Determination of Optimal Hyper- and Multispectral Image Channels by Spectral Fractal Structure

Veronika Kozma-Bognár and József Berke

Abstract

Multiband aerial mapping technology—for the high spectral and high spatial resolution images—also the phenomenal traditional aerial mapping techniques, which are more reliable, compared to data obtained during the foundation stages of the process. Over the last decade, aircraft data recording technology has developed considerably due to its applications in the field of research and has become an increasingly central theme of multiband and high spatial resolution integrated processing. This has a significant impact on assessment results. Using practical examples, the author's show that a properly selected spectral fractal structure based on data reduction and data selection procedures, significantly contributes to the hyper- and multispectral data cube optimum exploitation of additional information.

Introduction

The twenty-first century characterized the great extent of the affects of the rapidly growing developments in information technology, in the field of remote sensing science. In recent years, the development of remote sensing technologies has played a vitally more important role in a range of multispectral and high geometric resolution images. The multiband and high resolution aerial and satellite images of active applications and laser measurement technologies open up new perspectives on the object surface and the amount of accurate data collected. Presently, following a long delay, the technical development of sensor processing methods and applications, is now justified to clarify existing procedures, and new analytical methods. Despite the fact that the world has been introduced to several imaging and multiband integrated programs [14], in order to support the processing of images, the range of these remains quite limited. Much GIS software has been developed, for example the adaptation procedures for the main methods of hyperspectral image processing for existing multispectral shots. Hyper- and multispectral technology resulting from disparities (e.g. difference of the channel bandwidth, the geometric resolution), raises questions as to the accuracy of these methods. The procedures selected depend on its order, the software use, experience of the person doing the processing and the impact on the multiband images of the final outcomes.

Data processing of the data cube, which make up hundreds of spectral channels for part of preprocessing selection, which contain the information necessary for the implementation of the tasks. The channels of data cube can be charged with noise and/or correlated with each other [5, 17]. Band selection methods are used to filter redundant

V. Kozma-Bognár (✉)
Georgikon Faculty, Department of Meteorology and Water Management, University of Pannonia, Keszthely, Hungary
e-mail: kbv@georgikon.hu

J. Berke
Institute of Basic and Technical Sciences, Dennis Gabor College, Budapest, Hungary
e-mail: berke@gdf.hu

data [15] such as the common essence: principal component analysis [12], minimum noise fraction [7, 12], lambda-lambda R^2 model [13], stepwise discriminant analysis [16], derivative greenness vegetation indices [6]. The preprocessing, within dimension reduction methods, greatly affects the outcome of the evaluation, because inadequate preprocessing may be excluded from a band that has useful informations for our investigation.

During research and goal-oriented studies, we aimed to develop a method that allows for multiband aircraft images, which selects maximum information to be considered in the production of optimal shots. Another consideration is that when different methods are used together sensors produced high spectral resolution images and high spatial resolution data, which is workable and useable.

Materials and Methods

Research Sites

The research areas of the two samples were analyzed from recordings. Hyperspectral research took place in the rural Bakony landscape, which is located on the outskirts of Várvölgy (Hungary). The test areas [9] (the highest point coordinates: N: 46°51′20.76″, E: 17°18′06.53″, H: 247 m), were chosen which are mostly used for agricultural land use. The test areas also include easily identified target and other terrestrial objects (forest belt, driveway, dirt road), which later served as points of reference during this study (Fig. 1). The field data collection determined the number of categories and a total of eight categories are reported in the survey. The entire test area, situated in natural and artificial formations, was assessed, and data collection was carried out on the state of the plant population. Immediately following data collection, the research areas are finally reviewed and records are compiled to define the types of objects involved. The following categories have been set: corn, triticale, wheat, sunflower, uncultivated land, forests, driveway (asphalt), a dirt road.

Multispectral research was carried out 2 km from Keszthely (Hungary) at the Agrometeorological Research Station test area, where the corn plot experiments involving traffic originating pollution took place [10]. Located north of the station (the center of the test area N: 46°44′08.55″, E: 17°14′19.76″, H:114 m) a test area comprising of six plots (10 m × 10 m/plot) was established.

