

IMPROVING CLASSIFICATION OF MULTISPECTRAL IMAGES BASED ON SELECTED RELATIONS

L. Cohen¹, O. Almog¹, M. Shoshany^{1*}

¹ Civil Engineering Faculty, Department of Transportation and Geo-Information, Technion – Israel Institute of Technology, Haifa, Israel. (almogo, maximsh)@technion.ac.il; shonilior@gmail.com

Commission III, Workgroup IV

KEY WORDS: spectral relations, multi-spectral classification, spectral significant features.

ABSTRACT:

A novel classification technique based on definition of unique spectral relations (such as slopes among spectral bands) for all land cover types named (SSF Significant Spectral Features) is presented in the article.

A large slopes combination between spectral band pairs is calculated and spectral characterizations that emphasizes the best spectral land cover separation is sought. Increasing in dimensionality of spectral representations is balanced by the simplicity of calculations. The technique has been examined on data acquired by a flown hyperspectral scanner (AISA). The spectral data was narrowed into the equivalent 8 world-view2 channels. The research area was in the city of “Hadera”, Israel, which included 10 land cover types in an urban area, open area and road infrastructure. The comparison between the developed SSF technique and common techniques such as: SVM (Support Vector Machine) and ML (Maximum Likelihood) has shown a clear advantage over ML technique, while produced similar results as SVM. The poorest results of using SSF technique was achieved in an herbaceous area (70%). However, the simplicity of the method, the well-defined parameters it produces for interpreting the results, makes it intuitive over using techniques such as SVM, which is considered as a not explicit classifier.

1. INTRODUCTION

Technological advances in the field of visual sensors increased the availability of multi or hyper spectral information in a growing number of bands while continuing to decrease the band width. The advanced capabilities in visual acquisition sensors is main engine of mapping processes relying on multiband imageries, which enable extracting and characterizing objects features, while maintaining high levels of precision and accuracy. Increasing the volume of information had brought to new challenges and requires the development of appropriate algorithms. However, using all acquired information does not necessarily contribute to the improvement of image classification in pixel level or extracting image objects. Aside the need to choose best representative image bands, the spectral reflectance values are affected by the acquisition configuration and by the environmental conditions existed such as: flux density effect, topography, scene illumination and atmospheric behaviour. Being subjected to those unknown parameters enhances the difficulties in identifying land cover, and generally increases spectral confusion. Therefore, identification of significant spectral properties in land cover types is needed. Furthermore, the most common classification methods are statistical, and do not differentiate the information elements that contribute best to the spectral uniqueness increasing of land cover types. In cases of low relative spectral variance between types of land covers, an efficient separation will not be achieved.

Scientific literature divides the supervised classification capability into two main categories: (1) Parametric (explicit) statistical capabilities such as Maximum likelihood estimator, which calculate statistical parameters (average and co-variance) that describe each material group; (2) Non parametric (non-explicit) statistical capabilities such as Support Vector machine (SVM) that searches for hyperplane (in the dimensionality of the

spectral bands) that geometrically separates two material groups. Mostly, SVM is considered to reach high accuracies and reliabilities classification results in comparison to another classifiers. However, SVM classifier acts as a “black box” that its use does not contribute the understanding of the spectral feature relations (Auria and Moro, 2008). Moreover, both methods require a large set of verified training sample, that is not always available or is not reliable enough (Rasti et al, 2020).

In this research, we developed the SSF Spectral Significant Features classification method, that seeks and enhances the most unique features in different land covers signatures by: a) Increasing inner dimensionality with displaying the signatures in slopes domain based on first derivative calculation between each pair of spectral bands; b) Usage of Sum Of Differences algorithm to detect the bands which contribute most to increase spectral uniqueness compared to all other signatures.

