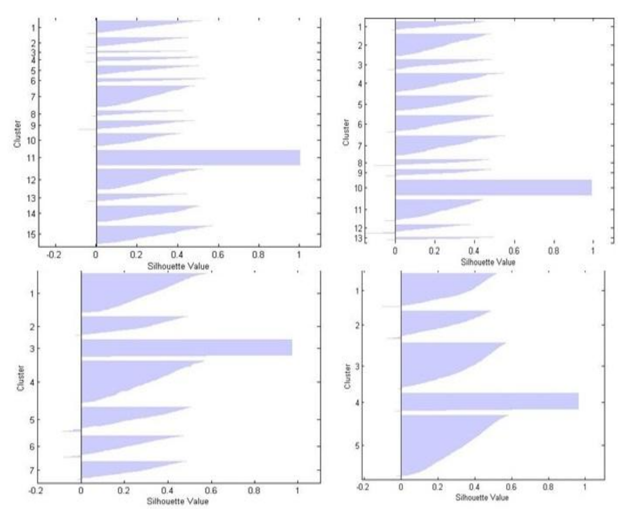
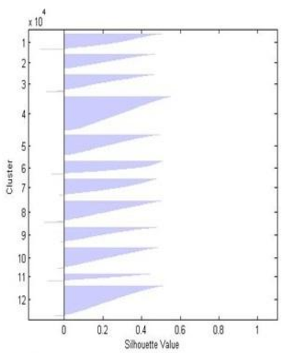


Figure 4.5 Cluster mapping at k=13, 15, 7 and 5

In bin interval having 10 clusters, Sum of distance is coming less than that of 8 clusters. Because of better separation 8 clusters in first interval has been considered. In figure 6, having k=15, 13 consistent mapping is not there. In k=7, 5 or less than 5 mapped portion is either irregular, less or having very high peak value with respect to average silhouette value.

- *Second bin interval:*
  In second bin interval the silhouettes of 8 clusters and 12 clusters are similar. In the case of 12 clusters sum of distance is less.



Silhouette in cluster 2 at k=12

```
IDX2 = kmeans(c12,12, 'dist','city', 'display','iter');
  iter  phase     num          sum
     1      1   46233      6112.55
     2      1    6091      5914.92
     3      1    3391      5844.96
     4      1    2866      5795.23
     5      1    2006      5765.12
     6      1    1546      5747.64
     7      1    1181      5737.42
     8      1     814      5732.92
     9      1     819      5728.93
    10      1     615       5725.5
    11      1     516      5723.05
    12      1     480      5721.52
    13      1     392      5720.64
    14      1     256      5720.02
    15      1     266      5719.45
    16      1      84      5719.37
    17      1      42      5719.34
    18      1      19      5719.33
    19      2       0      5719.33
19 iterations, total sum of distances = 5719.33
```

Sum of distance is 5719.33

Figure 4.6 Silhouette and sum of distance at k=12

- *Other less desired cases in second bin interval:*

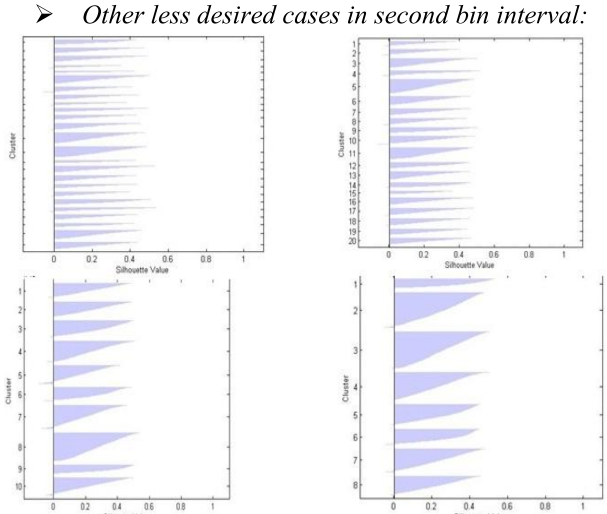


Figure 4.7 Silhouette at k=25, 20, 10 and 8

At k=25 and 20 mapping is very sharp and erratic. After moving to a lesser number of clusters, average value of silhouettes is almost equal. At k= 12 mapping is consistent and at k=10 again mapping is inconsistent. At k=8 mapping is good.

- *Third bin interval:*

At k=8, the clusters are giving best silhouettes. Sum of distance is 1630.97.

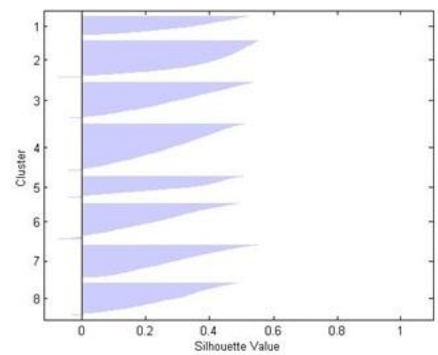
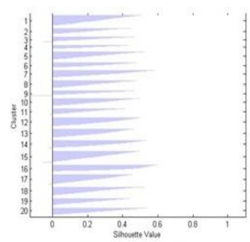


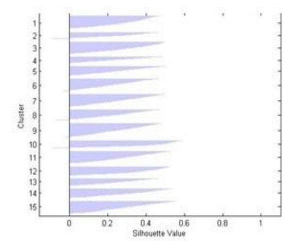
Figure 4.8 Silhouette at k= 8

Sum of distance = 1630.97.

