

ACCOUNTING FOR VARIANCE IN HYPERSPECTRAL DATA COMING FROM LIMITATIONS OF THE IMAGING SYSTEM

B. Shurygin^{a,*}, M. Shestakova^a, A. Nikolenko^a, E. Badasen^a, P. Strakhov^a

^a Dept. of Aerophysics and Space Research, MIPT, Dolgoprudny, Russia -
lodinn@lodinn.com

Commission VII, WG VII/4

KEY WORDS: Supervised classification, decorrelation, Mahalanobis distance, Signal-to-noise ratio (SNR), Bayesian approach, Principal component analysis, Pixel-wise calibration.

ABSTRACT:

Over the course of the past few years, a number of methods was developed to incorporate hyperspectral imaging specifics into generic data mining techniques, traditionally used for hyperspectral data processing. Projection pursuit methods embody the largest class of methods employed for hyperspectral image data reduction, however, they all have certain drawbacks making them either hard to use or inefficient. It has been shown that hyperspectral image (HSI) statistics tend to display “heavy tails” (Manolakis2003)(Theiler2005), rendering most of the projection pursuit methods hard to use. Taking into consideration the magnitude of described deviations of observed data PDFs from normal distribution, it is apparent that *a priori* knowledge of variance in data caused by the imaging system is to be employed in order to efficiently classify objects on HSIs (Kerr, 2015), especially in cases of wildly varying SNR. A number of attempts to describe this variance and compensating techniques has been made (Aiazzi2006), however, new data quality standards are not yet set and accounting for the detector response is made under large set of assumptions. Current paper addresses the issue of hyperspectral image classification in the context of different variance sources based on the knowledge of calibration curves (both spectral and radiometric) obtained for each pixel of imaging camera. A camera produced by ZAO NPO Lepton (Russia) was calibrated and used to obtain a test image. *A priori* known values of SNR and spectral channel cross-correlation were incorporated into calculating test statistics used in dimensionality reduction and feature extraction. Expectation-Maximization classification algorithm modification for non-Gaussian model as described by (Veracini2010) was further employed. The impact of calibration data coarsening by ignoring non-uniformities on false alarm rate was studied. Case study shows both regions of scene-dominated variance and sensor-dominated variance, leading to different preprocessing parameters and, ultimately, classification results. A multilevel system for denoting hyperspectral pushbroom scanners calibration quality was proposed.

1. INTRODUCTION

Hyperspectral imaging is a flexible tool employed in a wide range of remote sensing applications. Said flexibility comes with a number of drawbacks, one of which is an extremely high degree of data redundancy in hyperspectral images. Combined with large image size, data redundancy often requires dimensionality reduction, which in turn serves as a stepping stone for the further image processing. A lot of non-specific data mining techniques for feature extraction are being employed in order to preserve the maximum amount of information, the most well-known ones being projection pursuit (PP) methods such as principal component analysis (PCA) and independent component analysis (ICA). However, as was shown by Manolakis (Manolakis2003) and Theiler (Theiler2005), hyperspectral data tends to display non-normal distribution of pixel brightness, rendering the base assumption of ICA incorrect, while PCA does not account for noise contribution in data variance, leading to over-estimation of informativity of dark and noisy channels. A number of modifications and adaptations of these algorithms has been developed over the course of past years, most notable examples being noise-adjusted principal component (NAPC) transform, non-linear PCA (NLPCA) and maximum noise fraction (MNF). As mentioned by Wang (Wang2006), PCA-based approaches tend to conceal “many subtle material substances uncovered by very high spectral resolution hyperspectral imaging sensors”, however, they are still one of the most accessible and efficient tools in the field of hyperspectral remote sensing.

This success should be mainly attributed to the lack of information about the dataset - dimensionality reduction is most often required in cases where little to no spectra of the objects on the acquired scene are known *a priori* or obtained during the series of ground measurements. At present, a few other means to incorporate auxiliary information in the processing algorithms were made, including previously mentioned NAPC and its modification, INAPCA (interference and noise-adjusted principal components analysis (Chang1999)), which can attribute for noise fraction and interference in hyperspectral imaging systems. These approaches, however, are often inapplicable by the end users due to lack of information about the noise.

In this work, we analyze the impact of employing the pixel-wise radiometric and spectral calibration curves on the ability to extract the meaningful signal and filter the noise from the data during supervised classification and propose a method to estimate a noise fraction and covariance matrix based on sensor specifications.

