Adaptive Bayesian Contextual Classification Based on Markov Random Fields

Qiong Jackson, *Student IEEE*, David Landgrebe, *Life Fellow, IEEE*
School of Electrical & Computer Engineering
Purdue University, West Lafayette, IN USA
qiong@ecn.purdue.edu and landgreb@ecn.purdue.edu

Adaptive Bayesian Contextual Classification Based on Markov Random Fields

Qiong Jackson, *Student IEEE*, David Landgrebe, *Life Fellow, IEEE*
School of Electrical &Computer Engineering
Purdue University

***Abstract*-In this paper an Adaptive Bayesian Contextual classification procedure that utilizes both spectral and spatial interpixel dependency contexts in statistics estimation and classification is proposed. Essentially, this classifier is the constructive coupling of an adaptive classification procedure and a Bayesian contextual classification procedure. In this classifier, the joint prior probabilities of the classes of each pixel and its spatial neighbors are modeled by the Markov Random Field. Experiments with real hyperspectral data show that, starting with a small training sample set, this classifier can reach classification accuracies similar to that obtained by a pixelwise maximum likelihood classifier with a very large training sample set. Additionally, classification maps are produced which have significantly less speckle error.**

I. INTRODUCTION

Hyperspectral image data acquired by new generation sensors contain extremely rich spectral and spatial attributes, which offer the potential to discriminate more detailed classes with high classification accuracy using a conventional Maximum Likelihood Pixel (MLP) classifier. This classifier performs classification by maximizing the class conditional probability. However, two difficulties inhibit this potential. First, due to the limited training sample size, the class statistics estimated from the limited training sample sets are less accurate and the resulting classifier performance is more limited. Additionally, in a conventional MLP classifier, it is explicitly assumed that the spectral properties are independent of the properties of all other pixels. Consequently, the MLP classifier may have difficulty distinguishing the pixels that come from different land-cover classes but have very similar spectral properties. The result is often a snow-like classification map. Several studies [1] [2] have shown that a contextual classifier that utilizes both spectral and spatial contextual information are able to better discriminate between the pixels with similar spectral attributes but located in different regions, allow reduction of the speckle error, and improve the classification performance significantly. However, this type of classifier also faces the problem of the small training sample size where the class conditional probability density must be estimated in the analysis of hyperspectral data [1][2].

In [3], it has been demonstrated that an adaptive maximum likelihood pixel classifier (AMLP) may alleviate the small training sample problem by including semi-labeled samples along with the training samples during the process of statistics estimation. Essentially, this classifier is formed by adding a feedback loop to a conventional ML classifier such that the loop carries additional class information generated from semi-labeled samples (classification outputs). The key to successful performance of this classifier is to establish a positive feedback process so the statistics estimation and classification can improve each other at each iteration. We have shown in [3] that higher initial accuracy and a large number of semi-labeled samples can allow the establishment of this positive feedback and lead to faster convergence of classification accuracy. However, as with a conventional MLP classifier, performance of this adaptive MLP classifier is limited by using just spectral information.

In this paper, an adaptive Bayesian contextual classifier that utilizes both spectral and spatial interpixel dependency contexts in statistics estimation and classification is proposed. Essentially, the proposed classifier is the combination of a Bayesian contextual classifier and an adaptive classification procedure. In this classifier, only interpixel class dependency context is considered, and the joint prior probabilities of the classes of each pixel and its spatial neighbors are modeled by the Markov Random Field. As an adaptive classification procedure, the statistics estimation and classification are performed in a recursive manner. Because usually a contextual classifier achieves higher accuracy than a MLP classifier, the proposed classifier has several advantages over the adaptive MLP classifier. First of all, the positive feedback may be easier to be established. Secondly, it may converge faster. Third, the final accuracy can be higher with much less speckle error. Compared with a conventional one-pass contextual classifier, this approach should mitigate the small training sample problem in the analysis of hyperspectral data.

