Hybrid Model with Super Resolution and Decision Boundary Feature Extraction and Rule based Classification of High Resolution Data

Navjeet Kaur
M.Tech Research Scholar
Sri Guru Granth Sahib World University
Fatehgarh Sahib, Punjab, India

Simarjeet Kaur
Assistant Professor
Sri Guru Granth Sahib World University
Fatehgarh Sahib, Punjab, India

ABSTRACT
Data Mining is defined as the process of applying specific algorithms for extracting patterns from data. Data mining strategies can be grouped into Classification, Prediction, Estimation, Association Rule Mining and Clustering. Satellite images are used these days. With the help of high resolution image we can extract the features of an image. There are many types of satellites which provide accurate and high resolution images to individuals and organizations. High resolution images give detail information about rural, land, buildings, streets etc. Applications of High resolution images are defense, disaster management and medicals. Object identification is to identify an object which belongs to and feature extraction to extract features from the image like size, shape, texture and spectral information of the image. With the help of feature extraction techniques we can easily find the relevant information from the data and the data can easily be classified. In this hybrid error super resolution model and decision boundary feature extraction algorithms are used and then both the algorithms are compared. This gives the better accuracy with the help of kappa coefficient.

General Terms
Classification, Feature Extraction, Super Resolution.

Keywords
Feature extraction, Hybrid Model, Decision Boundary Feature Extraction, Spatial features, Gray level features.

1. INTRODUCTION
Due to the rapid growth of information, data mining is used these days. With the help of data mining the relevant information can be used. Data mining is defined as the process of applying specific algorithms for extracting patterns from data [1] [2]. Data Mining is of different forms like clustering, association rule mining, prediction and classification etc. Classification is a form of data analysis that extracts models describing important data classes. For example, a classification model can be build to categorize bank loan application as either safe or risky [3] [4]. Feature extraction means to extract the features of an image of high resolution data. Feature Extraction is an important stage that uses algorithms and techniques to detect and isolate various desired portions or features of a given image [10]. Feature extraction is of two types: spatial features describing the shape of the objects and grey level features describing properties related to the pixel values of the objects [5]. Hybrid Error model (HEM) [6] is a combination of statistical and structural model. In such cases statistical approach is utilize to recognize pattern primitives and structural approach is then used for the recognition of sub patterns and patterns itself. It is clear that single model used for classification doesn’t behave efficiently, so multiple methods to be combined together giving result to hybrid model. Super Resolution means increase the resolution of image also increases the details of the image with the help of this image quality also increases that helps to view the image. Discriminately informative features and discriminately redundant features can be extracted from the decision boundary itself. The discriminating informative feature vectors have a component, which is normal to the decision boundary at least at one point on the decision boundary. Further, discriminating redundant feature vectors are orthogonal to a vector normal to the decision boundary at every point on the decision boundary. Feature extraction can be performed by eliminating redundant features. Decision boundary feature matrix (DBFM) was defined in order to extract discriminately informative features and discriminately redundant features from the decision boundary. It can be shown that the rank of the Decision boundary feature matrix is the smallest dimension where the same classification accuracy can be obtained as in the original feature space [7].

2. RELATED WORK
Hough Transformation [2012] [8] is a technique which can be used to isolate features of a particular shape of an image. The desired features can be specified in parametric form, Hough transform is most commonly used for the detection of regular curves such as lines, circles, etc. Hough transform can be used in applications where a simple analytic description of a feature is not possible. The Hough transform is used in many of the applications, as manufactured part contains feature boundaries which can be described by regular curves. The advantage of the Hough transform technique is that it is tolerant of gaps in feature boundary descriptions and is relatively unaffected by image noise.

Yucheng et.al [2014] [9] Multi Resolution Techniques many classifiers based on the spectral analysis of individual pixels have their limitations, usually they produce -salt and pepper noisy results. Segmentation based on mean shift was used to gain the initial segments from remote sensing images. Vectorization method is used to generate polygons from the segmented image and features such as spectral, shape, texture etc. were extracted. This was feasible and efficient when applied to Quickbird satellite images.

F.Masulli et.al [2014] [9] Dynamic Programming model is a means of optimization through a recursive search to find a global optimum. This approach is applicable if a function can be expressed in terms of relationships between neighboring
pixels and involves a sequential decision making process. To optimize road extraction in SPOT imagery dynamic programming was used.

F.Masulli et.al [2014] [9] Model Template Matching method in this approach, a template describing the general characteristics of the feature is defined. Templates were fixed in the terms of attributes such as size, shape etc. Features were extracted by moving the template through the image and to examine the match at each location using a similarity measure to find optimal locations. The measure can include shape image constraints or external constraints.

3. METHODOLOGY
A high resolution satellite data image is taken (see figure 1). High resolution satellites are like Quickbird, IKONIS and IRIS etc. Next, the preprocessing can be done through predefined tools in Matlab. Preprocessing means remove noise, thinning and smoothness of the image. Then features can be extracted from the images. The features like color, size, shape etc. are extracted with the help of the algorithms hybrid error model with super resolution and decision boundary feature extraction. Then the rule based classification can be done. The results show better accuracy with the help of the kappa coefficient. This can be shown with the help graphs and comparison of algorithm.

Steps of Proposed work:
1. The first step is to select the dataset of high resolution satellite.
2. In the Preprocessing step remove the noise, thinning with the help of the predefined tools in the Matlab.
3. Feature extraction is next step after the preprocessing. Features can be extracted with the help of feature extraction techniques like hybrid error model with super resolution and decision boundary feature extraction.
4. After applying the techniques the image can be classified then rule based classification can be applied.
5. After applying all the steps performance can be evaluated with the help of the kappa coefficient.