Our assumption is, that combining the abovementioned with dynamic and iterative decision rules to classify imagery pixels, will produce high rates of precision and accuracy in classification products. Slope domain is adopted from derivative analysis (Cimtay, 2017) and multi resolution analysis methodology implemented mostly by wavelet transform (Almog et al, 2008). Wavelet coefficients are calculated based on the derivative (in multiple scales) between two adjacent values in the original intensity reflectance signature. Shoshany et al, 2009 compared classification results of different land covers based on different signatures domain. The original reflectance signatures were transformed into wavelet domain, PCA (Principal Component Analysis) and SAM (Spectral Angular Mapper). They showed that transforming spectral signatures into wavelet domain significantly increased the correct assignment rates of signature into its correct land cover type. That, due to averaging the signature band’s noise level. In this research we examined the accuracy and reliability of the proposed method by comparing its

* Corresponding author

performances to the most common classification methods: (1) SVM; (2) ML.

2. SPECTRAL RELATIONS FOR IMPROVING OBJECTS CLASSIFICATION

Usage of spectral relations in classification processes leads to decreasing of multiplications effects such as flux density on the radiometric calibration (as presented in Law and Nichol, 2004 and Almog et al, 2008). Mostly, it is used in calculation of vegetation indices or specific land cover. Scientific literature describes the usage of derivative operator as an emphasize of spectral separation between different land cover's signatures. Fu et al, 2006 presented increasing accuracy and reliability rates in classification processes using transformation of signatures into the derivative domain, by: (1) Increasing the statistical difference between spectral signatures by increasing the number of bands; (2) Improving spectral separation. Rud et al, 2006 used ratios characterizations in specific bands in shrubs signatures that emphasized the uniqueness in spectral features and improved classification accuracy significantly by transforming the reflectance signature into those ratios. Cimtay, 2017 presented increasing differentiation capabilities between same plant signatures acquired under different conditions by transforming the signatures into derivative domain. Contrarily, using derivative domain requires deep understanding of specific spectral features and large number of spectral combinations, as described by Torrecilla and Piera, 2009 and Sun et al, 2008

3. METHODOLOGY

The developed methodology is composed of two main stages: (1) Transforming the spectral signatures domain into slopes domain that emphasizes spectral uniqueness; (2) Detecting unique spectral bands (SSF) to achieve better separation between different types of land covers. The new spectral domain is less sensitive to additional noises and emphasizes uniqueness between spectral ratios.

The methodology has 7 stages, and illustrated in figure 1:

(a) Definition of typical signature and its statistical parameters for each land cover's type. Land cover types can be achieved from several resources such as: spectral measurements, spectral libraries, expert knowledge, or previous unsupervised classification process.

(b) Slope calculation between each band (i) and all other bands in the signature (j) for each typical land cover (k).

$$S_{kq} = \left| \frac{R_i - R_j}{\lambda_i - \lambda_j} \right| \quad (1)$$

Where, q – the combination index of the i^{th} and j^{th} spectral combination, where $i \neq j$

(c) Calculation of sum of differences between each land cover's signature (k) and other signatures from any slope combination (q).

$$SODkq = \sum_{i=1}^m (Skq - Slq) \quad (2)$$

(d) Ordering each combination signature according SODkq from higher to lower.

(e) Choosing the band combination with the maximum SOD for each land cover's signature (determined empirically).

(f) Transforming the spectral image into slope domain according to the only combination that represent unique spectral characteristics.

(g) Assigning each pixel (signature) to its correct land cover's type based on minimum sum of differences at the selected bands combination.

The research examined two main aspects: (1) Efficiency in using slopes combination rather than reflectance information; (2) Efficiency of using SSF technique in comparison to ML and SVM. For this purpose, classification processes based on 3 classifiers were applied on both reflectance data and spectral slopes.

