Efficient Clustering and on-board ROI-based Compression for Hyperspectral Radar

Rossella Giordano, Angela Lombardi, Pietro Guccione

Dipartimento di Ingegneria Elettrica e dell'Informazione
Politecnico di Bari
Email: giordanorossella88@gmail.com, angela.lombardi@poliba.it, pietro.guccione@poliba.it

***Abstract*—In recent years, hyperspectral sensors for remote sensing of the Earth have become very popular. Such systems are able to provide the user with images having both spectral and spatial information. The current hyperspectral spaceborne sensors are able to capture large areas with increased spatial and spectral resolution. For this reason, the volume of acquired data must be reduced on-board in order to avoid a low orbital duty cycle due to limited storage space. Recently, literature have focused the attention to efficient way of on-board data compression, since this is a challenge task due to the difficult environment (outer space), and due to the limited power and computing resources. The current work proposes a framework for on-board operations such as: automatic recognition of target types or detection of events in near real time, in regions of interest with an unsupervised classifier; the compression of specific regions with different bit rates compared to the remaining acquisition (background); the management of the data volume to be transmitted to the Ground Station. Experiments are shown using real data taken from AVIRIS airborne sensor in a harbor area.**

Keywords–Hyperspectral; ROI; clustering; on-board compression.

I. INTRODUCTION

A hyperspectral dataset is a data cube composed by a number of images equal to the number of bands in which the sensor has acquired and size equal to that of the captured area. Hyperspectral images are used in a wide variety of applications like: analysis of type of vegetation, monitoring of the forest condition (biomass, deforestation, changes with seasonal cycles), monitoring of coastline and glaciers (glaciers erosion), monitoring of the environment disasters (forest fires, oil spill, flooding) and so on [1].

The high dimensionality of hyperspectral data is very advantageous for the image analysis but, at the same time, the on-board storage space and the transmission bandwidth are limited resources. For this reason, it is necessary to set a proper system of on-board data compression before the transmission to the Ground Station.

Usually, hyperspectral images show high spatial correlation among neighbor pixels and correlation among bands. Such properties have been exploited by compression algorithms to remove the redundance. There are many ways of analysis for redundancy reduction described in literature: these are divided into transformed-based and predictors-based. The transformed-based methods use Discrete Wavelet Transform (DWT) or Discrete Cosine Transform (DCT) [2]. The wavelet transform method offers an efficient rate-distortion and for this reason it has been implemented within the JPEG2000 standard, which is currently the most advanced and efficient compression way for images [3].

The proposed framework aims: (i) at recognizing specific regions of interest (ROIs) using a someway commanded query (generated by the on-ground user) and by means of a computationally efficient clustering algorithm and (ii) to make a differential efficient compression of the regions on the basis of a hierarchy specified in the query (region of interest, medium interest, background). Comparison of the algorithm efficiency is carried on with current JPEG2000 standard in terms of generated data volume and distortion level on the regions of interest.

The rest of the paper is organized as follows: the adopted framework is presented in Section II, the results of the experiments are shown in Section III, and finally the conclusions are reported in Section IV.

II. GENERAL FRAMEWORK

The work has been carried out in two steps.

During the first step, it is supposed that the system is able to recognize specific regions of interest (particular topographic classes, specific events as disaster or limited areas), according to external commands or specific query. Starting from this nontrivial hypothesis, a clustering algorithm is applied to the acquired image to isolate the class or region of interest. The clustering resorts to methods already known in literature (to simplify the on-board process) such as the k-means clustering, but an automatic estimation of some critical parameters is applied for a quicker and more efficient on-board implementation.

The second step consists in the encoding and compression of the ROIs identified in the first step. The aim of these operations is to encode the ROIs with a higher bit rate since they represent the areas of interest for the user, while the remaining part of the image (that we shall call the background) using a lower bit rate.

The overall algorithm has, this way, certainly a distortion rate that is lower than that of a comparing standard as JPEG2000. However, over the specific area of interest, the distortion can be reduced even using the same average bit rate. An alternative comparison with JPEG2000 accounts for the same distortion and compares the data volume generate in the two cases.

