Soil Classification and Crop Suggestion using Image Processing

T. Abimala, S. Flora Sashya and K. Sripriya
Soil Classification & Crop Suggestion based on HSV, GLCM, Gabor Wavelet Techniques and Decision Tree Classifier in Image Processing

T. Abimala 1, S. Flora Sashya 2, K. Sripriya 3
1 Assistant Professor, Dept. of ICE, 2,3 Student, Dept. of Instrumentation and Control Engineering (ICE)
St. Joseph’s College of Engineering
1 abimala.rajendran@gmail.com 2 sashyaflora@gmail.com 3 sripriya2198@gmail.com

Abstract- This paper is intended to support agriculture by classifying 7 different types of soils like Clay, Clayey Peat, Clayey Sand, Humus Clay, Peat, Sandy Clay and Silty Sand, and in suggesting suitable crops that could be grown in those particular soils using image processing. Pre-processing is done by using Low Pass filter. HSV, GLCM, Gabor Wavelet algorithms are used for feature extraction. HSV, GLCM are used to perform colour based feature extraction. Gabor filters are used to perform texture based feature extraction. The features obtained from the test image are then compared with the features obtained from the images in the dataset. Matching of image features is achieved by training the Decision Tree classifier with statistical measurements like mean, standard deviation, skew and kurtosis. Finally the soil is predicted with the help of segmented images that are given as input for simulation using Matlab R2018a and is followed by crop suggestion.

Keywords- Low pass filter, HSV, GLCM, Gabor Wavelet technique, Decision Tree classifier.

I. INTRODUCTION

Agriculture is the backbone of India. And several times it becomes difficult to classify soils in different regions of the country with required accuracy. Our project proposes the idea of classifying soil and suggesting suitable crops using image processing. Coolpix Camera is used to take images of 7 different soils. Nearly 200 images are loaded into the dataset. Input image is subjected to pre-processing, feature extraction, classification, testing, and finally the result is produced. The input image is matched with a similar image in the dataset that is image retrieval is being done here. Since Gabor filters are used in our proposed system the efficiency is pretty high. By the process of colour based and texture based feature extraction the accuracy is improved. Our project further extends a helping hand for the farmers by predicting the water absorption rate.

II. PROBLEM DEFINITION

The existing method of soil classification and crop suggestion require manual involvement, human errors, and the results are uncertain. The method is also time consuming and invasive in nature. But our proposed system overcomes all these errors because it takes into account the physical properties of soil for classification and prediction.

III. METHODOLOGY

Following are the various levels involved in image processing:
1. Low level processing
2. Medium level processing
3. High level processing

Figure 1 shows the detailed processes involved in different levels of image processing.

LOW LEVEL PROCESSING:
1. Filtration
2. Enhancement
3. Sharpening
4. Noise reduction

MEDIUM LEVEL PROCESSING:
1. Feature Extraction
2. Classification

HIGH LEVEL PROCESSING:
1. Result Identification

Figure 3.1 Image processing classification process

- Low level processing (LLP): It involves image enhancement, removes noise using Gabor filter and resizes the image.
- Medium level processing (MLP): It involves image segmentation and classification.
- High level processing (HLP): It involves image identification.
Flow Diagram:

Figure 3.2 Flow Diagram of our proposed system

Filtration:

Filtration in our proposed system is done using Low Pass filter. Low Pass filter is used to pass signals with frequency lower than the cut-off frequency and attenuates all other signals with frequencies greater than the cut-off frequency. In our proposed system Low Pass filter is used to remove unwanted components and features from the signals so as to reduce noise in the signal. Low Pass filter is also used for shade correction, even brightening and for removing artifacts.

Image Enhancement:

There are two methodologies in image enhancement:

- Frequency domain Processing (FDP)
- Special domain Processing (SDP)

Figure 4 shows the image enhancement steps

FDP: It is basically achieved by filter operation based on Fourier transformation as given in equation no I.

\[ G(u, v) = H(u, v) F(U, V) \] ..........................(I)

Where:

- \( F(u, v) \) : Fourier Transformation
- \( H(u, v) \) : Filter function
- \( G(u, v) \) : Yields

SDP: It is based on manipulation of pixel in an image and it is achieved by equation no II

\[ g(x, y) = T[f(x, y)] \] .............................................(II)

Where:

\( f(x, y) \) = input image
\( g(x, y) \) = processed image
\( T \) = operator on f, divided over neighbor f(x,y)

Enhancement can be done by using gray level transformation,
histogram processing, arithmetic logic operation and special filtering.

