Hyperspectral remote sensing image classification based on decision level fusion

Peijun Du (杜培军)[1*], Wei Zhang (张 伟)[2], and Junshi Xia (夏俊士)[1***]**

[1]*Key Laboratory for Land Environment and Disaster Monitoring of State Bureau of Surveying and Mapping of China, China University of Mining and Technology, Xuzhou 221116, China*

[2]*Hebei Bureau of Surveying and Mapping, Shijiazhuang 050031, China*

[*]*Corresponding author: dupjrs@cumt.edu.cn;* [**]*corresponding author: cumtwzh@163.com;*

[***]*corresponding author: xiajunshi@126.com*

Received September 9, 2010; accepted November 6, 2010; posted online February 21, 2011

To apply decision level fusion to hyperspectral remote sensing (HRS) image classification, three decision level fusion strategies are experimented on and compared, namely, linear consensus algorithm, improved evidence theory, and the proposed support vector machine (SVM) combiner. To evaluate the effects of the input features on classification performance, four schemes are used to organize input features for member classifiers. In the experiment, by using the operational modular imaging spectrometer (OMIS) II HRS image, the decision level fusion is shown as an effective way for improving the classification accuracy of the HRS image, and the proposed SVM combiner is especially suitable for decision level fusion. The results also indicate that the optimization of input features can improve the classification performance.

OCIS codes: 100.4145, 280.4788.

doi: 10.3788/COL201109.031002.

Hyperspectral remote sensing (HRS) is viewed as one of the most evolving and most promising technologies for advanced earth observations in the 21st century. Traditionally, HRS image classification is implemented by a single classifier with the original hyperspectral data and other derived features as input. Examples of classifiers are the support vector machine (SVM), the maximum likelihood classifier (MLC), and the back-propagation neural network (BPNN). These methods have proven their effectiveness in many applications, however, there are still some problems. Firstly, each classifier has its own merits and limitations, and achieving the desired accuracy using a single classifier is often difficult[1,2]. Secondly, the adjacent wavebands of HRS data are highly correlated, thus the simultaneous use of all bands cannot assure high accuracy. Due to the limitations of both the classifiers and data, finding new solutions to improve the classification performance is necessary. Decision level fusion, using a specific criterion or algorithm to integrate the results of different classifiers, has shown great benefits in improving the classification accuracy of multi-source remote sensing images[3,4]. After a survey of HRS classification techniques and decision level fusion algorithms, some issues on HRS image classification based on decision level fusion are explored in this letter.

Many decision level fusion algorithms have been developed. After comparing their suitability and performance for remote sensing image classification, we selected three fusion strategies for this study: the improved evidence theory, the linear consensus, and the SVM combiner.

Evidence theory is also known as the Dempster-Shafer (D-S) evidence theory, which was first applied by Dempster and then developed by Shafer. Compared with the Bayesian theory, the D-S evidence theory assigns probability to sets and is able to handle the uncertainty caused by unknown factors[5]. The D-S evidence theory uses the discrimination framework, the confidence function, the likelihood function, and the probability allocation function to represent and process information. Supposing that $\Theta = \{C_1, C_2, \cdots, C_i, \cdots, C_M\}$ is the discrimination framework and M is the number of classes, the basic probability allocation function m is a function from 2^Θ to [0, 1] meeting the requirements of

$$\begin{cases} m(\phi) = 0 \\ \sum\limits_{A \subseteq \Theta} m(A) = 1. \end{cases} \tag{1}$$

If there are two or more different evidences, the orthogonal sum can be used to combine those evidences. Assuming that Z_1, Z_2, $\cdots$, and Z_n are the probability allocation functions corresponding to evidences F_1, F_2, $\cdots$, and F_n, the orthogonal sum $Z = Z_1 \oplus Z_2 \oplus \cdots \oplus Z_n$ is

$$Z(\phi) = 0, \tag{2}$$

$$Z(A) = K^{-1} \times \sum_{\cap A_i} \prod_{1 \le i \le n} Z_i\left(A_i\right), \tag{3}$$

$$K = \sum_{\cap A_i \neq \phi} \prod_{1 \le i \le n} Z_i(A_i). \tag{4}$$

When various evidences are inconsistent or contradictory to each other, the combined result of the D-S evidence theory may be unreasonable. A modified evidence combination algorithm was proposed and tested by Sun *et al.*[6], which proved superior to the traditional method in processing evidence that were contradicting and highly inconsistent. For a remote sensing image, different classifiers may generate different classification labels, resulting in generation of evidence with high contradiction, so the modified evidence combination is applied to the classification integration of HRS images. The