The motor vehicle emissions and negative environmental effects of chemicals that cause physiological tests to two characteristics of the contaminant were selected: carbon black, and cadmium. Accordingly, multispectral collection was evaluated for the following determined object types:
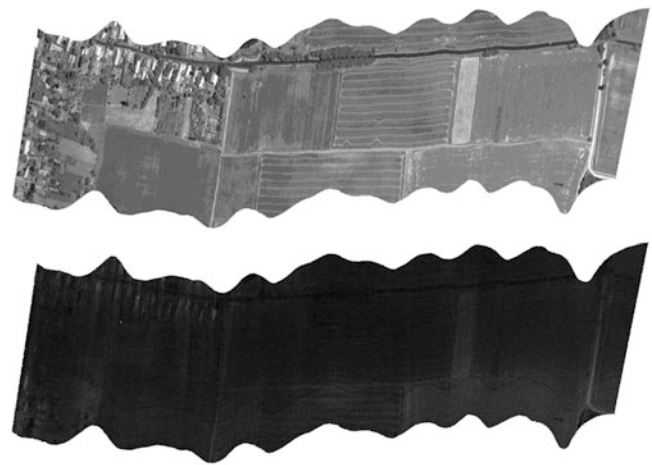

Fig. 1 Noise-free (254th band, *upper*) and noise (273rd band, *bottom*) AISA Dual Image Layer from Várvölgy test area, 06/19/2007
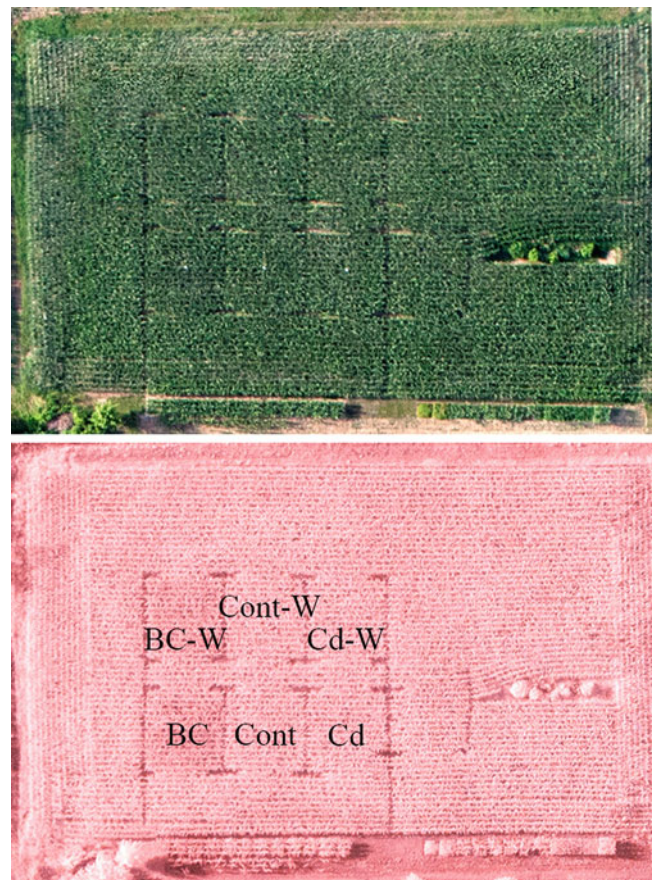

Fig. 2 Visible range (*upper*) and near infrared (*bottom*) image from test area of Keszthely, 06/21/2012

BC = Black Carbon polluted, BC-W = Black Carbon polluted and irrigated, Cont = Control, Cont-W = Control irrigated, Cd = Cadmium polluted, Cd-W = Cadmium polluted and irrigated (Fig. 2).

Table 1 Main parameters of data collection with the help of air-shot

	Visible data (VIS)	Near infrared data (NIR)	Hyperspectral data
Type of sensor	Canon 30D	Canon 30DIR	AISA Dual
Height of flight	400 m	400 m	1,200 m
Spectral band	400–700 nm	720–1,150 nm	400–2,450 nm (498 bands)
Geometrical resolution	0.1 m	0.1 m	1.0 m
Data recording	14 bit/pixel	14 bit/pixel	14 bit/pixel

The realization of an accurate study of the possible pre-processing for eliminating problems resulted in images of the total area of each plot being used as a criterion for selection, which were then analyzed.