II ADAPTIVE BAYESIAN CONTEXTUAL CLASSIFIER

In this section, the new adaptive Bayesian contextual classifier is developed that combines the adaptive procedure proposed in [3] with the Bayesian Contextual Iteration Conditional Modes (ICM) [4]. In this Bayesian Contextual or Maximum A Posterior Probability (MAP) classifier, the joint prior probabilities of the classes of each pixel and its spatial neighbors are modeled by the Markov Random Field. In other words, the spatial information is utilized in the MAP classifier. Compared to an MLP classifier, an MAP classifier performs classification by maximizing the posterior probability. In this new classifier, contextual information is incorporated into the process of a weighting factor computation and MAP classification. There are two reasons for this operation. One is to further emphasize the positive effect from the correctly classified semi-labeled samples and discourage the negative influence from the misclassified semi-labeled ones, and the second is to enhance the classification using contextual information in addition to the likelihood. In a manner similar to the adaptive procedure [3] and ICM [4], this new method is also an iterative process that achieves optimal statistics estimation and classification by starting with the initial estimate and classification based on training samples only, and repeating the following steps at

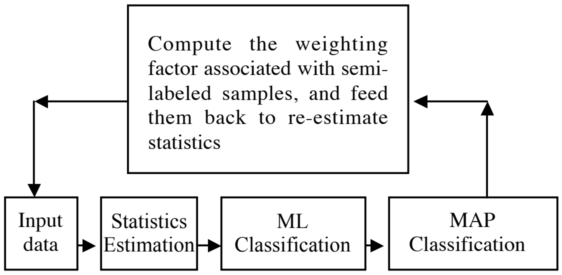


Fig. 1 A complete cycle of the Adaptive Bayesian Contextual Classifier

each iteration using training samples and semi-labeled samples. The flow chart in Fig. 1 illustrates one complete cycle of the adaptive contextual classifier. For notational purposes, in the following, the MLP and MAP classifiers at each cycle (except the first cycle) are denoted as ABC-ML and ABC-MAP classifiers, respectively. The MLP and MAP classifiers at the first cycle are the conventional ones because only training samples are used to estimate statistics at this cycle.

III. EXPERIMENTAL RESULTS AND DISCUSSION

For this experiment, the data is part of an airborne hyperspectral data flighline over the Washington DC mall. In this data set 191 bands in the 0.4 to 2.4 μm region of the visible and infrared spectrum are used. Since the data has high spatial resolution (about 5 meters), the testing samples and training samples are manually selected. There are 11 subclasses, and about 20 training samples per class selected. An average of 1200 test samples for each class are selected. Even though the training and testing samples can be identified in this case, selecting this many testing samples was a daunting task that took about 3 hours. By comparison, it only took about 15 minutes to select training samples. The desired scene classes are *Roof, Road, Path, Trees,* and *Grass*. However, two of these classes are spectrally multimodal and must be modeled by using several subclasses. Thus, five subclasses were used to form the class *Roof*, and two were used to model *Road*. In addition, the class *Shadow* was added so that the list of classes is suitably exhaustive.

This data set is a challenge to analyze for several reasons. First, classes are complex. There is a large diversity in the materials used in constructing rooftops, and consequently no single spectral response is representative of the class *Roof*. Even though some of the subclasses are spectrally quite different, some are quite similar. Subclasses of *Roof*, and *Road* are spectrally similar in that they may be made of similar materials, e.g, asphalt. Third, this data was collected during a dry season; most of lawns are not well grown and as a result, the class *Grass* and *Path* are difficult to differentiate, since some areas of grass are nearly bare soil, which is spectrally similar to the gravel of *Path*.

The classification accuracy at each iteration is graphed in Fig. 2. The following results may be observed: 1) The Adaptive Bayesian Contextual (ABC) classification procedure outperforms the Adaptive MLP classifier significantly. 2) After just three cycles the classification accuracy obtained by ABC-MAP converges with a net increment of about 13% over the initial MLP. 3) At each cycle, the ABC-MAP classifier achieves higher overall classification accuracy than the ABC-ML classifier does, and has less speckle errors (seen below). This indicates that contextual information does help to reduce the speckle error and accordingly improve classification performance. 4) During the first cycle the classification accuracy increment from the ABC-ML to ABC-MAP is about 3% However, the classification accuracy increase for the ABC-ML at the second cycle is more than twice that amount, i.e., about 8%. This indicates that using additional contextual information does improve the classification performance, but the improvement is limited. Essentially, the significant improvement of the classification performance is assumed to stem from better statistics estimates produced by the adaptive method.