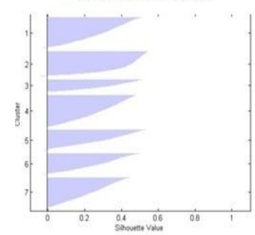
- *Other less favored cases in third bin interval:*



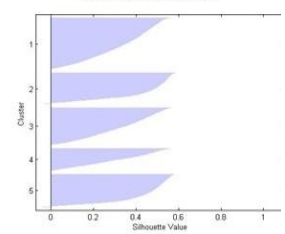
Silhouettes at k=20



Silhouettes at k=15



Silhouettes at k=7



Silhouettes at k=5

Figure 4.9 Silhouettes at k= 20,15,7 and 5

At k=20 and 15 peaks of all clusters are irregular and inconsistent. At k=8 clusters are better separated and at k= 5 and 7 they are partially separated. In third bin interval 8 clusters has been taken.

Fourth bin interval:

Optimum number of clusters are 6 and sum of distance is 1382.56.

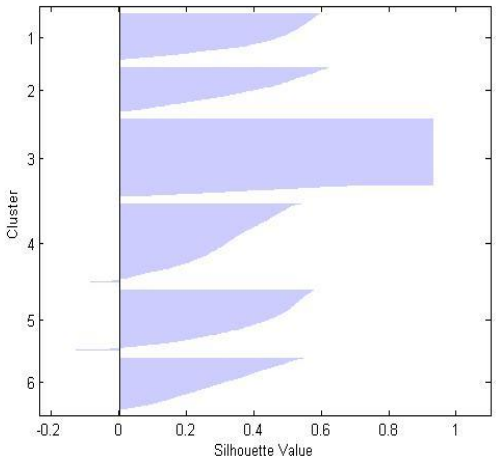
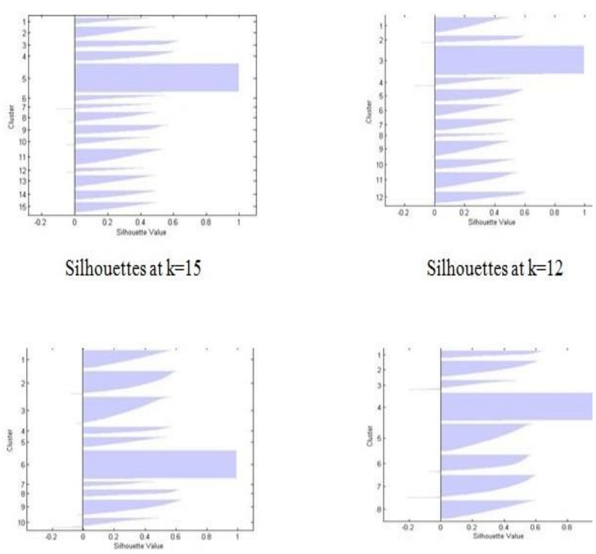


Figure 4.10 Number of clusters in Fourth bin interval

- *Other less desired cases in fourth bib interval:*



Silhouettes at k=15

Silhouettes at k=12

Silhouettes at k=10

Silhouettes at k=8

Figure 4.11 Silhouettes at k= 15, 12, 10 and 8

Cluster peaks are consistent at k=6 is better separated but at k= 15, 12, 10 clusters are not better separated. At K=8 clusters are better separated but peak values are very irregular

- *Combining the clusters of each bin interval:*

In figure 4.12, finally 6 clusters are optimum number of clusters after combining the clusters of each bin interval.

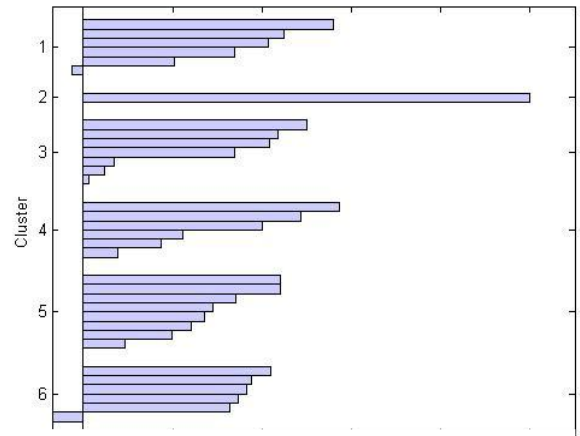
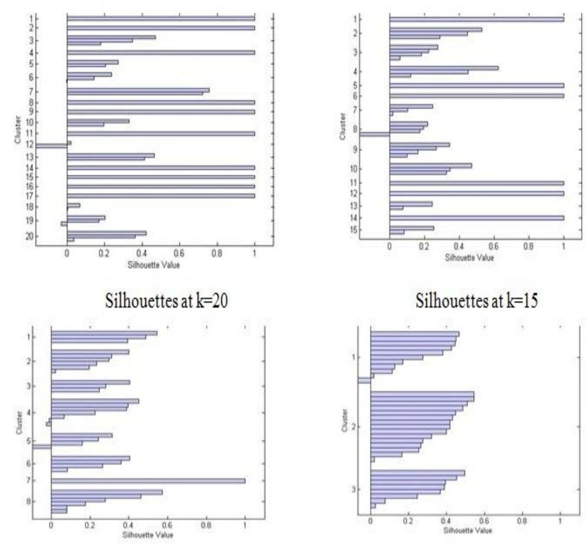


