Method of Hyperspectral Satellite Image Classification under Contaminated Training Samples Based on Dempster-Shafer's Paradigm

Mikhail A. Popov, Sofiya I. Alpert, Victor N. Podorvan, Maksym V. Topolnytskyi, Serhii I. Mieshkov

Abstract **-** The aim of this paper is to show that evidence theory of Dempster-Shafer may be successfully applied to supervised classification in remote sensing. A new method to classify hyperspectral satellite images under contaminated training samples based on the use of the evidence theory is proposed. The flowchart of the classification process is described. The algorithmic and software support is developed for the evidence theory-based method. The results of experiment with real hyperspectral image showed the effectiveness of the proposed method.

Keywords - contaminated training sample, evidence theory of Dempster-Shafer, hyperspectral satellite image, image classification.

I. INTRODUCTION

The process of solution of actual scientific and practical problems, using hyperspectral satellite images (HSI) as a rule includes a procedure of its classification [1]. The most accurate results are provided by supervised classification method, which uses a priori information about the characteristics of the classes. This information is extracted from the training sample. But, in fact to extract completely pure training samples is not possible. For example, the training subsample of the vegetation may contain spots of bare soil, paths, artificial objects, etc. In addition, the adequacy of the pixel signals can be influenced by the conditions of sunlight and atmospheric correction. Altogether, these factors lead to deterioration in the quality of training sample. In such cases we'll use the term "contaminated samples".

The impact of the training sample contamination on the classification accuracy was studied in number of publications. C. Ruwet and G. Haesbroeck have explored impact of contamination of training dataset in statistical clustering using the k-means algorithm [2]. Theoretical framework for determining the loss in the accuracy of remote sensing image classification was proposed and studied by D. Yan [3]. C. Persello and L. Bruzzone have investigated classification problems characterized by few training samples and affected by sample selection bias [4].

In general, it was shown that even small amount of contamination in a training sample can lead to appreciable errors. Thus, the problem of effective classification under contaminated training samples has not yet received complete solution and saves its relevance.

In the second half of the past century, A. Dempster and G. Shafer put forward new paradigm the essence of which is to simulate uncertainty by using a probability range rather than a single probability value [5]. This paradigm got embodied in the evidence theory, which has some advantages over traditional probability-theoretical approach and expands the range of practical applications [6]-[8]. In this study, we show that the methods of Dempster-Shafer evidence

M. A. Popov, Authors are with Scientific Centre for Aerospace Research of the Earth, National Academy of Sciences of Ukraine, Kyiv, Ukraine (e-mail: mpopov@casre.kiev.ua)

theory (DST) can be useful in solving the above problem image classification under contaminated training samples.

The paper is organized as follows. After the formulation of the problem we provide basics and notions of DST and describe the essence of the proposed method of classification of the HSI. With the necessary degree of detail there has been described the organization of the classification process and content of its procedures. In the final part the article considers the results of experimental studies of the accuracy assessment of the proposed method with various degrees of contamination of the training samples.

II. PROBLEM

The hyperspectral image (HSI) consists of a set of spectral images:

$$S_k = \{\pi_n, u_{nk}\}_{n=1}^{N_\pi}; \quad k = 1,2,...,K, \tag{1}$$

where S_k is the *k*-spectral image; K is the total number of spectral images (spectral bands); π_n is the *n*-th pixel; N_π is the total number of pixels in the HSI; u_{nk} is the *k*-component of full signal $\mathbf{u_n}$ of the pixel π_n.

The full signal $\mathbf{u_n}$ of an arbitrary pixel π_n convenient to consider as a vector with components u_{nk} in the spectral space Λ^K which created by axes $(\lambda_1,...,\lambda_k,...,\lambda_K)$, i. e.:

$$\mathbf{u_n} = \{u_{nk}\}_{k=1}^{K}; \quad u_n \in \Lambda^K. \tag{2}$$

Each π_n pixel of the HSI displays an object of some class; aim of pixel-wise classification is to determine the class of the pixel object π_n as accurately as possible, based on the analysis of the $\mathbf{u_n}$ signal.

In this paper, the problem of HSI pixel-wise supervised classification is solved under the following assumptions and limitations:

1. The nomenclature of the classes to which objects in the image can belong to is known, and training sample with representatives of each class is also available;
2. The results of classification must be unambiguous; i. e. is each pixel may belong to only one class it most closely resembles;
3. Quality of training samples is limited, within the training samples may be present extraneous elements and noise. In addition, forming a training sample, the expert can make mistakes.

