

# **A METHOD FOR COMBINING SPATIAL AND HYPERSPECTRAL INFORMATION FOR DELINEATION OF HOMOGENOUS MANAGEMENT ZONES**

**S.Cohen and O.Levi**

*Department of Industrial Engineering and Management,  
Ben-Gurion University of the Negev,  
P.O.B 653, Beer-Sheva 84105, Israel*

**Y.Cohen and V.Alchanatis**

*Institute of Agricultural Engineering, ARO, Israel,  
Bet-Dagan 50250, Israel*

## **ABSTRACT**

Hyperspectral (HS) remote sensing is a constantly developing field. New remote sensing applications of different fields constantly appear. The possibility of acquisition information about an object without physical contact is spanning new opportunities in many fields and for precision agricultural in particular. These opportunities demand constant improvement and development of new analysis approaches and algorithms, which will exploit the advantages of HS imaging. Though hyperspectral imaging has been transformed in the last 30 years, still most available HS data processing algorithms analyze the data based on the spectral information exclusively and do not treat the data as an image. As for the methods that use spatial and spectral information, most of them use the information serially, as a two-step processing technique. The first processing steps focus on the spectral information, and the second step focus on the spatial one. In this paper we present a classification method which uses both spectral and spatial information simultaneously. A comparison between classification results of different processing approaches is presented. A hyper spectral image of an experimental field of potatoes, where five different levels of fertilizers were applied, was used to test and compare the algorithms' performance. Three pre-processing models were applied on the hyperspectral image: calculation of a spectral index, principle component analysis and the full HS image. It was found that the principle components and the whole spectra performed better than the spectral index. Furthermore, from the comparison of the classification results, the new approach yielded better classification results than other similar methods reported in the literature.

**Keywords:** Nitrogen, management zones, hyperspectral images, classification, spatial/spectral processing

## INTRODUCTION

Modern agricultural crop's yield relies on close monitoring and adequate treatment. Efficient resources management can result in increased yield and prevent environmental damage. There are many examples for the importance of close monitoring, such as in nitrogen supply. Nitrogen is one of the most important ingredients in fertilization, which affect the chlorophyll level of the plant. When nitrogen supply surpasses vegetation's nutritional needs, the excess is eliminated by runoff and water infiltration leading to pollution of aquatic ecosystems (i.e., eutrophication) (Haboudane *et al.*,2002,Cohen *et al.*,2010). This nitrogen loss represents also an economic loss for the farmers. However, inappropriate reduction of nitrogen supply could result in reduced yields and lead to substantial yield losses. The common monitoring tools used nowadays are based on point sampling of biophysical and biochemical properties of the crop. When dealing with large commercial fields, point sampling does not represent the spatial variability.

Hyperspectral imaging plays a principal role in the remote sensing world for two decades (Goetz *et al.*,1985). Hyperspectral datasets unique form adds new challenges to basic computer vision problems, such as segmentation and classification. One obstacle set by hyperspectral imaging is its data size. Hyperspectral image is usually built from hundreds of images, each represents spectrum bandwidth. This enormous data size results in a long computational processing time. Different solutions were proposed in order to use the hyperspectral data. One is to use spectral indices, which are combined from a small number of wavelengths designed to identify certain reaction (Li *et al.*,2010). Second is the use of mathematical manipulations in order to reduce the hyperspectral image dimension with minimum loss of information. The common mathematical tool is principle component analysis (PCA). There are other methods to confront the dimensionality problem, such as analyzing the full hyperspectral image with the use of parallel computing, which can excel the running time. No matter what portion of the hyperspectral are used, most available hyperspectral data processing techniques focus on analyzing the spectral information without incorporating spatial information (Plaza *et al.*,2009). Hyperspectral data are usually not treated and processed as images, but as unordered listings of spectral measurements with no spatial arrangement (Tadjudin and Landgrebe,1998). The importance of analyzing both spectral and spatial information has been identified and there were some researches done in this direction but they were applied mostly for land-use classification (Plaza *et al.*,2009). Precision agriculture can benefit from methodologies that utilize the spectral as well as the spatial information in hyperspectral images as a tool to delineate management zones in the field scale. One common approach utilizing the spatial information is to perform two step processing. The first step is a spectral based analysis, and the second is refinement of the previous step results using spatial information (Landgrebe,2003). The objective of this study was to explore the contribution of the spatial information to the delineation of reliable management zones based on aerial hyperspectral images. For that a new classification method that uses the spatial and spectral information simultaneously