1671-7694/2011/031002(4)

detailed equations are as follows[6]:

$$k_{ij} = \sum_{\substack{A_i \bigcap A_j = \phi \\ A_i \in F_i, A_j \in F_j}} Z_i(A_i) Z_j(A_j), \tag{5}$$

$$\tilde{k} = \frac{1}{n(n-1)/2} \sum_{i<j} k_{ij}, \tag{6}$$

$$\varepsilon = \mathrm{e}^{-\tilde{k}}, \tag{7}$$

$$Z(A) = p(A) + K\varepsilon q(A), \quad A \neq \Phi, \Theta, \tag{8}$$

$$Z(\Theta) = p(\Theta) + K\varepsilon q(\Theta) + k(1-\varepsilon), \tag{9}$$

$$p(A) = \sum_{\substack{A_i \in F_i \\ \bigcap_{i=1}^{n} A_i = A}} Z_1(A_1) Z_2(A_2) \cdots Z_n(A_n), \tag{10}$$

$$q(A) = \frac{1}{n} \sum_{i=1}^{n} Z_i(A), \tag{11}$$

where ε is the confidence of the evidence, $\tilde{k}$ is the average contradiction level between two evidences, and K is the total contradiction level of all evidences. This evidence combination method can reduce the limitations caused by high evidence inconsistency.

For multiple classifier combination in remote sensing, each classifier result can be viewed as a piece of evidence. The probability allocation function is represented by the classification accuracy of a specific class. For example, if a pixel is classified to the ith class, the basic probability is $m(C_i) = P_i$, $m(\Theta) = 1 - P_i$, where P_i is the accuracy of the ith class given by the specific classifier. After the evidence combination is completed, the label with the maximum evidence confidence is selected as the final class.

The consensus theory is a popular method for multiple classifier combination and is suitable for integrating multiple outputs of the category probability generated by all member classifiers. Two commonly used models of the consensus theory are the linear consensus model and the logarithm consensus model. The principle of the linear consensus model is given by[7,8]

$$T_j(X) = \sum_{i=1}^{N} p_i(C_j|X) \lambda_{ij}. \tag{12}$$

Each classifier is regarded as an expert and the output element corresponding to X is its membership degree, confidence level, or probability to every class. $T_j(X)$ is the membership of the unlabeled pixel X to class j after combining multiple classifiers. $p_i(C_j|X)$ is the probability or the confidence level of X belonging to class j by the ith classifier. λ_{ij} is the classification accuracy (producer accuracy) of classifier i to class j and it represents the importance degree as a weight. Within the member classifiers, the SVM, BPNN, MLC, and the decision tree classifier (DTC) may generate the probability value directly, but the minimum distance classifier (MDC) should use the following transformation to calculate the probability value:

$$P(X \in C_i|X) = \frac{1/d_k(C_i|X)}{\sum\limits_{i=1}^{M} 1/d_k(C_i|X)}, \tag{13}$$

where $d_k(C_i|X)$ denotes the Mahalanobis distance between the spectral vector of the element X and the center of the ith class.

After this transformation, the decision level fusion can be used based on the class probability output of each classifier. In the decision level fusion, we can decide the class of element X according to the biggest membership degree derived from Eq. (12).

In the hierarchical classifier framework of SVM combiner for the decision level fusion, the probability output from each individual classifier is used as the input of the SVM combiner or classifier for the next level. The inaccuracy of class probability and inconsistency of inter-classifiers in the first level can be reduced by processing and classifying their probability outputs in the second level. The SVM and the BPNN classifiers usually have very good decision level fusion abilities, so the SVM classifier is adopted as the decision level combiner. The structure is shown in Fig. 1.

