

# Simplified Classification of Multispectral Image Fragments

A. Lorencs<sup>1</sup>, I. Mednieks<sup>1</sup>, J. Sinica-Sinavskis<sup>1</sup>

<sup>1</sup>*Institute of Electronics and Computer Science,  
Dzerbenes iela 14, LV-1006 Riga, Latvia  
mednieks@edi.lv*

**Abstract**—A simplified approach to classification of multispectral image fragments by their specific spectral features is presented. Application of this approach to discrimination of vegetation areas occupied by the Giant Hogweed species is described and compared with an approach based on calculation of the Consolidated Covariance Image. The proposed method is based on calculation of mean and standard deviation and successive thresholding within certain spectral bands that are found to be informative for the specific task by analysing the ground truth data. It is shown that the method provides close to perfect discrimination of Giant Hogweed from other vegetation areas represented in ground truth data (absence of commission errors together with clear identification of Giant Hogweed fragments in corresponding ground truth regions). Simplicity of the method provides for fast processing of multispectral images from large areas. The proposed approach is perspective for analysis of multispectral images in different application fields where it is possible to choose several informative spectral bands, e.g. in biomedical imaging.

**Index Terms**—Multispectral imaging, image classification, spectral features.

## I. INTRODUCTION

Multispectral imaging (MSI) becomes increasingly popular for analysis of objects in remote sensing [1]–[5], biomedicine [6]–[7] and some other research fields. It is characterized by acquisition of images usually in less than 20, generally non-contiguous, spectral bands in visible and near infrared wavelength ranges. One of the common processing tasks in MSI is to classify image pixels or regions into several classes of interest. While the number of bands in MSI is significantly lower than in hyperspectral imaging where it exceeds 100, the amount of data to be processed, especially in remote sensing applications, can be huge, and it is of major importance to simplify processing and exploit only data from limited number of spectral bands that may serve as features for classification. If the spectra of pixels of different classes overlap significantly, it is necessary to apply rather sophisticated techniques, e.g. Bayesian classification serving well for a lot of tasks related with observation of natural phenomena [8]. However, in certain cases it is possible to find characteristic spectral features and

distinguish between the classes using simpler rules. It is true especially in cases where classification is performed within a low number of classes, possibly after some simple preprocessing that has masked out parts of image that are spectrally really different and clearly out of interest.

The paper discusses one such simplified classification approach that have proved to work efficiently for particular application task of finding areas covered by invasive plants in remote sensing data. The approach however is generic and may serve for different tasks as well, e.g. for classification of skin lesions in biomedical MSI.

The problem of mapping invasive plants is important in many countries and its solution based on multispectral remote sensing was considered since satellite images became available [9]. Mapping and elimination of Giant Hogweed (*Heracleum sosnowskyi* Manden) causing stress to the natural ecosystem and being a human health hazard is of considerable importance in Latvia as well as several other Eastern European countries. The paper presents a method for processing of multispectral images and its application example related with mapping of separately growing Giant Hogweed areas in a 15-band multispectral image acquired using the airborne hyperspectral sensor.

## II. DATA ACQUISITION

Multispectral data used within this study were acquired by the Institute for Environmental Solutions ([www.videsinstituts.lv](http://www.videsinstituts.lv)) in one flight over the area in Latvia containing known sample fields containing hogweed and other characteristic types of vegetation (forests, meadows, cornfields etc.). Total analysed area is about 150 ha. To acquire the multispectral data, a hyperspectral imager CASI–1500 from ITRES Research (<http://www.itres.com>) was used, mounted in the Observer P-68 aircraft. Imager was configured to use a full field of view (40°) and to combine data into 15 spectral bands: pixel values for band 1 are related with average light intensity for wavelengths 427.3 nm ± 28.6 nm, band 2 for 479.8 nm ± 23.9 nm, band 3 for 518.0 nm ± 14.3 nm, band 4 for 551.4 nm ± 19.1 nm, band 5 for 597.9 nm ± 27.5 nm, band 6 for 633.7 nm ± 8.4 nm, band 7 for 671.9 nm ± 29.8 nm, band 8 for 710.1 nm ± 8.4 nm, band 9 for 728.0 nm ± 9.5 nm, band 10 for 742.3 nm ± 4.8 nm, band 11 for 761.4 nm ± 14.3 nm, band 12 for 804.4 nm ± 28.6 nm, band 13 for 866.4 nm ± 33.4 nm, band 14 for 922.4 nm ± 22.6 nm, and band 15 for 991.5 nm ±

44.1 nm. Geo-corrected multispectral image with a spatial resolution  $0.5 \text{ m} \times 0.5 \text{ m}$  was obtained by exploiting GPS data that were also collected during the flight. Radiometric and geometric corrections were made using the ITRES software. No atmospheric corrections or spectral transformations were made.