The diagram of the implemented algorithm is shown in Figure1.

A. Clustering algorithm for detection

Unsupervised classification methods are more suitable for automatic segmentation and identification of ROIs in hyperspectral images. Clustering algorithms, starting from a data

 ISBN: 978-1-61208-487-9

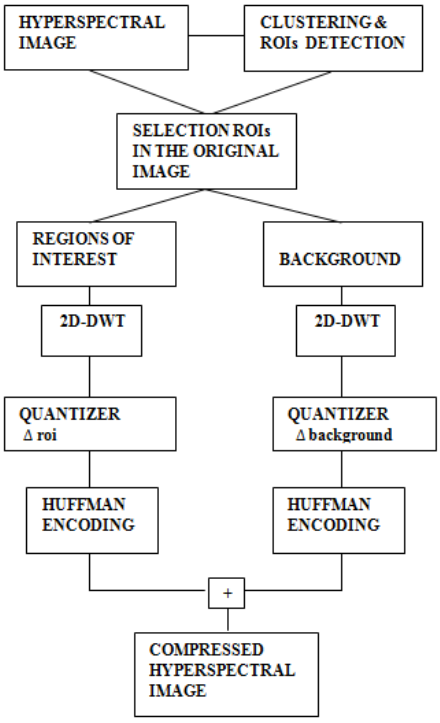

Figure 1. Block diagram of the implemented algorithm.

set, are able to identify significant patterns without knowning labels and, consequently, they do not need of training phases to improve the performances of an automatic classifier. For these reasons, clustering methods can bring significant benefits:

- it is not necessary to employ considerable resources to learn features of a large set of images to identify ROIs, making the application general purpose;
- by using the intrinsic characteristics of the image, the computational time is reduced, making possible a real time (or near-real time) automatic detection of ROIs;
- the characteristics of some patterns may change over time, but a clustering algorithm can easily follow such variations.

One of the most efficient clustering algorithm known in literature is the K-means [4]: through an iterative procedure, K partitions of the points in the initial set are identified by minimizing the sum of squares of distances between the centroid of each partition and all the other points of the set. A great advantage offered by this process is the high convergence speed, while the identification of the optimal parameter K represents a critical step.

In this work, the pixel spectral depth has been employed to make an unsupervised classification of the regions by means of the K-means algorithm. An automated method that exploits the statistical characteristics of the image [5] has been improved to fix the number of clusters K: from the three-dimensional co-occurrence matrices of the hyperspectral cube [6], information about the patterns of pixels of the entire 3D structure have been extracted to calculate the optimal value of K.

For a two-dimensional image quantized to n levels of gray, a co-occurrence matrix (GLCM) describes the spatial relationship among gray level values and it consists of a $n \times n$ array obtained by specifying the separation distance between two pixels (i.e., the offset) and the direction along which to cross the image. These two sets of information can be synthesized through a displacement vector $d = (dx, dy)$, where dx represents the offset in the x direction and dy is the offset value in the y direction of the input image. Each entry (i, j) in the matrix represents the total number of times that the pixel with value i occurred in the specified spatial relationship to a pixel with value j in the input image, hence the matrix provides information about the different combinations of pixel gray levels existing in an image.

Formally, for an image I of size $P \times Q$, the gray-level co-occurrence matrix C is defined as:

$$C(i,j) = \sum_{i=1}^{P} \sum_{j=1}^{Q} \begin{cases} 1, & \text{if } I(x,y) = i \wedge I(x+dx, y+dy) = j \\ 0, & \text{otherwise} \end{cases} \quad (1)$$

In hyperspectral images, it may not be appropriate to analyze independently the individual spectral bands. To extract both spatial and spectral information, co-occurrence matrices for volumetric data (3D GLCM) have been adopted. These matrices are able to capture the spatial dependence of gray-level values across spectral bands. A 3D GLCM is defined by specifying a displacement vector $d = (dx, dy, dz)$, where dx and dy are the same as described for 2D matrices, and dz represents the offset distance along the spectral axis of the hyperspectral image. However, while for a 2D GLCM only 4 displacement directions are possible, in the three-dimensional case, there are 26 admissible co-occurrence directions and only 13 of them differ from each other. The 13 directions and their corresponding displacement vectors are reported in Table I, where D is the offset distance between the pixel of interest and its neighbors, θ is measured in the XY plane in the positive x direction, and ϕ is the angle between the vector which identifies the pixel along the spectral direction and the XY plane. A graphical presentation of the pixel of interest X_0 and the 13 adjacent pixels X_i, $i = 1, \ldots, 13$ in the directions specified in Table I is shown in Figure 2.

