Performance of Spectral Angle Mapper and Parallelepiped Classifiers in Agriculture Hyperspectral Image

Sahar A. El_Rahman¹,²

¹ Electronics, Communication, and Computers Systems Engineering - Electrical Department
Faculty of Engineering-Shoubra, Benha University
Cairo, Egypt

² Computer Science Department, College of Computer and Information Sciences
Princess Nourah Bint Abdulrahman University
Riyadh, Saudi Arabia

Abstract—Hyperspectral Imaging (HSI) is used to provide a wealth of information which can be used to address a variety of problems in different applications. The main requirement in all applications is the classification of HSI data. In this paper, supervised HSI classification algorithms are used to extract agriculture areas that specialize in wheat growing and get a classified image. In particular, Parallelepiped and Spectral Angel Mapper (SAM) algorithms are used. They are implemented by a software tool used to analyse and process geospatial images that is an Environment of Visualizing Images (ENVI). They are applied on Al-Kharj, Saudi Arabia as the study area. The overall accuracy after applying the algorithms on the image of the study area for SAM classification was 66.67%, and 33.33% for Parallelepiped classification. Therefore, SAM algorithm has provided a better study area image classification.

Keywords—Accuracy Assessment; ENVI; Hyperspectral Imaging; Parallelepiped Classifier; Spectral Angel Mapper; Supervised Classification

I. INTRODUCTION

Hyperspectral imaging (HSI) has been an active development and research area. Hyperspectral imaging is poised to enter the mainstream of remote sensing because the commercial hyperspectral imaging system’s appearance. Remote sensing is the art and science of acquiring information about an area, object, or phenomenon by measuring the electromagnetic radiation emanating from the surface of the earth using indirect handle with the area, a phenomenon or object under implementation[1].

Hyperspectral images can be used in many various applications like as resource management, environmental monitoring, agriculture and mineral exploration as shown in Fig. 1 [2][3]. However, a recognition of the data and the nature restrictions and various processing strategies are considered an effective use of hyperspectral images.

HSI is described as a technique of spectral sensing that gathers hundreds of relatively small wave bands that supply spectral data to differentiate spectral unique objects or materials [4]. It is used in the analysis, measurement, and interpretation of the spectrum obtained from a particular object in a short, medium or long distance through remote sensing [5].

Fig. 1. Application of Hyperspectral Imaging [3]

II. HYPERSONTAL IMAGING

Hyperspectral images provide large spectral information that is more accurate and detailed over another remotely sensed data type. It has an enhanced ability in that the probability of detecting interested materials is increased. It also provides additional information necessary for materials identification and classification [6]. The HSI pixels compose spectral vectors represent the materials spectral characteristic in the location image [4].

Hyperspectral imagers use hundreds of wavelength channels that have the capability to detect and select unique materials characteristics and features, much like DNA or a fingerprint have unique structures and features [7].
A. Advantages of Hyperspectral Over Multispectral Data

Hyperspectral sensors are expected to improve the ability to observe the earth’s surface. For example, increasing the classification accuracy in contrast to multispectral imaging systems, where HSI gives an opportunity to extract additional exhaustive information over the data available using traditional multispectral imagery. The difference between hyperspectral and multispectral depends on the measurement type or on an arbitrary band number, where HSI systems provide exhaustive spectral information, which enables the analyst to classify and detect the pixels based on their spectral characteristics. However, in several cases, a multispectral imaging system has a spatial resolution higher than hyperspectral systems, however, it has less spectral channels [1].

The technologies of multispectral remote sensing have been usually utilized for remote sensing classification of vegetation since 1960s [8][9]. In a single observation, 3 to 6 spectral bands of data are generated by multi-spectral sensors which range from the visible end of the Electromagnetic Spectrum (EMS) to the NIR (Near Infrared) end [9]. This narrow window of spectral bands is the essential drawback of multi-spectral sensor. Through the previous decade, the improvement in spectrometer has helped to overcome the limitations of multispectral sensors. So, they are providing a good performance in identification, classification, and object detection of earth characteristics [10][11][12]. HSI sensor usually gathers further than 200 spectral bands that span between the visible part of EMS and SWIR (Short Wave Infrared) part. HSI sensor does not present only spectral data in detail consisting of hundreds of bands in one combination [9]. So, these features have resulted in new governmental and scholarly exploration of mapping and classification of vegetation and land cover with HSI application [10][13].