The objective of our research is to develop a simple and effective method to HSI classification for above assumptions and limitations. Since in developing this method we lean on the Dempster-Shafer's paradigm, so it is useful to outline the basic concepts of the evidence theory.

III. BASIC CONCEPTS OF DEMPSTER–SHAFER EVIDENCE THEORY

Suppose Ω is a set of hypotheses about membership of pixel or object. DST allows considering any subsets of the Ω set. Total number of such subsets can be 2^Ω (the empty set Ø and the set Ω also included in this number). Applied to image classification, it means that not only single classes (singletons) but also any combinations of classes may be considered.

In DST with every hypothesis A $(A \in 2^\Omega)$ is associated so-called mass $m(A)$. The $m(A)$ mass value represents the degree of belief allocated to the hypothesis *A*; it belongs to the interval

[0, 1] and satisfies the following conditions:

$$\sum_{A \subseteq 2^{\Omega}} m(A) = 1;$$

$$m(\varnothing) = 0.$$

The mass $m(A)$ also referred to the basic probability. Any subset A whose $m(A) > 0,$ is called the focal subset [5].

Having the basic probability distribution, one can compute the level of hypothesis support. For this purpose it was introduced two measures: belief function *Bel(A)* and plausibility function *Pl(A).*

Belief function (4) measures the minimum or necessary support for the hypothesis whereas plausibility function (5) reflects the maximum or potential support for that hypothesis. These two functions are respectively defined as follows:

$$Bel(A) = \sum_{B \subseteq A} m(B), \tag{3}$$

$$Pl(A) = \sum_{B \cap A \neq \varnothing} m(B). \tag{4}$$

An important advantage of the DST is the presence of simple rule of combining data provided by many sources [5]. Suppose one source appointed to the hypothesis *A* the mass value $m_1,$ and the other source independently appointed to the same hypothesis the mass value m_2. The rule of Dempster allows combining these masses; the resultant is given by the following equation:

$$m_1 \oplus m_2(A) = \frac{1}{1-K} \sum_{B_1 \cap B_2 = A} m_1(B_1) \cdot m_2(B_2), \tag{5}$$

where

$$K = \sum_{B_1 \cap B_2 = \varnothing} m_1(B_1) \cdot m_2(B_2). \tag{6}$$

The *K* value reflects the degree of conflict among the sources and so it is called the conflict coefficient. The range of values of the conflict coefficient lies within the interval [0, 1], where zero value indicates the absence of contradictory assessments of the sources. The more contradictions we have, the closer is the *K* value to 1.

P. Smets suggested a concept of pignistic probability [9], which proved to be very fruitful in decision-making sphere, including classification tasks. Pignistic probability *BetP* of focal subset *B* is expressed in the following way:

$$BetP(B) = \sum_{A \in 2^{\Omega}} m(A) \frac{|B \cap A|}{|A|}; \quad B \in \Omega. \tag{7}$$

The value of pignistic probability always lies within the confidence interval: $Bel(B) \leq BetP(B) \leq Pl(B)$. Like any other probabilities, pignistic probability is normalized to 1.

In this work, when we determine the class of pixel, we will rely on the criteria of maximum of pignistic probability:

$$D = \arg(\max_{A \in 2^{\Omega}}(BetP(A))), \tag{8}$$

where *D* is the decision taken on the set of the considered classes.

IV. CLASSIFICATION FRAMEWORK

Since, as mentioned in the formulation of the problem, information about the representatives of the classes under consideration is available to the expert, that's why we select the supervised classification method [1]. In this case, we propose the method for the HSI classification which is represented by the scheme in Fig. 1.