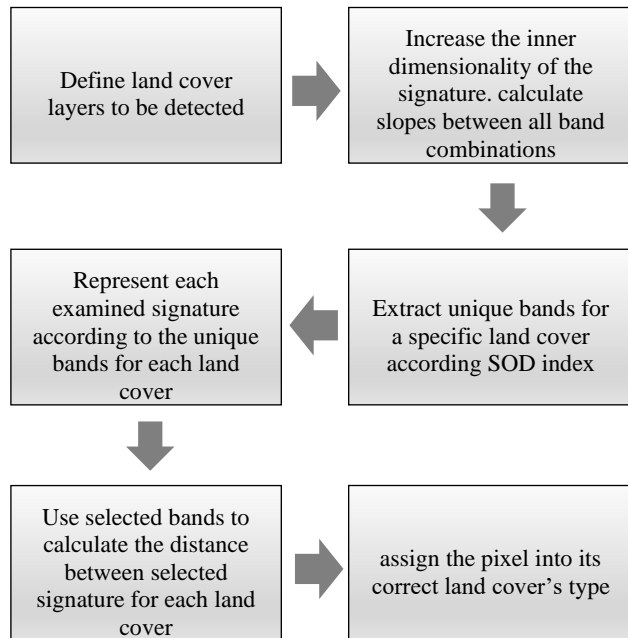


Figure 1. General methodology

4. EXPERIMENTAL CONFIGURATION

A flown hyperspectral scanner (AISA) was used as multispectral source for this research around "Hadera" city in the north of Israel. Weather conditions were clear. The AISA image has 156 spectral bands in the range of 400-2400 nanometers, with band width between 9-13 nanometers. The spectral data was narrowed into the equivalent 8 world-view2 channels (380, 465, 550, 600, 655, 710, 840 and 970 nm). Weather conditions were clear. The spatial resolution is approximately 1.7·1.7 m² per pixel. The scene represents a flat topographic area with characterizations of urban area and open area containing roads, shrubs, and soil. The selected land cover types are asphalt, trees, shrubs, grass, white roofs, bright soil, kurkar, concrete, dark soil and red tile roofs. Figure 2, presents the scene area and selected ROI (Region of Interest) for each land cover, created with ENVI software (©Harris ltd).

Figure 3 presents the average signatures of the 10 land cover's types calculated from the ROI's. it can be seen from figure 3, that the spectral distance between different land cover's type is low (trees and bushes, bright soil and red brick and asphalt and dark soil). This would lead to high confusion when assigning each signature to its correct land cover's type.



Figure 2. Scene area and ROI used as training samples

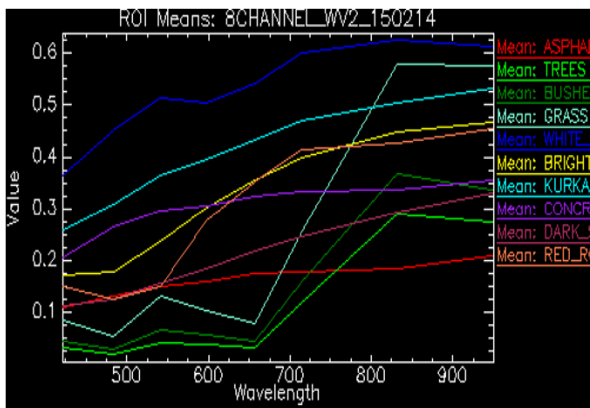


Figure 3. Mean spectral signature for each land cover

5. RESULTS

5.1 Synthetic experiment

Synthetic experiments presented in this research showed that controlled exposure of spectral library signatures to different acquiring effects flaws less the classification product of SSF method compared to the statistical method product. We show that the decrease of precision and accuracy indexes is slower in the SSF product against the statistical method product.

Next, we exam the classification products of common SVM and ML classifiers against the SSF classification product, both in the spectral and in slopes domain. The same training and validation areas were used to create a fair comparison.

Examination of classification products shows that SVM method got the most accurate results (86%), and the SSF method achieved 77%, much better than the ML classification product (69%).

Another comparison was made between the methods products in the slopes domain. Unlike the spectral domain results, the SSF method presents the highest results (88%), against 85% of SVM and 74% of ML. The substantial improvement is an outcome of: a) Increasing the number of bands in the slopes signatures that contributes to detection of significant features; b) Enlarging the range of signatures 'pseudo-reflectance' values, which increases the efficiency in defining classification according to supportive decision rules.

In addition to the above, we found that methods, which classification is based on the land cover spectral properties, is better compared to classification based on statistically calculated parameters.

It was also proven that the SSF classification method produces provides more than only efficient results, both general and relatively, comparing to common statistical methods. Furthermore, it was proven that representing land covers signatures based on selected features enhances their spectral uniqueness and the separability between them relatively and increased the accuracies and reliabilities in classification processes.

