Deep Learning Concept for Hyperspectral Imagery Classification

Sergey A. Stankevich, Iryna A. Piestova, Victor N. Podorvan

Abstract — The adaptation of deep learning concept for hyperspectral imagery classification is presented. The concept proposed a six-layer processing methodology, which includes a pre-processing, an optimal band selection procedure, both spectral feature-based and spatial feature-based land cover classification as well as land cover change detection process using formal schema.

Keywords — deep learning, hyperspectral imagery classification, band selection, change detection.

I. Introduction

Along with the technology development and the improvement of hyperspectral imagery both spectral and spatial resolution, the hyperspectral data requires a deeper analysis due to the increased amount and complexity of observable phenomena [1]. Representation and organization of multiple levels in order to express complex relationships among data defines a deep learning process. One of the promises of deep learning is replacing handcrafted features with efficient algorithms for unsupervised or semi-supervised feature learning and hierarchical feature extraction. Feature hierarchies are discovered considering that higher levels are formed by combining features from lower level [2].

II. Concept and methods

Deep learning is a branch of machine learning based on a set of algorithms that attempt to model high-level abstractions in data using multiple processing layers. For each level, models include complex structures or otherwise, composed of multiple non-linear transformations. This one must be developed according to the type of information desired, namely pixel-based, object-based or structure-based. In respect to the hyperspectral imagery analysis, the deep learning general methodology involves several sequential and parallel data processing levels [3]: pre-processing, band selection, preliminary classification – both spectral feature-based and spatial feature-based, class fusion as well as change detection if needs.

A. Pre-processing

Pre-processing level operations include geo-referencing, geometric, radiometric and atmospheric correction and can sometimes supplied by satellite imagery provider. In case of multiple images analysis initially they are matched using ground control points (GCP). Hyperspectral image digital numbers (DN) are converted into at-sensor spectral radiance using each-band calibration factors extracted from image metadata. Next, the atmospheric correction is performed and floating-point values of the surface spectral reflectance are obtained [4].

S. Stankevich, Scientific Centre for Aerospace Research of the Earth, Institute of Geological Science, National Academy of Sciences of Ukraine, Kiev, Ukraine (e-mail: st@casre.kiev.ua).
I. Piestova, Scientific Centre for Aerospace Research of the Earth, Institute of Geological Science, National Academy of Sciences of Ukraine, Kiev, Ukraine
V. Podorvan, Scientific Centre for Aerospace Research of the Earth, Institute of Geological Science, National Academy of Sciences of Ukraine, Kiev, Ukraine

B. Band Selection

These operations consist in selection of informative bands of hyperspectral image [5]. The $C(\lambda)$ informativity criterion is used for the hyperspectral imagery spectral bands selection in remote sensing applications [6]:

$$C(\lambda) = \frac{D(\lambda)}{4r^2(\lambda)} \cdot \log_2[1 + \psi(\lambda)] \tag{1}$$

where $D(\lambda)$ is Kullback-Leibler divergence in the spectral space, $r(\lambda)$ is an equivalent spatial resolution of current spectral bands set, $\psi(\lambda)$ is an equivalent signal-to-noise ratio in target detection by multi-dimensional optical signal, λ is a spectral bands subset.

The best of spectral bands subset selection of hyperspectral image in terms of (1) criterion is an optimal search problem in hyperspectral image spectral bands combinations space. This problem can be solved using one of the known methods, e. g. pseudo-gradient search [7].

C. Spectral feature-based classification

Spectral feature-based classification is the analysis of object or material features by spectral values. This one detects desired target spectra in image for remote sensing applications.

