DER (Dynamic Evidential Reasoning), applied to the classification of hyperspectral images

Cecilia Sanz
LIDI
(Laboratory of Research and Development in Computer Science)
Faculty of Computer Science - University of La Plata

***Abstract* - This paper describes a novel method for classification based on the evidential reasoning theory and the implementation presented by Peddle and Franklin. DER (Dynamic Evidential Reasoning) introduces some variations. It allows the incorporation of new evidence for the classifier in order to increment its accuracy, and it also defines a different decision rule. The inclusion of new evidence is a learning process where the precision of the classifier is analyzed using the Khat indicator. For this process, a set of samples belonging to a given, and *a priori* known, class is needed. This process changes the discriminate functions. The decision rule introduces two stages of decision. One stage is called "reject decision", where the maximum support value is analyzed, and the object to be classified is assigned to the unknown class if this support value is not "enough" (different approaches for the meaning of "enough" were studied). The other stage is called ambiguity decision, and the similarity between the maximum support and other supports of the rest of the classes are analyzed here.**

In this paper, an application of this method is presented, particularly, for the classification of different crops (during the growing season) in the area of Nebraska, using hyperspectral images. Some results are presented.

***Index Terms* – Hyperspectral Analysis, evidential reasoning, crop classification**

I. INTRODUCTION

The automate analysis of remote sensing images is characterized by the growing volumes of data and the integration of different kind of information (e.g. spectral and spatial). Hyperspectral scanners, are instruments that acquire multispectral images in many, very narrow, contiguous spectral bands. The images produced by these scanners typically contain from ten to hundreds of data channels, which enables the construction of an effectively continuous reflectance spectrum for every pixel in the scene [1]. However, all this information and other ancillary data will be useful, only if methods that allow the integration of these multiple data coming from diverse information sources are provided.

This paper presents a novel method, named DER, which allows the combination of pieces of evidences issued from several sources of information. It is based on the evidential reasoning method, and it introduces some variations.

One of its novelty aspects is a learning process, whose main goal is to allow the incorporation of new evidence for the classifier. It was proved that this learning process improve the accuracy of the classifier.

A different decision rule is also proposed by DER, which analyzes in two stages the final support for each class in order to choose the more appropriate class to be assigned to the unknown object. The first stage is called "Reject Stage" and the second one is called "Ambiguity Stage". These stages try to study in more detail the assignment of the unknown object to a class, in cases when there is not "enough" evidence for the class with maximum support, or when the evidence is distributed among the classes, and some of them have a similar support with the maximum one. This classifier has some advantages as compared with the conventional ones.

The method was applied to a crop type classification application, using hyperspectral images from the region of Nebraska, that were acquired by PRA (Photon Research Associates) group using a CASI sensor. The images used in this application were taken on the same date (August 9, 1998) and they show the crops in their growing season. Particularly, the goal of the application is to identify three kinds of crops: corn, soybean and sorghum. One more class (Road class) was added to the previous ones, in order to test the classifier with something not belonging to the crop classes. Nine spectral and four spatial sources of information were selected as input for the classifier were selected. The spectral ones were based on the calculation of NDVI index and band ratios, which give an idea of the absorption and reflection characteristics of the target. The spatial sources were some first order statistics (mean, standard deviation, kurtosis, and skewness, using a 3x3 window) used as texture indexes for the region. The texture images are calculated previously using ENVI software. It was not the objective of this research to find the best sources of information for the application, but to test the behavior of the classifier.

The remainder of this paper is organized as follows. Section II briefly reviews the evidential reasoning method. Section III describes DER classifier. Section IV presents the experiments carried out to evaluate the classifier as long as some results obtained. Finally, conclusions are drawn in Section V.

II. EVIDENTIAL REASONING METHOD

The evidential reasoning approach is based on the Dempster – Shafer theory [2]. Peddle and Franklin proposed and developed a system for evidential reasoning classification [3].

This theory provides a heuristic basis for combining evidence from diverse data sets. It defines a sets of hypothesis which is called the frame of discernment or universe of discourse (Θ). In remote sensing application the frame of discernment can be the set of all possible classes. Associated with each piece of evidence is a numerical magnitude of support and plausibility, corresponding to the amount of evidence (or mass) in favor of a given class and the mass that fails to refute that class, respectively [4]

Support(x): strength of evidence in favor of x

Plaussibility(x): amount of evidence which fail to refute x

where x ε Θ

The remaining task is to combined the evidence assigned from each source to each class, in order to obtain only one measure of support and plausibility for each class. Dempster orthogonal sum is used for this purpose.