**Feature Extraction:**
Features are the fundamental components of an object. It is used to distinguish one object from the other. Features are also referred to as descriptors. The process of obtaining features from an object is known as description of an object. In our proposed system feature extraction is done by two methods
1. Colour based feature extraction
2. Texture based feature extraction

**1. Colour Based Feature Extraction:**
Hue Saturation Value (HSV), Gray Level Co-occurrence Matrix (GLCM) are used for Colour based feature extraction in our proposed system. The existing system used few tests including Cone Penetration Test (CPT), Vane Shear Test (VST), Standard Penetration Test (SPT), Pressure Meter Test (PMT). But our proposed techniques are used to extract necessary features so as to suggest crops and the accuracy exceeds the bar which was set by the existing tests.

**Hue Saturation Value (HSV):**
Using this model, an object with a specific color can be detected and the influence of light intensity from the outside is reduced.

**Gray Level Co-occurrence Matrix (GLCM):**
Given an image composed of pixels, each with an intensity (a specific gray level), the GLCM is a tabulation of, how often different combinations of gray levels co-occur in an image or image section. Texture feature calculations use the contents of the GLCM to give a measure of the variation in intensity.

**2. Texture Based Feature Extraction:**

**Gabor Wavelet Technique:**
These are wavelets invented by Dennis Gabor using complex functions constructed to serve as a basis for Fourier transforms in information theory applications. The important property of the wavelet is that it minimizes the product of its standard deviations in the time and frequency domain.

![Figure 3.5 Gabor filters of five scales and eight directions](image-url)

**Texture Classification:**

**Average Texture Classification Rates for RW and SW with Pre-processing**

**SEQUENCE**

<table>
<thead>
<tr>
<th>SOIL</th>
<th>LOAMY</th>
<th>CLAY</th>
<th>SANDY</th>
<th>SILTY</th>
<th>PEATY</th>
</tr>
</thead>
<tbody>
<tr>
<td>L3E3</td>
<td>100</td>
<td>91.19</td>
<td>98.201</td>
<td>94.96</td>
<td>83.114</td>
</tr>
<tr>
<td>E3L3</td>
<td>100</td>
<td>92.25</td>
<td>93.679</td>
<td>94.24</td>
<td>91.317</td>
</tr>
<tr>
<td>L3S3</td>
<td>100</td>
<td>95.84</td>
<td>93.455</td>
<td>96.2</td>
<td>94.849</td>
</tr>
<tr>
<td>E3S3</td>
<td>100</td>
<td>95.35</td>
<td>93.993</td>
<td>93.3</td>
<td>94.365</td>
</tr>
<tr>
<td>E3E3</td>
<td>100</td>
<td>93</td>
<td>93.475</td>
<td>96.31</td>
<td>92.374</td>
</tr>
<tr>
<td>E3S3</td>
<td>100</td>
<td>91.7</td>
<td>98.164</td>
<td>92.83</td>
<td>91.043</td>
</tr>
<tr>
<td>S3E3</td>
<td>100</td>
<td>91.08</td>
<td>94.303</td>
<td>92.01</td>
<td>90.589</td>
</tr>
<tr>
<td>S3S3</td>
<td>100</td>
<td>94.39</td>
<td>97.918</td>
<td>94.66</td>
<td>93.973</td>
</tr>
</tbody>
</table>