One of the widely used algorithms such as support vector machine (SVM) [8] or spectral topological classifier (STC) [9] can be applied for classification by spectral features. The following equation described the classification rule for STC:

$$\Delta z = \frac{1}{m}\sum_{\lambda} |\,\mathrm{sign}(\Delta E_{\lambda}, \sigma_{\lambda})\,| + \frac{1}{2(m-1)}\sum_{\lambda} |\,\Delta\,\mathrm{sign}\left(\frac{\partial E_{\lambda}}{\partial \lambda}, \frac{\partial \sigma_{\lambda}}{\partial \lambda}\right)| \tag{2}$$

where, $\mathrm{sign}(E,\sigma) = \begin{cases} 1 \text{ if } E > \sigma \\ 0 \text{ if } |E| \le \sigma \\ -1 \text{ if } E < -\sigma \end{cases}$.

D. Spatial feature-based classification

There is an object-oriented classification of image targets by spatial features. Especially the Radon transform can be used to detect image targets by its shape [10] or the Hough transform can be used to detect linear patterns [11].

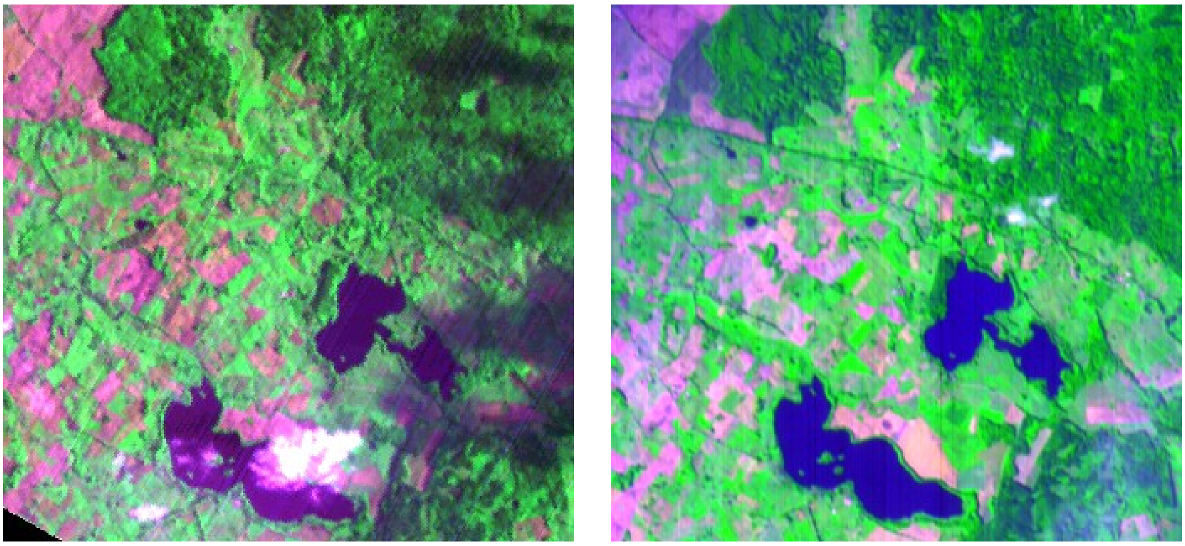
Both classification types generate the corresponding preliminary probability maps for image objects. They are provided by a fundamentally different identification features.

E. Change detection

In remote sensing applications changes are considered as land surface component alterations. Change detection is the process of identifying and quantifying differences in the objects state using multi-temporal satellite images [12]. Post-classification comparison separately classified multitemporal images, pixel by pixel. Engagement classification maps instead raw images minimizes impacts of atmospheric, sensor and environmental differences between multitemporal datasets, as well as provides a complete matrix of change information [13].

III. Results and discussion

The methodology described in the previous section was applied to the hyperspectral satellite image analysis of territory Lekuk, Poland. The PROBA-1 satellite image by Compact High Resolution Imaging Spectrometer (CHRIS) was investigated (Fig. 1). CHRIS instrument provides 19 spectral bands (fully programmable) in the VNIR range (400 – 1050 nm) at a 17 m ground sampling distance (GSD) [14]. Two multitemporal source hyperspectral images were acquired in similar conditions for the same phase of vegetation phenological cycle.