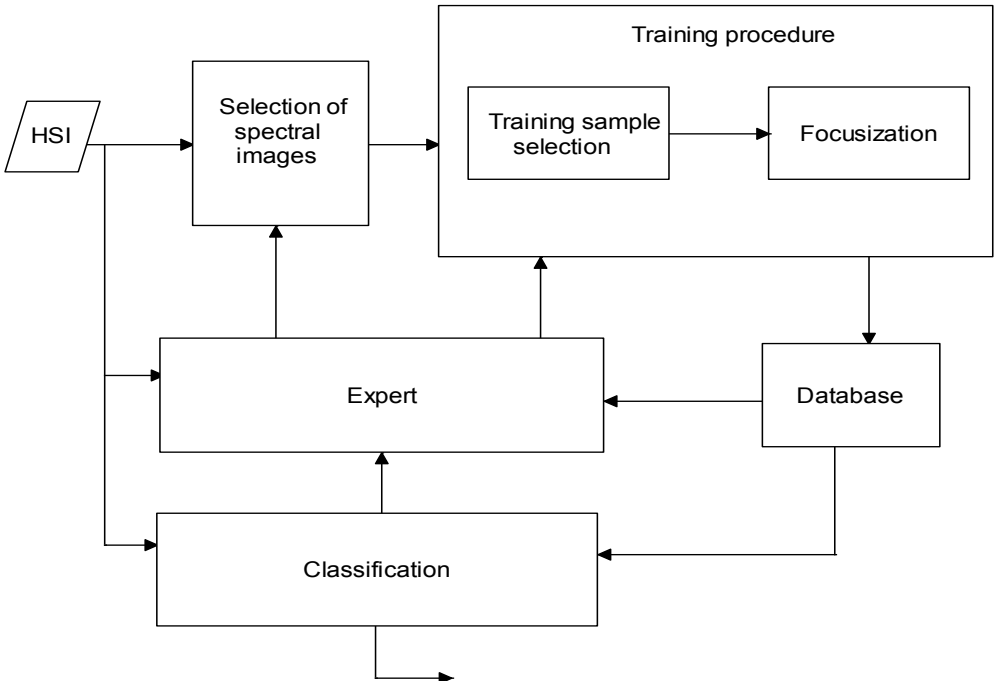


Fig. 1 Scheme of HSI supervised classification method

1) Selection of spectral images

Usually, the part of the spectral images of HSI has high level of noise. Moreover some spectral images of HSI are very similar to each other. Both very noisy and similar images have insufficient informativity, and therefore can be removed from consideration in the classification. Spectral images with significant level of noise can be determined and removed by an expert. The similarity of spectral images can be estimated using correlation coefficient.

Consider the basic procedures that conjointly implement the proposed method. Applying procedure of the selection of spectral images (the pseudo-code of algorithm for the selection of the spectral images is represented in Fig. 2) we are setting up the sub-list of spectral images $\{S_k\}$, the length of K^* which is substantially the shorter than the length K of the original list of spectral images $(k=1,2, ..., K^*; K^*<<K)$. The sub-list $\{S_k\}$ is the shorter, however it still retains the amount of information needed to deal with the task of classification successfully.

Upon completion of the spectral images selection procedure, we get a possibility to reduce the dimension of the spectral space R from K to K^* $(K^* << K)$ and to make HSI classification in the spectral space $\Lambda^{K^*} = (\lambda_1, ..., \lambda_k, ..., \lambda_{K^*})$. Therefore, instead of expression (3) the pixel signal u_n will be described below by the following expression:

$$\mathbf{u_n} = \{u_{nk}\}_{k=1}^{K^*}, \quad \mathbf{u_n} \in \Lambda^{K^*}. \tag{9}$$

It should be noted that the reduction number of correlated spectral images from K to K^* not only reduces the volume of further computations, but also satisfies the requirement of the mutual independence of the data sources that are combined by the rule of Dempster [5].

```
N: number of spectral images S_i
r_max : threshold
K_min: minimum number of spectral bands

i = i + 2: step size

For i = 1 to N
S is a vector of images S_i
Sort S
begin
write ("For spectral image S_i we calculate: a_i, cov_i, δ_i");
readln (a_i, cov_i, δ_i);
write ("For spectral image S_{i+1} we calculate: a_{i+1}, cov_{i+1}, δ_{i+1}");
readln (a_{i+1}, cov_{i+1}, δ_{i+1});
write ("For spectral images S_i and S_{i+1} we calculate: r_{i,i+1}");
readln (r_{i,i+1});
if r_{i,i+1} ≤ r_thr then we choose images S_i and S_{i+1}
else
begin
if δ_i > δ_{i+1} then we choose image S_i;
else: we choose image S_{i+1}
end
write ("we get a new number of spectral bands, K*");
end.
begin
if K* < K_min then we correct r_thr and / or K_min
else
begin
if K* > 2·K_min then write ("For the sub-list of spectral images with
length K* we calculate: r_{i,i+1}")
else: we use these K* spectral images for the following
classification
end
end.
```

Fig. 2 The pseudo-code of algorithm for the selection of the spectral images

2) Training procedure

A training data set is formed by the expert from pixels of the HSI. Suppose the HSI displays objects of *L* classes. Expert analyzes the image and selects compact groups of pixels to obtain the training subset for each of the *L* classes. Thus, we get a training sample *TS* that includes *L* training subsamples *ts*, i.e. $TS = \{ts_l\}_{l=1}^{L}$.

