Managing the Spectral-Spatial Mix in Context Classification Using Markov Random Fields

X. Jia, *Senior Member, IEEE* , and J. A. Richards, *Fellow, IEEE*

***Abstract*—A straightforward method is presented for determining the most appropriate weighting of the spectral and spatial contributions in the Markov random field approach to context classification. The spectral and spatial components are each normalized to fall in the range (0,1) after which the appropriate value for the weighting coefficient can determined simply, guided by an assessment of the importance of the spatial contribution. Experimental results are presented using an artificial data set and real data recorded by the Landsat Thematic Mapper and Airborne Visible/Infrared Imaging Spectrometer.**

***Index Terms*—Markov random fields, spatial context, thematic mapping.**

I. Introduction

THE MARKOV random field (MRF) process incorporates spatial information into a classification by modifying the usual form of a discriminant function through the addition of a term that recognizes spatial correlations. It is one of a number of spatial consistency-seeking techniques that also include probabilistic relaxation [1]. When based on the maximum-likelihood estimation for handling the spectral data, and the Gibbs distribution/Ising model basis for the spatial term, the MRF-based discriminant function for the class on pixel m is usually expressed [1] as follows:

$$\mathrm{g_c}(\mathbf{x}_m) = -\frac{1}{2}\ln|\Sigma_c| - \frac{1}{2}(\mathbf{x}_m - \mathbf{m}_\mathrm{c})^t \times \Sigma_c^{-1}(\mathbf{x}_m - \mathbf{m}_\mathrm{c}) - \sum_{\partial m}\beta\left[1 - \delta(\omega_\mathrm{c}, \omega_{\partial m})\right] \quad (1)$$

where $\mathbf{m}_c$ and Σ_c are the mean vector and covariance matrix of class c, respectively, $\omega_{\partial m}$ is the labeling on the pixels in a neighborhood surrounding pixel m, and $\beta > 0$ is a parameter with value fixed by the user.

To use (1), there needs to be an allocation of classes over the scene before the last term can be computed. Accordingly, an initial classification is performed and (1) is applied iteratively over the full scene until the labeling has stabilized.

The first term in (1) is the likelihood of a pixel belonging to class c determined from the spectral measurements only. Although other models may not be based on the assumption of a Gaussian class-conditional density in (1), they can be employed as long as they provide posterior probabilities or their equivalents. For example, a minimum-distance classifier leads to

$$\mathrm{g_c}(\mathbf{x}_m) = -(\mathbf{x}_m - \mathbf{m}_\mathrm{c})^t(\mathbf{x}_m - \mathbf{m}_\mathrm{c}) - \sum_{\partial m}\beta\left[1 - \delta(\omega_\mathrm{c}, \omega_{\partial m})\right]. \quad (2)$$

The data range for the first term in either (1) or (2) can vary largely and is difficult to predict because it is affected by a number of factors, such as radiometric resolution of the data, number of spectral bands, number of classes, and class model used.

The second term in (1) and (2) is a measure of spatial support for the pixel's likelihood of belonging to class c. When all the K neighbors are labeled as class c, we have the case of the strongest support. The second term will then be zero so that there is no reduction on the likelihood of belonging to class c for the central pixel under examination. When none of the neighbors is labeled as class c, corresponding to the weakest support, the spectral likelihood value will be reduced by the full value of second term, i.e., $K\beta$.

An optimal value for the parameter β needs to be determined for each application of MRF because an inappropriate balance of the spectral and spatial terms can yield poor results; however, that is not a straightforward task because of the unpredictable range of the first term. Often, β is found experimentally [2], using a set of trials on a labeled training set. There are, however, theoretically more rigorous methods for finding the parameters in applying MRF in classification, but the complexity involved in their determination can sometimes obviate the benefit of a theoretical basis. Salzenstein and Pieczynski [3] use iterative conditional estimation as a means for estimating the parameters, whereas Tso and Mather [4] develop an approach based on genetic algorithms to improve the efficacy of the use of MRF in classification. Serpico and Moser [5] employ a Ho-Kashyap optimization for determining the parameters. Another recent approach by Farag *et al.* [6] estimates the spectral-spatial mixing parameters by iterating the overall classification map until there is no further improvement in the log likelihood of class membership. Interestingly, they base the spectral term on a support vector algorithm. Similarly, Lakshmanan and Derin [7] determine the parameters iteratively while simultaneously segmenting (classifying) the image. Their method also copes with additive Gaussian noise. Although this method is theoretically appealing, the authors have to invoke simplifications in their optimization process to render the approach tractable.

Manuscript received September 28, 2007; revised November 29, 2007.

X. Jia is with the School of Information Technology and Electrical Engineering, University College, The University of New South Wales, Australian Defence Force Academy, Campbell ACT 2600, Australia.

