IDENTIFICATION MODEL FOR LARGE REMOTE SENSING DATASETS APPLIED TO ENVIRONMENTAL ANALYSIS WITHIN MEXICO

Ivan E. Villalon-Turrubiates, *Senior Member, IEEE*
Instituto Tecnológico y de Estudios Superiores de Occidente (ITESO), Universidad Jesuita de Guadalajara
Periférico Sur Manuel Gómez Morín 8585, 45604 Tlaquepaque Jalisco México
Telephone: (+52 33) 3669-3434+3873, E-mail: villalon@ieee.org

ABSTRACT

The classification procedure to identify remote sensing signatures from a particular geographical region can be achieved using an accurate identification model that is based on multispectral data and uses pixel statistics for the class description. This methodology is referred to as the Multispectral Identification Model. This paper presents this particular methodology applied to large remote sensing datasets (multispectral images obtained from the SPOT-5 satellite sensors) with the objective to perform environmental and land use analysis for regions within Mexico, taking advantage of high-performance computing techniques to improve the processing time and computational load. The results obtained uses real multispectral scenes (high-resolution optical images) to probe the efficiency of the classification technique.

***Index Terms*—** image classification, remote sensing, image processing, multispectral.

1. INTRODUCTION

Considerable progress has been made in the application of remote sensing techniques to both research and operational solutions for environmental analysis. Modern applied theory of image processing is now a mature and well-developed research field, presented and detailed in many works ([1] thru [7] are only some indicative examples).

Although the existing theory offers a manifold of statistical techniques to tackle with the particular environmental monitoring problems, in many applications areas there still remain some unresolved crucial theoretical and data processing problems. One of them is particularly related to the extraction of physical characteristics (e.g. water, land cover, vegetation, soil, humid content, and dry content) for applications in environmental analysis within regions of Mexico. Moreover, there is an important necessity for a model that could be implemented with a high-performance computing technique to reduce the processing time until a suitable value that could be considered as real time. Is important to clarify that the concept of real time is completely related to the time in which the processing of the information is needed to be available.

2. MULTISPECTRAL IMAGING

Multispectral imaging is a technology originally developed for space-based imaging. Multispectral remote sensing (MRS) images are the main type of products acquired by remote sensing radiometers. Usually, MRS systems have from 3 to 7 radiometers; each one acquires one digital image (also called scene) in a small band of visible spectra, ranging 450 nm to 12,500 nm [6].

The wavelengths for the spectral bands are defined as follows (the values are approximated, exact values depends on the particular MRS instruments [7]):

1. Blue: 450-520 nm.
2. Green: 520-600 nm.
3. Red: 600-690 nm.
4. Near-Infrared: 750-900 nm.
5. Mid-Infrared 1: 1,550-1,750 nm.
6. Mid-Infrared 2: 2,080-2,350 nm.
7. Thermal-Infrared: 10,400-12,500 nm.

For different purposes, combinations of spectral bands can be used. Most commonly, they are represented with red (R), green (G) and blue (B) channels (traditional photography), and are referred to as True-Color remote sensing images [6].

3. IMAGES FROM SPOT-5 SENSORS

SPOT Imagery (from its French acronym "Système Pour L'Observation de la Terre") is the worldwide distributor of geographic information products and services derived from the SPOT Earth observation satellites. A SPOT satellite image is a view of the Earth seen through one of the satellite's high-resolution imaging instruments. The technical characteristics of each instrument determine the resolution and spectral mode of the image. The acquired image is then processed to suit users' requirements in terms of geographic information. It is delivered in a standard format able to be integrated directly in current geographic information software packages [8].

The images used for this research are provided by SPOT-5 through its Mexican office SEMAR (from its Spanish acronym "Secretaría de Marina") under the ERMEXS program (from its Spanish acronym "Estación de Recepción México de la Constelación SPOT") [9].

The spatial resolution of the images is 10m (spectral mode Hi) for a *6000x6000* pixels image, and the spectral resolutions (3 spectral bands) corresponds to the following:

1. B1 – Green (G): 520-600 nm.
2. B2 – Red (R): 600-690 nm.
3. B3 – Near-Infrared (I): 750-900 nm.

4. MULTISPECTRAL IDENTIFICATION MODEL

The Multispectral Identification Model (MIM) classificatory rule is computationally simple and this study shows that it can result in a high classification accuracy that is comparable to other more computationally intensive algorithms. It is characterized by the mean and variance values of the remote sensing signatures (RSS) that are defined as classes, and the Euclidean distances based on the Pythagorean Theorem.

The training data for class segmentation requires the number of RSS to be classified (c); the means matrix **M** ($c \times c$ size) that contains the mean values μ_{cc}: ($0 \le \mu_{cc} \le 255$, gray-level) of the RSS classes for each of the spectral bands within the image; and the variances matrix **V** ($c \times c$ size) that contains the variances of the RSS classes for each band. The matrix **M** and **V** represents the weights of the classification process.

Next, the image is separated in all the different spectral bands and each (i, j)-*th* pixel is statistically analyzed calculating the means and variances from a neighborhood set of *5x5* pixels for each spectral band, respectively.