5.2 field experiments

Figure 4 illustrate spectral signature after transformed into slope domain (note that there are $8 \cdot 7/2 = 28$ combination bands). While reflectance intensity varies between 0.03-0.6, the slopes range varies between 0.2-4.

Red band (the 3rd band) has the strongest spectral variance between land cover in spectral domain, and the 26th band has the strongest variance in the slope domain. The most significant difference between the two domains is in the information content reflected in the variance representation. Most of spectral signature (excluding vegetation signatures) are characterized by large range of adjacent bands with low spectral variance. The signatures in slope domain show significant fluctuations excluding asphalt, dark soil and concrete. Spectral signatures show apparently significant spectral separation. Slope domain indicates of some combinations that has no separation in comparison to unique combinations. 26th combination indicates of a well separation between grass and all other land covers, 14th combination indicates a well separation of red roofs and 1st combination of white roofs.

SOD graphs (Figure 5) show an increasing expression of variability between bands that increase the separation between land covers. It can be seen, for example, that first (SOD) band indicates of a good separation between concrete to dark soil and asphalt, but a bad separation between soil and asphalt. However, other bands such as 6,12 and 17 indicate of a good separation between asphalt and dark soil. It is also can be seen that the spectral distance between trees and bushes in the SOD graphs increased slightly. This is in comparison to the reflectance signature of those land covers shown in figure 3.

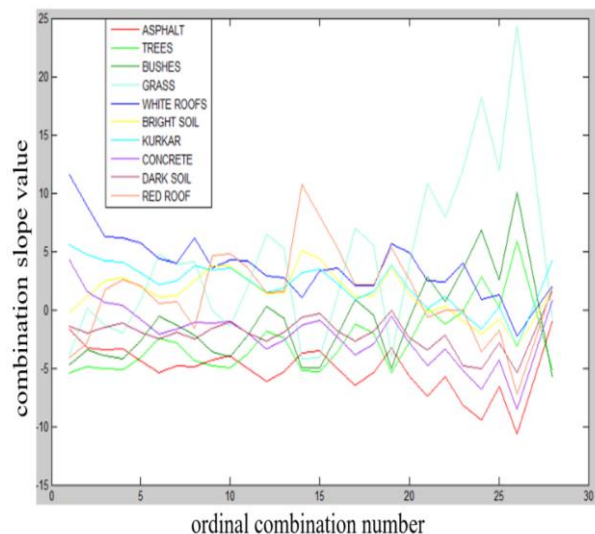


Figure 4. mean spectral signature in slope domain

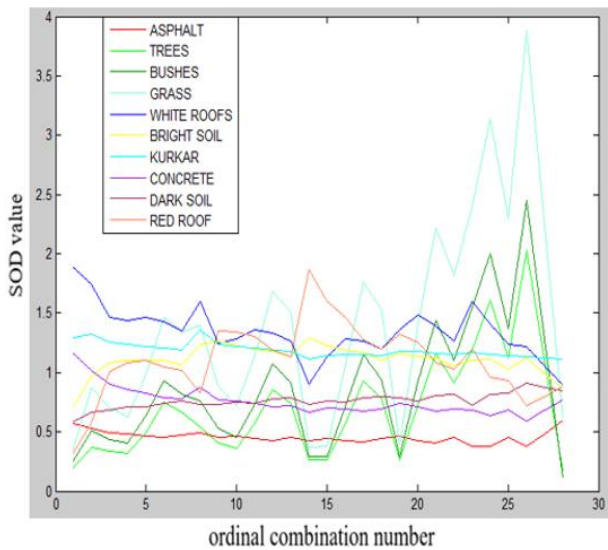
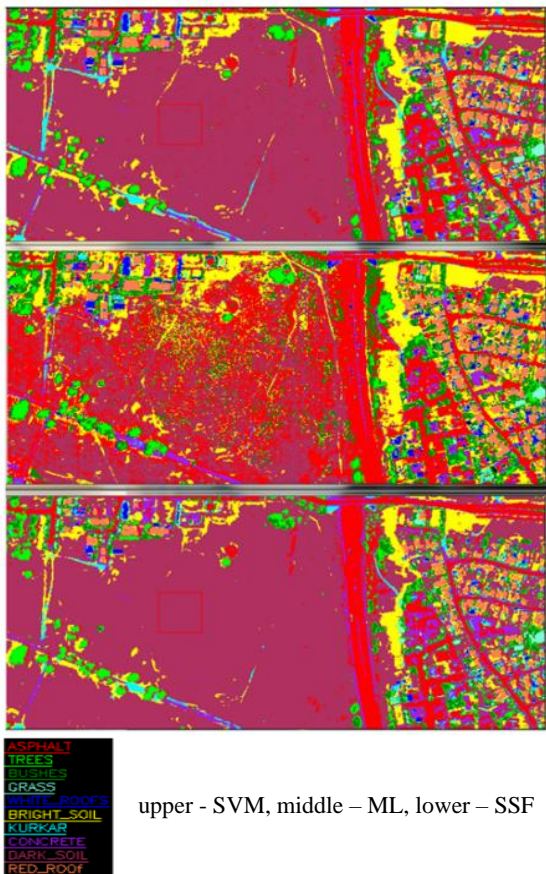