To obtain multiple classification results for the decision level fusion, providing every member classifier with identical or different input features is necessary. We designed four strategies to organize the input features. The first scheme is the most commonly used, in which the original hyperspectral data are used by all the classifiers. The second scheme is an improvement of the first one, in which all classifiers still use identical input features but the feature set consists of both the original data and the texture features derived from the original data. In the third scheme, all wavebands are divided into several groups based on the interband correlation analysis. Each group of data together with the texture features extracted is used by a specific classifier, which means that the feature inputs for multiple classifiers are different but every group of data should be a representative subset of the original data. In the fourth scheme, the first ten components derived from the maximum noise fraction (MNF) transformation of the original data and the texture features are used as input for all member classifiers. Figure 2 shows the flowchart of the four feature combination schemes.

Airborne hyperspectral operational modular imaging spectrometer (OMIS) II image was used in the experiment. The image size was 400 rows with 400 pixels on each row. After removing five bands that contained heavy noise, the image with 59 bands was used for classification. The training samples and testing samples were selected independently from the image. Figure 3 is the false color composite of the HRS image using bands 61, 21, and 11 as the red, green, and blue (R, G, and B) components, respectively.

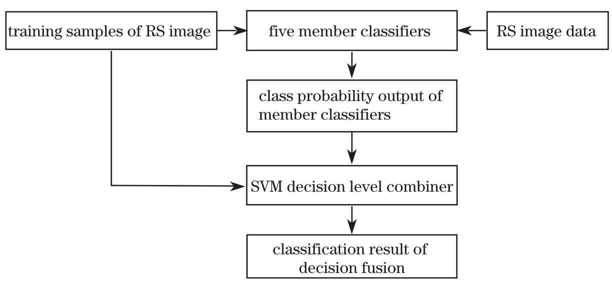

Fig. 1. Structure of decision level fusion using SVM combiner. RS: remote sensing.

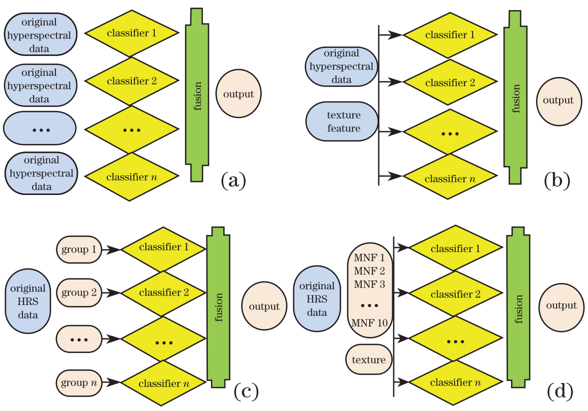

Fig. 2. Flowcharts of four feature input schemes. (a) Original data for all classifiers; (b) original data and textural feature for all classifiers; (c) different input feature sets for each classifier; (d) MNF components and texture for all classifiers.

Fig. 3. False color composite of original image.

In the decision level fusion, the members in the classifier ensemble include SVM, BPNN, MLC, DTC, and MDC. The mean and variance derived from the gray-level co-occurrence matrix (GLCM) of band 18, which had the biggest variance, were adopted as the texture feature in the second scheme. The grouping results, which are calculated based on the correlation of adjacent bands in the third scheme, are 1–7, 8–20, 21–32, 33–55, and 56–59. The results are classified separately using the five member classifiers.

The classification results are shown in Figs. 4–7. Table 1 summarizes the accuracy of different schemes.

For different classifier inputs, the second scheme, which includes the spectral and texture features, and the fourth scheme, which uses MNF transformation and texture features, obtain higher classification accuracy than using the original data alone. The worst classification accuracy