Furthermore, HSI handles narrow spectral bands through a spectral continuous range, and presents the all pixel spectra in the image, even as multi-spectral imagery deals with several images at somewhat narrow and discrete bands, as shown in Fig. 2.

The main advantage of HSI is the amount of spectral details it produces [14]. The main disadvantages are complexity and cost. Hyperspectral data analysis needs fast computers, large data storage capacities, and sensitive detectors. Also, the ways to programme hyperspectral satellite, which are found by one of the researchers, to arrange data on its own and transmits one and only the most significant images, as both storage and transmission of that large data could demonstrate the complexity and cost. The entire potential of HSI has not been explored yet [1][15].

III. CLASSIFICATION

Classification is a technique of information extraction, where it considered as one of the most often used techniques. Most of these techniques depend on the spectral reflectance property analysis of hyperspectral imagery and utilize specific techniques prepared to proceed several types of spectral analysis. The hyperspectral classification processes can be implemented using one of the two techniques: Unsupervised and Supervised [1][16].

A. Supervised and Unsupervised Classification

Supervised Classification, as shown in Fig. 3, needs to identify known prior as training sites out of a collection of personal experience, fieldwork, and map analysis. They are utilized for the classification algorithm training to the eventual land cover mapping of the rest of the images. Each pixel is then evaluated and assigned to the category (class) which has the maximum likelihood [17]. In the unsupervised, the pixels that have similar spectral features (covariance matrices, standard deviations, means, etc.) are grouped by the computer or algorithm into distinct classes as stated by several statistically defined specifications as shown in Fig. 4.

The Supervised hyperspectral algorithm uses the sample with recognized identity (i.e., information clusters with assigned pixels), and pixels with unknown identity are classified by the algorithm. The process begins with choosing and naming regions on the image by the user, which correspond to the clusters of concern. These clusters correspond to information clusters. Then, the algorithm of image classification will detect all analogous regions [18].
B. Supervised Algorithms

There are some supervised methods which have been designed to outline the problem of HSI classification. Various studies have applied algorithms and each algorithm provides different classification accuracy, but the most famous are Parallelepiped and SAM (Spectral Angle Mapper) algorithms.

1) Parallelepiped Classifiers: In this classifier, a common decision rule is used for classifying information. In the image data space, N dimensional parallelepiped classification is formed by the decision boundaries. From the mean of each specific cluster, a standard deviation threshold defines the dimensions of parallelepiped classification. A pixel is being classified and assigned to that class, if its value falls between the high threshold and the low threshold of every band in n bands. In ENVI, the pixel is allocated to the last class be matched, if the pixel value lies in more than one class. Areas are designated as unclassified that do not lie within any of the classes, as shown in Fig. 5 [20].

Parallelepiped classifier is considered as a very fast algorithm where every parallelepiped class demonstrates one information class. Each pixel with digital number into the bounds (range) determined by certain parallelepiped class will be chosen to the corresponding information class [21].

2) SAM (Spectral Angle Mapper) Classifier: SAM is an algorithm that allows quick mapping of the spectral symmetry of the spectra of the image to the spectra of the reference [18]. The spectra of the reference can either be a spectrum measured in a laboratory, a field spectrum or obtained straightway from the image. SAM classifier determines the spectral symmetry between both spectra, handling the two spectra as vectors in the space at the same dimension of the band number [22][23]. This can be easily explained, as shown in Fig. 6, where the spectra of the reference and the test are represented in a two-dimensional plot as two band data. The angle among every reference spectra and every test spectra is computed. The SAM program, in ENVI, assigns the angles to output channels, and then every pixel is allocated to the class defined by the reference spectrum. The class that is assigned to each pixel is saved in the output channel [18].