J. A. Richards is with the ANU College of Engineering and Computer Science, The Australian National University, Canberra ACT 0200, Australia (e-mail: John.richards@anu.edu.au).

Digital Object Identifier 10.1109/LGRS.2008.916076

In this letter, a convenient method is provided to manage the spectral-spatial mixing parameter in context classification.

II. Method

We have devised a simple heuristic method for estimating β in (1) or (2). As will be seen, it does not change the fundamental structure of the application of MRF as expressed in (1) and (2), and yet it is much easier to use than iterated methods that seek spectral-spatial consistency simultaneously and gives the same results as the more usual practical approach based on a number of experimental trials.

Rewrite (1) and (2) as

$$g_c(\mathbf{x}_m) = -e(c) - \beta a(c) \tag{3}$$

in which $e(c)$ is the spectral component of the discriminant function, and $a(c)$ is the spatial component. If we first normalize the spectral and spatial terms to the range (0,1), finding β is easier and more meaningful.

The spectral term is normalized over all classes

$$e(c)^* = \frac{e(c)}{\sum\limits_c e(c)}.$$

The spatial term is nomalized by the number K of members in the neighborhood chosen by

$$a(c)^* = \frac{a(c)}{K}.$$

We now search for the optimal weighting β^* of the normalized spectral and spatial terms

$$\mathrm{g_c}(\mathbf{x}_m) = -e(c)^* - \beta^* a(c)^*. \tag{4}$$

In (4), β^* directly indicates the proportion of the spatial information to be taken into account. Trials to find β^* can be narrowed greatly compared with the case for MRF in its original form because we often have some feeling for the relative importance of the spectral and spatial contributions.

Modification of (4) does not materially alter the algorithm. For example, we can multiply (3) throughout by an arbitrary constant A

$$g_c(\mathbf{x}_m) = -Ae(c) - A\beta a(c) = -Ae(c) - \beta' a(c).$$

Without affecting the decision rule, we can choose

$$A = \frac{1}{\sum\limits_{c=1}^{C} e(c)} \qquad \beta' = \frac{\beta^*}{K}$$

which leads to (4).

III. Experiments and Results

Fig. 8.9 of Richards and Jia [1] was used to demonstrate the β^* selection method above. The initial spectral classification accuracy is 88%.

Choosing the influence of the spatial term to improve labeling accuracy with MRF depends on the strength of the initial spectrally derived classification. Two sets of initial memberships were tested.

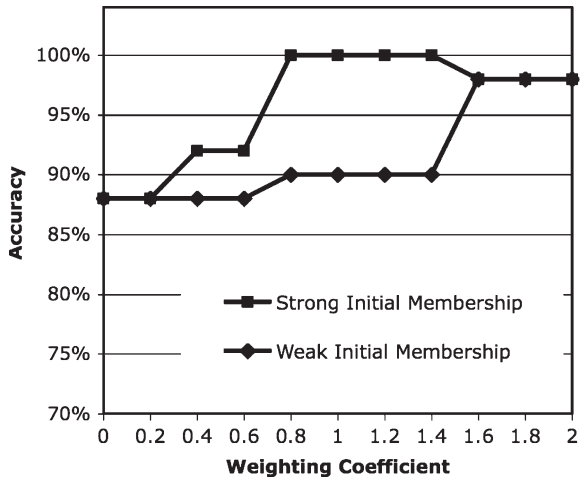

Fig. 1. Classification accuracy as a function of the weighting coefficient for an artificial image (Fig. 8.9 Richards and Jia [1]).

1) Correctly labeled pixels with posterior probabilities of 0.85 and wrongly labeled pixels with posterior probabilities of 0.65. This is called the strong initial membership case.
2) Correctly labeled pixels with posterior probabilities of 0.65 and wrongly labeled pixels with posterior probabilities of 0.85. This is called the weak initial membership case.

Three iterations were applied, given that the neighbors up to three pixels away are considered adequate. It can be seen in Fig. 1 that improvement in classification accuracy for case 1) starts even when a small proportion of the spatial term is applied. It reaches a maximum (100% accuracy) when the spatial weight is in the range 0.8–1.4. Higher weights are required for case 2), and the improved accuracy is limited to 98%. This is a result of the poor initial class membership for the correctly labeled pixels (and the good initial class membership for the wrongly labeled pixels). Here, the spatial term is more important, but the improvement is limited by the poor initial results.

Tests were also conducted on a 145 $\times$ 145 pixel portion of a Thematic Mapper (TM) image (with an eight-bit radiometric resolution) recorded over Tippecanoe County, IN. The data originally consisted of eight classes. We classified the data into the four largest classes (alfalfa/oats, soybeans, corn, and wheat) using a minimum-distance algorithm. All available pixels were allocated to those classes so that the performance of the algorithm could be inspected free of any unclassified regions. Even though some additional errors would be introduced into the original classification, that is not important in the context of examining the performance of procedures that seek to develop spatial consistency using neighborhood information. The numbers of training samples of 171, 185, 80, and 75, respectively, for the four classes were used for the spectrally based classification using a minimum distance algorithm.