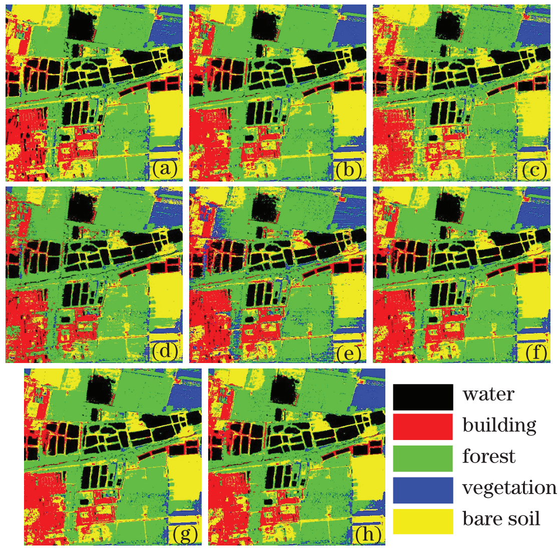

Fig. 4. Classification results of the first scheme. (a) SVM; (b) BPNN; (c) DTC; (d) MDC; (e) MLC; (f) evidence theory; (g) linear consensus; (h) SVM combiner.

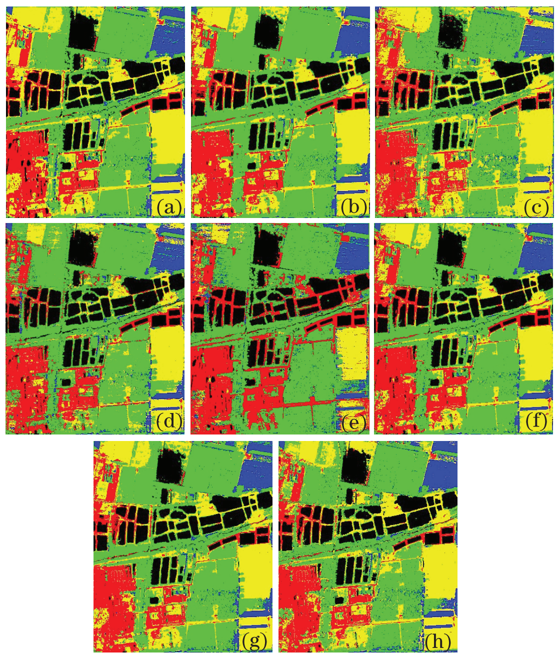

Fig. 5. Classification results of the second scheme. (a) SVM; (b) BPNN; (c) DTC; (d) MDC; (e) MLC; (f) evidence theory; (g) linear consensus; (h) SVM combiner.

Table 1. Classification Accuracy and Kappa Coefficient of Each Scheme

Classification Method	The First Scheme		The Second Scheme		The Third Scheme		The Fourth Scheme	
	Overall Accuracy (%)	Kappa	Overall Accuracy (%)	Kappa	Overall Accuracy (%)	Kappa	Overall Accuracy (%)	Kappa
SVM	90.72	0.88	91.21	0.89	86.11	0.83	91.70	0.90
BPNN	92.53	0.91	91.20	0.89	80.74	0.76	93.51	0.92
DTC	87.86	0.85	89.32	0.87	86.88	0.84	91.49	0.89
MDC	91.42	0.89	91.84	0.90	79.41	0.74	91.14	0.89
MLC	87.16	0.84	88.21	0.85	88.83	0.86	92.11	0.90
Evidence Theory	92.18	0.90	92.67	0.91	90.79	0.88	94.00	0.92
Linear Consensus	92.53	0.91	92.67	0.91	89.88	0.87	92.32	0.90
SVM Combiner	92.67	0.91	96.86	0.96	90.65	0.88	93.44	0.92

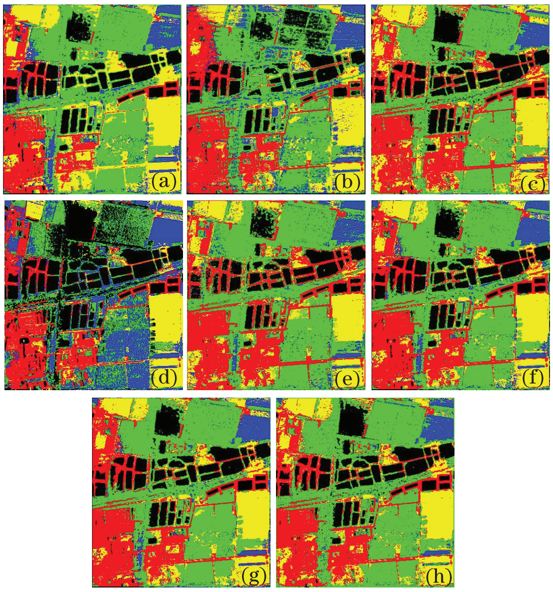

Fig. 6. Classification results of the third scheme. (a) SVM; (b) BPNN; (c) DTC; (d) MDC; (e) MLC; (f) evidence theory; (g) linear consensus; (h) SVM combiner.