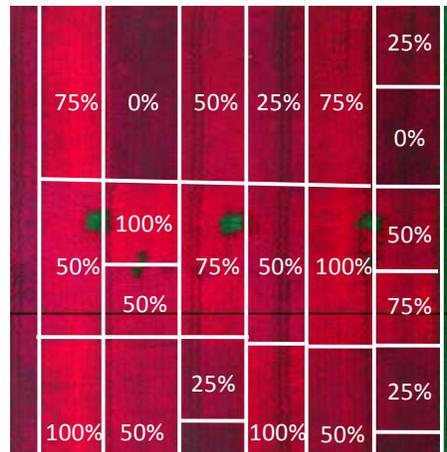
is introduced for hyperspectral classification of nitrogen levels. In addition an existing two-step analysis, introduced by Lark (1998), was adjusted and applied for hyperspectral analysis. Furthermore, the superiority of the complete spectral information over spectral indices was presented for the same purpose.

### DATA SET AND DATA PROCESSING

To evaluate the performance of the classification methods for agricultural applications a hyperspectral image of potato field, which was under different nitrogen treatments, was used. The hyperspectral image was acquired on May 25, 2007 using a push-broom AISA system in the range of 400-1000 nm, with 420 bands with spectral resolution of 1.3 nm. The image was acquired from a 500 meters height and has 1 m spatial resolution. Pre-processing of the image included selection of every second band of the original 420 bands and smoothing of the 210-bands spectra of the new cube, with a 15-points window (Alchanatis and Cohen,2011).

Different types of hyperspectral data were used: TCARI/OSAVI spectral index which was developed to estimate chlorophyll level in the field (Haboudane *et al.*,2002), principle components of the hyperspectral image (PCA) (Trevor. Hastie *et al.*,2009), which is very often used for dimensionality reduction of a hyperspectral image, and the raw data of the hyperspectral image. Each of the data type was classified using fuzzy C-means (FCM), spatially coherent regions multi-variate classification (SCR) developed by (Lark,1998) and the RDP-BD classification which uses RDP-BD segmentation algorithm (Levi *et al.*,2011).

Figure 1 presents a combination of the bands 750 nm (IR), 670 nm (Red) and 550 nm (Green) of the experimental plot overlaid by the borders of the nitrogen treatments. Treatment denoted 100% in the figure was applied with 400 kg ha<sup>-1</sup> of nitrogen. All other nitrogen treatments were applied with relative amounts to their percentage. More details on the experiment can be found in (Cohen *et al.*,2010).



**Figure 1. Combination of IR (750 nm), Red (670 nm) and Green (550 nm) bands of the experimental plot overlaid by the borders of the N treatments.**

## **CLASSIFICATION METHODS**

In this research a comparison between three different classification methods was done. The classification methods were differed in respect to their spatial and spectral information fusion. The first method used is Fuzzy C-means (FCM), which only uses spectral information. The next two methods are based on the FCM but differ in their spatial and spectral information fusion. The second method used is the spatially coherent regions generation (SCR) developed by Lark in (Lark,1998), which uses the spectral and spatial information serially. The SCR algorithm produces preliminary classification according to the spectral information and, as a second step, updates the classification in respect to the spatial information. The last method used is RDP-BD algorithm which uses the spatial and spectral information simultaneously for segmentation, which is then classified by FCM.

### **Fuzzy C-mean Classification Algorithm**

The FCM (Bezdek and Ehrlich,1984) is a soft clustering algorithm, where each object belongs to a cluster in a certain degree. As a first step, for each object, the distances to all the centers of groups are calculated. The object is classified according to the shortest distance. As a second step the centers of the groups are updated according to the objects that belong to them. These two steps repeat till there is no change in the classification.