Two different sets of testing data were used to check the operation of the algorithm. In the first, the testing data consisted

TABLE I
CLASSIFICATION ACCURACY FOR THE HOMOGENEOUS AREAS IN THE TM IMAGE

Standard MRF		Normalised MRF	
β	Accuracy	β*	Accuracy
0	80.77%	0	80.77%
1	80.89%	0.2	82.14%
2	81.46%	0.4	82.37%
3	81.91%	0.6	82.82%
4	82.25%	0.8	82.94%
5	82.25%	1.0	82.94%
10	82.59%	1.2	82.94%
20	82.94%	1.4	82.94%
30	83.28%	1.6	83.28%
40	83.28%	1.8	83.28%
50	83.28%	2.0	83.28%

of 121, 115, 88, and 44 homogeneous samples, respectively, of each of the four classes.

In the second testing set, all the known labeled data of 18 061 pixels were used for checking accuracy improvement with MRF. The main difference from the first set data is that the second set contains heterogeneous fields that include boundaries between classes. We chose these two different trials to assess the algorithm 1) at a per-field level with a well-trained classifier (because most context methods basically favor labeling with same-class neighbors) and 2) when field boundaries are involved in the testing data.

The results with the first testing set are shown in Table I. Both the conventional MRF method for determining β and our normalized process were applied. It can be seen that the value of β in the conventional approach has a much greater range requiring some time to find the right value. The weighting coefficient β^* is easier to select based on our feel for the significance of the spatial term. In the end, the same highest accuracy was achieved by both methods. The results were achieved after five iterations, which allow the neighbors up to five pixels away in each direction to contribute the central pixel's labeling. This was shown to be reasonable in an earlier study [8]. It can be seen that the improvement starts at a weight of 0.2 and is best when the weights are higher than 1.6 for homogenous testing fields.

Fig. 2 shows the results with the second testing set. The improvement is slightly peaked at a weighting coefficient of 1.6.

A 145 $\times$ 145 pixel Airborne Visible/Infrared Imaging Spectrometer (AVIRIS) data set (with 16-bit radiometric resolution) was used as well to test the new method. It covers an area of mixed agriculture and forestry in northwestern Indiana. The data were recorded in June 1992 with 220 bands. The ground truth consists of 16 classes, with 10 366 labeled pixels [9].

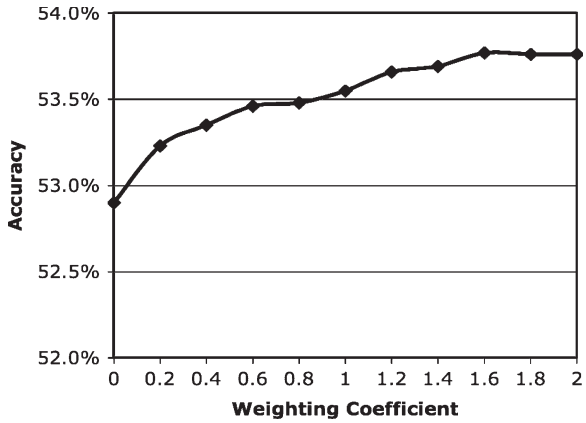

Fig. 2. Classification accuracy versus weighting coefficients for the heterogeneous areas in the TM image.

TABLE II
CLASSIFICATION ACCURACY FOR THE AVIRIS IMAGE

Standard MRF		Normalised MRF	
β	Accuracy	β*	Accuracy
0	79.20%	0	79.20%
1	83.11%	0.01	85.05%
10	85.08%	0.1	85.24%
50	85.15%	1	85.26%
100	85.22%	1.2	85.26%
200	85.26%	1.4	85.26%
300	85.26%	1.6	85.26%
400	85.26%	1.8	85.26%
500	85.26%	2	85.26%

Maximum-likelihood classification was performed using every fourteenth band. The initial spectral classification accuracy is 79.20%. Table II shows the improvement, with selection of the weighting coefficient by the method of this letter and by the conventional approach, after five iterations. It can be seen that the results are consistent with the TM data results, although with the normalized approach, much smaller values of β^* can provide accuracy improvement. Interestingly, that is related to the number of classes. With more classes (16 in the case of the AVIRIS trial compared with 4 for the TM data), the gap between the spectral likelihood of a pixel belonging to each class is smaller; therefore, the chance that the labeling is affected by the spatial term becomes higher.

Fig. 3 shows the initial (spectrally determined) labeling, and the outcome after MRF with a β^* of 1.6 is used over five iterations.

