Applications of Dempster-Shafer evidence theory to data processing in remote sensing

Igor Prokopenko[1,†], Sofiia Alpert[1,2,*,†] and Yuliia Petrova[1,†]

[1] *National Aviation University, Liubomyra Huzara Ave., 1, Kyiv, 03058, Ukraine*
[2] *Scientific Centre for Aerospace Research of the Earth of the Institute of Geological Science of the National Academy of Sciences of Ukraine, Olesia Honchara Str., 55-b, Kyiv, 01054, Ukraine*

Abstract

Nowadays various modern techniques and methods of remote sensing allow to identify, distinguish and investigate various objects, their main properties and connections. Remote-sensing techniques always require data processing. One of the most essential data processing procedures is the hyperspectral satellite image classification. It was noted that large volume of information causes a problem with the image classification procedure. Various spectral bands can give different probabilities of the same object belonging to a certain class. A lot number of spectral bands generates a multi-alternative classification problem. This problem demands a multi-alternative solution. Dempster-Shafer evidence theory can be used for solution of this multi-alternative classification problem. This theory can deal with ambiguous, partial, vague and controversial data. It was emphasized, that Dempster-Shafer evidence theory can be applied for image classification. It was considered an example of image classification applying Dempster-Shafer theory in this work. It also was analyzed two examples of applying the Dempster-Shafer evidence theory and Dempster combination rule to an object coordinate determination. Each sensor (radar) gives one coordinate of an object. The value of this coordinate lies within a confidence interval. Then all basic masses, belief functions, plausibility functions for all given intervals and all possible intersections of these intervals and belief intervals were calculated, applied main concepts of Dempster-Shafer evidence theory, accuracy and probability of failure-free operation of sensors. Then it was determined most likely coordinate of an object and its most probable confidence interval or intersection of these confidence intervals. Analyzing these two examples, we considered the relationship between reliability and basic mass.

Keywords

remote sensing, evidence theory, Dempster combination rule, data processing

1. Introduction

Recently, with the rapid development of science and technology, many new remote sensing methods, technologies and materials have appeared. These new remote sensing techniques and approaches are applied for environmental monitoring, agriculture problems, biological and geological tasks, land-cover classification and various ecological problems.

Nowadays a lot of remote sensing satellite image processing methods and techniques are known [1–3].

It should be noted that hyperspectral satellite images are most informative. They contain a large amount of information about the objects, which allows detecting, analyzing and classifying objects, recording changes and providing forecast estimates. This technique can be applied for the identification of objects by analyzing their unique spectral signatures.

It collects and processes information across the electromagnetic spectrum to obtain the spectrum for each pixel in a hyperspectral satellite image.

ADP'24: International Workshop on Algorithms of Data Processing, November 5, 2024, Kyiv, Ukraine
* Corresponding author.
† These authors contributed equally.
igorprok48@gmail.com (I. Prokopenko); sonyasonet87@gmail.com (S. Alpert); panijulia.p@gmail.com (Y. Petrova)
0000-0003-4169-3774 (I. Prokopenko); 0000-0002-7284-6502 (S. Alpert); 0000-0002-3768-7921 (Y. Petrova)

Hyperspectral satellite images provide unique additional information about the characteristics and properties of researched objects. It was emphasized, that hyperspectral images provide important and unique information about the all characteristics of researched objects [4, 5, 6].

The total number of spectral bands in hyperspectral sensors is up to several hundred.

Various remote sensing objects have different emissivity characteristics. In different spectral different objects have various spectral characteristics.

It should be noted that hyperspectral satellite images have some disadvantages.

A large data volume causes a problem with the classification procedure. Different spectral bands can give different estimates (probabilities) of the same object belonging to a certain class. In other words, a large number of spectral bands pose a multi-alternative classification problem, which, in turn, provides a multi-alternative solution to the problem in the form of a certain set of hypotheses.

It should be noted, that Dempster-Shafer evidence theory can be applied for solution of this multi-alternative classification problem.

2. Main concepts of Dempster-Shafer evidence theory

Dempster-Shafer evidence theory is a generalization of traditional probability theory. But unlike probability theory, Dempster-Shafer theory can process incomplete, imprecise and conflicting information. It can process ignorance and missing information.

This theory deal with probabilities of a collection of hypotheses, whereas a classic probability theory deals with only one single hypothesis.

That's why Dempster-Shafer evidence theory is more flexible approach than the probability theory.

Dempster-Shafer evidence theory allows to combine data obtained from different sources (experts) and provides a multi-alternative solution of the problem (in the form of a set of hypotheses) in the presence of inaccurate and contradictory input data.

In other words, using the Dempster combination rule, it is possible to process all expert opinions and obtain an integral (generalized) assessment.

Dempster-Shafer evidence theory or the theory of belief functions, originated as a mathematical approach for modeling uncertainty. This theory was developed in two stages by A. P. Dempster and G. Shafer.

The foundation of evidence theory is Dempster. He dealt with multivalued mappings and statistical inference. Dempster introduced a method for combining evidence from different experts (sources of information) to derive probabilistic conclusions. Then this method was formalized into Dempster's combination rule.