The training sample must satisfy a number of requirements [10], [11]. The key requirements include completeness, sufficiency and purity.

The completeness means that the classes must be presented in the training sample in approximately equal proportions.

The sufficiency supposes the presence of the minimum size of training sample $N_{TS}^{\min}$ which ensures the correctness of the classification accuracy estimation. Let's denote a size of training sample *TS* as N_{TS} , then sufficiency requirement is expressed by the inequality

$$N_{TS} \geq N_{TS}^{\min}. \tag{10}$$

There are some ways to calculate $N_{TS}^{\min}$ value, the simplest way is to use binomial model [10], [11], according to which the minimum size of training sample *TS* is calculated as:

$$N_{TS}^{\min} = \frac{Z_{\alpha}^2 \, P_0 \, (1 - P_0)}{b^2}, \tag{11}$$

where P_0 is a prior estimate of correct classification probability; Z_{α} is the normal score for the desired two-tailed probability α of Type I error; *b* is the desired absolute precision as a proportion.

The purity of a training sample is defined by the proportion of pixels of extraneous classes in the total number of pixels of the given sample. That may be so-called mixed pixels, the failed pixels and so on.

The set of pixel signals of training sample allows to represent each class by a set of intervals in the spectral space Λ^{K^*} . To do this, each of the axes $(\lambda_1, ..., \lambda_k, ..., \lambda_{K^*})$ of the spectral space Λ^{K^*} is divided into intervals according to the number *L* of the classes and each of the intervals gets a mark of corresponding class. The order of the interval formation for the class, for example for the class *l*, we'll describe using Fig. 3.

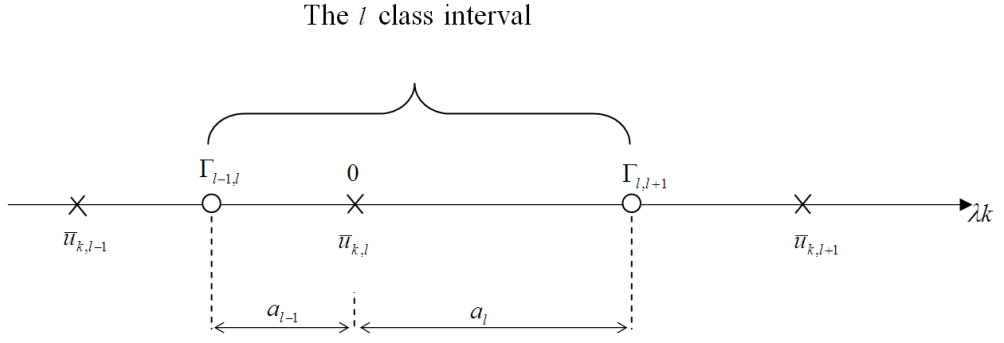


Fig. 3 Forming spectral intervals for classes

The point for the interval forming on the spectral axis λk is point *O*, its position is determined by the expectation signal $\mu_{k,l}$ of pixels belonging (according to the expert opinion) to the *l* class. The left and right boundaries of the interval are limited by points $\Gamma_{l-1,l}$ and $\Gamma_{l,l+1}$. Positions of these points are determined by segments a_{l-1} and a_l respectively.

Let $\sigma_{k,l-1}; \sigma_{k,l}; \sigma_{k,l+1}$ be standard deviations of the signals of pixels of (l-1), (l), and (l+1) classes respectively. Under such notations values a_{l-1} and a_l are calculated from the relationships:

$$\frac{a_l}{a_{l+1}} = \frac{\sigma_{k,l}}{\sigma_{k,l+1}} \text{ and } \frac{a_l}{a_{l-1}} = \frac{\sigma_{k,l}}{\sigma_{k,l-1}}. \quad (12)$$

In the same way, other intervals can be constructed for each of the classes represented in the HSI.

Upon completion the construction of the spectral intervals for each of the considered classes we conduct focusization procedure. By focusization of the class interval we mean the procedure of obtaining the list of focal pixel subsets, whose signals are located in this interval and the calculation procedure of the basic probabilities for them.

Having contaminated training sample we have any reason to assume that interval with certain class mark may contain not only the signals of the pixels of the same class, but the signals of the pixels of other classes as well. Therefore, forming focal subsets we need to consider a number of hypotheses about the class membership of the pixels.

One focal subset is built from a single hypothesis that the pixel whose signal is located within the interval, really belongs to the same class as the interval. Each of the other focal subsets comprises two hypotheses; one hypothesis states that the class membership pixel corresponds to a given interval, and another hypothesis states that the pixel belongs to another specific class.

