A New Method for Hyperspectral Data Classification

Hamid. Dehghani
Dept. of Electrical Engineering, Tarbiat Modaress Univercity
hdehghan@ modares.ac.ir .

Hassan. Ghassemian
Dept. of Electrical Engineering, Tarbiat Modaress Univercity
P. O. Box: 14115-111, Tehran, Iran
ghassemi@ modares.ac.ir

Abstract: As the number of spectral bands of high spectral resolution data increases, the capability to detect more detailed classes should also increase, and the classification accuracy should increase as well. Often, it is impossible to access enough training pixels for supervise classification. For this reason, the performance of traditional classification methods isn't useful. In this paper, we propose a new model for classification that operates based on decision fusion. In this classifier, learning is performed at two steps. In first step, only training samples are used and in second step, this classifier utilizes semilabeled samples in addition to original training samples. At the beginning of this method, spectral bands are categorized in several small groups. Information of each group is used as a new source and classified. Each of this primary classifier has special characteristics and discriminates the spectral space particularly. With using of the benefits of all primary classifiers, it is made sure that the results of the fused local decisions are accurate enough. In decision fusion center, some rules are used to determine the final class of pixels. This method is applied to real remote sensing data. Results show classification performance is improved, and this method may solve the limitation of training samples in the high dimensional data and the Hughes phenomenon may be mitigated.

Keywords: Remote Sensing, Hyperspectral, Data Analysis, Decision Fusion.

1. Introduction

Remotely sensed data are often used to determine the land cover composition of sites. This is made possible by the large amount of data acquired by different types of sensors such as multispectral and hyperspectral sensors. For information extraction from remote sensing images, it is necessary to classification them. For supervise classification, we are forced to use the training samples [1]. In supervised classification methods is assumed that the parameters of classifier can be estimated by the training samples. For accurate estimation and therefore accurate and reliable classification, enough training samples are necessary. For each class, in some classifiers such as MLC, the number of image bands determines the necessary minimum number of training samples. Since the number of image bands in the mentioned cases is large, even this minimum margin would be large and providing that, would be hard and expensive. This problem is sharp, if we have to use multisatellite sensors and multitemporal images [2]. To mitigate the small training sample problem, a new classifier is proposed in this paper. This classifier that is based on decision fusion, enhances estimation and hence improves classification accuracy by utilizing the classified samples (referred as semilabeled samples), in addition to the original training samples. This proposed adaptive classifier potentially has the following benefits:

A. The large number of semilabeled samples can enhance the estimation of the parameters, and therefore reduce the effect of the limited training samples problem.
B. The estimated parameters are more representative of the true class distribution.
C. This classifier is adaptive and can be improved when the new data of the considered scene is available.

2. Classification Based on Decision Fusion

We pursue an approach that can be used for high dimensional data classification as well as multisensor data classification. For this purpose, we consider all bands of image, careless their sensors. Hence we will deal with the classification problem of high dimensional data. With this strategy, even we can use the decision fusion rules for classifying the data of one sensor. To classify in this method, we follow the under mentioned flow chart.

Step. 1: The bands of image would be categorized according to different criterions such as minimum and maximum correlation. The number of bands in each group depends on total number of training samples.

Step. 2: The bands of each group are considered as the bands of a new source. The existing data in each source is used for primary classification. In this step, only existing training samples are used. In this paper two simultaneous three-layer back-propagation network and maximum likelihood classifiers are used.

Step. 3: For each source the output of these classifiers would be posterior probability, which specifies the degree of dependency of pixels to the classes of the given source. These posterior probabilities will be used for determining the class of pixels in a decision fusion center.

1) Decision Fusion Center Rule

Rule 1: Arithmetic Mean: In this rule, the arithmetic mean of the posterior probabilities related to each class is calculated [3].

$$C_j(X)=\sum_{i=1}^{N} p(w_j/x_i) \quad (1)$$

where N is the number of bands of image and $P(w_j|x_i)$ is posterior probability. The class, for which C is largest, is selected as the class of pixel.

Rule 2: geometric mean: The geometric mean of the posterior probabilities related to each class is calculated [4]. The considered pixel is assigned to the class that it's F is largest

$$F_j(X)=\prod_{i=1}^{N} p(w_j/x_i) \quad (2)$$

Rule 3: Neural network: The upper rules use specific law for combining the primary decisions and final deciding. But the Neural network method is nonparametric rule. Three-layer back-propagation network is used for this purpose. The inputs of this network are the posterior probabilities provided by primary classifiers [3].