Applied Tools

During the study Hungarian hyperspectral AISA Dual hyperspectral sensor aerial images were analysed. The AISA Dual hyperspectral sensor system is one of the most important in the European Union developed by the Finnish SPECIM Spectral Imaging Ltd. The family members of AISA camera combine the AISA Eagle and Hawk sensors. The container has been secured by the coordination of the VNIR spectral range in order to collect data from up to 498 spectral bands. The AISA Dual System is a compact, low-cost, aircraft-mounted, high-performance and highly operational characteristic of the imaging system. The AISA stable operation system provides us with high spectral and relatively high (max. 1.0 m) spatial resolution aerial images, which provide us with enough information for our Várvölgy field test study. The sample area covered was smaller in Keszthely due to the appropriate required level of assessment for recording the results of higher spatial resolution. The multispectral airborne imaging test area considerably surpassed the resolution of 1 m below the field (7–30 cm^2/pixel). The high resolution multispectral images for field production data provided by the sensors used in digital lines in the visible (VIS: 0.4–0.7 mm) range and the near infrared (NIR: 0.7–1.15 mm). Major collection tools were used for mapping the parameters in Table 1.

Spectral Fractal Structure

The starting point of the hyper- and multispectral images, using the characteristics of the data reference, began when the different plant populations spectral fractal dimension measurements were taken. The measurement program developed by the SFD Information Technology Ltd [1]. Was considered optimal for spectral bands used to select and validate the fractal dimension (SFD) values. Reviews of the images containing the tested crop areas were analyzed, and those values only measured SFD.

The SFD is a general fractional dimension [11] derived from the structure of the processing, which is a novel application of fractals. The spatial dimensional structure of the SFD outside of the spectral bands is also suitable for measuring the color structure [1, 4] and provides sufficient information to colors, shades and other fractal properties as well. Calculating the values of the SFD (two or more band images), the defined fractal dimension is applied to the measured data as a function (the total number of boxes valued spectral function of the spectral box) simple mathematical averages are calculated as follows [2]:

$$SFD_{measured} = \frac{n \times \sum_{j=1}^{S-1} \frac{\log(BM_j)}{\log(BT_j)}}{S-1} \qquad (1)$$

where,
n—the number of image layers or image channels
S—the spectral resolution (in bits)
BM_j—valued spectral pixel boxes containing j number of bits
BT_j—all possible spectral boxes for the number of j-bits

The possible number of boxes spectral j bits is calculated as follows:

$$BT_j = \left(2^S\right)^n \qquad (2)$$

Since the formula (1) metric [3], the evaluations of both hyper- and multispectral images can be used for exact measurements.