**RANDOM**

<table>
<thead>
<tr>
<th>SOIL</th>
<th>LOAMY</th>
<th>CLAY</th>
<th>SANDY</th>
<th>SILTY</th>
<th>PEATY</th>
</tr>
</thead>
<tbody>
<tr>
<td>L3E3</td>
<td>100</td>
<td>89.89</td>
<td>91.583</td>
<td>91.19</td>
<td>88.8</td>
</tr>
<tr>
<td>E3L3</td>
<td>100</td>
<td>89.55</td>
<td>91.126</td>
<td>90.91</td>
<td>88.102</td>
</tr>
<tr>
<td>L3S3</td>
<td>100</td>
<td>92.09</td>
<td>94.216</td>
<td>90.56</td>
<td>90.587</td>
</tr>
<tr>
<td>S3L3</td>
<td>100</td>
<td>91.6</td>
<td>93.161</td>
<td>90.44</td>
<td>90.843</td>
</tr>
<tr>
<td>E3E3</td>
<td>100</td>
<td>96.05</td>
<td>92.73 95</td>
<td>94.764</td>
<td></td>
</tr>
<tr>
<td>E3S3</td>
<td>100</td>
<td>93.92</td>
<td>96.735</td>
<td>95.14</td>
<td>93.244</td>
</tr>
<tr>
<td>S3E3</td>
<td>100</td>
<td>92.48</td>
<td>95.492</td>
<td>93.76</td>
<td>92.232</td>
</tr>
<tr>
<td>S3S3</td>
<td>100</td>
<td>94.84</td>
<td>96.993</td>
<td>95.27</td>
<td>94.413</td>
</tr>
</tbody>
</table>

**Average Texture Classification Rates for SW with Pre-processing Techniques**

**MEAN**

<table>
<thead>
<tr>
<th>SOIL</th>
<th>LOAMY</th>
<th>CLAY</th>
<th>SANDY</th>
<th>SILTY</th>
<th>PEATY</th>
</tr>
</thead>
<tbody>
<tr>
<td>L3E3</td>
<td>100</td>
<td>93.184</td>
<td>90.082</td>
<td>62.373</td>
<td>92.159</td>
</tr>
<tr>
<td>E3L3</td>
<td>100</td>
<td>94.623</td>
<td>91.014</td>
<td>95.238</td>
<td>89.627</td>
</tr>
<tr>
<td>L3S3</td>
<td>100</td>
<td>99.638</td>
<td>97.192</td>
<td>96.970</td>
<td>95.986</td>
</tr>
<tr>
<td>S3L3</td>
<td>100</td>
<td>96.801</td>
<td>95.829</td>
<td>96.086</td>
<td>95.776</td>
</tr>
<tr>
<td>E3E3</td>
<td>100</td>
<td>99.959</td>
<td>99.925</td>
<td>99.965</td>
<td>99.947</td>
</tr>
<tr>
<td>E3S3</td>
<td>100</td>
<td>99.996</td>
<td>99.985</td>
<td>99.988</td>
<td>99.990</td>
</tr>
<tr>
<td>S3E3</td>
<td>100</td>
<td>99.998</td>
<td>99.966</td>
<td>99.975</td>
<td>99.979</td>
</tr>
<tr>
<td>S3S3</td>
<td>100</td>
<td>99.996</td>
<td>99.995</td>
<td>99.990</td>
<td>99.990</td>
</tr>
</tbody>
</table>

**A_MEAN**

<table>
<thead>
<tr>
<th>SOIL</th>
<th>LOAMY</th>
<th>CLAY</th>
<th>SANDY</th>
<th>SILTY</th>
<th>PEATY</th>
</tr>
</thead>
<tbody>
<tr>
<td>L3E3</td>
<td>100</td>
<td>72.58</td>
<td>94.23</td>
<td>95.81</td>
<td>48.76</td>
</tr>
<tr>
<td>E3L3</td>
<td>100</td>
<td>71.62</td>
<td>94.71</td>
<td>91.18</td>
<td>66.34</td>
</tr>
</tbody>
</table>
### Average Texture Classification Rates for RW with Pre-processing Techniques