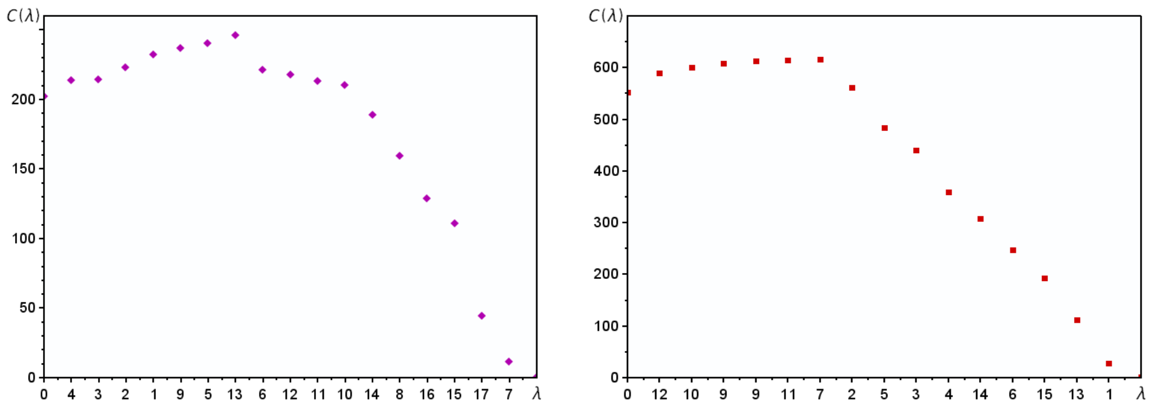


a *b*

Fig. 1 CHRIS hyperspectral satellite images (19 spectral bands, 17 m resolution, pseudo-natural color synthesized) of test site in Lekuk region (Poland): August 03, 2014 (*a*) and August 20, 2015 (*b*)

Pre-processing of the both images was performed, and then the two images were co-registered with pixel accuracy using GCP.

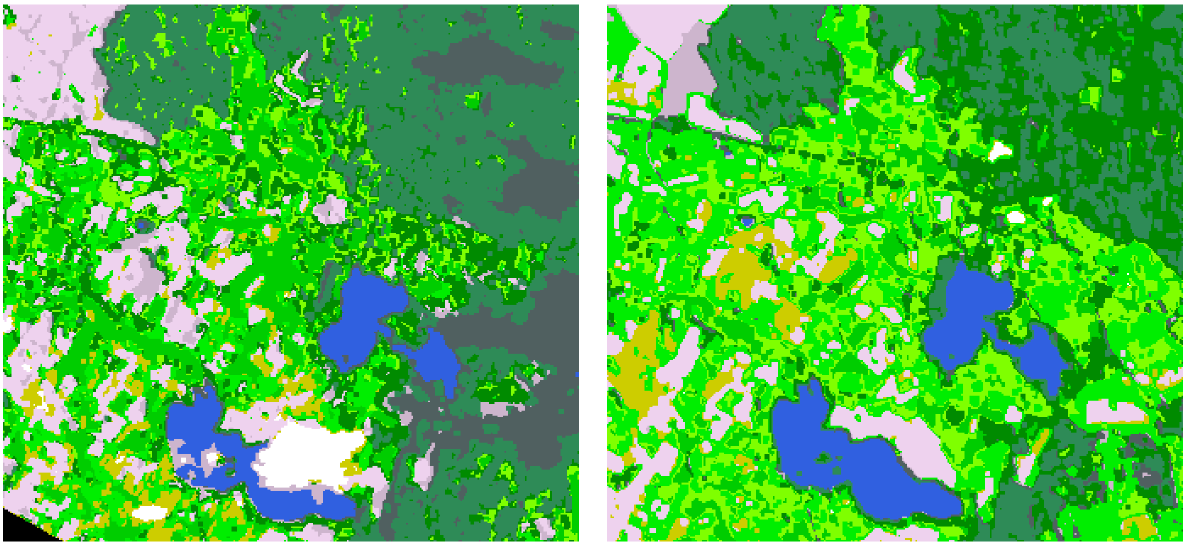
The process of optimal band selection using (1) criterion is illustrated by Fig. 2 diagram. Totally 11 spectral bands of 18 were selected for further analysis from the image of 2014 and 10 ones from 16 – from the image of 2015.