To compute the output of the classifier, the distances between the pixel statistics and the training data is calculated using Euclidean distances based on the Pythagorean Theorem for means and variances, respectively. The decision rule used by the MIM is based on the minimum distances gained between the weighted training data and the pixel statistics. All the process is performed with parallel computing techniques to improve the processing times.

The detailed stages of the computational algorithm of the MIM method for RSS classification of the MRS scenes is described as follows, and Figure 1 shows the detailed processing structure of the MIM classifier:

1. Set the number of RSS to classify.
2. Select one point on the MRS image for each class to be classified.
3. Separate the spectral bands from the true-color MRS image.
4. The selected points determine the training weights that consist of the means matrix **M** and the variances matrix **V**.
5. For each (i, j)-th pixel in the spectral bands, respectively, perform the following process:
   - Set a *5x5* pixel neighbourhood shift window **W**.
   - Determine the mean of the shift window **W**.
   - Determine the variance of the shift window **W**.
   - Calculate the Euclidean distances between the means and the training means for each band and for each class.
   - Calculate the Euclidean distances between the variance and the training variances for each band and for each class.
   - Select the minimum class distance for the means.
   - Select the minimum class distance for the variances.
   - Perform a comparison between the class distance for the mean and the class distance for the variance, and classify the pixel according to the minimum value and the class from which is obtained.

5. SIMULATION EXPERIMENT

The RSS are extracted from two different multispectral scenes using the method. Figure 2(a) and 2(b) shows the high-resolution *6000x6000* pixels images in TIFF format with three spectral bands [6].

Figure 2(c) and 2(d) shows the RSS map obtained applying the MIM method for the adopted ordered weight vector.

Three level RSS are selected for this particular simulation process:

– RSS relative to the wet zones of the MRS image.
– RSS relative to the humid zones of the MRS image.
– RSS relative to the dry zones of the MRS image.

6. CONCLUDING REMARKS

From the simulation results, is possible to deduce that the developed MIM classifier provides a high-accurate classification without unclassified zones because it uses more robust information in the processing (several spectral bands). The simplicity of the developed methodology and the use of parallel programming allows real-time implementation as a high-performance computing technique. The reported results show the qualitative analysis of the overall performance of the MIM method applied to environmental analysis of regions within Mexico using SPOT-5 images.

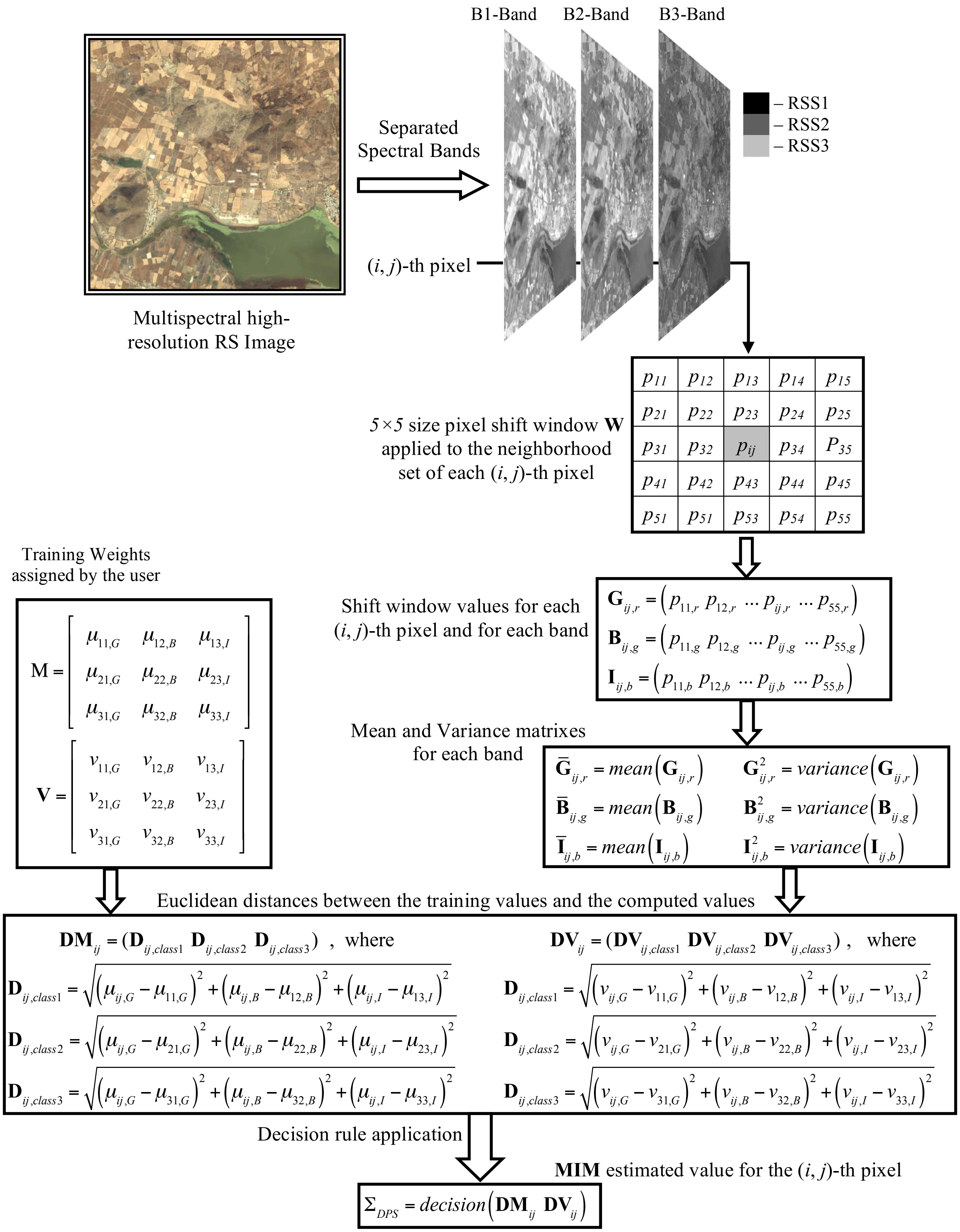
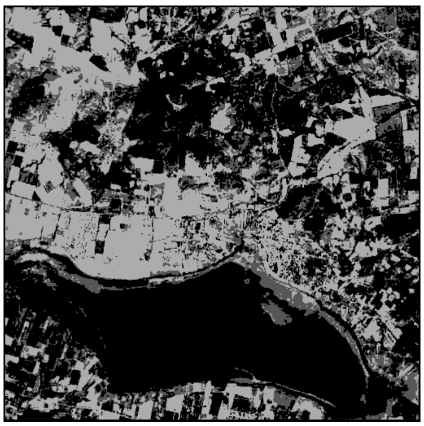