Each of the focal subsets has corresponding basic probability. To explain the rule for basic probability calculation, consider the following example.

Suppose, the signals of the Q_Σ pixels are located in the spectral interval of the l_1 class. The class membership of pixel has been evaluated by the expert as follows: Q_1 pixels belong to the class with the same name (i. e. (l_1)), Q_2 pixels belong to the l_2 class, and Q_3 pixels belong to the l_3 class.

In this case, the list of focal subsets for the spectral interval will include three subsets: $\{l_1\}$, $\{l_1, l_2\}$, $\{l_1, l_3\}$. Basic probabilities (masses) for these focal subsets are calculated as:

$$\left.\begin{aligned} m(\{l_1\}) &= \frac{Q_1}{Q_\Sigma} \\ m(\{l_1, l_2\}) &= \frac{Q_2}{Q_\Sigma} \\ m(\{l_1, l_3\}) &= \frac{Q_3}{Q_\Sigma} \end{aligned}\right\}, \quad (13)$$

where $Q_\Sigma = Q_1 + Q_2 + Q_3$.

Thus, the focusization procedure is completed and we have formed a database (DB), comprising:

- description (including boundaries and class marks) for each of the formed intervals within the spectral space Λ^{K^*};
- the focal subsets and its basic probabilities (masses) for each of the formed intervals within the space Λ^{K^*}.

Thus, the training procedure is completed and further classification procedure can be run.

3) Classification algorithm

Each pixel π_n of HSI displays an object of some class and the aim of pixel-wise classification is to determine the class of the pixel's object π_n as accurately as possible, based on the analysis of the $\mathbf{u_n}$ signal.

In pixel-wise approach, the pixels of HSI are analyzed and classified consistently and independently. Therefore, it is sufficient to consider the content of the classification procedure for the case of only one pixel, say of an arbitrary pixel π_n.

The algorithm that performs classification procedure consists of the following steps:

Step 1. Extract the known signal $\mathbf{u}_n = \{ u_{nk} \}_{k=1}^{K^*}$ of the pixel π_n.

Step 2. Knowing components u_{nk} of the vector signal $\mathbf{u_n}$ identify all spectral intervals within which corresponding components are located.

Step 3. Using information from DB compose the list of focal subsets and their basic probabilities (masses) for each of the spectral intervals, which have been identified in step 2.

Step 4. Applying the rule of Dempster (5), perform the calculation of the combined masses for the focal subsets from the list that was composed in step 3.

Step 5. Based on the results of calculations performed in the previous step and using the formula (7) calculate the pignistic probability BetP value for each of focal subsets – singletons.

Step 6. The calculated values of pignistic probability for focal subsets - singletons are ranked and further we define (using the criterion (8) of maximum pignistic probability) the most likely class membership for the pixel π_n.

This algorithm is applied sequentially to each pixel of the hyperspectral image and in the end we get the classified HSI in whole. Thus, the classification process is completed.

V. EXPERIMENT AND RESULTS

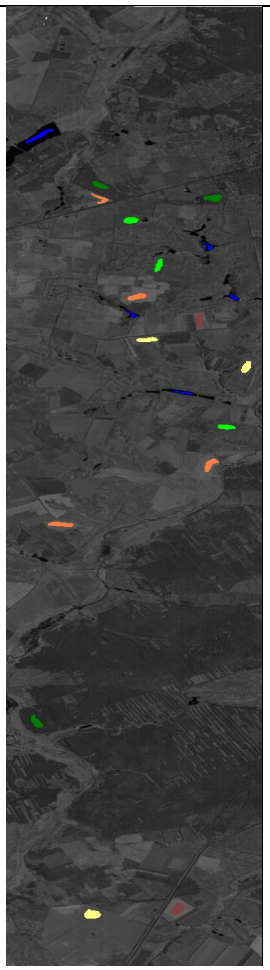
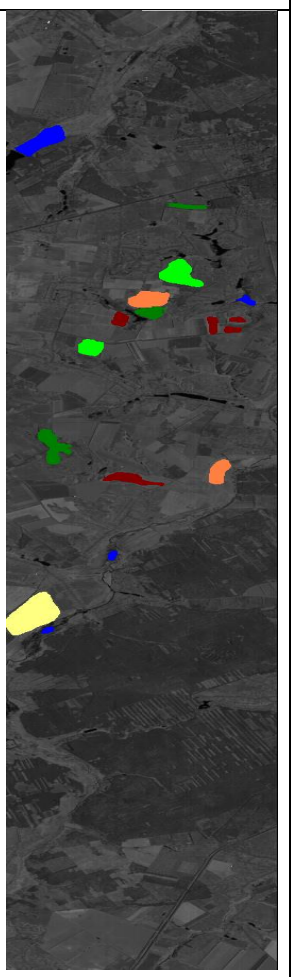
Accuracy assessment of the developed method was conducted on base EO1H1810252014211110KF hyperspectral image received 30.07.2014 by Hyperion Imaging Spectrometer of EO-1 satellite. A Hyperion HSI consists of 242 narrow spectral images which together cover the wavelength range from 400 to 2500 nm. All spectral images have the same spatial resolution of 30 meters. All spectral images have same spatial resolution 30 meters.