IV. METHODOLOGY

A. Study Area

As the Kingdom of Saudi Arabia is a desert with a dry climate, caring about its agriculture is important so it won’t be worse. Studying its agriculture and which areas are infected and what areas are good for growing crops might improve its agriculture in the future. Al-kharj is selected as a study area because it is one of the important provinces in the Kingdom of Saudi Arabia for exporting wheat to markets around the world that it is existing in southeast area of Riyadh [24]. The detected area of study (as shown in Fig. 7), is downloaded from United States Geological Survey website [25].

Al-Kharj image when downloaded has the following information: Exact location is in Path 165 and Row 43. Acquisition date was in year 2007 and since Julian day is identified month and can be calculated on the basis that the Julian day was 80 so it was acquired on 23/3/2007 so we can say that Al-Kharj was scanned in the spring season. It was composed of 242 files in .tiff format where each band is represented by a certain file. Consequently, those bands have to be combined to create a single image with all bands [25].

B. Methods

An adequate preprocessing of HSI is requisite to extract helpful information from it. There are a number of preprocessing methods that should be done before starting the classifying [26]. The usage of these methods depends on the downloaded image as shown in Fig. 8. First, we need to acquire the study area image. Then Preprocessing is needed on the image such as: collecting the bands into one image and applying wavelength for each band using Hyperion tool [27]. This tool is not found in ENVI, that it is utilized to support Hyperion data use in ENVI. Where the most basic functionality of this tool is converting Geo TIFF data sets into files of ENVI format which contain band information and wavelength [28]. Then, spatial and spectral subset is used to minimize the huge data and eliminate redundant data and atmospheric correction. After the preprocessing, the main processing takes place which is applied to two supervised classification algorithms: SAM, and Parallelepiped. Finally, these two algorithms will be

Fig. 4. Unsupervised Classification [19]

Fig. 5. Parallelepiped classifier

Fig. 6. SAM program

Fig. 7. Study Area
compared using accuracy assessment as described in the following sections.

![Image](image1.png)

Fig. 6. The plot of a test and a reference spectrums of a two band image [18]

![Image](image2.png)

Fig. 7. Hyperspectral image for Al-Kharj Kharj [25]

1) **Preprocessing**: The data have been subjected to gathering all the bands into one image using the Hyperion Tool. Subsetting has been done to minimize the complexity and the big volume of the data produced by HSI sensors where it requires a lot of time to process, and to provide the base for comprehensive analysis, it is subjected to atmospheric corrections. Gathering all bands into one image using the Hyperion tool is accomplished first, then the sub setting before the atmospheric correction [27].

Transforming the image into an ENVI format file using Hyperion Tool. Hyperion tool is found in an external file in ENVI software, where it will get the study area image ready in an ENVI standard format instead of separated bands as images files with TIFF extension. The output of this process is six files: two with .DAT extension, two with .HDR extension, one with .txt extension, one with .STA extension.

Sometimes the hyperspectral image downloaded contains lot of unneeded information. This information will cause an over processing and minimize the performance. These unwanted data in the study will be cut off. This method is called Sub setting. This will enable the processor to focus only on data needed for a study area and get better performance of HSI classification. The process can account for the differing data quality and discernment capabilities among spectral bands, and use the spatial and spectral information with each other [29].

In spatial sub setting, the wanted area of study might be a part of the downloaded image from the satellite. Spatial subsetting is applied to the image to resize it, as well as utilize the function of resizing to produce new images with any aspect ratio or size [30]. It is focused on choosing the area of study, but it must be selected in a square shape. Fig. 9 shows a subset by image dialog showing the selected subset area.

The result after sub setting is viewed in three displays. Fig. 10-a is for false color display, Fig. 10-b is for true color display and Fig. 10-c is for gray scale display.