Determination of the Optimal Conditions for SFD-Based Studies

Hyperspectral imaging system's service data could be highly noisy and redundant. For this reason data analysis, which considers noisy and redundancy as bands, we search the optimal bands, which adequately represent all of the image information. The Várvölgy test area, which has 359 available channels of data cubes in some bands, was observed as
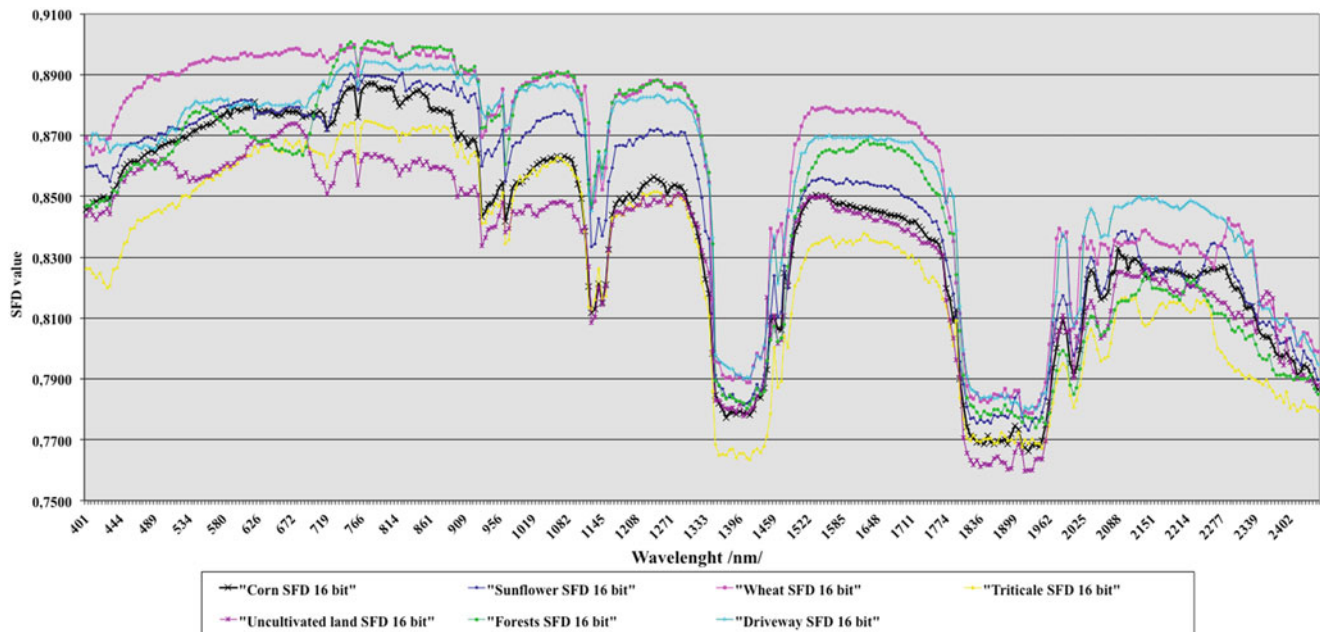
Fig. 3 16 bits SFD fingerprints of corn, sunflower, wheat, triticale, uncultivated areas, forests and driveway

having a strong correlation with literature, which offers different data reduction procedures, such as principal component analysis, and the correlation between the individual channels, or energy minimization. Many of these procedures do not contain the actual image data reduction, and can only be described as pieces of information about the probability of actual data on a number of independent bands (PCA-based methods). However, there is not always enough information available from the bands, or the information provided highly depends on the input data (e.g. depend on the rate of noisy bands and of their weight). In order to overcome these problems, a data processing method, which is not or is less sensitive to the input data procedure, should be used. Using this procedure, we have achieved results in SFD based structural analysis, using the correct application of noise results [10]. This task is also proposed to serve as an additional solution.

The first step is testing the hyperspectral image created by the SFD based fingerprint crops, which represents the channel or SFD values versus wavelength (Fig. 3), [8].

The SFD plant cultures per band and their measured values are sorted in descending order according to the channel numbers. Five plants of these five columns are sorted in the order in which adjacent value differences were formed, and these difference values were added and titled: "organized patterns of differences in the amount of SFD". Then the resulting curve approached the local maximum value of channel numbers, which are the channels for optimal recording of a particular crop (Fig. 4).

Figure 4 demonstrates the top 21 local places with a maximum of five plants of maximum position found for

the five following crops (corn, wheat, triticale, sunflower, uncultivated areas), which are the following: 2, 26, 49, 55, 62, 78, 83, 90, 98, 114, 122, 138, 166, 201, 215, 272, 285, 293, 301, 327, 350. The hyperspectral images of adjacent channels are usually from the same correlated sources of information, therefore, there was a condition placed on the obtained local maximum that they are considered not to fall within a range of at least two bands. We believe this series uses the most appropriate channels for the further phases of evaluation (e.g. classification). As a further alternative, aspects of the 359 channels are selected six channels, in which some are named as the minimum optimal channels. The SFD fingerprints based on the following six channels were considered to be optimal for the outermost bands between local maximum ratings: 65, 71, 128, 167, 243, and 288. We calculated the standard deviation amounts with the aim of more adjacent channels of the same values and, the standard deviation of the lowest channel for selection.