The Hyperion imager filmed the strip of the terrain at the junction of three districts of the Kiev's region: Makariv's, Kiev-Svyatoshin's and Fastiv's (shown in Table 1 in RGB representation). On the terrain there are objects of the following classes: water (1), coniferous forest (2), deciduous forest (including garden and forest plantations) (3), grain fields (4), vegetable fields (5), and harvested fields (6).

For hyperspectral images proceeding and analysis the special software was developed. With its help very noisy and similar spectral images have been removed. We used the following criterion: the correlation coefficient r_{thr} between neighboring spectral images should not exceed of 0.8. After that for further work 17 spectral images have been selected. The spectral bands in which these images have been obtained are enumerated in Table II; numbering of the spectral bands corresponds to the EO-1 technical documentation [12].

The spatial location of the training subsamples and the number of pixels in each of them are shown in Table I (left image). To assess the accuracy of the proposed method, examining subsamples (ground sites) were selected which have no intersections with the training subsamples (see right image in Table I). The sizes of training sample and examining sample have been calculated given the sufficiency requirement (the formula (11) and meeting the condition (10)).

TABLE I
CLASSES, TRAINING AND EXAMINING SUBSAMPLES

Classes of objects	Training subsamples		Examining subsamples	
	Number of pixels	Location in the Hyperion image	Number of pixels	Location in the Hyperion image
1. Water	335		1394	
2. Coniferous forest	397		1471	
3. Deciduous forest	344		1541	
4. Grain fields	525		1349	
5. Vegetable fields	440		2059	
6. Harvested fields	315		1347	
Total pixels	2356		9161	

In experimental research, we have tested the proposed method (below in tables and charts, this method is denoted as DST) in comparison with the SVM method [1]. We have selected the SVM method, because most researchers agree with the opinion that namely this method provides the most accurate classification [13], [14]. Moreover, SVM method is conveniently implemented in software in the ENVI image processing system [15]. In the experiment, we used 8 versions of the same initial training sample. The versions differ in the degree of contamination, specifically: 0, 10, 20, 30, 35, 40, 45, and 50%. Sequentially for each version, the Hyperion hyperspectral image was independently handled using both DST and SVM method. Each time the error matrix was formed, that allows evaluation of the classification accuracy.

The results of the experiment have demonstrated that when we use non-contaminated training sample our method shows overall accuracy 0.859 whereas the overall accuracy of the SVM was found to be 0.778 (Fig. 4).

The classification accuracy decreases monotonously with the increase of contamination degree of training sample. In the case of 50%-contaminated training sample we obtained the following results: the overall accuracy for the DST and SVM methods equals 0.629 and 0.475

respectively. Therefore, the accuracy of the DST method is better by 8-16% compared to the SVM method accuracy.

TABLE II
Selected spectral bands of Hyperion Imaging Spectrometer

Spectral Band	Average Wavelength, nm	Full Width at Half the Maximum, nm
21	559.09	10.93
33	681.20	10.33
36	711.72	10.53
43	782.95	10.88
77	912.45	11.05
88	1023.40	11.05
95	1094.09	10.99
111	1255.46	10.73
136	1507.73	11.11
140	1548.02	11.25
153	1679.20	11.55
183	1981.86	10.92
186	2012.15	10.91
198	2133.24	10.73
202	2173.53	10.63
217	2324.91	10.41
222	2375.30	10.41

In Fig. 5 we can see the distribution of the errors of omission for the considered classes in the case the non-contaminated training sample. The numbering of classes corresponds with the numbering in Table I.

The average omission and commission errors of Hyperion HSI classification by DST and SVM methods for the non-contaminated training sample case are shown in Table III. Their comparison shows, that the DST classification method is on 9% more efficient than the SVM method.