Le m, m2 be the mass from source 1 and source 2 over a set of labels A, the orthogonal sum m1 $\oplus$ m2 (which determines m', the mass assigned to a labeling proposition An) is defined as:

$$m'(An) = K^{-1} \sum_{Ai \cap Aj = An} m1(Ai)\, m2(Aj) \qquad \textbf{(1)}$$

$$K = 1 - \sum_{Ai \cap Aj = \phi} m1(Ai)\, m2(Aj) \qquad \textbf{(2)}$$

where k measures the conflict degree between the sources providing m1 and m2. When k=1, total conflict occurs and the sum cannot be applied. When k=0, the sources to be combined are totally agreeing [5].

The technique proposed by Peddle to compute evidence is based on a frequency approach, where two considerations are taken into account: a) the training data contains evidence for the set of classes, b) the frequency of occurrence for a given value in the training set represents the magnitudes of supports for those classes [6].

Finally, it is necessary to apply a final rule of decision to select the class to be assigned to the pixel (in remote sensing application) to be classified. Different approaches have been proposed for this stage (i.e. select the class with maximum support, or the one with maximum sum of support and plausibility, etc.). DER proposed a new rule for this stage.

III. DER CLASSIFIER

The method proposed in this paper, is a modified version of the evidential reasoning approach presented by Peddle and Franklin. It incorporates a learning process stage and it suggests a different decision rule to be applied to determine which class will be the one to assign to the unknown object.

A . Learning Process

The learning stage is a supervised process, where samples representatives of each class in the frame of discernment are needed. The user presents the samples belonging to a known a priori class to the classifier, and after that a classification process is carried out. The results are analyzed using the Khat indicator. More specifically, the value of Khat is compared with a precision parameter (α, where $\alpha \varepsilon$ [0,1]), in order to determine the necessity of incorporating the new evidence provided by the set of samples (in which case the frequency of occurrence of the samples values is modified for the corresponding source and class).

The determination of the value of α is being study, but as a consequence of performed experimentation, is clear that as its value increases there are more possibility of incorporating the evidence. This parameter is considered as a threshold for Khat. A value of 0.90 is recommended, and was the one used for the application presented here.

The learning process is executed until the evidence is not modified for non of the sample sets.

B. Decision Rule

The rule of decision analyze the class to be assigned to the unknown object in two steps. First, the class with maximum support measure (*ms*) is selected as the final one. The main objective of the reject phase is to determine if the evidence is "enough" to assign the object to this class. However, it is difficult to specify the meaning of "enough". One alternative studied here, was to establish a threshold value for the support measure. Good results were obtained for some values given to the threshold, however the selection of the appropriate value depends on the application and is based on the experimentation. Another possibility was studied, and it is the one used by DER. Two restriction of confidence are proposed to maintain *ms* class as the final for the unknown object: a) the support must not be zero and b) it must be the result of the evidence supplied by at least two sources of information.

The last issue obliges the algorithm to keep the amount of sources providing evidence for each class.

When one of these restrictions is verified for the *ms* class, we decide to assign the object being classified to the "Unknown Class". On the other hand, if none of them is true, the "Ambiguity stage" takes place.

In this other phase, the distance among the supports of the *ms* and the rest of the classes is evaluated.

If the values of supports are "near", there is no clear decision as to which class to assign the object to be classified. A new parameter (called *distance parameter*) is introduced here, and specify when two support measures are considered to be near. For the application presented in this paper a value of 0.01 is used.

To decide which class will be the one selected as the final, additional information is needed when the minimum distance encountered is below the distance parameter. The solution proposed at this step is to take into account not only the support measure, but the amount of sources supplying evidence for each class and a weight measure assigned to each pair source – class. This two indicators allows to measure the amount of sources agreeing in provide some support to one class, and how "confidence" is that source for the class. A combination of these three values are performed: Weight Index * (Amount Sources Index+Support Measure). All these values are normalized to obtained a real value between zero and one. Finally, the class with maximum value resulted from the above combination is selected.

On the other hand, *ms* class is selected if the minimum distance value is above the threshold.

IV. Experimental Results

The method proposed was applied to the classification of three different kinds of crops and roads appear in hyperspectral images obtained using CASI sensor. The images were provided by PRA as research material. Four geocorrected (UTM grid zone 14 coordinates) images from the area of Nebraska (USA) were used to carry out this experimentation.