Based on the above channels, for the purposes of ascertaining the optimal information content of the resulting bands, the area used as input for the spectral angle mapper method of carrying out classification was considered. In addition, other method to reduce the number of data channels was measured and determined by comparative analyzes and were carried out based on the accuracy of results. Giving the optimal channel selection procedures for comparing the different methods of data as input for completion of the surface identification categories. Subsequently these were evaluated to reduce the number of dimensions which best describes the result of the processing of our test crop.
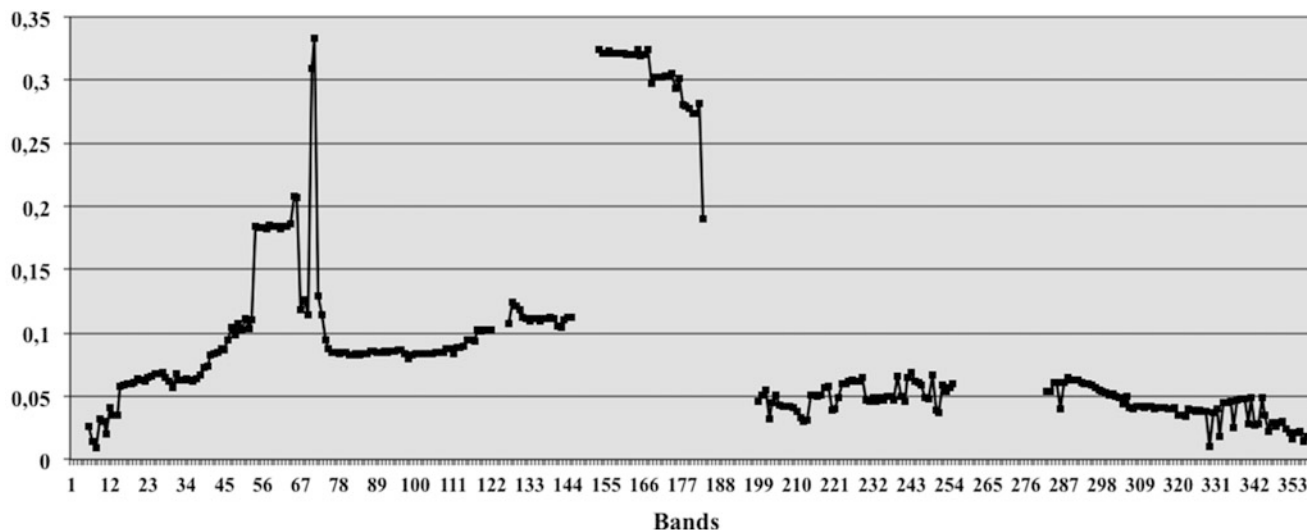
Fig. 4 Arranged in increasing order of the sum of the difference between the SFD
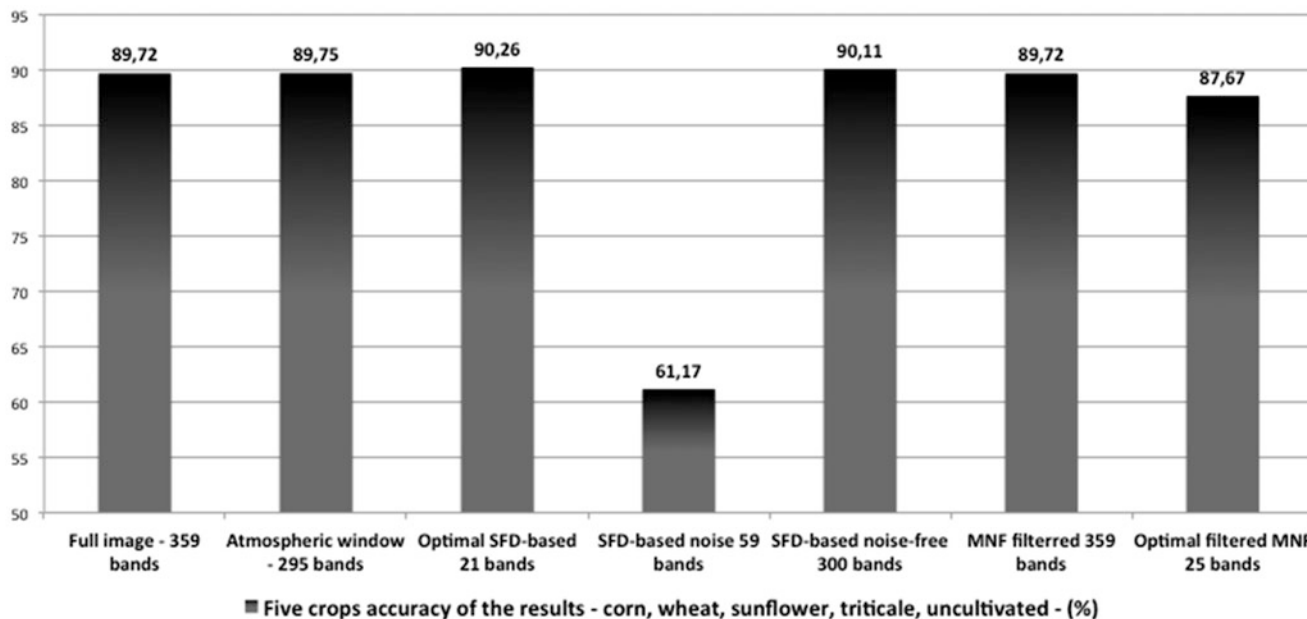