TABLE I. DISPLACEMENT VECTORS AND DIRECTIONS OF CO-OCCURRENCE MATRICES FOR VOLUMETRIC DATA.

Displacement vectors	Directions (θ, ϕ)
$(D, 0, D)$	$(0°, 45°)$
$(D, 0, 0)$	$(0°, 90°)$
$(D, 0, -D)$	$(0°, 135°)$
(D, D, D)	$(45°, 45°)$
$(D, D, 0)$	$(45°, 90°)$
$(D, D, -D)$	$(45°, 135°)$
$(0, D, D)$	$(90°, 45°)$
$(0, D, 0)$	$(90°, 90°)$
$(0, D, -D)$	$(90°, 135°)$
$(-D, D, D)$	$(135°, 45°)$
$(-D, D, 0)$	$(135°, 90°)$
$(-D, D, -D)$	$(135°, 135°)$
$(0, 0, D)$	$(-, 0°)$

In the implemented algorithm, an offset distance $D = 1$ between adjacent pixels has been considered, obtaining 13 3D co-occurrence matrices for the entire hyperspectral cube. The diagonal elements of such matrices represent pixel pairs with no gray level difference and contain information on clusters of

 ISBN: 978-1-61208-487-9

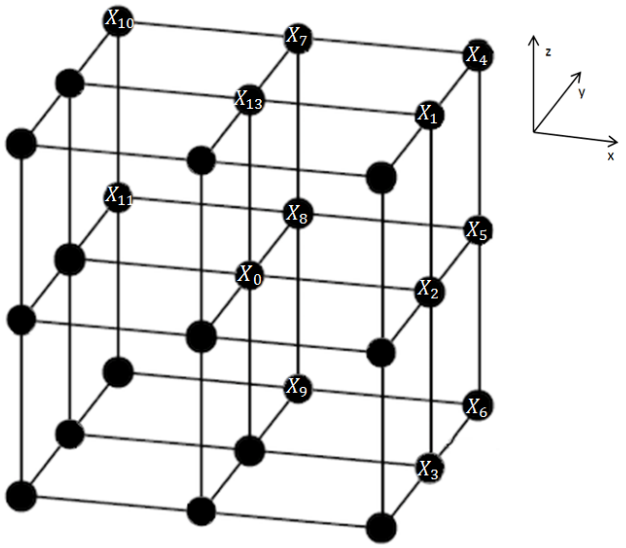

Figure 2. Pixel of interest X_0 and the adjacent pixels in the 13 allowed directions.

pixels. Then, for each matrix, the following operations were performed:

1) extraction of the main diagonal;
2) computation of the histogram of the values of the main diagonal;
3) detection of the local maximum of the histogram.

The gray value corresponding to the local maximum is an index of the number of clusters present in the considered co-occurrence direction, therefore the total number of clusters of the hyperspectral image is obtained as the maximum value among the 13 local maxima.

B. ROIs detection

After the identification of the K clusters, the various spectral signatures represented by the centroid vector are computed and remain saved in the on-board memory. Since data are not normally distributed, these classes are statistically discriminated through two nonparametric quantities: the median value and the interquartile range. These parameters are also used to specify the ROIs by the user.

In order to identify practical regions of interest, a rectangular window of fixed size is chosen to circumscribe the identified clusters. Therefore, a rectangular and fixed-size block is applied to the scanning of the image. Every time the window selects a region where the cluster of interest is in majority (i.e., over a given threshold), the rectangle is classified as ROI, else it is classified as background. The coordinates of the ROIs are saved for the recognition of these regions in the various spectral bands of the original image.