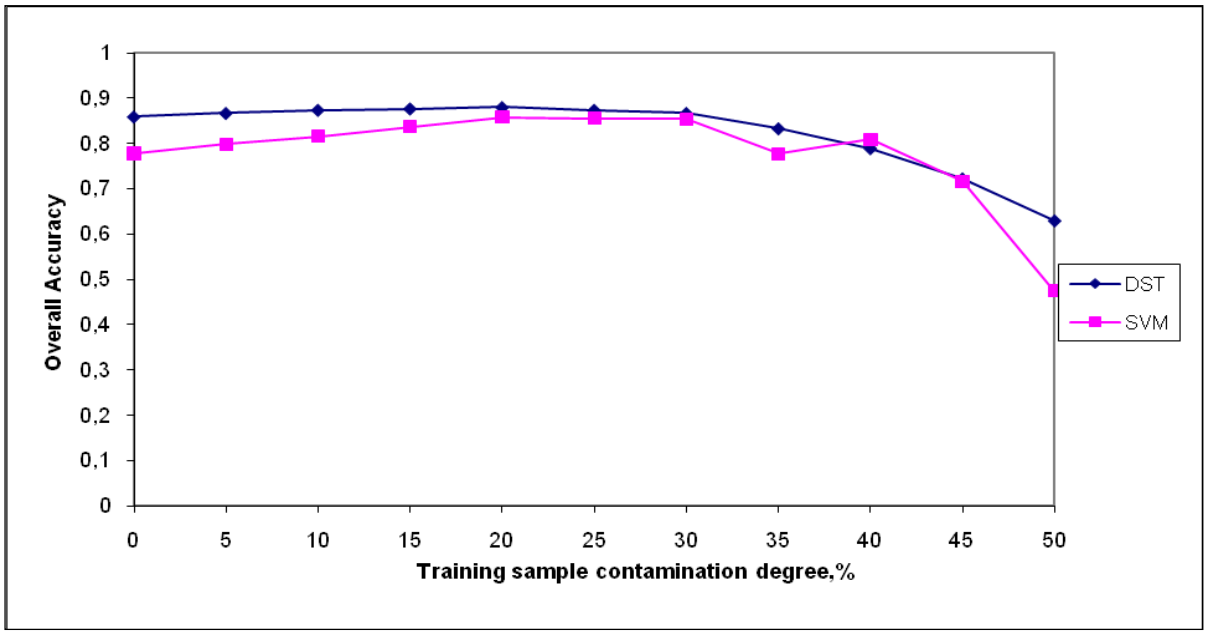


Fig. 4 Accuracy of DST and SVM methods as a function of the training sample contamination degree