The Spectral Subsetting is established to identify bad bands. The ones which will don't support the analysis will only cause an overload on the processor. ENVI headers have related ancillary data (spectral library name, band name, bad bands list, wavelength, Full Width Half Maximum (FWHM)) depending on the image data type. These bad bands could be identified in ENVI, where Edit dialog of bad bands list contains 242 bands where each good band is highlighted and the rest is bad bands. In most hyperspectral images, last and first bands are considered bad. In the area of study has 155 good bands and the rest are bad.

   a) **Atmospheric Correction**: Even a relatively clear atmosphere interacts with coming, and reflected solar energy. These interactions minimize the degree of coming energy arriving the ground for certain wavelengths. And also, they minimize the degree of reflected energy arriving an airborne or satellite sensor. So, in these regions, little useful information can be obtained from image bands. If atmospheric conditions are spatially variable, atmospheric effects may also differ between areas in a single scene, that's why we must work to rectify the image that is affected by the atmospheric gases. Also, ENVI can be used to rectify HSI image with atmospheric impact where it has multiple Atmospheric Correction methods [26][31].

Fig. 11 shows the influence of the atmosphere correction on the extent of absorption and reflectance of each pixel before and after atmosphere correction. In order to know the extent of the effect it we will view the spectral profile for a certain pixel in the image. The spectral profile for a certain pixel in the study area image before the atmospheric correction, as shown in Fig. 11-a. Output of spectral profile will be in the form of a curve, the top of each curve reflects the reversal, and the bottom of each curve reflects the absorption of each pixel, in zoom window there is the intersection of two lines point of this intersection represent the current pixel of the image and it will show the result of this pixel on spectral profile. If the pixel has a higher value reflection, it is better to detect materials [32].

Fig. 11 shows the image after applying atmospheric correction, where the spectral profile differs between the image before and after atmospheric correction. In Fig. 11-a, there is a lot of absorbed energy unlike Fig. 11-b which has a really high reflected energy.
Fig. 8. Hyperspectral imaging analysis

Fig. 9. Subset by Image Dialog where the red box contains the study area

Fig. 10. Spatial subset of Al-kharj study area
(a) False color  (b) True color  (c) Gray scale

Fig. 11. Spectral profile for a certain pixel
(a) Before atmospheric correction  (b) After atmospheric correction
2) Processing: In the processing phase, NDVI (Normalized Difference Vegetation Index) values are utilized in corrected image, to show the green vegetation amount exist in the pixel. As more green vegetation is indicated, the NDVI values also increase. Such a step is just to view the amount of green area in a study area. Then, the classification methods SAM and Parallelepiped are applied.

a) NDVI: The study area image is analysed to check the presence of vegetation and the density of the vegetation. The Vegetation Delineation tool enables us to fast recognize the vegetation presence and to see its vigor level. Also, the Wizard supplies helpful tools to create graphics that are used in briefings and reports.

For most applications of spectral processing, dealing with data with atmospheric correction due to produce more accurate results. Since the study area image is already corrected, there is no need to perform any correction. The NDVI equation:

\[
NDVI = \frac{(NIR - Red)}{(NIR + Red)}
\]  

(1)

The NDVI generates an image that ranges from -1 to 1. Pixels with no vegetation tend towards -1, while the pixels with vigorous vegetation tend towards 1. Examine results, as shown in Fig. 12.

b) SAM (Spectral Angle Mapper): SAM is a physically based spectral classifier which uses a n - D angle to identify the pixels to reference spectrum. It determines the spectral symmetry among two spectrums by computing the angle between them and handling them as vectors with dimensions equal to the band number. Such algorithm is comparatively insensitive to light and albedo effect when utilized in standardized reflectance data. The spectrum of endmember utilized by this algorithm can come from spectral libraries, or directly, they are extracted from the image (as an average spectrum of ROI) or ASCII files [2][20]. The steps followed in SAM are shown in Fig. 13-a. This work is based on spectral library for more accuracy. Therefore, it will match each pixel spectral signature in the image of the study area to the selected vegetation endmember spectral signature.