C. ROIs encoding

The DWT is often used in image processing applications and for compression. It is a decomposition f signals into lower resolution and details and it may be viewed as successive low-pass and high-pass filtering of a discrete time-domain signal. On the selected regions of interest, a 2D-DWT is performed for each band. If the considered image for a given band is composed by A rows and B columns, after the 2D-DWT operation, four sub-bands (C,H,D,V) images, each with A/2 rows and B/2 columns, are obtained. The sub-band C has the highest energy compared to the other sub-bands, since it corresponds to the low-pass horizontal and vertical coefficients of the wavelet transform. The 2D-DWT of a function $g(x, y)$ of size $A \times B$ is:

$$W_\delta(a,b) = \frac{1}{\sqrt{AB}} \sum_{y=0}^{B-1} \sum_{x=0}^{A-1} g(x,y)\delta_{a,b}(x,y) \tag{2}$$

$$\begin{aligned} W_\vartheta^k &= (j,a,b)\frac{1}{\sqrt{AB}} \sum_{y=0}^{B-1} \sum_{x=0}^{A-1} g(x,y)\vartheta_{j,a,b}^k(x,y) \\ k &= H, D, V \end{aligned} \tag{3}$$

$$\begin{aligned} \delta_{a,b}(x,y) &= \delta(x-a)\delta(x-b) \\ &= \sum_m h_\delta(m-2a)\sqrt{2}\delta(2x-m) \\ &\times \sum_n h_\delta(n-2b)\sqrt{2}\delta(2x-n) \end{aligned} \tag{4}$$

$$\begin{aligned} \vartheta_{j,a,b}^H &= 2^{\frac{j}{2}}\vartheta^H(2^j x - a, 2^j x - b) \\ &= 2^{\frac{j}{2}} \sum_m h_\vartheta(m-2a)\sqrt{2}\vartheta(2^{j+1}x - m) \\ &\times \sum_n h_\vartheta(n-2b)\sqrt{2}\vartheta(2^{j+1}x - n) \end{aligned} \tag{5}$$

where δ is a scaling function, $W_\delta(a,b)$ the approximation coefficients of the function $g(x,y)$, $W_\vartheta^k(j,a,b)$ the coefficients relative to the horizontal, diagonal and vertical details, h_ϑ and h_δ are called wavelet filters. In this work, we have used a biorthogonal wavelet filter [7] [8] after the verification that this wavelet is able to reduce the distortion in the final images, consequent to the approximation in using just the low-pass coefficients.

The block for the the DWT implementation is followed by an operation of uniform scalar dead-zone quantization (USDZQ) [9], in which each component C of each ROI, in relation to the priorities set by the user, is quantized with a different number of bits, which is anyway smaller than that of the original acquisition. The quantization indices are expressed as:

$$q[i] = \text{sign}(C[i])\lfloor \frac{|C[i]|}{\Delta} \rfloor \tag{6}$$

where $C[i]$ denotes the samples of sub-band C and Δ is the quantization step.

The rectangular windows (actually, the low-pass coefficient of the wavelet transform), after the identification of the assigned label (ROI or background, if just two level of priority are given) are quantized with a number of bits per pixel (bpp). The ROI are quantized with a number of bits per sample larger than the background windows.

For each quantized window an entropic coding is applied to decrease the length of the code to be transmitted to the

 ISBN: 978-1-61208-487-9

ground. Entropy encoding is a variable-length encoding that reduces the volume of data to be transmitted by eliminating, in principle, all the redundancy. In this way, the average number of bits per sample becomes very close to the theoretical limit, i.e., the source entropy [10]:

$$H(P) = \sum_{k=1}^{n} -p(k)\log_2 p(k) \quad (7)$$

where $p(k)$ is the probability of emission of symbol x_k.

In this case, the Huffman encoding [11] has been implemented. Huffman encoding is used for lossless data compression; it is based on the frequency of occurrence of the data. The principle is to use a lower number of bits to encode the data that occur more frequently. The average length of a Huffman code depends on the statistical frequency with which the source produces each symbol from its alphabet. Pixels within a region that must be encoded may be considered as the outcome of a discrete source, according to the fact that their values are outputs of a finite dictionary, since a previous fine quantization has been applied to the image [12]. For a fast implementation, the frequency of all pixels taken from all bands is used to estimate the probability of each symbol. As shown in [13], the Huffman encoding needs less execution time than the other entropy encoding algorithms.