a *b*

Fig. 2 Optimal band selection for CHRIS hyperspectral imagery of test site in Lekuk region (Poland): August 03, 2014 (*a*) and August 20, 2015 (*b*)

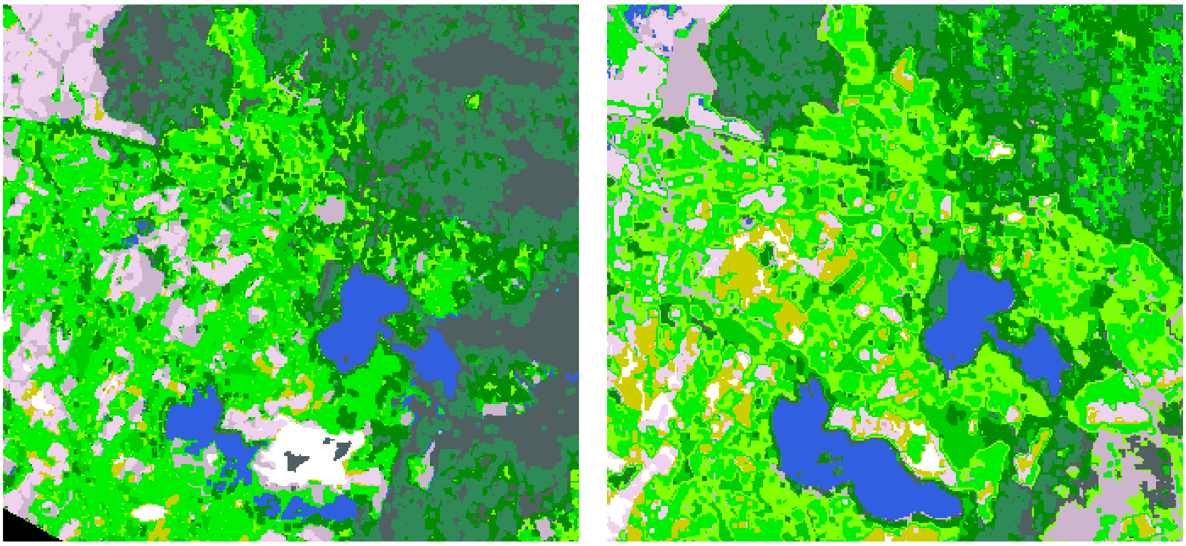
Spectral-feature classification of land cover into eight primary classes of woodlands and farmlands was performed using the SVM algorithm (Fig.3) and STC (Fig.4). The weather conditions greatly affect the classification results. It is best to use clean cloudless pictures. Otherwise, separate classes for clouds and cloud shadows were created. The area occupied by these classes does not participate in change detection process.



a *b*

Fig. 3 Land cover classification by SVM algorithm (12 unified classes) of test site in Lekuk region (Poland): August 03, 2014 (*a*) and August 20, 2015 (*b*)

Legend: – Unclassified, – Water, – Coniferous, – Deciduous, – Barren, – Cropland, – Sparse, – Meadow, – Grassland, – Cloud, – Shadows, – Wetland



a *b*

Fig. 4 Land cover classification by STC (12 unified classes) of test site in Lekuk region (Poland): August 03, 2014 (*a*) and August 20, 2015 (*b*)

Legend: – Unclassified, – Water, – Coniferous, – Deciduous, – Barren, – Cropland, – Sparse, – Meadow, – Grassland, – Cloud, – Shadows, – Wetland

The SVM algorithm is more resistant to image noise and radiometric distortions. Vice versa, the STC more accurately separates the classes with complex spectra.