In order to use the FCM the user should set in advance the number of clusters. In our experiment the image was classified into three groups: low, medium and high nitrogen level.

### **Spatially Coherent Regions Generation**

The classification algorithm proposed by Lark (Lark and Stafford,1997 ,Lark,1998) is based on the FCM algorithm. Lark's classification is performed in two steps: FCM classification and spatial smoothing. Based on the FCM classification results, a classification vector for each voxel in the image is available. The classification vector consists of a fitting degree to each cluster, where the classification is based exclusively on the spectral information. The smoothing step is done by updating the classification vector of each voxel according to the classification vector of its neighbors. The classification vector is updated by calculating a weighted average according to the neighbors' weights, which are complement to 1, e.g. if the self-vector is set with 0.2 the neighboring vector is set with 0.8. Decrease in the weight of the self-vector increase the smoothing magnitude. In this study two weights were empirically set to the self vector: 0.2 and 0.05. The smoothing radius is set according to the variogram range of the image. The range of the variogram calculated for the experimental site using Geostatistical analysis wizard of ArcGIS 10 (ESRI, Ltd.) and was found to be 18 m (coincided with the treatments width). The SCR algorithm was not designed originally for hyperspectral images but for multi-year yield data with few variants (Lark, 1998). This is the first attempt to use it for the classification of a hyperspectral image.

## **RDP-BD classification**

The RDP-BD classification is based on two steps. The first step is RDP-BD segmentation, where the RDP-BD algorithm uses both the spatial and spectral information (Levi *et al.*,2011). The second step is performing FCM classification over the segmentation results. In order to use the algorithm, one should set the value of the Penalty Weight ( $PW$ ) of the target function, which affects the segmentation results. As the  $PW$  increases the number of segments decreases and the resulted image is smoother with bigger variance.

## **COMPARISON METHODS**

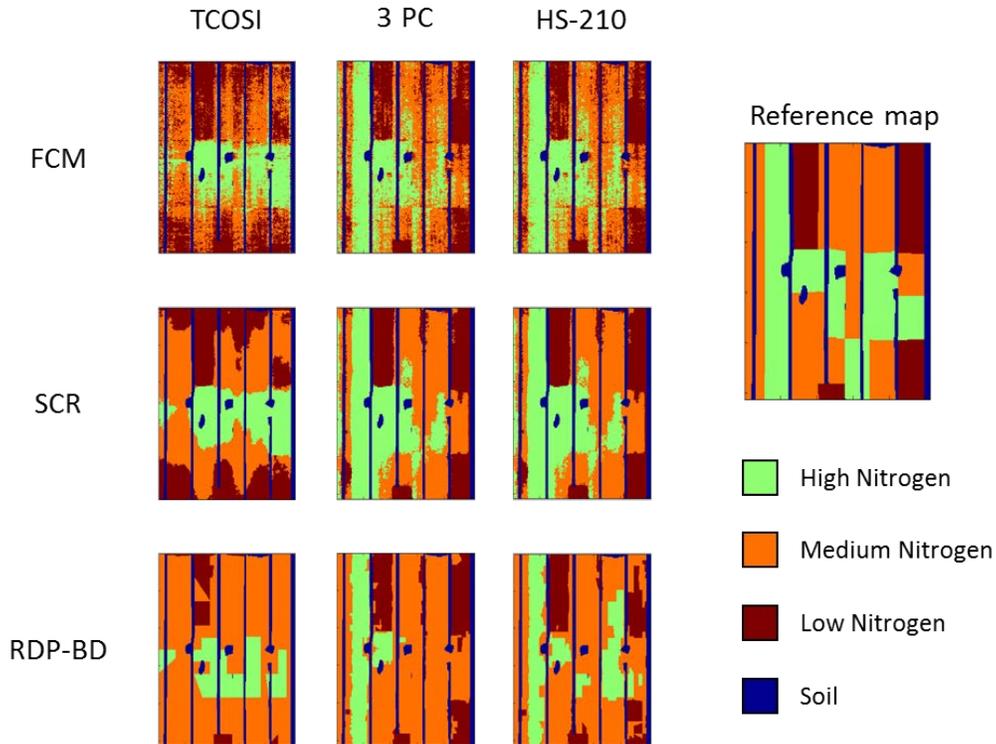
To evaluate the performance of the classification methods a reference map was built (Figure 2). The reference map was built by manually classify the image into four categories: soil, and low, medium and high nitrogen levels. The soil was excluded manually and the classification into nitrogen levels was done according to prior knowledge of the treatments locations with complementary visual inspection. Accordingly, all the three classification methods (FCM, SCR, and RDP-BD) were set to detect three classes and the classes were matched manually to nitrogen levels. Confusion matrices were calculated by overlaying the reference map and the classified images produced by each classification method. Based on the confusion matrices overall accuracy and kappa coefficient (Cohen,1968) were calculated and compared.