The decoding is performed through the reverse operations of encoding, quantization and 2D-DWT for each ROI and background. The information sent to the decoder, in addition to the entire coded image, are: the coordinates of the ROIs identified with the clustering algorithm, the components (H, D, V) relating to the DWT transform for each ROI and the dictionaries for each ROI and BG used in the Huffman encoding.

III. RESULTS

The purpose of this section is to show an example of unsupervised classification (clustering) of a harbor area followed by the compression of the ROIs to detect the ships near the monitored area in real time.

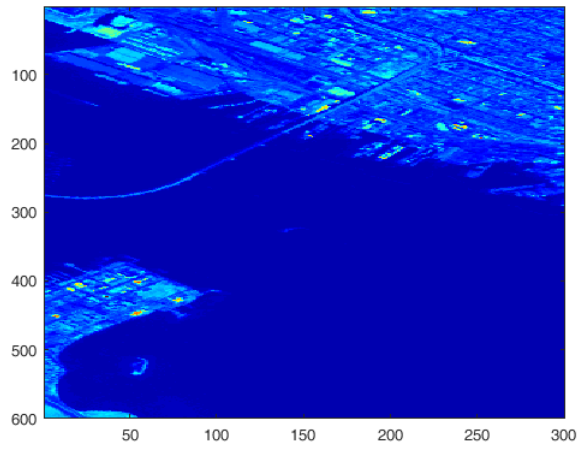

Figure 3. Input image.

An hyperspectral data set acquired by the sensor AVIRIS (mounted on airborne) [14], has been used to test the algorithm. It has 224 contiguous spectral channels with wavelengths from 400 to 2500 nanometers and covers a harbor area of San Diego (USA), as shown in Figure 3. Some bands of the spectrum in which the water vapor absorption is high, have been eliminated so the final extracted hyperspectral image has size 600×300 pixels with 181 bands. In each band, the reported value represents the spectral radiance.

Performing the method described in Section II-A, a number of classes $k = 5$ have been identified. The resulting clustered image is shown in Figure 4. It is important to note that the clustering of the image can be executed just one time, during the first monitoring of a given area.

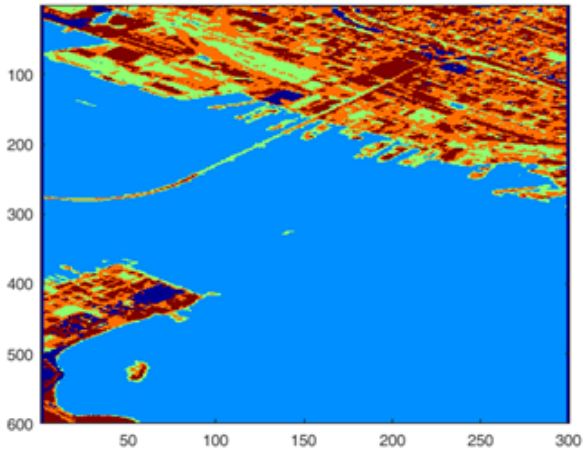

Figure 4. Result after the clustering algorithm.

The spectral signatures relating to the sea, cement and asphalt are supposed to be set by the user and stored in memory, with their median and IQR values. A rectangular window of size 100×100 pixels slides on the clustered image and the algorithm computes the number of pixels belonging to each class within the block window.

A block is classified as ROI if the pixels belonging to the "class sea" occupy from 10% to 90% of the total area and those belonging to "cement" and "asphalt" occupy from 2% to 40% (these values have been calculated through some experimental trials). Four ROIs have been identified whose coordinates are stored and sent to the decoder for a proper reconstruction. The rest of the image, not identified as ROI, is considered background. The bit rate to associate to ROIs and the bit rate to associate to the background have been estimated after an optimization procedure. Initially the maximum value of bpp is taken and then a series of combinations is considered, selecting the best one on the basis of the minimum distortion effect on the ROIs. The results in terms of entropy using different bpp combinations for each ROI, are shown in Table II.