This rule allowed to combine independent sources of data. Unlike traditional probability theory, it can combine conflicting or incomplete evidence from different sources and calculate degrees of belief.

Dempster-Shafer evidence theory does not require the strict assumptions of classic probability theory, particularly when there is insufficient information to assign precise probabilities. G. Shafer is a famous scientist in the fields of probability theory and statistics. He is known for his work on the theory of belief functions and the Dempster-Shafer evidence theory, which is a mathematical background for modeling incompleteness and uncertainty.

Shafer expanded Dempster's work in the scientific work "A Mathematical Theory of Evidence", which provided a more detail analysis of this theory. Shafer reinterpreted Dempster's work in terms of belief functions and introduced the concept of belief and plausibility measures.

Since Shafer's formalization, the Dempster-Shafer theory has gained attention in fields such as artificial intelligence, data fusion, decision-making and expert systems, because this theory can process uncertain data.

This theory can be applied as an alternative to traditional probabilistic models, especially in cases where data are incomplete or imprecise. It also can model imprecision [7, 8, 9].

Shafer's contributions have been influenced expanding the traditional scope of known traditional probability theory. He generalized Bayesian probability by degrees of belief. His work challenges the assumption that uncertainty must always be represented by probabilities and has led to new approaches in fields like expert systems, machine learning and pattern recognition.

Shafer also made important contributions to the decision theory, philosophy of statistics and the foundations of probability. His work intersects with the ideas of game theory and imprecise probability, offering alternative views on how to process uncertain and incomplete information.

Dempster-Shafer theory can process incomplete, uncertain and ambiguous information. Ω is a frame of discernment, so Ω is a set of hypotheses about membership of certain pixel; 2^{Ω} −number of all subsets of Ω. Let's note, that Ω and $\emptyset$ are included in this number too. $m(A)$ is the basic mass (basic probability) that represents the degree of belief allocated to the certain hypothesis A.

Basic mass satisfies the two conditions:

$$\sum_{A\subseteq 2^{\Omega}} m(A) = 1;\ m(\emptyset) = 0. \tag{1}$$

Let's note, that basic mass also satisfies such condition:

$$0 \le m(A) \le 1.$$

Subset A is called the focal subset, if basic mass $m(A) > 0$.

Let's consider main differences between Dempster-Shafer evidence theory and probability theory. The three conditions are met for the probability theory:

1. $P(\Omega) = 1$.
2. If $X \subset Y$, then the condition is necessarily satisfied:
$$P(X) \le P(Y).$$
3. $P(X) + P(\bar{X}) = 1$.
4. $\bar{X}$ is the complement of the set X, so
$$X \cap \bar{X} = \emptyset,$$
$$X \cup \bar{X} = \Omega.$$

The three conditions are met for the Dempster-Shafer evidence theory:

1. $m(\Omega) \ne 1$.
2. If $X \subset Y$, then the condition is not necessarily satisfied:
$$m(X) \le m(Y).$$
3. No relationship is demanded between $m(X)$ and $m(\bar{X})$.
4. $\bar{X}$ − the complement of the set X, so
$$X \cap \bar{X} = \emptyset,$$
$$X \cup \bar{X} = \Omega.$$

The probability theory requires complete knowledge of combined probabilities and the priory knowledge of probability distribution. The main limitation of probability theory cannot model imprecision and measure a body of evidence.

Dempster-Shafer evidence theory is more flexible approach than the probability theory. It is generalization of classical probability theory. It can process incomplete and vague information. Dempster-Shafer theory can deal with probabilities of a collection of hypotheses, whereas a traditional probability theory deals with only one single hypothesis. Dempster-Shafer evidence theory can deal with missing and ignorance data.

If we consider main differences between Dempster-Shafer evidence theory and probability theory, we can conclude, that Dempster-Shafer evidence theory has advantages over probability theory.

Let's note, that belief function $Bel(A)$ and plausibility function $Pl(A)$ shows the level of hypothesis support.

Belief function $Bel(A)$ measures the minimum or necessary support for the hypothesis. It is calculated by summing the basic probabilities over all nonempty subsets $B \leq A$.

The Belief function is defined as follows:

$$Bel(A) = \sum_{B \subseteq A} m(B). \tag{2}$$

Plausibility function $Pl(A)$ reflects the maximum or potential support for that hypothesis. The values of the plausibility function are a set of basic probabilities of all nonempty subsets B intersecting with the considered subset A

$$Pls(A) = \sum_{B \cap A \neq \emptyset} m(B). \tag{3}$$

The plausibility function $Pls(A)$ and belief function $Bel(A)$ are interconnected:

$$Pls(A) = 1 - Bel(\bar{A}), \tag{4}$$

where $\bar{A}$ – the complement of the set A, so

$$A \cap \bar{A} = \emptyset, \;\; A \cup \bar{A} = \Omega.$$

Belief function defines the lower boundary of the interval containing the exact value of the probability of the considered subset A. Plausibility function defines the upper boundary of the interval containing the exact value of the probability of the considered subset A:

$$Bel(A) \leq Prob(A) \leq Pl(A).$$

Let's note, that $[Bel(A),\ Pl(A)]$ is called the belief interval. The length of this belief interval shows the imprecision about the uncertainty value of A.