<table>
<thead>
<tr>
<th></th>
<th>SOIL LOAMY CLAY SANDY SILTY PEATY</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>MEAN</strong></td>
<td></td>
</tr>
<tr>
<td>L3S3</td>
<td>100</td>
</tr>
<tr>
<td>E3L3</td>
<td>100</td>
</tr>
<tr>
<td>S3L3</td>
<td>100</td>
</tr>
<tr>
<td>E3S3</td>
<td>100</td>
</tr>
<tr>
<td>S3S3</td>
<td>100</td>
</tr>
<tr>
<td><strong>AVG</strong></td>
<td>100</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>SOIL LOAMY CLAY SANDY SILTY PEATY</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>SKED</strong></td>
<td></td>
</tr>
<tr>
<td>L3E3</td>
<td>100</td>
</tr>
<tr>
<td>E3L3</td>
<td>100</td>
</tr>
<tr>
<td>L3S3</td>
<td>100</td>
</tr>
<tr>
<td>S3L3</td>
<td>100</td>
</tr>
<tr>
<td>E3E3</td>
<td>100</td>
</tr>
<tr>
<td>E3S3</td>
<td>100</td>
</tr>
<tr>
<td>S3E3</td>
<td>100</td>
</tr>
<tr>
<td><strong>AVG</strong></td>
<td>100</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>SOIL LOAMY CLAY SANDY SILTY PEATY</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>KURT</strong></td>
<td></td>
</tr>
<tr>
<td>L3E3</td>
<td>100</td>
</tr>
<tr>
<td>E3L3</td>
<td>100</td>
</tr>
<tr>
<td>L3S3</td>
<td>100</td>
</tr>
<tr>
<td>S3L3</td>
<td>100</td>
</tr>
<tr>
<td>E3E3</td>
<td>100</td>
</tr>
<tr>
<td>E3S3</td>
<td>100</td>
</tr>
<tr>
<td>S3E3</td>
<td>100</td>
</tr>
<tr>
<td><strong>AVG</strong></td>
<td>100</td>
</tr>
</tbody>
</table>

### Classification:
Classification is based on features of image and category of organized data. Basically classification method has two phases:
- Training phase
- Testing phase

Types of classification:
- Supervised classification
- Unsupervised classification

Statically process classification can be done by following six steps as shown in figure 5

**Figure 3.7 Steps of classification**
- Definition of classification classes
- Selection of feature
- Sampling of training data
- Estimation of universe statics
- Classification
- Verification of Result

The most basic classification techniques are:
- Multilevel slice classification
- Minimum distance classification
• Maximum distance
Other classification like expert system, fuzzy system etc. Distance may be based on nearest neighbour method, farthest neighbour method, Centroid method, Group average method, Wad method.

Classifier:
Decision Tree Classifier is used in our proposed system. Decision tree builds classification models in the form of a tree structure. It breaks down a data set into smaller and smaller subsets while at the same time an associated decision tree is incrementally developed. A decision node has two or more branches. Leaf node represents a classification or decision.

Decision Tree Classifier: A Concise Technical Overview

Decision Tree Induction Algorithm

INPUT: $S$, where $S = \text{set of classified instances}$
OUTPUT: Decision Tree
Require: $S \neq \emptyset$, num attributes $> 0$
1. procedure BUILD_TREE
2. repeat
3. $\text{maxGain} \leftarrow 0$
4. $\text{splitA} \leftarrow \text{null}$
5. $c \leftarrow \text{Entropy(Attributes)}$
6. for all Attributes $a$ in $S$
7. $\text{gain} \leftarrow \text{InformationGain}(a, e)$
8. if $\text{gain} > \text{maxGain}$ then
9. $\text{maxGain} \leftarrow \text{gain}$
10. $\text{splitA} \leftarrow a$
11. end if
12. end for
13. Partition($S$, splitA)
14. until all partitions processed
15. end procedure

Decision Tree Raising Algorithm

INPUT: $D$, where $D = \text{Unpruned Decision Tree}$
OUTPUT: Pruned Decision Tree
1. procedure PRUNE_TREE
2. for all Nodes $n$ in $D$
3. for all Children $c$ of $n$
4. if Replacing $n$ with $c$ does not lower accuracy of $D$ then
5. Replace $n$ with $c$
6. Reclassify former sibling nodes of $n$
7. end if
8. end for
9. end for
10. end procedure

Advantages and disadvantages of different classifiers:
IV. IMPLEMENTATIONS

Mat lab 2018(A)

MATLAB is a scientific programming language and provides strong mathematical and numerical support for the implementation of advanced algorithms. It is for this reason that MATLAB is widely used by the image processing and computer vision community. New algorithms are very likely to be implemented first in MATLAB, indeed they may only be available in MATLAB. We used Computer Vision and Image Processing Tools.