IV. DISCUSSION AND CONCLUSIONS

With the spectral and spatial components normalized, an appropriate value for the weighting coefficient for the spatial

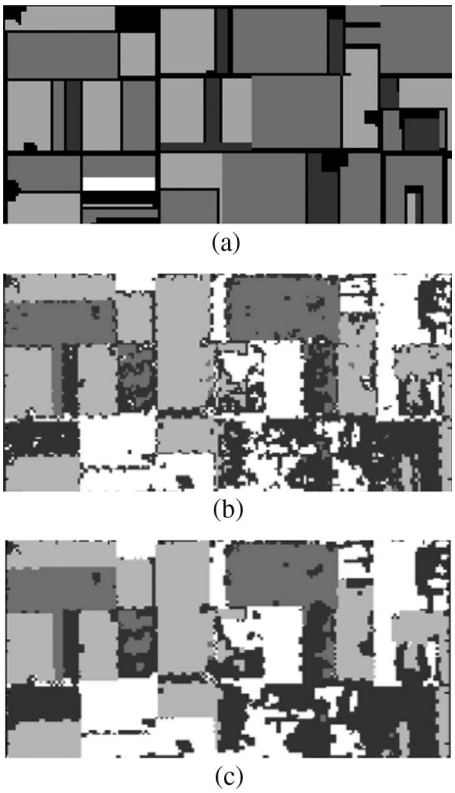

Fig. 3. Portion of the TM image. (a) Ground truth, (b) after original spectral classification, and (c) after application of MRF with weighting coefficient of 1.6. From dark gray to white, the classes are alfalfa/oats, soybeans, corn, and wheat, respectively. (Black pixels in (a) are the undefined background areas.)

term can easily be found. This modified version of MRF classification makes a systematic evaluation of the impact of the choice of weight on the ultimate classification results achieved possible.

In the MRF approach, the spatial component is implicitly based on the assumption that the neighboring pixels have the same class labels as the central pixel under examination. Thus, as the weight is increased, classification accuracy improves more in homogenous regions, but pixels at class boundaries are at the risk of overcorrection. When the spatial weighting is increased further, the overall accuracy may not be improved further, and beyond a certain critical level, classification accuracy may start to decrease as shown in the artificial data.

The optimal weighting coefficient is not only affected by the degree of the spatial correlation in the scene but also by the quality of the initial spectrally based classification. It is also weakly dependent on the number of classes, as inferred from our results above, because the degree of adjustment to the discriminant function for a given class will depend upon how many classes are represented among the neighboring pixels.

Spatial resolution is also a factor that influences spatial correlation and therefore affects how many iterations are required in the application of MRFs. For the current image with a spatial resolution of 20×20 m, three iterations are adequate, although five are generally preferred. It is recommended that the weight be determined by considering the spatial resolution of the sensor, scene correlation, and initial spectral classification uncertainty.

ACKNOWLEDGMENT

The TM and AVIRIS data and associated ground truth used were downloaded from http://dynamo.ecn.purdue.edu/~biehl/MultiSpec for which the authors thank Professor David Landgrebe.

REFERENCES

[1] J. A. Richards and X. Jia, *Remote Sensing Digital Image Analysis*, 4th ed. Berlin, Germany: Springer-Verlag, 2006.
[2] A. H. S. Solberg, T. Taxt, and A. K. Jain, "A Markov random field model for classification of multisource satellite imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 34, no. 1, pp. 100–113, Jan. 1996.
[3] F. Salzenstein and W. Pieczynski, "Parameter estimation in hidden fuzzy Markov random fields and image segmentation," *Graph. Models Image Process.*, vol. 59, no. 4, pp. 205–220, 1997.
[4] B. C. K. Tso and P. M. Mather, "Classification of multisource remote sensing imagery using a genetic algorithm and Markov random fields," *IEEE Trans. Geosci. Remote Sens.*, vol. 37, no. 3, pp. 1255–1260, May 1999.
[5] S. B. Serpico and G. Moser, "Weight parameter optimization by the Ho-Kashyap algorithm in MRF models for supervised image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 12, pp. 3695–3705, Dec. 2006.
[6] A. A. Farag, R. M. Mohaned, and A. El-Baz, "A unified framework for MAP estimation in remote sensing image segmentation," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 7, pp. 1617–1634, Jul. 2005.
[7] S. Lakshmanan and H. Derin, "Simultaneous parameter estimation and segmentation of Gibbs random fields using simulated annealing," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 11, no. 8, pp. 799–813, Aug. 1989.
[8] J. A. Richards and X. Jia, "A Dempster–Shafer relaxation approach to context classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, pt. 2, no. 5, pp. 1422–1431, May 2007.
[9] [Online]. Available: http://dynamo.ecn.purdue.edu/~biehl/MultiSpec