Figure 5. Sum of differences values for each band in slope domain for all land cover's types.



upper - SVM, middle – ML, lower – SSF

Figure 6. classification results

It can be seen in figure 6, that classification of land covers is affected directly by their reflectance variance. Transforming the reflectance information into band's ratios (slope combinations) should reduce the high variance and enabled a better separation between land cover's types. Our hypothesis was that using slope domain would lead to a better accuracy and reliability in land cover classification, in comparison to the usages of reflectance values implemented by each classifier (SVM, ML and SSF). We also examined the accuracy and reliability of SSF classifier in comparison to the other. Figure 6 depicts material classified maps

achieved by 3 classifiers (ML, SVM and SSF) implemented on the slopes domain. It can be seen in figure 6, that dark soil was confused with asphalt in ML classification but not in SVM or SSF one. Bright soil and red brick were also confused when implemented on ML classifier. However, it seems that trees and bushes were not confused using all 3 classifiers.

Table 1 presents a comparison between the 3 classifiers implemented both on reflectance and slopes domains.

	ML SPEC	ML SLOPE	SVM SPEC	SVM SLOPE	SSF SPEC	SSF SLOPE
ASPHALT	98.08	100.00	100.00	100.00	35.77	99.62
TREES	62.50	90.63	95.31	98.96	88.02	89.58
BUSHES	84.65	48.25	67.54	62.72	76.75	73.68
GRASS	66.67	68.63	81.37	83.33	66.67	65.69
WHITE ROOFS	23.73	90.96	87.57	91.53	84.75	83.62
BRIGHT_SOIL	79.76	92.15	80.36	87.61	67.67	97.89
KURKAR	92.86	20.71	92.86	98.57	89.29	91.43
CONCRETE	53.33	82.67	64.00	60.00	90.67	69.33
DARK_SOIL	36.30	45.93	86.91	71.36	90.37	87.65
RED ROOF	93.72	100.00	98.74	100.00	96.23	93.31

Table 1. mean accuracy of classification process of 3 classifiers based on spectral and slopes data

The classification results based on SSF and SVM appeared to be similar, while the classification results based on ML was much poorer. It is seen that ML classifier scored in 8 out of 10 land cover's types better results when was implemented on slopes domain in comparison to the reflectance domain. However, kurkar and bushes classification based on ML slope domain decreased significantly in comparison to the reflectance domain (from 92% to 20% and from 84% to 48% correspondingly). Dark soil achieved extremely poor results in ML classifier in both reflectance and slope domains (36% and 45% correspondingly). It can be seen in figure 5, that in SOD domain the dark soil signature is similar to concrete, despite that only in first band the separation between them is higher.