In this work two different examinations were made. The first comparison was related to the different data sets: raw hyperspectral image (HS-210), 3-PCs of the hyperspectral image and TCARI/OSAVI spectral index (TCOSI). The second comparison was related to the different classification methods: spectral based classification (FCM) spatially smoothed spectral classification (SCR) and simultaneous spectral and spatial segmentation and FCM classification (RDP-BD).

## **RESULTS**

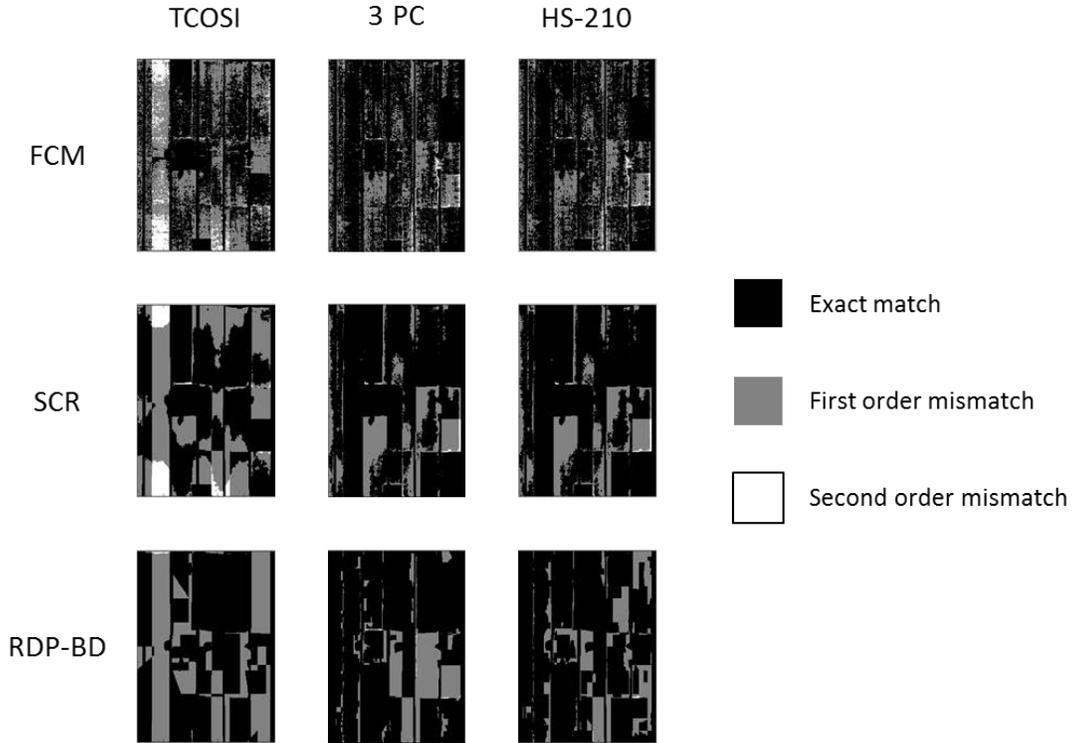
Figure 2 illustrates the classification results of the different methods and data bases. The different colors represent different treatment classes: high nitrogen (green), medium nitrogen level (orange), low nitrogen level (dark red) and soil (blue) which was extracted manually. For the SCR classification result with a weight of 0.05 is presented and for the RDP-BD the 15&140 PWs combination is presented. Both represent the best results of each method. A visual comparison with the reference map leads to two main insights. First, the TCOSI provides inferior accuracy compared with the 3-PCs and the HS-210. There is a big difference between the reference map and the TCOSI's classification. For example, all 3 classification methods, which were based on the TCOSI, classified the high nitrogen treatment to the left of the map as medium to low nitrogen levels. The second main observation is that the integration of spatial analysis with the spectral classification divided the experimental field into big segments, each associated with one class. Except for the TCOSI result, in the two complete