Ground truth regions for major pixel classes were marked on the image by the person familiar with the study area: Hogweed (marked as Hn, where n is the sequence number), Trees (Tn), Crops (Cn), Grass (Gn), Grass-cut (GCn), Road-cement (Rn), Road-gravel (RGn), and Soil (Sn). Processed image visualized from the spectral bands similarly to human perception is presented in Fig. 1. It is noticed that the image contains also pixels not related with one of major classes so it is necessary to mask them out and limit analysis only to vegetation areas possibly including hogweed pixels.

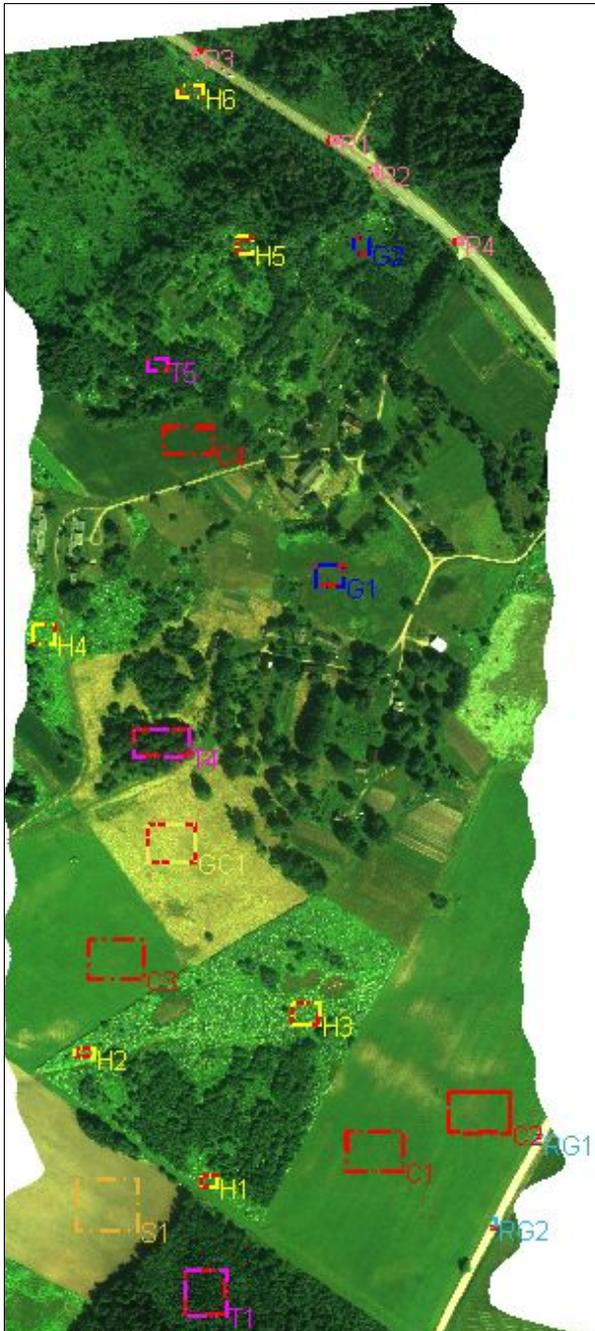


Fig. 1. Central part of the processed image visualized as RGB image from 3 spectral bands (band 7 mapped to red colour, band 4 to green colour, band 1 to blue colour), with marked design regions for pixel classes.