Fig. 5 Procedures to reduce the total number of dimensions—the result of plant cultures

The classified images were taken and the results of the error matrix were calculated with the accuracy of the results for each of the crops (Fig. 5), and the average of the five crops, i.e., seven methods were prepared using input data series (full image—359 bands; atmospheric window-based bands—295 bands; optimal SFD-based minimal bands—6 bands; SFD-based noise-free bands—300 bands; SFD-based noise bands—59 bands; optimal filtered MNF bands—25 bands, filtered MNF bands full image—359 bands). The obtained results clearly show that the SFD value-based methods for optimal bands selected crops combined results

of the highest accuracy. We note that in practice, the most accepted and used noise reduction method (MNF = Minimum Noise Fraction) to carry out the image after the input of the hit accuracy was the same as the accuracy of the unfiltered image results. This can be explained by the effect of the MNF noise reduction AISA Dual image sensor, we recognized some visible band, and however, the accuracy of the results was invariant.

We have developed the channel selection procedure based on the SFD values of the multispectral images used to optimize the selection of images. In investigation, the

Table 2 Maximum SFD values of the multispectral images

	Visible SFD values 3 bands	Near Infrared SFD values 3 bands
Track 1	2.1361	1.9418
Track 2	2.1416	1.9187
Track 3	2.1398	1.9400
Track 4	2.1328	1.9215
Track 5	2.1273	1.8406
Average	2.1355	1.9125

SFD values were deduced from the size of the actual information content of images. Airborne imaging of one-one workflow (e.g. flight track) recording usually occurs in the Keszthely sample area where they were looking to provide the best and most reliable information for the object-finding study.

The multispectral images of the preprocessing, post processing only recorded SFD maximum purchasing values per sensor and per flight (Table 2), as further investigations of these images gave the most reliable results (hit accuracy as well as other measured parameters and correlation studies).

Comparision of Hyperspectral Data Selection Procedures

In the scope of our research different classification methods for analyzing hyperspectral images were selected. We have chosen the optimal minimum and maximum bands based on spectral fractal structures for optimal selection to discuss the effectiveness and the accuracy of the results of each classification method.