hyperspectral data results, the segments and their associated classes are with general accordance to the reference map (qualitative measures are given further in this section). The only difference between the two methods is their borders feature. The SCR provides smooth borders while the RDP-BD divides the images with straight lines.



**Figure 2. A reference map of nitrogen levels and classification results of the different data bases and methods**

Figure 3 illustrates the classification accuracy level of each classification method and data set. Each image in Figure 3 is the matching result between the classified image and the reference map. Black represents exact match, meaning that the classification is accurate. Gray represent a first order mismatch, meaning that there is one class difference between the reference map and the results, i.e. high or low nitrogen levels were classified as medium level. The white color represents a second order mismatch (considerable mistake), i.e. high and low nitrogen levels were switched. In overall, there are few second order mismatches mainly in the classifications which were based on TCOSI.



**Figure 3. Classification match and first and second mismatches compared to the reference map**

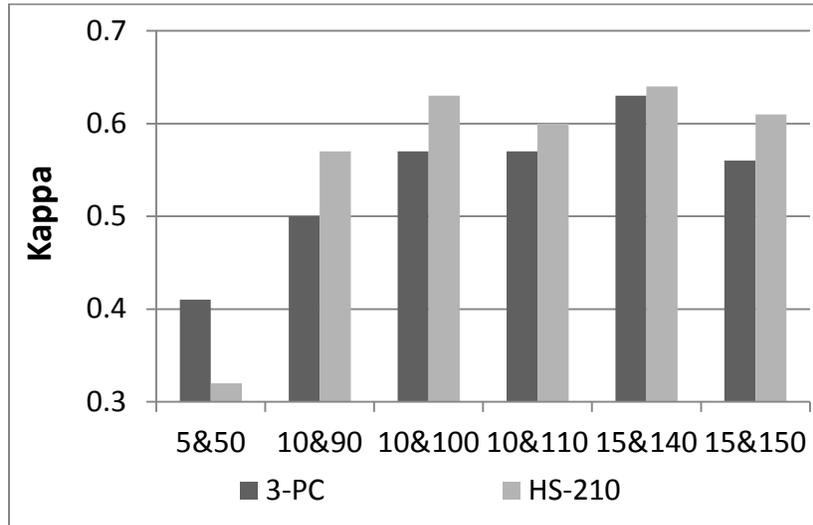
Beyond the visual inspection the overall accuracy and kappa coefficients provide quantitative measures for the classifications performance (Table 1). For SCR and RDP-BD accuracy and kappa measures are provided to the best results. The 3-PC and HS-210 resulted in better accuracies than the TCOSI using all three classification methods. Kappa (k) coefficient indicates for poor classification performance obtained by the TCOSI (less than 0.4) (Landis and Koch,1977). The performance of other two data sets is moderate for the FCM (k=0.4-0.6) and substantial (k>0.6) (Landis and Koch,1977). The SCR and RDP-BD present comparable results.