Figure 4.12 Optimum clusters in final cluster

Other less desired cases



Silhouettes at k=20

Silhouettes at k=15

Silhouettes at k=8

Silhouettes at k=3

Figure 4.13 Silhouettes at k= 20,15,8 and 3

In figure 4.13, Silhouettes at k=20, 15, 8, 3 are irregular and inseparable. So these cases have been avoided.

After getting final number of clusters, their index value has been mapped on the generated code book according to Euclidean distance. Classified Image is reconstructed from new code book.

Case1

Figure 4.14 Resultant grayscale Image after final clustering with 6 clusters

Case2

Other less favored cases when we done the clustering without considering the better separated silhouettes. Here separation of silhouettes is very important for clusters to be separated and not to be merged. Started from maximum number of clusters and then move towards lesser number of clusters some more satisfactory results can be found. Some more results have been derived in some other number of clusters given below

Figure 4.15.1 Result with 7 clusters in final Clustering

Figure 4.15.2 Result with three clusters in final clustering

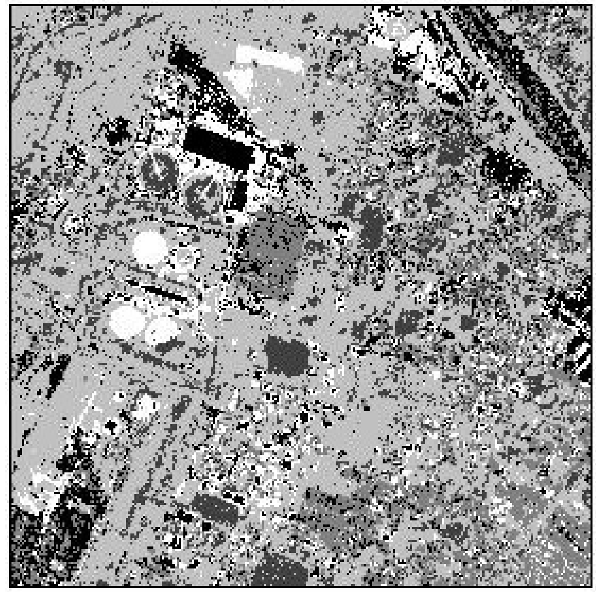
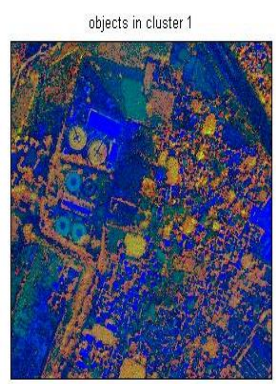


Figure 4.15.3 Result with 5 clusters in final clustering

In case 2 algorithmic constrain is not taken into account. The clustering is done with random value of cluster. In case 1, when constrain of present scheme has taken into account, more features can be identified. After using colors in clusters features are more identifiable. Different colors have been applied for different cluster.

- *Color image in different clusters*

Classified image is colored as generated from case 1. Six colors have applied in six clusters each.



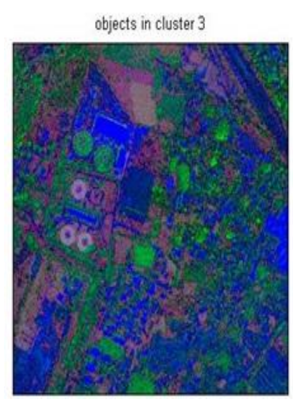


Figure 4.16.1 image in cluster 1

Figure 4.16.2 image in cluster 3

Other less desired cases

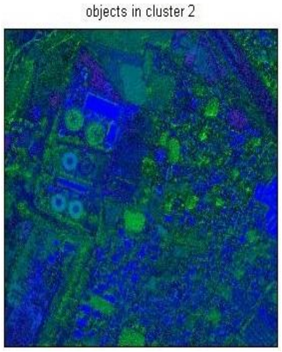


Figure 4.17.1 image in cluster 2

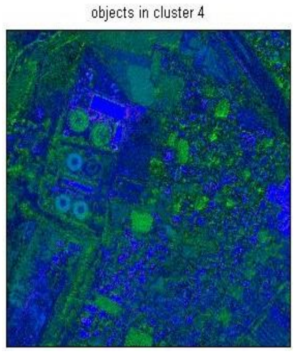


Figure 4.17.2 image in cluster 4

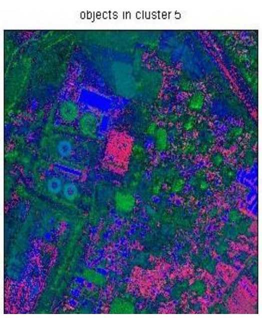


Figure 4.17.3 image in cluster 5

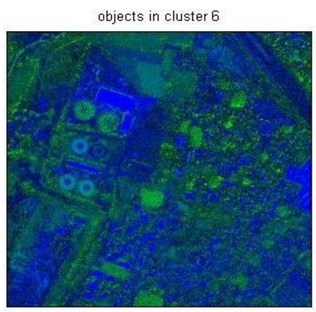


Figure 4.17.4 image in cluster 6

According to figure 4.16.1 and figure 4.16.2 maximum features are extracted in cluster 1 and in cluster 3. In these two clusters, the features can be clearly identified but in other clusters features are not completely visible. In figure 4.17.1 and 4.17.2 ground is completely blue so roads and ground are not completely distinguished. In figure 4.17.3 roads, ground, vegetation and water bodies are also identifiable. Result of unsupervised classification with same number of classes and iterations has also been derived and shown in figure 4.18 from ISODATA algorithm.