We investigated five supervised classification methods, the basic aim of the selection was to investigate methods that are primarily examining the development of multispectral images, and however the development of a unique hyperspectral data classification method is also present in the comparison. An additional consideration is that the most commonly used hyperspectral data processing software methods are able to be studied, since the users, especially professionals, primarily engage with this form. The above concerns the following five metrics based on the classification methods selected: Parallelepiped (PA), Minimum Distance (MD), Mahalanobis (MA), Maximum Likelihood (ML), Spectral Angle Mapper (SAM). One independent samples (region of interest = ROI) was selected on each of the nine tested areas by the input parameters of the classification. The five classes consist of nine samples, whereas in four cases (areas 1 and 3 corn, areas 2 and 5 triticale, areas 4 and 6 wheat, area 7 uncultivated area, areas 8 and 10 sunflower) two plots grow the same crop. We have also added a
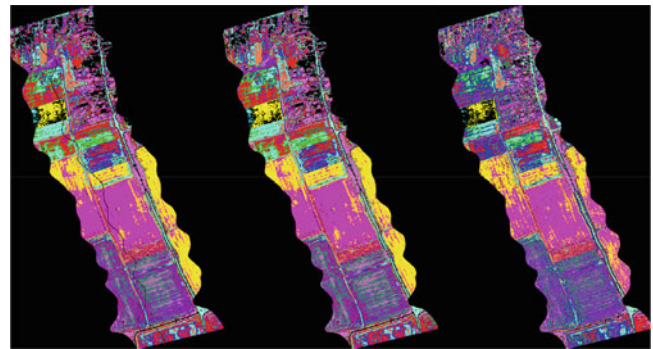

Fig. 6 Spectral Angle Mapper image metrics based on graded results (based on full image, noise-free channels of SFD and six optimal channels)

further three samples (forest, asphalt driveway, dirt road), which were not the subject of our study, but were found between the fields in the former nine samples.

The aforementioned 6 minimum and 21 maximum optimal channels were definded by SFD spectrum curves. In addition to the studies have been investigating the full 359 channels of recording and the selected bands without noise on the basis of the SFD.

The accuracy of the results was measured on the classified images (Fig. 6). Within each ROI, the numbers of pixels correctly classified were determined by the ROI intensity values. The accuracy of the results in percentage was calculated from the number of correctly classified pixels divided by the total ROI, and then multiplied by one hundred. The average price for the five crops selected images classification processing are based on the accuracy of results summarized in Table 3.

The result analysis clearly shows that the classification of the proposed channels by the SFD fingerprints (without noise, the minimum and maximum optimal channels), based on the choice of any metrics for corn, wheat, uncultivated land and sunflower crops, allowed greater accuracy results than the classification of the 359 channels. In the case of triticale classification of four metrics (MA, ML, MD and SAM) obtained similar results. In conclusion, the amount of PA gave the least accuracy (69.91 %) among the five classification procedures, which is explained by the fact that the PA does not develop specifically classified hyperspectral images. The classifications of the ML clearly performed best (97.61 %), which is consistent with the results of most of the multispectral image processing methods obtained from literature [5, 6]. However we did not expect such accuracy as a consequence of considerable noise of hyperspectral images. Interestingly, the highest accurate individual aggregate results were obtained from the Mahalanobis classification for noise-free image (98.61 %). Another important finding, following the removal of the noisy bands by the SFD spectrum curves, the accuracy of

Table 3 Average accuracy of results in the selected classification methods

Average accuracy of results					
Type of supervised classification	Full image (359 bands)	SFD-based noise-free (300 bands)	Optimal SFD-based minimal (6 bands)	Optimal SFD-based (21 bands)	SUM
Mahalanobis	96.21	98.61	86.09	92.74	93.41
Maximum likelihood	98.16	98.49	96.08	97.71	97.61
SAM	75.91	76.66	75.23	74.40	75.55
Parallelepiped	65.35	73.45	68.00	72.84	69.91
Minimum distance	80.61	80.56	82.68	79.46	80.83
Average	83.40	85.56	81.61	83.43	

the 300 channels classification was measured 2 % higher (2.16 %) than the classification of the 359 channels.

The result is superior from the aspect that we obtained, in this case, a significant increase in the high accuracy of the results. The 21 bands classification demonstrated almost the same (slightly better) as the 359 bands, which in our case is important because 21 bands based classification has been running for a significantly shorter time than the (computing capacity) 359 bands based classification. The 6 bands classification accuracy of the result was just below (1.79 %) in the overall accuracy of 359 bands. Knowing that the running time has an almost two fold magnitude difference, this is an excellent result for images classifications with high-performance computing workloads.