3. Proposed Model

The new model has under additional steps [5]:

Step. 3: After primary classification, the class of all pixels is determined in each source.

Step. 4: Then the image is observed thoroughly and the pixels, which absolute majority primary classifiers have agreed on their class, are determined and marked.

Step. 5: The number of marked samples in each class is determined and the class that having the minimum number of marked samples is chosen.

Step. 6: New training or semilabeled samples for classes are selected from the marked pixels, the number of which would be the same as the minimum number specified in the previous step.

Step. 7: In the decision fusion center we use a three-layer back-propagation neural network to make the final decision for the class of pixels. To train this network we use new training samples in addition to original training samples. The functionality diagram of this classifier is shown in "Fig. 1".

4. Experiments

For comparison the different methods, an experiment was developed. The multispectral data used in this experiment, is an agricultural segment of Indiana State. This image has provided in 12 bands and its radiometric resolution is 8 bits. These bands are presented in table.1. Training and testing regions have selected from this image in 8 classes and shown in "Fig. 2(a). In this image, 18 pixels per class used as training samples.

Accuracy and reliability are the important measures that are applied in this paper..

Accuracy: α=n/A , Reliability: β=n/B (3)

n is the number of test samples are correctly classified, A is the total test pixels and B is the total test pixels.

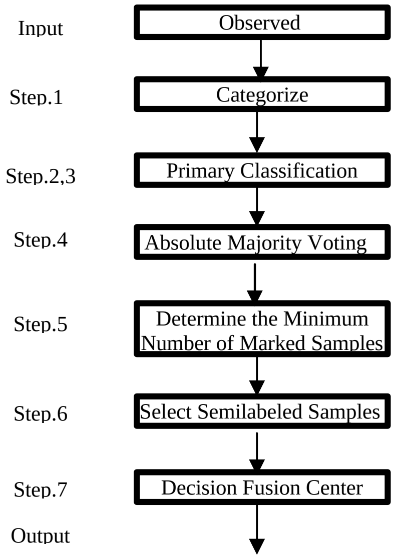

Fig.1. The diagram of adaptive classifier

We used all bands together, and classification was performed with MLC and three-layer back-propagation neural network classifiers. In neural network classifier for facilitating the training process, we used gray code instead of gray level. Hence the total number of input neurons was 96, the number of output neurons was 8, and network with 16 hidden units was completed. The result of these classifiers have presented in table.2. In decision fusion methods, initially it was necessary that the observed bands were categorized according to minimum and maximum correlation. The results of this process were three new sources that every one has 4 bands (presented in table.3). For primary classification in each source, both three-layer back-propagation neural network and maximum likelihood classifiers were used, simultaneously. For facilitating the training process, gray code was used. Hence the number of input neurons was 32, the number of output units was 8, and the number of hidden neurons was 16. In the rule 3 and proposed method, we used a three-layer back-propagation network in decision fusion center. Because we have three sources and each source was classified with two methods, therefore the number of input neurons was 3*2*8=48. The number of output neurons was 8, and for hidden layer, 36 neurons were selected. The results of classification methods have presented in table.4

5. Conclusions

As the number of spectral bands of high spectral resolution data increases, the capability to detect more detailed classes should also increase, and the classification accuracy should increase as well. Often the number of training samples used for supervised classification techniques is limited, thus limiting the precision with which class characteristics can be estimated. As the number of spectral bands becomes large, the limitation on performance imposed by the

Table 1. The band of image.

Band	1	2	3	4	5	6
Spectral bound (µm)	0.46 to 0.49	0.48 to 0.51	0.50 to 0.54	0.52 to 0.57	0.54 to 0.60	0.58 to 0.65
	7	8	9	10	11	12
	0.61 to 0.70	0.72 to 0.92	1.00 to 1.40	1.50 to 1.80	2.00 to 2.60	9.30 to 11.7

Table 2. The results of experimentation for non-decision fusion methods

Method	Accuracy	Reliability
Maximum likelihood	68.06	74.38
Neural network	69.15	72.03

Table 3. The new sources and their bands after grouping

Criterion	Maximum correlation	Minimum correlation
New source	Bands	Bands
Source-1	1,2,3,7	1,6,8,10
Source-2	4,5,6,12	2,4,11,12
Source-3	8,9,10,11	3,5,7,9

Table 4. The results of experimentation for the methods based on decision fusion

Criterion	Maximum correlation		Minimum correlation	
Rule	Accuracy	Reliability	Accuracy	Reliability
Arithmetic mean	69.63	70.50	88.15	84.25
Geometric mean	50.73	51.47	83.04	79.47
Network	68.06	71.39	87.63	83.23
Adaptive	**72.50**	**73.55**	**91.38**	**87.50**

limited number of training samples can become severe. In this paper, decision fusion techniques have been used to develop a supervise classification scheme for high dimensional data analysis.