**Table 1. Overall accuracy of the three classification methods using the three different data-sets; \*OA stands for overall accuracy**

	TCOSI		3-PC		HS-210	
	OA*	Kappa	OA	Kappa	OA	Kappa
<b>Fuzzy C-mean</b>	0.59	0.25	0.71	0.54	0.71	0.54
<b>SCR</b>	0.57	0.33	0.77	0.63	0.77	0.63
<b>RDP-BD</b>	0.55	0.19	0.78	0.63	0.78	0.64
<b>Average</b>	0.57	0.26	0.75	0.60	0.75	0.60

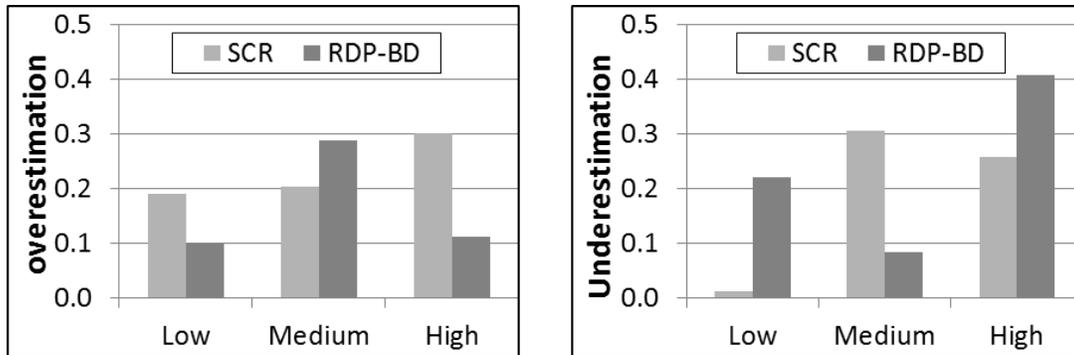
An exhaustive examination of various PWs combinations used to apply the RDP-BD classification shows (Figure 4) that out of the 6 combinations, the result

of 5&50 combination is much lower and it might be explained by its lower PWs. All other combinations provided similar results. The SCR is also sensitive to the weight of the self-vector which controls the smoothing magnitude. When SCR was applied with a weight of 0.2 the classification results was very similar to the FCM (Overall accuracy of 0.73 vs. FCM overall accuracy of 0.71) and the smoothing effect was negligible (data not shown).



**Figure 4. Kappa coefficient for the 6 PWs combinations of the RDP-BD**

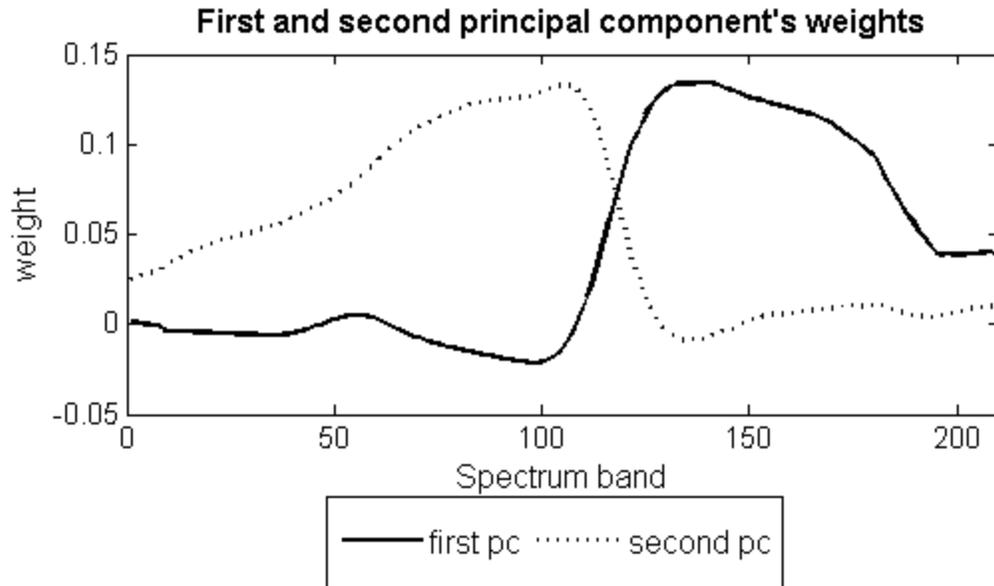
Further examination of the classification performance of the SCR and the RDP-BD was conducted through the calculation of under- and over-estimation of individual nitrogen level (HS-210; Figure 5). Despite their comparable overall accuracy and kappa they are different in classification of individual classes. The SCR suffers from higher overestimation especially for the low and high nitrogen levels and the RPD-BD suffers from more underestimation of these two levels. In general these opposite errors entail similar implications. Overestimation of low nitrogen level and underestimation of high level may cause to suboptimal application of nitrogen and yield loss while the opposite errors cause to over application of nitrogen, damage to the yield quality (less dry matter in the tubers) and nitrogen leaching to the soil.