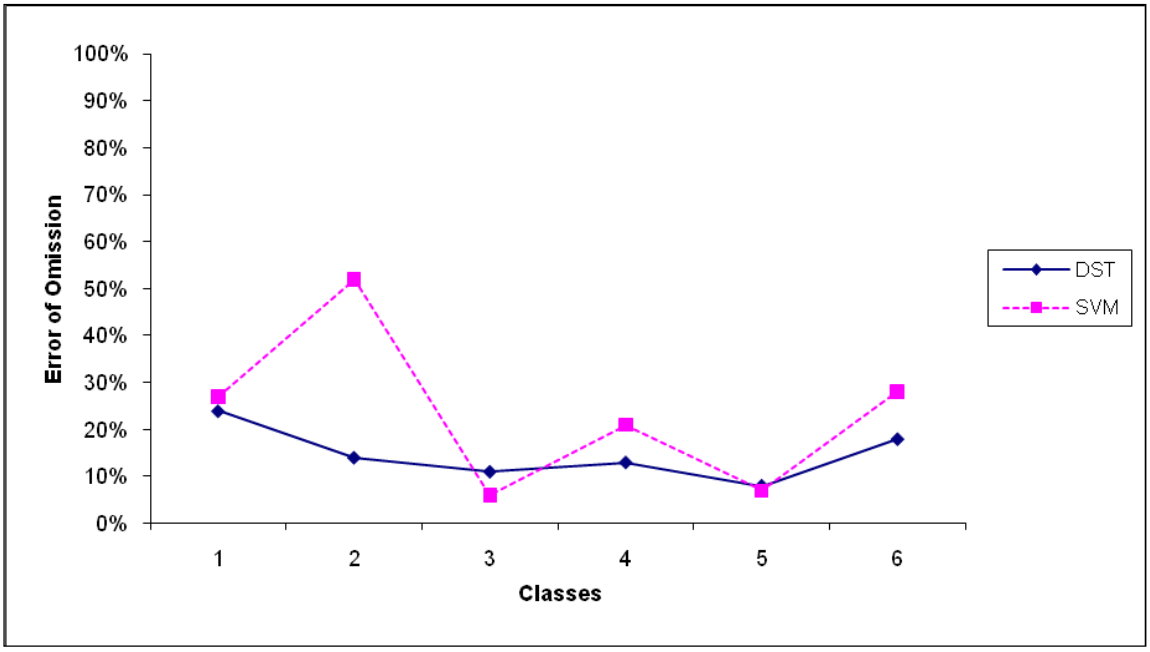


Fig. 5 Error of Omission distribution for the considered classes (non-contaminated training sample case)

TABLE III

Average Omission and Commission Errors

Method	Average Omission Error	Average Commission Error
DST	15%	13%
SVM	24%	22%

VI. CONCLUSION

Thus, in this paper we have proposed a new method of supervised classification of hyperspectral images especially for prevalent cases when the training sample is contaminated, including significant contamination. The method is based on the use of Dempster-Shafer paradigm.

The flowchart of the classification process is described. The algorithmic and software support are developed for the evidence theory-based method. The results of experiments with real hyperspectral image showed sufficiently high accuracy of the proposed method.

Further development of the proposed method the authors associate, primarily, with the improvement of the procedure of classes' description formulation procedure. In particular, it is advisable to provide the optimization of the class intervals on spectral axes, and explore other possible approaches to the focusization procedure, etc. The actual task is also the development of recommendations to use the proposed classification method in various thematic applications.

REFERENCES

[1] B. Tso, Classification Methods for Remotely Sensed Data, London: Tailor & Francis, 2001, 332 p.
[2] C. Ruwet, G. Haesbroeck, "Impact of contamination on training and test error rates in statistical clustering", Taylor & Francis: STM, Behavioural Science and Public Health Titles, 2011, Vol. 40 (3), pp. 394-411.
[3] D. Yan, P. Gong, A. Chen, L. Zhong, "Classification under Data Contamination with Application to Remote Sensing Image Mis-registration", Journal: Computing Research Repository – CORR, 2011, Vol. abs/1101.3.
[4] C. Persello, L. Bruzzone, "Active and Semisupervised Learning for the Classification of Remote Sensing Images", IEEE Transactions on Geoscience and Remote Sensing, 2014, Vol. 52, No. 11, pp. 6937-6956.
[5] G. Shafer, A Mathematical Theory of Evidence, Princeton: Princeton University Press, 1976, 297 p.
[6] B.A. Momani, S. Mcclean, P. Morrow, "Using Dempster–Shafer to incorporate knowledge into satellite image classification", Artif. Intel. Review, 2006, Vol. 25, pp. 161–178.
[7] A. Taroun, J.B. Yang, "Dempster-Shafer Theory of Evidence: Potential usage for decision making and risk analysis in construction project management", The Built & Human Environment Review, 2011, Vol. 4. Special Issue 1, pp. 155-166.
[8] M.A. Popov, M.V. Topolnytskyi, "Multispectral / hyperspectral satellite image classification algorithm based on the Dempster-Shafer evidence theory", Journal of Information, Control and Management Systems, 2014, Vol. 12, No. 2, pp. 165-176.
[9] P. Smets, M. Henrion, R.D. Shachter, L.N. Kanal, J.F. Lemmer, "Constructing the pignistic probability function in a context of uncertainty", Uncertainty in Artificial Intelligence, 1990, Vol. 5, pp. 29–40.
[10] W.G. Cochran, Sampling Techniques. 3rd Ed, NY: John Wiley & Sons, 1977, 428 p.
[11] G. Congalton, K. Green, Assessing the Accuracy of Remotely Sensed Data: Principles and Practices. 2nd Ed., Boca Raton: CRC Press, 2009, 183 p.
[12] https://directory.eoportal.org/web/eoportal/satellite-missions/e/eo-1
[13] C.J.C. Burges, "A tutorial on support vector machines for pattern recognition", Data Mining and Knowledge Discovery, 1998, Vol. 2, No. 2, pp. 121–167.
[14] C. Li, J. Wang, L. Wang, L. Hu, P. Gong, "Comparison of Classification Algorithms and Training Sample Sizes in Urban Land Classification with Landsat Thematic Mapper Imagery", Remote Sensing, 2014, 6, pp. 964-983.
[15] ENVI Ver. 4.7, December 2009 // Visual Information Solutions.