In table.4, the results on minimum and maximum correlation criterions for categorizing the bands in image report that the minimum correlation is the better criterion for creation the new sources. In maximum correlation state, the information of these bands has maximum overlapping in comparison with minimum correlation state. Therefore the results of primary classification of minimum correlation sources are more accurate and hence in decision fusion center the final decisions about classes of pixels are made more accurate. The methods and rules in decision fusion center will be desired that aren't sensitive to the band categorizing criterions.

The comparison of the results in table.3 and table.4, show the effectiveness of the new model. This improvement occurred, because for training the network in decision fusion center, the large number of semilabeled samples was used and on the other hand, this samples are from a larger portion of the entire data set, hence the weights of this network are more assured. In other words, the extracted information by primary classifiers can help to create the more accurate decision fusion center. Certainly we can use the marked samples to complete the training the primary classifiers. Probably, after doing this work, the provided results are more accurate. Of course we must care to time of process

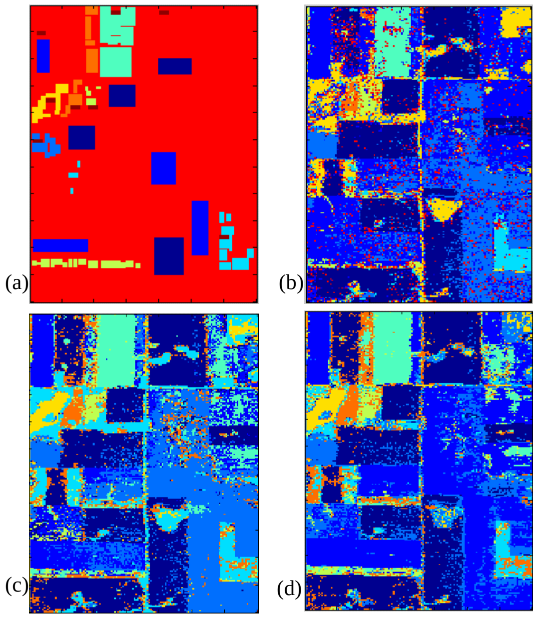

Fig.2. The result class maps. (a) ground truth map, training (brown color) and test regions. (b) The classification result of 12 bands by neural network. (c) Thematic map produced by new model with maximum correlation Criterion. (d) Class map produced by new model with minimum correlation Criterion.

and error propagation. In this case, if mislabeled pixels are belonged to semilabeled samples error strongly is propagated in all parts of model.

REFERENCES

[1] Richards A. John, 1999. *Remote sensing digital image analysis*. Third edition, Springer.

[2] Byeungwoo Jeon and landgrebe D., 1999. Decision fusion approach for multitemporal classification. IEEE Trans. Geosci. Remote Sens., vol. 37, no.3, pp.1227-1233.

[3] Giorgio G., 1999. Combination of neural and statistical algorithms for supervised classification of remote sensing images. pattern recognition letters.

[4] Benediktsson A., 1999. Classification of multisource and hyperspecrtal data based on decision fusion. IEEE Trans. Geosci. Remote Sens. vol. 37, no.3, pp.1367-1377.

[5] H. Ghassemian, H. Dehghani, Sept. 2003 "AN ADAPTIVE HYPERSPECTRAL CLASSIFICATION MODEL WITH A LIMITED TRAINING SAMPLES", Proceeding of ISPRS2003 International Workshop, Challenges in Geospatil Analysis.

[6] Jackson Q. and Landgrebe D., 2001. An adaptive classifier design for high-dimensional data analysis with a limited training data set. IEEE Trans. Geosci. Remote Sens., vol. 39, no.12,pp.2664-2679.

[7] Brumby P., 2002, Evolving feature extraction algorithms for hyperspectral and fused imagery. Space and Remote Sensing Sciences.

[8] Michael D. Richard and Richard P. Lippmann, 1991, Neural network classifiers estimate bayesian a posteriori probabilities, Neural computation 3, pp.461-483.

[9] David Landgrebe, "Hyperspectral Image Data Analysis", Dept. of EE, Purdue University, 2002.