The result of study area classification is represented in Fig. 13-b after SAM classifier is applied, where the wheat is represented by the yellow color, and wheat (tan) is represented by the green color.

c) Parallelepiped Classification: It is a widely supervised algorithm. The bands of the image are utilized to define the pixels of the training area for every band based on least and most pixel value. Though, it is more accurate than other algorithms classification, it is not more widely utilized. Due to considerable unclassified pixels are left and likewise it have an overlapping among pixels of the training area. The candidate pixel values are compared to lower and upper boundaries [33]. This method needs the training areas that are selected by the user for utilization as the classifying base. When these data are collected, a different number of classification routines are ready to identify the concerning pixels and carry out the classification.

- **Defining ROIs (Regions of interest)**

ROIs are images parts, selected either graphically or by other ways like a threshold. They are with irregular shape and are normally utilized to make statistics for masking, classification, and different related processes [26].

Number of ROI depends on the study area image work on. So, if the image has lots of features such as: mountains, sea, ocean grass, building. You will add region of interest based on the features you can see in the image. When the image represented by false color has red circles where they represent pivot which is an agricultural area, based on the knowledge on how pivot is represented by circles, we know it is agriculture. Since pivot is represented by red, we can say that everything with red color is agriculture, but the brightness of this red color is different from place to place and this is because the density of green land is different between them [34].

- **Parallelepiped Classification in ENVI**

Using ROI represented above, the classes are created utilizing parallelepiped classifier. The default attributes and varied standard deviations from the mean of these regions of interest are used [17][21]. The steps followed in Parallelepiped, is shown in Fig. 14-a. Based on the ROI defined, it will search for the min, max digital values for each region in ROI in order to classify any pixel that is between the values as it belongs to this region.

The result of applying parallelepiped classification on the image has classified it only to the agricultural area where other areas were ignored, as shown in Fig. 14-b.

3) Post processing: Classified images require post-processing to perform generalizing for classes, which is called Majority Analysis. It is utilized in changing any spurious pixel within a considerable single class to that class. As shown in Fig. 15, the results after applying the majority of the two classified images. Whereas, the image is smothered but details such as certain pivot with a high density of green land has been lost. So, results before applying majority will be certified.

Fig. 12. Vegetation delineation and stress detection for processing complete
Fig. 13. SAM classification

Fig. 14. Parallelepiped Classification
V. RESULTS

The image of the area of study (Al-Kharj in Kingdom of Saudi Arabia) has been classified using SAM classifier as shown in Fig. 13-b and by using Parallelepiped classifier as shown in Fig. 14-b, the validation or accuracy assessment is a significant stage in the remote sensing image process. Where it defines the data value of the resulting information to the user [35]. The aggregate accuracy is studied by taking the sum of No. of pixels classified right and divided by the total pixel number [36]. The correct class of the pixels is defined by the ground truth ROIs. The overall accuracy after applying it on the image of the study area for SAM classification was 66.67%, and 33.33% for Parallelepiped classification. Therefore, SAM provides better classification for the image of the area of study.

VI. CONCLUSION

This work has achieved the potential utilization of Parallelepiped and SAM classification algorithms integrated with EO-1 Hyperion imagery analysis to extract all areas of the wheat in the study region, that is Al-Kharj in Kingdom of Saudi Arabia, as it is considered as one of the regions greatly export wheat to the market. The Parallelepiped and SAM classifiers were implemented by using the identical training set and points of validation chosen on the gained EO-1 Hyperion image, that allow an explicit performance comparison of them.

SAM has better results usually because it is based on comparing pixels to reference objects using spectral signature. Parallelepiped differs from SAM, because it is depending on checking the digital value of the pixel if it ranges within the Min-Max values of the drawing ROIs.

ACKNOWLEDGMENT

The author would like to thank all the participants involved in this work at CCIS - PNU, especially to Abeer R. Alwalid, Ghadah A. Aljarallah, Manal I. Al Manna, Nada I. Al Rashed and Wateen A. Aliady.

REFERENCES