As a first step, the learning process was tested using a set of samples for each class. Table I shows the results obtained for the first training samples set, Khat indicator is below the precision parameter (0.90), as a consequence the samples are incorporated to the system and the evidence is modified. The second test sample set is presented then to the classifier, and a value of 0.92 was obtained for Khat, no evidence is incorporated to the system. Only those one table is included in this paper as an example of the learning process behavior. However 7 sets of training samples were used, and in the second iteration over the sets no changes were made to the evidence (all sets give a Khat value higher than α), therefore the learning process finished.

In the first pass of this stage the average for Khat was 0.88, and in the second pass the average obtained for Khat was 0.97, in conclusion the learning process improve the results for the training samples test. However, it is necessary to evaluate yet the precision of DER for test samples, different from the selected for the learning phase.

DER was evaluated then with test samples sets, and its decision rule was compared with the one that selects the maximum support. The accuracy results of the classification tests carried out are shown in Table II. The column labeled as DER presents the average accuracy obtained with this method after the learning process stage. The second column (MS) shows the average accuracy for the algorithm, but with the decision rule that selects the class with maximum support as the labeled class for the unknown object. Experiments proved the validity of the proposed method which yields better results than the one to was compared. In addition to that, it has the advantages of the original evidential reasoning method, that is, it can handle data set with higher dimensions, allows the integration of different kind of data from diverse sources , and provides an explicit mechanism to handle data inconsistencies, errors, or uncertainty.

TABLE I
SET I – CONFUSION MATRIX
C 1: SOYBEAN CLASS - C 2: CORN CLASS - C 3: SORGHUM CLASS
C 4: ROAD CLASS - C 5: UNKNOWN CLASS

	C 1	C 2	C 3	C 4	C 5	Total
C 1	74	6	5	0	0	85
C 2	1	19	22	0	0	42
C 3	0	7	56	0	0	63
C 4	9	3	0	37	0	49
C 5	0	0	0	0	0	0
Total	84	35	83	37	0	239

Overview Accuracy: 0.77% - **Khat:** 0.69%

TABLE II
CLASSIFICATION ACCURACY (%) ACHIEVED BY RED AND THE EVIDENTIAL REASONING METHOD USING THE RULE OF MAXIMUM SUPPORT

	DER	MS
Soybean	96.62	70.43
Corn	87.30	56.56
Sorghum	91.89	72.10
Road	99.51	90.82
Average	93.83	72.47

V. Discussion and Conclusions

The method DER is a modified evidential reasoning method, which incorporates a learning process phase in order to add new evidence for the classifier. Apart from that, it defines a rule of decision that evaluates in two stages the class with maximum support and reject or accept it as the final class for the unknown object. One of these steps is called “Reject Stage” and the other “Ambiguity Stage”, in the last one if the class is reject, then the amount of sources index and the weight index are taken into account also.

Some encouraging results were presented here for the application described. However, there are some aspects that required further study. In relation with DER method, the distance parameter will be analyzed in more detail. Its computational cost is being studied [7], as a consequence a parallel version is being developed. The application identifies crops in a particular evolution stage, we are interested in extend this to study crops in all their stages of evolution over time.

Acknowledgment

The author wish to thank PRA for providing the material for this research, R. Jordan, A. De Giusti.

References

[1] T. M. Lillesand, R.W. Kiefer, Remote Sensing and Image Interpretation, Third Edition. John Wiley & Sons, 1994.

[2] G. Shafer, A Mathematical Theory of Evidence, Princeton University Press, Princenton, N.J., pp. 297

[3] D.Peddle, ““Mercury⊕: An evidential reasoning image classifier,” Computers & Geosciences, vol. 21, no. 10, pp.1163-1175, 1995

[4] D.Peddle, “An Empirical comparison of evidential reasoning, linear discriminant analysis, and maximum likelihood algorithms for alpine land cover classification,” Canadian J. of Remote Sensing, vol.19, no.1, pp. 31-44, 1993.

[5] L. Mascarilla, C.l Frélicot, “Combining Rejection – Based Pattern Classifiers,” Univerté de La Rochelle, Laboratoire d´Informatique et d´Imagerie Industrielle.

[6] D.Peddle, ““Knowledge Formulation for Supervised Evidential Classification,” Photogrammetric Engineering & Remote Sensing, vol.61, no. 4, pp. 409-417, 1995.

[7] C. Sanz, A. De Giusti, “A distributed solution for dynamic evidential reasoning applied to the classification of hyperspectral images”, The 5th World Multi-Conference on Systemics, Cybernetics and Informatics, July, 2001 , unpublished.