Since the objects under study haven't expressed geometric features, in this research the spatial-based processing performed as a post-classification by clumping.

After mapping of changes occurring at different-time land cover classifications of study area

(Fig. 3, Fig. 4) the spatial distribution of their importance has been obtained [15]. The expert pair-class estimate matrix is given in Table I.

TABLE I
EXPERT CLASS-CHANGE ESTIMATE MATRIX

Land cover	0	1	2	3	4	5	6	7	8	9	10	11
0. Unclassified	0	0	0	0	0	0	0	0	0	0	0	0
1. Water	0	4	7	7	5	5	6	6	6	0	0	6
2. Coniferous	0	1	4	4	1	2	2	3	3	0	0	2
3. Deciduous	0	1	4	4	1	2	2	3	3	0	0	2
4. Barren	0	3	7	7	4	5	5	6	6	0	0	6
5. Cropland	0	3	6	6	3	4	5	6	6	0	0	6
6. Sparse	0	2	6	6	3	3	4	4	5	0	0	4
7. Meadow	0	2	5	5	2	2	4	4	4	0	0	4
8. Grassland	0	2	5	5	2	2	3	4	4	0	0	4
9. Cloud	0	0	0	0	0	0	0	0	0	0	0	0
10. Shadows	0	0	0	0	0	0	0	0	0	0	0	0
11. Wetland	0	2	6	6	2	2	4	4	4	0	0	4

Evaluation of changes made based on the analysis of the impact occurring, or possible changes to the ecological status of the territory. The expert chooses the code of change, which is subdivided into 7 estimates: three negative, three positive and one neutral.

Both positive and negative changes were subdivided into three levels each for easy visualization, as shown in Fig. 5.

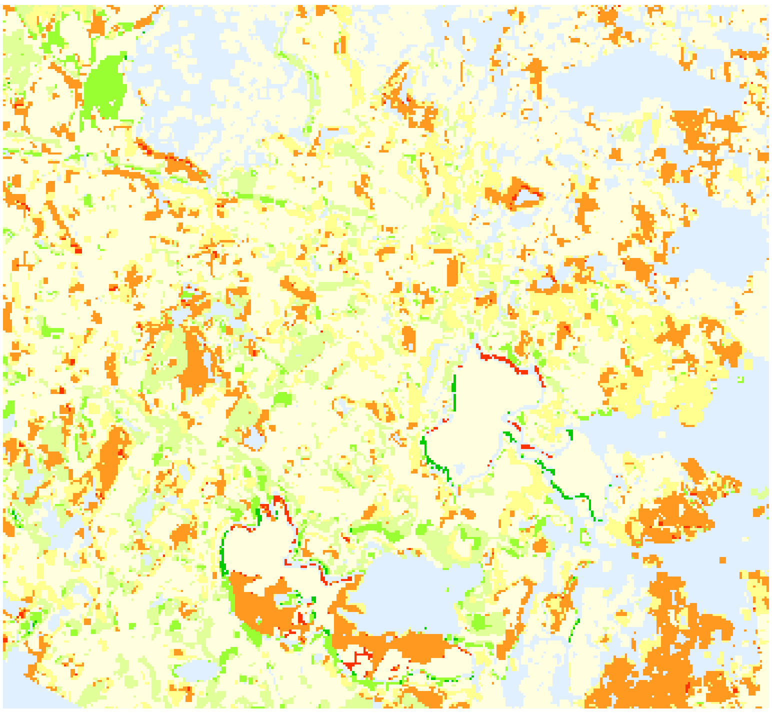


Fig. 5 Spatial distribution of land cover change estimates of test site in Lekuk region (Poland) during 2014-2015: in levels of importance.
Legend: – strong negative, – moderate negative, – weak negative, – neutral, – weak positive, – moderate positive, – strong positive

In accordance with the methodology described in paragraph 2, the process of classification of hyperspectral imaging is illustrated by dataflow diagram (Fig. 6). It consists of the following layers: level 1 – input data, level 2 – pre-processing, level 3 – band selection, level 4 – classification, level 5 – change detection, level 6 – output.