4.1 Comparison of results from other classification methods

- K-nearest Neibour classifier

Each of the characteristics in training set has been considered as a different dimension in some space. The similarity of two points has been considered to be the distance between them in this space under some metric. The way in which the algorithm decides which of the points from the training set are similar enough to be considered when choosing the class to predict for a new observation is to pick the k closest data points to the new observation, and to take the most common class among these. This is why it is called the k Nearest Neighbors algorithm. The algorithm can be summarized as

Figure 4.18 Unsupervised classifications with 6 classes result derived by ISODATA Algorithm

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.

It is a supervised algorithm so initial training set should be given carefully. In this current data we have used training set as training=[0 0 0;0 1 0;1 0 0;1 0 1;2 0 3;7 2 0]. Selection of these training set is random. Numbers of classes are six. After application of this algorithm we got following result.

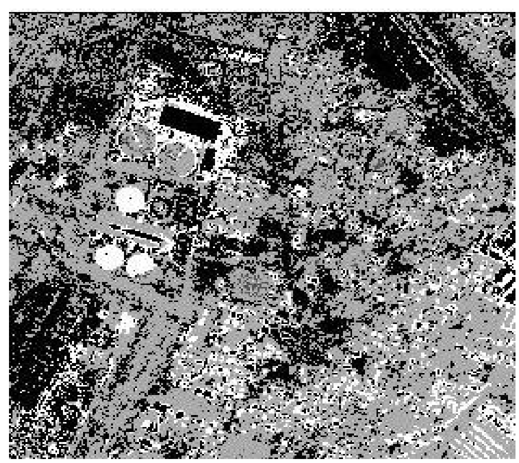


Figure 4.19 Classification result by K-NEAREST NEIBOUR algorithm

The computational time is calculated, which is coming out to be 4.068579 seconds.

Computational time in proposed scheme's result was 2.586537 seconds.

Table 4.2 Confusion matrix generated by KNN classified result

	Actual values			
Predicted values	5524	0	92	48
	3	92	37	0
	2	56	7060	35
	240	0	25	1471

Total number of correctly classified pixels = 14147.
Number of wrongly classified pixels=538.
Total number of pixels=14685.
Accuracy of classification=96.3364%
Accuracy is better than the proposed scheme.
Accuracy has been calculated by NAÏVE BAYES classifier

- Support vector machine classification

In machine learning, support vector machines (SVMs, also support vector networks) are supervised learning models with associated learning algorithms that analyze data and recognize patterns, used for classification and regression analysis. SVM training algorithm builds a model that assigns new examples into one category or the other, making it a non-probabilistic binary linear classifier. In addition to performing linear classification, SVMs can efficiently perform a non-linear classification using what is called the kernel trick, implicitly mapping their inputs into high-dimensional feature spaces. Support vector machine (SVM) can be used when data has exactly two classes. An SVM classifies data by finding the best hyper plane that separates all data points of one class from those of the other class. The *best* hyper plane for an SVM is the one with the largest margin between the two classes. Margin is the maximal width of the slab parallel to the hyper plane that has no interior data points. The support vectors are the data points that are closest to the separating hyper plane; these points are on the boundary of the slab. The following figure illustrates these definitions, with + indicating data points of type 1 and – indicating data points of type –1.

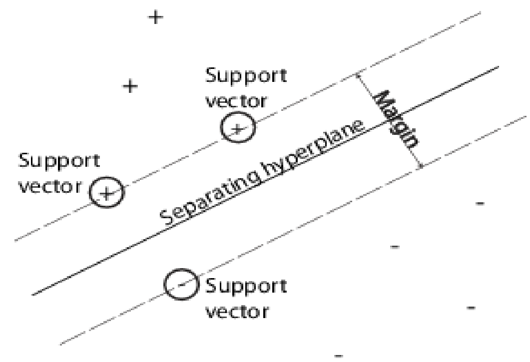


Figure 4.20 Hyper plane, margin selection between classes and support vectors

Nonlinear Transformation with Kernels
Some binary classification problems do not have a simple hyper plane as a useful separating criterion. For those problems, there is a variant of the mathematical approach that retains nearly all the simplicity of an SVM separating hyper plane.
This approach uses these results from the theory of reproducing kernels:
There is a class of functions $K(x, y)$ with the following property. There is a linear space S and a function φ mapping x to S such that
$K(x, y) = \langle\varphi(x), \varphi(y)\rangle$.
The dot product takes place in the space S.
This class of functions includes:

- Polynomials: For some positive integer d,
  $K(x, y) = (1 + \langle x, y\rangle)^d$.
- Radial basis function (Gaussian): For some positive number σ,
  $K(x, y) = \exp(-\langle(x-y), (x-y)\rangle/(2\sigma^2))$.
- Multilayer PERCEPTRON (neural network): For a positive number p_1 and a negative number p_2,

$K(x, y) = \tanh(p1\langle x,y\rangle + p2)$.
Using Support Vector Machines
In support vector machine classification method, classification performed only for two classes. A process of developing SVM algorithm for more than two classes has been developed. Although SVM can classify linearly separable data more efficiently but some kernel method are also available to first transform data from linear to kernel space and try to draw a hyper plane between them. Support vectors are touching points on margin of hyper plane. Multi class SVM approach is one verses the rest approach. It is a supervised method in which first data has been trained according to selected classes and then classified. For that we have SVMTRAIN and SVMCLASSIFY function.

Figure 4.21 Classification result by SVM method

Result is obtained by taking supervised Training Set= [0 0 0;.4103 .3197 .4784;0 .1206 .5529;0 0 1;.3103 .2057 .5529;.6061 .4925 .2627]. Numbers of misclassified pixels are the most in this method because of wrong training set. Computational time is coming out to be 256.814851 seconds.

Table 4.3 Confusion matrix of SVM classified data

	Actual values					
Predicted values	217	0	0	0	0	15
	0	3	0	0	0	4
	0	0	25	0	0	12
	0	0	0	439	0	8
	0	0	0	4	791	59
	0	4	3	14	105	12982

Total number of rightly classified pixels=14457.
Total number of wrongly classified pixels=228.
Total number of pixels=14685.
Accuracy of classification=98.4474%.
So accuracy is high in SVM classification but computational time is also high.

- Neural network classifier

The term "neural network" usually refers to models employed in artificial intelligence. Neural network models which emulate the central nervous system are part of theoretical neuroscience and computational neuroscience. Neural networks are typically organized in layers. Layers are made up of a number of interconnected 'nodes' which contain an 'activation function'. Patterns are presented to the network via the 'input layer', which communicates to one or more 'hidden layers' where the actual processing is done via a system of weighted 'connections'. In back propagation, 'learning' is a supervised process that occurs with each cycle or 'epoch' (i.e. each time the network is presented with a new input pattern) through a forward activation flow of outputs, and the backwards error propagation of weight adjustments. More simply, when a neural network is initially presented with a pattern it makes a random 'guess' as to what it might be. It then sees how far its answer was from the actual one and makes an appropriate adjustment to its connection weights

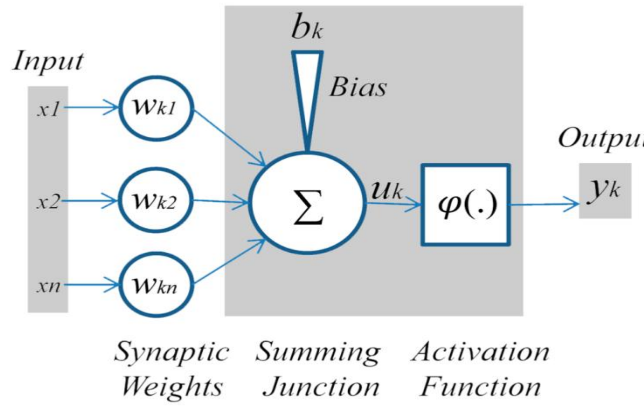


Figure 4.22 Neural networks architecture

In above figure x1, x2, x3,… are input vectors.w1,w2,w3,… are weights corresponding to them. bk are biases and ϕ is activation function to decide output classes.

P=[0 0 0;0 1 0;1 0 0;1 0 1;2 0 3;7 2 0] has been used as training set to train the neural network and then neural network has simulated for code book data. Result of this classification is good because it is not drawing any boundary between classes.

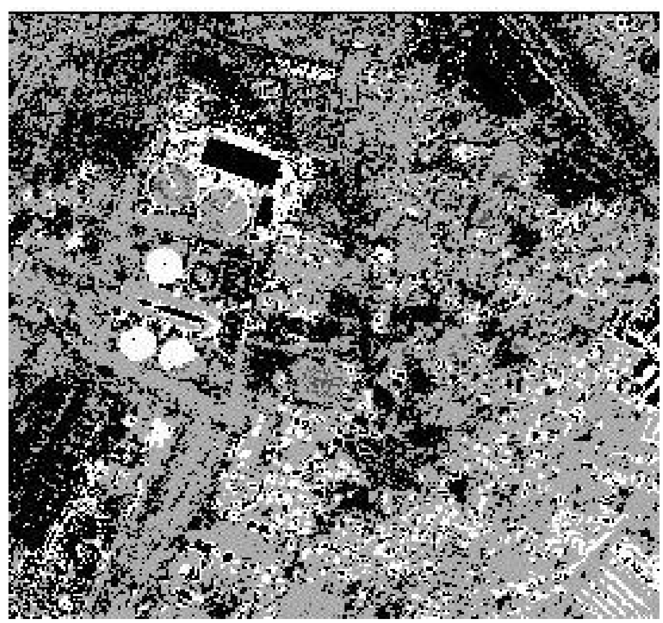


Figure 4.23 Classification results by ANN classifier

Computational time is coming out to be .487458 seconds.