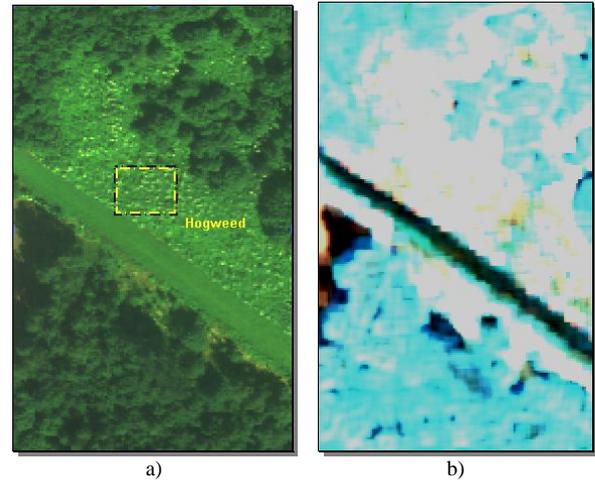


Fig. 2. Enlarged image of the area containing a hogweed region H1 (a) and the corresponding CCIm (b).

### III. METHODS

Zoomed-in area related with hogweed is shown in Fig. 2(a). It is seen that hogweed areas are speckled and characterized by a large variance of pixel values in visualized spectral bands. Therefore initially application of the Consolidated Covariance Image [10] (CCIm) to detection of hogweed areas was considered. It allows the image to be transformed to the form where the increased covariance around a pixel results in brighter colours of that pixel. CCIm of the area was calculated based on processing image fragments of size  $5 \times 5$  pixels; covariance values shown with red colour were obtained from initial pixel values in bands 1..7, values shown with green colour were obtained from bands 8..10, and values shown with blue colour were obtained from bands 11..15. Obtained CCIm is illustrated in Fig. 2(b). It is seen that, indeed, hogweed areas are nearly white. However, forest regions including shades are also characterized by large variance of pixel values for the involved spectral bands; the same applies to places related with transitions from one pixel class to another. Therefore application of CCIm requires some further processing to filter out such artifacts. This approach is also characterized by rather large amount of calculations and a possibility to apply a simpler approach was investigated.

To deal with pixels related with vegetation, different vegetation indices are used [11]. For our task, a simple ratio vegetation index (RVI) was calculated as a ratio of pixel values from bands 12 and 7, and thresholding used to separate pixels of interest. Fig. 3 illustrates the results obtained when the threshold is chosen such that pixels not related with green vegetation are masked out.

Vegetation present in the processed multispectral image falls into 4 categories: hogweed, trees, crops and grass. Each pixel in the image is represented by a vector of 15 average light intensity values from the mentioned spectral bands. Ground truth regions for each of these classes are marked on the image and used as design regions of the classifier.

Each marked region was divided into fragments with size  $7 \times 7$  pixels without overlapping. For a fixed spectral band  $\lambda$ , within a set of pixel values from each fragment, two characteristics were calculated, namely mean value  $\bar{x}_\lambda$  and standard deviation  $\sigma_\lambda$  of light intensity. As these values for

different fragments within a marked region generally are not equal, it is important to consider minimum and maximum values of these characteristics for each region.

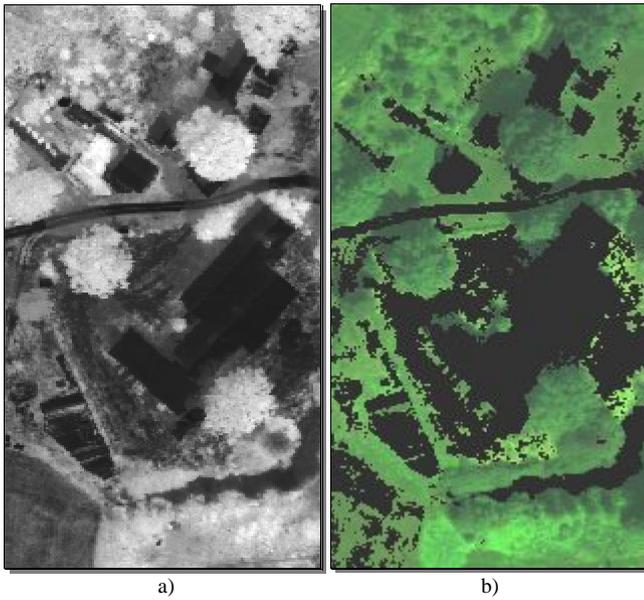


Fig. 3. Masking of non-vegetation pixels using ratio vegetation index (RVI): (a) RVI image fragment; (b) thresholding result.