A distortion measure commonly used to evaluate the compressed image quality compared to the original one is the distortion index (DI), defined as:

$$DI = \frac{1}{\frac{1}{ABZ}\sum_{x=0}^{A}\sum_{y=0}^{B}\sum_{z=0}^{Z}[I_{or}(x,y,z) - I_{rec}(x,y,z)]^2} \quad (8)$$

TABLE II. ENTROPY VALUES FOR EACH IDENTIFIED ROI

ROI1 bpp	ROI2 bpp	ROI3 bpp	ROI4 bpp	BG bpp	Entropy-ROI1	Entropy-ROI2	Entropy-ROI3	Entropy-ROI4	Entropy-BG
6	6	6	6	4	2.6756	2.4787	2.1821	3.2062	1.5267
10	10	10	10	4	6.0831	5.7822	5.4600	6.9111	1.5267
12	12	12	12	4	8.0538	7.7570	7.4218	8.9016	1.5267
12	12	12	12	10	8.0538	7.7570	7.4218	8.9016	5.9627

Where A, B, Z represent the three dimensions of image, I_{or} is the original image and I_{rec} is the reconstructed image. Figure 5 shows the results obtained in terms of distortion index and transmitted data volume, in several cases of compression. To make a comparison, the standard JPEG2000 has been applied to the same image. As shown, there is a considerable difference in the distortion level between the implemented method and JPEG2000 compression. This difference can be explained by the fact that the distortion has been calculated only in the region of interest, instead of averaging the result of the distortion that would occur between the ROIs and the BG. The reason of this choice is that the comparison would be otherwise unfavorable to the proposed framework, but at the same time it would be not consistent with the purpose of the algorithm, i.e., the transmission of a reduced data volume to the ground. From the two curves in Figure 5 it is possible to note that, for the same distortion index, the data volume to be transmitted to the Ground Station with the implemented method applied to the ROI is much lower than that obtained with JPEG2000. On the other hand, a drawback of the algorithm is that it compress the part of the image that is not of interest to the user (BG) with a low quality level.

In Figure 6, it is shown an example of reconstructed image (for the 50 band), where the four ROIs are quantized with 10 bpp and the background with 4 bpp.

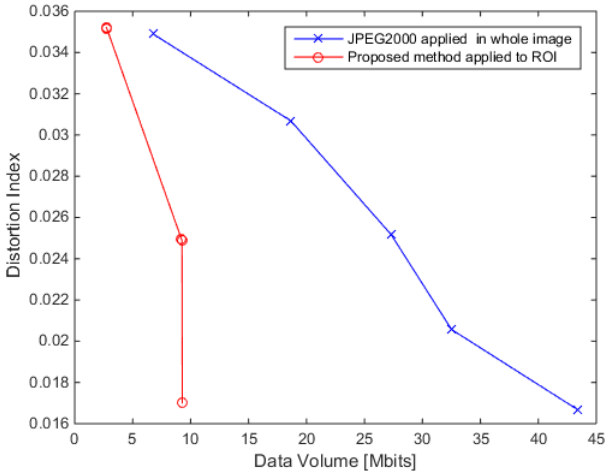

Figure 5. Relationship between data volume and distortion index in JPEG200 (blue) and the proposed method (red).

IV. CONCLUSIONS

The proposed method allows a reduction of the on-board data volume produced by an hyperspectral acquisition system and a better distortion index, at least in the regions of interest, when compared to existing systems in use (e.g., JPEG2000).

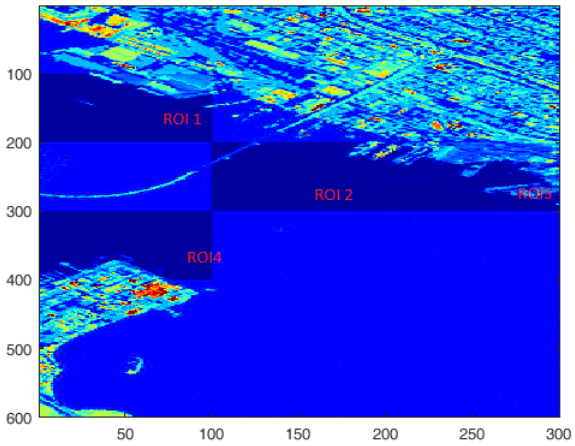

Figure 6. Reconstructed image (relating to 50 band of the cube), 10 bpp for ROIs and 4 bpp for background