Main properties of plausibility function $Pl(A)$ and belief function $Bel(A)$:

$$Bel(\Omega) = 1; \tag{5}$$

$$Pls(\Omega) = 1; \tag{6}$$

$$Bel(A) \leq Pls\ (A), \quad \forall A \subseteq \Omega; \tag{7}$$

$$Bel(\bar{A}) = 1 - Pls(A), \qquad \forall A \subseteq \Omega; \tag{8}$$

$$Bel(A) + Bel(\bar{A}) \leq 1, \;\; \forall A \subseteq \Omega. \tag{9}$$

The main advantage of the Dempster-Shafer evidence theory is the presence of a simple rule for combining data from different experts [8, 9]. This Dempster's combination rule is applied for combining data from different experts or other sources.

Suppose that one expert assigned mass m_1 to the class A and another expert independently assigned mass m_2 to the same class.

Then, the combined assessment of the mass of the class A is defined as follows:

$$m(A) = \frac{1}{1-K} \sum_{B_1 \cap B_2 = A} m_1(B_1) \cdot m_2(B_2), \tag{10}$$

$$K = \sum_{B_1 \cap B_2 = \emptyset} m_1(B_1) \cdot m_2(B_2), \tag{11}$$

where K is called conflict coefficient. The value of K reflects the degree of conflict among the sources or experts.

Conflict coefficient also satisfies next condition:

$$0 \leq K \leq 1.$$

The less contradictions we have, the closer is the K value to 0.

3. Example of application of Dempster-Shafer evidence theory to image classification

The process of solution of actual scientific and practical problems, using hyperspectral satellite images as a rule includes a procedure of its classification.

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.

The hyperspectral image consists of a set of spectral images:

$$S_k = \{\pi_n, u_{nk}\}_{n=1}^{N_\pi}; \quad k = 1,2,\dots,K. \tag{12}$$

where S_k is the k-spectral image; K is the total number of spectral images; π_n is the n-th pixel; N_π is the total number of pixels in the hyperspectral image; u_{nk} is the k-component of full signal u_n of the pixel π_n.

Full signal of a pixel is considered as a vector with components u_{nk} in the spectral space:

$$u_n = \{u_{nk}\}_{k=1}^{K}. \tag{13}$$

Each π_n pixel of the hyperspectral image displays an object of some class and the aim of pixel-wise classification, using the Dempster-Shafer evidence theory is to determine the class of the pixel object π_n as accurately as possible, based on the analysis of the u_n signal.

Dempster-Shafer evidence theory can be applied for image classification. The pixels of satellite image are classified independently.

Let's consider an example of hyperspectral satellite image classification applying Dempster-Shafer evidence theory and Dempster's combination rule. The set of pixel signals of training sample represents each class by a set of intervals in the spectral space. Each of the axes of this spectral space is divided into intervals according to the number S of the classes. Each of these intervals gets a mark of corresponding class.

The position of the interval is set by the average values of the signals of pixels of certain class in the certain spectral band. For example, $\bar{u}_{k,s}$ and $\bar{u}_{k,s+1}$ are average values of the signals of pixels of s-class and $s+1$-class respectively for k-spectral band (see Figure 1).

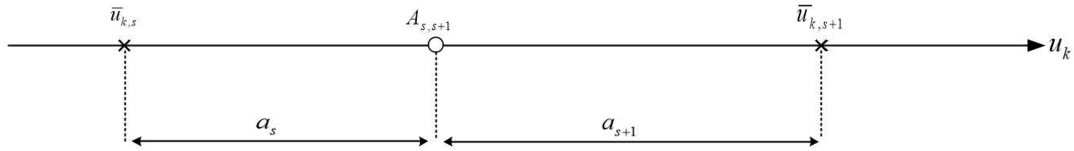


Figure 1: Construction of the spectral intervals for classes.

The point $A_{s,s+1}$ divides s-class and $s+1$-class.

The position of point $A_{s,s+1}$, intervals a_s and a_{s-1} are defined from the next proportion:

$$\frac{a_s}{a_{s+1}} = \frac{\delta_{k,s}}{\delta_{k,s+1}},$$

where $\delta_{k,s}$, $\delta_{k,s+1}$ are variances (standard deviations) of the signals of pixels of s-class and $s+1$-class.

Then we should conduct a focalization procedure. Focalization is 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.

Supposing, that interval with certain class mark contains not only the signals of the pixels of the same class, but the signals of the pixels of other classes as well. That's why, forming focal subsets we consider a number of hypotheses about the class membership of the pixels. Each focal subset is assigned a basic probability.