Table 4.4 Confusion matrix of NEURAL NETWORKS classifier

	Predicted values			
Actual values	5523	0	95	46
	8	91	31	2
	10	64	7033	46
	228	1	26	1481

Here total number of rightly classified pixels=14128.
Total number of wrongly classified pixels=557.
Total number of pixels=14685.
Accuracy of classification=96.207%.
Here accuracy is not as good as in proposed scheme.

- Fuzzy logic based C-MEANS classification algorithm

In fuzzy clustering, every point has a degree of belonging to class, as in fuzzy logic, rather than to completely belonging to one class. Thus, points on the edge of a cluster may be having lesser degree than the points inside the cluster. Any point x has a set of coefficients giving the degree of being in the kth cluster $w_k(x)$. With fuzzy c-means, the centroid of a cluster is the mean of all points, weighted by their degree of belonging to the cluster.

$$c_k = \frac{\sum_x w_k(x)^m x}{\sum_x w_k(x)^m}.$$

The degree of belonging, $w_k(x)$, is related inversely to the distance from x to the cluster center as calculated on the previous pass. It also depends on a parameter m that controls how much weight is given to the closest center. The fuzzy c-means algorithm is very similar to the k-means algorithm

Fuzzy logic is a form of many-valued logic that deals with approximate, rather than fixed and exact reasoning. Compared to traditional binary logic (where variables may take on true or false values), fuzzy logic variables may have a truth value that ranges in degree between 0 and 1. Fuzzy logic has been extended to handle the concept of partial truth, where the truth value may range between completely true and completely false. It is like K-MEANS clustering but only difference is membership grade associated with it

The algorithm minimizes intra-cluster variance as well, but has the same problems as k-means; the minimum is a local minimum, and the results depend on the initial choice of weights.

Figure 4.24 Result of FUZZY C-MEANS classification

This result is coming out to be fine and it is as good as result coming out from proposed scheme but computational time is very high here. Computational time is 19.992756 seconds. In proposed scheme computational time was only 2.586537 seconds. Initial centroid is taken randomly so result may differ each time.

Table 4.5 Confusion matrix by FUZZY classification of data

	Predicted values					
Actual values	2900	99	13	0	136	0
	81	2819	52	0	107	2
	11	46	2406	16	0	0
	0	0	2	561	0	85
	10	47	4	0	3417	11
	0	0	0	15	122	172

Total number of rightly classified pixels=13826.
Total number of wrongly classified pixels=859.

Total number of pixels=14685.
Classification Accuracy =94.1505%.
Here accuracy is too low as compare to other classifier and computational time is also high. This is also less efficient that proposed scheme.

4.2 comparison chart of different classification result
Accuracy comparison of classifiers

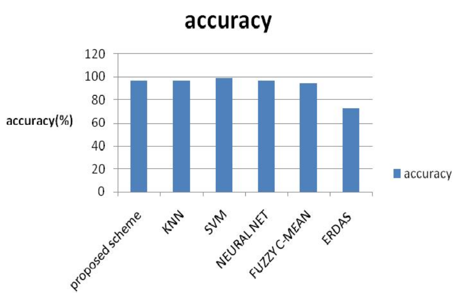


Figure 4.25 Accuracy comparison of proposed scheme with others

Computational time comparison of classifiers

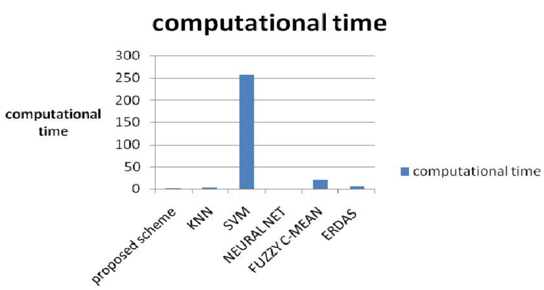


Figure 4.26 Computational time comparison of proposed scheme with other classifier

Comparison of Kappa statistics

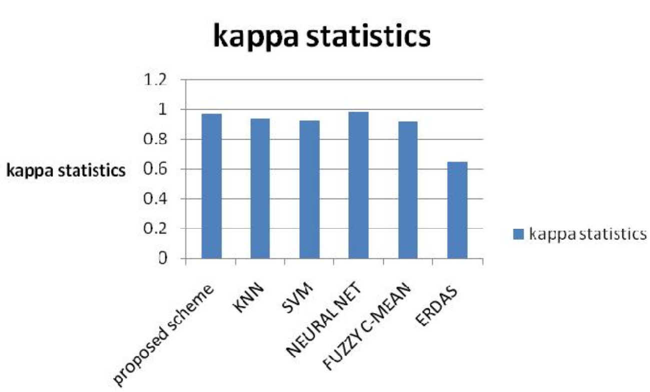


Figure 4.27 Kappa statistics comparison of proposed scheme with others

Accuracy of SVM classifier is higher but computational time is also high. Kappa coefficient is also lesser than proposed scheme, neural network and fuzzy classifier. In proposed algorithm all results are optimum such as accuracy and computational time. Neural network classifier is also giving almost similar results as proposed scheme. Fuzzy classifier is giving very good results in terms of accuracy and feature extraction but computational time is also very high. In fuzzy classifier computational time is high. In KNN classifier all parameters are also optimum.