2. NOISE SOURCES IN HYPERSPECTRAL DATA

2.1 Noise types

In most hyperspectral imaging systems, the only source of a random noise is the photosensitive matrix, often presented by a charge-coupled device (CCD). Signal measured by a camera consists of 4 main components - the light intensity intended to be measured, thermal noise background (dark current), shot noise, and read-out noise. Signal-to-noise ratio for such system is given by the

*Corresponding author

equation

$$SNR = \frac{N_{sig}}{\sqrt{N_{sig} + N_{dark} + N_{read}^2}},$$

where N_{sig} is the number of light-induced electrons detected, N_{dark} is the number of thermal electrons detected, and N_{read} accounts for the readout noise. As it can be seen, depending on signal level expected SNR may vary, being significantly lower for regions with low incident light intensity and/or low quantum efficiency of the CCD in question.

Another major contribution to the decline in image quality caused by imaging system limitations is improper accounting for the image forming, including blur, infamous "smile effect", striping and other artifacts.

2.2 Calibration quality levels

For the given hyperspectral sensor, full calibration data would include spectral response curves and radiometric response depending on exposure time on per-pixel basis. However, this data is usually not available for the end user. Instead, simplifications are made. Most commonly, spectral response curves are substituted with channel central wavelengths and widths; sometimes, a single background and/or radiometric scaling coefficient are assigned for the whole photosensitive matrix in the stead of a matrix of coefficients. Depending on a system design, that may lead to the multiple artifacts, as it was mentioned in the previous section. In this article, we propose a grade system for the calibration data quality and availability for the end user, from best to worst:

1. Spectral and radiometric response curves are available for each pixel, auxiliary flight data allows to match acquisition parameters with ones used in pre-flight calibration.
2. Spectral and radiometric response curves were measured on per-pixel basis and simplified to the few parameters, namely dark current, radiometric calibration curve slope, pixel central wavelength and width of spectral response curve.
3. Channel width is not measured, central wavelength and radiometric coefficients are available on per-pixel basis.
4. Calibration data (central wavelength, spectral width, dark current and radiometric coefficients) is available on per-channel basis.
5. Channel width is not measured, only central wavelength and radiometric coefficients are available on per-channel basis.
6. Channel width is not measured, central wavelength are known on per-channel basis, radiometric coefficients are available only for the matrix as a whole.

Normally, end user would operate with the levels 3-4 of calibration data. For this work, we had access to level 1 quality of calibration data according to the scheme above, which was subsequently brought down to level 2. In the course of the presented research, we imitated levels 3-5 of calibration data as well, extracting them from the selfsame level 1 dataset. Extraction was done by fitting data using MATLAB and omitting parts of the data set wherever necessary. Initial data set comprised 87 millions of data points (approximately 330 Gb), level 2 data included 580000 parameters (2.2 Gb), level 3 data had one parameter less than level 2 data for each pixel, resulting in 1.67 Gb of calibration data. Finally, levels 4 and 5 yielded approximately 1 Mb of calibration data each. Level 6 contains a negligible amount of data and results in the fastest, but the most inaccurate pre-processing

of an image.

It can be seen that level 1 data is highly impractical, levels 2 and 3 are usable, but in many cases, levels 4 through 6 would be preferred because of less stored data and significantly faster processing.

3. INCORPORATION OF CALIBRATION DATA INTO CLASSIFICATION ALGORITHMS

3.1 Preparation of calibration data for testing

From the equation for SNR in section 2.1, lower signal values result in higher noise fraction. For hyperspectrometers with varying channel width (prism-based), this means that areas with finer spectral resolution tend to be intrinsically more noisy than areas with coarser spectral resolution, and for the normal operating mode of hyperspectral camera, when SNR is photon noise-limited, the SNR ratio is expected to be proportional to the square root of channel width under assumption there is no other *a priori* knowledge about signal levels in different spectral ranges. Other way of estimating it is the inverse square root of the radiometric calibration coefficient, which is also more precise since it takes into account possible vignetting and other sources of signal dwindling towards the edges of the CCD matrix.

The camera we used in this research was a pushbroom hyperspectrometer made by ZAO NPO Lepton (Russia) and it had the following parameters:

- Spectral range 400–1000 nm;
- Channel width 2–9 nm;
- 290 spectral channels;
- 500 pixels cross-track;
- Field of view 3.2°.