One focal subset is formed from a single hypothesis that the pixel whose signal is located within the interval, belongs to the same class as the interval. Each of the other focal subsets includes two hypotheses. One hypothesis states that the class membership of the pixel corresponds to a given interval, and another hypothesis states that pixel does not correspond to a given interval. If pixel does not correspond to a given interval, it belongs to another specific class. Other intervals can be formed for each of the classes represented in the hyperspectral satellite image in the same way.

Suppose, the signals of the M pixels are located in the spectral interval of the s_1 −class.

Then expert assesses the class membership of pixels:

1. M_1 pixels are assigned to the s_1 − class.
2. M_2 pixels are assigned to the s_2 − class.
3. M_3 pixels are assigned to the s_3 − class.

The list of focal subsets for the spectral interval includes such subsets:

$$\{s_1\};$$
$$\{s_1,\ s_2\};$$
$$\{s_1,\ s_3\}.$$

Then basic massed for these focal subsets are defined as follows:

1. $m(\{s_1\}) = \frac{M_1}{M}$.
2. $m(\{s_1,\ s_2\}) = \frac{M_2}{M}$.
3. $m(\{s_1,\ s_3\}) = \frac{M_3}{M}$.

Let's note, that $M = M_1 + M_2 + M_3$.

Each pixel of hyperspectral satellite image displays an object of some class. Main purpose of classification procedure is to determine the class of the pixel's object, based on the analysis of the pixel signal.

Let's note, that pixels of satellite image are analyzed and classified independently. Therefore, it should consider the classification procedure for only one arbitrary pixel.

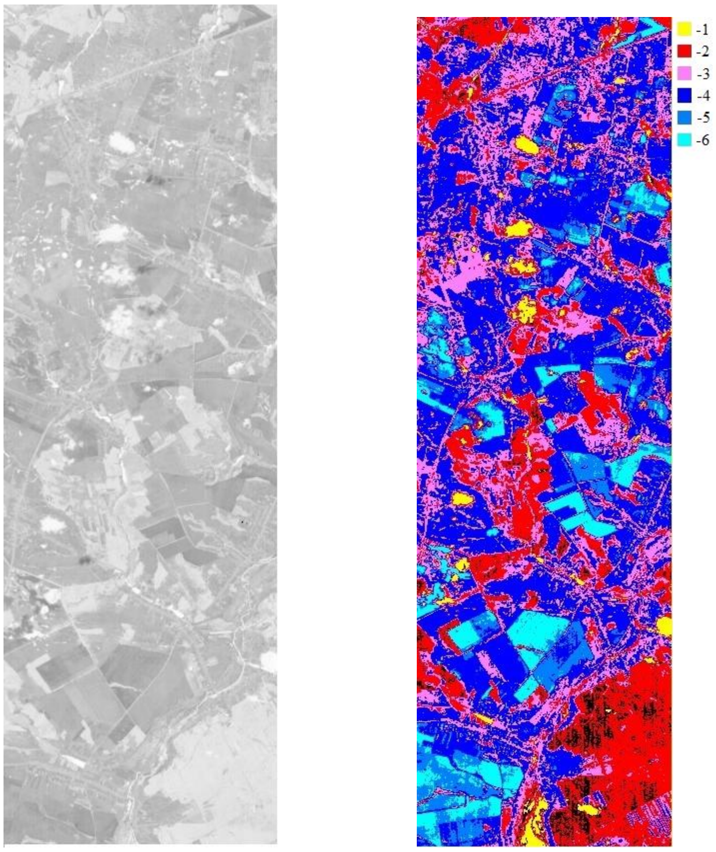
The classification procedure involves such steps:

1. The known signal $u_n = \{u_{nk}\}_{k=1}^{K}$ of the pixel π_n is retrieved.
2. Analyzing components u_{nk} of the vector signal u_n we should form spectral intervals within which corresponding components are located.
3. Focal subsets and their basic masses for each of the spectral intervals are composed.
4. The calculation of the combined basic masses for the focal subsets was conducted, applying Dempster's combination rule.
5. The calculated values of combined basic masses for all focal subsets are ranked. Then the most likely class membership for the pixel is defined, applying the criterion of maximum basic probability.

This classification procedure can be applied sequentially to each pixel of the hyperspectral satellite image and in the end of this procedure we get the classified hyperspectral satellite image in whole.

The results of classification of the hyperspectral image EO1H1810252013112110KF of the Kyiv region in April 2013, obtained by the EO-1 satellite system (see Figure 2, a) by the Dempster-Shafer method are shown in Figure 2, b [9, 4].

The six classes of objects were identified: 1 – deciduous forest, 2 – orchards, 3 – uncultivated land, 4 – meadows, 5 – cereal fields, 6 – vegetable fields.



a) b)

Figure 2: Hyperspectral image Hyperion EO1H1810252013112110KF (April 2013) (a) and classification result (b).

4. Examples of applying the Dempster-Shafer evidence theory to an object coordinate determination

4.1. First example

Let's consider an example of applying the Dempster's combination rule to determine the coordinates of an object (target). Most likely coordinate of an object locates in the interval or intersection of intervals with maximum value of basic mass, belief function and plausibility [10, 11, 12].