V. Conclusion and Future scope

By comparing the results obtain from different classifiers it is deduced that in term of accuracy neural network, support vector machine, KNN and proposed scheme are giving almost similar accuracy. SVM classifier and fuzzy classifier are not giving desired computational time. When dataset is large then computational time will be higher. KNN and fuzzy classifiers are also giving similar results. Neural network classifiers have good accuracy and less computational time.

Table 5.1 Comparison of classification parameter of different classifiers

	Accuracy	Computational time	Kappa value	Features extraction priority
By Proposed scheme	95.95%	2.586537 seconds	0.9674	1
By KNN Classifier	96.34%	4.068579 seconds	0.9390	3
By SVM Classifier	98.45%	256.81485 seconds	0.9220	5
By NEURAL NET. classifier	96.21%	0.487458 seconds	0.9370	4
By FUZZY C-MEAN classifier	94.15%	19.992756 seconds	0.9430	2
By ERDAS classifier	72.00%	6.000000 seconds	0.6432	2

References

[1] Abraham A., Jain L. and Kacprzyk, , 2002. "Studies in Fuzziness and Soft Computing, Springer Verlag" Germany, Chapter 1, pp. 1-35.
[2] Andrew Moore, 1998."K - means and Hierarchical Clustering", Pearson education, pp 115-186.
[3] Assefa M., and Jonathan D. September 2002." Photogrammetric Engineering & Remote Sensing Vol. 68, No. 9, pp. 905-911".
[4] Carson, C., et al june 1999."Blobworld:A system for region-based image indexing and retrieval "proc. Third Int. conf. on visual information system.
[5] Chen, Tse-Wei et al. 2008."Fast image segmentation based on K-means clustering with histograms in HSV color space".

[6] Chia-Feng Juang, and Chun-I Lee,(2007) “A Fuzzified Neural Fuzzy Inference Network for Handling Both Linguistic and numerical information Simultaneously”, Neuro Computing, doi:10.1016/j.neucom.

[7] Chiang, M.M., & Mirkin, B. (2007). “Experiments for the number of clusters in K-means”,

[8] Springer-Verlay Berlin Heidelberg.

[9] Dau, T., Kollmeier, B., Kohlrausch, A. ,1997.”Modeling auditory processing of amplitude modulation”,I+II.J. Acoust. Soc. Am. 102 (5), 2892±2919.

[10] Derleth, R.P.,1999.” Temporal and compressive properties of the normal and impaired auditory system”. Ph.D. thesis,Universit at Oldenburg.

[11] Elona, T., and Koivistoinen, H.,”On autonomous K-means clustering ”, in proceeding of 15th international symposium on methodologies for intelligent systems. May 2005.pp.228-236.

[12] Francis, I.F., Anderson, T.R.1997.“Binaural phoneme recognition using the auditory image model and cross-correlation”. In: Proc. ICASSP pp. 1231±1234.

[13] Gelin, P., Junqua, J.-C., 1999.“Techniques for robust speech recognition in the car environment”. In: Proc. Eurospeech , Budapest, Hungary, Vol. 6, pp. 2483±2486.

[14] Grabianowski, Ed, 2006. "How Speech Recognition Works.", 10 November.

[15] Hitoshi Iyatom, and Masafumi hagiwara, 2004. “Adaptive Fuzzy Inference Neural network”, Pattern Recognition, doi:10.1016/j.patcog.2004.04.003.

[16] Hurley, C. B. (2004).“Clustering visualizations of multidimensional data”, Journal of Computational and Graphical Statistics, 13(4), 788-806.

[17] Ingrouille, M., (2001). Sepals. “In *Encyclopedia of life sciences* (Vol. 10, p. 1038) Wiley

[18] InterScience.” New York, NY: John Wiley & Sons, Inc.

[19] Inselberg, A, & Dimsdale, B. (1990). “Parallel coordinates: A tool for visualizing multi-,

[20] dimensional geometry”, *IEEE Computer Society*, 1990.

[21] Jain, A., and Vailaya A., 1996. “Image Retrieval using Color and Shape, Pattern Recognition”, vol. 29, no. 8, pp. 1233-1244.

[22] James Buckley and Eslami, 2002.“An introduction to Fuzzy logic and fuzzy sets”, [6] www.citeseer.org [7] www.wseas.com

[23] Jang and Sun, 1998. “Neuro Fuzzy and Soft Computing”, Prentice Hall pp 35-68.

[24] Kal-YI Huang, 2002. Photogrammetric Engineering & Remote Sensing Vol. 68, No. 1, pp. 33-40.

[25] Kasabov N,1998. Evolving Fuzzy neural networks: algorithms, applications and biological motivation, in: Yamakawa T, Matsumoto G (Eds.), Methodologies for the Conception, Design and Application of Soft Computing, World Scientific, Singapore. pp. 271–274.

[26] Kaufman L. and Rousseeuw P.J., 1990. “Finding Groups in Data: An Introduction to Cluster Analysis”, John Wiley & Sons.