Conclusion

In the hyperspectral image data cube can accurately identify surface objects with the use of the SFD curves. The bands of the local maximum of the SFD fingerprints accurately representing the total information content of the image data. In some cases these bands represent more reliable data selection than other mathematical methods based on the eigenvalue calculation. The image classification of noise-free channels which are defined on our SFD spectrum curves, based on the choice of five metrics (PA, MA, ML, SAM and MD) for corn, wheat, uncultivated land and sunflower crops allowed greater accuracy of results than the classification of the entire 359 bands. The classification on the basis of the noise-free, optimal channels which are defined on our SFD spectrum curves, in almost all investigated metrics showed a higher accuracy of the results than the entire image (359 channels) using significantly less computing capacity. In the preprocessing of multispectral images, from the maximum spectral fractal structure selected images provide the most reliable accuracy for the measured results and for the other correlation studies.

References

1. Authors Internet site of parameter Spectral Fractal Dimension, http://www.digkep.hu/sfd/index.htm.
2. J. Berke, "Measuring of Spectral Fractal Dimension," Journal of New Mathematics and Natural Computation, vol. 3/3, pp. 409–418. DOI: 10.1142/S1793005707000872.
3. J. Berke, "Measuring of Spectral Fractal Dimension," Advances in Systems Computing Sciences and Software Engineering, Springer-Verlag pp. 397-402., 2006. ISBN 10 1-4020-5262-6.
4. J. Berke, "Using Spectral Fractal Dimension in Image Classification," Innovations and Advances in Computer Sciences and Engineering, Springer Science+Business Media B.V. 2010, DOI: 10.1007/978-90-481-3658-2_41, [CISSE'08 Conf., 2008].
5. C.-I. Chang, "Hyperspectral data exploitation theory and applications", John Wiley and Sons, New Jersey, 2007. ISBN: 978-0-471-74697-3
6. C.D. Elvidge, Z. Chen, "Comparison of broadband and narrowband red and near-infrared vegetation indices," Remote Sensing of Environment, 54, pp. 38– 48, 1995.
7. A.A. Green, M. Berman, P. Switzer, M.D. Craig, "A transformation for ordering multispectral data in terms of image quality with implications for noise removal," IEEE Transactions on Geoscience and Remote Sensing, Vol. 26, No. 1, pp. 65-74, 1988.
8. V. Kozma-Bognár, J. Berke, "New Evaluation Techniques of Hyperspectral Data. Journal of Systemics," Cybernetics and Informatics. Vol. 8. Number 5. pp. 49-53, ISSN: 1690-4524. http://www.iiisci.org/journal/SCI/, 2010.
9. V. Kozma-Bognar J. Berke, G. Martin, "Application possibilities of aerial and terrain data evaluation in particulate pollution effects," EGU General Assembly, Wien, 2012.
10. V. Kozma-Bognár, "Investigation of Hyperspectral Image Processing and Application in Agriculture," Ph.D. dissertation, University of Pannonia, 2012.
11. B.B. Mandelbrot, "The fractal geometry of nature". W.H. Freeman and Company, New York, 1983
12. J.A. Richards, "Remote Sensing Digital Image Analysis", Springer-Verlag, Berlin, p. 240, 1999.
13. SAS Institute Inc., Measuring the Internal Consistency of a Test, Using PROC CORR to Compute Cronbach's Coefficient Alpha, SAS Communications, 20:4, TT2–TT5, 1995.
14. R.A. Schowengerdt, "Remote Sensing Models and Methods for Image Processing", Elsevier, 2007. ISBN 13: 978-0-12-369407-2.
15. P.S. Thenkabail, E.A. Enclona, M.S. Ashton, B. Van Der Meer, "Accuracy assessments of hyperspectral waveband performance for

vegetation analisys applications," Remote Sensing of Environment 91 pp 354-376, 2004.

16. P.S. Thenkabail, R.B. Smith, E. De-Pauw, "Evaluation of narrow-band and broadband vegetation indices for determining optimal hyperspectral wavebands for agricultural crop characterization,"

Photogrammetric Engineering and Remote Sensing, 68(6), pp. 607–621, 2002.

17. P.K. Varshney, M.K. Arora, "Advanced Image Processing Techniques for Remotely Sensed Hyperspectral Data", Springer-Verlag, New York., 2004.