Dark current background was highly uniform over the matrix, whereas radiometric multiplicative calibration coefficients were not (fig. 1). For the further research, this level 2 calibration data

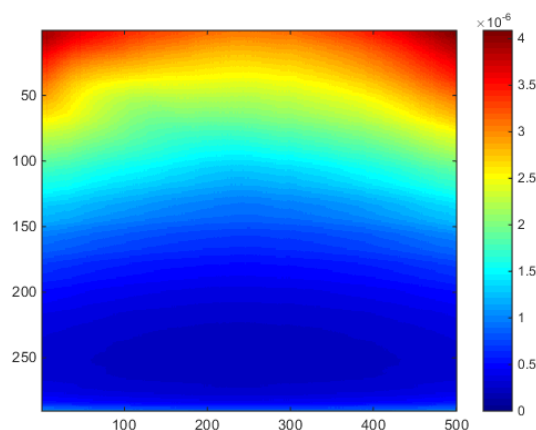


Figure 1: Radiometric calibration coefficients for the camera used in research.

was brought down to the level 4 by averaging radiometric response over channels. Relative error

$$err_i^j = \frac{C_i^j - \bar{C}_i}{\bar{C}_i},$$

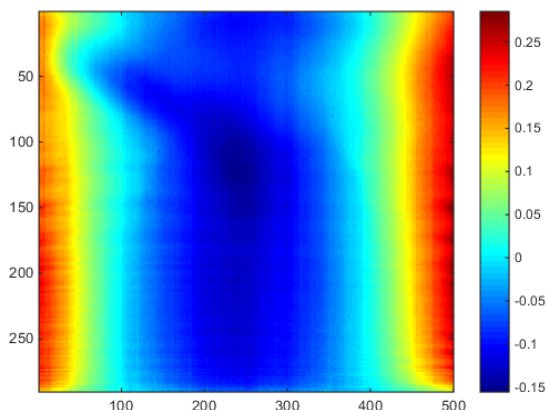


Figure 2: Relative error in radiometric calibration coefficients after whitening.

where C is the radiometric calibration coefficient, i denotes channel number, j denotes cross-track coordinate, is presented on the figure 2. Notably, shown distribution of relative error results in darker edges of the scene - an artifact commonly found in hyperspectral images. In a similar manner, spectral calibration was also handled.

Distribution of the radiometric calibration coefficients shown on figure 1 reveals yet another problem - even with constant imaged object brightness over the whole spectral range, signal values would vary up to 8 times due to different channel widths and quantum efficiency of the matrix depending on the spectral region. This means that regardless of the actual hyperspectral scene, channels in the blue spectral region are expected to have at least 4 times lower SNR than channels in the green spectral region, in addition, they have also significantly higher cross-correlation.

3.2 Utilization of calibration data in the feature-weighted Mahalanobis distance

We propose to explicitly compensate for these factors during supervised classification. For that, we utilize a feature-weighted Mahalanobis distance similar to one described in (Wolfel2005), but using our estimations for noise fraction as weights:

$$D = \sqrt{(L - L^k)^T F S^{-1} (L - L^k)},$$

where L is the pixel radiance, L^k is the mean radiance for the user-selected class k , S is the covariance matrix of the training set, and F is a noise and interference-whitening matrix identical to one described in (Chang1999).

This approach requires the prior knowledge of a noise covariance matrix. Since the noise is signal-dependent, we propose to further modify it so that F in the equation above is substituted for gF , where F is calculated under assumption of signal uniformity over the spectral range and g is a diagonal weighting matrix accounting for the variance of noise caused by signal non-uniformity. That selection allows for the fast calculations since both F and S^{-1} are independent of data point and g is diagonal. Following the MNF ideology, we propose g of a form

$$g_{i,j} = \begin{cases} \sqrt{DN_i^k - DN_{dark}}, & \text{if } i = j \\ 0, & \text{otherwise,} \end{cases}$$

where DN stands for digital number - raw signal before applying radiometric calibration coefficients. It is worth noting that the gFS^{-1} kernel version with g being considered an identity