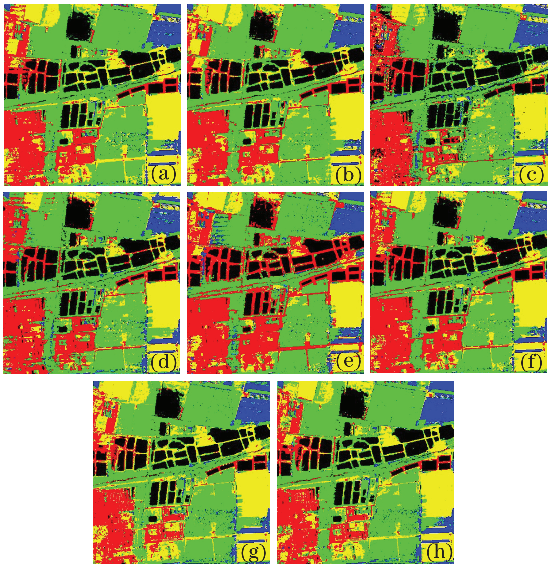

Fig. 7. Classification results of the fourth scheme. (a) SVM; (b) BPNN; (c) DTC; (d) MDC; (e) MLC; (f) evidence theory; (g) linear consensus; (h) SVM combiner.

is generated by the grouping scheme based on interband correlation, which is the third scheme. Therefore, we can conclude that both the textural features and the MNF transformation have positive effects on classification, and that the MNF transformation may help in overcoming the Hughes phenomena to a certain extent. However, if only the relevant part of the whole band set is used, the classification accuracy is not satisfactory.

For the experimental results of the three decision level fusion strategies, the overall accuracy of decision level fusion is comparable to or higher than the best single member classifier. Out of all the classification schemes, the SVM combiner for decision level fusion obtains the highest accuracy, or quite close to the highest accuracy, which means that it is a very good tool for decision level fusion.

In conclusion, we present some ideas and experiments on HRS image classification based on decision level fusion. Experiments have been done to compare three different decision level fusion methods using different input feature sets. By comparing the performance of four kinds of feature combination schemes, the SVM combiner and the improved evidence theory are found to obtain high accuracy when using the original data together with the textural feature data and components from the MNF transformation together with the textural feature data, respectively. We therefore conclude that the decision level fusion, especially the SVM combiner, is effective in HRS image classification, and that the introduction of textural features is helpful to improve further the classification accuracy. In our future research, the application scopes, as well as the advantages and disadvantages of each decision level fusion strategy will be compared and analyzed.

This work was supported by the National "863" High Technology Research and Development Program of China (No. 2007AA12Z162), the National Natural Science Foundation of China (No. 40871195), and the Scientific Research Foundation for the Returned Overseas Chinese Scholars, Ministry of Education of China.

References

1. D. Lu and Q. Weng, Int. J. Remote Sensing **28,** 823 (2007).
2. H. Su and Y. Sheng, Chin. Opt. Lett. **8,** 811 (2010).
3. H. T. X. Doan and G. M. Foody, Int. J. Remote Sensing **28,** 4609 (2007).
4. Y. Zhang, B. He, and X. Li, Chin. Opt. Lett. **7,** 781 (2009).
5. C. Liu, J. Dai, W. Zhong, Y. Wang, and D. Xia, Pattern Recognition and Artificial Intelligence (in Chinese) **16,** 213 (2003).
6. Q. Sun, X. Ye, and W. Gu, Acta Electron. Sin. (in Chinese) **28,** (8) 117 (2000).
7. S. Zhao, P. Li, and X. Feng, Journal of Zhengzhou Institute of Surveying and Mapping (in Chinese) **24,** 247 (2007).
8. J. A. Benediktsson, J. R. Sveinsson, and P. H. Swain, IEEE Trans. Geosci. Remote Sensing **35,** 833 (1997).