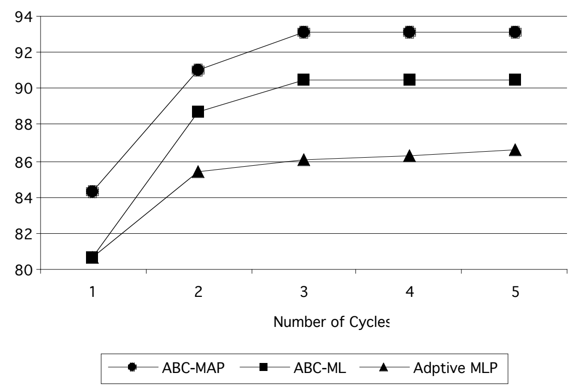
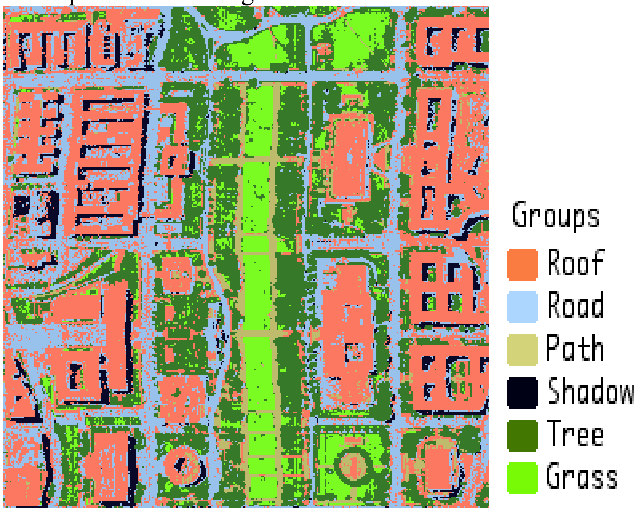
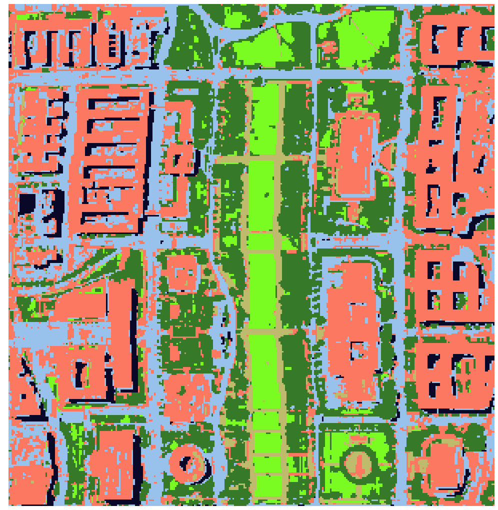