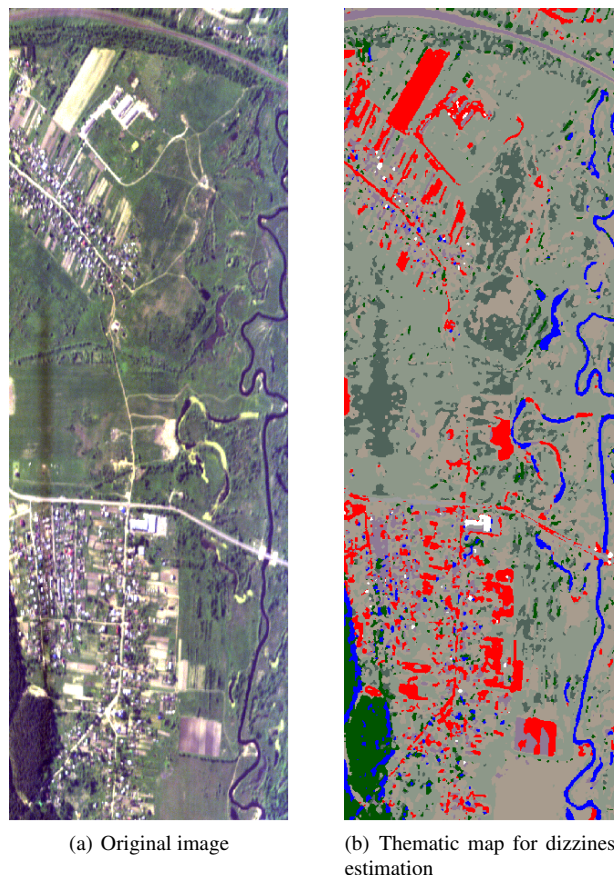


Figure 3: Image selected for processing

matrix is an *a priori* generated weighting matrix, while version with DN modifies the very concept of Mahalanobis distance and makes it not scale-invariant (brighter areas are considered more informative).

For the research presented in this paper, modified version of Mahalanobis distance with DN-based noise whitening matrix calculation was used.

4. DATA AND RESULTS

For analysis of calibration data availability impact on the quality of supervised classification while employing the proposed approach we have used a hyperspectral image obtained by a camera presented above. That image was obtained on 23 May, 2014. A fragment of the image used for analysis as well as the thematic map produced using the threshold-based decision tree are presented on the figure 4. Namely, channels with central wavelengths closest to 450, 500, 550, 600, 670, 750 and 860 nm were selected, maximum likelihood classifier was employed and the result was spatially filtered. This way, an artificially produced reference map representing spectral regions with important spectral features was acquired. While this result does not pertain an absolute scale of measuring classifier performance, it does guarantee that the supertypes of objects, such as vegetation, man-made materials, soil or water should be appropriately detected. Calibration data level 2 was used to pre-process the hyperspectral image, while level 4 of data is enough to perform all the necessary computations during classification. For classification purposes, a training set for 10 classes was provided. It included 5 types of man-made materials, water, soil and 3 types of vegetation. Classification was done using unmodified Mahalanobis distance and feature-weighted Mahalanobis distance. Results are

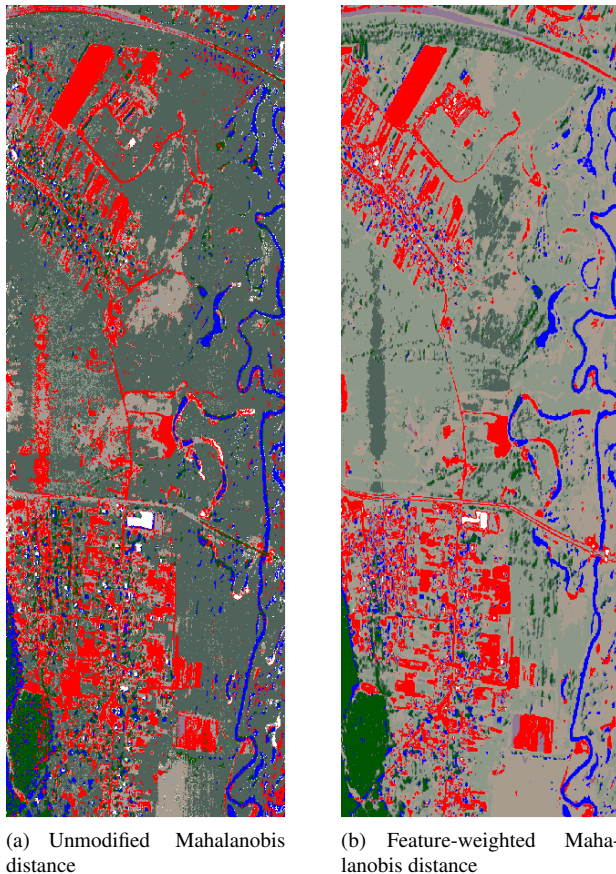


Figure 4: Classification results

presented on figure 4. Both of the classification results were not post-processed. It is clearly seen that the proposed modification led to significant reduction of contribution of noisy channels into final result. While using non-modified Mahalanobis distance, huge part of the grass-like vegetation was erroneously attributed to forest-like vegetation. After combining 10 classes into 3 large clusters - water, vegetation, man-made materials - it was revealed that the non-modified Mahalanobis distance classifier yielded 17.3% cases of wrong detection, while for modified one this value is as low as 2.2%.