Suppose we have 3 sensors (radars):

1. Sensor 1 with confidence interval A.
2. Sensor 2 with confidence interval B.
3. Sensor 3 with confidence interval C.

In this case, the basic mass, confidence interval, accuracy of each sensor (radar) and reliability are determined by the passport data. Each radar gives one coordinate of an object. The value of this coordinate locates within a confidence interval (see Figure 3).

Let's note, that different sensors have different accuracy, basic masses, confidence intervals and reliability. We find the most likely confidence interval where the object's coordinate locates, taking into account the accuracy and reliability of the sensors. Reliability is the probability of failure-free operation. Accuracy of the sensor is determined by the length of the confidence interval. It should be noted that the smaller the confidence interval in which the object coordinate value is located, the higher the value of the basic mass and accuracy of this sensor.

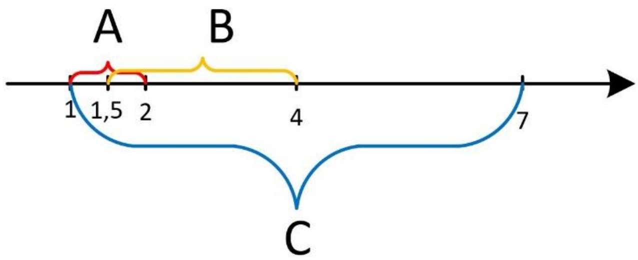


Figure 3: Confidence intervals A, B, C and their intersections.

Initial conditions of the problem are as follows:

1. I sensor: confidence interval $A \equiv (1,2)$; $p_1 = 0.5 -$ probability of failure-free operation;
2. II sensor: confidence interval $B \equiv (1.5,4)$; $p_2 = 1 -$ probability of failure-free operation;
3. III sensor: confidence interval $C \equiv (1,7)$; $p_3 = 0.6 -$ probability of failure-free operation.

Then basic masses are defined as the ratio of the probability of failure-free operation (reliability) and accuracy:

$$m_1(\{A\}) = \frac{p_1}{|A|} = \frac{0.5}{1} = 0.5;$$
$$m_2(\{B\}) = \frac{p_2}{|B|} = \frac{1}{2.5} = 0.4;$$
$$m_3(\{C\}) = \frac{p_3}{|C|} = \frac{0.6}{6} = 0.1.$$

Then values of the basic masses are normalized:

$$m_{1n}(\{A\}) = \frac{0.5}{|0.5 + 0.4 + 0.1|} = 0.5;$$
$$m_{2n}(\{B\}) = \frac{0.4}{|0.5 + 0.4 + 0.1|} = 0.4;$$
$$m_{3n}(\{C\}) = \frac{0,1}{|0.5 + 0.4 + 0.1|} = 0.1.$$

Then we should calculate belief functions and plausibility functions for intervals A, B, C and all possible intersections of these intervals.

1. Belief functions for intervals A, B, C are defined as follows:
$$Bel(\{A\}) = m_{1n}(\{A\}) = 0.5;$$
$$Bel(\{B\}) = m_{2n}(\{B\}) = 0.4;$$
$$Bel(\{C\}) = m_{3n}(\{C\}) = 0.1.$$
2. Plausibility functions for intervals *A*, *B*, *C* are defined as follows:
$$Pls(\{A\}) = m_{1n}(\{A\}) + m_{3n}(\{C\}) = 0.5 + 0.1 = 0.6;$$
$$Pls(\{B\}) = m_{2n}(\{B\}) + m_{3n}(\{C\}) = 0.4 + 0.1 = 0.5;$$
$$Pls(\{C\}) = m_{3n}(\{C\}) = 0.1.$$
3. Basic mass for intersection $A \cap B$ is defined as follows:
$$m_{12}(\{A \cap B\}) = m_{1n}(\{A\}) \cdot m_{2n}(\{B\}) = 0.5 \cdot 0.4 = 0.2.$$
In this case conflict coefficient $K = 0$, because all intersections of sets are not empty. Basic mass for intersection $A \cap B \cap C$ is defined as follows:
$$m_{123}(\{A \cap B \cap \mathrm{C}\}) = m_{12}(\{A \cap B\}) \cdot m_{3n}(\{C\}) = 0.2 \cdot 0.1 = 0.02.$$
In this case conflict coefficient $\widetilde{K} = 0$ too, because all intersections of sets are not empty. Belief functions and plausibility functions for intersection $A \cap B \cap C$ are defined as follows:
$$Bel(\{A \cap B \cap \mathrm{C}\}) = m_{123}(\{A \cap B \cap \mathrm{C}\}) = 0.02;$$
$$Pls(\{A \cap B \cap \mathrm{C}\}) = m_{1n}(\{A\}) + m_{2n}(\{B\}) + m_{3n}(\{C\}) = 0.5 + 0.4 + 0.1 = 1.$$
4. Basic mass for intersection $A \cap C$ is defined as follows:

$$m(\{A \cap C\}) = m_{1n}(\{A\}) = 0.5.$$

Belief functions and plausibility functions for intersection $A \cap C$ are defined as follows:

$$Bel(\{A \cap C\}) = Bel(\{A\}) = m_{1n}(\{A\}) = 0.5;$$

$$Pls(\{A \cap C\}) = Pls(\{A\}) = m_{1n}(\{A\}) + m_{3n}(\{C\}) = 0.5 + 0.1 = 0.6.$$