System Requirements
- Windows 7 (or) higher
- 64 bit operating system
- Disk Space
  - 2 GB for MATLAB only,
  - 4–8 GB for a typical installation.
- Minimum 2GB RAM needed
- No specific graphic cards required

V. RESULT ANALYSIS

Processing of Images:

<table>
<thead>
<tr>
<th>INPUT IMAGE</th>
<th>ENHANCED IMAGE</th>
<th>FEATURE EXTRACTED (HSV) IMAGE</th>
<th>GRAY SCALE IMAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image1.png" alt="Image" /></td>
<td><img src="image2.png" alt="Image" /></td>
<td><img src="image3.png" alt="Image" /></td>
<td><img src="image4.png" alt="Image" /></td>
</tr>
<tr>
<td><img src="image5.png" alt="Image" /></td>
<td><img src="image6.png" alt="Image" /></td>
<td><img src="image7.png" alt="Image" /></td>
<td><img src="image8.png" alt="Image" /></td>
</tr>
<tr>
<td><img src="image9.png" alt="Image" /></td>
<td><img src="image10.png" alt="Image" /></td>
<td><img src="image11.png" alt="Image" /></td>
<td><img src="image12.png" alt="Image" /></td>
</tr>
<tr>
<td><img src="image13.png" alt="Image" /></td>
<td><img src="image14.png" alt="Image" /></td>
<td><img src="image15.png" alt="Image" /></td>
<td><img src="image16.png" alt="Image" /></td>
</tr>
</tbody>
</table>

Output Obtained:

<table>
<thead>
<tr>
<th>INPUT IMAGE</th>
<th>TYPE OF SOIL</th>
<th>CROPS SUGGESTED</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image1.png" alt="Image" /></td>
<td>CLAY</td>
<td>1. Broccoli 2. Spinach 3. Carrot</td>
</tr>
<tr>
<td><img src="image2.png" alt="Image" /></td>
<td>SANDY CLAY</td>
<td>1. Lettuce 2. Cabbage greens</td>
</tr>
<tr>
<td><img src="image3.png" alt="Image" /></td>
<td>SILTY SAND</td>
<td>1. Potatoes 2. Wheat 3. Sugar Beet</td>
</tr>
<tr>
<td><img src="image4.png" alt="Image" /></td>
<td>CLAYEY PEAT</td>
<td>1. Rhododendron rhododendron 2. Sedges 3. Sphagnum moss</td>
</tr>
<tr>
<td><img src="image5.png" alt="Image" /></td>
<td>HUMUS CLAY</td>
<td>1. Crucifers 2. Flowering plants</td>
</tr>
</tbody>
</table>

Figure 5.1 Pre-processing of images

Figure 5.2 Obtained output

VI. CONCLUSION

The proposed system has added features like crop suggestion, prediction of water absorption by plants which couldn’t be found in existing systems. Accuracy is more because feature extraction is based on colour and texture of the input images. Thus our proposed idea will make sure to help farmers, agriculture activists in efficient soil classification and crop suggestion.

ACKNOWLEDGMENT

We wish to express our deep sense of gratitude to Miss T. Abimala, Assistant Professor, Department of Instrumentation and Control Engineering, St. Joseph’s College of Engineering, Chennai, for her excellent guidance, valuable suggestion that greatly helped us to complete this paper successfully.
REFERENCES


[12] Wavelet Based Features for Texture Classification P.S. Hiremath*, S. Shivashankar †, Dept. of P.G. Studies and Research in Computer Science, Gulbarga University


[16] Determining Composition of Grain Mixtures by Texture Classification Based on Feature Distributions. Timo Ojala, Matti Pietikäinen, and Jarkko Nisula

[17] Characteristics of Soil by Randy Scott, Consulting Rosarian Woodbridge

[18] Loam From Wikipedia, the free encyclopaedia

[19] Classification of Textures Based on Features Extracted from Pre-processing Images on Random Windows by B.V. Ramana Reddy, A. Suresh, M. Radhika Mani, and V. Vijaya Kumar