These features make the algorithm suitable for real- time (or near real-time) detection of events and recognition of targets. This system works by performing three principal steps:

1) a clustering algorithm that automatically segments the image;
2) an user-oriented choice of the segments and, through this choice, the identification of ROIs;
3) a simpler and faster compression algorithm than the JPEG2000. This is a ROI-based method and it uses wavelet transform and entropy encoding.

To improve the general framework, future developments could include:

- automatic selection of some critical parameters such as the proper number of bits per pixel for ROIs and BG;
- parallelization of the algorithm and its implementation on a Graphic Processor Unit (GPU).
- tests in different scenarios, such as fast identification of oil spills and detection of emergency situations.

ACKNOWLEDGMENT

The research has been carried on in partnership with industrial companies, in the framework of the National Research Project (PON) “Apulia Space”, ID $PON03PE\_00067\_6$.

REFERENCES

[1] H .F. Grahn and P. Geladi, Techniques and Applications of Hyperspectral Image Analysis Chichester, U.K.: Wiley, 2007.

[2] S. Mallat, A Wavelet Tour of Signal Processing, San Diego, CA: Academic, 1998.

 ISBN: 978-1-61208-487-9

[3] D. S. Taubman and M. W. Marcellin, JPEG2000: Image Compression Fundamentals, Standards, and Practice, Kluwer, 2001.

[4] A. K. Jain, M. N. Murty, and P. J. Flynn, "Data clustering: a review," ACM Computing Surveys, Vol. 31(3), 1999, pp. 264-323.

[5] K. Koonsanit, and C. Jaruskulchai, "A simple estimation the number of classes in satellite imagery," ICT and Knowledge Engineering (ICT & Knowledge Engineering), 2011 9th International Conference on IEEE, January 2012, pp. 124-128.

[6] A. S. Kurani, D. H. Xu, J. Furst, and D. S. Raicu, "Co-occurrence matrices for volumetric data," 7th IASTED International Conference on Computer Graphics and Imaging, Kauai, USA, August 2004, pp. 447-452.

[7] A. Cohen, I. Daubechies, and J. C. Feauveau, "Biorthogonal bases of compactly supported wavelets," Commun. Pure and Appl. Math., Vol. 45(5), May 1992, pp. 485-560.

[8] M. Antonini, M. Barlaud, P. Mathieu, and I. Daubechies, "Image coding using wavelet transform," IEEE Trans. Image Process., Vol. 1(2), Apr.1992, pp. 205-220.

[9] J. Yu, "Advantages of Uniform Salar Dead-Zone Quantization in Image Coding, Communications, Circuits and Systems," ICCCAS 2004, Vol.2, June 2004, pp. 805-808.

[10] C. E. Shannon, "A mathematical theory of communication," ACM SIGMOBILE Mobile Computing and Communications Review, Vol. 5(1), 2001, pp.3-55.

[11] D. A. Huffman, "A method for the construction of minimum-redundancy codes," Proceedings of the IRE, Vol. 40(9), 1952, pp. 1098-1101.

[12] M. Mitzenmacher, "On the Hardness of Finding Optimal Multiple Preset Dictionaries," IEEE Transaction on Information Theory, Vol. 50(7), July 2004, pp. 1536-1539.

[13] A. Shahbahrami, R. Bahrampour, M. S. Rostami, and M. A. Mobarhan, "Evaluation of Huffman and Arithmetic Algorithms for Multimedia Compression Standards," International Journal of Computer Science, Engineering and Applications (IJCSEA), Vol. 1, August 2011, pp. 34-47.

[14] National Aeronautics and Space Administration (NASA). Airborne visible/infrared imaging spectometer (aviris), retrieved: February, 2016. URL http://aviris.jpl.nasa.gov/

 ISBN: 978-1-61208-487-9