Minimum mean values and standard deviations of light intensity within fragments, obtained for hogweed design regions (see Fig. 4), feature 3 local extremes in spectral bands 4, 12 and 14. It was noticed that intensity values for hogweed were higher than corresponding values for trees, and for majority of cases higher than corresponding values for grass and crops as well. Analysis of graphs show that particular interest should be focused on spectral bands with wavelengths  $\lambda_1 = 550nm$ ,  $\lambda_2 = 800nm$  and  $\lambda_3 = 925nm$ . In Fig. 4, local minima  $\sim_h(\lambda_j)$  for hogweed design regions are indicated. Thresholds  $\Delta_1$ ,  $\Delta_2$  and  $\Delta_3$  are defined then as follows:

$$\Delta_1 = \sim_h(\lambda_1) - v_1 = 2410 - 10 = 2400, \quad (1)$$

$$\Delta_2 = \sim_h(\lambda_2) - v_2 = 7310 - 10 = 7300, \quad (2)$$

$$\Delta_3 = \sim_h(\lambda_3) - v_3 = 4010 - 10 = 4000, \quad (3)$$

where the value  $v_j = 10$  is chosen to guarantee that values  $\sim_h(\lambda_j)$  for hogweed fragments are higher than corresponding threshold  $\Delta_j$ . Obviously, it should be taken into account that for some fragments of trees, grass or crops, mean value  $\sim_h(\lambda_j)$  may be higher than the corresponding threshold value  $\Delta_j$ . Therefore two additional thresholds are defined:

$$\Delta_4 = \dagger_h(\lambda_1) - v_4 = 350, \quad (4)$$

$$\Delta_5 = \dagger_h(\lambda_2) - v_5 = 570, \quad (5)$$

where  $\dagger_h(\lambda_j)$  are minimum standard deviation values for hogweed fragments for corresponding wavelengths.

The main construction principle of the classifier for identification of the hogweed fragments is formulated as follows: to guarantee that each fragment of the image related with hogweed will be characterized by such values  $\sim_h(\lambda_j)$  and  $\dagger_h(\lambda_j)$  that exceed corresponding thresholds but fragments containing trees, grass or crops will not meet at least one of these 5 conditions. Construction of classifier  $C_1$  is explained below.

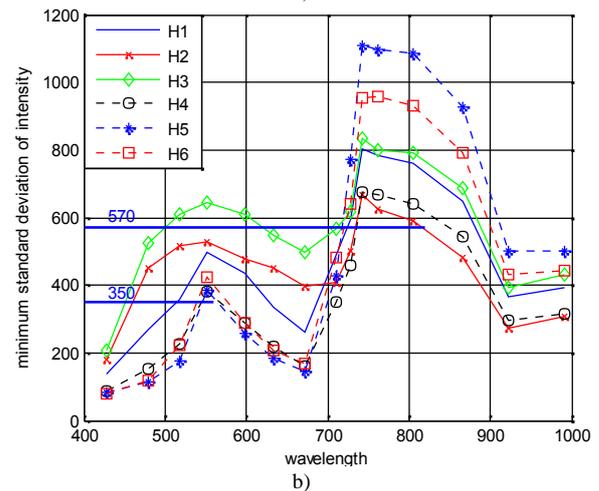
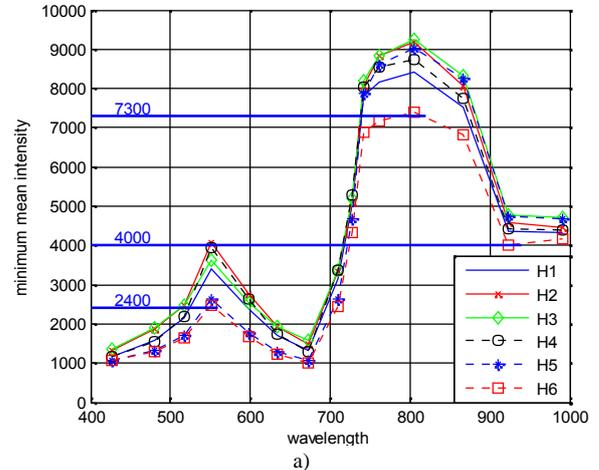


Fig. 4. Minimum mean values (a) and standard deviations (b) for spectral bands in hogweed design regions with thresholds used for classification identified.