In Total, it is seen that using SVM classifier comparing to ML classifier achieves higher accuracies. Moreover, SVM classifier scored similarly in both domains with slight advantage to the slope domain over the reflectance one. Note that both concrete and dark soil achieved relatively low results exactly as bushes in both reflectance and slope domain using SVM classifier. This is due to the SVM classifier characteristics, which searches for a well separating boundary between each pair of land cover's type. Both reflectance and slope domains scored similarly when were classified by SSF classifier. As expected, concrete scored better when was implemented on reflectance domain. it turned out that for concrete, a unique ratio between any band's combination in slope domain could not be founded to increase the spectral distance. On the other side, Asphalt scored significantly better results (increased from 35% to 99%) when was implemented on slopes domain, due to spectral distance increasing after the signature was transformed into slope domain.

When comparing SVM to SSF, it is seen that trees, grass, white roofs and kurkar using SVM classifier is better than using SSF. For bushes, bright soil, concrete and dark soil the trend is oppositely. Red roofs and white roofs scored similarly when implemented on both classifiers.

Understanding the characteristics of each domain can be used as associated rules in classification processes. From table 1, it can be studied that for achieving high accuracy results in classifying concrete it is better to use SSF classifier in reflectance domain. However, for distinguishing between bushes and trees it is better

to use ML classifier in reflectance domain. For classifying white roofs use SVM classifier in slope domain and for bright soil use SSF classifier in slope domain.

6. SUMMARY

Using relative reflectance data such as the slope feature emphasizes spectral characterisations and reduces acquisition conditions effects. Classifying pixels using statistical classifier (such as Maximum Likelihood) based on the relative reflectance (slopes) improved the total accuracy and reliability rates. In this research we developed a new classifier named Significant Spectral Features, which achieved similar classification results in comparison to SVM classifier. While SVM is considered as a not explicit classifier, SSF is an explicit classifier and is easy to use for feature extraction processes.

Remote sensing mapping has developed significantly in recent years. Increasing the number of spectral images and spectral band enabled producing more spatial databases and spatial comprehension in large scales. Better understanding of classification processes by improving spectral feature extraction as was shown in this article would lead to much accurate products in many fields and enable a better utilization of spatial sources.

REFERENCES

- Almog O., Shoshany M., Alchantis V., Qizel F., 2008. Improving Hyperspectral Classification Based on Wavelet Decomposition, The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences Vol. XXXVII Part B7, Beijing.
- Cimtay Y, Ilk H., 2017. A Novel Derivative-Based Classification Method for Hyperspectral Data Processing. Digital Image Proceedings and Computer Graphics Volume: 15 | Number: 4 | 2017 | Special Issues.
- Auria L., Moro R.A., 2008. Support Vector Machines as a Technique for Solvency Analysis, DIW German Institute for Economic Research, ISSN print Edition pp: 1433-0210, ISSN electronic edition 1619-4535.
- Fu Z., Caelli T., Liu N., Robles-Kelly A., 2006. Boosted Band Ratio Feature Selection for Hyper spectral Image Classification, NICTA, RSISE, Australian National University, Canberra, ACT 0200, Australia.
- Law K.H., Nichol J., 2004, Topographic Correction for Differential Illumination Effects on IKONOS Satellite Imagery, ISPRS XXXV Congress.
- Rasti B., Hong D., Hang R., Ghamisi P., Kang X., Chanussot J., Benediktsson J., 2020. Feature Extraction for Hyperspectral Imagery: The Evolution from Shallow to Deep. IEEE GRSM DRAFT 2020.
- Rud R., Shoshany M., Alchantis V., Cohen Y., 2006. Application of Spectral Features' Ratios for Improving Classification in Partially Calibrated Hyperspectral Imagery: A Case Study of Separating Mediterranean Vegetation Species J Real-Time Image Proc, DOI 10.1007/s11554-006-0015-8.
- Shoshany M., Almog O, Alchanatis V., 2009. Wavelet Decomposition for Reducing Flux Density Effects on Hyperspectral Classification. IEEE Geoscience and Remote Sensing Letters. 2009 vol. 6. No. 1. pp: 38 – 41.
- Sun Y., Liu X., Wu Y., Liao C., 2008. Identifying Hyperspectral Characters, of Wetlands Species Using In-Situ Data, The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences Vol. XXXVII Part B7, Beijing.
- Torrecilla E., Pira J., 2009. Derivative Analysis of Hyperspectral Oceanographic Data, Advances in Geo-Science and Remote Sensing, pp: 597-618.