4. RESULTS AND DISCUSSIONS
This (see Figure 2) describes classification after using the hybrid error model with super resolution. Different Colors describe the different regions like sand, soil, vegetation, land, water.
This pie chart (see Figure 3) after applying the hybrid error model with super resolution. This describes the percent area covered by land, soil, rural, urban, water and vegetation. The dark blue color shows the land area covered, the sky blue color shows the rural area, the sea green color shows soil area, the yellow color shows urban area covered, orange color shows water area and the dark red color shows vegetation area covered.

This (see Figure 4) describes the classification after applying the decision boundary feature extraction. Different colors show different parts land, soil, urban, rural, water and vegetation. It shows how much area covered by these.
Fig 4: After applying the DBFE

This pie chart (see Figure 5) describes the percent area covered by different regions land, rural, soil, urban, water and vegetation. The dark blue color shows the land area covered, the sky blue color shows the rural area, the sea green color shows soil area, the yellow color shows urban area covered, the orange color shows water area and the dark red color shows vegetation area covered.

Fig 5: Pie chart after applying the DBFE

5. EXPERIMENTAL RESULTS

The table (see Table 1) shows the comparison of statistical, structural and hybrid error model and this table shows that hybrid error model gives the better results than the single model. This can be shown with the help of the kappa coefficient.
Table 1 Comparison of Statistical, Structural with hybrid error model

<table>
<thead>
<tr>
<th>Methods</th>
<th>Veg.</th>
<th>Urban</th>
<th>Barren</th>
<th>Rural</th>
<th>Water</th>
<th>Soil</th>
<th>Overall</th>
<th>Kappa</th>
</tr>
</thead>
<tbody>
<tr>
<td>HYSRM</td>
<td>94.82</td>
<td>80.11</td>
<td>90.18</td>
<td>95.42</td>
<td>90.18</td>
<td>88.46</td>
<td>91.41</td>
<td>.9217</td>
</tr>
<tr>
<td>Statistical</td>
<td>87.64</td>
<td>77.12</td>
<td>81.43</td>
<td>72.67</td>
<td>82.11</td>
<td>80.06</td>
<td>82.01</td>
<td>.8360</td>
</tr>
<tr>
<td>Structural</td>
<td>88.88</td>
<td>71.31</td>
<td>81.73</td>
<td>85.24</td>
<td>72.19</td>
<td>79.76</td>
<td>86.16</td>
<td>.8292</td>
</tr>
</tbody>
</table>

The graph (see Figure 6) given below shows that the hybrid error model shows better accuracy than the statistical and structural model. The hybrid model gives better results than single model.

![Graph showing comparison](image)

Fig 6: Comparison of HYSRM, Statistical and Structural

The table (see Table 2) compares the hybrid error model with super resolution and decision boundary feature extraction with the help of kappa coefficient. The kappa coefficient gives the better results. Hybrid error model gives better accuracy than DBFE.

Table 2 Comparison of HYSRM and Decision

<table>
<thead>
<tr>
<th>Methods</th>
<th>Veg.</th>
<th>Urban</th>
<th>Barren</th>
<th>Rural</th>
<th>Water</th>
<th>Soil</th>
<th>Overall</th>
<th>Kappa</th>
</tr>
</thead>
<tbody>
<tr>
<td>HYSRM</td>
<td>94.82</td>
<td>80.11</td>
<td>90.18</td>
<td>95.42</td>
<td>90.18</td>
<td>88.46</td>
<td>91.41</td>
<td>.9217</td>
</tr>
<tr>
<td>Decision</td>
<td>81.67</td>
<td>68.12</td>
<td>78.01</td>
<td>88.28</td>
<td>79.10</td>
<td>78.18</td>
<td>80.62</td>
<td>.8192</td>
</tr>
</tbody>
</table>

The graph (see Figure 7) shows the comparison results of hybrid error model and decision boundary feature extraction with the help of kappa coefficient. It shows that hybrid error model gives better results than the decision boundary feature extraction.
6. CONCLUSIONS AND FUTURE SCOPE

Data mining scheme is developed by integrating classification process with Hybrid Super Resolution Model and Decision Boundary Feature Extraction algorithm. Here, a high resolution satellite image is used to perform classification. Classification is based on feature extracted by using Hybrid Super Resolution Model and DBFE. These algorithms provide best results for classification. Rule based classification is implemented on feature extracted from high resolution satellite data. The proposed work gives best Kappa Coefficient and better accuracy then the previous work. This also gives the better accuracy for HYSRM than the decision boundary feature extraction. A new approach for natural or man-made object extraction in urban areas from high resolution satellite data is proposed. The experimental results on real Google Earth imagery demonstrate that it is an effective way to detect objects by learning from object’s features. It is fully depend on supervised learning. The results prove that the approach effectively extracts the major objects from the Google Earth’s image and the result can be uploaded successfully on Google Earth. Work done can be used for the defense purposes as they can be helpful for the army personnel to identify and have knowledge about the defense areas beforehand.

By using this approach we are able to classify the satellite image according to different areas like water, urban, vegetation, barren and land region with different colors assigned to each feature’s pixels. Accurate results have been obtained in these images, two of which have been discussed. The algorithm shows promising results as all the features have been identified using this algorithm accurately irrespective of type of satellite image being processed.

7. FUTURE SCOPE

Further, this algorithm can be used for identification of more features in the satellite image. It can also be used in hyper spectral satellite images to further classify/identify different types of each usual object using training data set. The pixels of two or more images can be mixed to obtain a new image. The future scope of the research includes proposing certain modification to the algorithm so that the Kappa coefficient can be improved further.

8. REFERENCES