[27] Kuo RJ, Chen CH, Hwang YC,2001. “An intelligent stock trading decision Support system through integration of genetic algorithm based Fuzzy neural network and Artificial neural network", Fuzzy Sets and Systems, 118(1): 21-45.

[28] Kuo RJ, Lee LC, Lee CF,1998. “Intelligent stock decision support system. Through artificial neural networks and Fuzzy Delphi", Comput. Intell. Finance, 6: 24-34.

[29] Le CY,1985. "Une procedure dapprentissage pour reseau a seuil assymetrique", Cognitive, 85: 599-604.

[30] Lee CC,1990. “Fuzzy logic in control systems: Fuzzy logic controller", IEEE Trans. Systems Man Cybernet Part I, 20(2): 404-418.

[31] Ma, W.Y, and Manjunath, B.,1997.”Ne Tra: A text book for Navigation large image database”.Proc. SPIE Int. Conf. on image processing,pp.568-571.

[32] Mirnaser Mirbagheri,2008. Analysis Factors influencing on the Economical development, Physic –Mathematical. Sci Series, BDU, 4: 90-96.

[33] Mirnaser Mirbagheri, Tagiev N,2009. Reviewing and Analyzing the Economic growth and Economic structure of Iran, Transactions of National Academy of Sciences of Azerbaijan series of Economic, 1: 126-130.

[34] Mirbagheri 929 Moody J, Levin U, Rehfuss S,1993. "Predicting the US index of industrial production", Neural Network World, 3: 791-794.

[35] N.K. Kasabov, Q. Song, DENFIS, 2002. “dynamic evolving neural-fuzzy inference system and its application for time-series prediction”, IEEE Trans.Fuzzy Syst.14–154.

[36] Ortega, M.,1998. “supporting ranked boolean similarity queries in MARS”, IEEE Trans. On knowledge and data engineering, vol.10,no. 6, pp. 905-925.

[37] Qi M,2001. "Predicting US recessions with leading indicators and neural network Models", Inter. J. Forecasting, 17: 383-401.

[38] Sadeghi MR, Afsar A, Sohrabi B,2008. Inventory Lot-sizing with supplier. Selection using hybrid intelligent algorithm,Science Direct, Applied Soft Computing, 8: 1523-1529.

[39] Smeulders, A.W.M., et al, 2000. “Content Based Image Retrieval at the End of the Early Years”, IEEE Trans. on PAMI, vol. 22, no. 12, pp. 1-32, December.

[40] Smith, J.R, and Chang, S.F.,1996.”Visual seek: A fully Automated content based Image query system” proc. ACM Multimedia conf.,Boston,MA,1996.

[41] Stamatios V. Kartalopoulous, 2003.Under- standing neural networks and fuzzy logic”, Prentice hall.

[42] Stockman, G., and Shapiro, L., 2001. “Computer Vision”, Prentice Hall.

[43] Sugar, C. A., & James, G. M. (2003). “Finding the number of clusters in a dataset: An

[44] information-theoretic approach”, *Journal of the American Statistical Association*, 98, 463, 750-762.

[45] Sural S. , Qian G. and Pramanik S., 2002. “A Histogram with Perceptually Smooth Color Transition for Image Retrieval”, Proc. Fourth Int. Conf. on CVPRIP, Durham.

[46] Swain M. and Ballard D., 1991. “Color Indexing”, Int. Journal of Computer Vision, vol.7, no. 1, pp. 11-32.

[47] Swanson NR, White H,1997a. "A model selection approach to real time. Macroeconomic forecasting using linear models and artificial neural networks", Rev. Econ. Stat., 39: 540-550.

[48] Swanson NR, White H,1997. "Forecasting economic time series using adaptive versus non adaptive and linear versus nonlinear econometric models", Inter. J. Forecasting, 13: 439-461.

[49] Tkacz G,2001. "Neural network forecasting of Canadian GDP growth ", Inter. J. Forecasting, 17: 57-69.

[50] Venalbes, W. N., Smith, D. M. et al. (2005). “*An introduction to R, notes on R: A*

[51] *programming environment for data analysis and graphics* (Version 1.9.1)”,

[52] Bristol, United Kingdom: Network Theory Limited.

[53] Wang J.Z. , Li Jia , Wiederhold Gio, 2001. “SIMPLIcity: Semantics-sensitive Integrated Matching for Picture LIbraries,” IEEE Trans. on PAMI, vol. 23, no. 9 pp.

[54] Ward, M. O., Rundensteiner, E. A., Yang, J., Doshi, P. R., & Rosario, G. (n.d.).

[55] “Interactive poster: XmdvTool: Interactive visual data exploration system for highdimensional

[56] data sets”.

[57] Yaonan Wang, , 2003. “Intelligent Information Processing Technology”, Higher Education Press, Beijing.

[58] Zhang G, Patuwo BE, Hu MY,1998. "Forecasting with artificial neural networks: the state of the art", Inter. J. Forecasting, 14: 35-62.

[59] Zheng, c., and Wang, p., 2000.” A new method for image segmentation based on intensity and heu clustering”. Proc. Of International Conference on Pattern Recognition, 3,3-8.

[60] Zurada JM,1992. Introduction to Artificial Neural Systems, PWS Pub. Co

.