Fig. 2. Progression of the classification accuracy

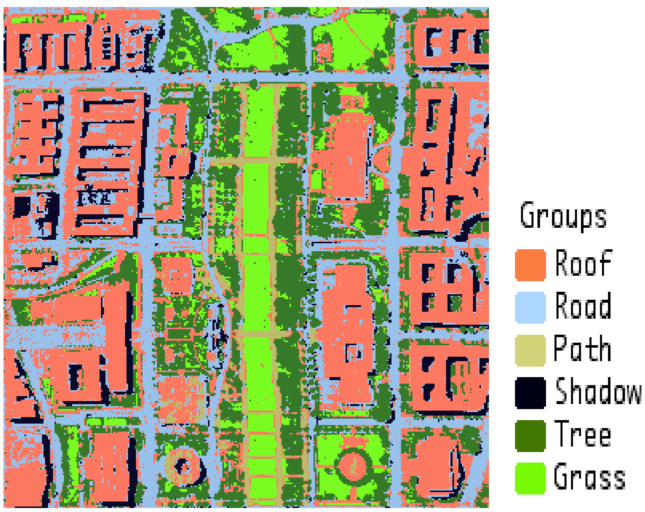
In addition to the classification accuracy, the segmented images are another way to assess the quality of the classification. Classification maps are provided in Fig. 3a through 3d. During the initial cycle, with limited training samples the initial statistics estimates are not very precise. It may be seen in Fig. 3a that classification errors occur in many places. These errors are mostly due to incorrectly estimated statistics and, to a lesser extent, the spectral similarity between classes. For instance, there is a great deal of similarity in spectral response between information classes *Roof* and *Road*, between *Path* and *Grass*, and between *Tree* and *Grass*. This type of error is reduced by the MAP classifier. However, errors of the first type still remain. In some areas the MAP classifiers create additional errors beyond those generated by the ML, leading to the loss of details. During the third cycle, both types of error have been greatly reduced by ABC-MAP. As a result, the objects in the classification map are well defined, clean and visually pleasant as indicated in Fig. 3d. On the contrary, even with good statistics, the adaptive MLP could not completely differentiate between the classes with similar spectral responses. As a result, there are still speckle errors in the classification map as shown in Fig. 3c.



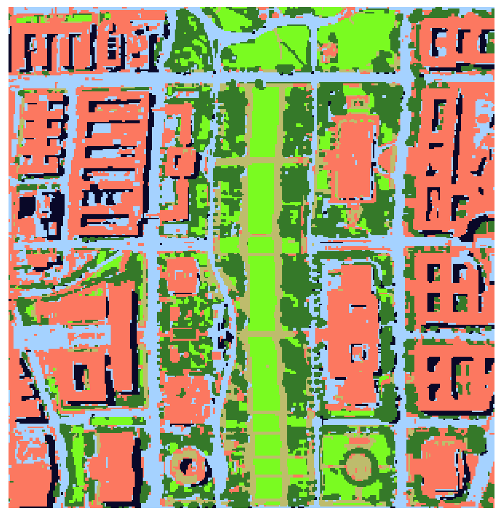
(a) MLPC at the first cycle



(b) MAPC at the first cycle



(c) Adaptive MLPC at the third cycle



(d) ABC-MAP at the third cycle

Fig. 3 Classification maps generated by four classifiers

III CONCLUSION

In this paper, an Adaptive Bayesian Contextual classification procedure based on Markov Random Fields is developed. In this procedure, the adaptive classifier and the Bayesian contextual classifier that is approximated by ICM [4] are integrated. As a result, the advantages of both classifiers are incorporated. The experimental results with hyperspectral data further reveal the benefits of this classification procedure. Starting with a limited training sample set, this method is able to steadily raise classification accuracy and eventually drive it close to the optimal value. The total improvement in the classification accuracy is significant and the convergence rate is fast even though a simple sub-optimal contextual classifier is used. This is significant because the classifier ICM [4] has a reputation of slow convergence when it is used alone.

REFERENCES

[1] B. Jeon and D. A. Landgrebe, " Spatio-temporal contextual classification of remotely sensed multispectral data," Proc. of 1990 IEEE Intern. Conf. on Syst., Man, and Cybern., Los Angeles, CA, pp. 342-344

[2] Yonhong Jhung and Philip H. Swain, " Bayesian contextual classification based on modified M-estimates and markov random fields", *IEEE Trans. Geosci. Remote Sensing*, vol.34, no. 1, pp. 68-75, Jan. 1996

[3] Q. Jackson and D. Landgrebe, "An Adaptive Classifier Design for High-Dimensional Data Analysis with a Limited Training Data Set," *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 39, No. 12, pp. 2664-2679, December 2001

[4] J. Besag, " On the statistical analysis of dirty pictures," *J. Royal Statist. Soc.*, vol. 68, pp.259-302, 1986