5. Basic mass for intersection $B \cap C$ is defined as follows:

$$m(\{B \cap C\}) = m_{2n}(\{B\}) = 0.4.$$

Belief functions and plausibility functions for intersection $A \cap C$ are defined as follows:

$$Bel(\{B \cap C\}) = Bel(\{B\}) = m_{2n}(\{B\}) = 0.4;$$

$$Pls(\{B \cap C\}) = Pls(\{B\}) = m_{2n}(\{B\}) + m_{3n}(\{C\}) = 0.4 + 0.1 = 0.5.$$

6. Belief intervals for intervals *A*, *B*, *C* and all possible intersections of these intervals are defined as follows:
   Belief interval for *A:* $[Bel(\{A\}), Pls(\{A\})] \equiv [0.5; 0.6]$;
   Belief interval for *B:* $[Bel(\{B\}), Pls(\{B\})] \equiv [0.4; 0.5]$;
   Belief interval for *C:* $[Bel(\{C\}), Pls(\{C\})] \equiv [0.1; 0.1] \equiv \{0,1\}$;
   Belief interval for $A \cap B \cap$ C: $[Bel(\{A \cap B \cap \text{C}\}), Pls(\{A \cap B \cap \text{C}\})] \equiv [0.02; 1]$; $m(\{A \cap B \cap \text{C}\}) = 0.02$.
   Belief interval for $A \cap$ C: $[Bel(\{A \cap \text{C}\}), Pls(\{A \cap \text{C}\})] \equiv [0.5; 0.6]$; $m(\{A \cap \text{C}\}) = 0.5$.
   Belief interval for $B \cap$ C: $[Bel(\{B \cap \text{C}\}), Pls(\{B \cap \text{C}\})] \equiv [0.4; 0.5]$; $m(\{B \cap \text{C}\}) = 0.4$.

So, interval A and intersection $A \cap C$ are assigned maximum values of basic masses (basic probabilities), belief functions and plausibility functions.

Then we can make a conclusion, most likely coordinate of an object lies within a intersection of confidence intervals A and C: $A \cap C$.

4.2. Second example

Let's consider another example of applying the Dempster's combination rule to determine the coordinates of an object (target).

Initial confidence intervals A, B, C will be same as initial confidence intervals in I example (see Figure 3).

But reliabilities (probabilities of failure-free operations) of these 3 sensors will be another.

1. I sensor: confidence interval $A \equiv (1, 2)$; $p_1 = 0.2 -$ probability of failure-free operation.
2. II sensor: confidence interval $B \equiv (1.5, 4)$; $p_2 = 1 -$ probability of failure-free operation.
3. III sensor: confidence interval $C \equiv (1, 7)$; $p_3 = 0.6 -$ probability of failure-free operation.

Then basic masses are defined as the ratio of the probability of failure-free operation (reliability) and accuracy:

$$m_1(\{A\}) = \frac{p_1}{|A|} = \frac{0.2}{1} = 0.2;$$

$$m_2(\{B\}) = \frac{p_2}{|B|} = \frac{1}{2.5} = 0.4;$$

$$m_3(\{C\}) = \frac{p_3}{|C|} = \frac{0.6}{6} = 0.1.$$

Then values of the basic masses are normalized:

$$m_{1n}(\{A\}) = \frac{0.2}{|0.2 + 0.4 + 0.1|} \approx 0.3;$$

$$m_{2n}(\{B\}) = \frac{0.4}{|0.2 + 0.4 + 0.1|} \approx 0.6;$$

$$m_{3n}(\{C\}) = \frac{0.1}{|0.2 + 0.4 + 0.1|} \approx 0.1.$$

Then we should calculate belief functions and plausibility functions for intervals *A*, *B*, *C* and all possible intersections of these intervals.