### Preprocessing

- $I$  – analysed image with 15 spectral bands
- obtain  $I_3$  from  $I$  by taking data for spectral bands 4, 12 and 14 only
- obtain  $DRF$  by dividing design regions from  $I_3$  into  $7 \times 7$  pixel fragments
- calculate mean and standard deviation of intensity values for non-overlapping fragments of  $DRF$
- calculate thresholds  $\Delta_1, \Delta_2, \Delta_3, \Delta_4, \Delta_5$  from mean and standard deviation values according to (1)–(5)
- calculate threshold  $\Delta_6 = 2.75$  as the mean value of minimum vegetation indices in hogweed  $DRFs$ .

### Classifier $C_1$

- calculate ratio vegetation index image  $RVI = I_{12}/I_7$  (ratio of values from bands 12 and 7 for each pixel)
- calculate binary image  $VIT = (RVI > \Delta_6)$

– calculate  $I3result = I3 .* imdilate(VIT)$ , i.e., multiply image  $I3$  pixel-by-pixel by the morphologically dilated image  $VIT$

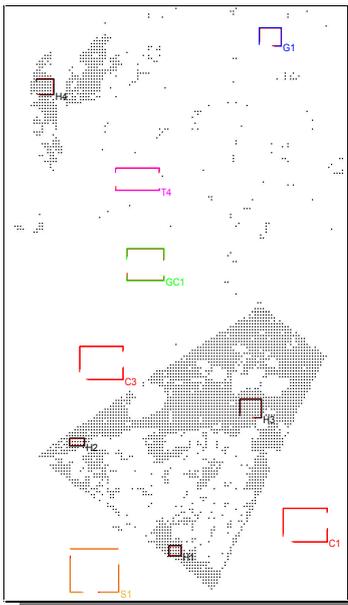


Fig. 5. Hogweed identification results obtained using classifier  $C_1$  in the lower left part of the processed multispectral image.

– calculate values for each fragment of  $I3result$  :

- mean values  $m1, m2, m3$
- standard deviations  $s1, s2$

– calculate result  $RES$  as binary image containing a logical 0 for a fragment of  $I3result$ , if  $m1 > \Delta_2$  &  $m2 > \Delta_2$  &  $m3 > \Delta_3$  &  $s1 > \Delta_4$  &  $s2 > \Delta_5$  and 1 otherwise.

Result  $RES$  obtained using classifier  $C_1$  is presented in Fig. 5 for enlarged lower left part of the multispectral image. It is a binary image where zeroes (black) refer to the detected hogweed fragments. Design regions are also marked. It is noticed that there are no hogweed fragments detected in design areas of different classes. Hogweed areas which can be clearly identified visually in the multispectral image are detected also in the results. On the other hand, there are hogweed fragments detected in other parts of the image where they are not noticed visually in multispectral image; that may indicate possible commission errors.

Colour versions of images can be downloaded from [12].

#### IV. CONCLUSIONS

Classifier  $C_1$  constructed on the basis of RVI and data from the design regions provides relatively precise identification of hogweed fragments of the processed multispectral image.

Although the classification rules exploit data from only 4 spectral bands, it was sufficient to obtain qualitative detection results of hogweed. No commission errors were observed in design regions of different classes.

Quite similar detection results of the hogweed fragments was obtained using the classifier  $C_2$  built on the basis of forming CCI<sub>m</sub>.

It remains an open question whether these classifiers will show similar characteristics when applied to multispectral images obtained in similar conditions using the same

equipment. Construction of classifiers on the basis of fragment size  $9 \times 9$  pixels did not affect the hogweed detection result. As we had no detailed field data about the vegetation presented in the multispectral image, we cannot be sure that all fragments where hogweed was detected contain these plants in reality. The same reason does not allow us to identify commission errors and calculate meaningful identification accuracy estimates.