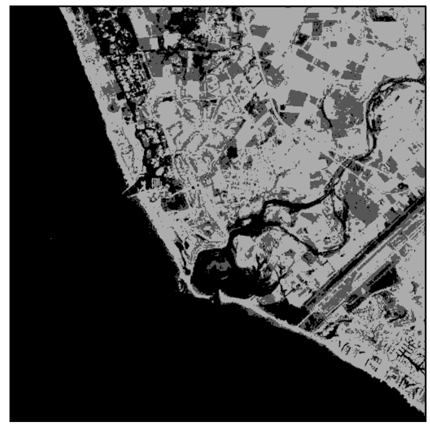
FIG. 1. PROCESSING STRUCTURE OF THE MIM METHOD.

(A) ORIGINAL MULTISPECTRAL IMAGE, SCENE 1.

(B) ORIGINAL MULTISPECTRAL IMAGE, SCENE 2.



(C) IDENTIFICATION MODEL APPLIED USING MIM, SCENE 1.



(D) IDENTIFICATION MODEL APPLIED USING MIM, SCENE 2.

FIG. 2. SIMULATION RESULTS FOR RSS EXTRACTION FROM MULTISPECTRAL REMOTE SENSING IMAGES.

ACKNOWLEDGMENT

The author would like to thank the **Instituto Tecnológico y de Estudios Superiores de Occidente (ITESO)** of Mexico for the resources provided for this research under the project titled "*Desarrollo de modelos adaptivos para el procesamiento digital de señales multiespectrales de percepción remota y su implementación como software de alto desempeño*".

REFERENCES

[1] S.W. Perry, H.S. Wong and L. Guan, *Adaptive Image Processing: A Computational Intelligence Perspective*. U.S.A.: CRC Press, 2002.

[2] P.M. Mather, *Computer Processing of Remotely-Sensed Images*. U.S.A.: John Wiley & Sons, 2004.

[3] R. Porter, D. Eads, D. Hush, and J. Theiler, "Weighted order statistics classifiers with large rank-order margin", in *Proceedings of the 20th International Conference on Machine Learning*, Washington U.S.A., 2003.

[4] I.E. Villalon-Turrubiates, "Weighted Pixel Statistics for Multispectral Image Classification of Remote Sensing Signatures: Performance Study", *Proceedings of the 5rd IEEE International Conference on Electrical Engineering, Computing Science and Automatic Control (CCE),* Mexico City, pp. 534-539, 2008.

[5] I. E. Villalon-Turrubiates, "Dynamical Characterization of Land-Use Classification using Multispectral Remote Sensing Data for Guadalajara Region", *Proceedings of the 2014 IEEE International Geoscience and Remote Sensing Symposium (IGARSS): Energy and our Changing Planet,* Quebec Canada, pp. 145-148, 2014.

[6] H. Hough, *Satellite Surveillance*. Loompanics Unlimited, U.S.A., 1992.

[7] J. R. Jensen, *Introductory Digital Image Processing: A Remote Sensing Perspective*. U.S.A.: Prentice-Hall, 2005.

[8] Satellite Imagery and Smart Mapping Solutions, http://www.spotimage.com, 2017.

[9] Estación de Recepción México de la Constelación SPOT – Secretaría de Marina SEMAR / Armada de México, http://www.ermexs.siap.gob.mx, 2012.