1. Belief functions for intervals A, B, C are defined as follows:
$$Bel(\{A\}) = m_{1n}(\{A\}) \approx 0.3;$$
$$Bel(\{B\}) = m_{2n}(\{B\}) \approx 0.6;$$
$$Bel(\{C\}) = m_{3n}(\{C\}) \approx 0.1.$$
2. Plausibility functions for intervals A, B, C are defined as follows:
$$Pls(\{A\}) = m_{1n}(\{A\}) + m_{3n}(\{C\}) = 0.3 + 0.1 = 0.4;$$
$$Pls(\{B\}) = m_{2n}(\{B\}) + m_{3n}(\{C\}) = 0.6 + 0.1 = 0.7;$$
$$Pls(\{C\}) = m_{3n}(\{C\}) = 0.1.$$
3. Basic mass for intersection $A \cap B$ is defined as follows:
$$m_{12}(\{A \cap B\}) = m_{1n}(\{A\}) \cdot m_{2n}(\{B\}) = 0.3 \cdot 0.6 = 0.18 \approx 0.2.$$
In this case conflict coefficient $K = 0$, because all intersections of sets are not empty.
Basic mass for intersection $A \cap B \cap C$ is defined as follows:
$$m_{123}(\{A \cap B \cap \mathrm{C}\}) = m_{12}(\{A \cap B\}) \cdot m_{3n}(\{C\}) = 0{,}2 \cdot 0{,}1 = 0{,}02.$$
In this case conflict coefficient $\widetilde{K} = 0$ too, because all intersections of sets are not empty.
Belief functions and plausibility functions for intersection $A \cap B \cap C$ are defined as follows:
$$Bel(\{A \cap B \cap \mathrm{C}\}) = m_{123}(\{A \cap B \cap \mathrm{C}\}) = 0.02;$$
$$Pls(\{A \cap B \cap \mathrm{C}\}) = m_{1n}(\{A\}) + m_{2n}(\{B\}) + m_{3n}(\{C\}) = 0.3 + 0.6 + 0.1 = 1.$$
4. Basic mass for intersection $A \cap C$ is defined as follows:
$$m(\{A \cap C\}) = m_{1n}(\{A\}) = 0.3.$$
Belief functions and plausibility functions for intersection $A \cap C$ are defined as follows:
$$Bel(\{A \cap C\}) = Bel(\{A\}) = m_{1n}(\{A\}) = 0.3;$$
$$Pls(\{A \cap C\}) = Pls(\{A\}) = m_{1n}(\{A\}) + m_{3n}(\{C\}) = 0.3 + 0.1 = 0.4.$$
5. Basic mass for intersection $B \cap C$ is defined as follows:
$$m(\{B \cap C\}) = m_{2n}(\{B\}) = 0.6.$$
Belief functions and plausibility functions for intersection $A \cap C$ are defined as follows:
$$Bel(\{B \cap C\}) = Bel(\{B\}) = m_{2n}(\{B\}) = 0.6;$$
$$Pls(\{B \cap C\}) = Pls(\{B\}) = m_{2n}(\{B\}) + m_{3n}(\{C\}) = 0.6 + 0.1 = 0.7.$$
6. Belief intervals for intervals A, B, C and all possible intersections of these intervals are defined as follows:
Belief interval for *A:* $[Bel(\{A\}), Pls(\{A\})] \equiv [0.3; 0.4]$;
Belief interval for *B:* $[Bel(\{B\}), Pls(\{B\})] \equiv [0.6; 0.7]$;
Belief interval for *C:* $[Bel(\{C\}), Pls(\{C\})] \equiv [0.1; 0.1] \equiv \{0{,}1\}$;
Belief interval for $A \cap B \cap \mathrm{C}$: $[Bel(\{A \cap B \cap \mathrm{C}\}), Pls(\{A \cap B \cap \mathrm{C}\})] \equiv [0.02; 1]$; $m(\{A \cap B \cap \mathrm{C}\}) = 0.02$.
Belief interval for $A \cap \mathrm{C}$: $[Bel(\{A \cap \mathrm{C}\}), Pls(\{A \cap \mathrm{C}\})] \equiv [0.3; 0.4]$; $m(\{A \cap \mathrm{C}\}) = 0.3$.
Belief interval for $B \cap \mathrm{C}$: $[Bel(\{B \cap \mathrm{C}\}), Pls(\{B \cap \mathrm{C}\})] \equiv [0.6; 0.7]$; $m(\{B \cap \mathrm{C}\}) = 0.6$.

So, interval B and intersection $B \cap C$ are assigned maximum values of basic masses (basic probabilities), belief functions and plausibility functions. Then we can make a conclusion, most likely coordinate of an object locates within the intersection of confidence intervals B and C: $B \cap C$, because initial value of reliability (probability of failure-free operation) of II sensor is maximum ($p_2 = 1$), and initial values of reliability (probability of failure-free operation) of I and III sensor are smaller ($p_1 = 0.2;\ p_3 = 0.6$).

Analyzing these two examples, we can consider the relationship between probability of failure-free operation and basic mass. The higher the initial value of probability of failure-free operation, the higher the value of basic mass of the sensor and it's belief interval.

It was noted, that determination of confidence interval or intersection of confidence intervals with maximum value of basic mass depends on sensor specifications.
Changes in sensor specifications affect the choice of confidence interval with maximum value of basic mass and coordinate determination of an object (target).

5. Conclusions

Nowadays Remote sensing is one of the most popular techniques of obtaining information about the properties of objects by data collected from sensors or unmanned aerial vehicles (UAV). Different techniques of remote sensing can distinguish various objects, study these objects and make predictive estimations. Remote sensing provides data about objects based on analysis of electromagnetic radiation emitted or reflected from these researched objects. Remote sensing is used in various fields. It is applied for ecology, hydrology, geology, geophysics, geography, agriculture, oceanography. Modern remote-sensing techniques involves data processing. It was noted, that one of the most important data processing procedures is the image classification. The process of solution of various remote sensing tasks, using hyperspectral satellite images includes a classification procedure. Each pixel of the hyperspectral image displays an object of some class and the aim of pixel-wise classification is to determine the class of this pixel object.