Classifier  $C_2$  is considerably more sophisticated therefore we recommend usage of  $C_1$ .

#### ACKNOWLEDGMENT

The authors would like to thank the Institute for Environmental Solutions for making the multispectral data available, Dr. James Burger for introduction to the problem and sharing the preprocessed data, as well as M.Menke for his assistance.

#### REFERENCES

- [1] H. P. White, A. Abuelgasim, "Monitoring environmental remediation: hyperspectral mapping of re-vegetated areas affected by smelting operations in Sudbury, Canada", in *IEEE GRSS Workshop on Hyperspectral Image and Signal Process. Conf. (WHISPERS 2010)*, Reykjavik, Iceland, 2010, pp.1–4.
- [2] M. Dalponte, L. Bruzzone, D. Gianelle, "Fusion of hyperspectral and LIDAR remote sensing data for classification of complex forest areas", *IEEE Trans. Geosci. Remote Sens.*, vol. 46, pp. 1416–1427, 2008. [Online]. Available: <http://dx.doi.org/10.1109/TGRS.2008.916480>
- [3] P. Krahwinkler, J. Rossmann, "Tree species classification based on the analysis of hyperspectral remote sensing data", *30th EARSeL Symposium Remote Sensing for Science Education and Natural and Cultural Heritage*, pp. 321–328, 2010.
- [4] M. Dalponte, L. Bruzzone, D. Gianelle, "Tree species classification in the southern Alps with very high geometrical resolution multispectral and hyperspectral data", in *IEEE GRSS Workshop Hyperspectral Image and Signal Process. Conf. (WHISPERS 2011)*, Lisbon, Portugal, 2011, pp.1–4.
- [5] J. Mullerova, J. Pergl, P. Pysek, "Remote sensing as a tool for monitoring plant invasions: Testing the effects of data resolution and image classification approach on the detection of a model plant species *Heracleum mantegazzianum* (giant hogweed)", in *Int. Journal of Applied Earth Observation and Geoinformation*, pp. 55–65, 2013. [Online]. Available: <http://dx.doi.org/10.1016/j.jag.2013.03.004>
- [6] T. Shi, C. A. DiMarzio, "Multispectral method for skin imaging: development and validation", in *Applied Optics*, vol. 46, no. 36, pp. 8619–8626, 2007. [Online]. Available: <http://dx.doi.org/10.1364/AO.46.008619>
- [7] B. D'Alessandro, A. P. Dhawan, "Transillumination imaging for blood oxygen saturation estimation of skin lesions", *IEEE Trans. Biomedical Engineering*, vol. 59, no. 9, pp. 2660–2667, 2012. [Online]. Available: <http://dx.doi.org/10.1109/TBME.2012.2209647>
- [8] R. Dinuls, G. Erins, A. Lorencs, I. Mednieks, J. Sinica-Sinavskis, "Tree species identification in mixed Baltic forest using LiDAR and multispectral data", in *IEEE. Selected Topics Appl. Earth Observ. Remote Sensing*, vol. 5, no. 2, pp. 594–603, 2012.
- [9] G. L. Anderson, J. H. Everitt, A. J. Richardson, D. E. Escobar, "Using satellite data to map false broomweed (*Ericameria austrotexana*) infestations on south Texas rangelands", in *Weed Technol.*, vol. 7, pp. 865–871, 1993.
- [10] R. Dinuls, A. Lorencs, I. Mednieks, "Using consolidated covariance image for discrimination of habitats", in *Proc. 13th Biennial Baltic Electronics Conf.*, Tallinn, Estonia, pp. 299–302, 2012. [Online]. Available: <http://dx.doi.org/10.1109/BEC.2012.6376876>
- [11] R. D. Jackson, A. R. Huete, "Interpreting vegetation indices", in *Preventive Veterinary Medicine*, vol. 11, no. 3–4, pp. 185–200, 1991. [Online]. Available: [http://dx.doi.org/10.1016/S0167-5877\(05\)80004-2](http://dx.doi.org/10.1016/S0167-5877(05)80004-2)
- [12] *Simplified Classification of Multispectral Image Fragments*, Institute of Electronics and Computer Science. [Online]. Available: <http://www.edi.lv/lv/petijumu-virzieni/petijumi/hogweed>