**Figure 5. Overestimation (left) and underestimation (right) of individual nitrogen levels using the SCR and RDP-BD based on HS-210**

### DISCUSSIONS AND FUTURE DIRECTIONS

In this work the effect of the hyperspectral data set and the potential contribution of the spatial analysis on nitrogen level classification performance were examined. Comparison between the different data sets has shown that the TCOSI provided the poorest performance (Table 1). The 3-PC and the complete spectral curve (HS-210) demonstrated comparable and significantly better performance. The TCOSI is a spectral index which is based only on four specific bands while the 3-PC and the HS-210 use the whole spectral range. These results correspond with other studies that have shown the superiority of the whole spectral range upon using spectral indices or few selected bands in evaluating LAI, biomass, water content and chlorophyll and nitrogen content (Bonfil et al., 2005; Cohen et al., 2010; Kempeneers et al., 2005; Lee et al., 2004; Pimstein et al., 2007). A closer look at bands' weights composing the 2 first PCs further supports this assumption as it shows that the dominant bands are spread over the entire spectrum (Figure 6). In this study a simple classification was used. Better classification can be improved by using chemometric methods like the partial least square regression and wavelet analysis (Alchanatis and Cohen, 2011) but an appropriate incorporation of these methods to classify a hyperspectral image is still a challenge.



**Figure 6. Weights of each spectrum band when composing the two main PCs.**

All of the studies that analyzed continuous spectral range for biophysical and biochemical properties of the canopy used point spectral measurements and not hyperspectral images (Alchanatis and Cohen, 2011). In this study the contribution of spatial analysis to spectral analysis of continuous spectrum of canopy extracted from a hyperspectral image was exemplified. Spatial analysis assists with delineating homogeneous zones in a field and classifying them according to their biophysical/biochemical properties. Three levels of incorporation of spatial analysis were compared: no spatial analysis, spatial analysis for retrospective smoothing of spectral classification and a simultaneous spectral-spatial segmentation. Classification measures per-pixel showed an improvement of the spectral and spatial analysis over the spectral classification (Table 1). But the more pronounced contribution was segmentation of the images into homogeneous zones which fit to the desire to divide a field into homogeneous zones for variable rate fertilization. The comparison between the SCR and the RPD-BD did not reveal an advantage of the simultaneous integration of the spectral and spatial information over the two step analysis.

Both methodologies are sensitive to the weights that control the segmentation or smoothing magnitude. Since there is no automatic way to set the optimal weights for SCR and *PWs* for the RPD-BD appropriate application would require iterative operations. For the RPD-BD the variance change rate was successfully used for optimal selection of *PWs*. A similar approach might be used for the SCR.

The spatial analysis contribution to the division of the field according to nitrogen levels was proved based on a small experimental field. In the future it should be examined for other biophysical and biochemical properties like biomass and water content and for a field scale. These methodologies would assist with testing the efficiency of variable rate nitrogen fertilization as on one hand they

utilize the informative spectral data to properly classify nitrogen level in the pixel scale and on the other hand delineate management zones in the field scale.

### **ACKNOWLEDGMENT**

This project was supported by the Israeli Space Agency, Israeli Ministry of Science and also by Research Grant Award No. CA-9102-06 from BARD-AAFC - The United States - Israel Binational Agricultural Research and Development Fund and Agriculture and Agri-Food, Canada.

The authors wish to express their appreciation of the vital contributions of Gadi Hadar and Ran Ferdman, potato growers from Kibbutz Ruhama, who provided land, seed and pesticide, and managed the irrigation for the project at no cost. The field experiments could not have been performed without the collaboration of Zion Dar from the Israeli extension and Yossi Sofer from Haifa Chemicals, Ltd.