Output level of Fig.6 is already suitable for use as a final data product remote sensing applications.

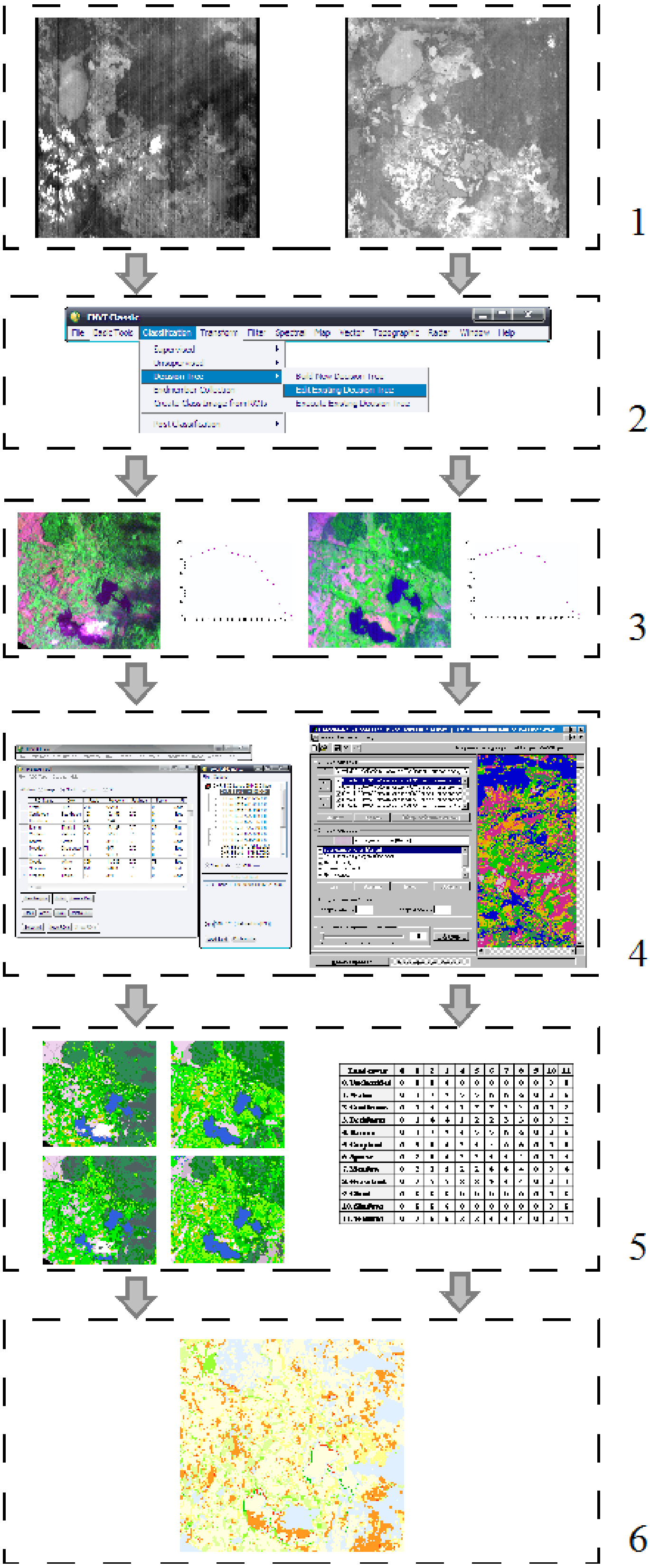


Fig. 6 Dataflow diagram for hyperspectral imagery classification based on deep learning concept

IV. Conclusion

Thus the deep learning concept for hyperspectral imagery classification methodology is proposed. The methodology includes six levels of hyperspectral imagery analysis, from raw hyperspectral image up to final data product.

The proposed concept is the basis for formalizing the hyperspectral imagery classification process. Such formalization is very important in developing the information systems for Earth's surface monitoring or the similar geospatial services [16].