It was emphasized, that hyperspectral images provide important and unique information about the all characteristics of researched objects. But it was noted, that large data volume causes a problem with the classification procedure. Different spectral bands can give different assessments or probabilities of the same object belonging to a certain class, because a lot of spectral bands poses a multi-alternative classification problem, which provides a multi-alternative solution to the problem.

It should be noted, that Dempster-Shafer evidence theory can be applied for solution of this multi-alternative classification problem.

It was considered and analyzed main concepts of Dempster-Shafer evidence theory. It is generalization of classical probability theory and it can overcome some limitations of known probability theory.

This theory can process conflicting, incomplete and ambiguous information. It was proposed to apply the Dempster's combination rule for classification of vague and contradictory data from different experts. Dempster-Shafer theory can deal with missing information, ignorance data and probabilities of a collection of hypotheses.

It was noted, that Dempster-Shafer evidence theory can be applied for image classification.

It was considered an example of hyperspectral satellite image classification applying Dempster-Shafer evidence theory and Dempster's combination rule. It was considered results of classification of the hyperspectral image in this work.

It also was considered two examples of applying the Dempster-Shafer evidence theory to an object coordinate determination, applying data for 3 sensors (radars). Each radar gives one coordinate of an object. The value of this coordinate locates within a confidence interval. It was noted, that various sensors (radars) have different confidence intervals, accuracy and reliability (probability of failure-free operation).

Then all basic masses, belief functions, plausibility functions for intervals *A, B, C,* all possible intersections of these intervals and belief intervals were calculated, applying main concepts of Dempster-Shafer theory, basic masses, accuracy and probability of failure-free operation of sensors. Then it was determined most likely coordinate of an object and its most probable confidence interval or intersection of these confidence intervals.

Analyzing these two considered examples, we can show the relationship between reliability (probability of failure-free operation) and basic mass. The higher the initial value of reliability (probability of failure-free operation), the higher the value of basic mass of the sensor and it's belief interval.

It was noted, that determination of confidence interval or intersection of confidence intervals with maximum value of basic mass depends on sensor specifications.

References

[1] J. Zhu, X. Wang, Y. Song, Evaluating the reliability coefficient of a sensor based on the training data within the framework of evidence theory, IEEE Access 6 (2018) 30592–30601. doi: 10.1109/ACCESS.2018.2816915.

[2] W. Jiang, S. Wang, X. Liu, H. Zheng, B. Wei, Evidence conflict measure based on OWA operator in open world, PLoS ONE 12 (2017). doi: 10.1371/journal.pone.0177828.

[3] D. Li, Y. Deng, A new correlation coefficient based on generalized information quality, Open Access Journal 7 (2019) 175411–175419. doi: 10.1109/ACCESS.2019.2957796.

[4] M. O. Popov, V. M. Podorvan, S. I. Alpert, A method for hyperspectral satellite image classification using Dempster's combination rule, Space science and technology 1 (2015) 25–37.

[5] B. C. Wintle, H. Fraser, B. C. Wills, A. E. Nicholson, F. Fidler, Verbal probabilities: Very likely to be somewhat more confusing than numbers, PLoS ONE 14 (2019). doi: 10.1371/journal.pone.0213522.

[6] S. Alpert, The new approach to applying the Dezert-Smarandache theory in land-cover classification in UAV-based remote sensing, Management of Development of Complex Systems 49 (2022) 33–39. doi: 10.32347/2412-9933.2022.49.33-39.

[7] F. Shi, X. Su, H. Qian, N. Yang, W. Han, Research on the fusion of dependent evidence based on rank correlation coefficient, Sensors 17 (2017) 2362. doi: 10.3390/s17102362.

[8] X. Su, P. Xu, S. Mahadevan, Y. Deng, On consideration of dependence and reliability of evidence in Dempster–Shafer theory, Journal of Information and Computational Science 11 (2014) 4901–4910.

[9] M. Popov, S. Alpert, V. Podorvan, M. Topolnytskyi, S. Mieshkov, Method of hyperspectral satellite image classification under contaminated training samples based on Dempster-Shafer's paradigm, Central European Researchers Journal 1 (2015) 86–97.

[10] 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 6 (2014) 964–983.

[11] C. Persello, L. Bruzzone, Active and semisupervised learning for the classification of remote sensing images, IEEE Transactions on Geoscience and Remote Sensing 52 (2014) 6937–6956.

[12] G. Sun, X. Guan, X. Yi, J. Zhao, Conflict evidence measurement based on the weighted separate union kernel correlation coefficient, Open Access Journal 6 (2018) 30458–30472. doi: 10.1109/ACCESS.2018.2844201.