### **REFERENCE**

Alchanatis, V., Cohen, Y., 2011. Spectral and Spatial Methods for Hyperspectral Image Analysis for estimation of Biophysical and Biochemical Properties of Agricultural Crops. In: Thenkabail, P., Lyon, J.G., Huete, A. (Eds.), *Hyperspectral Remote Sensing of Vegetation*. Taylor and Francis Press, pp. 265-307.

Bezdek, J.C., Ehrlich, R., 1984. FCM: The fuzzy c-means clustering algorithm. *Comput. Geosci.* 10, 191-203.

Cohen, J., 1968. Weighted kappa: Nominal scale agreement provision for scaled disagreement or partial credit. *Psychol. Bull.* 70, 213.

Cohen, Y., Alchanatis, V., Zusman, Y., Dar, Z., Bonfil, D., Karnieli, A., Zilberman, A., Moulin, A., Ostrovsky, V., Levi, A., 2010. Leaf nitrogen estimation in potato based on spectral data and on simulated bands of the VEN $\mu$ S satellite. *Precision Agriculture* 11, 520-537.

Donoho, D.L., Huo, X., 2002. Beamlets and Multiscale Image Analysis. *Multiscale and Multiresolution Methods: Theory and Applications*, 149.

Goetz, A.F.H., Vane, G., Solomon, J.E., Rock, B.N., 1985. Imaging spectrometry for earth remote sensing. *Science* 228, 1147.

Haboudane, D., Miller, J.R., Tremblay, N., Zarco-Tejada, P.J., Dextraze, L., 2002. Integrated narrow-band vegetation indices for prediction of crop chlorophyll content for application to precision agriculture. *Remote Sens. Environ.* 81, 416-426.

Kim, M.S., Daughtry, C., Chappelle, E., McMurtrey, J., Walthall, C., 1994. The use of high spectral resolution bands for estimating absorbed photosynthetically

- active radiation (A par). In: Anonymous Proceedings of the 6th Symposium on Physical Measurements and Signatures in Remote Sensing, 299-306.
- Landgrebe, D.A.,2003. Signal Theory Methods in Multispectral Remote Sensing. Wiley-Interscience.
- Landis, J.R., Koch, G.G.,1977. The measurement of observer agreement for categorical data. *Biometrics* , 159-174.
- Lark, R.,1998. Forming spatially coherent regions by classification of multi-variate data: an example from the analysis of maps of crop yield. *Int. J. Geogr. Inf. Sci.* 12, 83-98.
- Lark, R., Stafford, J.,1997. Classification as a first step in the interpretation of temporal and spatial variation of crop yield. *Ann. Appl. Biol.* 130, 111-121.
- Le Maire, G., Francois, C., Dufrene, E.,2004. Towards universal broad leaf chlorophyll indices using PROSPECT simulated database and hyperspectral reflectance measurements. *Remote Sens. Environ.* 89, 1-28.
- Levi, O., Cohen, S., Mhabary, Z.,2011. Effective Hyper-Spectral Image Segmentation Using Multi-scale Geometric Analysis. *International Journal of Computer Information Systems and Industrial Management Applications* 3, 836.
- Li, F., Miao, Y., Hennig, S.D., Gnyp, M.L., Chen, X., Jia, L., Bareth, G.,2010. Evaluating hyperspectral vegetation indices for estimating nitrogen concentration of winter wheat at different growth stages. *Precision Agriculture* 11, 335-357.
- Plaza, A., Benediktsson, J.A., Boardman, J.W., Brazile, J., Bruzzone, L., Camps-Valls, G., Chanussot, J., Fauvel, M., Gamba, P., Gualtieri, A.,2009. Recent advances in techniques for hyperspectral image processing. *Remote Sens. Environ.* 113, S110-S122.
- Tadjudin, S., Landgrebe, D.,1998. Classification of high dimensional data with limited training samples.
- Trevor. Hastie, Robert. Tibshirani, JH (Jerome H.) Friedman,2009. *The Elements of Statistical Learning*. Springer.