References

[1] M. Dalponte, L.Bruzzone, L. Vescovo, D. Gianelle, “The role of spectral resolution and classifier complexity in the analysis of hyperspectral images of forest areas,” *Remote Sensing of Environment*, vol. 113, pp. 2345–2355, November 2009.

[2] C. Vaduva, I. Gavat, M. Datcu, “Deep learning in very high resolution remote sensing image information mining communication concept,” *Proceedings of the 20th European Signal Processing Conference (EUSIPCO)*, Bucharest, Romania, pp. 2506–2510, August 2012.

[3] J. Li, P.R. Marpu, A. Plaza, J.M. Bioucas-Dias, J.A. Benediktsson, “Generalized composite kernel framework for hyperspectral image classification,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 51 (9), pp. 4816–4829, March 2013.

[4] G. Shaw, D. Manolakis, “Signal processing for hyperspectral image exploitation,” IEEE *Signal Processing Magazine*, vol. 19 (1), pp. 12–16, January 2002.

[5] G. Petrie, P. Heasler, T. Warner, “Optimal band selection strategies for hyperspectral data sets,” *Proceedings of IEEE International Geoscience and Remote Sensing Symposium*, Seattle, USA, vol. 3, pp. 1582–1584, July 1998.

[6] S.A. Stankevich, “Quantitative estimation of hyperspectral aerospace imagery informativity in remote sensing applications,” *Reports of the National Academy of Sciences of Ukraine*, no. 8, pp. 53–58, August 2006.

[7] S.A. Stankevich, “Hyperspectral aerospace imagery spectral bands optimal selection in remote sensing applications,” *Space Science and Technology*, vol. 13, No. 2, pp. 25–28, April 2007.

[8] J. Shawe-Taylor, N. Cristianini, “An Introduction to Support Vector Machines and Other Kernel-based Learning Methods,” Cambridge University Press, 204 p, March 2000.

[9] M.A. Popov, S.A. Stankevich, V.N. Podorvan, “Software and technology for land cover classification by hyperspectral satellite imagery,” *Proceedings of First Scientific Conference “Earth and Space Sciences for Society”*, Kiev, Ukraine, p. 122, June 2007.

[10] F. Peyrin, R. Goutte, “Image invariant via the Radon transform,” *Proceedings of the IEEE International Conference on Image Processing and its Applications*, Maastricht, Netherlands, pp. 458-461, April 1992.

[11] J. Illingworth, J. Kittler, “A survey of the Hough transform,” *Computer Vision, Graphics and Image Processing*, vol. 44 (1), pp. 87–116, October 1988.

[12] M. Hussain, D. Chen, A. Cheng, H. Wei, D. Stanley, “Change detection from remotely sensed images: From pixel-based to object-based approaches,” *ISPRS Journal of Photogrammetry and Remote Sensing*, vol.80, pp. 91–106, June 2013.

[13] D. Lu, P. Mausel, E. Brondizio, E. Moran, “Change detection techniques,” *Int. J. Remote Sensing*, vol. 25, pp. 2365–2407, June 2004.

[14] R. Duca, F. Del Frate, “Hyperspectral and multiangle CHRIS-PROBA images for the generation of land cover maps,” *IEEE Transactions on Geoscienceand Remote Sensing*, vol. 46 (10), pp. 2857–2866, October 2008.

[15] S.A. Stankevich, A.A. Kozlova, “Long-term land cover change computer-aided mapping by remote sensed imagery,” *Proceedings of the International Conference on Information and Digital Technologies (IDT 2015)*, Žilina, Slovakia, pp. 327–329, July 2014.

[16] J. Zhang, Y. Zeng, W. Xie, T. Wang, “Decision making based on Earth observation technology,” In: Deren Li, Jie Shan, Jianya Gong (Eds.), “Geospatial Technology for Earth Observation”. Springer, pp. 497–530, September 2009.