5. CONCLUSION

We have proposed and tested MNF-based approach to hyperspectral image classification, which is based on having a complete set of calibration information for the imaging system. It has been shown that supervised classification quality is significantly improved when employing such data, both from perspective of dizziness and erroneous detection as compared to human-produced map with the usage of the known wavelengths containing important spectral features. Proposed approach does not significantly increase computation time needed for the classification, however, it requires a large amount of information about the calibration data. In this research, amount of data needed to perform the image preprocessing was about 2.2 Gb, with estimated relative error up to 27% in case of not having a complete set of data and using averaged values instead.

ACKNOWLEDGEMENTS

These results are obtained under funding support from the Federal Target Program Research and Development in Priority Ar-

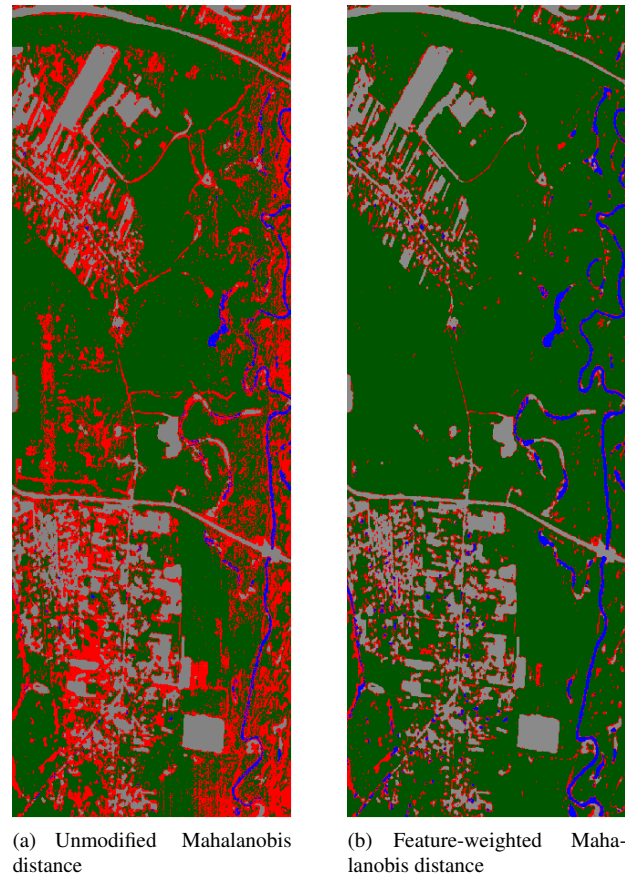


Figure 5: Classification results after combining classes into 3 large clusters, erroneous detection cases are shown in red.

eas of Development of the Russian Scientific and Technological Complex for 2014-2020 (Grant Agreement No. 14.575.21.0028, unique identification number RFMEFI57514X0028).

REFERENCES

- Aiazzi, B., 2006. Noise modelling and estimation of hyperspectral data from airborne imaging spectrometers. In: *Ann. Geophys.*, vol. 49, no. 1, pp.1-9.
- Chang, C., 1999. Interference and noise-adjusted principal components analysis. In: *IEEE Transactions on Geoscience and Remote Sensing*, vol. 37, p. 5, pp. 2387-2396.
- Kerr, G., 2015. Quality assessment for hyperspectral airborne systems. Diss. Humboldt-Universität zu Berlin, Mathematisch-Naturwissenschaftliche Fakultät, Berlin, Germany. <http://edoc.hu-berlin.de/dissertationen/kerr-gregoire-henry-gerard-2015-07-30/PDF/kerr.pdf> (6 Jan 2016).
- Manolakis, D., 2003. Hyperspectral image processing for automatic target detection applications. In: *Lincoln Laboratory Journal*, 14(1), pp.79-116.
- Theiler, J., 2005. Characterizing Non-Gaussian Clutter and Detecting Weak Gaseous Plumes in Hyperspectral Imagery. In: *Defense and Security*, International Society for Optics and Photonics, pp. 182-193.
- Veracini, T., 2010. Robust hyperspectral image segmentation based on a non-gaussian model. In: *2nd "International Workshop on Cognitive Information Processing (CIP)"*, (pp. 192-197), IEEE.

Wang, J., 2006. Independent component analysis-based dimensionality reduction with applications in hyperspectral image analysis. In: *IEEE Transactions on Geoscience and Remote Sensing*, vol. 44, no. 6, pp. 1586-1600.

Wolfel, M., 2005. Feature weighted Mahalanobis distance: improved robustness for Gaussian classifiers. In: *13th European Signal Processing Conference*, IEEE.